Self-Assembling Peptide Nanostructures for Controlled Cell-Material Interactions

Dissertation zur Erlangung des Grades
„Doktor der Naturwissenschaften“ (Dr. rer. nat.)
im Promotionsfach Chemie

am Fachbereich 09
Chemie, Pharmazie, Geographie und Geowissenschaften
der Johannes Gutenberg-Universität Mainz
vorgelegt von

Kübra Kaygisiz

geboren in Bad Kreuznach

Mainz, August 2023

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

MAX-PLANCK-INSTITUT FÜR POLYMERFORSCHUNG

The work in this thesis was carried out in the period from February 2020 to August 2023 under the supervision of [redacted] at the Max Planck Institute for Polymer Research, Mainz.

Reviewer 1: [redacted]
Reviewer 2: [redacted]
Dean: Prof. [redacted]

Date of Submission: 24.08.2023
Date of Oral Examination: 02.11.2023

Mainzer Dissertation D77

ABSTRACT

The ability of peptides to self-assemble into complex hierarchical supramolecular structures is the basis for their diverse structural functions in Nature. This inspired researchers to design peptides from bottom-up to employ them as versatile scaffolds for cell-interactive materials in biomedical contexts. An example for applying this concept is utilizing peptide nanofibrils for therapeutic retroviral gene delivery. However, it is difficult to *a priori* design the bioactivity of a peptide, as even small changes in the peptide's sequence can drastically impact its self-assembly behavior and physicochemical properties on multiple length scales. The combination of experimental and computational approaches holds great promise to overcome these challenges in the quest for functional materials design for biomedical applications.

This thesis provides fundamental mechanistic insights for the interactions between peptide nanofibrils with cells and viruses together with strategies for peptide materials design in the context of clinical applications. Viral infectivity as a biological readout was chosen not only because of its clinical relevance but also because it offers a relatively simple readout to unravel the multi-parameter, multi-scale challenges for the identification and optimization of biologically active supramolecular peptide assemblies. The investigation of the peptide–cell–material interaction in this work can be divided into four main parts. The thesis begins with (I) optimizing the sequence of the enhancing factor C (EF-C) peptide that binds virions and facilitates gene delivery applications. This is followed by a (II) comprehensive investigation of essential physicochemical features, which are subsequently employed in (III) machine-learning assisted discovery of novel peptides. Finally, (IV) the application of peptides is explored with a focus on translational clinical aspects.

First, a peptide library was established by designing and synthesizing 163 short peptides based on the known infectivity-enhancing self-assembling peptide EF-C. To systematically study the structure-property-activity relationship of the obtained bioactive sequences, a data-mining strategy was applied to correlate various physicochemical, bioinformatic and biological properties of the peptides. Thereby, it was shown that peptides forming μm-sized β-sheet-rich aggregates are essential for promoting interactions of retroviruses with cells. These key-properties are also observed in short amyloidal peptides from diverse pathological and functional origin which emphasizes the universality of this requirement. To gain mechanistic insights of the peptide aggregation process and supramolecular fibril formation, molecular dynamics simulations were performed and revealed that fibril–fibril supramolecular networks are observed for non-dynamic, stable fibrils with low solvent accessible surface areas, *i.e.* high fibril surface hydrophobicity. Within the EF-C based library, peptides with an alternating amphiphilic sequence and a high ratio of hydrophobic amino acids exhibit a strong propensity to form μm-sized aggregates. This sequence pattern proved to be a reliable strategy for optimizing the sequence of EF-C and rationally design novel bioactive peptide fibrils.

Next, a machine learning workflow was performed on the EF-C based library to inverse-engineer *de novo* minimalistic peptides *via* continuous vector representations. The conceptual idea behind this approach is that relevant sequence information underlying infectivity enhancement can be extracted and decoded from a higher-dimensional vector space and this information can be applied for the design of new sequences. With this approach, a new

sequence space of self-assembling peptides was discovered. The results show that infectivity enhancement is governed by both electrostatic and hydrophobic contributions. As these newly found peptides display unexpected sequences that are not related to the original EF-C peptide, nor the derived peptide library and they were also not found in Nature, this machine-learning approach presents an efficient strategy to discover entirely new self-assembling peptide sequences with aggregation propensity and high bioactivity.

In the next step, these newly acquired insights about the impact of molecular design on aggregation properties were employed to develop biodegradable peptide fibrils suitable for *ex vivo* clinical applications. Specifically, peptide amphiphiles for viral gene delivery were evaluated. In contrast to the metabolically very stable amyloid-type peptides that contain an alternating amphiphilic sequence pattern, peptide amphiphiles are characterized by their block-type design and flexibility to customize fibril surface properties. By systematically studying a library composed of 35 peptide amphiphiles, it was found that peptide amphiphiles that form small aggregates can promote virus-cell interactions while being biodegradable. Via molecular dynamics simulations it was shown that their aggregation is dependent on the mobility of monomers in fibrils and surface hydrophobicity that is governed by hydrophobicity, flexibility, and order of amino acids within the peptide sequence. The findings from the PA library are well in line with amyloid-type peptides and emphasize the general importance of the key-physicochemical properties for promoting viral infectivity, *i.e.* positively charged, µm-sized aggregates formed by hydrophobic, β-sheet rich structures.

Finally, to showcase how these properties can be effectively harnessed to regulate interactions between fibrils and cells, a photo-decomposable peptide was engineered. The decomposition was achieved by incorporating a photosensitive linker into the peptide backbone that causes fragmentation of the alternating amphiphilic peptide sequence. This, in turn, results in decomposition of the β-sheet ordered supramolecular assembly and loss of its cell-adhesive properties. Due to the external trigger light, this process can be precisely controlled in space and time, allowing the manufacturing of cell-adhesive regions on surfaces.

Overall, the powerful combination of physicochemical and computational methods can circumvent the individual limitations of both approaches. The findings presented in this thesis showcase the bottom-up design and discovery of new functional peptide nanomaterials for biomedical applications.

ZUSAMMENFASSUNG

Die Fähigkeit von Peptiden, sich zu komplexen hierarchischen supramolekularen Nanostrukturen anzuordnen, ist eine wichtige Grundlage für ihre strukturelle und funktionelle Diversität in der Natur. Dies hat die Forschung dazu inspiriert, Peptide nach einem „Bottom-up“ Ansatz zu entwerfen, um sie als vielseitige Gerüststrukturen für zellinteraktive Materialien im biomedizinischen Kontext einzusetzen. Ein Beispiel für die Anwendung dieses Konzepts ist die Verwendung von Peptid-Nanofibrillen für die therapeutische retrovirale Genübertragung. Es ist jedoch eine Herausforderung, die Bioaktivität eines Peptids rational vorherzusagen, da selbst kleine Änderungen in der Peptidsequenz die Selbstassemblierung und physikochemischen Eigenschaften auf verschiedenen Längenskalen drastisch beeinflussen können. Die Kombination von experimentellen und computergestützten Ansätzen ist ein besonders vielversprechender Ansatz, um diese Herausforderungen auf dem Weg zu einem hochentwickelten Materialdesign zu meistern.

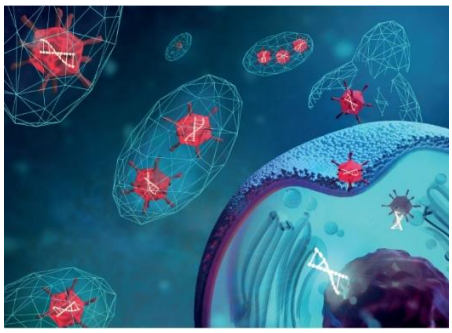
In dieser Arbeit werden grundlegende mechanistische Erkenntnisse über die Wechselwirkungen zwischen Peptid-Nanofibrillen mit Zellen und Viren sowie Strategien für das Design von Peptidmaterialien im Zusammenhang mit klinischen Anwendungen vorgestellt. Die Verstärkung der viralen Infektion ist eine biologische Untersuchungsmethode, die nicht nur wegen ihrer klinischen Relevanz gewählt wurde, sondern auch, weil sie eine Multiparameter- und Multiskalendimension für die Identifizierung und Optimierung biologisch aktiver supramolekularer Peptidstrukturen bietet. Die Untersuchung der Interaktionen zwischen Peptiden und Zellen ist in dieser Arbeit in vier Hauptteile gegliedert. Die Arbeit beginnt mit (I) der Optimierung der Sequenz des bekannten Peptids „Enhancing Factor C“ (EF-C), die Viren binden kann und die Genübertragung verstärken kann. Darauf folgt (II) eine umfassende Untersuchung der wesentlichen physikochemischen Merkmale, die anschließend in (III) der durch maschinelles Lernen unterstützten Entwicklung neuer Peptidsequenzen eingesetzt werden. Schließlich wird (IV) die Anwendung von Peptiden mit dem Schwerpunkt auf ihre klinischen Aspekte erforscht.
Zunächst wird eine Peptidbibliothek aufgebaut, indem kurze Peptide entworfen und hergestellt werden, die auf Derivaten eines bekannten infektiösitätssteigernden, selbstassemblierenden Peptids EF-C basieren. Um die Struktur-Eigenschafts-Wirkungs-Beziehung der bioaktiven Sequenz systematisch zu untersuchen, wurde eine Data-Mining-Strategie angewandt und verschiedene physikochemische, bioinformatische und infektiösitätssteigernde Eigenschaften von 163 Peptiden charakterisiert. Auf diese Weise wurde herausgefunden, dass es nicht die Fibrillen selbst sondern ihre µm-großen Aggregate sind, die reich an β-Faltblatt-Strukturen sind und die Wechselwirkung zwischen Retroviren und Zellen fördern. Diese Schlüsseleigenschaften werden auch bei kurzen amyloiden Peptiden verschiedener pathogener und funktioneller Herkunft beobachtet, was die mögliche Universalität dieser Merkmale unterstreicht. Innerhalb der EF-C-basierten Bibliothek zeigen Peptide mit einer alternierenden amphiphilen Sequenz und einem hohen Anteil an hydrophoben Aminosäuren eine starke Neigung zur Bildung von Aggregaten in µm-Größe. Dieses Sequenzmuster erwies sich als zuverlässige Strategie zur Optimierung der EF-C-Sequenz und zum rationalen Design bioaktiver Peptidfibrillen. Um Einblicke in den mechanistischen Ursprung der µm-großen Aggregation von supramolekularen Fibrillen zu gewinnen, wurden Molekulardynamik Simulationen durchgeführt. Die Ergebnisse zeigen, dass supramolekulare Fibril-Netzwerke bei nicht-

dynamischen, stabilen Fibrillen mit geringer lösungsmittelzugänglicher Oberfläche, d.h. hoher Hydrophobizität der Fibrillenoberfläche, beobachtet werden.
Als Nächstes wurde der experimentelle Datensatz der EF-C-basierten Bibliothek für einen maschinellen Lernprozess genutzt, um *de novo* minimalistische Peptide zu entwickeln. Genauer gesagt wurde eine kontinuierliche Vektordarstellung als Ansatz gewählt, um Peptidsequenzen zu abstrahieren. Die konzeptionelle Idee hinter diesem Ansatz ist, dass relevante Sequenzinformationen, die der Steigerung der Infektiosität zugrunde liegen, aus einem höherdimensionalen Vektorraum extrahiert und in neue Sequenzen dekodiert werden können. Auf diese Weise wurde ein neuer Sequenzraum von selbstorganisierenden Peptiden entdeckt, der zeigt, dass die Infektionsverstärkung sowohl durch elektrostatische als auch durch hydrophobe Beiträge bestimmt wird. Diese neu gefundenen Peptide basieren zum Teil auf unerwarteten Sequenzen, die nicht dem Trainingset der ursprünglichen Bibliothek ähneln und nach bestem Wissen auch nicht in der Natur zu finden sind. Dieser Ansatz des maschinellen Lernens stellt also eine hocheffiziente Methode zur Entdeckung völlig neuer bioaktiver, selbstorganisierender Peptide dar.
Die klinische *ex vivo*-Anwendung von EF-C basierten Peptiden wird allerdings durch deren schlechte Bioabbaubarkeit limitiert. Die neu gewonnenen Erkenntnisse über den Einfluss des molekularen Designs auf die Aggregationseigenschaften wurden daher genutzt, um biologisch abbaubare Peptidfibrillen für den viralen Gentransport zu entwickeln. Im Gegensatz zu den nur schwer abbaubaren Peptiden des Amyloid-Typs, die ein alternierendes amphiphiles Sequenzmuster enthalten, zeichnen sich Peptidamphiphile durch ihr blockartiges Design und ihre Flexibilität bei der Anpassung der Oberflächeneigenschaften der Fibrillen aus. Durch die systematische Untersuchung einer Bibliothek von 35 Peptidamphiphilen wurde herausgefunden, dass insbesondere kleine Aggregate die Interaktion zwischen Retroviren und Zellen fördern und gleichzeitig von Zellen effizient abgebaut werden können. Mittels Molekulardynamiksimulationen wurde festgestellt, dass die Aggregation von der Mobilität der Monomere in den Fibrillen und der Oberflächenhydrophobizität abhängt, die wiederum von der Hydrophobizität, der Flexibilität und der Reihenfolge der Aminosäuren innerhalb der Peptidsequenz bestimmt wird. Die Ergebnisse dieser Peptidamphiphil-Bibliothek stimmen gut mit den Peptiden des Amyloid-Typs überein und unterstreichen die allgemeine Bedeutung der physikochemischen Schlüsseleigenschaften, die identifiziert wurden. Für die Förderung der viralen Infektionsverstärkung sind positiv geladene, µm-große Aggregate, die von hydrophoben, β-faltblattreichen Strukturen gebildet werden, essenziell.
Schließlich wird gezeigt, dass diese Eigenschaften effektiv genutzt werden können, um die Wechselwirkungen zwischen Fibrillen und Zellen zu regulieren. Dies wird über eine stimulus-responsive Gruppe im Peptidrückgrat erreicht, die eine lichtgetriebene Fragmentierung der alternierenden amphiphilen Peptidsequenz ermöglicht, was schließlich zur Zersetzung des supramolekularen Aufbaus mit β-Faltblattanordnung und zum Verlust der zelladhäsiven Eigenschaften führt. Durch den optischen Trigger kann dieser Prozess räumlich und zeitlich genau gesteuert werden, was die Herstellung von klar definierten zelladhäsiven Oberflächen ermöglicht.

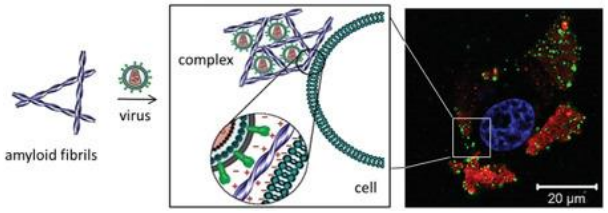
Insgesamt liefert diese Arbeit grundlegende Erkenntnisse und Designstrategien für funktionelle Peptid-Nanomaterialien. Die leistungsstarke Kombination von physikalisch-chemischen und computergestützten Methoden kann die Beschränkungen der jeweils einzelnen Methoden überwinden, um neue Peptidmaterialien für künftige biomedizinische Anwendungen zu entdecken.

GRAPHICAL ABSTRACTS

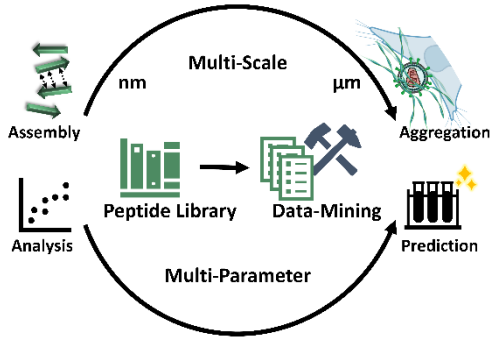
MATERIALS PROMOTING VIRAL GENE DELIVERY .....PAGE 37 / 75



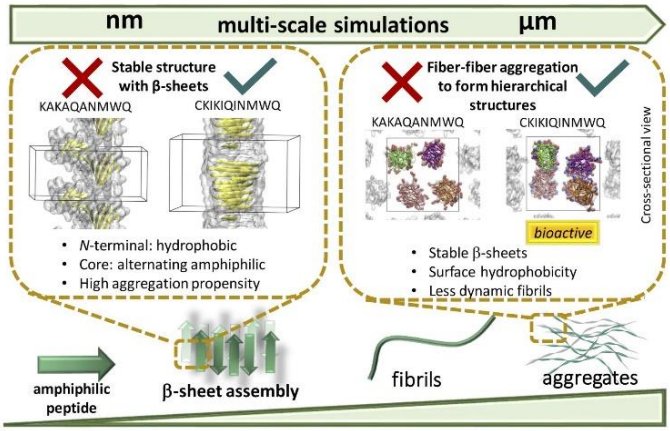
SUPRAMOLECULAR PEPTIDE NANOFIBRILS WITH OPTIMIZED SEQUENCES AND MOLECULAR STRUCTURES FOR EFFICIENT RETROVIRAL TRANSDUCTION..... PAGE 37 / 123



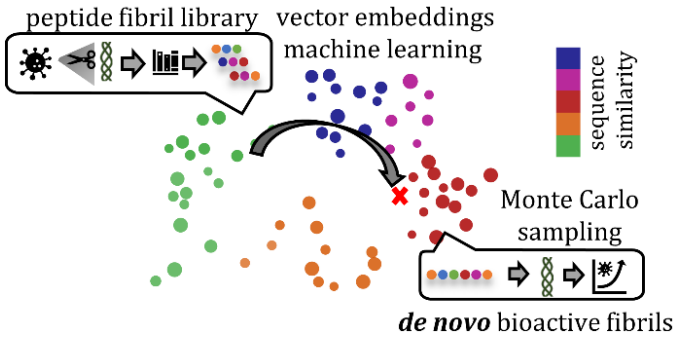
DATA-MINING UNVEILS STRUCTURE-PROPERTY-ACTIVITY CORRELATION OF VIRAL INFECTIVITY ENHANCING SELF-ASSEMBLING PEPTIDES..... PAGE 40 / 173



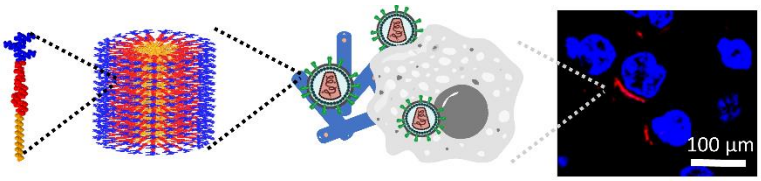
UNRAVELING STRUCTURE-PROPERTY RELATIONSHIPS IN SELF-ASSEMBLING PEPTIDE NANOFIBRILS VIA SIMULATIONS ..... PAGE 43 / 265



INVERSE DESIGN OF VIRAL INFECTIVITY-ENHANCING PEPTIDE FIBRILS FROM CONTINUOUS PROTEIN-VECTOR EMBEDDINGS ..... PAGE 46 / 351

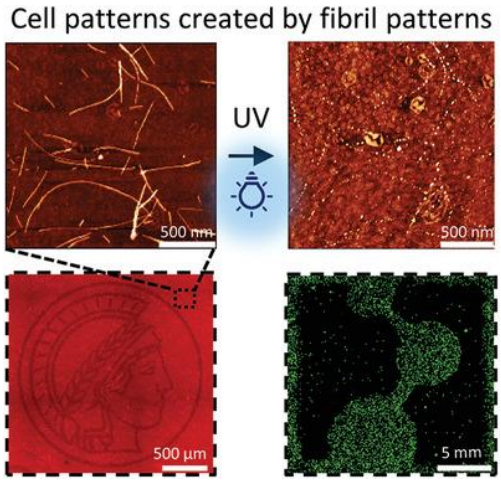


PEPTIDE AMPHIPHILES AS BIODEGRADABLE ADJUVANTS FOR EFFICIENT RETROVIRAL GENE DELIVERY..... PAGE 49 / 385



Peptide Amphiphiles: Biodegradable Viral Delivery Agents

PHOTOINDUCED AMYLOID FIBRIL DEGRADATION FOR CONTROLLED CELL PATTERNING..... ..... PAGE 52 / 463



CONTENTS

ABSTRACT I
ZUSAMMENFASSUNG III
GRAPHICAL ABSTRACTS V
CONTENTS VII
LIST OF ABBREVIATIONS IX
1. INTRODUCTION AND BASIC CONCEPTS 1
Supramolecular Materials for Biomedical Applications 1
Supramolecular Chemistry: Complexity by Design 3
Peptide Self Assembly 6
Supramolecular Interactions encoded in Amino Acids 6
Secondary Structure arising from the Amino Acids Order 8
Structural Diversity of the Amyloid Structure 9
Peptide Nanostructures as Materials for Cell-Applications 12
Interactions of Fibrils with Cells and Virions 12
Structure Controls Cell-Interactions 16
Design and Discovery of Functional Peptide Nanostructures 19
Screening Endogenous Peptides 19
Rational Design Strategies 20
Synthesis and Characterization of Peptide Fibrils 21
Computational Modeling and Prediction of Peptide Sequence and Structure 27
2. MOTIVATION AND OBJECTIVES 32
3. RESULTS AND DISCUSSION 37
3.1 Materials Promoting Viral Gene Delivery 37
3.2 Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction 37
3.3 Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides 40
3.4 Unraveling Structure-Property Relationships in Self-Assembling Peptide Nanofibrils via Simulations 43
3.5 Inverse Design of Viral Infectivity-Enhancing Peptide Fibrils from Continuous Protein-Vector Embeddings 46
3.6 Peptide Amphiphiles as Biodegradable Adjuvants for Efficient Retroviral Gene Delivery 49

3.7 Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning..... 52
4. CONCLUSION AND FUTURE PERSPECTIVES..... 55
5. REFERENCES..... 60
6. THESIS PUBLICATIONS..... 75
6.1 Materials Promoting Viral Gene Delivery..... 75
6.2 Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction.....123
Supporting Information.....136
6.3 Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides.....173
Supporting Information.....191
6.4 Unraveling Structure-Property Relationships in Self-Assembling Peptide Nanofibrils via Simulations.....265
Supporting Information.....313
6.5 Inverse Design of Viral Infectivity-Enhancing Peptide Fibrils from Continuous Protein-Vector Embeddings.....351
Supporting Information.....363
6.6 Peptide Amphiphiles as Biodegradable Adjuvants for Efficient Retroviral Gene Delivery 385
Supporting Information.....413
6.7 Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning.....463
Supporting Information.....474
7. APPENDIX.....503
Scientific Contributions to Peer-Reviewed Journals and Conferences.....503
List of Thesis Publications.....505
Declaration of Originality in Academic Work.....506
Curriculum Vitae.....507
ACKNOWLEDGEMENTS.....513

LIST OF ABBREVIATIONS

Amino acid code:

A	Alanine	**I**	Isoleucine	**R**	Arginine
C	Cysteine	**K**	Lysine	**S**	Serine
D	Aspartic acid	**L**	Leucine	**T**	Threonine
E	Glutamic acid	**M**	Methionine	**V**	Valine
F	Phenylalanine	**N**	Asparagine	**W**	Tryptophan
G	Glycine	**P**	Proline	**Y**	Tyrosine
H	Histidine	**Q**	Glutamine		

AFM	atomic force microscopy
AI	artificial intelligence
AL1	sequence IGSNVVTWYQQL
ATR	attenuated total reflection
CAR	chimeric antigen receptor
CCD	charge-coupled device
CD	circular dichroism
CD4	cluster of differentiation 4
CLEM	correlative light electron microscopy
CLSM	confocal laser scanning microscopy
Cryo-EM	cryogenic electron microscopy
DIC	*N*,*N*′-di*iso*propylcarbodiimide
DMSO	dimethyl sulfoxide
ECM	extracellular matrix
EF-C	enhancing factor C, (sequence QCKIKQIINMWQ)
EIPA	5-(*N*-ethyl-*N*-isopropyl) amiloride
EPI-X4	endogenous peptide inhibitor of C-X-C chemokine receptor type 4
ESI	electron spray ionization
Fmoc	fluorenyl methoxycarbonyl
FRET	Förster resonant energy transfer
FT-IR	fourier transform infrared spectroscopy
gp120	glycoprotein 120
H^0	hydrophobic amino acid with uncharged side chain
HeLa	Henrietta Lacks
HFIP	hexafluoro*iso*propanol
HIV-1	human immunodeficiency virus 1
HPLC	high-pressure liquid chromatography

IUPAC	international union of pure and applied chemistry
MALDI-ToF	matrix-assisted laser desorption ionization assisted time of flight
MD	molecular dynamics
ML	machine learning
MLV	murine leukemia virus
MS	mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
P^-	hydrophilic amino acid with negatively charged side chain
P^+	hydrophilic amino acid with positively charged side chain
P^0	hydrophilic amino acid with uncharged side chain
PA	peptide amphiphile
PAP	prostatic acidic phosphatase
PCA	principal component analysis
Rg	radius of gyration
RLU	relative light units
RMSF	root-mean-square fluctuation
SASA	solvent accessible surface area
SELFIES	self-referencing embedded string
SEVI	semen derived enhancer of virus infection
SMILES	simplified molecular-input line-entry system
SPPS	solid phase peptide synthesis
TEM	transmission electron microscopy
ThT	thioflavin–T
t-SNE	t-distributed stochastic neighbor embeddings
UPy	ureido-pyrimidinone
UV	ultraviolett

1. INTRODUCTION AND BASIC CONCEPTS

Supramolecular Materials for Biomedical Applications

Quality of life has been steadily increasing over the past decades.[1] While this is certainly a positive development, the longer life expectancy creates an aging society that is more susceptible to cancer and neurodegenerative diseases.[2] Pressured by this concerning threat the United Nations set agendas for addressing these challenges. In their roadmap for Sustainable Development they plan to *"ensure healthy lives and promoting well-being for all at all age"* by 2030.[3]

Personalized medicine is a technology that can address these challenges. It has not only the potential to revolutionize how hereditary diseases and oncology can be treated but is being successfully applied in recent years. One example is the emerging technology of genetic engineering of cells. These so-called gene therapies operate by modulating the genetic information in affected cells of a patient so that they acquire functions that are beneficial for curing the disease.[4] In this way, gene therapy directly contributes to advancing the human ability to cure previously untreatable hereditary diseases, such as spinal muscular atrophy,[5] and acquired diseases like cancer.[6] Since the first commercial production in 2003[7] and its first U.S. Food and Drug Administration approval of Kymriah as treatment for leukemia in 2017[8] the number of gene therapies has been rising each year (**Figure 1A**).[9] In the overwhelming number of these applications the cells of the patient are isolated and genetically modified *ex vivo* by inserting a therapeutic gene to express the desired properties, *e.g.* display chimeric antigen receptors (CAR) on T cells that recognizes tumor surface antigens and triggers the specific elimination of the tumor cell after re-administration into the patient (**Figure 1B**).[10] Hence, the CAR T cell therapy promises to be more effective and have fewer side effects than traditional treatments like chemotherapy.[10] A very recently approved gene therapy even addresses tissue engineering challenges, by locally enabling wound repair in patients with skin blistering epidermolysis bullosa genetic disorder through delivering the gene for expressing the missing human type VII collagen protein.[11]

These biomedical manufacturing processes require the use of precisely defined materials as adjuvants, *e.g.* to support the cellular environment[12] and assist the gene delivery.[13] This creates an urgent demand for potent adjuvants in therapeutical applications.

Non-pathogenic viral particles are frequently applied vectors for gene therapy[14,15] and vaccines[16,17] due to their inherent infectivity that ensures the transport and integration of foreign genes into cells.[18] Lentiviruses and γ-retroviruses cannot replicate and are less likely to cause inflammatory responses. Hence, they are the preferred choice for the safe integration of genetic information, in fundamental research as well as in most of the current clinical trials.[19–21] However, these viral vectors display low cell attachment resulting in low transduction efficiencies which is one of the major challenges limiting therapeutic efficacy.[22,23] To promote the cell-material interaction of viral particles, transduction-enhancing additives, such as polymers, lipids, peptides amongst others, have been applied (**Figure 1C**).[13]

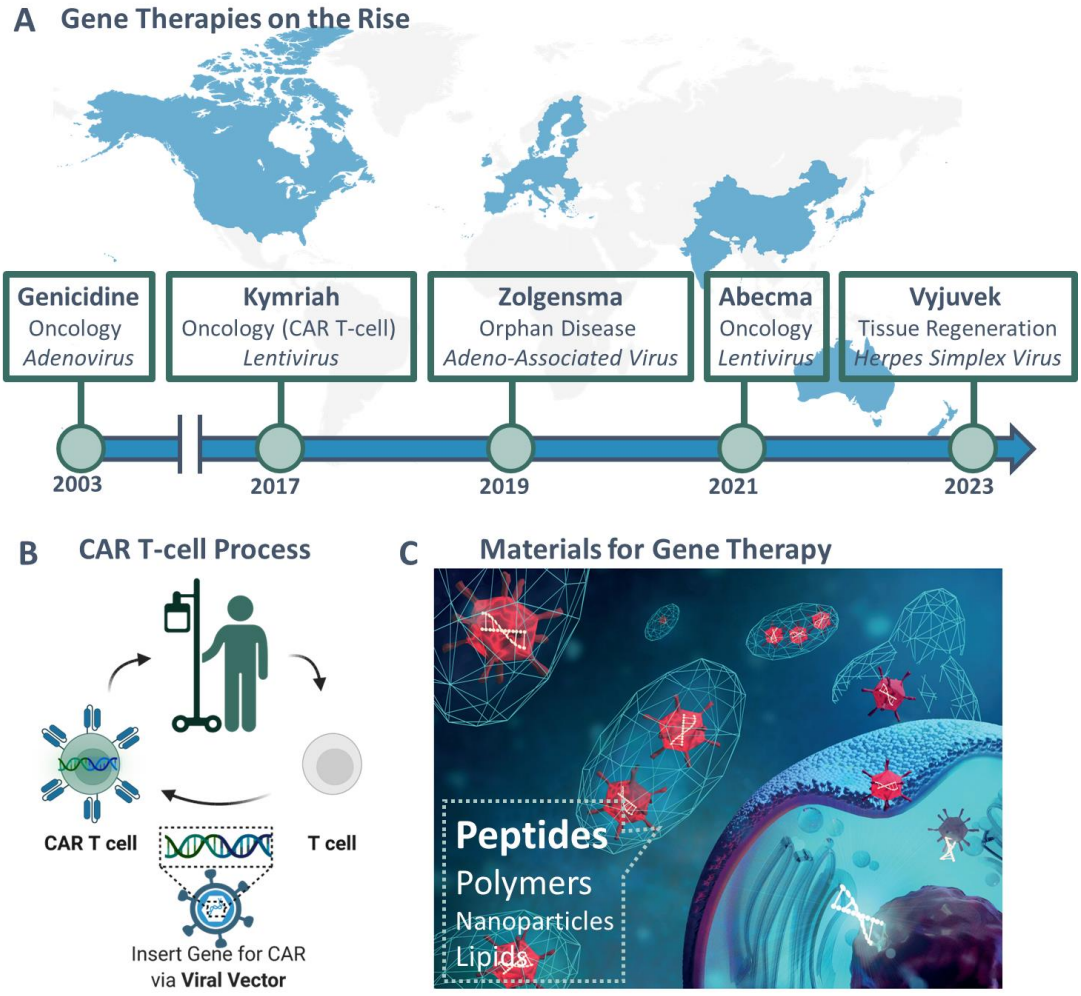


Figure 1 Schematic illustration of the significance of gene therapy. **A** World map visualizing geographic locations of current gene therapy approvals. Timeline highlights milestones in the development of different gene therapies: The first approval was in 2003 with gendicine from National Medical Product Administration in China.[7] The first U.S. Food and Drug Administration was in 2017 with Kymriah, a CAR-T cell therapy.[8] Recent approvals such as Vyjuvek in May 2023 started on combining tissue engineering and wound regeneration with gene therapeutic methods.[11] **B** Simplified visualization of a typical *ex vivo* CAR T-cell therapy process. Created with permission from BioRender.com **C** Artwork demonstrates that synthetic materials, such as polymer, lipids and peptides can be exploited to promote gene delivery and facilitate biomedical applications. Reproduced from Ref 13. This image is licensed under CC BY 3.0 (https://creativecommons.org/licenses/by/3.0/).

There is a significant gap between the demand and availability of tailored materials for specific healthcare applications, *i.e.* personalized medicine. At first glance, using naturally occurring materials for these biomedical applications seems appealing due to their inherent biocompatibility and intrinsic complexity that can provide cellular recognition motifs and hierarchical structures across various scales.[24] However, providing the ideal environment to facilitate viral vector delivery and support cell culture proves to be a challenge when utilizing naturally occurring materials. The main problem arises from the potential risk of pathogen transmission, as purification methods for biological materials are still limited.[24] Additionally, the natural origin of these materials leads to inherent variability between batches, further impeding their mass-scale production and use.[24]

To overcome the challenges, synthetic alternatives have been developed. Pioneering studies in this direction primarily aimed to precisely mimic specific aspects of materials from biological systems. However, these attempts often fell short in replicating the key features of (I) ordered nanostructures, (II) structural polymorphism, and (III) complexity exhibited by biological materials.[25]

Noteworthy, in Nature, all of these features are governed by self-organization processes that enable pattern formation and information transfer within complex systems. These self-organization processes can be observed across various length scales ranging from global ecosystems,[26] to bird flocking,[27] cellular networks and even down to the atomic level in structural biology (**Figure 2A**).[28] All of them have in common that the properties of their individual parts sum up to give rise to new behavior in the organized state. Small effects are amplified by cooperativity which is the main strategy of biology to create diverse materials at different length scales.[29] Exploiting these cooperative effects of molecular interactions on the nanoscopic scale can also enable new avenues for synthetic systems in the quest for complexity levels rivaling those found in Nature.

Supramolecular Chemistry: Complexity by Design

Supramolecular materials are formed through non-covalent intermolecular interactions that enable order across length scales. This allows for the creation of materials that exhibit controlled complexity and dynamics, resembling natural systems to a certain extent.[30] After decades of controversial discussions, pioneering concepts in the field of supramolecular chemistry have emerged by Hermann Staudinger, who received the Nobel Prize in 1953 for his discoveries in the field of macromolecular chemistry.[31] Staudinger's concept of macromolecules questioned the dogma that molecules cannot exceed a certain size and established the concept of the *polymer*, that is composed of *monomers* covalently bound to each other.[32] This discovery is the basis of the technological success in the modern world, and essential for modern everyday life. Jean-Marie Lehn further expanded upon this concept by introducing the idea of the "*chemistry beyond the molecule"*: monomers can also form a polymer via *non-covalent* interactions, called supramolecular chemistry as stated in his Nobel Lecture.[33] Just like in classical covalent polymers, supramolecular materials are composed of small molecules as repeating units, that autonomously organize, *self-assemble*, to higher ordered structures that are not connected through covalent bonds.[29] The self-assembly process is driven by molecular forces, such as van der Waals, hydrogen bonding, hydrophobic interactions, ionic interactions and π–π stacking (**Figure 2B**).[29] In its most basic concept, the self-assembly process can be reduced to its thermodynamic aspect, namely to the energy gain of the free enthalpy according to Gibbs-Helmholtz equation (eqn. 1).[34]

$$\Delta G = \Delta H - T\Delta S \qquad \text{(eqn. 1)}$$

ΔH and the ΔS describe the change in enthalpy and entropy, respectively, during a chemical reaction or molecular interaction. In general, for self-assembly to occur, entropy must decrease (ΔS > 0), whereas enthalpy (ΔH) can increase,[34] *e.g.* the self-assembly of the hepatitis B virus capsid protein is driven by ΔH being 2 kcal/mol, and ΔS*T being 17 kcal/mol at room temperature.[35] In contrast to covalent polymers, supramolecular materials are held together by a variety of interactions with different binding energies (**Figure 2C**).[36] This diversity is the origin of complexity on the molecular scale and the starting point for tunable, adaptable and transient

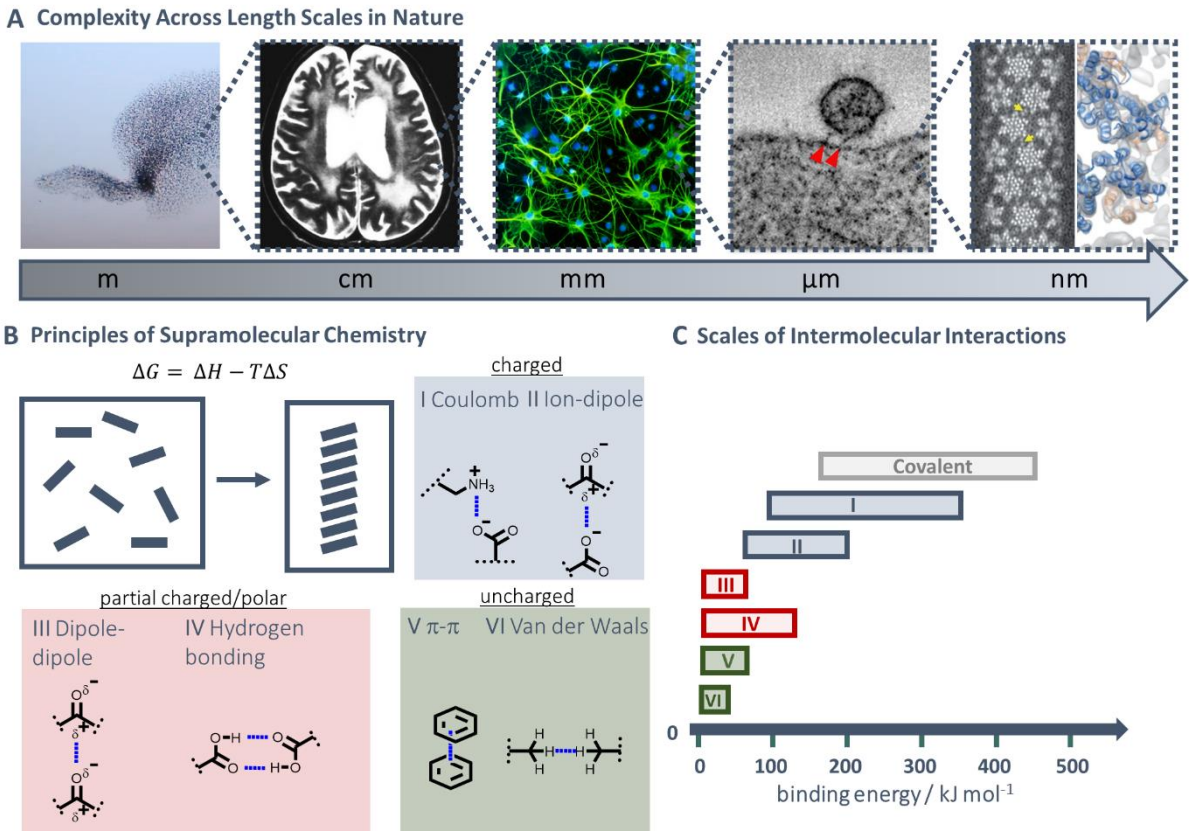


Figure 2 Self-assembly across multiple length scales. **A** Natural systems organize in a non-covalent way. Selected images serve as illustrative examples of self-assembly and are not directly connected to each other. Complex patterns occur across several length scales from bird flocking (m),[300] brain tissue (cm),[301] and neural cells (mm),[302] HIV-1 infection where viral entry is enabled by non-covalent supramolecular interactions (µm)[303] and finally the HIV-1 capsid that displays a regular tubular assembly of proteins (nm).[28] Images adapted from Refs 300, 301, 302 and 28 with permissions from 2022 AAAS, 1996 Elsevier, 2019 Springer Nature, and 2013 Springer Nature, respectively. Reference 303 is licensed under creative commons (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/). **B** Principles of self-assembly at the nanometer scale. The gain in free energy ΔG is the driving force for molecular organization and is driven by non-covalent interactions, mainly by (I) Coulomb, (II) Ion-dipole, (III) dipole-dipole, (IV) hydrogen bonding, (V) π-π and (VI) van der Waals interactions. **C** Graphical representation of the binding energy range of the interaction modes (I) – (VI) in comparison to covalent interaction mode.[36,304]

structure formation on the macroscopic level.[37] Meijer, Stupp, and Aida recognized the potential lying in supramolecular chemistry and introduced another dimension to supramolecular polymerization by transferring the concepts of covalent polymers to the field of non-covalent polymers. Some of the breakthrough discoveries in this field are the creation of functional soft materials with UPy as non-covalent crosslinkers in covalent polymers[38] for modular functional biological scaffolds (**Figure 3A-I**),[39] living supramolecular polymerization (**Figure 3A-II**),[40] and tissue engineering with peptide amphiphiles that yield nanofibers in water and form a gel that directs cell growth once injected in tissue (**Figure 3A-III**).[41]

While structural control across length scales is well studied for economically important synthetic polymers, there is a lack of knowledge how molecular features transfer to functional systems across length scales by self-assembly in supramolecular polymers. This knowledge gap becomes even more significant when we acknowledge that self-assembly is not a novel concept to Nature. Many processes in Nature are driven by transient self-assembly of biomaterials and supramolecular chemistry is a prerequisite to dynamically built and decompose structures. Proteins as one of the most fundamental polymers in biology are involved most prevalent in such self-assembly processes.[42] The cell cytoskeleton and extracellular matrix consist of various

protein assemblies that form intricate networks to provide cellular attachment (**Figure 3B-I**),[43] adaptive structures (**Figure 3B-II**),[44] and stable scaffolds (**Figure 3B-III**).[45] One prominent

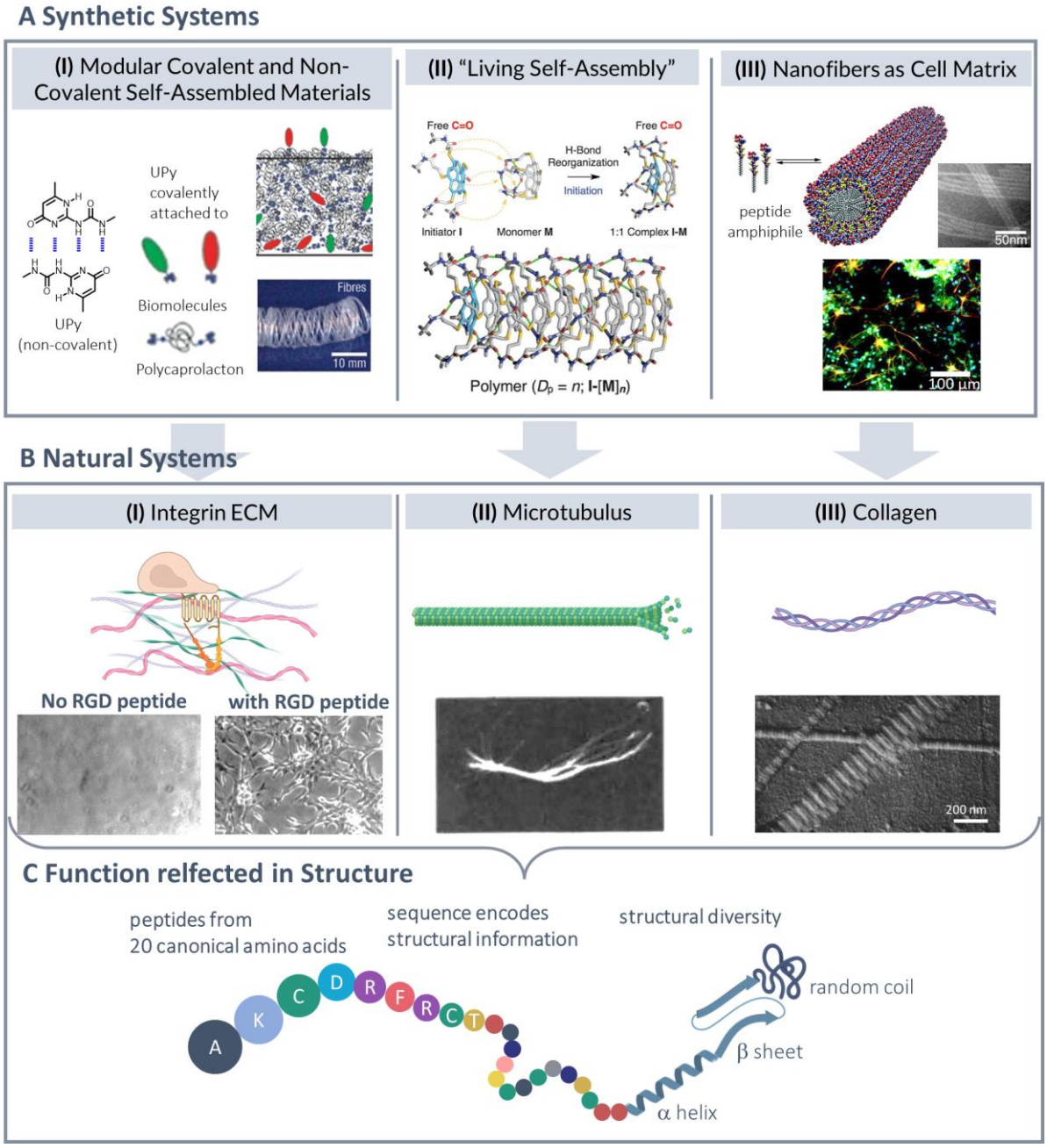


Figure 3 Selected examples highlighting supramolecular materials with function in synthetic and natural systems. **A** Synthetic systems providing biomimetic materials. **(I)** Modular bioscaffolds created by self-assembly of UPy functionalized with biomolecules to support cell-attachment. Reproduced from Ref 39 with permission from 2005 Springer Nature. **(II)** A supramolecular chain-growth polymerization that is enabled by H-bond reorganization. Reproduced from Ref 40 with permission from 2015 AAAS. **(III)** Hierarchical self-assembly of synthetic peptide amphiphiles act as fibrillar scaffolds for cell growth. Reproduced from Ref 41 with permission from 2001 AAAS. **B** Examples showcasing selected roles of proteins in biology: **(I)** Widefield microscopy image of myoblast cells that do not adhere on alginate hydrogel without RGD-peptide motif. Reproduced from Ref 43 with permission from 2003 Elsevier. **(II)** Electron microscopy image of microtubules protein self-assembly that allow adaptive creation and decomposition of cellular structure. Reproduced from Ref 44 with permission from 1984 Springer Nature. **(III)** Electron microscopy image of collagen type I. Reproduced from Ref 45 with permission from 1993 Elsevier. **C** Function and structure of peptides are encoded in the amino acids composition and order. Peptides can adapt various secondary structure with different energy landscapes due to the diversity in supramolecular interaction modes from the amino acids. Schemes in B and C created with permission from BioRender.com.

example are microtubules as essential structure forming polymeric parts of cells, they dynamically grow and shorten depending on the chemical environment by self-assembly of proteins.[46] A protein sequence can adapt multiple morphological appearances with only certain secondary structure conformations that enable correct function. The reasons for polymorphism are controversially discussed[47] and one possible origin is believed to be the chemical diversity of amino acid interactions that allow a broad variation of energetic landscapes (**Figure 3C**).[48] Hence, amino acids are Nature's preferred choice for encoding structural diversity which enables diversity in function[47] and peptides are therefore an ideal biomaterial for constructing complexity in synthetic materials.[49]

Peptide Self Assembly

Supramolecular Interactions encoded in Amino Acids

In contrast to other synthetic supramolecular materials, self-assembly of peptides is usually taking place in water, which is particularly important for creating biocompatible systems for medical applications.[50] Peptide self-assembly is governed by specific interactions of the building blocks, which are characterized by the polarity and charge of the side chain groups in each amino acids (**Figure 4A**).

There are twenty canonical genetically-encoded L-amino acids, which can be represented with the single letter code. Each amino acid can be assigned to a characteristic charge and polarity, depending on the chemical properties of the side chain at the solution pH (**Figure 4A**).[51] Some of the nonpolar amino acids are further characterized by their aromatic groups such as phenylalanine (F), tryptophan (W), and tyrosine (Y) and some by the small size of the side chains

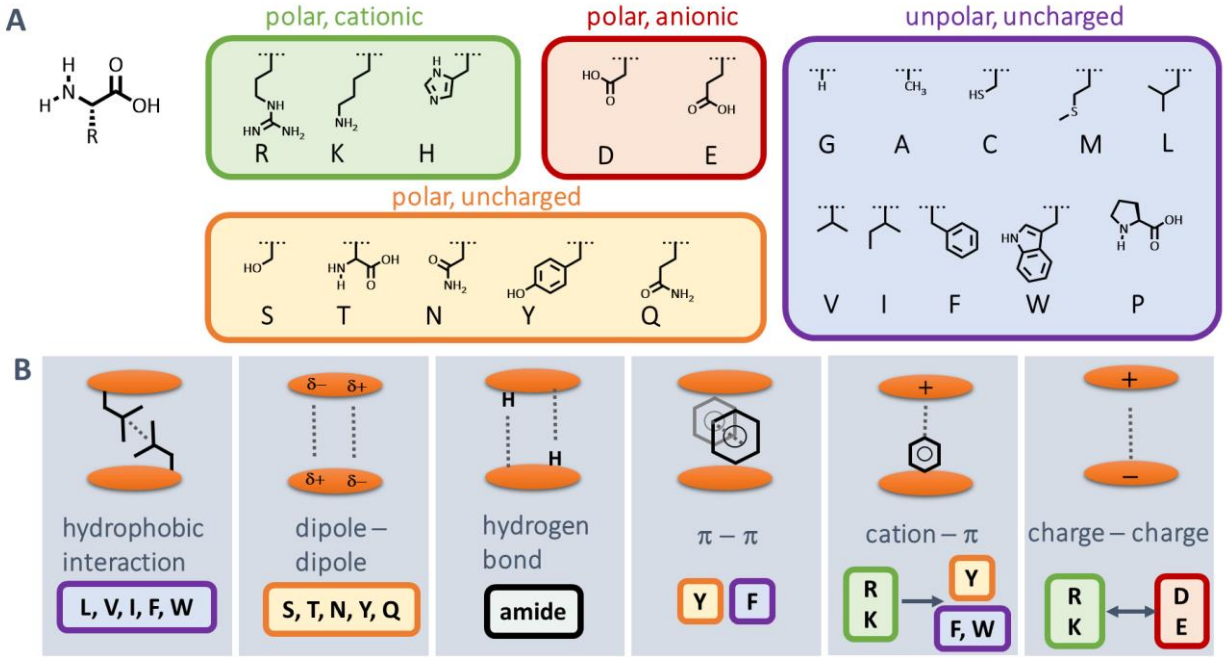
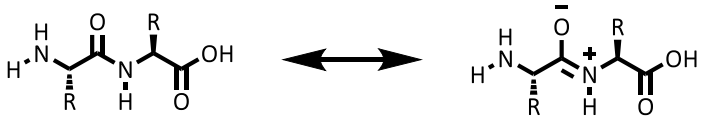


Figure 4 Overview of the 20 canonical amino acids and their interaction modes. **A** The canonical amino acids consist of an amine and a carboxyl group and a side chain residue, indicated by "*R*". The side chain residues "*R*" can be categorized according to their charge and polarity and are shown here according to Kyte-Doolittle classification.[305] **B** Supramolecular interaction modes between amino acids and selected amino acids mainly contributing to the corresponding interaction. The amide backbone is the main contributor to hydrogen bond interactions.

with a van der Waals volume smaller than 100 Å^3 such in alanine (A), glycine (G), and serine (S).[52] Depending on the chemical nature of the side chain, each amino acid can contribute with different preferred supramolecular interaction modes (**Figure 4B**).
A covalent amide bond that is formed between at least two amino acids' conjugation of the carboxylic acid and the amine group results in peptides. The partial double bond character of the C–N bond creates a rotational barrier as well as partial charges on the oxygen and nitrogen. In turn, this mesomeric effect is the origin of the peptide backbone's stiffness that enables intermolecular hydrogen bonding and dipolar interactions (**Scheme 1**).



Scheme 1 Mesomery of a peptide backbone causes a negative charge at the oxygen atom and a positive charge at the nitrogen atom and introduces a partial double bond character and thus a rotational barrier.

The dipole-dipole interactions arising from the polar peptide bond in the backbone are a common driving force for intermolecular association. Dipole-dipole interactions of the side chains can provide additional points for interaction, *e.g.* the amide group in glutamine (**Figure 4B**). Other side chain groups such as amines as in lysine or carboxylic acids as in glutamic acid, can contribute with electrostatic Coulomb interactions, either in an attractive or repulsive way. Because peptides are typically dissolved in water, the Coulomb interaction strength of charged groups (**Figure 2C**) is weakened by the strong dipole-charge interactions of peptides with water.[53] Therefore, the Coulomb interactions and eventually the charge of the molecule is highly dependent on the local environment of the charged group within the peptide sequence and assembly. For example, charged groups that are buried in a hydrophobic environment are less accessible for water and can interact more strongly with other charged sites, whereas in the presence of ions, binding to the charged groups is screened and they therefore interact less with each other. In general, these charge screening effects can hinder or facilitate the self-assembly process, depending on the charge and local environment in the peptide.[54] Interactions of water with polar functional groups are strongly dominated by hydrogen bonding interactions, which in turn decreases the intermolecular hydrogen bonding strength that can be in this way similarly influenced as shown for Coulomb interactions.[50] Uncharged amino acids with aliphatic and aromatic side chains are responsible for hydrophobic interactions, that are described by van der Waals and π-π stacking interactions, respectively (**Figure 4B**). The origin of hydrophobic interactions is the higher binding energy (enthalpy) of molecules with each other than with water. This in turn favors the release of water molecules into the solute and increases the entropy of a system. Thereby hydrophobic effects are the result of entropy or enthalpy gain of aggregated molecules that require less water molecules for solvation than the separate molecules.[55] Although the origins of π–π interactions are discussed controversially,[56,57] its impact on peptide self-assembly in biological and synthetic systems cannot be denied. It is suggested that the local electron dispersion[56] and the conformational prearrangement of the geometrically flat π-systems,[57] in aromatic side chains such as phenylalanine improve and stabilize the assembly by π-stacking.[58,59] Overall, hydrophobic amino acids usually interact strongly with each other in water and their interaction is rising at elevated temperatures due to the gain in entropy upon water molecule release,[60] which enables to create peptide assemblies with a desired robustness as stable supramolecular polymers. Because hydrophobic interaction contributions grow with the interfacial size of a molecule,[61] and act on range long distances,[62]

this mode is of particular interest for supramolecular peptide assembly, yet challenging to explain in multi-scale hierarchical assemblies.[50]

Secondary Structure arising from the Amino Acids Order

It has been agreed upon that the function of a protein is intimately linked to its three-dimensional structure, which is achieved through the folding of secondary and tertiary structures.[63] A secondary structure describes a regular local substructure on a peptide backbone chain and is the basis of hierarchical supramolecular structure formation. One of the most common examples in peptide self-assembly is based on β-sheet structures (**Figure 5A**). The ability to form a certain secondary structure is determined by the amino acid composition and order.[64] For example, the aromatic amino acids F, W, Y, I, and V are known to preferentially adopt β-sheet conformation.[65] Small, uncharged amino acids such as glycine or alanine in β-sheet structures are known to demand less steric limitations and introduce flexibility and thus their amount can balance the dynamics of the self-assembled structure. In contrast I, L, and V have larger aliphatic hydrophobic side chains, which introduce a higher rigidity.[65]
The contributions of individual amino acids in a sequence add up to influence the final supramolecular structure. In biological contexts, proteins with regions of amino acids that favor β-sheet formation will dominate the overall folding.[58] For example, in their seminal study with the prominent example of Amyloid Aβ42, Gazit and co-workers showed that the regions prone to β-sheet formation can be pinpointed and reduced to a minimum sequence, the diphenylalanine FF that can form β-sheet rich amyloid type sequences (**Figure 5B**).[58] However, predicting the secondary structure of a given peptide sequence is not straightforward, as

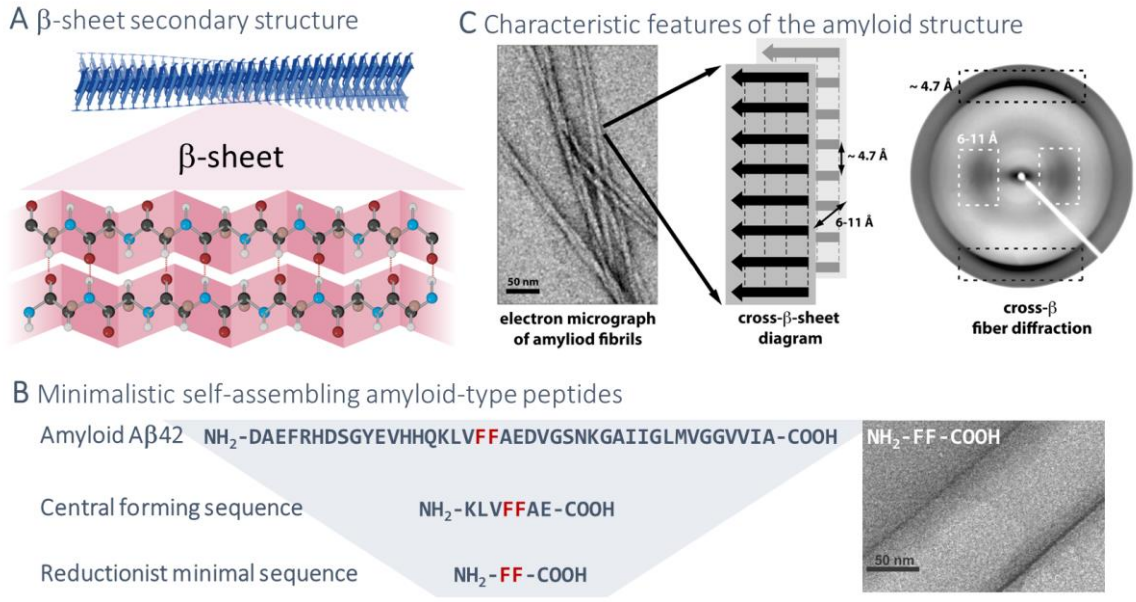


Figure 5 Secondary structure of peptides illustrated at the example of the amyloid structure. **A** Cartoon representation of a peptide fibril and its β-sheet intermolecular order. Created with permission from BioRender.com. **B** Sequence reduction of amyloid Aβ42 to its minimal β-sheet forming sequence, the diphenylalanine. Electron microscopy image shows fibrillar morphology of the diphenylalanine sequence. Scale bar 50 nm. Reproduced from Ref 58 with permission from 2003 AAAS. **C** The intermolecular order of the amyloid structure is characterized by fibrillar morphology (electron micrograph, scale bar 50 nm) composed of two β sheets that are arranged perpendicular to each other with a typical meridional distance of 4.7 Å and equatorial distance of 6-11 Å, as observed in the diffraction pattern for the cross-β sheet. Reproduced from Ref 69 with permission from 2010 Elsevier.

different ordering of the same amino acid sequences produces different structures when self-assembled in solution.[66,67]

Structural Diversity of the Amyloid Structure

While the diversity of supramolecular interaction modes of amino acids makes it challenging to *a priori* predict secondary structures, it makes peptides highly interesting candidates for the creation of complex nanostructures via a reductionist approach.[29] Many research groups have focused on short peptides that form β-sheet-rich structures to tackle this challenge.[68]
Perhaps, the most prominent example for β-sheet forming peptides is the amyloid structure that is characterized by high aspect ratio fibrillar structures with a characteristic meridional interspacing of 4.7 Å and an equatorial interfibrillar distance of approximately 6–11 Å (**Figure 5C**).[69] Amyloid fibrils are known for their extraordinary stability and mechanical stiffness in the range of 2–3 GPa[70] which is attributed to the cross β-sheet-rich structures, that are mainly stabilized by hydrophobic interactions and hydrogen bonds.[71] Short self-assembling peptides that form amyloid structures often arrange in the so-called sterical zipper configuration,[72] which increases their stability[73] and is in many cases the structural basis for their bioactivity.[74] In general, amyloid structures are associated with misfolded proteins that cause neurodegenerative diseases[75–77] but are also known to serve various purposes in organisms such as in bacteria biofilms,[78,79] as scaffolds for hormone storage[80] and melanin polymerization.[81]
While β-sheets can be observed in biology most common in amyloids, it has to be noted, that it is not limited to this, and β-sheet secondary order can for example also be observed in synthetic peptide-hybrid systems, such as peptide amphiphiles (PAs). PAs consist of a polar peptide and a long aliphatic tail.[41,82,83] The hydrophobic moiety, most commonly a fatty acid, triggers phase separation and assembly into various morphologies. The secondary structure of these systems is controlled by the specific peptide sequence and the solubility in water is governed by the terminal hydrophilic amino acids. For example, nanotubular PA assemblies have been pioneered by Stupp and coworkers and have been extensively investigated in terms of their physicochemical properties and for various cell-applications (**Figure 3A-III**).[84–91] Just like amyloids, a peptide amphiphile can show different morphologies when assembled (polymorphism) due to the different assembly pathways and kinetics.[92,93]
It has been generally accepted that the amyloid formation is governed by a typical nucleation–elongation model, in which monomeric soluble peptides first aggregate into β-sheets consisting of oligomeric species, which subsequently grow into fibrillar structures with long-range order (**Figure 6A**).[94] Differences in the assembly kinetics lead to lag-times, which are of particular interest for the investigation of protein misfolding in amyloid structures. In fact, naturally occurring amyloid structures are typically divided into pathological and functional contexts[95,96] and it has been hypothesized that they differ from each other in their assembly kinetics.[97–100] Small sized, intermediate oligomeric species rather than mature, thermodynamically stable amyloid fibrils have been associated with cytotoxicity[101–103] along with peptides that self-assemble slowly and prevail for a long time (~ days). During this time, these oligomers can also act as seeds for further protein nucleation.[97,98] In contrast, peptides that assemble fast (~ minutes) and show no cytotoxicity are considered “functional” amyloids.[99,100] Atomic force microscopy measurements revealed that the polymorphic structures in mature amyloid fibers display different degrees of stiffness,[104] which was in turn related to the internal β-sheet

content.[105] Depending on the pathway of structure formation, a peptide can assemble into fibrillar, crystalline or β-sheet rich, amorphous aggregates that lack long-range fibrillar or crystalline order.[106–109] For example, if a peptide assembles into amyloid fibers it has less β-sheet content and a smaller Young's modulus compared to amyloid crystals of the same peptide.[110] This difference in stiffness can be traced back to the cooperative effects arising in structures of higher symmetry.[110]

There are hypotheses that a higher prevalence of crystalline (highly ordered) fibril aggregates in pathological amyloids disrupt cellular membranes whereas unordered fibrillar aggregates found in functional amyloids have less exposed terminal ends and are thus less toxic.[111]

Several decades of extensive research on the origins of pathological and functional amyloids,[112] recognized peptide hydrophobicity, electric charges, and the presence of anions in the buffer solution as important factors for amyloid formation.[113–116] However, comparably less research has focused on understanding the fibril-fibril interactions that determine their aggregation into larger particles and accumulation. Although no conclusive evidence has been reported yet, it is suggested that certain fibril surface properties drives lateral aggregation. These properties include hydrophobic interactions,[117–119] ion-salt bridges,[114] and polar interactions.[120,121] It is hypothesized that the absence of these interactions is connected to the dispersion of fibrils, while their presence may contribute to the aggregation of fibrils into larger structures (**Figure 6B**).[120,122–124] Fibrils with a hydrophobic surface tend to expel water, making them less solvated and more prone to interact with other fibrils.[122,123] The hydrophobic effect is for example also observed in protein misfolding, where aggregation is triggered at interfaces like lipid droplets in the cytosol[125] or macroscopic hydrophobic surfaces.[126] On the other hand, fibrils displaying polar amino acids on their surface can leverage polar interactions. For example, hydrogen bonding between polar amino acids from different fibrils promotes lateral aggregation stabilizing this superassembly.[120,124] However, in cases where hydrophilic interactions with water dominate, *e.g.* when there are many charged groups on the fibril surface, the structure becomes well-hydrated, hindering fibril-fibril interactions.[120] This solvation by water can be overcome by the presence of multivalent ions, such as phosphate ions from a buffer, which induce ion-bridging interactions. These interactions counteract charge repulsion and water interactions with cationic charged groups on the fibril surface, thus triggering fibril-fibril aggregation.[114]

To date, studies investigating amyloid aggregation have primarily focused on the nm-scale regimes, *i.e.* fibril formation[127–129] or µm-scale fibril alignment.[130–139] However, there is a lack of understanding of the formation of µm-sized aggregates comprised of unordered fibrils. Despite their importance in biological activity, the specific factors determining their (dis)order remain unknown to date and are subject to further research.

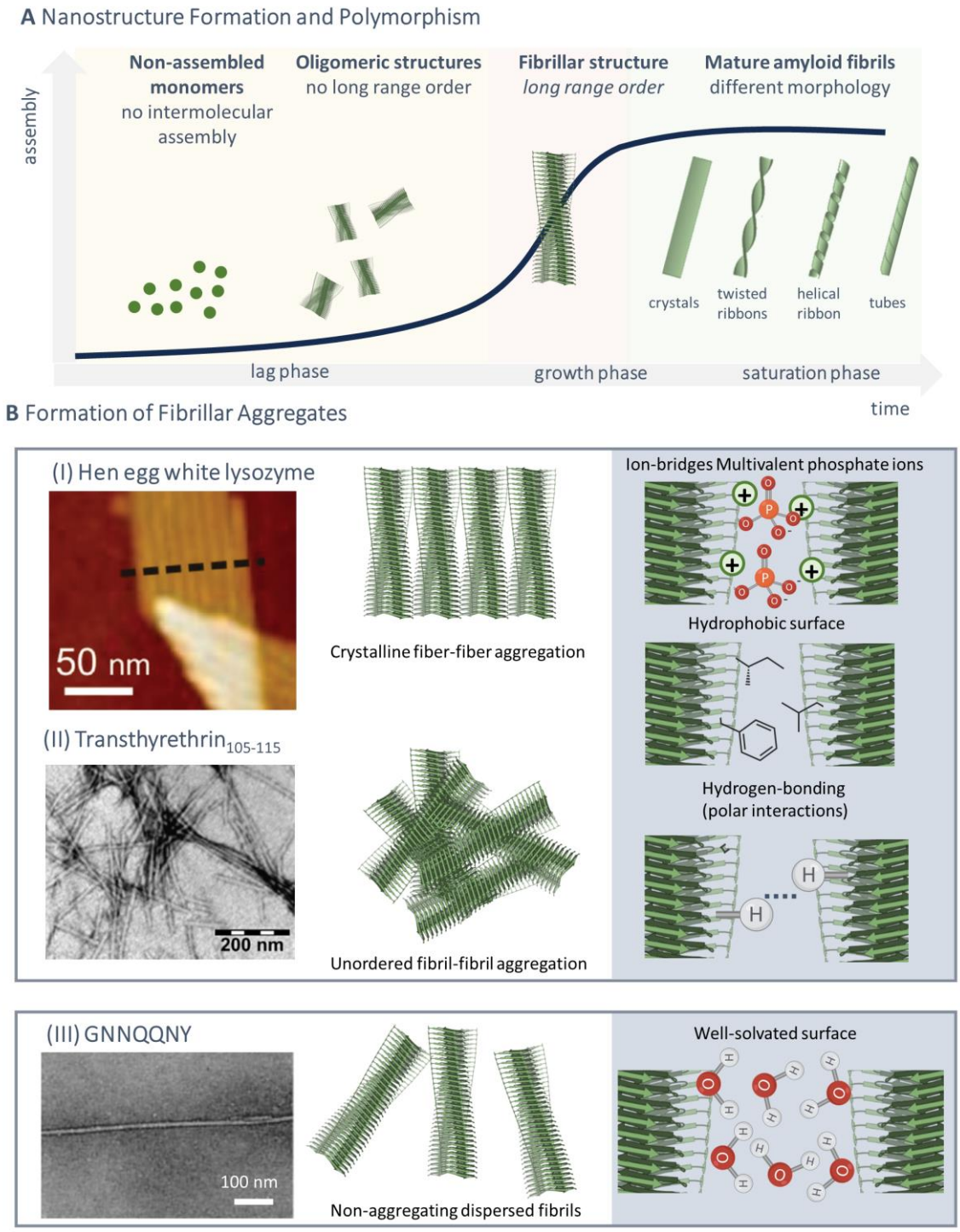


Figure 6 Overview of a simplified amyloid fibril formation pathway and schematic representation of current hypotheses on lateral fibril-fibril interaction modes. **A** self-assembly of peptide monomers into the amyloid structure proceeds via a nucleation and growth mechanism in three main phases. In the initial lag phase monomers assemble into β-sheet oligomers but they do not have long-range order. Upon addition of further monomers to the oligomers the structures evolve into fibrils (growth phase) and yield the mature amyloid structure that can be observed in various polymorphic structures in the saturation phase. Reproduced from Ref 48 with permission from 2018 Wiley and created with permission from BioRender.com. **B** Microscopy measurements and schematic visualization of selected examples for amyloid fibril-fibril aggregation types: Atomic force microscopy of crystalline aligned fiber-fiber aggregation (top), electron microscopy of unordered fibril-fibril aggregation (middle) and electron microscopy of dispersed, non-aggregating fibrils (bottom). Reproduced from Ref 137, 306, 307 with permission from 2011 ACS, 2008 Elsevier, and 2001 National Academy of Sciences. (I) The hen egg white lysozyme forms highly aligned fibril-fibril aggregates (atomic force microscopy scale bar 50 nm)[137], (II) whereas a transthyrethrin fragment forms unordered fibril-fibril aggregates (electron microscopy, scale bar 200 nm).[306] Fibril-fibril interaction is typically governed by ion-bridging, hydrophobic, or hydrogen-bonding interactions. (III) The amyloidogenic GNNQQNY peptide is mostly observed as isolated and non-aggregated fibrils (electron microscopy, scale bar 100 nm). This is believed to be due to a well-solvated surface that hinders attractive fibril-fibril interactions.

Peptide Nanostructures as Materials for Cell-Applications

Over millions of years of evolution, Nature has evolved peptide nanostructures as materials with inherent bioactivity and compatibility. Self-assembling peptides possess the remarkable ability to form both stable and dynamic structures in aqueous environments and as a result, they hold great potential to fulfill unmatched needs as supramolecular biomaterials for therapeutic applications.[140] For example, in the treatment of rare skin diseases, recent therapeutic strategies combine tissue engineering techniques with gene delivery using conventional static polymeric materials as matrices.[11] However, the outcomes of such applications are expected to be much more impactful when supramolecular materials are used instead because they can adapt to the tissue and are specifically tailored for efficient gene delivery and optimal cell growth.[13]

Inspired by naturally occurring amyloid proteins that can form complex multi-scale materials with cell adhesive properties, researchers designed synthetic amyloid-type peptides. These synthetic peptides have found applications as scaffolds for 3D cell culture in tissue and neuronal regeneration, as well as drug delivery, including viral gene delivery.[49,141,142] Synthetic amyloid-type peptides are not related to pathological amyloids with regards to their sequence and can be chemically tailored to tune their stability and bioactivity.[77,143,144] The main advantage and motivation to use these peptides as biomaterials for cell-applications is their bioactivity, their precise and automated synthesis, the programmability of their sequence that ensures facile preparation of hierarchical yet dynamic supramolecular materials, which in turn are typically biocompatible. Sequence design for example allows to create amyloid fibrils with different cell-adhesiveness.[145] For example, amyloid-like nanofibrils with a positive net charge and a high content of β-sheet structures can support neuronal outgrowth when injected into facial nerves of living mice.[145]

Interactions of Fibrils with Cells and Virions

Amyloid type peptides cannot only efficiently interact with cells but also with lenti- and retroviruses. This mediator role enhances the infectivity of these viral vectors, which was first discovered for amyloid Aβ.[146] Subsequently, similar effects have been observed in a different context: amyloid structures present in seminal fluids boosted human immunodeficiency virus 1 (HIV-1) infection.[147] Specifically, these amyloid structures are formed by the prostatic acidic phosphatase (PAP) regions $PAP_{248-286}$ and PAP_{85-120}, as well as semen derived enhancer of virus infection (SEVI).[148,149] These findings quickly attracted interest in exploring amyloids that boost viral infectivity beyond their initial pathogenic context[149–153] and apply them as adjuvants in therapeutic applications, *e.g* by enhancing the retroviral gene transfer.[154]

In fact, functional amyloids are promising materials for supporting retroviral and lentiviral gene therapy due to their inherent biocompatibility and low cytotoxicity.[67,155,156] The first example for potential clinical applications in this regard was the 12-mer peptide enhancing factor C (EF-C) derived from the residues 417–428 of the HIV-1 envelope glycoprotein 120 (gp120).[156] This peptide, QCKIKQIINMWQ, commercially known as Protransduzin, forms positively charged amyloid fibrils and significantly enhances virus-cell interactions, outcompeting other transduction enhancers available in the market (**Figure 7A**).[156] The biocompatibility and facile

preparation of the amyloid nanostructures make them appealing for various therapeutic applications, for example in addition to gene therapy[156] can be used as additives in vaccines.[157]

Of course, these developments also sparked curiosity about the mode of action and the specific prerequisites for efficient virus-cell interactions. Investigations by Münch and coworkers found several features that are important for efficient virus delivery. In the seminal study on infectivity enhancing amyloid fibrils it was found that non pre-aggregated Aβ fibrils are inactive,[146] which was confirmed in later studies with other peptides[158] and led to the conclusion that activity emerges only from the supramolecular self-assembly of the peptides into fibrillar structures. The presence of amyloidal fibrils that can adapt high-aspect ratio structures up to hundreds of nanometers might not only serve to entangle virus particles that have a diameter of around 80-100 nm, but also offer plenty interactions points for cellular protrusions which have been identified to actively transport these fibrillar structures towards the cells (**Figure 7B**).[159]

A common characteristic for the aforementioned transduction enhancing amyloid-type peptides is their high isoelectric point, which leads to a positive net charge of basic amino acid side chains under physiological conditions. Lysine as a positively charged amino acid can interact with the negatively charged membranes of cells and viruses which leads to colocalization of the two and increases the chances of infection.[156,160,161] In turn, no enhancement of viral infectivity is found for fibrils lacking positive charges due to the replacement of cationic charged amino acids with uncharged alanine,[161] as well as negatively charged fibrils.[162] Interestingly, these negatively charged fibrils can still interact with virus particles but fail to attach to cell membranes (**Figure 7C**).[162] This suggests that the positive charge is important for promoting virus-cell colocalization.

After transportation of positively charged fibrils to the cellular membrane Schütz, Münch and coworkers found that the fibrils enter the cells via macropinocytosis.[163] This was studied for the uptake of EF-C fibrils in TZM-bl cells, which were pretreated with various uptake inhibitors (**Figure 7D**).[163] Macropinocytosis is an endocytosis uptake mechanism, which cells employ to internalize a large amount of foreign material by encapsulation in endosomal vacuoles.[164] Correlative light electron microcopy measurements showed that EF-C peptide fibrils are taken up in these endosomal vacuoles and degraded in these compartments (**Figure 7E**).[163] Presumably, upon degradation of the fibril, the virus and subsequently its genetic payload are released in the nucleus, which initiates the transduction process.[165]

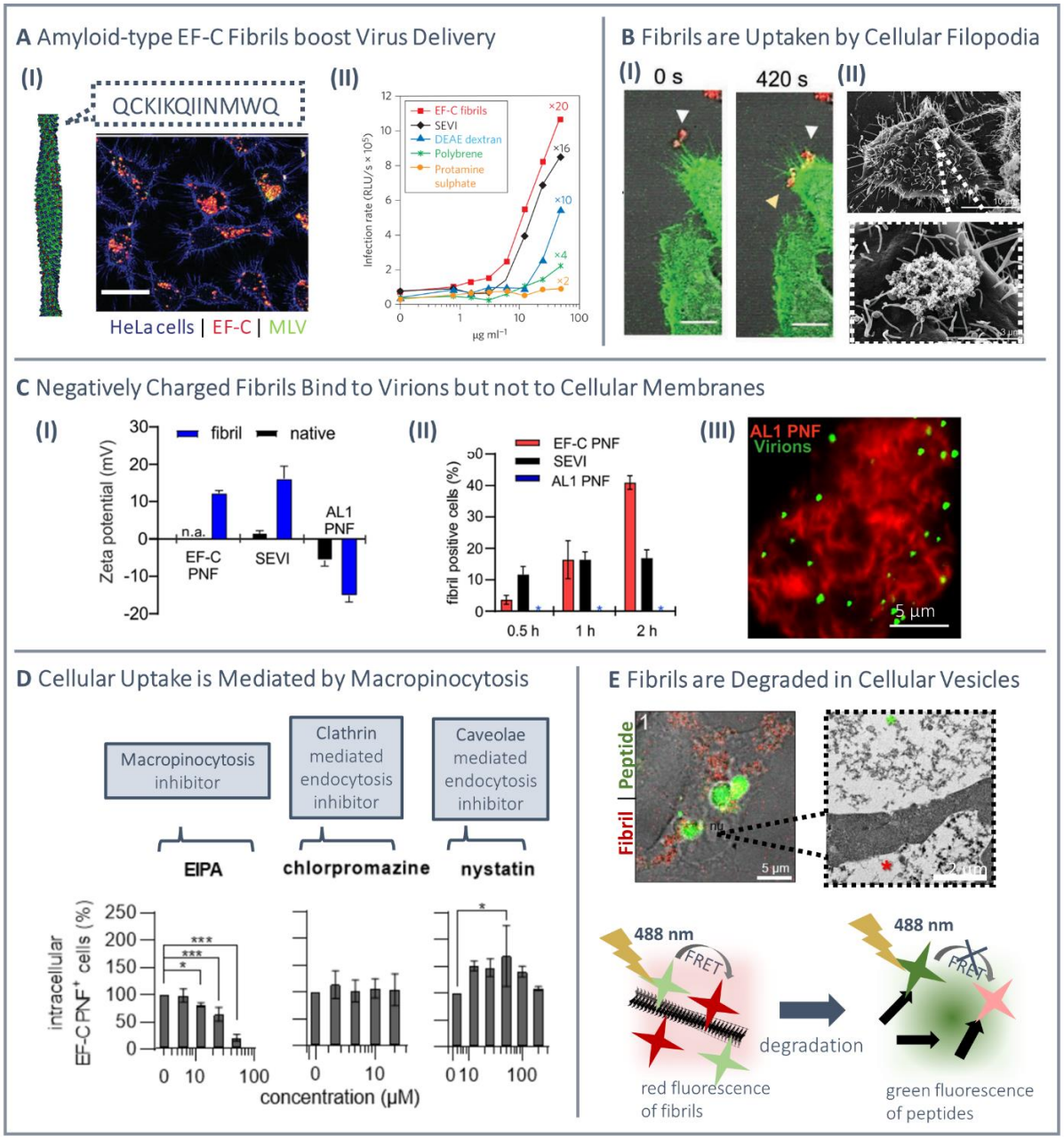


Figure 7 Summary of important findings on amyloid fibrils that promote viral infectivity. **A (I)** Schematic visualization of fibrillar 12-mer self-assembling peptide EF-C (QCKIKQIINMWQ) and merged confocal microscopy measurements of the interaction between rhodamine labelled EF-C fibrils (red), HeLa cell membrane (blue) and yellow-fluorescent protein labelled murine leukemia virus (MLV, green) scale bar 15 µm.[154] **(II)** Comparison of viral infectivity enhancement efficiency of EF-C with traditional polymeric enhancers. Reproduced from Ref 156 with permission from 2013 Springer Nature. **B (I)** Time lapse fluorescence microscopy of HeLa cellular filopodia (green) actively pulling EF-C fibrils (red).[159] **(II)** Electron microscopy of EF-C fibrils entangled with HeLa cell filopodia, top scale bar 10 µm, bottom scale bar 3 µm. Reproduced from Ref 159 with permission from 2021 Wiley. **C** In contrast to EF-C and SEVI, the peptide AL1 have a **(I)** negative Zeta-potential and **(II)** does not attach to cells as evaluated by flow cytometry, **(III)** yet AL1 fibrils (red) can capture virions (green) as observed via confocal microscopy, scale bar 5 µm. Reproduced from Ref 308 with permission from 2021 ACS. **D** Flow cytometry measurements quantifying uptake of dye-labelled EF-C in TZM-bl cells in presence of 5-(*N*-Ethyl-*N*-isopropyl)-Amiloride (EIPA), chlorpromazine, and nystatin that are macropinocytosis, clathrin and caveolae mediated endocytosis inhibitors. Reproduced from Ref 163. **E** Correlative light electron microscopy (CLEM) of intracellular EF-C fibrils that are labelled with the Förster resonant energy transfer (FRET) fluorophore pairs ATTO488 (donor, green) and ATTO647 (acceptor, red). In fibrillar state both dyes are in proximity which enables FRET and red fluorescence is observed upon excitation at 488 nm. However, upon degradation of the fibril FRET cannot take place and the peptides appear green at 488 nm excitation. Therefore, CLEM shows that the fibrils are degraded in cellular vesicles upon internalization of TZM-bl cells. Reproduced from Ref 163.

Based on these studies, the proposed mode of action (**Figure 8**) for infectivity enhancing amyloid type fibrils suggests that a positive net charge and fibril formation is important for binding to negatively charged viral particles. The amyloid-virus complex is subsequently actively transported to the cellular membrane by filopodia[159] and is taken up via macropinocytosis.[163] The fibrils are then degraded in cellular vesicles, which likely releases the virus payload so that viral gene delivery takes place. In this way amyloid-type peptide fibrils facilitate viral infection by overcoming the electrostatic repulsion between virions and cells, which is the rate-limiting step.[166]

Overall, certain features such as a positive charge of amyloid-type fibrils have been identified for the literature known infectivity enhancing peptides EF-C,[154] SEVI,[148] and Aβ.[146] These pioneering studies give rise to questions about the generalizability of these observations to a broader range of different peptides. Recent findings imply that certain peptide characteristics other than charge may be important for efficient cell interactions. For example, hydrophobic interactions have been identified as a prerequisite for efficient gene delivery for polyplexes and polymers in non–viral applications[167,168] and binding of peptides to cells.[169] Moreover, multivalent binding of the amyloid fibrils SEVI and Aβ to the cell surface have been identified to play an important role,[170] yet have not been studied systematically.

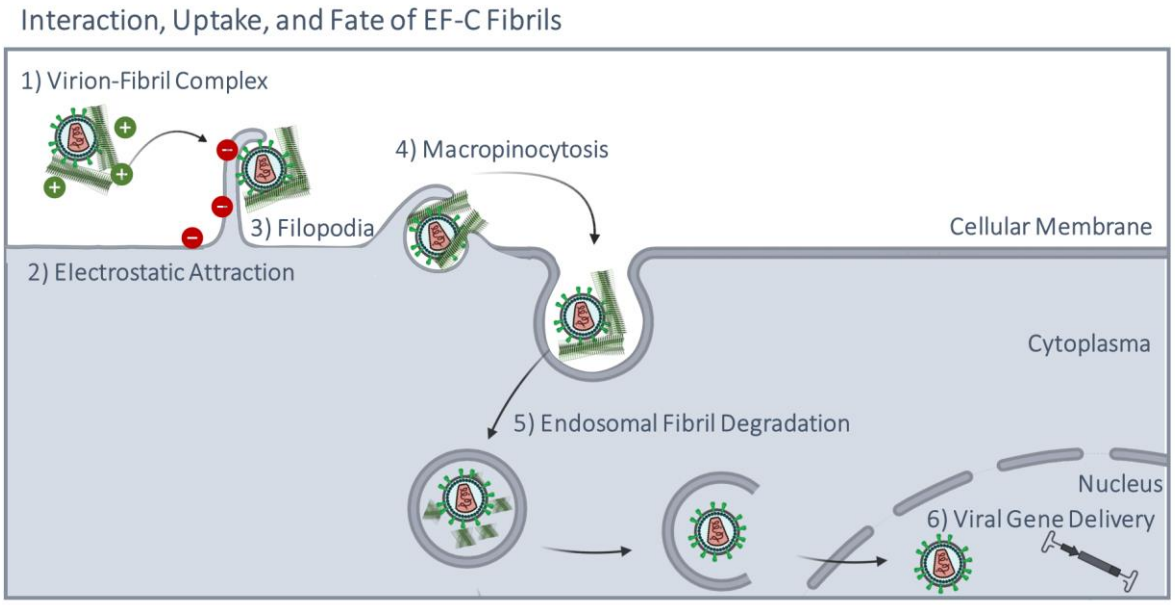


Figure 8 Graphical summary of reported findings on the interaction, uptake, and fate of EF-C fibrils that deliver HIV-1 to HeLa cells. **1)** First amyloid fibrils capture virions from solution and form a complex. **2)** The positively charged fibrils are attracted by the negatively charged cell-membrane[162] and **3)** actively recruited by cellular filopodia to the cell surface.[159] **4)** The uptake is mediated via macropinocytosis[163] and **5)** fibrils are degraded within cellular vesicles.[163] **6)** Finally, the HIV-1 delivers its gene to the nucleus.[156] Schematic visualization is not true to scale. Created with permission from BioRender.com.

Structure Controls Cell-Interactions

Cellular signaling is performed in a dynamic environment, in which both the cell influences the environment and the environmental dynamics affects the cell.[171] Therefore, natural matrices such as the ECM are constantly reshaped by various stimuli.[172] The ECM is indeed an excellent example of how the reciprocity of fibrillar morphologies has crucial functions in living cells by facilitating material—cell interactions.[173–175] These interactions can be modeled by short self-assembling peptides as demonstrated by the hierarchical, complex structures from the assembly of simple peptide building blocks that form amyloid-type structures[176] that can for example direct cell alignment.[177] A common strategy applied for cell-supportive materials is to integrate bioactive peptide molecules that can act as cell-adhesive cues.[39] For example, the bioactive epitope RGD derived from fibronectin has been attached to supramolecular peptide fibrils to instruct cellular adhesion that was used to create a areas of increased cell-density (**Figure 9A-I**).[178] Another common strategy is to change the stiffness of the substrate, *e.g.* a polymeric gel to tune the cellular focal adhesion and spreading (**Figure 9A-II**) and thereby change cell morphology.[179]

In recent years, much emphasis has been put on triggering self-assembly of peptides into fibrillar structures to redirect intracellular biological pathways.[180,181] However, for translational applications not only the presence but also their controlled degradation of biomaterials to prevent toxicity or off-target effects. Therapeutic peptide fibrils should provide degradability either via an external trigger that controls their structural integrity, or by enabling metabolization by cells.[182]
Environmental conditions, such as changes in pH, peptide concentration, and solvent are simple handles to induce fibril decomposition, because the non-covalent supramolecular interactions are highly sensitive to these changes.[183] Another way is to disrupt the supramolecular interactions by adding chaotropic additives.[184] However, manipulating the local environment in the presence of cells is often challenging for these conditions and therefore, external triggers have been developed to fragment peptides sequences or change steric conformation in a controlled way. Examples of such external triggers include photo-sensitive reactive groups like azobenzene, hydrazone, and *o*-nitrobenzyl ester groups (**Figure 9B-I**).[173,185–188] In general, these photo responsive manipulation strategies are well-established in tissue-engineering to achieve control over cellular behavior, *e.g.* by local softening of hydrogels (**Figure 9B-II**).[189–194] Perhaps one of the most desired achievements in tissue engineering is the creation of controlled cell-patterns.[195] Techniques for cell-patterning allow to stimulate cell migration and communication locally.[196–198] Yet, there are only a few reports on how photo-responsive strategies can be applied in peptide fibrils to modulate cellular responses,[173,178,188,199] and photo-responsive peptides have not been applied yet for cell-patterning. This lack of relevant studies can most likely be traced back to the high intrinsic stability of amyloid-type peptide fibrils, which also limits efficient *in vitro* biodegradation.[200]

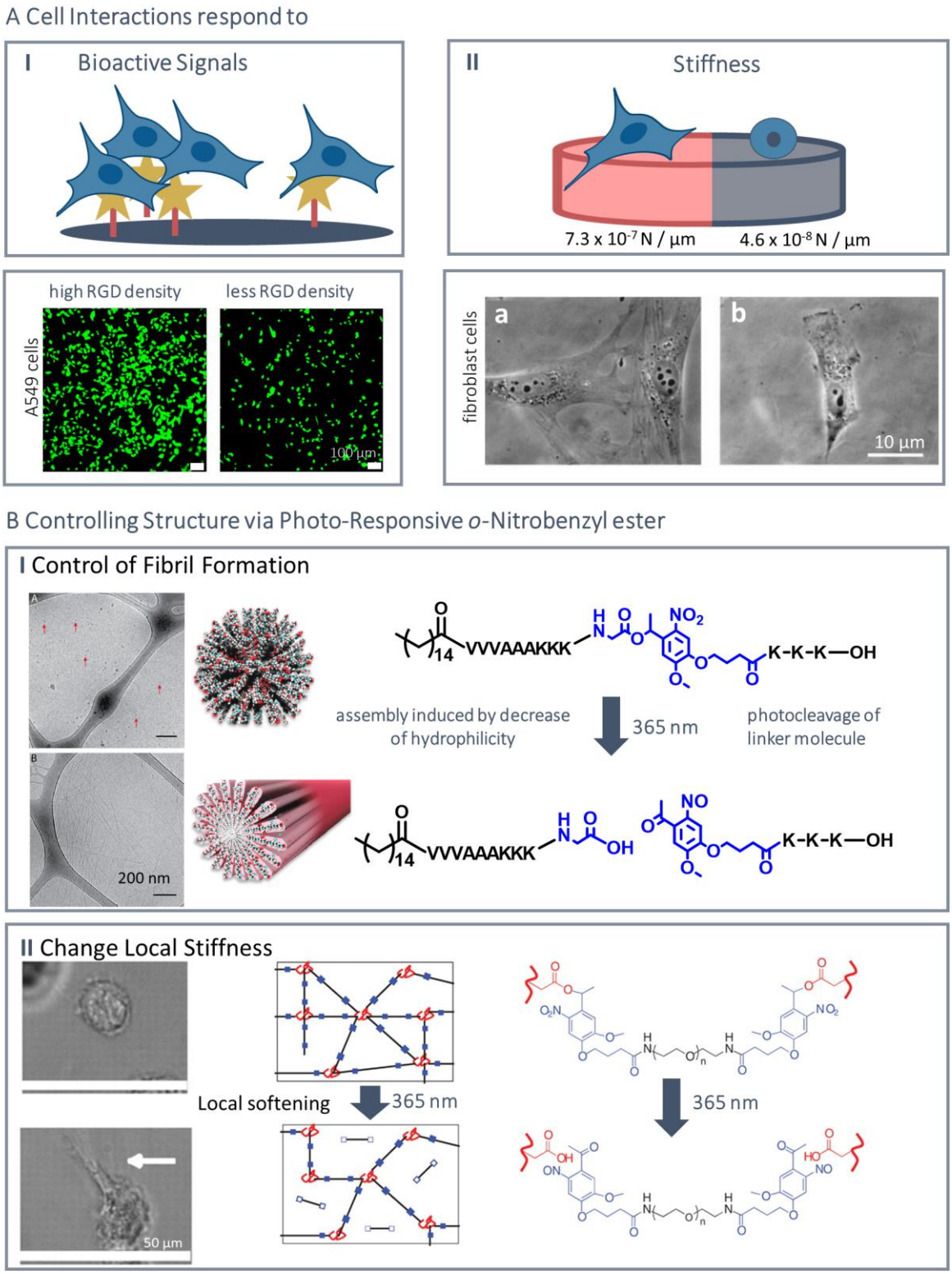


Figure 9 Overview on selected methods to control and guide cell attachment and spreading. **A-I** Cell attachment can be controlled with bioactive peptide epitopes such as RGD. A549 cells grow denser at substrates with high RGD density. Reproduced from Ref 178 with permission from 2021 ACS. **A-II** Cell morphology can be influenced by substrate stiffness. Fibroblasts adapt a spread morphology at stiffer matrices compared to softer substrates. Reproduced from Ref 179 with permission from 1997 National Academy of Sciences. Created with permission from BioRender.com. **B** Examples showcasing how the photo cleavable reactive *o*-nitrobenzyl ester group can be utilized to change local material properties. **B-I** Photo cleavage of C-terminal charged amino acids induces fibril formation as observed via electron microscopy, scale bar 200 nm. Reproduced from Ref 199 with permission from 2015 ACS. **B-II** Hydrogels crosslinked with *o*-nitrobenzyl esters be softened upon photo cleavage by irradiation at 365 nm. Upon softening cells that are embedded in the gel change their morphology and spread. Reproduced from Ref 193 with permission from 2009 Science.

An intriguing perspective on degradability is presented by other types of peptide fibrils such as peptide amphiphiles.[87] The morphology of peptide amphiphile-based nanostructures can be destabilized when there is an imbalance between the assembly-driving hydrophobic moiety and the solubility-providing charged peptide component (**Figure 9B-I**).[199] Most commonly, endogenous fatty acids such as palmitic acids are used as hydrophobic moieties, which renders the degradation products biocompatible.[201–203] PAs are taken up via lipid-raft mediated endocytosis without affecting cell-viability,[204–206] which makes them attractive materials for future *ex vivo* gene delivery applications.

Overall, the interactions of fibrils with cells are governed by an interplay of various parameters at multiple length scales. The influence of peptide sequences on fibril properties and the fibril properties' influence on cell interactions creates a complex landscape that offers great opportunity for rational material design. Particularly interesting is the exploration of the significance of certain peptide properties in viral gene delivery and finding ways to tune the degradability of fibrillar structures. Synthetic short self-assembling peptides are particularly suitable for investigating these questions, as they provide an experimentally and theoretically accessible way to study the effects of individual amino acid changes on fibril formation.[207,208]

Design and Discovery of Functional Peptide Nanostructures

New functional peptide materials are often discovered by serendipity. While this often reveals unexpected structures, the preferred way from a chemist's perspective is to foster an understanding that enables a rational and systematic design. The main approaches to find new peptide sequences as building blocks in functional nanostructures are (I) taking inspiration from biology, (II) chemical engineering, and (III) using computational methods (**Figure 10**). Each of these approaches has its strengths and weaknesses that will be briefly discussed in the following sections.

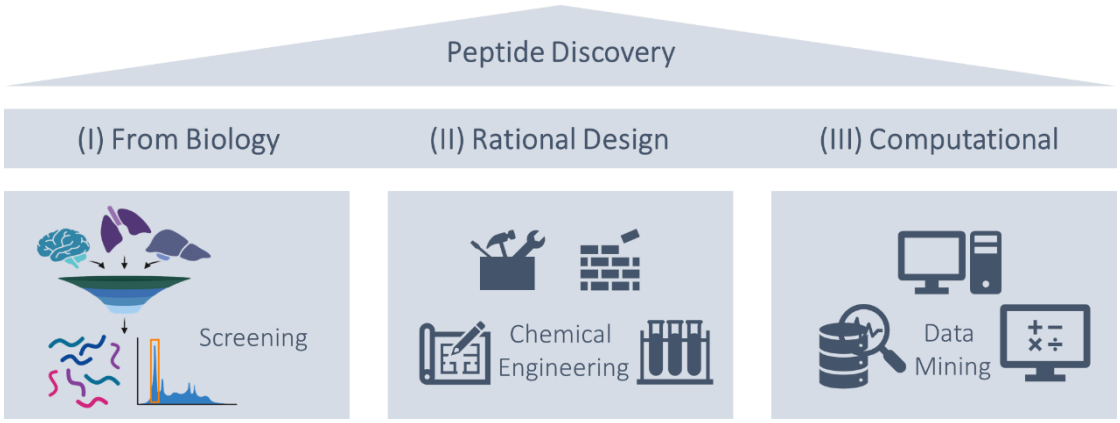


Figure 10 Strategies to find new functional peptides for self-assembly and bioactivity. Created with permission from BioRender.com.

Screening Endogenous Peptides

Bioinspired discovery approaches can rely on finding natural structures such as peptide sequences and protein fragments that have been optimized by Nature over millions of years of evolution. For example, extensive works on screening the biological active peptide sequences are conducted in the collaborative research center 1279 "human peptidome" that extracts and analyzes large protein and peptide libraries derived from human body fluids, tissues as well as microbiomes.[209] Generally, the workflow applies proteolysis, liquid chromatography for

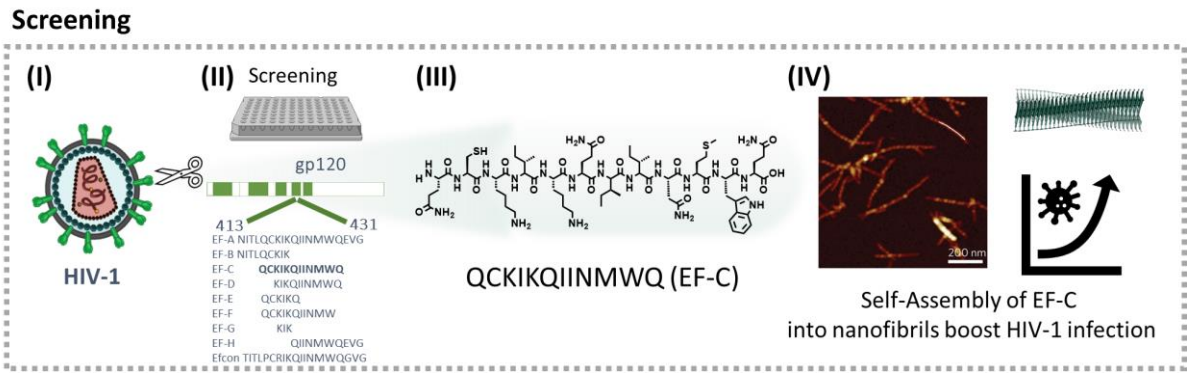


Figure 11 Graphical illustration of a peptide screening process from endogenous sources exemplified for the discovery of EF-C. **(I)** The HIV-1 glycoprotein (gp) 120 is fragmented and **(II)** high-throughput infection screening assays identify fraction $gp120_{413\text{-}431}$ as the active fraction. **(III)** Further reduction of the fraction isolates the sequence EF-C (QCKIKQIINMWQ) as the most potent peptide sequence. **(IV)** The EF-C peptide spontaneously assembles in water to amyloid-type fibrils and promote HIV-1 infection. Reproduced from Ref 156 with permission from 2013 Springer Nature.

separation and mass-spectrometry for characterization of proteins from natural origin. The proteins are repeatedly cleaved into short peptide fragments until a minimum bioactive peptide length is found.
Thereby, a wide range of peptides with different bioactivity has been identified in the past, such as antimicrobial peptide Angie-1 derived from the protein angiogenin in human serum,[210] antiviral peptide EPI-X4 derived albumin in hemofiltrate,[211] and infectivity enhancing amyloid-type peptides derived from semen fluid.[147]
Screening for certain bioactivity in biologic samples offers the possibility of surprising serendipitous findings, as previously shown by Münch and coworkers with the discovery of the short-self assembling peptide EF-C that has been isolated from the HIV-1 envelope glycoprotein 120 (gp120) (**Figure 11**) via high-throughput screening[154] and provided the basis for intensive research on amyloid-type peptides as retroviral transduction enhancers.[212]

Rational Design Strategies

Screening processes to find active peptides by serendipity from biological origin bind a lot of resources as they require a tedious workflow and strong collaborative efforts.
Therefore, it is important to develop alternative design strategies that could potentially provide higher chances to find active compounds. A popular approach is to rationally design peptides by chemical engineering of known active compounds. This allows to study property–activity relationships and have a high chance for creating new active compounds of the same type or sequence relation. Furthermore, this approach allows to use the chemical toolbox, *e.g.* including responsive trigger molecules that are known to change structure or conformation upon a certain stimulus. Supramolecular peptide materials offer significant advantages for rational design due to their facile sequence tuning at the level of individual amino acids, *i.e.* by Merrifield solid phase synthesis.[213] Unlike other covalent polymeric materials, which require the development of synthetic strategies for specific reactive groups or polymerization techniques, peptides can be readily modified and provide convenient access for systematic studies.[214]
For example, by designing derivatives of EF-C, peptides similar to the original cell-interactive peptide are created and used to study possible structure–activity relationship. Thereby, it was found that an increasing β-sheet order of the peptide fibrils correlates with neuronal cell attachment *in vitro* and *in vivo* (**Figure 12-IV**).[145]

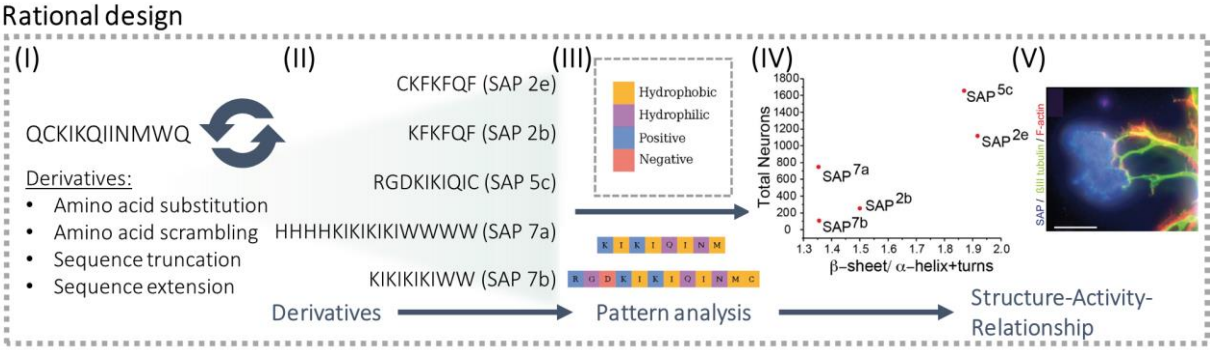


Figure 12 Rational design strategies for bioactive peptides illustrated with EF-C. **(I)** Sequence derivatives created in a systematic way by amino acid modifications yield **(II)** a peptide library that can be **(III)** investigated on recurring amino acid motifs, *e.g.* coarse-grained patterns. **(IV)** Structure-activity relationships with sequence properties such as β-sheet and bioactivity such as **(V)** neuronal cell attachment combined with pattern analysis yield insights in design rules for active peptides. Confocal microscopy shows merged image of neuronal cell tubulins (green) and actin (red) grasping peptide fibrils (blue), scale bar 20 µm. Reproduced from Ref 256 with permission from 2019 Wiley.

Although rational design through chemical engineering offers a direct way to discover active peptides and explore structure-property relationships, it can be time-consuming and expensive as each compound must be synthesized separately and thereby only provide insights into a limited chemical space. As a result, there are great inherent limitations to find new and unexpected peptide sequences using this method.

Synthesis and Characterization of Peptide Fibrils

To evaluate interactions qualitatively and quantitatively between peptides and cells, a broad spectrum of experimental methods is required. This is necessary because peptide self-assembly takes place at several length and time scales and the combination of different characterization techniques ensures a thorough investigation on these scales. In this section, the fundamental principles of frequently applied methods for peptide synthesis and characterization will be highlighted.

Short model peptides are typically synthesized via the well-established solid phase peptide synthesis (SPPS) using an Fmoc (fluorenyl methoxycarbonyl) protecting group protocol (**Figure 13A**).[213] This approach allows a sequential introduction of amino acids and building the peptides from the *C*- to *N*-terminus. The *C*-terminal amino acid is covalently immobilized on a polymeric bead, and each introduced amino acid is *N*-terminal protected with the base-labile Fmoc group. The coupling of amino acids with suitable coupling agents, typically *N*,*N*′-

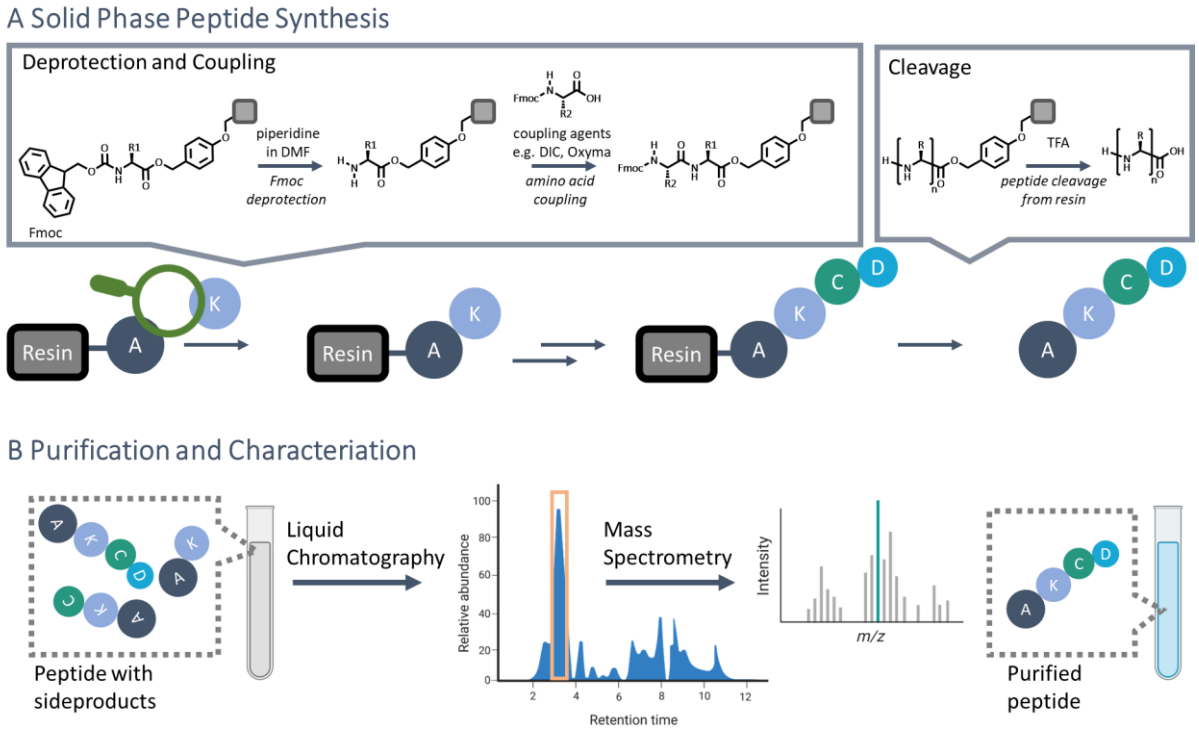


Figure 13 Schematic summary of a typical workflow in solid phase peptide synthesis. **A** Amino acids are covalently attached to a polymeric resin which enables subsequent amino acid coupling and thus sequence programmability. Prior to each coupling step, the Fmoc protecting group at the *N*-terminus is cleaved in basic conditions so that the next amino acid that is *C*-terminal preactivated with suitable coupling agents such as DIC and Oxyma can be coupled to the resin-bound amino acids. Each step is repeated until the desired peptide is created. In the final step, the peptide is cleaved from the resin under acidic conditions. **B** The crude product from solid phase synthesis is purified via reverse phase high pressure liquid chromatography and chemically identified by mass spectrometry. Created with permission from BioRender.com.

di*iso*propylcarbodiimide (DIC) as an activator and Oxyma as the activator base is performed after Fmoc-deprotection of the *N*-terminal amino acid in each step. The process is iterated until the final peptide sequence is achieved and released from the solid phase bead via acidification. Modern protocols involve developed methods such as microwave assisted synthesis that increase the reaction speed and efficiency.[215] While SPPS is not recommended for long peptides due to loss in yield with each additional coupling step, the procedure enables precise automatic syntheses of peptides shorter than 50 amino acids.[216]

After SPPS, the obtained peptides need to be purified which is typically done by high-pressure liquid chromatography (HPLC) coupled with mass spectrometry (MS) (**Figure 13B**). This allows to obtain the materials in high purity.[178]
Dissolving peptides with aggregation-prone sequence in water can yield self-assembled structures depending on the assembly pathway and preparation protocol. Some fibril-forming peptides tend to have a high aggregation propensity due to strong intermolecular hydrogen-bonding and hydrophobic interactions even at low molecular concentrations. Predissolving peptides in a polar aprotic solvent such as dimethyl sulfoxide (DMSO) or hexafluoro*iso*propanol (HFIP) can help to break the intermolecular hydrogen-bonds and reduce hydrophobic contributions, that in turn hinders peptide supramolecular structure formation. Diluting these solvents with enough water can restore the intermolecular interactions, which is why predissolving and diluting at the required conditions in water is a typical preparation protocol for controlled self-assembly of peptides.[178]
The supramolecular structure of peptides can be characterized via various microscopic and spectroscopic methods (**Figure 14**).

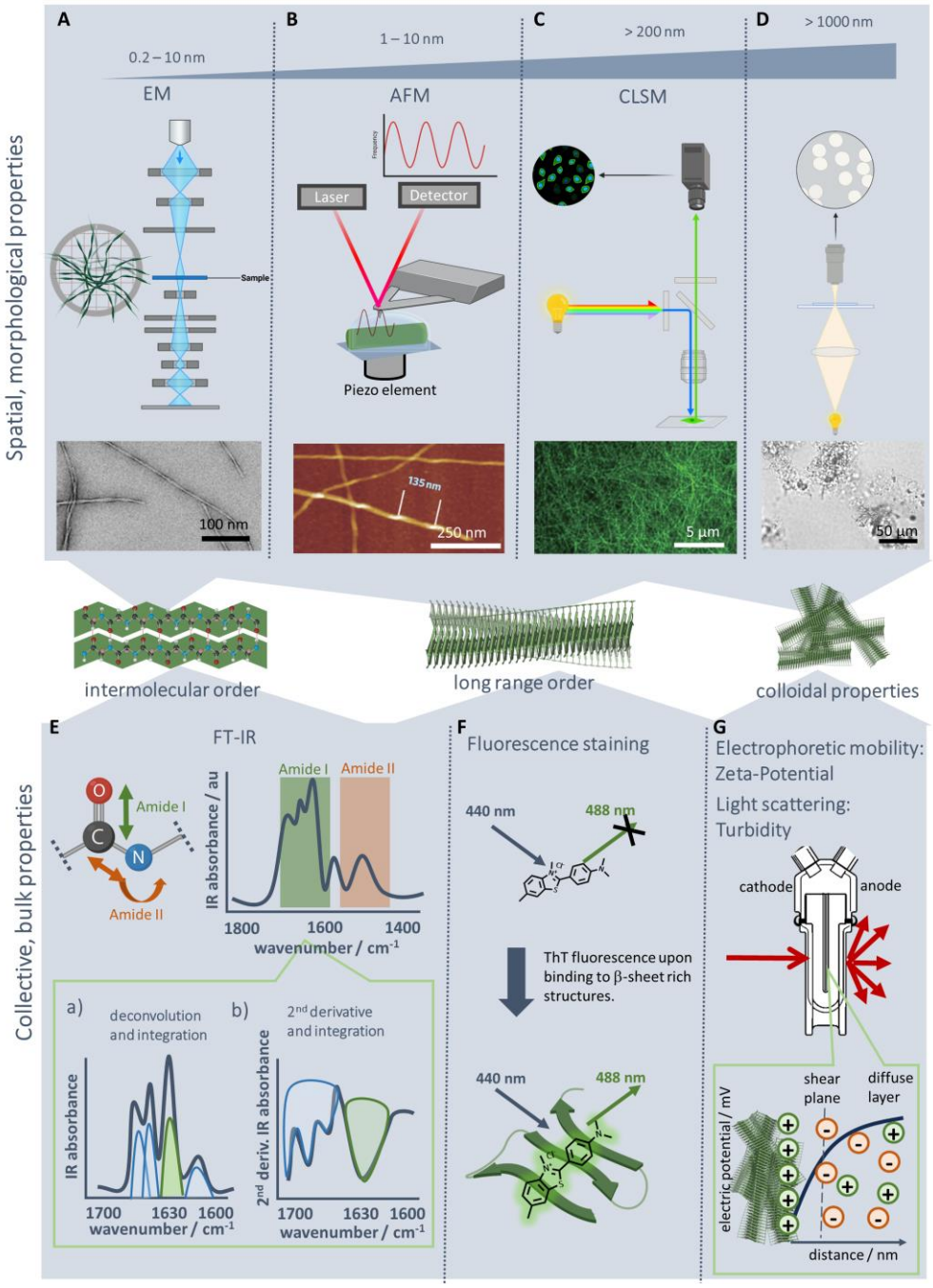


Figure 14 Overview on common characterization methods for self-assembling peptide fibrils with **A – D** spatial, morphological characterization and **E – G** collective, bulk property characterization. **A** Electron microscopy (EM) methods can resolve structures up to the atomic level. Micrograph displays amyloid Aβ structure, scale bar 100 nm. Reproduced from Ref 309 with permission from 2019 Springer Nature. **B** Atomic force microscopy (AFM) scans the surface topology and can resolve structures up to nm-length scale. Measurement shows β-lactoglobulin structure, scale bar 250 nm. Reproduced from Ref 131 with permission from 2010 Springer Nature. **C** Confocal laser scanning microscopy applies pinholes to resolve structures in the 100 nm length scale as shown for synthetic fiber $BPmocF_3$. Reproduced from Ref 310 with permission from 2016 Springer Nature. **D** Widefield microscopy for µm-sized structures as shown for protein aggregation. Reproduced from Ref 311 with permission from 2023 ACS. **E** FT-IR can be applied to evaluate intermolecular interaction and secondary structure by amide I and amide II region absorbance. The β-sheet content can be quantified by **a)** deconvolution and **b)** 2nd derivative and integration of the relevant peaks. **F** Scheme shows Thioflavin T structure that emits light at 488 nm after irradiation at 440 nm only upon binding to β-sheet fibrillar structures. **G** Colloidal bulk properties can be evaluated by light scattering (turbidity) and electrophoretic mobility (Zeta-potential). The Zeta-potential measures the electrophoretic double layer that is the shear plane between the immobilized and diffuse layer of ions. Schemes created with permission from BioRender.com.

While spectroscopic methods such as circular dichroism (CD), fourier transform infrared spectroscopy (FT-IR), and fluorescence staining methods give insights into bulk properties, microscopic methods such as transmission electron microscopy (TEM), atomic force microscopy (AFM), confocal laser scanning microscopy (CLSM) can be used to spatially resolve structures (**Figure 14A**).

The principle of TEM is based on the transmission of electrons through a thin specimen. The electrons interact with the specimen structure, and the resulting transmission pattern is magnified by electromagnetic lenses and typically projected on a charge-coupled device (CCD) sensor that converts the electron signal into a high-resolution image.[217] Typical morphologies that are determined via TEM for self-assembled peptides are non-fibrillar structures such as irregular amorphous aggregates and fibrillar structures such as crystalline fibrils, ribbons and sheet-like structures (**Figure 6A**). Overall, TEM is a high-resolution method for qualitative nanoscale characterization, but it lacks reliable quantification of fibril polymorphism and information on intermolecular interactions like β-sheet content. Further the drying process of the peptide structures on the specimen can cause structural changes such as fibers sticking more closely to each other, which does not resemble the initial structure *in situ*. To capture the actual *in situ* structure, Cryo-EM methods are applied. In this method, the specimen is shock-frosted to preserve its native structure, which allows for the nanoscale analysis of biological samples[218] such as the interaction of peptide fibrils with cells and virions. For example, cryo-EM was applied to resolve the cross-β-sheet structure of an infectivity-enhancing peptide at a sub nanometer resolution and the variation of the amount of hydrophobic and polar contact modes within a fiber was discovered to be the molecular basis for polymorphism.[121]

Similarly, AFM is often applied for nanoscopic morphological characterization of peptide fibrils (**Figure 14B**).[219] In contrast to EM methods, AFM mechanically screens the surface of a sample with an atomically small tip. The interactions between the tip and the sample cause changes in the tip's motion that can be detected to construct a topographical image of the sample's surface. Thus, AFM can provide information about surface features at the molecular scale, such as height variations and material properties. By chemically functionalizing the tip and coupling spectroscopic methods such as FT-IR, opportunities for further characterization such as determining the hydrophobicity and secondary structure of the nanostructure are possible.[220] Liquid-mode AFM allows *in situ* measurements of the native structure of peptide nanostructures that are applied on a substrate and covered in an aqueous medium to evaluate time-resolved assembly and decomposition processes.[221]

To visualize the spatiotemporal interactions of peptide fibrils with living cells *in situ*, confocal laser scanning microscopy (CLSM, **Figure 14C**) and widefield microscopy (**Figure 14D**) are the methods of choice. These microscopy methods rely on absorption of visible light or on fluorescence upon excitation. In contrast to widefield microscopy, CLSM has a higher resolution and is thus commonly applied as a powerful tool to study assembly processes in solution in real-time.[222] In principle, CLSM is based on excitation of fluorescent dye probes in a sample by a laser beam and capturing the fluorescence emission by using a pinhole that acts as an aperture and eliminates emission that does not originate from the focal plane. Thereby, a high axial resolution d can be achieved, that is theoretically limited by the abbe diffraction limit (eqn. 2)[223]

$$d = 2\lambda/N_A^2 \qquad \text{(eqn. 2)}$$

with N_A being the numerical aperture depending on the instrumental and refractive index (immersion oil typically 1.45).[223] For a peptide dye-labeled with a fluorescent probe with emission wavelength of 514 nm (*i.e.* fluorescein) the theoretically axial resolution is 488 nm. By utilizing super resolution techniques such as Airy Scan, a rotational pinhole disk,[224] the numerical aperture can be decreased which gives access to even higher spatiotemporal resolution, typically in the range of 100 nm and up to microseconds imaging time per frame that enable structural analysis of dynamic self-assembly processes.[225]

To comprehensively evaluate self-assembled peptide structures, it is important to assess both their spatially resolved morphologies and bulk properties. This is particularly important for amyloid-type peptides as they tend to display polymorphism (**Figure 6A**). For a fast and easy evaluation of the structural properties of a peptide's secondary structure CD and FT-IR measurements are commonly applied.

The principle of FT-IR is based on the distinct absorption energy that can be assigned to stretching and bending vibrational modes characteristic for every chemical bond and secondary structure conformation (**Figure 14E**).[226] For Self-assembled peptides, a common strategy is to measure the attenuated total reflectance (ATR) FT-IR of lyophilized peptides to avoid noise from water and for a concentration independent measure.[226] The quantification of the ratio of intermolecular structures can be conducted via spectral deconvolution[227] or by integration of the second derivative[228] of the spectra in the amide II (1500–1600 cm^{-1}) and amide I (1600–1700 cm^{-1}) regions for characteristic β-sheet peaks at approximately 1520 cm^{-1} and 1615–1640 cm^{-1} (**Figure 14E**).[229]

CD measures the difference in absorbance of circularly polarized light that is caused by the chirality and structural conformation of a sample. The secondary structure of self-assembled peptides can be assigned to characteristic Cotton effects for example the $n \rightarrow \pi^*$ energy transition of H-bonding interaction of β-sheets at 195 nm and 220 nm.[230] In contrast to ATR FT-IR, CD is always measured in solution. It is not suited to characterize colloidally aggregating peptide samples, that disturb the absorbance measurement by light scattering.[231]

Additional characterization of amyloid structural features can be conducted via the fluorescent dye Thioflavin-T (ThT) or the commercialized derivative of ThT, Proteostat.[232] The ThT dye binds to β-sheet rich structures, which immobilizes the axial rotation of the molecule that can be excited at 450 nm and emits at 482 nm (**Figure 14F**).[232] As the ThT-fluorescence can also be influenced by a variety of other conditions such as amyloid polymorphism,[233] charged amino acids,[234] and aromatic amino acid,[235] large fluorescence variations and false-positive and false-negative results can occur. Therefore, the ThT-fluorescence assay should only be applied as a qualitative measure and always accompanied with other complementary methods for secondary structure evaluation such as FT-IR and CD.

Zeta-potential measurements combine light-scattering and electrophoretic measurements to characterize the bulk properties of colloidal aggregation and their particle surface charge (**Figure 14G**). The derived count rate of scattered light detected in a backscattering setup at 173° is suited for evaluating information on turbidity of concentrated samples of peptide fibrils.[106,236,237] The mobility of particles in an electric field can be determined by phase shifts in light scattering and is employed in Zeta-potential measurements to calculate the charge of the

shear plane that determines the overall surface charge of peptide fibrils and aggregates (**Figure 14G**).[238] However, limitations arise from insoluble, non-self-assembled peptides that can also contribute to the count rate of scattered light and the aggregation state of peptide fibrils can significantly change the Zeta-potential.[239–241] To avoid misinterpretation, zeta-potential measurement should be accompaniment by complementary structural analysis such as TEM and widefield microscopy.

By combining spectroscopic and microscopic methods, an in-depth analysis of local properties of peptide fibrils is possible. For example, FT-IR combined with AFM can spatially resolve the intermolecular assembly order at the nanoscale,[129] and MALDI-ToF Mass Spectrometry (MS) Imaging performs mass spectrometry measurements while scanning a surface thereby directly visualizing spatial distributions of molecular species on a substrate (**Figure 15**).[178]

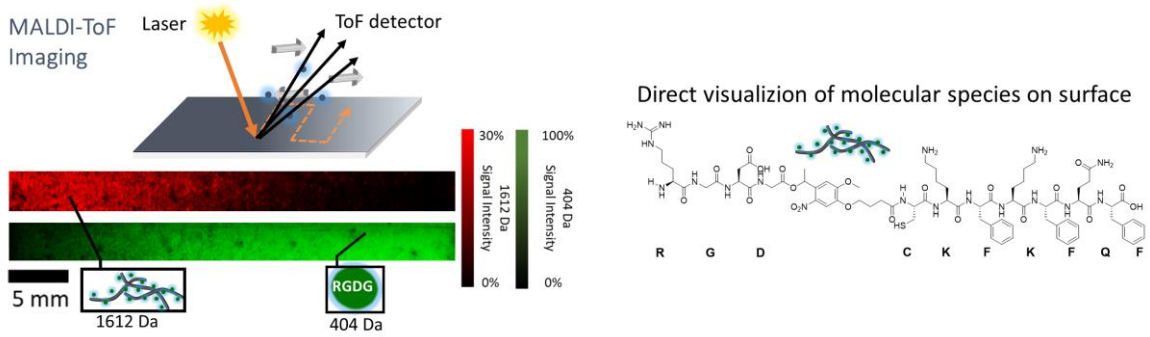


Figure 15 Matrix-assisted laser desorption ionization mass spectrometry time of flight (MALDI-MS-ToF) imaging demonstrated for a substrate that is coated with two different peptide species. Mass spectrometry allows direct molecular species identification and scanning the surface with a laser spatially resolved the distribution of the molecules. Reproduced from Ref 178. This image is licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

Finally, the effects of peptide fibrils on cells are typically studied via microscopic and fluorescence-based assay. The delivery of materials can be studied in specialized cell-lines that can express marker molecules upon successful delivery. This principle is employed for example, to quantify HIV-1 infection in TZM-bl cells. The TZM-bl cell line is a genetically engineered version of HeLa cells that produces β-galactosidase upon genetic code modification by HIV-1.[242] The β-galactosidase amounts can be easily evaluated in a high-throughput manner via a luminescence assay that is based on chemiluminescence upon enzymatic degradation of a 1,2 dioxetane-based substrate (**Figure 16A**).[243] Fluorescence based assays can also be applied to determine cell-viability. A commonly employed technique is staining cells with calcein, which selectively enters living cells and reacts to a fluorescence species upon reacting with intracellular esterase. This method enables the visualization of the distribution of living cells, providing valuable insights into their spatial arrangement and viability (**Figure 16**B).[244]
To quantify the viability of cells, staining methods such as the MTT assay can be used. In this method, colorless MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) reacts by mitochondrial reductase of living cells into formazan, that absorbs light at 500-600 nm (**Figure 16C**).[245]
Overall, interactions between peptide fibrils, virions, and cells occur across various length scales, which is why various complementary methods are required to comprehensively characterize their properties.

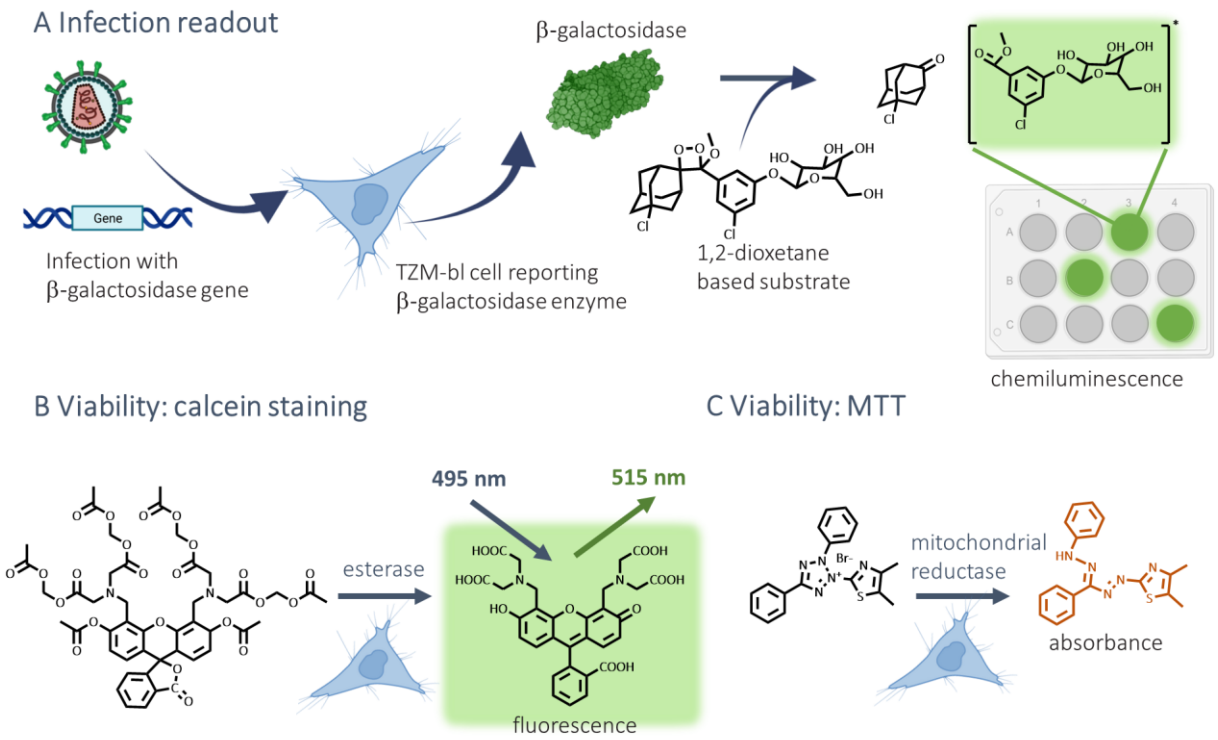


Figure 16 Overview of selected biological assay to evaluate effects of peptide fibrils on cells. **A** Infection enhancement of virions is evaluated by a chemiluminescence assay that exploits β-galactosidase enzyme that is expressed upon the introduction of the corresponding gene from the virus. β-Galactosidase degrades a 1,2 dioxoetane based substrate into an excited state intermediate that decays by irradiation of green light. **B** The cellular viability can be evaluated by calcein staining. Calcein is taken up into cells and only living cells produce esterase that catalyzes the formation of a fluorescent compound. **C** The MTT viability assay takes advantage of mitochondrial activity of living cells to transform the colorless MTT compound into formazan that appears as yellow crystals. Created with permission from BioRender.com.

It is challenging to rationally anticipate the design of peptides with a certain self-assembly or bioactivity *a priori*. This is due of the cooperative nature of supramolecular assembly that is sensitive to changes of the peptide sequence that affects aggregation propensity and bioactivity. Therefore, in addition to experimental techniques for analyzing and quantifying structural properties, computational models are essential to mechanistically explain the rapid self-assembly process and to trace back multi-scale dependencies to the molecular level.

Computational Modeling and Prediction of Peptide Sequence and Structure

Computational methods, including machine learning and molecular dynamics simulations, are a rapidly expanding field in peptide discovery.[246] They offer important capabilities for supramolecular peptide nanostructure design, including (I) the prediction of the aggregation behavior of specific sequences, (II) mechanistic insights into aggregation, solvent interactions, and three-dimensional structural properties, and (III) the generation of entirely new sequences with desired properties (**Figure 17**).[247] Theoretical models are essential to understand molecular contributions within complex hierarchical structures that are formed at various time and length scales. They also allow to model an ensemble of structural conformations to gain insights into structural dynamics.[248] Thereby, they are complementary to experimental characterization methods such as TEM that collect static information of the structure. Computer simulations, with their ability to model self-assembly interactions, have revolutionized protein research by

enabling three-dimensional structural predictions.[249] One of the reasons why modelling and artificial intelligence (AI) can be applied here is because the experimental results of the past decades in characterizing proteins through techniques like crystallography and electron microscopy are publicly accessible in databases.

One of the initial motivations behind computationally analyzing the aggregation behavior of peptides emerged from the interest to understand protein aggregation in pathogenic contexts.[250,251] Most aggregation analysis tools focus on predicting sequence aggregation propensity (**Figure 17A**)[252] or identifying sequence pattern similarities with known aggregating peptides (**Figure 12**).[253] To perform these analyses, bioinformatic techniques are applied that classify amino acid side chain's physical, chemical and structural properties to certain descriptors. For example, the most frequently considered basic descriptors are hydrophobicity, net charge, and van der Waals volume that are determining each amino acid to a different extent.[254] A combination of descriptors through mathematical operation can yield new descriptors, which makes it possible to create a large number of complex descriptors.[255]

Pattern analysis aims to identify recurring motifs in peptides with a target property in mind to both foster a mechanistic understanding and also to predict new active sequences. To reduce the combinatorial space, coarse-graining according to certain amino acid classification schemes, *e.g.* based on hydrophobicity and charge are commonly applied.[253,256]

Traditional aggregation prediction software relies on experimentally determined properties of individual amino acid side chains, and applies these properties to the entire peptide sequence. As an example, the aggregation prediction tool "Aggrescan" assesses the distribution of specific amino acids classified as "aggregation prone" within a sequence, relative to their neighboring amino acid environment.[257] The capability of amino acids to undergo aggregation is based on empirical studies on Amyloid Aβ.[257] This enables the identification of "hot spot" regions in the sequence that are more likely to contribute to aggregation (**Figure 17A-I**).[257] Similarly, other tools such as Tango,[250] Waltz[258] and APPNN[259] apply the free energy potential of hydrogen bonds, position specific solvation, and a neural network correlation of bioinformatic amino acids properties, respectively (**Figure 17A-II**). While these tools perform well to give away a general tendency for aggregation, they cannot predict the structure of aggregated proteins or peptides.

The three-dimensional structure of peptides nanostructures can be modelled by molecular dynamics (MD) simulations. This method applies simplified force fields to calculate the interaction between molecules. MD simulation for peptide self-assembly typically employ coarse-grained and atomistic methods (**Figure 17B**).

Atomistic simulations are typically used for structures with known conformations and utilized to determine the secondary structure and stability of the self-assembled system.[248] However, atomistic calculations of large molecules, such as typical short peptides, demand immense computational power, particularly when starting from an initially disordered state.[246] Consequently, for peptide systems, coarse-grained models have been widely adopted.[246] By reducing the resolution and complexity of the system, these models allow for more efficient simulations and enable the calculation of long time (μs to ms) and large length scales (~ millions of atoms).[246] One common coarse-graining approach involves representing a group of atoms as a single bead that captures its essential properties such as charge and polarity. This bead categorization can be customized to answer different questions. For example, the Martini model summarizes four heavy atoms into one bead, *i.e.* representing the amide backbone of a peptide as one bead.[260] This simplification enables high-throughput screening of peptide self-assembly and was applied, for instance, to predict aggregation propensity of hundreds of

dipeptides[261] and tripeptides[262] in a cost-effective manner. However, coarse-graining methods lack the capability to reproduce specific intermolecular interactions.[263–266]
To mitigate this limitation, backmapping coarse-grained simulations with atomistic simulations is employed. This approach reduces information loss and enables the modeling of peptide self-assembly with greater accuracy.[267] Backmapping can be used to assess the stability of peptide assemblies by calculating the root-mean-square fluctuation (RMSF) of the peptide backbone. A low RMSF corresponds to minimal structural variation within a given timeframe, indicating a stable structure. For instance, by studying various amyloid polymorphs it was observed that stable structures exhibit a buried hydrophobic core and saturated, non-frustrated hydrogen bonding interactions.[268] Another method to quantify fibril stability involves calculating the proportion of β-sheet secondary structures within a fibrillar assembly using atomistic MD simulations.[248,269]

While computational methods have long been used in peptide design and discovery to model molecular interactions and predict secondary structures, recently there is a growing interest in the application of machine learning approaches. Machine learning methods hold the great potential to surpass the natural pace of evolution, which took millions of years to create proteins capable of precise folding and function. In 1997, the software Rosetta was one of the pioneering tools to model the protein structural conformation.[270] Rosetta is based on physical force fields which implies assumptions on certain parameters such as electrostatics as important, which, in turn, limits the model in view of parameters that can be applied.[271] Roughly 20 years later in 2020, the deep-learning model AlphaFold2, was reported to precisely predict the three-dimensional structure of any given amino acid sequence without any pre-assumptions made.[249] Prior to these developments, the determination of protein structures was mostly limited to laborious crystallography and Cryo-EM analysis.[249] More recently in 2022, ProtGPT2, a machine learning tool was introduced.[272] ProtGPT2 can perform one of most challenging tasks in biochemistry, namely the inverse design of proteins with a desired target function.[272] This large language model can generate *de novo* protein sequences because it was trained on an immense database with approximately 45 million sequences. It is important to remember that while these models revolutionize protein and peptide discovery, they are powered by the collective efforts and tedious labor of thousands of research groups collecting protein sequences and structure information for over more than 50 years and making it publicly accessible.[272]

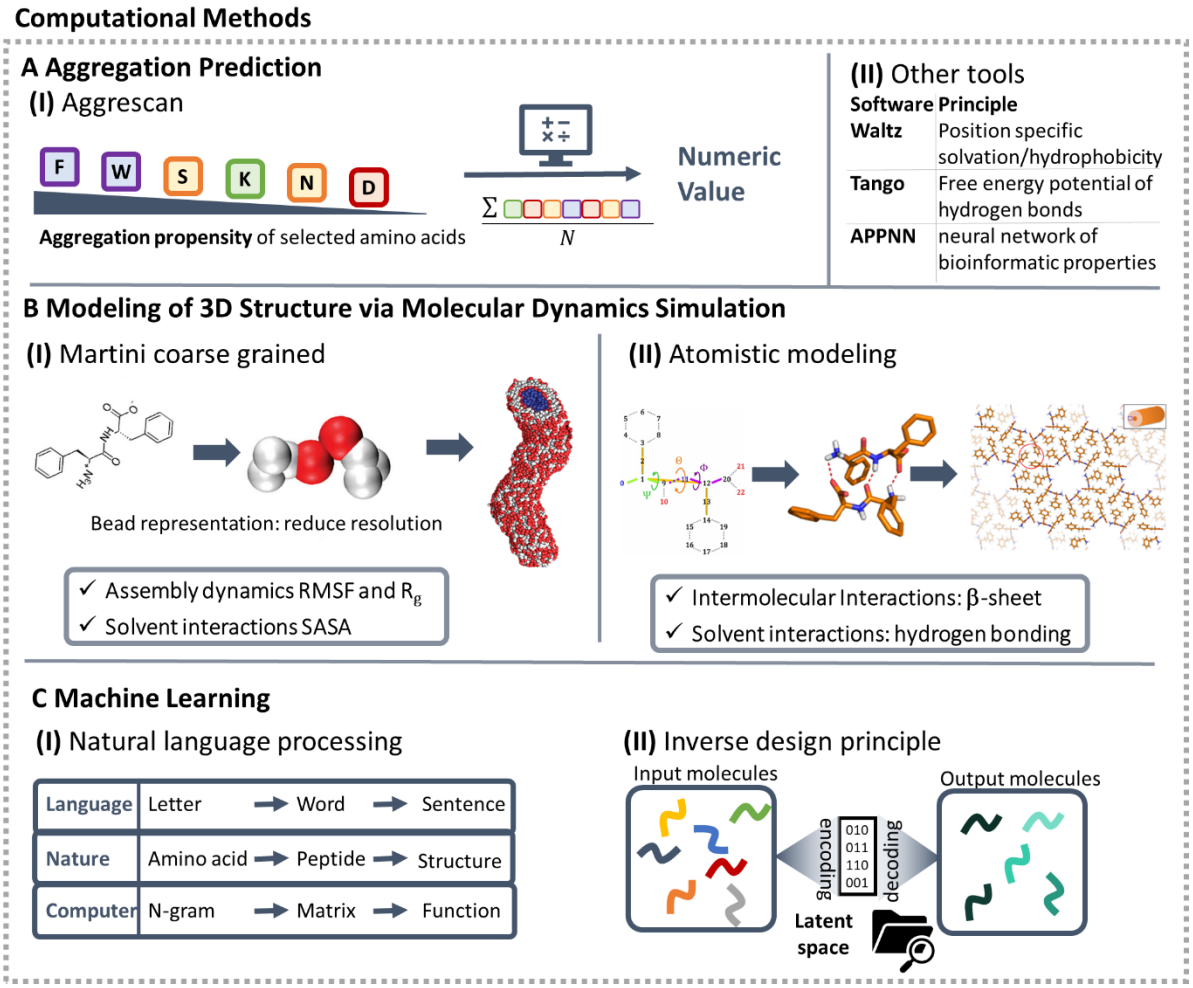


Figure 17 Summary of selected bioinformatic and computational methods to discover and model self-assembling peptides. **A (I)** Traditional aggregation methods such as Aggrescan extrapolate the single amino acids properties such as aggregation propensity of side chains to the overall peptide structure.[257] **(II)** Other tools like Waltz,[258] Tango,[250] and APPNN[259] apply a similar approach by using the hydrophobicity, free energy potential of hydrogen bonds, and a complex variety of amino acid bioinformatic properties, respectively. **B** One bead per four heavy atoms coarse-grained MD simulations compared to all-atom simulations of diphenylalanine self-assembly. The conversion from atomistic to CG representations comes with a drastic reduction of resolution that reduces computational cost. CG enables the study of assembly dynamics over long time scales and solvent interactions for the entire structure (solvent accessible surface area, SASA), however it cannot determine the exact conformational analysis and intermolecular order such as β-sheet and solvent interacting sites such as hydrogen-bonding, that is, in turn, covered by atomistic simulations. **B-I** Reproduced from Ref 261 with permission from 2011 ACS. **B-II** Reproduced from Ref 312 with permission from 2013 ACS. **C (I)** Natural language processing machine-learning tools are based on the same principle as language. The combination of letters gives rise to words, that can give a meaning if assembled in a specific order in a sentence. This can be transferred to chemical systems such as amino acids creating peptides that have a function if the assembled correctly in a structure. N-grams are the letters and amino acids equivalents in computer language, N-grams can create abstract vectorized representations in a matrix that reports a function. **(II)** Illustration depicting principle of inverse engineering of molecules by natural language processing. A dataset is represented in a uniform manner and encodes their corresponding information in a higher dimensional latent space to find patterns. These patterns can be then decoded to actual molecules.[280]

To enable proper training of an algorithm the molecules of interest must be represented in a uniform way. There has long been a desire to establish a universal language for representing chemical structures in words, such as the IUPAC nomenclature.[273] With the advent of computers, the development of approaches for representing molecules and polymers in a

machine-readable format has rapidly advanced with prominent examples of string representations including SMILES,[274] BigSMILES[275] and SELFIES.[276] Due to recent developments in large language modeling in chemisty[277] the importance of such universal presentation dramatically increased.[278] Proteins and peptides are an evident starting point, as they are already represented with one-letter code by convention. Leveraging machine learning techniques holds great promise to enable rapid and cost-efficient pre-screening of new peptides with desired target properties.[279] This is crucial because screening by generating random sequences becomes unfeasible as the sequence space expands exponentially with the number of residues in a peptide.

A popular approach in data-driven AI methods is training a machine-learning model in the latent space to find underlying hidden connection between datasets.[280] The underlying concept behind this approach is that a desired property can be encoded by peptides that may be far apart in sequence space but close in a higher dimensional space. This higher-dimensional space considers various bioinformatic features beyond the primary sequence alone.[281] In this framework, the sequence is initially encoded to train a model in a higher-dimensional space and the learned relationships are utilized in a decoding step to inverse engineer a sequence (**Figure 17C**).[280] Word embedding models such as Word2Vec have been initially developed for natural language processing[282] and use the peptide's primary sequence as an input information to encode complex sequence order to train a model.[283] To universally encode diverse sequence lengths across various datasets, N-gram models are widely used in computational linguistics to obtain machine-processable representations of a word. Due to the reliable representation without loss of information on amino acids distribution and retaining an interpretable structure N-grams found application in encoding and predicting amyloid-peptides.[284] To model physicochemical properties of proteins the N-gram representation can be supplemented with further information. An example of this is ProtVec, a vector model that was trained on a protein database with more than 300,000 sequences including their bioinformatic information.[285] High-dimensional vector spaces are challenging to visualize. To address this, dimension reduction techniques like principal component analysis (PCA) and t-distributed stochastic neighbor embeddings (t-SNE) can be employed. These methods allow for the visualization of multi-dimensional features by reducing them to a two-dimensional plot, enabling the evaluation of distance relationships in a human-readable manner.[286]

For the rational design of new peptides, simple computational models with only few parameters are preferred because they maintain explainability and facilitate the extraction of fundamental design principles for further research.[287] Machine learning continues to become more powerful each year, and will likely be soon established as an elementary part of the toolbox in supramolecular peptide design.

2. MOTIVATION AND OBJECTIVES

Custom medical treatments tailored for individual patients hold immense potential to revolutionize our approach to combating diseases. Personalized medicine has evolved from a visionary concept to approved medications with practical applications. For instance, in the field of gene therapy, notable progress has been made in tissue engineering in a recent market approval that enables treatment of patients with skin blistering genetic disorders.[11]

Many of these biomedical manufacturing processes rely on retroviruses as stable and safe gene delivery vehicles. However, they suffer from low cell attachment and limited infectivity.[18] To address these challenges and improve therapeutic efficacy, various adjuvants including polymers, lipids, and peptides have been utilized to promote the interaction between viral particles and cells. In **Chapter 6.1**, we provide a comprehensive review on the available transduction enhancers and compare the properties of various materials to deliver viral vectors. We identify a gap between the demand for complex, tailored materials in therapeutic applications and their actual availability. Supramolecular biomaterials, particularly self-assembling peptides, offer versatile possibilities for therapeutic applications yet have remained largely unexplored.

For instance, the self-assembling amyloid-type peptide EF-C has been found to promote lenti– and retroviral gene delivery. This peptide was successfully tested in clinical settings which sparked interest in using peptide fibrils as adjuvants in gene therapy.[154] The positive charge of EF-C fibrils helps to overcome electrostatic repulsions between cells and virions.[162] Furthermore, it has been observed that cells actively pull the fibrils to the cell surface via protrusions,[159] take them up via macropinocytosis,[163] and degrade them within cellular vesicles.[163] However, the exact structure-property-activity relationship of EF-C fibrils in this context and whether there are universal design rules that can be generalized still remain unclear. Answering these questions will not only benefit the future design of peptide materials in the context of gene therapy but also provide fundamental insight into the bioactivity of amyloid structures.

Therefore, the aim of this thesis is to gain an understanding of design principles and characteristic features of supramolecular peptide fibrils for therapeutic applications. To this end, I focus on four main objectives that are: **(I) optimizing** the EF-C structure to improve transduction efficiency, **(II)** systematic **understanding** of important molecular and physicochemical features for improving cell-virus interactions **(III) predicting** and testing new peptides based on the acquired insights and finally **(IV) applying** selected peptides with focus on translational clinical aspects.

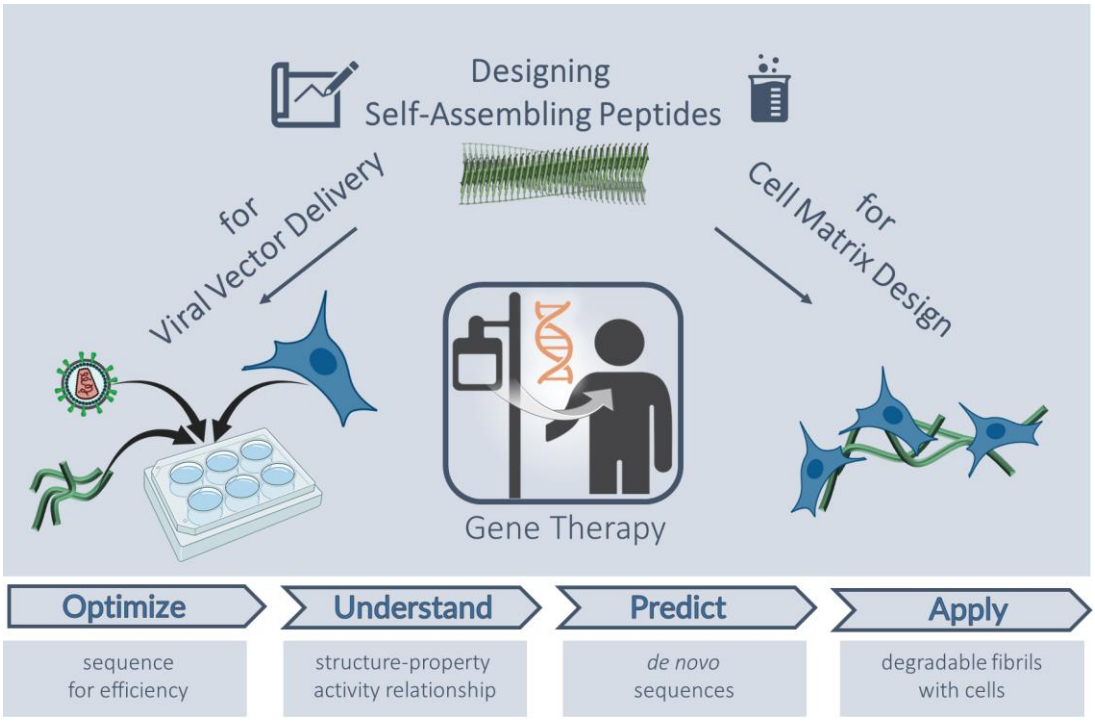


Figure 18 Motivation and design of this thesis. The aim of this thesis is to unlock the potential of self-assembling peptides for biomedical applications in gene therapy, *i.e.* for viral vector delivery and as cell-supportive matrix scaffolds. To this end, I divide this work into four main objectives, that are the optimization of the sequence, the study of the structure-property-activity relationship, prediction of new sequences and their application with focus on clinically relevant aspects of degradability. Created with permission from BioRender.com.

(I) Optimizing sequences to increase transduction efficiency

In the first objective (**Chapter 3.2**), we will create a peptide library of derivatives based on EF-C and investigate how their sequence impacts the physicochemical properties and transduction enhancement capabilities. The sequence optimization process will focus on creating shorter peptides with equal or superior properties compared to the original sequence. By designing shorter peptides with optimized sequences, we aim to maximize the desired bioactive properties while minimizing costs associated with the production of these materials.

(II) Understanding the underlying structure–property–activity relationship

In the second project, the peptide library established in (I) will be expanded and diversified to form the basis for a structure-property-activity relationship study (**Chapter 3.3**). The overall aim is to gain systematic insights into important molecular and physicochemical features for cell-virus interaction. As the relevant processes and self-assembly happen across different length scales, various techniques for the physicochemical characterization of the new materials are

used to tackle this multi-scale problem. By exploring similar structures covering a wide range of activities, we can gain insights into how these small structure changes affect their physicochemical properties and bioactivity. This, in turn, allows us to decipher the underlying design rules and universal physicochemical features that govern infectivity enhancement. Since the self-assembly of short peptides occurs on experimentally inaccessible time scales, we will perform molecular dynamics simulations (**Chapter 3.4**). This computational approach will help us to understand self-assembly, stability and multi-scale aggregation on a mechanistic level for selected peptides in our library. By elucidating the mechanisms involved, both through experimental and simulation studies, we will not only gain a thorough understanding of infectivity enhancing peptides but also shed light on multi-scale self-assembly of short peptides forming amyloid-type structures.

(III) Predicting *de novo* active sequences

The third objective employs the EF-C based library to predict new infectivity enhancing peptides. While machine learning has proven to be an effective method in predicting protein structure and function, reports on its application in the discovery of short-self assembling peptides are limited so far. In **Chapter 3.5** we will leverage natural language processing models on the EF-C based peptide library, to train an algorithm that identifies common bioinformatic properties for active peptides in the latent space. Decoding these hidden features into actual peptide sequences enables us to find novel peptides with high therapeutic potential. This approach will be applied on a limited data set with 163 peptides that represents a typical situation in early research stages of clinical studies. Overall, this study will evaluate new data-driven methods for rapidly expanding the sequence space of active peptides.

(IV) Applying degradable fibrils with cells

For clinical applications, it is of utmost importance that the materials introduced into a patient, *e.g.* peptide nanofibrils, can be efficiently degraded by cells into non-toxic metabolites. Especially with stable amyloid-like fibrils this has been a challenge so far. To address this requirement, we will design new peptides and examine their degradability via two different approaches.

In **Chapter 3.6** we will investigate peptide amphiphiles as potential degradable transduction enhancers. These peptide derivatives form fibrillar structures but are not related to amyloid-type peptides. Specifically, the peptide amphiphile C_{16}-VVVAAAKKK-NH_2 developed by Stupp and coworkers will provide the basis for the investigation.[288] Peptide amphiphile self-assembly is primarily driven by the hydrophobic tail, which resides inside the fibrillar core, while the peptide part is presented on the exterior of the fibril. By incorporating less hydrophobic and smaller amino acids like alanine and glycine, it may be possible to alter the hydrophobicity and flexibility of the fibril surface in a way that enables control over fibril-fibril aggregation and thus biodegradability. Changing the sequence of short peptides has the advantage of deconvoluting contributions of single amino acids on self-assembly, activity and degradability. To study these effects in a systematic way, we will create sequence variations and characterize them by experiments and via molecular dynamics simulations.

In **Chapter 3.7** we further broaden our approach by incorporating a light-sensitive linker into a peptide backbone. Specifically, we use the sequence optimized EF-C derivative CKFKFQF and a photo-responsive *o*-nitrobenzyl linker. It will be particularly interesting to test how the

structural integrity and bioactivity of the peptide fibril can be controlled using the external light trigger. It was shown that the presence of peptide fibrils[178] and substrate stiffness[179] can guide cell-attachment and growth. However, to date, there has been no reports that show how changes in peptide morphology affect cell-adhesion.
Since translational applications are one of the ultimate goals of infectivity enhancing peptide fibrils, important aspects for clinical practice such as degradability must be addressed. The objectives in **Chapter 3.6** and **3.7** aim to address these challenges related to fibril stability from chemical engineering perspective by designing materials that provide optimized degradability (**Chapter 3.6**) and can be decomposed in a stimuli-responsive manner (**Chapter 3.7**).

Implications

To unlock the full potential of amyloid-type peptides as biomaterials for cellular applications, it is essential to gain a comprehensive understanding of and control over their structure across various length scales. In this thesis, my aim is to investigate the intricate relationships between molecular and physicochemical factors influencing cell-virus interactions, investigate predictive methods for *de novo* peptides, and optimize the design of peptide fibrils towards clinical translational applications. In a broader context, I aim to understand how to create functional biomaterials from non-functional biomolecules via self-assembly and how changes on the molecular scale propagate to greater length scales and modulate their (bio)functional response. The experimental readout for activity, viral infectivity enhancement, offers a multi-parameter, multi-scale challenge for the optimization of biologically active supramolecular peptide assemblies and is therefore an interesting starting point to address these questions. The development of complex materials requires cross-disciplinary efforts that involve integrating molecular simulations, chemical synthesis and engineering, physical characterization, and biological evaluation.

Overall, this thesis contributes to the development of functional peptide fibrils as materials supporting cell-attachment and enabling efficient gene delivery for biomedical applications.

3. RESULTS AND DISCUSSION

Chapter 3 provides a concise summary of the main findings and **Chapter 6** includes the full research publications associated with the objectives of this thesis.

3.1 Materials Promoting Viral Gene Delivery

Gene therapy promises to treat serious diseases such as cancer and has successfully emerged from purely fundamental research into clinical applications over the past 20 years. The majority of clinical trials and approved gene therapies for human treatment apply biotechnologically engineered virions as the key technology to safely deliver the therapeutic genes. The clinical applications are mainly limited by inefficient virion-cell interaction that results in low transduction efficiency. To promote the interaction between virions and cells, adjuvants such as polymers, lipids, nanoparticles, and peptides are applied in basic research and in clinics.

In this comprehensive review we summarize, highlight, and compare these selected viral delivery systems. We conclude that many of these materials can efficiently and safely boost infectivity, mainly by overcoming electrostatic repulsion and subsequent colocalization of the virions with the cells for efficient interaction and uptake of virions into cells. Traditionally, many reports focus on covalent polymeric materials as adjuvants but recently supramolecular biomaterials such as self-assembling peptides have attracted increasing interest. Nanostructures that form via non-covalent bonds offer versatile possibilities for therapeutic applications as they can adapt to the local environment and in the case of peptides are often inherently biocompatible and biodegradable. Therefore, self-assembling peptides are one of the most promising candidates to address the urgent need for highly efficient gene delivery with biocompatible materials, yet remain largely unexplored.

This review has been published by Kaygisiz, K. and Synatschke, C. V. in Biomaterials Science, **2020**, 8, 6113–6156, https://doi.org/10.1039/D0BM01367F and reprinted in **Chapter 6.1**, page 75.[13] The contribution of Kübra Kaygisiz to this study was to provide the concept for the article, perform the literature review, and write the original draft. A detailed overview of the contributions of all authors is summarized in **Chapter 6.1**, page 75.

3.2 Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction

High-throughput screenings in biologic samples enable possibilities for surprising, serendipitous findings. One example of such a seminal discovery was shown by Münch and coworkers for the 12-mer peptide enhancing factor C (EF-C, QCKIKQIINMWQ from the HIV-1 envelope glycoprotein $120_{417-428}$ (gp120).[154] This short-self assembling peptide was the first example demonstrating that functional amyloids are potential candidates for clinical applications in gene therapy and sparked fundamental interest for research on amyloid-type peptides as retroviral transduction enhancers.[212]

However, it remains unclear what molecular parameters and amino acid motifs in EF-C are responsible for the activity. Identifying minimal sequences is essential for a fundamental

understanding of bioactivity in amyloid nanomaterials. Therefore, in this study we report the optimization of the EF-C peptide sequence with respect to its retroviral transduction enhancement.

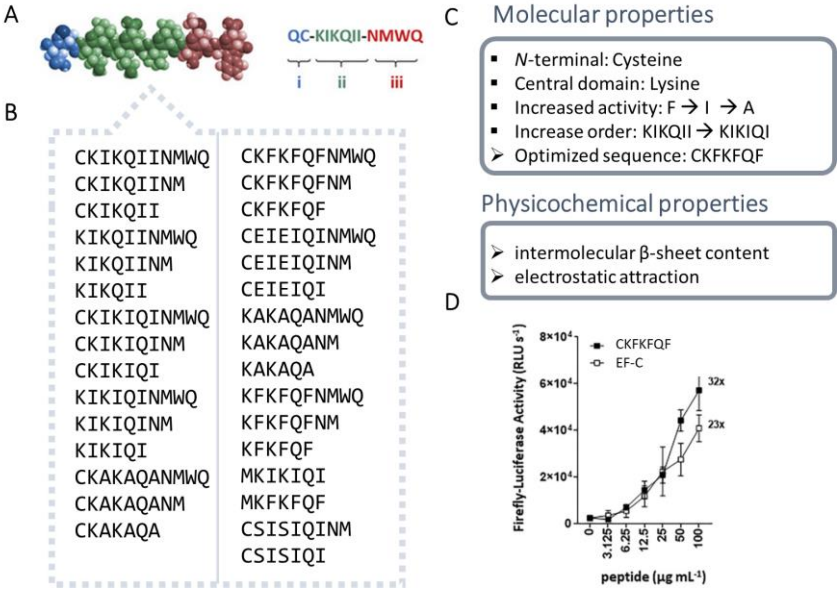


Figure 19 Summary of study design and main findings for sequence optimization of the EF-C peptide. **A** Schematic model of EF-C with important domains (i) QC, (ii) KIKQII, and (iii) NMWQ. **B** Peptide library based on sequence modifications of EF-C domains. **C** Main findings on molecular design and physicochemical properties important for retroviral infectivity enhancement. **D** Comparison of activity for EF-C and its sequence optimized derivative CKFKFQF. Luminescence readout of firelfly luciferase of β-galactosidase detection (RLU, relative light units). Figure A and D reproduced from Ref 67. These images are licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

Based on chemical intuition and previous knowledge on bioactive peptide motifs we divided the sequence EF-C (QCKIKQIINMWQ) in three smaller subunits: the *N*-terminal "QC" fragment, the central amphiphilic "KIKQII" fragment and a *C*-terminal "NMWQ" fragment (**Figure 19A**). It was assumed that the hydrophobic interactions of isoleucine (I) residues in the amphiphilic domain drives self-assembly while the cationic, polar amino acid lysine (K) provides electrostatic attractive sites with negatively charged virus and cell-membrane. Further the C-terminal domain was assumed to play an important role as well because it was previously identified as the binding site of the HIV-1 gp120 to CD4 receptors in cell membranes enabling viral entry.[289] To investigate the relevance of these sequence domains, we applied a rational design strategy that includes substitution of hydrophobic amino acid (I) in less (A) and higher hydrophobic (F) ones, replacing positively charged amino acids (K) with negatively charged amino acids (E), stepwise truncation of the *C*-terminal and *N*-terminal fragments, and scrambling amino acid order in the central domain to achieve an more regular alternating amphiphilic sequence pattern. 31 new peptides have been synthesized and their properties and activity were investigated via transmission electron microscopy, zeta-potential, attenuated total reflection (ATR) FT-IR, quantitative retroviral β-galactosidase transduction assays, and confocal fluorescence microscopy that provide information on fibril morphology, overall surface charge, secondary structure, and fibril-virus-cell interaction, respectively (**Figure 19B**).

Comparing the infectivity enhancement of peptides with similar sequences qualitatively by their charge and morphology and a quantitatively by their β-sheet content allows to identify a structure-activity relationship. An overall positive zeta-potential, and a high monomer to β-sheet rich fibril conversion have been found to correlate with infectivity enhancement. These

properties are more pronounced for peptides with regular alternating amphiphilic patterns (KIKIQI instead of KIKQII), with amino acids residues of high hydrophobicity (I, F instead of A) and positively charged amino acids (K instead of E). Notably, it was found that the *C*-terminal fragment "NMWQ" is not required to observe these properties and infectivity enhancement whereas the *N*-terminal cysteine significantly boosts fibril formation and activity. The sequence derivatives of EF-C emphasize the importance of structural features, *i.e.* sequence order, charge and hydrophobicity, and identify amino acids which are important (C) and irrelevant (NMWQ) for activity. Thereby, EF-C could be reduced to CKFKFQF, that shows comparable absolute activity but with reduced length of 60% compared to the original sequence (QCKIKQIINMWQ), which drastically cuts materials cost (**Figure 19C**). CKFKFQF did not only performed well in clinically relevant T-cell transduction assays (**Figure 19D**) but also provides an *N*-terminal thiol groups that can be used as reactive sites for chemical functionalization.

This study shows that designing self-assembling sequences with desired bioactivity is challenging, as a small change in sequence can drastically alter physicochemical properties on multiple length scales. Furthermore, the results suggest, that positive charge and a β-sheet rich fibrillar structure are required to bind to negatively charged viral particles and cellular membranes. However, because the investigated dataset of 31 peptides is relatively small and does not cover sufficiently diverse properties, a thorough and systematic investigation is required to make general claims for amyloid peptides as shown in **Chapter 3.3**.

This study has been published by Sieste, S., Mack, T., Lump, E., Hayn, M., Schütz, D., Röcker, A., Meier, C., Kaygisiz, K., Kirchhoff, F., Knowles, T. P. J., Ruggeri, F. S., Synatschke, C. V., Münch, J., Weil, T. in Advanced Functional Materials **2021**, 31, 2009382. https://doi.org/10.1002/adfm.202009382 and reprinted in **Chapter 6.2**, page 123.[67]
The contribution of Kübra Kaygisiz to this study was to conduct time-dependent zeta potential measurements and revise the manuscript. A detailed overview of the contributions of all authors is summarized in **Chapter 6.2**, page 123.

3.3 Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides

The peptide EF-C and a small library composed of 31 of its derivatives have shown that amyloid-type peptide fibrils can efficiently enhance retroviral gene transfer (**Chapter 3.2**). However, the underlying physicochemical codependences and structural patterns remain largely unexplored.

In this study, we address these questions and reveal fundamental requirements for infectivity enhancing peptides. Data-mining techniques can be utilized to systematically study structure-property-activity relationships and efficiently identify underlying patterns in a database.

First, we expanded the peptide library of EF-C derivatives to obtain a large collection of active and inactive peptides. The new sequences were designed by chemical intuition and knowledge gained from the sequence optimization study.[67] Specifically, sequences have been generated by truncations and extensions at the N- ("QC") and C-terminal ("NMWQ") domains, by substitutions of amino acids, point mutations by substitutions with alanine, sequence scrambling, and random sequences not related to EF-C. This approach yielded a peptide library consisting of 163 sequences with sequence lengths varying from 4 to 19 amino acids (**Figure 20A**).

In a second step, the peptides in the library were thoroughly characterized by multiple experimental and bioinformatic features across length scales. These features include for example, bioactivity, structure formation, intermolecular order and bioinformatic properties that were measured by a quantitative HIV-1 infection assay, transmission electron microscopy, zeta-potential measurements, and infrared spectroscopy or calculated from the sequence. In a next step, the data was evaluated with regard to their interdependent relationships using multi-parameter linear regression. To perform this type of property-activity correlation, it is important to use peptides with similar sequences, that differ greatly in their activity, so that the activity can be traced back to physicochemical properties and not to specific amino acid motifs. For example, we find distinct differences for the formation of fibrillar aggregates when comparing CKIKIQI (aggregating, active) and CKIKQII (non-aggregating, non active) (**Figure 20B**). Overall, this approach shows that infectivity is enhanced not by the amyloid fibrils *per se*, but by their µm-sized aggregates that are rich in β-sheet structures. Additionally, we find that this aggregation behavior is preferred for peptide fibrils with a hydrophobic surface (**Figure 20C**). We further confirm the validity of the structure-property relationship by rationally designing new active sequences based on alternating amphiphilic pattern with hydrophobic amino acids (**Figure 20D**). Alternating amphiphilic sequence pattern provide a better intermolecular packing and causes increase of β-sheet order which, in turn, correlates with bioactivity. Noteworthy, alternating amphiphilic sequences resemble highly regular zippers, conformations that render traditional amyloid structures with stability.[73,74,121] In addition, we show the broad applicability of the identified property-activity relationship with amyloid peptides not related to EF-C from different sources, both pathological and functional origin that follow this property–activity relationship (**Figure 20E**).

By combining complementary physicochemical techniques that characterize peptides across length scales from nano to micrometers, we circumvent the individual scale limitations of each method and show that even for relatively small datasets universal features and patterns underlying a certain bioactivity, *i.e.* infectivity enhancement can be extracted. Data-mining not only creates a deeper understanding of these peptides but also enables the rational design of peptide sequences for therapeutic applications. On a broader context, our findings show that fibril surface hydrophobicity is driving both virus-cell interaction and fibril-fibril µm-sized aggregation. The hydrophobic contributions may serve as a common ground for cell-interactions of amyloid nanomaterials beside charge-driven interactions. We believe that our study lays the experimental foundation for a comprehensive understanding of this significant material class. In summary, this work demonstrates that fundamental insights into self-assembly processes and promising design strategies of functional nanomaterials can be elucidated by the powerful combination of physicochemical and computational methods.

Data-mining performs well at rational screening for structure-activity property relationship if the datasets sequences are similar but their corresponding properties different. However, this approach cannot be applied to design bioactive peptides that are sequence-wise not related to

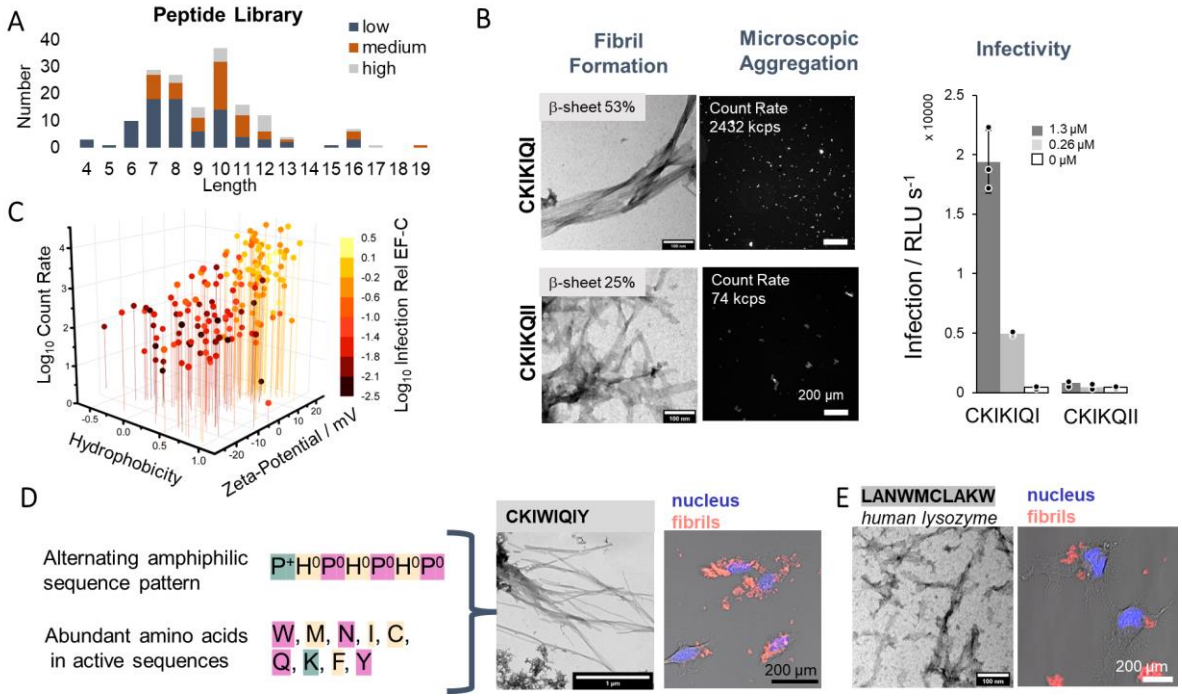


Figure 20 Overview on main findings of data mining infectivity enhancing peptides. **A** Amino acid length distribution of the peptide library consisting 163 peptides based on EF-C. The colors indicate low, medium and high infectivity enhancement. **B** Comparing physicochemical and bioactivity of peptides CKIKIQI and CKIKQII exemplify the importance of µm-sized fibrillar aggregation for activity. Overview shows from left to right transmission electron micrographs, scale bar 100 nm, fluorescence microscopy images, scale bar 200 µm and β-galactosidase infection readout (RLU, relative light units). **C** Multi-parameter plot showing correlation of $\log_{10}$ infection enhancement (color scale) with $\log_{10}$ count rate of scattered light, calculated hydrophobicity, and zeta-potential. **D** Pattern analysis reveals an alternating amphiphilic sequence pattern and abundant amino acids in infectivity enhancing peptides (H^0 hydrophobic, P^0 hydrophilic, P^+ hydrophilic, positively charged amino acid side chains). Retranslating this pattern and amino acids into peptides is a reliable method to find new active sequences as shown exemplarily for CKIWIQIY via electron and confocal fluorescence microscopy (scale bar 1 µm and 200 µm, respectively). **E** Amyloid-type peptide derived from human lysozyme analysed via electron and confocal fluorescence microscopy (scale bar 1 µm and 200 µm, respectively). This peptide is not related to EF-C but shares the key-properties for infectivity enhancement. Reproduced from Ref 290. These images are licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

the original peptide library. To expand the peptide sequence space, machine-learning is a suitable tool and is discussed in **Chapter 3.5**.
The identified structure-property activity relationship raises questions regarding the mechanistic origins of aggregation in peptide fibrils. Specifically, understanding why certain peptide fibrils form aggregates while others don't, requires investigating the fiber–fiber interactions in greater detail. For these interfibrillar interactions, assembly dynamics, and fibril-water as well as fibril–fibril interactions on very short timescales (~ns) have to be considered that go beyond the capabilities of experimental methods. Therefore, we performed molecular dynamics simulations to bridge the gap between structural information and the physicochemical properties in **Chapter 3.4**.

This study has been published by Kaygisiz, K.; Rauch-Wirth, L.; Dutta, A.; Yu, X.; Nagata, Y.; Bereau, T.; Münch, J.; Synatschke, C. V.; Weil, T. in Nature Communications, 2023, accepted, DOI: 10.1038/s41467-023-40663-6 and reprinted in **Chapter 6.3**, page 173.[290]

The contribution of Kübra Kaygisiz to this study was conceiving the idea of multi-parameter and multi-scale investigation of a structure-property-activity relationship. She designed and carried out all experiments except for the assays related to virus transduction and sum frequency generation spectroscopy. Kübra Kaygisiz analyzed all data and performed the formal analysis, wrote the python scripts for amino acids analysis, and wrote the original draft. A detailed overview of the contributions of all authors is summarized in **Chapter 6.3**, page 173.

3.4 Unraveling Structure-Property Relationships in Self-Assembling Peptide Nanofibrils via Simulations

Fibril-fibril interaction causes µm-sized aggregation and was found to be a prerequisite for fibril-virus-cell interactions (**Chapter 3.3**). Understanding these relationships is not only important to design new therapeutic materials but also to shed light on naturally occurring amyloid fiber aggregation, *i.e.* in neurodegenerative diseases. However, the molecular determinants for lateral fiber aggregation are experimentally inaccessible because they happen on multiple length scales and in very short time frames.

To break down the mechanistic reasons for macroscopic aggregation of supramolecular fibrils, we performed molecular dynamics simulations of selected peptides from our previous works using coarse-grained Martini[291] and atomistic CHARMM36 force fields[292] (**Chapter 3.2** and **Chapter 3.3**).

We were most interested in peptides showing a similar sequence but different morphology and aggregation behavior. To study this, a two-step process was applied. In the first step, the focus was on modeling the self-assembly of monomeric peptides into fibrils. Various aspects were evaluated, including the differences in fibril radius, peptide mobility within a fiber using root mean square fluctuation, fibril hydrophobicity through solvent accessible surface area analysis, and fibril stability by assessing β-sheet content. In the second step, coarse-grained fibrils were positioned in close proximity, and the potential interactions between fibrils were examined by monitoring changes in fiber-fiber distance over time, which gives us insight on fibril aggregation observed in experiments. For the evaluation of the self-assembly process coarse grained Martini models were used because the self-assembly processes require long simulation times (~ns) and involve many monomers (*i.e.* 80 peptides in 1000 nm^3). The coarse-grained models were converted back into an all-atom representation to examine the intermolecular order (β-sheet content) in an existing fiber, as well as the hydrogen bonding interactions between peptide fibrils and water. This approach ensures a computationally efficient workflow while maintaining atomistic accuracy.

Thereby, we aim to find mechanistic explanations for observations made by experiments regarding: **(I)** why the presence of *N*-Terminal cysteine improves fibril formation and activity whereas activity is almost lost for *N*-Terminal methionine substituted peptides, **(II)** how an alternating amphiphilic sequence domain (“KIKIQI”) differs from the original central domain (“KIKQII”) in terms of self-assembled fibril structure (see **Chapter 3.2** and **6.2**), and **(III)** which molecular features facilitate fibril-fibril aggregation (see **Chapter 3.3 and 6.3**).

In summary, we find that **(I)** a moderately hydrophobic *N*-terminal residue, such as cysteine weakens peptide-water interactions and thus promotes self-assembly, while using a strong hydrophobic residue such as methionine causes a hydration frustration due to the unfulfilled tendency of methionine to be buried in the center of the assembly. This hydration frustration results in misalignment of the structure, which is the origin of less β-sheet structures within the fibrils, thus leading to a lower stability and activity (**Figure 21A**). **(II)** Retaining an alternating order of hydrophobic and hydrophilic residues within the central domain increases the stability of β-sheet self-assembled fibrils. Hydrophobic amino acids in the central domain such as

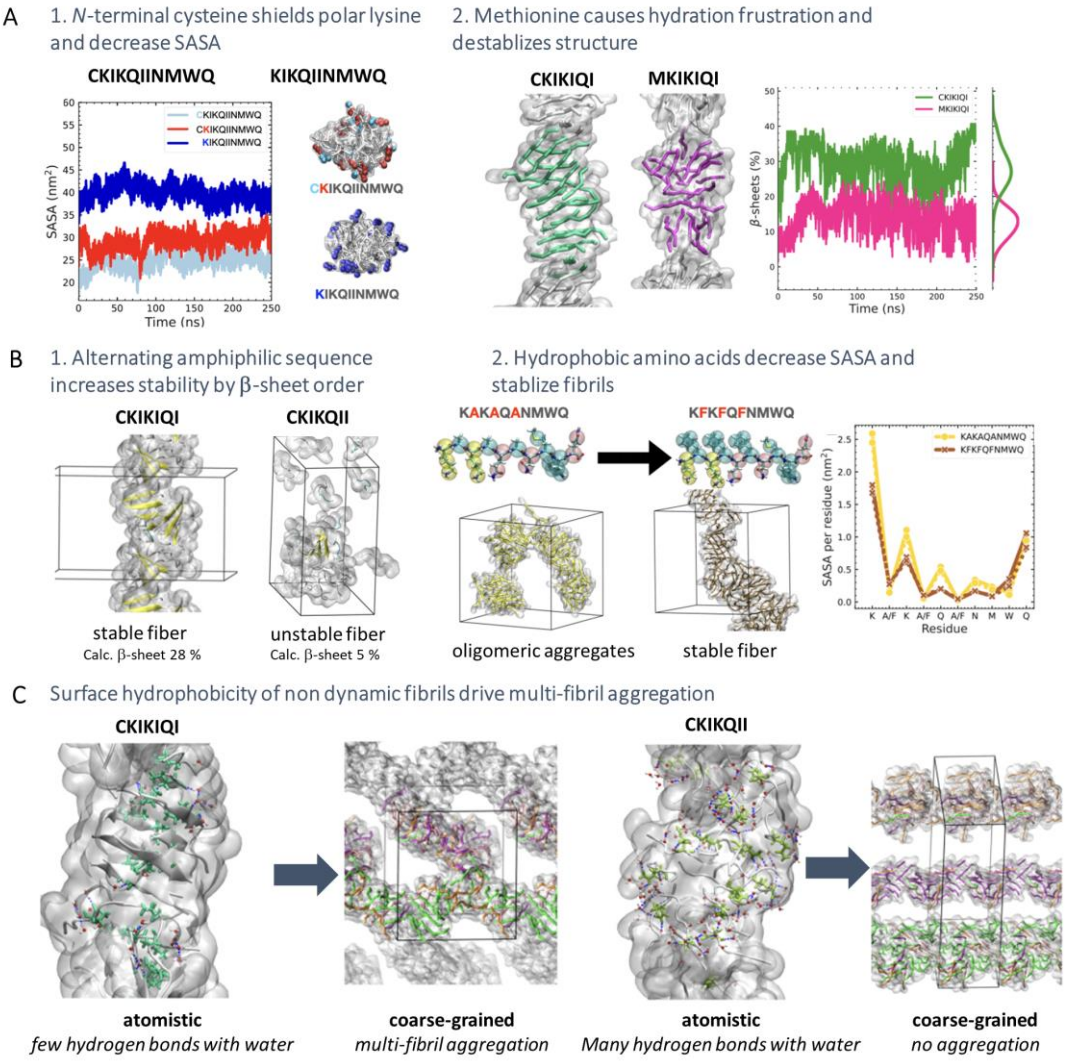


Figure 21 Main findings from molecular dynamics simulations demonstrated with selected peptides. **(I)** The role of *N*-terminal cysteine is shown by comparing solvent accessible surface area (SASA) of CKIKQIINMWQ and KIKQIINMWQ. Cysteine shields the polar lysine at 2nd position from water interactions shown by low SASA for CKIKQIINMWQ, and thereby stabilize the structure. Noteworthy, if the hydrophobicity of *N*-terminal substituent is too large, *i.e* for methionine, the nanostructure is less stable due to hydration frustration as shown by decreased β-sheet content for MKIKIQI. **(II)** In the central core domain, it is important to have an alternating amphiphilic sequence order to reduce SASA and increase β-sheet content, which stabilizes the fibril, as shown for CKIKIQI and CKIKQII. Increasing hydrophobicity of amino acids side chains (*i.e.* A → F) decreases SASA and further stabilizes the self-assembled structure. **(III)** Multi-fibril aggregation is determined by the water-interaction of the fibril surface as exemplarily shown for CKIKIQI that only forms few hydrogen bond interactions with water and has a low SASA (not shown) thus interacting well with other fibrils to a multi-fibril network. In contrast, CKIKQII has a high SASA (not shown) and forms many hydrogen bonds with water, hence it is well-hydrated and observed mostly as distinct fibrils. Figures are extracted from a *manuscript to be submitted* by Iscen, A.; Kaygisiz, K, Synatschke, C. V., Weil, T., Kremer, K. and shown in **Chapter 6.4**.

phenylalanine (F) and isoleucine (I) decrease the solvent accessible surface area and stabilize the structure better than less hydrophobic amino acids like alanine (A) (**Figure 21B**). **(III)** The quality of fiber-fiber interactions is determined by the self-assembly dynamics and fibril solvent accessible surface area, *i.e.* surface hydrophobicity. Fibrils that are dynamic (less stable, high radius of gyration, low β-sheet content) and interact well with water (high solvent accessible surface area), tend not to interact with other fibrils and are thus unlikely to aggregate (**Figure 21C**).

In summary by performing molecular dynamics simulations for many peptides with similar sequences but different physicochemical properties, we show how amino acid hydrophobicity and position affect morphology and biological properties of the fibrils on multiple length scales and can extract generalizable mechanistic rules for aggregation of amyloid-type structures: When peptides possess lower aggregation propensities, the self-assembled fibrils show lower tendency to form β-sheet rich structures and do not aggregate with other fibers in solution. In other words, increasing the mobility of the peptides within the supramolecular fibril, *i.e.* exchanging peptides within the fibril, and the attractive interactions with water molecules decreases fibril-fibril aggregation. In turn, peptides with substantial β-sheet content and low mobility exhibit the highest level of aggregation and thereby biological activity for retroviral gene transfer. Importantly, each finding from the molecular dynamics simulations is well in line and validated by experimental data on the morphology (electron microscopy), β-sheet content (ATR-FT-IR) and colloidal aggregation (light scattering). Thereby, we unexpectedly find that the dynamics of the individual fibrils are different than of fibrillar aggregates, which highlights the importance of multi-scale simulations. We believe that this will have future implications for considerations in supramolecular complex systems.

This study demonstrates that molecular features propagate across length scales and determine the fate of mesoscopic aggregates. These insights are relevant for optimizing molecular design, *e.g.* for controlling microscopic properties in therapeutic adjuvants.

This study is part of a *manuscript to be submitted* by Iscen, A., Kaygisiz, K., Synatschke, C. V., Weil, T.; Kremer, K. and shown in **Chapter 6.4**, page 265.

The contribution of Kübra Kaygisiz to this work is the conceptualization and selection of peptides for the simulations. She performed the experimental work and revised the manuscript. A detailed overview of the contributions of all authors is summarized in **Chapter 6.4**, page 265.

3.5 Inverse Design of Viral Infectivity-Enhancing Peptide Fibrils from Continuous Protein-Vector Embeddings

The discovery of new bioactive self-assembling peptides via traditional methods such as screening and rational design as demonstrated in **Chapter 3.2** and **Chapter 3.3** is challenging because of the multi-scale and multi-parameter complexity that is inherent to supramolecular assembly processes. At the same time, research on machine-learning assisted drug discovery has gained tremendous attraction in recent years. While big technological companies such as DeepMind, Google, and Microsoft are working on predicting protein structures and functions, there is much less knowledge on the digital discovery of supramolecular materials based on short peptides. The main reasons for this disparity is data availability as there are large databased for proteins but not for small self-assembling peptides.

Hence, a data driven approach for peptide discovery requires a solid base of experimental data. For the current objective we use the sequence and bioactivity information of the previously established library (**Chapter 3.3**), that is composed of 163 fully characterized sequences regarding their physicochemical properties and bioactivity, in the form of infectivity enhancement. Our aim was to predict *de novo* minimalistic 6-mer sequences whose sequences are not related to the original data set, but that still self-assemble and display bioactivity.

To this end, we applied an inverse-engineering method based on natural language processing in a three-step machine-learning approach. First, we converted the peptide sequences in a universal machine-readable way to train a regression model. For this, we chose the continuous vector representations from 3-mers that have been developed by Asgari and Mofrad[285] and that is known to retain important bioinformatic information encoded within the sequences (**Figure 22A-1**). We trained a linear regression model with these sequence vectors together with the corresponding experimental infectivity data. The established supervised machine-learning model was cross-validated on the training set and performed well in predicting infectivity for a given sequence from the training set (Pearson correlation coefficient R = 0.84). Secondly, after establishing a suitable model we explored the sequence space of 6-amino acid peptides via Monte Carlo for cost-efficient screening (**Figure 22A-2**). We chose 6-mer sequences because they represent the shortest peptides reported for infectivity enhancement to date and are not part of the training set. Consequently, discovery of active 6-mer peptides by traditional methods is especially challenging. By applying the trained model on peptide sequences generated via 16 independently performed Monte Carlo sampling runs we select 16 cationic peptides that have been predicted to be highly active.

The selection was further narrowed down by aggregation prediction. Eight of the 16 peptides were also predicted to aggregate by computational tools such as Aggrescan (**Figure 22A-3**). The experimental characterization of the 16 peptides revealed that 25% of the total cohort and 50% of those for which aggregation was predicted enhanced retroviral infectivity. Notably, these newly designed sequences are the shortest reported active peptides for enhancing infectivity reported to date, and bear no similarity in their sequences to the training set as verified by N-gram similarity scores.

This machine-learning approach allowed us to expand the sequence space in a direction that would have been unintuitive from the perspective of traditional rational design. For example, the model identified hydrophobic peptides containing large amounts of W and C as promising

candidates. Strikingly, peptides containing these amino acids are the first hydrophobic peptide fibrils with a negative surface charge that can enhance infectivity (**Figure 22B**). This is theoretically predicted from the property-activity relationship equation found in **Chapter 3.3** for the training set (**Figure 22C**), however, was so far not found experimentally.

One advantage of employing a continuous vector representation-based approach is the ability to encode both sequence and physicochemical information of a peptide into a numerical vector. The encoded vectors can be utilized in a simple way to train a machine learning (ML) model. In contrast to other methods such as deep learning, the vector embeddings apply a linear regression model that ensures that the trained model is in principle explainable. Furthermore, by employing Monte Carlo sampling we can efficiently screen a vast sequence space, saving time and lowering costs. This screening process allows us to identify novel active peptide sequences, which exhibit unexpected structural and property characteristics that differ significantly from the training set.

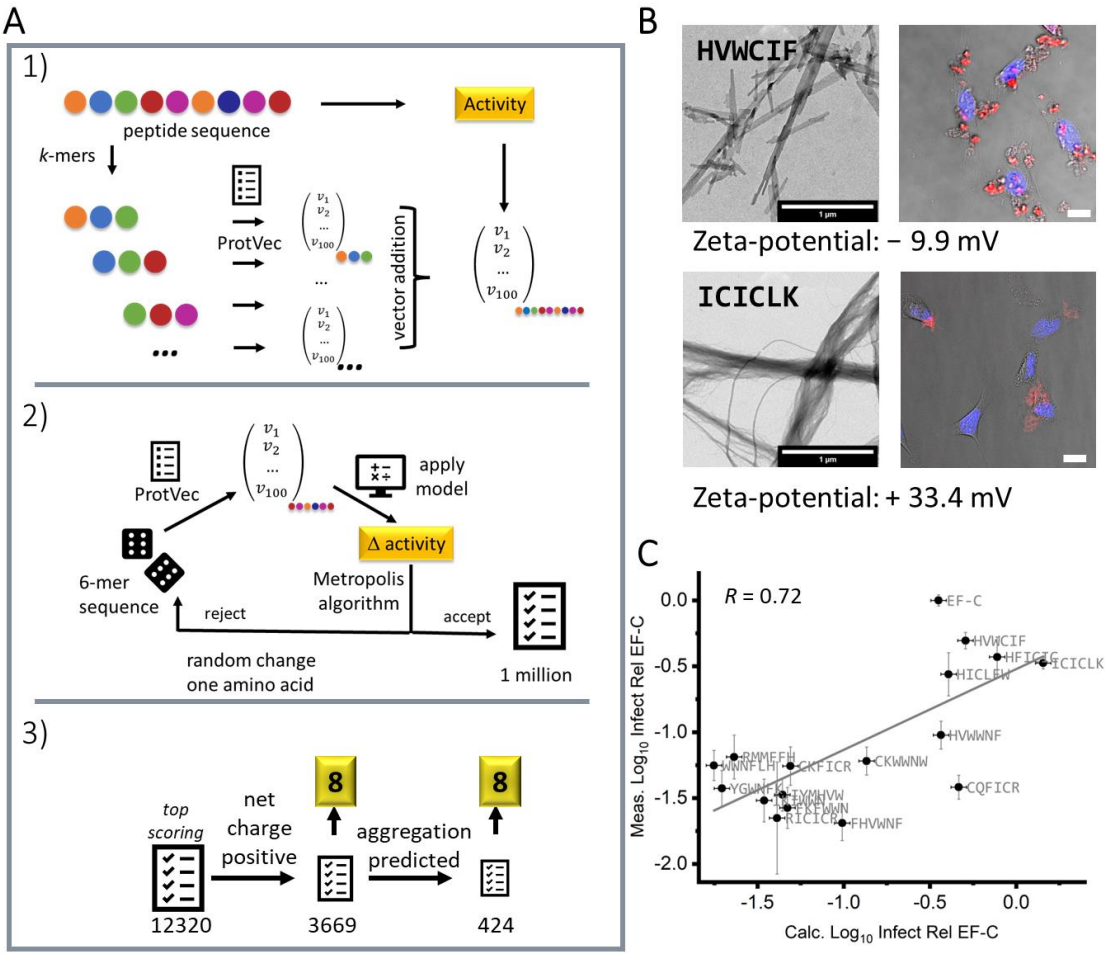


Figure 22 Summary of a machine-learning assisted peptide discovery. **A** Schematic summary of the 3-step prediction process employed to find *de novo* 6-mer peptides for infectivity enhancement. 1) Vector representation of experimentally tested peptide sequence by creating 3-mers and assigning them to ProtVec library.[285] 2) Supervised machine-learning linear regression model applied to Monte Carlo generated 6-mer sequences. 3) Preselecting Monte Carlo predicted peptides by charge and aggregation. 16 *de novo* 6-mer peptides have been experimentally tested. **B** Exemplary peptides that were tested positive for enhancing infectivity. Both peptides form fibrils (transmission electron microscopy, scale bar 1 µm) and colocalize with cell membranes (confocal fluorescence microscopy, scale bar 20 µm) but show different Zeta-potentials of opposite signs. **C** Structure-property activity relationship found from data-mining (**Chapter 3.3**) applied on *de novo* peptides show good correlation (R = 0.72) and indicate a similar mode of action in a new sequence space. Reproduced from Ref 293. These images are licensed under CC BY 3.0 (https://creativecommons.org/licenses/by/3.0/).

Overall, this study represents the first ML approach for the prediction of bioactive self-assembling peptides and demonstrates the powerful capabilities of natural language processing methods. Thereby, this study provides a rapid and cost-efficient method for screening self-assembling peptides for a desired type of bioactivity. In the future, our method has the potential to speed up *de novo* peptide discovery in the early stages of research.

This study has been published by Kaygisiz, K.; Dutta, A.; Rauch-Wirth, L.; Synatschke, C. V.; Münch, J.; Bereau, T.; Weil, T. in Biomaterials Science, 2023, 11 (15), 5251–5261, https://doi.org/10.1039/D3BM00412K and reprinted in **Chapter 6.5**, page 351.[293]

The contribution of Kübra Kaygisiz to this study was to conceptualize, design and carry out all the experimental work except for the virus transduction and cell-viability assays. She analyzed all the data, performed the formal analysis, wrote python scripts for amino acids analysis, performed and analyzed the aggregation prediction, and wrote the original draft. A detailed overview of the contributions of all authors is summarized in **Chapter 6.5**, page 351.

3.6 Peptide Amphiphiles as Biodegradable Adjuvants for Efficient Retroviral Gene Delivery

To enhance retroviral transduction efficiency in *ex vivo* gene therapy settings, adjuvants are required. At the same time, these adjuvants must also possess the ability to be easily eliminated or degraded after transduction to mitigate the risk of venous embolism when reintroducing the transduced cells into the patient's bloodstream. While self-assembling peptides outcompete traditional polymeric adjuvants in their efficiency[154] and would be ideal candidates as adjuvants in clinical applications, they also form potentially harmful μm-sized aggregates (**Chapter 3.3**). To date, no peptide system with efficient biodegradability has been reported in this context.

Based on the lessons learned on from our previous structure-property activity relationship studies (**Chapter 3.3** and **Chapter 3.4**), we designed self-assembling peptides that are efficient enhancers of retroviral transduction while being degradable by cells within the time frame of application.

Of particular interest for the design of these sequences are optimized surface properties that must find the sweet spot of providing enough fiber-fiber aggregation and to a certain degree dispersion of fibrils. Microscopic aggregation is necessary for the capture and delivery of viruses as observed in the structure-property-activity studies above (**Chapter 3.2** and **Chapter 3.3**), however when aggregates are too large this can counteract *in vitro* degradability. In **Chapter 3.4** we found that fibrils with less dynamic structure and high surface hydrophobicity tend to aggregate. To finetune the surface properties of fibrils, it is advantageous to use a design that enables flexible control of the exterior of the peptide fibril surface while maintaining stability of the fibrillar structure. These prerequisites are provided by peptide amphiphiles (PAs). PAs consist of a hydrophobic part, typically a fatty acid residue, and a hydrophilic peptide part. Certain peptide sequences that contain β-sheet prone amino acids can induce a fibrillar morphology. Because the hydrophobic moiety always tends to escape water contact by being buried in the center, it stabilizes the fibrillar structure. This allows great flexibility in modulating the peptide part that faces towards the fibril surface.

Therefore, we selected PAs to study how different sequences and hydrophobic moieties affect fiber formation, aggregation, virus infectivity, and cell degradability. To this end, we created a PA library consisting of 35 different compounds that are derived from the well-established PA C_{16}-VVVAAAKKK-NH_2 (**Figure 23A**).[288]

First, we find that C_{16}-VVVAAAKKK-NH_2 shows high activity as a transduction enhancer comparable to the commercially available gold standard adjuvants RetroNectin, Vectofusin-1 and EF-C (**Figure 23B**). The PAs in the library enhance viral transduction efficiency if they form β-sheet fibrils and aggregate into positively charged μm-sized clusters. They facilitate the adhesion of the virus to the cell membrane and thereby show a similar mode of action and structure-property activity relationship as reported for amyloid-type peptides for viral transduction. However, their fundamentally different constitution gives rise to distinct aggregation properties. Micrometer sized aggregates of an area below 10 μm^2 are enhancing infectivity while being degraded much more efficiently compared to larger aggregates. The organization and hydrophilicity of amino acids within the peptide sequence are key factors in

governing the morphologies of aggregates and biodegradability of the PAs. In particular, the presence of block-type amphiphilic peptide motifs (*i.e.* VVVAAAKKK-NH_2) in the PAs ensures effective interaction between cells and viruses, while also enabling biodegradability. Conversely, alternating amphiphilic peptide sequences (*i.e.* FQFKFKC-NH_2) result in the formation of large aggregates that are not degraded by cells during the transduction experiment (**Figure 23C**).
Similar to the findings for amyloid-type fibrils (**Chapter 3.4**) we applied coarse grained molecular dynamics simulations to study molecular features underlying multi-fibril aggregation. We find that the molecular origins of reduced aggregation in some peptides can be traced back to a more dynamic self-assembled structure (high root mean square fluctuation) and higher solvent accessible surface area (less hydrophobicity). In other words, PA's multifiber aggregation can be influenced by fiber dynamics, that is, in turn, influenced by choosing the sequence, *i.e.* increase of charged amino acids at the *C*-terminus, including small flexible amino acids close to the *N*-terminus, and flexible, unsaturated fatty acids.

With the extensive study of the PA library, we can validate these findings on multifiber aggregation with a variety of different compounds and find besides to the well-established PA C_{16}-VVVAAAKKK-NH_2, five new PAs that are: eicosapentaenoic acid-VVVAAAKKK-NH_2, C_{16}-*thioester*-CVVVAAAKKK-NH_2, C_{16}-*ester*-SVVVAAAKKK-NH_2, C_{16}-ALAAGKK-NH_2, and C_{16}-AKAVGK-NH_2 as potential adjuvants with high *in vitro* degradability and high transduction efficiency.

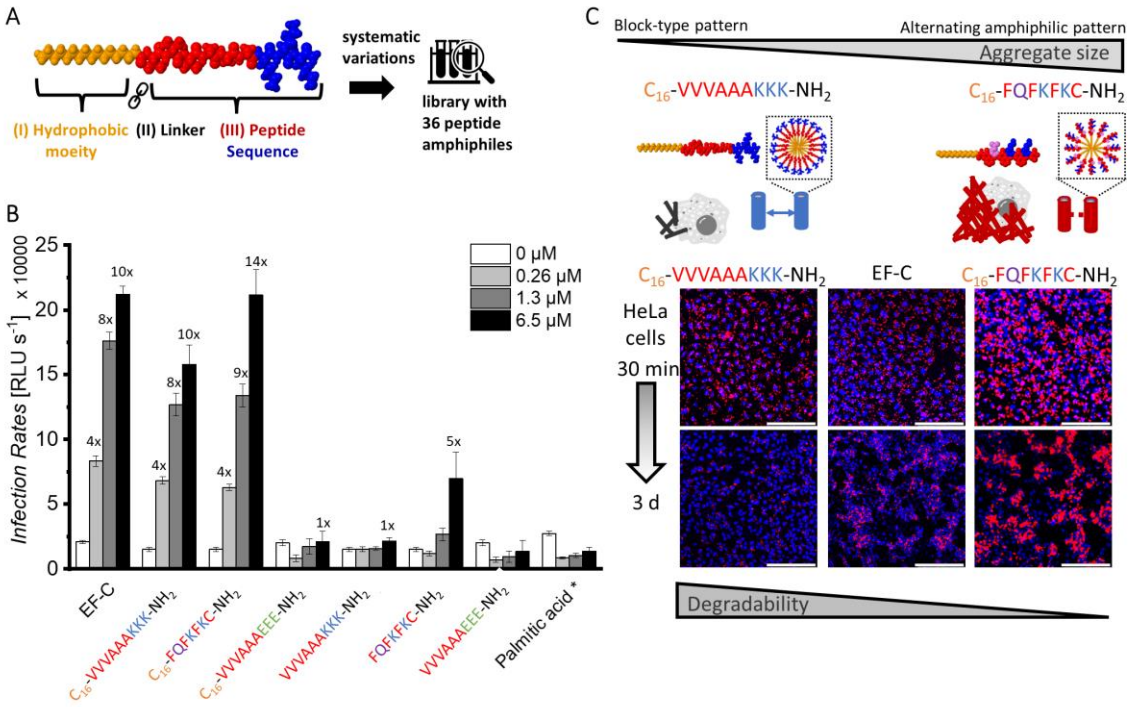


Figure 23 Main findings of applying peptide amphiphiles as biodegradable transduction enhancers. **A** Schematic representation of the strategy to create a library based on the peptide amphiphiles C_{16}-VVVAAAKKK-NH_2 by systematic variations of the hydrophobic moiety, the linker chemistry, and the peptide sequence. **B** Infectivity enhancement evaluated by β-galactosidase assay, showing that peptide amphiphile C_{16}-VVVAAAKKK-NH_2 is comparable to EF-C regarding its activity (RLU, relative light units). **C** Biodegradability from HeLa cells *in vitro* is depending on the aggregate size that is, in turn, determined by the sequence design. Block-type sequences such as in C_{16}-VVVAAAKKK-NH_2 form much smaller aggregates compared to an alternating amphiphilic pattern such as in C_{16}-FQFKFKC-NH_2. Figures are extracted from a *manuscript in revision* by Kaygisiz, K.; Rauch-Wirth, L.; Iscen, A.; Hartenfels, J.; Kremer, K.; Münch, J.; Synatschke, C. V.; Weil, T. and shown in **Chapter 6.6**.

Our study establishes PAs as promising additives in retroviral gene transfer that combine the best of two worlds: high transduction efficiency comparable to commercially available adjuvants and good biodegradability.

This study is part of a *manuscript in revision* by Kaygisiz, K.; Rauch-Wirth, L.; Iscen, A.; Hartenfels, J.; Kremer, K.; Münch, J.; Synatschke, C. V.; Weil, T. and shown in **Chapter 6.6**, page 385.
The findings of this study have been registered for European patent application under the number EP23171438.7.

The contribution of Kübra Kaygisiz to this study was to conceptualize the study and conceive the idea for infectivity enhancement and biodegradability of PAs. She carried out all the experiments related to peptide design, synthesis and characterization and *in vitro* peptide degradability, analyzed all data, performed the formal analysis, and wrote the original draft. A detailed overview of the contributions of all authors is summarized in **Chapter 6.6**, page 385.

3.7 Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning

Designing peptide sequences that can be degraded easily by cells after their successful applications (**Chapter 3.6**) is not only interesting for therapeutic *ex vivo* virus delivery, but also for tissue engineering applications such as applied in recent gene therapy techniques.[11]
Based on the systematic analysis of our peptide libraries (**Chapters 3.2** – **Chapter 3.4**) we know that β-sheet rich peptides are necessary for cell-interaction. Accordingly, decomposing β-sheet structures should decrease cell-interactions. Since they show a high intrinsic stability, breaking down amyloid-type β-sheet structures is a difficult endeavor. It is even more challenging to do this by selective disassembly of specific structures with spatial resolution.

Here, we present a chemically engineered amyloid-type peptide fibril that can be decomposed in a controlled fashion via an external trigger. To create such an adaptive peptide material, we selected a photo-triggered mechanism because this allows for spatiotemporal control with high precision. Specifically, we incorporated a well-established photocleavable *o*-nitrobenzyl linker in a sequence optimized derivative of EF-C, CKFKFQF, that forms peptide fibrils and has been successfully applied as surface coatings on substrates to stimulate neuronal cell growth[145] and instruct cell attachment.[178] By including the photocleavable linker in the peptide backbone (CKFK-linker-FQF, **Figure 24A**) we can fragment the amphiphilic peptide upon exposure to

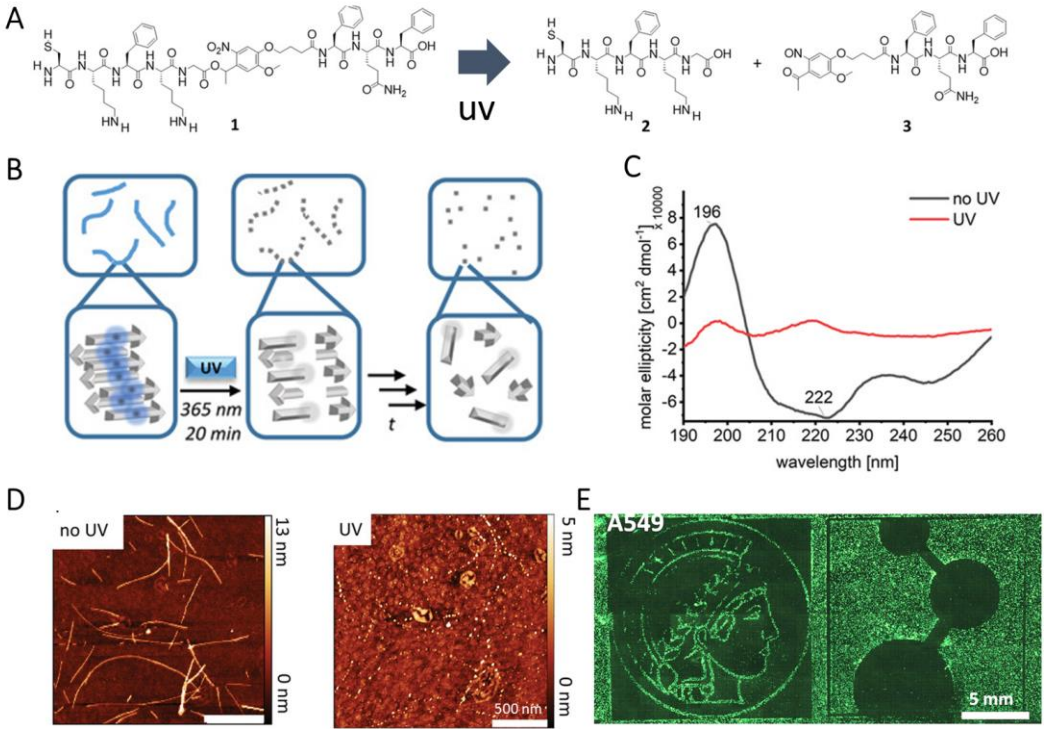


Figure 24 Main findings of utilizing photo-degradable peptides for controlled cell attachment. **A** Molecular structure of the peptide CKFK-linker-FQF (1), that can be cleaved in two fragments (2, 3) upon irradiation with UV-light. **B** Schematic demonstration of the effects of photo-cleavage on self-assembled peptide fibril morphology. Irradiation destabilizes the fibril and results in decomposition. **C** Circular dichroism spectra of peptide secondary structure before and after irradiation that shows loss of β-sheet conformation. **D** Atomic force microscopy in solution shows loss of morphology after *in situ* irradiation, scale bar 500 nm. **E** Fluorescence microscopy image of A549 cells stained with calcein (green) attaching only to non-irradiated peptide fibrils. The irradiation was performed with a photomask after coating a substrate with the peptide fibrils (1), scale bar 5 mm. Reproduced from Ref 294. These images are licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

light, into a charged, polar (CKFK) and a hydrophobic, uncharged (linker-FQF) part. This fragmentation on the molecular scale destabilizes the self-assembled structure on a mesoscopic scale (**Figure 24B**).
We found that the peptide CKFK-linker-FQF, like the unsubstituted CKFKFQF, can form amyloid-like fibrils. However, the CKFK-linker-FQF peptide exhibits a strong response to UV-light. Upon UV irradiation, the fibril undergoes structural changes, losing its secondary structure, fibril morphology, and bioactivity as shown via infrared spectroscopy, circular dichroism (**Figure 24C**), electron microscopy, atomic force microscopy (**Figure 24D**), Thioflavin-T fluorescence staining and cell attachment assays. By exposing surface coatings of these fibrils to UV-light through a photomask well-defined pattern of intact and decomposed fibrils can be created. The areas that remain unexposed retain their capacity to support cell adhesion in culture, in contrast to the regions exposed to light where the supportive fibril morphology for cell attachment is disrupted (**Figure 24E**). We validate the molecular fragmentation on the surface by direct detection of intact and fragmented molecules via MALDI-ToF-MS Imaging and we visualize the loss of β-sheet rich fibril morphology by staining the surface with an amyloid sensitive dye and subsequent fluorescence microscopy. Neither the irradiated nor the non-irradiated species are toxic to cells at the conditions we applied as evaluated by cell viability assays. Consequently, this study demonstrates the potential of photoresponsive peptide nanofibrils for arranging single cells precisely on surfaces, as exemplified with A549, Chinese Hamster Ovary, and Raw Dual type cells.
Overall, we demonstrate that photoresponsive amyloid-like fibrils can be applied as adaptable functional materials for the precise organization of cells on surfaces. This study demonstrates that that β-sheet rich structures are directly connected to cell-interactions which is well in line with previous findings for viral infectivity enhancement (**Chapter 3.2** – **Chapter 3.6**) and opens new avenues for other application settings, *e.g.* for tissue engineering. The ability to have spatiotemporal control over the structural integrity of these matrices enables the creation of tailored and customizable four-dimensional cell culture in a dynamic environment.

This study has been published by Kaygisiz, K.; Ender, A. M.; Gačanin, J.; Kaczmarek, L. A.; Koutsouras, D. A.; Nalakath, A. N.; Winterwerber, P.; Mayer, F. J.; Räder, H.; Marszalek, T.; Blom, P. W. M.; Synatschke, C. V.; Weil, T. in Macromolecular Bioscience, 2023, 23 (2), 2200294. https://doi.org/10.1002/mabi.202200294 and reprinted in **Chapter 6.7**, page 463.[294]

The contribution of Kübra Kaygisiz to this study was conceiving the idea of photo-induced fibril decomposition and conceptualization of the study. She designed and carried out experiments related to linker synthesis and characterization and peptide design, synthesis, and characterization. Further, she analyzed data and performed the formal analysis. She prepared parts of fibril and agarose coatings, performed and analyzed parts of the irradiation studies and cell-patterning assays, and wrote the original draft. Kübra Kaygisiz and Adriana M. Ender contributed equally to this work. A detailed overview of the contributions of all authors is summarized in **Chapter 6.7**, page 463.

4. CONCLUSION AND FUTURE PERSPECTIVES

Novel gene therapeutic strategies are a key component of personalized medicine. However, when it comes to accessibility for a larger population they fall short compared to vaccines, for instance. One of the reasons for this is due to the costly manufacturing processes of the tailor-made materials that are required. To increase accessibility, new cell engineering techniques and chemical strategies are necessary that establish suitable matrices for efficient cell-material interactions.

Amyloid fibrils have been exclusively studied in pathogenic contexts for a long time. Recently, their potential as functional materials have been discovered and explored for a variety of biotechnological applications. In this thesis I demonstrate a comprehensive understanding of

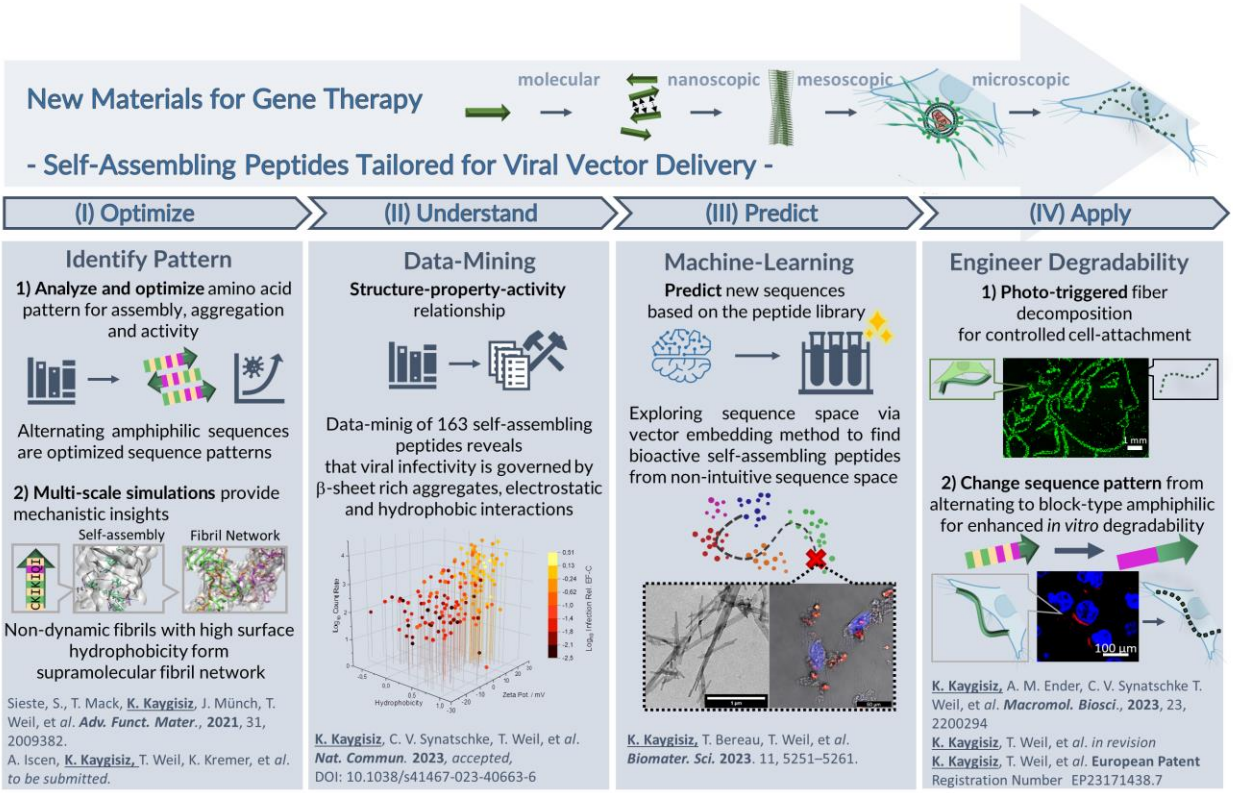


Figure 25 The aim of this thesis is to create new materials for gene therapy by using self-assembling peptides that promote viral gene delivery. A summary of the main findings of this thesis is shown. **(I) 1)** Creating a peptide library based on self-assembling peptide enhancing factor C (EF-C) identify a sequence optimized peptide and find that β-sheet rich peptide fibrils correlate with infectivity. **2)** Applying multi-scale simulations provide a mechanistic insight in self-assembly process and demonstrate that a more regular alternating amphiphilic sequence order form less dynamic fibrils with higher surface hydrophobicity that lead to fibril-fibril interactions. **(II)** Extensive data-mining of sequence properties of 163 peptides in a EF-C based library shows that viral infectivity is governed by µm-sized aggregation of β-sheet rich structures. These properties are predominantly found in alternating amphiphilic sequences with high fractions of hydrophobic amino acids. **(III)** Using this peptide library in a vector embedding machine-learning approach, new sequence spaces of self-assembling bioactive short peptides are identified. Electron microscopy scale bar 1 µm, Confocal fluorescence microscopy, fibrils (red), cell nucleus (blue), scale bar 50 µm **(IV)** Finally, **1)** applying chemically engineered photo-degradable peptide sequences enables spatiotemporal control over β-sheet morphology and cell-attachment. Fluorescence microscopy, cells (green), scale bar 1 mm. **2)** Applying peptide amphiphiles with block-type hydrophobicity not only promotes infectivity but also gives access on modulation of fibrils surface properties and control over aggregation properties. It is found that i*n vitro* cell degradability is enhanced if peptide amphiphiles are engineered to form small aggregates.

cell-material interactions and apply these insights to create new materials from the molecular level (peptide sequence design) to mesoscopic assembly properties (fibril formation), to microscopic features (µm-sized aggregation) (**Figure 25**).

In summary, I could demonstrate that the sequence-optimization and data-mining approach combined with the molecular dynamic simulations enables understanding of the emergence of hierarchical structure from supramolecular interactions of monomers across several length scales. Short peptides present a distinct advantage in effectively deciphering molecular sequence contributions (**Chapter 3.2**) and allow us to identify alternating amphiphilic sequence patterns with a high fraction of hydrophobic amino acids as a common sequence motif for the formation of β-sheet rich peptide fibrils (**Chapter 3.3**). The stability of the fibrillar assembly and hydrophobicity of the fibril surface regulates fibril-fibril interactions that are the origin of µm-sized aggregate formation (**Chapter 3.4**). A multi-parameter correlation reveals that these structural features are essential for promoting cell-virus interactions and thereby enhance retroviral transduction (**Chapter 3.3**). Without these key-structural features cell-material interactions are less efficient. We demonstrated the loss of cell-adhesiveness by decomposing β-sheet rich structures in a controlled manner by fragmentation of an alternating amphiphilic peptide using photochemistry (**Chapter 3.7**). For advancing clinical application of peptide fibrils, we evaluated peptide amphiphiles as a new material class of degradable transduction enhancers. Peptide amphiphiles show viral gene delivery efficiency comparable to the gold-standard EF-C. However, in contrast, they can be significantly better degraded by cells within the timeframe of their application (**Chapter 3.6**). This opens new avenues for clinical applications as the limited biodegradability of EF-C posed a major challenge for translational applications. In early stages of clinical research typically a small compound library of similar sequences is available. We demonstrate that utilizing a vector embedding based machine-learning approach can expand the sequence space of bioactive peptides in a time and cost-efficient way (**Chapter 3.5**).

While this interdisciplinary work primarily focused on peptide materials for viral gene delivery, the findings from these studies have implications on several research fields such as **(I)** amyloid misfolding, **(II)** material engineering for cells, and **(III)** clinical applications.

(I) Because of the structural complexity of proteins, it is challenging to understand the connection between molecular and mesoscale behavior in amyloid protein misfolding. The peptide fibrils studied in this thesis open possibilities to deconvolute molecular and structural properties as they exhibit distinct aggregation behavior while being similar on the molecular level. In experiments, we observe that some peptide fibrils form ordered and some form unordered fibrillar aggregates (**Chapter 6.3, p 224**). However, the precise parameters governing the different aggregation phenomena have not yet been studied from a mechanistic perspective. As µm-sized crystalline amyloidal aggregates are found in misfolded proteins, studying their emergence from single molecules up to multiple fibers could add valuable insights to understand their pathogenicity. A hypothesis worth exploring is whether differences in intermolecular interaction strengths affect the macroscopic order. For example, pathologic fibrils may form highly regular ordered fibrillar aggregates over long time frames (days to months)[97,98] due to weakly interacting monomers, while functional amyloids show strongly intermolecular interactions and rapidly aggregate[99,100] which kinetically hinders the formation of crystalline aggregates.

I believe that investigating the differences between fiber-fiber interaction may be the key to the understanding what makes pathological and functional amyloid materials different. First steps towards the importance of studying fibril-fibril interactions have been taken in this thesis. For example, molecular dynamics simulations show that the dynamics of the individual peptide fibrils is different compared to aggregated fibrils (**Chapter 3.4**). Although the dynamic nature of supramolecular peptide fibrils is one of the characteristic properties that distinguishes them from covalent polymers, there remain many unanswered questions *e.g.* how dynamic properties affect function and how peptide mobility affects biological activity. Answers to these questions are of utmost interest for future rational design of supramolecular complex systems.

(II) From a materials engineering perspective, the multi-scale complexity of sequences is intriguing, as even small changes can strongly impact their macroscopic properties. This provides an intriguing opportunity for fast and simple tunability of function using molecularly similar peptides. In the present thesis we show that the dynamics of the self-assembled structures and fibril surface hydrophobicity controls fiber–fiber aggregation across length-scales (**Chapter 3.4**). This provides a solid basis for the future application of coassembled peptides. Using a coassembly strategy is interesting as it allows a modular approach for rapid and customized functionalization of peptide fibrils, *e.g.* introducing hydrophobic and hydrophilic patches, to fine-tune the properties of the fibrils. Unpublished data from my work shows that physicochemical properties can be changed by coassembling a hydrophobic, self-assembling and a hydrophilic, non-assembling peptide. This is not only interesting from a fundamental perspective but also for biological questions: In a previous thesis by J. Müller (Ulm University) it was reported that fibril-virus attachment depends on the virus surface properties, *i.e.* glycoprotein density.[295] Coassembled peptides with fine-tuned fibril surface hydrophobicity offer exciting prospects for studying virus surface properties that may change viral gene delivery efficiency and may even enable us to target specific types of viruses.

Furthermore, short peptides that self-assemble into dynamic structures have recently been explored in the context of liquid-liquid phase separation.[296] These structures offer potential to act as traceless adjuvants, yet remain largely unexplored for gene delivery. In this thesis, the first applications for virus delivery with fibrous droplets of an unsaturated fatty acids modified peptide (**Chapter 6.6, p 446**) are reported. The development of these fibrous droplets or liquid-liquid phase separated droplets opens promising translational perspectives for viral and non-viral gene delivery in the future.

(III) One of the most exciting contemporary developments are machine-learning approaches that have the potential for designing sequences that cannot be found in Nature and thus give access to properties and biological activities beyond human imagination. Data-driven efforts in peptide self-assembly promise a quantitative understanding how molecular changes impact the functional properties of a material, which in turn can be used to predict appropriate molecular modifications to change supramolecular function. In this thesis both aspects are demonstrated – data-mining and simulations that reliably model multi-scale assembly and aggregation behavior (**Chapter 3.3** and **3.4**) as well as machine learning that predicts unexpected sequences not related to any previously known peptide (**Chapter 3.5**). This is especially relevant for drug discovery, where computational predictions offer a faster and more inexpensive exploration of the chemical space of peptides compared to traditional methods. In the future, a promising approach might be the combination of simulations with machine learning, eliminating the need

for an experimental basis to achieve accurate prediction.[297] Recent works[298] in the field have demonstrated the potential of this approach, which holds promise for expanding its application to cell-interactive materials once there is a suitable modeling basis provided. Artificial intelligence can now predict protein 3D structure[249] which only has been possible due to openly sharing protein data over the past decades. This thesis contributes to collecting and sharing experimental data for bioactive, short-self-assembling peptides, materials that have not been covered in such a systematic and extensive way before.

Finally, in a broader context, the findings presented in this thesis showcase the bottom-up design and discovery of materials for technological applications in the life-sciences. This thesis shows that structural complexity emerges from small self-assembling peptide units that result in diverse biophysical properties of the self-assembled nanostructures "*that are not present in the components of a complex mixture but come to light only as a result of interactions between molecules*".[30]

The polymorphism observed for the amyloidal peptides is perhaps an example par excellence for multi-scale complexity from the assembly process of peptide materials driven towards energetic equilibrium state. As there is a growing demand for adaptive molecular systems in the life-sciences, developing complex materials with structural control in a non-equilibrium state holds great promise for future cell-related applications.[299] Consequently, this thesis provides a step towards an understanding of materials that collectively and structurally transfer information across length scales from molecular to microscopic level, all of which are vital for the next-generation supramolecular materials.

5. REFERENCES

(1) James, C.; Devaux, M.; Sassi, F. Inclusive Growth and Health. *OECD Heal. Work. Pap.* **2017**, *103*, 48. https://doi.org/10.1787/93d52bcd-en.

(2) Roth, G. A.; Al., E. Global, Regional, and National Age-Sex-Specific Mortality for 282 Causes of Death in 195 Countries and Territories, 1980–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* **2018**, *392* (10159), 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7.

(3) UN General Assembly. Transforming our world : the 2030 Agenda for Sustainable Development, A/RES/70/1 https://sdgs.un.org/sites/default/files/publications/21252030 Agenda for Sustainable Development web.pdf (accessed Jun 5, 2023).

(4) Naldini, L. Gene Therapy Returns to Centre Stage. *Nature* **2015**, *526* (7573), 351–360. https://doi.org/10.1038/nature15818.

(5) FDA approves innovative gene therapy to treat pediatric patients with spinal muscular atrophy, a rare disease and leading genetic cause of infant mortality https://www.fda.gov/news-events/press-announcements/fda-approves-innovative-gene-therapy-treat-pediatric-patients-spinal-muscular-atrophy-rare-disease (accessed Apr 15, 2020).

(6) Reinhard, K.; Rengstl, B.; Oehm, P.; Michel, K.; Billmeier, A.; Hayduk, N.; Klein, O.; Kuna, K.; Ouchan, Y.; Wöll, S.; Christ, E.; Weber, D.; Suchan, M.; Bukur, T.; Birtel, M.; Jahndel, V.; Mroz, K.; Hobohm, K.; Kranz, L.; Diken, M.; Kühlcke, K.; Türeci, Ö.; Sahin, U. An RNA Vaccine Drives Expansion and Efficacy of Claudin-CAR-T Cells against Solid Tumors. *Science* **2020**, *367* (6476), 446–453. https://doi.org/10.1126/science.aay5967.

(7) Pearson, S.; Jia, H.; Kandachi, K. China Approves First Gene Therapy. *Nat. Biotechnol.* **2004**, *22* (1), 3–4. https://doi.org/10.1038/nbt0104-3.

(8) U.S. Food & Drug Administrations, Approval Letter - KYMRIAH (tisagenlecleucel) https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/kymriah-tisagenlecleucel (accessed Jan 30, 2023).

(9) Reardon, S. "It's a Vote for Hope": First Gene Therapy for Muscular Dystrophy Nears Approval, but Will It Work? *Nature* **2023**. https://doi.org/10.1038/d41586-023-01799-z.

(10) Sterner, R. C.; Sterner, R. M. CAR-T Cell Therapy: Current Limitations and Potential Strategies. *Blood Cancer J.* **2021**, *11* (4), 69. https://doi.org/10.1038/s41408-021-00459-7.

(11) Krystal Biotech, I. Package Insert - VYJUVEK https://www.fda.gov/vaccines-blood-biologics/vyjuvek (accessed Jun 5, 2023).

(12) Lou, J.; Mooney, D. J. Chemical Strategies to Engineer Hydrogels for Cell Culture. *Nat. Rev. Chem.* **2022**, *6* (10), 726–744. https://doi.org/10.1038/s41570-022-00420-7.

(13) Kaygisiz, K.; Synatschke, C. V. Materials Promoting Viral Gene Delivery. *Biomater. Sci.* **2020**, *8* (22), 6113–6156. https://doi.org/10.1039/D0BM01367F.

(14) High, K. A.; Roncarolo, M. G. Gene Therapy. *N. Engl. J. Med.* **2019**, *381* (5), 455–464. https://doi.org/10.1056/NEJMra1706910.

(15) Goker, F.; Larsson, L.; Del Fabbro, M.; Asa'ad, F. Gene Delivery Therapeutics in the Treatment of Periodontitis and Peri-Implantitis: A State of the Art Review. *Int. J. Mol. Sci.* **2019**, *20* (14), 3551. https://doi.org/10.3390/ijms20143551.

(16) Ku, M.-W.; Charneau, P.; Majlessi, L. Use of Lentiviral Vectors in Vaccination. *Expert Rev. Vaccines* **2021**, *20* (12), 1571–1586. https://doi.org/10.1080/14760584.2021.1988854.

(17) Van Tendeloo, V. F. I.; Van Broeckhoven, C.; Berneman, Z. N. Gene-Based Cancer Vaccines: An Ex Vivo Approach. *Leukemia* **2001**, *15* (4), 545–558. https://doi.org/10.1038/sj.leu.2402069.

(18) Thomas, C. E.; Ehrhardt, A.; Kay, M. A. Progress and Problems with the Use of Viral Vectors for Gene Therapy. *Nat. Rev. Genet.* **2003**, *4* (5), 346–358. https://doi.org/10.1038/nrg1066.

(19) Merten, O. W.; Charrier, S.; Laroudie, N.; Fauchille, S.; Dugué, C.; Jenny, C.; Audit, M.; Zanta-Boussif, M. A.; Chautard, H.; Radrizzani, M.; Vallanti, G.; Naldini, L.; Noguiez-Hellin, P.; Galy, A. Large-Scale Manufacture and Characterization of a Lentiviral Vector Produced for Clinical Ex Vivo Gene Therapy Application. *Hum. Gene Ther.* **2011**, *22* (3), 343–356. https://doi.org/10.1089/hum.2010.060.

(20) Joglekar, A. V.; Sandoval, S. Pseudotyped Lentiviral Vectors: One Vector, Many Guises. *Hum. Gene Ther. Methods* **2017**, *28* (6), 291–301. https://doi.org/10.1089/hgtb.2017.084.

(21) Morgan, M. A.; Galla, M.; Grez, M.; Fehse, B.; Schambach, A. Retroviral Gene Therapy in Germany with a View on Previous Experience and Future Perspectives. *Gene Ther.* **2021**, *28* (9), 494–512. https://doi.org/10.1038/s41434-021-00237-x.

(22) Labbé, R. P.; Vessillier, S.; Rafiq, Q. A. Lentiviral Vectors for T Cell Engineering: Clinical Applications, Bioprocessing and Future Perspectives. *Viruses* **2021**, *13* (8), 1528. https://doi.org/10.3390/v13081528.

(23) Vormittag, P.; Gunn, R.; Ghorashian, S.; Veraitch, F. S. A Guide to Manufacturing CAR T Cell Therapies. *Curr. Opin. Biotechnol.* **2018**, *53*, 164–181. https://doi.org/10.1016/j.copbio.2018.01.025.

(24) Lutolf, M. P.; Hubbell, J. A. Synthetic Biomaterials as Instructive Extracellular Microenvironments for Morphogenesis in Tissue Engineering. *Nat. Biotechnol.* **2005**, *23* (1), 47–55. https://doi.org/10.1038/nbt1055.

(25) Langer, R.; Vacanti, J. P. Tissue Engineering. *Science* **1993**, *260* (5110), 920–926. https://doi.org/10.1126/science.8493529.

(26) Castelvecchi, D.; Gaind, N. Climate Modellers and Theorist of Complex Systems Share Physics Nobel. *Nature* **2021**, *598* (7880), 246–247. https://doi.org/10.1038/d41586-021-02703-3.
(27) Attanasi, A.; Cavagna, A.; Del Castello, L.; Giardina, I.; Grigera, T. S.; Jelić, A.; Melillo, S.; Parisi, L.; Pohl, O.; Shen, E.; Viale, M. Information Transfer and Behavioural Inertia in Starling Flocks. *Nat. Phys.* **2014**, *10* (9), 691–696. https://doi.org/10.1038/nphys3035.
(28) Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C.; Zhang, P. Mature HIV-1 Capsid Structure by Cryo-Electron Microscopy and All-Atom Molecular Dynamics. *Nature* **2013**, *497* (7451), 643–646. https://doi.org/10.1038/nature12162.
(29) Whitesides, G. M.; Grzybowski, B. Self-Assembly at All Scales. *Science* **2002**, *295* (5564), 2418–2421. https://doi.org/10.1126/science.1070821.
(30) Stoddart, J. F. Editorial: From Supramolecular to Systems Chemistry: Complexity Emerging out of Simplicity. *Angew. Chem. Int. Ed.* **2012**, *51* (52), 12902–12903. https://doi.org/10.1002/anie.201209124.
(31) Ranby, B. The Concept of Macromolecules-Emergence, Development and Acceptance. In *Macromolecular Concept and Strategy for Humanity in Science, Technology and Industry*; Springer Berlin Heidelberg, 1996; pp 3–13. https://doi.org/10.1007/978-3-642-61036-3_2.
(32) Staudinger, H. Über Polymerisation. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.* **1920**, *53* (6), 1073–1085. https://doi.org/10.1002/cber.19200530627.
(33) Lehn, J. -M. Supramolecular Chemistry—Scope and Perspectives Molecules, Supermolecules, and Molecular Devices (Nobel Lecture). *Angew. Chemie Int. Ed. English* **1988**, *27* (1), 89–112. https://doi.org/10.1002/anie.198800891.
(34) Atkins, P.; De Paula, J. *Atkins' Physical Chemistry*; Great Britain by Oxford University Press, 2006.
(35) Ceres, P.; Zlotnick, A. Weak Protein−Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids. *Biochemistry* **2002**, *41* (39), 11525–11531. https://doi.org/10.1021/bi0261645.
(36) Rossow, T.; Seiffert, S. Supramolecular Polymer Networks: Preparation, Properties, and Potential. *Adv. Polym. Sci.* **2015**, *268*, 1–46. https://doi.org/10.1007/978-3-319-15404-6_1.
(37) Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. Reversible Polymers Formed from Self-Complementary Monomers Using Quadruple Hydrogen Bonding. *Science* **1997**, *278* (5343), 1601–1604. https://doi.org/10.1126/science.278.5343.1601.
(38) Lange, R. F. M.; Van Gurp, M.; Meijer, E. W. Hydrogen-Bonded Supramolecular Polymer Networks. *J. Polym. Sci. Part A Polym. Chem.* **1999**, *37* (19), 3657–3670. https://doi.org/10.1002/(SICI)1099-0518(19991001)37:19<3657::AID-POLA1>3.0.CO;2-6.
(39) Dankers, P. Y. W.; Harmsen, M. C.; Brouwer, L. A.; Van Luyn, M. J. A.; Meijer, E. W. A Modular and Supramolecular Approach to Bioactive Scaffolds for Tissue Engineering. *Nat. Mater.* **2005**, *4* (7), 568–574. https://doi.org/10.1038/nmat1418.
(40) Kang, J.; Miyajima, D.; Mori, T.; Inoue, Y.; Itoh, Y.; Aida, T. A Rational Strategy for the Realization of Chain-Growth Supramolecular Polymerization. *Science* **2015**, *347* (6222), 646–651. https://doi.org/10.1126/science.aaa4249.
(41) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. *Science* **2001**, *294* (5547), 1684–1688. https://doi.org/10.1126/science.1063187.
(42) Swenson, C. S.; Velusamy, A.; Argueta-Gonzalez, H. S.; Heemstra, J. M. Bilingual Peptide Nucleic Acids: Encoding the Languages of Nucleic Acids and Proteins in a Single Self-Assembling Biopolymer. *J. Am. Chem. Soc.* **2019**, *141* (48), 19038–19047. https://doi.org/10.1021/jacs.9b09146.
(43) Rowley, J. A.; Madlambayan, G.; Mooney, D. J. Alginate Hydrogels as Synthetic Extracellular Matrix Materials. *Biomaterials* **1999**, *20* (1), 45–53. https://doi.org/10.1016/S0142-9612(98)00107-0.
(44) Mitchison, T.; Kirschner, M. Dynamic Instability of Microtubule Growth. *Nature* **1984**, *312* (5991), 237–242. https://doi.org/10.1038/312237a0.
(45) Baselt, D. R.; Revel, J. P.; Baldeschwieler, J. D. Subfibrillar Structure of Type I Collagen Observed by Atomic Force Microscopy. *Biophys. J.* **1993**, *65* (6), 2644–2655. https://doi.org/10.1016/S0006-3495(93)81329-8.
(46) Gardner, M. K.; Charlebois, B. D.; Jánosi, I. M.; Howard, J.; Hunt, A. J.; Odde, D. J. Rapid Microtubule Self-Assembly Kinetics. *Cell* **2011**, *146* (4), 582–592. https://doi.org/10.1016/j.cell.2011.06.053.
(47) Fried, S. D.; Fujishima, K.; Makarov, M.; Cherepashuk, I.; Hlouchova, K. Peptides before and during the Nucleotide World: An Origins Story Emphasizing Cooperation between Proteins and Nucleic Acids. *J. R. Soc. Interface* **2022**, *19* (187). https://doi.org/10.1098/rsif.2021.0641.
(48) Adamcik, J.; Mezzenga, R. Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape. *Angew. Chem. Int. Ed.* **2018**, *57* (28), 8370–8382. https://doi.org/10.1002/anie.201713416.
(49) Sato, K.; Hendricks, M. P.; Palmer, L. C.; Stupp, S. I. Peptide Supramolecular Materials for Therapeutics. *Chem. Soc. Rev.* **2018**, *47* (20), 7539–7551. https://doi.org/10.1039/C7CS00735C.
(50) Krieg, E.; Bastings, M. M. C.; Besenius, P.; Rybtchinski, B. Supramolecular Polymers in Aqueous Media. *Chemical Reviews*. American Chemical Society January 4, 2016, pp 2414–2477. https://doi.org/10.1021/acs.chemrev.5b00369.
(51) White, S. H.; Wimley, W. C. Hydrophobic Interactions of Peptides with Membrane Interfaces. *Biochim. Biophys. Acta - Rev. Biomembr.* **1998**, *1376* (3), 339–352. https://doi.org/10.1016/S0304-4157(98)00021-5.
(52) PERKINS, S. J. Protein Volumes and Hydration Effects. The Calculations of Partial Specific Volumes, Neutron Scattering Matchpoints and 280-Nm Absorption Coefficients for Proteins and Glycoproteins from Amino

Acid Sequences. *Eur. J. Biochem.* **1986**, *157* (1), 169–180. https://doi.org/10.1111/j.1432-1033.1986.tb09653.x.
(53) Rehm, T. H.; Schmuck, C. Ion-Pair Induced Self-Assembly in Aqueous Solvents. *Chem. Soc. Rev.* **2010**, *39* (10), 3597. https://doi.org/10.1039/b926223g.
(54) von Gröning, M.; de Feijter, I.; Stuart, M. C. A.; Voets, I. K.; Besenius, P. Tuning the Aqueous Self-Assembly of Multistimuli-Responsive Polyanionic Peptide Nanorods. *J. Mater. Chem. B* **2013**, *1* (15), 2008. https://doi.org/10.1039/c3tb00051f.
(55) Biedermann, F.; Nau, W. M.; Schneider, H.-J. The Hydrophobic Effect Revisited-Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. *Angew. Chem. Int. Ed.* **2014**, *53* (42), 11158–11171. https://doi.org/10.1002/anie.201310958.
(56) Grimme, S. Do Special Noncovalent π–π Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008**, *47* (18), 3430–3434. https://doi.org/10.1002/anie.200705157.
(57) Martinez, C. R.; Iverson, B. L. Rethinking the Term "Pi-Stacking." *Chem. Sci.* **2012**, *3* (7), 2191. https://doi.org/10.1039/c2sc20045g.
(58) Reches, M.; Gazit, E. Casting Metal Nanowires Within Discrete Self-Assembled Peptide Nanotubes. *Science* **2003**, *300* (5619), 625–627. https://doi.org/10.1126/science.1082387.
(59) Gazit, E. Self Assembly of Short Aromatic Peptides into Amyloid Fibrils and Related Nanostructures. *Prion* **2007**, *1* (1), 32–35. https://doi.org/10.4161/pri.1.1.4095.
(60) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437* (7059), 640–647. https://doi.org/10.1038/nature04162.
(61) Tuñón, I.; Silla, E.; Pascual-Ahuir, J. L. Molecular Surface Area and Hydrophobic Effect. *"Protein Eng. Des. Sel.* **1992**, *5* (8), 715–716. https://doi.org/10.1093/protein/5.8.715.
(62) Meyer, E. E.; Rosenberg, K. J.; Israelachvili, J. Recent Progress in Understanding Hydrophobic Interactions. *Proc. Natl. Acad. Sci.* **2006**, *103* (43), 15739–15746. https://doi.org/10.1073/pnas.0606422103.
(63) Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. In *Advances in Protein Chemistry*; Academic Press, 1959; Vol. 14, pp 1–63. https://doi.org/10.1016/S0065-3233(08)60608-7.
(64) Guzzo, A. V. The Influence of Amino Acid Sequence on Protein Structure. *Biophys. J.* **1965**, *5* (6), 809–822. https://doi.org/10.1016/S0006-3495(65)86753-4.
(65) Fujiwara, K.; Toda, H.; Ikeguchi, M. Dependence of Alpha-Helical and Beta-Sheet Amino Acid Propensities on the Overall Protein Fold Type. *BMC Struct. Biol.* **2012**, *12* (1), 18. https://doi.org/10.1186/1472-6807-12-18.
(66) Cui, H.; Cheetham, A. G.; Pashuck, E. T.; Stupp, S. I. Amino Acid Sequence in Constitutionally Isomeric Tetrapeptide Amphiphiles Dictates Architecture of One-Dimensional Nanostructures. *J. Am. Chem. Soc.* **2014**, *136* (35), 12461–12468. https://doi.org/10.1021/ja507051w.
(67) Sieste, S.; Mack, T.; Lump, E.; Hayn, M.; Schütz, D.; Röcker, A.; Meier, C.; Kaygisiz, K.; Kirchhoff, F.; Knowles, T. P. J.; Ruggeri, F. S.; Synatschke, C. V.; Münch, J.; Weil, T. Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction. *Adv. Funct. Mater.* **2021**, *31* (17), 2009382. https://doi.org/10.1002/adfm.202009382.
(68) Ulijn, R. V.; Jerala, R. Peptide and Protein Nanotechnology into the 2020s: Beyond Biology. *Chem. Soc. Rev.* **2018**, *47* (10), 3391–3394. https://doi.org/10.1039/C8CS90055H.
(69) Greenwald, J.; Riek, R. Biology of Amyloid: Structure, Function, and Regulation. *Structure*. Cell Press October 13, 2010, pp 1244–1260. https://doi.org/10.1016/j.str.2010.08.009.
(70) Adamcik, J.; Ruggeri, F. S.; Berryman, J. T.; Zhang, A.; Knowles, T. P. J.; Mezzenga, R. Evolution of Conformation, Nanomechanics, and Infrared Nanospectroscopy of Single Amyloid Fibrils Converting into Microcrystals. *Adv. Sci.* **2021**, *8* (2), 2002182. https://doi.org/10.1002/advs.202002182.
(71) Zapadka, K. L.; Becher, F. J.; Gomes dos Santos, A. L.; Jackson, S. E. Factors Affecting the Physical Stability (Aggregation) of Peptide Therapeutics. *Interface Focus* **2017**, *7* (6), 20170030. https://doi.org/10.1098/rsfs.2017.0030.
(72) Schmidt, A.; Annamalai, K.; Schmidt, M.; Grigorieff, N.; Fändrich, M. Cryo-EM Reveals the Steric Zipper Structure of a Light Chain-Derived Amyloid Fibril. *Proc. Natl. Acad. Sci.* **2016**, *113* (22), 6200–6205. https://doi.org/10.1073/pnas.1522282113.
(73) Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C. Molecular Basis for Amyloid Fibril Formation and Stability. *Proc. Natl. Acad. Sci.* **2005**, *102* (2), 315–320. https://doi.org/10.1073/pnas.0406847102.
(74) Heerde, T.; Bansal, A.; Schmidt, M.; Fändrich, M. Cryo-EM Structure of a Catalytic Amyloid Fibril. *Sci. Rep.* **2023**, *13* (1), 4070. https://doi.org/10.1038/s41598-023-30711-y.
(75) Glenner, G. G.; Wong, C. W. Alzheimer's Disease and Down's Syndrome: Sharing of a Unique Cerebrovascular Amyloid Fibril Protein. *Biochem. Biophys. Res. Commun.* **1984**, *122* (3), 1131–1135. https://doi.org/10.1016/0006-291X(84)91209-9.
(76) Maltsev, A. V.; Bystryak, S.; Galzitskaya, O. V. The Role of β-Amyloid Peptide in Neurodegenerative Diseases. *Ageing Res. Rev.* **2011**, *10* (4), 440–452. https://doi.org/10.1016/j.arr.2011.03.002.
(77) Cherny, I.; Gazit, E. Amyloids: Not Only Pathological Agents but Also Ordered Nanomaterials. *Angew. Chem. Int. Ed.* **2008**, *47* (22), 4062–4069. https://doi.org/10.1002/anie.200703133.
(78) Romero, D.; Aguilar, C.; Losick, R.; Kolter, R. Amyloid Fibers Provide Structural Integrity to Bacillus Subtilis Biofilms. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (5), 2230–2234. https://doi.org/10.1073/pnas.0910560107.
(79) Taglialegna, A.; Lasa, I.; Valle, J. Amyloid Structures as Biofilm Matrix Scaffolds. *J. Bacteriol.* **2016**, *198* (19),

2579–2588. https://doi.org/10.1128/JB.00122-16.
(80) Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; Eisenberg, D.; Rivier, J.; Sawchenko, P.; Vale, W.; Riek, R. Functional Amyloids as Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009**, *325* (5938), 328–332. https://doi.org/10.1126/science.1173155.
(81) Watt, B.; van Niel, G.; Raposo, G.; Marks, M. S. PMEL: A Pigment Cell-Specific Model for Functional Amyloid Formation. *Pigment Cell Melanoma Res.* **2013**, *26* (3), 300–315. https://doi.org/10.1111/pcmr.12067.
(82) Dehsorkhi, A.; Castelletto, V.; Hamley, I. W. Self-assembling Amphiphilic Peptides. *J. Pept. Sci.* **2014**, *20* (7), 453–467. https://doi.org/10.1002/psc.2633.
(83) Zhang, S. Fabrication of Novel Biomaterials through Molecular Self-Assembly. *Nat. Biotechnol.* **2003**, *21* (10), 1171–1178. https://doi.org/10.1038/nbt874.
(84) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. *Science* **2001**, *294* (5547), 1684–1688. https://doi.org/10.1126/science.1063187.
(85) Chen, C. H.; Palmer, L. C.; Stupp, S. I. Self-Sorting in Supramolecular Assemblies. *Soft Matter* **2021**, *17* (14), 3902–3912. https://doi.org/10.1039/D1SM00113B.
(86) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. *Science* **2001**, *294* (5547), 1684–1688. https://doi.org/10.1126/science.1063187.
(87) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Peptide-Amphiphile Nanofibers: A Versatile Scaffold for the Preparation of Self-Assembling Materials. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (8), 5133–5138. https://doi.org/10.1073/pnas.072699999.
(88) Silva, G. A.; Czeisler, C.; Niece, K. L.; Beniash, E.; Harrington, D. A.; Kessler, J. A.; Stupp, S. I. Selective Differentiation of Neural Progenitor Cells by High-Epitope Density Nanofibers. *Science* **2004**, *303* (5662), 1352–1355. https://doi.org/10.1126/science.1093783.
(89) Shah, R. N.; Shah, N. A.; Lim, M. M. D. R.; Hsieh, C.; Nuber, G.; Stupp, S. I. Supramolecular Design of Self-Assembling Nanofibers for Cartilage Regeneration. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (8), 3293–3298. https://doi.org/10.1073/pnas.0906501107.
(90) Sleep, E.; Cosgrove, B. D.; McClendon, M. T.; Preslar, A. T.; Chen, C. H.; Sangji, M. H.; Pérez, C. M. R.; Haynes, R. D.; Meade, T. J.; Blau, H. M.; Stupp, S. I. Injectable Biomimetic Liquid Crystalline Scaffolds Enhance Muscle Stem Cell Transplantation. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (38), E7919–E7928. https://doi.org/10.1073/pnas.1708142114.
(91) Webber, M. J.; Tongers, J.; Renault, M.-A.; Roncalli, J. G.; Losordo, D. W.; Stupp, S. I. Development of Bioactive Peptide Amphiphiles for Therapeutic Cell Delivery. *Acta Biomater.* **2010**, *6* (1), 3–11. https://doi.org/10.1016/j.actbio.2009.07.031.
(92) Roth, P.; Meyer, R.; Harley, I.; Landfester, K.; Lieberwirth, I.; Wagner, M.; Ng, D. Y. W.; Weil, T. Supramolecular Assembly Guided by Photolytic Redox Cycling. *Nat. Synth.* **2023**, 1–9. https://doi.org/10.1038/s44160-023-00343-1.
(93) Korevaar, P. A.; Newcomb, C. J.; Meijer, E. W.; Stupp, S. I. Pathway Selection in Peptide Amphiphile Assembly. *J. Am. Chem. Soc.* **2014**, *136* (24), 8540–8543. https://doi.org/10.1021/ja503882s.
(94) Adamcik, J.; Mezzenga, R. Proteins Fibrils from a Polymer Physics Perspective. *Macromolecules* **2012**, *45* (3), 1137–1150. https://doi.org/10.1021/ma202157h.
(95) Knowles, T. P. J.; Buehler, M. J. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* **2011**, *6* (8), 469–479. https://doi.org/10.1038/nnano.2011.102.
(96) Al-Garawi, Z. S.; Morris, K. L.; Marshall, K. E.; Eichler, J.; Serpell, L. C. The Diversity and Utility of Amyloid Fibrils Formed by Short Amyloidogenic Peptides. *Interface Focus* **2017**, *7* (6), 20170027. https://doi.org/10.1098/rsfs.2017.0027.
(97) Dear, A. J.; Meisl, G.; Šarić, A.; Michaels, T. C. T.; Kjaergaard, M.; Linse, S.; Knowles, T. P. J. Identification of On- and off-Pathway Oligomers in Amyloid Fibril Formation. *Chem. Sci.* **2020**, *11* (24), 6236–6247. https://doi.org/10.1039/C9SC06501F.
(98) Michaels, T. C. T.; Šarić, A.; Curk, S.; Bernfur, K.; Arosio, P.; Meisl, G.; Dear, A. J.; Cohen, S. I. A.; Dobson, C. M.; Vendruscolo, M.; Linse, S.; Knowles, T. P. J. Dynamics of Oligomer Populations Formed during the Aggregation of Alzheimer's Aβ42 Peptide. *Nat. Chem.* **2020**, *12* (5), 445–451. https://doi.org/10.1038/s41557-020-0452-1.
(99) Pellarin, R.; Caflisch, A. Interpreting the Aggregation Kinetics of Amyloid Peptides. *J. Mol. Biol.* **2006**, *360* (4), 882–892. https://doi.org/10.1016/j.jmb.2006.05.033.
(100) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional Amyloid Formation within Mammalian Tissue. *PLoS Biol.* **2006**, *4* (1), 0100–0107. https://doi.org/10.1371/journal.pbio.0040006.
(101) Cremades, N.; Cohen, S. I. A.; Deas, E.; Abramov, A. Y.; Chen, A. Y.; Orte, A.; Sandal, M.; Clarke, R. W.; Dunne, P.; Aprile, F. A.; Bertoncini, C. W.; Wood, N. W.; Knowles, T. P. J.; Dobson, C. M.; Klenerman, D. Direct Observation of the Interconversion of Normal and Toxic Forms of α-Synuclein. *Cell* **2012**, *149* (5), 1048–1059. https://doi.org/10.1016/j.cell.2012.03.037.
(102) Klein, W. .; Stine, W. .; Teplow, D. . Small Assemblies of Unmodified Amyloid β-Protein Are the Proximate Neurotoxin in Alzheimer's Disease. *Neurobiol. Aging* **2004**, *25* (5), 569–580. https://doi.org/10.1016/j.neurobiolaging.2004.02.010.
(103) Stefani, M.; Dobson, C. M. Protein Aggregation and Aggregate Toxicity: New Insights into Protein Folding, Misfolding Diseases and Biological Evolution. *J. Mol. Med.* **2003**, *81* (11), 678–699.

https://doi.org/10.1007/s00109-003-0464-5.
(104) Zhang, S.; Andreasen, M.; Nielsen, J. T.; Liu, L.; Nielsen, E. H.; Song, J.; Ji, G.; Sun, F.; Skrydstrup, T.; Besenbacher, F.; Nielsen, N. C.; Otzen, D. E.; Dong, M. Coexistence of Ribbon and Helical Fibrils Originating from HIAPP 20–29 Revealed by Quantitative Nanomechanical Atomic Force Microscopy. *Proc. Natl. Acad. Sci.* **2013**, *110* (8), 2798–2803. https://doi.org/10.1073/pnas.1209955110.
(105) Ruggeri, F. S.; Adamcik, J.; Jeong, J. S.; Lashuel, H. A.; Mezzenga, R.; Dietler, G. Influence of the β-Sheet Content on the Mechanical Properties of Aggregates during Amyloid Fibrillization. *Angew. Chem. Int. Ed.* **2015**, *54* (8), 2462–2466. https://doi.org/10.1002/anie.201409050.
(106) Yoshimura, Y.; Lin, Y.; Yagi, H.; Lee, Y. H.; Kitayama, H.; Sakurai, K.; So, M.; Ogi, H.; Naiki, H.; Goto, Y. Distinguishing Crystal-like Amyloid Fibrils and Glass-like Amorphous Aggregates from Their Kinetics of Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (36), 14446–14451. https://doi.org/10.1073/pnas.1208228109.
(107) Maurer-Stroh, S.; Debulpaep, M.; Kuemmerer, N.; de la Paz, M. L.; Martins, I. C.; Reumers, J.; Morris, K. L.; Copland, A.; Serpell, L.; Serrano, L.; Schymkowitz, J. W. H.; Rousseau, F. Exploring the Sequence Determinants of Amyloid Structure Using Position-Specific Scoring Matrices. *Nat. Methods* **2010**, *7* (3), 237–242. https://doi.org/10.1038/nmeth.1432.
(108) Adachi, M.; So, M.; Sakurai, K.; Kardos, J.; Goto, Y. Supersaturation-Limited and Unlimited Phase Transitions Compete to Produce the Pathway Complexity in Amyloid Fibrillation. *J. Biol. Chem.* **2015**, *290* (29), 18134–18145. https://doi.org/10.1074/jbc.M115.648139.
(109) Ke, P. C.; Zhou, R.; Serpell, L. C.; Riek, R.; Knowles, T. P. J.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fändrich, M.; Otzen, D. E.; Chapman, M. R.; Dobson, C. M.; Eisenberg, D. S.; Mezzenga, R. Half a Century of Amyloids: Past, Present and Future. *Chem. Soc. Rev.* **2020**, *49* (15), 5473–5509. https://doi.org/10.1039/C9CS00199A.
(110) Adamcik, J.; Ruggeri, F. S.; Berryman, J. T.; Zhang, A.; Knowles, T. P. J.; Mezzenga, R. Evolution of Conformation, Nanomechanics, and Infrared Nanospectroscopy of Single Amyloid Fibrils Converting into Microcrystals. *Adv. Sci.* **2021**, *8* (2), 2002182. https://doi.org/10.1002/advs.202002182.
(111) Milanesi, L.; Sheynis, T.; Xue, W.-F.; Orlova, E. V.; Hellewell, A. L.; Jelinek, R.; Hewitt, E. W.; Radford, S. E.; Saibil, H. R. Direct Three-Dimensional Visualization of Membrane Disruption by Amyloid Fibrils. *Proc. Natl. Acad. Sci.* **2012**, *109* (50), 20455–20460. https://doi.org/10.1073/pnas.1206325109.
(112) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27–68. https://doi.org/10.1146/annurev-biochem-061516-045115.
(113) Frederix, P. W. J. M.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. Exploring the Sequence Space for (Tri-)Peptide Self-Assembly to Design and Discover New Hydrogels. *Nat. Chem.* **2015**, *7* (1), 30–37. https://doi.org/10.1038/nchem.2122.
(114) Strandberg, E.; Schweigardt, F.; Wadhwani, P.; Bürck, J.; Reichert, J.; Cravo, H. L. P.; Burger, L.; Ulrich, A. S. Phosphate-Dependent Aggregation of [KL]n Peptides Affects Their Membranolytic Activity. *Sci. Rep.* **2020**, *10* (1), 12300. https://doi.org/10.1038/s41598-020-69162-0.
(115) Marek, P. J.; Patsalo, V.; Green, D. F.; Raleigh, D. P. Ionic Strength Effects on Amyloid Formation by Amylin Are a Complicated Interplay among Debye Screening, Ion Selectivity, and Hofmeister Effects. *Biochemistry* **2012**, *51* (43), 8478–8490. https://doi.org/10.1021/bi300574r.
(116) Konstantoulea, K.; Louros, N.; Rousseau, F.; Schymkowitz, J. Heterotypic Interactions in Amyloid Function and Disease. *FEBS J.* **2022**, *289* (8), 2025–2046. https://doi.org/10.1111/febs.15719.
(117) Do, T. D.; Kincannon, W. M.; Bowers, M. T. Phenylalanine Oligomers and Fibrils: The Mechanism of Assembly and the Importance of Tetramers and Counterions. *J. Am. Chem. Soc.* **2015**, *137* (32), 10080–10083. https://doi.org/10.1021/jacs.5b05482.
(118) Eckes, K. M.; Mu, X.; Ruehle, M. A.; Ren, P.; Suggs, L. J. β Sheets Not Required: Combined Experimental and Computational Studies of Self-Assembly and Gelation of the Ester-Containing Analogue of an Fmoc-Dipeptide Hydrogelator. *Langmuir* **2014**, *30* (18), 5287–5296. https://doi.org/10.1021/la500679b.
(119) Lu, H.; Bellucci, L.; Sun, S.; Qi, D.; Rosa, M.; Berger, R.; Corni, S.; Bonn, M. Acidic PH Promotes Refolding and Macroscopic Assembly of Amyloid β (16–22) Peptides at the Air–Water Interface. *J. Phys. Chem. Lett.* **2022**, *13* (29), 6674–6679. https://doi.org/10.1021/acs.jpclett.2c01171.
(120) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the Cross-β Spine of Amyloid-like Fibrils. *Nature* **2005**, *435* (7043), 773–778. https://doi.org/10.1038/nature03680.
(121) Heerde, T.; Schütz, D.; Lin, Y.-J.; Münch, J.; Schmidt, M.; Fändrich, M. Cryo-EM Structure and Polymorphic Maturation of a Viral Transduction Enhancing Amyloid Fibril. *Nat. Commun.* **2023**, *14* (1), 4293. https://doi.org/10.1038/s41467-023-40042-1.
(122) De Mello, L. R.; Porosk, L.; Lourenço, T. C.; Garcia, B. B. M.; Costa, C. A. R.; Han, S. W.; De Souza, J. S.; Langel, Ü.; Da Silva, E. R. Amyloid-like Self-Assembly of a Hydrophobic Cell-Penetrating Peptide and Its Use as a Carrier for Nucleic Acids. *ACS Appl. Bio Mater.* **2021**, *4* (8), 6404–6416. https://doi.org/10.1021/acsabm.1c00601.
(123) Zhao, R.; So, M.; Maat, H.; Ray, N. J.; Arisaka, F.; Goto, Y.; Carver, J. A.; Hall, D. Measurement of Amyloid Formation by Turbidity Assay—Seeing through the Cloud. *Biophys. Rev.* **2016**, *8* (4), 445–471. https://doi.org/10.1007/s12551-016-0233-7.
(124) Mossou, E.; Teixeira, S. C. M.; Mitchell, E. P.; Mason, S. A.; Adler-Abramovich, L.; Gazit, E.; Forsyth, V. T. The

Self-Assembling Zwitterionic Form of <scp>L</Scp> -Phenylalanine at Neutral PH. *Acta Crystallogr. Sect. C Struct. Chem.* **2014**, *70* (3), 326–331. https://doi.org/10.1107/S2053229614002563.

(125) Colebc, N. B.; Murphy, D. D.; Grider, T.; Rueter, S.; Brasaemle, D.; Nussbaum, R. L. Lipid Droplet Binding and Oligomerization Properties of the Parkinson's Disease Protein α-Synuclein. *J. Biol. Chem.* **2002**, *277* (8), 6344–6352. https://doi.org/10.1074/jbc.M108414200.

(126) Maltseva, D.; Chatterjee, S.; Yu, C.-C.; Brzezinski, M.; Nagata, Y.; Gonella, G.; Murthy, A. C.; Stachowiak, J. C.; Fawzi, N. L.; Parekh, S. H.; Bonn, M. Fibril Formation and Ordering of Disordered FUS LC Driven by Hydrophobic Interactions. *Nat. Chem.* **2023**, 1–9. https://doi.org/10.1038/s41557-023-01221-1.

(127) Schwarze, B.; Huster, D. How Single Site Mutations Can Help Understanding Structure Formation of Amyloid β 1–40. *Macromol. Biosci.* **2023**, 2200489. https://doi.org/10.1002/mabi.202200489.

(128) Tycko, R. Physical and Structural Basis for Polymorphism in Amyloid Fibrils. *Protein Sci.* **2014**, *23* (11), 1528–1539. https://doi.org/10.1002/pro.2544.

(129) Ruggeri, F. S.; Longo, G.; Faggiano, S.; Lipiec, E.; Pastore, A.; Dietler, G. Infrared Nanospectroscopy Characterization of Oligomeric and Fibrillar Aggregates during Amyloid Formation. *Nat. Commun.* **2015**, *6*, 1–9. https://doi.org/10.1038/ncomms8831.

(130) Paparcone, R.; Cranford, S. W.; Buehler, M. J. Self-Folding and Aggregation of Amyloid Nanofibrils. *Nanoscale* **2011**, *3* (4), 1748. https://doi.org/10.1039/c0nr00840k.

(131) Adamcik, J.; Jung, J.-M.; Flakowski, J.; De Los Rios, P.; Dietler, G.; Mezzenga, R. Understanding Amyloid Aggregation by Statistical Analysis of Atomic Force Microscopy Images. *Nat. Nanotechnol.* **2010**, *5* (6), 423–428. https://doi.org/10.1038/nnano.2010.59.

(132) Bolisetty, S.; Adamcik, J.; Mezzenga, R. Snapshots of Fibrillation and Aggregation Kinetics in Multistranded Amyloid β-Lactoglobulin Fibrils. *Soft Matter* **2011**, *7* (2), 493–499. https://doi.org/10.1039/C0SM00502A.

(133) Ridgley, D. M.; Ebanks, K. C.; Barone, J. R. Peptide Mixtures Can Self-Assemble into Large Amyloid Fibers of Varying Size and Morphology. *Biomacromolecules* **2011**, *12* (10), 3770–3779. https://doi.org/10.1021/bm201005k.

(134) Knowles, T. P. J.; Oppenheim, T. W.; Buell, A. K.; Chirgadze, D. Y.; Welland, M. E. Nanostructured Films from Hierarchical Self-Assembly of Amyloidogenic Proteins. *Nat. Nanotechnol.* **2010**, *5* (3), 204–207. https://doi.org/10.1038/nnano.2010.26.

(135) Usov, I.; Adamcik, J.; Mezzenga, R. Polymorphism in Bovine Serum Albumin Fibrils: Morphology and Statistical Analysis. *Faraday Discuss.* **2013**, *166* (0), 151. https://doi.org/10.1039/c3fd00083d.

(136) Ridgley, D. M.; Barone, J. R. Evolution of the Amyloid Fiber over Multiple Length Scales. *ACS Nano* **2013**, *7* (2), 1006–1015. https://doi.org/10.1021/nn303489a.

(137) Lara, C.; Adamcik, J.; Jordens, S.; Mezzenga, R. General Self-Assembly Mechanism Converting Hydrolyzed Globular Proteins Into Giant Multistranded Amyloid Ribbons. *Biomacromolecules* **2011**, *12* (5), 1868–1875. https://doi.org/10.1021/bm200216u.

(138) Meisl, G.; Rajah, L.; Cohen, S. A. I.; Pfammatter, M.; Šarić, A.; Hellstrand, E.; Buell, A. K.; Aguzzi, A.; Linse, S.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. Scaling Behaviour and Rate-Determining Steps in Filamentous Self-Assembly. *Chem. Sci.* **2017**, *8* (10), 7087–7097. https://doi.org/10.1039/C7SC01965C.

(139) Qin, S.-Y.; Pei, Y.; Liu, X.-J.; Zhuo, R.-X.; Zhang, X.-Z. Hierarchical Self-Assembly of a β-Amyloid Peptide Derivative. *J. Mater. Chem. B* **2013**, *1* (5), 668–675. https://doi.org/10.1039/C2TB00105E.

(140) Webber, M. J.; Appel, E. A.; Meijer, E. W.; Langer, R. Supramolecular Biomaterials. *Nat. Mater.* **2016**, *15* (1), 13–26. https://doi.org/10.1038/nmat4474.

(141) Das, A. K.; Gavel, P. K. Low Molecular Weight Self-Assembling Peptide-Based Materials for Cell Culture, Antimicrobial, Anti-Inflammatory, Wound Healing, Anticancer, Drug Delivery, Bioimaging and 3D Bioprinting Applications. *Soft Matter* **2020**, *16* (44), 10065–10095. https://doi.org/10.1039/D0SM01136C.

(142) Reynolds, N. P. Amyloid-like Peptide Nanofibrils as Scaffolds for Tissue Engineering: Progress and Challenges (Review). *Biointerphases* **2019**, *14* (4), 040801. https://doi.org/10.1116/1.5098332.

(143) Balasco, N.; Diaferia, C.; Morelli, G.; Vitagliano, L.; Accardo, A. Amyloid-Like Aggregation in Diseases and Biomaterials: Osmosis of Structural Information. *Front. Bioeng. Biotechnol.* **2021**, *9*, 130. https://doi.org/10.3389/fbioe.2021.641372.

(144) Jackson, M.; Hewitt, E. Why Are Functional Amyloids Non-Toxic in Humans? *Biomolecules* **2017**, *7* (4), 71. https://doi.org/10.3390/biom7040071.

(145) Schilling, C.; Mack, T.; Lickfett, S.; Sieste, S.; Ruggeri, F. S.; Sneideris, T.; Dutta, A.; Bereau, T.; Naraghi, R.; Sinske, D.; Knowles, T. P. J.; Synatschke, C. V.; Weil, T.; Knöll, B. Sequence-Optimized Peptide Nanofibers as Growth Stimulators for Regeneration of Peripheral Neurons. *Adv. Funct. Mater.* **2019**, *29* (24), 1809112. https://doi.org/10.1002/adfm.201809112.

(146) Wojtowicz, W. M.; Farzan, M.; Joyal, J. L.; Carter, K.; Babcock, G. J.; Israel, D. I.; Sodroski, J.; Mirzabekov, T. Stimulation of Enveloped Virus Infection by β-Amyloid Fibrils. *J. Biol. Chem.* **2002**, *277* (38), 35019–35024. https://doi.org/10.1074/jbc.M203518200.

(147) Münch, J.; Rücker, E.; Ständker, L.; Adermann, K.; Goffinet, C.; Schindler, M.; Wildum, S.; Chinnadurai, R.; Rajan, D.; Specht, A.; Giménez-Gallego, G.; Sánchez, P. C.; Fowler, D. M.; Koulov, A.; Kelly, J. W.; Mothes, W.; Grivel, J. C.; Margolis, L.; Keppler, O. T.; Forssmann, W. G.; Kirchhoff, F. Semen-Derived Amyloid Fibrils Drastically Enhance HIV Infection. *Cell* **2007**, *131* (6), 1059–1071. https://doi.org/10.1016/j.cell.2007.10.014.

(148) Arnold, F.; Schnell, J.; Zirafi, O.; Sturzel, C.; Meier, C.; Weil, T.; Standker, L.; Forssmann, W.-G.; Roan, N. R.; Greene, W. C.; Kirchhoff, F.; Münch, J. Naturally Occurring Fragments from Two Distinct Regions of the

Prostatic Acid Phosphatase Form Amyloidogenic Enhancers of HIV Infection. *J. Virol.* **2012**, *86* (2), 1244–1249. https://doi.org/10.1128/JVI.06121-11.

(149) Roan, N. R.; Müller, J. A.; Liu, H.; Chu, S.; Arnold, F.; Stürzel, C. M.; Walther, P.; Dong, M.; Witkowska, H. E.; Kirchhoff, F.; Münch, J.; Greene, W. C. Peptides Released by Physiological Cleavage of Semen Coagulum Proteins Form Amyloids That Enhance HIV Infection. *Cell Host Microbe* **2011**, *10* (6), 541–550. https://doi.org/10.1016/j.chom.2011.10.010.

(150) Arnold, F.; Schnell, J.; Zirafi, O.; Sturzel, C.; Meier, C.; Weil, T.; Standker, L.; Forssmann, W.-G.; Roan, N. R.; Greene, W. C.; Kirchhoff, F.; Münch, J. Naturally Occurring Fragments from Two Distinct Regions of the Prostatic Acid Phosphatase Form Amyloidogenic Enhancers of HIV Infection. *J. Virol.* **2012**, *86* (2), 1244–1249. https://doi.org/10.1128/jvi.06121-11.

(151) Zhang, L.; Jiang, C.; Zhang, H.; Gong, X.; Yang, L.; Miao, L.; Shi, Y.; Zhang, Y.; Kong, W.; Zhang, C.; Shan, Y. A Novel Modified Peptide Derived from Membrane-Proximal External Region of Human Immunodeficiency Virus Type 1 Envelope Significantly Enhances Retrovirus Infection. *J. Pept. Sci.* **2014**, *20* (1), 46–54. https://doi.org/10.1002/psc.2587.

(152) Lee, Y.-H.; Ramamoorthy, A. Semen-Derived Amyloidogenic Peptides-Key Players of HIV Infection. *Protein Sci.* **2018**, *27* (7), 1151–1165. https://doi.org/10.1002/pro.3395.

(153) Castelletto, V.; Hamley, I. W. Amyloid and Hydrogel Formation of a Peptide Sequence from a Coronavirus Spike Protein. *ACS Nano* **2022**, *16* (2), 1857–1867. https://doi.org/10.1021/acsnano.1c10658.

(154) Yolamanova, M.; Meier, C.; Shaytan, A. K.; Vas, V.; Bertoncini, C. W.; Arnold, F.; Zirafi, O.; Usmani, S. M.; Müller, J. A.; Sauter, D.; Goffinet, C.; Palesch, D.; Walther, P.; Roan, N. R.; Geiger, H.; Lunov, O.; Simmet, T.; Bohne, J.; Schrezenmeier, H.; Schwarz, K.; Ständker, L.; Forssmann, W.-G.; Salvatella, X.; Khalatur, P. G.; Khokhlov, A. R.; Knowles, T. P. J.; Weil, T.; Kirchhoff, F.; Münch, J. Peptide Nanofibrils Boost Retroviral Gene Transfer and Provide a Rapid Means for Concentrating Viruses. *Nat. Nanotechnol.* **2013**, *8* (2), 130–136. https://doi.org/10.1038/nnano.2012.248.

(155) Kirti, S.; Patel, K.; Das, S.; Shrimali, P.; Samanta, S.; Kumar, R.; Chatterjee, D.; Ghosh, D.; Kumar, A.; Tayalia, P.; Maji, S. K. Amyloid Fibrils with Positive Charge Enhance Retroviral Transduction in Mammalian Cells. *ACS Biomater. Sci. Eng.* **2019**, *5* (1), 126–138. https://doi.org/10.1021/acsbiomaterials.8b00248.

(156) Yolamanova, M.; Meier, C.; Shaytan, A. K.; Vas, V.; Bertoncini, C. W.; Arnold, F.; Zirafi, O.; Usmani, S. M.; Müller, J. A.; Sauter, D.; Goffinet, C.; Palesch, D.; Walther, P.; Roan, N. R.; Geiger, H.; Lunov, O.; Simmet, T.; Bohne, J.; Schrezenmeier, H.; Schwarz, K.; Ständker, L.; Forssmann, W.-G.; Salvatella, X.; Khalatur, P. G.; Khokhlov, A. R.; Knowles, T. P. J.; Weil, T.; Kirchhoff, F.; Münch, J. Peptide Nanofibrils Boost Retroviral Gene Transfer and Provide a Rapid Means for Concentrating Viruses. *Nat. Nanotechnol.* **2013**, *8* (2), 130–136. https://doi.org/10.1038/nnano.2012.248.

(157) Shrimali, P. C.; Chen, S.; Dreher, R.; 2#, M. K. H.; Buck, J.; Kim, D.; Rudra, J. S.; Jackrel, M. E.; Rudra, J.; Jackrel, M. Amyloidogenic Propensity of Self-Assembling Peptides and Their Adjuvant Potential for Use as DNA Vaccines. *bioRxiv* **2022**, 2022.09.09.507367. https://doi.org/10.1101/2022.09.09.507367.

(158) Münch, J.; Rücker, E.; Ständker, L.; Adermann, K.; Goffinet, C.; Schindler, M.; Wildum, S.; Chinnadurai, R.; Rajan, D.; Specht, A.; Giménez-Gallego, G.; Sánchez, P. C.; Fowler, D. M.; Koulov, A.; Kelly, J. W.; Mothes, W.; Grivel, J.-C.; Margolis, L.; Keppler, O. T.; Forssmann, W.-G.; Kirchhoff, F. Semen-Derived Amyloid Fibrils Drastically Enhance HIV Infection. *Cell* **2007**, *131* (6), 1059–1071. https://doi.org/10.1016/j.cell.2007.10.014.

(159) Schütz, D.; Rode, S.; Read, C.; Müller, J. A.; Glocker, B.; Sparrer, K. M. J.; Fackler, O. T.; Walther, P.; Münch, J. Viral Transduction Enhancing Effect of EF-C Peptide Nanofibrils Is Mediated by Cellular Protrusions. *Adv. Funct. Mater.* **2021**, *31* (40), 2104814. https://doi.org/10.1002/adfm.202104814.

(160) Kirti, S.; Patel, K.; Das, S.; Shrimali, P.; Samanta, S.; Kumar, R.; Chatterjee, D.; Ghosh, D.; Kumar, A.; Tayalia, P.; Maji, S. K. Amyloid Fibrils with Positive Charge Enhance Retroviral Transduction in Mammalian Cells. *ACS Biomater. Sci. Eng.* **2019**, *5* (1), 126–138. https://doi.org/10.1021/acsbiomaterials.8b00248.

(161) Roan, N. R.; Münch, J.; Arhel, N.; Mothes, W.; Neidleman, J.; Kobayashi, A.; Smith-McCune, K.; Kirchhoff, F.; Greene, W. C. The Cationic Properties of SEVI Underlie Its Ability To Enhance Human Immunodeficiency Virus Infection. *J. Virol.* **2009**, *83* (1), 73–80. https://doi.org/10.1128/JVI.01366-08.

(162) Schütz, D.; Read, C.; Groß, R.; Röcker, A.; Rode, S.; Annamalai, K.; Fändrich, M.; Münch, J. Negatively Charged Peptide Nanofibrils from Immunoglobulin Light Chain Sequester Viral Particles but Lack Cell-Binding and Viral Transduction-Enhancing Properties. *ACS Omega* **2021**, *6* (11), 7731–7738. https://doi.org/10.1021/acsomega.1c00068.

(163) Schütz, D. Mechanistic Studies on Transduction Enhancing EF-C Peptide Nanofibrils and Optimization for Gene Therapeutic Applications, Ulm University, 2021.

(164) Mercer, J.; Schelhaas, M.; Helenius, A. Virus Entry by Endocytosis. *Annu. Rev. Biochem.* **2010**, *79* (1), 803–833. https://doi.org/10.1146/annurev-biochem-060208-104626.

(165) Marsh, M.; Helenius, A. Virus Entry: Open Sesame. *Cell* **2006**, *124* (4), 729–740. https://doi.org/10.1016/j.cell.2006.02.007.

(166) Meier, C.; Weil, T.; Kirchhoff, F.; Münch, J. Peptide Nanofibrils as Enhancers of Retroviral Gene Transfer. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2014**, *6* (5), 438–451. https://doi.org/10.1002/wnan.1275.

(167) Zhang, Z.; Qiu, N.; Wu, S.; Liu, X.; Zhou, Z.; Tang, J.; Liu, Y.; Zhou, R.; Shen, Y. Dose-Independent Transfection of Hydrophobized Polyplexes. *Adv. Mater.* **2021**, *33* (25), 2102219. https://doi.org/10.1002/adma.202102219.

(168) Hack, F. J.; Cokca, C.; Städter, S.; Hülsmann, J.; Peneva, K.; Fischer, D. Indole, Phenyl, and Phenol Groups:

The Role of the Comonomer on Gene Delivery in Guanidinium Containing Methacrylamide Terpolymers. *Macromol. Rapid Commun.* **2021**, *42* (8), 2000580. https://doi.org/10.1002/marc.202000580.

(169) Yang, Y.; Jalali, S.; Nilsson, B. L.; Dias, C. L. Binding Mechanisms of Amyloid-like Peptides to Lipid Bilayers and Effects of Divalent Cations. *ACS Chem. Neurosci.* **2021**, *12* (11), 2027–2035. https://doi.org/10.1021/acschemneuro.1c00140.

(170) Capule, C. C.; Brown, C.; Olsen, J. S.; Dewhurst, S.; Yang, J. Oligovalent Amyloid-Binding Agents Reduce SEVI-Mediated Enhancement of HIV-1 Infection. *J. Am. Chem. Soc.* **2012**, *134* (2), 905–908. https://doi.org/10.1021/ja210931b.

(171) Purvis, J. E.; Lahav, G. Encoding and Decoding Cellular Information through Signaling Dynamics. *Cell* **2013**, *152* (5), 945–956. https://doi.org/10.1016/j.cell.2013.02.005.

(172) Mierke, C. T. Bidirectional Mechanical Response Between Cells and Their Microenvironment. *Front. Phys.* **2021**, *9*, 619. https://doi.org/10.3389/fphy.2021.749830.

(173) Sur, S.; Matson, J. B.; Webber, M. J.; Newcomb, C. J.; Stupp, S. I. Photodynamic Control of Bioactivity in a Nanofiber Matrix. *ACS Nano* **2012**, *6* (12), 10776–10785. https://doi.org/10.1021/nn304101x.

(174) Wheeldon, I.; Farhadi, A.; Bick, A. G.; Jabbari, E.; Khademhosseini, A. Nanoscale Tissue Engineering: Spatial Control over Cell-Materials Interactions. *Nanotechnology* **2011**, *22* (21), 212001. https://doi.org/10.1088/0957-4484/22/21/212001.

(175) Ye, G.; Bao, F.; Zhang, X.; Song, Z.; Liao, Y.; Fei, Y.; Bunpetch, V.; Heng, B. C.; Shen, W.; Liu, H.; Zhou, J.; Ouyang, H. Nanomaterial-Based Scaffolds for Bone Tissue Engineering and Regeneration. *Nanomedicine* **2020**, *15* (20), 1995–2017. https://doi.org/10.2217/nnm-2020-0112.

(176) Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J. Biomimetic Peptide Self-Assembly for Functional Materials. *Nat. Rev. Chem.* **2020**, *4* (11), 615–634. https://doi.org/10.1038/s41570-020-0215-y.

(177) Wu, E. C.; Zhang, S.; Hauser, C. A. E. Self-Assembling Peptides as Cell-Interactive Scaffolds. *Adv. Funct. Mater.* **2012**, *22* (3), 456–468. https://doi.org/10.1002/adfm.201101905.

(178) Ender, A. M.; Kaygisiz, K.; Räder, H.-J.; Mayer, F. J.; Synatschke, C. V.; Weil, T. Cell-Instructive Surface Gradients of Photoresponsive Amyloid-like Fibrils. *ACS Biomater. Sci. Eng.* **2021**, *7* (10), 4798–4808. https://doi.org/10.1021/acsbiomaterials.1c00889.

(179) Pelham, R. J.; Wang, Y. Cell Locomotion and Focal Adhesions Are Regulated by Substrate Flexibility. *Proc. Natl. Acad. Sci.* **1997**, *94* (25), 13661–13665. https://doi.org/10.1073/pnas.94.25.13661.

(180) Chagri, S.; Ng, D. Y. W.; Weil, T. Designing Bioresponsive Nanomaterials for Intracellular Self-Assembly. *Nat. Rev. Chem.* **2022**, *6* (5), 320–338. https://doi.org/10.1038/s41570-022-00373-x.

(181) Pieszka, M.; Han, S.; Volkmann, C.; Graf, R.; Lieberwirth, I.; Landfester, K.; Ng, D. Y. W.; Weil, T. Controlled Supramolecular Assembly Inside Living Cells by Sequential Multistaged Chemical Reactions. *J. Am. Chem. Soc.* **2020**, *142* (37), 15780–15789. https://doi.org/10.1021/jacs.0c05261.

(182) Lin, Y.-A.; Ou, Y.-C.; Cheetham, A. G.; Cui, H. Rational Design of MMP Degradable Peptide-Based Supramolecular Filaments. *Biomacromolecules* **2014**, *15* (4), 1419–1427. https://doi.org/10.1021/bm500020j.

(183) Nespovitaya, N.; Gath, J.; Barylyuk, K.; Seuring, C.; Meier, B. H.; Riek, R. Dynamic Assembly and Disassembly of Functional β-Endorphin Amyloid Fibrils. *J. Am. Chem. Soc.* **2016**, *138* (3), 846–856. https://doi.org/10.1021/jacs.5b08694.

(184) Cao, P.; Raleigh, D. P. Analysis of the Inhibition and Remodeling of Islet Amyloid Polypeptide Amyloid Fibers by Flavanols. *Biochemistry* **2012**, *51* (13), 2670–2683. https://doi.org/10.1021/bi2015162.

(185) Peddie, V.; Abell, A. D. Photocontrol of Peptide Secondary Structure through Non-Azobenzene Photoswitches. *J. Photochem. Photobiol. C Photochem. Rev.* **2019**, *40*, 1–20. https://doi.org/10.1016/j.jphotochemrev.2019.05.001.

(186) Deeg, A. A.; Schrader, T. E.; Kempter, S.; Pfizer, J.; Moroder, L.; Zinth, W. Light-Triggered Aggregation and Disassembly of Amyloid-like Structures. *ChemPhysChem* **2011**, *12* (3), 559–562. https://doi.org/10.1002/cphc.201001012.

(187) Nakamura, K.; Tanaka, W.; Sada, K.; Kubota, R.; Aoyama, T.; Urayama, K.; Hamachi, I. Phototriggered Spatially Controlled Out-of-Equilibrium Patterns of Peptide Nanofibers in a Self-Sorting Double Network Hydrogel. *J. Am. Chem. Soc.* **2021**, *143* (46), 19532–19541. https://doi.org/10.1021/jacs.1c09172.

(188) Wegner, S. V.; Sentürk, O. I.; Spatz, J. P. Photocleavable Linker for the Patterning of Bioactive Molecules. *Sci. Rep.* **2016**, *5* (1), 18309. https://doi.org/10.1038/srep18309.

(189) Luo, Y.; Shoichet, M. S. A Photolabile Hydrogel for Guided Three-Dimensional Cell Growth and Migration. *Nat. Mater.* **2004**, *3* (4), 249–253. https://doi.org/10.1038/nmat1092.

(190) Matsuda, T.; Sugawara, T. Development of Surface Photochemical Modification Method for Micropatterning of Cultured Cells. *J. Biomed. Mater. Res.* **1995**, *29* (6), 749–756. https://doi.org/10.1002/jbm.820290611.

(191) Ryan, D.; Parviz, B. A.; Linder, V.; Semetey, V.; Sia, S. K.; Su, J.; Mrksich, M.; Whitesides, G. M. Patterning Multiple Aligned Self-Assembled Monolayers Using Light. *Langmuir* **2004**, *20* (21), 9080–9088. https://doi.org/10.1021/la048443u.

(192) Zhang, S.; Yan, L.; Altman, M.; Lässle, M.; Nugent, H.; Frankel, F.; Lauffenburger, D. A.; Whitesides, G. M.; Rich, A. Biological Surface Engineering: A Simple System for Cell Pattern Formation. *Biomaterials* **1999**, *20* (13), 1213–1220. https://doi.org/10.1016/S0142-9612(99)00014-9.

(193) Kloxin, A. M.; Kasko, A. M.; Salinas, C. N.; Anseth, K. S. Photodegradable Hydrogels for Dynamic Tuning of Physical and Chemical Properties. *Science* **2009**, *324* (5923), 59–63.

https://doi.org/10.1126/science.1169494.
(194) Yavitt, F. M.; Kirkpatrick, B. E.; Blatchley, M. R.; Anseth, K. S. 4D Materials with Photoadaptable Properties Instruct and Enhance Intestinal Organoid Development. *ACS Biomater. Sci. Eng.* **2022**, acsbiomaterials.1c01450. https://doi.org/10.1021/acsbiomaterials.1c01450.
(195) Kane, R. S.; Takayama, S.; Ostuni, E.; Ingber, D. E.; Whitesides, G. M. Patterning Proteins and Cells Using Soft Lithography. *Biomaterials* **1999**, *20* (23–24), 2363–2376. https://doi.org/10.1016/s0142-9612(99)00165-9.
(196) Li, Y.; Xiao, Y.; Liu, C. The Horizon of Materiobiology: A Perspective on Material-Guided Cell Behaviors and Tissue Engineering. *Chem. Rev.* **2017**, *117* (5), 4376–4421. https://doi.org/10.1021/acs.chemrev.6b00654.
(197) d'Alessandro, J.; Barbier--Chebbah, A.; Cellerin, V.; Benichou, O.; Mège, R. M.; Voituriez, R.; Ladoux, B. Cell Migration Guided by Long-Lived Spatial Memory. *Nat. Commun.* **2021**, *12* (1), 4118. https://doi.org/10.1038/s41467-021-24249-8.
(198) Brückner, D. B.; Arlt, N.; Fink, A.; Ronceray, P.; Rädler, J. O.; Broedersz, C. P. Learning the Dynamics of Cell–Cell Interactions in Confined Cell Migration. *Proc. Natl. Acad. Sci.* **2021**, *118* (7), e2016602118. https://doi.org/10.1073/pnas.2016602118.
(199) Matson, J. B.; Navon, Y.; Bitton, R.; Stupp, S. I. Light-Controlled Hierarchical Self-Assembly of Polyelectrolytes and Supramolecular Polymers. *ACS Macro Lett.* **2015**, *4* (1), 43–47. https://doi.org/10.1021/mz500677q.
(200) Schleeger, M.; Vandenakker, C. C.; Deckert-Gaudig, T.; Deckert, V.; Velikov, K. P.; Koenderink, G.; Bonn, M. Amyloids: From Molecular Structure to Mechanical Properties. *Polymer (Guildf).* **2013**, *54* (10), 2473–2488. https://doi.org/10.1016/j.polymer.2013.02.029.
(201) Dehsorkhi, A.; Hamley, I. W.; Seitsonen, J.; Ruokolainen, J. Tuning Self-Assembled Nanostructures Through Enzymatic Degradation of a Peptide Amphiphile. *Langmuir* **2013**, *29* (22), 6665–6672. https://doi.org/10.1021/la401025r.
(202) Mazza, M.; Patel, A.; Pons, R.; Bussy, C.; Kostarelos, K. Peptide Nanofibres as Molecular Transporters: From Self-Assembly to in Vivo Degradation. *Faraday Discuss.* **2013**, *166* (0), 181. https://doi.org/10.1039/c3fd00100h.
(203) Jun, H.-W.; Yuwono, V.; Paramonov, S. E.; Hartgerink, J. D. Enzyme-Mediated Degradation of Peptide-Amphiphile Nanofiber Networks. *Adv. Mater.* **2005**, *17* (21), 2612–2617. https://doi.org/10.1002/adma.200500855.
(204) Missirlis, D.; Khant, H.; Tirrell, M. Mechanisms of Peptide Amphiphile Internalization by SJSA-1 Cells in Vitro. *Biochemistry* **2009**, *48* (15), 3304–3314. https://doi.org/10.1021/bi802356k.
(205) Missirlis, D.; Teesalu, T.; Black, M.; Tirrell, M. The Non-Peptidic Part Determines the Internalization Mechanism and Intracellular Trafficking of Peptide Amphiphiles. *PLoS One* **2013**, *8* (1), e54611. https://doi.org/10.1371/journal.pone.0054611.
(206) Schnorenberg, M. R.; Bellairs, J. A.; Samaeekia, R.; Acar, H.; Tirrell, M. V.; LaBelle, J. L. Activating the Intrinsic Pathway of Apoptosis Using BIM BH3 Peptides Delivered by Peptide Amphiphiles with Endosomal Release. *Materials* **2019**, *12* (16), 2567. https://doi.org/10.3390/ma12162567.
(207) Wagoner, V. A.; Cheon, M.; Chang, I.; Hall, C. K. Impact of Sequence on the Molecular Assembly of Short Amyloid Peptides. *Proteins Struct. Funct. Bioinforma.* **2014**, *82* (7), 1469–1483. https://doi.org/10.1002/prot.24515.
(208) Betush, R. J.; Urban, J. M.; Nilsson, B. L. Balancing Hydrophobicity and Sequence Pattern to Influence Self-Assembly of Amphipathic Peptides. *Pept. Sci.* **2018**, *110* (1), e23099. https://doi.org/10.1002/bip.23099.
(209) Deutsche Forschungsgemeinschaft. DFG - GEPRIS - SFB 1279: Exploiting the Human Peptidome for Novel Antimicrobial and Anticancer Agents https://gepris.dfg.de/gepris/projekt/316249678?language=en (accessed Jul 16, 2023).
(210) Noschka, R.; Gerbl, F.; Löffler, F.; Kubis, J.; Rodríguez, A. A.; Mayer, D.; Grieshober, M.; Holch, A.; Raasholm, M.; Forssmann, W.-G.; Spellerberg, B.; Wiese, S.; Weidinger, G.; Ständker, L.; Stenger, S. Unbiased Identification of Angiogenin as an Endogenous Antimicrobial Protein With Activity Against Virulent Mycobacterium Tuberculosis. *Front. Microbiol.* **2021**, *11*, 618278. https://doi.org/10.3389/fmicb.2020.618278.
(211) Buske, C.; Kirchhoff, F.; Münch, J. EPI-X4, a Novel Endogenous Antagonist of CXCR4. *Oncotarget* **2015**, *6* (34), 35137–35138. https://doi.org/10.18632/oncotarget.6037.
(212) Meier, C.; Weil, T.; Kirchhoff, F.; Münch, J. Peptide Nanofibrils as Enhancers of Retroviral Gene Transfer. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2014**, *6* (5), 438–451. https://doi.org/10.1002/wnan.1275.
(213) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154. https://doi.org/10.1021/ja00897a025.
(214) Sinha, N. J.; Langenstein, M. G.; Pochan, D. J.; Kloxin, C. J.; Saven, J. G. Peptide Design and Self-Assembly into Targeted Nanostructure and Functional Materials. *Chem. Rev.* **2021**, *121* (22), 13915–13935. https://doi.org/10.1021/acs.chemrev.1c00712.
(215) Erdélyi, M.; Gogoll, A. Rapid Microwave-Assisted Solid Phase Peptide Synthesis. *Synthesis (Stuttg).* **2002**, No. 11, 1592–1596. https://doi.org/10.1055/s-2002-33348.
(216) Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S. High-Efficiency Solid Phase Peptide Synthesis (HE - SPPS). *Org. Lett.* **2014**, *16* (3), 940–943. https://doi.org/10.1021/ol4036825.
(217) Egerton, R. F. The Transmission Electron Microscope. In *Physical Principles of Electron Microscopy*; Springer International Publishing: Cham, 2016; pp 55–88. https://doi.org/10.1007/978-3-319-39877-8_3.

(218) Guo, F.; Jiang, W. Single Particle Cryo-Electron Microscopy and 3-D Reconstruction of Viruses; 2014; pp 401–443. https://doi.org/10.1007/978-1-62703-776-1_19.
(219) Simone Ruggeri, F.; Habchi, J.; Cerreta, A.; Dietler, G. AFM-Based Single Molecule Techniques: Unraveling the Amyloid Pathogenic Species. *Curr. Pharm. Des.* **2016**, *22* (26), 3950–3970. https://doi.org/10.2174/1381612822666160518141911.
(220) Ruggeri, F. S.; Longo, G.; Faggiano, S.; Lipiec, E.; Pastore, A.; Dietler, G. Infrared Nanospectroscopy Characterization of Oligomeric and Fibrillar Aggregates during Amyloid Formation. *Nat. Commun.* **2015**, *6* (1), 7831. https://doi.org/10.1038/ncomms8831.
(221) Ruggeri, F. S.; Šneideris, T.; Vendruscolo, M.; Knowles, T. P. J. Atomic Force Microscopy for Single Molecule Characterisation of Protein Aggregation. *Arch. Biochem. Biophys.* **2019**, *664*, 134–148. https://doi.org/10.1016/j.abb.2019.02.001.
(222) Kubota, R.; Tanaka, W.; Hamachi, I. Microscopic Imaging Techniques for Molecular Assemblies: Electron, Atomic Force, and Confocal Microscopies. *Chem. Rev.* **2021**, *121* (22), 14281–14347. https://doi.org/10.1021/acs.chemrev.0c01334.
(223) Kubitscheck, U. Principles of Light Microscopy. In *Fluorescence Microscopy*; Wiley, 2013; pp 33–95. https://doi.org/10.1002/9783527671595.ch2.
(224) Huff, J. The Airyscan Detector from ZEISS: Confocal Imaging with Improved Signal-to-Noise Ratio and Super-Resolution. *Nat. Methods* **2015**, *12* (12), i–ii. https://doi.org/10.1038/nmeth.f.388.
(225) Kubota, R.; Nakamura, K.; Torigoe, S.; Hamachi, I. The Power of Confocal Laser Scanning Microscopy in Supramolecular Chemistry: In Situ Real-time Imaging of Stimuli-Responsive Multicomponent Supramolecular Hydrogels. *ChemistryOpen* **2020**, *9* (1), 67–79. https://doi.org/10.1002/open.201900328.
(226) Ruysschaert, J.-M.; Raussens, V. ATR-FTIR Analysis of Amyloid Proteins. In *Peptide Self-Assembly - Methods in Molecular Biology*; Nilsson, B. L., Doran, T. M., Eds.; Springer New York, 2018; pp 69–81. https://doi.org/10.1007/978-1-4939-7811-3_3.
(227) Zandomeneghi, G.; Krebs, M. R. H.; McCammon, M. G.; Fändrich, M. FTIR Reveals Structural Differences between Native β-Sheet Proteins and Amyloid Fibrils. *Protein Sci.* **2009**, *13* (12), 3314–3321. https://doi.org/10.1110/ps.041024904.
(228) Ruggeri, F. S.; Mannini, B.; Schmid, R.; Vendruscolo, M.; Knowles, T. P. J. Single Molecule Secondary Structure Determination of Proteins through Infrared Absorption Nanospectroscopy. *Nat. Commun.* **2020**, *11* (1), 1–9. https://doi.org/10.1038/s41467-020-16728-1.
(229) Adochitei, A.; Drochioiu, G. Rapid Characterization of Peptide Secondary Structure by FT-IR Spectroscopy. *Rev. Roum. Chim.* **2011**, *56* (8), 783–791.
(230) Bulheller, B. M.; Rodger, A.; Hirst, J. D. Circular and Linear Dichroism of Proteins. *Phys. Chem. Chem. Phys.* **2007**, *9* (17), 2020. https://doi.org/10.1039/b615870f.
(231) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2007**, *1* (6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.
(232) Naiki, H.; Higuchi, K.; Hosokawa, M.; Takeda, T. Fluorometric Determination of Amyloid Fibrils in Vitro Using the Fluorescent Dye, Thioflavine T. *Anal. Biochem.* **1989**, *177* (2), 244–249. https://doi.org/10.1016/0003-2697(89)90046-8.
(233) Sidhu, A.; Vaneyck, J.; Blum, C.; Segers-Nolten, I.; Subramaniam, V. Polymorph-Specific Distribution of Binding Sites Determines Thioflavin-T Fluorescence Intensity in α-Synuclein Fibrils. *Amyloid* **2018**, *25* (3), 189–196. https://doi.org/10.1080/13506129.2018.1517736.
(234) Arad, E.; Green, H.; Jelinek, R.; Rapaport, H. Revisiting Thioflavin T (ThT) Fluorescence as a Marker of Protein Fibrillation – The Prominent Role of Electrostatic Interactions. *J. Colloid Interface Sci.* **2020**, *573*, 87–95. https://doi.org/10.1016/j.jcis.2020.03.075.
(235) Namioka, S.; Yoshida, N.; Konno, H.; Makabe, K. Residue-Specific Binding Mechanisms of Thioflavin T to a Surface of Flat β-Sheets within a Peptide Self-Assembly Mimic. *Biochemistry* **2020**, *59* (30), 2782–2787. https://doi.org/10.1021/acs.biochem.0c00280.
(236) Malm, A. V.; Corbett, J. C. W. Improved Dynamic Light Scattering Using an Adaptive and Statistically Driven Time Resolved Treatment of Correlation Data. *Sci. Rep.* **2019**, *9* (1), 1–11. https://doi.org/10.1038/s41598-019-50077-4.
(237) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys. Rev.* **2016**, *8* (4), 409–427. https://doi.org/10.1007/s12551-016-0218-6.
(238) Thomas Luxbacher. *The Zeta Potential for Solid Surface Analysis*; Anton Paar GmbH, 2014.
(239) Martinek, T. A.; Hetényi, A.; Fülöp, L.; Mándity, I. M.; Tóth, G. K.; Dékány, I.; Fülöp, F. Secondary Structure Dependent Self-Assembly of β-Peptides into Nanosized Fibrils and Membranes. *Angew. Chemie* **2006**, *118* (15), 2456–2460. https://doi.org/10.1002/ange.200504158.
(240) Toksoz, S.; Mammadov, R.; Tekinay, A. B.; Guler, M. O. Electrostatic Effects on Nanofiber Formation of Self-Assembling Peptide Amphiphiles. *J. Colloid Interface Sci.* **2011**, *356* (1), 131–137. https://doi.org/10.1016/j.jcis.2010.12.076.
(241) Chen, Y.; Gan, H. X.; Tong, Y. W. PH-Controlled Hierarchical Self-Assembly of Peptide Amphiphile. *Macromolecules* **2015**, *48* (8), 2647–2653. https://doi.org/10.1021/ma502572w.
(242) Takeuchi, Y.; McClure, M. O.; Pizzato, M. Identification of Gammaretroviruses Constitutively Released from Cell Lines Used for Human Immunodeficiency Virus Research. *J. Virol.* **2008**, *82* (24), 12585–12588. https://doi.org/10.1128/jvi.01726-08.
(243) Bronstein, I.; Fortin, J.; Stanley, P. E.; Stewart, G. S. A. B.; Kricka, L. J. Chemiluminescent and Bioluminescent

Reporter Gene Assays. *Anal. Biochem.* **1994**, *219* (2), 169–181. https://doi.org/10.1006/abio.1994.1254.
(244) Allen, T. M.; Cleland, L. G. Serum-Induced Leakage of Liposome Contents. *Biochim. Biophys. Acta - Biomembr.* **1980**, *597* (2), 418–426. https://doi.org/10.1016/0005-2736(80)90118-2.
(245) Mosmann, T. Rapid Colorimetric Assay for Cellular Growth and Survival: Application to Proliferation and Cytotoxicity Assays. *J. Immunol. Methods* **1983**, *65* (1–2), 55–63. https://doi.org/10.1016/0022-1759(83)90303-4.
(246) Tuttle, T. Computational Approaches to Understanding the Self-Assembly of Peptide-Based Nanostructures. *Isr. J. Chem.* **2015**, *55* (6–7), 724–734. https://doi.org/10.1002/ijch.201400188.
(247) Navarro, S.; Ventura, S. Computational Methods to Predict Protein Aggregation. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102343. https://doi.org/10.1016/j.sbi.2022.102343.
(248) Lee, O.-S.; Stupp, S. I.; Schatz, G. C. Atomistic Molecular Dynamics Simulations of Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers. *J. Am. Chem. Soc.* **2011**, *133* (10), 3677–3683. https://doi.org/10.1021/ja110966y.
(249) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.
(250) Fernandez-Escamilla, A.-M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. Prediction of Sequence-Dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nat. Biotechnol.* **2004**, *22* (10), 1302–1306. https://doi.org/10.1038/nbt1012.
(251) Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. Rationalization of the Effects of Mutations on Peptide Andprotein Aggregation Rates. *Nature* **2003**, *424* (6950), 805–808. https://doi.org/10.1038/nature01891.
(252) Conchillo-Solé, O.; de Groot, N. S.; Avilés, F. X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: A Server for the Prediction and Evaluation of "Hot Spots" of Aggregation in Polypeptides. *BMC Bioinformatics* **2007**, *8* (1), 65. https://doi.org/10.1186/1471-2105-8-65.
(253) Dutta, A.; Bereau, T.; Vilgis, T. A. Identifying Sequential Residue Patterns in Bitter and Umami Peptides. *ACS Food Sci. Technol.* **2022**, *2* (11), 1773–1780. https://doi.org/10.1021/acsfoodscitech.2c00251.
(254) Betts, M. J.; Russell, R. B. Amino-Acid Properties and Consequences of Substitutions. In *Bioinformatics for Geneticists*; John Wiley & Sons, Ltd: Chichester, UK, 2007; pp 311–342. https://doi.org/10.1002/9780470059180.ch13.
(255) FAUCHÈRE, J.-L.; CHARTON, M.; KIER, L. B.; VERLOOP, A.; PLISKA, V. Amino Acid Side Chain Parameters for Correlation Studies in Biology and Pharmacology. *Int. J. Pept. Protein Res.* **1988**, *32* (4), 269–278. https://doi.org/10.1111/j.1399-3011.1988.tb01261.x.
(256) Schilling, C.; Mack, T.; Lickfett, S.; Sieste, S.; Ruggeri, F. S.; Sneideris, T.; Dutta, A.; Bereau, T.; Naraghi, R.; Sinske, D.; Knowles, T. P. J.; Synatschke, C. V.; Weil, T.; Knöll, B. Sequence-Optimized Peptide Nanofibers as Growth Stimulators for Regeneration of Peripheral Neurons. *Adv. Funct. Mater.* **2019**, *29* (24), 1809112. https://doi.org/10.1002/adfm.201809112.
(257) De Groot, N. S.; Castillo, V.; Graña-Montes, R.; Ventura, S. AGGRESCAN: Method, Application, and Perspectives for Drug Design. *Methods Mol. Biol.* **2012**, *819*, 199–220. https://doi.org/10.1007/978-1-61779-465-0_14.
(258) Maurer-Stroh, S.; Debulpaep, M.; Kuemmerer, N.; de la Paz, M. L.; Martins, I. C.; Reumers, J.; Morris, K. L.; Copland, A.; Serpell, L.; Serrano, L.; Schymkowitz, J. W. H.; Rousseau, F. Exploring the Sequence Determinants of Amyloid Structure Using Position-Specific Scoring Matrices. *Nat. Methods* **2010**, *7* (3), 237–242. https://doi.org/10.1038/nmeth.1432.
(259) Família, C.; Dennison, S. R.; Quintas, A.; Phoenix, D. A. Prediction of Peptide and Protein Propensity for Amyloid Formation. *PLoS One* **2015**, *10* (8), e0134679. https://doi.org/10.1371/journal.pone.0134679.
(260) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824. https://doi.org/10.1021/jp071097f.
(261) Frederix, P. W. J. M.; Ulijn, R. V.; Hunt, N. T.; Tuttle, T. Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J. Phys. Chem. Lett.* **2011**, *2* (19), 2380–2384. https://doi.org/10.1021/jz2010573.
(262) Frederix, P. W. J. M.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. Exploring the Sequence Space for (Tri-)Peptide Self-Assembly to Design and Discover New Hydrogels. *Nat. Chem.* **2015**, *7* (1), 30–37. https://doi.org/10.1038/nchem.2122.
(263) Lee, O.-S.; Cho, V.; Schatz, G. C. Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics. *Nano Lett.* **2012**, *12* (9), 4907–4913. https://doi.org/10.1021/nl302487m.
(264) Kang, M.; Cui, H.; Loverde, S. M. Coarse-Grained Molecular Dynamics Studies of the Structure and Stability of Peptide-Based Drug Amphiphile Filaments. *Soft Matter* **2017**, *13* (42), 7721–7730. https://doi.org/10.1039/C7SM00943G.
(265) Fu, I. W.; Markegard, C. B.; Chu, B. K.; Nguyen, H. D. Role of Hydrophobicity on Self-Assembly by Peptide Amphiphiles via Molecular Dynamics Simulations. *Langmuir* **2014**, *30* (26), 7745–7754.

https://doi.org/10.1021/la5012988.
(266) Sasselli, I. R.; Syrgiannis, Z.; Sather, N. A.; Palmer, L. C.; Stupp, S. I. Modeling Interactions within and between Peptide Amphiphile Supramolecular Filaments. *J. Phys. Chem. B* **2022**, *126* (3), 650–659. https://doi.org/10.1021/acs.jpcb.1c09258.
(267) Lombardi, L. E.; Martí, M. A.; Capece, L. CG2AA: Backmapping Protein Coarse-Grained Structures. *Bioinformatics* **2016**, *32* (8), 1235–1237. https://doi.org/10.1093/bioinformatics/btv740.
(268) van der Kant, R.; Louros, N.; Schymkowitz, J.; Rousseau, F. Thermodynamic Analysis of Amyloid Fibril Structures Reveals a Common Framework for Stability in Amyloid Polymorphs. *Structure* **2022**, *30* (8), 1178-1189.e3. https://doi.org/10.1016/j.str.2022.05.002.
(269) Iscen, A.; Schatz, G. C. Peptide Amphiphile Self-Assembly. *EPL (Europhysics Lett.* **2017**, *119* (3), 38002. https://doi.org/10.1209/0295-5075/119/38002.
(270) Simons, K. T.; Kooperberg, C.; Huang, E.; Baker, D. Assembly of Protein Tertiary Structures from Fragments with Similar Local Sequences Using Simulated Annealing and Bayesian Scoring Functions. *J. Mol. Biol.* **1997**, *268* (1), 209–225. https://doi.org/10.1006/jmbi.1997.0959.
(271) Baek, M.; Baker, D. Deep Learning and Protein Structure Modeling. *Nat. Methods* **2022**, *19* (1), 13–14. https://doi.org/10.1038/s41592-021-01360-8.
(272) Ferruz, N.; Schmidt, S.; Höcker, B. ProtGPT2 Is a Deep Unsupervised Language Model for Protein Design. *Nat. Commun.* **2022**, *13* (1), 4348. https://doi.org/10.1038/s41467-022-32007-7.
(273) Hepler-Smith, E. Systematic Flexibility and the History of the IUPAC Nomenclature of Organic Chemistry. *Chem. Int.* **2015**, *37* (2). https://doi.org/10.1515/ci-2015-0232.
(274) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31–36. https://doi.org/10.1021/ci00057a005.
(275) Lin, T.-S.; Coley, C. W.; Mochigase, H.; Beech, H. K.; Wang, W.; Wang, Z.; Woods, E.; Craig, S. L.; Johnson, J. A.; Kalow, J. A.; Jensen, K. F.; Olsen, B. D. BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules. *ACS Cent. Sci.* **2019**, *5* (9), 1523–1531. https://doi.org/10.1021/acscentsci.9b00476.
(276) Krenn, M.; Ai, Q.; Barthel, S.; Carson, N.; Frei, A.; Frey, N. C.; Friederich, P.; Gaudin, T.; Gayle, A. A.; Jablonka, K. M.; Lameiro, R. F.; Lemm, D.; Lo, A.; Moosavi, S. M.; Nápoles-Duarte, J. M.; Nigam, A.; Pollice, R.; Rajan, K.; Schatzschneider, U.; Schwaller, P.; Skreta, M.; Smit, B.; Strieth-Kalthoff, F.; Sun, C.; Tom, G.; Falk von Rudorff, G.; Wang, A.; White, A. D.; Young, A.; Yu, R.; Aspuru-Guzik, A. SELFIES and the Future of Molecular String Representations. *Patterns* **2022**, *3* (10), 100588. https://doi.org/10.1016/j.patter.2022.100588.
(277) White, A. D. The Future of Chemistry Is Language. *Nat. Rev. Chem.* **2023**, *7* (7), 457–458. https://doi.org/10.1038/s41570-023-00502-0.
(278) Grisoni, F. Chemical Language Models for de Novo Drug Design: Challenges and Opportunities. *Curr. Opin. Struct. Biol.* **2023**, *79*, 102527. https://doi.org/10.1016/j.sbi.2023.102527.
(279) Wu, C.; Gao, R.; Zhang, Y.; De Marinis, Y. PTPD: Predicting Therapeutic Peptides by Deep Learning and Word2vec. *BMC Bioinformatics* **2019**, *20* (1), 456. https://doi.org/10.1186/s12859-019-3006-z.
(280) Strieth-Kalthoff, F.; Sandfort, F.; Segler, M. H. S.; Glorius, F. Machine Learning the Ropes: Principles, Applications and Directions in Synthetic Chemistry. *Chemical Society Reviews*. Royal Society of Chemistry September 7, 2020, pp 6154–6168. https://doi.org/10.1039/c9cs00786e.
(281) Ofer, D.; Brandes, N.; Linial, M. The Language of Proteins: NLP, Machine Learning & Protein Sequences. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1750–1758. https://doi.org/10.1016/j.csbj.2021.03.022.
(282) Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; Kuksa, P. Natural Language Processing (Almost) from Scratch. *J. Mach. Learn. Res.* **2011**, *12*, 2493–2537. https://doi.org/10.5555/1953048.2078186.
(283) Tshitoyan, V.; Dagdelen, J.; Weston, L.; Dunn, A.; Rong, Z.; Kononova, O.; Persson, K. A.; Ceder, G.; Jain, A. Unsupervised Word Embeddings Capture Latent Knowledge from Materials Science Literature. *Nature* **2019**, *571* (7763), 95–98. https://doi.org/10.1038/s41586-019-1335-8.
(284) Burdukiewicz, M.; Sobczyk, P.; Rödiger, S.; Duda-Madej, A.; Mackiewicz, P.; Kotulska, M. Amyloidogenic Motifs Revealed by N-Gram Analysis. *Sci. Rep.* **2017**, *7* (1), 12961. https://doi.org/10.1038/s41598-017-13210-9.
(285) Asgari, E.; Mofrad, M. R. K. Continuous Distributed Representation of Biological Sequences for Deep Proteomics and Genomics. *PLoS One* **2015**, *10* (11), e0141287. https://doi.org/10.1371/journal.pone.0141287.
(286) Van Der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9* (86), 2579–2605.
(287) Bereau, T. Computational Compound Screening of Biomolecules and Soft Materials by Molecular Simulations. *Model. Simul. Mater. Sci. Eng.* **2020**, *29* (2), 023001. https://doi.org/10.1088/1361-651X/abd042.
(288) Capito, R. M.; Azevedo, H. S.; Velichko, Y. S.; Mata, A.; Stupp, S. I. Self-Assembly of Large and Small Molecules into Hierarchically Ordered Sacs and Membranes. *Science* **2008**, *319* (5871), 1812–1816. https://doi.org/10.1126/science.1154586.
(289) Franke, R.; Hirsch, T.; Eichler, J. A Rationally Designed Synthetic Mimic of the Discontinuous CD4-Binding Site of HIV-1 Gp120. *J. Recept. Signal Transduct.* **2006**, *26* (5–6), 453–460. https://doi.org/10.1080/10799890600923179.
(290) Kaygisiz, K.; Rauch-Wirth, L.; Dutta, A.; Yu, X.; Nagata, Y.; Bereau, T.; Münch, J.; Synatschke, C. V.; Weil, T. Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides. *Nat. Commun.*, *accepted* **2023**. https://doi.org/10.1038/s41467-023-40663-6.

(291) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9* (1), 687–697. https://doi.org/10.1021/ct300646g.

(292) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ 1 and χ 2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273. https://doi.org/10.1021/ct300400x.

(293) Kaygisiz, K.; Dutta, A.; Rauch-Wirth, L.; Synatschke, C. V.; Münch, J.; Bereau, T.; Weil, T. Inverse Design of Viral Infectivity-Enhancing Peptide Fibrils from Continuous Protein-Vector Embeddings. *Biomater. Sci.* **2023**, *11* (15), 5251–5261. https://doi.org/10.1039/D3BM00412K.

(294) Kaygisiz, K.; Ender, A. M.; Gačanin, J.; Kaczmarek, L. A.; Koutsouras, D. A.; Nalakath, A. N.; Winterwerber, P.; Mayer, F. J.; Räder, H.; Marszalek, T.; Blom, P. W. M.; Synatschke, C. V.; Weil, T. Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning. *Macromol. Biosci.* **2023**, *23* (2), 2200294. https://doi.org/10.1002/mabi.202200294.

(295) Müller, J. A. The Versatile Role of Semen and Seminal Amyloid in Virus Infection, Ulm University, 2016.

(296) Wei, W.; Petrone, L.; Tan, Y.; Cai, H.; Israelachvili, J. N.; Miserez, A.; Waite, J. H. An Underwater Surface-Drying Peptide Inspired by a Mussel Adhesive Protein. *Adv. Funct. Mater.* **2016**, *26* (20), 3496–3507. https://doi.org/10.1002/adfm.201600210.

(297) Shao, L.; Ma, J.; Prelesnik, J. L.; Zhou, Y.; Nguyen, M.; Zhao, M.; Jenekhe, S. A.; Kalinin, S. V.; Ferguson, A. L.; Pfaendtner, J.; Mundy, C. J.; De Yoreo, J. J.; Baneyx, F.; Chen, C.-L. Hierarchical Materials from High Information Content Macromolecular Building Blocks: Construction, Dynamic Interventions, and Prediction. *Chem. Rev.* **2022**, *122* (24), 17397–17478. https://doi.org/10.1021/acs.chemrev.2c00220.

(298) Batra, R.; Loeffler, T. D.; Chan, H.; Srinivasan, S.; Cui, H.; Korendovych, I. V.; Nanda, V.; Palmer, L. C.; Solomon, L. A.; Fry, H. C.; Sankaranarayanan, S. K. R. S. Machine Learning Overcomes Human Bias in the Discovery of Self-Assembling Peptides. *Nat. Chem.* **2022**, *14* (12), 1427–1435. https://doi.org/10.1038/s41557-022-01055-3.

(299) Merindol, R.; Walther, A. Materials Learning from Life: Concepts for Active, Adaptive and Autonomous Molecular Systems. *Chemical Society Reviews*. Royal Society of Chemistry September 21, 2017, pp 5588–5619. https://doi.org/10.1039/c6cs00738d.

(300) Pennisi, E. How do giant flocks of birds stay in such tight formations? https://www.science.org/content/article/how-do-giant-flocks-birds-stay-such-tight-formations (accessed Jun 12, 2023).

(301) Price, R. W. Neurological Complications of HIV Infection. *Lancet (London, England)* **1996**, *348* (9025), 445–452. https://doi.org/10.1016/S0140-6736(95)11035-6.

(302) Kedersha, N. Microscopic Sensors Reveal the Brain's Chemical Chatter. *Nature* **2019**, *571* (7765), 301–301. https://doi.org/10.1038/d41586-019-02153-y.

(303) Ladinsky, M. S.; Gnanapragasam, P. N. P.; Yang, Z.; West, A. P.; Kay, M. S.; Bjorkman, P. J. Electron Tomography Visualization of HIV-1 Fusion with Target Cells Using Fusion Inhibitors to Trap the Pre-Hairpin Intermediate. *Elife* **2020**, *9*, 1–18. https://doi.org/10.7554/eLife.58411.

(304) Varshey, D. B.; Sander, J. R. G.; Friščić, T.; MacGillivray, L. R. Supramolecular Interactions. In *Supramolecular Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, 2012. https://doi.org/10.1002/9780470661345.smc003.

(305) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157* (1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0.

(306) Gras, S. L.; Tickler, A. K.; Squires, A. M.; Devlin, G. L.; Horton, M. A.; Dobson, C. M.; MacPhee, C. E. Functionalised Amyloid Fibrils for Roles in Cell Adhesion. *Biomaterials* **2008**, *29* (11), 1553–1562. https://doi.org/10.1016/j.biomaterials.2007.11.028.

(307) Balbirnie, M.; Grothe, R.; Eisenberg, D. S. An Amyloid-Forming Peptide from the Yeast Prion Sup35 Reveals a Dehydrated β-Sheet Structure for Amyloid. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (5), 2375–2380. https://doi.org/10.1073/pnas.041617698.

(308) Schütz, D.; Read, C.; Groß, R.; Röcker, A.; Rode, S.; Annamalai, K.; Fändrich, M.; Münch, J. Negatively Charged Peptide Nanofibrils from Immunoglobulin Light Chain Sequester Viral Particles but Lack Cell-Binding and Viral Transduction-Enhancing Properties. *ACS Omega* **2021**, *6* (11), 7731–7738. https://doi.org/10.1021/acsomega.1c00068.

(309) Kollmer, M.; Close, W.; Funk, L.; Rasmussen, J.; Bsoul, A.; Schierhorn, A.; Schmidt, M.; Sigurdson, C. J.; Jucker, M.; Fändrich, M. Cryo-EM Structure and Polymorphism of Aβ Amyloid Fibrils Purified from Alzheimer's Brain Tissue. *Nat. Commun.* **2019**, *10* (1), 1–8. https://doi.org/10.1038/s41467-019-12683-8.

(310) Onogi, S.; Shigemitsu, H.; Yoshii, T.; Tanida, T.; Ikeda, M.; Kubota, R.; Hamachi, I. In Situ Real-Time Imaging of Self-Sorted Supramolecular Nanofibres. *Nat. Chem.* **2016**, *8* (8), 743–752. https://doi.org/10.1038/nchem.2526.

(311) Wittmer, Y.; Jami, K. M.; Stowell, R. K.; Le, T.; Hung, I.; Murray, D. T. Liquid Droplet Aging and Seeded Fibril Formation of the Cytotoxic Granule Associated RNA Binding Protein TIA1 Low Complexity Domain. *J. Am. Chem. Soc.* **2023**. https://doi.org/10.1021/jacs.2c08596.

(312) Jeon, J.; Mills, C. E.; Shell, M. S. Molecular Insights into Diphenylalanine Nanotube Assembly: All-Atom Simulations of Oligomerization. *J. Phys. Chem. B* **2013**, *117* (15), 3935–3943. https://doi.org/10.1021/jp308280d.

6. THESIS PUBLICATIONS

6.1 Materials Promoting Viral Gene Delivery

Authors: Kübra Kaygisiz, Christopher V. Synatschke*
* corresponding author
Published in: Biomaterials Science, **2020**, *8* (22), 6113–6156.
DOI: 10.1039/D0BM01367F
Date of Publication: 29th of September 2020

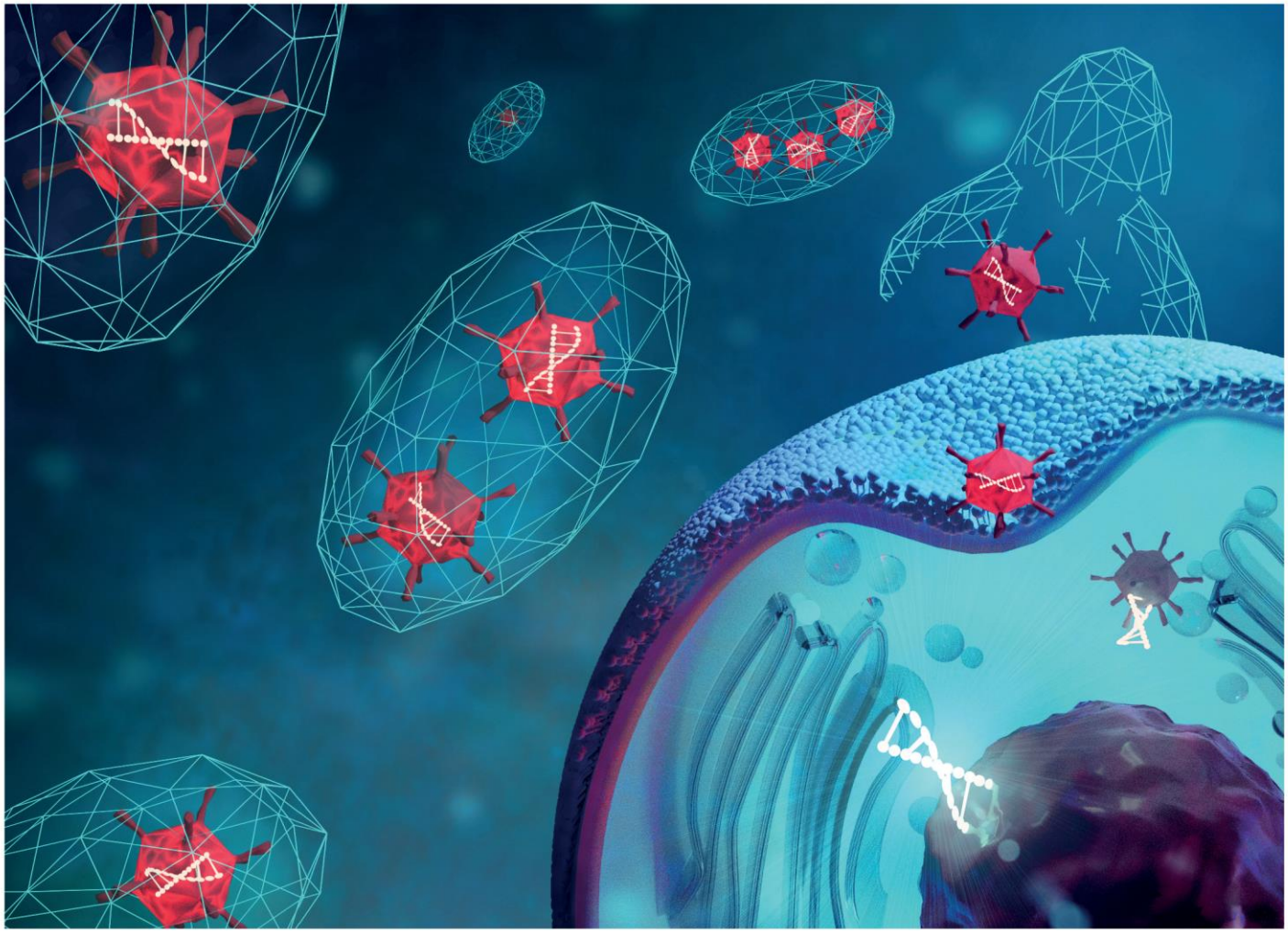
Author Contributions:
Kübra Kaygisiz provided the concept for the article, performed the literature review, and wrote the original draft. Christopher V. Synatschke supervised and revised the manuscript.

Copyrights: The following part is published in Biomater. Sci., **2020**, *8* (22), 6113–6156, https://doi.org/10.1039/D0BM01367F and licensed under CC BY 3.0 (https://creativecommons.org/licenses/by/3.0/). Further distribution of the material is possible if appropriate credit, a link provided to the license, and indications of changes are made.

Graphical abstract:



This review discusses the progress in developing materials that enhance viral transduction, including polymers, peptides, lipids, nanoparticles, and small molecules.

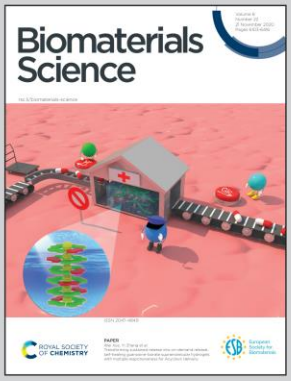


Showcasing a review article from researchers of the Max Planck Institute for Polymer Research, Mainz, Germany.

Materials promoting viral gene delivery

This review summarizes the progress in developing materials that enhance viral transduction, including polymers, peptides, lipids, nanoparticles, and small molecules, pointing out therapeutic applications in research and clinical context.

As featured in:



See Kübra Kaygisiz and Christopher V. Synatschke, *Biomater. Sci.*, 2020, **8**, 6113.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/biomaterials-science

Registered charity number: 207890

Biomaterials Science

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Biomater. Sci.*, 2020, **8**, 6113

Materials promoting viral gene delivery

Kübra Kaygisiz and Christopher V. Synatschke *

Therapeutic viral gene delivery is an emerging technology which aims to correct genetic mutations by introducing new genetic information to cells either to correct a faulty gene or to initiate cell death in oncolytic treatments. In recent years, significant scientific progress has led to several clinical trials resulting in the approval of gene therapies for human treatment. However, successful therapies remain limited due to a number of challenges such as inefficient cell uptake, low transduction efficiency (TE), limited tropism, liver toxicity and immune response. To adress these issues and increase the number of available therapies, additives from a broad range of materials like polymers, peptides, lipids, nanoparticles, and small molecules have been applied so far. The scope of this review is to highlight these selected delivery systems from a materials perspective.

Received 14th August 2020,
Accepted 27th September 2020

DOI: 10.1039/d0bm01367f

rsc.li/biomaterials-science

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

The promise of viral gene delivery

Due to their inherent infectivity, non-pathogenic viral particles (VPs) are highly interesting vectors for gene delivery. In therapeutic applications they replace or deactivate disease-causing genes or introduce new genes to treat a disease.[1] Since the first commercial production in 2003[2] the amount of gene therapies entering clinical trials and eventually earning approval is growing each year[3,4] with promising future prospects.[5] Since 2017, the European Medicine Agency (EMA) and the U.S. Food and Drug Administration (FDA) approved eight new therapeutics,[6,7] amongst them for example Luxturna®, the first adeno-associated virus-based gene therapy[8] and recently Zolgensma® a one-time injection gene therapy to treat spinal muscular atrophy.[9]

Novel tailor-made gene therapeutic strategies pave the way towards personalized medicine, in which the patient's individual genetic circumstances are considered. Viral gene delivery promises to enable various new treatments and current research is directed at areas such as oncolytic virotherapy,[10] tissue regeneration and formation (bones,[11] spinal cord,[12]

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. E-mail: synatschke@mpip-mainz.mpg.de



Kübra Kaygisiz

Kübra Kaygisiz studied Chemistry at the Johannes Gutenberg University of Mainz, where she received her Bachelor's (2017) and Master's (2019) degree. After completing her Master's thesis on photo-patterned peptide gradients, she continued her work in the group of Prof. Tanja Weil at the Max Planck Institute for Polymer Research for her PhD studies. Her research interests focus on peptides as viral transduction enhancers and biofunctional self-assembled nanostructures.



Christopher V. Synatschke

Christopher Synatschke studied chemistry at the University of Bayreuth, Germany, and in 2013 received his PhD with distinction (summa cum laude) working on therapeutic applications of polyelectrolyte nanostructures in the group of Prof. Axel Müller. Then, as a Feodor Lynen Postdoctoral Fellow at Northwestern University, USA, he explored peptide–polymer hybrid biomaterials in the group of Prof. Samuel Stupp. After returning to Germany in 2017, he joined the group of Prof. Tanja Weil at the Max Planck Institute for Polymer Research where he now heads the "Biomaterials" group. His research interests include bioinspired, nanostructured materials to control cellular behavior.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

eye,[13] cardiac muscles[14]), treatment of hemophilia,[15] neurodegenerative diseases[16] and cystic fibrosis.[17]

Unfortunately, therapies in clinical studies have faced serious setbacks through side effects such as serious liver injuries.[9] A lethal case due to high dose adenovirus administration in 1999[18] raised attention to safety issues like immunogenicity, range of infectable host cells (tropism) as well to biological limitations, such as loading capacity of genetic information and batch-to-batch variation of viral vectors (VV).

Consequently, alternatives to VVs have become popular in the last 20 years. Non-viral gene delivery tries to mimic VVs by packaging genetic information in synthetic carriers like cationic polymers or lipids or by using physical methods for direct delivery to host cells as reviewed elsewhere.[19] The advantages of these approaches are a lower immunogenicity due to the lack of preformed antibodies and an easier targeted delivery owing to the modular design of synthetic carriers.[20] However, non-viral gene delivery methods often require more laborious preparation and suffer from low efficacy compared to viral methods.[21] Viruses, optimized through millions of years of evolution, still continue to be the most efficient gene carriers also for therapeutic approaches and are hence used in more than 70% of clinical trials[22] and in the majority of approved gene therapy drugs.[7]

Viral vectors

Viruses are highly ordered nanoassemblies of nucleic acids and proteins and can infect a broad range of organisms from bacteria to humans, both in pathogenic and non-pathogenic contexts.[23,24] Therapeutic VVs are obtained by replacing original gene sequences with beneficial ones.[25,26] These nucleic acids are packaged in a virus capsid, which is composed of proteins displaying functional groups like amines on the surface and can interact with specific receptors on cell membranes. Some virus capsids are additionally enveloped by a lipid bilayer.[27] The virus mediated delivery of nucleic acids (transduction) consists of several steps involving diffusion and translocation through the cell membrane, followed by disruption of the endosome (endosomal escape), release and integration of viral genomes to the nucleus.[26,28]

In clinical trials, adenoviruses (Ad) are the most frequently used VVs for gene transfer. Ad provides high transduction efficiency (TE) to a broad range of dividing and non-dividing cells.[29] Ad tropism is mainly determined by capsid proteins, hexones and fibers,[30,31] which contain fiber knob domains for coxsackievirus and adenovirus receptor (CAR) interaction.[32] The Ad serotype 5 (Ad5) is one of the most common vectors in clinical trials, however, they accumulate in the liver and show strong immunogenicity due to high prevalence of pre-existing immunity to Ad.[33]

Adeno-associated viruses (AAV) have emerged as promising vectors with low immunogenicity and long-term stability, resulting in the recently approved therapies, marketed as Luxturna® and Zolgensma®.[34,35]

As part of the retrovirus (RV) family, lentiviruses (LV) have been increasingly applied in recent years due to their inherent ability to infect both dividing and non-dividing cells, efficient integration into host genetic information and possibility for large scale production.[36–39]

Finally, oncolytic viruses are attracting growing interest as versatile cancer therapeutics.[40–45] They can selectively replicate in cancer cells, eventually inducing cell-lysis and further activating antitumor immune response.[46,47] Promising oncolytic vectors are genetically modified from adenovirus, herpes virus, reovirus and measles virus.[48–51]

The properties of the most common VVs discussed in this review are summarized in Table 1.

Limitations

Despite intense research efforts, VVs still face several limitations. An immune response by the host can lead to rapid inactivation of viral particles and clearance from the blood stream. Subsequent inflammatory processes can lead to liver injury or even multiorgan failure, in severe cases.[60,61] The production of neutralizing antibodies is a lasting humoral immune response, which hinders any further viral gene delivery.[62] To minimize an immune response, low vector doses are used which in turn results in reduced efficiency.

Moreover, the entry to cells can be limited by the natural tropism of VVs. For example, Ad cannot transduce to CAR-negative $CD34^+$ stem cells.[63] Further, transduction can also be hindered by low passive diffusion rates to cell membrane,[64] poor translocation to the cellular endosome, slow endosomal escape and inefficient nuclear genome integration.[65,66]

Strategies to promote viral gene delivery

Over several decades various techniques to promote the delivery of VVs have been developed. These can be categorized into (i) physical methods, (ii) genetic bioengineering of viruses, and (iii) chemical methods, namely material additives. The main aims of these delivery strategies include enhancing transduction efficiency (TE), long-term release of VPs, reduction of

Table 1 Overview of most common VVs for gene delivery studies

Name	Avg. R_h	Zeta potential	Enveloped	Genetic payload	Ref.
Ad	100 nm	−30 mV	No	37 kb	47, 52 and 53
AAV	29 nm	−9 mV	No	4.7 kb	54–56
LV	166 nm	−18 mV	Yes	8 kb	57,59

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

immune response, broadening of tropism and targeted delivery.[6,67]

Physical methods

Physical methods such as microinjection,[68,69] microfluidic,[70,71] sonication,[72,73] centrifugation,[74] cellular deformation,[75] laser irradiation[76,77] or electroporation[78,79] induce or facilitate cell entry of vectors by mechanical force, mainly through disrupting the plasma membrane. They are commonly applied in non-viral gene delivery because they transport the genetic information easily without the need for carriers and the otherwise low infectivity. These methods are also viable for viral gene delivery with electroporation being a noteworthy exception.[80] A downside of using VVs with physical methods is the lack of protection from immune response as well as the invasive and cell-damaging procedure. Very comprehensive overviews of membrane disruptive methods for cargo delivery have recently been published elsewhere.[81,82]

Genetic engineering

Genetic engineering of virus capsids is laborious but gives access to vectors with higher safety profile, modified tropism, and no batch-to-batch variability. Especially in recent years, highly efficient vectors could be obtained using genetic manipulation.

Strategies include amino acids mutations, peptide domain insertions and incorporation of chemical functional groups[83–85] as recently reviewed for AAV.[86–88] For example, introduction of unnatural amino acid residues such as azides on the capsid surface can give access to selective click chemistry.[89,90] Further, genetically modified viruses can enhance TE,[91,92] broaden the tropism,[63,93] enable fluorescent imaging,[85] escape neutralizing antibodies,[94,95] generate stimuli responsive vectors,[96] *e.g.* by light[97,98] or enzymes,[99] and target cells.[100,101] Comprehensive recent reviews on bioengineering AAV can be found elsewhere.[102,103]

Chemical methods

Viral gene delivery can be enhanced by synthetic additives mostly by promoting attractive virus–cell interactions. To this end, VPs are either chemically modified on the outer sphere or non-covalently bound to additives.

Covalent attachment of additives to virus capsid yields stable and irreversible modifications.[104] However, the covalent attachment might interfere with transduction and deactivate viruses by shielding epitopes on the capsid surface necessary for adhesion.[105]

Chemical modification of the virus surface requires mild conditions to maintain the integrity of VPs.[106] Due to the abundance of lysine in the capsid protein,[107] most bioconjugation methods are mild amine functionalizations at physiological conditions, *e.g.* using *N*-hydroxysuccinimide (NHS).[108] Other approaches apply azide–alkyne click chemistry[109,110] and utilize thiol-displaying bioengineered capsids for Michael addition.[111]

Non-covalent attachment is simpler and generally obtained *via* attractive electrostatic interactions or incorporation in various scaffolds. This can result in different architectures such as coatings,[112–114] complexes,[115–120] capsules,[121–125] and matrices[126–128] as illustrated in Fig. 1. Table 2 provides a summary of these architectures and highlights their respective benefits and drawbacks.

Enhanced interactions between cells, VPs and additives can be obtained *via* various pathways.

For cationic delivery agents, binding is mainly driven by attractive electrostatic interactions.[129] This results in directed diffusion and colocalization of viral particles and cell membranes, thereby facilitating cell entry. Some materials can further enhance the fusion of viral and cellular membranes.[130]

Cationic as well as anionic additives with high molecular weight can sediment with VPs onto cells in culture, which increases TE due to increased contact between VPs and cells.[131,132]

VVs can also be immobilized on surfaces prior to cellular adhesion.[133,134]

Chemically facilitated cell entry can also be achieved by manipulating the host cell itself. Especially small cytostatic molecules can affect viral infectivity and TE by intervening with cell cycle processes.[135]

Applying delivery systems able to bypass receptor mediated cell entry can broaden tropism. For example, CAR is downregulated in cancer cells, which makes oncolytic Ad delivery challenging.[136] Delivery systems can improve addressing VPs to CAR negative cells and circumvent CAR-mediated endocytosis.[116,137–140]

A drawback of chemical delivery systems are cytotoxic reactions due to high additive concentrations.[114,141] Furthermore, non-degradable additives, *e.g.* high molecular weight polymers can result in unwanted accumulation and health risks *in vivo*.[142]

Delivery systems

Viral gene delivery systems including polymers, peptides, lipids, nanoparticles, and small molecules display diverse modes of action on a wide range of size scales (Fig. 2). The main focus of this review is to highlight materials for promoting viral gene delivery.

Polymers

Polymers are the most extensively studied delivery systems for VVs. They typically consist of covalently connected repeating monomeric units and reach a molecular mass of several kDa. This section provides an overview on common and emerging polymeric materials of natural and synthetic origin.

Research on polymeric delivery agents began with DEAE-D as one of the first reagents used for VV delivery to mammalian cells in 1965.[143] Due to its advantageous ease of use and high transduction efficiency, it has long been a popular delivery agent for various virus types.[144,145] Nevertheless, one of

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

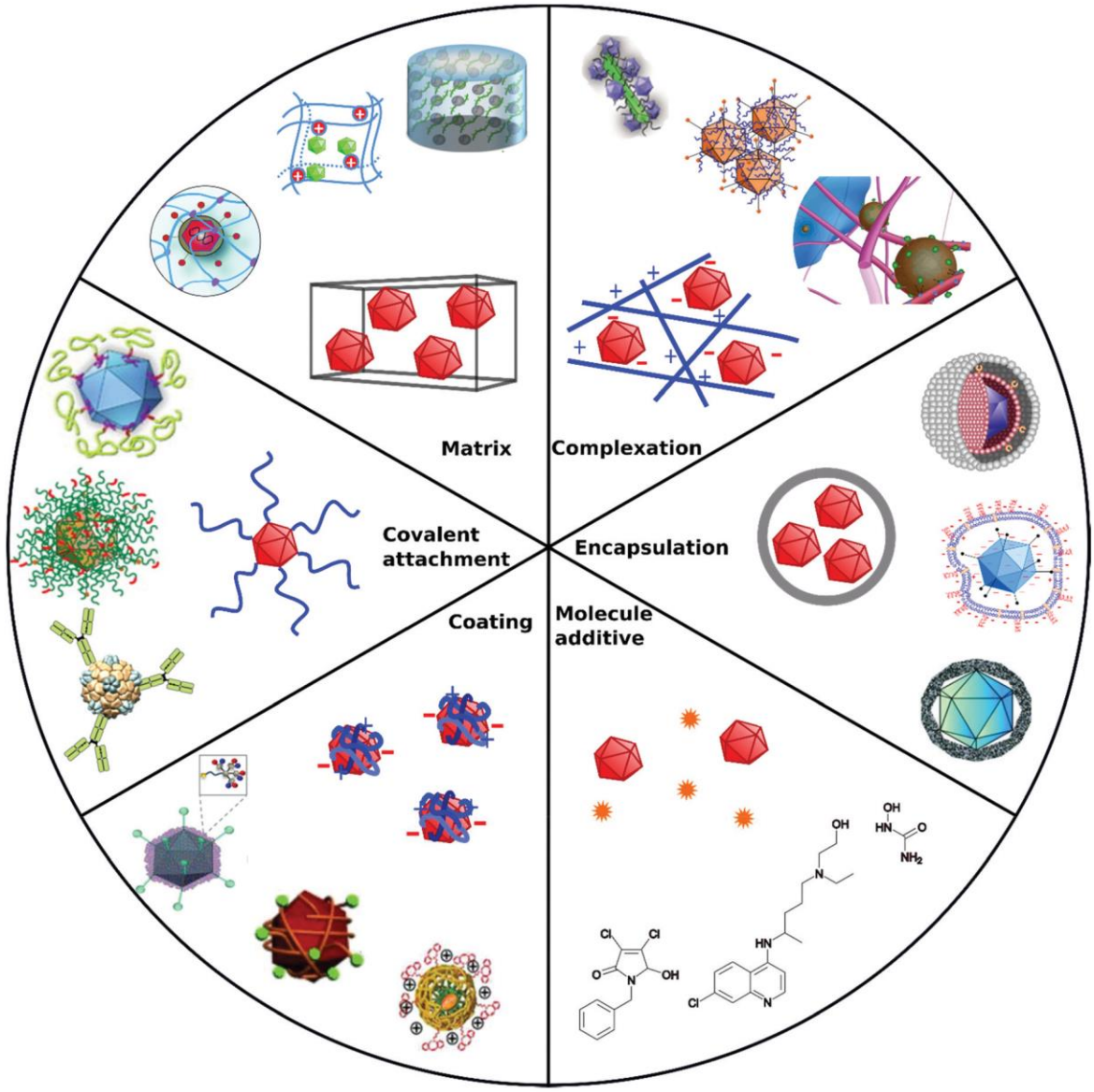


Fig. 1 Overview of different architectural formations for chemical delivery methods of VVs. Figure insets in clockwise order starting from top adapted with permissions from: ref. 274. Copyright 2014 The Royal Society of Chemistry, ref. 307. Copyright 2013 Elsevier, ref. 348. Copyright 2020 Springer Nature, ref. 421. Copyright 2019 American Chemical Society, ref. 422. Copyright 2014 Elsevier, ref. 112. Copyright 2012 John Wiley & Sons, ref. 208. Copyright 2019 John Wiley & Sons, ref. 209. Copyright 2012 John Wiley & Sons, ref. 254. Creative Commons BY 4.0. 2020 John Wiley & Sons, ref. 378. Copyright 2020 American Chemical Society, ref. 311. Copyright 2015 Elsevier, ref. 138. Copyright 2016 Elsevier, ref. 263. Copyright 2019 The Royal Society of Chemistry, ref. 291. Creative Commons BY 4.0. 2020 MDPI, ref. 282. Copyright 2012 Elsevier.

the major issues with DEAE-D is cellular toxicity at high concentrations needed for efficient transduction,[143] which resulted in a declining interest for its use as a delivery agent for VVs.

Beside DEAE-D, the cationic polymer polybrene was heavily used as an enhancer of viral delivery *in vitro* in early days.[146] Due to its cost efficiency as well as its simple and safe handling it remains a popular additive,[147] *e.g.* for enhancing delivery of Ad to human mesenchymal stem cells (hMSCs)[148] and as an additive in ultrasound-enhanced RV delivery to the retina.[72]

In general, it can be summarized that polycationic compounds like DEAE-D, polybrene, poly-L-lysine and protamine sulfate enhance TE, whereas polyanionic compounds like heparin, pyran or cyclodextrin modified with carboxylic groups inhibit TE and reverse the effect of polycationic additives, when combined.[146,149–151] However, anionic dextran sulfate in certain concentrations could promote focus formation of Rous sarcoma virus and thereby slightly enhance TE.[149] In contrast to these findings, dextran sulfate inhibited TE of HIV-1[152] and neutralized enhancing effects of cationic additives.[149] The anionic polysaccharide chondroitin sulfate C also showed

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 2 Overview of chemical delivery approaches for VVs

Approach	Main formation pathway	Main function	Benefits	Limitations	Examples	Ref.
Covalent attachment	Bioconjugation chemistry on capsid proteins	Target cell receptors by specific capsid modifications, shield from immune response	Permanent, stable	Interference with capsid functions	Polymers, peptides,	104–106 and 108–110
Coating	Electrostatic interactions, precipitation	Overcome charge repulsion	Maintain virus capsid functions	Non-permanent, elimination due to cationic charge *in vivo*	Calcium phosphate, silica, flexible polymers	112–114
Complexation	Electrostatic interactions	Overcome charge repulsion	Maintain virus capsid functions	Non-permanent, elimination due to cationic charge *in vivo*	Peptide fibrils, iron nanoparticles	115–120
Encapsulation	Incorporation in capsules	Shield from immune response, overcome charge repulsion	Stealth VPs	Insulation of capsid functional domains	Liposomes, polymers	121–125
Matrix	Coincubation/ coincorporation/ cogelation	Local administration Implantable or injectable matrix	Spatiotemporal release	Degradability, unwanted accumulation	Hydrogels	126–128
Molecule additive	No formation with VPs	Manipulate cellular processes	Synergistic mode of action, easy to apply	Cytotoxicity, no directed VP diffusion to cell	Cytostatic drugs	135

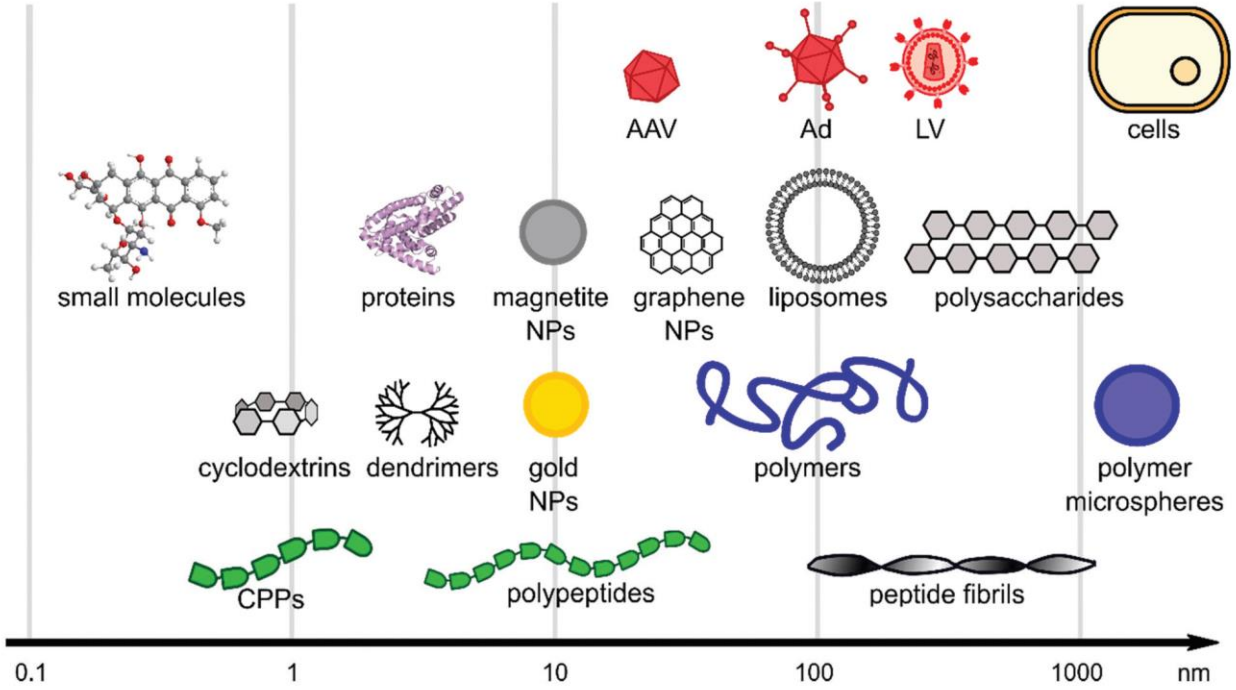


Fig. 2 Size scales of selected materials of interest for viral delivery. The *x*-axis is in logarithmic scale and schematic figures point out size range of material classes.

unexpected increase of TE in combination with polybrene *in vitro*. This observation was explained by sedimentation of larger polymer–virus complexes on cells in contrast to free diffusion of unmodified viruses in the supernatant solution.[131,153]

In order to yield VVs with improved safety profile, *e.g.* protect them from neutralizing antibodies and evade immune response, polymers have been coated[154,155] or covalently bound[156–158] to viruses. The safety of polymer-coated VVs has been increased by changing tropism and retargeting cells.[156] Further, natural and synthetic hydrogels from polymers have been investigated for spatially controlled VV delivery.[159,160] Amongst them macroporous hydrogels were utilized for long-term release of VVs,[161–164] and thermoresponsive polymer gels were applied as injectable systems.[127,163] A selected overview of general structural motifs of the polymers and a summary of all

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

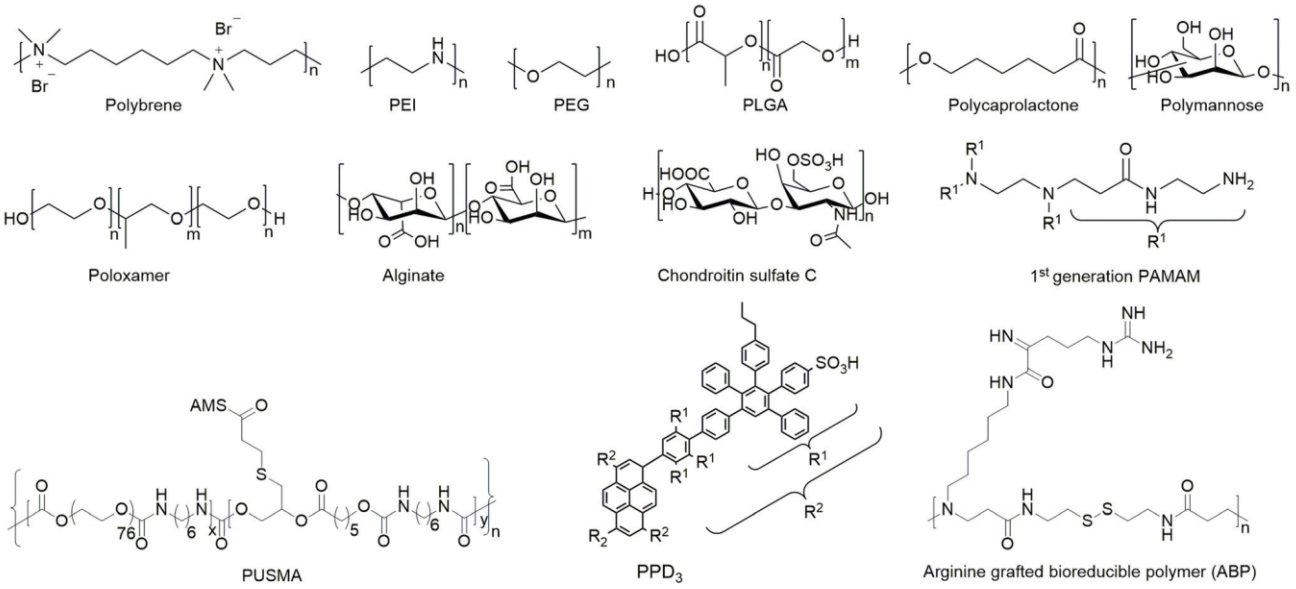


Chart 1 Overview of general structural motifs of selected polymers.

polymers discussed in the following subsections are provided in Chart 1 and Table 3, respectively.

PEG. Polyethylene glycol (PEG) is one of the most versatile polymers used in vast technical and consumer related areas due to its stability and biocompatibility.[165] For more than half a century, the pharmaceutic industry has employed PEG for shielding proteins from humoral immune response to prolong circulation time.[166–169] Consequently, viruses coated with PEG (PEGylated) were first developed to circumvent the immune response *in vivo*. An extensive review of PEG utilized for viral delivery was published in 2010.[170]

Covalent attachment of functional PEG derivatives to Ad capsid proteins, first introduced by O'Riordan in 1999,[106] protected the virus from neutralizing antibodies[106,157] and displayed reduced clearance rates by Kupffer cells.[171] However, PEG-coated stealthy VPs are hindered in cell membrane interaction, which can be tackled by bifunctional multipurpose PEG linkers. Cell specific bioactive peptides attached to these PEG linkers introduced targeted high affinity interaction while showing low immunogenicity even in systemic administration.[172,173] Protection from neutralizing antibodies has been shown for PEGylated AAV at certain cross-linking ratios.[174] Interestingly, TE of VPs varied depending on the grafting efficiency of the activation agents for covalent attachment of PEG and their susceptibility to induce aggregation of VPs.[175] Genetic modifications of AAV resulting in expression of thiol groups on the capsid surface extended bioconjugation methods with thioether or disulfide bonds, *e.g.* for selective conjugation of NHS-functionalized PEG.[83] Moreover orthogonal conjugation of PEG derivatives can also be achieved by click chemistry, when genetically modified AAV capsids presenting azide moieties are used.[176]

PEGylated Ad with low molecular weight of 5 kDa were successfully applied for accumulation in tumor tissue and showed enhanced permeability and retention (EPR) effect when injected intraveneously.[177] Noteworthy, in other studies, PEGylation of Ad with 20 kDa PEG instead of 5 kDa resulted in reduced liver transduction and hepatotoxicity.[178,179] PEGylation also shields Ad from pre-existing Ad antibodies. This is relevant for applications in vaccines or gene therapy due to possible pre-existing immunity against Ad in humans.[180] *In vitro* TE of PEGylated Ad was reduced compared to *in vivo*. This observations was traced back to higher cellular uptake with integrin and proteoglycan interaction, which is more strongly induced by pharmacodynamic force *in vivo*.[181] Beside neutralization of Ad by the serum, blood cells can also inhibit Ad activity after systemic administration.[182] Addressing this need, PEGylated Ad has been used to circumvent endothelial cell activation from blood cells and thereby improve TE and safety.[183]

However, a crucial issue is the so-called PEG-dilemma. Concurrent to EPR effect, PEGylated viruses show lower cellular uptake and lower endosomal escape eventually leading to reduced TE. In order to overcome this problem, PEG was combined with other polymers as reviewed elsewhere for non-viral gene delivery.[184] Copolymers from PEG include spherical poly-DL-lactide particles,[121] pH-sensitive poly(L-histidine-*co*-L-phenylalanine),[185] micellar poly(disulfide amine),[186,187] degradable gelatin[161] to name just a few examples. Fan *et al.* reported an elaborate design of a β-cyclodextrin–PEI-MMP-cleavable-peptide–PEG polymer (CDPCP, Fig. 3). This polymer design aimed to overcome liver accumulation by using PEG 5 kDa, and further to electrostatically coat Ad by grafting PEG onto a high-molecular weight, branched PEI-cyclodextrin copolymer *via* matrix metalloproteinase (MMP) sensitive peptide linkers

 This journal is © The Royal Society of Chemistry 2020

View Article Online

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 3 Overview of polymer sequences, properties, and VP-polymer fabrication methods for the presented polymeric delivery systems

Abbreviation	Compound	Properties	Fabrication of VP-peptide	Ref.
DEAE-D	Diethylaminoethyl-dextran	c., h.	Inc. (poliovirus/avian sarcoma virus/simian virus 40)	143, 149, 144, 149 and 145
Polybrene	1,5-Dimethyl-1,5-diazaundecamethylene polymethobromide	c.	Inc. (sarcoma virus/RV/Ad)	146, 147 and 148
PEG	Polyethylene glycol	h.	Cov. conj. (Ad/AAV)	157, 173, 177, 178, 83, 174 and 176
PELA	Poly-DL-lactide-poly(ethylene glycol)	Amph.	Inc. in microspheres (Ad)	121
PEG*b*PHF	Poly(ethylene glycol)-*b*-poly(L-histidine-*co*-L-phenylalanine)	pH sensitive, h.	Inc., nanoplex formation (Ad)	185
Poly-histidine/PEG		h., pH sensitive	Inc. (AAV)	282
HA–PEG	Hyaluronic acid cross-linked with PEG	Porous scaffold	Inc. (LV)	162
CDPCP	β-Cyclodextrin–PEI-MMP-cleavable-PEG (MMP-cleavable = GPLGIAGQC)	c.	Inc. (Ad)	138
APP	PEGylated and taxol-conjugated polymeric arginine grafted poly(disulfide amine)	c., h.	Inc. (Ad)	187
PEGDA/PLL	Poly(ethylene glycol) diacrylate blended with PLL	h. matrix	Inc. (LV)	283
PEI	25 kDa poly(ethylene imine)	c.	Inc. (Ad)	114
PEI–DEG-bis-NPC	2 kDa PEI cross-linked with diethylene glycol	h., c.	Inc. (Ad)	200
PEI–CyD–FA	600 Da PEI cross-linked with cyclodextrin and folic acid	c., h.	Inc. (Ad)	199
PEI–DA	PEI conjugated with deoxycholic acid	c., h.	Inc. (Ad)	207
rPEI	1.8 kDa PEI cross-linked with cystamine	Branched polymers	Inc. (Ad)	196
PCDP	PEI cross-linked with cystamine derivative	Bioreducible, c.	Inc. (Ad)	203
PEI–pGMA	PEI-*b*-poly(glycidyl methacrylate)	c.	Inc. (LV)	210
PEI–DPA	PEI functionalized with 3-(3,4-dihydroxy-phenyl) propionic acid (catechol groups)	c., h.	Inc. (AAV)	209
PEI–DBCO	1.8 kDa PEI–dibenzocyclooctyl	c., h.	Cov. conj. (LV)	208
PEI/hyaluronic acid		c., a.	Layer by layer deposition (Ad)	201
PEI/chondroitin sulfate		c., a.	Layer by layer deposition (measles virus)	202
PEG-*g*-PEI	2 kDa PEG grafted on 25 kDa PEI	c.	Inc. (Ad)	139
Poloxamer 407	PEO_{101}–PPO_{56}–PEO_{101}	Viscous oil/gel	Inc. (Ad/LV/AAV)/ coinjection (Ad/LV)	230, 232, 233, 238, 229, 231 and 160
Poloxamer 407/ polycarbophil		Viscous oil/gel	Inc. (Ad)	239
Poloxamer PF68 and T908		Viscous oil/gel	Inc. (AAV)	236 and 237
LentiBOOST	Poloxamer 338 (PEO_{141}–PPO_{44}–PEO_{141})	Viscous oil/gel	Inc. (LV)	240–245
PAMAM, EGFR targeting peptide, PEG	Polyamidoamine G3, G4, G5, peptide: CYHWYGYTPQNVI, 2 kDa PEG	c., dendrimer	Inc. (Ad)	140
PAMAM, antibody, PEG	PEGylated polyamidoamine G4, 63 kDa, Erbitux®	c., dendrimer	Inc. (Ad)	251
PPD3	Polyphenylene 3	Amph., dendrimer	Inc. (Ad)	116
PPD3-dendron	One quarter of amphiphilic polyphenylene 3	Amph. dendron	Inc. (Ad)	254
PCL	Poly(ε-caprolactone)	Nanofibers	Electrospinning (Ad)	164
PCL/ELP	PCL blended with elastin like pentapeptide $(VPGVG)_{128}$	Nanofibers	Electrospinning (AAV)	223
PCL–AAV protein tagged	80 kDa PCL tagged with AAV protein binding to AAV	Microspheres or electrospun fibers	PCL–AAV protein binding	224
PLGA	Poly(lactide-*co*-glycolic acid)	p. (LA)/h. (GA)	Enc. (Ad)/lyophilization (LV, Ad)	155, 133 and 215
PLL/PLGA	Poly (lactic-*co*-glycolic) acid and poly-L-lysine		Inc. in PLL and enc. with PLGA (Ad)	213
PLGA/PLL	75/25 DL-PLGA 9.4 kDa/PLL 56 kDa	p. (LA)/h. (GA), cat., h. (PLL)	Inc. (Ad)	213
PLGA/PEG	50/50	p. (LA)/h. (GA), h. (PEG)	Cov. conj. of PEG–SPA and enc. in PLGA (Ad)	214
Alginate	Linear copolymer of [D-mannuronate (β1→4) L-guluronate (α1→4)]$_n$	h. gel	Emulsion (Ad)/cogelation (Ad)/cogelation (AAV, LV)/ cogelation (LV)	258, 259, 294 and 260–262
Alginate/poloxamer 407			Enc. (rAAV)	122
Chitosan	Poly β-(1→4)-linked D-glucosamine	c., linear polymer	Inc. (MLV)	269
Chitosan/β-glycerol phosphate	Crosslinked hydrogel	p., c.	Cogelation (LV)	127
Polymannose	Poly (1→6)-linked α-D-mannose	Linear and branched polymer network	Cov. conj. oxidative *via* $NaIO_4$ (Ad)/cov. conj. reductive amination (Ad)	266 and 265

View Article Online
Review Biomaterials Science

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 3 *(Contd.)*

Abbreviation	Compound	Properties	Fabrication of VP-peptide	Ref.
Cellulose-*g*-P (QDMAEMA)	Cellulose-grafted poly(*N*,*N*-dimethylaminoethyl methacrylate)	c., brush polymer	Inc. (cowpea chlorotic mottle virus and norovirus-like particles)	274
Polyaminoglycoside	Poly hydroxyethyl disulfide diglycidyl ether and tobramycin	Branched c. polymer	Inc. (LV)	276
β-Cyclodextrin		h. and p.	Cov. conj. (TMV)	272
α-Cyclodextrin	α-Cyclodextrin with pluronic PF68 and chondroitin sulfate or hyaluric acid	c., gel mixtures	Inc. (rAAV)	271
EGDE 3,3′	Ethylene glycol diglycidyl ether (EGDE) and 3,3′-diamino-*N*-methyl dipropylamine (3,3′)	Branched c. copolymer	Inc. (Ad)	295
Polydopamine			Inc. (AAV) on upside down polydopamine coated surface	290
Catecholamines	Polynorepinephrine or polydopamine		Seeding onto surface (AAV)	289
pHPMA	Poly(*N*-(2-hydroxypropyl)methacrylamide)	Polyvalent, h.	Inc. (Ad)/cov. conj. (Ad)	280, 156, 158 and 279
PUSMA	PEG cross-linked with 1,6-hexamethylene diisocyanate and epsilon caprolactone sulfamethazine	Biodegradable, thermoresponsive sol–gel polymer	Inc. (Ad)	163
Poly-arginine-*g*-polydisulfide amine		c.	Inc. (Ad)	114
Polyester urethane urea	Copolymer of polycaprolactone diol, butyl diisocyanate, and putrescine blended PEG	Nanofibers	Electrospinning (AAV)	222
Polystyrene	Polystyrene coated with methyl methacrylate and divinylbenzene	Nanocups	Inc. (VV)	293
pNaSS	Poly(ε-caprolactone) grafted poly(sodium styrene sulfonate)	c. film	Inc. and immobilization (Ad)	225
HEMA/APMA	Hydroxyethyl methacrylate (HEMA) with aminopropyl methacrylamide (APMA)	c. hydrogel	Inc. (rAAV)	291
PVA–VEA	Vinyl ether acrylate-functionalized poly(vinyl alcohol)	pH degradable hydrogel network	Inc. (Ad)	292
Poly(2-ethyl-2-oxazoline)		Thermoresponsive, h.	Cov. conj. *via* EDC/NHS (hepatitis B VLP)	278
Polyketal	Cross-linked amino ketal methacrylamide and ketal bis methacrylamide mixed with siRNA	Acid degradable	Polymerized on Eosin-5-isothiocynate conjugated AAV	284 and 285
DNA-aptamer	sgc8-aptamer	Polynucleotide	Cov. conj. (AAV)	288
	5′-ATCTAACTGCTGCGCCGCCGGG AAAATACTGTACGGTTAGA-3′	Polynucleotide	Cov. conj. (MS2 bacteriophage)	286
	Example: TGTGCCAAAGAGAGTGGTGGGGGGGTGG-GCGGAACTCGCG	Polynucleotide	Inc. (vesicular stomatitis virus)	287

Abbreviations: cationic (c.), hydrophilic (h.), hydrophobic (p.), amphiphilic (amph.), anionic (a.), incubation (inc.), encapsulation (enc.) covalent conjugation (cov. conj.), adenovirus (Ad), adeno-associated virus (AAV), lentivirus (LV), retrovirus (RV), tobacco mosaic virus (TMV), murine leukemia viruses (MLV).

(GPLGIAGQC). These peptides are cleaved in the environment of the tumor and uncover positive charge of PEI thus enhancing cellular uptake and Ad gene delivery. Thereby CDPCP avoids liver accumulation, inhibits Ad blood cell interaction and prolongs circulation time.[138]

These advances achieved in stimuli-responsive targeted delivery have the potential to be a valuable tool for systemic administration of viruses, *e.g.* of oncolytic viruses to tumor tissue, in future clinic applications. However, despite extensive use of PEG as a delivery system, in the 2010s, unexpected immune response and accelerated blood clearance of PEGylated therapeutics were reported.[188,189] Consequently there is a need for alternative polymeric delivery systems.[190]

PEI. Polyethylene imine (PEI), commercialized as jetPEI®, is a commonly used transfection agent in non-viral gene delivery and enhances TE by assisting cellular uptake and endosomal escape.[191,192] The secondary and tertiary amines in the backbone of the polymer yield positive charges at physiologic pH and promote attractive electrostatic interactions with anionic particles. After cell intake PEI can deliver its cargo directly into the cytoplasm by accelerated endosomal rupture without entering the degradative lysosomes due to its proton sponge effect induced buffering capacity and osmotic swelling at lower intracellular pH.[193–195] PEI can be applied in linear or branched forms for viral delivery.[196] One major drawback especially of high molecular weight PEI (25 kDa) is its cytotoxicity.[114,141] Attempts have been made to overcome this disadvantage by copolymerization. Grafting PEG onto PEI,[139,197] cross-linking with cyclodextrin,[198,199] introducing bioreducible disulfide groups[196] or diethylene glycol[200] improved toxicity, hemolytic activity, and enhanced transduction of Ad to CAR-negative cells. Lower immunogenicity was achieved through approaches such as layer-by-layer deposition of PEI with hyaluronic acid[201] or chondroitin sulfate[202] and by cross-linking PEI with cystamine.[203]

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

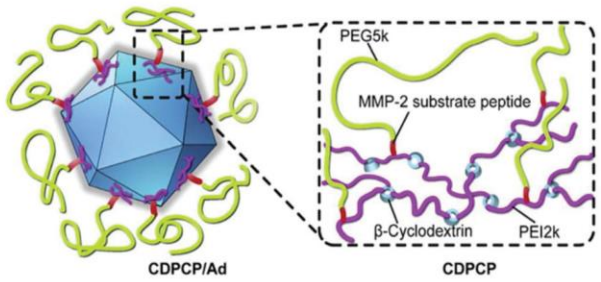


Fig. 3 Schematic diagram of Ad conjugated to enzymatically cleavable PEG–PEI–β-cyclodextrin (CDPCP). Adapted with permission from ref. 138. Copyright 2016 Elsevier.

Moreover cholesterol conjugated to PEI can facilitate cell entry *via* an energy- and endocytosis-independent membrane translocation pathway, originally known from cell penetrating peptides.[204,205] The amphiphilic combination of bile acids like deoxycholic acid and the cationic PEI efficiently transport cargo[206] and further enhance TE of Ad in CAR negative cells.[207]

An elaborate approach of PEI assisted gene delivery was applied by Pan *et al.* for genetic manipulation of primary T-cell to target malignant tissue as new potent immunotherapeutics. In a stepwise process, they first labelled T-cells with azide groups by metabolic incorporation of azide-glucose onto the membrane, then adding PEI-dibenzocyclooctyl-complexed LV. Conjugation of azide-dibenzocyclooctyl *via* click chemistry resulted in colocalization of complexed LV with T-cells, which facilitate interaction and robustly enhance TE in T-cells (Fig. 4A). *In vivo* experiments with tumor-bearing mice displayed significantly longer survival times than the control group or polybrene-assisted delivery.[208]

Delivery systems with spatially resolved gene expression have the potential to mimic the complexity of biological systems. An interesting example is branched PEI, which was functionalized with catechol groups and coated onto AAV. In this way, viral particles were obtained which could be placed and immobilized on tissue culture plates by a micropipette (Fig. 4B). Increased amounts of polymer coatings led to more stable immobilization resisting several washing steps and allowing to pattern viral particles. Thus, spatially resolved gene expression and cell manipulation could be achieved.[209] Recently, spatially defined transduction has been further simplified by coating poly(glycidyl methacrylate) copolymerized PEI as viral affine substrate. LVs were immobilized simply *via* electrostatic interactions and colocalized with HUVECs by incubation on this platform, resulting in enhanced TE and maintained cell functions (Fig. 4C). Further, patterned transduction and GFP expression could be generated by covering the substrate with a mask during fabrication.[210]

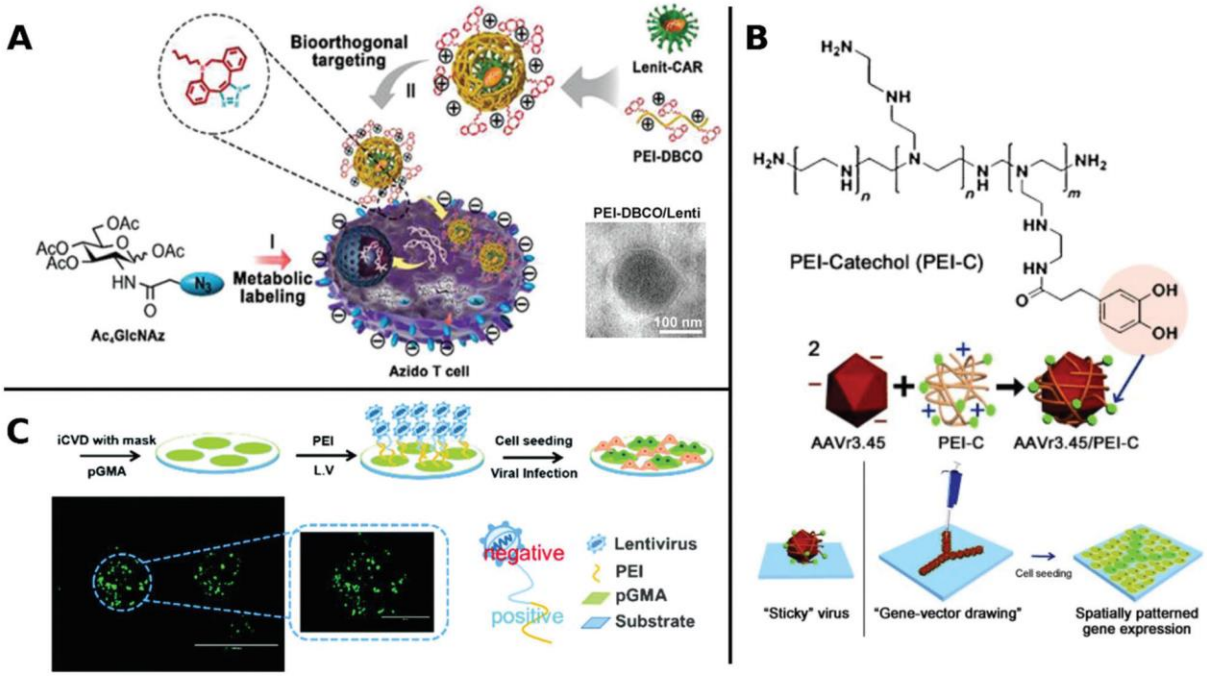


Fig. 4 A. Schematic illustration of PEI-DBCO/azide-glucose system in T-cells. LV are coated with PEI-DBCO, which is conjugated to azide-glucose pretreated T-cells *via* click-chemistry. Inlet TEM-image shows morphology of coated LV (scale bar: 100 nm). Adapted with permission from ref. 208. Copyright 2019 John Wiley & Sons. B. Formation of AAV/PEI-C hybrid particles and schematic illustration of immobilization and patterning of sticky vectors. Adapted with permission from ref. 209. Copyright 2012 John Wiley & Sons. C. Schematic fabrication process of viral delivery platform by subsequently coating of pGMA, PEI and LV attached by electrostatic interactions. Representative image of spatially patterned GFP-expression of transduced HUVECs (scale bar: 200 μm, inlet image: 400 μm). Adapted with permission from ref. 210. Copyright 2019 The Royal Society of Chemistry.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

In recent years, PEI has evolved from a simple polymeric additive to a popular element in copolymeric formulations for enhancing gene delivery, as copolymers can mediate inherent drawbacks of PEI such as high toxicity.

Polyesters. Poly(lactic-*co*-glycolic acid) (PLGA) is an FDA approved copolymer of lactic acid and glycolic acid and commonly used as microspheres for controlled release of drugs and in tissue engineering.[211] PLGA encapsulation of cargo is typically achieved by an emulsion process without the need for charge induced adhesion.[211] However, cationic stabilizers like chitosan and PLL facilitate encapsulation by inducing attractive electrostatic interactions between negatively charged cargo, such as VPs, and PLGA.[212,213]

While Ad encapsulated in PLGA displayed greater cell viability and lower immunogenicity *in vivo*, it did not enhance TE.[155] The TE could be enhanced by PEGylation[214] before the encapsulation step and by surface modification of PLGA spheres with polysaccharides.[12]

In a tissue engineering approach, LV or Ad was immobilized in a PLGA scaffold and ensured long-term transgene expression over a period of four weeks *in vivo*. LV or Ad were incorporated in the PLGA scaffold by consecutive lyophilization of both components together with sucrose. Additional fibronectin and collagen could further enhance binding but not TE for LV delivery.[133] Zhu *et al.* incorporated Ad in PLGA scaffolds for bone regeneration *in vivo*. The electrospun PLGA yielded a nanofibrous scaffold, which enabled spatially defined, long-term gene expression of Ad encoding hBMP2 to promote osteogenic differentiation. This technique was able to cover more than 80% of the bone defects that could not be repaired in the control groups within eight weeks.[215] Another electrospinning approach coated oncolytic vaccinia virus onto PLGA nanofibers to maintain their antitumor activity at the tumor tissue site.[216] Furthermore, melt-processing was reported as a solvent- and additive-free approach for fabrication of PLGA–VP scaffolds. This homogenous dispersed composite materials for prolonged VP release are created by heating up PLGA and VP.[217]

Another example of polyesters is polycaprolactone (PCL) which is widely used as long-term implants in biomedical applications due to their good mechanical properties, slow degradability and biocompatibility.[218] In recent years, electrospinning techniques emerged as a promising process to create PCL fibrous structures for sustained and localized gene delivery.[219,220] In order to incorporate cargo to the fibers, coaxial spinning techniques can be utilized, in which an inner jet with cargo molecules and outer jet with polymer solution overcome the limitations of single stream electrospinning.[221]

In one such approach, Ad was encapsulated in PCL in a core–shell fashion for subsequent porogen-mediated release. Pores of approx. 3 μm size were created by leaching out incorporated low molecular weight PEG particles. The fibrous scaffolds enabled efficient local TE and reduced macrophage activation to attached cells.[164] Similar results were obtained with PCL containing copolymers of polyester urethane urea and polyester ether urethane urea electrospun together with low and high molecular weight PEG and AAV to yield various fibrous nanostructures for sustained delivery to cardiac tissue over a period of 2 months *in vitro* (Fig. 5A).[222]

Blending PCL with elastine-like pentapeptide (ELP) together with AAV resulted in nanofibrous scaffolds with adjustable degradability dependent on the polymer ratio for controlled AAV release.[223] Moreover, cysteine tagged proteins binding selectively to AAV, were attached to maleimide-displaying PCL microspheres or electrospun fibers and gave access to a therapeutic platform for intramuscular injection or subcutaneous implantation with reduced off-target (Fig. 5B).[224]

Recently, poly(ε-caprolactone)-grafted poly(sodium styrene sulfonate) films immobilized rAAV and enhanced gene delivery to hard-to-transduce hMSCs *in vivo*, thus enabling less invasive, effective treatment of focal cartilage lesions.[225]

PLGA and PCL were successfully applied as highly versatile platforms mainly as electrospun fibers and microspheres with controllable spatiotemporal gene delivery, especially considering applications in tissue engineering. Due to its biocompatibility and slow degradation rate both materials, PLGA and PCL has proven to be a reliable scaffold and a promising material especially for local, long-term gene delivery *in vivo.*

Poloxamers. Poloxamers, or Pluronics as a trade name, are commercially available, FDA-approved, thermoresponsive ABA-type triblock copolymers PEO–PPO–PEO consisting of two blocks of hydrophilic poly(ethylene oxide) (PEO) and one block of hydrophobic poly(propylene oxide) (PPO).[226]

Poloxamers are sold in various block lengths and commonly used in pharmaceutical and cosmetic applications as surfactants with no safety concerns.[227] More recently, poloxamers have also been applied as gels for gene delivery, as they exhibit thermal gelation behavior at concentrations higher than 20%.[228] For example, suspensions of VPs cooled at 4 °C can be gelated at physiological temperature around 37 °C, which gives access to homogeneous, injectable, spatially precise viral delivery *in vivo*.[229] However, a potential adverse effect of poloxamer injection concerning the temporal accumulation of microclots in vessels and organs has been reported.[142]

Several studies have shown enhanced TE of localized viral delivery by poloxamer 407 (PEO_{101}–PPO_{56}–PEO_{101}) amongst others to vascular smooth muscle cells *in vitro*,[230] to arteries *in vivo*,[231] to the central nervous system,[232] to solid tumors *in vivo*,[229] and for treatment of spinal cord injury.[160]

Rey-Rico *et al.* applied a series of poloxamers for enhanced AAV delivery to hMSCs even in the presence of host cell receptor binding inhibitory heparin or VP neutralizing antibodies. Prior to gelation, the VPs were encapsulated in poloxamer micelles to shield from adverse conditions, and cell viability was maintained to 100% over a period of 21 days during gene transfer in culture.[233] Long-term expression of AAVs to hMSCs have previously been reported for fibrin or RAD16-1 encapsulated viruses for shorter periods compared to poloxamer encapsulation.[234,235] Linear poloxamer PF68 and four branched poloxamer T908 encapsulated rAAV in a micellar architecture and enhanced spatiotemporal gene delivery efficiency to osteoarthritis chondrocytes *in vitro.*[236,237]

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

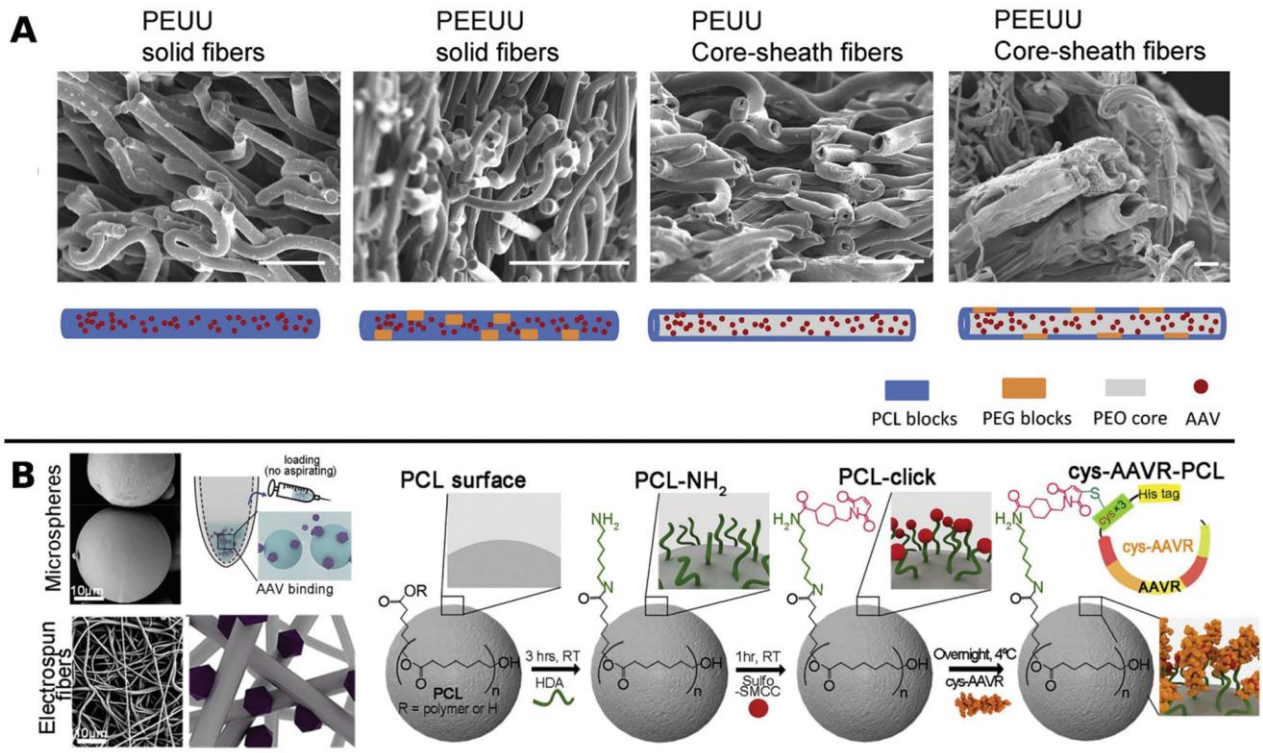


Fig. 5 A. SEM images at the cross sections of freeze dried polyester urethane urea (PEUU) and polyester ether urethane urea (PEEUU) scaffolds and schematic illustration of rAAV encorporation strategy (scale bar: 10 μm). Adapted with permission from ref. 222. Copyright 2017 Elsevier. B. SEM images representing PCL microspheres (top) and electrospun fibers (bottom) (scale-bar: 10 μm) with schematic diagram of AAV–PCL complexation. Reaction scheme shows steps to fabricate AAV tagged PCL. Adapted with permission from ref. 224. Copyright 2019 Elsevier.

Recently, treatment of cartilage damage was applied in a clinically relevant *in vivo* minipig model with rAAV delivered by poloxamer 407 (Fig. 6). Controlled release of the therapeutic SOX9 from the poloxamer gel improved repair of full-thickness chondral defects.[238]

To further enhance adhesiveness of poloxamer 407 for local viral delivery to organ surfaces, it was blended with >1% polycarbophil, a polyacrylic acid cross-linked with divinyl glycol. Addition of polycarbophil changed the sol–gel transition temperature and led to higher adhesiveness due to numerous carboxylic groups, which can easily form bonds with surrounding molecules. Spatially resolved and stable Ad delivery to heart tissue was shown *in vitro* and *in vivo* without having adverse impact to TE or cell viability.[239] Poloxamer 338 (PEO_{141}–PPO_{44}–PEO_{141}), also traded as synperonic F108, enhanced LV delivery in difficult-to-transduce T-cells more efficiently than polybrene. In combination with polybrene the TE was further elevated, which was explained by the distinct modes of each adjuvant, polybrene compensating electrostatic repulsion and poloxamer 338 fluidization the membrane.[240]

More recently, poloxamer 338 was commercialized as LentiBOOST™, and applied for LV delivery to CD34+ stem cells,[241–243] CD4+ and CD8+ stem cells,[244] and T-cells.[245] Enhanced TE, non-toxicity and clinical relevance were emphasized in all of these studies. The evolving progress in poloxamers as adjuvants for viral delivery in clinical application holds great promise for improved gene-based treatments.

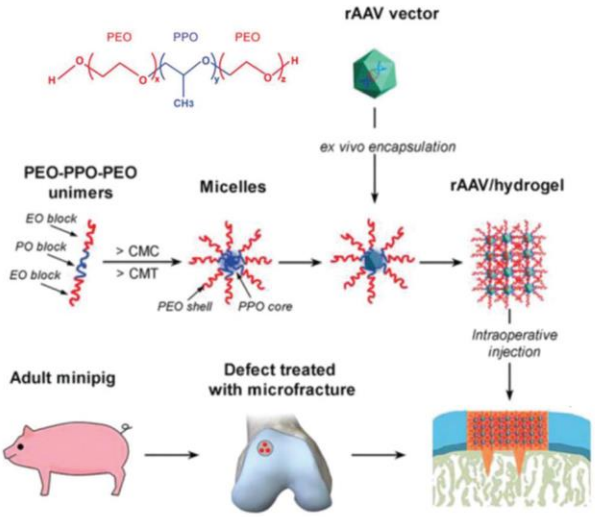


Fig. 6 Structure of poloxamer 407 and flowchart to generate rAAV poloxamer hydrogel systems for controlled release in knee implantations for minipigs. Adapted with permission from ref. 238. Creative Commons BY 4.0. 2019 John Wiley & Sons.

Dendrimers. Dendrimers are branched, radially symmetric molecules with a high density of functional surface moieties,

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

monodisperse size and internal cavities, which makes them an attractive nanomaterial as a vector carrier.[246] A general overview of various dendrimers used in non-viral gene delivery can be found elsewhere.[247]

Early works focused on inhibiting virus infections[248] or were used to control assembly[249] of viruses without any relation to gene delivery.

Polyamidoamine (PAMAM) is a common cationic dendrimer frequently used in non-viral gene delivery[250] and was applied for the first time to enhance TE of Ad by Vetter *et al.*[140] An epidermal growth factor receptor (EGFR) targeting peptide sequence was coupled to the dendrimer *via* a PEG-linker and coated on Ad, which resulted in increased specific infection to EGFR overexpressing tumor cells. The PAMAM complexed Ad showed lower cellular toxicity and higher TE than branched PEI.[140] EGFR targeting antibodies attached to PEGylated PAMAM also showed enhanced TE of Ad and tumor suppression *in vivo*.[251]

Further, polyphenylene dendrimers (PPD) were applied for viral delivery. PPD have amphiphilic properties forming a rigid globular architecture in solution with functional end groups organized like a shell in the periphery.[252] These features make PPD a promising protein-mimicking drug delivery agent.[253] Recently, Wu *et al.* reported PPD3 with engineered amphiphilic surface patches mimicking a protein corona, which enabled non-electrostatic interactions with Ad. The capsid proteins were complexed by PPD due to amphiphilic interactions, which resulted in enhanced TE to CAR negative cells and protection from neutralizing antibodies and the coagulation factor X.[116] Further remodelling of PPD3 led to one amphiphilic dendron branch, which was one quarter of the original PPD3 size (Fig. 7). This minimized dendron branch allowed Ad binding and non-covalent post-modification of viral capsids while maintaining advantageous properties of PPD3.[254] One disadvantage in using dendrimers is their laborious synthesis. Branched polymers carry a large number of functional groups, similar to dendrimers, but are often easier to synthesize. One example is the cationic copolymer EGDE 3,3′ containing the monomer ethylene glycol diglycidyl ether (EGDE) and 3,3′-diamino-*N*-methyl dipropylamine (3,3′), which induced higher TE of Ad to bladder cancer cells compared to the transfection reagent PEI.[255]

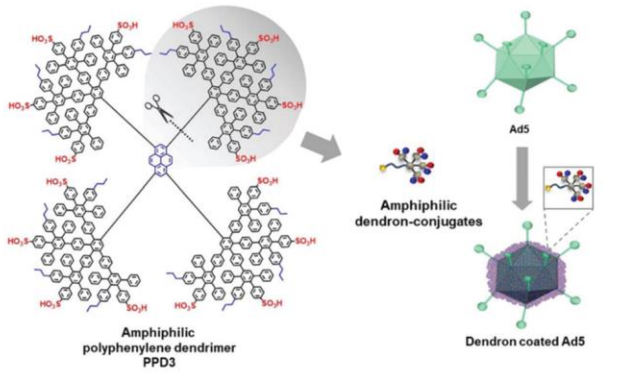


Fig. 7 Structure of amphiphilic polyphenylene dendron derived by desymmetrization of PPD3 dendrimer. Schematic chart shows dendron coated Ad5. Adapted with permission from ref. 254. Creative Commons BY 4.0. 2020 John Wiley & Sons.

Polysaccharides. Biomaterials from polysaccharides provide a versatile and biocompatible tool for clinical applications. Most polysaccharides used for viral delivery are hydrogels building a matrix for spatiotemporal release of VVs. Matrix-mediated viral delivery was recently reviewed elsewhere.[256] The following sections mainly cover saccharides in hydrogel applications.

Alginate is a linear copolymer consisting of β-D-mannuronate and α-L-guluronate and is commonly used as a hydrogel in food additives or pharmaceutical applications. Even though unmodified alginate hydrogels cannot interact specifically with mammalian cells,[257] a lot of progress has been made using such gels in controlled VV release in recent years. Ad encapsulated in alginate circumvent immune response *in vivo*[258] and augment long-term oncolytic Ad infection to tumor cells.[259] Distinct hydrogel-rAAV capsules were formed by tuning the composition and cross-linking temperature for alginate and poloxamer 407 containing systems. All of these rAAV capsules showed enhanced targeted delivery to hMSCs.[122] The spatiotemporal release kinetics of VVs can be stunted or enhanced by hydrogel physiochemical properties.[260,261] For example fabrication methods determine hydrogel mesh size, matrix affinity interactions and degradability and thus result in varying release kinetics for different sizes of VVs. LV delivery from $CaCO_3$ or $CaCl_2$ gelated alginate microcapsules generated by a microfluidic technique correlated with mechanical properties like hydrogel network mesh size and degradation rate and was controllable by the gelation method.[261] Moreover, low molecular weight alginate gels show faster degradation and release of LV, compared to more sustained release from high molecular weight alginate.[260] On-chip fabrication of alginate microgels gave access to blends of adjustable alginate formulations in a bottom up approach (Fig. 8A). Depending on the composition of different alginate molecular weights the degradation and release time of LV could be controlled *in vivo*.[262] Building on these results, Madrigal *et al.* reported that physical properties of various formulated alginates impact AAV and LV delivery to a different extent (Fig. 8B). LV release was tuned by initial strength and degradation rate of alginate gel, whereas AAV delivery remained unchanged independent of formulation. This result was interpreted as release of AAV by diffusive transport, whereas LV release is mainly controlled by degradation rate of alginate, which may be due to mesh/VP size relation. Consequently fast degradable alginate gels lead to higher TE for both AAV and LV.[263]

Mannose receptors were highly expressed in liver tissue and endocytosis-mediated cell entry is enhanced in the presence of mannose.[264] In order to target hepatocellular carcinoma, polymannose was covalently attached to Ad surface either by reductive amination[265] or by oxidation with sodium periodate.[266] The polymannose–Ad conjugate showed enhanced gene delivery to hepatocellular carcinoma cells both *in vitro* and *in vivo*.[266]

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

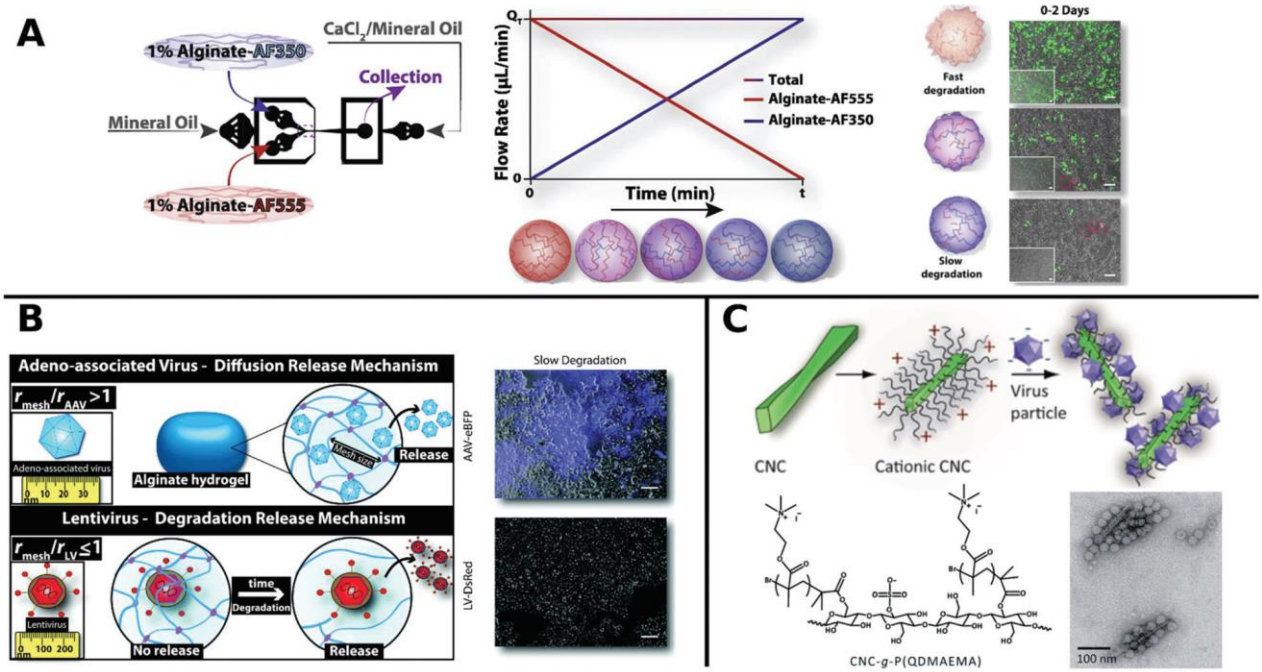


Fig. 8 A. Generation of gradient alginate microgel suspension *via* microfluidic blending. Gradual replacement of degradable alginate-AF555 formulation with non-degradable alginate AF350 leads to microgel composites with controllable degradation and LV release rates as observed in promoted GFP expression in HEK-293T cells by merged phase-contrast/fluorescent photomicrographs (scale bar: 100 µm). Adapted with permission from ref. 262. Copyright 2018 Elsevier. B. Illustration depicting the diffusion controlled release of AAV and degradation limited release of LV from alginate hydrogels. Merged phase-contrast/fluorescent images show gen-expression achieved after first day for AAV and LV in slow degradable alginate hydrogels (scale bar: 100 µm). Adapted with permission from ref. 263. Copyright 2019 The Royal Society of Chemistry. C. Schematic diagram shows the formation of VPs complexed with colloidal cellulose nanocrystals (CNC). TEM micrographs show CCMV complexed with CNC-*g*-P(QDMAEMA) (scale bar: 100 nm). Adapted with permission from ref. 274. Copyright 2014 The Royal Society of Chemistry.

Chitosan is a non-toxic, biodegradable cationic polymer, that shows neuroprotective effects after spinal cord injury by sealing nerve cell membranes[267] and promotes peripheral nerve regeneration.[268] Cross-linked with β-glycerol phosphate, chitosan yields compact fibrous hydrogels with increased charge density and binding to anionic particles as well as long-term gene expression over a period of seven days of encapsulated LV to dorsal root ganglia neurons.[127] Furthermore, chitosan has been used to replace the function of the viral envelope in non-infectious murine leukemia virus (MLV), which led to increased infectivity and transduction.[269] Hyaluric acid (HA) is a highly abundant linear polysaccharide in the extracellular matrix. HA applied as an injectable *in situ* forming scaffold can generate macroporous structures which are attractive vehicles for localized long-term release of viral particles. In a recent report, HA scaffolds of various pore sizes for LV delivery in mice mammary fat were compared with each other. Void spaces in HA scaffolds were created by different fabrication techniques. Nanoporous structures were achieved by cross-linking HA particles with PEG precursors, while macroporous architectures were created through *in situ* assembly of HA particles with PEG particles or enzymatic degradation of included PEG particles. Open, macroporous HA–PEG hydrogels displayed increased host cell infiltration and yielded higher TE compared to nanoporous hydrogels.[162]

Cyclodextrins (CD) are composed of 6–8 (α–γ) cyclic arranged glucose subunits forming a hydrophobic interior and a hydrophilic exterior toroid-shaped oligosaccharide, which is commonly used in pharmaceutical applications for drug delivery.[270] CDs have been applied as hydrogels in combination with other polymers[138,271] or as supramolecular linkers[272] for gene delivery as recently reviewed elsewhere.[273] Examples include Ad delivery to tumor microenvironment with a responsive polymer design containing PEG, PEI, MMP-sensitive peptide and β-CD.[138] Further, polypseudorotaxane gels based on either HA or chondroitin sulfate combined with α-CD were used to encapsulate and release rAAV to hMSCs. α-CD enhanced the viscoelasticity and storage modulus of the gels at physiological temperature and prolonged the permanence at the application site.[271] CD were also covalently attached to TMV surface in a supramolecular strategy to enable facile host–guest interactions with adamantyl moieties of imaging agents or chemotherapeutic drugs.[272]

In order to overcome anionic surface charges of some polysaccharides, modifications and combination approaches with other polymers have been made. For example, cellulose nanocrystals were surface-modified by atom-transfer radical polymerization with poly(*N*,*N*-dimethylaminoethyl methacrylate) to yield a brush-like cationic polymer (Fig. 8C). This polymer showed high-affinity virus binding.[274]

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Bioreducible, branched polyaminoglycosides involving the antibiotic aminoglycoside tobramycin were proposed by Xu *et al.* as a transfection reagent.[275] Recently, they reported, LV–polyaminoglycoside complexes, which efficiently induce cell apoptosis to glioma cells by facilitating cellular uptake *via* endocytosis pathways.[276]

Polysaccharides are highly suitable biopolymers for viral gene delivery, mainly as hydrogels that adapt to the tissue environment. Release of VPs is achieved by degradation into non-toxic components. Combination of polysaccharides with polymers and functional molecules can further expand the range of material properties and influence controllable and efficient delivery of VPs in future applications.

Miscellaneous. In recent years, several additional polymers have been explored as VV delivery agents. In this section, miscellaneous emerging polymeric delivery systems are briefly introduced.

Several examples of hydrophilic polymers with similar stealth properties to PEG have been reported. For example, poly(2-oxazoline)s are thermoresponsive, hydrophilic polymers which gained increasing attraction for biomedical applications.[190,277] Recently, hepatitis B-like viral particles grafted with poly(2-ethyl-2-oxazoline)s were reported to reduce antigenic reaction.[278]

Similarly, poly(*N*-(2-hydroxypropyl)methacrylamide) (pHPMA) was conjugated through lysine residues on Ad surface and was able to shield from neutralizing antibodies, enhanced plasma circulation, decreased hepatotoxicity and indirect accumulation in tumor site. However, coating with pHPMA prevented efficient cellular uptake and reduced TE.[156,158,279] Francini *et al.* applied a new conjugation strategy with pHPMA bearing diazonium reactive groups, which can be bioconjugated onto a variety of functional amino acid residues of oncolytic Ad resulting in more efficient coupling and dense coverage of surface. While immunogenicity of conjugated VPs is low, TE is severely reduced due to low cellular uptake and delayed unpackaging of the vector. This study illustrates the necessary balance between efficient screening of the virus from being recognized by the immune system and sufficient cellular uptake in order to achieve effective transduction.[280] To take advantage of benefits from different systems, hybrids consisting of polymers and peptides have been investigated for enhancing TE. For example, arginine grafted onto poly(disulfide amine) were able to enhance TE, while being bio reducible and less immunogenic, when coated onto Ad.[114,281] AAV in hydrogels from PEG incorporated with poly-L-histidine showed ratio-controlled and pH-dependant swelling and TE (Fig. 9A).[282] PEG diacrylate matrices (PEGDA) blended with PLL can improve long-term, localized and efficient LV transduction when implanted *in vivo*.[283] Further, Kwon *et al.* introduced elaborate viral/non-viral chimeric systems by siRNA or DNA encapsulating polyketals assembling in a core–shell structure around AAV. This design enabled simultaneous gene transduction in a stimuli-responsive fashion by using of polyketals that are degraded in the acidic endosomal environment. Chimeric systems are promising platforms for obtaining syner-

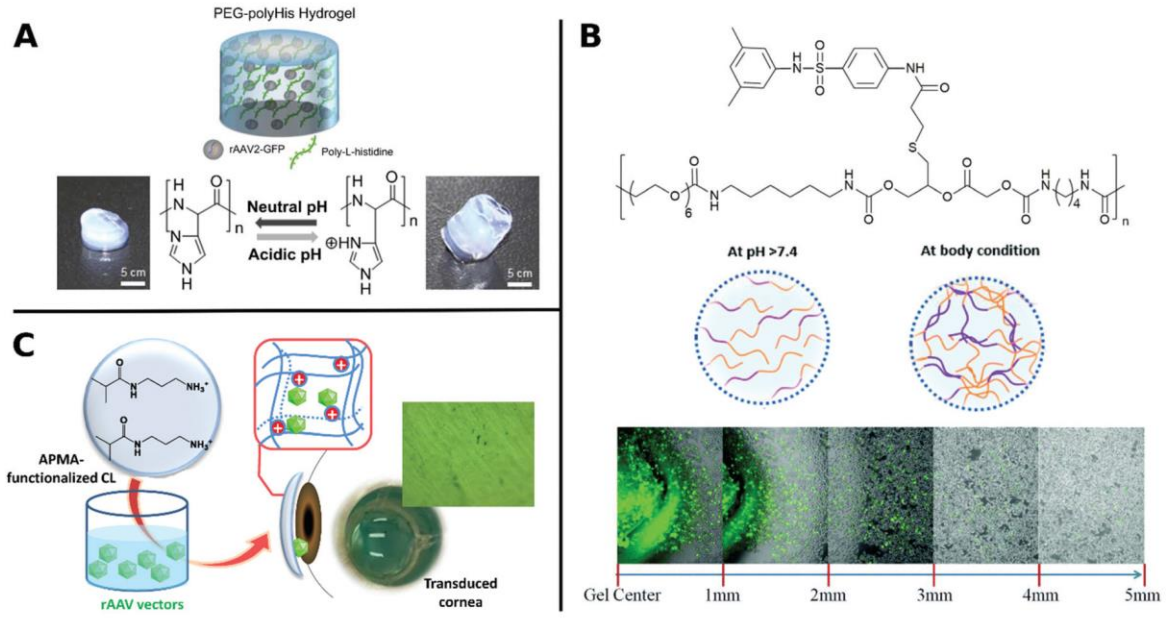


Fig. 9 A. Schematic illustration of rAAV embedded into PEG-polyHis hydrogel incubated under neutral (pH 7.4) or acidic (pH 6.0) conditions. Protonation of amine groups in polyHis under acidic conditions result in increased water uptake and swelling of the hydrogel. Adapted with permission from ref. 282. Copyright 2012 Elsevier. B. Structure of PUSMA and schematic illustration of sol–gel phase transition state at physiological conditions. Merged optical/fluorescence image of distance-dependent release of GFP-expressing Ad from PUSMA hydrogels. Adapted with permission from ref. 163. Copyright 2019 The Royal Society of Chemistry. C. Illustration of functionalization of APMA-hydrogel contact lenses for sustained VV delivery to cornea and picture of X-Gal stained bovine cornea after seven days in direct contact with rAAV encapsulated hydrogels. Adapted with permission from ref. 291. Creative Commons BY 4.0. 2020 MDPI.

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

gistic therapeutic effects by simultaneous expression and silencing of multiple genes. In cancer therapy, for example, simultaneous upregulation of pro-apoptotic mediators by AAV delivery and silencing of pro-survival genes by siRNA, can result in significantly more effective treatment.[284,285] Beside classical polymer systems, DNA-aptamers have also been utilized for targeted gene delivery. DNA-aptamers were covalently attached to viral capsid and selectively targeted Jurkat T cells and delivered cargo through an endocytic pathway.[286] DNA aptamers can further improve biocompatibility of viruses by shielding them from neutralizing antibodies and enhancing *in vivo* circulation rate.[287] Reducible disulfide linkages were utilized to covalently attach AAV to multiple DNA-aptamers, which were cleaved by intracellular glutathione and facilitate release of AAV in the cell, thereby enhancing TE.[288]

Polymers can also be used as 2D coatings to promote transduction by colocalization of viruses with host cells. So-called substrate mediated delivery was enabled *via* AAV capturing adhesive catecholamine surfaces (Fig. 10A).[289] Adhesive polydopamine-coated substrates were further improved in an simple upside down arrangement to adhere inverted quasi spherical droplets containing human neural stem cells (hNSCs) and AAVs (Fig. 10B). TE was enhanced by shortened path lengths and increased contact frequencies between cells and vectors due to the large contact angle of droplets on polydopamine coated surfaces.[290]

Several hydrophilic polymers have been applied as 3D materials, namely hydrogels, often for slow release of VPs. The biodegradable multiblock sulfamethazine and PEG-containing polyurethane (PUSMA) exhibit pH and thermoresponsive behavior and was used as an injectable hydrogel for *in vivo* oncolytic Ad delivery. The sol–gel transition occurred below pH 8.0 through a hydrophilic to hydrophobic transition leading to the formation of a porous hydrogel network. This hydrogel enabled spatiotemporal Ad delivery to injection site under physiological conditions (Fig. 9B).[163] An interesting example for the rAAV delivery to the cornea used hydrogels from a copolymer of hydroxyethyl methacrylate (HEMA) and aminopropyl methacrylamide (APMA) which can be worn as contact lenses (Fig. 9C). This transparent hydrogel network allowed high vector loading and controlled, long-term gene expression over a period of 14 days while being able to correct refractive errors.[291]

Microgels from vinyl ether acrylate-functionalized poly(vinyl alcohol) (PVA–VEA) and thiolated PVA–VEA were fabricated by a microfluidic technology and Michael-type cross-linking reaction to yield pH-degradable injectable spheres for efficient Ad delivery to tumor sites. The antitumor treatment could be reinforced by the addition of the bromodomain inhibitor JQ1. Accumulation of oncolytic Ad at the tumor site was enabled by pH-dependent controlled release from microgels. This combination approach paves the way for treatment of tumors by synergistic viral, chemo and immunotherapy as a promising system in clinical applications.[292]

Polystyrene coated with methyl methacrylate and divinylbenzene were formed into nanocups <500 nm by a template and applied as cavitation mediated carrier for oncolytic VV. Physical stimuli, *e.g.* ultrasound after intravenous injection resulted in enhanced transport and antitumor activity to treatment site.[293]

Several new polymeric systems and composites are emerging. Careful choice of the polymer and conjugation strategy allows the user to tailor pharmaceutical properties such as circulation time, controlled release of viral cargo and targeting of cells or tissues for next generation gene therapy.

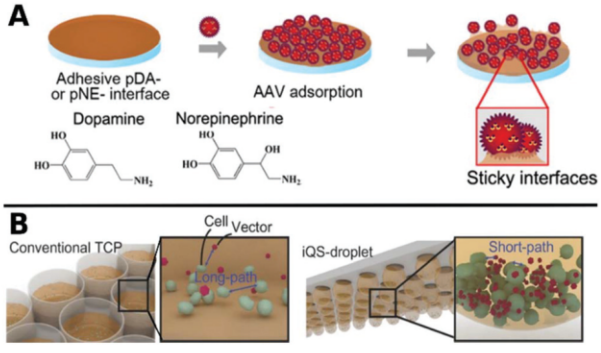


Fig. 10 A. Schematic illustrations of adsorption of AAV on catecholamine coated substrates. Adapted with permission from ref. 289. Copyright 2014 American Chemical Society. B. Schematic illustration demonstrating the promotion of cell–AAV interactions in an inverted quasi-spherical (iQS) droplet on polydopamine coated substrates due to shorter pathways compared to conventional tissue culture plates (TCPs). Adapted with permission from ref. 290. Copyright 2017 John Wiley & Sons.

Peptides

Due to their natural abundance and their ability for interactions with cells and viruses, bioactive peptides are highly interesting auxiliary agents for VV delivery.[296] Tailor-made peptides can be easily produced in large scale, are biocompatible and biodegradable.

Peptides applied for non-viral gene delivery were summarized recently elsewhere.[297] When used as enhancers for viral gene delivery three different classes of peptides have been applied so far: cell penetrating peptides (CPP), fibrils formed from self-assembling peptides and proteins.

The formation of peptide–virion complexes is typically achieved either by attractive electrostatic interactions between negatively charged viral particles and positively charged peptides or by bioconjugation of the respective peptide to the capsid. With the large structural variety offered by peptide sequences, it is not surprising that many different types of peptides have been reported as promoting viral gene delivery. Table 4 provides an overview of the reported peptides, their sequences, and physicochemical properties. Furthermore, the types of VVs that have been reported in combination with these peptides are highlighted.

CPPs. Cell penetrating peptides (CPPs), also called protein transduction domains (PTD), typically contain between 5 and 30 amino acids, have an overall positive charge and can facili-

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 4 Overview of peptide sequences, properties and VP-peptide fabrication methods for the presented CPPs, fibrils and proteins

Name	Structure	Properties	Fabrication of VP-peptide	Ref.
$Tat_{47–57}$	YGRKKRRQRRR	h., c.	Inc. (Ad)/cov. conj. (Ad)/bioengineering (Ad)	311,315,318
$Tat_{48–57}$	GRKKRRQRRR	h., c.	Cov. conj. (Ad)	309
$Tat_{48–60}$	GRKKRRQRRRPPQ	h., c.	Cov. conj. (Ad)/bioengineering (Ad)	310, 316, 320 and 137
Tat HA2	CRRRQRRKKRGGDIMGEWGNEIFGAIAGFLG	Amph., c.	Bioengineering (AAV)/inc. (AAV)	322,328
OM-pBAEs	CRRR-PEG-CRRR	h., c.	Inc. (Ad)	312
R5	RRRRR	h., c.	Cov. conj. (Ad)	319
R8	RRRRRRRR	h., c.	Cov. conj. (Ad5/Ad/TYMV)/inc. (Pol., Ad)	311,316–318,320,329
R9	RRRRRRRRR	h., c.	Cov. conj. (Ad)/—	309 and 327
HP4	RRRRPRRRTTRRRR	h., c.	Inc. (Ad)	119
K7	KKKKKKK	h., c.	Bioengineering (Ad)/—	137, 365, 366 and 326
Pep1	KETWWETWWTEWSQPKKKRKV	Amph., c.	Cov. conj. (Ad)/inc. (Ad)	311,318,320
$Pen/Antp_{62–77}$	RQIKIWFQNRRMKWKKGG	Amph., c.	Cov. conj. (Ad)/inc. (Ad)	308,309,311,318,320,328
Pro	VRLPPPVRLPPPVRLPPP	Amph., c.	Cov. conj. (Ad5)	316 and 317
n.n	NRPDSAQFWLHHGGGSLLGRMKGA	c.	Cov. conj. (hepatitis B)	314
12.51	TARGEHKEEELI	c.	Bioengineering (Ad)	313
THR	THRPPMWSPVWP	Amph., c.	Inc. (AAV)	321
KH27K	KHHHHHHHHHHHHHHHHHHHHHHHHHHHHK	h., c.	Inc. (polyomavirus)	329
FUSO	CGLFEALLELLESLWELLLEA	a.	Inc. (polyomavirus)	329
LAH4	KKALLALALHHLAHLALHLALALKKA	Amph., c.	Inc. (AAV2 and AAV8, Pol)/inc. (LV)	328,329,351
Vectofusin-1	KKALLHAALAHLLALAHHLLALLKKA	Amph., α-helix	Inc. (LV)	130 and 354
$PAP_{248–286}$	GIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMY	Amyloid, c.	Inc. (HIV-1)	336
$PAP_{85–120}$	IRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL	Amyloid, c.	Inc. (HIV-1)	337
$SEM1_{45–107}$	GQHYSGQKGKQQTESKGSFSIQYTYHVDANDHDQSRKSQ-QYDLNALHKTTKSQRHLGGSQQLL	Amyloid, c.	Inc. (HIV-1)	338
EF-C	QCKIKQIINMWQ	Amph., c. amyloid	Inc. (RV, LV)	118
EP2	QCKIKQIINMWQEVG	Amyloid, c.	Inc. (HIV-1)	343
EP3	NITLQCKIKQIINMWQEVG	Amyloid, c.	Inc. (HIV-1)	344
P13	Ac-NWFDITNWLWYIK-NH_2	Amyloid, c.	Inc. (HIV-1)	339
P16	Ac-NWFDITNWLWYIKKKK-NH_2	Amyloid, c.	Inc. (HIV-1)	339
P16-D	Ac-NWFAITNWLWYIKKKK-NH_2	Amyloid, c.	Inc. (HIV-1)	341
CD4bs-M	XGXSGGDPEIVTXKXXLTRDGGN (X = ε-aminohexanoic acid)	Amyloid, c.	Inc. (HIV-1)	346
Fmoc-SAP	Fmoc-DDIKVAVK	Gel, c.	Coinjection (LV)	348
DPF1	INMWQG	Fibrils	Inc. (HIV-1)	342
RAD16-1	Ac-RADARADARADARADA-CONH2	Fibrils	Cogelation (AAV)	234
KK, KY	KYKGAIIGNIK, KYRSGAITIGY	Amyloid		347
α-Syn	140 amino acids, consecutive KTKEGV	Amyloid	Inc. (RV)	117
Protamine sulfate	Mixture of polypeptides	c.	Inc. (RV, Ad, LV)	243, 356 and 357
Retronectin	574 amino acids, 63 kDa	c.	Inc. on precoated substrates (LV, RV)	359 and 362–364
PEG–PLL	$PEG_{12000}–PLL_{48}$/20 kDa PEG dendrimer–PLL	c.	Inc. (RV)/inc. (LV)	368 and 369
PLL	Poly-L-lysine 150–300 kDa	c.	Inc. (VLP)	367
Fibrin	Fibrinogen and thrombin mixture	Fibrous gel	Cogelation (LV, Ad/AAV)	372, 374, 235 and 373
Collagen	Collagen type 1	Fibrous gel	Cogelation (LV)	126
	Collagen/hydroxyapatite ($Ca_5(PO_4)_3(OH)$)	Gel	Cogelation (LV)	126
Gelatin	Hydrolyzed collagen	Gel	Cov. conj: (Ad)/cogelation (Ad)	385 and 386
Serum proteins	HSA, LDL, Transferrin	c.	Inc. (AAV8/HIV-1)	375, 377 and 376
SELP	GAGAGS and GVGVP blocks	Gel	Inc. (Ad/GLV-1h68)	382, 383 and 384

Abbreviations: cationic (c.), hydrophilic (h.), amphiphilic (amph.), anionic (a.), incubation (inc.), covalent conjugation (cov. conj.), adenovirus (Ad), adeno-associated virus (AAV), lentivirus (LV), retrovirus (RV), tobacco mosaic virus (TMV), polyomavirus (Pol).

This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

tate efficient entry into cells. CPPs can exhibit various secondary structures from random coil to α-helix,[298] and typically have a large amount of basic amino acids like arginine or lysine and have an amphiphilic or hydrophilic character. CPP motifs are widespread in Nature, *e.g.* in heparin-, RNA- and DNA-binding proteins, signalling peptides or viral proteins and allow receptor independent cell-entry.[299] Since the discovery of the first CPP derived from HIV-1 binding protein Tat in 1988,[300,301] more than 1000 CPPs predominantly from natural origin or found by phage display are known today.[302]

Their ability to transport variable cargo across cellular membranes has made CPPs a facile tool for the delivery of DNA and RNA, liposomes, nanoparticles, proteins, and drugs as reviewed elsewhere[303,304] and of viral nanoparticles for efficient transduction as comprehensively reviewed.[305,306] In this segment, we highlight the most important developments in CPPs for facilitating transport of VP into target cells.

The most straightforward and easy way to obtain CPP–virus complexes is by co-incubation, making use of electrostatic interactions between positively charged CPPs and negatively charged VP. This method can be applied independently of the virus type, and ratio of CPP to VP can be easily adjusted in solution. However, the formation of CPP–VP complexes cannot be controlled, there is batch-to-batch variability and the highly positively charged CPP–VP complexes are prone to aggregation in physiologic conditions, for example, in the presence of electrolytes or serum (Fig. 11).[307] A more stable but demanding approach is to covalently connect CPPs to the vector capsid. Early studies have demonstrated both approaches: the utilization of electrostatically bound Pen and Ad complexes to facilitate viral gene transfer in muscle cells[308] and covalently bound Tat–Ad conjugates for delivery to tumor cell lines.[309] By covalently attaching Tat to exposed lysine residues of the vector capsid *via* an MHS linker, transduction efficiency was further improved.[310] Similar observations were made for Tat, Pen, R8 and Pep1 when they were covalently attached to PEGylated adenoviruses (Fig. 11).[311]

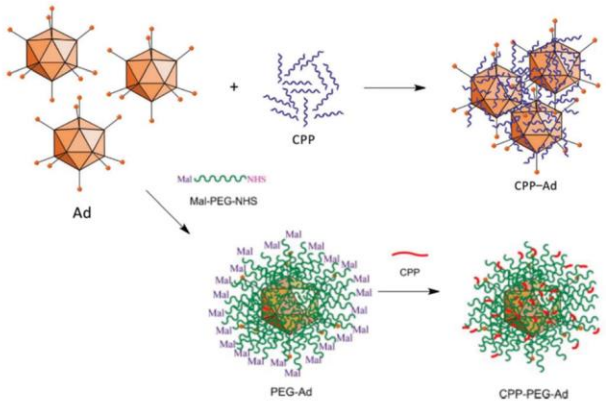


Fig. 11 Approaches for producing CPP–Ad complexes (top) and conjugates (bottom). Complex formation results from electrostatic interactions between positively charged CPPs and negatively charged VPs. Covalent conjugations of CPPs are conducted on viral capsids, *e.g. via* bifunctional PEG linkers. Adapted with permission from ref. 307 and 311. Copyright 2013 and 2015 Elsevier.

Coating of CRRR-*co*-PEG-*co*-CRRR onto an oncolytic Ad (Fig. 12A) enhanced transduction in tumor sites with a longer blood circulation time and lower liver sequestration was achieved.[312]

In addition to CPPs derived from natural peptides, screening for suitable sequences by phage display has become a powerful method to discover new CPPs for enhanced transduction and targeted virus delivery.[313,314]

It was further possible to broaden the tropism of Ad by gene transfer to otherwise non-transducable CAR-negative cells. To this end, the surface knobs were modulated with Tat peptides,[137,315] Tat peptides were attached to surface bound lysine residues[316,317] or simply incubated in solution.[318] Other examples for hard-to-transduce cell types that are successfully targeted with CPPs include Ad delivery to resistant stem cells and various cancer cells which was efficiently achieved by addition of arginine-rich HP4 derived from herring protamine.[119] Furthermore, by decorating CPPs onto capsids *via* hydrazone chemistry, plant viruses like CPMV and TYMV were able to transduce otherwise non-infectable mammalian cells.[319,320]

Specific CPPs capable of crossing the blood brain barrier (BBB), can facilitate the delivery of CPP virus-complexes to the central nervous system after systemic application (Fig. 12B).[321] In an interesting approach AAV containing a brain derived neurotrophic factor fused with Tat were delivered intranasally to the central nervous system to act as an antidepressant in mice.[322]

Frequently used CPPs for VV delivery are various Tat peptide fragments, oligoarginines, and penetratin (Pen). These peptides show different transduction enhancing properties depending on the viral particle, host cell, and CPP concentration.[318] This difference in enhancement is believed to result from structural properties of the CPPs as well as their respective cellular entry mechanism. In general, various endocytic and non-endocytic cell entry pathways are controversially discussed in the literature.[323,324]

Regarding the structure, one requirement for the electrostatic stabilization of the virus peptide complex, as well as adhering to and crossing the negatively charged lipid bilayer of cells, is a high amount of positive charges in the peptide.[318] Among peptides with positive charge those rich in arginine showed greater membrane permeability than CPPs with high amount of other cationic amino acids like lysine, histidine or ornithine.[325,326] Beside the charge, hydrophilicity can also influence interaction with the membrane. It is believed that the higher performance of hydrophilic peptides like Tat and oligoarginines is due to stronger interactions with heparan sulfate proteoglycans of the cell membrane, whereas amphipathic peptides like Pep1 interfere with electrostatic binding of proteoglycans, thus resulting in lower transduction efficiency.[318] Interestingly, a higher CPP concentration for VV delivery did not necessarily lead to a linear increase in transduction efficiency as shown for Tat, Pen, oligoarginine and

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

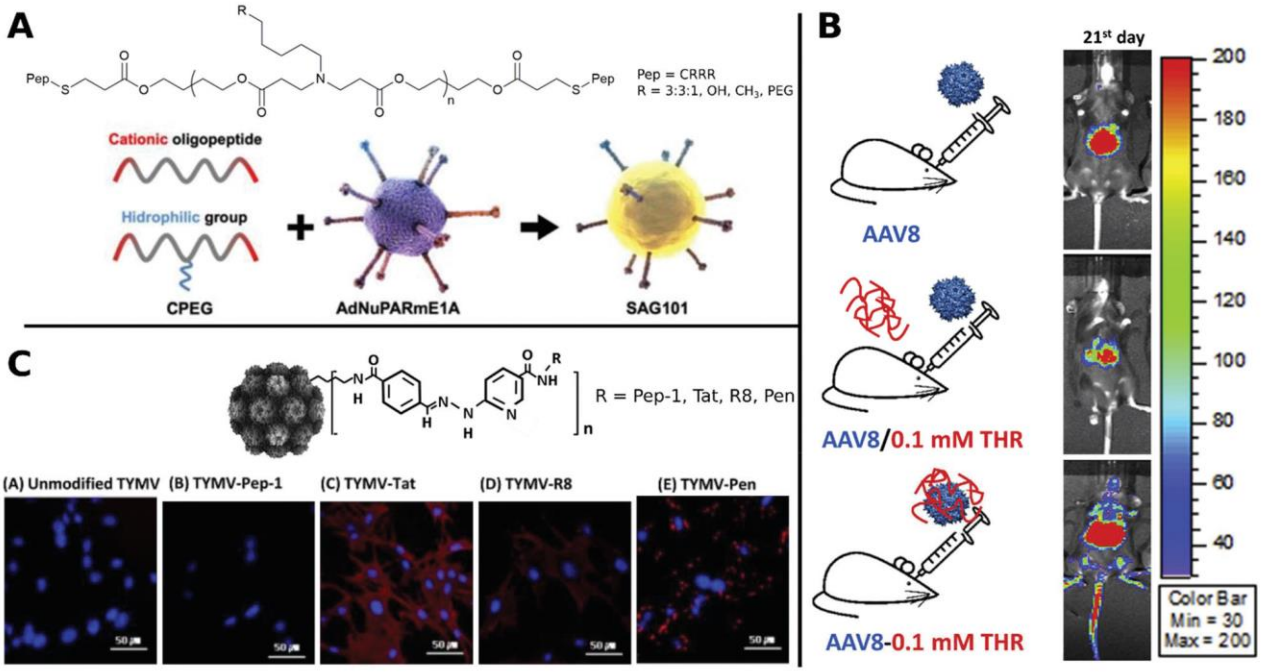


Fig. 12 Selection of CPPs. A. Schematic representation of oncolytic Ad coating with oligopeptides. Adapted with permission from ref. 312. Creative Commons BY 4.0. 2020 Ivyspring International Publisher. B. Schematic representation of three groups of AAV8 administration: AAV8 only, AAV8/0.1 mM THR (not incubated) and AAV8–0.1 mM THR (incubated). Luminescence expression images were taken on 21st day after systemic administration. Adapted with permission from ref. 321. Copyright 2018 Elsevier. C. Hydrazone conjugated TYMV–peptide. Confocal analysis of transfected BHK21 cells stained with Hoechst 33258 dye (blue) with CPP modified and unmodified TYMV visualized with AF594 (red) (scale bar: 50 μm). Adapted with permission from ref. 320. Copyright 2018 Elsevier.

Pep1.[316,318] Chirality also significantly affects cell permeability. A 9-mer of poly-D-arginine showed five-fold higher cellular uptake than poly-L-arginine.[327]

Mechanistic pathways for the TE of AAV–CPP complexes were investigated by delivering AAV to non-permissive cell lines and blocking specific receptors. The tested peptides Pen, Tat-HA2 and LAH4 facilitate energy dependent and independent endocytosis as well as receptor-mediated pathways for the internalization of AAV.[328] Further, Tat–Ad promoted cellular uptake *via* heparane sulfate receptors on the membrane surface, while the oligoarginine adenovirus conjugate R8-Ad was more dependent on chondroitin sulfate B receptors. This lays the foundation for CPP-dependent virus delivery.[316]

In a comparative study, TYMV was bioconjugated to Tat, R8, Pen or Pep-1. Improved efficiency was observed for Tat, R8 and Pen, whereas Pep-1 showed no change in transduction. This was traced back to different internalization routes and distributions of CPP–virus complexes in the cytoplasm and visualized by confocal images (Fig. 12C).[320]

Recently, Váňová *et al.* reported further mechanistic explanations for different TEs by investigating the influence of various CPPs on the activity and stability of VPs. They found that KH27K, FUSO, R8 and LAH4 affected the stability of VPs in different ways. KH27K promotes the aggregation and enlargement of VPs, while LAH4 destabilizes VPs but still enhances infection by concentrating them onto the cell surfaces (Fig. 13A).[329] This study provides an example for enhanced TE *via* virus disassembly, which was suggested in an earlier report covering Pen, Tat and R8[311] and provides the basis for further explanations of observed differences between CPPs.[320]

Peptide fibrils. The intrinsic ability of some peptides to self-assemble into higher order structures allow for numerous applications as functional biomaterials.[330,331] In the last ten years, peptide fibrils caused a paradigm shift in viral gene delivery from using mostly spherical particles to anisotropic, fibrillar nanocarriers, due to their convenient handling and high TE.

In contrast to CPPs, fibrillar peptides typically have a distinct secondary structure and cannot necessarily cross cell membranes. Enhanced viral gene delivery is achieved by colocalization of electrostatically complexed peptides and viral particles with the cellular membrane. Due to the high aspect ratio and rigidity of fibrils it is assumed that they cannot adapt to and cover virus surfaces thoroughly, thus creating excess positive charges in the virus–fibril complex for electrostatic interaction with cell membranes.[332]

The most thoroughly investigated fibrils for viral gene delivery are amyloids. Amyloid fibrils are highly stable and rigid, which makes them interesting materials for applications where long-term stability or tolerance of harsh conditions are required. For a long time, amyloid-forming peptides have exclusively been associated with diseases like plaque formation in Alzheimer's disease.[333,334] However, these structures were also found in non-pathogenic contexts as functional amyloids,

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

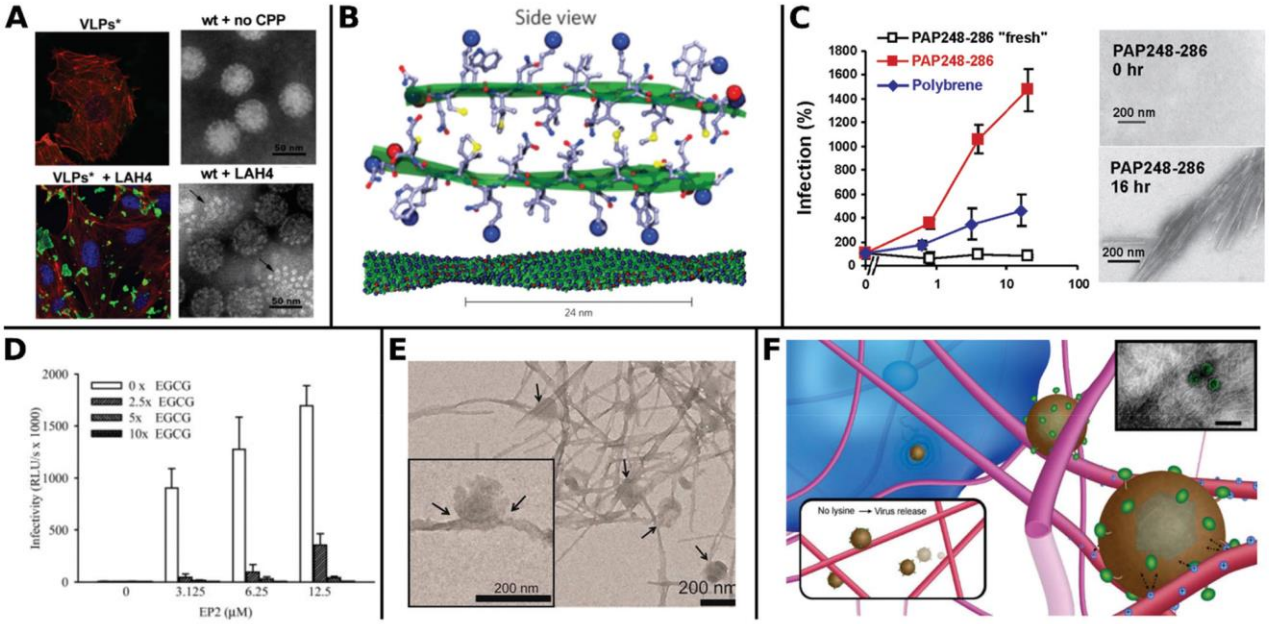


Fig. 13 Overview of fibril virus interaction and bioactivity. A. Left panels: Confocal analysis of 3T6 cells (actin = red, nucleus = blue) after association with MPyV virus like particles (green) without (upper image) and with LAH4 peptide (bottom image). Right panels: TEM images of virions in buffer were stable and compact (upper image), whereas coincubation with LAH4 led to virions disassembling into pentamers (bottom image) (scale bar: 50 μm). Adapted with permission from ref. 329. Copyright 2020 Elsevier. B. Molecular model of EF-C amyloid peptide fibril. Adapted with permission from ref. 118. Copyright 2013 Springer Nature. C. Overnight incubated PAP248–286 enhances HIV-1 infection of TZM-bl cells more efficiently than non-incubated fresh PAP248–286 or Polybrene. TEM images show structural change to peptides before and after incubation for 16 h (scale bar = 200 nm). Adapted with permission from ref. 345. Copyright 2007 Elsevier. D. Decrease of HIV-1 R5 infectivity after EGCG addition on EP2 peptides. Adapted with permission from ref. 344. Copyright 2014 John Wiley & Sons. E. TEM image of virions adhered to α-Syn fibrils (scale bar = 200 nm). Adapted with permission from ref. 117. Copyright 2018 American Chemical Society. F. Schematic representation of electrostatic interaction between LV and fibrillar peptide. Charge based immobilization of LV is achieved through peptide functionalization with additional Lysine to give Fmoc-DDIKVAVK. Inset shows TEM image of nanofibrous network with green colored LV (TEM scale bar = 200 nm). Adapted with permission from ref. 348. Copyright 2020 Springer Nature.

for example on bacteria surfaces[335] and in seminal fluids.[336] In the latter case, prostatic acidic phosphatase (PAP), a semen derived enhancer of virus infection (SEVI) with two distinct regions $PAP_{248–286}$ and $PAP_{85–120}$, and semenogelins (SEM) was found to form amyloidal structures and enhance HIV-1 infection.[336–338] These very first findings sparked further research in amyloid peptides for viral gene delivery.

Transduction enhancing peptides can also be found in virus envelopes. The optimization of the glycoprotein fragments $gp120_{417–428}$ (EF-C, commercialized as Protransduzin) and $gp41_{671–682}$ (P13 and P16) from the HIV-1 envelope and transmembrane protein, respectively, form small amyloid fragments that assemble into cationic nanofibrils of several hundred nanometers in length and a few nanometers in diameter (Fig. 13B). These short fragments showed higher TE compared to Tat, polybrene, DEAE and SEVI in various cell lines including difficult to transduce TZM-bl cells, while being cost effective and convenient in handling.[118,339] Fluorescent dyes coupled with free amino groups in EF-C have been introduced as a tool to study plasma stability and *in vivo* biodistribution while maintaining the structural and functional properties of non-labeled EF-C.[340] Point mutations in the peptide sequences had a great impact on TE as shown for P16-D, where a substitution of aspartic acid with alanine resulted in increased activity.[341] Sequence variations of $gp120_{417–428}$ yielded the 6-mer DPF1 and longer sequences EP2 and EP3 which accelerate amyloid formation of SEVI and SEM and enhanced TE.[342–344]

The fibrillar structure is necessary for enhanced VV delivery (Fig. 13C)[345] as shown by the following example: addition of epigallocatechin gallate, a main compound of green tea, which inhibits amyloid formation resulted in lower TE (Fig. 13D).[344] Amyloids can also help to broaden the tropism of viruses. For instance HIV-1 was able to transduce in hardly infectable CD4 negative cells with the addition of amyloidal CD4bs-M peptides.[346]

Moreover, the cofibrillation of amyloid α-Syn with positively charged polymers like poly-L-lysine or chitosan enhanced TE of amyloid fibrils as recently reported by Maji and coworkers. The authors suggested that these positively charged additives increased electrostatic interactions and local density of virions on cell surface, thereby facilitating the transduction (Fig. 13E).[117]

Besides using derivates of naturally occurring amyloids for vector delivery, amyloids can also be computationally designed as shown for KK and KY, which enhanced DNA delivery to

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

mammalian cells by forming DNA–amyloid complex and overcoming charge repulsion, as discussed for VVs.[347] Since synthetic amyloid fibrils have so far only been applied in *ex vivo* studies, the question of fibril stability and degradability *in vivo* remains open.

Hydrogels formed from fibrillar peptides are emerging as VV delivery scaffolds, due to their simple handling and spatiotemporal release of viruses. For instance, Fmoc-SAP derived from the epitope IKVAV enabled localized delivery of LV after implantation into the central nervous system of mice. The modification of the sequence with an additional lysine at the C-terminus increased electrostatic interactions and immobilized the VP to obtain focal gene delivery to the site of injection (Fig. 13F).[348] The peptide hydrogel RAD16-1 (commercialized as PuraMatrix™) is frequently used in cell culture. It displays favourable nano-structural and biomechanical properties and promotes the proliferation of various mammalian cells.[349] Rey-Rico *et al.* reported RAD16-1 hydrogels for durable genetic modifications to stem cells by enhanced localized AAV delivery over a period of 21 days *in vivo*.[234]

The LAH4 peptide family has attracted growing interest in recent years as a DNA transfection reagent as well as for VV delivery.[328,350,351] One commercialized derivative of the LAH4 family is Vectofusin-1.[351] This 26-mer cell penetrating peptide is the first transduction enhancer with α-helical fibrillar structure. Studies have shown enhanced TE of Vectofusin-1 compared to other delivery agents such as Tat, Pen, LAH4 derivatives, KH27K, R8, FUSO, polybrene, and retronectin.[328,351–355] Interestingly, a variation in the histidine sequence order or peptide length resulted in a significant change of bioactivity. A minimum length of 21 amino acids of the LAH4 peptide family was found to be necessary for successful vector delivery.[355]

Polypeptides. Polypeptides for VV delivery are mainly derived from natural sources and are typically used as scaffolds for enhanced delivery by colocalization.

Protamine sulfate is an early examples for viral transduction enhancers, originally derived from salmon sperm and clinically used for treatment of heparin overdose. This substance is actually a mixture of several polycationic peptides and facilitates virus–cell interactions by charge neutralization similar to other enhancers like polybrene, DEAE dextran or poly-L-lysine.[356,357] In comparative transduction studies of LV delivery to CD34+ stem cells, protamine sulfate showed a TE similar to Vectofusin-1 and higher than that of Retronectin.[243] TE could be enhanced even further by combining protamine sulfate with chondroitin sulfate.[358]

Retronectin is a 574 amino acid long recombinant polypeptide based on the fibronectin fragment CH296. It is one of the first commercially available enhancers for VV delivery and has already been tested in clinical trials.[359–361] A typical workflow for transduction involves the coating of Retronectin on a solid surface, the addition of viral particles and cells followed by spin centrifugation.[362] To overcome the limitations of the laborious preparation on surfaces and the necessity of centrifugation, various alternative approaches have been developed. The coating of epoxy-modified paramagnetic beads with Retronectin enables the effective capture of retroviral particles from solution and provides access to remote-controllable transduction.[363] *In vivo* LV delivery to stem cells *via* intra-bone marrow injection in mice showed enhanced TE when Retronectin was co-injected with the virus.[364]

Poly-L-lysine (PLL) is a popular additive in cell cultures for reinforcing adhesion of cells to culture dish. It is used as short oligomers up to polypeptides of several 100 kDa and can either be covalently attached to virus surface or simply coated on surfaces.[365–367] PLL has been investigated as a replacement for polybrene in viral delivery by creating the block copolymer PEG_{12000}–PLL_{48} which was coated on RV through electrostatic interactions. In contrast to polybrene, this copolymer augments RV transduction to delicate primary cultured brain cells without cytotoxic effects.[368] Further optimization of PEG–PLL compositions were made by functionalizing four-arm PEG acrylates with PLL of different molecular weights ranging from 1–70 kDa. Increased TE was observed for increased molecular weights of PLL.[369] Moreover, PLL has been applied as a mimic of the viral envelope and enhanced cellular internalization and transduction of virus-like particles.[367] Recently, modifications of PLL_{120} with *p*-toluylsulfonyl arginine resulted in change of conformation from random coil to α-helix, which gave higher DNA-transfection rates due to higher cell-permeability.[370] Layer-by-layer deposition – referred to as protein surface precipiation here – of PLL and dextran sulfate sodium salt were applied onto Ad surface to obtain nanospheres that protect Ad from antigenic reactions and thereby enhanced TE (Fig. 14A).[371] These recent findings highlight the importance of material deposition techniques and conformational structure for viral gene delivery.

Enhanced viral delivery with spatiotemporal control can be achieved by fibrous protein scaffolds, such as fibrin. For example, when VPs were incorporated in fibrin, a local, long-term release of Ad,[372] AAV[235,373] and LV[374] *in vitro* and *in vivo* was observed at the injection site up to 14 days. Fibrin scaffolds could be further improved by the addition of hydroxyapatite nanoparticles, resulting in stronger trapping as well as slower release of VPs from scaffolds[374] (Fig. 14B) and gels.[126]

Vupputuri *et al.* reported that protein impurities in cell culture can lead to increase or decrease of viral TE. By carefully adjusting the amount of protein added to viral particles, they could improve TE of polymer–virus complexes. Proteins bound to VPs can increase TE due to sedimentation of heathier particles, however at the same time cellular uptake pathway is changed by larger VPs thus protein addition has to be carefully balanced.[132] This observation have far-reaching consequences for comparability of studies concerning TE, when taking batch-to-batch variability of culture medium and VP purification into account. For example, proteins from the human serum can also augment AAV transduction. Cationic human serum albumin (cHSA), an abundant endogenous transporter protein in blood, increased transduction of an *in vivo* hemophilia B mouse model fivefold.[375] PEGylated cHSA improved TE of HIV-1 to TZM-bl cells while displaying lower immuno-

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

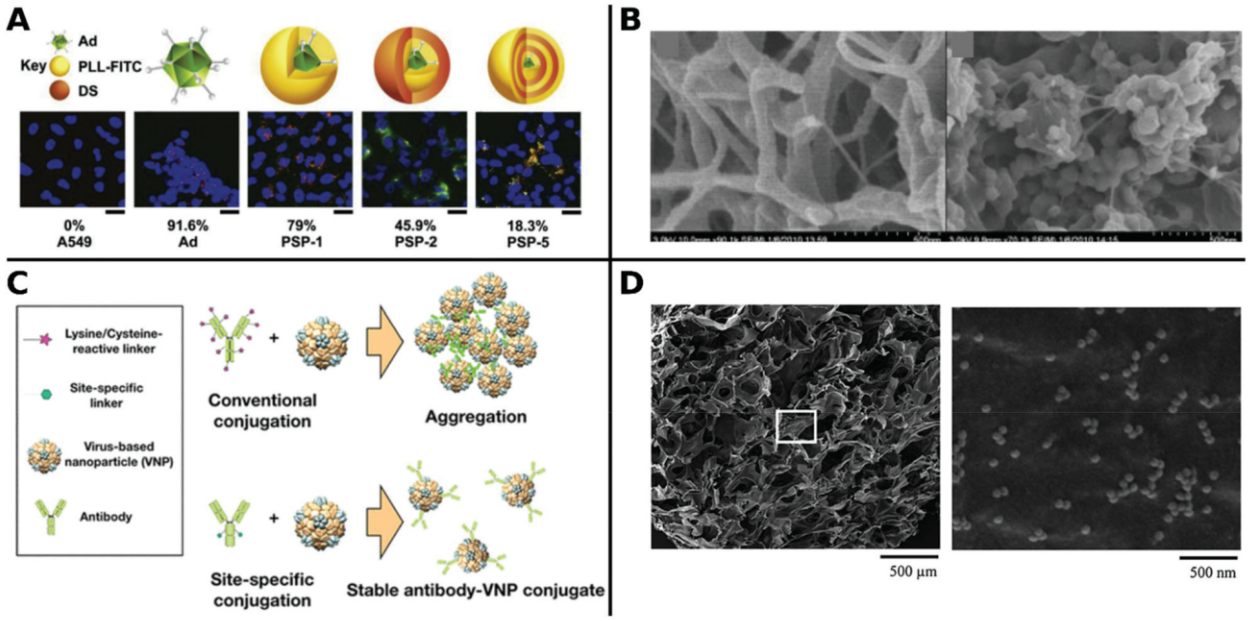


Fig. 14 A. Schematic representation of naked Ad and layer-by-layer PLL/dextran sulfate sodium nanosphere coated Ad (PSP). Fluorescence images show gene expression with percentual transduction rates (scale bar: 15 µm). Adapted with permission from ref. 371. Copyright 2019 Elsevier. B. SEM images of fibrin scaffold with LV incorporated without (left) and with hydroxylapatite (right) (scale bar: 500 nm). Adapted with permission from ref. 374. Copyright 2012 Elsevier. C. Schematic illustration of site-specific antibody conjugation on VP capsid surface. Adapted with permission from ref. 378. Copyright 2020 American Chemical Society. D. SEM images of Ad immobilized on gelatin sponges (scale bar: left 500 µm, right 500 nm). Adapted with permission from ref. 385. Copyright 2013 John Wiley & Sons.

genicity.[376] Combinations of HSA, low density lipoproteins (LDL) or transferrin yield AAV–serum protein complexes which enhanced transduction to liver *in vivo*, prevented binding of AAV to other proteins, and suppressed nonspecific binding to cells.[377] Finally, site-specific antibody conjugation to TYMV, CPMV and bacteriophage *via* azide–alkyne click chemistry was reported by Park *et al.* (Fig. 14C).[378]

These examples highlight that natural proteins can efficiently enhance viral delivery. However, concerns of potential immunogenicity of xenogeneic materials exist.[379] Nature-inspired, synthetic silk elastin-like protein polymers (SELPs) may circumvent this problem. SELPs mainly consist of repeating amino acid blocks based on the typical silk (GAGAGS) and mammalian elastin (GVGVP) motifs and were first reported for Ad release in cancer gene therapy in 2004.[380,381] SELPs showed enhanced, local, long-term Ad delivery up to four weeks in various studies.[382,383] In a comparative study with poloxamer 407, SELPs showed promising results with a superior TE, lower toxicity and immunoreactivity.[384] Furthermore, collagen, an abundant fibrous protein in the extracellular matrix, showed long-term localized gene expression *in vivo* up to 4 weeks when it was co-gelated with LV.[126] Organic–inorganic hybrid systems, such as collagen hydrogels combined with hydroxyapatite nanoparticles retain LV over a period of four weeks *in vivo* and increase local delivery efficiency even more.[126] Moreover, gelatin gels, which consist mainly of hydrolysed collagen, were also applied for viral delivery. Gelatin sponges and Ad were functionalized with biotin to link them after avidin addition. The virus tethered in a virus–biotin–avidin–biotin–gelatin arrangement showed enhanced spatiotemporal TE *in vivo* after implantation for bone regeneration (Fig. 14D).[385] Local long-term gene expression was also shown for oncolytic Ad delivery from gelatin gels *in vivo* up to 20 days. These gels could further protect Ad from immune response and prevent unwanted delivery to healthy sites.[386]

Lipids

Lipids are amphiphilic molecules typically consisting of a hydrophilic head group and a hydrophobic tail. Lipids are the main component of cellular membranes, virus envelopes and extracellular vesicles, which are Nature's endogenous nanocarriers to deliver biological information from cell to cell.[387] The unique ability of lipids in aqueous media to form bilayers in a vesicular structure (liposomes) and spherical monolayers (micelles) enables them to encapsulate both hydrophilic and hydrophobic cargo.[388,389] After encapsulation the cargo is protected from degradation as well as immune reaction[390] and can further address specific targets by functionalization at the liposome's outer sphere.[391] Since their first discovery in the 1960s,[392] liposomes have been extensively used as delivery agents in clinical trials and pharmaceutical industry.[393,394]

In non-viral gene delivery, lipids are currently the gold standard. For example, liposomes were applied as DNA and RNA nanocarriers[395,396] as well as mimics of the viral envelope for gene delivery.[397] Recent reviews on non-viral application of lipids for gene delivery can be found elsewhere.[398,399] Lipids

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 5 Overview of lipids and lipid formulations, properties and fabrication methods

Name	Structure	Properties	Fabrication of VP lipid complexes	Ref.
Chol	Cholesterol	z., fuso.	Inc. (Ad)	405
DOTAP/Chol	(1,2-Dioleoyloxypropyl)-*N*,*N*,*N*-trimethylammonium chloride/cholesterol	c.	Lipid bilayer wrapping (Ad)/inc. in preformed liposomes (Ad)/self-assembly around VP (Ad)	409, 390, 53 and 416
DOTAP/DOPE	DOTAP/(1,2-dioleoyl-*sn-glycero*-3-phosphoethanolamine)	c./fuso., z.	Self-assembly around VP (Ad)	409
DOTAP/DOPE/Chol		c./fuso. z.	Inc. (dry film or extrusion method) or assembly around VP (RV)	410
DMPC/Chol	Dimyristoyl phosphatidylcholine/cholesterol	z.	Self-assembly around VP (Ad)	53 and 409
DMPC/Chol/apoA-1	DMPC/Chol/apolipoprotein A-1	z.	Inc. in preformed liposomes (Ad)	124
DMPC/Chol/DSPE-PEG		z.	Self-assembly around VP (Ad)	53
TMAG/DLPC/DOPE	*N*-(α-Trimethylammonio-acetyl)didodecyl-D-glutamate chloride/dilauroyl phosphatidylcholine/DOPE	z.	Inc. in preformed liposomes (AAV, rAd)	407 and 417
DC-Chol/DOPE	3 beta [*N*-(*N*′,*N*″-dimethylaminoethane)-carbamoyl]cholesterol/DOPE	z.	Inc. in preformed liposomes (RV, Ad + siRNA)	401 and 423
Lipofectamine (DOSPA/DOPE (3/1))	2,3-Dioleyloxy-*N*-[2(sperminecarboxamido)ethyl]-*N*,*N*-dimethyl-1-propaninium trifluoroacetate/DOPE	c., fuso.	Inc. in preformed lipid-DNA complex (AAV)/inc. in preformed liposomes (RV/HSV/Reovirus)	415, 400, 418 and 420
Lipofectin (DOTMA/DOPE (1/1))	*N*-[1-(2,3-Dioleyloxy)propyl]-*N*,*N*,*N*-trimethylammonium chloride/DOPE	c., fuso.	Inc. in preformed liposomes (RV)	400 and 414
Lecithin/Chol (4/1)		z., fuso.	Inc. in preformed liposomes (alphavirus)	421
Lysolecithin	L-α-Lysophosphatidylcholine	z., fuso	Preconditioning of host cells prior to LV injection	424–427
PC/PG/Chol	Phosphatidylcholine/phosphatidylglycerol/Chol	z.	Reverse-evaporation unilamellar vesicle (MLV)	406
PEGDE	[1,2-Distearoyl-*sn-glycero*-3-phosphoethanolamine-*N*-[methoxy(poly-ethylene glycol)-2000]		Inc. in preformed liposomes (Ad)	432 and 433
DSPE-PEG-biotin	(1,2-Distearoyl-*sn-glycero*-3-phosphoethanolamine-*N*-[biotinyl(polyethylene glycol)2000])		Inc. in preformed liposomes (RV)	435
PC/CHEMS/Chol-PEG-MMP-peptide	PC/CHEMS/Chol-PEG2000-GPLGIAGQC	a., MMP responsive	Inc. in preformed liposomes with Ca^{2+} ions (Ad)	413
DOPE/CHEMS (3/2)	DOPE/cholesteryl hemisuccinate	a., fuso	Inc. in preformed liposomes with Ca^{2+} ions (Ad)	412
CHEMS/PC/Chol (4/5/1)	CHEMS/phosphatidylcholine (PC)/Chol	a.	Inc. in preformed liposomes with Ca^{2+} ions (Ad)	123
Lecithin/Chol/DSPE-PEG/folate		a.	Inc. in preformed liposomes (Ad)	422

Abbreviations: cationic (c.), hydrophilic (h.), zwitterionic (z.), anionic (a.), incubation (inc.), fusogenic (fuso), adenovirus (Ad), adeno-associated virus (AAV), lentivirus (LV), retrovirus (RV), herpes simplex virus (HSV), murine leukemia viruses (MLV).

for viral gene delivery, which are discussed in the following section are summarized in Table 5.

Formulation. In viral gene delivery, the most frequently used lipids (Chart 2) are cationic phospholipids (*e.g.*, DOTAP, DOSPA and DOTMA), zwitterionic helper lipids (*e.g.*, DOPE) and sterols (*e.g.*, cholesterol (Chol)).[400] The internalization efficiency into cells is highly dependent on the formulation of lipid mixtures.[388,400] Usually the negatively charged viral particles are encapsulated in liposomes and subsequently taken up into cells.[150,400] For example, commercialized transfection agents like Lipofectamine™ and Lipofectin™ contain the mixed formulations DOSPA/DOPE (3/1) and DOTMA/DOPE (1/1), respectively.[400,401] The zwitterionic helper lipid DOPE is added to the formulations to enhance TE by direct fusion with the cell-membrane and bypassing cellular endocytosis, thereby avoiding inefficient endosomal escape.[402] Addition of cholesterol changes at certain concentrations the order of phospholipid assembly and results in more rigid liposomes.[403] However, when pure Chol is applied, it increases infectivity of Ad by forming Chol–Ad clusters *via* hydrophobic interactions with capsid proteins.[404,405]

Formation and release. Complexes of lipids and VP are typically obtained by simple mixing of the VPs in aqueous solution with preformed liposomes on a thin film and subsequent incubation.[406,407] However, concerns regarding this method have been raised due to toxic and transduction ineffective aggregate formation.[408] Alternative encapsulation routes such as spontaneous self-assembly around VPs (Fig. 15A) have been demonstrated to circumvent these problems with partial success.[53,409,410]

In general, it is assumed that TE of VPs is increased by accelerated cell entry due to the small size and positive charge of liposomal carriers[53,410] and further promoted by the so-called proton sponge effect, which has recently been critically

 This journal is © The Royal Society of Chemistry 2020

View Article Online

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

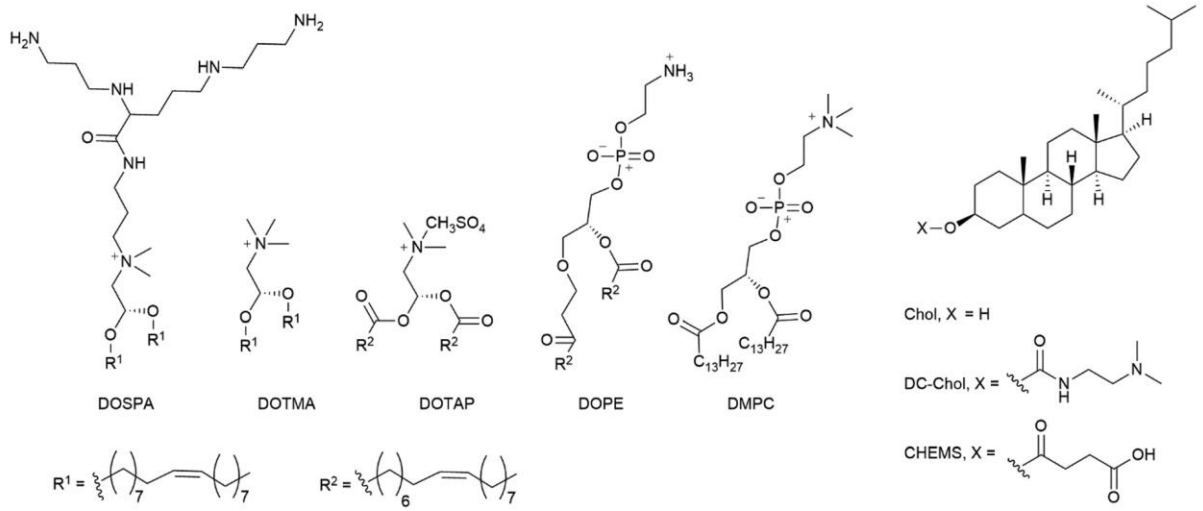


Chart 2 Overview of most frequently used lipids for viral gene delivery.

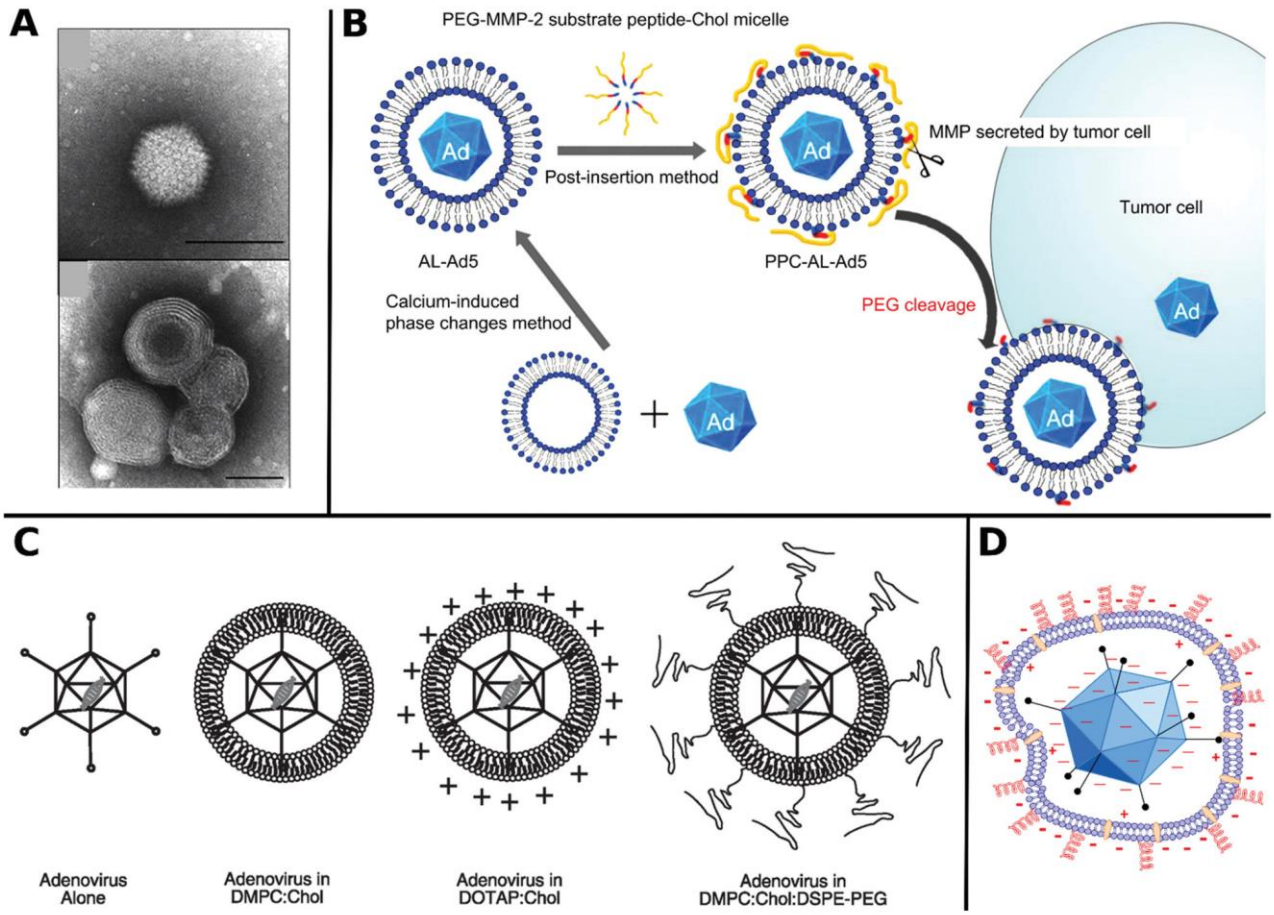


Fig. 15 A. TEM images of naked Ad (top) and DOTAP/Chol encapsulated Ad (bottom) (scale bar: 100 nm). Adapted with permission from ref. 409. Copyright 2008 American Chemical Society. B. Schematic illustration of enzyme-responsive liposome-encapsulated Ad5 for specific delivery in tumor cells. Adapted with permission from ref. 413. Copyright 2013 Elsevier. C. Schematic representation of lipid encapsulated Ad. Adapted with permission from ref. 53. Copyright 2008 John Wiley & Sons. D. Schematic diagram of electrostatic interactions between zwitterionic phospholipid lecithin with negatively charged Ad5. Cholesterol is shown between lipid bilayer and PEG chains on the outer sphere as red spirals. Adapted with permission from ref. 422. Copyright 2014 Elsevier.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

reviewed.[195] The proton sponge effect is caused by an increase of cationic charge at endosomal acidic pH, leading to osmotic swelling and accelerated liposome dissociation in the cytoplasm.[129,411] Apart from this, various supporting strategies have been developed to facilitate the endosomal escape by creating stimuli-responsive liposomes, for example, by employing pH-triggered release of Ad from DOPE/CHEMS[412] or MMP-cleavable enzyme-responsive liposomes (Fig. 15B).[413]

Most studies were conducted with complexes from cationic liposomes and VPs. Enhanced TE, shielding from immune response and broadened tropism of RV,[400,401,410,414] AAV,[415] Ad,[124,416,417] HSV,[418,419] as well as recently for oncolytic applications with reovirus[420] and alphavirus[421] were reported. Moreover, liposome vesicles from zwitterionic, cationic and PEG-lipid formulations can act as an artificial envelope for non-enveloped viruses and thereby enhance targeted delivery and prolong blood circulation lifetime (Fig. 15C and D).[53,412,422] Further co-envelopment of Ad and siRNA yielded a dual active hybrid vector for multiple gene delivery.[423]

Not all applications with lipids for virus delivery require preformed VP–lipid complexes. Parsons *et al.* applied lysolecithins as surfactants on host cells prior to LV administration for cystic fibrosis treatment in mice. In this way, a more effective, less immunogenic and persistent transduction was achieved already with a single injection.[424–427] Maguire *et al.* were the first to report and isolate AAV in extracellular vesicles, which were released from AAV producer cells during the reproduction process.[428] These naturally encapsulated AAV showed enhanced TE and high biocompatibility at various application sites *in vivo* and has been reviewed elsewhere.[429]

Modified lipids. One drawback of cationic liposomes in systemic administrations is unwanted interaction with serum proteins and subsequent clearance.[430] To tackle this issue, lipids have been covalently attached to PEG to achieve stealthy carriers with longer circulation times and protection from blood plasma protein adsorption.[166,431] These systems also enable tumor[413,432–434] and cell targeting.[435] In a different approach, PEG spacers were used as surface functionalization tools for liposomes, *e.g.* for covalent attachment of the CPP Tat or Pen.[391,436]

Anionic lipids. Anionic liposomes, on the other hand, were introduced in the last decade and emerged as carriers for VVs with enhanced TE and low toxicity and immunogenicity.[123,437] They are more compatible with the abundance of negatively charged molecules in a physiological environment and therefore less prone to recognition by neutralizing antibodies. Due to the same charge polarity of VPs and anionic liposomes calcium-ions forming bridged complexes were required for encapsulation.[123,437] Surprisingly, enhanced TE is possible despite the anionic charge[438] even in a CAR-independent manner for encapsulated Ad.[123,422,437] The mechanism that promotes interactions of anionic liposomes with plasma membranes is still not fully understood. Possible reasons, such as longer interaction time due to reduced clearance and toxicity,[437,439,440] faster release of cargo,[441,442] or receptor mediated mechanisms[443] have been discussed. A thorough understanding of mechanisms promoting interactions of anionic liposomes with VPs is yet to be achieved.[444,445] This is of significant importance, especially with regard to oncolytic viral gene delivery, where an elevated presence of anionic lipids is observed on cancer cell membrane surface.[446,447]

Nanoparticles

Nanoparticles (NPs) are nanoscale materials, which have attracted growing interest in recent years due to their versatility and unique size-scaling properties. For example, NPs made from noble metals such as gold display strong surface plasmon absorption, making them an interesting material for local phototherapy and as an imaging agent *in vivo*.[448,449] Among manifold preparation techniques to obtain uniform, functional NPs, template-based strategies are well established. VPs are attractive templates due to the uniform, highly precise 3D structure of the virus capsid which are decorated with chemical functional groups.[450]

NPs have also emerged in recent years as an invaluable asset for viral gene delivery. In general three approaches can be distinguished: (1) NPs and VP administered together without specific association between them; (2) VPs and NPs that are covalently bound, *e.g.* in a conjugate; and (3) non-covalent interactions such as electrostatic binding between VPs and NPs. Depending on the NP–VP size ratio, various architectures can be achieved. If the NPs size exceed the VPs size *e.g.* iron oxide NPs combined with AAV[120] VPs are decorated on the outer sphere of NPs. However, if the VPs are larger than NPs typically, NPs and VPs are combined either by coating VP surface *via* incubation and precipitation or by covalent attachment through functional groups on VP surface. For example, mineral shells for VPs can be created by precipitation from a solution of an inorganic salt and exhibit new biological and physical properties impacting stability, host cell adsorption and entry mechanism.[451]

Covalent attachment of NPs to VPs is accessible by abundant lysine residues on capsid proteins and by genetically engineered functional groups, *e.g.* thiol- or azide-moieties.

Frequently applied NPs for viral gene delivery are made from calcium phosphate, gold, silica, graphene or iron oxide, are discussed in the following sections and summarized in Table 6.

Calcium phosphate. Calcium phosphate ($Ca_3(PO_4)_2$) is highly abundant in human bone and teeth and one of the most important biominerals in Nature.[452] As a common material in humans and animals, $Ca_3(PO_4)_2$ has a significant role also for viral infectivity, as reported for avian sarcoma virus. It is suggested that this virus occurs in aves in mineralized state and promotes transmission and bypassing of the zoonotic barrier to humans by a highly effective adsorption and $Ca_3(PO_4)_2$ mediated internalization pathway into host cells.[113] At neutral conditions, $Ca_3(PO_4)_2$ forms a solid precipitate around VPs, which can be degraded under acidic condition present in the endosome.[453] $Ca_3(PO_4)_2$ is an ideal candidate as a biomineral due to its biocompatible properties and degradability into non-toxic ions after cellular entry (Fig. 16). It

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 6 Overview of nanoparticles utilized for viral delivery, properties and fabrication methods

Name	Structure	Properties	Fabrication of VP–NP vectors	Ref.
$Ca_3(PO_4)_2$	Calcium phosphate	Shell like encapsulation	Coprecipitation by inc. in $CaCl_2$ and subsequent biomineralization through addition of Na_2HPO_4 (Ad5)	112
$Ca_3(PO_4)_2$, DOPA, PEG-DPPE	Calcium phosphate, dioleoylphosphatydic acid, 1,2-dipalmitoyl-sn-*glycero*-3-phosphoethanolamine-*N*-[azido (polyethylene glycol)-2000]	3 layered encapsulation	Coprecipitation in Ca-rich medium, inc. by assembly of lipid bilayers of DOPE, coassembly of PEG-DPPE with DOPE (oAd)	434
$CaCO_3/MnCO_3$		Shell like encapsulation	Coprecipitation from $CaCl_2$ and $MnCl_2$ with Na_2CO_3 (oAd)	454
Gold nanorods		NP = *ca.* 60 × 30 × 1 nm	Coinjection (oAd)	463
Gold NPs	Funct. with PEG/RGD or azide-group	NP = 14 nm	Cov. conj. (Ad)	457
Gold NPs	Au-sulfo-NHS	NP = 1.4 nm	Cov. conj. (Ad)	456
Silica gel	SiO_2 gel polymerized from tetraethoxysilane	Macroporous hydrogel	Cogelation from sol–gel drying process (Ad)	128
Silica cloak	SiO_2	Anionic	Coprecipitation from silic acid in aqueous media on Ad pretreated with PLL	125
SPIONs	Fe_3O_4 (superparamagnetic iron oxide nanoparticles)	NP = 10 nm	Cov. conj. (Ad)/inc. (Ad/measles virus/LV	498, 490, 505, 494, 484 and 493
SPIONs–COOH	Fe_3O_4 functionalized with carboxylic acid	NP = 5 nm	Cov. conj. (AAV)	489, 502 and 503
SPIONs/heparin	Heparin coated SPION	NP = 100 nm	Inc. (AAV)	120
SPIONs/chitosan	Chitosan coated SPION	NP = 50–150 nm	Inc. (Ad)	497
SPIONs/PEG2000	PEG2000 cross-linked SPION	NP = 117 nm	Inc. (Ad)	492
SPIONs/PEI or polybrene or Si-PEI	PEI/polybrene/Si-PEI coatings on SPION	NP = 28 nm/ 64 nm/101 nm	Inc. (oAd)	491
NiNTA-biotin–streptavidin SPIONs	Fe_3O_4 particles conjugated with streptavidin–biotin–nitrilotriacetic acid–Ni	NP = 120 nm	Hexahistidine coding AAV interact with Ni-ions chelated on NPs.	115
PEI–GO–PEG–FA	Polyethyleneimine–graphene oxide sheets–polyethylene glycol–folic acid	NP = 25 nm	Inc. (measles virus)	481
Carbon dots	Carbon dots, citric acid, branched PEI	NP = 11–36 nm	Inc. (rAAV)	482

Abbreviations: incubation (inc.), adenovirus (Ad), oncolytic adenovirus (oAd), adeno-associated virus (AAV), lentivirus (LV).

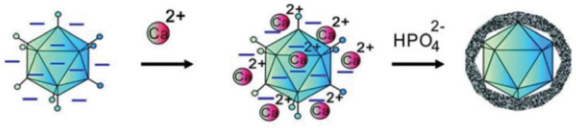


Fig. 16 Schematic illustration of *in situ* mineralization of Ad5 by precipitation with $Ca_3(PO_4)_2$. Adapted with permission from ref. 112. Copyright 2012 John Wiley & Sons.

has been applied as a virus shell by biomineralization and enabled enhanced CAR-independent infectivity, expanded tropism and shielding from neutralizing antibodies.[112,113] Chen *et al.* described a combinatorial design in which oncolytic Ad was subsequently coated with $Ca_3(PO_4)_2$, DOPA lipids and PEG to form a spherical three layered charged shell with almost neutral charge. This design masked the viral surface and lowered unwanted immune response and liver toxicity while improving TE to tumor tissue.[434] Similar, easy to fabricate virus–inorganic complexes from $Ca_3(PO_4)_2$ are promising for systemic administration due to low administration dosage and prolonged circulation lifetime.[112] Recently, Huang *et al.* proposed a more elaborate calcium-based shell for oncolytic Ad delivery, which was created from calcium and manganese carbonates. This shell prolonged *in vivo* circulation of VPs, protected them from immune response, significantly reduced their liver accumulation while still resulting in high tumor accumulation. In the acidic tumor microenvironment, Mn^{2+}-ions were released and converted endogeneous H_2O_2 into O_2. Thereby oncolytic Ad duplication activity and antitumor effects were enhanced. Due to the increase of O_2 concentration, real time, label-free monitoring with magnetic resonance imaging was possible.[454]

Inorganic particles from $Ca_3(PO_4)_2$ can provide stealth biomineral shells by spontaneous coating of VPs surface and thereby have an influence on tropism, TE and biocompatibility. Combinations with other inorganic particles enable precise tuning of the release properties and give access to tailored carriers.

Gold. Gold NPs have been explored for various biomedical applications like diagnostics, radiotherapy or gene delivery due to their easy fabrication, biocompatibility, and functionalization.[455]

Gold NPs were applied as covalent conjugates with VPs and reported either to retain or enhance TE[456,457] or attenuate infectivity,[458–461] depending on the capsid modification. Covalent conjugates were mainly achieved by attachment to the abundant lysine residues of capsid proteins. However,

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

modifications with high gold amounts interfered with capsid protein functions, which are important for virus infectivity.[456] Nevertheless, enhanced transduction was shown with various positively charged PEGylated Gold NPs functionalized with RGD epitope or azide group and coincubated with Ad, which formed spherical particles and exhibit higher TE to hMSCs compared to PEI or Lipofectamine coated Ad (Fig. 17).[457]

A further interesting clinical application of Gold NPs is photothermal therapy,[462] which has been efficiently addressed in a combinatorial therapeutic approach with VPs.[456] Oncolytic Ad applied in combination with gold nanorods exhibit hyperthermia-induced enhancement of receptor-mediated endocytosis and TE *in vivo*, enabling locally induced tumor cell lysis.[463]

Silica. Silica NPs are biodegradable minerals which are typically fabricated in a sol–gel process starting from tetramethyl orthosilicate or tetraethyl orthosilicate and are broadly employed as food additives or in biomedical applications.[464,465] Silica coatings for VPs are promising candidates to broaden the tropism of viral gene delivery, since a size dependent uptake by tumor cells is known for silica nanoparticles.[466] Recently, Ad coated with silica showed enhanced TE in glioma while reducing liver accumulation and immunogenicity.[125] Other formulations of silica have also been tested. Silica-based gels are synthesized through a simple sol–gel process, resulting in highly porous materials, suitable as implantable viral depots with long-term release. Typically, the degradation of silica gels is in the range of weeks and depends on the amount of water which is present. For example oncolytic Ad were encorporated in a biodegradable high water content silica-based gel and spatiotemporal release of oncolytic Ad in implanted tumor tissue site was achieved by gradually degradation *in vivo.* The virus could be stored in the silica gels and infectivity was preserved for a period of at least two years *in vitro* at 4 °C.[128]

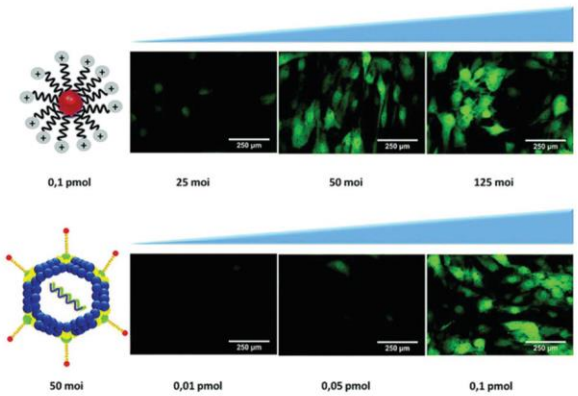


Fig. 17 Optimization process of Ad conjugated to gold NPs by maintaining gold NP concentration fixed at 0.1 pmol and varying amount of VPs (moi, multiplicity of infectivity, top). The best ratio at 50 moi was kept and gold NP concentration was varied (bottom) to observe enhanced TE at higher gold NP concentrations. Adapted with permission from ref. 457. Copyright 2019 The Royal Society of Chemistry.

Graphene. Graphene-based NPs display unique physiochemical properties, like electrical conductivity or photostability and have emerged as an abundant and low-cost resource for nanomedical application in recent years.[467] For example, they have been applied in tumor treatment,[468] photothermal therapeutic approaches,[469–471] virus detection[472] and have been used as drug[473,474] and gene delivery agents.

Since pioneering works from Kostarelos *et al.*[475,476] research efforts have been focussed on graphene as non-viral transfection agents, often in combination with grafted cationic polymers as adhesion promoters.[477–480]

Recently, graphene have started to attract interest as viral delivery platforms. Graphene oxide (GO) sheets decorated with PEG-folate were shown to target susceptible cancer cells.[468] Based on these results, oncolytic measles virus was encapsulated by GO sheets functionalized with PEI and PEG-folate (Fig. 18). The encapsulated measles virus was protected against neutralizing antibodies, showed high tumor accumulation *in vivo*, exhibited higher infectivity and TE to cancer cells than virus alone and folate-free GO–PEI–PEG sheets. However, the release mechanism and *in vivo* stability of the VP–GO sheet cluster remain unexplored.[481]

Another carbon-based material, carbon dots, spherical particles of few nanometer-size diameter, have been applied for viral gene delivery for the first time.[482] This relatively new material class, mostly derived from graphene, has been reported to show size-, charge-, and aggregation-dependent toxicity.[483] In order to stimulate cartilage repair, rAAV coding for the highly chondroreparative SOX9 transcription factor was complexed with various functionalized carbon dots, which resulted in significant increase of transduction to hMSCs with high cell viability. Especially carbon dots functionalized with 2-citric acid, poly(ethylene glycol) monomethyl ether and *N*,*N*-dimethylethylenediamine displayed efficient release and TE of rAAV.[482]

Iron oxide. Iron oxide NPs have an outstanding position among inorganic NPs for viral gene delivery and have been intensively studied as an interface technique between physical and chemical delivery. They have especially attracted interest as the main component for remotely guided magnetic delivery vehicles. Gene delivery by using a gradient magnetic field, also

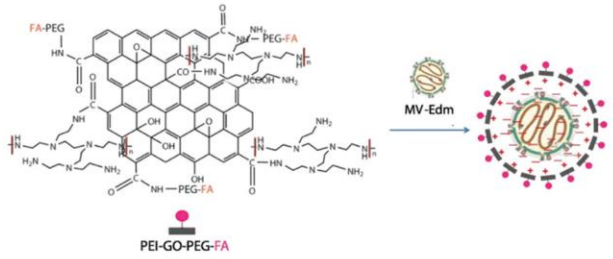


Fig. 18 Structure of PEI–GO–PEG–FA and schematic illustration of GOS/measles virus (MV-Edm) complexation for cancer therapy. Adapted with permission from ref. 481. Creative Commons BY 4.0. 2019 Springer Nature.

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

called magnetofection, has been reported to enhance TE of non-viral vectors and VVs after employing external magnetic fields for *in vitro* and *in vivo* applications[484–486] and have been successfully commercialized for example as AdenoMag (OZ Biosciences),[487] fluidMAG-Heparin (Chemicell)[120] or Endorem (Guerbet).[488]

Magnetite (Fe_3O_4) as superparamagnetic iron oxide nanoparticles (SPIONs) have been reported in several studies as a vehicle to enhance gene delivery efficiencies, *e.g.* for delivery of AAV,[115,489] Ad,[487,490] oncolytic Ad,[491,492] LV,[484,493] and measles virus.[494] They are biocompatible and degradable and allow surface modification with various functionalities.[495,496]

Coating SPIONs with biopolymers like heparin[120] or chitosan[497] can further augment TE with lower vector dosage and enable faster transduction (Fig. 19A). Moreover, polymeric coatings such as PEI or PEG on SPIONs as well as liposome encapsulation are reported to be highly beneficial for improving TE of Ad with reduced dose (Fig. 19B).[484,492,498] For example, PEI coatings give access to positively charged particles boosting cell interaction and endosomal escape by their proton sponge effect without affecting the superparamagnetic properties.[499] In a comparative study PEI, polybrene, or silica–PEI were coated onto iron oxide NPs and resulted in improved potency of oncolytic Ad. The highest antitumor activity *in vivo* was observed for silica–PEI coated iron oxide NPs, which was attributed to a higher shielding ability against neutralizing antibodies and higher magnetophoretic mobility.[491] In another study, various commercially available SPIONs were compared to each other in terms of susceptibility for aggregation and TE of attached alphavirus. Positively-charged NPs displayed rapid and effective isolation and concentration of VPs, resulting in higher TE than negatively charged NPs.[500]

In general, VPs can be coated with modified SPIONs through simple incubation due to electrostatic interactions to obtain hybrid particles.[487,492] Beside this, covalent attachment of VPs onto SPIONs is regularly achieved by EDC–NHS chemistry coupling carboxy-functionalized SPIONs to the lysine residues on viral capsid.[489,498]

Another, more laborious route to attach VPs onto SPIONs surfaces can be achieved by using a chelating approach, with a linker unit consisting of biotin-functionalized nickel ions conjugated to streptavidin modified SPIONs, which further bind to hexahistidine displaying AAV particles (Fig. 19C). This hybrid particle can transduce into otherwise non-permissive hNSCs by employing magnetic force.[115] Recently, precise manual virus stamping to single cells in culture as well as to mouse brain was reported by Schubert *et al.* This technique employs VPs covalently bound to silica-coated SPIONs which are loaded in a micropipette tip and brought into contact with the desired cells by a magnetic field (Fig. 19D). This method enables the delivery of multiple virus types to a single cell for multiple virus engineering or to track genetically modified single cells.[501]

An interesting example to implement magnetofection with phototherapeutic tumor treatment was shown for AAV encoding a photosensitive apoptosis-inducing protein. AAV conju-

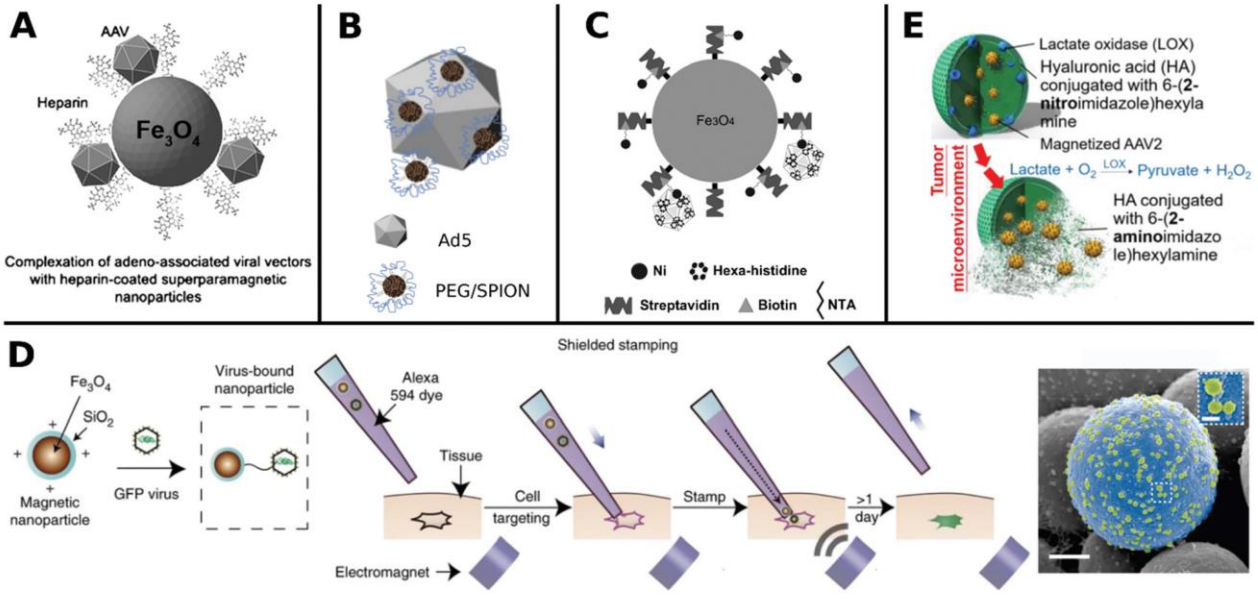


Fig. 19 A. Schematic illustration of AAV bound to heparin-coated SPIONs. Adapted with permission from ref. 120. Copyright 2011 Elsevier. B. Schematic illustrations of PEG_8-functionalized SPION NPs-labeled Ad5 capsids. Adapted with permission from ref. 498. Copyright 2012 Elsevier. C. Schematic illustration of streptavidin functionalized SPIONs chelated with biotin bound hexahistidine/AAV. Adapted with permission from ref. 115. Copyright 2011 Elsevier. D. Schematic representations of the main steps for single-cell virus stamping and colored SEM image of magnetic NP (blue) bound to LV (green) (scale bar: 500 nm, inset: 100 nm). Adapted with permission from ref. 501. Copyright 2018 Springer Nature. E. Tumor bio-reduction-activated NPs incorporating magnetized virus delivery. In the reductive tumor microenvironment the carrier payload dissociates and AAV conjugated SPIONs can be tracked by MRI. Adapted with permission from ref. 503. Copyright 2018 American Chemical Society.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

gated to iron oxide NPs were remotely guided and delivered to the tumor site with microscale precision, where phototoxicity was then triggered by irradiation.[489,502,503] Recently, Tseng *et al.* further developed this approach with an elaborate hybrid design. AAV encoding the photosensitive apoptosis-inducing protein was conjugated to SPIONs and loaded to a nitroimidazole conjugated hyaluronic acid NP containing lactate oxidase. This design employed the lactate-rich tumor microenviroment to subsequently initiate H_2O_2 production by lactate oxidase, resulting in reduction and disassembly of 2-nitroamidazole-hyaluronic acid matrix and finally release of SPIONs attached to AAV into tumor site (Fig. 19E). This enzymatically-controlled release enabled targeted tumor cell lysis with an combinatorial approach of functional materials.[503]

SPIONs are approved as contrast agents for magnetic resonance imaging,[504] which makes them a clinically promising vehicle for targeted delivery of VPs. One of main advantages of using magnetic NPs for VP delivery is the reduced vector dosage and short exposure times for efficient transduction due to controlled colocalization of cells and VPs. Magnetic NPs enable to deliver VPs to otherwise non-permissible cells and may enable the crossing of biophysical barriers *in vivo* in the future.

Small molecules

Small molecules are defined as organic compounds with a molecular mass below 900 Da and a physical size below 1 nm in pharmaceutical research. Today, the majority of drugs used as medicine in everyday life are small molecules, which are characterized by their high bioavailability and effective mode of action at the target site.[81] Research on small molecules that promote viral gene delivery has become more wide-spread in recent years.

The aim of adding small molecules during a transduction procedure is to achieve satisfactory TE with low vector doses to minimize side-effects and to enable transduction to cell types that are difficult to infect. The underlying mode of action of small molecules to achieve this aim is fundamentally different to previously discussed materials, which were interacting with VPs, forming a complex, aggregate or conjugate which in turn facilitates the interaction with cells. Here, small molecules interfere with multiple cellular processes, for example, by affecting gene transcription, cell entry mechanisms or DNA replication and thereby facilitating viral gene expression. The focus in small molecule research for viral gene delivery has been on dual mode drugs, which are promising due to their therapeutic as well as transduction enhancing properties. These molecules are usually simply added to VPs for application, even though covalent conjugation to VPs were also shown exemplarily.[506] In the following sections important findings are highlighted and summarized in Table 7 and Chart 3.

In early works, Miller *et al.* showed that DNA-damaging agents like UV-irradiation or cisplatin increased TE of AAV to non-dividing cells.[507] Later they introduced less cytotoxic DNA-synthesis inhibitors like aphidicolin or hydroxyurea and topoisomerase inhibitors such as etoposide or camptothecin as enhancers of AAV transduction to fibroblasts.[508] Groschel *et al.* demonstrated that these cytostatic agents etoposide, camptothecin, taxol, and aphidicolin, interfere with various mechanisms of cell cycle progression and arrest the cell cycle at G_2/M phase, which is a checkpoint to control DNA replication before nuclei division. Arresting the cell cycle in this phase boosted the early step of HIV virus replication and thereby enhanced TE.[135] Another well-known chemotherapeutic drug, doxorubicin, greatly enhanced rAAV infection to airway cells as well as neuronal cells with only mild cytotoxic effects at tested concentrations. Doxorubicin promoted rAAV accumulation at the nucleus, probably through proteasome inhibiting effects resulting in accelerated intracellular viral nuclear transport, but without affecting cell entry mechanisms.[509,510] The exact mechanism remained unclear and it was suggested that enhanced TE is more likely resulting from stimulated capsid uncoating and increased reverse transcription of viral genome than by increased cell entry.[135] Further cytokine stimulated downregulation of proteasome activity is believed to enhance TE.[511] Evidence for this hypothesis was found by addition of cytokines or proteasome inhibitors, particularly FDA approved bortezomib, resulting in enhanced TE of even low dosed LV to stem cells.[511–513]

Furthermore, the proteasome inhibitor MG132 (carbobenzoxy-L-leucyl-L-leucyl-L-leucinal) enhanced TE of rAAV to airway epithelia cells by modulating the ubiquitin–proteasome system in the endosomal environment.[514] Moreover, adenosine 5′-triphosphate-binding cassette transporter proteins, which efflux chemotherapeutic drugs and HIV protease inhibitors, play a significant role in reducing TE of LV into $CD34^+$ stem cells. Enhanced TE of up to sixfold was observed after the addition of Verapamil, an adenosine 5′-triphosphate-binding cassette-inhibitor.[515]

However, the identified chemotherapeutic compounds have huge disadvantages including damaging DNA and are therefore not ideal candidates for clinical use to healthy cells. In order to identify additional transduction enhancing drugs, high-throughput screens were conducted.

The first library screening for compounds promoting LV transduction identified 30 potential enhancers of viral transduction from 1280 investigated pharmaceutically active drugs. The non-DNA damaging, protein kinase activator phorbol 12-myristate 13-acetate (PMA) was found as a very promising candidate to enhance TE of LV into hard to transduce stem cells. Similar to previously discovered agents, PMA displayed interference with the cell cycle and inhibition of proliferation, suggesting that the transduction efficiency is influenced by processes within the cell rather than by promoted cell entry.[516]

Another high-throughput screening with focus on enhancing TE of rAAV to HeLa cells investigated 2396 pharmaceutically known, mostly FDA approved compounds from which 13 compounds were identified as capable potentiators. In this screening, five main enhancer groups were identified: epipodophyllotoxins, inducers of DNA damage, effectors of epigenetic modification, anthracyclines, and proteasome inhibitors.

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 7 Overview of drugs presented for viral delivery, properties and fabrication methods

Name	Medical use	Assumed mechanism of action	VPs and cell line	Ref.
Camptothecin	Cytostatic drug	Topoisomerase inhibitor/G_2/M cell cycle arrest	AAV to fibroblasts/HIV to HeLa	508 and 135
Taxol	Cytostatic drug	G_2/M cell cycle arrest	HIV to HeLa/AAV to HeLa	135 and 506
Hydroxyurea	Cytostatic drug	DNA-synthesis inhibitor	AAV to fibroblasts	508
Aphidicolin	Cytostatic drug	DNA-synthesis inhibitor	AAV to fibroblasts/HIV to HeLa	135 and 508
Verapamil	Calcium channel blocker	Adenosine 5′-triphosphate-binding cassette-inhibitor	Vesicular stomatitis virus-G to hematopoietic progenitor cells	515
Bortezomib	Cytostatic drug	Proteasome inhibitor	Lymphocytic choriomeningitis virus (LCMV) to T-cells/rAAV to HeLa, U87, HepG2, NHF1	513 and 517
MG132		Proteasome inhibitor	rAAV to human airway epithelia	514
Phorbol 12-myristate 13-acetate	Proteinkinase activator		LV to HSC	516
Doxorubicin	Cytostatic drug	Proteasome inhibitior	rAAV to neuronal cells/rAAV to human airway epithelia	510, 509 and 527
Teniposide	Cytostatic drug	Topoisomerase inhibitior	rAAV to HeLa, U87, HepG2, NHF1	517
Etoposide	Cytostatic drug	Topoisomerase inhibitor	rAAV to HeLa/AAV to fibroblasts/HIV to HeLa	517, 508 and 135
Bleomycin	Cytostatic drug, glycopeptide antibiotic	DNA damage	rAAV to HeLa, U87, HepG2, NHF1	517
Parthenin	Sesquiterpene lactone		rAAV to HeLa, U87, HepG2, NHF1	517
RH-1	Diaziridinylbenzoquinone		rAAV to HeLa, U87, HepG2, NHF1	517
Vorinostat	Cytostatic drug	HDAC inhibition	rAAV to HeLa, U87, HepG2, NHF1	517
Nanaomycin A	DNMT3B inhibitor	DNMT3B inhibitor	rAAV to HeLa, U87, HepG2, NHF1	517
Menogaril	Anthracycline		rAAV to HeLa, U87, HepG2, NHF1	517
Pyrromycin			rAAV to HeLa, U87, HepG2, NHF1	517
Daunorubicin		DNA intercalation, topoisomerase inhibition, polymerase inhibition, free radical damage to DNA	rAAV to HeLa, U87, HepG2, NHF1	517
Physalin B			rAAV to HeLa, U87, HepG2, NHF1	517
Siomycin	Thiazole antibiotic		rAAV to HeLa, U87, HepG2, NHF1	517
Tetrocarcin A	Microbial metabolite	BCL-2 inhibitor	rAAV to HeLa, U87, HepG2, NHF1	517
6-Gingerol	Phytoactive material, anticancer activity	Bypassing endocytic receptor mediated pathway, interfering with cell cycle	AAV2- to A375 and 526 malignant melanoma cells	58
Pyrrol based derivatives		Inhibition of production of IFNβ and various interferon-stimulated genes	Oncolytic Rhabodviruses to resistant cancer cells (*e.g.* CT26)	518
Prostaglandine E_2	Abortifacient agent, labor induction		LV to HSCs	520–522
(+)-JQ1	Cytostatic drug	Bromodomain inhibitor	Ad to A549, HeLa, Jurkat and P388D1 cells	524
Hydroxychloroquine	Antimalaria medication	TLR9 inhibition	AAV to murine and human tissue	523
Rosuvastatin	Statin medication	Promoter of low-density lipoprotein-receptor	Vesicular stomatitis virus (LV) to natural killer cells	526
Embelin	Cytostatic drug and anti-inflammatory effects	Inhibitor of X chromosome-linked inhibitor-of-apoptosis protein XIAP	Oncolytic vaccinia virus to T cells and NK cells	525

The highest level of transduction, with approx. 50–100 times compared with virus-only transduction, was observed for the chemotherapeutic epipodophyllotoxin teniposide. It was observed that the TE varied significantly for the tested substances depending on cell type, which indicates an underlying correlation between mode-of-action and cell type. In unison with previous studies, the authors suggest that rAAV can be enhanced by two pathways, whose detailed mechanisms remain unknown: the inhibition of proteasomes and the induction of DNA damage.[517]

A further screening for enhancing interferon sensitized oncolytic virus transduction to resistant cancer cells identified pyrrole-based molecules as potent enhancers. This was also feasible with AAV and Ad. High activities up to thousand-fold enhancement of viral transduction, tolerability and plasma stability were found for systemic administration in mice. The mode of action of these molecules was attributed to an inhibition of interferon-β production, which gives antiviral properties to resistant cancer cells.[518,519]

In recent years, further pharmaceutically applied drugs were identified as promising viral gene delivery agents.

Prostaglandine E_2 enhanced LV transduction to HSCs by acting during the endocytosis phase *via* a yet unclear mechanism.[520–522]

Lee *et al.* reported synergistic effects in terms of inducing apoptosis to malignant melanoma cells when AAV2 encoding a pro-apoptotic protein and the phytoactive compound 6-gingerol were combined. 6-Gingerol was previously reported to be

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

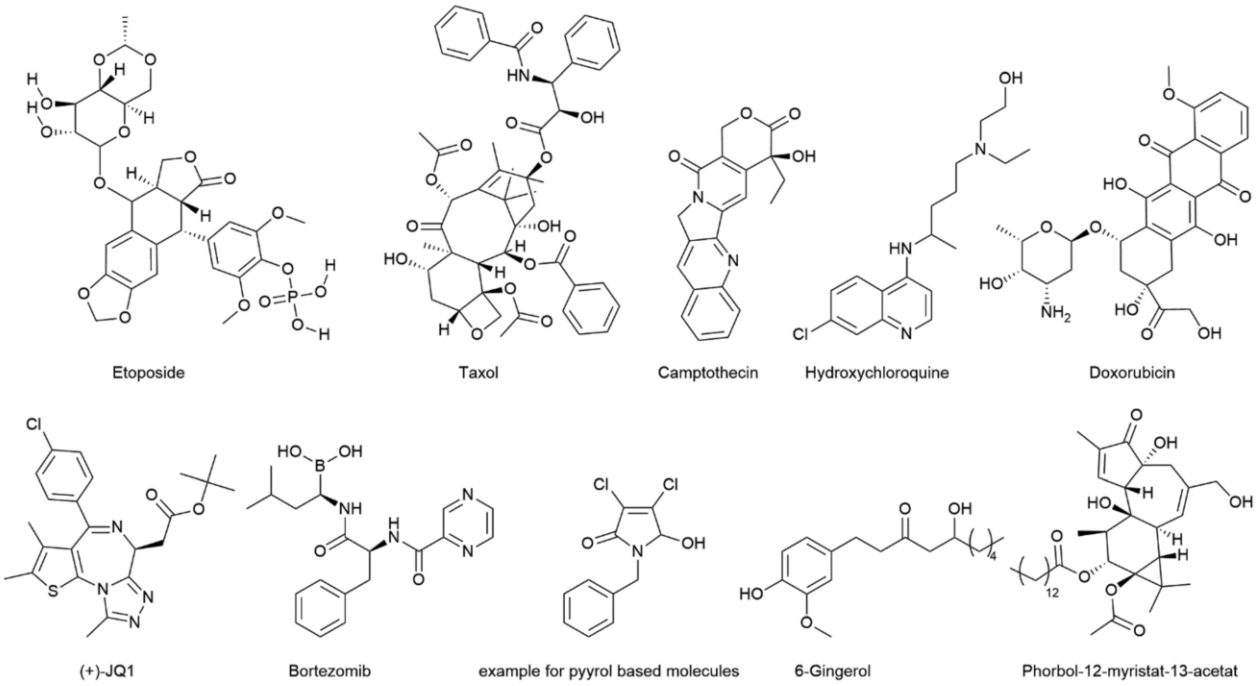


Chart 3 Overview of selected drugs as enhancers of viral gene transfer.

effective against cancer and for the inhibition of angiogenesis. The cytostatic activity has its source in the control of the cell cycle and subsequently inducing apoptosis. The authors proposed that AAV2 can bypass its usual endocytic receptor-mediated pathway in the presence of 6-gingerol and thereby enhancing apoptosis in melanoma cells more efficiently.[58]

Furthermore, hydroxychloroquine, an approved antimalaria drug, has been applied for promoting AAV transduction in murine and human tissue *in vitro* and *in vivo*, *e.g.* in the retina with no adverse effects. Among the multimodal effects this drug might have on TE, one mode of action was proposed as an inhibition of TLR9 gene expression, which is a known sensor for anti-viral response of cells.[523]

Intervention in cell regulation processes was similarly shown for Ad infection which was enantioselectively promoted by the bromodomain inhibitor (+)-JQ1. This molecule blocks acetylated lysine interaction with the bromodomain family proteins, which play a significant role in cell regulation and gene transcription. In this way, Ad infectivity and gene expression were elevated by suppression of regulating bromodomains, whose mechanisms still need to be elucidated.[524]

Embelin, used as an adjuvant in oncolytic virotherapy was recently reported by Wang *et al.* Lymphoma cell lysis was induced by mitigating antiviral immunity against oncolytic virus upon addition of Embelin. The authors proposed that the infiltration of the oncolytic virus was facilitated by the disruption of Interleukin-6/STAT3 signalling and further promoted by the inherent ability of Embelin to induce apoptotic death to tumor cells.[525]

An interesting approach was reported by Gong *et al.* who enhanced TE by increasing low-density lipoprotein-receptors expression with Rosuvastatin. Vesicular stomatitis virus was able to efficiently target and transduce in natural killer cells after addition of Rosuvastatin. Naturally, untreated killer cells lack of sufficient low-density lipoprotein-receptors, which are necessary for virus interaction. Statins, especially Rosuvastatin, can induce low-density lipoprotein expression and thereby facilitate virus–cell interaction. Additionally, geranylgeranyl-pyrophosphate suppresses cytotoxic effects of statins, resulting in a highly effective and powerful tool to enable genetic modification of natural killer cells. Genetically modified natural killer cells can, for example, express tumor antigen-specific receptors, which may pave the way for targeted treatment of malignant cells.[526]

Small molecules, which are beneficial for disease treatment, eliciting host immune response while enabling efficient transduction are actively being pursued for clinical application. Pharmaceutically active drugs that have been found so far, are paving the way especially for oncolytic gene therapy and can be included easily in *ex vivo* gene transfer protocols. However, further research is indispensable for investigation of unwanted side effects of drugs combined with viruses during delivery and should be carefully evaluated in therapeutic application.

From development to usage

Most of the introduced materials for promoting viral gene transfer have not yet been translated to the clinic and are often not commercially available. Here, we will highlight materials that are emerging from the development stage to actual clini-

 This journal is © The Royal Society of Chemistry 2020

View Article Online
Biomaterials Science Review

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

cal usage. Various material classes which have the potential for application in medical formulations are currently investigated in preclinical studies. For example, *in vivo* studies in animals utilized materials like PLGA,[133,155,215,216] PEG,[163,181,283] PEI,[208] poloxamers,[160,229,231,232,238,239] dendrimers,[116,251,253] polysaccharides,[258,259,262,265] gelatin,[385,386] collagen,[126] fibrin,[374] human serum albumin,[375,377] calcium carbonate,[454] gold nanorods,[454] silica implants,[128] graphene,[481] magnetic NPs,[491,501,505] and small molecules.[517,518,523]

Commercially available kits make transduction enhancers more broadly available and are thus attractive tools for use in research. For instance, several studies about CAR T-cell manufacturing employed kits of the peptidic enhancers Vectofusin®-1 (Miltenyi Biotec) and Protransduzin™ (JPT Peptide Technologies).[353,354,528,529] However, popular materials in research cannot seamlessly be transferred to the clinical context. Low cell viability in high concentrations limits the application of Polybrene® (Abbott Laboratories Corp.) in drug formulations.[530] More efficient enhancers than polybrene[531] are traded as jetPEI® (Polyplus transfection) and Lipofectamine™ (Thermo Fisher Scientific) and can be applied to a broad range of VVs. Kits for specific viruses are available, but their exact composition is a trade secret, *e.g.*: for LVs (ViraDuctin™ (Cell Biolabs), TransDux™ (System Biosciences), ViralPlus (Applied Biological Materials Inc.), SureENTRY (Qiagen)), as well as for Ad in the form of AdenoBoost™ (Sirion Biotech). High transduction rates with low-titer viruses are enabled with polycationic magnetic beads, commercialized as MACSductin™ (Miltenyi Biotec), ViroMag and AdenoMag (OZ Biosciences), Lenti-X™ Accelerator (Takara Bio) and MISSION® ExpressMag® (Sigma-Aldrich). In recent years, a growing number of such kits were made available in GMP grade addressing the demand for systematic investigation of transduction agents in a clinical context.

Clinical trials use RetroNectin® (Takara Bio) in particular as the gold standard for *ex vivo* transduction of lenti- and retroviruses, even though a high multiplicity of infection and a time consuming coating protocol is needed.[361,532–534] Poloxamers such as LentiBoost™ (Sirion Biotech)[535] are commonly used in formulations for *in vivo* clinical trials as enhancers and to prevent uncontrolled binding of VVs to injection devices.[536–539] While the majority of viral gene therapy drugs approved up until 2017 do not use additional transduction enhancers in their formulations, those approved since then have made use of agents such as protamine sulfate (*ex vivo*, Zynteglo),[540,541] Poloxamer 188 (*in vivo*, Zolgensma and Luxturna),[542,543] and human serum albumin (*ex vivo*, Yescarta, Kymriah).[544,545]

Conclusions and outlook

Viral gene delivery has come a long way and is now becoming increasingly therapeutically relevant. The rising number of approvals for clinical use in recent years holds a great promise for treatment of currently uncurable diseases. However, the full potential of gene therapy cannot be realized until safe and efficient delivery systems are available. In this review, we highlighted the most important approaches for viral gene delivery from a materials perspective. In particular, the high biocompatibility of peptidic delivery systems, and the superior shielding effects of polymer coatings and lipid encapsulation need to be emphasized. Both hydrogels and magnetic NPs allow spatially defined virus release, often in a time-controlled manner. Especially the local targeted delivery of VVs *via* magnetic nanoparticles to tumor sites has enormous potential in oncolytic virotherapy. The performance of delivery systems is determined by many experimental parameters, *e.g.* virus type, cell type, concentration, and incubation time. However, detailed large-scale comparative studies of different delivery agents are still missing. Therefore, it is often difficult or even impossible to draw conclusions on the relative performance of the delivery agents. In current viral gene delivery systems, immunogenicity, native tropism, unspecific delivery, and poor efficiency at low vector doses are remaining hurdles to be overcome. The materials introduced here have the potential to establish hybrid VVs for clinical applications. However, clinical data of administration to humans remains scarce.

Gene delivery as a therapeutic approach is one of the major medicinal and biotechnological advances of this century. We expect that novel bioengineered viruses promoted by synthetic materials are destined for future efficient and safe clinical applications.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank Nicole Kirsch-Pietz for proof-reading the manuscript. Open Access funding provided by the Max Planck Society.

Notes and references

1 C. E. Thomas, A. Ehrhardt and M. A. Kay, *Nat. Rev. Genet.*, 2003, **4**, 346–358.
2 S. Pearson, H. Jia and K. Kandachi, *Nat. Biotechnol.*, 2004, **22**, 3–4.
3 Gene Therapy Clinical Trials Worldwide, http://www.abedia.com/wiley/years.php, (accessed 15 April 2020).
4 R. Goswami, G. Subramanian, L. Silayeva, I. Newkirk, D. Doctor, K. Chawla, S. Chattopadhyay, D. Chandra, N. Chilukuri and V. Betapudi, *Front. Oncol.*, 2019, **9**, 297.
5 Grand View Research, *Gene Therapy Market Size, Share & Trends Analysis Report By Indication (Acute Lymphoblastic Leukemia, Large B-cell Lymphoma), By Vector Type, By Region, And Segment Forecasts, 2020–2027*, 2020.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

6 K. A. High and M. G. Roncarolo, *N. Engl. J. Med.*, 2019, **381**, 455–464.
7 C.-C. Ma, Z.-L. Wang, T. Xu, Z.-Y. He and Y.-Q. Wei, *Biotechnol. Adv.*, 2020, **40**, 107502.
8 FDA approves novel gene therapy to treat patients with a rare form of inherited vision loss, https://www.fda.gov/news-events/press-announcements/fda-approves-novel-gene-therapy-treat-patients-rare-form-inherited-vision-loss, (accessed 15 April 2020).
9 FDA approves innovative gene therapy to treat pediatric patients with spinal muscular atrophy, a rare disease and leading genetic cause of infant mortality, https://www.fda.gov/news-events/press-announcements/fda-approves-innovative-gene-therapy-treat-pediatric-patients-spinal-muscular-atrophy-rare-disease, (accessed 15 April 2020).
10 M. Zheng, J. Huang, A. Tong and H. Yang, *Mol. Ther.–Oncolytics*, 2019, **15**, 234–247.
11 C.-H. Lu, Y.-H. Chang, S.-Y. Lin, K.-C. Li and Y.-C. Hu, *Biotechnol. Adv.*, 2013, **31**(8), 1695–1706.
12 A. M. Thomas and L. D. Shea, *J. Controlled Release*, 2013, **170**, 421–429.
13 P. R. Gupta and R. M. Huckfeldt, *J. Neural Eng.*, 2017, **14**, 051002.
14 M. G. Katz, A. S. Fargnoli, R. D. Williams and C. R. Bridges, *Hum. Gene Ther.*, 2013, **24**, 914–927.
15 S. Rangarajan, L. Walsh, W. Lester, D. Perry, B. Madan, M. Laffan, H. Yu, C. Vettermann, G. F. Pierce, W. Y. Wong and K. J. Pasi, *N. Engl. J. Med.*, 2017, **377**, 2519–2530.
16 L. Samaranch, A. Pérez-Cañamás, B. Soto-Huelin, V. Sudhakar, J. Jurado-Arjona, P. Hadaczek, J. Ávila, J. R. Bringas, J. Casas, H. Chen, X. He, E. H. Schuchman, S. H. Cheng, J. Forsayeth, K. S. Bankiewicz and M. D. Ledesma, *Sci. Transl. Med.*, 2019, **11**, eaat3738.
17 U. Griesenbach, K. M. Pytel and E. W. F. W. Alton, *Hum. Gene Ther.*, 2015, **26**, 266–275.
18 E. Marshall, *Science*, 1999, **286**, 2244–2245.
19 N. A. Helal, A. Osami, A. Helmy, T. McDonald, L. A. Shaaban and M. I. Nounou, *Pharmazie*, 2017, **72**, 627–651.
20 M. Ogris and E. Wagner, *Hum. Gene Ther.*, 2011, **22**, 799–807.
21 H. Kamiya, H. Tsuchiya, J. Yamazaki and H. Harashima, *Adv. Drug Delivery Rev.*, 2001, **52**, 153–164.
22 K. Lundstrom, *Diseases*, 2018, **6**, 42.
23 C. A. Suttle, *Nat. Rev. Microbiol.*, 2007, **5**, 801–812.
24 T. Watanabe and Y. Kawaoka, *Clin. Transl. Immunol.*, 2020, **9**, e1114.
25 T. Dull, R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono and L. Naldini, *J. Virol.*, 1998, **72**, 8463–8471.
26 M. A. Kay, J. C. Glorioso and L. Naldini, *Nat. Med.*, 2001, **7**, 33–40.
27 W. Lucas, in *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd, Chichester, UK, 2010.
28 D. Bouard, N. Alazard-Dany and F.-L. Cosset, *Br. J. Pharmacol.*, 2009, **157**, 153–165.
29 R. G. Crystal, *Hum. Gene Ther.*, 2014, **25**, 3–11.
30 T. J. Wickham, P. Mathias, D. A. Cheresh and G. R. Nemerow, *Cell*, 1993, **73**, 309–319.
31 M. C. Dechecchi, A. Tamanini, A. Bonizzato and G. Cabrini, *Virology*, 2000, **268**, 382–390.
32 V. N. Krasnykh, J. T. Douglas and V. W. van Beusechem, *Mol. Ther.*, 2000, **1**, 391–405.
33 N. C. Di Paolo, N. van Rooijen and D. M. Shayakhmetov, *Mol. Ther.*, 2009, **17**, 675–684.
34 N. F. Nidetz, M. C. McGee, L. V. Tse, C. Li, L. Cong, Y. Li and W. Huang, *Pharmacol. Ther.*, 2020, **207**, 107453.
35 M. F. Naso, B. Tomkowicz, W. L. Perry and W. R. Strohl, *BioDrugs*, 2017, **31**, 317–334.
36 D. Escors and K. Breckpot, *Arch. Immunol. Ther. Exp.*, 2010, **58**, 107–119.
37 A. V. Joglekar and S. Sandoval, *Hum. Gene Ther: Methods.*, 2017, **28**, 291–301.
38 G. B. Li and G. X. Lu, *Mol. Biotechnol.*, 2009, **43**, 250–256.
39 O. W. Merten, S. Charrier, N. Laroudie, S. Fauchille, C. Dugué, C. Jenny, M. Audit, M. A. Zanta-Boussif, H. Chautard, M. Radrizzani, G. Vallanti, L. Naldini, P. Noguiez-Hellin and A. Galy, *Hum. Gene Ther.*, 2011, **22**, 343–356.
40 S. K. Totsch, C. Schlappi, K.-D. Kang, A. S. Ishizuka, G. M. Lynn, B. Fox, E. A. Beierle, R. J. Whitley, J. M. Markert, G. Y. Gillespie, J. D. Bernstock and G. K. Friedman, *Oncogene*, 2019, **38**, 6159–6171.
41 D. M. Patel, P. M. Foreman, L. B. Nabors, K. O. Riley, G. Y. Gillespie and J. M. Markert, *Hum. Gene Ther.: Clin. Dev.*, 2016, **27**, 69–78.
42 Q. Lan, S. Xia, Q. Wang, W. Xu, H. Huang, S. Jiang and L. Lu, *Front. Med.*, 2020, **14**, 160–184.
43 Y.-N. Zhang, S.-B. Wang, S.-S. Song, P.-Y. Hu, Y.-C. Zhou, Y.-P. Mou and X.-Z. Mou, *Biotechnol. Lett.*, 2020, **42**, 865–874.
44 J. W. Choi, Y. S. Lee, C. O. Yun and S. W. Kim, *J. Controlled Release*, 2015, **219**, 181–191.
45 N. Martínez-Vélez, M. Garcia-Moure, M. Marigil, M. González-Huarriz, M. Puigdelloses, J. Gallego Pérez-Larraya, M. Zalacaín, L. Marrodán, M. Varela-Guruceaga, V. Laspidea, J. J. Aristu, L. I. Ramos, S. Tejada-Solís, R. Díez-Valle, C. Jones, A. Mackay, J. A. Martínez-Climent, M. J. García-Barchino, E. Raabe, M. Monje, O. J. Becher, M. P. Junier, E. A. El-Habr, H. Chneiweiss, G. Aldave, H. Jiang, J. Fueyo, A. Patiño-García, C. Gomez-Manzano and M. M. Alonso, *Nat. Commun.*, 2019, **10**, 1–10.
46 T. Shi, X. Song, Y. Wang, F. Liu and J. Wei, *Front. Immunol.*, 2020, **11**, 683.
47 C. S. Lee, E. S. Bishop, R. Zhang, X. Yu, E. M. Farina, S. Yan, C. Zhao, Z. Zeng, Y. Shu, X. Wu, J. Lei, Y. Li, W. Zhang, C. Yang, K. Wu, Y. Wu, S. Ho, A. Athiviraham, M. J. Lee, J. M. Wolf, R. R. Reid and T.-C. He, *Genes Dis.*, 2017, **4**, 43–63.
48 K. P. Kicielinski, E. A. Chiocca, J. S. Yu, G. M. Gill, M. Coffey and J. M. Markert, *Mol. Ther.*, 2014, **22**, 1056–1062.
49 E. Galanis, S. N. Markovic, V. J. Suman, G. J. Nuovo, R. G. Vile, T. J. Kottke, W. K. Nevala, M. A. Thompson,

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

J. E. Lewis, K. M. Rumilla, V. Roulstone, K. Harrington, G. P. Linette, W. J. Maples, M. Coffey, J. Zwiebel and K. Kendra, *Mol. Ther.*, 2012, **20**, 1998–2003.
50 S. Aref, K. Bailey and A. Fielding, *Viruses*, 2016, **8**, 294.
51 K. Harrington, D. J. Freeman, B. Kelly, J. Harper and J.-C. Soria, *Nat. Rev. Drug Discovery*, 2019, **18**, 689–706.
52 A. Ricobaraza, M. Gonzalez-Aparicio, L. Mora-Jimenez, S. Lumbreras and R. Hernandez-Alcoceba, *Int. J. Mol. Sci.*, 2020, **21**, 3643.
53 R. Singh, B. Tian and K. Kostarelos, *FASEB J.*, 2008, **22**, 3389–3402.
54 J. F. Wright, T. Le, J. Prado, J. Bahr-Davidson, P. H. Smith, Z. Zhen, J. M. Sommer, G. F. Pierce and G. Qu, *Mol. Ther.*, 2005, **12**, 171–178.
55 J. L. Madrigal, R. Stilhano and E. A. Silva, *Tissue Eng., Part B*, 2017, **23**, 347–361.
56 B. Balakrishnan and G. Jayandharan, *Curr. Gene Ther.*, 2014, **14**, 86–100.
57 M. A. Croyle, S. M. Callahan, A. Auricchio, G. Schumer, K. D. Linse, J. M. Wilson, L. J. Brunner and G. P. Kobinger, *J. Virol.*, 2004, **78**, 912–921.
58 J. H. Lee, Y. Kim, Y.-E. Yoon, Y.-J. Kim, S.-G. Oh, J.-H. Jang and E. Kim, *New Biotechnol.*, 2017, **37**, 194–199.
59 O. S. Kumru, Y. Wang, C. W. R. Gombotz, B. Kelley-Clarke, W. Cieplak, T. Kim, S. B. Joshi and D. B. Volkin, *J. Pharm. Sci.*, 2018, **107**, 2764–2774.
60 Y. Zhang, N. Chirmule, G. P. Gao, R. Qian, M. Croyle, B. Joshi, J. Tazelaar and J. M. Wilson, *Mol. Ther.*, 2001, **3**, 697–707.
61 D. Ghosh and M. A. Barry, *J. Virol.*, 2005, **79**, 13667–13672.
62 F. H. E. Schagen, M. Ossevoort, R. E. M. Toes and R. C. Hoeben, *Crit. Rev. Oncol. Hematol.*, 2004, **50**, 51–70.
63 K. Kawabata, F. Sakurai, N. Koizumi, T. Hayakawa and H. Mizuguchi, *Mol. Pharm.*, 2006, **3**, 95–103.
64 A. S. Chuck, M. F. Clarke and B. O. Palsson, *Hum. Gene Ther.*, 1996, **7**, 1527–1534.
65 J. S. Bartlett, R. Wilcher and R. J. Samulski, *J. Virol.*, 2000, **74**, 2777–2785.
66 J. E. Ziello, Y. Huang and I. S. Jovin, *Mol. Med.*, 2010, **16**, 222–229.
67 D. L. Puhl, A. R. D'Amato and R. J. Gilbert, *Brain Res. Bull.*, 2019, **150**, 216–230.
68 C. K. Yun, J. W. Hwang, T. J. Kwak, W. J. Chang, S. Ha, K. Han, S. Lee and Y. S. Choi, *Lab Chip*, 2019, **19**, 580–588.
69 D. A. Amado, J. M. Rieders, F. Diatta, P. Hernandez-Con, A. Singer, J. T. Mak, J. Zhang, E. Lancaster, B. L. Davidson and A. S. Chen-Plotkin, *Mol. Ther.*, 2019, **27**, 465–478.
70 N. Moore, J. R. Chevillet, L. J. Healey, C. McBrine, D. Doty, J. Santos, B. Teece, J. Truslow, V. Mott, P. Hsi, V. Tandon, J. T. Borenstein, J. Balestrini and K. Kotz, *Sci. Rep.*, 2019, **9**, 1–11.
71 Y. Deng, M. Kizer, M. Rada, J. Sage, X. Wang, D. J. Cheon and A. J. Chung, *Nano Lett.*, 2018, **18**, 2705–2710.
72 C. H. Peng, L. C. Woung, K. H. Lu, C. Y. Tsai, S. D. Lee, C. S. Huang, T. C. Lin, K. H. Chien and D. K. Hwang, *J. Chin. Med. Assoc.*, 2018, **81**, 830–836.
73 L. Qian, B. Thapa, J. Hong, Y. Zhang, M. Zhu, M. Chu, J. Yao and D. Xu, *J. Thorac. Dis.*, 2018, **10**, 1099–1111.
74 A. B. Bahnsonz, J. T. Dunigan, B. E. Baysal, T. Mohney, R. Wayne Atchison, M. T. Nimgaonkar, E. D. Ball and J. A. Barranger, *J. Virol. Methods*, 1995, **54**, 131–143.
75 Y. Chen, S. Aslanoglou, G. Gervinskas, H. Abdelmaksoud, N. H. Voelcker and R. Elnathan, *Small*, 2019, **15**, 1904819.
76 S. Y. Chang, Y. H. Park, N. T. Carpena, T. T. Pham, P. S. Chung, J. Y. Jung and M. Y. Lee, *Lasers Med. Sci.*, 2019, **34**, 367–375.
77 M. Tsukakoshi, S. Kurata, Y. Nomiya, Y. Ikawa and T. Kasuya, *Appl. Phys., B*, 1984, **35**, 135–140.
78 E. Neumann, M. Schaefer-Ridder, Y. Wang and P. H. Hofschneider, *EMBO J.*, 1982, **1**, 841–845.
79 H. Song, R. A. Bush, Y. Zeng, H. Qian, Z. Wu and P. A. Sieving, *Mol. Ther.–Methods Clin. Dev.*, 2019, **13**, 77–85.
80 A. L. Coulberson, N. V. Hud, J. M. LeDoux, I. D. Vilfan and M. R. Prausnitz, *J. Controlled Release*, 2003, **86**, 361–370.
81 M. P. Stewart, R. Langer and K. F. Jensen, *Chem. Rev.*, 2018, **118**, 7409–7531.
82 X. Du, J. Wang, Q. Zhou, L. Zhang, S. Wang, Z. Zhang and C. Yao, *Drug Delivery*, 2018, **25**, 1516–1525.
83 F. Kreppel, J. Gackowski, E. Schmidt and S. Kochanek, *Mol. Ther.*, 2005, **12**, 107–117.
84 S. J. White, S. A. Nicklin, H. Büning, M. J. Brosnan, K. Leike, E. D. Papadakis, M. Hallek and A. H. Baker, *Circulation*, 2004, **109**, 513–519.
85 J. S. Chandran, P. S. Sharp, E. Karyka, J. M. D. C. Aves-Cruzeiro, I. Coldicott, L. Castelli, G. Hautbergue, M. O. Collins and M. Azzouz, *Sci. Rep.*, 2017, **7**, 14766.
86 E. J. Lee, C. M. Guenther and J. Suh, *Curr. Opin. Biomed. Eng.*, 2018, **7**, 58–63.
87 H. Büning and A. Srivastava, *Mol. Ther.–Methods Clin. Dev.*, 2019, **12**, 248–265.
88 R. E. Kelemen, S. B. Erickson and A. Chatterjee, in *Methods in Molecular Biology*, Humana Press Inc., New York, NY, 2018, vol. 1728, pp. 313–326.
89 R. Schubert, S. Herzog, S. Trenholm, B. Roska and D. J. Müller, *Nat. Protoc.*, 2019, **14**, 3205–3219.
90 R. E. Kelemen, R. Mukherjee, X. Cao, S. B. Erickson, Y. Zheng and A. Chatterjee, *Angew. Chem., Int. Ed.*, 2016, **55**, 10645–10649.
91 N. Gabriel, S. Hareendran, D. Sen, R. A. Gadkari, G. Sudha, R. Selot, M. Hussain, R. Dhaksnamoorthy, R. Samuel, N. Srinivasan, A. Srivastava and G. R. Jayandharan, *Hum. Gene Ther: Methods.*, 2013, **24**, 80–93.
92 J. A. Sullivan, L. M. Stanek, M. J. Lukason, J. Bu, S. R. Osmond, E. A. Barry, C. R. O'Riordan, L. S. Shihabuddin, S. H. Cheng and A. Scaria, *Gene Ther.*, 2018, **25**, 205–219.
93 F. Sakurai, H. Mizuguchi and T. Hayakawa, *Gene Ther.*, 2003, **10**, 1041–1048.
94 M. Nakatake, H. Kurosaki, N. Kuwano, K. Horita, M. Ito, H. Kono, T. Okamura, K. Hasegawa, Y. Yasutomi and T. Nakamura, *Mol. Ther.–Oncolytics*, 2019, **14**, 159–171.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

95 Z. Chai, J. Sun, K. M. Rigsbee, M. Wang, R. J. Samulski and C. Li, *J. Controlled Release*, 2017, **262**, 348–356.
96 M. J. Brun, E. J. Gomez and J. Suh, *J. Controlled Release*, 2017, **267**, 80–89.
97 Y. Wang, S. Li, Z. Tian, J. Sun, S. Liang, B. Zhang, L. Bai, Y. Zhang, X. Zhou, S. Xiao, Q. Zhang, L. Zhang, C. Zhang and D. Zhou, *Nucleic Acids Res.*, 2019, **47**, e114.
98 E. J. Gomez, K. Gerhardt, J. Judd, J. J. Tabor and J. Suh, *ACS Nano*, 2016, **10**, 225–237.
99 N. N. Thadani, J. Yang, B. Moyo, C. M. Lee, M. Y. Chen, G. Bao and J. Suh, *ACS Synth. Biol.*, 2020, **9**, 461–467.
100 R. C. Münch, A. Muth, A. Muik, T. Friedel, J. Schmatz, B. Dreier, A. Trkola, A. Plückthun, H. Büning and C. J. Buchholz, *Nat. Commun.*, 2015, **6**, 6246.
101 Y. Lv, F. J. Xiao, Y. Wang, X. H. Zou, H. Wang, H. Y. Wang, L. S. Wang and Z. Z. Lu, *BMC Biotechnol.*, 2019, **19**, 23.
102 C. Li and R. J. Samulski, *Nat. Rev. Genet.*, 2020, **21**, 255–272.
103 C. Domenger and D. Grimm, *Hum. Mol. Genet.*, 2019, **28**, 1–3.
104 A. M. Wen and N. F. Steinmetz, *Chem. Soc. Rev.*, 2016, **45**, 4074–4126.
105 P. S. Banerjee, P. Ostapchuk, P. Hearing and I. Carrico, *J. Am. Chem. Soc.*, 2010, **132**, 13615–13617.
106 C. R. O'Riordan, A. Lachapelle, C. Delgado, V. Parkes, S. C. Wadsworth, A. E. Smith and G. E. Francis, *Hum. Gene Ther.*, 1999, **10**, 1349–1358.
107 Z. Chen, N. Li, S. Li, M. Dharmarwardana, A. Schlimme and J. J. Gassensmith, *WIREs Nanomed. Nanobiotechnol.*, 2016, **8**, 512–534.
108 F. Kreppel and S. Kochanek, *Mol. Ther.*, 2008, **16**, 16–29.
109 M. A. Bruckman, G. Kaur, L. A. Lee, F. Xie, J. Sepulveda, R. Breitenkamp, X. Zhang, M. Joralemon, T. P. Russell, T. Emrick and Q. Wang, *ChemBioChem*, 2008, **9**, 519–523.
110 Y. Chu, Y. H. Oum and I. S. Carrico, *Virology*, 2016, **487**, 95–103.
111 I. Yildiz, I. Tsvetkova, A. M. Wen, S. Shukla, M. H. Masarapu, B. Dragnea and N. F. Steinmetz, *RSC Adv.*, 2012, **2**, 3670–3677.
112 X. Wang, Y. Deng, S. Li, G. Wang, E. Qin, X. Xu, R. Tang and C. Qin, *Adv. Healthcare Mater.*, 2012, **1**, 443–449.
113 H. Zhou, G. Wang, X. Wang, Z. Song and R. Tang, *Angew. Chem., Int. Ed.*, 2017, **56**, 12908–12912.
114 P. H. Kim, T. il Kim, J. W. Yockman, S. W. Kim and C. O. Yun, *Biomaterials*, 2010, **31**, 1865–1874.
115 E. Kim, J. S. Oh, I. S. Ahn, K. I. Park and J. H. Jang, *Biomaterials*, 2011, **32**, 8654–8662.
116 Y. Wu, L. Li, L. Frank, J. Wagner, P. Andreozzi, B. Hammer, M. D'Alicarnasso, M. Pelliccia, W. Liu, S. Chakrabortty, S. Krol, J. Simon, K. Landfester, S. L. Kuan, F. Stellacci, K. Müllen, F. Kreppel and T. Weil, *ACS Nano*, 2019, **13**, 8749–8759.
117 S. Kirti, K. Patel, S. Das, P. Shrimali, S. Samanta, R. Kumar, D. Chatterjee, D. Ghosh, A. Kumar, P. Tayalia and S. K. Maji, *ACS Biomater. Sci. Eng.*, 2019, **5**, 126–138.
118 M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständker, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff and J. Münch, *Nat. Nanotechnol.*, 2013, **8**, 130–136.
119 J. I. Youn, S. H. Park, H. T. Jin, C. G. Lee, S. H. Seo, M. Y. Song, C. W. Lee and Y. C. Sung, *Cancer Gene Ther.*, 2008, **15**, 703–712.
120 J.-H. Hwang, S. Lee, E. Kim, J.-S. Kim, C.-H. Lee, I.-S. Ahn and J.-H. Jang, *Int. J. Pharm.*, 2011, **421**, 397–404.
121 D. Xia, L. B. Feng, X. L. Wu, G. D. Xia and L. Xu, *Mol. Med. Rep.*, 2015, **12**, 2336–2342.
122 P. Díaz-Rodríguez, A. Rey-Rico, H. Madry, M. Landin and M. Cucchiarini, *Int. J. Pharm.*, 2015, **496**, 614–626.
123 Z. Zhong, S. Shi, J. Han, Z. Zhang and X. Sun, *Mol. Pharm.*, 2010, **7**, 105–115.
124 K.-H. Park, C.-O. Yun, O.-J. Kwon, C.-H. Kim, J.-R. Kim and K.-H. Cho, *Hum. Gene Ther.*, 2010, **21**, 579–587.
125 A. A. Sapre, G. Yong, Y. san Yeh, L. E. Ruff, J. S. Plaut, Z. Sayar, A. Agarwal, J. Martinez, T. N. Nguyen, Y. T. Liu, B. T. Messmer, S. C. Esener and J. M. Fischer, *J. Controlled Release*, 2019, **297**, 48–59.
126 S. Shin and L. D. Shea, *Mol. Ther.*, 2010, **18**, 700–706.
127 S. S. McMahon, N. Nikolskaya, S. N. Choileáin, N. Hennessy, T. O'Brien, P. M. Strappe, A. Gorelov and Y. Rochev, *J. Gene Med.*, 2011, **13**, 591–601.
128 L. Kangasniemi, M. Koskinen, M. Jokinen, M. Toriseva, R. Ala-Aho, V. M. Kähäri, H. Jalonen, S. Ylä-Herttuala, H. Moilanen, U. H. Stenman, I. Diaconu, A. Kanerva, S. Pesonen, T. Hakkarainen and A. Hemminki, *Gene Ther.*, 2009, **16**, 103–110.
129 A. Elouahabi and J. M. Ruysschaert, *Mol. Ther.*, 2005, **11**, 336–347.
130 D. Fenard, D. Ingrao, A. Seye, J. Buisset, S. Genries, S. Martin, A. Kichler and A. Galy, *Mol. Ther.–Nucleic Acids*, 2013, **2**, e90.
131 N. Landázuri and J. M. Le Doux, *J. Gene Med.*, 2004, **6**, 1304–1319.
132 S. Vupputuri, S. Karode, B. J. Neely and J. D. Ramsey, *J. Virol. Methods*, 2013, **192**, 1–11.
133 S. Shin, D. M. Salvay and L. D. Shea, *J. Biomed. Mater. Res., Part A*, 2010, **93**, 1252–1259.
134 C. A. Gersbach, S. R. Coyer, J. M. Le Doux and A. J. García, *Biomaterials*, 2007, **28**, 5121–5127.
135 B. Groschel and F. Bushman, *J. Virol.*, 2005, **79**, 5695–5704.
136 J. T. Douglas, M. Kim, L. A. Sumerel, D. E. Carey and D. T. Curiel, *Cancer Res.*, 2001, **61**, 813–817.
137 S. Kurachi, K. Tashiro, F. Sakurai, H. Sakurai, K. Kawabata, K. Yayama, H. Okamoto, S. Nakagawa and H. Mizuguchi, *Gene Ther.*, 2007, **14**, 1160–1165.
138 G. Fan, M. Fan, Q. Wang, J. Jiang, Y. Wan, T. Gong, Z. Zhang and X. Sun, *Acta Biomater.*, 2016, **30**, 94–105.
139 K. Singarapu, I. Pal and J. D. Ramsey, *J. Biomed. Mater. Res., Part A*, 2013, **101**, 1857–1864.

 This journal is © The Royal Society of Chemistry 2020

View Article Online

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

140 A. Vetter, K. S. Virdi, S. Espenlaub, W. Rödl, E. Wagner, P. S. Holm, C. Scheu, F. Kreppel, C. Spitzweg and M. Ogris, *Mol. Pharm.*, 2013, **10**, 606–618.

141 S. M. Moghimi, P. Symonds, J. C. Murray, A. C. Hunter, G. Debska and A. Szewczyk, *Mol. Ther.*, 2005, **11**, 990–995.

142 J. Raymond, A. Metcalfe, I. Salazkin and A. Schwarz, *Biomaterials*, 2004, **25**, 3983–3989.

143 J. S. Pagano and A. Vaheri, *Arch. Gesamte Virusforsch.*, 1965, **17**, 456–464.

144 P. K. Vogt, *Virology*, 1967, **33**, 175–177.

145 J. H. McCutchan and J. S. Pagano, *J. Natl. Cancer Inst.*, 1968, **41**, 351–357.

146 J. S. Manning, A. J. Hackett and N. B. Darby, *Appl. Microbiol.*, 1971, **22**, 1162–1163.

147 H. E. Davis, J. R. Morgan and M. L. Yarmush, *Biophys. Chem.*, 2002, **97**, 159–172.

148 C. Zhao, N. Wu, F. Deng, H. Zhang, N. Wang, W. Zhang, X. Chen, S. Wen, J. Zhang, L. Yin, Z. Liao, Z. Zhang, Q. Zhang, Z. Yan, W. Liu, D. Wu, J. Ye, Y. Deng, G. Zhou, H. H. Luu, R. C. Haydon, W. Si and T.-C. He, *PLoS One*, 2014, **9**, e92908.

149 K. Toyoshima and P. K. Vogt, *Virology*, 1969, **38**, 414–426.

150 A. Fasbender, J. Zabner, M. Chillón, T. O. Moninger, A. P. Puga, B. L. Davidson and M. J. Welsh, *J. Biol. Chem.*, 1997, **272**, 6479–6489.

151 A. Leydet, C. Moullet, J. P. Roque, M. Witvrouw, C. Pannecouque, G. Andrei, R. Snoeck, J. Neyts, D. Schols and E. De Clercq, *J. Med. Chem.*, 1998, **41**, 4927–4932.

152 M. Baba, R. Pauwels, J. Balzarini, J. Arnout, J. Desmyter and E. De Clercq, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 6132–6136.

153 N. Landázuri, D. Krishna, M. Gupta and J. M. Le Doux, *Biotechnol. Prog.*, 2007, **23**, 480–487.

154 M. Chillón, J. H. Lee, A. Fasbender and M. J. Welsh, *Gene Ther.*, 1998, **5**, 995–1002.

155 S. J. Beer, C. B. Matthews, C. S. Stein, B. D. Ross, J. M. Hilfinger and B. L. Davidson, *Gene Ther.*, 1998, **5**, 740–746.

156 K. D. Fisher, Y. Stallwood, N. K. Green, K. Ulbrich, V. Mautner and L. W. Seymour, *Gene Ther.*, 2001, **8**, 341–348.

157 Y. Eto, J. Q. Gao, F. Sekiguchi, S. Kurachi, K. Katayama, H. Mizuguchi, T. Hayakawa, Y. Tsutsumi, T. Mayumi and S. Nakagawa, *Biol. Pharm. Bull.*, 2004, **27**, 936–938.

158 N. K. Green, C. W. Herbert, S. J. Hale, A. B. Hale, V. Mautner, R. Harkins, T. Hermiston, K. Ulbrich, K. D. Fisher and L. W. Seymour, *Gene Ther.*, 2004, **11**, 1256–1263.

159 S. K. Seidlits, R. M. Gower, J. A. Shepard and L. D. Shea, *Expert Opin. Drug Delivery*, 2013, **10**, 499–509.

160 H.-F. Wu, J.-S. Cen, Q. Zhong, L. Chen, J. Wang, D. Y. B. Deng and Y. Wan, *Biomaterials*, 2013, **34**, 1686–1700.

161 J. A. Shepard, F. R. Virani, A. G. Goodman, T. D. Gossett, S. Shin and L. D. Shea, *Biomaterials*, 2012, **33**, 7412–7421.

162 A. Ehsanipour, T. Nguyen, T. Aboufadel, M. Sathialingam, P. Cox, W. Xiao, C. M. Walthers and S. K. Seidlits, *Cell. Mol. Bioeng.*, 2019, **12**, 399–413.

163 T. M. D. Le, B.-K. Jung, Y. Li, H. T. T. Duong, T. L. Nguyen, J. W. Hong, C.-O. Yun and D. S. Lee, *Biomater. Sci.*, 2019, **7**, 4195–4207.

164 I.-C. Liao, S. Chen, J. B. Liu and K. W. Leong, *J. Controlled Release*, 2009, **139**, 48–55.

165 M. J. Roberts, M. D. Bentley and J. M. Harris, *Adv. Drug Delivery Rev.*, 2012, **64**, 116–127.

166 A. L. Klibanov, K. Maruyama, V. P. Torchilin and L. Huang, *FEBS Lett.*, 1990, **268**, 235–237.

167 Y. Inada, M. Furukawa, H. Sasaki, Y. Kodera, M. Hiroto, H. Nishimura and A. Matsushima, *Trends Biotechnol.*, 1995, **13**, 86–91.

168 F. Martin, A. Huang, B. Uziely, B. Kaufman and T. Safra, *Cancer Res.*, 1994, **54**, 987–992.

169 A. Abuchowski, J. R. McCoy, N. C. Palczuk, T. van Es and F. F. Davis, *J. Biol. Chem.*, 1977, **252**, 3582–3586.

170 P. Wonganan and M. A. Croyle, *Viruses*, 2010, **2**, 468–502.

171 R. Alemany, K. Suzuki and D. T. Curiel, *J. Gen. Virol.*, 2000, **81**, 2605–2609.

172 H. Romanczuk, C. E. Galer, J. Zabner, G. Barsomian, S. C. Wadsworth and C. R. O'Riordan, *Hum. Gene Ther.*, 1999, **10**, 2615–2626.

173 K. I. Ogawara, M. G. Rots, R. J. Kok, H. E. Moorlag, A. M. Van Loenen, D. K. F. Meijer, H. J. Haisma and G. Molema, *Hum. Gene Ther.*, 2004, **15**, 433–443.

174 G. K. Lee, N. Maheshri, B. Kaspar and D. V. Schaffer, *Biotechnol. Bioeng.*, 2005, **92**, 24–34.

175 H. T. Le, Q. C. Yu, J. M. Wilson and M. A. Croyle, *J. Controlled Release*, 2005, **108**, 161–177.

176 T. Yao, X. Zhou, C. Zhang, X. Yu, Z. Tian, L. Zhang and D. Zhou, *Molecules*, 2017, **22**, 1155.

177 J. Q. Gao, Y. Eto, Y. Yoshioka, F. Sekiguchi, S. Kurachi, T. Morishige, X. Yao, H. Watanabe, R. Asavatanabodee, F. Sakurai, H. Mizuguchi, Y. Okada, Y. Mukai, Y. Tsutsumi, T. Mayumi, N. Okada and S. Nakagawa, *J. Controlled Release*, 2007, **122**, 102–110.

178 K. Doronin, E. V. Shashkova, S. M. May, S. E. Hofherr and M. A. Barry, *Hum. Gene Ther.*, 2009, **20**, 975–988.

179 S. E. Hofherr, E. V. Shashkova, E. A. Weaver, R. Khare and M. A. Barry, *Mol. Ther.*, 2008, **16**, 1276–1282.

180 A. Wortmann, S. Vöhringer, T. Engler, S. Corjon, R. Schirmbeck, J. Reimann, S. Kochanek and F. Kreppel, *Mol. Ther.*, 2008, **16**, 154–162.

181 H. Mok, D. J. Palmer, P. Ng and M. A. Barry, *Mol. Ther.*, 2005, **11**, 66–79.

182 M. Lyons, D. Onion, N. K. Green, K. Aslan, R. Rajaratnam, M. Bazan-Peregrino, S. Phipps, S. Hale, V. Mautner, L. W. Seymour and K. D. Fisher, *Mol. Ther.*, 2006, **14**, 118–128.

183 S. E. Hofherr, H. Mok, F. C. Gushiken, J. A. Lopez and M. A. Barry, *Hum. Gene Ther.*, 2007, **18**, 837–848.

184 H. Hatakeyama, H. Akita and H. Harashima, *Adv. Drug Delivery Rev.*, 2011, **63**, 152–160.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

185 J. W. Choi, S. J. Jung, D. Kasala, J. K. Hwang, J. Hu, Y. H. Bae and C. O. Yun, *J. Controlled Release*, 2015, **205**, 134–143.
186 K. Nam, H. Y. Nam, P. H. Kim and S. W. Kim, *Biomaterials*, 2012, **33**, 8122–8130.
187 D. Kasala, S. H. Lee, J. W. Hong, J. W. Choi, K. Nam, Y. H. Chung, S. W. Kim and C. O. Yun, *Biomaterials*, 2017, **145**, 207–222.
188 J. J. F. Verhoef, J. F. Carpenter, T. J. Anchordoquy and H. Schellekens, *Drug Discovery Today*, 2014, **19**, 1945–1952.
189 Y. Zhao, C. Wang, L. Wang, Q. Yang, W. Tang, Z. She and Y. Deng, *Eur. J. Pharm. Biopharm.*, 2012, **81**, 506–513.
190 T. T. Hoang Thi, E. H. Pilkington, D. H. Nguyen, J. S. Lee, K. D. Park and N. P. Truong, *Polymer*, 2020, **12**, 298.
191 S.-M. Zou, P. Erbacher, J.-S. Remy and J.-P. Behr, *J. Gene Med.*, 2000, **2**, 128–134.
192 W. T. Godbey, K. K. Wu and A. G. Mikos, *J. Controlled Release*, 1999, **60**, 149–160.
193 O. Boussif, F. Lezoualc'h, M. A. Zanta, M. D. Mergny, D. Scherman, B. Demeneix and J. P. Behr, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 7297–7301.
194 A. Akinc, M. Thomas, A. M. Klibanov and R. Langer, *J. Gene Med.*, 2005, **7**, 657–663.
195 L. M. P. Vermeulen, S. C. De Smedt, K. Remaut and K. Braeckmans, *Eur. J. Pharm. Biopharm.*, 2018, **129**, 184–190.
196 J. W. Choi, J. P. Nam, K. Nam, Y. S. Lee, C. O. Yun and S. W. Kim, *Biomacromolecules*, 2015, **16**, 2132–2143.
197 T. Blessing, M. Kursa, R. Holzhauser, R. Kircheis and E. Wagner, *Bioconjugate Chem.*, 2001, **12**, 529–537.
198 S. H. Pun, N. C. Bellocq, A. Liu, G. Jensen, T. Machemer, E. Quijano, T. Schluep, S. Wen, H. Engler, J. Heidel and M. E. Davis, *Bioconjugate Chem.*, 2004, **15**, 831–840.
199 H. Yao, S.-C. Chen, Z. Shen, Y.-C. Huang, X. Zhu, X. Wang, W. Jiang, Z.-F. Wang, X.-W. Bian, E.-A. Ling, H. Kung and M. C. Lin, *Curr. Med. Chem.*, 2013, **20**, 2601–2608.
200 J. Han, D. Zhao, Z. Zhong, Z. Zhang, T. Gong and X. Sun, *Nanotechnology*, 2010, **21**, 105106.
201 C. Yoshihara, K. Hamada and Y. Koyama, *Oncol. Rep.*, 2010, **23**, 733–738.
202 K. Nosaki, K. Hamada, Y. Takashima, M. Sagara, Y. Matsumura, S. Miyamoto, Y. Hijikata, T. Okazaki, Y. Nakanishi and K. Tani, *Mol. Ther.–Oncolytics*, 2016, **3**, 16022.
203 Y. Na, J. P. Nam, J. W. Hong, E. Oh, H. C. Shin, H. S. Kim, S. W. Kim and C. O. Yun, *J. Controlled Release*, 2019, **305**, 75–88.
204 S. Y. Chae, H. J. Kim, M. S. Lee, Y. L. Jang, Y. Lee, S. H. Lee, K. Lee, S. H. Kim, H. T. Kim, S.-C. Chi, T. G. Park and J. H. Jeong, *Macromol. Biosci.*, 2011, **11**, 1169–1174.
205 A. Bajaj, P. Kondaiah and S. Bhattacharya, *Bioconjugate Chem.*, 2008, **19**, 1640–1651.
206 D. Lee, D. Kim, H. Mok, J. H. Jeong, D. Choi and S. H. Kim, *Pharm. Res.*, 2012, **29**, 2213–2224.
207 C.-H. Lee, D. Kasala, Y. Na, M. S. Lee, S. W. Kim, J. H. Jeong and C.-O. Yun, *Biomaterials*, 2014, **35**, 5505–5516.
208 H. Pan, P. Li, G. Li, W. Li, B. Hu, H. He, Z. Chen, F. Wang, L. Liu, Y. Gong, Y. Han, Y. Luo, M. Zheng, Y. Ma, L. Cai and Y. Jin, *Adv. Funct. Mater.*, 2019, **29**, 1807528.
209 E. Kim, I. T. Song, S. Lee, J.-S. Kim, H. Lee and J.-H. Jang, *Angew. Chem., Int. Ed.*, 2012, **51**, 5598–5601.
210 S.-H. Kim, S. J. Yu, I. Kim, J. Choi, Y. H. Choi, S. G. Im and N. S. Hwang, *Chem. Commun.*, 2019, **55**, 2317–2320.
211 H. K. Makadia and S. J. Siegel, *Polymer*, 2011, **3**, 1377–1397.
212 M. N. V. Ravi Kumar, U. Bakowsky and C. M. Lehr, *Biomaterials*, 2004, **25**, 1771–1777.
213 C. B. Matthews, G. Jenkins, J. M. Hilfinger and B. L. Davidson, *Gene Ther.*, 1999, **6**, 1558–1564.
214 H. Mok, J. W. Park and T. G. Park, *Pharm. Res.*, 2007, **24**, 2263–2269.
215 Y. Zhu, D. Li, K. Zhang, L. Jiang, C. Shi, J. Fangteng, C. Zheng, B. Yang and H. Sun, *J. Biomed. Nanotechnol.*, 2017, **13**, 437–446.
216 N. Badrinath, Y. Il Jeong, H. Y. Woo, S. Y. Bang, C. Kim, J. Heo, D. H. Kang and S. Y. Yoo, *Int. J. Pharm.*, 2018, **552**, 437–442.
217 P. W. Lee, S. Shukla, J. D. Wallat, C. Danda, N. F. Steinmetz, J. Maia and J. K. Pokorski, *ACS Nano*, 2017, **11**, 8777–8789.
218 K. Deshmukh, M. Basheer Ahamed, R. R. Deshmukh, S. K. Khadheer Pasha, P. R. Bhagat and K. Chidambaram, in *Biopolymer Composites in Electronics*, Elsevier, Amsterdam, NL, 2017, pp. 27–128.
219 S. Stojanov and A. Berlec, *Front. Bioeng. Biotechnol.*, 2020, **8**, 130.
220 Y. Zheng, Y. Wu, Y. Zhou, J. Wu, X. Wang, Y. Qu, Y. Wang, Y. Zhang and Q. Yu, *ACS Appl. Mater. Interfaces*, 2020, **12**, 7905–7914.
221 Z. M. Huang, C. L. He, A. Yang, Y. Zhang, X. J. Han, J. Yin and Q. Wu, *J. Biomed. Mater. Res., Part A*, 2006, **77**, 169–179.
222 X. Gu, Y. Matsumura, Y. Tang, S. Roy, R. Hoff, B. Wang and W. R. Wagner, *Biomaterials*, 2017, **133**, 132–143.
223 S. Lee, J. S. Kim, H. S. Chu, G. W. Kim, J. I. Won and J. H. Jang, *Acta Biomater.*, 2011, **7**, 3868–3876.
224 S. H. Kim, S. Lee, H. Lee, M. Cho, D. V. Schaffer and J. H. Jang, *Mol. Ther.–Nucleic Acids*, 2019, **18**, 432–443.
225 J. K. Venkatesan, C. Falentin-Daudré, A. Leroux, V. Migonney and M. Cucchiarini, *Tissue Eng., Part A*, 2020, **26**, 450–459.
226 Wyandotte Chemicals Corporation, *U.S*, 2674619, 1954.
227 V. C. McLain, *Int. J. Toxicol.*, 2008, **27**, 93–128.
228 K. W. Chun, J. B. Lee, S. H. Kim and T. G. Park, *Biomaterials*, 2005, **26**, 3319–3326.
229 Y. Wang, S. Liu, C. Y. Li and F. Yuan, *Cancer Res.*, 2005, **65**, 7541–7545.
230 K. L. March, J. E. Madison and B. C. Trapnell, *Hum. Gene Ther.*, 1995, **6**, 41–53.

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

231 L. J. Feldman, C. J. Pastore, N. Aubailly, M. Kearney, D. Chen, M. Perricaudet, P. G. Steg and J. M. Isner, *Gene Ther.*, 1997, **4**, 189–198.
232 P. M. Strappe, D. W. Hampton, B. Cachon-Gonzalez, J. W. Fawcett and A. Lever, *Eur. J. Pharm. Biopharm.*, 2005, **61**, 126–133.
233 A. Rey-Rico, J. K. Venkatesan, J. Frisch, I. Rial-Hermida, G. Schmitt, A. Concheiro, H. Madry, C. Alvarez-Lorenzo and M. Cucchiarini, *Acta Biomater.*, 2015, **27**, 42–52.
234 A. Rey-Rico, J. K. Venkatesan, J. Frisch, G. Schmitt, A. Monge-Marcet, P. Lopez-Chicon, A. Mata, C. Semino, H. Madry and M. Cucchiarini, *Acta Biomater.*, 2015, **18**, 118–127.
235 H. H. Lee, A. M. Haleem, V. Yao, J. Li, X. Xiao and C. R. Chu, *Tissue Eng., Part A*, 2011, **17**, 1969–1978.
236 A. Rey-Rico, J. Frisch, J. K. Venkatesan, G. Schmitt, I. Rial-Hermida, P. Taboada, A. Concheiro, H. Madry, C. Alvarez-Lorenzo and M. Cucchiarini, *ACS Appl. Mater. Interfaces*, 2016, **8**, 20600–20613.
237 A. Rey-Rico, J. K. Venkatesan, G. Schmitt, S. Speicher-Mentges, H. Madry and M. Cucchiarini, *Mol. Pharm.*, 2018, **15**, 2816–2826.
238 H. Madry, L. Gao, A. Rey-Rico, J. K. Venkatesan, K. Müller-Brandt, X. Cai, L. Goebel, G. Schmitt, S. Speicher-Mentges, D. Zurakowski, M. D. Menger, M. W. Laschke and M. Cucchiarini, *Adv. Mater.*, 2020, **32**, 1906508.
239 J. M. Caronia, D. W. Sorensen, H. M. Leslie, J. H. van Berlo and S. M. Azarin, *Biotechnol. Bioeng.*, 2019, **116**, 2353–2363.
240 I. Höfig, M. J. Atkinson, S. Mall, A. M. Krackhardt, C. Thirion and N. Anastasov, *J. Gene Med.*, 2012, **14**, 549–560.
241 I. Hauber, N. Beschorner, S. Schrödel, J. Chemnitz, N. Kröger, J. Hauber and C. Thirion, *Hum. Gene Ther: Methods.*, 2018, **29**, 104–113.
242 Y. Jang, Y.-S. Kim, M. M. Wielgosz, F. Ferrara, Z. Ma, J. Condori, L. E. Palmer, X. Zhao, G. Kang, D. J. Rawlings, S. Zhou and B. Y. Ryu, *Gene Ther.*, 2020, 1–12.
243 J. W. Schott, D. León-Rico, C. B. Ferreira, K. F. Buckland, G. Santilli, M. A. Armant, A. Schambach, A. Cavazza and A. J. Thrasher, *Mol. Ther.–Methods Clin. Dev.*, 2019, **14**, 134–147.
244 M. Delville, T. Soheili, F. Bellier, A. Durand, A. Denis, C. Lagresle-Peyrou, M. Cavazzana, I. Andre-Schmutz and E. Six, *Mol. Ther.–Methods Clin. Dev.*, 2018, **10**, 341–347.
245 B. Simon, D. C. Harrer, C. Thirion, B. Schuler-Thurner, G. Schuler and U. Uslu, *J. Immunol. Methods*, 2019, **472**, 55–64.
246 E. Abbasi, S. Aval, A. Akbarzadeh, M. Milani, H. Nasrabadi, S. Joo, Y. Hanifehpour, K. Nejati-Koshki and R. Pashaei-Asl, *Nanoscale Res. Lett.*, 2014, **9**, 247.
247 J. Yang, Q. Zhang, H. Chang and Y. Cheng, *Chem. Rev.*, 2015, **115**, 5274–5300.
248 Z. Mhlwatika and B. Aderibigbe, *Molecules*, 2018, **23**, 2205.
249 M. A. Kostiainen, O. Kasyutich, J. J. L. M. Cornelissen and R. J. M. Nolte, *Nat. Chem.*, 2010, **2**, 394–399.
250 G. Navarro, G. Maiwald, R. Haase, A. L. Rogach, E. Wagner, C. T. de Ilarduya and M. Ogris, *J. Controlled Release*, 2010, **146**, 99–105.
251 A. R. Yoon, D. Kasala, Y. Li, J. Hong, W. Lee, S. J. Jung and C. O. Yun, *J. Controlled Release*, 2016, **231**, 2–16.
252 R. E. Bauer, C. G. Clark and K. Müllen, *New J. Chem.*, 2007, **31**, 1275–1282.
253 R. Stangenberg, Y. Wu, J. Hedrich, D. Kurzbach, D. Wehner, G. Weidinger, S. L. Kuan, M. I. Jansen, F. Jelezko, H. J. Luhmann, D. Hinderberger, T. Weil and K. Müllen, *Adv. Healthcare Mater.*, 2015, **4**, 377–384.
254 J. Wagner, L. Li, J. Simon, L. Krutzke, K. Landfester, V. Mailänder, K. Müllen, D. Y. W. Ng, Y. Wu and T. Weil, *Angew. Chem., Int. Ed.*, 2020, **59**, 5712–5720.
255 L. M. Kasman, S. Barua, P. Lu, K. Rege and C. Voelkel-Johnson, *Mol. Pharm.*, 2009, **6**, 1612–1619.
256 D. Steinhauff and H. Ghandehari, *Bioconjugate Chem.*, 2019, **30**, 384–399.
257 J. A. Rowley and D. J. Mooney, *J. Biomed. Mater. Res.*, 2002, **60**, 217–223.
258 G. Sailaja, H. HogenEsch, A. North, J. Hays and S. K. Mittal, *Gene Ther.*, 2002, **9**, 1722–1729.
259 J. W. Choi, E. Kang, O. J. Kwon, T. J. Yun, H. K. Park, P. H. Kim, S. W. Kim, J. H. Kim and C. O. Yun, *Gene Ther.*, 2013, **20**, 880–892.
260 R. S. Stilhano, J. L. Madrigal, K. Wong, P. A. Williams, P. K. M. Martin, F. S. M. Yamaguchi, V. Y. Samoto, S. W. Han and E. A. Silva, *J. Controlled Release*, 2016, **237**, 42–49.
261 J. L. Madrigal, R. S. Stilhano, C. Siltanen, K. Tanaka, S. N. Rezvani, R. P. Morgan, A. Revzin, S. W. Han and E. A. Silva, *J. Mater. Chem. B*, 2016, **4**, 6989–6999.
262 J. L. Madrigal, S. N. Sharma, K. T. Campbell, R. S. Stilhano, R. Gijsbers and E. A. Silva, *Acta Biomater.*, 2018, **69**, 265–276.
263 J. L. Madrigal, S. Shams, R. S. Stilhano and E. A. Silva, *Biomater. Sci.*, 2019, **7**, 645–656.
264 S. Magnusson and T. Berg, *Biochem. J.*, 1989, **257**, 651–656.
265 S. Espenlaub, A. Wortmann, T. Engler, S. Corjon, S. Kochanek and F. Kreppel, *J. Gene Med.*, 2008, **10**, 1303–1314.
266 Z. Liu, F. Ke, C. Duan, H. Lan, J. Li, C. Gao, J. Li and Z. Zhong, *Bioconjugate Chem.*, 2013, **24**, 1387–1397.
267 Y. Cho, R. Shi and R. B. Borgens, *J. Exp. Biol.*, 2010, **213**, 1513–1520.
268 Q. Guo, C. Liu, B. Hai, T. Ma, W. Zhang, J. Tan, X. Fu, H. Wang, Y. Xu and C. Song, *J. Biomed. Mater. Res., Part B*, 2018, **106**, 787–799.
269 R. Keswani, K. Su and D. W. Pack, *J. Controlled Release*, 2014, **192**, 40–46.
270 J. Szejtli, *Chem. Rev.*, 1998, **98**, 1743–1753.
271 A. Rey-Rico, H. Babicz, H. Madry, A. Concheiro, C. Alvarez-Lorenzo and M. Cucchiarini, *Int. J. Pharm.*, 2017, **531**, 492–503.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

272 L. Chen, X. Zhao, Y. Lin, Y. Huang and Q. Wang, *Chem. Commun.*, 2013, **49**, 9678–9680.
273 A. Rey-Rico and M. Cucchiarini, *Polymer*, 2019, **11**, 514.
274 H. Rosilo, J. R. McKee, E. Kontturi, T. Koho, V. P. Hytönen, O. Ikkala and M. A. Kostiainen, *Nanoscale*, 2014, **6**, 11871–11881.
275 Y. Huang, X. Ding, Y. Qi, B. Yu and F. J. Xu, *Biomaterials*, 2016, **106**, 134–143.
276 W. Fan, M. Shao, J. Zhang, G. Jin, F. Liu and F.-J. Xu, *Adv. Funct. Mater.*, 2019, **29**, 1807104.
277 M. Bauer, C. Lautenschlaeger, K. Kempe, L. Tauhardt, U. S. Schubert and D. Fischer, *Macromol. Biosci.*, 2012, **12**, 986–998.
278 S. Y. Fam, C. F. Chee, C. Y. Yong, K. L. Ho, A. R. Mariatulqabtiah, H. Y. Lau and W. S. Tan, *Int. J. Mol. Sci.*, 2019, **20**, 4903.
279 K. D. Fisher, N. K. Green, A. Hale, V. Subr, K. Ulbrich and L. W. Seymour, *J. Drug Targeting*, 2007, **15**, 546–551.
280 N. Francini, D. Cochrane, S. Illingworth, L. Purdie, G. Mantovani, K. Fisher, L. W. Seymour, S. G. Spain and C. Alexander, *Bioconjugate Chem.*, 2019, **30**, 1244–1257.
281 T. Kim, M. Ou, M. Lee and S. W. Kim, *Biomaterials*, 2009, **30**, 658–664.
282 Y. F. Zeng, S. J. Tseng, I. M. Kempson, S. F. Peng, W. T. Wu and J. R. Liu, *Biomaterials*, 2012, **33**, 9239–9245.
283 P. Shrimali, M. Peter, A. Singh, N. Dalal, S. Dakave, S. V. Chiplunkar and P. Tayalia, *Biomater. Sci.*, 2018, **6**, 3241–3250.
284 S. K. Cho and Y. J. Kwon, *Biomaterials*, 2012, **33**, 3316–3323.
285 C. A. Hong, S. K. Cho, J. A. Edson, J. Kim, D. Ingato, B. Pham, A. Chuang, D. A. Fruman and Y. J. Kwon, *ACS Nano*, 2016, **10**, 8705–8714.
286 G. J. Tong, S. C. Hsiao, Z. M. Carrico and M. B. Francis, *J. Am. Chem. Soc.*, 2009, **131**, 11174–11178.
287 D. Muharemagic, M. Labib, S. M. Ghobadloo, A. S. Zamay, J. C. Bell and M. V. Berezovski, *J. Am. Chem. Soc.*, 2012, **134**, 17168–17177.
288 Y. Wu, L. Zhang, C. Cui, S. Cansiz, H. Liang, C. Wu, I.-T. Teng, W. Chen, Y. Liu, W. Hou, X. Zhang and W. Tan, *J. Am. Chem. Soc.*, 2018, **140**, 2–5.
289 E. Kim, S. Lee, S. Hong, G. Jin, M. Kim, K. I. Park, H. Lee and J.-H. Jang, *ACS Appl. Mater. Interfaces*, 2014, **6**, 8288–8294.
290 S.-H. Kim, M. Lee, M. Cho, I.-S. Kim, K. I. Park, H. Lee and J.-H. Jang, *Macromol. Biosci.*, 2017, **17**, 1700148.
291 F. Alvarez-Rivera, A. Rey-Rico, J. K. Venkatesan, L. Diaz-Gomez, M. Cucchiarini, A. Concheiro and C. Alvarez-Lorenzo, *Pharmaceutics*, 2020, **12**, 335.
292 H. Qiao, X. Chen, Q. Wang, J. Zhang, D. Huang, E. Chen, H. Qian, Y. Zhong, Q. Tang and W. Chen, *Biomater. Sci.*, 2020, **8**, 2472–2480.
293 R. Myers, C. Coviello, P. Erbs, J. Foloppe, C. Rowe, J. Kwan, C. Crake, S. Finn, E. Jackson, J. M. Balloul, C. Story, C. Coussios and R. Carlisle, *Mol. Ther.*, 2016, **24**, 1627–1633.
294 J. L. Madrigal, S. Shams, R. S. Stilhano and E. A. Silva, *Biomater. Sci.*, 2019, **7**, 645–656.
295 L. M. Kasman, S. Barua, P. Lu, K. Rege and C. Voelkel-Johnson, *Mol. Pharm.*, 2009, **6**, 1612–1619.
296 M. Rabenstein and Y.-K. Shin, *Biochemistry*, 1995, **34**, 13390–13397.
297 Z. Kang, Q. Meng and K. Liu, *J. Mater. Chem. B*, 2019, **7**, 1824–1841.
298 K. Numata, Y. Horii, K. Oikawa, Y. Miyagi, T. Demura and M. Ohtani, *Sci. Rep.*, 2018, **8**, 1–17.
299 F. Milletti, *Drug Discovery Today*, 2012, **17**, 850–860.
300 A. D. Frankel and C. O. Pabo, *Cell*, 1988, **55**, 1189–1193.
301 M. Green and P. M. Loewenstein, *Cell*, 1988, **55**, 1179–1188.
302 P. Agrawal, S. Bhalla, S. S. Usmani, S. Singh, K. Chaudhary, G. P. S. Raghava and A. Gautam, *Nucleic Acids Res.*, 2016, **44**, D1098–D1103.
303 B. Gupta, T. Levchenko and V. Torchilin, *Adv. Drug Delivery Rev.*, 2005, **57**, 637–651.
304 S. M. Farkhani, A. Valizadeh, H. Karami, S. Mohammadi, N. Sohrabi and F. Badrzadeh, *Peptides*, 2014, **57**, 78–94.
305 J. Váňová, A. Hejtmánková, M. H. Kalbáčová and H. Španielová, *Materials*, 2019, **12**, 2671.
306 R. E. Taylor and M. Zahid, *Pharmaceutics*, 2020, **12**, 225.
307 A. S. Nigatu, S. Vupputuri, N. Flynn, B. J. Neely and J. D. Ramsey, *J. Pharm. Sci.*, 2013, **102**, 1981–1993.
308 J. P. Gratton, J. Yu, J. W. Griffith, R. W. Babbitt, R. S. Scotland, R. Hickey, F. J. Giordano and W. C. Sessa, *Nat. Med.*, 2003, **9**, 357–362.
309 F. Kühnel, B. Schulte, T. Wirth, N. Woller, S. Schäfers, L. Zender, M. Manns and S. Kubicka, *J. Virol.*, 2004, **78**, 13743–13754.
310 Y. Yoshioka, R. Asavatanabodee, Y. Eto, H. Watanabe, T. Morishige, X. Yao, S. Kida, M. Maeda, Y. Mukai, H. Mizuguchi, K. Kawasaki, N. Okada and S. Nakagawa, *Life Sci.*, 2008, **83**, 747–755.
311 A. S. Nigatu, S. Vupputuri, N. Flynn and J. D. Ramsey, *Biomed. Pharmacother.*, 2015, **71**, 153–160.
312 P. Brugada-Vilà, A. Cascante, M. Á. Lázaro, C. Castells-Sala, C. Fornaguera, M. Rovira-Rigau, L. Albertazzi, S. Borros and C. Fillat, *Theranostics*, 2020, **10**, 2744–2758.
313 T. V. Nguyen, S. S. Anguiano-Zarate, W. E. Matchett, M. E. Barry and M. A. Barry, *Virology*, 2018, **514**, 118–123.
314 B. K. Gan, C. Y. Yong, K. L. Ho, A. R. Omar, N. B. Alitheen and W. S. Tan, *Sci. Rep.*, 2018, **8**, 8499.
315 T. Han, Y. Tang, H. Ugai, L. E. Perry, G. P. Siegal, J. L. Contreras and H. Wu, *Virol. J.*, 2007, **4**, 103.
316 Y. Eto, Y. Yoshioka, R. Asavatanabodee, S. Kida, M. Maeda, Y. Mukai, H. Mizuguchi, K. Kawasaki, N. Okada and S. Nakagawa, *Peptides*, 2009, **30**, 1548–1552.
317 S. Kida, Y. Eto, Y. Yoshioka, S. Nakagawa, K. Kawasaki and M. Maeda, *Protein Pept. Lett.*, 2010, **17**, 164–167.
318 A. S. Nigatu, S. Vupputuri, N. Flynn, B. J. Neely and J. D. Ramsey, *J. Pharm. Sci.*, 2013, **102**, 1981–1993.

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

319 Z. Wu, K. Chen, I. Yildiz, A. Dirksen, R. Fischer, P. E. Dawson and N. F. Steinmetz, *Nanoscale*, 2012, **4**, 3567.
320 D. Kim, Y. Lee, T. W. Dreher and T.-J. Cho, *Virus Res.*, 2018, **252**, 13–21.
321 X. Zhang, T. He, Z. Chai, R. J. Samulski and C. Li, *Biomaterials*, 2018, **176**, 71–83.
322 X. Ma, P. Liu, X. Zhang, W. Jiang, M. Jia, C. Wang, Y. Dong, Y. Dang and C. Gao, *Sci. Rep.*, 2016, **6**, 22404.
323 E. Koren and V. P. Torchilin, *Trends Mol. Med.*, 2012, **18**, 385–393.
324 S. Pujals, J. Fernández-Carneado, C. López-Iglesias, M. J. Kogan and E. Giralt, *Biochim. Biophys. Acta, Biomembr.*, 2006, **1758**, 264–279.
325 S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda and Y. Sugiura, *J. Biol. Chem.*, 2001, **276**, 5836–5840.
326 D. J. Mitchell, L. Steinman, D. T. Kim, C. G. Fathman and J. B. Rothbard, *J. Pept. Res.*, 2000, **56**, 318–325.
327 P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 13003–13008.
328 Y. Liu, Y. J. Kim, M. Ji, J. Fang, N. Siriwon, L. I. Zhang and P. Wang, *Mol. Ther.–Methods Clin. Dev.*, 2014, **1**, 12.
329 J. Váňová, A. Hejtmánková, J. Žáčková Suchanová, P. Sauerová, J. Forstová, M. Hubálek Kalbáčová and H. Španielová, *Int. J. Pharm.*, 2020, **576**, 119008.
330 S. Das, R. S. Jacob, K. Patel, N. Singh and S. K. Maji, *Biomacromolecules*, 2018, **19**, 1826–1839.
331 J. Chen and X. Zou, *Bioact. Mater.*, 2019, **4**, 120–131.
332 C. Meier, T. Weil, F. Kirchhoff and J. Münch, *WIREs Nanomed. Nanobiotechnol.*, 2014, **6**, 438–451.
333 G. G. Glenner and C. W. Wong, *Biochem. Biophys. Res. Commun.*, 1984, **122**, 1131–1135.
334 W. M. Wojtowicz, M. Farzan, J. L. Joyal, K. Carter, G. J. Babcock, D. I. Israel, J. Sodroski and T. Mirzabekov, *J. Biol. Chem.*, 2002, **277**, 35019–35024.
335 D. Romero, C. Aguilar, R. Losick and R. Kolter, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 2230–2234.
336 J. Münch, E. Rücker, L. Ständker, K. Adermann, C. Goffinet, M. Schindler, S. Wildum, R. Chinnadurai, D. Rajan, A. Specht, G. Giménez-Gallego, P. C. Sánchez, D. M. Fowler, A. Koulov, J. W. Kelly, W. Mothes, J. C. Grivel, L. Margolis, O. T. Keppler, W. G. Forssmann and F. Kirchhoff, *Cell*, 2007, **131**, 1059–1071.
337 F. Arnold, J. Schnell, O. Zirafi, C. Sturzel, C. Meier, T. Weil, L. Standker, W.-G. Forssmann, N. R. Roan, W. C. Greene, F. Kirchhoff and J. Münch, *J. Virol.*, 2012, **86**, 1244–1249.
338 N. R. Roan, J. A. Müller, H. Liu, S. Chu, F. Arnold, C. M. Stürzel, P. Walther, M. Dong, H. E. Witkowska, F. Kirchhoff, J. Münch and W. C. Greene, *Cell Host Microbe*, 2011, **10**, 541–550.
339 L. Zhang, C. Jiang, H. Zhang, X. Gong, L. Yang, L. Miao, Y. Shi, Y. Zhang, W. Kong, C. Zhang and Y. Shan, *J. Pept. Sci.*, 2014, **20**, 46–54.
340 S. Rode, M. Hayn, A. Röcker, S. Sieste, M. Lamla, D. Markx, C. Meier, F. Kirchhoff, P. Walther, M. Fändrich, T. Weil and J. Münch, *Bioconjugate Chem.*, 2017, **28**, 1260–1270.
341 H. Zhang, X. He, Y. Shi, Y. Yu, S. Guan, X. Gong, H. Yin, Z. Kuai and Y. Shan, *RSC Adv.*, 2016, **6**, 82082–82087.
342 J. Chen, R. Ren, F. Yu, C. Wang, X. Zhang, W. Li, S. Tan, S. Jiang, S. Liu and L. Li, *Biophys. J.*, 2017, **113**, 1425–1439.
343 J. Chen, R. Ren, S. Tan, W. Zhang, X. Zhang, F. Yu, T. Xun, S. Jiang, S. Liu and L. Li, *PLoS One*, 2015, **10**, e0144522.
344 S. Tan, L. Li, L. Lu, C. Pan, H. Lu, Y. Oksov, X. Tang, S. Jiang and S. Liu, *FEBS Lett.*, 2014, **588**, 1515–1522.
345 J. Münch, E. Rücker, L. Ständker, K. Adermann, C. Goffinet, M. Schindler, S. Wildum, R. Chinnadurai, D. Rajan, A. Specht, G. Giménez-Gallego, P. C. Sánchez, D. M. Fowler, A. Koulov, J. W. Kelly, W. Mothes, J.-C. Grivel, L. Margolis, O. T. Keppler, W.-G. Forssmann and F. Kirchhoff, *Cell*, 2007, **131**, 1059–1071.
346 A. Groß, K. Rödel, B. Kneidl, N. Donhauser, M. Mössl, E. Lump, J. Münch, B. Schmidt and J. Eichler, *ChemBioChem*, 2015, **16**, 446–454.
347 C. Kokotidou, S. V. R. Jonnalagadda, A. A. Orr, G. Vrentzos, A. Kretsovali, P. Tamamis and A. Mitraki, *Biomolecules*, 2020, **10**, 1–31.
348 A. L. Rodriguez, T.-Y. Wang, K. F. Bruggeman, R. Li, R. J. Williams, C. L. Parish and D. R. Nisbet, *Nano Res.*, 2016, **9**, 674–684.
349 S. Zhang, T. C. Holmes, C. M. DiPersio, R. O. Hynes, X. Su and A. Rich, *Biomaterials*, 1995, **16**, 1385–1393.
350 A. Kichler, C. Leborgne, J. Marz, O. Danos and B. Bechinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1564–1568.
351 D. Fenard, S. Genries, D. Scherman, A. Galy, S. Martin and A. Kichler, *J. Virol. Methods*, 2013, **189**, 375–378.
352 C. Piovan, V. Marin, C. Scavullo, S. Corna, E. Giuliani, S. Bossi, A. Galy, D. Fenard, C. Bordignon, G. P. Rizzardi and C. Bovolenta, *Mol. Ther.–Methods Clin. Dev.*, 2017, **5**, 22–30.
353 C. Radek, O. Bernadin, K. Drechsel, N. Cordes, R. Pfeifer, P. Sträßer, M. Mormin, A. Gutierrez-Guerrero, F. Cosset, A. D. Kaiser, T. Schaser, A. Galy, E. Verhoeyen and I. C. D. Johnston, *Hum. Gene Ther.*, 2019, **30**, 1477–1493.
354 A. Jamali, L. Kapitza, T. Schaser, I. C. D. Johnston, C. J. Buchholz and J. Hartmann, *Mol. Ther.–Methods Clin. Dev.*, 2019, **13**, 371–379.
355 S. Majdoul, A. K. Seye, A. Kichler, N. Holic, A. Galy, B. Bechinger and D. Fenard, *J. Biol. Chem.*, 2016, **291**, 2161–2169.
356 K. Cornetta and W. F. Anderson, *J. Virol. Methods*, 1989, **23**, 187–194.
357 S. M. Arcasoy, J. D. Latoche, M. Gondor, B. R. Pitt and J. M. Pilewski, *Gene Ther.*, 1997, **4**, 32–38.
358 J. Y. Lee and H. H. Lee, *Cytotechnology*, 2018, **70**, 193–201.
359 H. Hanenberg, X. L. Xiao, D. Dilloo, K. Hashino, I. Kato and D. A. Williams, *Nat. Med.*, 1996, **2**, 876–882.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

360 A. Quintás-Cardama, R. K. Yeh, D. Hollyman, J. Stefanski, C. Taylor, Y. Nikhamin, G. Imperato, M. Sadelain, I. Rivière and R. J. Brentjens, *Hum. Gene Ther.*, 2007, **18**, 1253–1260.
361 G. Enblad, H. Karlsson, G. Gammelgård, J. Wenthe, T. Lövgren, R. M. Amini, K. I. Wikstrom, M. Essand, B. Savoldo, H. Hallböök, M. Höglund, G. Dotti, M. K. Brenner, H. Hagberg and A. Loskog, *Clin. Cancer Res.*, 2018, **24**, 6185–6194.
362 K. Dodo, H. Chono, N. Saito, Y. Tanaka, K. Tahara, I. Nukaya and J. Mineno, *PLoS One*, 2014, **9**, e86275.
363 B. Heemskerk, A. Jorritsma, R. Gomez-Eerland, M. Toebes, J. B. A. G. Haanen and T. N. M. Schumacher, *Hum. Gene Ther.*, 2010, **21**, 1335–1342.
364 H.-J. Lee, Y.-S. Lee, H.-S. Kim, Y.-K. Kim, J.-H. Kim, S.-H. Jeon, H.-W. Lee, S. Kim, H. Miyoshi, H.-M. Chung and D.-K. Kim, *Biologicals*, 2009, **37**, 203–209.
365 H. Mizuguchi, T. Sasaki, K. Kawabata, F. Sakurai and T. Hayakawa, *Biochem. Biophys. Res. Commun.*, 2005, **332**, 1101–1106.
366 K. Tashiro, A. Kondo, K. Kawabata, H. Sakurai, F. Sakurai, K. Yamanishi, T. Hayakawa and H. Mizuguchi, *Biochem. Biophys. Res. Commun.*, 2009, **379**, 127–132.
367 J. D. Ramsey, H. N. Vu and D. W. Pack, *J. Controlled Release*, 2010, **144**, 39–45.
368 H. Katakura, A. Harada, K. Kataoka, M. Furusho, F. Tanaka, H. Wada and K. Ikenaka, *J. Gene Med.*, 2004, **6**, 471–477.
369 M. Skoumal, S. Seidlits, S. Shin and L. Shea, *Biotechnol. Bioeng.*, 2016, **113**, 2033–2040.
370 H. Fang, Z. Guo, L. Lin, J. Chen, P. Sun, J. Wu, C. Xu, H. Tian and X. Chen, *J. Am. Chem. Soc.*, 2018, **140**, 11992–12000.
371 H. Ran, G. Quan, Y. Huang, C. Zhu, C. Lu, W. Liu, X. Pan and C. Wu, *Biochem. Biophys. Res. Commun.*, 2019, **508**, 791–796.
372 A. Breen, P. Dockery, T. O'Brien and A. Pandit, *J. Biomed. Mater. Res., Part A*, 2009, **89**, 876–884.
373 C. Schmidt, D. Bezuidenhout, P. Zilla and N. H. Davies, *J. Biomater. Appl.*, 2014, **28**, 1408–1418.
374 M. E. Kidd, S. Shin and L. D. Shea, *J. Controlled Release*, 2012, **157**, 80–85.
375 M. Wang, J. Sun, A. Crosby, K. Woodard, M. L. Hirsch, R. J. Samulski and C. Li, *Gene Ther.*, 2017, **24**, 49–59.
376 D. Palesch, F. Boldt, J. A. Müller, K. Eisele, C. M. Stürzel, Y. Wu, J. Münch and T. Weil, *ChemBioChem*, 2016, **17**, 1504–1508.
377 X. Pei, T. He, N. E. Hall, D. Gerber, R. J. Samulski and C. Li, *Virology*, 2018, **518**, 95–102.
378 J. Park, P. L. Chariou and N. F. Steinmetz, *Bioconjugate Chem.*, 2020, **31**, 1408–1416.
379 J. A. Gustafson and H. Ghandehari, *Adv. Drug Delivery Rev.*, 2010, **62**, 1509–1523.
380 J. Cappello, J. Crissman, M. Dorman, M. Mikolajczak, G. Textor, M. Marquet and F. Ferrari, *Biotechnol. Prog.*, 1990, **6**, 198–202.
381 Z. Megeed, M. Haider, D. Li, B. W. O'Malley, J. Cappello and H. Ghandehari, *J. Controlled Release*, 2004, **94**, 433–445.
382 A. Hatefi, J. Cappello and H. Ghandehari, *Pharm. Res.*, 2007, **24**, 773–779.
383 D. L. Price, P. Li, C.-H. Chen, D. Wong, Z. Yu, N. G. Chen, Y. A. Yu, A. A. Szalay, J. Cappello, Y. Fong and R. J. Wong, *Head Neck*, 2016, **38**, 237–246.
384 R. Price, J. Gustafson, K. Greish, J. Cappello, L. McGill and H. Ghandehari, *Int. J. Pharm.*, 2012, **427**, 97–104.
385 W.-W. Hu, Z. Wang and P. H. Krebsbach, *J. Tissue Eng. Regener. Med.*, 2016, **10**, E63–E72.
386 B. K. Jung, E. Oh, J. W. Hong, Y. Lee, K. D. Park and C. O. Yun, *Biomaterials*, 2017, **147**, 26–38.
387 R. A. Haraszti, M. C. Didiot, E. Sapp, J. Leszyk, S. A. Shaffer, H. E. Rockwell, F. Gao, N. R. Narain, M. DiFiglia, M. A. Kiebish, N. Aronin and A. Khvorova, *J. Extracell. Vesicles*, 2016, **5**, 32570.
388 D. Bitounis, R. Fanciullino, A. Iliadis and J. Ciccolini, *ISRN Pharm.*, 2012, **2012**, 1–11.
389 A. D. Miller, *Angew. Chem., Int. Ed.*, 1998, **37**, 1768–1785.
390 P. Yotnda, D. H. Chen, W. Chiu, P. A. Piedra, A. Davis, N. S. Templeton and M. K. Brenner, *Mol. Ther.*, 2002, **5**, 233–241.
391 Y. L. Tseng, J. J. Liu and R. L. Hong, *Mol. Pharmacol.*, 2002, **62**, 864–872.
392 A. D. Bangham and R. W. Horne, *J. Mol. Biol.*, 1964, **8**, 660–668.
393 G. Bozzuto and A. Molinari, *Int. J. Nanomed.*, 2015, **10**, 975.
394 B. Balakrishnan and E. David, *J. Biosci.*, 2019, **44**, 1–8.
395 I. Tariq, S. R. Pinnapireddy, L. Duse, M. Y. Ali, S. Ali, M. U. Amin, N. Goergen, J. Jedelská, J. Schäfer and U. Bakowsky, *Eur. J. Pharm. Biopharm.*, 2019, **135**, 72–82.
396 E. Baghdan, S. R. Pinnapireddy, B. Strehlow, K. H. Engelhardt, J. Schäfer and U. Bakowsky, *Int. J. Pharm.*, 2018, **535**, 473–479.
397 B. Malaekeh-Nikouei, L. Gholami, F. Asghari, S. Askarian, S. Barzegar, M. Rezaee and R. Kazemi Oskuee, *Colloids Surf., B*, 2018, **165**, 252–261.
398 J. Buck, P. Grossen, P. R. Cullis, J. Huwyler and D. Witzigmann, *ACS Nano*, 2019, **13**, 3754–3782.
399 I. A. Khalil, M. A. Younis, S. Kimura and H. Harashima, *Biol. Pharm. Bull*, 2020, **584**, 584–595.
400 C. P. Hodgson and F. Solaiman, *Nat. Biotechnol.*, 1996, **14**, 339–342.
401 C. D. Porter, K. V. Lukacs, G. Box, Y. Takeuchi and M. K. L. Collins, *J. Virol.*, 1998, **72**, 4832.
402 S. Fletcher, A. Ahmad, E. Perouzel, M. R. Jorgensen and A. D. Miller, *Org. Biomol. Chem.*, 2006, **4**, 196–199.
403 P. L. G. Chong, W. Zhu and B. Venegas, *Biochim. Biophys. Acta, Biomembr.*, 2009, **1788**, 2–11.
404 N. Imelli, O. Meier, K. Boucke, S. Hemmi and U. F. Greber, *J. Virol.*, 2004, **78**, 3089–3098.
405 S. Worgall, T. S. Worgall, K. Kostarelos, R. Singh, P. L. Leopold, N. R. Hackett and R. G. Crystal, *Mol. Ther.*, 2000, **1**, 39–48.

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

406 D. V. Faller and D. Baltimore, *J. Virol.*, 1984, **49**, 269–272.
407 M. Mizuno and J. Yoshida, *J. Cancer Res.*, 1998, **89**, 352–354.
408 C. Qiu, M. B. de Young, A. Finn and D. A. Dichek, *Hum. Gene Ther.*, 1998, **9**, 507–520.
409 R. Singh, K. T. Al-Jamal, L. Lacerda and K. Kostarelos, *ACS Nano*, 2008, **2**, 1040–1050.
410 R. K. Keswani, I. M. Pozdol and D. W. Pack, *Mol. Pharm.*, 2013, **10**, 1725–1735.
411 R. Mo, Q. Sun, N. Li and C. Zhang, *Biomaterials*, 2013, **34**, 2773–2786.
412 J. Van den Bossche, W. T. Al-Jamal, A. Yilmazer, E. Bizzarri, B. Tian and K. Kostarelos, *Biomaterials*, 2011, **32**, 3085–3093.
413 Y. Wan, J. Han, G. Fan, Z. Zhang, T. Gong and X. Sun, *Biomaterials*, 2013, **34**, 3020–3030.
414 C. L. Innes, P. B. Smith, R. Langenbach, K. R. Tindall and L. R. Boone, *J. Virol.*, 1990, **64**, 957–961.
415 C. Meunier-Durmort, R. Picart, T. Ragot, M. Perricaudet, B. Hainque and C. Forest, *Biochim. Biophys. Acta, Biomembr.*, 1997, **1330**, 8–16.
416 H. S. Shi, L. P. Yang, W. Wei, X. Q. Su, X. P. Li, M. Li, S. T. Luo, H. L. Zhang, L. Lu, Y. Q. Mao, B. Kan and L. Yang, *J. Transl. Med.*, 2013, **11**, 86.
417 A. Natsume, M. Mizuno, Y. Ryuke and J. Yoshida, *Jpn. J. Cancer Res.*, 2000, **91**, 363–367.
418 T. Shikano, H. Kasuya, T. T. Sahin, N. Nomura, A. Kanzaki, M. Misawa, Y. Nishikawa, T. Shirota, S. Yamada, T. Fujii, H. Sugimoto, N. Kanazumi, S. Nomoto, S. Takeda and A. Nakao, *Curr. Cancer Drug Targets*, 2011, **11**, 111–122.
419 X. Fu and X. Zhang, *Mol. Ther.*, 2001, **4**, 447–453.
420 F. Sakurai, S. Inoue, T. Kaminade, T. Hotani, Y. Katayama, E. Hosoyamada, Y. Terasawa, M. Tachibana and H. Mizuguchi, *Int. J. Pharm.*, 2017, **524**, 238–247.
421 Y. Wang, H. Huang, H. Zou, X. Tian, J. Hu, P. Qiu, H. Hu and G. Yan, *Mol. Pharm.*, 2019, **16**, 779–785.
422 N. Mendez, V. Herrera, L. Zhang, F. Hedjran, R. Feuer, S. L. Blair, W. C. Trogler, T. R. Reid and A. C. Kummel, *Biomaterials*, 2014, **35**, 9554–9561.
423 A. Yilmazer, B. Tian and K. Kostarelos, *PLoS One*, 2014, **9**, e114985.
424 M. Limberis, D. S. Anson, M. Fuller and D. W. Parsons, *Hum. Gene Ther.*, 2002, **13**, 1961–1970.
425 A. G. Stocker, K. L. Kremer, R. Koldej, D. S. Miller, D. S. Anson and D. W. Parsons, *J. Gene Med.*, 2009, **11**, 861–867.
426 P. Cmielewski, D. S. Anson and D. W. Parsons, *Respir. Res.*, 2010, **11**, 84.
427 P. Cmielewski, N. Farrow, S. Devereux, D. Parsons and M. Donnelley, *Exp. Lung Res.*, 2017, **43**, 426–433.
428 C. A. Maguire, L. Balaj, S. Sivaraman, M. H. W. Crommentuijn, M. Ericsson, L. Mincheva-Nilsson, V. Baranov, D. Gianni, B. A. Tannous, M. Sena-Esteves, X. O. Breakefield and J. Skog, *Mol. Ther.*, 2012, **20**, 960–971.
429 B. György and C. A. Maguire, *WIREs Nanomed. Nanobiotechnol.*, 2018, **10**, e1488.
430 T. Ishida, H. Harashima and H. Kiwada, *Biosci. Rep.*, 2002, **22**, 197–224.
431 M. L. Immordino, F. Dosio and L. Cattel, *Int. J. Nanomed.*, 2006, **1**, 297–315.
432 S. Y. Han, Y. J. Lee, H. I. Jung, S. W. Lee, S. J. Lim, S. H. Hong and J. S. Jeong, *Exp. Mol. Med.*, 2008, **40**, 427–434.
433 L. Yang, L. Wang, X. Q. Su, L. Wang, X. C. Chen, D. Li, S. T. Luo, H. S. Shi, L. J. Chen and Y. S. Wang, *Cancer Gene Ther.*, 2010, **17**, 49–57.
434 J. Chen, P. Gao, S. Yuan, R. Li, A. Ni, L. Chu, L. Ding, Y. Sun, X. Y. Liu and Y. Duan, *ACS Nano*, 2016, **10**, 11548–11560.
435 N. G. Mukherjee, L. Andrew Lyon and J. M. Le Doux, *Nanotechnology*, 2009, **20**, 065103.
436 V. P. Torchilin, R. Rammohan, V. Weissig and T. S. Levchenko, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8786–8791.
437 Z. Zhong, J. Han, Y. Wan, Z. Zhang and X. Sun, *Mol. Pharm.*, 2011, **8**, 673–682.
438 K. K. Son, D. Tkach and D. H. Patel, *Biochim. Biophys. Acta, Biomembr.*, 2000, **1468**, 11–14.
439 L. F. Neves, J. Duan, A. Voelker, A. Khanal, L. McNally, J. Steinbach-Rankins and B. P. Ceresa, *J. Microencapsulation*, 2016, **33**, 391–399.
440 S. Resina, P. Prevot and A. R. Thierry, *PLoS One*, 2009, **4**, e6058.
441 Y. Nie, L. Ji, H. Ding, L. Xie, L. Li, B. He, Y. Wu and Z. Gu, *Theranostics*, 2012, **2**, 1092–1103.
442 M. Kapoor and D. J. Burgess, *Int. J. Pharm.*, 2012, **432**, 80–90.
443 L. Billiet, J. P. Gomez, M. Berchel, P. A. Jaffrès, T. Le Gall, T. Montier, E. Bertrand, H. Cheradame, P. Guégan, M. Mével, B. Pitard, T. Benvegnu, P. Lehn, C. Pichon and P. Midoux, *Biomaterials*, 2012, **33**, 2980–2990.
444 M. Jedynak, R. Worch, M. Podsiadła-Białoskórska, J. Chroboczek and E. Szołajska, *Biochim. Biophys. Acta, Biomembr.*, 2018, **1860**, 2215–2223.
445 M. Mazzon and J. Mercer, *Cell. Microbiol.*, 2014, **16**, 1493–1502.
446 R. F. A. Zwaal, P. Comfurius and E. M. Bevers, *Cell. Mol. Life Sci.*, 2005, **62**, 971–988.
447 S. Ran and P. E. Thorpe, *Int. J. Radiat. Oncol., Biol., Phys.*, 2002, **54**, 1479–1484.
448 C. Burda, X. Chen, R. Narayanan and M. A. El-Sayed, *Chem. Rev.*, 2005, **105**, 1025–1102.
449 S. S. Lucky, K. C. Soo and Y. Zhang, *Chem. Rev.*, 2015, **115**, 1990–2042.
450 W. Chen, G. Wang and R. Tang, *Nano Res.*, 2014, **7**, 1404–1428.
451 X. Wang, X. Liu, Y. Xiao, H. Hao, Y. Zhang and R. Tang, *Chem. – Eur. J.*, 2018, **24**, 11518–11529.
452 H. L. Jang, K. Jin, J. Lee, Y. Kim, S. H. Nahm, K. S. Hong and K. T. Nam, *ACS Nano*, 2014, **8**, 634–641.

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

453 G. Wang, R. Y. Cao, R. Chen, L. Mo, J. F. Han, X. Wang, X. Xu, T. Jiang, Y. Q. Deng, K. Lyu, S. Y. Zhu, E. De Qin, R. Tang and C. F. Qin, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7619–7624.
454 L.-L. Huang, X. Li, J. Zhang, Q. R. Zhao, M. J. Zhang, A.-A. Liu, D.-W. Pang and H.-Y. Xie, *Nano Lett.*, 2019, **19**, 8002–8009.
455 P. Ghosh, G. Han, M. De, C. K. Kim and V. M. Rotello, *Adv. Drug Delivery Rev.*, 2008, **60**, 1307–1315.
456 M. Everts, V. Saini, J. L. Leddon, R. J. Kok, M. Stoff-Khalili, M. A. Preuss, C. L. Millican, G. Perkins, J. M. Brown, H. Bagaria, D. E. Nikles, D. T. Johnson, V. P. Zharov and D. T. Curiel, *Nano Lett.*, 2006, **6**, 587–591.
457 Y. Hernandez, R. González-Pastor, C. Belmar-Lopez, G. Mendoza, J. M. De La Fuente and P. Martin-Duque, *RSC Adv.*, 2019, **9**, 1327–1334.
458 J. Kim, M. Yeom, T. Lee, H.-O. Kim, W. Na, A. Kang, J.-W. Lim, G. Park, C. Park, D. Song and S. Haam, *J. Nanobiotechnol.*, 2020, **18**, 54.
459 I. Papp, C. Sieben, K. Ludwig, M. Roskamp, C. Böttcher, S. Schlecht, A. Herrmann and R. Haag, *Small*, 2010, **6**, 2900–2906.
460 M. Sametband, S. Shukla, T. Meningher, S. Hirsh, E. Mendelson, R. Sarid, A. Gedanken and M. Mandelboim, *MedChemComm*, 2011, **2**, 421.
461 M.-C. Bowman, T. E. Ballard, C. J. Ackerson, D. L. Feldheim, D. M. Margolis and C. Melander, *J. Am. Chem. Soc.*, 2008, **130**, 6896–6897.
462 V. P. Zharov, J. W. Kim, D. T. Curiel and M. Everts, *Nanomedicine*, 2005, **1**, 326–345.
463 B. K. Jung, Y. K. Lee, J. Hong, H. Ghandehari and C. O. Yun, *ACS Nano*, 2016, **10**, 10533–10543.
464 Y. Wang, Q. Zhao, N. Han, L. Bai, J. Li, J. Liu, E. Che, L. Hu, Q. Zhang, T. Jiang and S. Wang, *Nanomedicine*, 2015, **11**, 313–327.
465 R. Peters, E. Kramer, A. G. Oomen, Z. E. Herrera Rivera, G. Oegema, P. C. Tromp, R. Fokkink, A. Rietveld, H. J. P. Marvin, S. Weigel, A. A. C. M. Peijnenburg and H. Bouwmeester, *ACS Nano*, 2012, **6**, 2441–2451.
466 M. Ekkapongpisit, A. Giovia, C. Follo, G. Caputo and C. Isidoro, *Int. J. Nanomed.*, 2012, **7**, 4147–4158.
467 X. Huang, X. Qi, F. Boey and H. Zhang, *Chem. Soc. Rev.*, 2012, **41**, 666–686.
468 J. Tian, Y. Luo, L. Huang, Y. Feng, H. Ju and B. Y. Yu, *Biosens. Bioelectron.*, 2016, **80**, 519–524.
469 J. T. Robinson, S. M. Tabakman, Y. Liang, H. Wang, H. Sanchez Casalongue, D. Vinh and H. Dai, *J. Am. Chem. Soc.*, 2011, **133**, 6825–6831.
470 K. Yang, L. Hu, X. Ma, S. Ye, L. Cheng, X. Shi, C. Li, Y. Li and Z. Liu, *Adv. Mater.*, 2012, **24**, 1868–1872.
471 L. Feng, X. Yang, X. Shi, X. Tan, R. Peng, J. Wang and Z. Liu, *Small*, 2013, **9**, 1989–1997.
472 F. Liu, K. S. Choi, T. J. Park, S. Y. Lee and T. S. Seo, *BioChip J.*, 2011, **5**, 123–128.
473 Z. Liu, J. T. Robinson, X. Sun and H. Dai, *J. Am. Chem. Soc.*, 2008, **130**, 10876–10877.
474 X. Sun, Z. Liu, K. Welsher, J. T. Robinson, A. Goodwin, S. Zaric and H. Dai, *Nano Res.*, 2008, **1**, 203–212.
475 D. Pantarotto, R. Singh, D. McCarthy, M. Erhardt, J.-P. Briand, M. Prato, K. Kostarelos and A. Bianco, *Angew. Chem., Int. Ed.*, 2004, **43**, 5242–5246.
476 R. Singh, D. Pantarotto, D. McCarthy, O. Chaloin, J. Hoebeke, C. D. Partidos, J. P. Briand, M. Prato, A. Bianco and K. Kostarelos, *J. Am. Chem. Soc.*, 2005, **127**, 4388–4396.
477 Z. Wang, H. Shen, S. Song, L. Zhang, W. Chen, J. Dai and Z. Zhang, *J. Nanosci. Nanotechnol.*, 2017, **18**, 2286–2293.
478 M. Vincent, I. De Lázaro and K. Kostarelos, *Gene Ther.*, 2017, **24**, 123–132.
479 J. Chen, Q. Wang, J. Zhou, W. Deng, Q. Yu, X. Cao, J. Wang, F. Shao, Y. Li, P. Ma, M. Spector, J. Yu and X. Xu, *Nanoscale*, 2017, **9**, 10820–10831.
480 H. N. Abdelhamid, M. Dowaidar and Ü. Langel, *Microporous Mesoporous Mater.*, 2020, **302**, 110200.
481 M. Xia, D. Luo, J. Dong, M. Zheng, G. Meng, J. Wu and J. Wei, *J. Exp. Clin. Cancer Res.*, 2019, **38**, 408.
482 W. Meng, A. Rey-Rico, M. Claudel, G. Schmitt, S. Speicher-Mentges, F. Pons, L. Lebeau, J. K. Venkatesan and M. Cucchiarini, *Nanomaterials*, 2020, **10**, 855.
483 J. Fan, M. Claudel, C. Ronzani, Y. Arezki, L. Lebeau and F. Pons, *Int. J. Pharm.*, 2019, **569**, 118521.
484 F. Scherer, M. Anton, U. Schillinger, J. Henke, C. Bergemann, A. Krüger, B. Gänsbacher and C. Plank, *Gene Ther.*, 2002, **9**, 102–109.
485 S. Huth, J. Lausier, S. W. Gersting, C. Rudolph, C. Plank, U. Welsch and J. Rosenecker, *J. Gene Med.*, 2004, **6**, 923–936.
486 H. Mok and M. Zhang, *Expert Opin. Drug Delivery*, 2013, **10**, 73–87.
487 C. Sapet, C. Pellegrino, N. Laurent, F. Sicard and O. Zelphati, *Pharm. Res.*, 2012, **29**, 1203–1218.
488 P. G. Kyrtatos, P. Lehtolainen, M. Junemann-Ramirez, A. Garcia-Prieto, A. N. Price, J. F. Martin, D. G. Gadian, Q. A. Pankhurst and M. F. Lythgoe, *JACC Cardiovasc. Interv.*, 2009, **2**, 794–802.
489 S. J. Tseng, K. Y. Huang, I. M. Kempson, S. H. Kao, M. C. Liu, S. C. Yang, Z. X. Liao and P. C. Yang, *ACS Nano*, 2016, **10**, 10339–10346.
490 S. M. Shalaby, M. K. Khater, A. M. Perucho, S. A. Mohamed, I. Helwa, A. Laknaur, I. Lebedyeva, Y. Liu, M. P. Diamond and A. A. Al-Hendy, *Fertil. Steril.*, 2016, **105**, 1638–1648.
491 N. Tresilwised, P. Pithayanukul, P. S. Holm, U. Schillinger, C. Plank and O. Mykhaylyk, *Biomaterials*, 2012, **33**, 256–269.
492 J. W. Choi, J. W. Park, Y. Na, S. J. Jung, J. K. Hwang, D. Choi, K. G. Lee and C. O. Yun, *Biomaterials*, 2015, **65**, 163–174.
493 S. Castellani, C. Orlando, A. Carbone, S. Di Gioia and M. Conese, *Genes*, 2016, **7**, 103.
494 S. I. Kadota, T. Kanayama, N. Miyajima, K. Takeuchi and K. Nagata, *J. Virol. Methods*, 2005, **128**, 61–66.

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

495 T. N. Britos, C. E. Castro, B. M. Bertassoli, G. Petri, F. L. A. Fonseca, F. F. Ferreira and P. S. Haddad, *Mater. Sci. Eng., C*, 2019, **99**, 171–179.
496 A. A. Belanova, N. Gavalas, Y. M. Makarenko, M. M. Belousova, A. V. Soldatov and P. V. Zolotukhin, *Oncol. Res. Treat.*, 2018, **41**, 139–143.
497 S. R. Bhattarai, S. Y. Kim, K. Y. Jang, K. C. Lee, H. K. Yi, D. Y. Lee, H. Y. Kim and P. H. Hwang, *Nanomedicine*, 2008, **4**, 146–154.
498 J. Yun, A. M. Sonabend, I. V. Ulasov, D. H. Kim, E. A. Rozhkova, V. Novosad, S. Dashnaw, T. Brown, P. Canoll, J. N. Bruce and M. S. Lesniak, *J. Clin. Neurosci.*, 2012, **19**, 875–880.
499 J. W. Park, K. H. Bae, C. Kim and T. G. Park, *Biomacromolecules*, 2011, **12**, 457–465.
500 B. Kurena, A. Vežāne, D. Skrastiņa, O. Trofimova and A. Zajakina, *J. Virol. Methods*, 2017, **245**, 28–34.
501 R. Schubert, S. Trenholm, K. Balint, G. Kosche, C. S. Cowan, M. A. Mohr, M. Munz, D. Martinez-Martin, G. Fläschner, R. Newton, J. Krol, B. G. Scherf, K. Yonehara, A. Wertz, A. Ponti, A. Ghanem, D. Hillier, K. K. Conzelmann, D. J. Müller and B. Roska, *Nat. Biotechnol.*, 2018, **36**, 81–88.
502 Z. X. Liao, I. M. Kempson, Y. C. Fa, M. C. Liu, L. C. Hsieh, K. Y. Huang and L. F. Wang, *Bioconjugate Chem.*, 2017, **28**, 1702–1708.
503 S.-J. Tseng, I. M. Kempson, K.-Y. Huang, H.-J. Li, Y.-C. Fa, Y.-C. Ho, Z.-X. Liao and P.-C. Yang, *ACS Nano*, 2018, **12**, 9894–9902.
504 P. Reimer and T. Balzer, *Eur. Radiol.*, 2003, **13**, 1266–1276.
505 F. Scherer, M. Anton, U. Schillinger, J. Henke, C. Bergemann, A. Krüger, B. Gänsbacher and C. Plank, *Gene Ther.*, 2002, **9**, 102–109.
506 F. Wei, K. I. McConnell, T. K. Yu and J. Suh, *Eur. J. Pharm. Sci.*, 2012, **46**, 167–172.
507 I. E. Alexander, D. W. Russell and A. Dusty Miller, *J. Virol.*, 1994, **68**, 8282–8287.
508 D. W. Russell, I. E. Alexander and A. D. Miller, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 5719–5723.
509 L. N. Zhang, P. Karp, C. J. Gerard, E. Pastor, D. Laux, K. Munson, Z. Yan, X. Liu, S. Godwin, C. P. Thomas, J. Zabner, H. Shi, C. W. Caldwell, R. Peluso, B. Carter and J. F. Engelhardt, *Mol. Ther.*, 2004, **10**, 990–1002.
510 T. Zhang, J. Hu, W. Ding and X. Wang, *Neurochem. Int.*, 2009, **55**, 521–528.
511 F. R. Santoni De Sio, P. Cascio, A. Zingale, M. Gasparini and L. Naldini, *Blood*, 2006, **107**, 4257–4265.
512 V. Leuci, G. Mesiano, L. Gammaitoni, C. Cammarata, S. Capellero, M. Todorovic, N. Jordaney, P. Circosta, A. Elia, M. Lesnikova, G. E. Georges, W. Piacibello, F. Fagioli, A. Cignetti, M. Aglietta and D. Sangiolo, *J. Biotechnol.*, 2011, **156**, 218–226.
513 M. Basler, C. Lauer, U. Beck and M. Groettrup, *J. Immunol.*, 2009, **183**, 6145–6150.
514 D. Duan, Y. Yue, Z. Yan, J. Yang and J. F. Engelhardt, *J. Clin. Invest.*, 2000, **105**, 1573–1587.
515 B. M. Davis, L. Humeau, V. Slepushkin, G. Binder, L. Korshalla, Y. Ni, E. Oluwakemi Ogunjimi, L. F. Chang, X. Lu and B. Dropulic, *Blood*, 2004, **104**, 364–373.
516 J. M. Johnston, G. Denning, R. Moot, D. Whitehead, J. Shields, J. M. Le Doux, C. B. Doering and H. T. Spencer, *Gene Ther.*, 2014, **21**, 1008–1020.
517 S. C. Nicolson, C. Li, M. L. Hirsch, V. Setola and R. J. Samulski, *J. Virol.*, 2016, **90**, 7019–7031.
518 M. H. Dornan, R. Krishnan, A. M. Macklin, M. Selman, N. El Sayes, H. H. Son, C. Davis, A. Chen, K. Keillor, P. J. Le, C. Moi, P. Ou, C. Pardin, C. R. Canez, F. Le Boeuf, J. C. Bell, J. C. Smith, J.-S. Diallo and C. N. Boddy, *Sci. Rep.*, 2016, **6**, 26786.
519 J. S. Diallo, F. Le Boeuf, F. Lai, J. Cox, M. Vaha-Koskela, H. Abdelbary, H. MacTavish, K. Waite, T. Falls, J. Wang, R. Brown, J. E. Blanchard, E. D. Brown, D. H. Kirn, J. Hiscott, H. Atkins, B. D. Lichty and J. C. Bell, *Mol. Ther.*, 2010, **18**, 1123–1129.
520 E. Zonari, G. Desantis, C. Petrillo, F. E. Boccalatte, M. R. Lidonnici, A. Kajaste-Rudnitski, A. Aiuti, G. Ferrari, L. Naldini and B. Gentner, *Stem Cell Rep.*, 2017, **8**, 977–990.
521 K. E. Masiuk, R. Zhang, K. Osborne, R. P. Hollis, B. Campo-Fernandez and D. B. Kohn, *Mol. Ther.–Methods Clin. Dev.*, 2019, **13**, 390–398.
522 G. C. Heffner, M. Bonner, L. Christiansen, F. J. Pierciey, D. Campbell, Y. Smurnyy, W. Zhang, A. Hamel, S. Shaw, G. Lewis, K. A. Goss, O. Garijo, B. E. Torbett, H. Horton, M. H. Finer, P. D. Gregory and G. Veres, *Mol. Ther.*, 2018, **26**, 320–328.
523 L. C. Chandler, A. R. Barnard, S. L. Caddy, M. I. Patrício, M. E. McClements, H. Fu, C. Rada, R. E. MacLaren and K. Xue, *Mol. Ther.–Methods Clin. Dev.*, 2019, **14**, 77–89.
524 B. Lv, J. Li, M. Li, Y. Zhuo, K. Ren, E. Li and G. Yang, *Sci. Rep.*, 2018, **8**, 1–10.
525 P. Wang, Y. Wu, C. Yang, G. Zhao, Y. Liu, G. Cheng and S. Wang, *OncoTargets Ther.*, 2020, **13**, 1421–1429.
526 Y. Gong, R. G. J. Klein Wolterink, I. Janssen, A. J. Groot, G. M. J. Bos and W. T. V. Germeraad, *Mol. Ther.–Methods Clin. Dev.*, 2020, **17**, 634–646.
527 Z. Yan, R. Zak, Y. Zhang, W. Ding, S. Godwin, K. Munson, R. Peluso and J. F. Engelhardt, *J. Virol.*, 2004, **78**, 2863–2874.
528 L. S. Vermeer, L. Hamon, A. Schirer, M. Schoup, J. Cosette, S. Majdoul, D. Pastré, D. Stockholm, N. Holic, P. Hellwig, A. Galy, D. Fenard and B. Bechinger, *Acta Biomater.*, 2017, **64**, 259–268.
529 A. Castro, M. Drosch, M. Liesenfeld, T. Kaan, A. Schilz, H. Wenschuh, W. Forsmann and K. Kuhlke, *Cytotherapy*, 2019, **21**, e14.
530 P. Lin, D. Correa, Y. Lin and A. I. Caplan, *PLoS One*, 2011, **6**, e23891.
531 M. Themis, S. Forbes, L. Chan, R. Cooper, C. Etheridge, A. Miller, H. Hodgson and C. Coutelle, *Gene Ther.*, 1998, **5**, 1180–1186.
532 P. F. Kelly, S. Radtke, C. von Kalle, B. Balcik, K. Bohn, R. Mueller, T. Schuesler, M. Haren, L. Reeves,

Open Access Article. Published on 29 September 2020. Downloaded on 7/31/2023 4:30:32 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

J. A. Cancelas, T. Leemhuis, R. Harris, A. D. Auerbach, F. O. Smith, S. M. Davies and D. A. Williams, *Mol. Ther.*, 2007, **15**, 211–219.

533 A. Aiuti, *Science*, 2002, **296**, 2410–2413.

534 K. L. Shaw, E. Garabedian, S. Mishra, P. Barman, A. Davila, D. Carbonaro, S. Shupien, C. Silvin, S. Geiger, B. Nowicki, E. M. Smogorzewska, B. Brown, X. Wang, S. de Oliveira, Y. Choi, A. Ikeda, D. Terrazas, P.-Y. Fu, A. Yu, B. C. Fernandez, A. R. Cooper, B. Engel, G. Podsakoff, A. Balamurugan, S. Anderson, L. Muul, G. J. Jagadeesh, N. Kapoor, J. Tse, T. B. Moore, K. Purdy, R. Rishi, K. Mohan, S. Skoda-Smith, D. Buchbinder, R. S. Abraham, A. Scharenberg, O. O. Yang, K. Cornetta, D. Gjertson, M. Hershfield, R. Sokolic, F. Candotti and D. B. Kohn, *J. Clin. Invest.*, 2017, **127**, 1689–1699.

535 S. S. De Ravin, S. Anaya O'Brien, N. Kwatemaa, N. Theobald, S. Liu, J. Lee, L. Kardava, T. Liu, F. Goldman, S. Moir, J. Bleesing, B. Neven, J. Puck, M. J. Cowan, E. Mamcarz, S. Gottschalk, M. M. Meagher, L. Notarangelo, E. Kang, X. Wu and H. L. Malech, *Blood*, 2019, **134**, 608–608.

536 R. E. MacLaren, M. Groppe, A. R. Barnard, C. L. Cottriall, T. Tolmachova, L. Seymour, K. R. Clark, M. J. During, F. P. M. Cremers, G. C. M. Black, A. J. Lotery, S. M. Downes, A. R. Webster and M. C. Seabra, *Lancet*, 2014, **383**, 1129–1137.

537 A. M. Maguire, F. Simonelli, E. A. Pierce, E. N. Pugh, F. Mingozzi, J. Bennicelli, S. Banfi, K. A. Marshall, F. Testa, E. M. Surace, S. Rossi, A. Lyubarsky, V. R. Arruda, B. Konkle, E. Stone, J. Sun, J. Jacobs, L. Dell'Osso, R. Hertle, J. Ma, T. M. Redmond, X. Zhu, B. Hauck, O. Zelenaia, K. S. Shindler, M. G. Maguire, J. F. Wright, N. J. Volpe, J. W. McDonnell, A. Auricchio, K. A. High and J. Bennett, *N. Engl. J. Med.*, 2008, **358**, 2240–2248.

538 L. A. George, S. K. Sullivan, A. Giermasz, J. E. J. Rasko, B. J. Samelson-Jones, J. Ducore, A. Cuker, L. M. Sullivan, S. Majumdar, J. Teitel, C. E. McGuinn, M. V. Ragni, A. Y. Luk, D. Hui, J. F. Wright, Y. Chen, Y. Liu, K. Wachtel, A. Winters, S. Tiefenbacher, V. R. Arruda, J. C. M. van der Loo, O. Zelenaia, D. Takefman, M. E. Carr, L. B. Couto, X. M. Anguela and K. A. High, *N. Engl. J. Med.*, 2017, **377**, 2215–2227.

539 J. R. Mendell, Z. Sahenk, V. Malik, A. M. Gomez, K. M. Flanigan, L. P. Lowes, L. N. Alfano, K. Berry, E. Meadows, S. Lewis, L. Braun, K. Shontz, M. Rouhana, K. R. Clark, X. Q. Rosales, S. Al-Zaidy, A. Govoni, L. R. Rodino-Klapac, M. J. Hogan and B. K. Kaspar, *Mol. Ther.*, 2015, **23**, 192–201.

540 A. A. Thompson, M. C. Walters, J. Kwiatkowski, J. E. J. Rasko, J.-A. Ribeil, S. Hongeng, E. Magrin, G. J. Schiller, E. Payen, M. Semeraro, D. Moshous, F. Lefrere, H. Puy, P. Bourget, A. Magnani, L. Caccavelli, J.-S. Diana, F. Suarez, F. Monpoux, V. Brousse, C. Poirot, C. Brouzes, J.-F. Meritet, C. Pondarré, Y. Beuzard, S. Chrétien, T. Lefebvre, D. T. Teachey, U. Anurathapan, P. J. Ho, C. von Kalle, M. Kletzel, E. Vichinsky, S. Soni, G. Veres, O. Negre, R. W. Ross, D. Davidson, A. Petrusich, L. Sandler, M. Asmal, O. Hermine, M. De Montalembert, S. Hacein-Bey-Abina, S. Blanche, P. Leboulch and M. Cavazzana, *N. Engl. J. Med.*, 2018, **378**, 1479–1493.

541 European Medicines Agency Committee for Medicinal Products for Human Use, *Assessment Report Zynteglo*, 2019.

542 AveXis, Package Insert Zolgensma, https://www.fda.gov/media/126109/download, (accessed 13 September 2020).

543 Spark Therapeutics, Package Insert LUXTURNA, https://www.fda.gov/media/109906/download, (accessed 13 September 2020).

544 Kite Pharma, Package Insert YESCARTA, https://www.fda.gov/media/108377/download, (accessed 13 September 2020).

545 Novartis, Package Insert KYMRIAH, https://www.fda.gov/media/107296/download, (accessed 13 September 2020).

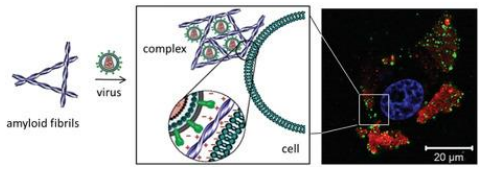
 This journal is © The Royal Society of Chemistry 2020

6.2 Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction

Authors: Stefanie Sieste[#], Thomas Mack[#], Edina Lump, Manuel Hayn, Desiree Schütz, Annika Röcker, Christoph Meier, Kübra Kaygisiz, Frank Kirchhoff, Tuomas P. J. Knowles, Francesco S. Ruggeri,* Christopher V. Synatschke,* Jan Münch,* Tanja Weil*
[#] shared first authorship * corresponding author
Published in: Advanced Functional Materials **2021**, *31* (17), 2009382.
DOI: 10.1002/adfm.202009382
Date of Publication: 22nd of February 2021
Author Contributions:
Stefanie Sieste designed, synthesized and characterized peptide nanofibrils, conducted confocal microscopy, wrote and revised the manuscript. Thomas Mack designed, synthesized and characterized peptide nanofibrils, wrote and revised the manuscript. Edina Lump and Manuel Hayn conducted infection assays. Desiree Schütz conducted confocal microscopy. Annika Röcker conducted cell viability studies. Christoph Meier discussed the concept and results. Kübra Kaygisiz conducted time-dependent zeta potential measurement and revised the manuscript. Frank Kirchhoff acquired funding, designed and discussed the concept and results, and revised the manuscript. Tuomas P. J. Knowles wrote and revised the manuscript. Francesco S. Ruggeri designed and revised the concept and results, wrote and corrected the manuscript. Christopher V. Synatschke designed and discussed the concept and results, wrote and revised the manuscript. Jan Münch acquired funding for the project, designed and discussed the concept and results, wrote and revised the manuscript. Tanja Weil acquired funding for the project, discussed the concept and results, wrote and revised the manuscript.

Copyrights: The following part is published in *Adv. Funct. Mater.* **2021**, *31* (17), 2009382 https://doi.org/10.1002/adfm.202009382. and licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/). Further distribution of the material is possible if appropriate credit, a link provided to the license, and indications of changes are made.

Graphical abstract:



Enhancing factor-C (EF-C), a fragment of the human immunodeficiency virus-1 glycoprotein, is previously shown to self-assemble into nanofibrils and enhance viral transduction. Herein, a peptide library based on EF-C is synthesized through multiparameter and multiscale optimization. It is evaluated for fibril formation, secondary/quaternary structure, virion-fibril-cell interaction, and critical parameters for enhanced transduction can be determined in a structure–activity fashion.

Check for updates

FULL PAPER

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction

Stefanie Sieste, Thomas Mack, Edina Lump, Manuel Hayn, Desiree Schütz, Annika Röcker, Christoph Meier, Kübra Kaygisiz, Frank Kirchhoff, Tuomas P. J. Knowles, Francesco S. Ruggeri, Christopher V. Synatschke,* Jan Münch,* and Tanja Weil**

Amyloid-like peptide nanofibrils (PNFs) are abundant in nature providing rich bioactivities and playing both functional and pathological roles. The structural features responsible for their unique bioactivities are, however, still elusive. Supramolecular nanostructures are notoriously challenging to optimize, as sequence changes affect self-assembly, fibril morphologies, and biorecognition. Herein, the first sequence optimization of PNFs, derived from the peptide enhancing factor-C (EF-C, QCKIKQIINMWQ), for enhanced retroviral gene transduction via a multiparameter and a multiscale approach is reported. Retroviral gene transfer is the method of choice for the stable delivery of genetic information into cells offering great perspectives for the treatment of genetic disorders. Single fibril imaging, zeta potential, vibrational spectroscopy, and quantitative retroviral transduction assays provide the structure parameters responsible for PNF assembly, fibrils morphology, secondary and quaternary structure, and PNF-virus-cell interactions. Optimized peptide sequences such as the 7-mer, CKFKFQF, have been obtained quantitatively forming supramolecular nanofibrils with high intermolecular β-sheet content that efficiently bind virions and attach to cellular membranes revealing efficient retroviral gene transfer.

1. Introduction

Peptide nanofibrils (PNFs) are ubiquitous in nature. They are formed by spontaneous self-assembly and exert many vital functions in living organisms. Peptide nanofibrils consisting of regular cross-β sheet structures, often termed amyloid-like structures,[1–4] feature many outstanding properties such as high rigidity, tensile strength, and enhanced resistance towards biodegradation.[5,6] They play both functional and pathological roles in many living organisms: Serving as templates for melanin biosynthesis,[7] for hormone storage,[8] participating in spermatozoa selection and clearance,[9] or constituting the fibrillar curli structures of *Escherichia coli* promoting their surface adhesion.[10] Their formation has also been associated with numerous human diseases (around 50), such as Alzheimer's disease or Diabetes

S. Sieste, T. Mack, K. Kaygisiz, Dr. C. V. Synatschke, Prof. T. Weil
Department Synthesis of Macromolecules
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
E-mail: synatschke@mpip-mainz.mpg.de; weil@mpip-mainz.mpg.de
S. Sieste, T. Mack, Prof. T. Weil
Institute of Inorganic Chemistry I
Ulm University
Albert-Einstein-Allee 11, 89081 Ulm, Germany
Dr. E. Lump, M. Hayn, D. Schütz, Dr. A. Röcker,
Prof. F. Kirchhoff, Prof. J. Münch
Institute of Molecular Virology
Ulm University Medical Center
Meyerhofstraße 1, Ulm 89081, Germany
E-mail: jan.muench@uni-ulm.de
The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adfm.202009382.

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adfm.202009382

Dr. C. Meier
Institute of Organic Chemistry III/Macromolecular Chemistry
Ulm University
Albert-Einstein-Allee 11, 89081 Ulm, Germany
Prof. T. P. J. Knowles, Prof. F. S. Ruggeri
Department of Chemistry
University of Cambridge
Cambridge CB2 1EW, UK
E-mail: simone.ruggeri@wur.nl
Prof. T. P. J. Knowles
Cavendish Laboratory
University of Cambridge
Cambridge CB3 0HE, UK
Prof. F. S. Ruggeri
Laboratory of Organic and Physical Chemistry
Wageningen University
Stippeneng 4, Wageningen 6703 WE, The Netherlands
Prof. J. Münch
Core Facility of Functional Peptidomics
Ulm University Medical Center
Albert-Einstein-Allee 47, 89081 Ulm, Germany

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

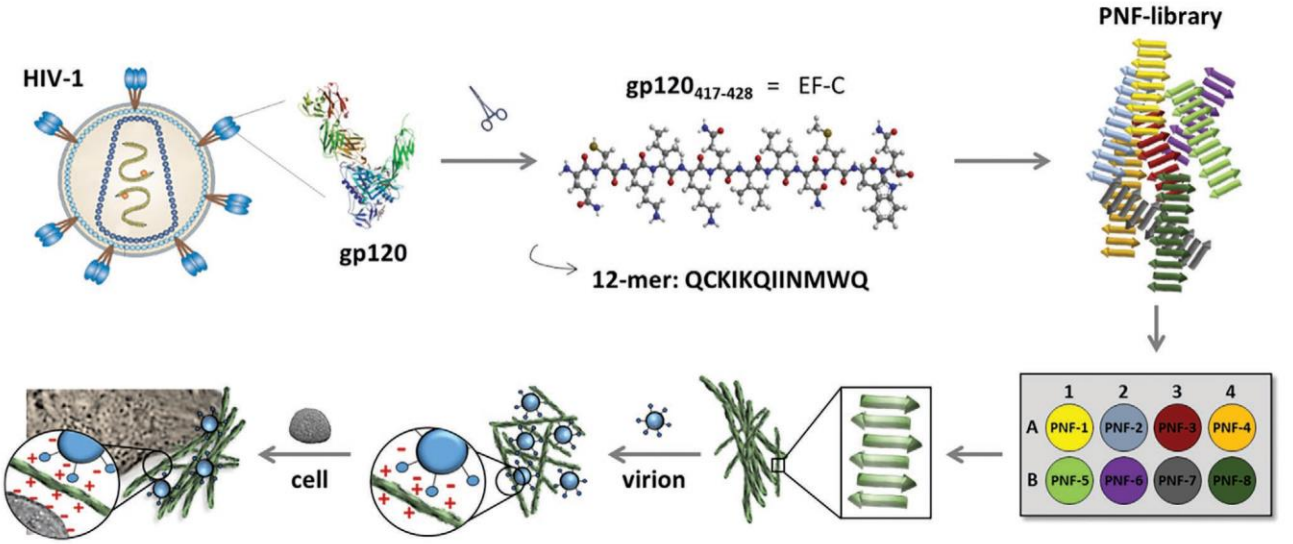


Figure 1. Schematic illustration of the structure–activity study of transduction enhancing PNFs. EF-C is a bioactive fragment of the HIV-1 glycoprotein gp120 that forms PNFs and enhances viral transduction. Through multiparameter and multiscale optimization, a library of EF-C derived peptides was synthesized. The library was evaluated for fibril formation, structural parameters, virion-PNF-cell interactions, and viral transduction enhancement.

mellitus type 2.[11] We have previously shown that PNF formation is also common in semen from healthy human individuals because semen contains an abundance of amyloidogenic fragments of prostate acidic phosphatase and the semenogelins 1 and 2.[12,13] PNFs likely play important roles in innate immunity and in eliminating poor-quality sperm upon deposition in the female reproductive tract.[9] In addition, semen PNFs enhance the infectivity of human immunodeficiency virus-1 (HIV-1), herpesviruses 1 and 2 and Ebola virus, and may increase the transmission potential of these sexually transmitted pathogens.[14,15]

The rich bioactivity of endogenous amyloids has inspired scientists to elucidate structural features crucial for their broad functions. Moreover, understanding their specific interactions of PNF with cellular membranes and viruses is not only of great fundamental interest but also essential to design new nanomaterials for medicine, for example, to enhance retroviral gene transfer.[16–18] Retroviral gene transfer is the method of choice for the delivery of genetic material into cells offering great perspectives for the treatment of genetic disorders, malignancies, and infectious diseases.[19,20] Until now, γretroviral and lentiviral vectors (herein both referred to as retroviral vectors) have been used in more than 478 gene therapy clinical trials[21] allowing stable, long-lasting gene expression. HIV-1 based lentiviral vectors even allow transduction of non-dividing cells,[22,23] which is essential to treat, that is, metachromatic leukodystrophy or inherited primary immunodeficiencies with a significant clinical benefit.[24,25] However, major challenges in retroviral gene delivery need to be overcome such as low transduction efficiencies due to: i) low concentrations of transducing particles in the viral vector stock; and ii) low efficiency of virion attachment to the target cell. Nowadays, retroviral gene delivery is often promoted by a combination of spin-infection and utilization of RetroNectin, a long polypeptide, consisting of 574 amino acid and a fibronectin derivative that colocalizes with virions and cells.[26–29] The large scale use of RetroNectin is limited by its high costs and even more importantly, it has to be coated onto plates or the inner surface of culture bags to capture virions, requiring several washing steps[30–32] which impedes in vivo applications.[33]

PNFs have been reported as powerful alternatives to RetroNectin as they enhance retroviral gene transfer with a similar efficiency but do not require coating.[18,34,35] We discovered the self-assembling 12-mer peptide EF-C (QCKIKQIINMWQ) corresponding to residues 417–428 of the HIV envelope protein gp120 (**Figure 1**).[18] EF-C PNFs efficiently concentrate virus stock solutions and increase cell attachment as well as retroviral gene transfer into hematopoietic stem cells and T cells by different types of vectors in vitro and ex vivo, which holds various opportunities for gene therapy approaches, in particular chimeric antigen receptor T cell (CAR-T) immunotherapy.[18,34–36] Compared to other PNFs and known transduction enhancers, EF-C reveals one of the highest enhancing effects on retroviral gene transfer and has been commercialized under the brand name Protransduzin.[34]

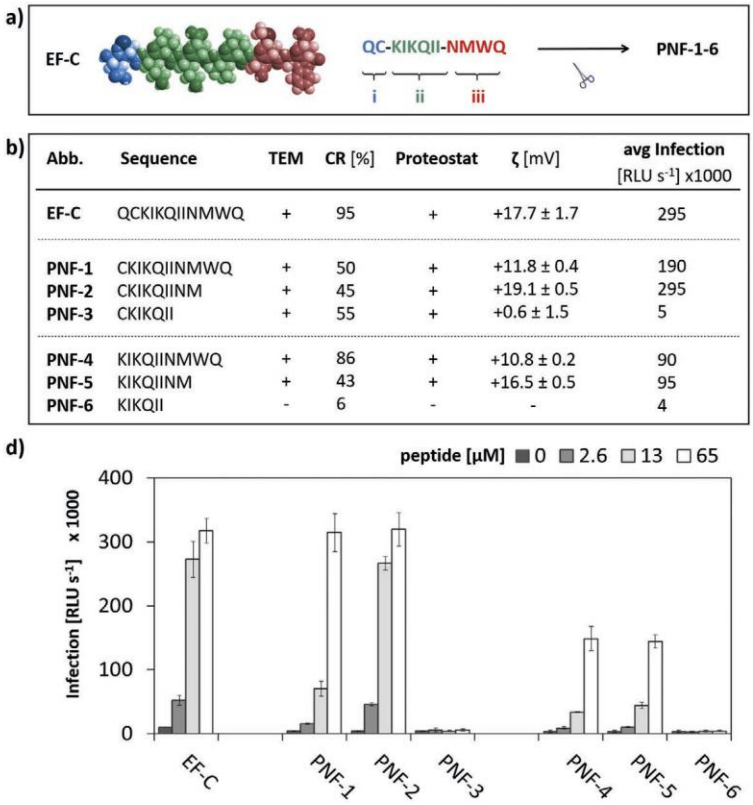
Very little is known about the structural features of bioactive PNFs such as EF-C, which severely limits rational approaches to further optimize their bioactivity. To the best of our knowledge, no structure–activity study of supramolecular nanofibrils has been reported yet. This limitation is due to the fact that supramolecular nanostructures are notoriously challenging to optimize as multiple parameters affect self-assembly, fibril morphologies, and biorecognition and the involved structures cover several length scales from the molecular to the mesoscopic regime. A limited number of studies have investigated small sequence variations in amyloid PNFs on transduction enhancement.[37–40] However, rational approaches that unravel structure–morphology–bioactivity relationships of PNF are still elusive. Herein, we report the first sequence optimization of PNFs for enhanced retroviral gene transduction via a rational multiparameter-based approach. The amino acids and peptide motifs within PNFs responsible for self-assembly, fibrils

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de



b)

Abb.	Sequence	TEM	CR [%]	Proteostat	ζ [mV]	avg Infection [RLU s^{-1}] x1000
EF-C	QCKIKQIINMWQ	+	95	+	+17.7 ± 1.7	295
PNF-1	CKIKQIINMWQ	+	50	+	+11.8 ± 0.4	190
PNF-2	CKIKQIINM	+	45	+	+19.1 ± 0.5	295
PNF-3	CKIKQII	+	55	+	+0.6 ± 1.5	5
PNF-4	KIKQIINMWQ	+	86	+	+10.8 ± 0.2	90
PNF-5	KIKQIINM	+	43	+	+16.5 ± 0.5	95
PNF-6	KIKQII	-	6	-	-	4

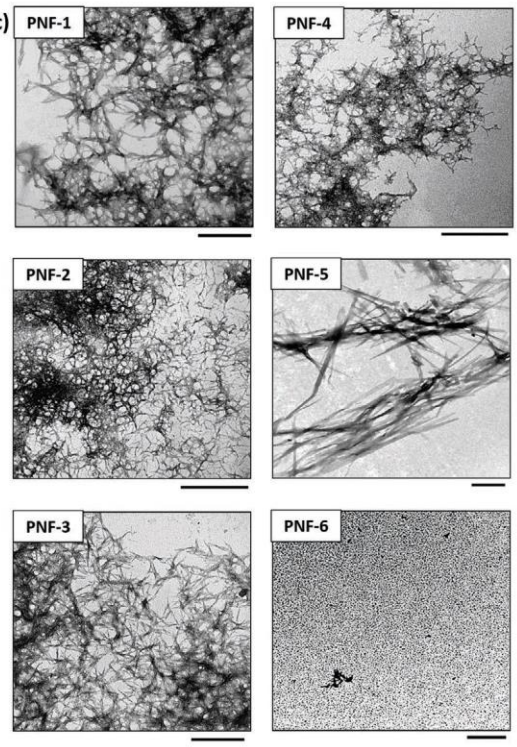


Figure 2. PNF library generation upon sequence truncation of EF-C. a) EF-C was subdivided into three domains: i) QC; ii) KIKQII; and iii) NMWQ. b) Impact of sequence variations on peptide self-assembly (with + = fibrils and − = no aggregate formation detected), monomer-to-fibril conversion rates (CR) and transduction enhancement. Proteostat emission indicates accessible cross β-sheet binding pockets in the fibril core and indicated PNFs with + = Proteostat activity and − = no Proteostat fluorescence observed. ζ-potential values for PNF solutions. Average (avg) RLU s^{-1} in infectivity assay for 13–65 µM. c) Representative TEM micrographs of PNF-1 to PNF-5 reveal the formation of fibrillar aggregates. No aggregates were observed for PNF-6. The scale bars represent 500 nm. d) HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Shown are mean values derived from triplicate infections ± standard deviation. RLU s^{-1}; relative light units per second.

morphology, secondary and quaternary structure, and virus and/or cell binding were identified. We believe that unravelling the structural parameters of PNFs responsible for their unique bioactivities paves the way to a fundamental understanding of the rich bioactivity of PNFs and ultimately addresses the great need for improved nanomaterials for clinical applications, for example, in gene therapy.

EF-C served as a starting point for multiparameter optimization. Its sequence was first divided into smaller fragments to identify amino acids and amino acid patterns responsible for assembly and interaction with virions or cellular membranes. Subsequently, several focused peptide libraries were synthesized as depicted in **Figure 2**a. EF-C features: i) a short N-terminal *QC* fragment; ii) an amphiphilic *KIKQII* domain; and iii) a C-terminal *NMWQ* sequence. The C-terminal *NMWQ* domain represents a fragment of the β20/21 strand in the gp120 HIV-1 envelope protein region that has been proposed to constitute the primary HIV binding site for the CD4 receptors of immune cells[41] and to play a crucial role for the transduction enhancement of PNFs. The hexameric *KIKQII* fragment features an amphiphilic pattern with three hydrophobic isoleucine (I) and two positively charged lysine (K) residues. The lipophilic (I) residues could promote self-assembly, whereas the hydrophilic and positively charged (K) residues provide solubility for stabilizing the superstructures in aqueous media.[15,17]

PNF formation was achieved by dilution from a dimethyl sulfoxide stock solution to PBS buffer and aggregation was confirmed by transmission electron microscopy (TEM). Conversion rates (CR) from the soluble peptide monomers to supramolecular PNFs were quantified as published previously.[42] ζ-potential measurements were performed to determine the charge of PNF assemblies; secondary and quaternary structures of PNFs were characterized by infrared spectroscopy and intermolecular β-sheet structures and were confirmed by the blue-shifted emission of Proteostat[43] (Figure S1, Table S1, Supporting Information). The transduction enhancing properties of the PNFs were evaluated in infection experiments using HIV-1 particles as prototype retroviral vectors. Virions were exposed to increasing concentrations of PNFs and the mixtures were used to infect TZM-bl reporter cells,[44] which are widely used to study HIV infection rates. These cells are genetically engineered to express the HIV receptors CD4 and CXCR4/CCR5 and contain a stably integrated β-galactosidase gene under the control of the viral long terminal repeat promotor region. Upon successful infection, the HIV-1 transactivator Tat protein is expressed and turns on the viral promotor resulting in β-galactosidase production, which

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

was quantified in a chemiluminescence-based assay performed 72 h post infection. None of the peptides tested in this study affected cell viability within the concentrations used for transduction studies (Figure S2, Supporting Information).

2. Results

2.1. PNF Library Generation upon Sequence Truncation

EF-C instantaneously forms positively charged PNFs (ζ = +17.7 mV) with high conversion rates (CR: 95%) and it is used as an internal standard to compare the infectivity enhancement of the peptide nanofibrils in different sets of experiments, that is, on different days. Several studies have shown that even short peptide sequences can exhibit robust fibril formation.[45,46] We therefore envisioned the identification of EF-C derived peptides with fewer than 13 amino acids that retain high transduction enhancement. After deleting the N-terminal glutamine (Q) and stepwise truncation of the C-terminal residues, PNF-1 (CKIKQIINMWQ), PNF-2 (CKIKQIINM), and PNF-3 (CKIKQII) still formed fibrils but with moderate CRs (45–55%) independent of the length of the C-terminal NMWQ fragment (Figure 2b,c). The transduction efficiencies of PNF-1 and PNF-2 with positive ζ-potentials were comparable to EF-C (at 65 μM, corresponding to 200 μg mL^{-1} EF-C, Figure 2d), whereas the almost neutral (ζ = + 0.6 mV) PNF-3 without the NMWQ fragment was inactive.

After deletion of the (QC) sequence, PNF-4 (KIKQIINMWQ) still assembled with high CR (86%) and PNF-5 (KIKQIINM) with moderate CR (43%, Figure 2b,c). A time-dependent measurement of the ζ-potential of PNF-4 and PNF-16 over 30 days revealed no significant changes in the respective ζ-potential values (Figure S3, Supporting Information). Interestingly, positively charged PNF-4 and PNF-5, which lack the cysteine residue, were considerably less active than PNF-1 and PNF-2 despite their positive ζ-potentials and comparable CRs. PNF-6 (KIKQII) did not form fibrils and showed no activity. These data gave a first indication that significant transduction enhancement requires: i) PNF formation; ii) positive ζ-potentials; and iii) an N-terminal cysteine residue.

2.2. Optimized Amphiphilic Pattern Improves Fibril Formation

In polar solvents, the morphology of intermolecular β-sheet rich aggregates is driven by the formation of a bilayer structure, where hydrophobic domains form the fibril core and the hydrophilic side chains arrange at the surface of the fibrillar backbone.[18] In order to improve the packing of the peptides in the PNFs, we altered: i) the arrangement of the hydrophilic and lipophilic amino-acids within the peptide sequences; ii) replaced the lipophilic (I) residues by other hydrophobic amino acids; and iii) identified minimal sequences with high CRs and transduction enhancement.[47,48] The central amphiphilic PNF-6 (KIKQII) fragment of the EF-C sequence reveals two structural defects in its sequence as the second (I) residue of KIKQII faces an opposite direction to the other two (I)-residues and the polar (Q) is located adjacent to the unpolar (I) and also opposite to the polar (K) residues (**Figure 3**a, blue model). By switching the positions of (Q) and its neighboring (I) in the QII fragment to an IQI sequence, a strictly alternating spatial arrangement of hydrophilic and lipophilic residues was obtained (Figure 3a, red model). In contrast to the PNF-1-6 series consisting of the KIKQII fragment, PNF-7 (CKIKIQINMWQ), PNF-8 (CKIKIQINM), and PNF-9 (CKIKIQI) now provide a facial amphiphilic character, where the positively charged and nonpolar residues are oriented not only in parallel but also perpendicular to the fibril axis. PNF-7 to PNF-9 with positive ζ-potentials (ζ = + 10.7–22.3 mV) were formed with moderate to high CRs (54–87%) and showed the ability to enhance retroviral infection at 65 μM, comparable to EF-C (Figure 3b–d; Figure S4, Supporting Information).

The corresponding (C)-free derivatives PNF-10 (KIKIQINMWQ) and PNF-11 (KIKIQINM) with the optimized KIKIQI patterns also assembled with high CRs (64–93%) into PNFs with positive ζ-potentials (ζ = + 17.2–41.9 mV) that were however less active than EF-C in promoting retroviral infection. Interestingly, PNF-9 with only seven amino acids assembled into bioactive and positively charged PNFs (ζ = +17.4 mV, CR = 54%), whereas neutral PNF-3 (ζ = +0.6 mV) was inactive and PNF-6 and PNF-12 lacking the N-terminal (C) did not form fibrils (Figures 2c and 3c). These results indicate that the optimized amphiphilic *CKIKIQI* pattern of PNF-9 with N-terminal (C) represents the shortest sequence for forming positively charged PNFs displaying transduction enhancement in the range of EF-C.

Next, the (I) residues in PNF-9 were replaced by the less bulky and less hydrophobic alanine (A) and the sterically more demanding aromatic phenylalanine (F) residues. PNF-13 (CKAKAQANMWQ), PNF-14 (CKAKAQANM), and PNF-15 (CKAKAQA) assembled into negatively charged PNFs (ζ = −3.7 to −9.1 mV) with high (PNF-13, 92%) to moderate CRs (PNF-14, PNF-15, 38–44%) that did not enhance viral transduction (Figure 3b–d). It has been reported that the aggregates of peptides with similar sequences can significantly vary in their ζ-potential, often preventing a direct correlation between the peptide sequence and ζ-potential.[49–51] The analogous lack in enhancing transduction was found after removing the N-terminal (C) of all three alanine derivatives resulting in PNF-S1 (KAKAQANMWQ), PNF-S2 (KAKAQANM), and PNF-S3 (KAKAQA) (Figure S5, Supporting Information). In contrast, the (F)-containing derivatives PNF-16 (CKFKFQFNMWQ), PNF-17 (CKFKFQFNM), and PNF-18 (CKFKFQF) formed positively charged PNFs (ζ = + 14.1–22.3 mV) with high CRs (88–96%) (Figure 3b–d) that all showed significant transduction enhancement. Removal of the N-terminal (C) afforded PNF-S4 (KFKFQFNMWQ) and PNF-S5 (KFKFQFNM) displaying high transduction enhancement, whereas PNF-S6 (KFKFQF) was inactive (Figure S5, Supporting Information).

2.3. An N-Terminal Cysteine Improves Fibril Formation and Enhances Transduction

The impact of the N-terminal accessible thiol-residues was furthermore investigated by substitution of (C) in PNF-9 by the sulfur-containing amino acid methionine (M) in PNF-S7 (MKIKIQI). Both PNF-9 and PNF-S7 formed positively charged

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

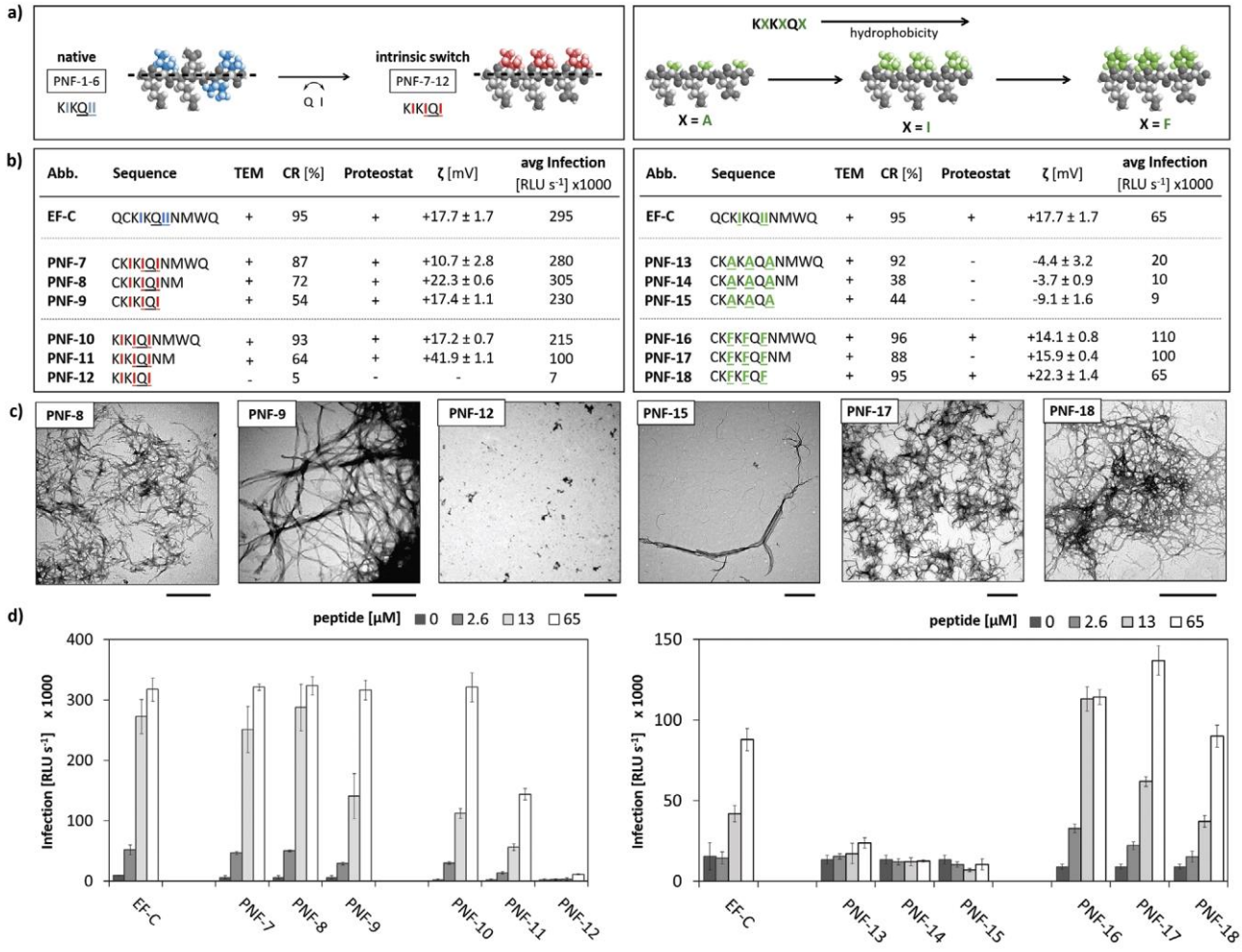


Figure 3. Improving PNF aggregating and transduction enhancement. a) Optimization of the amphiphilic fragment (II) of EF-C. Hydrophilic amino acids are given in grey and hydrophobic in color. b) Tabular summary of the features of PNF-7 to PNF-18 including PNF formation as observed in TEM measurements with + = fibrils and − = no aggregate formation detected. Monomer-to-fibril conversion rates (CR) were determined in a fluorescence-based assay. Proteostat emission indicates accessible cross β-sheet binding pockets in the fibril core and was assessed for PNFs with + = Proteostat activity and − = no Proteostat fluorescence observed. ζ-potential values for PNF solutions. Average (avg) RLU s^{-1} in infectivity assay for 13–65 µM. c) Representative TEM micrographs of PNF-8,-9,-15,-17,-18 reveal fibrillar aggregate morphologies. No aggregates were observed for PNF-12. Scale bars: 500 nm. d) Transduction enhancement assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF given as mean values from triplicate infections ± standard deviation. RLU s^{-1}; relative light units per second.

nanofibrils of comparable morphology as confirmed by TEM measurements and both derivatives assembled with similar CRs (PNF-9, 54% and PNF-S7, 51%) (Figure S6, Supporting Information). However, PNF-S7 was significantly less bioactive than PNF-9 at all tested concentrations. The substitution of (C) by (M) in the shortest bioactive sequence PNF-18 afforded PNF-S8 (MKFKFQF), which was inactive (Figure S6, Supporting Information). The cysteine groups in the bioactive peptides could in principle undergo oxidation reactions under air, which could enhance the formation of disulfide bridges.[52] Therefore, we have studied bioactivity of PNF-8 in the presence of the reducing agent TCEP. Our data shows that the application of TCEP has no significant influence on bioactivity (Figure S7, Supporting Information). We conclude that in our study, N-terminal (C) leads to robust fibril formation and enhances infectivity also in case of short peptide sequences.

2.4. Unraveling the Structure–Activity Relationship

We have shown that the formation of bioactive PNFs required moderate to high CR as well as an overall positive ζ-potential. However, PNFs with similar positive charges and CRs, such as PNF-2 and PNF-5 could still reveal significantly different transduction enhancing activities (Figures 2 and 3). Therefore, CR and positive ζ-potential represent critical features, but additional, yet unknown parameters exist that account for the observed variability in the biological activity of PNFs. To identify these parameters, we first correlated the sequence and transduction enhancement of PNFs with their secondary and quaternary structure that was obtained by FTIR spectroscopy (**Figure 4**) from freeze-dried PNF solutions. It has been shown that the properties of amyloid structures in a liquid and dried environment remain highly conserved.[53–55] In proteins, the IR spectra

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

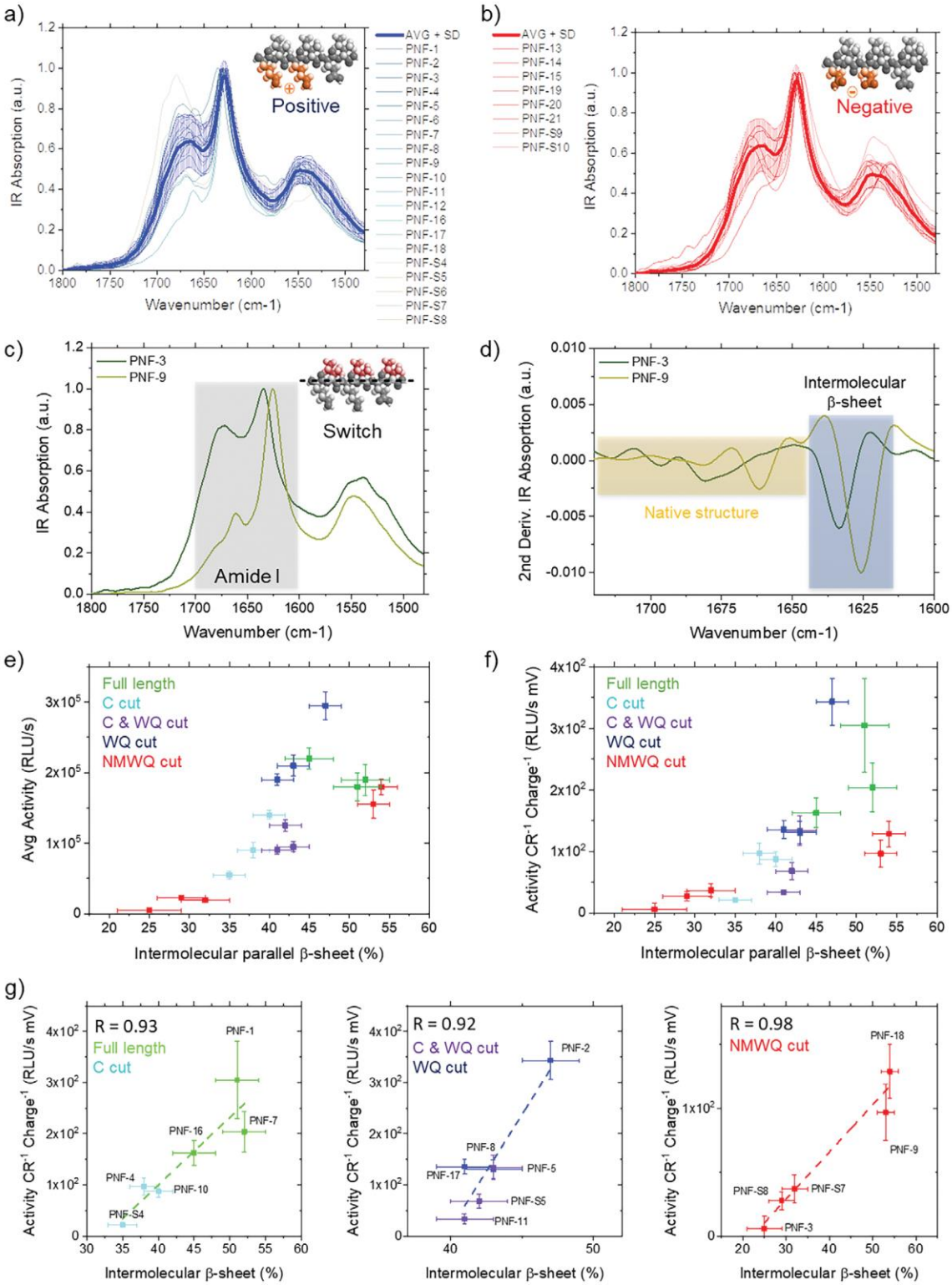


Figure 4. FTIR secondary structure evaluation and infectivity enhancement as a function of intermolecular β-sheet for all PNFs. FTIR spectra of a) positively and b) negatively charged PNFs. c) Modification of the amide band I upon intrinsic switch. d) Method of quantification of intramolecular native versus intermolecular aggregated β-sheet structure. Secondary structure versus e) average infection enhancing activity and f) average infection enhancing activity normalized by charge and conversion rate. g) Linear fit of activity versus intermolecular β-sheet for the groups of peptides with chemical similarity. Full length: PNF-1,-7,-16; C cut: PNF-4,-10,-S4; C and WQ cut: PNF-5,-11,-S5; WQ cut: PNF-2,-8,-17; NMWQ cut: PNF-3,-9,-18,-S7,-S8.

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are dominated by the amide band I, whose shape is intimately related to the secondary and quaternary structure of the protein backbone. A first qualitative view of the spectra shows that the amide band I of almost all PNFs exhibited a maximum in the range between 1615–1630 cm^{-1} (Figure 4a,b). The high absorption in this spectroscopic range demonstrates the presence of a high content of intermolecular β-sheet structures, stabilized by a tight network of hydrogen bonds, similar to amyloidogenic aggregates.[56–58] Then, the integration of the deconvolved amide band I by a second derivative analysis allowed robust quantification of each structural contribution (α-helix, random coil, native and aggregated β-sheet) (Figure 4c,d).[59,60] The percentage of native intramolecular secondary structural elements (1635–1700 cm^{-1}) was quantified and correlated versus the percentage of intermolecular hydrogen bonded β-sheet (1610–1635 cm^{-1}) for each PNF (Figure 4d).

The PNF library showed a wide variation of secondary structures as indicated by the heterogeneity of shapes of the amide band I of the PNFs (Figure 4a,b). The amide I band of the different negatively charged PNFs varied significantly (Figure 4b), and with similar standard deviation to the one of the positive PNFs (Figure 4a). However, there were no major differences in the secondary and quaternary structures of negatively charged PNFs that would explain the observed significant differences in their transduction enhancement compared to positively charged PNFs and the EF-C sequence. This observation is in line with previous reports indicating that the ζ-potential of the PNFs is the first factor determining their transduction enhancing features.[17,18,61] However, positively charged PNFs revealed high variations in their structure and ability to boost transduction. An initial qualitative investigation of the spectra of these PNFs indicated that sequence modifications altered their secondary and quaternary structure, which also had an impact on their transduction enhancing features. For instance, the intrinsic switch of the core (Q) and (I) of PNF-3 (CKIKQII) leads to the formation of PNF-9 (CKIKIQI) with a higher content of intermolecular β-sheet structures that displayed improved transduction enhancing features (Figure 4c,d). The corresponding spectra showed that both in the presence and absence of the N-terminal cysteine (C), the intrinsic switch KIKQII → KIKIQI allowed us to modulate with a higher degree of freedom the secondary structure of the PNFs upon the stepwise removal of WQ and NM amino acids (Figure S8, Supporting Information).

2.5. The Intermolecular β-Sheet Content of PNFs Modulates their Transduction Efficiency

Next, we assessed how the molecular structure of the positively charged PNFs affects their impact on transduction efficiency. To reduce the intrinsic variability in the biological assay, the average transduction enhancement was calculated (Table S2, Supporting Information) and plotted as average biological activity against the content of the measured intermolecular β-sheet structures (Figure 4e). These data showed an unexpected and significant correlation as the biological activity increased steadily as a function of the intermolecular β-sheet content of the PNFs. As both charges and the CR are important parameters contributing to transduction enhancement, we normalized our results against CR and ζ-potential. Normalization for CR considers only the fraction of peptide monomers participating in PNF formation, whereas normalization for charge recognizes that only positively charged PNFs facilitate virus transport into cells (Figure 4f; Figure S9, Supporting Information). Interestingly, even after CR and charge normalization, a direct correlation between infectivity enhancement and the presence of intermolecular β-sheet content of the PNFs was observed.

To unravel the functional relationship between secondary structure and bioactivity, we also considered that the side chain absorption of the amino acids W, Q, and N can significantly contribute to the amide band I of a short peptide (≈5–15%). Thus, we grouped the peptides into three groups that are either chemically identical, that is, in case of structural isomers, or consist of amino acids, whose side chains do not affect the amide band I absorption. The plot of the β-sheet content versus infectivity for each of these groups revealed that these physical quantities are correlated by a statistically significant linear relationship (Figure 4g, R = 0.92–0.98). These data clearly demonstrate that the content of the intermolecular β-sheet structures in the PNFs is a fundamental parameter modulating their infectivity (Figure 4g).

We observed above that (C) promotes enhanced transduction rates (Figure 2). The structural analysis revealed that deletion of C-terminal cysteine leads to a net average reduction of the intermolecular β-sheets content (≈15%) in the PNFs (Figure 4g, light blue PNF-4,-10,-S4 versus green PNF-1,-7,-16). Interestingly, the influence of the N-terminal (C) is not of a hydrophobic nature, as an exchange of (C) with the chemically similar (M) decreased the intermolecular β-sheets content and reduced infectivity in the corresponding peptides libraries (Figure 4g, red PNF-9, 18 to PNF-S7/-S8). Therefore, we found that the main role of the N-terminal (C) is to generate amyloid-like PNFs with a higher content of hydrogen bonded β-sheets and thus more stable PNFs with a higher degree of order,[45,62] which obviously enhances their infectivity.

Deletion of both the N-terminal (C) and the C-terminus had an impact on the content of intermolecular β-sheet, which also affected transduction. We observed that the N-terminal (C) or part of the C-terminus (NMWQ) is a necessary prerequisite for high infectivity of the EF-C sequence. Truncation of the C-terminal (WQ) sequence only slightly affected PNF infectivity (Figure 4f, green versus blue), and this minor difference in secondary structure was attributed to the absorption of the (WQ) side chains that were removed in the corresponding series PNF-2,-5,-8,-17,-S5. The removal of both (C) and C-terminal (WQ) sequence slightly reduced the intermolecular β-sheet content (≈5%), which also diminished infectivity enhancement (Figure 4g, blue PNF-4,-10,-S4 versus violet PNF-1,-7,-16), whereas deletion of (C) and the full C-terminus significantly reduced the intermolecular β-sheet content (≈20%), which resulted in an almost complete loss of the transduction enhancing features (Figure S10, Supporting Information).

As mentioned previously, the optimization of the amphiphilic pattern also resulted in significant changes in the secondary structure (Figure 4c,d). This effect was smaller for the sequences

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

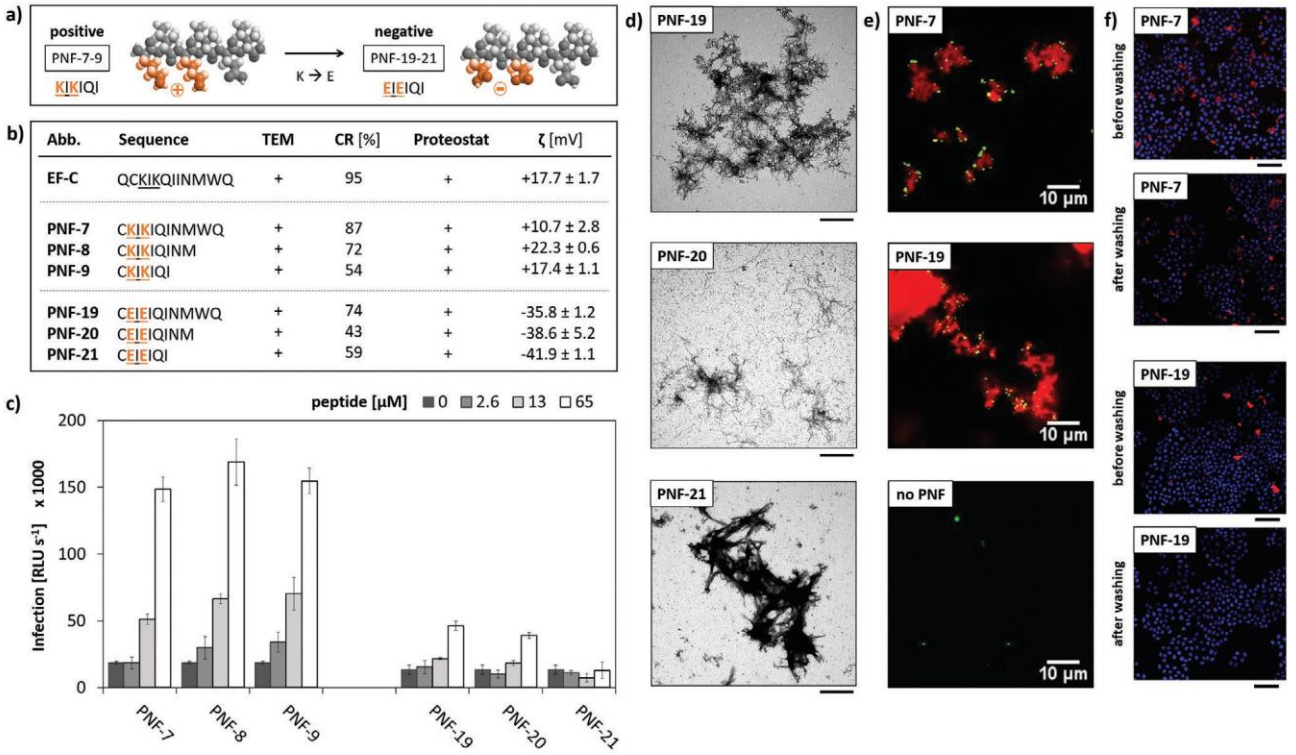


Figure 5. Variation of PNF charge from positive (bioactive) to negative (inactive). Both (K)-and (E)-rich fibrils bind virions, but the latter did not increase infectivity. a) Substitution of positively charged Lysine (K) residues by negatively charged glutamic acid (E) in the peptide sequences (PNF-7 to PNF-9) yielded corresponding negatively charged derivatives (PNF-19 to PNF-21) to elucidate the impact of PNF charge on self-assembly and infectivity enhancement. b) Tabular list of peptide derivatives PNF-19 to PNF-21 with glutamic acid residues and a positively charged lysine derivative set PNF-7 to PNF-9 including the ability to form PNFs as observed in TEM measurements with + = fibrils detected. Monomer-to-fibril conversion rates (CR) were determined in a fluorescence-based assay. High values indicate a large percentage of peptides participating in fibril formation. Proteostat emission indicates accessible cross β-sheet binding pockets in the fibril core and was assessed for PNFs with + = Proteostat activity and − = no Proteostat fluorescence observed. ζ-potential values for PNF solutions confirm the negative charge of fibrillar aggregates with glutamic acid residues. c) Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Values above the columns depict the change from baseline (0 μM). Shown are mean values derived from triplicate infections ± standard deviation. RLU s^{-1}; relative light units per second. d) Representative TEM micrographs of PNF-19 to PNF-21 reveal fibrillar aggregate morphologies. The scale bars represent 500 nm. e) Representative LSM micrographs of Proteostat-labelled PNFs (red) incubated with YFP-labelled viral particles (green) revealed formation of mesoscale virus-PNF clusters for PNF-7 and PNF-19. f) Representative LSM micrographs of Proteostat-labelled PNFs (red) and TZM-bl cells (nucleus in blue). PNFs were added to cells for 30 min and analyzed. After 1 h, cells were washed and again analyzed. PNF-7 revealed association with cellular membranes while PNF-19 did not associate with cells. The scale bars represent 100 μm.

close in length to EF-C and it became more pronounced for the highly truncated sequences, that is, those lacking the NMWQ fragment (Figure 4g, red). In these (C)-containing peptides, the intrinsic switch enabled the modulation of the secondary structure and their transduction enhancing features (Figure 4g, red). By optimizing the internal amphiphilic pattern of PNF-3, which afforded PNF-9, both β-sheet content and transduction efficiency received a significant boost (Figure 4g), which is in good agreement with an optimized packing of the molecules due to the elimination of a polarity "defect" in the structure. The effects delivered by the intrinsic switch were of similar importance, when considering the (F)-derivative peptides (PNF-18). This structure–activity relationship represents the first result clearly indicating that the molecular structure, and in particular the content of intermolecular β-sheet content, is a fundamental parameter determining the infectivity enhancing properties of peptide nanofibrils.

2.6. Negatively Charged PNFs Bind Virions but Not Cells

The positively charged (K) residues could be involved in interactions with negatively charged viral membranes and/or the negatively charged lipid bilayers of cellular membranes. In this way, positively charged PNFs could enhance virus transduction by reducing the electrostatic repulsion forces between viral and cellular membranes.[15,17] In order to elucidate the interaction partner of PNFs, negatively charged peptide sequences based on bioactive PNF-7 to PNF-9 were prepared in which the hydrophilic, polar (K) residue was replaced by glutamic acid (E). The corresponding derivatives PNF-19 (CEIEIQINMWQ), PNF-20 (CEIEIQINM), and PNF-21 (CEIEIQI) formed PNFs with negative zeta potentials with good CRs (43–79%), almost comparable to their corresponding positive (K)-derivatives PNF-7 to PNF-9. The CR of the negative (E)-peptides PNF-19 to PNF-21, but their biological activity was an order of magnitude smaller (**Figure 5**).

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Next, we studied the formation of mesoscale virus-PNF clusters by applying fluorescent laser scanning microscopy (LSM). PNFs were stained by the red-emitting fluorophore Proteostat and yellow fluorescent protein (YFP)-labelled retroviral particles (green) were used. Colocalization of PNFs and virions was observed for all investigated PNFs as depicted in Figure 5e and Figure S11, Supporting Information, indicating that despite their negative charge, PNF-19 and PNF-21 were still able to bind virions, although to a lower extent as their positively charged analogues. The interaction of PNFs with cellular membranes was then assessed by incubating Proteostat-labelled PNFs with TZM-bl cells for 30 min. The samples were analyzed by confocal microscopy before washing confirming the presence of PNF. After 1 h, cells were washed with PBS and examined again. PNF clusters were removed that did not significantly interact with the cellular surface. In contrast to (*K*)-containing PNFs, negatively charged PNFs based on optimized peptide sequences interacted with virions but they did not associate with the cellular membrane (Figure S12, Supporting Information), which most likely explains their lack of activity in the transduction assay (Figure 5c). The (K) residues in the bioactive PNFs reveal enhanced virus binding and facilitate colocalization of the virus-PNF complexes at cellular membranes. Interestingly, a set of peptides PNF-S9 and PNF-S10, where the charged residues were replaced with serine (S) (Figure S13, Supporting Information), show intermediate activity roughly between that of the highly active (K)-containing PNFs and the non-active (E)-containing PNFs. Colocalization of these PNFs and viral particles indicated poor attachment to cellular membranes confirming the overall importance of positively charged residues to enhance viral transduction efficacy. We conclude that a positive ζ-potential is crucial for the transduction enhancing properties of PNFs but it is not essential for PNF assembly and interaction with virions.

2.7. The Shortest Transduction Enhancing Sequence PNF-18 is a Potent Enhancer of Lenti- and γ-Retroviral Transduction

The peptide PNF-18 is a highly efficient transduction enhancer providing a short sequence of only seven amino acids. We selected this hit peptide for a comparative analysis with the EF-C 12-mer using multiple viral vectors and cell lines (**Figure 6**a). A lentiviral vector encoding a luciferase gene and carrying the G-protein of the vesicular stomatitis virus (VSV-G), a viral fusion protein often used in gene transfer approaches, was generated. p24 antigen of the viral pseudotype preparation (2 or 14 ng) were exposed to EF-C and PNF-18, and used to inoculate HEK 293T cells. Transduction efficiencies were determined 3 days later by quantifying luciferase activities in cellular lysates. Both PNFs substantially increased transduction (Figure 6b). At 100 μg mL^{-1}, PNF-18 is more active than EF-C, with maximal transduction enhancement of 32-fold (2 ng p24) and 17× (14 ng p24), as compared to 23× (2 ng p24) and 10× (14 ng p24), respectively, for EF-C. We also generated a lentiviral vector, SEW LUC2, by co-transfection of a crippled lentiviral genome encoding luciferase, a gag-pol and a VSV-G expression plasmid. The resulting vector was treated with EF-C and PNF-18, and mixtures were added to HeLA (TZM-bl) cells and murine fibroblast 3T3 cells. Again, transduction rates were determined by quantifying luciferase 3 days later. As shown in Figure 6c, both PNF dose-dependently increased transduction of both cell lines analyzed. PNF-18 revealed maximal enhancement of 121-fold in TZM-bl cells and 66-fold in 3T3 cells. Since one of the most advanced clinical applications of PNF is to increase retroviral transduction of T cells in CAR T cell therapy,[34,63] we finally evaluated PNF-18 in this setting. For this, a GFP encoding retroviral vector carrying the Env protein of the Gibbon ape leukemia virus (GALV-Env) which is commonly used in CAR T cell therapy was generated, without applying any optimization protocols or virion concentration by ultracentrifugation. Next, the vector was exposed to PNF-18 and added to primary T cells isolated from two independent blood donors. Transduction rates were analyzed 3 days later by determining the percentage of GFP + cells by flow cytometry (Figure 6d,e). In the absence of PNFs, hardly any transduced GFP+ cells were detected. However, with an increasing concentration of PNF-18, the number of transduced cells gradually increased to 9.5% for donor A, and 7.2% for donor B (Figure 6d), corresponding to 63 or 48-fold increased transduction rates as compared to the control containing no PNFs (Figure 6e). These results show, that fibrils derived from the 7-mer peptide PNF-18 substantially improved lentiviral gene transfer rates into adherent and primary suspension cells, independently of the viral fusion protein and the vector types used, rendering PNF-18 as a novel highly versatile and easy to produce transduction enhancer.

3. Conclusions

Supramolecular nanostructures are notoriously challenging to optimize as rational approaches face multiparameter and multiscale challenges. Herein, we have reported the first sequence optimization of peptides forming bioactive amyloid-like PNFs that enhance retroviral gene transduction based on single fibril analysis and quantitative infection data. The impact of different amino acid patterns and sequences responsible for assembly, fibril morphologies, virus and cell membrane binding were resolved. We found that the motif NMWQ in QCKIKQIINMWQ was not essential for their transduction enhancing properties. Optimizing the amphiphilic six amino acid sequence (KIKQII) afforded (CKFKFQF) with quantitative peptide-to-fibril CRs. The N-terminal (C) also affected fibril morphology by further increasing the intermolecular β-sheet content, which also contributed to high transduction efficacy in particular in short sequences. The structure–activity relationship showed a linear correlation of the content of intermolecular β-sheet structures and transduction enhancement of the PNFs. The (*K*)-residues were found important for binding both virions and cell surface structures by reducing the electrostatic repulsion forces between viral and cellular membranes to colocalize virions at the cellular membrane, which is required for transduction enhancement. Interestingly, negatively charged PNFs with optimized sequence patterns could still bind virions but the formed negative PNF-virion complexes did not co-localize with cellular membranes.

The nine-mers PNF-8 and PNF-17 as well as the seven-mer PNF-18 (CKFKFQF) represent a minimum sequence

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

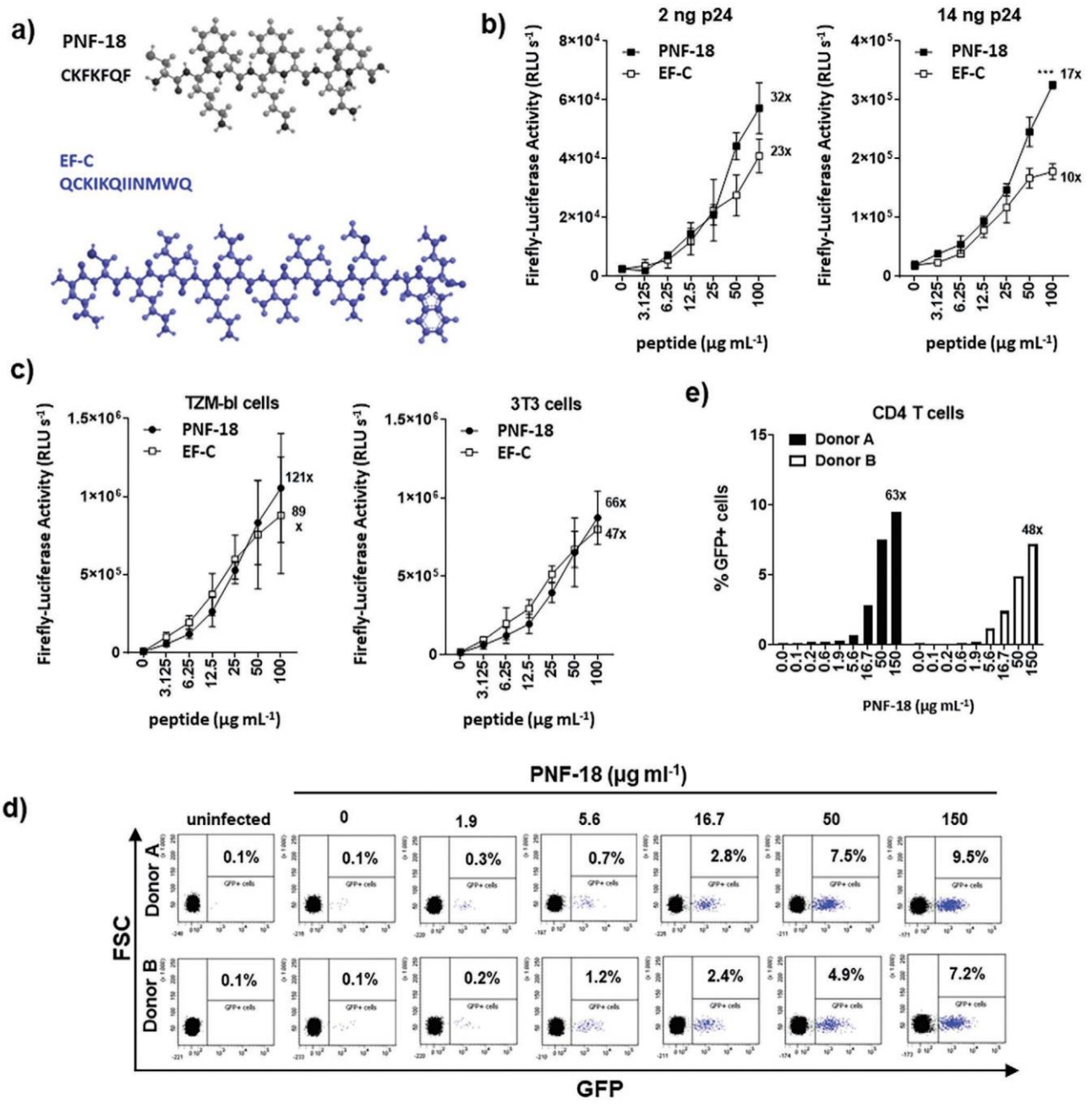


Figure 6. Fibrils derived from the shortened peptide PNF-18 are potent enhancer of lenti- and γ-retroviral transduction. a) Structure of PNF-18 and the parental EF-C peptide. b) Transduction efficiencies of a low (left) or median (right) dose of VSV-G pseudotyped HIV-1 NL4-3 particles encoding a luciferase. Viral particles were exposed to the indicated concentrations of EF-C and PNF-18 and used to inoculate HEK 293T cells. Luciferase activities were determined three days later, shown are mean values derived from triplicate infection ± standard deviation (SD). c) Transduction efficiencies of a VSV-G pseudotyped lentiviral vector (SEW LUC2) in TZM-bl or 3T3 cells. The SEW LUC2 vector preparation was incubated with the PNF and used to infect TZM-bl and 3T3 cells. Luciferases activities were determined 3 days later, shown are mean values derived from triplicate infections ± SD. d) Dot plots showing GFP+ CD4 T cells obtained from two independent donors after transduction with GALV pseudotyped γ-retrovirus at indicated concentrations of PNF-18 at three days post transduction. e) Percentage of GFP positive cells shown in (d). Numbers above data points give the *n*-fold enhancement of transduction as compared to controls containing no PNF.

combining PNF formation with high CRs and high infectivity due to: 1) high content of cross-β-sheet structures; 2) ability to bind virions; and 3) interact with cellular membranes. PNF-18 provides a more than 40% reduced sequence length compared to the original 12-mer EF-C sequence promoting optimized lentiviral HIV infection and enhanced retroviral transduction with high efficiency outperforming the original EF-C sequence. PNF-18 is less expensive to synthesize, it can be purified and

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

produced without batch-to-batch variations and it further provides reactive thiol groups to introduce further chemical modifications, that is, via maleimide chemistry. We envision that negatively charged (E)-based PNFs capable of binding viral particles without interacting with cellular membranes could offer exciting opportunities to develop antiviral nanomaterials as well as PNFs targeting specific cell types as positive charges usually mediate unspecific electrostatic interactions with cellular membranes.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

S.S. and T.M. contributed equally to this work. The project was funded by the German Research Foundation (DFG) – project number 316249678 – CRC 1279 (A03, A05, C01). T.W., J.M., and F.K. acknowledge funding by the Volkswagen Foundation (project 89943) and J.M. by the Leibniz Association (SAW project). The authors thank Jasmina Gačanin for assisting in peptide synthesis and Nicole Kirsch-Pietz in proofreading.

Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data available on request from the authors.

Keywords

peptide nanofibrils, retroviral gene transfer, self-assembly, structure–activity relationship

Received: November 3, 2020
Revised: January 20, 2021
Published online: February 22, 2021

[1] T. P. J. Knowles, J. F. Smith, G. L. Devlin, C. M. Dobson, M. E. Welland, *Nanotechnology* **2007**, *18*, 044031.
[2] T. Lührs, C. Ritter, M. Adrian, D. Riek-loher, B. Bohrmann, H. Do, D. Schubert, R. Riek, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17342.
[3] L. C. Serpell, *Biochim. Biophys. Acta, Mol. Basis Dis.* **2000**, *1502*, 16.
[4] G. Wei, Z. Su, N. P. Reynolds, P. Arosio, I. W. Hamley, E. Gazit, R. Mezzenga, *Chem. Soc. Rev.* **2017**, *46*, 4661.
[5] I. Cherny, E. Gazit, *Angew. Chem., Int. Ed.* **2008**, *47*, 4062.
[6] Z. Xu, R. Paparcone, M. J. Buehler, *Biophys. J.* **2010**, *98*, 2053.
[7] D. M. Fowler, A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch, J. W. Kelly, *PLoS Biol.* **2006**, *4*, 0100.
[8] S. K. Maji, D. Schubert, C. Rivier, S. Lee, J. E. Rivier, R. Riek, *PLoS Biol.* **2008**, *6*, e17.
[9] N. R. Roan, N. Sandi-Monroy, N. Kohgadai, S. M. Usmani, K. G. Hamil, J. Neidleman, M. Montano, L. Ständker, A. Röcker, M. Cavrois, J. Rosen, K. Marson, J. F. Smith, C. D. Pilcher, F. Gagsteiger, O. Sakk, M. O'Rand, P. V. Lishko, F. Kirchhoff, J. Münch, W. C. Greene, *Elife* **2017**, *6*, e24888.
[10] M. R. Chapman, L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark, S. J. Hultgren, *Science* **2002**, *295*, 851.
[11] T. P. J. Knowles, M. Vendruscolo, C. M. Dobson, *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 384.
[12] N. R. Roan, H. Liu, S. M. Usmani, J. Neidleman, J. A. Muller, A. Avila-Herrera, A. Gawanbacht, O. Zirafi, S. Chu, M. Dong, S. T. Kumar, J. F. Smith, K. S. Pollard, M. Fandrich, F. Kirchhoff, J. Munch, H. E. Witkowska, W. C. Greene, *J. Virol.* **2014**, *88*, 7221.
[13] S. M. Usmani, O. Zirafi, J. A. Müller, N. L. Sandi-Monroy, J. K. Yadav, C. Meier, T. Weil, N. R. Roan, W. C. Greene, P. Walther, K. P. R. Nilsson, P. Hammarström, R. Wetzel, C. D. Pilcher, F. Gagsteiger, M. Fändrich, F. Kirchhoff, J. Münch, *Nat. Commun.* **2014**, *5*, 3508.
[14] J. Münch, E. Rücker, L. Ständker, K. Adermann, C. Goffinet, M. Schindler, S. Wildum, R. Chinnadurai, D. Rajan, A. Specht, G. Giménez-Gallego, P. C. Sánchez, D. M. Fowler, A. Koulov, J. W. Kelly, W. Mothes, J. C. Grivel, L. Margolis, O. T. Keppler, W. G. Forssmann, F. Kirchhoff, *Cell* **2007**, *131*, 1059.
[15] A. Röcker, N. R. Roan, J. K. Yadav, M. Fändrich, J. Münch, *Chem. Commun.* **2018**, *54*, 7557.
[16] K. Kaygisiz, C. V. Synatschke, *Biomater. Sci.* **2020**, *8*, 6113.
[17] C. Meier, T. Weil, F. Kirchhoff, J. Münch, *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.* **2014**, *6*, 438.
[18] M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständker, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff, J. Münch, *Nat. Nanotechnol.* **2013**, *8*, 130.
[19] E. Montini, D. Cesana, M. Schmidt, F. Sanvito, M. Ponzoni, C. Bartholomae, L. S. Sergi, F. Benedicenti, A. Ambrosi, C. Di Serio, C. Doglioni, C. Von Kalle, L. Naldini, *Nat. Biotechnol.* **2006**, *24*, 687.
[20] L. Naldini, U. Blömer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, D. Trono, *Science* **1996**, *272*, 263.
[21] S. L. Ginn, A. K. Amaya, I. E. Alexander, M. Edelstein, M. R. Abedi, *J. Gene Med.* **2018**, *20*, e3015.
[22] P. Kumar, C. Woon-Khiong, *Curr. Gene Ther.* **2011**, *11*, 144.
[23] T. Sakuma, M. A. Barry, Y. Ikeda, *Biochem. J.* **2012**, *443*, 603.
[24] A. Biffi, E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, C. Baldoli, S. Martino, A. Calabria, S. Canale, F. Benedicenti, G. Vallanti, L. Biasco, S. Leo, N. Kabbara, G. Zanetti, W. B. Rizzo, N. A. L. Mehta, M. P. Cicalese, M. Casiraghi, J. J. Boelens, U. Del Carro, D. J. Dow, M. Schmidt, A. Assanelli, V. Neduva, C. Di Serio, E. Stupka, J. Gardner, C. Von Kalle, C. Bordignon, F. Ciceri, A. Rovelli, M. G. Roncarolo, A. Aiuti, M. Sessa, L. Naldini, *Science* **2013**, *341*, 1233158.
[25] L. W. Seymour, A. J. Thrasher, *Nat. Biotechnol.* **2012**, *30*, 588.
[26] H. Hanenberg, K. Hashino, H. Konishi, A. R. Hock, I. Kato, D. A. Williams, *Hum. Gene Ther.* **1997**, *2206*, 2193.
[27] T. Moritz, P. Dutt, X. Xiao, D. Carstanjen, T. Vik, H. Hanenberg, D. A. Williams, *Blood* **1996**, *88*, 855.
[28] A. Quintás-Cardama, R. K. Yeh, D. Hollyman, J. Stefanski, C. Taylor, Y. Nikhamin, G. Imperato, M. Sadelain, I. Rivière, R. J. Brentjens, *Hum. Gene Ther.* **2007**, *18*, 1253.
[29] P. Zhou, J. Lee, P. Moore, K. M. Brasky, *Hum. Gene Ther.* **2001**, *12*, 1843.
[30] K. Dodo, H. Chono, N. Saito, Y. Tanaka, K. Tahara, I. Nukaya, J. Mineno, *PLoS One* **2014**, *9*, e86275.
[31] C. H. J. Lamers, P. van Elzakker, S. C. L. van Steenbergen, S. Sleijfer, R. Debets, J. W. Gratama, *Cytotherapy* **2008**, *10*, 406.
[32] T. Tsukamoto, S. Okada, *J. Virol. Methods* **2017**, *248*, 234.

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED FUNCTIONAL MATERIALS
www.afm-journal.de

[33] H. J. Lee, Y. S. Lee, H. S. Kim, Y. K. Kim, J. H. Kim, S. H. Jeon, H. W. Lee, S. Kim, H. Miyoshi, H. M. Chung, D. K. Kim, *Biologicals* **2009**, *37*, 203.
[34] A. Castro, M. Drosch, M. Liesenfeld, T. Kaan, A. Schilz, H. Wenschuh, W. Forsmann, K. Kuhlke, *Cytotherapy* **2019**, *21*, e14.
[35] D. Fenard, D. Ingrao, A. Seye, J. Buisset, S. Genries, S. Martin, A. Kichler, A. Galy, *Mol. Ther.–Nucleic Acids* **2013**, *2*, e90.
[36] D. Schuetz, S. Rode, C. Read, J. Mueller, B. Glocker, K. Sparrer, O. Fackler, P. Walther, J. Muench, bioRxiv https://doi.org/10.1101/2020.10.01.321810.
[37] J. Chen, R. Ren, S. Tan, W. Zhang, X. Zhang, F. Yu, T. Xun, S. Jiang, S. Liu, L. Li, *PLoS One* **2015**, *10*, e0144522.
[38] J. Chen, R. Ren, F. Yu, C. Wang, X. Zhang, W. Li, S. Tan, S. Jiang, S. Liu, L. Li, *Biophys. J.* **2017**, *113*, 1425.
[39] S. Tan, L. Li, L. Lu, C. Pan, H. Lu, Y. Oksov, X. Tang, S. Jiang, S. Liu, *FEBS Lett.* **2014**, *588*, 1515.
[40] H. Zhang, X. He, Y. Shi, Y. Yu, S. Guan, X. Gong, H. Yin, Z. Kuai, Y. Shan, *RSC Adv.* **2016**, *6*, 82082.
[41] R. Franke, T. Hirsch, J. Eichler, *J. Recept. Signal Transduction* **2006**, *26*, 453.
[42] C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil, B. Knöll, *Adv. Funct. Mater.* **2019**, *29*, 1809112.
[43] S. Navarro, S. Ventura, *Biotechnol. J.* **2014**, *9*, 1259.
[44] E. J. Platt, K. Wehrly, S. E. Kuhmann, B. Chesebro, D. Kabat, *J. Virol.* **1998**, *72*, 2855.
[45] J. Adamcik, F. S. Ruggeri, J. T. Berryman, A. Zhang, T. P. J. Knowles, R. Mezzenga, *Adv. Sci.* **2020**, *8*, 2002182.
[46] V. Castelletto, P. Ryumin, R. Cramer, I. W. Hamley, M. Taylor, D. Allsop, M. Reza, J. Ruokolainen, T. Arnold, D. Hermida-Merino, C. I. Garcia, M. C. Leal, E. Castaño, *Sci. Rep.* **2017**, *7*, 43637.
[47] R. J. Betush, J. M. Urban, B. L. Nilsson, *Pept. Sci.* **2018**, *110*, e23099.
[48] F. T. Senguen, N. R. Lee, X. Gu, D. M. Ryan, T. M. Doran, E. A. Anderson, B. L. Nilsson, *Mol. BioSyst.* **2011**, *7*, 486.
[49] Y. Chen, H. X. Gan, Y. W. Tong, *Macromolecules* **2015**, *48*, 2647.
[50] S. Toksoz, R. Mammadov, A. B. Tekinay, M. O. Guler, *J. Colloid Interface Sci.* **2011**, *356*, 131.
[51] T. A. Martinek, A. Hetényi, L. Fülöp, I. M. Mándity, G. K. Tóth, I. Dékány, F. Fülöp, *Angew. Chem.* **2006**, *118*, 2456.
[52] R. Honda, *Biophys. J.* **2018**, *114*, 885.
[53] F. S. Ruggeri, P. Flagmeier, J. R. Kumita, G. Meisl, D. Y. Chirgadze, M. N. Bongiovanni, T. P. J. Knowles, C. M. Dobson, *ACS Nano* **2020**, *14*, 5213.
[54] K. Baginska, J. Makowska, W. Wiczk, F. Kasprzykowski, L. Chmurzynski, *J. Pept. Sci.* **2008**, *14*, 283.
[55] A. Adochitei, G. Drochioiu, *Rev. Roum. Chim.* **2011**, *56*, 783.
[56] V. Cabiaux, R. Brasseur, R. Wattiez, P. Falmagne, J. M. Ruysschaert, E. Goormaghtigh, *J. Biol. Chem.* **1989**, *264*, 4928.
[57] E. Goormaghtigh, V. Cabiaux, J. Ruysschaert, *Eur. J. Biochem.* **1990**, *193*, 409.
[58] D. M. Byler, H. Susi, *Biopolymers* **1986**, *25*, 469.
[59] F. S. Ruggeri, B. Mannini, R. Schmid, M. Vendruscolo, T. P. J. Knowles, *Nat. Commun.* **2020**, *11*, 2945.
[60] F. S. Ruggeri, G. Longo, S. Faggiano, E. Lipiec, A. Pastore, G. Dietler, *Nat. Commun.* **2015**, *6*, 7831.
[61] N. R. Roan, J. Münch, N. Arhel, W. Mothes, J. Neidleman, A. Kobayashi, K. Smith-McCune, F. Kirchhoff, W. C. Greene, *J. Virol.* **2009**, *83*, 73.
[62] F. S. Ruggeri, J. Adamcik, J. S. Jeong, H. A. Lashuel, R. Mezzenga, G. Dietler, *Angew. Chem., Int. Ed.* **2015**, *54*, 2462.
[63] A. Jamali, L. Kapitza, T. Schaser, I. C. D. Johnston, C. J. Buchholz, J. Hartmann, *Mol. Ther.–Methods Clin. Dev.* **2019**, *13*, 371.

© 2021 The Authors. Advanced Functional Materials published by Wiley-VCH GmbH

16163028, 2021, 17, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adfm.202009382 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supporting Information

© 2021 Wiley-VCH GmbH

ADVANCED FUNCTIONAL MATERIALS

Supporting Information

for *Adv. Funct. Mater.,* DOI: 10.1002/adfm.202009382

Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction

Stefanie Sieste, Thomas Mack, Edina Lump, Manuel Hayn, Desiree Schütz, Annika Röcker, Christoph Meier, Kübra Kaygisiz, Frank Kirchhoff, Tuomas P. J. Knowles, Francesco S. Ruggeri, Christopher V. Synatschke,* Jan Münch,* and Tanja Weil**

Supplementary Information for

Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction

Stefanie Sieste, Thomas Mack, Edina Lump, Manuel Hayn, Desiree Schütz, Annika Röcker, Christoph Meier, Kübra Kaygisiz, Frank Kirchhoff, Tuomas P. J. Knowles, Francesco S. Ruggeri, Christopher V. Synatschke,* Jan Münch,* and Tanja Weil**

S. Sieste, T. Mack, K. Kaygisiz, Dr. C. V. Synatschke, Prof. T. Weil
Department Synthesis of Macromolecules
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
Email: weil@mpip-mainz.mpg.de, synatschke@mpip-mainz.mpg.de

S. Sieste, T. Mack, Dr. C. Meier, Prof. T. Weil
Institute of Inorganic Chemistry I
Ulm University
Albert-Einstein-Allee 11, 89081 Ulm, Germany

Dr. E. Lump, M. Hayn, D. Schütz, Dr. A. Röcker, Prof. F. Kirchhoff, Prof. J. Münch
Institute of Molecular Virology
Ulm University Medical Center
Meyerhofstraße 1, 89081 Ulm, Germany
Email: jan.muench@uni-ulm.de

Prof. T. P. J. Knowles, Prof. F. S. Ruggeri
Department of Chemistry
University of Cambridge
Cambridge CB2 1EW, UK
Email: fsr26@cam.ac.uk

Prof. T. P. J. Knowles
Cavendish Laboratory
University of Cambridge
Cambridge CB3 0HE, UK

Prof. F. S. Ruggeri
Laboratory of Physical Chemistry, Stippeneng 4, 6703 WE, Wageningen University, the Netherlands
Email: simone.ruggeri@wur.nl

Prof. J. Münch
Core Facility of Functional Peptidomics
Ulm University Medical Center
Albert-Einstein-Allee 47, 89081 Ulm, Germany

Methods

Materials for peptide synthesis and characterization

PyBOP, HBTU, Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ile-OH, Fmoc-Cys(Trt)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Met-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Ala-OH, Fmoc-Phe-Wang resin, Fmoc-Ile-Wang resin, Fmoc-Gln(Trt)-Wang resin and Fmoc-Met-Wang resin were purchased from Novabiochem®. *N*-ethyldiisopropylamine for synthesis (DIPEA) and potassium chloride (KCl) were obtained from Merck. Piperidine (≥ 99.5% for peptide synthesis) and trifluoroacetic acid (TFA, ≥ 99.9%) were obtained from Carl Roth. Dimethylformamide (DMF, peptide synthesis), diethyl ether and acetonitrile (HiPerSolv Chromanorm for HPLC-gradient grade) were purchased from VWR Chemicals Prolabo. Dimethylsulfoxid (DMSO, ACS reagent, ≥ 99.9%) was purchased from Honeywell, Riedel-de Haën®. Vivaspin 500 tubes (3 kDa MWCO) were purchased from Sartorius. Syringe filters Chromafil®Xtra RC-20/13 (0.20 µm) were obtained from Machery-Nagel. Uranyl acetate was purchased from Merck. Fluorescamine was purchased from PanReac AppliChem. PBS and α-Cyano-4-hydroxycinnamic acid were purchased from Sigma Aldrich. Proteostat® was purchased from Enzo Life Sciences. All chemicals were used as received unless explicitly stated otherwise. Some PNF sequences were purchased from Phtd Peptides industrial Co. limited with purity of ≥ 95% (marked with ** in Supporting Table S3).

1. Synthesis of PNF derivatives

The peptides were synthesised according to standard Fmoc solid phase peptide synthesis on a microwave assisted peptide synthesizer (Liberty I by CEM or a Biotage Initiator+) in a 0.1 mmol scale. The C- to N-terminal coupling was achieved by using Wang resins preloaded with the first Fmoc-protected amino acid. Fmoc-deprotections were carried out with DMF solutions containing 20% or 25% v/v piperidine. Coupling reactions with 5 equiv. Fmoc-protected amino acid solutions in DMF at 75 °C for 5 or 10 min were catalysed by addition of 5 equiv. HBTU or PyBOP and 10 equiv. DIEA. Cleavage of the resin and final deprotection of the side chains were performed with a mixture of TFA, TIS and H_2O at a ratio of 95:2.5:2.5 by shaking at RT for 2h. The cleaved peptides were precipitated in cold diethyl ether and lyophilized. The final products were analysed with mass spectrometry measured on a Bruker Reflex III matrix assisted laser desorption/ionization time of flight (MALDI-TOF) MS device. 4-Hydroxy-α-cyanocinnaminic acid served as matrix.

2. Purification of PNF derivatives

All peptides were purified via HPLC using an optimized binary gradient (0.1% TFA water-acetonitrile as mobile phase). For preparative scale a Shimadzu system was used (LC-20AP, CBM-20A, SPD-20A) with a reversed phase C18 column (LiChrospher, Merck). Purity was verified by MALDI-TOF-MS (Bruker Reflex III) and analytical HPLC using a 1260 Infinity Quarternary LC System (Agilent Technologies) with an analytical ChroCART® 125-4

column (LiChrospher, Merck). All data can be found in Figurer S12 and Table S3 (Supporting Information).

3. Preparation of PNF fibrils

The peptides were dissolved in DMSO (c = 10 mg/mL, stored at 4°C before usage) and added dropwise to PBS buffer to a final incubation concentration of 650 µM or 1 mg/mL as indicated in the text. Other concentrations described in the text were achieved by further dilution of preformed PNFs.

4. TEM characterization of PNFs

5 µL of pre-assembled PNF fibrils (1 mg/mL) were placed on copper grids which were coated with a thin electron-transparent Formvar-layer and freshly etched with oxygen plasma before use. After 5 min incubation time, the grids were washed three times with ddH_2O (double-distilled water) followed by sample staining with 2% uranyl acetate solution for 5 min. After additional three washing steps excess solvent was removed by filter paper. Measurements were performed on a Jeol 1400 electron microscope with 100 kV acceleration voltage or with an EM 109 transmission electron microscope (Zeiss) at an acceleration voltage of 80 kV.

5. ATR FT-IR

For ATR FT-IR spectroscopy measurements, 50 µL of the respective 1 mg/mL PNF solution was lyophilized and the resulting powder used for measurement. All spectra were recorded on a Bruker Tensor 27 spectrometer with a diamond crystal as ATR element (PIKE MiracleTM,

spectral resolution 4 cm^{-1}). Every sample was measured with 20 scans. Data were analyzed with OriginPro (OriginLab) software.

6. Zeta potential

50 µL of preformed PNF fibrils (1 mg mL^{-1}) were diluted to 1 mL in an aqueous solution of 1 mM KCl. The dilution ratio was adjusted, if the count rate was not sufficient. The zeta potential was derived from the electrophoretic mobility of the peptides and measured using a Zeta Nanosizer ZS (Malvern Instruments) with 1 mL disposable folded capillary cells (Zetasizer Nano series, Malvern). Each measurement was performed in triplicates.

7. Laser scanning microscopy

For cell-binding studies 60 000 TZM-bl cells were seeded in 8-well IBIDI slides and stained the following day with Hoechst for 5 min at 37°C. 180 µL preformed PNF fibrils were stained with 20 µL diluted Proteostat Amyloid plaque detection kit solution according to suppliers' protocols (Enzo Life Science) to get a final concentration of 5 µg/ml. After 5 min, fibril mixtures were transferred to TZM-bl cells. The interaction of fibril clusters with cells was monitored after 30 min incubation time on a Leica confocal microscope with laser excitation wavelengths of 405 nm (Hoechst) and 561 nm (Proteostat). After 1h incubation, samples were washed three times with PBS and analysed again. For MLV-binding studies 180 µl preformed and Proteostat stained fibrils were mixed with 20 µl MLV-YFP and incubated for 10 min. to get a final concentration of 50 µg/ml. Then the mixture was added to an 8-well IBIDI slide and analysed after sedimentation with laser excitation of 488 nm (YFP) and 561 nm (Proteostat).

8. Conversion rate

For the determination of the amount of peptide monomer converting into fibrils, we established a fluorescence-based assay using amine reactive dye fluorescamine.[1] Each PNF was incubated in PBS for 1 d with a concentration of 1 mg/mL (200 µL). 100 µL sample was centrifuged using a Vivaspin 500 tube (3kDa MWCO) to separate free monomers from fibrils (16.06 g, 4°C, 45 min). The other 100 uL of the sample served as a reference (original sample). Both the filtrate and the original sample were lyophilized and then dissolved in 25 µL DMSO. In a black 384-well-plate (Greiner Bio-one) 10 µL of DMSO sample (original sample and filtrate) was added and mixed with 3 uL of fluorescamine solution (3 mg/mL, DMSO). Fluorescence enhancement was measured with an excitation wavelength of λ_{ex}= 365 nm and an emission wavelength of λ_{em}= 470 nm with multiple reads 3x3 (Infinite® M1000 PRO microplate reader) after 20 mins of incubation. All values were calculated as n-fold fluorescence enhancement (DMSO-only as a reference was set to 1). The conversion rate CR was defined according to the following equation:

$$CR = 100 - \frac{100 \times Fluorescence\ Intensity\ (Filtrate)}{Fluorescence\ Intensity\ (Original)}\ [\%]$$

9. Proteostat® assay

Fluorescence spectra were recorded on an Infinite® M1000 PRO microplate reader (Tecan). 9 µL of sample aliquots (incubated for 1d in PBS at 1 mg/mL) were placed in black UV Star® 384 microliter well-plates (Greiner bio-one). A ProteoStat® solution was prepared according to manufacturers' protocol and diluted hundred-fold in PBS. After addition of 1 µL of Proteostat® solution to every sample and 10 mins of incubation time, the fluorescence

emission was recorded at 603 nm upon excitation at 550 nm with multiple reads per well (3x3).

10. Effect of PNF on HIV-1 infection

HIV-1 NL4-3 (R5) virus stocks were generated by calcium phosphate transfection as described before.[2] 10000 TZM-bl cells (a HeLa cell line derivative that expresses high amounts of human CD4, CCR5 and CXCR4 and contain a ß-galactosidase gene under the control of the HIV-1 long terminal repeat (LTR) promotor) were seeded into 96-well plates in 100 µL DMEM. On the next day, 80 µL fresh medium was added to the cells. 40 µL dilutions of EF-C or PNF solutions were prepared in PBS and mixed 1:1 (v/v) with 40 µL virus (pre-diluted 1:30 (v/v) in cell culture medium) by resuspending. After 10 min incubation to allow binding of virus particles to fibrils, 20 µL of virus-fibril mixture was added to 180 µL of cells. 3 days post infection, viral infectivity was determined using a galactosidase screen kit from Tropix as recommended by the manufacturer. β-Galactosidase activities were quantified as relative light units (RLU) per second with an Orion Microplate luminometer (Berthold).

11. CellTiter-Glo assay

TZM-bl cells were seeded in 96 well plates (~6000 cells/well). On the following day, 50 µL of fibril samples (130 µM) were added and serial diluted to the final concentrations of 65, 13, 2.6 and 0 µM. 50 µL of CellTiter-Glo mix was added and shaken for 2 mins. After 10 mins

50 µL were transferred to white plates and luminescence was recorded using an Orion microplate luminometer (Berthold).

12. Generation of lenti- and γ-retroviral vectors

800 000 HEK293T cells were seeded in 2 ml DMEM one day prior to the transfection. Cells were cotransfected with 250 µl OptiMEM containing 2.5 µg DNA and 7.5µl TransIT-LT1. For GALV pseudotyped γ-retrovirus, plasmids were used in a ratio 3:1:2.5 (pcmE26-gfp : GALV : pCsGPpA-ed). For VSV-G pseudotyped pSEW LUC2, plasmids were used in a ratio 2:2:1 (pSEW-luc: VSV-G : delta8.9). 64 h after transfection, virus was harvested, centrifuged for 3 min at 1300 rpm and stored at -80°C. VSV-G pseudotyped HIV-1 NL4-3 Firefly-Luciferase reporter virus stocks were generated using the calcium phosphate transfection method. For this, env-deficient proviral NL4-3 plasmid DNA and VSV-G env plasmid DNA were inoculated with $CaCl_2$ and ultrapure water before being mixed with 1x HBS. After incubation for 10 min at room temperature, the 200 µL of the transfection mix were added to 650,000 HEK293T cells/well seeded one day prior to the transfection in 6W plates. 16 hours post transfection, transfection mix was removed from the cells and cell culture medium was replaced with fresh medium. 48 hours post transfection, virus stocks were harvested, characterized by p24 ELISA and frozen at -80°C. HIV-1 NL4-3 (R5) virus stocks were generated by calcium phosphate transfection as described before.[3]

13. Effect of PNF on VSV-G pseudotyped lentiviral transduction

To determine the effect of PNF on VSV-G pseudotyped HIV-1 or VSV-G pseudotyped pSEW LUC2 transduction, 10 000 cells were seeded in 96W microtiter plates. The next day, virus stocks were inoculated for 10 minutes with either EF-C or PNF-13 to allow binding of virus particles to fibrils. Indicated fibril concentrations were achieved during this inoculation step.

180 µL cells were subsequently infected with 20 µL of these PNF/VS mixtures in triplicates to achieve the indicated final concentrations of p24/mL on the cells. Three days post infection, infection rates were determined using the Firefly luciferase reporter assay (Promega). Firefly-Luciferase activities were quantified as relative light units (RLU) per second with an Orion Microplate luminometer (Berthold).

14. Effect of PNF on GALV pseudotyped γ-retroviral transduction

PBMCs were isolated using lymphocyte separation medium (Biocoll separating solution; Biochrom, Berlin, Germany). CD4 +T cells were isolated using the RosetteSep Human CD4 +T Cell Enrichment Cocktail (Stem Cell Technologies, Vancouver, Canada). Primary cells were cultured in RPMI-1640 containing 10% fetal calf serum (FCS), 2 mM glutamine, streptomycin (100 µg ml^{-1}), penicillin (100 U ml^{-1}), and 66 IU ml^{-1} = 10 ng ml^{-1} interleukin 2 (IL-2). Before infection, primary cells were stimulated for three days with phytohemagglutinin (PHA) (1 µg ml^{-1}). 20 000 CD4 +T cells were seeded into 96-well plates in 100 µl RPMI. Next day, 80 µl fresh medium was added to the cells. Dilutions of EF-C or PNF-13 were prepared in 80µl volume and mixed with 80µl virus (virus was diluted 1:8 in RPMI without FCS). After 10 min incubation to allow binding of virus particles to fibrils, 20µl of virus-fibril mixture was added to 180 µl cells. 3 days post infection percentage of GFP+ cells were determined by flow cytometry.

Supporting Tables

Table S1. Qualitative fibril formation determined via transmission electron microscopy ("+" for fibrils, and "-" for no fibrils/aggregates). Ability to bind Proteostat ("Prot bind.," "+" for ≥10 n-fold fluorescence enhancement and "-" for ≤10 n-fold fluorescence enhancement). Zeta potential calculated by electrophoretic mobility. Conversion rate ("CR") as the percentage of monomers converting into a fibril. Entries marked with * have previously been published.[1]

PNF name	sequence	TEM	Prot bind.	Zeta potential [mV]	CR [%]
PNF-1	CKIKQIINMWQ	+	+	+11.8 ± 0.4	50
PNF-2	CKIKQIINM	+	+	+19.1 ± 0.5	45*
PNF-3	CKIKQII	+	+	+0.6 ± 1.5	55*
PNF-4	KIKQIINMWQ	+	+	+10.8 ± 0.2	86
PNF-5	KIKQIINM	+	+	+16.5 ± 0.5	43
PNF-6	KIKQII	-	-	-	6
PNF-7	CKIKIQINMWQ	+	+	+10.7 ± 2.8*	87*
PNF-8	CKIKIQINM	+	+	+22.3 ± 0.6*	72*
PNF-9	CKIKIQI	+	+	+17.4 ± 1.1*	54*
PNF-10	KIKIQINMWQ	+	+	+17.2 ± 0.7	93*
PNF-11	KIKIQINM	+	+	+41.9 ± 1.1*	64*
PNF-12	KIKIQI	-	-	-	5*
PNF-13	CKAKAQANMWQ	+	-	-4.4 ± 3.2	92
PNF-14	CKAKAQANM	+	-	-3.7 ± 0.9	38
PNF-15	CKAKAQA	+	-	-9.1 ± 1.6	44
PNF-16	CKFKFQFNMWQ	+	+	+14.1 ± 0.8	96
PNF-17	CKFKFQFNM	+	-	+15.9 ± 0.4	88
PNF-18	CKFKFQF	+	+	+22.3 ± 1.4*	95*
PNF-19	CEIEIQINMWQ	+	+	-35.8 ± 1.2	74
PNF-20	CEIEIQINM	+	+	-38.6 ± 5.2	43
PNF-21	CEIEIQI	+	+	-41.9 ± 1.1	59
PNF-S1	KAKAQANMWQ	+	+	-19.6 ± 1.7	65
PNF-S2	KAKAQANM	+	n.d.	n.d.	n.d.
PNF-S3	KAKAQA	+	n.d.	n.d.	n.d.
PNF-S4	KFKFQFNMWQ	+	-	+27.4 ± 0.3	94

PNF-S5	KFKFQFNM	+	-	+20.1 ± 1.9*	91
PNF-S6	KFKFQF	+	-	-	83*
PNF-S7	MKIKIQI	+	-	10.1 ± 0.9	51
PNF-S8	MKFKFQF	+	-	+8.8 ± 0.5	75
PNF-S9	CSISIQINM	+	+	-31.7 ± 2.8	86
PNF-S10	CSISIQI	+	+	-31.8 ± 1.5	82

Table S2. Average RLU/s in infectivity assays of the peptides over multiple experiments at 13 and 65 μM.

PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-1	CKIKQIINMWQ	1.8E+5
PNF-7	CKIKIQINMWQ	1.9E+5
PNF-16	CKFKFQFNMWQ	2.2E+5
PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-4	KIKQIINMWQ	0.9E+5
PNF-10	KIKIQINMWQ	1.4E+5
PNF-S4	KFKFQFNMWQ	0.55E+5
PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-3	CKIKQII	0.05E+5
PNF-9	CKIKIQI	1.55E+5
PNF-18	CKFKFQF	1.8E+5
PNF-S7	MKIKIQI	0.19E+5
PNF-S8	MKFKFQF	0.23E+5
PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-2	CKIKQIINM	2.95E+5
PNF-8	CKIKIQINM	2.1E+5
PNF-17	CKFKFQFNM	1.9E+5
PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-5	KIKQIINM	0.95E+5
PNF-11	KIKIQINM	0.9E+4
PNF-S5	KFKFQFNM	1.25E+5
PNF name	**sequence**	Avg Infection [RLU s^{-1}] (13-65 μM)
PNF-6	KIKQII	0.04E+5
PNF-12	KIKIQI	0.1E+5
PNF-S6	KFKFQF	0.1E+5

Table S3. Collected fractions for all peptides after purification with HPLC, see Supplementary Fig. 12. Calculated mass and assignments for all signals found in mass spectra with an intensity of at least 10 % in relation to the main peak. Entries marked with * have previously been published.[1] PNF marked with ** have been purchased from Phtd Peptides industrial Co. limited with purit ≥ 95%.

SAP name	fraction collected [min]	calc. exact mass [g mol^{-1}]	m/z found
PNF-1	10.90 to 11.89	1403.74	1404.75 $[M+H]^+$, 1416.75 [unknown], 1420.75 $[M+O+H]^+$, 2807.48 $[2M-2H+H]^+$, 2823.48 $[2M-2H+O+H]^+$
PNF-2*	9.81 to 10.79	1089.60	976.57 $[M-ASN+H]^+$, 1090.61 $[M+H]^+$, 1112.59 $[M+Na]^+$, 1196.65 $[M+Anisyl+H]^+$, 2065.15 [unknown], 2179.20 $[2M-2H+H]^+$, 2201.18 $[2M-2H+Na]^+$
PNF-3*	9.63 to 10.05	844.52	845.53 $[M+H]^+$, 867.51 $[M+Na]^+$, 1034.57 $[M+CHCA+H]^+$, 1688.03 $[2M-2H+H]^+$, 1710.01 $[2M-2H+Na]^+$
PNF-4	10.42 to 11.02	1300.73	1301.74 $[M+H]^+$
PNF-5	9.55 to 10.12	986.59	987.60 $[M+H]^+$, 1003.60 $[M+O+H]^+$
PNF-6	9.09 to 9.65	741.51	742.52 $[M+H]^+$, 764.50 $[M+Na]^+$
PNF-7*	10.29 to 11.28	1403.74	1404.75 $[M+H]^+$, 1420.74 $[M+O+H]^+$, 1426.73 $[M+Na]^+$, 1436.74 [unknown], 1452.73 [unknown], 2807.47 $[2M-2H+H]^+$
PNF-8*	10.25 to 10.67	1089.60	976.57 $[M-ASN+H]^+$, 1090.61 $[M+H]^+$, 1003.60 $[M+O+H]^+$, 2065.16 [unknown], 2179.20 $[2M-2H+H]^+$
PNF-9	9.96 to 10.48	844.52	845.53 $[M+H]^+$, 857.53 [unknown], 867.51 $[M+Na]^+$, 1688.03 $[2M-2H+H]^+$, 1710.02 $[2M-2H+Na]^+$, 1725.99 [unknown]
PNF-10*	10.65 to 11.29	1300.73	1301.74 $[M+H]^+$, 1317.74 $[M+O+H]^+$, 1323.55 $[M+Na]^+$, 1349.73 [unknown]
PNF-11*	9.96 to 10.49	986.59	987.60 $[M+H]^+$, 1003.60 $[M+O+H]^+$

PNF-12*	9.82 to 10.29	741.51	742.52 $[M+H]^+$, 764.50 $[M+Na]^+$
PNF-13	8.95 to 9.68	1277.60	1278.61 $[M+H]^+$
PNF-14**	n.d.	963.46	n.d.
PNF-15**	n.d.	718.38	n.d.
PNF-16	10.99 to 11.98	1505.69	1506.70 $[M+H]^+$, 1522.70 $[M+O+H]^+$, 1528.69 $[M+Na]^+$
PNF-17	10.84 to 11.48	1191.56	1192.56 $[M+H]^+$, 1204.56 [unknown]
PNF-18*	10.78 to 11.38	946.47	947.48 $[M+H]^+$, 1136.52 $[M+CHCA+H]^+$, 1892.94 $[2M-2H+H]^+$
PNF-19**	n.d.	1405.64	n.d.
PNF-20**	n.d.	1091.50	n.d.
PNF-21**	n.d.	846.42	n.d.
PNF-S1	8.16 to 9.15	1174.59	1175.60 $[M+H]^+$, 1191.60 $[M+O+H]^+$, 1223.59 [unknown]
PNF-S2**	n.d.	860.45	n.d.
PNF-S3**	n.d.	615.37	n.d.
PNF-S4	11.16 to 11.80	1402.69	1403.69 $[M+H]^+$, 1419.69 $[M+O+H]^+$, 1435.69 [unknown], 1451.68 [unknown]
PNF-S5*	10.09 to 11.08	1088.55	1089.55 [M+H]+, 1105.55 [M+O+H]+, 1111.54 $[M+Na]^+$
PNF-S6*	10.64 to 11.41	843.46	844.47 [M+H]+, 866.45 [M+Na]+
PNF-S7	10.25 to 10.79	872.55	873.56 $[M+H]^+$, 889.55 $[M+O+H]^+$, 895.54 $[M+Na]^+$, 911.52 $[M+K]^+$
PNF-S8	9.99 to 10.98	974.50	975.51 $[M+H]^+$, 991.51 $[M+O+H]^+$
PNF-S9**	n.d.	1007.48	n.d.
PNF-S10**	n.d.	762.39	n.d.

Supporting Figures

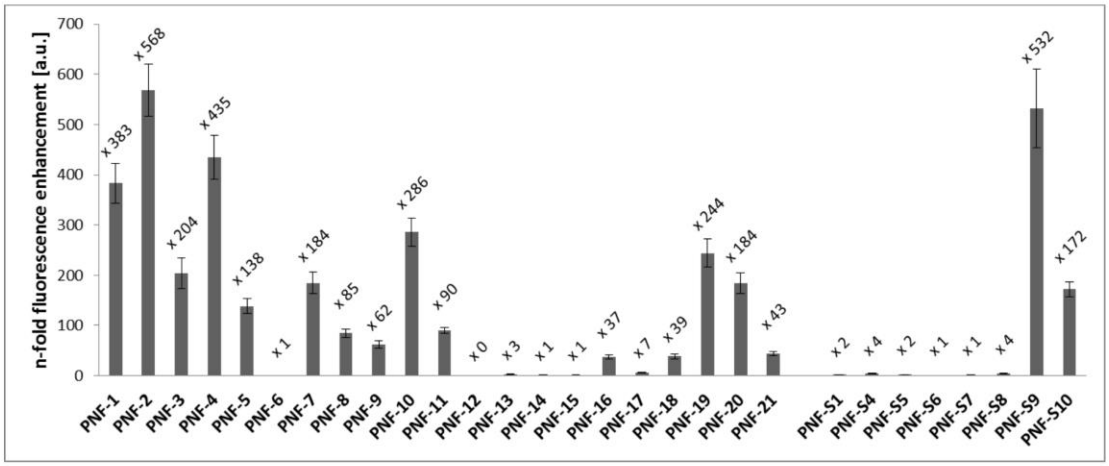


Figure S1: Proteostat® assay showing n-fold fluorescence enhancement of preformed PNF samples at 1 mg mL^{-1}.

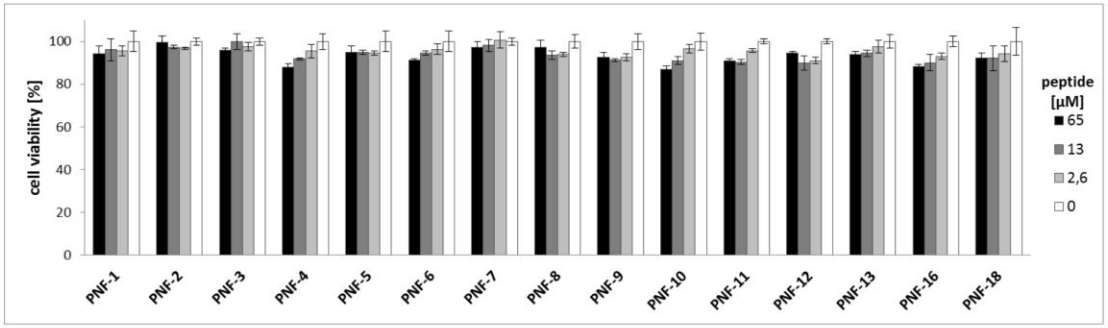


Figure S2: CellTiter-Glo cell viability assay of TZM-bl cells treated with increasing concentrations of PNFs. Values shown indicate mean values derived from biological triplicate ± standard deviation.

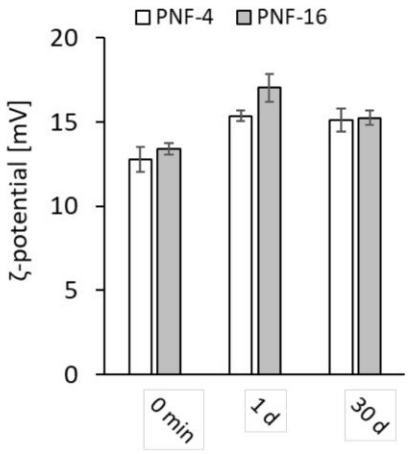


Figure S3: Time-dependent ζ-potential measurements of PNF-4 and PNF-16.

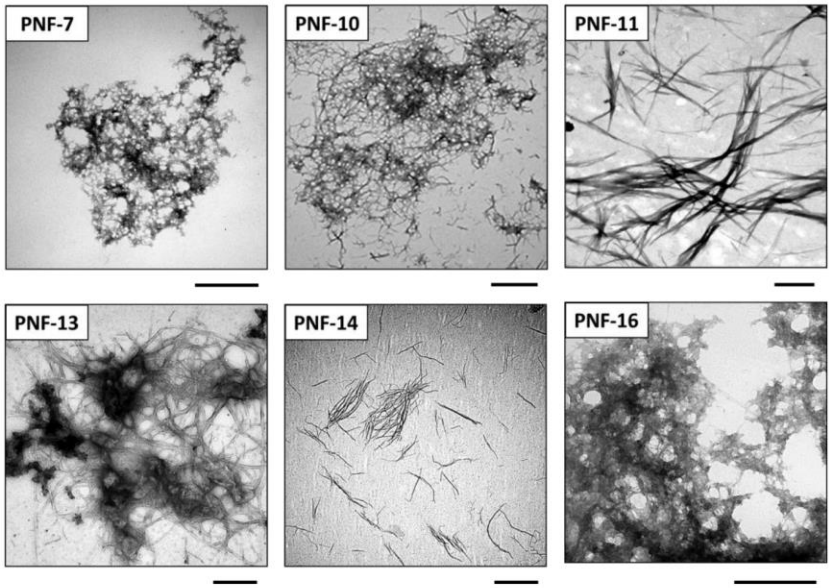
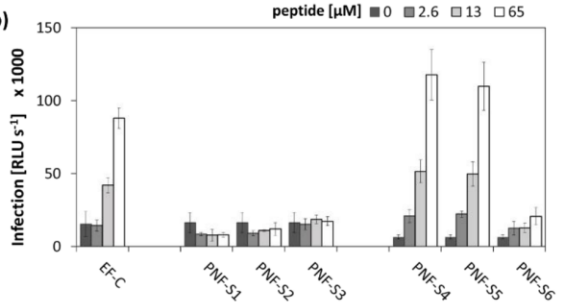


Figure S4: Representative TEM micrographs of PNF assemblies. Scale bars represent 500 nm.

a)

Abb.	Sequence	TEM	CR [%]	ζ [mV]
EF-C	QCKIKQIINMWQ	+	95	+17.7 ± 1.7
PNF-S1	KAKAQANMWQ	+	65	-19.6 ± 1.7
PNF-S2	KAKAQANM	n.d.	n.d.	n.d.
PNF-S3	KAKAQA	n.d.	n.d.	
PNF-S4	KFKFQFNMWQ	+	94	+27.4 ± 0.3
PNF-S5	KFKFQFNM	+	91	+20.1 ± 1.9*
PNF-S6	KFKFQF	-	83*	-

b)



c)

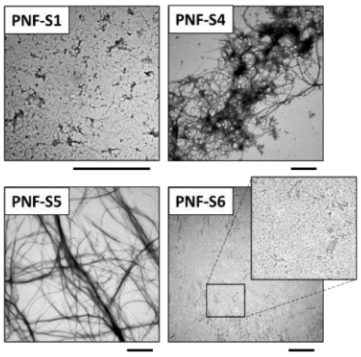


Figure S5: a) Tabular list of alanine and phenylalanine containing peptide derivatives, including the ability to form PNFs for PNF-S1, PNF-S4 and PNF-S5 as observed in TEM measurements with + = fibrils and - = no aggregate formation detected. Monomer-to-fibril conversion rates (CR) were determined in a fluorescence-based assay. High values indicate a large percentage of peptides participating in fibril formation. Zeta potential values for PNF solutions. b) Transduction enhancement assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Shown are mean values derived from triplicate infections ± standard deviation. RLU/s; relative light units per second. c) Representative TEM micrographs of PNF assemblies. Scale bars represent 500 nm. Entries marked with * have previously been published.[1]

a)

Abb.	Sequence	TEM	CR [%]	ζ [mV]
PNF-9	CKIKIQI	+	54*	+17.4 ± 1.1
PNF-S7	MKIKIQI	+	51	+10.1 ± 0.9
PNF-18	CKFKFQF	+	95*	+22.3 ± 1.4
PNF-S8	MKFKFQF	+	75*	+8.8 ± 0.5

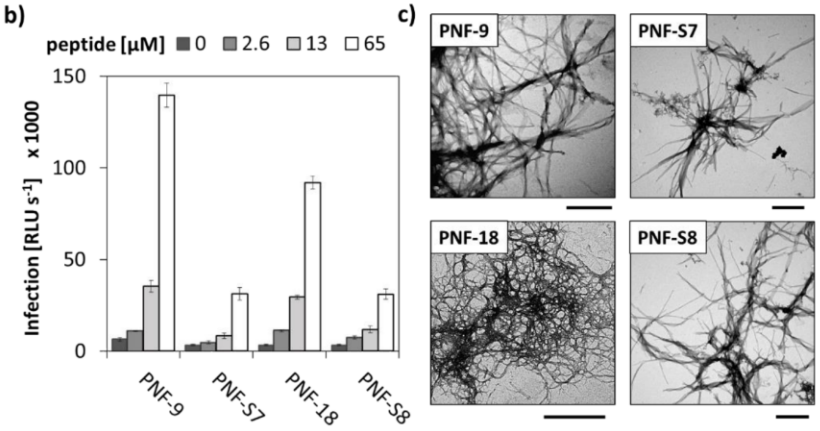


Figure S6: a) Tabular list of cysteine and methionine containing peptide derivatives, including the ability to form PNFs as observed in TEM measurements with + = fibrils and - = no aggregate formation detected. Monomer-to-fibril conversion rates (CR) were determined in a fluorescence-based assay. High values indicate a large percentage of peptides participating in fibril formation. Zeta potential values for PNF solutions. b) Transduction enhancement assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Shown are mean values derived from triplicate infections ± standard deviation. RLU/s; relative light units per second. c) Representative TEM micrographs of PNF assemblies. Scale bars represent 500 nm. Entries marked with * have previously been published.[1]

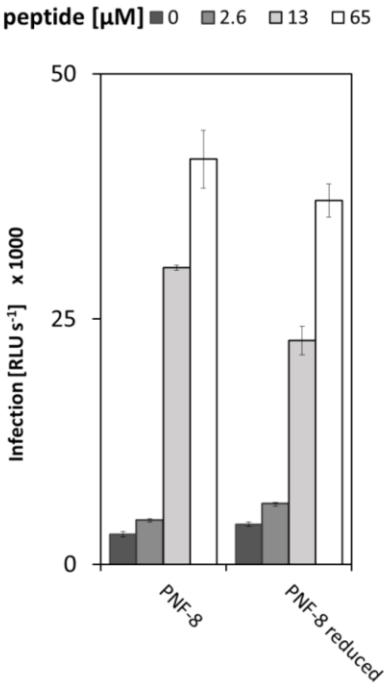


Figure S7: Transduction enhancement assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Shown are mean values derived from triplicate infections ± standard deviation. RLU/s; relative light units per second. PNF-8 and PNF-8 treated with 1 eq TCEP before inducing self-assembly.

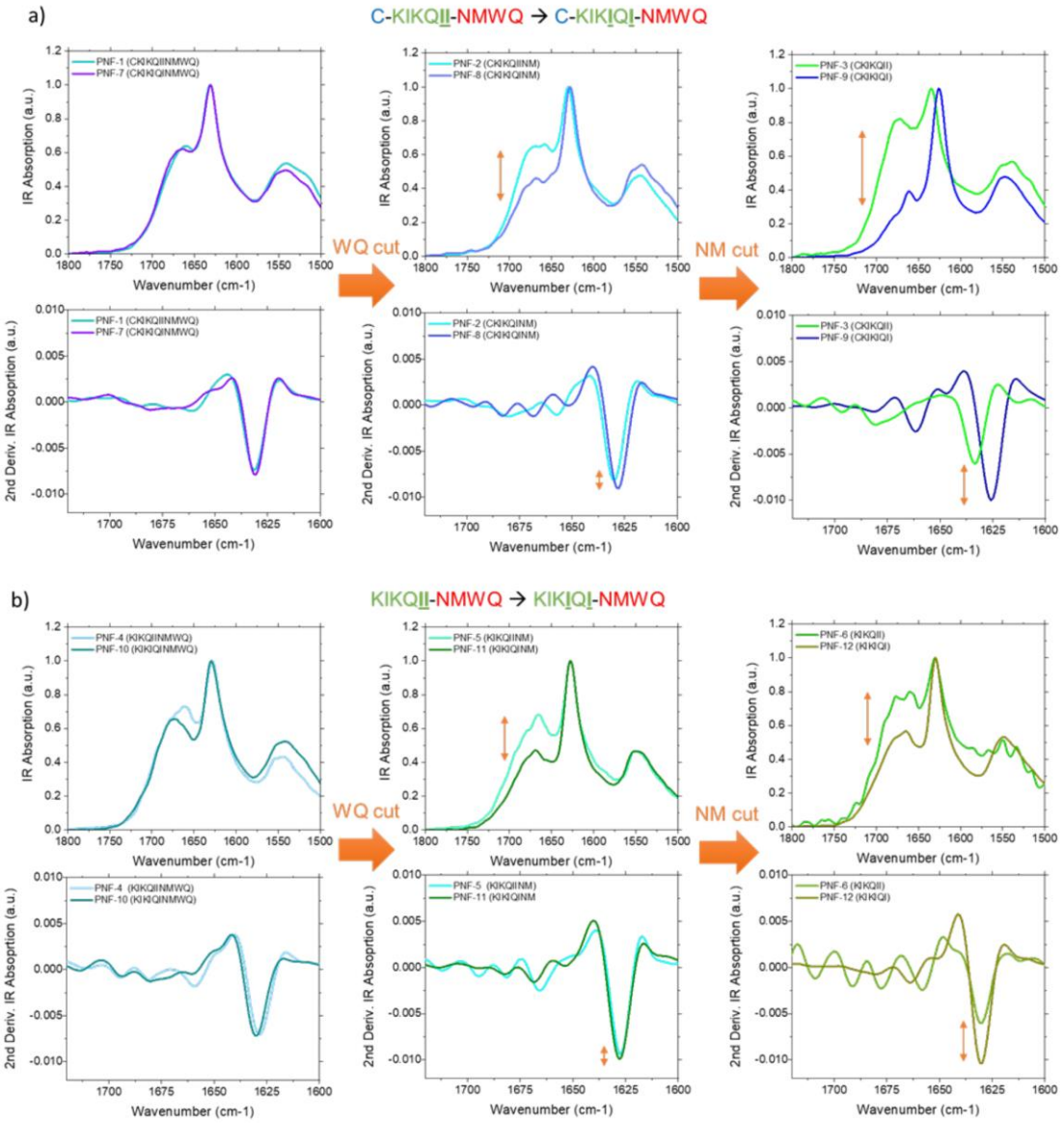


Figure S8: Structural variability upon intrinsic switch in a) presence and b) absence of the N-terminal cysteine.

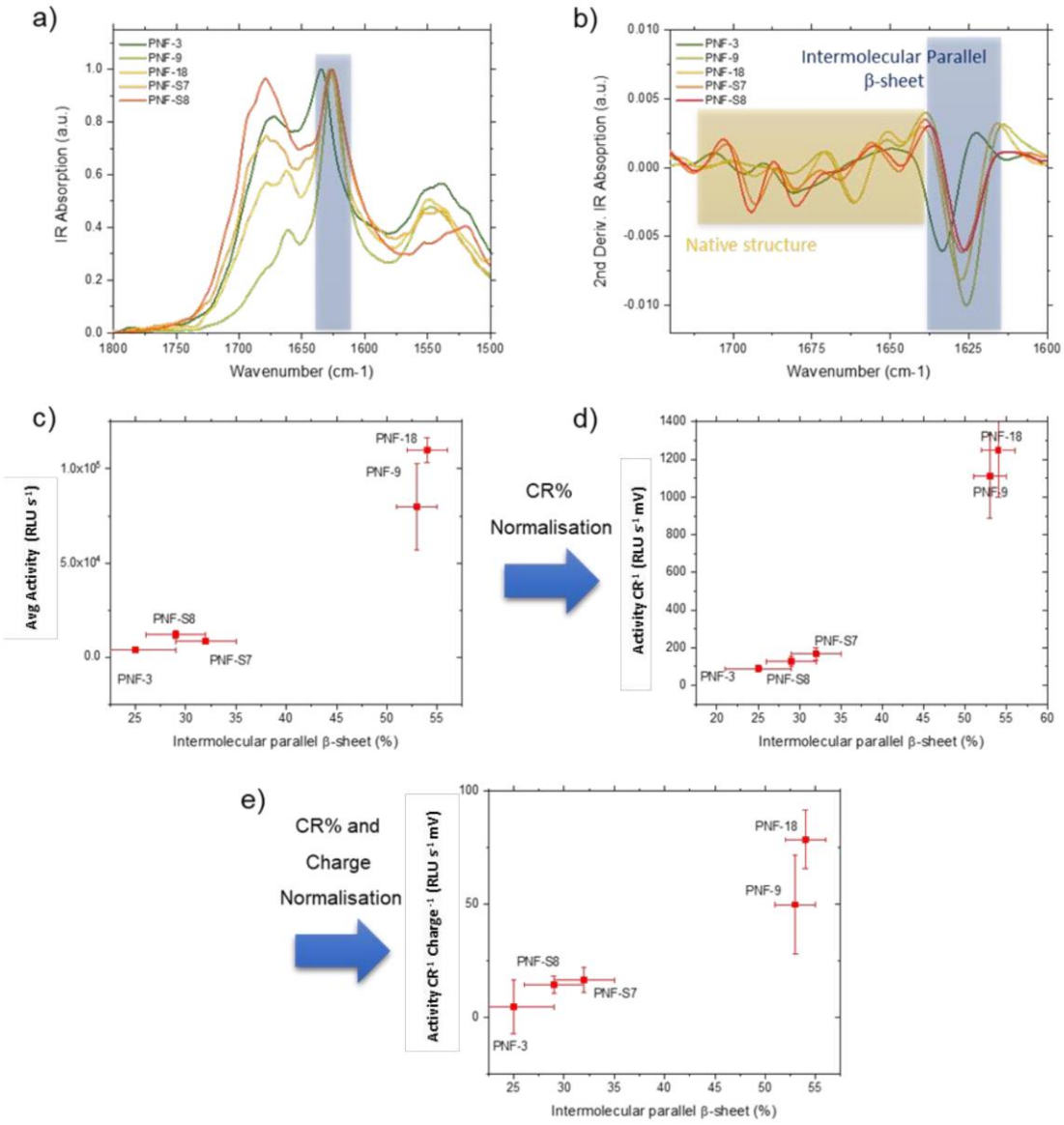


Figure S9: Structural variability upon intrinsic switch in a) presence and b) absence of the N-terminal cysteine.

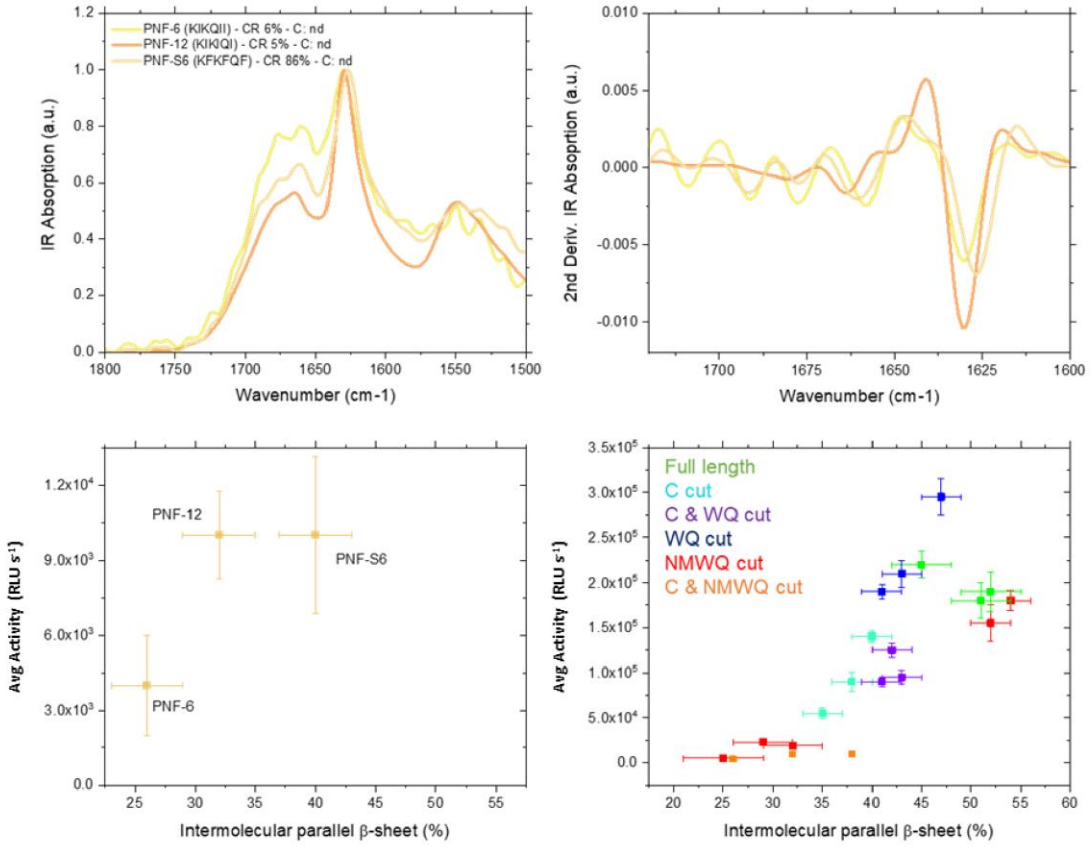


Figure S10: Structural variability vs. activity for PNF-6, -12, S6.

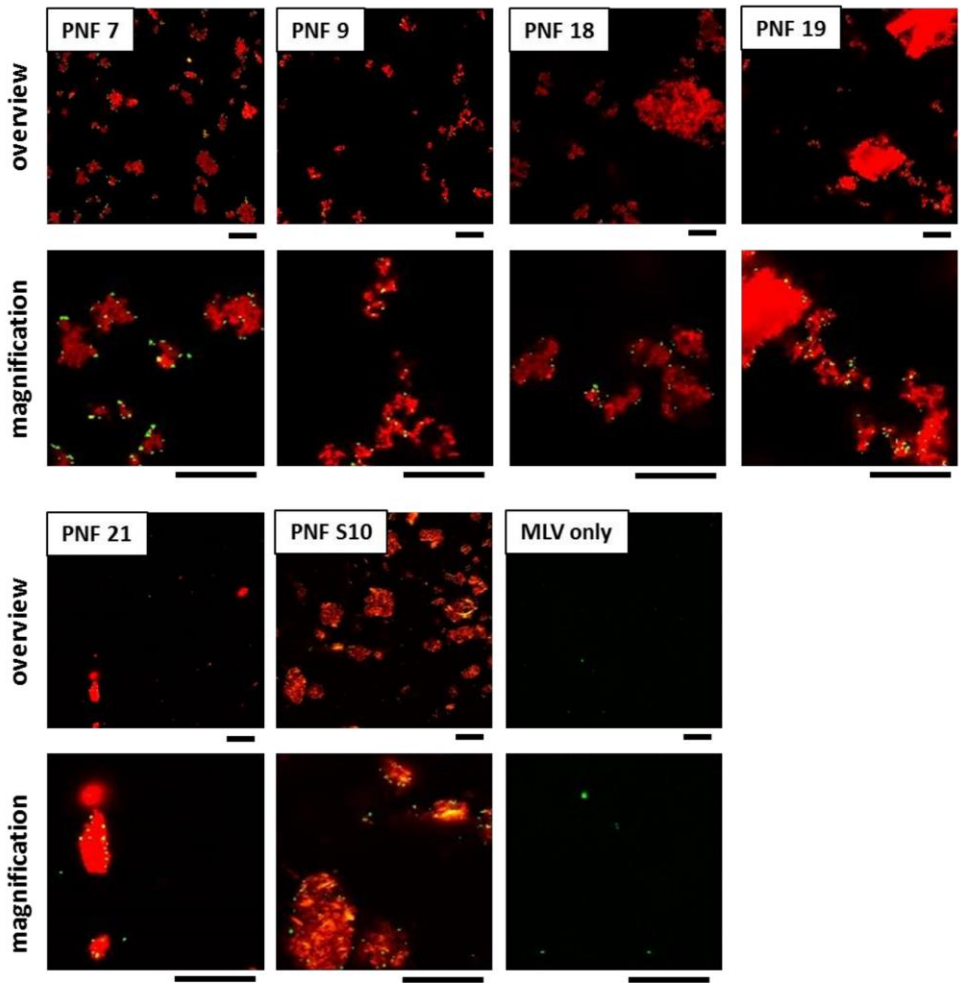


Figure S11: Laser scanning micrographs of different Proteostat® stained PNFs (red) incubated with MLV-YFP (green) to investigate attractions between fibrils and viral particles. Scale bars represent 20 µm.

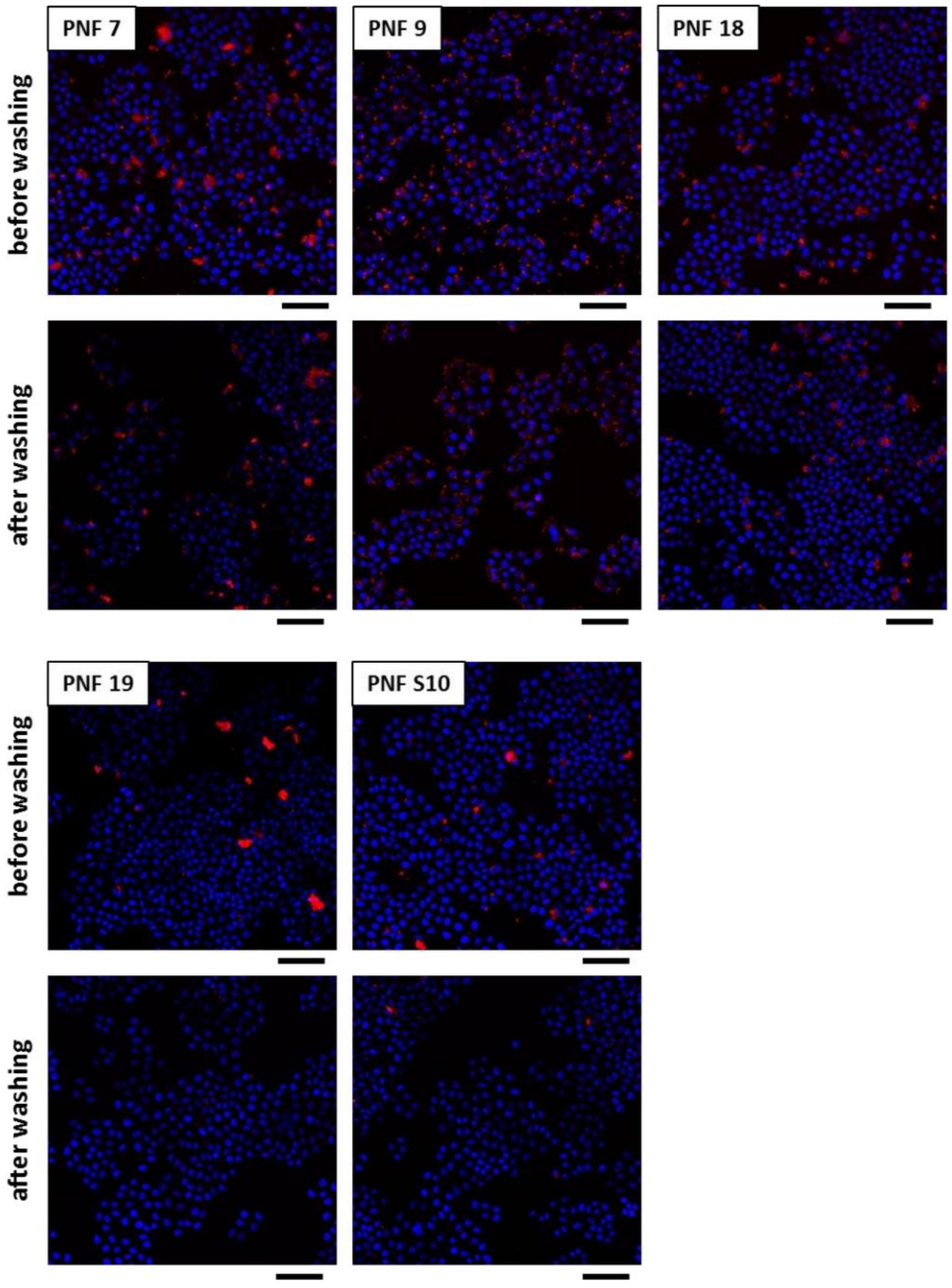


Figure S12: Laser scanning micrographs of Proteostat® stained PNFs (red) incubated on Hoechst stained TZM-bl cells (blue) for determination of PNF cell attachment before and after washing with PBS. Scale bars represent 100 µm.

a)

Abb.	Sequence	TEM	CR [%]	ζ [mV]
PNF-S9	CSISIQINM	+	86	-31.7 ± 2.8
PNF-S10	CSISIQI	+	82	-31.8 ± 1.5

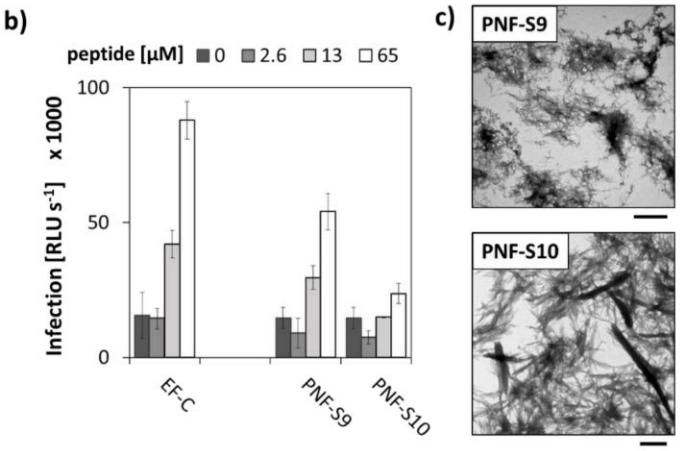


Figure S13: a) Tabular list of serine containing peptide derivatives, including the ability to form PNFs as observed in TEM measurements with + = fibrils and - = no aggregate formation detected. Monomer-to-fibril conversion rates (CR) were determined in a fluorescence-based assay. High values indicate a large percentage of peptides participating in fibril formation. Zeta potential values for PNF solutions. b) Transduction enhancement assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PNF. Shown are mean values derived from triplicate infections ± standard deviation. RLU/s; relative light units per second. c) Representative TEM micrographs of PNF assemblies. Scale bars represent 500 nm.

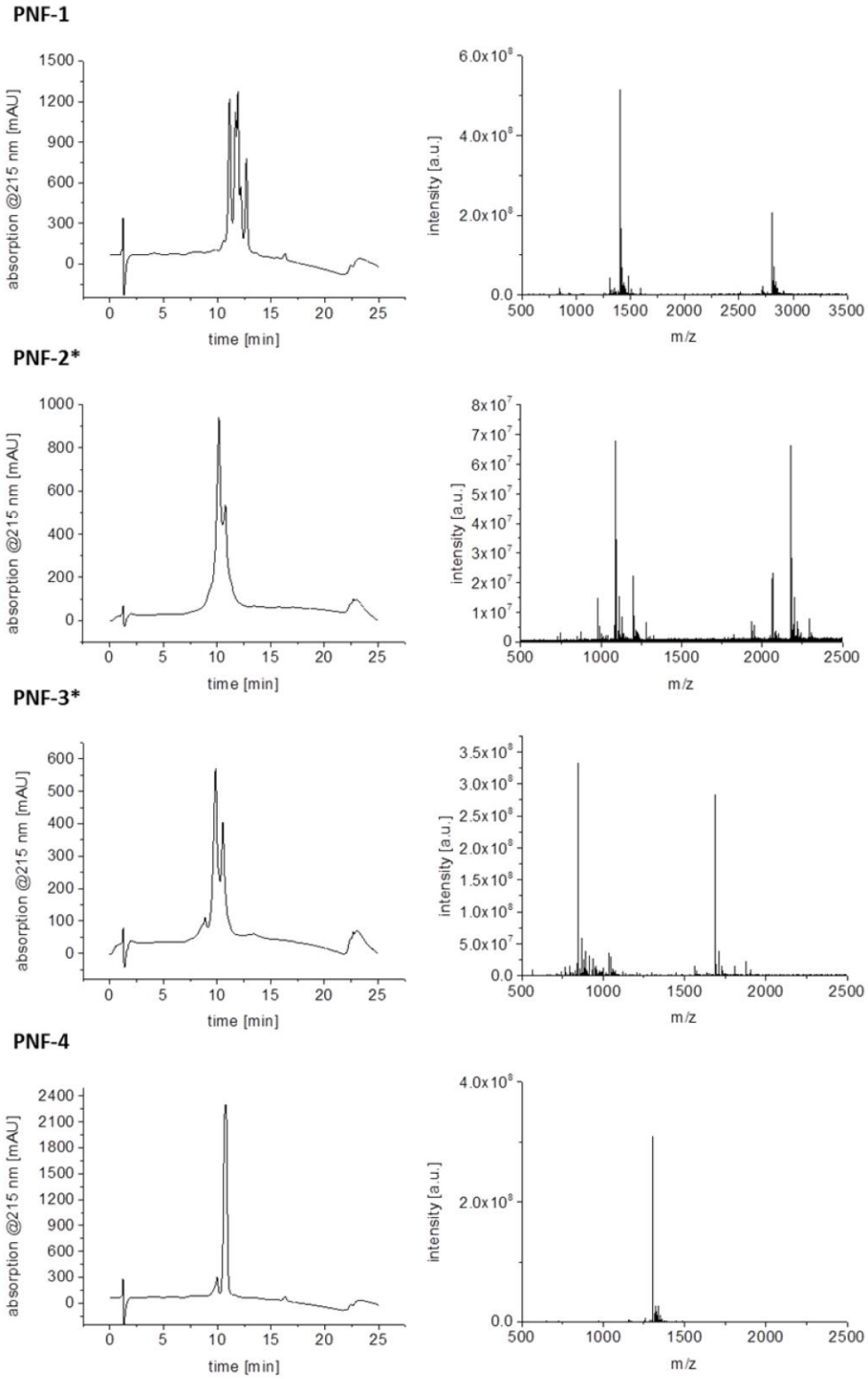


Figure S14: ctnd.

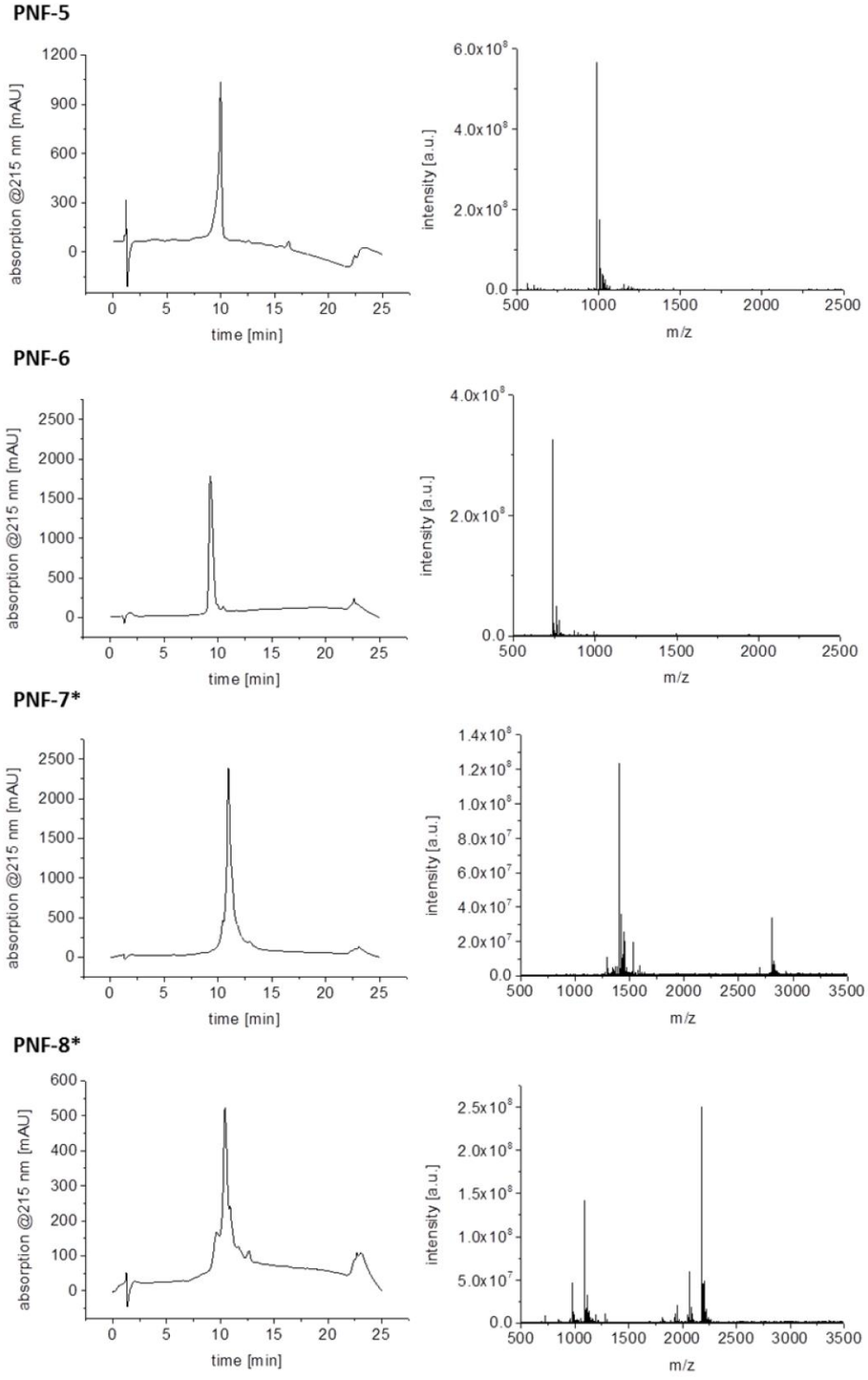


Figure S14: ctnd.

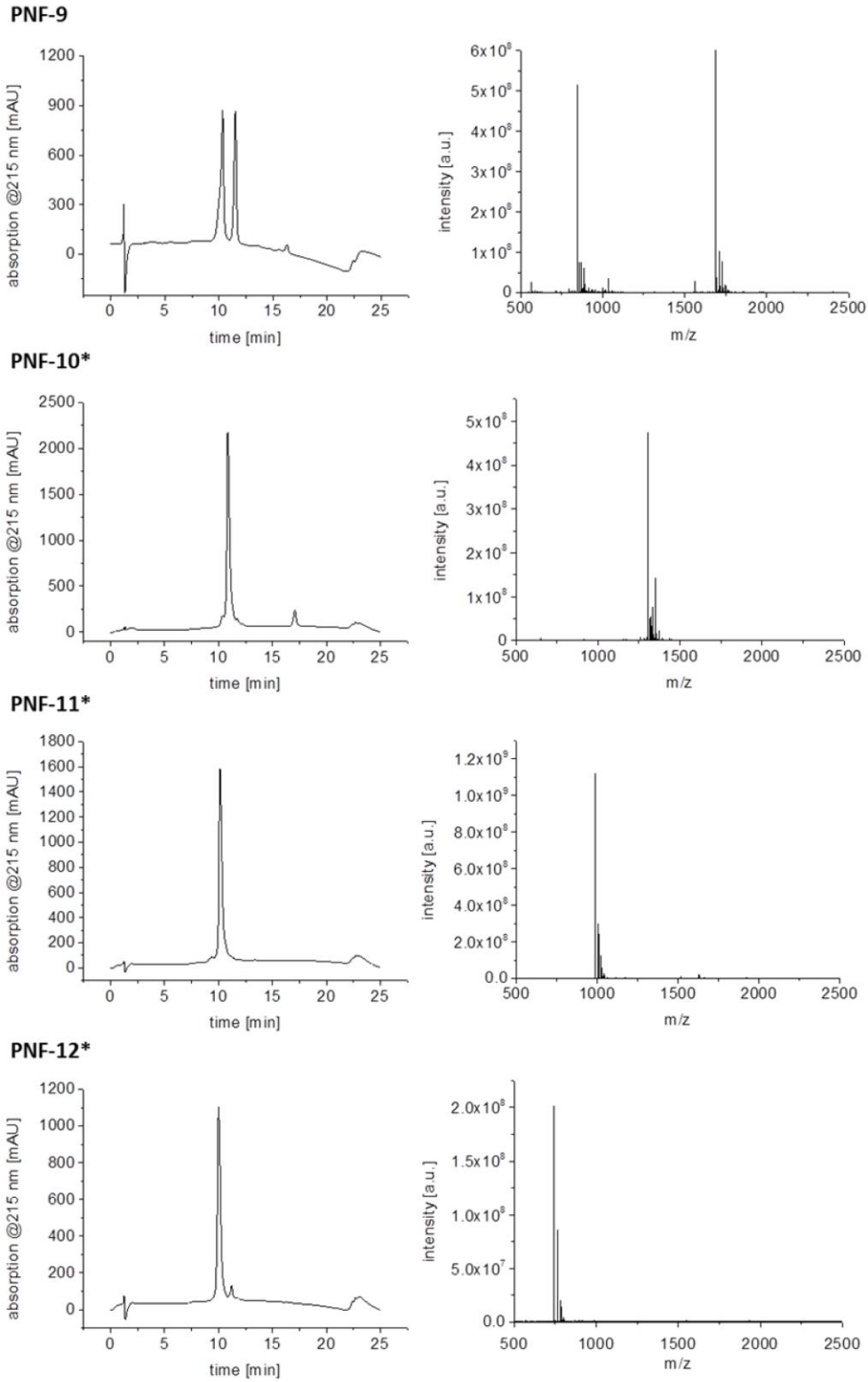


Figure S14: ctnd.

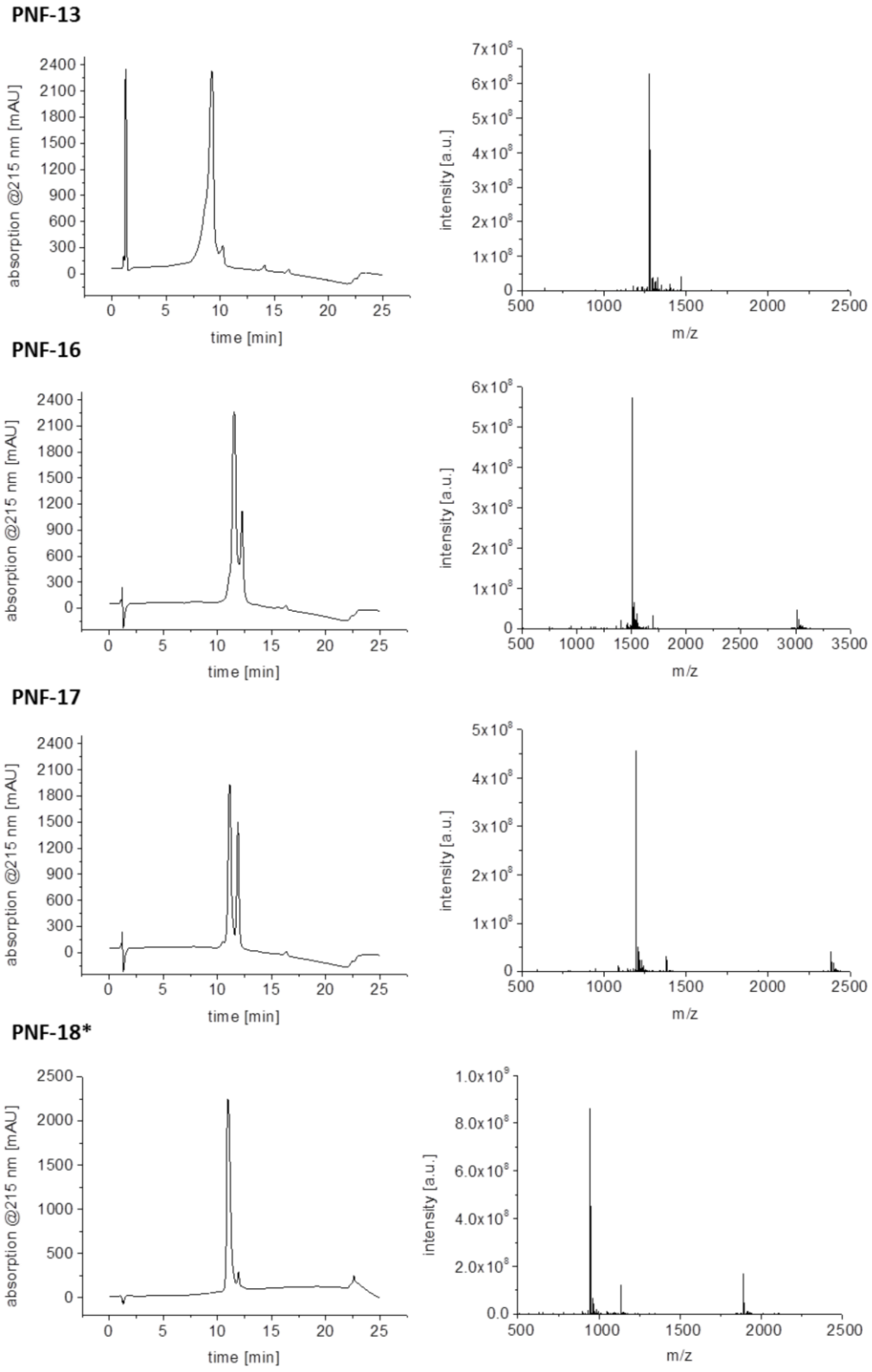


Figure S14: ctnd.

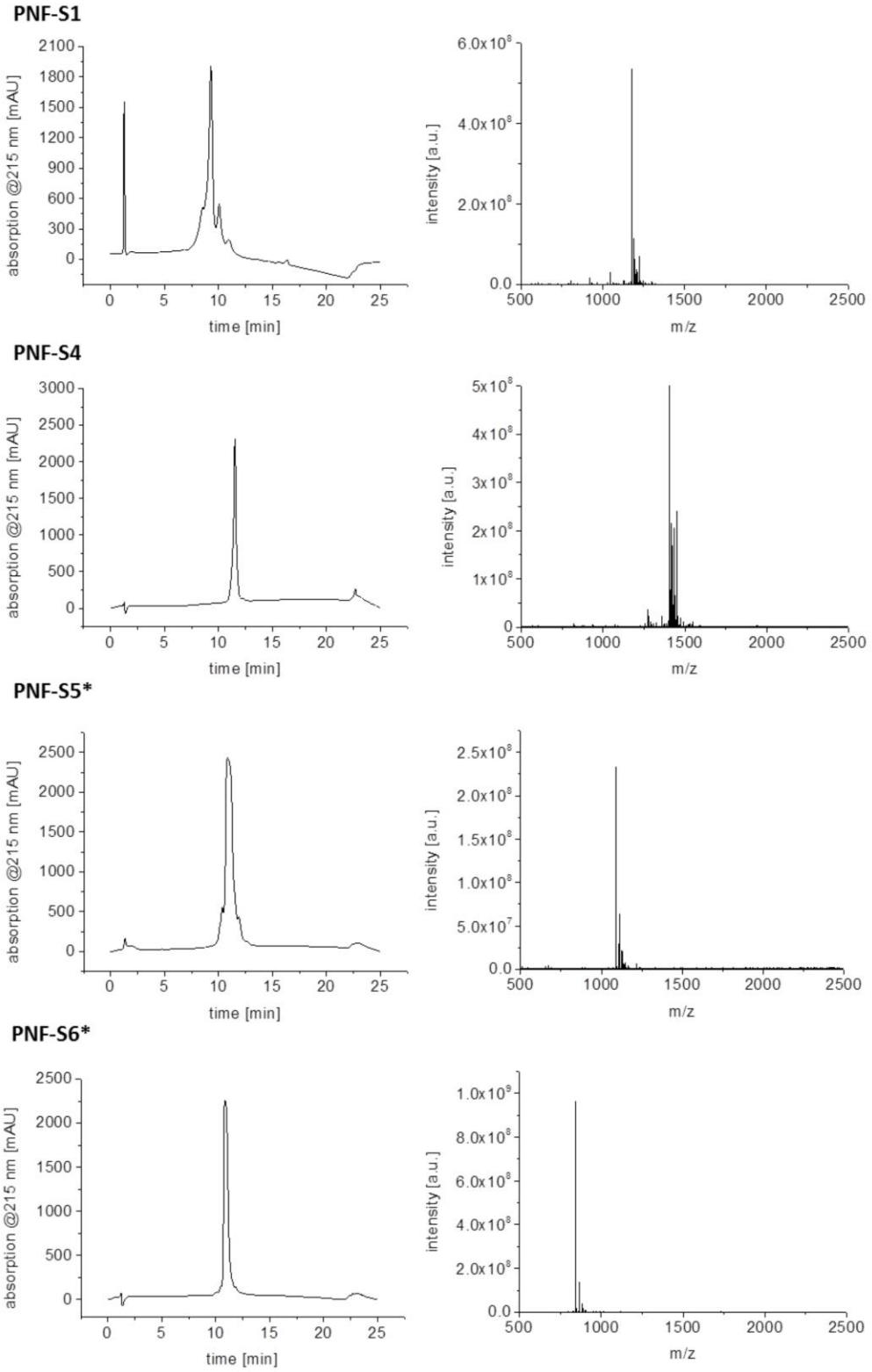


Figure S14: ctnd.

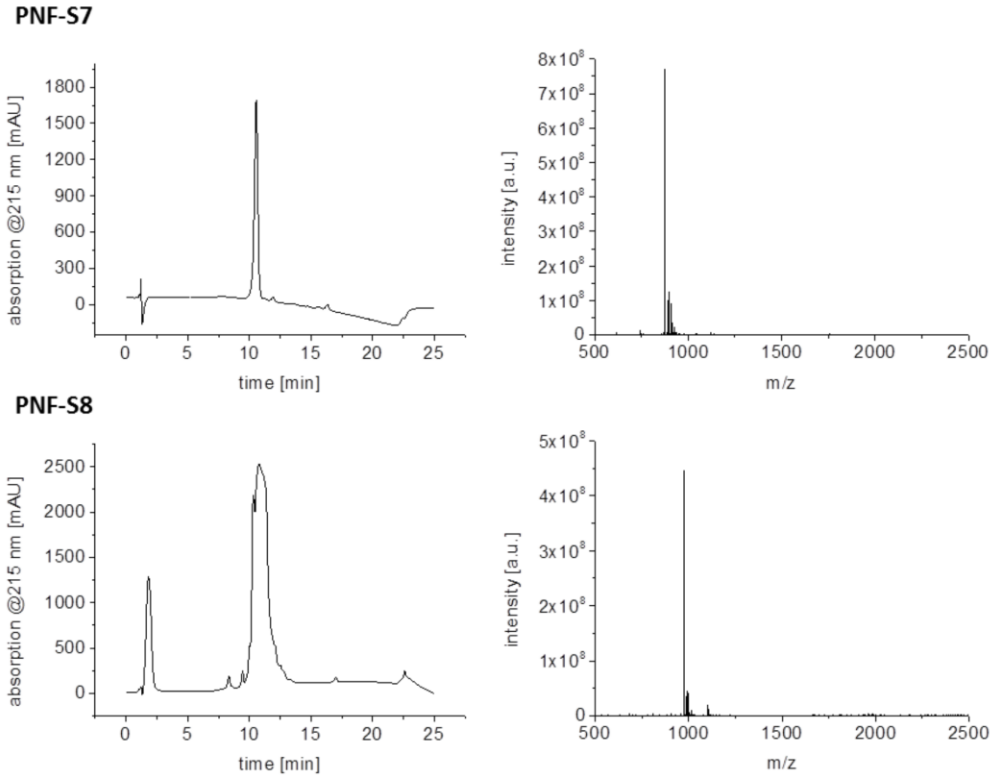


Figure S14: High performance liquid chromatograms of PNFs and corresponding MALDI spectra of purified fractions. Entries marked with * have previously been published.[1]

References

[1] C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil, B. Knöll, *Adv. Funct. Mater.* **2019**, *29*, 1809122.

[2] M. Yolamanova, *Ph.D. Thesis*, Univeristy Ulm, **2016**.

[3] A. Papkalla, J. Münch, C. Otto, F. Kirchhoff, *J. Virol.* **2002**, *76*, 8455.

6.3 Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides

Authors: Kübra Kaygisiz, Lena Rauch-Wirth, Arghya Dutta, Xiaoqing Yu, Yuki Nagata, Tristan Bereau, Jan Münch, Christopher V. Synatschke*, Tanja Weil*
* corresponding author
Accepted for publication in: Nature Communications
DOI: DOI: 10.1038/s41467-023-40663-6
Author Contributions:
Kübra Kaygisiz conceived the idea, conceptualized the study, designed, and carried out all experiments except assays related to virus transduction and SFG. She analyzed all data, and performed the formal analysis, wrote python scripts for amino acids analysis, and wrote the original draft. Lena Rauch-Wirth designed, carried out the experiments, analyzed data, and performed the formal analysis related to virus transduction, and revised the manuscript. Arghya Dutta wrote parts of the python script and analyzed data related to pattern analysis and revised the manuscript. Xiaoqing Yu measured and analyzed SFG experiments, supervised by Yuki Nagata. Xiaoqing Yu and Yuki Nagata revised the manuscript. Tristan Bereau conceptualized and supervised multi-parameter regression and pattern analysis, provided the funding resources, and revised the manuscript. Jan Münch conceptualized, supervised the assays related to virus transduction, provided the funding resources, and revised the manuscript. Christopher V. Synatschke conceptualized, supervised and revised the manuscript. Tanja Weil conceptualized, supervised the entire study, provided the funding resources, and revised the manuscript.

Copyrights: The following part is accepted for publication in Nat. Commun. **2023**, DOI: 10.1038/s41467-023-40663-6 and licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/). Further distribution of the material is possible if appropriate credit, a link provided to the license, and indications of changes are made.

nature communications

Article

https://doi.org/10.1038/s41467-023-40663-6

Data-mining unveils structure–property–activity correlation of viral infectivity enhancing self-assembling peptides

Received: 27 February 2023

Accepted: 1 August 2023

Published online: xx xx 2023

Check for updates

Kübra Kaygisiz[1], Lena Rauch-Wirth[2], Arghya Dutta[3,5], Xiaoqing Yu[4], Yuki Nagata[4], Tristan Bereau[3,6], Jan Münch[2], Christopher V. Synatschke[1] ✉ & Tanja Weil[1] ✉

Gene therapy via retroviral vectors holds great promise for treating a variety of serious diseases. It requires the use of additives to boost infectivity. Amyloid-like peptide nanofibers (PNFs) were shown to efficiently enhance retroviral gene transfer. However, the underlying mode of action of these peptides remains largely unknown. Data-mining is an efficient method to systematically study structure–function relationship and unveil patterns in a database. This data-mining study elucidates the multi-scale structure–property–activity relationship of transduction enhancing peptides for retroviral gene transfer. In contrast to previous reports, we find that not the amyloid fibrils themselves, but rather μm-sized β-sheet rich aggregates enhance infectivity. Specifically, microscopic aggregation of β-sheet rich amyloid structures with a hydrophobic surface pattern and positive surface charge are identified as key material properties. We validate the reliability of the amphiphilic sequence pattern and the general applicability of the key properties by rationally creating new active sequences and identifying short amyloidal peptides from various pathogenic and functional origin. Data-mining—even for small datasets—enables the development of new efficient retroviral transduction enhancers and provides important insights into the diverse bioactivity of the functional material class of amyloids.

UNCORRECTED PROOF

The stable and safe integration of genetic information is an emerging challenge in cell engineering, e.g., for gene therapy. Retroviral gene transfer is the commonly used method to integrate genetic payloads into cells. Retroviral vectors can deliver therapeutic transgenes with low immunogenicity and hold great promise for the treatment of genetic disorders in recent ex vivo clinical trials[1] and have been approved for chimeric antigen receptor (CAR) T-cell therapy[2]. However, low transduction efficiencies of the viral vectors at target cells

[1]Department Synthesis of Macromolecules, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. [2]Institute of Molecular Virology, Ulm University Medical Center, Meyerhofstraße 1, 89081 Ulm, Germany. [3]Department Polymer Theory, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. [4]Department Molecular Spectroscopy, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. [5]Present address: Institute of Biochemistry II, Faculty of Medicine, Goethe University, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany. [6]Present address: Institute for Theoretical Physics, Heidelberg University, Philosophenweg 19, 69120 Heidelberg, Germany.
✉e-mail: synatschke@mpip-mainz.mpg.de; weil@mpip-mainz.mpg.de

Article https://doi.org/10.1038/s41467-023-40663-6

resulting from low virus–cell attachment rates are still a major challenge that limits therapeutic efficacy[3,4].

Various materials such as polymers, lipids, and peptides have been introduced as additives to increase viral transduction efficiency mainly by facilitating colocalization of viruses and cells[5]. Interestingly, the transduction-enhancing ability of amyloid peptide nanofibers (PNF) was first found in seminal fluid in the context of human immunodeficiency virus 1 (HIV-1) infection[6]. This finding resulted in intense research on peptide fibrils to boost viral infectivity[7–11]. For example, the 12-mer peptide enhancing factor C (EF-C, QCKIKQIINMWQ) corresponding to residues 417–428 of the HIV-1 envelope glycoprotein 120 (gp120) was found in a random screening approach and spontaneously self-assembled into positively charged amyloid fibrils, which strongly enhance retroviral transduction[12]. Due to their versatile material properties, amyloid nanostructures have attract broad interest for various therapeutic applications, such as gene therapy[12] or as additives in vaccines[13].

Many peptides and proteins have a propensity to assemble into hierarchically ordered amyloid structures and Nature uses these structures in both, pathogenic and functional contexts[14,15]. A hypothesis that distinguishes pathogenic from functional amyloids is based on the kinetics of their assembly process. It has been proposed that oligomeric species rather than mature, thermodynamically stable amyloid fibrils are related to cytotoxicity[16–18]. Oligomers that prevail for a long time in solution can interact and also refold other monomeric proteins thus acting as multipliers for proteins misfolding and aggregation[19,20]. In contrast, reports on functional amyloids typically show very fast and quantitative assembly without long lag times, thus preventing the formation of cytotoxic oligomers[21,22]. The tensile strength and stability of amyloid fibrils are attributed to their high cross β-sheet structures, which are mainly stabilized by hydrogen bonding and hydrophobic interactions[23]. The defined molecular structure of the peptide building blocks enables modifications of the sequence at the level of single amino acids, which gives access to tuneable assembly features and resulting materials properties from the nano- to the microscale (structure–property), resulting in programmable bioactivity (property–activity)[24].

The observation that certain functional amyloids are biocompatible and exhibit low cytotoxicity combined with virus-binding properties makes them a promising class of materials for retro- and lentiviral gene therapy[12,25,26]. In previous studies, we and others have shown via electron and fluorescence microscopy that virus-particles associate to fibrillar structures[12,26,27]. However, a fundamental understanding of the corresponding molecular as well as nanoscopic property–activity relationship is limited by inherent multi-scale and multi-parameter complexity of the supramolecular self-assembling systems. Data-mining and regression models are efficient tools to study structure–function relationship and unveil underlying design principles and patterns in a database[28,29]. However, so far, the relationship between properties and bioactivity of peptides has been elucidated mainly for antimicrobial[30] or antiviral peptides[31], while they are still elusive for self-assembling nanostructured peptides with transduction enhancing properties.

We have designed and synthesized a peptide library consisting of 163 sequences to study structure–property–activity relationships for infectivity enhancement of nanostructures composed of short peptides. Based on the known and commercially available self-assembling 12-mer peptide EF-C (Fig. 1A)[12], a smaller library was reported by us containing optimized peptide sequences with improved bioactivity[26,32]. However, a fundamental understanding of the structure–property–activity relationship of amyloid peptides and their unique infectivity is still lacking. Therefore, we expand the peptide library by scrambling and substituting amino acids, point mutations, peptide sequence truncations, extensions with bioactive epitopes, and random peptide sequences (Table S2). The resulting peptide library includes peptides with sequence lengths ranging from 4 to 19 amino acids (Table S3, Fig. 1B), whose bioactivity, aggregation, and nanostructure formation were characterized in a quantitative HIV-1 infection assay as well as by transmission electron microscopy (TEM), light scattering count rate, zeta-potential measurements, and Thioflavin T (ThT) fluorescence assays. Collecting this multi-parameter dataset provided the physicochemical properties for morphology, microscopic aggregation, surface charges, and β-sheet amyloid structures, respectively. Together with bioinformatic parameters, these acquired data were applied in a property–activity correlation analysis using multiple linear regression. Microscopic aggregation of already formed amyloid structures with positive surface charges was identified as the key material characteristics for boosting viral transduction. We demonstrated the general understanding and universal applicability of the key properties for pathogenic and functional amyloids, unrelated to EF-C. To the best of our knowledge, this is the first structure–property–activity relationship of infectivity-enhancing amyloid peptide nanomaterials, that enables not only an understanding but also the rational design and discovery of therapeutic and pathogenic peptide sequences. We envision that the presented study provides a fundamental understanding of this important structural class.

Results

Characterization of a peptide library based on EF-C derivatives

To unveil the structure–property–activity relationship of infectivity enhancing peptides, we performed a data-mining approach. This procedure involves data-generation by creating a peptide library, processing and analyses of the peptide measurements, establishing a model for the structure–property–activity relationship, and finally interpretation and prediction based on the established relationship. The peptide library was synthesized via routine solid phase peptide synthesis according to Fmoc-protocol[33] and their purities were confirmed by liquid chromatography mass spectrometry (LC-MS, Fig. S35). Unless stated otherwise, the sequences were obtained in >95% purity by LC-MS from commercial suppliers (Table S3). To induce fibril formation, the peptides were first dissolved in dimethyl sulfoxide (DMSO) and then diluted in phosphate-buffered saline (PBS). The physicochemical properties of the peptides were analyzed on multiple length scales (Fig. 1C). On nanometer scale, the morphologies of molecular assemblies were investigated by TEM and were categorized in (1) fibrillar and (2) non–fibrillar structures (SI Section 1.1, Fig. S1, Fig. S33). Fibril formation was further assessed by the amyloid sensitive ThT dye. Enhancement of fluorescence intensity was accounted to ThT binding to β–sheet rich fibrils (SI Section 1.1)[34]. Both fibril formation and ThT-activity of peptides were evaluated as binary features (fibrils/no fibrils and ThT-active/not ThT-active, Table S5). β–sheet content was determined via Fourier transform infrared spectroscopy (FT–IR) analysis (Fig. S34) and efficiency to assemble into nanostructures was determined via the conversion rate (CR, Table S5) assay as reported earlier[26]. On the micrometer scale, zeta-potential measurements were conducted to gain information on surface charge of peptide assemblies. Another parameter taken from the zeta-potential measurements was the derived count rate of scattered light, which gives information on turbidity of the sample due to aggregation of fibrillar assemblies (SI Section 1.1, Table S5)[35]. The microscopic aggregation of PNF was further characterized for selected examples via brightfield and fluorescence microscopy (Fig. S32). In this study, we use the term microscopic aggregation to describe the fiber-fiber interaction into less ordered micrometer-sized particles and exclude the highly ordered μm-sized crystalline alignments that have been previously studied for amyloid aggregation[36–38]. Bioinformatic parameters for the library are calculated with amino acid properties to yield global peptide properties like hydrophobicity or isoelectric point (SI Section 1.2, Table S4)[39,40].

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

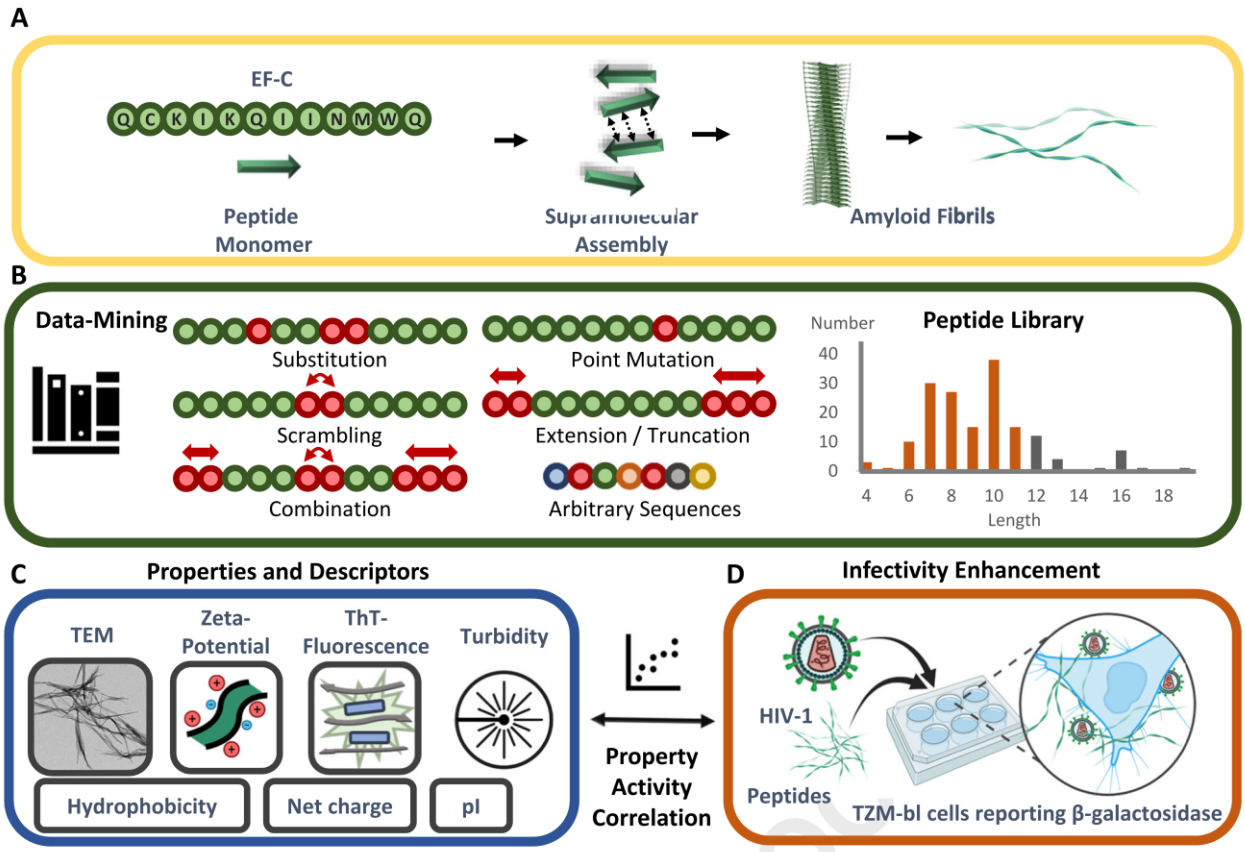


Fig. 1 | Schematic visualization of this work to identify property-activity relationships of peptides enhancing retroviral transduction enhancement. A EF-C is a 12-mer peptide sequence, derived from the glycoprotein gp120 of the human immunodeficiency virus 1 (HIV-1), that can assemble via supramolecular interactions into amyloidal fibrils with high cross-β sheet content. **B** Amino acids of the sequence of EF-C were systematically substituted, point mutated, or scrambled. Sequence extensions and truncations have been conducted by repeating terminal amino acids or introducing bioactive peptides. Randomly generated peptides were added to the library with no connection to EF-C. The resulting library consists of 163 peptides, mostly comprising fewer than 12 residues (Table S3). **C** Experimental methods, such as TEM (Fig. S33), zeta-potential, ThT fluorescence (Table S5), turbidity analysis (Table S8) as well as bioinformatic characterization such as hydrophobicity, net charge and isoelectric point (pI) (Table S4) were applied to correlate with (**D**) the infectivity enhancement of HIV-1 on TZM-bl cells by the peptides, as quantified via a chemiluminescence based assay (Table S5). HIV-1 and peptides were incubated together before they were added to cells. Created with BioRender.com. Source data for **B** is provided within the Source Data file.

As an indicator of biological activity, the transduction enhancement of the self-assembled peptides was assessed by determining HIV-1 infection rates of TZM–bl cells (Fig. 1D). This cell type is a widely applied model cell line for high throughput analysis of HIV-1 infection, because it expresses the HIV-1 receptors CD4 and CCR5 and reports β–galactosidase upon successful infection[41]. HIV-1 was mixed with different concentrations of peptides and subsequently used to infect TZM–bl cells. The infection rates were assessed with a chemiluminescence–based assay, which allowed quantification of the infection rates relative to virus-only infectivity. For example, a 1.3 μM concentration of EF-C fibrils enhanced HIV-1 infection by a factor of 42 ± 7, as compared to infection rates of the virus only control (Table S5). In this study, EF-C at 1.3 μM concentration was selected as a reference, to compare quantitative infection data obtained for every peptide among various measurements (Table S5). Additionally, the cell viability was determined for a representative selection of 77 peptides at a concentration of 1.3 μM and revealed no toxicity (Figs. S2, S3).

The herein investigated peptide library based on EF-C derivatives consists of peptides, which are similar in sequence (Table S2), but show a wide variety of biological activity (*i.e.* infectivity enhancement), which allow to study structure–property–activity relationships. Peptides showing an *n*–fold infectivity enhancement of <4–fold relative to virus-only infectivity or <10% relative to EF-C are considered as inactive, which is the case for 49% of the peptide library (Table S5).

Microscopically aggregated amyloid fibrils and positive surface charge highly correlate with transduction enhancement

To elucidate the structure–property–activity relationship, we applied linear regression on the experimental and bioinformatic descriptors and evaluated the Pearson correlation coefficient (Pearson *R*). Notably, for most of the descriptors, the logarithm of the infection enhancement relative to EF-C (Log_{10} Infection Rel. EF-C) correlates with a higher Pearson *R* compared to absolute infection enhancement values. To enable linear regression, we therefore applied logarithmic scale for the infectivity relative to EF-C for all the following correlations.

Fibril formation (Fig. 2A, Fig. S5 $R = 0.59$), ThT–fluorescence (Fig. 2B, Fig. S5 $R = 0.40$), and hydrophobicity according to Fauchère hydropathy scale (Fig. S5, $R = 0.40$) correlate strongly with infectivity enhancement. Our data indicate that a high hydrophobicity is important to achieve quantitative assembly of peptides into fibrils (Fig. S6). The importance of the latter is further demonstrated with a representative selection of 80 peptides (Table S5), which we additionally characterized regarding the conversion of monomers to fibrils (Fig. S5, conversion rate, $R = 0.50$, Fig. S7 B, Table S5).

Previous reports indicated that the zeta-potential representing the surface charge of the fibrils plays a crucial role for transduction enhancement[6,12,25,27]. Plotting the zeta-potential against Log_{10} Infection relative to EF–C shows indeed the strongest linear correlation ($R = 0.67$, Fig. 2C, Fig. S5, Table S5). Another comparably high

Article https://doi.org/10.1038/s41467-023-40663-6

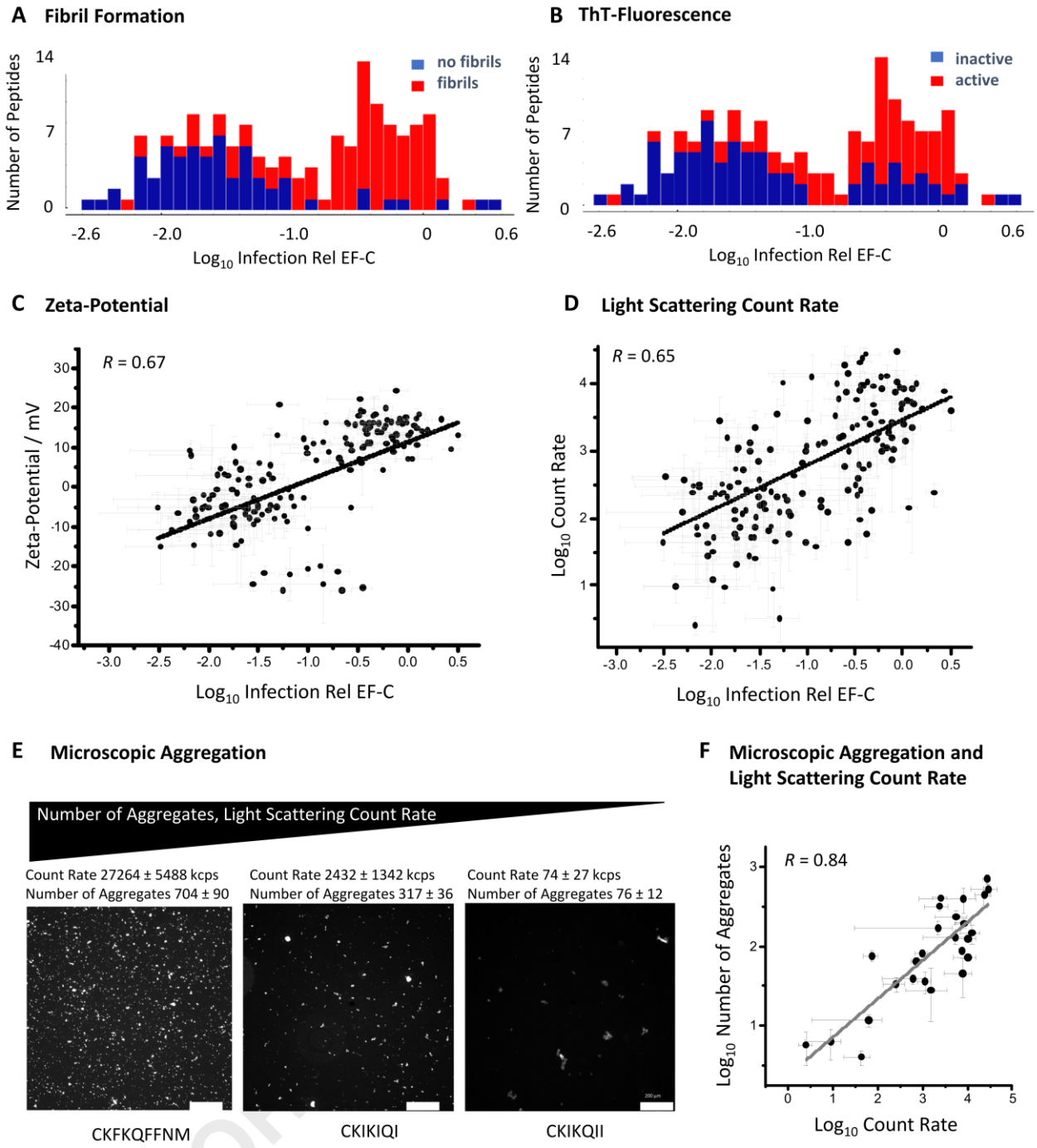


Fig. 2 | Correlation of experimental parameters with Log_{10} Infection Rel EF-C of a library containing 163 peptides. A Histogram of Log_{10} Infection Rel EF-C showing fibril forming (red) and non-fibril forming (blue) peptide distribution analyzed via TEM (Fig. S33). **B** Histogram of Log_{10} Infection Rel EF-C showing ThT-active (red) and ThT-inactive (blue) peptide distribution (Table S5). **C** Zeta-potential plotted against Log_{10} Infection Rel EF-C. Mean values are displayed with error bars indicating standard deviation from triplicate measurements (Table S5). Linear fit with Pearson correlation factor R = 0.67. **D** Count rate of scattered light plotted against Log_{10} Infection Rel EF-C. Linear fit with Pearson correlation factor R = 0.65. Mean values are displayed with error bars indicating standard deviation from triplicate measurements (Table S5). **E** Selection of peptides with high and low count rate showing relationship between number of particles and count rate of scattered light (Table S8). Peptide fibrils were preincubated at 1 mg/mL one day and diluted to 0.1 mg/mL with ThT solution (50 μM) just before measurement (Fig. S32). Number of aggregated fibrils (> 10 $μm^2$) were quantified via fluorescence microscopy in an area of 1330 μm × 1330 μm in triplicate measurements via automated counting (3D object counter ImageJ), scale bar 200 μm. **F** Log_{10} Count rate from zeta-potential measurements of representatively selected peptides plotted against averaged number of aggregated particles in an area of 1330 μm × 1330 μm (Table S8). Mean values are displayed with error bars indicating standard deviation from triplicate measurements, linear fit (solid line) demonstrates a high correlation with a Pearson R of 0.84. Source data for **A**–**D** and **F** is provided within the Source Data file.

Article https://doi.org/10.1038/s41467-023-40663-6

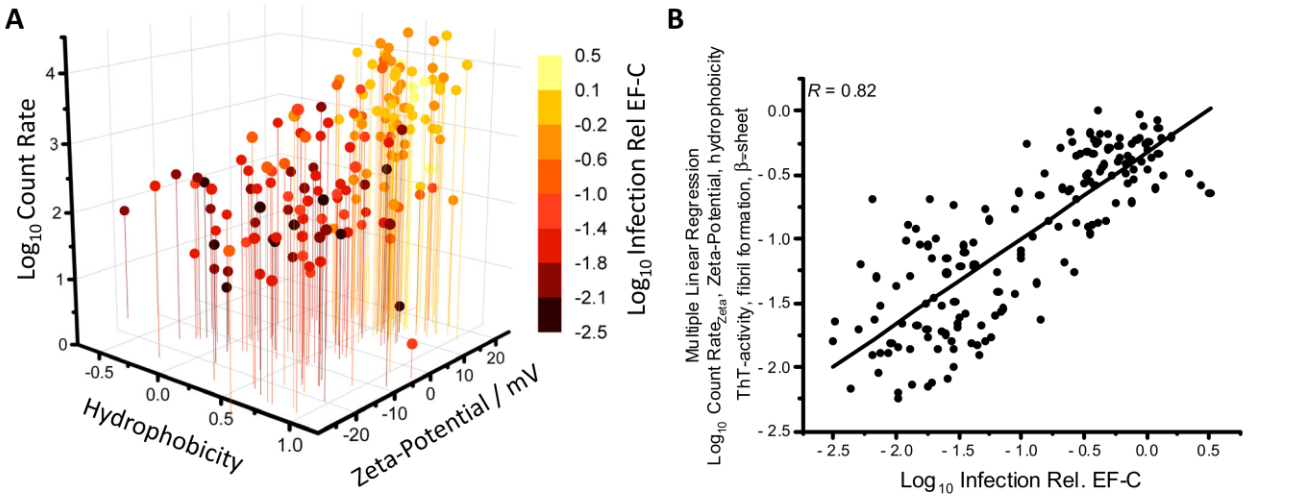


Fig. 3 | Multiple parameter analysis to evaluate property-activity relationship between physicochemical and bioinformatic descriptors and infectivity enhancement. A Four-dimensional plot visualizes that a positive zeta-potential (Table S5), high hydrophobicity (Table S4), and high count rate of scattered light (Table S5) are corresponding to infectivity enhancement (color coded Log_{10} Infection Rel. EF-C, Table S5). **B** Multiple linear regression (constrained linear least-squares problems solution) with the equation: y = A0 + A1*x1 + A2*x2 + [...]+ A5 * x5 with best performing single parameters of 163 peptides based on EF-C (Fig. S5, Table S4, S5): Log_{10} Count Rate, zeta-potential, hydrophobicity, ThT-activity, fibril formation, β-sheet and Log_{10} Infection Rel EF-C as dependent target variable show an improved linear correlation by 22% compared to the best correlating feature zeta-potential. Linear Fit with Pearson correlation factor R is 0.82. The found relationship is Log_{10} Infect Rel Ef-C = −2.33462 + 0.02128 (zeta-potential) + 0.29879 (Log_{10} Count Rate) + 0.27355 (fibril formation) + 0.26241 (hydrophobicity) + 0.10744 (ThT-activity) + 0.00356 (β-sheet). Source data for Fig. 3 A, B is provided within the Source Data file.

correlating property is microscopic aggregation of the peptides, which was measured via the count rate of scattered light during zeta-potential measurement (Log_{10} Count Rate, $R = 0.65$ Fig. 2D, Fig. S5, Table S5, Table S8). The timescale for the aggregation of peptides is within seconds as can be observed via TEM and turbidity from light scattering (Fig. S4). Although there is no suitable model for correlating the scattering intensity of light with the size of PNFs (SI Section 1.1), our data shows that the logarithmic number of μm−sized aggregated fibrils determined by fluorescence microscopy (Fig. 2E, Fig. S32) correlates linearly with the logarithmic scattering intensity (Fig. 2F, Table S8), which further validates the relationship between the microscopic aggregation and infectivity.

Despite the high correlation of zeta-potential and count rate with infectivity enhancement (Fig. 2C, D, Table S5), some data points deviate from the regression line. Moreover, not all bioactive peptides also form fibrils—for example, 7 infectivity-enhancing peptides from the library (Fig. 2A) form β-sheet rich amorphous aggregates instead. Interestingly, these peptides show a high calculated hydrophobicity due to large amounts of tryptophan (W) and tyrosine (Y) residues (Fig. S16), which indicates that multiple parameters are codependent of each other and cannot be covered from single parameter models to describe the property−activity relationship sufficiently.

Synergistic effect of hydrophobicity, charge, and microscopic aggregation

We hypothesized that several conditions must be fulfilled to observe peptide−virus−cell interactions, which ultimately leads to enhanced viral transduction. So, in a next step, we considered multiple descriptors at once to identify the relevant physicochemical properties. Plotting the best correlating continuous features (zeta-potential, count rate, hydrophobicity) against Log_{10} infection relative to EF-C directly visualizes that a positive zeta-potential, high count rate and high hydrophobicity are coinciding with infectivity enhancement (Fig. 3A, Table S5).

In fact, applying multiple linear regression on the best correlating features revealed a superior Pearson correlation parameter ($R = 0.82$) compared to the single features alone, an increase by 22% compared to best correlating single feature zeta-potential ($R = 0.67$) (Fig. 3B, Table S5). A strong effect of multi-parameter considerations is observed for the β−sheet content, which was previously identified as a critical feature for infectivity in a small data set[26]. In our extended peptide library, we observe that the presence of β-sheet rich structures is crucial for activity only if we apply multi-parameter regression ($R = 0.91$, Fig. S10D, E) but not from single-parameter correlations ($R = 0.29$, Fig. S5, Table S5). This is because peptides with similar β−sheet content are not enhancing transduction efficiency to the same extent (Fig. S7) because they either do not show positive surface charge, do not form fibrils or do not aggregate (Fig. S10A−C).

Applying multi-parameter regression has its strongest impact on correlation of the bioinformatic parameters. For example, the highest single-parameter bioinformatic correlation can be found for hydrophobicity ($H_{fauchere}$, $R = 0.40$, Fig. S5). However, in combination with all other calculated descriptors the correlation coefficient increases by roughly 70% (Fig. S11, R = 0.67, Table S4). The advantage in applying multiple linear regression is, that each contribution is weighted and reveals the importance of combined properties for the activity. In this way a statistically significant ($p < 0.05$) correlation between zeta-potential, count rate, fibril formation, and infectivity enhancement (Fig. S10D) for the experimental parameters and hydrophobicity ($H_{fauchere}$), aliphatic index (AI), hydrophobic moment index (HMI), instability index (II), graph shape index (GSI), upsilon steric parameter (USP), polarizability (P) and normalized van der Waals Volume (Norm vdW Vol.) (Fig. S11) for the bioinformatic parameters are identified. We propose that next to hydrophobicity, also other global peptide features (P, AI, HMI, USP, Norm. vdW Vol., GSI and II) correlate significantly with infectivity enhancement ($p < 0.05$) but with a comparably small correlation coefficient (Figs. S5 S10, and S11). These parameters calculate different aspects of the amino acid side chain hydrophobicity, charge, sterical demand, and the sequence amphiphilicity (SI Section 1.2).

A differentiated perspective on the mode of action

To showcase the underlying codependences of the internal order (β−sheet and fibril formation), microscopic aggregation, and zeta-

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

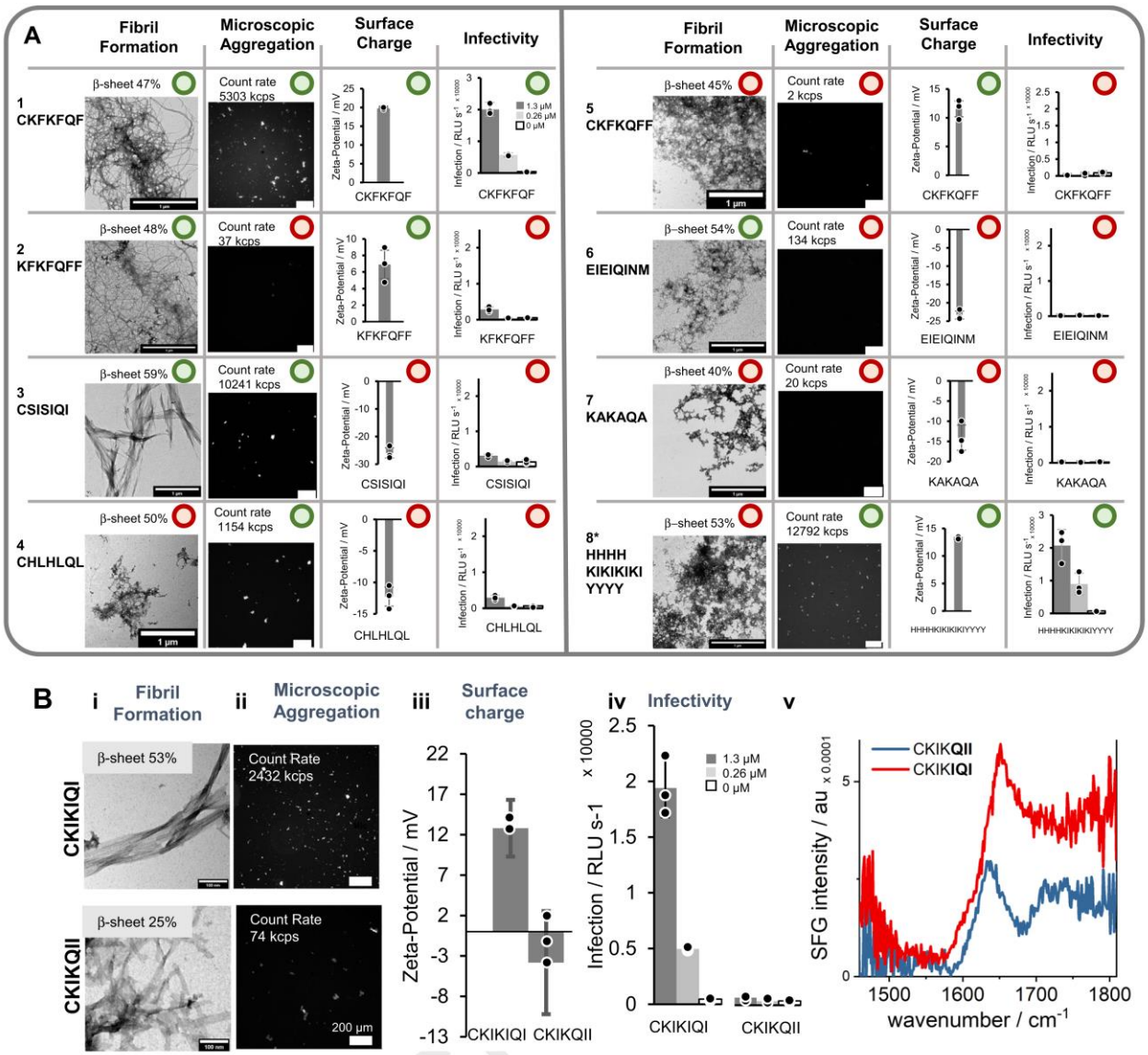


Fig. 4 | Evaluation of global property–activity relationship for infectivity enhancement with selected peptides as case studies to gain insight into the mode of action. Internal order is evaluated by TEM (Fig. S33, Table S5, fibril formation) and FT-IR (Fig. S34, Table S5 β-sheet content), surface charge by zeta-potential and infectivity is shown in absolute infection rates of peptides at 1.3, 0.26 and 0 μM (Table S5). All displayed error bars indicate standard deviation from triplicate measurements. **A** Overview of all possible combinations of physicochemical features, which correlate most strongly with infectivity enhancement with examples for each possible combination. 1 The majority of peptides (75%), which enhance infectivity have high internal order, aggregate into μm-sized aggregates, and have a positive zeta-potential, as shown exemplarily for the peptide CKFKFQF. 2 Without microscopic aggregation there is no infectivity enhancement even though other prerequisites are fulfilled such as fibril formation and positive zeta-potential shown for KFKFQFF. 3 Peptide fibrils which aggregate to clusters but have a negative surface charge are not enhancing viral infection, shown for CSISIQI. However, if the peptides are highly hydrophobic slight infection enhancement can be observed (e.g. for EIEIQINMWQ, Fig. S14). Peptides which do not fulfill at least two of the features (assembly, aggregation, charge) cannot enhance infectivity as exemplarily shown for 4 (CHLHLQL or CKFKFQF with EGCG, 5 CKFKQFF 6 EIEIQINM, 7 KAKAQA. 8* Peptides with a high amount of hydrophobic amino acids (W, Y) can enhance transduction although they are not assembling into fibrils if they form β-sheet rich, positively charged aggregates e.g. shown for HHHHKIKIKIKIYYYY. **B** Comparison of physicochemical properties and activity of CKIKIQI and CKIKQII. i TEM micrographs and β-sheet content determined by FT-IR (Fig. S34), scale bar 1 μm. ii Fluorescence microscopy of ThT-stained peptide fibrils with their respective light-scattering count rate (Table S5), scale bar 200 μm. iii Zeta-potential iv Absolute infection rates of peptides at 1.3, 0.26 and 0 μM concentration. v SFG spectroscopy at the amide I and II region. Source data for **A**, **B** is provided within the Source Data file.

potential in the observed property–activity relationship, we exemplarily selected individual peptides (Fig. 4A, B, SI Section 6).

The most prevalent case (38% of the library, 75% of all active peptides, Table S5) for infectivity enhancement are peptides assembling into fibrils with β-sheet secondary structure, which further aggregate into positively charged μm–sized aggregates (for example CKFKFQF, Fig. 4A–1).

The importance of these parameters is shown by the lack of activity when removing one of the properties (Fig. 4A). For example, if peptides assemble into positively charged fibrils but do not aggregate into μm–sized clusters (e.g. KFKFQFF, Fig. 4A–2), no strong infectivity enhancement can be observed. In our library, cases where an assembly into fibrils but no aggregation into μm–sized particles occur, are mostly observed for peptides, which do not contain *N*–terminal

Article https://doi.org/10.1038/s41467-023-40663-6

cysteine at the first position (Fig. S13, SI Section 5.1). For example, the peptides CKIKIQI and KIKIQIC only differ in the position of cysteine within their sequence, but CKIKIQI shows stronger aggregation and infectivity enhancement than KIKIQIC (Fig. S14). In contrast to lysine, cysteine is an amino acid with a hydrophobic residue, and might increase the surface hydrophobicity of the fibrillar structure when positioned at the *N*–terminus. The thiol group of the cysteine side chain can further form disulfide bonds, which might increase inter-fibril attachment and aggregation.

Small changes in molecular sequence can induce a strong change in microscopic aggregation properties as observed for example for CKIKQII and CKIKIQI. Sum–frequency generation (SFG) measurements can give information on water interaction of fibrils, which is dependent on surface hydration and hydrophobicity properties of fibrils[42,43]. To investigate changes in water–fibril interaction the peptides CKIKQI and CKIKQII were selected as examples since they are composed of same amino acids in different order, form fibrils and differ significantly in their aggregation behavior (Fig. 4B). The SFG measurements show that while both peptides accumulate at the air/water interface, the signal intensity of CKIKIQI is higher than CKIKQII. This indicates that CKIKIQI is more ordered or may be present at higher concentrations at the air/water interface (Fig. 4B). Both effects can be traced back to a higher surface hydrophobicity and lateral interaction of fibrils in μm–sized aggregates, which in turn are a prerequisite for infectivity enhancement.

Noteworthy, not every parameter in this relationship must be fulfilled to observe any infectivity enhancement. For example, while the majority (90%) of infectivity-enhancing peptides (Infection relative to EF-C > 10%) forms fibrils, some of the active peptides show non–fibrillar morphologies (Fig. 2A). These peptides contain a high amount (> 20%) of hydrophobic amino acids such as W or Y and form amorphous amyloid aggregates with high β–sheet content (Table S5). Additionally, they fulfill other prerequisites such as positive zeta-potential and microscopic aggregation (Fig. S16). Fibril formation is strongly connected to aggregation as it likely facilitates entanglement but appears not to be an indispensable prerequisite for infectivity enhancement (Figs. 4A–8, S16, and S19).

As already extracted from the global correlation, the surface charge plays a key role. Only weak infectivity enhancement is observed for peptides with negative zeta-potential even if the fibrils with high β–sheet content are aggregating (Fig. 4A–3, Table S5). However, as it also can be extracted from the single parameter correlation with zeta-potential (Fig. 2A, Table S5), five peptides with negative zeta-potential show a moderate infectivity enhancement larger than 10% Rel EF-C (Log_{10} Infect > −1) (Fig. S15). We hypothesize that these peptide fibrils can slightly enhance infectivity because of their extraordinarily high hydrophobicity (> 0.6) and microscopic aggregation of fibrils, which favor interactions with viruses and cells despite the electrostatic repulsion. The importance of considering hydrophobic interactions in addition to electrostatic attraction is further emphasized by a recent report from us, in which we applied machine learning approach to target different sequence spaces[44]. An overall lower hydrophobicity—be it by substituting phenylalanine (F) or isoleucine (I) amino acids with alanine (A) or leaving out hydrophobic amino acids—results in loss of activity (SI Section 5.3–5.4, Figs. S17–S20), further emphasizing its impact.

Activity is also lost, when more than one of the criterions (fibril formation, positive zeta-potential, aggregation) we identified is not fulfilled (Fig. 4A4–7). Peptides that only show μm–sized aggregation (Fig. 4A–4), positive zeta-potential (Fig. 4A–5), or β–sheet rich–fibrils (Fig. 4A–6) do not enhance viral infectivity. To illustrate that the activity of a peptide can be solely traced back to physicochemical properties while the sequence stays same, we added a chaotropic substance (Fig. S21)[45]. Adding epigallocatechin gallate (EGCG) to CKFKFQF drastically decreases infectivity enhancement, β–sheet, and surface charge while microscopic aggregation still takes place (Fig. S21). This once more demonstrates that positive surface charge of microscopically large aggregates with a β–sheet structure are prerequisites for activity.

In summary, our findings show that higher hydrophobicity of assembled peptides drives microscopic aggregation, which is also a prerequisite for virus–cell membrane interaction and colocalization additional to electrostatic attraction. Thereby, we conclude that multiple conditions including (I) a positive surface charge, (II) ordered structure, (III) microscopic aggregation are required for highly efficient retroviral transduction enhancement (Fig. 5).

Design rules: alternating, amphiphilic sequences patterns

We wondered if a sequence pattern exists, that underlies the identified key physicochemical properties and results in infectivity enhancement. To find common design rules which are preserved in active peptides, the sequences in the library were simplified by coarse–graining the amino acids according to their side chain hydropathy (Hydrophilic P, Hydrophobic H) and charge (−, 0, +). The distribution of recurring patterns in highly active peptides (Infect. Rel. EF-C > 70%) was counted (SI Section 7, Fig. S22A)[46]. The majority (55%) of the active sequences and 44 peptides in the library (163 peptides, Table S3) were composed of alternating amphiphilic sequences $P^+H^0P^0H^0P^0H^0P^0$ or $P^+H^0P^+H^0P^0H^0P^0$ (Fig. 6A).

Interestingly, an alternating amphiphilic sequence does not guarantee high transduction efficiency: 8 out of 44 peptides have an alternating amphiphilic sequence but do not show transduction enhancement. This is likely due to their low sequence hydrophobicity (Fig. S23A). To systematically study the impact of hydrophobicity on activity, amino acids from highly active sequences (W, M, N, I, C, Q, K, F and Y, Fig. S22B, Fig. 6A, B) and from weakly active peptides (S, G, R, H, T, V, P, A and L, Figs. S22B and S23B)[46] were used to representatively create peptides which share the same amphiphilic pattern ($P^+H^0P^0H^0P^0H^0P^0$ or $P^+H^0P^+H^0P^0H^0P^0$) and test their infectivity enhancement. Noteworthy, we observed that most of the herein-found amino acids from highly active sequences and weakly active sequences have been known to promote and break β-sheet fibril formation, respectively (SI Section 7).

This approach was exceptionally effective in generating new and active sequences: Eight of the 10 newly created peptides composed of the identified pattern and amino acids prevalent in highly active sequences show infectivity enhancement (Fig. 6B, Table S6). Out of these only the W–rich peptide RFWFWFW does not form fibrils (Figs. 6B and S25) but fulfills all other important requirements for infectivity enhancement—especially a high hydrophobicity (Fig. 6B). Reducing the number of W by one decreases hydrophobicity (1.58 → 1.18) and results in loss of the infectivity enhancing properties as shown for (RINFWFW, Fig. 6B and E). A high amount of W is related to more disordered fibrils but increased activity of the peptides—as also observed for other examples such as KIKIWIW → KIWIWIW (Fig. S25) or C–terminal extension with W (Fig. S16 and S20). These examples once more emphasize the codependences of multiple descriptors for infectivity-enhancing peptides and enable the creation of new active sequences in a reliable fashion. Strikingly, in this way a cysteine- and glutamine-rich peptide KCQCQCQ was found, which showed activity in the range of EF-C while being half the length of EF-C (Fig. 6B, Table S6).

The active peptides all share the identified key–properties for infectivity enhancement, which emphasizes the importance of an amphiphilic sequence pattern with sufficiently hydrophobic amino acids and cationic amino acids for aggregation of self–assembled peptide fibrils.

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

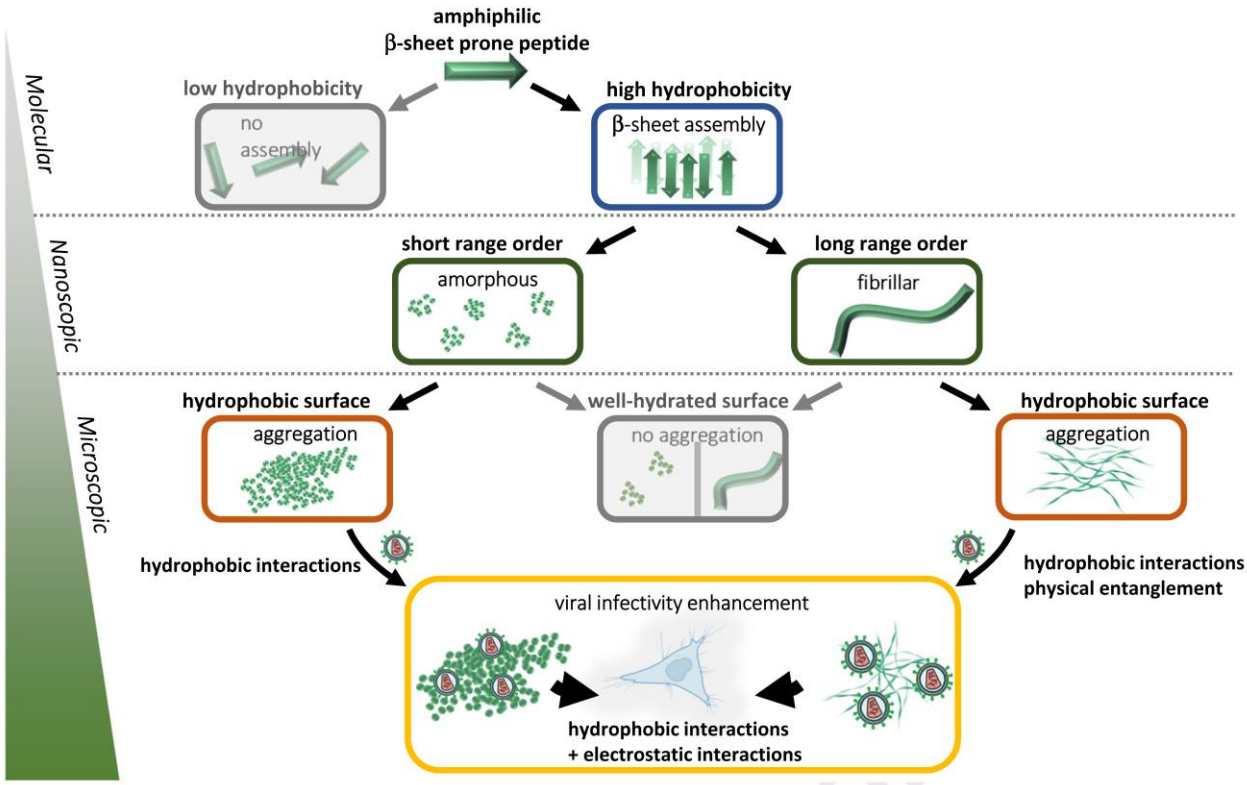


Fig. 5 | Schematic overview of proposed mode of action over multiple scales from molecular to microscopic level. Molecular—Amphiphilic peptides with certain hydrophobicity can form β-sheet rich structures. Nanoscopic—Depending on the hydrophobicity of the β-sheet structures, short-range ordered amorphous aggregates or long-range ordered fibrils can form and further aggregate into μm-sized particles. Microscopic—β-sheet rich aggregates can bind viruses and interact with cellular membrane by electrostatic and hydrophobic interactions. Created with BioRender.com.

Key properties for infectivity enhancement translate to functional and pathogenic amyloids

The fact that aggregation is necessary to observe transduction enhancement may shed light on amyloids in pathogenic contexts. In Nature, microscopic aggregation of amyloidal fibrils is well known as so-called amyloid–plaques, which results in deposition of non-soluble proteins in neurodegenerative diseases as well as facilitating adhesion of bacterial biofilms on substrates[14]. Amyloid plaques from amyloid Aβ occur in several neurodegenerative diseases and may have a connection to neuroinflammation processes. It is speculated that amyloid Aβ is produced by the immune system to combat infections[47,48]. In a downstream process, these amyloid fibrils might form micron–sized aggregates (amyloid plaques) that facilitate infections, leading to stronger infectious reactions—which eventually produce more amyloid fibrils in a vicious cycle. We hypothesized that the property–activity findings from the EF-C library can be translated to other amyloid peptides from functional and pathogenic contexts. To test key–properties for infectivity enhancement, we selected amyloidal peptides from protein databases and literature reports based on published data from transmission electron micrographs and ThT-fluorescence assays (Table S1).

Three of the 24 peptides that were selected, namely ISKLEYSNFSVRY, LANWMCLAKW and VHDCVNITIK, showed infectivity enhancement of retroviral transduction in the range of 30–40% relative to EF-C (Fig. 7A, Table S7). These newly found transduction-enhancing peptides have a positive zeta-potential (Fig. 7B, Table S7, + 9.3 mV, +14.7 mV, +9.3 mV, respectively), microscopically aggregate (Fig. 7B, Fig. S29-B, Table S7), assemble into fibrils (Figs. S36, S37, and 7C–E) with a high β–sheet amount (Figs. S30 and S39, Table S7) and are ThT-active (Fig. S29–A). We studied the interaction of fibrils with HeLa cells by applying confocal fluorescent laser scanning microscopy. The interaction of the amyloidal fibrils with HeLa cells is clearly visible in Fig. 7C–E and shows that these newly discovered infectivity-enhancing peptides act in a comparable fashion as EF-C by associating with cellular membranes and facilitating viral uptake through colocalization[12].

Interestingly, four fibril-forming peptides IKFLSVN, GYVIIK, RQGNINIVA, and MKVIFLKDVKG, that are derived from peroxiredoxin III[49], serum amyloid P[50], insects' chorion[51], and ribosomal protein[52], respectively, have positive zeta-potential in a similar range, but do not enhance infectivity. This is because these peptides do not aggregate to microscopic clusters as visible from low light scattering count rates (Figs. 7B and S29–B) and microscopy data (Fig. S31).

In summary, studying endogenous amyloidal peptides confirms the identified key–properties for transduction enhancement and further emphasizes the importance of microscopic fibrillar aggregates beside the electrostatic interactions to mediate binding of the positively charged fibrils to the negatively charged cell surface.

Discussion

In this study, we aimed to understand the mode of action of short peptides acting as adjuvants to increase viral infectivity. We hypothesized that certain physicochemical properties are responsible for the infection-enhancing properties of peptides. To test this, we applied a data-mining approach that includes the creation of a suitable dataset, the evaluation of corresponding data, and interpretation of found relationships. Therefore, we created a peptide library based on known infectivity-enhancing peptide EF-C. This enables tracing back differences in activity to specific physicochemical properties because sequences stay relatively similar.

Since the interactions between fibrils, viruses and cells occur on multiple length scales we focus on applying complementary methods

Article https://doi.org/10.1038/s41467-023-40663-6

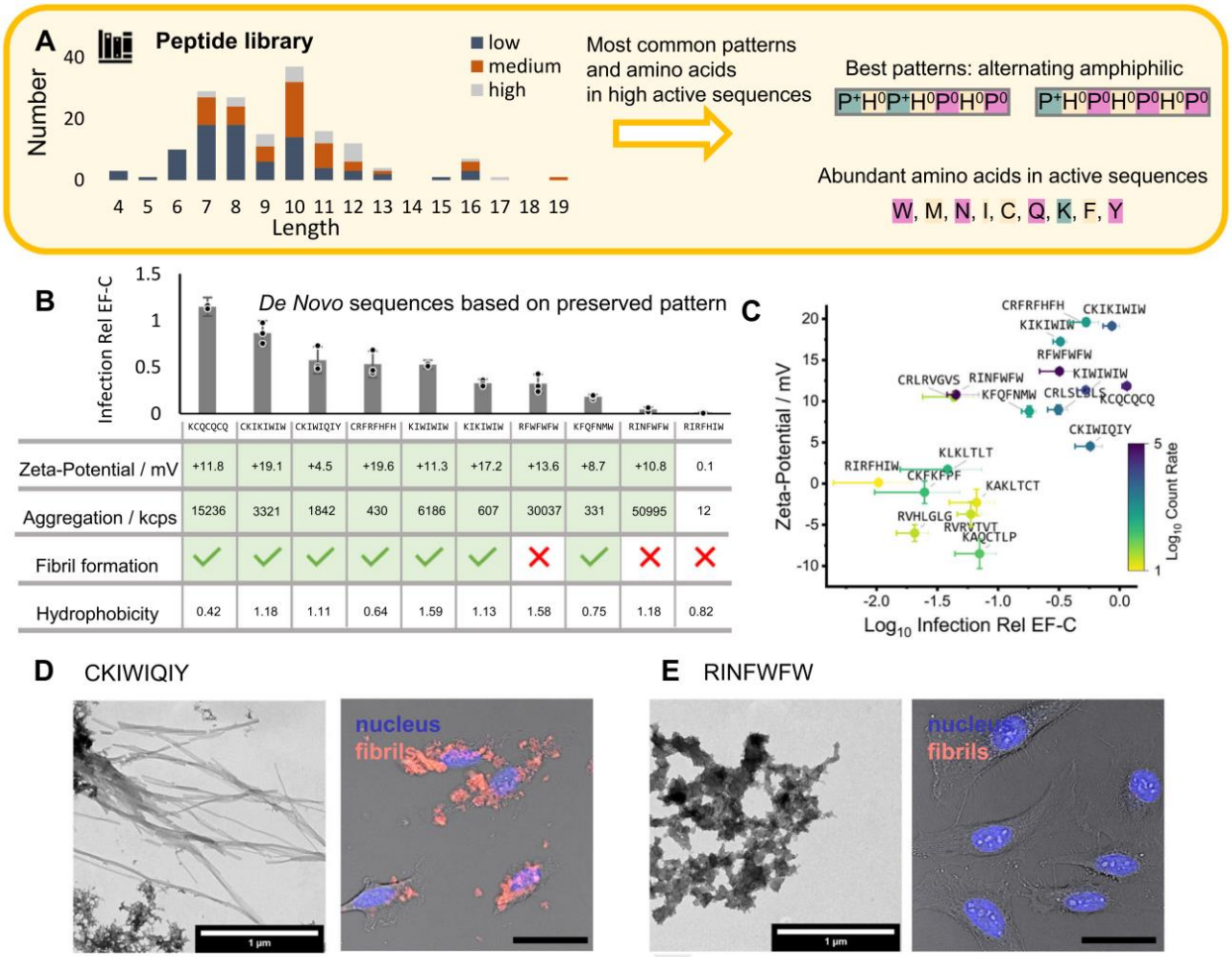


Fig. 6 | Preserved pattern in infectivity enhancing peptides can be applied to create new active sequences confirming property-activity relationship found for original library. A Coarse graining infectivity enhancement in low (< 10%) medium (10-70%) and high (> 70%, Infection Rel EF-C) and coarse graining amino acids according to their hydrophilicity (H = hydrophobic; P = hydrophilic) and side chain charge at pH 7.4 (+ = positive, - = negative, 0 = neutral) enables identification of most abundant sequence pattern in highly active peptides (Table S5) that is an alternating amphiphilic motif $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$. Source code is available under the following link https://doi.org/10.5281/zenodo.8079726[46]. **B** New peptides are created by combining the preserved pattern $P^+H^0P^+H^0P^0H^0P^0$ with the identified best-performing amino acids. 8 of 10 newly created peptides are enhancing infectivity (> 10% Infection Rel EF-C at 1.3 µM, Table S6) and overall confirm the identified property-activity relationship. Error bars indicate standard deviation from triplicate measurements. **C** Graphical visualization of the most important property-activity relationship zeta-potential plotted versus Log_{10} Infection relative to EF-C. The color code represents the Log_{10} Count Rate from light scattering, which can be traced back to microscopic aggregation (Table S6). Mean values are displayed with error bars indicating standard deviation from triplicate measurements. **D** Selected representative example for fibril morphology (TEM, Fig. S25, scale bar 1 µm) and cell interaction (confocal microscopy, scale bar 50 µm) of an active peptide (CKIWIQIY), which is based on the preserved pattern. Representative confocal image shows merged measurement from transmission, Proteostat (red, fibrils) and Hoechst (blue, cell nuclei) channels. **E** Selected representative example for amorphous morphology (TEM, scale bar 1 µm) and no cell interaction (confocal microscopy, scale bar 50 µm) of an inactive peptide (RINFWFW), which is based on the preserved pattern. Confocal image shows merged measurement from transmission, Proteostat (red, fibrils) and Hoechst (blue, cell nuclei) channels. CKIWIQIY and RINFWFW were both susceptible for amyloid sensitive dye staining (Fig. S24), but RINFWFW does not attach to cells in accordance with its low activity. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample, confocal images have been prepared in biological triplicates and recorded at least at three different sites. Source data for **A**–**C** is provided within the Source Data file.

to characterize multiple physicochemical properties for a large number of peptides, thus circumventing scale limitations of individual characterization methods. For example, TEM is a qualitative high-resolution method for structural characterization on a nanometer scale, but it cannot quantify reliably fibril polymorphism and it lacks information on intermolecular interactions, such as β–sheet content, which was therefore determined via FT–IR spectroscopy. It is well-known that amyloid polymorphism can greatly influence its bioactivity[53,54]. However, the systematic study of different fibril morphologies is outside the scope of this study. To detect the formation of β-sheet rich fibrils in bulk, we utilized the amyloid-sensitive ThT dye. However, it should be noted that ThT-fluorescence has been associated with false-positive results (SI Section 1.1). All these methods cannot give information on microscopic aggregation properties, for which we applied microscopy and light scattering. Bioinformatic properties gave information for elucidating general effects of structural changes on infectivity enhancement. Multiparameter analyses enabled us to correlate multi-scale dependencies of interaction of materials with cells and viruses.

We observed that the interactions of peptides with cells and viruses are mostly driven by three codependent properties: (I) ordered structure (II) microscopic aggregation and (III) positive surface charge. While the importance of zeta-potential and fibril formation was already suggested for EF-C PNF in previous reports[12,26,55], microscopic

Article https://doi.org/10.1038/s41467-023-40663-6

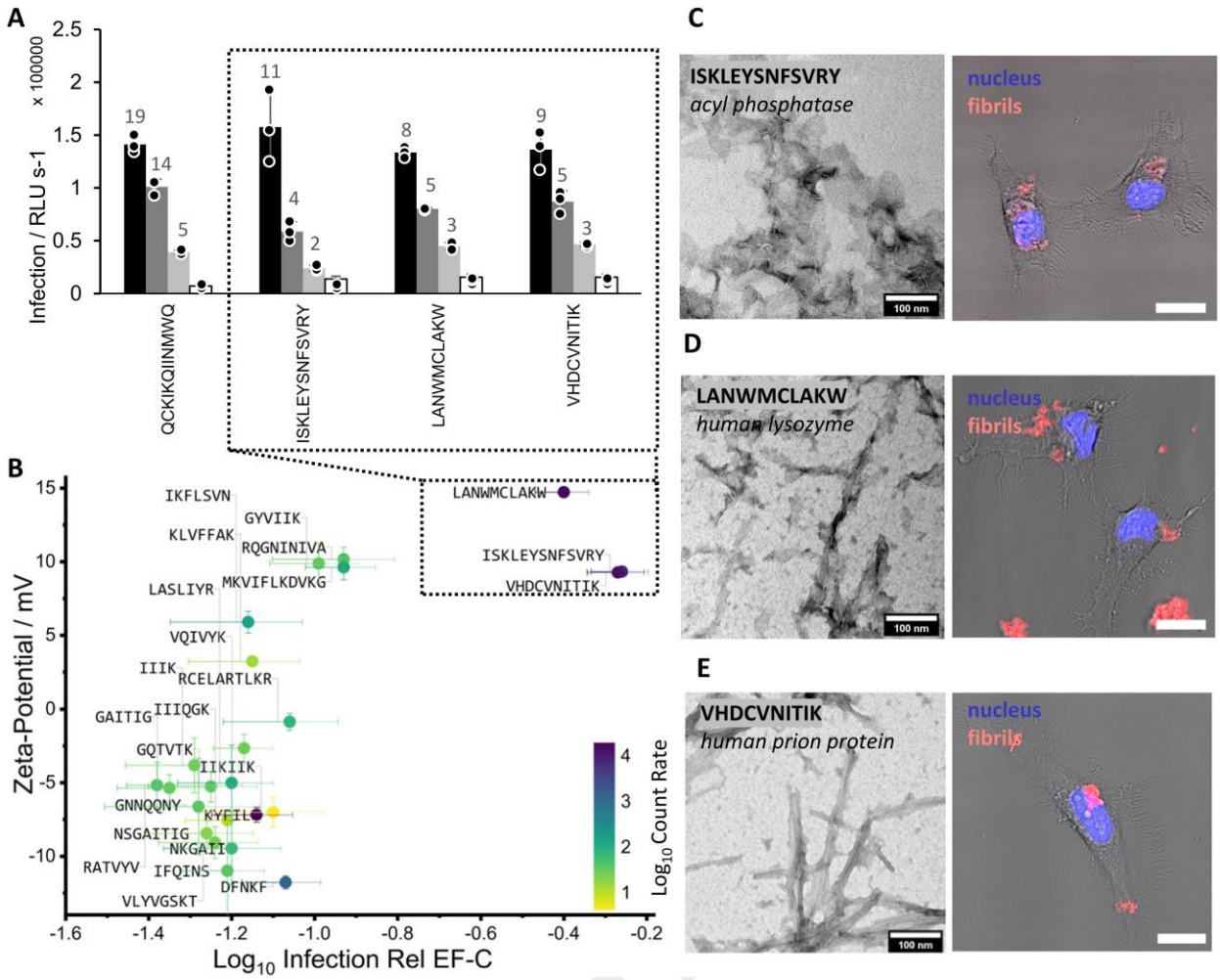


Fig. 7 | Infectivity enhancing properties of selected amyloid peptides from pathogenic and functional origin. A Infection rates of EF-C, and peptides ISKLEYSNFSVRY, LANWMCLAKW, VHDCVNITIK at 6.5, 1.3, 0.26, 0 μM concentration. Numbers at the end of bars indicate n-fold infectivity enhancement compared to virus-only infectivity (0 μM) (Table S7). Error bars indicate standard deviation from triplicate measurements. **B** Zeta-potential plotted against Log_{10} Infection Rel EF-C. Color indicates Log_{10} Count Rate by scattered light. Three peptides (box) that show identified important physicochemical properties that have positive zeta-potential, show fibril formation, and aggregation are enhancing infectivity (Table S7). Mean values are displayed with error bars indicating standard deviation from triplicate measurements. TEM micrograph (1 mg/mL, PBS, Fig. S36, S37 scale bar 100 nm) and confocal microscopy (scale bar 20 μm) of acyl phosphatase derived peptide ISKLEYSNFSVRY (**C**), human lysozyme derived peptide LANWMCLAKW (**D**) and human prion protein derived peptide VHDCVNITIK (**E**). 20 μg/mL preassembled peptides (stained with Proteostat, red) were incubated with Henrietta Lacks (HeLa) cells (nucleus stained with Hoechst, blue). For single channels see Fig. S28. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample, confocal images have been prepared in biological triplicates and recorded at least at three different sites. Source data for **A**, **B** is provided within the Source Data file.

aggregation—which we found to correlate strongly with infectivity—was not systematically studied nor identified as a necessary feature before.

We hypothesize that μm-sized aggregates can accelerate infectivity by physical entanglement with cellular protrusions, which are actively transporting the fibril/virus complexes to the cell surface, where virions can enter the cells[55]. Interestingly, for aggregates larger than 10 μm^2 it is not the size, but rather the number of aggregates correlating with infectivity enhancement. Fibrils aggregating to few very large aggregates might simply not offer enough particles for virus–cell interactions or may be too big to get efficiently engaged by protrusions. In contrast, fibrils aggregating to many small μm-sized aggregates provide a high number of dispersed particles for viruses to attach to and subsequently interact with cell membranes. Noteworthy, there are no obvious changes in fibril formation and aggregation properties of the fibrils in the presence of virions, based on our own data and earlier reports[12,25]. The importance of well-distributed aggregates rather than aggregate size is emphasized by findings from cumulative surface area of total aggregates, where we do not observe any difference between the surface area of "active" aggregates and "inactive" aggregates (SI Section 4.3, Fig. S8). Interestingly, the formation of certain aggregate size and number appears to be an intrinsic feature of the non-covalent self-assembly properties of the molecular sequence, *e.g.* after mechanical downsizing aggregates by pressure or ultrasonication, their original size is recovered immediately in solution (Fig. S9), preventing size-dependent activity investigations of the same peptide.

The strong correlation observed for μm-sized aggregation further emphasizes a fundamental common principle for infectivity enhancement of peptides: The hydrophobic surface properties may not only be a driving force for peptides to form fibrils[56], increase fibril-fibril association[57,58] which leads to μm-sized aggregates, but also a prerequisite to bind and capture viral particles. An interplay between hydrophobicity, electric charge, and anions in buffer is well-known to

Article https://doi.org/10.1038/s41467-023-40663-6

control the formation of hierarchical structures, such as amorphous aggregates or amyloid fibrils[59–61]. Hydrophobicity as a prerequisite for efficient gene delivery was recently shown for polyplexes and polymers for non–viral applications[62,63]. The interaction of amphiphilic peptides with cellular lipid membranes was reported to be facilitated by electrostatic and hydrophobic interactions of the peptides[64] and multivalent binding to the amyloid surface reduced binding affinity to cells and infectivity of semen-derived enhancer of viral infection (SEVI) amyloid fibrils[65]. Thus, we conclude that, a high hydrophobicity can facilitate interaction of peptide fibrils with cell membranes and thereby increase infectivity enhancement as shown systematically in our peptide library.

Although, an alternating amphiphilic sequence pattern with hydrophobic and cationic amino acids is well-known for amyloids[66], and it is found in our library prevalent for active sequences, activity is not limited to these sequence motifs as long as the identified physicochemical properties are fulfilled. For example, we found the same relationship between microscopic aggregation and infectivity enhancement for structurally different endogenous short self-assembling amyloid peptides as exemplified by the three newly discovered highly effective peptides ISKLEYSNFSVRY, LANWMCLAKW, and VHDCVNITIK. Interestingly, these peptides are derived from different biological contexts. While the sequences ISKLEYSNFSVRY and LANWMCLAKW occur naturally in acyl phosphatase and human lysozyme, the peptide VHDCVNITIK is a sequence from the pathogenic human prion protein[52]. It is important to note that while these protein fragments can enhance infectivity, their physicochemical properties cannot be simply extrapolated to the parent protein, as shown exemplary for lysozyme and LANWMCLAKW (Fig. S40). However, it is not recommended to use peptides derived from the original pathogenic origin for therapeutic purposes because they may act as potential seeds for disease-related aggregation events. Amyloid fibrils that aggregate expose fewer terminal ends that are known to disrupt cellular membranes[67]. The fibrillar aggregates reported in this study do not show toxicity but good cell viability for HeLa cells (Fig. S2) as well as other cell types for similar fibrils, as reported elsewhere[32,68,69]. Microscopic aggregation has been previously observed in the non-diseased state of endogenous semen protein prostatic acid phosphatase[70], which suggests a broader biological significance of micrometer-sized aggregation for amyloids beyond simple disease-related associations.

Pioneering works have put much efforts in identifying amyloid fiber evolution mechanisms on a molecular and nm-size level, e.g. for oligo and protofilament formation[71–73]. On the inter-fiber mesoscale lateral aggregation was traced back H bonds[74] and VdW polar interaction modes[75] and to hydrophobic interactions[76,77]. Interestingly, experiments and simulations showed that dynamic, flexible fibrils are less prone to align into multi-fiber strands[37,78]. However, the exact structural determinants for the μm–sized aggregation of amyloid fibrils still remain elusive, and the majority of studies, reporting on mechanistic insights of amyloid formation, focus on aggregation in the nm–length scales (fibril formation)[79] and rarely on μm–sized aggregation. One of the reasons for the lack of studies that deal with understanding the multi-scale aggregation phenomena stems from the difficulty of relating molecular information to microscopic properties, which is especially challenging for self-assembly processes occurring within short time frames. Our data indicates that high β–sheet content of amyloidal peptides, and hydrophobicity are not only necessary for interaction with viral particles and cells but also favor aggregation of fibrils into μm–sized fibrillar aggregates. The higher surface hydrophobicity of CKIKIQI compared to CKIKQII which is observed from SFG measurements (Fig. 4B–v) may be arising from a better intermolecular β–sheet packing of peptide monomers in the fibrils due to localization of the more hydrophobic amino acid isoleucine in the center of the peptide. Furthermore, a high propensity to form β–sheet structures does not necessarily lead to lateral aggregation of fibrils into microscopic aggregates[52]. For example, the sequence GNNQQNY forms β–sheet rich amyloid fibrils but with well hydrated external faces, which presumably hinder them to form microscopic aggregates (Table S7)[75]. The interplay between polar and non-polar interactions, so-called hydrophobic patches, at the fibril surface can determine the fate of ordered or less ordered fiber–fiber aggregation across length scales[36,80] and might explain why some amyloid fibrils form unordered large ($>10\ \mu m^2$) aggregates while others stay isolated or form well-aligned bundles. Our data shows that regular arrangement of polar and non-polar amino acids facilitates stable β-sheet rich structures and SFG measurements indicate that hydrophobic surface interface dominates for microscopically aggregating peptides. Therefore, surface hydrophobicity likely is a key driving force for formation and μm-sized aggregation[81] of amyloidal fibrils and an important characteristic besides charge–driven interactions to associate with viruses and cells[82]. The fewer water-accessible areas may result in higher fibril aggregation which may be at the same time the driving force for cell and virus interaction. However, structural evidence of these interactions is needed to support these hypotheses, which is beyond the scope of this study. Due to the lack of experimental methods to show the assembly dynamics and peptide–water interaction on very short timescales (~ns) which is necessary to explain molecular dynamics, our current efforts focus on molecular simulations which may connect structural information to physicochemical properties underlying intermolecular interactions which eventually result into microscopic aggregation.

The specific factors that determine the ordered or unordered aggregation of micrometer-sized particles are currently unknown, but they are important not only for understanding biological activity but also for a wide range of diseases. We believe that multi-scale and multiparameter considerations in mechanistic origins of microscopical aggregation are not only important for studying pathogenesis, such as neurodegenerative diseases but also for infectivity enhancement and understanding possible connections between them.

This study represents the first investigation of the multi-scale structure–property–activity correlation of self-assembling amyloidal peptides for in vitro retroviral transduction enhancement. In contrast to previous reports, we find that not the amyloid fibrils itself but rather their μm-sized aggregated states facilitate retroviral gene delivery. By applying a data-mining approach, we could perform a multi-parameter regression and unveil crucial co-dependencies in the peptide properties that govern their activity and act on length scales from nano to micrometers: Amphiphilic, hydrophobic sequences form microscopic aggregates with a high β–sheet content and a positively charged surface, all of that required for efficient retroviral transduction enhancement. The fibril-fibril aggregation and the fibril-virus interaction are both driven by the hydrophobicity and amphiphilicity of the peptide sequence. Importantly, we exploited these structure–property–activity findings to rationally create novel active sequences. Moreover, by successfully applying our findings to screen for active amyloid peptides from endogenous proteins, we validated the universal applicability of our study. In that way, our study contributes to a fundamental understanding on the emergence of bioactivity of microscopically aggregating amyloid fibrils.

Methods

Materials

Fmoc-protected amino acids and Wang resin were purchased from Novabiochem®. OymaPure®, *N*-ethyldi*iso*propylamine, piperidine and trifluoroacetic acid were obtained from Carl Roth. Dimethylformamide (DMF for peptide synthesis), diethyl ether and dimethylsulfoxid was purchased from Acros Organics. Acetonitrile was purchased from Fisher Scientific. Uranyl acetate and fluorescamine were purchased from Merck. Lysozyme was purchased from Amresco. PBS, Thioflavin T

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

and α-cyano-4-hydroxycinnamic acid were purchased from Sigma Aldrich. Proteostat® was purchased from Enzo Life Sciences. All chemicals are listed in Supplementary Data 1 and were used as received unless explicitly stated otherwise.

Data mining

The herein applied data-mining procedure covers data generation, processing, data extraction and evaluation, model development, prediction, and interpretation to evaluate the identified property–activity relationship.

The first step involved data generation, where 163 distinct peptides based on an infectivity-enhancing peptide EF-C (QCKIKQIINMWQ) were synthesized (Table S2), six physicochemical and one biological properties were measured (see SI Section 1.1 and method section below). The collected data was then processed, as summarized in Table 1. Twelve bioinformatic properties including peptide net charge, hydrophobicity, isoelectric point, aliphatic index, hydrophobic moment index, boman index, instability index, graph shape index, upsilon steric parameter, smoothed upsilon steric parameter, polarizability, and normalized van der Waals volume were calculated via the "peptides" package in R (see SI Section 1.2)[40].

The next step was extraction and evaluation of the collected data: For an initial screening every descriptor was summarized in a table with the software Excel (Version 2304 Build 16.0.16327.20200, Microsoft 365) and imported to the data-mining software Orange3 (Orange3-v.3.34.0, Anaconda) to find the best correlating descriptor combinations via the functions "scatter plot → find informative projections" and "pearson correlations". The Pearson correlation coefficients for every descriptor were calculated via the function "correlation plot" of the software OriginPro 2021 v.9.8.0.200 (OriginLab). The property activity relationship was identified using the function "constraint multiple regression" analysis (constrained linear least-squares problems solution with the equation: y = A0 + A1*x1 + A2*x2 + [...]+ A5 * x5) of the software OriginPro 2021 v.9.8.0.200 (OriginLab). The model was evaluated using various statistical metrics such as the Pearson correlation coefficient and the statistically significant descriptors were identified by $p < |t|$, $t = 0.05$.

The established linear regression model was validated with de novo peptides predicted from a pattern analysis technique and with existing peptides from a database screening.

The pattern analysis aims to find amino acid patterns and amino acids that are prevalent in high-active peptides (see SI Section 7). Briefly, amino acids were first classified into four categories based on their hydrophilicity and charge at pH 7.4 using the Kyte–Doolittle hydropathy scale[83]. The infectivity enhancement was categorized in low (< 4-fold virus only, <10% relative to EF-C), medium (10–70% relative to EF-C), high (> 70% relative to EF-C). All possible 7-mer patterns were then generated from these coarse-grained amino acids and matched against existing patterns in the peptide library. To identify patterns that are more prevalent in high-active sequences compared to medium- or low-active sequences, an activity index of a pattern p ($AI(p)$, Eq. (1)) was calculated by subtracting the sum of its relative numbers in medium- and low-active sequences from its relative number in high-active sequences as in:

$$AI(p) = \frac{N_{high}(p)}{\sum N_{high}} - \frac{N_{medium}(p)}{\sum N_{medium}} - \frac{N_{low}(p)}{\sum N_{low}} \quad (1)$$

For the amino acid analysis, the absolute abundance of amino acids in one category (low, medium, high) in the whole peptide library was first determined. The relative abundance of amino acids "*Rel AA*" was then calculated by determining the number "*N*" of peptides containing a certain amino acid "*AA*" relative to the total amount of peptides of one category (Eq. (2)).

$$Rel\,AA = \frac{N(peptides\ containing\ AA)_{category}}{\sum \text{peptide}_{category}} \quad (2)$$

All pattern and amino acid analyses were conducted using Python scripts. Source code is provided[46].

The prediction and evaluation of new peptides were conducted by retranslating the coarse-grained best-performing patterns with best-performing amino acids and worst-performing amino acids to yield 10 peptide sequences designed to be active and 8 peptides designed to be inactive. To avoid human bias, the selection of amino acids from the pool of best-performing and worst-performing ones was conducted randomly.

For the data-base screening 538 short (< 20 amino acids) peptides across various sources (176 peptides from PDB[84], 260 peptides from a report[52] and 102 peptides from Waltz database[85]) were collected by their bioinformatic properties (hydrophobicity, isoelectric point, net charge, calculated the "peptides" package in R[40] v 4.0.3) and by any reported information on structure formation (TEM, FT-IR and ThT-fluorescence). A list of all considered peptides from the above-mentioned databases is summarized in Supplementary Data 2. Ultimately, 24 different peptides (SI Section 8) were selected based on fibril formation by TEM and preferably by net charge (> 0), hydrophobicity (> 0.2), β-sheet formation by ThT fluorescence.

Solid-phase peptide synthesis and characterization

Peptide sequences were purchased from Phtd Peptides industrial Co. limited with purity of ≥ 95%. Some peptides were synthesized from us (marked with * in Table S3) by using an automated microwave peptide synthesizer (CEM, Liberty Blue™) at a 0.1 mmol scale using Fmoc protected Wang resin according to the standard coupling strategy. Briefly, the resin was swollen in DMF for 1 h and subsequently the Fmoc protecting group was cleaved with a piperidine solution (20 vol% in DMF) by microwaving at 155 W, 75 °C for 15 s and at 30 W, 90 °C for 50 s. The resin was washed three times with DMF and Fmoc protected amino-acid (5 equiv. relative to the resin loading capacity), DIC (5 equiv.) and Oxyma Pure® (10 equiv.) were dissolved in DMF and added to the reaction vessel. After microwaving at 170 W, 75 °C for 15 s and at 30 W, 90 °C for 110 s the resin was washed with DMF. Repeating this procedure for all required amino acids yields the desired peptide sequence. In the final step, the Fmoc protecting group was cleaved by microwaving at 155 W, 75 °C for 15 s and at 30 W, 90 °C for 50 s with piperidine solution (20% in DMF) and the resin was washed with DCM. The peptide was cleaved off the resin through treatment with 5 mL of trifluoroacetic acid/water/tri*iso*propylsilane (95% / 2.5% / 2.5%) for 2 h and precipitated in cold diethyl ether (40 mL). The precipitate was centrifuged at 3434 x g for 15 min and dissolved in water and 0.1% TFA and purified via HPLC using a gradient of water and acetonitrile containing 0.1% TFA or 0.1% NH_4OH for peptides containing mainly basic or acidic amino acids, respectively. HPLC runs for purification were conducted in preparative scale using a Shimadzu system with the following modules: DGU-20A5R, LC-20AP, CBM-20A, SPD-M20A, SIL-10AP, FRC-18A and a Phenomenex Gemini® 5 µm NX-C18, 110 Å, 150 ×30 mm, 5 µm pore size with a flow rate of 25 mL/min. The peptides were lyophilized to obtain a white solid. Purity (Fig. S33 and Table S3, Supporting Information) was analyzed by MALDI-TOF-MS (rapifleX MALDI-TOF/TOF, Bruker MALDI Synapt G2-SI, Waters) and HPLC-ESI-MS on a Shimadzu LC-2020 Single Quadrupole MS instrument equipped with the modules LC-20AD, SIL-20ACHT, SPD-20A, CTO-using a Kinetex EVOC18 100 Å LC 50 × 2.1 mm column with 2.6 µm pore size. An acetonitrile/water mixture with additional 0.1% formic acid was used as eluent.

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

Table 1 | Summary of pre-processing methods to obtain output data for data-mining from the readout data of experimental descriptors

Descriptor	Method	Readout	Pre-processing	Output
Infectivity enhancement	β-galactosidase assay	β-galactosidase luminescence (infection rate intensity) in relative light units per second (RLU / s)	Calculation of n-fold infection in presence of peptides relative to virus only infection in absence of peptides. Calculation of ratio between n-fold infection of peptides relative to n-fold infection of EF-C (Infection Rel. EF-C).	Logarithmic infectivity relative to EF-C (Log_{10} Rel Infect EF-C)
Fibril formation	TEM	Micrographs	Qualitative evaluation of high aspect ratio nanostructures (fibrils)	Fibril formation observed (1) or not observed (0)
Monomer to nanostructure conversion	Conversion Rate	Fluorescamine fluorescence intensity	Fluorescence intensity ratio before and after filtration	Percentual fluorescence intensity ratio (conversion rate)
Amyloid structure	ThT-fluorescence	Thioflavin-T fluorescence intensity	N-fold fluorescence intensity of peptides compared to buffer control sample	ThT-active (N-fold >2), or ThT-inactive (N-Fold <2).
β-sheet content	FT-IR	Infrared absorbance (1300 and 1800 cm^{-1}) of peptide samples	Data-smoothing, normalization, calculation of second derivate of absorbance intensity, calculation of peak integral ratios between peaks at 1625-1635 cm^{-1} (β-sheet) and approx. 1650-1700 cm^{-1} (native structure).	Percentual β-sheet content
Charge	Zeta-Potential	Zeta-Potential in mV	-	Zeta-Potential in mV
Microscopic aggregation	Light Scattering Count Rate	Light Scattering Count Rate in kilo counts per second (kcps)	-	Light Scattering Count Rate in kilo counts per second (kcps)
Microscopic aggregation	Widefield Microscopy	Microscopy images	Count area size and number of aggregates (> 10 μm^2) via 3D objects counter (ImageJ)	Aggregate area size and number in a microscopy image.

Nanofiber–formation

To induce fibril–formation, peptides were dissolved in DMSO (10 mg/mL stock solution) and diluted in PBS (DBPS w/o calcium chloride and magnesium chloride, Sigma Aldrich) to a final incubation concentration of 650 µM or 1 mg/mL as indicated in the text. Other concentrations described in the text were achieved by further dilution of preformed PNFs. The solution was incubated at room temperature for 24 h in order to obtain fibrils. Literature-derived PNFs were prepared according to literature procedures (see SI Section 7).

Amyloid fibril characterization

ATR–FT-IR. For ATR FT–IR spectroscopy measurements, 200 µL of the respective 1 mg/mL PNF solution was lyophilized and the resulting powder used for measurement. All spectra were recorded on a Bruker Tensor 27 spectrometer with a diamond crystal as ATR element (PIKE MiracleTM, spectral resolution 2 cm^{-1}). Every sample was measured with 64 scans. Data were analyzed with OriginLab software and calculation of β–sheet content was conducted according to a previous report[26]. **TEM.** For TEM measurements 5 µL of PNF fibrils (1 mg/mL) were placed on copper grids coated with carbon and formvar layer (300 mesh, Plano GmbH). After 10 min incubation time, the grids were staining with 4% uranyl acetate solution for 2.5 min and washed with water. Measurements were performed on a Jeol 1400 electron microscope with 120 kV acceleration voltage. **CR Assay.** For the determination of the amount of peptide monomer converting into fibrils the amine reactive dye fluorescamine was used analogous to a previous reported protocol[26]. Briefly, 100 µL of the preformed PNF solution was centrifuged in a Vivaspin 500 tube (3 kDa MWCO) to separate fibers from free peptide monomer (16200 x g, 4 °C, 45 min). The filtrate and 100 µL of the non-filtered PNF sample were lyophilized and dissolved in 30 µL DMSO. 10 µL of the filtrate and non-filtered DMSO samples were submitted in a black UV Star® 384 microliter well–plates (Greiner bio–one) and 3 µL of the fluorescamine solution (10 mg/mL, DMSO) was added to measure the fluorescence enhancement. After 20 min of incubation at RT, fluorescence was recorded on an Infinite® M1000 PRO microplate reader (Tecan) at λem = 470 nm upon excitation at λex = 365 nm with 10 nm bandwidths and multiple reads per well (3×3). Measurements were repeated in triplicates and averaged with standard deviation. All values were calculated as n-fold fluorescence enhancement (DMSO only as a reference was set to 1). The conversion rate CR was defined according to the equation Eq. (3).

$$CR = 100 - \left(\frac{100 \times fluorescence\ intensity\ (filtrate)}{fluorescence\ intensity\ (original)}\right)[\%] \qquad (3)$$

ThT–assay. 4 µL of 1 mg/mL peptide fibrils solution is added to a 20 µL, 50 mM ThT–solution in PBS. The mixture is incubated for 15 min at RT to allow binding of ThT dye with β–sheet rich structures. For reference PBS (4 µL) was added instead of peptide fibril solution. The samples were placed in black UV Star® 384 microliter well–plates (Greiner bio–one). Fluorescence spectra were recorded on an Infinite® M1000 PRO microplate reader (Tecan) at λem = 488 nm upon excitation at λex = 440 nm with 10 nm bandwidths and multiple reads per well (3×3). Measurements were repeated in triplicates and averaged with standard deviation. A peptide fibril was considered as ThT-active if the fluorescence intensity was at least twice as strong compared to the control (PBS containing 10% DMSO). **Zeta-Potential.** For the determination of the surface charge of aggregated peptides zeta-potential measurements were conducted with a Zetasizer Nano ZS, Malvern Instruments. Unless stated otherwise the peptides were dissolved from DMSO (10 mg/mL) in PBS (pH 7.4) to concentration of 1 mg/mL and incubated for 1d at RT. Just before the measurement the 60 µL of the peptides were further diluted in 600 µL KCl (aq, 1 mM) in a 1 mL disposable folded capillary cells (DTS–1060, Zetasizer Nano series,

ACCEPTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

Malvern). The zeta potential of the peptides was derived from the electrophoretic mobility based on the Smoluchowski formula. All measurements were averaged over two individually performed experiments and conducted in triplicates for each sample. **Aggregate analysis**. The derived count rate of scattered light at 633 nm, 173° from the zeta-potential measurement was used as information on the light scattering intensity and turbidity of the sample as an indicator for microscopic aggregation[35]. Furthermore, fluorescence microscopy was conducted by staining 1 μL preformed fibrils (1 mg/mL) with 9 μL ThT (50 μM). Brightfield microscopy was performed simultaneously to evaluate aggregation also for non ThT-active fibrils. The measurements were performed on a Leica DMi8 microscope with 10x air objective and equipped with a Leica MC170 HD camera with a filter setting λ_{em} = 527/30 nm upon excitation at λ_{ex} = 480/40 nm) and processed with the Leica LASX software. To quantify size and number of microscopic aggregation of particles > 10 μm^2 at least 3 independently measured images were analyzed via the 3D object counter analysis tool of ImageJ, v 1.53f51. **Sum–frequency generation**. The SFG experiments were carried out with a femtosecond Ti: Sapphire amplified laser system (Spitfire Ace, Spectra–Physics, 800 nm, 40 fs, 1 kHz) with 5 W output power. We used 2 W to pump an optical parametric amplifier (TOPAS, light conversion) with a non–collinear difference frequency generation stage to generate broadband IR pulses. Another 1 W of the laser output passed through an etalon to generate narrowband visible pulses (10 cm^{-1}). The incident angles of the visible and IR beams were 36° and 41° with respect to the surface normal, respectively. The visible (13 μJ) and IR (5 μJ) pulses were overlapped spatially and temporally at the sample position. Subsequently, the generated SFG pulse was guided into a spectrograph (Acton SP 300i, Princeton Instruments), and detected with an EMCCD (Newton, Andor Technology). To avoid IR absorption of water vapor, we purged N_2 for the measurement. All spectra were collected in the ssp (denoting s-, s- and p-polarized SFG, visible and IR, respectively) polarization combination. The spectra were normalized to the non–resonant signal taken from z–cut quartz after subtracting a background spectrum.

Virus–Peptide Interaction. TZM-bl cell line and R5-tropic HIV-1 stock plasmid (pBRNL4.39-92TH014) were obtained from the National Institutes of Health AIDS Research and Reference Reagent Program. Human wild-type HeLa cell line were obtained from abcam, ab260075, LOT: GR3292155-1. The cell lines were not further authenticated, and no commonly misidentified cell lines are used within this study. Infectious R5–tropic HIV-1 stocks and infectivity assays were prepared analogous to a previous report[26]. Briefly, 10,000 TZM-bl cells were seeded into 96-well plates in 180 μL DMEM one day prior to infection experiment. Peptides were incubated two days prior to experiment at 650 μM concentration in PBS to generate fibrils. One day before the experiment the peptide samples were diluted to 130 μM in PBS. On the day of the infection experiment, 40 μl of 130 μM, 26 μM, 5.2 μM and 0 μM peptide samples were mixed 1:1 (v/v) with 40 μL of HIV-1 (prediluted 1:125 (v/v) in cell culture medium), resulting in fibril concentration of 65 μM, 13 μM, 2.6 and 0 μM during virion treatment. After 10 min incubation to allow binding of virus particles to peptide fibrils, 20 μL of the virus-peptide fibril mixture were added to 180 μL of cells resulting in final cell culture peptide fibril concentrations of 6.5 μM, 1.3 μM, 0.26 and 0 μM, respectively. The effect of peptide fibrils (final concentration on cells 6.5, 1.3, 0.26, 0 μM) on HIV-1 infection was studied 3 days post infection via a luminescence assay for β–galactosidase, which was reported by TZM–bl cells upon infection. For the property–activity correlation a peptide concentration of 1.3 μM was used as a standard condition and EF-C was always used as a reference to compare infectivity enhancement between peptides. The HIV-1 infection assay was conducted in three technical replicates and reproduced at least once.

The cell viability after addition of peptides to TZM–bl cells was studied via the CellTiter–Glo assay. To this end, 10,000 cells were seeded and on the next day serial diluted peptides were added. After 3 days the supernatant was removed and 100 μL CellTiter–Glo Reagent 1:1 diluted in PBS was added. After 10 min 50 μL was transferred to white microplate and luminescence was recorded by Orion microplate luminometer.

Confocal laser scanning microscopy studies were performed for the visualization of the cell–peptide interaction. HeLa cells were seeded one day prior to conducting the assay (40,000/well) in an 8–well IBIDI slide. 4 μL of the preformed peptide fibrils (1 mg/mL) were diluted with 4 μL Proteostat (Enzo Life Science, 1 μL stock in 999 μL PBS) and further diluted with medium to receive a final peptide concentration of 20 μg/mL. The nucleus of the HeLa cells was stained with Hoechst 33342 (NucBlue™, Thermo Fisher Scientific). The peptide solution mixture was transferred to the HeLa cells and incubated for 30 min at 37 °C before washing three times with PBS. The interaction of fibril clusters with cells was monitored after 30 min incubation time on a Stellaris 8 confocal laser scanning microscope (Leica) equipped with a 20x air objective and laser excitation wavelength of 405 nm (Hoechst) and 561 nm (Proteostat).

Statistics and reproducibility

All virus assays were conducted in technical triplicates. All tested peptides for the virus infectivity assay were conducted at four different concentrations (6.5 μM, 1.3 μM, 0.26 μM, 0 μM). To verify reproducibility of the experimental findings the infection assays were at least repeated independently in biological duplicates. The physicochemical characterization methods for peptide self-assembly and morphology, that are determined via CR Assay, ThT–Assay and Zeta-Potential have been conducted in technical triplicates with two independent repetitions for Zeta-Potential and ThT-Assay measurements. Unless stated otherwise the aggregate analysis via widefield microscopy and TEM analysis have been conducted once with at least three microscopy images recorded for each peptide sample. ATR–FT-IR measurements have been conducted once for each peptide sample. The sum–frequency generation experiments have been conducted in two independent repetitions. To ensure comparability within individual measurements, the reference peptide "EF-C" was included in all assays. All replication attempts were successful.

No statistical method was used to predetermine sample size. No data were excluded from the analyses and the experiments were not randomized because it is not relevant for this study, as infectivity enhancement is independent from the order of samples tested. The investigators were not blinded to allocation during experiments and outcome assessment because it is not relevant for this study, as the results from infection assays are read out automated via a multi-plate reader.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The authors declare that the main data supporting the findings of this study are available within the paper and its associated Supplementary Information. Source data for manuscript and supplementary figures are provided with this paper. Other databases that are used for this study are "The Protein Data Bank" (PDB, https://www.rcsb.org/)[84], a report by Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L (https://doi.org/10.1038/nbt1012)[52] and the WALTZ-DB 2.0 database (http://waltzdb.switchlab.org/home)[85]. A summary of peptides considered for evaluation from these databases is provided in a supplementary file. Source data are provided with this paper.

Author Proof

Article https://doi.org/10.1038/s41467-023-40663-6

Code availability

The computer codes and the corresponding input data for amino acid and peptide pattern analyses are openly available at the following data repository https://doi.org/10.5281/zenodo.8079726[46].

References

1. Yoder, K. E., Rabe, A. J., Fishel, R. & Larue, R. C. Strategies for targeting retroviral integration for safer gene therapy: advances and challenges. *Front. Mol. Biosci.* **8**, 418 (2021).
2. U.S. Food & Drug Administrations, Approval Letter - KYMRIAH (tisagenlecleucel). https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/kymriah-tisagenlecleucel (2017).
3. Labbé, R. P., Vessillier, S. & Rafiq, Q. A. Lentiviral vectors for T cell engineering: clinical applications, bioprocessing and future perspectives. *Viruses* **13**, 1528 (2021).
4. Vormittag, P., Gunn, R., Ghorashian, S. & Veraitch, F. S. A guide to manufacturing CAR T cell therapies. *Curr. Opin. Biotechnol.* **53**, 164–181 (2018).
5. Kaygisiz, K. & Synatschke, C. V. Materials promoting viral gene delivery. *Biomater. Sci.* **8**, 6113–6156 (2020).
6. Münch, J. et al. Semen-derived amyloid fibrils drastically enhance HIV infection. *Cell* **131**, 1059–1071 (2007).
7. Arnold, F. et al. Naturally occurring fragments from two distinct regions of the prostatic acid phosphatase form amyloidogenic enhancers of HIV infection. *J. Virol.* **86**, 1244–1249 (2012).
8. Roan, N. R. et al. Peptides released by physiological cleavage of semen coagulum proteins form amyloids that enhance HIV infection. *Cell Host Microbe* **10**, 541–550 (2011).
9. Zhang, L. et al. A novel modified peptide derived from membrane-proximal external region of human immunodeficiency virus type 1 envelope significantly enhances retrovirus infection. *J. Pept. Sci.* **20**, 46–54 (2014).
10. Lee, Y.-H. & Ramamoorthy, A. Semen-derived amyloidogenic peptides-Key players of HIV infection. *Protein Sci.* **27**, 1151–1165 (2018).
11. Castelletto, V. & Hamley, I. W. Amyloid and hydrogel formation of a peptide sequence from a coronavirus spike protein. *ACS Nano* **16**, 1857–1867 (2022).
12. Yolamanova, M. et al. Peptide nanofibrils boost retroviral gene transfer and provide a rapid means for concentrating viruses. *Nat. Nanotechnol.* **8**, 130–136 (2013).
13. Shrimali, P. C. et al. Amyloidogenic propensity of self-assembling peptides and their adjuvant potential for use as DNA vaccines. *bioRxiv* 2022.09.09.507367 https://doi.org/10.1101/2022.09.09.507367 (2022).
14. Knowles, T. P. J. & Buehler, M. J. Nanomechanics of functional and pathological amyloid materials. *Nat. Nanotechnol.* **6**, 469–479 (2011).
15. Al-Garawi, Z. S., Morris, K. L., Marshall, K. E., Eichler, J. & Serpell, L. C. The diversity and utility of amyloid fibrils formed by short amyloidogenic peptides. *Interface Focus* **7**, 20170027 (2017).
16. Cremades, N. et al. Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell* **149**, 1048–1059 (2012).
17. Klein, W., Stine, W. & Teplow, D. Small assemblies of unmodified amyloid β-protein are the proximate neurotoxin in Alzheimer's disease. *Neurobiol. Aging* **25**, 569–580 (2004).
18. Stefani, M. & Dobson, C. M. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **81**, 678–699 (2003).
19. Dear, A. J. et al. Identification of on- and off-pathway oligomers in amyloid fibril formation. *Chem. Sci.* **11**, 6236–6247 (2020).
20. Michaels, T. C. T. et al. Dynamics of oligomer populations formed during the aggregation of Alzheimer's Aβ42 peptide. *Nat. Chem.* **12**, 445–451 (2020).
21. Pellarin, R. & Caflisch, A. Interpreting the aggregation kinetics of amyloid peptides. *J. Mol. Biol.* **360**, 882–892 (2006).
22. Fowler, D. M. et al. Functional amyloid formation within mammalian tissue. *PLoS Biol.* **4**, 0100–0107 (2006).
23. Zapadka, K. L., Becher, F. J., Gomes dos Santos, A. L. & Jackson, S. E. Factors affecting the physical stability (aggregation) of peptide therapeutics. *Interface Focus* **7**, 20170030 (2017).
24. Sinha, N. J., Langenstein, M. G., Pochan, D. J., Kloxin, C. J. & Saven, J. G. Peptide design and self-assembly into targeted nanostructure and functional materials. *Chem. Rev.* **121**, 13915–13935 (2021).
25. Kirti, S. et al. Amyloid fibrils with positive charge enhance retroviral transduction in mammalian cells. *ACS Biomater. Sci. Eng.* **5**, 126–138 (2019).
26. Sieste, S. et al. Supramolecular peptide nanofibrils with optimized sequences and molecular structures for efficient retroviral transduction. *Adv. Funct. Mater.* **31**, 2009382 (2021).
27. Schütz, D. et al. Negatively charged peptide nanofibrils from immunoglobulin light chain sequester viral particles but lack cell-binding and viral transduction-enhancing properties. *ACS Omega* **6**, 7731–7738 (2021).
28. Fernandez-Lozano, C., Gestal, M., Munteanu, C. R., Dorado, J. & Pazos, A. A methodology for the design of experiments in computational intelligence with multiple regression models. *PeerJ* **4**, e2721 (2016).
29. Yamashita, H. et al. Machine learning-based amino acid substitution of short peptides: acquisition of peptides with enhanced inhibitory activities against α-amylase and α-Glucosidase. *ACS Biomater. Sci. Eng.* **6**, 6117–6125 (2020).
30. Giguère, S. et al. Machine learning assisted design of highly active peptides for drug discovery. *PLOS Comput. Biol.* **11**, e1004074 (2015).
31. Chowdhury, A. S., Reehl, S. M., Kehn-Hall, K., Bishop, B. & Webb-Robertson, B. J. M. Better understanding and prediction of antiviral peptides through primary and secondary structure feature importance. *Sci. Rep.* **10**, 1–8 (2020).
32. Schilling, C. et al. Sequence-optimized peptide nanofibers as growth stimulators for regeneration of peripheral neurons. *Adv. Funct. Mater.* **29**, 1809112 (2019).
33. Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
34. Krebs, M. R. H., Bromley, E. H. C. & Donald, A. M. The binding of thioflavin-T to amyloid fibrils: Localisation and implications. *J. Struct. Biol.* **149**, 30–37 (2005).
35. Yoshimura, Y. et al. Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl Acad. Sci. USA.* **109**, 14446–14451 (2012).
36. Adamcik, J. et al. Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. *Nat. Nanotechnol.* **5**, 423–428 (2010).
37. Bolisetty, S., Adamcik, J. & Mezzenga, R. Snapshots of fibrillation and aggregation kinetics in multistranded amyloid β-lactoglobulin fibrils. *Soft Matter* **7**, 493–499 (2011).
38. Ridgley, D. M., Ebanks, K. C. & Barone, J. R. Peptide mixtures can self-assemble into large amyloid fibers of varying size and morphology. *Biomacromolecules* **12**, 3770–3779 (2011).
39. Fauchère, J.-L., Charton, M., Kier, L. B., Verloop, A. & Pliska, V. Amino acid side chain parameters for correlation studies in biology and pharmacology. *Int. J. Pept. Protein Res.* **32**, 269–278 (2009).
40. Osorio, D., Rondón-Villarreal, P. & Torres, R. Peptides: a package for data mining of antimicrobial. *Pept. R. J.* **7**, 4 (2015).
41. Takeuchi, Y., McClure, M. O. & Pizzato, M. Identification of gammaretroviruses constitutively released from cell lines used for human immunodeficiency virus research. *J. Virol.* **82**, 12585–12588 (2008).
42. Fu, L., Ma, G. & Yan, E. C. Y. In situ misfolding of human islet amyloid polypeptide at interfaces probed by vibrational sum frequency generation. *J. Am. Chem. Soc.* **132**, 5405–5412 (2010).

ACCEPTED PROOF

43. Hosseinpour, S. et al. Structure and dynamics of interfacial peptides and proteins from vibrational sum-frequency generation spectroscopy. *Chem. Rev.* **120**, 3420–3465 (2020).
44. Kaygisiz, K. et al. Inverse design of viral infectivity-enhancing peptide fibrils from continuous protein-vector embeddings. *Biomater. Sci.* https://doi.org/10.1039/D3BM00412K (2023).
45. Zhao, J. et al. Epigallocatechin-3-gallate (EGCG) inhibits fibrillation, disaggregates amyloid fibrils of α-synuclein, and protects PC12 cells against α-synuclein-induced toxicity. *RSC Adv* **7**, 32508–32517 (2017).
46. Kaygisiz, K. et al. Data-mining unveils structure-property-activity correlation of viral infectivity enhancing self-assembling peptides. *Zenodo.org* https://doi.org/10.5281/ZENODO.8079727 (2023).
47. Fulop, T. et al. Can an infection hypothesis explain the beta amyloid hypothesis of Alzheimer's Disease? *Front. Aging Neurosci.* **10**, (2018).
48. Hur, J. Y. et al. The innate immunity protein IFITM3 modulates γ-secretase in Alzheimer's disease. *Nature* **586**, 735–740 (2020).
49. Medini, K. et al. Controlling gelation with sequence: Towards programmable peptide hydrogels. *Acta Biomater.* **43**, 30–37 (2016).
50. Kurnellas, M. P., Adams, C. M., Sobel, R. A., Steinman, L. & Rothbard, J. B. Amyloid fibrils composed of hexameric peptides attenuate neuroinflammation. *Sci. Transl. Med.* **5**, 179ra42 (2013).
51. Maurer-Stroh, S. et al. Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods* **7**, 237–242 (2010).
52. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22**, 1302–1306 (2004).
53. Aubrey, L. D. et al. Quantification of amyloid fibril polymorphism by nano-morphometry reveals the individuality of filament assembly. *Commun. Chem.* **3**, 125 (2020).
54. Qiang, W., Yau, W.-M., Lu, J.-X., Collinge, J. & Tycko, R. Structural variation in amyloid-β fibrils from Alzheimer's disease clinical subtypes. *Nature* **541**, 217–221 (2017).
55. Schütz, D. et al. Viral transduction enhancing effect of EF-C peptide nanofibrils is mediated by cellular protrusions. *Adv. Funct. Mater.* **31**, 2104814 (2021).
56. Betush, R. J., Urban, J. M. & Nilsson, B. L. Balancing hydrophobicity and sequence pattern to influence self-assembly of amphipathic peptides. *Pept. Sci.* **110**, e23099 (2018).
57. De Mello, L. R. et al. Amyloid-like self-assembly of a hydrophobic cell-penetrating peptide and its use as a carrier for nucleic acids. *ACS Appl. Bio Mater.* **4**, 6404–6416 (2021).
58. Zhao, R. et al. Measurement of amyloid formation by turbidity assay—seeing through the cloud. *Biophys. Rev.* **8**, 445–471 (2016).
59. Frederix, P. W. J. M. et al. Exploring the sequence space for (tri-) peptide self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37 (2015).
60. Strandberg, E. et al. Phosphate-dependent aggregation of [KL]n peptides affects their membranolytic activity. *Sci. Rep.* **10**, 12300 (2020).
61. Marek, P. J., Patsalo, V., Green, D. F. & Raleigh, D. P. Ionic strength effects on amyloid formation by amylin are a complicated interplay among debye screening, ion selectivity, and hofmeister effects. *Biochemistry* **51**, 8478–8490 (2012).
62. Zhang, Z. et al. Dose-independent transfection of hydrophobized polyplexes. *Adv. Mater.* **33**, 2102219 (2021).
63. Hack, F. J. et al. Indole, phenyl, and phenol groups: the role of the comonomer on gene delivery in guanidinium containing methacrylamide terpolymers. *Macromol. Rapid Commun.* **42**, 2000580 (2021).
64. Yang, Y., Jalali, S., Nilsson, B. L. & Dias, C. L. Binding mechanisms of amyloid-like peptides to lipid bilayers and effects of divalent cations. *ACS Chem. Neurosci.* **12**, 2027–2035 (2021).
65. Capule, C. C., Brown, C., Olsen, J. S., Dewhurst, S. & Yang, J. Oligovalent amyloid-binding agents reduce SEVI-mediated enhancement of HIV-1 infection. *J. Am. Chem. Soc.* **134**, 905–908 (2012).
66. West, M. W. et al. De novo amyloid proteins from designed combinatorial libraries. *Proc. Natl Acad. Sci. USA.* **96**, 11211–11216 (1999).
67. Milanesi, L. et al. Direct three-dimensional visualization of membrane disruption by amyloid fibrils. *Proc. Natl Acad. Sci.* **109**, 20455–20460 (2012).
68. Kaygisiz, K. et al. Photoinduced amyloid fibril degradation for controlled cell patterning. *Macromol. Biosci.* **23**, 2200294 (2023).
69. Ender, A. M. et al. Cell-instructive surface gradients of photoresponsive amyloid-like fibrils. *ACS Biomater. Sci. Eng.* **7**, 4798–4808 (2021).
70. Usmani, S. M. et al. Direct visualization of HIV-enhancing endogenous amyloid fibrils in human semen. *Nat. Commun.* **5**, 1–8 (2014).
71. Chiti, F. & Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
72. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).
73. Knowles, T. P. et al. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **318**, 1900–1903 (2007).
74. Mossou, E. et al. The self-assembling zwitterionic form of *L*-phenylalanine at neutral pH. *Acta Crystallogr. Sect. C. Struct. Chem.* **70**, 326–331 (2014).
75. Nelson, R. et al. Structure of the cross-β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
76. Do, T. D., Kincannon, W. M. & Bowers, M. T. Phenylalanine oligomers and fibrils: The mechanism of assembly and the importance of tetramers and counterions. *J. Am. Chem. Soc.* **137**, 10080–10083 (2015).
77. Eckes, K. M., Mu, X., Ruehle, M. A., Ren, P. & Suggs, L. J. β sheets not required: combined experimental and computational studies of self-assembly and gelation of the ester-containing analogue of an Fmoc-dipeptide hydrogelator. *Langmuir* **30**, 5287–5296 (2014).
78. Paparcone, R., Cranford, S. W. & Buehler, M. J. Self-folding and aggregation of amyloid nanofibrils. *Nanoscale* **3**, 1748 (2011).
79. Nguyen, P. H., Sterpone, F. & Derreumaux, P. Aggregation of disease-related peptides. in *Progress in Molecular Biology and Translational Science* vol. 170 435–460 (Elsevier, 2020).
80. Adamcik, J. & Mezzenga, R. Adjustable twisting periodic pitch of amyloid fibrils. *Soft Matter* **7**, 5437 (2011).
81. Elsawy, M. A. et al. Modification of β-sheet forming peptide hydrophobic face: effect on self-assembly and gelation. *Langmuir* **32**, 4917–4923 (2016).
82. Wu, Y. et al. Patchy amphiphilic dendrimers bind adenovirus and control its host interactions and in vivo distribution. *ACS Nano* **13**, 8749–8759 (2019).
83. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
84. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
85. Louros, N. et al. WALTZ-DB 2.0: an updated database containing structural information of experimentally determined amyloid-forming peptides. *Nucleic Acids Res.* **48**, D389–D393 (2020).

Acknowledgements

The study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 316249678—SFB 1279 (A02, A03, A05, C01). A.D. acknowledges support by BiGmax, the Max

UNCORRECTED PROOF

Article https://doi.org/10.1038/s41467-023-40663-6

Planck Society's Research Network on Big-Data-Driven Materials-Science. T.B. acknowledges support from the Emmy Noether program of the Deutsche Forschungsgemeinschaft (DFG).

Author contributions

KK conceived the idea, designed, and carried out the experiments, analyzed data, performed the formal analysis, and wrote the original draft. LRW designed, and carried the experiments related to virus transduction, analyzed data, performed the formal analysis, and revised the manuscript. AD, XY, YN analyzed data, performed the formal analysis, and revised the manuscript. TB, CVS, JM, and TW conceptualized and supervised, provided the funding resources, and revised the manuscript. The final manuscript was approved by all authors.

Competing interests

J.M. and L.R.W. hold patents and K.K., L.R.W., C.V.S., J.M. and T.W. submitted a patent for application of peptide fibrils for retroviral transduction enhancement. The remaining authors declare no competing interests. The authors have no relevant non-financial interests to disclose.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-40663-6.

Correspondence and requests for materials should be addressed to Christopher V. Synatschke or Tanja Weil.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

UNCORRECTED PROOF

Supporting Information

Supplementary Information for

Data-Mining Unveils Structure–Property–Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides

Kübra Kaygisiz, Lena Rauch-Wirth, Arghya Dutta, Xiaoqing Yu, Yuki Nagata, Tristan Bereau, Jan Münch, Christopher V. Synatschke, and Tanja Weil**

K. Kaygisiz, Dr. C. V. Synatschke, Prof. T. Weil
Department Synthesis of Macromolecules
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
Email: synatschke@mpip-mainz.mpg.de, weil@mpip-mainz.mpg.de

L. Rauch-Wirth, Prof. J. Münch
Institute of Molecular Virology
Ulm University Medical Center
Meyerhofstraße 1, 89081 Ulm, Germany
Email: jan.muench@uni-ulm.de

Dr. A. Dutta, Prof. T. Bereau
Department Polymer Theory
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany

Dr. A. Dutta (*present address*)
Institute of Biochemistry II, Faculty of Medicine
Goethe University
Theodor-Stern-Kai 7, 60590 Frankfurt, Germany

X. Yu, Dr. Y. Nagata
Department Molecular Spectroscopy
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany

Prof. T. Bereau (*present address*)
Institute for Theoretical Physics
Heidelberg University
Philosophenweg 19, 69120 Heidelberg, Germany

*Corresponding authors: Christopher V. Synatschke, synatschke@mpip-mainz.mpg.de and Tanja Weil, weil@mpip-mainz.mpg.de

Content

1. REMARKS ON EXPERIMENTAL PARAMETERS AND BIOINFORMATIC DESCRIPTORS 3

1.1 EXPERIMENTAL PARAMETERS 3
1.2 BIOINFORMATIC PARAMETERS 5

2. CELL VIABILITY OF REPRESENTATIVELY SELECTED PEPTIDES 7

3. TIMESCALE OF PEPTIDE AGGREGATION 8

4. SINGLE PARAMETER CORRELATIONS 9

4.1 CORRELATION MATRIX 9
4.2 CORRELATION OF INFECTION RATES WITH HYDROPHOBICITY AND FORMATION OF β-SHEET RICH NANOSTRUCTURES 10
4.3 CORRELATION OF INFECTION RATES AND COUNT RATE WITH μM-SIZED AGGREGATE FORMATION 11

5. MULTIPARAMETER CORRELATIONS 14

6. SELECTED EXAMPLES SHOWCASING STRUCTURE–PROPERTY–ACTIVITY RELATIONSHIP 17

6.1 *N*-TERMINAL CYSTEINE LEADS TO MORE AGGREGATION AND INCREASES INFECTIVITY ENHANCEMENT 17
6.2 PEPTIDES WITH HIGH HYDROPHOBICITY CAN ENHANCE INFECTIVITY EVEN IF THEY HAVE NEGATIVE ZETA-POTENTIAL OR DO NOT FORM FIBRILS 19
6.3 PEPTIDES SUBSTITUTED WITH LOW HYDROPHOBIC AMINO ACID ALANINE DO NOT ENHANCE INFECTIVITY 20
6.4 *C*-TERMINAL SEQUENCE EXTENSIONS WITH HYDROPHOBIC AMINO ACIDS INCREASE INFECTIVITY ENHANCEMENT 22
6.5 THE CHAOTROPIC SUBSTANCE EGCG CHANGES PHYSICOCHEMICAL PROPERTIES AND HIGHLIGHTS IMPORTANCE OF SURFACE CHARGE AND INTERNAL ORDER 23

7. FURTHER INFORMATION ON PATTERN AND AMINO ACID ANALYSIS 24

IMPACT OF SEQUENCE ORDER IS GREATER FOR SHORTER SEQUENCES COMPARED TO LONG SEQUENCES 29

8. AMYLOID PEPTIDES FROM FUNCTIONAL AND PATHOGENIC CONTEXTS 30

PREPARATION OF PEPTIDE FIBRILS DESCRIBED IN LITERATURE 31

9. PEPTIDE LIBRARY: TABLES 34

BIOINFORMATIC DESCRIPTORS 40
EXPERIMENTAL DESCRIPTORS 44
MICROSCOPY IMAGES OF μM-SIZED PEPTIDE AGGREGATION 51
TEM MICROGRAPHS OF PEPTIDE LIBRARY 52
FT-IR OF PEPTIDE LIBRARY 55
LC-MS MEASUREMENT OF SYNTHESIZED PEPTIDES 61
TEM OF LITERATURE DERIVED AMYLOID PEPTIDES 62
FT-IR OF PEPTIDES DESIGNED VIA PATTERN ANALYSIS 64
FT-IR OF LITERATURE DERIVED PEPTIDES 66

SUPPLEMENTARY REFERENCES 69

1. Remarks on experimental parameters and bioinformatic descriptors

1.1 Experimental parameters

Transmission electron microscopy (TEM) measurements. TEM measurements were used to control if peptides assemble. The assemblies were categorized in non-fibrillar structures (no assembly, amorphous and micellar aggregates) and fibrillar structures (fibrils, bundles, sheets) as shown in **Figure S1**. A peptide was defined as non-fibrillar if no fibrillar structure was observed once on the entire TEM-grid but only amorphous aggregates. In contrast, peptides defined as fibrillar may also include amorphous structures. Amongst nonfibrillar assemblies also amorphous aggregates are summarized. Amorphous aggregates are defined as structures with high β-sheet content which lack long-range order, that is crystallinity or fibril formation. It is generally assumed that these structures might be oligomers from on or off pathway amyloid assembly.[1–4]

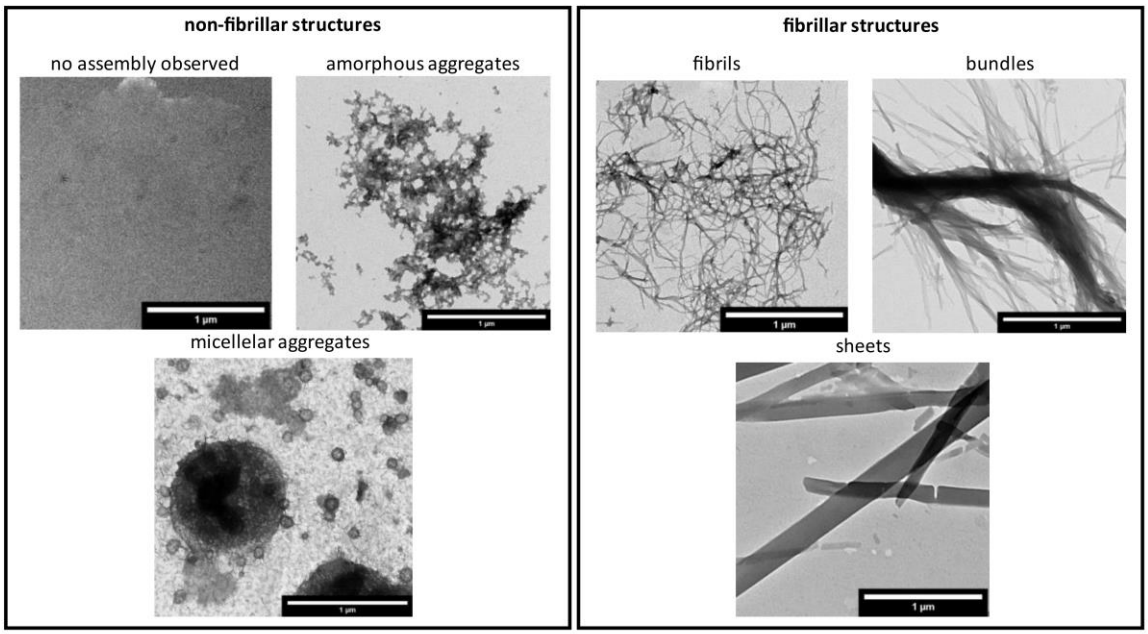


Figure S1 Overview of observed morphologies using representative examples and categorization into non-fibrillar structures (no assembly, amorphous aggregates, micellar aggregates) and fibrillar structures (fibrils, bundles, sheets).

Thioflavin T (ThT) fluorescence measurement. The ThT-fluorescence assay is used as a standard method to show the presence of amyloidal secondary structure. ThT binds to β-sheet rich structures, which immobilizes molecular rotation and enables excitation at 450 nm and emission at 482 nm.[5,6] However, it has to be kept in mind, that this assay is prone to show large variations in fluorescence intensity depending on structural polymorphism,[7] electrostatic interactions with positive charged particles,[8] and in the presence of aromatic amino acid such as phenylalanine (F), tryptophane (W) and tyrosine (Y).[5,9] Consequently, in this study a classification into ThT-active (> 2 fold compared to control without peptide) and non-active peptides will be considered rather than comparing exact ThT fluorescence intensities (**Table S5, S6, S7**).

Turbidity measurements via light scattering. Most of the peptide fibrils investigated in this study aggregate into microscopically large clusters with visible turbidity as can be also observed in bright field and fluorescence microscopy (**Figure 2**, **Figure S31**, **Figure S32**). In general, there are two main

ways to determine turbidity of samples, this can be via transmittance and via scattering. Transmittance is typically measured above 400 nm where the peptide itself is not absorbing and is frequently applied as it is an easy and fast method.[10,11] However, transmittance contributes not only from µm-sized aggregates but also nm-sized aggregates such as distinct fibrils, which makes hard to distinguish between fibril formation and micrometer sized aggregation. Although dynamic light scattering is a frequently applied method for determination of particle size, it cannot be conducted in solutions with anisotropic aggregates in micrometer size range, since the Stokes-Einstein relation is only applicable for homogeneous anisotropic spherical particles.[12] Further models describing scattering intensity and particles sizes, such as Mie or Rayleigh cannot be applied to these self-assembling peptides containing heterogeneous mixture of aggregates ranging from nano- to micrometer in size.[10] Since the fibrils formed by the short-self assembling sequences are highly polymorphic (TEM, **Figure S33**) and form heterogenous aggregate size (microscopy, **Figure S32**), it is not straightforward to apply DLS to determine polydispersity of aggregates. We assume that the heterogeneity of fibril aggregation in our peptide library originates from fast self-assembly kinetics and the resulting turbidity is observed already by bare eye within seconds after inducing assembly. Therefore, we used the derived count rate of scattered light to gain information on turbidity of fibrils solutions.[1,13] Light scattering rates detected in a backscattering setup at 173° is suitable for evaluating concentrated samples.[12] Still this method bears the risk that insoluble particles (other than assembled and aggregated) are contributing to the scattering events. Therefore, the scattering measurement is accompaniment by complementary characterization methods such as TEM (for nanostructure verification, **Figure S33**) and microscopy (for macroscopic verification, **Figure S32**).

Zeta Potential. The zeta-potential describes the electrophoretic mobility of a particle in an electric field. For the case of supramolecularly assembled peptides, this method provides information about the surface charge of peptide fibrils and aggregates as well as information about aggregate size because this also plays a role for the final mobility. The net charge of a peptide does not necessarily correlate with the measured zeta-potential of the molecularly dissolved peptide or an aggregate of peptide. It is known that peptides with similar sequence and same net charge can differ significantly in their zeta-potential depending on the pH and aggregation state of the fibrils.[14–16]

Infectivity Data. All infectivity data discussed in this study are referring to an *in vitro* peptide concentration of 1.3 µM on cells (**Table S5**). The virus stock concentration is same for all peptides within one experiment. The infectivity enhancement of a peptide was determined via a luminescence assay, which detects β galactosidase expressed by TZM-bl cells upon infection of HIV-1. To evaluate the infectivity enhancement capability of a peptide, first the background luminescence (cell-only without virus or peptide) was subtracted and set relative the corresponding virus-only infectivity, which was simultaneously measured on each well plate. Then the infectivity was calculated relative to EF-C infectivity at 1.3 µM, which was applied as a reference peptide to ensure comparability of infectivity data among different experiments. Noteworthy, the logarithm of the infection enhancement relative to EF-C (Log Infection Rel. EF-C) correlated with a higher Pearson *R* with peptide properties compared to absolute infection enhancement values (**Figure S5**). To enable linear regression, we therefore applied logarithmic scale on infectivity relative to EF-C for all the correlations. The asymmetric error bars for the logarithmic plot were calculated according to eq (S1) and eq (S2) for positive σ^+ and negative σ^-errors, respectively.

$$\sigma^+ = \log_{10}(\text{Infection Rel EF-C} + \Delta\,\text{Infection Rel EF-C}) - \log_{10}\text{Infection Rel EF-C} \quad \text{eq (S1)}$$

$$\sigma^{-} = \log_{10} \text{Infection Rel EF-C} - \log_{10} (\text{Infection Rel EF-C - } \Delta \text{ Infection Rel EF-C}) \quad - \text{eq (S2)}$$

1.2 Bioinformatic parameters

A selection of bioinformatic parameters were calculated for all peptides of the peptide library according to previously reported protocols.[17,18] The descriptors include peptide net charge, hydrophobicity, isoelectric point, aliphatic index, hydrophobic moment index, boman index, instability index, graph shape index, upsilon steric parameter, smoothed upsilon steric parameter, polarizability, normalized van der Waals volume (**Table S4**). It must be noted that the bioinformatic descriptors only evaluate the primary sequence of peptides and proteins and not aggregation propensities. **Net charge.** The sum of the charges of amino acid side chains at a certain pH is the net charge. The net charge was calculated (pH = 7.4, pKscale = EMBOSS) with the "peptides" package in R.[18] **Hydrophobicity.** The hydrophobicity gives information about average hydrophobicity of the peptide calculated by hydrophobicity of amino acid side chains. The calculation of the hydrophobicity was applied according to Fauchère[17] as an experimental based on water–octanol partition coefficients and Kyte-Doolittle[19] as a calculation based hydrophobicity scales. The calculation of hydrophobicity was conducted via the "peptides" package in R.[18] **Isoelectric point (pI)** The pI describes the pH at which all negative and positive charged side chains of a peptide are statistically balanced (total net charge neutral). Since all peptides are incubated at pH 7.4 in phosphate buffer the side chains of the peptides with pI higher than 7.4 are positively charged. Peptides with pI lower than 7.4 are accordingly negative charged. The pI was calculated with the pKscale EMBOSS via "peptides" package in R.[18] **Aliphatic index.** The aliphatic index indicates the thermostability by evaluating the amount of aliphatic amino (alanine (A), isoleucine (I), leucine (L) and valine (V)) acids in a protein.[20] The aliphatic index was calculated via the "peptides" package in R.[18] **Hydrophobic moment index.** The hydrophobic moment index is proposed to measure amphiphilicity perpendicular to the axis of a peptide structure.[21] The calculate of the hydrophobic moment index (angle = 100, window = 11) was conducted with the "peptides" package in R.[18] **Boman index.** The sum of solubility for amino acid side chains normalized by the number of amino acids gives an estimation of the peptide-protein or peptide-membrane interaction. A boman index, which is negative or nearly 0, is assumed to be connected to high membrane interaction, e.g. for antimicrobial activity.[22] The boman index is calculated with the "peptides" package in R.[18] **Instability index.** The *in vivo* stability of a protein is estimated by prevalence of pairs of amino acid in a peptide. Thereby the direct proximity of certain amino acids in stable and unstable peptides is evaluated and gives back information on composition in a simplified way. A peptide with a instability index below 40 is considered to be stable.[23] The instability index is calculated with the "peptides" package in R.[18] **Graph shape index.** The graph shape index measures the sterical demand (branching, symmetry, complexity) of the amino acid side chains. The index for a peptide was calculated by dividing through number of amino acids after adding up the respective single amino acid graph shape descriptors.[17] **Upsilon steric parameter.** The upsilon steric parameter measures steric effects by considering the minimal van der Waals radius of amino acid side chains. The index for a peptide was calculated by dividing through number of amino acids after adding up the respective single amino acid descriptors.[17] **Polarizability.** The polarizability is a bulk parameter describing the molecular volume and models dispersion forces. It is related to molar refractivity, which gives information on electrical properties of a molecule. The index for the peptide was calculated by dividing through number of amino acids after adding up the respective single amino acid descriptors.[17] **Normalized van der Waals volume.** The van der Waals volume is another bulk parameter correlating with polarizability and was calculated for a peptide by

dividing through number of amino acids after adding up the respective single amino acid side chain van der Waals volume.[17]

Linear Regression. For conducting multiple linear regression, the constrained multiple linear regression tool (v.1.10) of OriginPro, Version 2021 was applied. As a dependent variable (y) the Log Infection Rel EF-C was selected, the independent variables (A) were the features of interest. The selected model for regression analysis is constrained linear least-squares problems solution with the following general equation: y = A0+A1*x1+A2*x2+[…].

2. Cell Viability of Representatively Selected Peptides

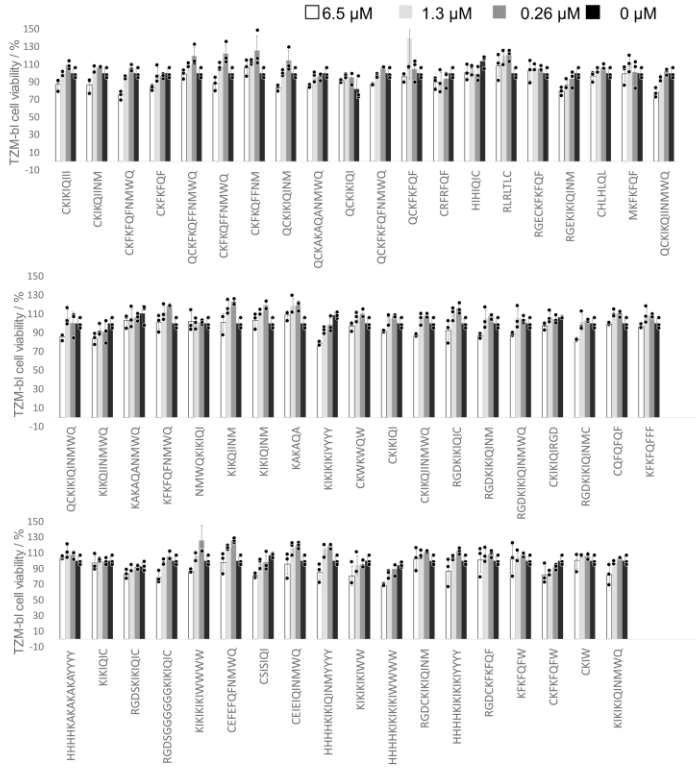


Figure S2 Cell Viability normalized to 100% metabolic activity as determined by CellTiterGlo Assay of 57 representatively selected peptides from the library. The selection represents infectivity enhancing and not enhancing peptides (**Table S5**) and show that cell viability (> 80%) is maintained for all peptides at 1.3 µM concentration. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S2 is provided.

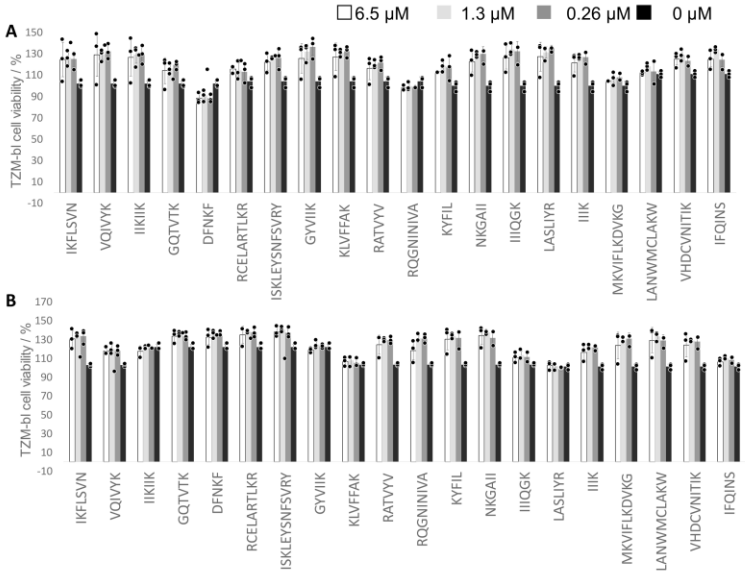


Figure S3 Cell Viability normalized to 100% metabolic activity determined by CellTiterGlo Assay of 20 peptides reported in literature.[33,34,43–48,35–42] Cell viability (> 80%) is maintained for all peptides at 1.3 µM concentration regardless of the preparation method, which is **A** PNF formation method by dilution of 10 mg/mL peptide stock in DMSO in PBS to 1 mg/mL or **B** preparation according to literature reported protocol (**SI Section 8**). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S3 is provided.

3. Timescale of peptide aggregation

Fibril formation and µm-sized aggregation occur immediately upon incubation of peptides in PBS as demonstrated exemplary for CKFKFQF with TEM micrographs (**Figure S4**A, B) and by turbidity as observed via bare eye (data not shown), and light scattering exemplary shown for CKFKFQF (**Figure S4**C) and QCIKIQINMWQ (**Figure S4**D). No significant change in light scattering count rate is observed regardless of measurement directly after sample preparation (T_0) or after incubation for 1 day (T_{1d}) or 2 months (T_{2m}), indicating that the fibril formation and aggregation is completed within the timeframe of sample preparation (within seconds).

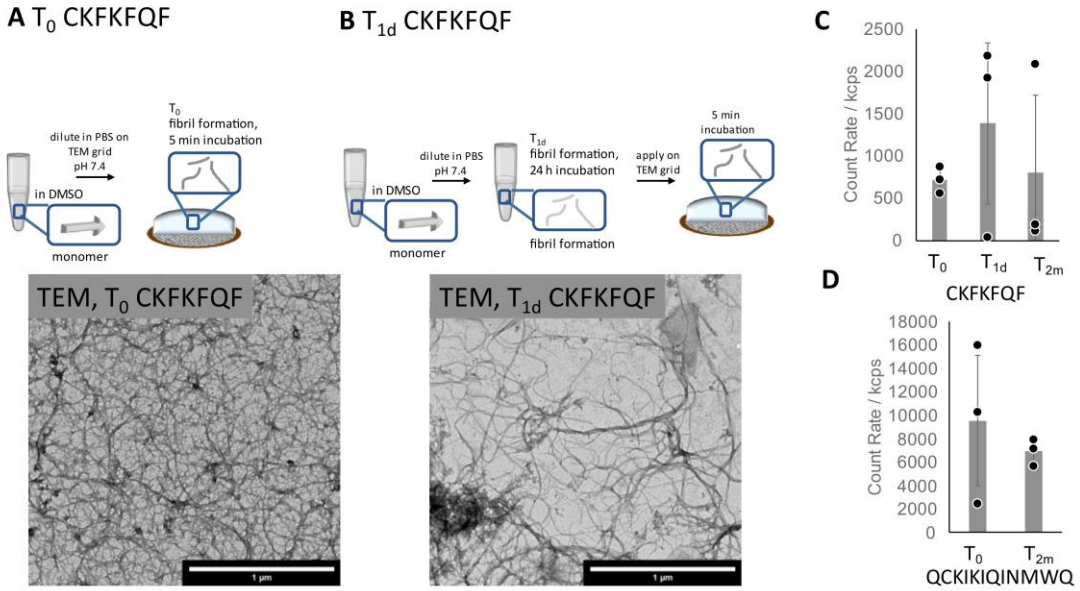


Figure S4 Time dependent fibril-formation and aggregation of selected peptides CKFKFQF and QCKIKIQINMWQ. **A** Schematic illustration and TEM micrograph of T_0 incubation condition. The peptide monomer, dissolved in DMSO (10 mg/mL), was added to a PBS droplet on a TEM grid and incubated at 1 mg/mL concentration for 5 min. Fibril formation is observed via TEM (scale bar 1 µm). **B** Schematic illustration and TEM micrograph of T_{1d} incubation condition. The peptide monomer, dissolved in DMSO (10 mg/mL), was added to PBS and incubated at 1 mg/mL concentration for 24 h. Then, the peptide fibrils were incubated on a TEM grid for 5 min. Fibril formation was observed via TEM (scale bar 1 µm). **C** The count rate of scattered light was determined from Zeta-Potential measurements for CKFKFQF (0.1 mg/mL) after T_0, T_{1d}, and T_{2m}. For T_0 incubation condition, the peptide was diluted from DMSO (10 mg/mL) in PBS (1 mg/mL) and immediately further diluted in 1 mM KCl to 0.1 mg/mL for measurement, which tookapproximately 3 min. For T_{1d} and T_{2m} measurements, the sample was incubated at 1 mg/mL concentration for one day and two months, respectively, and diluted in 1 mM KCl to 0.1 mg/mL just before measurement. Error bars indicate standard deviation from triplicate measurements. **D** The count rate of scattered light was determined from Zeta-Potential measurements for QCIKIQINMWQ (0.1 mg/mL) after T_0, and T_{2m}. Created with BioRender.com. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S4 C, D is provided.

4. Single Parameter Correlations

4.1 Correlation matrix

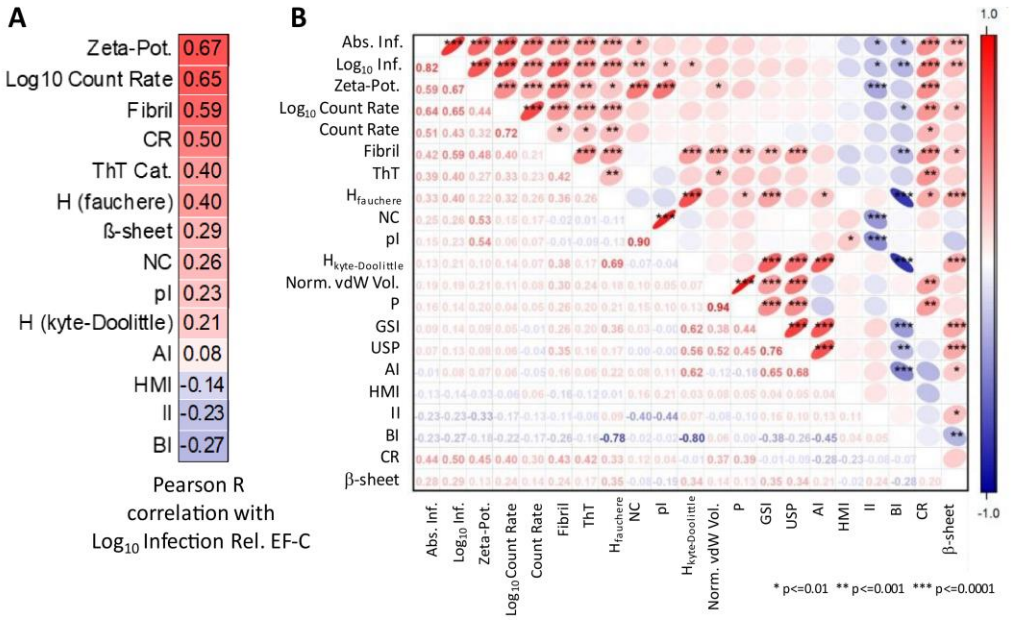


Figure S5 A Pearson *R* correlation factor of Log_{10} Infection Rel EF-C values with experimental and calculated properties of 163 peptides of an EF-C based library. **B** Correlation matrix of physicochemical and bioinformatic properties of a peptide library based on EF-C (**Table S4**, **Table S5**). All columns except of Conversion Rate (CR) contain data of 163 peptides. Conversion rate was determined for 80 representative subset of peptides (**Table S5**). Abbreviations: Absolute Infectivity at 1.3 µM peptide concentration (Abs. Inf.), and the logarithmic infectivity relative to EF-C (Log_{10} Inf.), the derived count rate of scattered light from zeta-potential measurements (Log_{10} Count Rate), zeta potential (Zeta-Pot.), fibril formation (Fibril) as evaluated qualitatively from TEM measurements (**Figure S33**), conversion of monomers to assembled structure (Conversion Rate, CR), percentual β–sheet content from fourier transform infrared spectroscopy (FT-IR, **Figure S34**) measurements (β–sheet), categorical ThT fluorescence (a peptide is ThT-active, if *n*-fold ThT fluorescence is >2, **Table S5**), Net charge (NC), hydrophobicity according to Fauchère scale ($H_{fauchere}$) and Kyte-Doolittle scale ($H_{Kyte-Doolittle}$), isoelectric point (pI), aliphatic index (AI), hydrophobic moment index (HMI), boman index (BI), instability index (II), graph shape index (GSI), upsilon steric parameter (USP), polarizability (P), and normalized van der Waals Volume (Norm vdW Vol.) **Table S4**. Significance level is determined via Pearson two-sided t-distribution test and indicated by p-values *p<= 0.01, **p<= 0.001, ***p<= 0.0001. Source data for Figure S5 is provided.

4.2 Correlation of infection rates with hydrophobicity and formation of β-sheet rich nanostructures

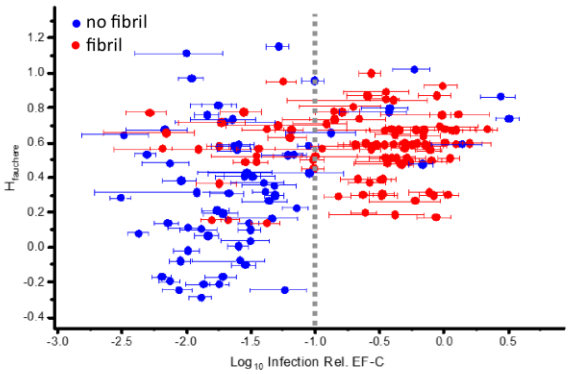


Figure S6 Graphical representation of calculated hydrophobicity (**Table S4**) plotted versus $\log_{10}$ Infection Rel. EF-C (**Table S5**). The color code represents fibril formation (red) and no fibril formation (blue) as observed via TEM (**Figure S33**, **Table S5**). Interestingly, there are no active peptides (dashed line, Infection Rel EF-C <10% or Log Infection Rel EF-C <-1) with a hydrophobicity lower than 0.17. From these active peptides 90% are forming fibrils. The majority (69%) of peptides with a hydrophobicity higher than 0.17 are also forming fibrils, illustrating a positive correlation between hydrophobicity, fibril formation and infectivity enhancement. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S6 is provided.

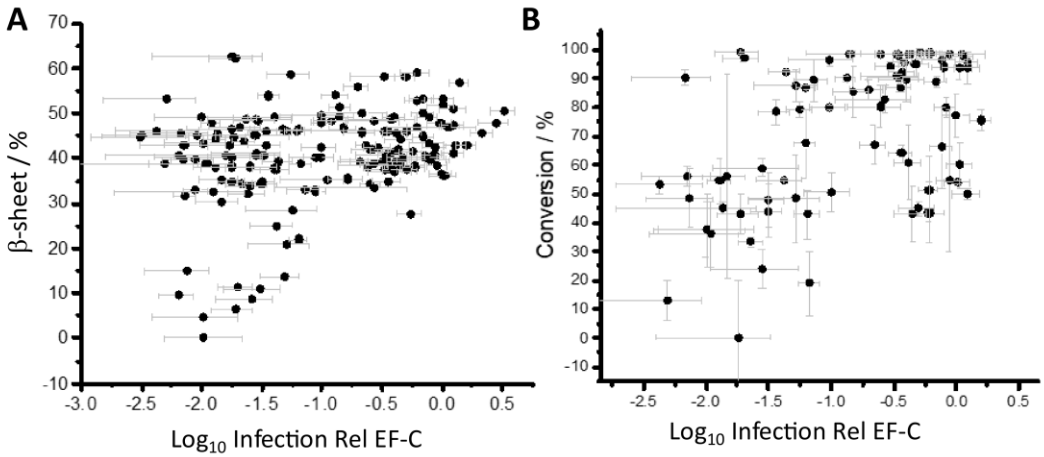


Figure S7 Characterization of β-sheet and monomer to assembly conversion of the peptide library (**Figure S34**, **Table S5**). Error bars indicate standard deviation from triplicate measurements. **A** β-sheet content obtained from attenuated total reflection (ATR)-FT-IR measurements plotted against $\log_{10}$ infection relative to EF-C. **B** Conversion of monomers to assembled structure contain data of 80 peptides plotted against $\log_{10}$ infection relative to EF-C. Source data for Figure S7 is provided.

4.3 Correlation of infection rates and count rate with µm-sized aggregate formation

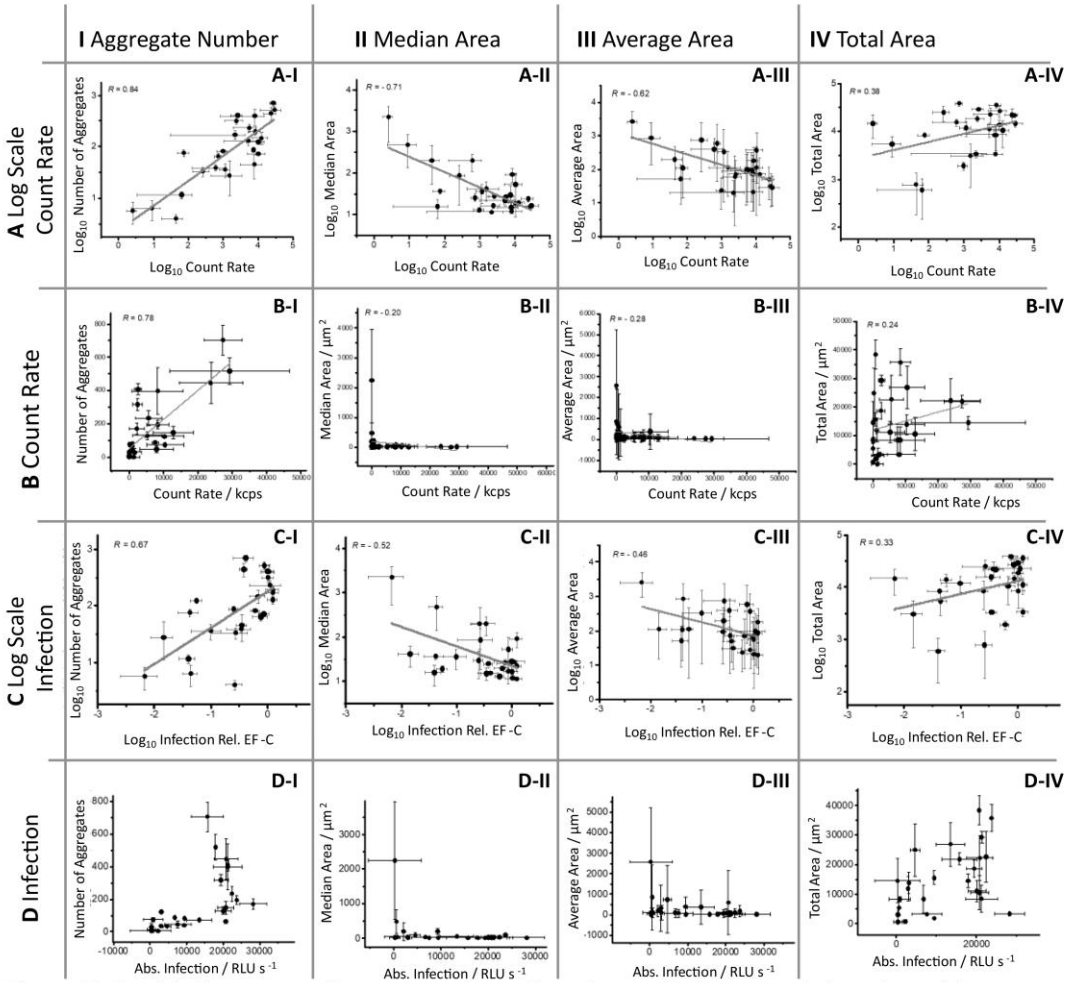


Figure S8 Graphical summary of aggregate size and number analysis as correlation plots with count rate and infection rate. **A-I** Log_{10} number of aggregates plotted against Log_{10} count rate of scattered light display a high correlation coefficient R = 0.84. **A-II** Log_{10} median area and **A-III** Log_{10} average area of aggregates plotted against Log_{10} count rate of scattered light display a negative correlation (R = – 0.71 and – 0.62, respectively) indicating that smaller aggregates are contributing stronger to count rate than larger ones. **A-IV** Plotting the Log_{10} total area against the Log_{10} count rate of scattered light show no good correlation (R = 0.38). **B-I** Number of aggregates plotted against count rate of scattered light display a high correlation coefficient R = 0.78. **B-II** Median area and **B-III** average area of aggregates plotted against the count rate of scattered light display no linear relationship and thus no good correlation (R = – 0.20 and – 0.28, respectively) indicating that log_{10} scale is a better measure to evaluate aggregates. **B-IV** Plotting the total area of aggregates against the count rate of scattered light show no good correlation (R = 0.24). **C-I** Log_{10} number of aggregates plotted against Log_{10} Infection Rel. EF-C display a good correlation coefficient R = 0.67. **C-II** Log_{10} median area and **C-III** Log_{10} average area of aggregates plotted against Log_{10} Infection Rel. EF-C display a negative correlation (R = – 0.52 and – 0.46, respectively) indicating that smaller aggregates are more active than larger ones. **C-IV** Plotting the Log_{10} total area against the Log_{10} Infection Rel. EF-C show no good correlation (R = 0.33). **D-I** Number of aggregates **D-II** median area and **D-III** average area of aggregates plotted against Absolute Infection (RLU, relative light units) display a non-linear relationship indicating that log_{10} scale is a better measure to evaluate aggregates. **D-IV** Plotting the total area of aggregates against the count rate of scattered light show no correlation. All aggregate analyses were performed in triplicate measurements (Std. Dev.) by automated 3D objects counting function (ImageJ) from widefield microscopy images (**Figure S32**), area 1330 µm × 1330 µm. The count rate and infection rate data are retrieved from **Table S8**. Mean values are displayed with error bars indicating standard deviation from triplicate measurements. Source data for Figure S8 is provided.

The light scattering count rate as well as the infectivity enhancement correlate with increasing number of visible aggregates (**Figure S8**I, R = 0.84 for Log_{10} Count Rate correlation with Log_{10} Aggregate Number (**Figure S8**AI). In contrast, having few larger aggregates (> 100 μm^2) does not enhance infectivity as can be observed from the negative correlation between infection rates and count rates in **Figure S8**II, III. We hypothesize that either too few virions can associate with the larger fibril aggregates or that many smaller aggregates can transduces many cells more efficiently than few larger aggregates. This is supported by the correlation of the cumulative surface area of the particles that, instead of the full distribution, only considers total surface area (**Figure S8**IV), and shows no correlation with the count rate or infectivity.

Mechanical downsizing of µm-sized aggregates can be achieved by applying pressure on a peptide solution which is entrapped between two glass slides (**Figure S9**AII, BII, C). However, upon releasing the physical pressure, the initial aggregate size is recovered within seconds (**Figure S9**AIII, BIII, C). This fast recovery likely originates from the fast non-covalent self-assembly of the peptide structures. This is emphasized also by ultrasonication experiments, which show comparable aggregation before and after ultrasonication (**Figure S9**D) and similar infectivity values for peptides treated and not treated with ultrasonication (**Figure S9**E). Thus, it is not possible to tune the aggregate sizes under comparable conditions for the same peptide sequence by simple mechanical forces.

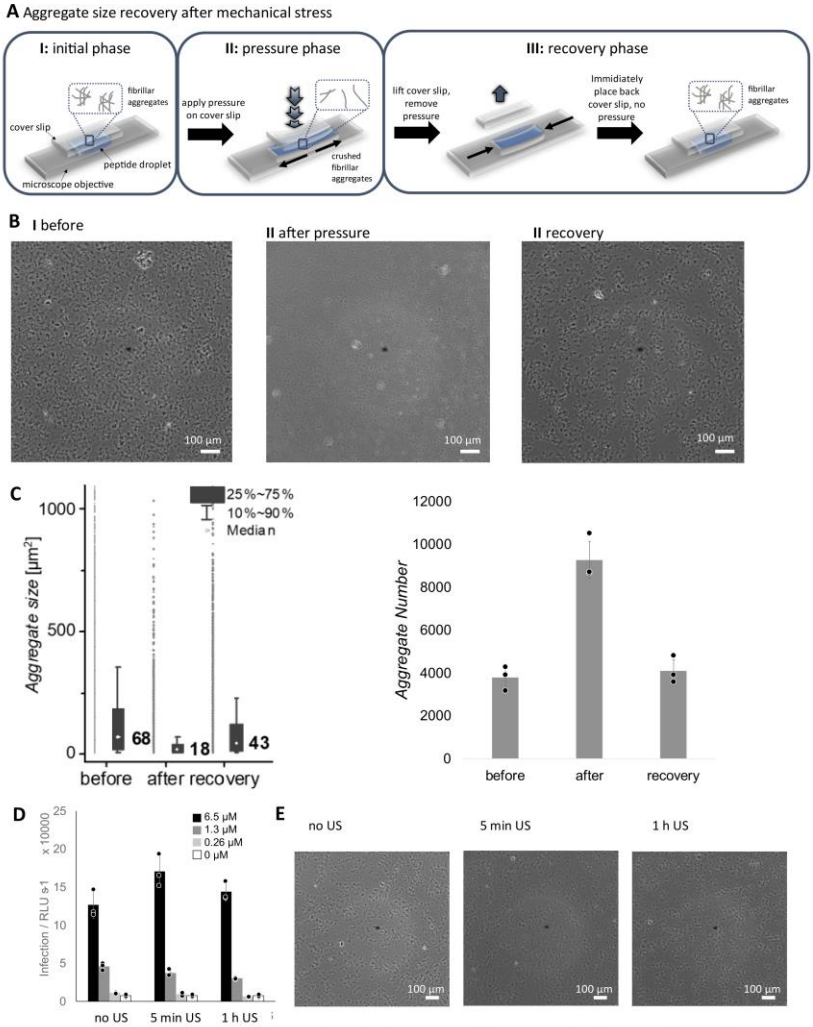
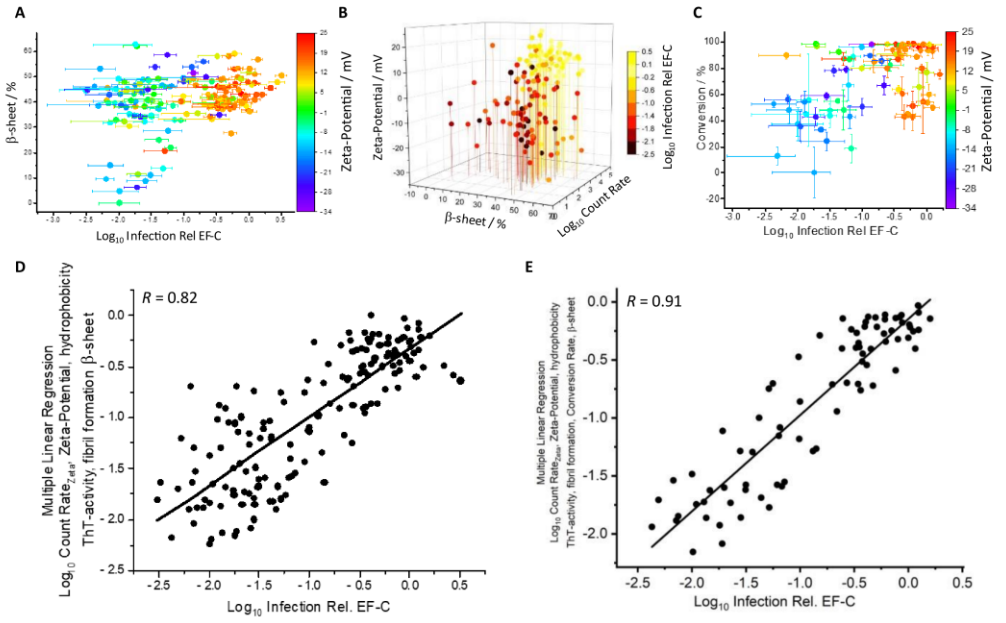


Figure S9 Changes in aggregate size upon mechanical stress. **A** Schematic illustration of **I** aggregates composed of self-assembling peptide fibrils placed on a microscope objective and covered by a glass cover slip. **II** Applying pressure on glass cover slip mechanically downsizes aggregates, **III** however, once the pressure is released (*e.g.* by lifting the cover slip) the initial aggregate size is recovered. **B** Brightfield microscopy measurements (objective 20x) of CKFKFQF (1 mg/mL, 30 µL) covered by a glass cover slip, **I** before applying pressure, **II** after applying pressure and **III** after releasing pressure by lifting cover slip, scale bar 100 µm. **C** Aggregate size distribution for objects > 10 µm^2 in an area of 1331 µm × 1331 µm. Box plots show median with whiskers indicating 10-90% and boxes indicating 25-75% of all data displayed on the left side to each box. In the bar plot mean values are displayed with error bars indicating standard deviation from triplicate measurements. **D** Applying ultrasonication (US) on CKFKFQF (1 mg/mL in PBS) for 5 min or 1 h does not affect infection rates or **E** aggregate size and number compared to no ultrasonication, brightfield microscopy scale bar 100 µm. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S9 C, D is provided.

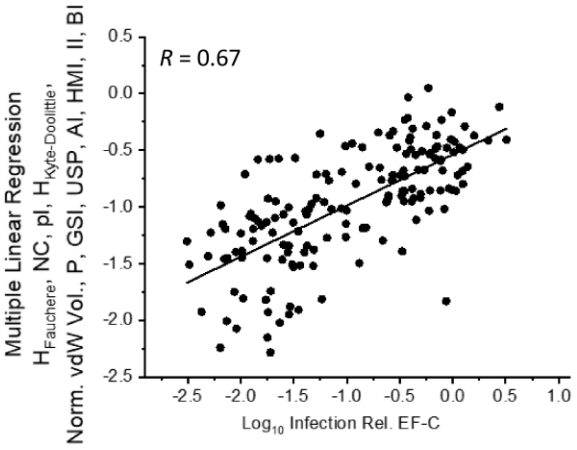
5. Multiparameter Correlations



	Value	Standard Error	t-Value	P >|t|
Intercept	-2.33472	0.16267	-14.3527	0
Zeta-Pot.	0.02128	0.00311	6.83608	1.79E-10
Log_{10} Count Rate	0.29879	0.04746	6.29522	3.05E-09
Fibril	0.27355	0.09097	3.00697	0.00308
$H_{Fauchere}$	0.26241	0.12751	2.0579	0.04129
ThT-activity	0.10774	0.07976	1.35076	0.17876
β-sheet	0.00356	0.00323	1.10357	0.2715

	Value	Standard Error	t-Value	P >|t|
Intercept	-2.21225	0.18525	-11.9422	0
Zeta-Pot.	0.01798	0.00283	6.34982	1.79E-08
Fibril	0.47174	0.10511	4.48813	2.71E-05
Log_{10} Count Rate	0.20446	0.0478	4.27703	5.81E-05
ThT-activity	0.19678	0.10179	1.93317	0.0572
β-sheet	0.00698	0.00374	1.86557	0.06623
$H_{Fauchere}$	0.06023	0.13112	0.45933	0.6474
Conversion	2.65E-04	0.00173	0.15333	0.87857

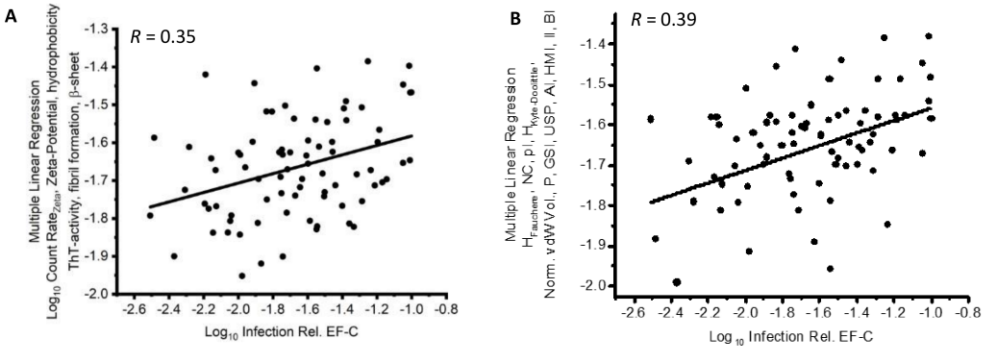
Figure S10 Multiparameter analysis of assembly (**Figure S33**), secondary structure (**Figure S34**), surface charge and microscopic aggregation (**Table S5**). **A** β-sheet content plotted against Log_{10} Infection Rel. EF-C. The color bar indicates the zeta-potential. Error bars indicate standard deviation from triplicate measurements. **B** Peptides with a high β-sheet content, which are microscopically aggregating (high count rate) and have a positive zeta-potential are enhancers of viral transduction (top right corner). The color bar indicates Log_{10} Infection Rel. EF-C. **C** Conversion of monomers to assembled structures (Conversion Rate, CR) plotted against Log_{10} Infection Rel. EF-C of 80 selected peptides (**Table S5**). The color bar indicates the zeta-potential. Error bars indicate standard deviation from triplicate measurements. Note, that for one peptide (CKFKQFF) with -2.2 Log_{10} Infection Rel to EF.C shows no infectivity enhancement despite its positive zeta-potential and high CR. For this peptide the turbidity is very low and amorphous aggregates rather than fibril formation is observed, which indicates no microscopic aggregation. **D** Multiple linear regression (*constrained linear least-squares problems solution) with the equation: y =* A0+A1*x1+A2*x2+[…]+ A5 * x5 and zeta-pot, Log_{10} Count Rate, Fibril formation, $H_{fauchere}$, ThT-activity and β-sheet as independent variables (A, listed in the order of ascending P variable) and Log_{10} Infection Relative EF-C as dependent target variable (y). Linear Fit with Pearson correlation factor 0.82. Statistically significant parameters ($p < 0.05$) are highlighted. **E** Multiple linear regression (constrained linear least-squares problems solution) with the equation: y = A0+A1*x1+A2*x2+[…]+ A7 * x7 and Fibril formation, zeta-pot, Log_{10} Count Rate, β-sheet, $H_{fauchere}$, ThT-activity and Conversion Rate as independent variables (A, listed in the order of ascending P variable) and Log_{10} Infection Relative EF-C as dependent target variable (y) with 79 peptides from the library which include CR as well as FT-IR data. Linear Fit with Pearson correlation factor 0.91. Significance level is determined via Pearson two-sided t-distribution test and indicated by p-values. Statistically significant parameters ($p < 0.05$) are highlighted. Source data for Figure S10 is provided.



	Value	Standard Error	t-Value	Prob>|t|
Intercept	-3.63904	0.839	-4.33736	2.63E-05
$H_{Fauchere}$	1.79563	0.34155	5.2573	4.95E-07
P	-30.14357	6.52932	-4.61665	8.32E-06
AI	-0.00839	0.0021	-3.98661	1.04E-04
HMI	-1.15218	0.30454	-3.78336	2.23E-04
USP	3.99456	1.36645	2.92331	0.004
Norm. vdW Vol.	0.70321	0.24673	2.8501	0.00499
GSI	0.6973	0.26492	2.63217	0.00937
II	-0.00259	0.00127	-2.04129	0.04297
NC	0.12281	0.0659	1.86361	0.06433
pI	0.09049	0.05484	1.6499	0.10106
BI	0.10473	0.08246	1.27011	0.20601
$H_{Kyte-Doolittle}$	-0.05017	0.14137	-0.35489	0.72317

Figure S11 Multiple linear regression (constrained linear least-squares problems solution) with the equation: y = A0+A1*x1+A2*x2+[…]+ A12 * x12. The independent variables are Net charge (NC), hydrophobicity according to Fauchère scale ($H_{fauchere}$) and Kyte-Doolittle scale ($H_{Kyte-Doolittle}$), isoelectric point (pI), aliphatic index (AI), hydrophobic moment index (HMI), boman index (BI), instability index (II), graph shape index (GSI), upsilon steric parameter (USP), polarizability (P), normalized van der Waals Volume (Norm vdW Vol.), **Table S4**. The independent variables (A) are listed in the order of ascending P variable. Log_{10} Infection Relative EF-C is the dependent target variable (y) and linear fit with Pearson correlation factor is 0.67. Significance level is determined via Pearson two-sided t-distribution test and indicated by p-values. Statistically significant parameters ($p < 0.05$) are highlighted. Source data for Figure S11 is provided.

To avoid overfitting of the multiple linear regression control fits were conducted with non-active peptides and showed no statistically significant correlation (**Figure S12**).



	Value	Standard Error	t-Value	Prob>|t|
Intercept	-1.88885	0.1721	-10.9754	0
Log_{10} Count Rate	0.08681	0.06463	1.34317	0.18338
ThT-activity	0.12609	0.09898	1.2739	0.20674
$H_{Fauchere}$	0.13761	0.12161	1.13159	0.26151
Zeta-Pot.	0.00351	0.00411	0.85224	0.39687
Fibril	0.05948	0.10263	0.57958	0.56398
β-sheet	-0.00119	0.00315	-0.37622	0.70784

	Value	Standard Error	t-Value	Prob>|t|
Intercept	-2.31348	0.7412	-3.12127	0.00264
HMI	-0.48614	0.25335	-1.91886	0.0592
P	-7.86118	4.9337	-1.59336	0.11572
$H_{Fauchere}$	0.45682	0.30324	1.50646	0.13658
AI	-0.00162	0.0016	-1.01192	0.31516
Norm. vdW Vol.	0.16416	0.18077	0.90807	0.36705
USP	1.05354	1.16697	0.9028	0.36982
GSI	0.17629	0.19697	0.89501	0.37394
pI	0.03747	0.04735	0.79131	0.43152
NC	0.03404	0.05421	0.62789	0.53218
II	3.70E-04	9.64E-04	0.38393	0.70223
BI	0.01619	0.06266	0.2584	0.79688
$H_{Kyte-Doolittle}$	-0.01675	0.11812	-0.14179	0.88767

Figure S12 Multiparameter correlation of experimental and bioinformatic features of 81 peptides, which are not enhancing infectivity (**Table S5**, Infection Rel EF-C < 10%) conducted to check for overfitting. As expected, there are no statistically significant correlations found. **A** Multiple linear regression (constrained linear least-squares problems solution) with the equation: y = A0+A1*x1+A2*x2+[…]+ A5 * x5 and zeta-pot, Log10 Count Rate, Fibril formation, $H_{fauchere}$, ThT-activity and β-sheet as independent variables (A, listed in the order of ascending P variable) and Log10 Infection Relative EF-C as dependent target variable (y). Linear Fit with Pearson correlation factor 0.35. **B** Multiple linear regression (constrained linear least-squares problems solution) with the equation: y = A0+A1*x1+A2*x2+[…]+ A12 * x12. The independent variables (A) are all bioinformatic parameters and listed in the order of ascending P variable, **Table S4**. Log10 Infection Relative EF-C is the dependent target variable (y) and linear fit with Pearson correlation factor is 0.39. Significance level is determined via Pearson two-sided t-distribution test and indicated by p-values. Statistically significant parameters ($p < 0.05$) are highlighted. Source data for Figure S12 is provided.

6. Selected examples showcasing structure–property–activity relationship

The thresholds for classification were set as described in the following. Internal order was found if fibril formation is observed via TEM (**Figure S33**) and characteristic β-sheet peak at 1630 cm^{-1} is observed via FT-IR (**Figure S34**). Microscopic aggregation is positive for peptides showing count rate of scattered light above 100 kcps (**Table S5**) and observable aggregates via brightfield or fluorescence microscopy. The surface charge was assigned via the zeta-potential and considered as positive above a threshold of + 5 mV. Peptides were assigned to be infectivity enhancing if the Infection Rel EF-C at 1.3 µM is greater than 0.1 (> 4 fold of virus only infectivity).

Note, that since the selected thresholds are set arbitrarily and the infectivity enhancement of the peptides are continuous multiparameter dependent properties not all peptides can be covered in the illustrated schemes.

6.1 *N*-terminal cysteine leads to more aggregation and increases infectivity enhancement

The presence of *N*-terminal cysteine has a great impact on aggregation and infectivity enhancement as observed systematically in our library. Peptides which do not contain *N*-terminal cysteine at the first

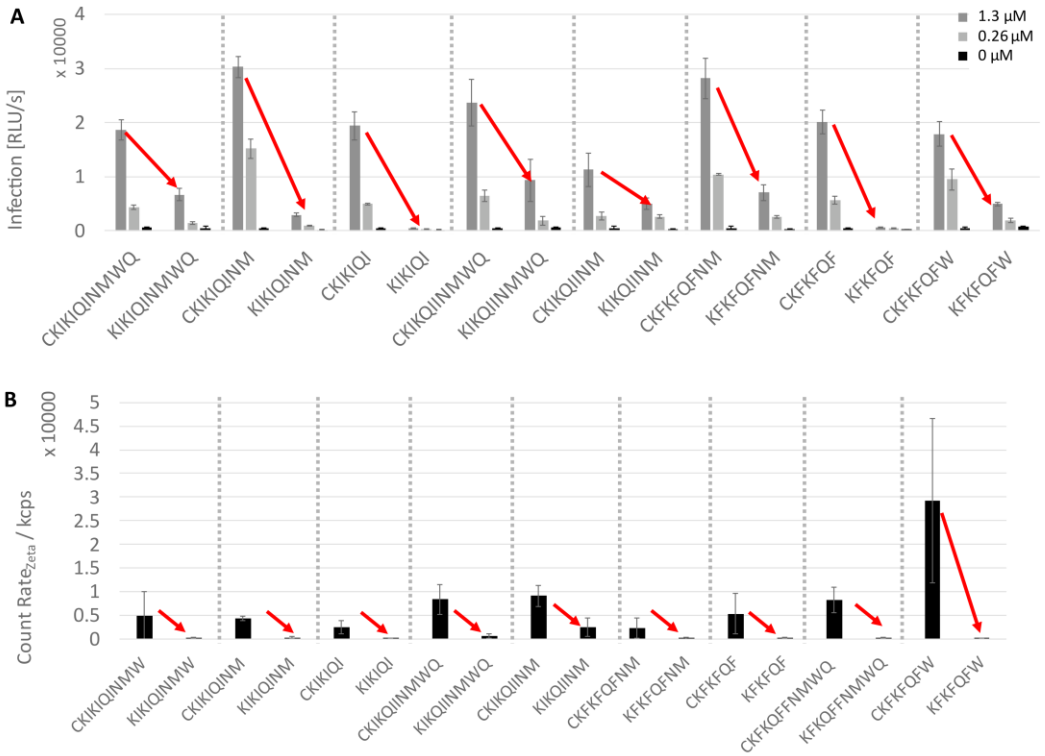


Figure S13 A Comparison of peptides' infection enhancement with and without *N*-terminal cysteine at conc. 1.3, 0.26 and 0 µM (**Table S5**). **B** Count Rate of scattered light of peptides with and without *N*-terminal cysteine (**Table S5**). Red arrows indicate the overall trend showing that infection enhancement and count rate decreases in the absence of an *N*-terminal cysteine. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S13 is provided.

position (**Figure S13**) are less prone to aggregate into µm-sized particles and are inefficient at enhancing infectivity.

Comparing sequences CKIKIQI and KIKIQIC, which switch the *N*-terminal cysteine to the *C*-terminus show, a stronger infectivity enhancement for CKIKIQI (**Figure S14**A). Remarkably, there is no significant change in the intermolecular β-sheet content (ATR-IR **Figure S14**C). However, assembled fibrils of CKIKIQI aggregate more strongly (TEM **Figure S14**E), which results in larger numbers of µm-sized aggregates (ThT-stained peptides via fluorescence microscopy, Count rate of scattered light **Figure S14**B, D, F). We hypothesize, that the higher hydrophobicity of cysteine compared to lysine residues at the N-terminus facing outwards from the fibril, lead to higher fibril-fibril aggregation which might eventually also facilitate the interaction of virus and cells.

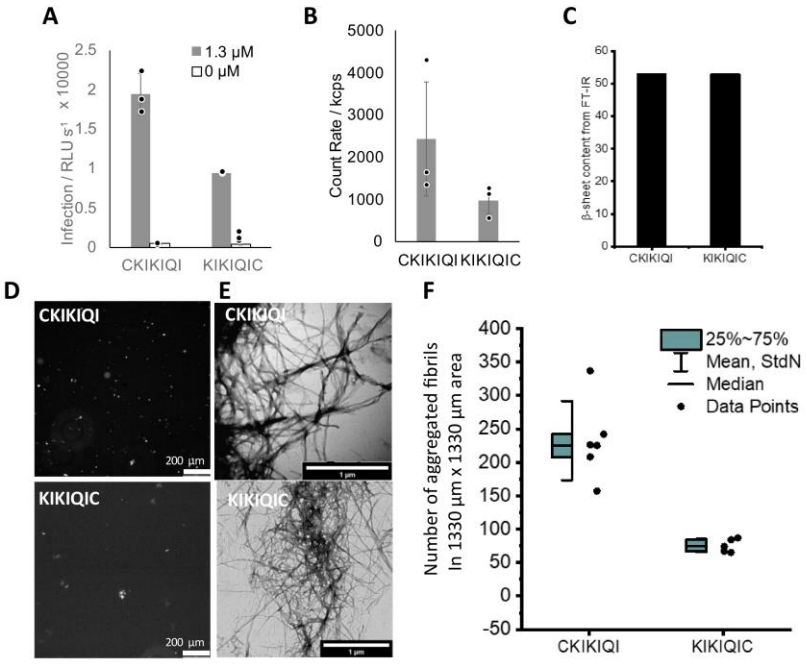


Figure S14 Comparing experimental data for CKIKIQI and KIKIQIC. **A** Infectivity enhancement, **B** Count Rate, **C** β-sheet content, **D** fluorescence microscopy of ThT-stained fibrils (scale bar 200 µm) and **E** TEM micrographs (scale bar 1 µm). **F** Number of ThT-stained µm-sized aggregates (>10 µm² in 1330 µm × 1330 µm) determined by fluorescence microscopy. Box plots show median with whiskers indicating mean value with standard deviations and boxes indicating 25-75% of all data displayed on the right side to each box. The bar plot error bars indicate std dev. from triplicate measurements. Source data for Figure S14 is provided.

6.2 Peptides with high hydrophobicity can enhance infectivity even if they have negative zeta-potential or do not form fibrils

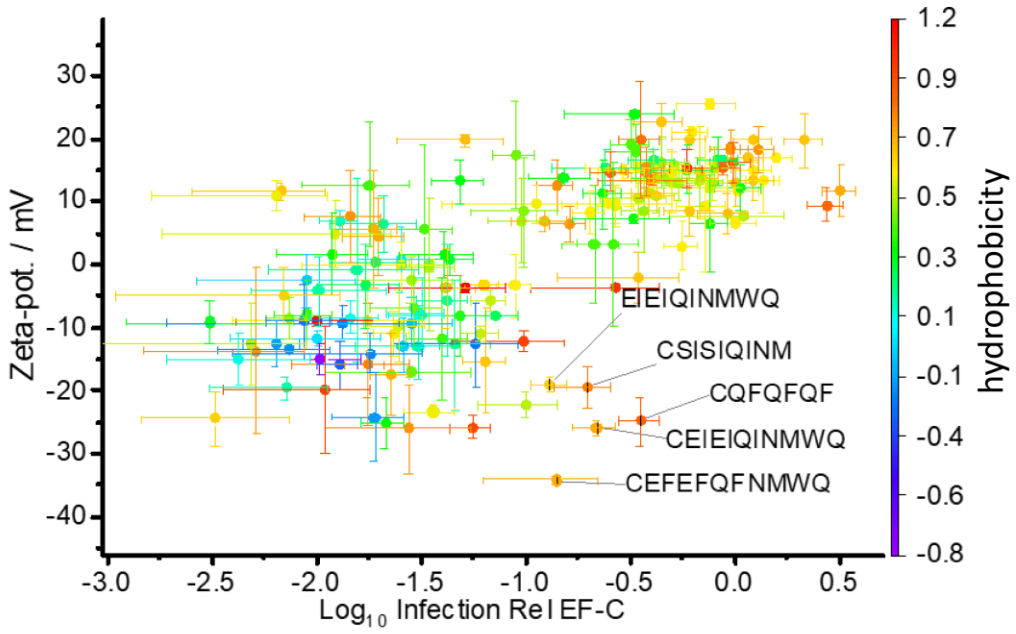
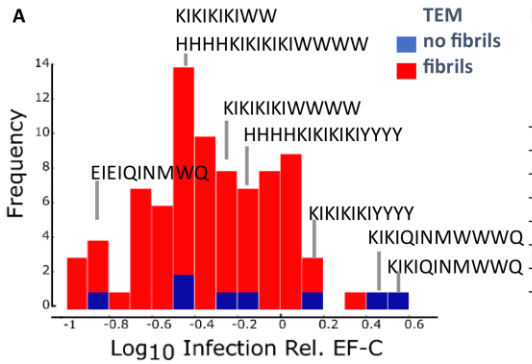


Figure S15 Peptides with negative zeta-potential (marked) slightly enhance infectivity if they have high (>0.6) calculated hydrophobicity (Fauchère, **Table S4**) and fulfill other prerequisites such as fibril formation and microscopic aggregation (**Table S5**). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S15 is provided.



B

sequence	Log_{10} Infection Rel EF-C	Zeta Pot. / mV	Count Rate / kcps	β-sheet / %	Hydro phob.
KIKIQINMWWQ	0.51	11.8	3893	50	0.74
KIKIQINMWWWQ	0.44	9.45	7596	48	0.86
KIKIKIKIYYYY	0.14	13.3	9918	57	0.59
HHHHKIKIKIKIYYYY	-0.17	13.2	12793	53	0.48
KIKIKIKIWWWW	-0.23	15.2	12421	38	1.02
HHHHKIKIKIKIWWWW	-0.42	14.5	23757	49	0.80
KIKIKIKIWW	-0.43	15.4	9436	49	0.77
EIEIQINMWQ	-0.88	-19.13	1264	54	0.66

Figure S16 Peptides which are enhancing infectivity (> 10% relative to EF-C, 4-fold relative to virus only, **Table S5**), which do not form fibrils as analyzed by TEM (**Figure S33**). **A** Distribution plot of log_{10} Infection Rel EF-C > -1. Peptide sequences, which do not form fibrils (blue bars) are indicated. **B** Properties of indicated peptides: All peptides which do not form fibrils but are active, aggregate and show positive zeta-potential and have a high calculated hydrophobicity resulting from the high amount of hydrophobic amino acids (W or Y, > 20%) except for the peptide EIEIQINMWQ, which shows negative zeta-potential and with a log_{10} infectivity of -0.88 (13 % Rel EF-C) is only weakly active. Source data for Figure S16 is provided.

6.3 Peptides substituted with low hydrophobic amino acid alanine do not enhance infectivity

Substituting hydrophobic residues such as phenylalanine or isoleucine with alanine results in a dramatic loss of infectivity enhancement (**Figure S17**). As shown exemplarily for CKIKIQINMWQ and CKAKAQANMWQ, the alanine substitution results in loss of self-assembly capabilities and µm-sized aggregation which can be traced back to reduced hydrophobicity of the side chain residues (**Figure S18)**.

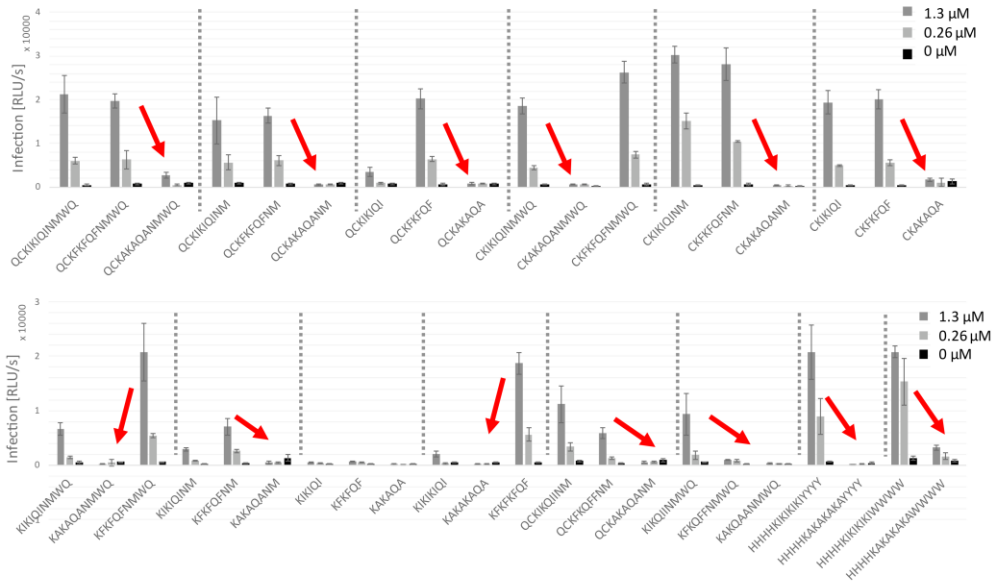


Figure S17 Infection rates of alanine, isoleucine and phenylalanine substituted amphiphilic peptide sequences at 1.3, 0.26 and 0 µM concentration. Dashed lines indicate one group of peptides, where the core sequence order stays same, arrows indicate the change of infection rates relative to 1.3 µM, phenylalanine derivatives. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S17 is provided.

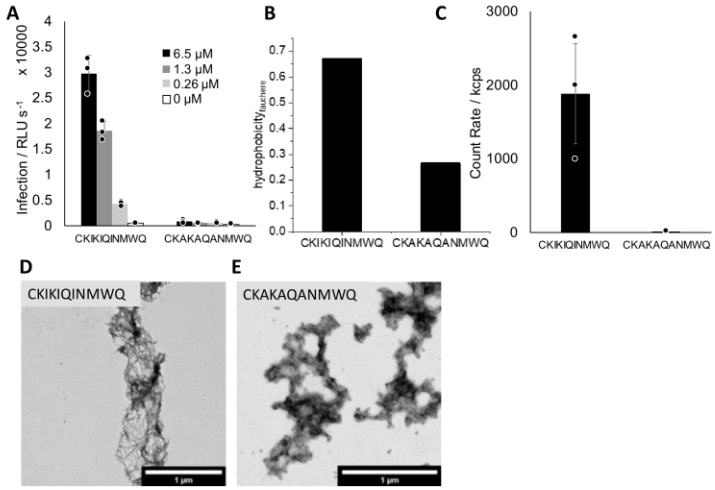


Figure S18 Comparing experimental data for CKIKIQINMWQ and CKAKAQANMWQ. **A** Infectivity enhancement, **B** Hydrophobicity calculated for the sequences according to Fauchère hydropathy scale, **C** Count Rate of scattered light **D** fibril morphology of CKIKIQINMWQ by TEM (scale bar 1 µm) and **E** amorphous aggregates morphology of CKAKAQANMWQ by TEM (scale bar 1 µm). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S18 is provided.

The reduced aggregation propensity for alanine substituted peptides and its direct connection to loss in infectivity enhancement is further emphasized with the examples HHHHKIKIKIKIWWWW / HHHHKAKAKAKAWWWW (**Figure S19**A-G) and HHHHKIKIKIKYYYY / HHHHKAKAKAKAYYYY (**Figure S19**H-N), because other than aggregation and infectivity enhancement no other physicochemical properties are changing for these sequences.

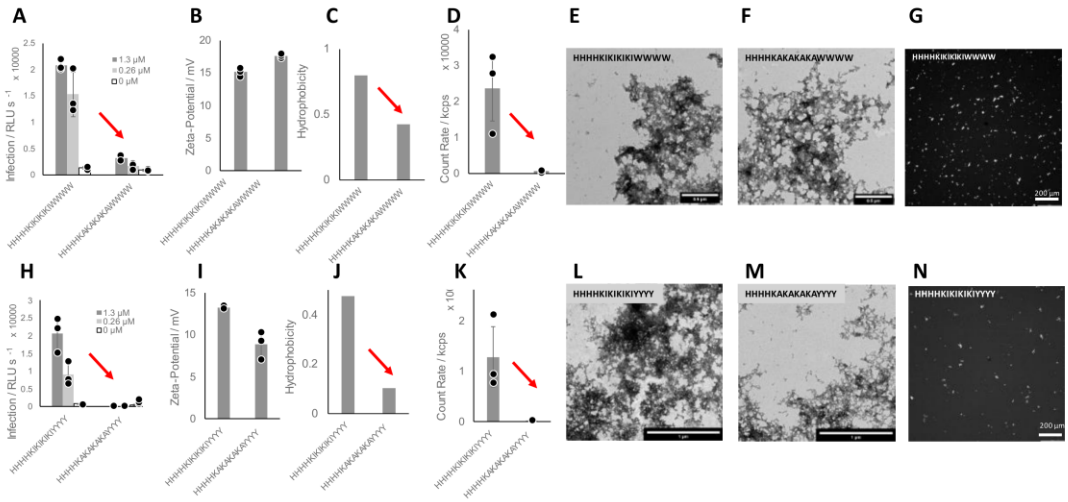


Figure S19 Alanine substitution reduces µm-sized aggregation and thus infectivity enhancement as shown exemplarily for HHHHKIKIKIKIWWWW / HHHHKAKAKAKAWWWW (A-G) and HHHHKIKIKIKYYYY / HHHHKAKAKAKAYYYY (H-N). **A, H** Infectivity data at 1.3 µM and 0.26 µM concentration, **B, I** zeta-potential, **C, J** Hydrophobicity calculated for the sequences according to Fauchère hydropathy scale, **D, K** Count Rate of scattered light **E, F, L, M** amorphous aggregate morphology by TEM (scale bar 1 µm) for HHHHKIKIKIKYYYY, HHHHKAKAKAKAWWWW, HHHHKIKIKIKYYYY and HHHHKAKAKAKAYYYY, respectively. **G, N** fluorescence microscopy of ThT-stained aggregates (scale bar 200 µm) for HHHHKIKIKIKIWWWW and HHHHKIKIKIKYYYY, respectively. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S19 is provided.

6.4 *C*-terminal sequence extensions with hydrophobic amino acids increase infectivity enhancement

If an alternating amphiphilic core sequence, which does not self-assemble is extended with hydrophobic amino acids (I), (F), (W) at the *C*-terminus, fibril formation and infectivity enhancement are observed (**Figure S20**A, B, D). The introduction of hydrophobic amino acids in an already assembling, active sequence (CKIKIQI) is not significantly affecting self-assembly or activity (**Figure S20**C).

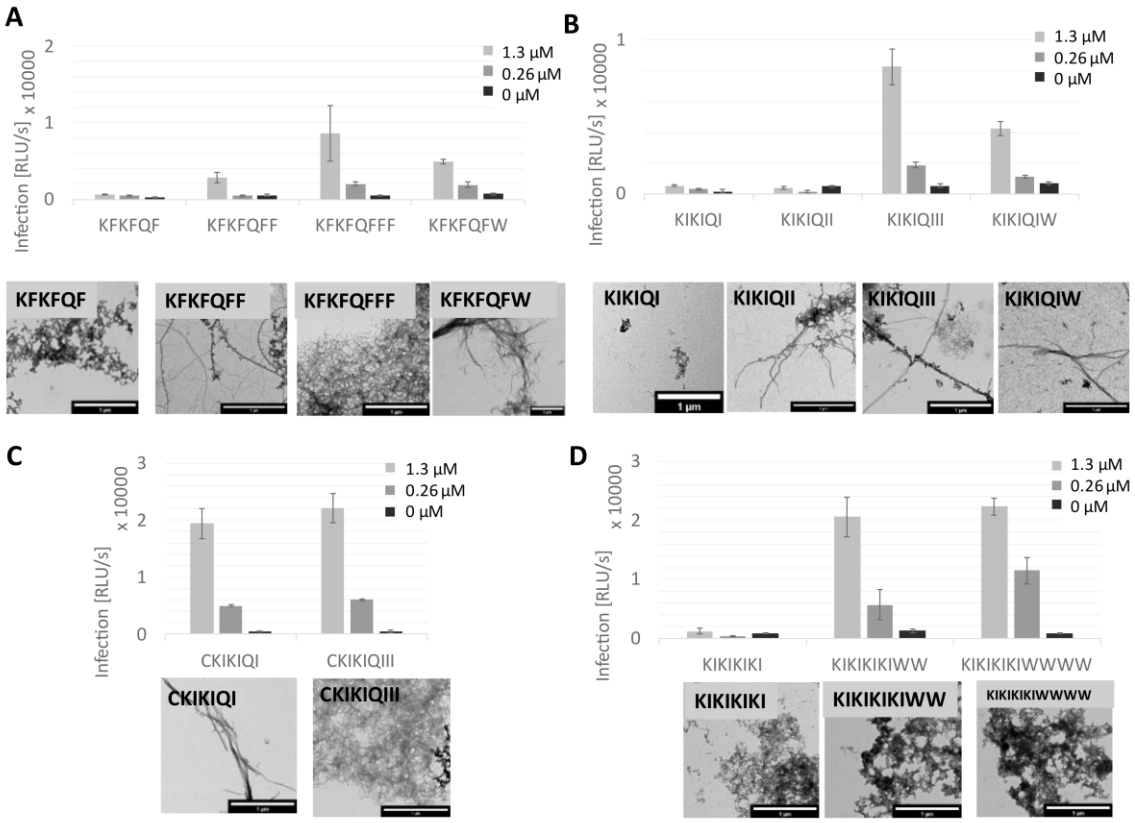


Figure S20 Hydrophobic amino acids at the *C*-terminus result in fibril formation and higher infectivity. Absolute Infection Rates of triplicate measurements of peptide concentrations at 1.3, 0.26, 0 µM. TEM measurements scale bar = 1µm. **A** Extension of the sequence KFKFQF with (F) and (W). **B** Extension of the sequence KFKFQF with (I) and (W). **C** Extension of the sequence CKIKIQI with (I). **D** Extensions of the sequence KIKIKIKI with (W). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S20 is provided.

6.5 The chaotropic substance EGCG changes physicochemical properties and highlights importance of surface charge and internal order

The addition of chaotropic agents can disrupt the β-sheet structure in amyloids.[24,25] This leads to a loss of transduction enhancing effects as was investigated through the addition of epigallocatechin gallate (EGCG) to the peptide CKFKFQF (**Figure S21**A, B) — well in accordance with previous findings for EF-C and SEVI.[26,27] While cell-viability stays remains unaffected by the addition of EGCG to CKFKFQF (data not shown), the fibrillar morphology is lost (**Figure S21**D), which is well in line with previous findings from other groups.[27] Also the intermolecular structure is changing as shown by ATR-IR measurements in which an increase of unordered structural elements can be identified (**Figure S21**C). Beside the loss of fibrillar morphology, the strongest change in physicochemical properties is the decrease of zeta-potential. However, the microscopically formed aggregates do not disappear after the addition of EGCG as clearly visible in brightfield microscopy (**Figure S21**D) and by the negligible change in scattered light count rate (**Figure S21**F). This observation was also previously observed for $PAP_{248–286}$, an amyloidal peptide, showing that addition of chaotropic substances decrease internal order but not microscopic aggregation.[28] These experiments underline that the transduction enhancing properties result synergistically from fibril formation, a strong positive zeta-potential and aggregation propensity into µm-sized particles.

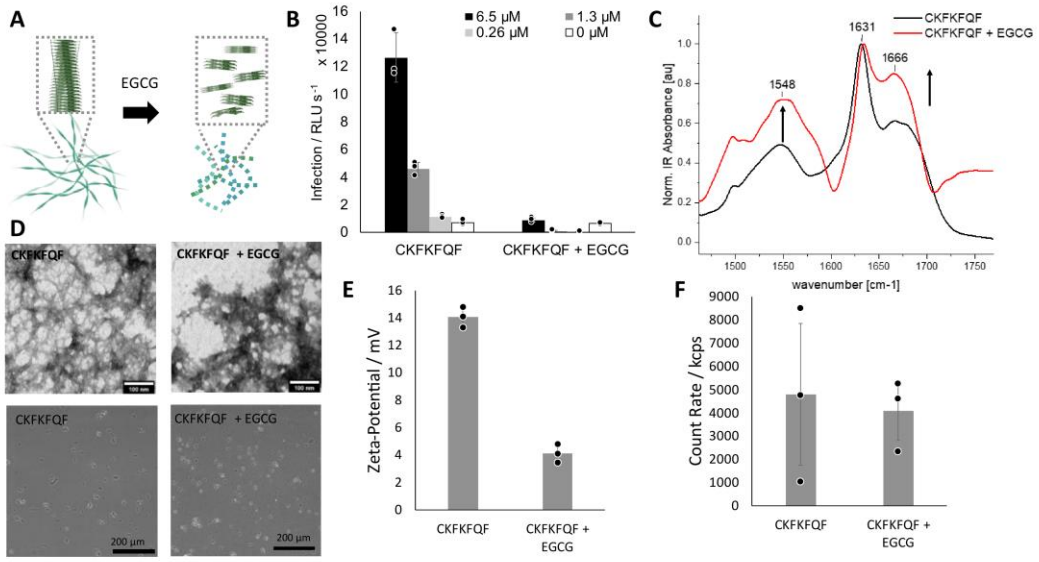


Figure S21 Effect of addition of chaotropic substance epigallocatechin-3-gallate (EGCG) to transduction enhancing peptide, shown exemplarily for CKFKFQF. EGCG is added as one molar equivalents to the preassembled peptide fibrils. **A** Schematic visualization of amyloid fibril disruptive property of EGCG. Created with Biorender.com **B** Infection rates for CKFKFQF with and without EGCG addition. **C** ATR-FT-IR measurement show an increase of unordered structural elements (1548 cm^{-1} and 1666 cm^{-1}) upon addition of EGCG. **D** TEM measurements of CKFKFQF before and after the addition of EGCG, scale bar 100 nm. Brightfield microscopy image of fibrillar aggregates before and after addition of EGCG, scale bar 200 µm. **E** Zeta-potential measurement before and after addition of EGCG. **F** No significant change is visible for the derived count rate of scattered light during DLS measurement for CKFKFQF (1 mg/mL, 1057 µM) with and without EGCG (1057 µM). Control indicates buffer with EGCG (1057 µM) only. Error bars indicate standard deviation from triplicate measurements. Created with BioRender.com. Source data for Figure S21 B, C, E, and F is provided.

7. Further information on pattern and amino acid analysis

To find recurring patterns in active peptides, the sequences in the library (**Table S3**) were simplified by coarse-graining the amino acids and the activity.[29] To this end, the amino acids were categorized in



Figure S22 Schematic visualization of pattern analysis and amino acid analysis workflow. **A** Pattern analysis **A1)** First, amino acids are coarse-grained into 4 categories according to their hydrophilicity and charge at pH 7.4 by Kyte-Doolittle classification.[19] **A2)** The coarse-grained amino acids were combined into all possible 7-mer patterns (16384) and matched with existing patterns in the peptide library (163, **Table S3**) were counted. As shown exemplarily for the sequence QCKIKQIINMWQ (EF-C) for matching it is not important at which position in the sequence the pattern occurs. Therefore, several patterns can match one sequence. **A3)** To identify patterns, which are more prevalent in high active sequences compared to medium or low active sequences the activity index (AI, eq. I) was calculated. To this end, first the thresholds for activity were set relative to EF-C. Low active sequences (81 peptides) have an infectivity < 10% rel. to EF-C, which is below 4 fold virus only activity, medium active sequences (55 peptides) have an infectivity between 10 – 70% rel. to EF-C and high active sequences (27 peptides) are defined to have an infectivity enhancement >70% rel. to EF-C (>29 fold of virus only infectivity). For the calculation of the *AI* the relative number *N* of pattern *p* in high-active sequences were subtracted from the relative *N* of *p* in medium and low active sequences to find pattern which are mostly prevalent in high active sequences and less prevalent in medium and low. **A4)** The highest *AI* was found for the alternating amphiphilic pattern $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$ which can be found in 55% and 40 % of all high-active peptides, respectively. **B** Amino acid analysis **B1)** To find amino acids mostly prevalent in high active peptides first the absolute abundance of amino acids in one category (low, medium, high) in the whole peptide library (163 sequences, **Table S3**) was determined. **2)** The relative abundance of amino acids (Rel AA, eq. II) was calculated by determining the number of peptides containing a certain amino acid relative to total amount of peptides of one category. **3)** W, M, N, I, C, Q, K, F, Y are more prevalent in high and medium active sequences than in low active sequences (highlighted box) as determined by the ratio Rel. AA high / low or medium / low > 1. Note, that all pattern and amino acid analyses were conducted with custom made python scripts. Source data for Figure S22 A4 and source code for Figure S22 A, B is provided.[30]

their two basic parameters, that are hydrophilicity and charge according to their side group properties by Kyte-Doolittle classification (H^0 hydrophobic non-charged, P^0 hydrophobic non-charged, P^- hydrophilic negatively charged, P^+ hydrophilic positively charged, **Figure S22**A).[19] The infectivity enhancement was categorized in low (< 4 fold virus only, < 10% relative to EF-C), medium (10 – 70% relative to EF-C), high (> 70% relative to EF-C). The coarse-graining approach simplifies and reduces the number of sequences to enable observation of recurring patterns while still considering the two important physicochemical properties hydrophilicity and charge.

In **Figure S22**A a schematic overview of the pattern and amino acid analysis is shown. For the identification of patterns, coarse-grained sequences from 2-mers to 7-mers were evaluated via a

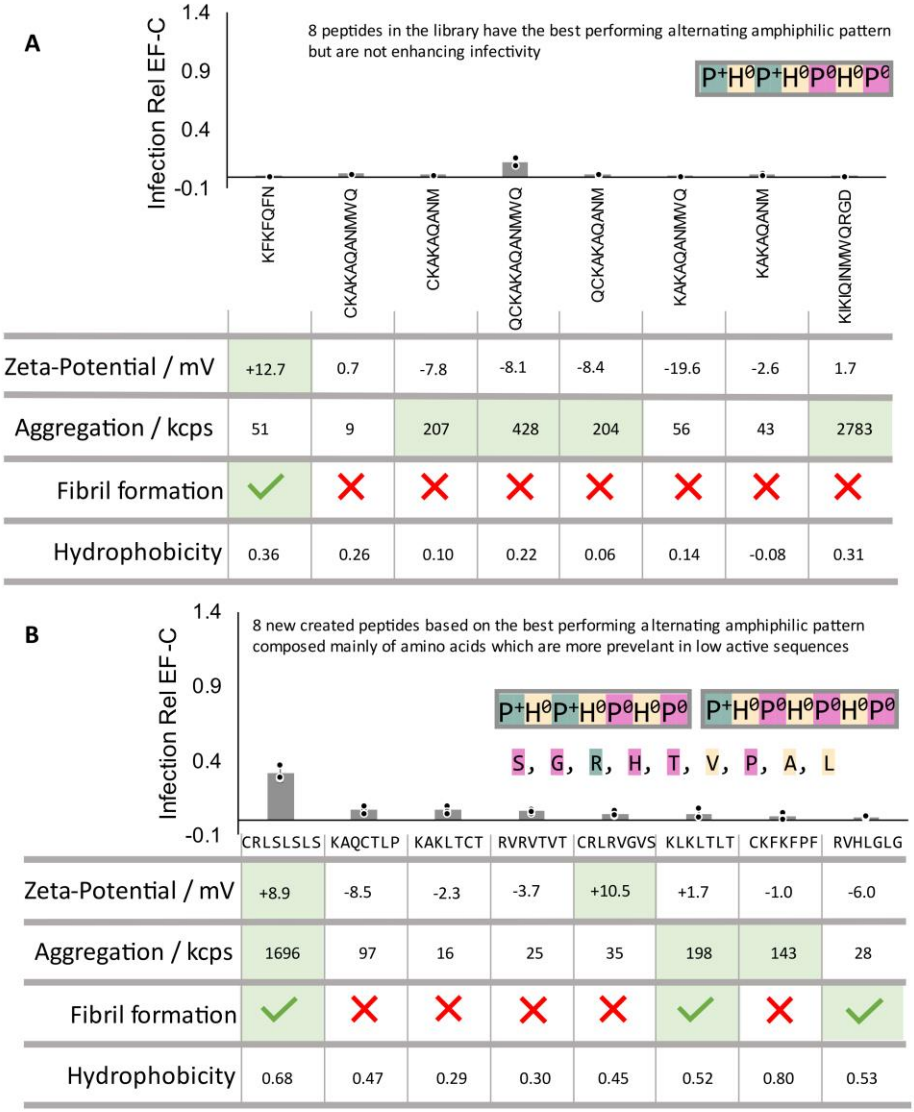


Figure S23 Peptides sharing the best performing pattern $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$, which are not enhancing infectivity. **A** 8 of the peptides in the library have the pattern $P^+H^0P^+H^0P^0H^0P^0$ but do not enhance infectivity (**Table S6**). **B** Creating peptides based on the identified patterns $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$ but mainly composed of amino acids prevalent in low active sequences result in weakly active peptides (**Table S6**). Fibril formation determined via TEM (**Figure S25**). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S23 A, B is provided.

python script. A sequence length of 7-mer was selected because all the infectivity enhancing peptides in the library contain more than 7 amino acids. The highest prevalence was found for the alternating amphiphilic pattern $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$ which can be found in 55% and 40 % of all high-active peptides, respectively. Interestingly, the importance for a strictly alternating, amphiphilic pattern is more pronounced for short 7-mer sequences than for 9-mer or 11-mer sequences as shown in direct comparison for sequence scrambled, *C*-terminal truncated versions of (I) and (F) substituted peptides (**Figure S26**).

6 of the 8 peptides in the library which share the pattern $P^+H^0P^+H^0P^0H^0P^0$ but do not enhance infectivity are composed of less hydrophobic amino acids alanine, which drastically reduces hydrophobicity of the peptides, and the ability to form fibrils and aggregate into positively charged clusters. Alanine is not found in high active peptides (**Figure S22**B1) and also reduces infectivity if applied in newly created peptides (**Figure S23**B**)**. The other two peptides in the library have the pattern $P^+H^0P^+H^0P^0H^0P^0$ but do not fulfill the key physicochemical properties for infectivity enhancement, that is for KFKFQFN aggregation into μm-sized aggregates and for KIKIQINMWQRGD fibril formation and a positive zeta-potential (**Figure S23**A**, Table S6)**.

By applying the amino acid analysis (**Figure S22**B) we find the amino acids W, M, N, I, C, Q, K, F and Y present in a higher frequency in highly active sequences “active amino acids” compared to amino acids S, G, R, H, T, V, P, A and L which are predominantly present in weakly active peptides “inactive amino acids” (**Figure S22**B). Interestingly, most of the “active amino acids” W, M, I, C, Q, F and Y are known to promote β-sheet amyloid fibril formation except of N, K (β-sheet breaking), whereas the “inactive amino acids” are either neutral (R, A, D, G) or β-sheet breaking (E, H, P, S) except of T, L, V (β-sheet promoting).[31,32]

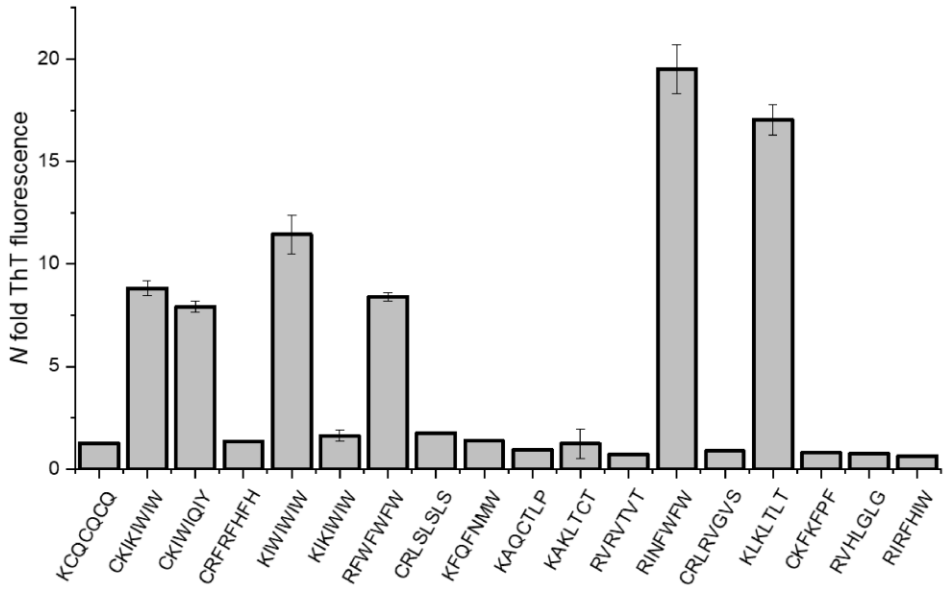


Figure S24 N-fold ThT fluorescence of newly created peptides relative to ThT fluorescence in PBS, excitation 440/10 nm, emission 488/10. The sequences are based on preserved sequence pattern $P^+H^0P^+H^0P^0H^0P^0$ and $P^+H^0P^0H^0P^0H^0P^0$ and listed according to their activity. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S24 is provided.

By applying amino acids prevalent in low active sequences show, sequences are created that also have low activity, despite using high active patterns. The only exception to this trend was observed for the peptide CRLSLSLS (medium active) (**Table S6**, **Figure S23**B**)**. Note, that the amino acids V, L, S, T, P are not strongly represented in the peptide library, which may distort the reliability for classification of these amino acids (**Figure S22**B1). All the non-active sequences created by applying the best performing pattern but mainly composed with amino acids prevalent in low active sequences, as expected, do not fulfill one or more key-physicochemical property (**Figure S23**B**)**. Remarkably, these non-active sequences show on average also lower sequence hydrophobicity compared to the positive examples (**Figure 6**B**, Table S6**). Thus, they demonstrate that not only the sequence pattern, but also physicochemical properties originating from the selected amino acids, play a role in creating active peptides.

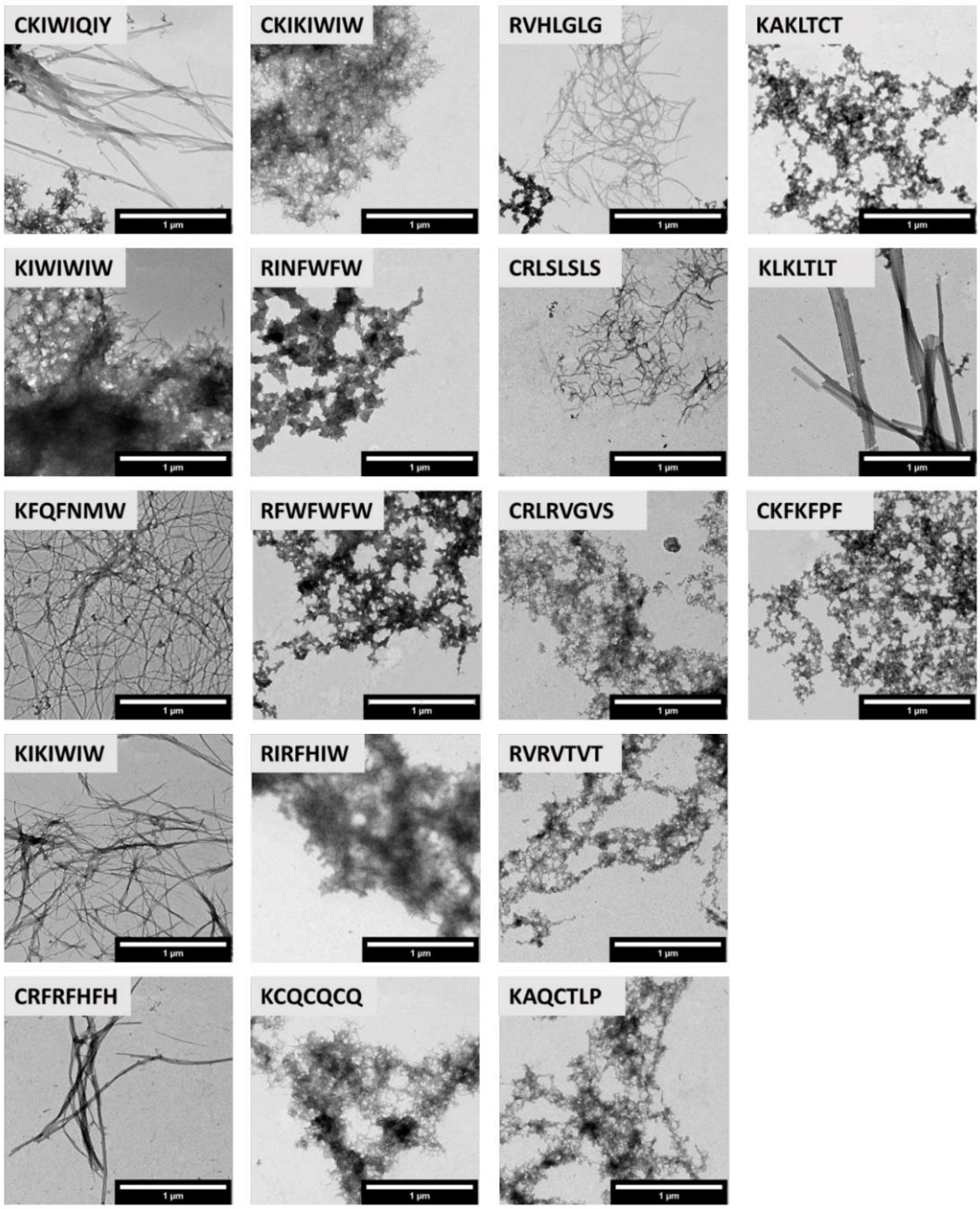


Figure S25 TEM micrographs for all newly created peptides based on best performing sequence pattern $P^{+}H^{0}P^{+}H^{0}P^{0}H^{0}P^{0}$ or $P^{+}H^{0}P^{0}H^{0}P^{0}H^{0}P^{0}$, scale bar 1 µm. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

Impact of sequence order is greater for shorter sequences compared to long sequences

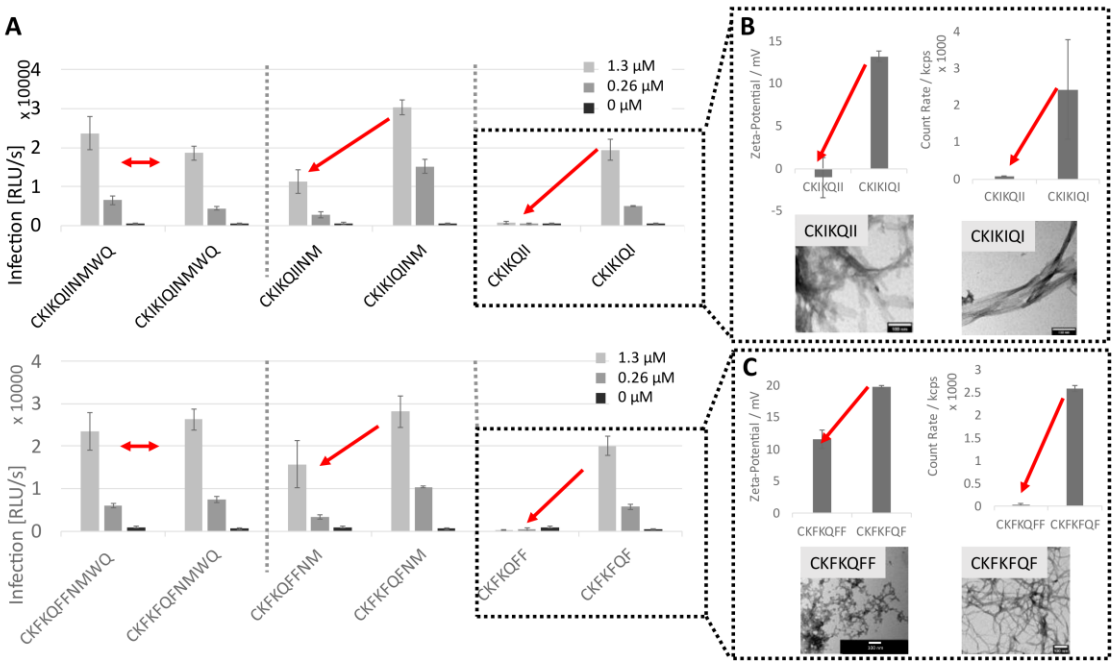


Figure S26 Impact of sequence order on infectivity enhancing properties. **A** Absolute infection of peptides at concentrations 1.3 µM, 0.26 µM and 0 µM of isoleucine (I) containing peptides which were scrambled (…QII…) → (…IQI…) and C-terminal truncated (Top) and phenylalanine (F) containing peptides which were scrambled (…QFF) → (…FQF…) and C-Terminal truncated (Bottom). With decreasing sequence length, the importance of alternating sequence order for infectivity enhancement is more pronounced (CKIKQII → CKIKIQI, CKFKQFF → CKFKFQF). **B, C:** Sequences with alternating sequence order have higher zeta-potential (left) and higher Count Rates (right), and form more ordered fibrils (bottom, TEM scale bar 100 nm). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S26 is provided.

8. Amyloid peptides from functional and pathogenic contexts

Table S1 Selected amyloid peptides from protein data bases and literature reports. Note, that the peptide KLVFFAK, derived from amyloid ab was reported for retroviral transduction enhancement, however with very low efficiencies (> 4 fold compared to virus only).[33] "+" indicates reported ThT fluorescence (ThT) or fibril formation, "0" indicates no data reported. All peptides were synthesized via Fmoc SPPS protocol and purchased from Phtd Peptides industrial Co. limited with purity of ≥ 95%.

sequence	ThT	fibril	pI	charge	$h_{fauchere}$	derived from	original context	ref.
IKFLSVN	+	+	9.7	0.94	0.70	peroxiredoxin III (Prx III)	antioxidant enzyme	34
VQIVYK	+	+	9.3	0.94	0.67	tau 623–628	microtubule-associated protein	35
IIKIIK	+	+	10.8	1.94	0.87	S. aureus phenol-soluble modulin PSMα3	bacterial amyloid biofilm stability	36
GQTVTK	+	+	9.7	0.94	0.09	nucleoprotein of SARS-CoV-2	RNA binding protein	37
DFNKF	0	+	6.3	-0.06	0.24	human calcitonin hormone	medullary thyroid carcinoma	38
RCELARTLKR	0	+	11.3	2.87	0.09	human lysozyme 5–14	antimicrobial enzyme	39
ISKLEYSNFSVRY	0	+	9.1	0.94	0.39	Acyl phosphatase 86-98	hydrolase of acid anhydrides	39
GYVIIK	+	+	9.3	0.94	0.80	Serum amyloid P 213–218	associated with multiple sclerosis	35
KLVFFAK	+	+	10.6	2.00	0.69	Amyloid Aβ 16-22	Alzheimer's disease	33
RATVYV	+	+	9.3	0.94	0.49	VQIVYK (tau 623–628)	microtubule-associated protein	Waltz
RQGNINIVA	0	+	10.5	0.94	0.30	Insects' chorion S36	protective function in the eggshell	40
KYFIL	0	+	9.3	0.94	1.05	-	-	41
NKGAII	+	+	9.7	0.94	0.39	Alzheimer's Aβ 27-32	Alzheimer's disease	42
IIIQGK	0	+	9.7	0.94	0.69	-	-	43
LASLIYR	0	+	9.3	0.94	0.77	Mammalia melanocyte protein PMEL 610-616	control of melanin assembly	40
IIIK	0	+	9.7	0.94	1.10	-	-	44
MKVIFLKDVKG	0	+	10.5	1.94	0.47	N-terminal domain of ribosomal protein L9 (Beta 1)	protein synthesis	39
LANWMCLAKW	0	+	8.5	0.87	0.97	human lysozyme 25-34	antimicrobial enzyme	39
VHDCVNITIK	0	+	7.3	-0.02	0.56	human prion protein 176-185	High expression in nervous system, misfolding associated with neurodegenerative diseases	39
IFQINS	0	+	6.1	-0.06	0.755	human lysozyme	antimicrobial enzyme	45
NSGAITIG	0	+	13.8	1	0.44	Adenovirus Fiber Shaft type 2, 385 -392	receptor-binding head domain	46
GAITIG	0	+	13.8	1	0.69	Adenovirus Fiber Shaft type 2, 387 -392	receptor-binding head domain	46
VLYVGSKT	0	+	9.7	1	0.54	α-Synuclein 37–44	Parkinson's disease	47
GNNQQNY	+	+	5.9	0	-0.18	Yeast Protein Sup35 7-13	Terminate cellular translation	48

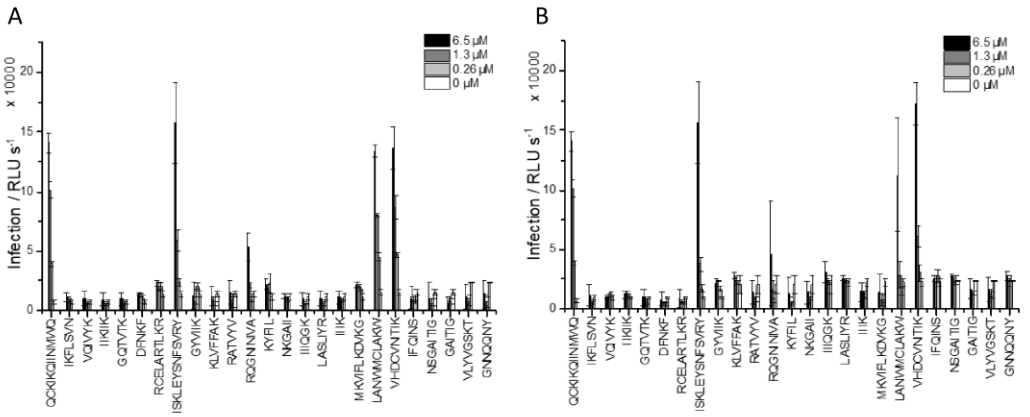


Figure S27 Overview of infectivity enhancing effect of amyloidal peptides derived from functional or pathogenic contexts. Absolute infection rates in RLU/s determined via β-galactosidase based luminescence assay for peptides at different concentration (6.5 µM, 1.3 µM, 0.26 µM) to TZM-bl cells and infected with HIV-1. 0 µM represents the virus only infectivity without the addition of peptide. **A** displays the infectivity rates for peptides prepared by diluting them from DMSO stock solution to PBS (1 mg/mL) and incubation for 1d at rt prior to dilution according to the final concentration (6.5 µM, 1.3 µM, 0.26 µM). **B** displays the infectivity rates for peptides prepared according to literature described procedure (**SI Section 8**). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S27 is provided.

Preparation of peptide fibrils described in literature

Unless stated otherwise all peptides were incubated at rt for 1 w before characterization or evaluation in an infectivity assay. IKFLSVN was dissolved in 50 mM Tris buffer, pH 7.5 containing 100 mM NaCl to a final concentration of 14 mg/mL.[34] VQIVYK was dissolved in 50 mM PB, pH 7.4 to a final concentration of 1 mM.[49] IIKIIK was prepared as a 20 mM DMSO stock before diluting to 6 mM in ultrapure water.[36] GQTVTK was prepared as a 1582 µM solution using 20 mM Tris, 300mM NaCl, pH 8.0 buffer and incubated at 37°C for 1 d under constant agitation.[37] DFNKF was first dissolved as a 100 mM stock solution in DMSO and then diluted to 2 mM in 0.02 M NaCl, 0.01 M Tris, pH 7.2 to induce fibril formation.[50] RCELARTLKR was prepared as a 10 mg/mL in 50 mM PB, pH 7.2.[39] ISKLEYSNFSVRY was incubated in 25% (v/v) TFE, 50 mM in acetate buffer, pH 5.5 at a concentration 0.4 mg/mL.[51] GYVIIK was first prepared as a 4 mg/mL DMSO stock solution and then diluted in 100 mM MES buffer, pH 7.4 to a concentration of 1.3 mg/mL.[52] KLVFFAK was prepared as a 4 mg/mL solution in 50 mM PBS with 1 M NaCl, pH 2. The solution was constantly agitated for 3d at 37 °C.[33] RATVYV was dissolved in ultrapure water to a final concentration of 1 mM and was incubated for 4 weeks at 25°C.[53] RQGNINIVA was dissolved in water with adjusted pH 7.5 and incubated at a concentration of 0.5 mM for 4 weeks at rt.[2] KYFIL was prepared in PBS, pH 7.4 at a concentration of 3 % wt.[41] NKGAII was dissolved in 20 mM ammonium acetate buffer, pH 7.0 to the final concentration of 250 µM.[42] IIIQGK was dissolved in 25 mM Hepes buffer, pH 7.4 and incubated at a concentration of 8 mM.[43] LASLIYR was incubated in water, pH 7.5 at a concentration of 0.5 mM at rt for 4 weeks.[2] IIIK was dissolved in water to a concentration of 4 mM and incubated for 1 week at rt.[44] MKVIFLKDVKG was prepared as 1 mM solution in water at pH 5.4.[54] LANWMCLAKW and VHDCVNITIK were diluted to a concentration of 1 mM in 50 mM PB, pH 7.2.[51] IFQINS was prepared at 1.5 mM in water, pH 6.5.[45] NSGAITIG was incubated at 10 mg/mL concentration in water for 6 weeks, GAITIG was incubated at 20 mg/mL concentration in water for 6 weeks.[46] VLYVGSKT[47] and GNNQQNY[48] were dissolved at 20 mg/mL concentration in water.

Note, that for 20 of the 24 selected peptides fibril formation under the literature described preparation protocol was observed via TEM (**Figure S37**). Noteworthy, 12 of the selected peptides formed also fibrils upon dilution from DMSO stock solutions in PBS to 1 mg/mL (**Figure S36**). For the infectivity assay, the peptide fibrils were diluted in phosphate-buffered solution regardless of the initial preparation protocol. No significant difference in activity was observed for both preparation methods (**Figure S27**).

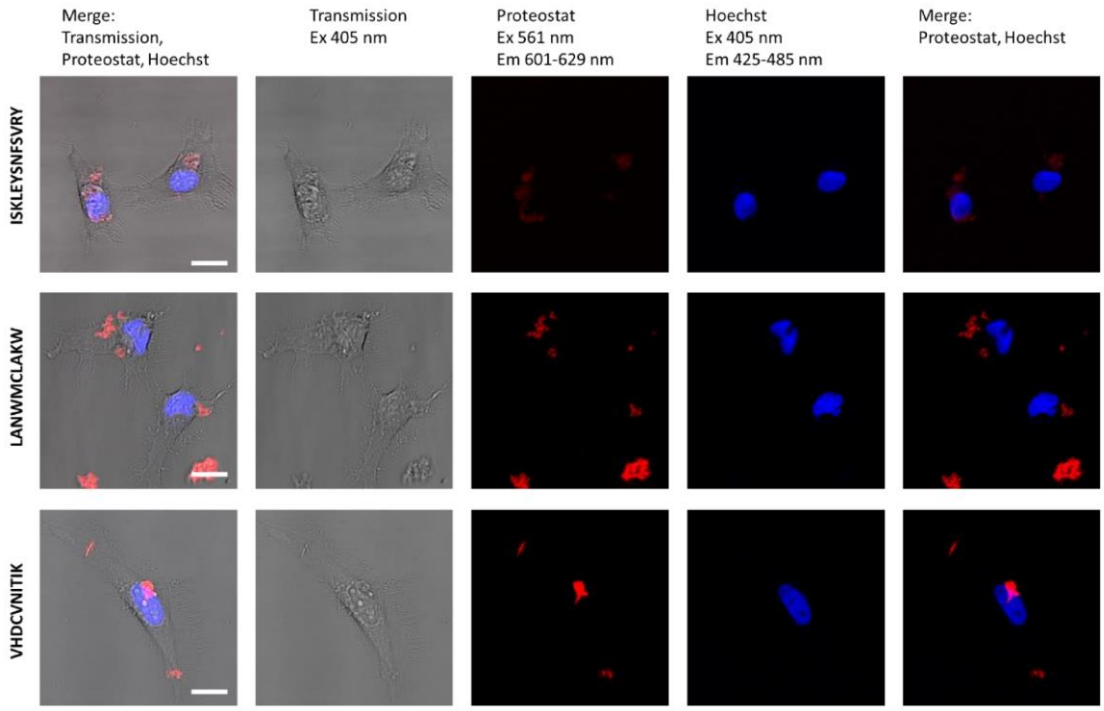


Figure S28 Fibril-cell colocalization shown by confocal microscopy. Acyl phosphatase derived peptide ISKLEYSNFSVRY, human lysozyme derived peptide LANWMCLAKW and human prion protein derived peptide VHDCVNITIK were added to HeLa cells. To this end, 20 µg/mL preassembled peptides (stained with Proteostat, red) were incubated 30 min with HeLa cells (nucleus stained with Hoechst, blue), washed three times with PBS and incubated for further 30 min before imaging, scale bar 20 µm. Note, that ISKLEYSNFSVRY cannot be stained as strongly fluorescent as the other two peptides (**Figure S29**A) therefore it also appears weaker fluorescent in the confocal microscopy. Confocal images have been prepared in biological triplicates and recorded at least at three different sites.

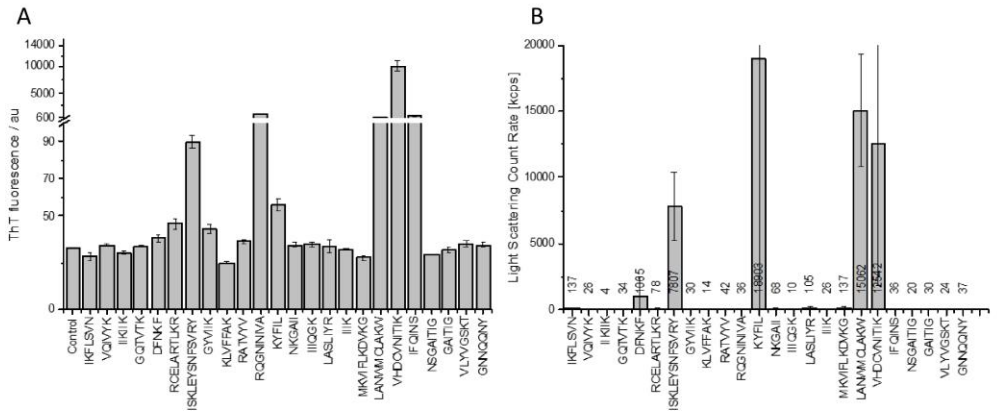


Figure S29 Thioflavin-T (ThT) fluorescence and Count Rate of scattered light of literature known peptides prepared by diluting from 10 mg/mL DMSO stock solution in PBS (1 mg/mL). **A** The ThT-Assay was conducted by adding ThT (50 µM in PBS) to the peptide in a 1:5 (peptide: ThT, v:v) ratio. The control is the background fluorescence of ThT without peptide (50 µM ThT added to PBS with 10% DMSO in 1:5 (PBS: ThT, v:v) ratio). **B** The derived count rate of scattered light was determined at 633 nm, 173° angle for peptides diluted from PBS (1 mg/mL) in KCl (1mM) to a final concentration of 0.1 mg/mL. Error bars indicate standard deviation from triplicate measurements. Source data for Figure S29 is provided.

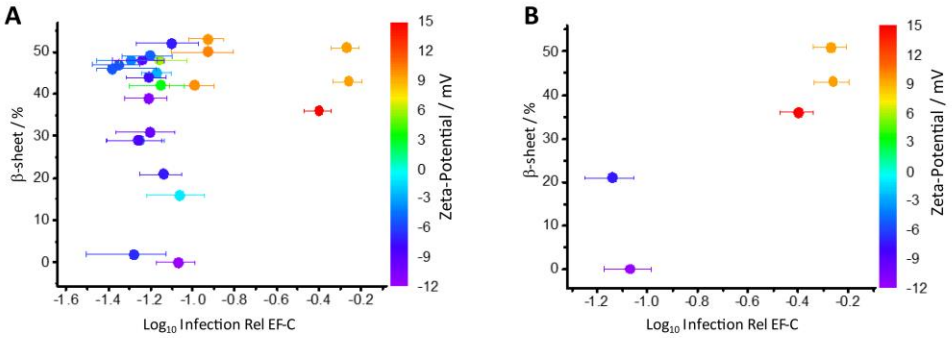


Figure S30 β-sheet content plotted against Log Infection Rel EF-C of 24 literature reported peptides. Color scheme represents zeta-potential. **A** All peptides regardless of the count rate are shown. **B** Peptides above a critical count rate of 100 kcps is shown. All tested peptides with a high β-sheet content and a positive zeta-potential are enhancers of viral transduction (top right corner). Error bars indicate standard deviation from triplicate measurements. Source data for Figure S30 is provided.

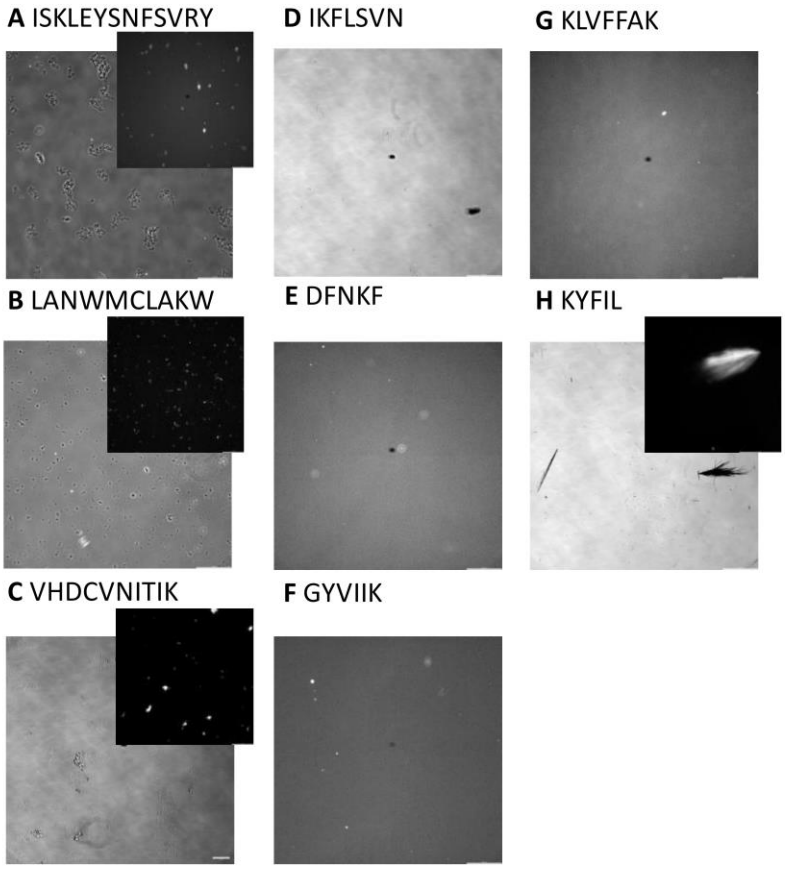


Figure S31 Brightfield and fluorescence microscopy images of selected peptides: **A-C** ISKLEYSNFSVRY, LANWMCLAKW and VHDCVNITIK as infectivity enhancing peptides show fluorescence upon addition of ThT (small image details) and microscopic aggregation (> 10 µm) as observed via brightfield microscopy. **D-G** IKFLSVN, DFNKF, GYVIIK, KLVFFAK do not enhance infectivity, are not ThT active (**Figure S29**) and do not aggregate microscopically in agreement with scattering results. **H** KYFIL is forming large fibrous aggregates, which are ThT active. This peptide is not enhancing infectivity due to its negative zeta-potential (scale bar 200 µm). Microscopy measurements were conducted once with at least three microscopy images recorded for each peptide sample.

9. Peptide library: tables

Table S2 EF-C derived peptides sorted according to their structural relationship.

N-Truncations	C-Truncations	Scrambling	Alternating	Substitutions
RGDSKIKIQIC	QCKFKFQFNMWQ	KAKAQANMWQ	QCKIKQIINM	QCKIKIQINMWQ
RGDKIKIQIC	QCKFKFQFNM	NMWQKAKAQA	QCKIKIQINM	QCKFKFQFNMWQ
KIKIQIC	QCKFKFQF			QCKAKAQANMWQ
		KFKFQFNMWQ	CKIKQIINM	
ILKNLSRSRKIKIQINMWQ	CKFKFQFNMWQ	NMWQKFKFQF	CKIKIQINM	QCKIKIQINM
RGDKIKIQINMWQ	CKFKFQFNMW			QCKFKFQFNM
QCKIKIQINMWQ	CKFKFQFNM	KIKQIINMWQ	KIKQIINM	QCKAKAQANM
CKIKIQINMWQ	CKFKFQF	NMWQKIKQII	KIKIQINM	
MKIKIQINMWQ				QCKIKIQI
KIKIQINMWQ	KFKFQFNMWQ	KIKIQINMWQ	QCKIKQII	QCKFKFQF
KIKIKIQINMWQ	KFKFQFNMW	NMWQKIKIQI	QCKIKIQI	QCKAKAQA
	KFKFQFNM			
CKIKIQINMW	KFKFQFN	KIKIQIC	CKIKQII	CKIKIQINMWQ
KIKIQINMW	KFKFQF	CKIKIQI	CKIKIQI	CKAKAQANMWQ
				CKFKFQFNMWQ
RGDCKIKIQINM	KFKFQFW	KIKIQINMWA	KIKIQI	CEFEFQFNMWQ
QCKIKIQINM	KFKFQF	KIKIQINMAQ	KIKQII	CEIEIQINMWQ
CKIKIQINM	KFKFQFFF	KIKIQINAWQ	KIKIQII	MKIKIQINMWQ
KIKIQINM	KFKFQFF	KIKIQIAMWQ	KIKIKIQI	
		KIKIQINMWQ		CKIKIQINMW
KIKIKIQI	QCKFKQFFNMWQ	KIKIQIMWNQ	KIKQIIN	CKFKFQFNMW
RGDKIKIQI	QCKFKQFFNM	KIKIQIMQWN	KIKIQIN	
QCKIKIQI	QCKFKQFF			CKIKIQINM
CKIKIQI		KIKIQIRGD	NMWQKIKQII	CKFKFQFNM
KIKIQI	CKFKQFFNMWQ	RGDKIKIQI	NMWQKIKIQI	CEIEIQINM
	CKFKQFFNM			CSISIQINM
QCKIKQIINMWQ	CKFKQFF	CKIKIQIRGD	CKIKQIIRGD	MKIKIQINM
CKIKQIINMWQ		RGDKIKIQIC	CKIKIQIRGD	CKAKAQANM
KIKQIINMWQ	QCKAKAQANMWQ			
	QCKAKAQANM	RGDKIKIQINMC	KAKQAANMWQ	CKIKIQI
QCKIKQIINM	QCKAKAQA	RGDCKIKIQINM	KAKAQANMWQ	CKWKWQW
CKIKQIINM				CKYKYQY
KIKQIINM	CKAKAQANMWQ	RGDKIKIQINMWQ	QCKFKQFFNMWQ	CRFRFQF
	CKAKAQANM	KIKIQINMWQRGD	QCKFKFQFNMWQ	CRFRVPF
QCKIKQII	CKAKAQA			CHLHLQL
CKIKQII			CKFKQFFNMWQ	CKFKFQF
KIKQII	KAKAQANMWQ		CKFKFQFNMWQ	CQFQFQF
	KAKAQANM			CSISIQI
QCKAKAQANMWQ	KAKAQA		KFKQFFINMWQ	MKFKFQF
CKAKAQANMWQ			KFKQFFNMWQ	CEIEIQI
KAKAQANMWQ	QCKIKIQINMWQ		KFKFQFNMWQ	CKAKAQA
	QCKIKIQINM			CEFEFQF
QCKAKAQANM	QCKIKIQI		QCKFKQFFNM	
CKAKAQANM			QCKFKFQFNM	KIKIQINMWQ
KAKAQANM	MKIKIQINMWQ			KAKAQANMWQ
	CKIKIQINMWQ		CKFKQFFNM	KFKFQFNMWQ
KAKAKAQA	CKIKIQINMW		CKFKFQFNM	EIEIQINMWQ
QCKAKAQA	CKIKIQINM			
CKAKAQA	CKIKIQI		QCKFKQFF	KIKIQINM
KAKAQA	CKIKIQIII		QCKFKFQF	KFKFQFNM
				KAKAQANM
QCKFKFQFNMWQ	KIKIQINMWQRGD		CKFKQFF	EIEIQINM
CKFKFQFNMWQ	KIKIQINMWQ		CKFKFQF	KIEIQINM
KFKFQFNMWQ	KIKIQINMWWWQ			EIKIQINM
	KIKIQINMWWQ		KFKFQFF	
QCKFKFQFNM	KIKIQINMW		KFKFQF	KIKIQI
CKFKFQFNM	KIKIQINM		KFKFQFFF	KAKAQA
KFKFQFNM	KIKIQIN			KFKFQF
	KIKIQI			KYKYQY
RGDCKFKFQF	KIKIQII			KWKWQW
QCKFKFQF	KIKIQIII			EFEFQF
CKFKFQF	KIKIQIW			
KFKFKFQF				KFKFQFW
KFKFQF	RGDCKIKIQINM			KIKIQIW

	RGDKIKIQINMC			
QCKFKQFFNMWQ	RGDKIKIQINM			KIKIQIC
CKFKQFFNMWQ	RGDKIKIQINMWQ			HIHIQIC
KFKQFFNMWQ				RLRLTLC
	RGDKIKIQI			
QCKFKQFFNM	RGDKIKIQIC			KIKIKIQI
CKFKQFFNM				KAKAKAQA
	KIKIKIKIWWWW			KLKLKLQL
QCKFKQFF	KIKIKIKIWW			KFKFKFQF
CKFKQFF	KIKIKIKI			KVKVKVQV
KLKLQLL	QCKIKQIINMWQ			KIKIKIKI
KLKQLL	QCKIKQIINM			EIEIEIEI
	QCKIKQII			KFKFEFEF
KIKIKIKIYYYY				
HHHHKIKIKIKIYYYY	CKIKQIINMWQ			
	CKIKQIINM			QCKIKQIINMWQ
CKFKFQFW	CKIKQII			QCKFKQFFNMWQ
KFKFQFW				
	KIKQIINMWQ			QCKIKQIINM
	KIKQIINM			QCKFKQFFNM
	KIKQIIN			QCKAKAQANM
	KIKQII			
				QCKIKQII
	CEIEIQINMWQ			QCKFKQFF
	CEIEIQINM			
	CEIEIQI			CKIKQIINMWQ
				CKFKQFFNMWQ
	EIEIQINMWQ			
	EIEIQINM			CKIKQIINM
				CKFKQFFNM
	RGDKIKIQINM			
	RGDKIKIQINMC			CKIKQII
				CKFKQFF
	KFKFQFFF			
	KFKFQFF			KIKQIINMWQ
				KFKQFFNMWQ
	KIKIQIII			KAKQAANMWQ
	KIKIQII			KFKQFFINMWQ
	KVKVQVVV			KIKQII
	KVKVQVV			KLKQLL
				KIKIQII
				KFKFQFF
				HLHLPLL
				KLKLQLL
				KVKVQVV
				KIKIQIII
			Random Sequences	KFKFQFFF
			HGDKCHGDKC	KVKVQVVV
			KKFQKKFQ	
			PPFHPPPFHP	RGECKFKFQF
			MDQMDQMDQMDQMDQ	RGDCKFKFQF
			FDPFDPFDP	
			TKTLTKTL	RGEKIKIQINM
			FKFDKFKFDK	RGDKIKIQINM
			KVKGVGK	
			SISISRRI	HHHHKIKIKIKIYYYY
			HRRHFRHKITKKK	HHHHKAKAKAKAYYYY
			KNERIKNERI	HHHHKIKIQINMYYYY
			KIRGKFEKED	
			HGDKCHGDKC	HHHHKIKIKIKIWWWW
			KKFQKKFQ	HHHHKAKAKAKAWWWW

Table S3 Overview of all peptides investigated in this study. Sequences indicated with references were reported previously by us (Schilling et al.[29] and Sieste et al.[55]) or are derived from literature reports. Sequences indicated with * have been synthesized by us (see **Figure S35** for LC-MS of peptides, which were not reported before by us) other sequences have been commercially obtained with purity of ≥ 95 % from Phtd Peptides industrial Co. limited.

Sequence	Report	Synthesis
CKIKIQIII	[29]	*
CEIEIQI	[55]	
CKIKQIINM	[29]	
CKFKFQFNMWQ	[55]	
CKFKFQFNM	[55]	
CKFKFQF	[29]	
KFKFQFNMW		*
KFKFQFNM	[29]	
KFKFQFN		*
CKAKAQANMWQ	[55]	
CKAKAQANM	[55]	
QCKFKQFFNMWQ		
QCKFKQFFNM		
QCKFKQFF		
CKFKQFFNMWQ		
CKFKQFFNM		
CKFKQFF		
QCKIKIQINM		
QCKAKAQANMWQ		
QCKAKAQANM		
QCKAKAQA		
QCKIKIQI		
QCKIKQIINM		
QCKIKQII		
QCKFKFQFNMWQ		
QCKFKFQFNM		
QCKFKFQF		
CRFRFQF		
HIHIQIC		
RLRLTLC		
HLHLPLL		
RGECKFKFQF		
RGEKIKIQINM		
KYKGAIIGNIK		
HGDKCHGDKC		
RPRGLLLGNLR		
KKFQKKFQ		
PPFHPPPFHP		
MDQMDQMDQMDQMDQ		
FDPFDPFDP		
TKTLTKTL		
FKFDKFKFDK		
KVKGVGK		
SISISRRI		
HRRHFRHKITKKK		
KNERIKNERI		
KIRGKFEKED		
CKFQC		
MKFM		
CKFC		
RGDKIRGDKI		
KNDKND		
HGEHGE		
HGEHGEHGE		
CRFRVPF		
CHLHLQL		
CETMYDKILKNLSRSR		
MKFKFQF	[55]	
QCKIKQIINMWQ	[55]	
QCKIKIQINMWQ	[55]	
KIKIQINMWQ	[29]	

NMWQKAKAQA		*
NMWQKFKFQF		*
KVKVKVQV		*
KIKQIINMWQ	55	
KAKAQANMWQ	55	
KFKFQFNMWQ	55	
KFKQFFNMWQ		*
KAKQAANMWQ		*
NMWQKVGTPL		*
NMWQKIKQII		*
NMWQKIKIQI		*
KIKQIINM	55	
KIKQIIN		*
KIKIQINMW		*
KIKIQINM	29	
KIKIQIN		*
KAKAQA	55	
KIKIQI	29	
KFKFQF	29	
KIKIQINMWA		*
EIEIEIEI		*
KIKIQINMAQ		*
KIKIQINAWQ		*
KIKIQIAMWQ		*
KIKIKIKIYYYY		*
HHHHEIEIEIEI		*
KYKYQY		*
EFEFQF		*
CEFEFQF		
CKFKFQFNMW		*
CKYKYQY		*
KWKWQW		*
CKWKWQW		*
KLKQLL		*
CKIKIQINMWQ	29	
RGDKIKIQI		*
CKIKIQINM	29	
CKIKIQI	29	
CKIKQII	29	
CKIKQIINMWQ	29	
KIKIQIRGD	29	*
RGDKIKIQIC	29	
RGDKIKIQINM	29	
RGDKIKIQINMWQ	29	*
KFKFEFEF		*
KIKQII	55	
KIEIQINM		*
CKIKIQIRGD	29	*
CKIKQIIRGD		
RGDKIKIQINMC	29	*
CQFQFQF		
KIKIQII		*
KFKFQFFF	29	*
KIKIKIQI	29	*
KAKAKAQA		*
KLKLKLQL	29	*
KFKFKFQF	29	*
KVKVQVV		*
KLKLQLL		*
KFKFQFF		*
KIKIQIII	29	*
KVKVQVVV		*
HHHHKAKAKAKAYYYY		*
KIKIQIC		
RGDSKIKIQIC		
RGDSGGGGGGKIKIQIC		
ILKNLSRSRKIKIQIC		
KIKIKIKIWWWW		*

KIKIKIKI		*
CEFEFQFNMWQ		
CSISIQI	55	
CEIEIQINMWQ	55	
CEIEIQINM	55	
CSISIQINM	55	
KAKAQANM	55	
CKAKAQA	55	
HHHHKIKIQINMYYYY		*
KIKIKIKIWW	29	*
HHHHKIKIKIKIWWWW	29	*
MKIKIQINM		
RGDCKIKIQINM		*
MKIKIQINMWQ		*
HHHHKIKIKIKIYYYY		*
RGDCKFKFQF		
KIKIQIW		*
KFKFQFW		*
CKFKFQFW		
EIKIQINM		*
IKVAVKIKIQINM		*
HHHHKAKAKAKAWWWW		*
CKIW		*
CKIKIQINMW		*
ILKNLSRSRKIKIQINMWQ		
EIEIQINM	29	
KIKIQINMWWQ		*
KIKIQINMWWWQ		*
KFKQFFINMWQ		*
KIKIQIMWNQ		*
KIKIQIMQWN		*
KIKIQINMWQRGD		*
EIEIQINMWQ		
KIKIKIQINMWQ		*
IKFLSVN	34	
VQIVYK	35	
IIKIIK	36	
GQTVTK	37	
DFNKF	38	
RCELARTLKR	39	
ISKLEYSNFSVRY	39	
GYVIIK	35	
KLVFFAK	33	
RATVYV	Waltz	
RQGNINIVA	40	
KYFIL	41	
NKGAII	42	
IIIQGK	43	
LASLIYR	40	
IIIK	44	
MKVIFLKDVKG	39	
LANWMCLAKW	39	
VHDCVNITIK	39	
IFQINS	45	
NSGAITIG	46	
GAITIG	46	
VLYVGSKT	47	
GNNQQNY	48	
CKIWIQIY		
KIWIWIW		
KFQFNMW		
KIKIWIW		
CKIKIWIW		
RVHLGLG		
CRLSLSLS		
CRLRVGVS		
RVRVTVT		
RINFWFW		

RFWFWFW		
RIRFHIW		
CRFRFHFH		
KAQCTLP		
KCQCQCQ		
KAKLTCT		
KLKLTLT		
CKFKFPF		

Bioinformatic descriptors

Table S4 Summary of bioinformatic descriptors calculated for the peptide library: Net charge (NC), hydrophobicity according to Fauchère scale ($H_{fauchere}$) and Kyte-Doolittle scale ($H_{Kyte\text{-}Doolittle}$), isoelectric point (pI), aliphatic index (AI), hydrophobic moment index (HMI), boman index (BI), instability index (II), graph shape index (GSI), upsilon steric parameter (USP), polarizability (P), normalized van der Waals Volume (Norm vdW Vol.).

Sequence	NC	$H_{fauchere}$	$H_{kyte\text{-}Doolittle}$	pI	AI	HMI	BI	II	GSI	USP	P	Norm. vdW Vol.
CKIKIQIII	1.87	0.93	1.52	9.68	216.67	0.14	-1.03	-19.41	3.12	0.86	0.19	3.99
CEIEIQI	-2.13	0.78	0.79	3.61	167.14	0.20	0.45	126.37	2.72	0.82	0.17	3.71
CKIKQIINM	1.87	0.60	0.34	9.68	130.00	0.51	0.54	16.33	2.63	0.81	0.18	3.92
CKFKFQFNMWQ	1.87	0.67	-0.58	9.68	0.00	0.18	1.27	-11.43	1.95	0.63	0.20	4.08
CKFKFQFNM	1.87	0.59	-0.22	9.68	0.00	0.12	1.19	-16.19	2.21	0.70	0.22	4.55
CKFKFQF	1.87	0.67	-0.06	9.68	0.00	0.21	0.92	-23.67	2.28	0.68	0.23	4.80
KFKFQFNMW	1.94	0.67	-0.60	10.81	0.00	0.24	1.07	-16.19	2.01	0.63	0.20	4.28
KFKFQFNM	1.94	0.48	-0.56	10.81	0.00	0.17	1.50	-19.46	2.26	0.71	0.23	4.82
KFKFQFN	1.94	0.37	-0.91	10.81	0.00	0.25	2.05	-23.67	2.25	0.70	0.23	4.87
CKAKAQANMWQ	1.87	0.27	-0.85	9.68	27.27	0.15	1.58	9.09	1.50	0.59	0.13	2.50
CKAKAQANM	1.87	0.10	-0.56	9.68	33.33	0.07	1.58	8.89	1.66	0.64	0.14	2.62
QCKFKQFFNMWQ	1.87	0.60	-0.83	9.68	0.00	0.46	1.62	3.81	1.92	0.64	0.19	4.07
QCKFKQFFNM	1.87	0.51	-0.55	9.68	0.00	0.40	1.63	2.57	2.14	0.70	0.22	4.49
QCKFKQFF	1.87	0.56	-0.49	9.68	0.00	0.42	1.50	0.71	2.19	0.68	0.22	4.69
CKFKQFFNMWQ	1.87	0.67	-0.58	9.68	0.00	0.46	1.27	10.10	1.95	0.63	0.20	4.08
CKFKQFFNM	1.87	0.59	-0.22	9.68	0.00	0.46	1.19	10.12	2.21	0.70	0.22	4.55
CKFKQFF	1.87	0.67	-0.06	9.68	0.00	0.54	0.92	10.16	2.28	0.68	0.23	4.80
QCKIKIQINM	1.87	0.52	-0.04	9.68	117.00	0.10	1.04	-24.01	2.52	0.79	0.18	3.93
QCKAKAQANMWQ	1.87	0.23	-1.08	9.68	25.00	0.15	1.91	2.88	1.50	0.59	0.13	2.63
QCKAKAQANM	1.87	0.07	-0.85	9.68	30.00	0.08	1.98	1.46	1.65	0.64	0.14	2.76
QCKAKAQA	1.87	0.01	-0.86	9.68	37.50	0.04	1.93	-0.68	1.56	0.61	0.13	2.52
QCKIKIQI	1.87	0.57	0.15	9.68	146.25	0.09	0.77	-32.51	2.66	0.80	0.19	3.98
QCKIKQIINM	1.87	0.52	-0.04	9.68	117.00	0.44	1.04	8.16	2.52	0.79	0.18	3.93
QCKIKQII	1.87	0.57	0.15	9.68	146.25	0.48	0.77	7.70	2.66	0.80	0.19	3.98
QCKFKFQFNMWQ	1.87	0.60	-0.83	9.68	0.00	0.18	1.62	-15.93	1.92	0.64	0.19	4.07
QCKFKFQFNM	1.87	0.51	-0.55	9.68	0.00	0.09	1.63	-21.11	2.14	0.70	0.22	4.49
QCKFKFQF	1.87	0.56	-0.49	9.68	0.00	0.07	1.50	-28.89	2.19	0.68	0.22	4.69
CRFRFQF	1.87	0.67	-0.23	10.53	0.00	0.24	3.59	-2.20	2.40	0.68	0.25	5.19
HIHIQIC	0.09	1.00	0.87	7.52	167.14	0.21	-0.17	151.74	3.13	0.82	0.19	3.96
RLRLTLC	1.87	0.70	0.60	10.53	167.14	0.23	2.34	36.09	2.46	0.77	0.20	4.18
HLHLPLL	0.16	1.11	1.03	7.71	222.86	0.13	-1.48	36.09	2.72	0.76	0.17	3.23
RGECKFKFQF	1.87	0.31	-0.88	9.67	0.00	0.33	2.72	14.34	1.98	0.61	0.21	4.35
RGEKIKIQINM	1.94	0.20	-0.71	10.79	106.36	0.18	2.45	-28.64	2.34	0.73	0.18	3.89
KYKGAIIGNIK	2.94	0.28	-0.18	10.64	115.45	0.24	0.45	18.16	2.19	0.64	0.15	3.26
HGDKCHGDKC	0.02	-0.02	-1.70	7.42	0.00	0.16	3.34	3.88	1.65	0.55	0.14	2.93

RPRGLLLGNLR	2.94	0.35	-0.38	12.80	141.82	0.28	2.71	21.82	1.97	0.61	0.16	3.15
KKFQKKFQ	3.94	-0.10	-2.13	11.28	0.00	0.71	3.42	8.75	2.07	0.69	0.23	4.85
PPFHPPPFHP	0.16	0.82	-1.04	7.71	0.00	0.08	0.34	99.54	2.79	0.28	0.10	0.48
MDQMDQMDQMDQMDQ	-5.06	0.08	-1.70	3.14	0.00	0.29	3.97	9.33	1.84	0.74	0.17	3.72
FDPFDPFDP	-3.06	0.58	-0.77	3.34	0.00	0.34	1.91	92.82	2.40	0.49	0.13	1.98
TKTLTKTL	1.94	0.31	-0.38	10.81	97.50	0.59	1.44	8.75	2.64	0.67	0.16	3.49
FKFDKFKFDK	1.94	0.17	-1.14	10.33	0.00	0.17	2.77	-13.36	2.25	0.70	0.22	4.82
KVKGVGK	2.94	-0.08	-0.59	11.10	82.86	0.52	0.96	-44.06	1.86	0.51	0.13	2.13
SISISRRI	1.94	0.41	0.26	12.50	146.25	0.60	3.16	104.43	2.65	0.75	0.17	3.63
HRRHFRHKITKKK	7.27	-0.21	-2.47	12.83	30.00	0.38	5.82	86.63	2.59	0.70	0.23	4.92
KNERIKNERI	1.94	-0.29	-2.18	10.77	78.00	0.19	5.80	0.51	2.32	0.76	0.20	4.33
KIRGKFEKED	0.94	-0.24	-1.98	9.44	39.00	0.25	4.51	2.79	1.99	0.66	0.18	4.07
CKFQC	0.79	0.73	0.08	8.23	0.00	0.25	1.11	-7.08	1.99	0.66	0.19	3.89
MKFM	0.94	0.82	0.68	9.70	0.00	0.61	-0.53	7.50	2.38	0.74	0.24	4.88
CKFC	0.79	0.97	0.98	8.23	0.00	0.54	0.00	7.50	2.09	0.66	0.19	3.88
RGDKIRGDKI	1.94	-0.19	-1.56	10.77	78.00	0.23	4.67	-41.94	2.00	0.63	0.16	3.54
KNDKND	-0.06	-0.79	-3.63	6.45	0.00	0.20	6.97	-5.82	1.70	0.73	0.15	3.50
HGEHGE	-1.83	-0.17	-2.37	5.33	0.00	0.22	3.51	-63.93	1.52	0.46	0.13	2.81
HGEHGEHGE	-2.72	-0.17	-2.37	5.32	0.00	0.17	3.51	-67.57	1.52	0.46	0.13	2.81
CRFRVPF	1.87	0.72	0.24	10.53	41.43	0.26	2.65	63.60	2.67	0.59	0.20	3.44
CHLHLQL	0.09	0.95	0.57	7.52	167.14	0.17	-0.17	101.71	2.44	0.81	0.19	3.96
CETMYDKILKNLSRSR	1.87	0.19	-0.90	9.40	73.13	0.78	3.39	86.63	2.21	0.72	0.18	3.90
MKFKFQF	1.94	0.63	-0.14	10.81	0.00	0.20	0.76	-23.67	2.36	0.70	0.24	5.08
QCKIKQIINMWQ	1.87	0.60	-0.40	9.68	97.50	0.50	1.14	8.47	2.23	0.72	0.17	3.60
QCKIKIQINMWQ	1.87	0.60	-0.40	9.68	97.50	0.19	1.14	-18.34	2.23	0.72	0.17	3.60
KIKIQINMWQ	1.94	0.59	-0.38	10.81	117.00	0.24	0.94	-16.47	2.34	0.73	0.17	3.68
NMWQKAKAQA	1.94	0.14	-1.19	10.81	30.00	0.32	1.87	9.00	1.47	0.58	0.13	2.51
NMWQKFKFQF	1.94	0.58	-0.89	10.81	0.00	0.37	1.52	-13.57	1.97	0.64	0.20	4.25
KVKVKVQV	2.94	0.21	0.20	11.10	145.00	0.10	0.75	-39.71	2.74	0.72	0.17	2.43
KIKQIINMWQ	1.94	0.59	-0.38	10.81	117.00	0.58	0.94	15.70	2.34	0.73	0.17	3.68
KAKAQANMWQ	1.94	0.14	-1.19	10.81	30.00	0.20	1.87	9.00	1.47	0.58	0.13	2.51
KFKFQFNMWQ	1.94	0.58	-0.89	10.81	0.00	0.22	1.52	-13.57	1.97	0.64	0.20	4.25
KFKQFFNMWQ	1.94	0.58	-0.89	10.81	0.00	0.54	1.52	10.11	1.97	0.64	0.20	4.25
KAKQAANMWQ	1.94	0.14	-1.19	10.81	30.00	0.41	1.87	32.68	1.47	0.58	0.13	2.51
NMWQKVGTPL	0.94	0.56	-0.46	9.70	68.00	0.54	0.57	-16.47	1.94	0.51	0.12	2.03
NMWQKIKQII	1.94	0.59	-0.38	10.81	117.00	0.68	0.94	15.70	2.34	0.73	0.17	3.68
NMWQKIKIQI	1.94	0.59	-0.38	10.81	117.00	0.39	0.94	-16.47	2.34	0.73	0.17	3.68
KIKQIINM	1.94	0.48	0.08	10.81	146.25	0.61	0.77	17.13	2.73	0.83	0.19	4.11
KIKQIIN	1.94	0.37	-0.19	10.81	167.14	0.72	1.22	18.14	2.79	0.84	0.19	4.06
KIKIQINMW	1.94	0.68	-0.03	10.81	130.00	0.25	0.43	-19.41	2.43	0.74	0.17	3.65
KIKIQINM	1.94	0.48	0.08	10.81	146.25	0.19	0.77	-23.09	2.73	0.83	0.19	4.11
KIKIQIN	1.94	0.37	-0.19	10.81	167.14	0.28	1.22	-27.81	2.79	0.84	0.19	4.06
KAKAQA	1.94	-0.21	-0.98	10.81	50.00	0.18	1.87	8.33	1.53	0.60	0.13	2.30
KIKIQI	1.94	0.53	0.37	10.81	195.00	0.28	0.31	-34.12	2.99	0.85	0.20	4.25

KFKFQF	1.94	0.53	-0.48	10.81	0.00	0.26	1.28	-29.28	2.36	0.69	0.25	5.19
KIKIQINMWA	1.94	0.64	0.15	10.81	127.00	0.24	0.20	-31.50	2.31	0.72	0.16	3.30
EIEIEIEI	-4.06	0.58	0.50	3.37	195.00	0.14	0.95	269.83	2.88	0.85	0.17	3.89
KIKIQINMAQ	1.94	0.39	-0.11	10.81	127.00	0.22	0.99	-4.13	2.47	0.78	0.18	3.69
KIKIQINAWQ	1.94	0.49	-0.39	10.81	127.00	0.24	0.99	-16.47	2.24	0.71	0.15	3.25
KIKIQIAMWQ	1.94	0.68	0.15	10.81	127.00	0.11	0.09	-16.47	2.31	0.71	0.16	3.40
KIKIKIKIYYYY	3.93	0.59	-0.23	10.20	130.00	0.21	0.26	-9.51	3.01	0.80	0.23	5.08
HHHHEIEIEIEI	-3.61	0.43	-0.73	5.32	130.00	0.17	2.18	183.22	2.91	0.80	0.19	4.15
KYKYQY	1.93	0.11	-2.53	9.74	0.00	0.15	2.84	-4.23	2.36	0.69	0.25	5.48
EFEFQF	-2.06	0.65	-0.35	3.61	0.00	0.23	1.70	-4.23	2.25	0.69	0.23	4.86
CEFEFQF	-2.13	0.77	0.06	3.61	0.00	0.19	1.28	-2.20	2.18	0.68	0.21	4.52
CKFKFQFNMW	1.87	0.76	-0.29	9.68	0.00	0.19	0.84	-13.57	1.99	0.63	0.20	4.10
CKYKYQY	1.86	0.32	-1.81	9.19	0.00	0.13	2.25	-2.20	2.28	0.68	0.23	5.05
KWKWQW	1.94	0.76	-2.33	10.81	0.00	0.21	1.61	8.33	0.89	0.34	0.10	2.25
CKWKWQW	1.87	0.87	-1.64	9.68	0.00	0.17	1.20	8.57	1.02	0.38	0.11	2.27
KLKQLL	1.94	0.48	0.02	10.81	195.00	0.62	0.31	19.50	2.19	0.83	0.20	4.25
CKIKIQINMWQ	1.87	0.67	-0.12	9.68	106.36	0.19	0.74	-14.06	2.29	0.72	0.17	3.57
RGDKIKIQI	1.94	0.16	-0.69	10.79	130.00	0.32	2.73	-38.28	2.43	0.73	0.17	3.82
CKIKIQINM	1.87	0.60	0.34	9.68	130.00	0.14	0.54	-19.41	2.63	0.81	0.18	3.92
CKIKIQI	1.87	0.68	0.67	9.68	167.14	0.23	0.09	-27.81	2.81	0.82	0.19	3.99
CKIKQII	1.87	0.68	0.67	9.68	167.14	0.60	0.09	18.14	2.81	0.82	0.19	3.99
CKIKQIINMWQ	1.87	0.67	-0.12	9.68	106.36	0.50	0.74	15.18	2.29	0.72	0.17	3.57
KIKIQIRGD	1.94	0.16	-0.69	10.79	130.00	0.24	2.73	-28.84	2.43	0.73	0.17	3.82
RGDKIKIQIC	1.87	0.30	-0.37	9.67	117.00	0.31	2.33	-33.45	2.36	0.72	0.17	3.68
RGDKIKIQINM	1.94	0.19	-0.71	10.79	106.36	0.18	2.62	-29.50	2.35	0.73	0.18	3.80
RGDKIKIQINMWQ	1.94	0.31	-0.94	10.79	90.00	0.31	2.47	-23.42	2.10	0.67	0.16	3.52
KFKFEFEF	-0.06	0.49	-0.45	6.53	0.00	0.15	1.60	-10.04	2.33	0.69	0.24	5.08
KIKQII	1.94	0.53	0.37	10.81	195.00	0.74	0.31	19.50	2.99	0.85	0.20	4.25
KIEIQINM	-0.06	0.52	0.13	6.41	146.25	0.20	0.93	77.14	2.69	0.83	0.18	3.99
CKIKIQIRGD	1.87	0.30	-0.37	9.67	117.00	0.19	2.33	-24.96	2.36	0.72	0.17	3.68
CKIKQIIRGD	1.87	0.30	-0.37	9.67	117.00	0.53	2.33	7.21	2.36	0.72	0.17	3.68
RGDKIKIQINMC	1.87	0.30	-0.44	9.67	97.50	0.26	2.30	-26.21	2.30	0.73	0.17	3.68
CQFQFQF	-0.13	0.89	0.06	5.92	0.00	0.19	0.91	-34.51	2.18	0.68	0.22	4.56
KIKIQII	1.94	0.71	0.96	10.81	222.86	0.29	-0.43	-27.81	3.16	0.87	0.19	4.21
KFKFQFFF	1.94	0.84	0.34	10.81	0.00	0.12	0.22	-19.46	2.51	0.69	0.26	5.37
KIKIKIQI	2.94	0.50	0.35	11.10	195.00	0.11	0.31	-44.31	3.00	0.85	0.20	4.28
KAKAKAQA	2.94	-0.24	-1.00	11.10	50.00	0.09	1.87	8.75	1.54	0.60	0.13	2.33
KLKLKLQL	2.94	0.45	0.00	11.10	195.00	0.10	0.31	-3.56	2.20	0.83	0.20	4.28
KFKFKFQF	2.94	0.50	-0.50	11.10	0.00	0.10	1.28	-38.25	2.37	0.69	0.25	5.23
KVKVQVV	1.94	0.38	0.79	10.81	165.71	0.25	0.07	-30.57	2.86	0.73	0.17	2.10
KLKLQLL	1.94	0.66	0.56	10.81	222.86	0.25	-0.43	18.76	2.24	0.85	0.19	4.21
KFKFQFF	1.94	0.71	-0.01	10.81	0.00	0.27	0.67	-23.67	2.44	0.69	0.25	5.29
KIKIQIII	1.94	0.85	1.40	10.81	243.75	0.12	-1.00	-23.09	3.29	0.89	0.19	4.19
KVKVQVVV	1.94	0.49	1.21	10.81	181.25	0.11	-0.45	-25.50	2.96	0.73	0.16	1.87

HHHHKAKAKAKAYYYY	4.38	0.10	-1.65	10.20	25.00	0.17	2.14	47.31	2.28	0.65	0.20	4.00
KIKIQIC	1.87	0.68	0.67	9.68	167.14	0.24	0.09	-27.81	2.81	0.82	0.19	3.99
RGDSKIKIQIC	1.87	0.27	-0.41	9.67	106.36	0.32	2.43	-4.27	2.27	0.70	0.16	3.49
RGDSGGGGGGKIKIQIC	1.87	0.17	-0.41	9.67	68.82	0.34	1.24	32.06	1.47	0.45	0.10	2.26
ILKNLSRSRKIKIQIC	4.87	0.39	-0.08	11.75	146.25	0.52	2.17	51.73	2.49	0.79	0.18	3.94
KIKIKIKIWWWW	3.94	1.02	-0.10	11.28	130.00	0.23	-0.57	-40.36	2.03	0.57	0.14	2.92
KIKIKIKI	3.94	0.41	0.30	11.28	195.00	0.18	0.32	-65.54	3.04	0.85	0.20	4.39
CEFEFQFNMWQ	-2.13	0.73	-0.51	3.61	0.00	0.16	1.49	2.24	1.89	0.63	0.18	3.90
CSISIQI	-0.13	0.95	1.56	5.92	167.14	0.19	-0.53	8.57	2.65	0.77	0.14	3.08
CEIEIQINMWQ	-2.13	0.74	-0.05	3.61	106.36	0.17	0.97	84.05	2.23	0.72	0.15	3.39
CEIEIQINM	-2.13	0.67	0.43	3.61	130.00	0.11	0.82	100.51	2.55	0.81	0.17	3.70
CSISIQINM	-0.13	0.81	1.03	5.92	130.00	0.09	0.07	8.89	2.50	0.77	0.15	3.22
KAKAQANM	1.94	-0.08	-0.94	10.81	37.50	0.11	1.94	8.75	1.64	0.64	0.14	2.65
CKAKAQA	1.87	0.04	-0.49	9.68	42.86	0.15	1.42	8.57	1.56	0.60	0.13	2.32
HHHHKIKIQINMYYYY	2.38	0.51	-1.09	9.61	73.13	0.20	1.59	46.19	2.85	0.77	0.23	4.84
KIKIKIKIWW	3.94	0.77	0.06	11.28	156.00	0.23	-0.21	-50.43	2.43	0.68	0.16	3.51
HHHHKIKIKIKIWWWW	4.39	0.80	-0.88	11.28	97.50	0.24	0.74	-12.97	2.27	0.60	0.16	3.36
MKIKIQINM	1.94	0.56	0.28	10.81	130.00	0.11	0.42	-19.41	2.69	0.82	0.19	4.14
RGDCKIKIQINM	1.87	0.30	-0.44	9.67	97.50	0.33	2.30	-19.13	2.30	0.73	0.17	3.68
MKIKIQINMWQ	1.94	0.64	-0.17	10.81	106.36	0.16	0.64	-14.06	2.34	0.74	0.18	3.75
HHHHKIKIKIKIYYYY	4.38	0.48	-0.98	10.20	97.50	0.22	1.36	10.17	3.00	0.78	0.23	4.98
RGDCKFKFQF	1.87	0.29	-0.88	9.67	0.00	0.32	2.91	-22.06	1.99	0.62	0.20	4.25
KIKIQIW	1.94	0.78	0.19	10.81	167.14	0.25	-0.06	-27.81	2.56	0.73	0.17	3.64
KFKFQFW	1.94	0.77	-0.54	10.81	0.00	0.24	0.77	-23.67	2.02	0.59	0.21	4.45
CKFKFQFW	1.87	0.87	-0.16	9.68	0.00	0.22	0.51	-19.46	1.99	0.60	0.20	4.20
EIKIQINM	-0.06	0.52	0.13	6.41	146.25	0.19	0.93	11.60	2.69	0.83	0.18	3.99
IKVAVKIKIQINM	2.94	0.57	0.88	11.10	172.31	0.36	-0.24	-25.64	2.81	0.80	0.17	3.26
HHHHKAKAKAKAWWWW	4.39	0.43	-1.55	11.28	25.00	0.22	1.52	24.18	1.54	0.48	0.12	2.38
CKIW	0.87	1.15	0.55	8.55	97.50	0.67	-0.75	-13.73	1.96	0.58	0.13	2.80
CKIKIQINMW	1.87	0.76	0.22	9.68	117.00	0.20	0.26	-16.47	2.36	0.73	0.17	3.53
ILKNLSRSRKIKIQINMWQ	4.94	0.39	-0.51	12.55	123.16	0.52	2.29	45.14	2.29	0.75	0.18	3.79
EIEIQINM	-2.06	0.57	0.18	3.61	146.25	0.16	1.09	111.83	2.65	0.83	0.17	3.86
KIKIQINMWWQ	1.94	0.74	-0.43	10.81	106.36	0.30	0.64	-14.06	2.13	0.67	0.16	3.35
KIKIQINMWWWQ	1.94	0.86	-0.47	10.81	97.50	0.16	0.39	-12.06	1.95	0.61	0.14	3.07
KFKQFFINMWQ	1.94	0.69	-0.40	10.81	35.45	0.49	0.93	10.10	2.17	0.67	0.20	4.23
KIKIQIMWNQ	1.94	0.59	-0.38	10.81	117.00	0.06	0.94	-11.67	2.34	0.73	0.17	3.68
KIKIQIMQWN	1.94	0.59	-0.38	10.81	117.00	0.22	0.94	-11.67	2.34	0.73	0.17	3.68
KIKIQINMWQRGD	1.94	0.31	-0.94	10.79	90.00	0.49	2.47	-16.89	2.10	0.67	0.16	3.52
EIEIQINMWQ	-2.06	0.66	-0.30	3.61	117.00	0.21	1.19	91.46	2.28	0.73	0.16	3.48
KIKIKIQINMWQ	2.94	0.56	-0.27	11.10	130.00	0.09	0.83	-26.21	2.46	0.75	0.18	3.80

Experimental descriptors

Table S5 Summary of experimental descriptors measured for the peptide library: Absolute Infectivity at 1.3 µM peptide concentration (Abs. Inf.) and the corresponding standard deviation from triplicate measurements (Std. Dev.), the infectivity relative to EF-C (Inf. rel. EF-C) and the logarithmic infectivity relative to EF-C (Log_{10} Inf. rel EF-C), percentual β–sheet content from FT-IR measurements (**Figure S34**), ThT-fluorescence relative to control (Rel. ThT) and categorical ThT fluorescence (ThT. Cat is 1, if Rel. ThT >2), zeta potential (Zeta-pot.) and the derived count rate of scattered light from zeta-potential measurements (Count Rate), Conversion of monomers to assembled structure (Conversion Rate CR) and fibril formation (1 = fibrils present, 0 non-fibrillar structures) as evaluated qualitatively from TEM measurements (**Figure S33**).

Sequence	Abs. Inf. / RLU s^{-1}	Std. Dev. / RLU s^{-1}	Inf. rel. EF-C	Std. Dev.	Log_{10} Inf. rel. EF-C	ß-sheet / %	Rel. ThT	ThT Cat.	Zeta pot. / mV	Std. Dev. / mV	Count Rate / kcps	Std. Dev. / kcps	Log_{10} Count Rate	CR/ %	Std. Dev. CR / %	Fibril
CKIKIQIII	22150	2532	0.98	0.20	-0.01	37	4.1	1	16.2	3.2	4953	4191	3.7			1
CEIEIQI	633	578	0.03	0.03	-1.55	45	2.2	1	-26.1	7.1	52	25	1.7	59	3	1
CKIKQIINM	11280	3038	0.50	0.16	-0.30	46	14.7	1	12.9	4.8	9058	2261	4.0	45	2	1
CKFKFQFNMWQ	26320	2463	1.16	0.23	0.06	47	2.4	1	17.0	0.3	143	56	2.2	98	2	1
CKFKFQFNM	28183	3701	1.24	0.27	0.09	41	2.4	1	15.1	0.9	2224	2193	3.3	94	6	1
CKFKFQF	20067	2227	1.23	0.25	0.09	47	2.3	1	19.7	0.2	5303	4282	3.7	96	1	1
KFKFQFNMW	6690	1272	0.45	0.12	-0.35	38	1.3	0	22.6	2.8	610	162	2.8			1
KFKFQFNM	7067	1516	0.47	0.13	-0.33	40	1.2	0	13.9	1.5	132	141	2.1	95	4	1
KFKFQFN	270	135	0.02	0.01	-1.74	38	1.1	0	12.7	10.0	51	9	1.7			1
CKAKAQANMWQ	600	134	0.04	0.01	-1.36	39	1.9	0	0.7	2.3	9	7	1.0	92	0	0
CKAKAQANM	430	114	0.03	0.01	-1.50	34	1.2	0	-7.8	1.4	207	146	2.3	48	10	0
QCKFKQFFNMWQ	30897	2664	0.25	0.07	-0.61	42	25.4	1	9.6	0.4	18911	16193	4.3	99	0	1
QCKFKQFFNM	5847	75	0.43	0.10	-0.37	42	12.7	1	13.7	0.4	10385	1136	4.0	99	0	1
QCKFKQFF	170	1010	0.01	0.01	-1.91	32	2.3	1	4.9	5.2	225	280	2.4			1
CKFKQFFNMWQ	23487	142	0.61	0.16	-0.21	59	27.7	1	8.6	4.1	8210	2689	3.9	99	1	1
CKFKQFFNM	15720	4394	0.41	0.16	-0.39	42	46.0	1	15.2	3.0	27264	5488	4.4			1
CKFKQFF	260	5573	0.01	0.00	-2.17	45	1.3	0	11.6	1.4	3	1	0.4	90	3	0
QCKIKIQINM	15250	156	0.40	0.16	-0.40	37	2.5	1	11.4	0.3	9524	5626	4.0	90	7	1
QCKAKAQANMWQ	2767	5347	0.07	0.02	-1.14	33	1.5	0	-8.1	0.4	428	487	2.6	89	7	0
QCKAKAQANM	563	671	0.01	0.01	-1.83	35	1.4	0	-8.4	2.6	204	224	2.3	56	35	0
QCKAKAQA	807	176	0.03	0.01	-1.60	34	1.3	0	0.7	5.2	1319	1729	3.1			0
QCKIKIQI	3577	272	0.11	0.04	-0.95	35	2.0	1	9.5	3.2	12427	14002	4.1			1
QCKIKQIINM	11187	1058	0.35	0.12	-0.46	38	23.3	1	10.4	3.2	20551	16126	4.3			1
QCKIKQII	207	3309	0.01	0.00	-2.19	41	3.0	1	11.0	2.5	301	199	2.5			1
QCKFKFQFNMWQ	19753	151	0.62	0.12	-0.21	47	1.9	0	21.1	0.8	1208	302	3.1	99	0	1
QCKFKFQFNM	16353	1583	0.51	0.10	-0.29	37	2.3	1	15.7	0.7	3670	2822	3.6	99	2	1
QCKFKFQF	20303	1679	0.77	0.16	-0.12	40	2.0	1	12.2	3.7	2669	932	3.4			1
CRFRFQF	7508	2292	0.34	0.20	-0.46	40	3.5	1	-2.1	3.9	391	338	2.6	98	0	1
HIHIQIC	4621	1135	0.27	0.16	-0.57	48	12.5	1	-3.7	0.5	261	128	2.4	83	5	1
RLRLTLC	1334	826	0.05	0.03	-1.29	21	1.9	0	19.7	0.9	126	117	2.1	87	5	1
HLHLPLL	133	167	0.01	0.01	-2.00	49	1.1	0	-9.1	0.5	200	207	2.3	38	10	0

RGECKFKFQF	7411	193	0.33	0.18	-0.48	46	2.4	1	24.0	0.3	219	213	2.3	99	0	1
RGEKIKIQINM	7838	1054	0.24	0.11	-0.61	38	3.4	1	15.3	3.0	595	686	2.8	80	7	1
KYKGAIIGNIK	80	1260	0.00	0.00	-2.51	45	1.4	0	-9.1	3.6	44	21	1.6			0
HGDKCHGDKC	263	121	0.01	0.01	-1.99	5	1.2	0	-11.6	1.2	12	10	1.1	37	13	0
RPRGLLLGNLR	717	229	0.05	0.02	-1.32	14	1.3	0	-8.2	6.5	210	156	2.3			0
KKFQKKFQ	423	220	0.03	0.02	-1.54	40	1.2	0	-9.2	4.0	74	37	1.9			0
PPFHPPPFHP	263	234	0.02	0.01	-1.75	47	1.4	0	-15.8	9.7	59	26	1.8			0
MDQMDQMDQMDQMDQ	63	143	0.00	0.01	-2.37	46	1.1	0	-15.0	4.0	10	4	1.0	53	3	0
FDPFDPFDP	347	75	0.02	0.00	-1.63	49	1.3	0	-10.9	3.6	258	235	2.4			0
TKTLTKTL	317	31	0.02	0.01	-1.67	34	1.3	0	-25.1	4.1	515	356	2.7			0
FKFDKFKFDK	513	55	0.05	0.02	-1.33	47	1.5	0	-12.7	10.3	45	22	1.7			0
KVKGVGK	287	213	0.03	0.01	-1.59	9	1.8	0	-12.8	4.2	51	30	1.7			0
SISISRRI	313	133	0.03	0.03	-1.55	49	1.0	0	-17.0	0.7	25	14	1.4	24	7	0
HRRHFRHKITKKK	150	289	0.01	0.01	-1.87	45	1.1	0	-9.2	2.7	10	4	1.0	45	2	0
KNERIKNERI	143	126	0.01	0.00	-1.89	38	1.3	0	-16.0	3.7	322	166	2.5	54	8	0
KIRGKFEKED	643	20	0.06	0.03	-1.24	29	3.4	1	-12.7	6.7	214	151	2.3			0
CKFQC	497	364	0.02	0.00	-1.64	44	1.5	0	-17.4	6.5	130	106	2.1	33	2	0
MKFM	383	57	0.02	0.00	-1.76	35	1.1	0	-16.0	5.2	131	159	2.1			0
CKFC	240	46	0.01	0.01	-1.96	39	1.4	0	-19.9	10.0	235	136	2.4	36	11	0
RGDKIRGDKI	163	156	0.01	0.00	-2.13	15	1.1	0	-13.6	0.4	286	367	2.5	48	10	0
KNDKND	230	85	0.01	0.01	-1.98	43	1.3	0	-15.0	25.8	52	29	1.7			0
HGEHGE	140	116	0.01	0.00	-2.19	10	1.6	0	-12.6	3.5	241	285	2.4			0
HGEHGEHGE	360	35	0.02	0.01	-1.72	6	1.2	0	-24.2	7.2	340	318	2.5	43	7	0
CRFRVPF	377	117	0.02	0.01	-1.70	11	0.9	0	4.4	6.0	74	90	1.9	97	1	0
CHLHLQL	2934	106	0.10	0.05	-1.01	50	3.2	1	-12.3	1.5	1155	381	3.1	97	2	0
CETMYDKILKNLSRSR	363	670	0.02	0.01	-1.72	62	1.1	0	0.4	0.3	1099	1415	3.0	99	0	0
MKFKFQF	1427	127	0.06	0.01	-1.20	22	1.0	0	-3.1	0.9	124	68	2.1	87	1	1
QCKIKQIINMWQ	21312	70	1.00	0.25	0.00	52	3.8	1	6.6	0.4	2606	936	3.4	77	7	1
QCKIKIQINMWQ	21221	3959	1.04	0.27	0.02	36	9.1	1	11.7	0.4	8271	7432	3.9	94	4	1
KIKIQINMWQ	6650	4304	0.30	0.07	-0.53	42	5.9	1	14.9	0.9	2196	1431	3.3	94	1	1
NMWQKAKAQA	1137	1144	0.04	0.01	-1.38	37	5.2	1	-5.6	8.9	242	186	2.4			1
NMWQKFKFQF	7323	244	0.27	0.23	-0.57	33	1.0	0	9.6	3.0	14045	10746	4.1			1
KVKVKVQV	467	6103	0.02	0.01	-1.76	43	5.8	1	-3.4	6.8	94	74	2.0			0
KIKQIINMWQ	9363	178	0.35	0.16	-0.46	41	4.9	1	15.4	0.8	616	488	2.8	90	2	1
KAKAQANMWQ	193	3859	0.01	0.00	-2.15	43	1.0	0	-19.6	1.7	56	45	1.7	56	4	0
KFKFQFNMWQ	20677	104	0.76	0.24	-0.12	43	2.8	1	25.4	0.8	735	170	2.9	96	0	1
KFKQFFNMWQ	1000	5282	0.09	0.02	-1.05	33	1.6	0	-3.4	5.1	237	94	2.4			0
KAKQAANMWQ	343	79	0.03	0.02	-1.51	11	1.2	0	-12.9	5.2	690	585	2.8			0
NMWQKVGTPL	277	162	0.02	0.01	-1.61	32	2.2	1	-9.6	5.7	27	8	1.4			0
NMWQKIKQII	283	76	0.03	0.02	-1.60	38	1.3	0	-0.2	3.8	982	389	3.0			0
NMWQKIKIQI	2300	180	0.21	0.04	-0.69	39	5.0	1	8.3	3.4	10024	8218	4.0			1
KIKQIINM	5053	145	0.45	0.12	-0.35	44	1.2	0	16.3	4.1	2481	1952	3.4	43	9	1
KIKQIIN	323	1079	0.04	0.01	-1.40	38	1.1	0	-11.7	9.8	64	61	1.8			0
KIKIQINMW	17523	90	2.16	0.44	0.33	46	4.6	1	19.7	4.2	240	92	2.4			1

KIKIQINM	2947	1917	0.36	0.08	-0.44	41	1.6	0	8.5	6.4	262	143	2.4	64	3	1
KIKIQIN	2153	339	0.27	0.09	-0.58	42	1.8	0	3.3	12.9	43	26	1.6			1
KAKAQA	147	595	0.02	0.01	-1.74	40	1.2	0	-14.1	3.1	21	12	1.3	0	20	0
KIKIQI	550	111	0.07	0.01	-1.17	39	1.1	0	-5.6	1.6	108	80	2.0	19	11	0
KFKFQF	687	60	0.06	0.01	-1.21	46	1.2	0	-11.0	1.3	183	152	2.3	67	12	0
KIKIQINMWA	7677	60	0.69	0.15	-0.16	50	1.5	0	14.7	7.3	1101	467	3.0			1
EIEIEIEI	200	961	0.02	0.01	-1.75	63	2.0	0	-8.9	4.9	538	681	2.7			1
KIKIQINMAQ	2383	152	0.21	0.06	-0.67	46	1.6	0	3.1	92.6	5859	8262	3.8			1
KIKIQINAWQ	6013	567	0.54	0.12	-0.27	28	2.5	1	13.1	3.0	1494	1899	3.2			1
KIKIQIAMWQ	4350	783	0.39	0.07	-0.41	37	2.4	1	11.3	4.0	541	620	2.7			1
KIKIKIKIYYYY	30950	289	1.38	0.26	0.14	57	1.3	0	13.3	5.1	9918	3446	4.0			0
HHHHEIEIEIEI	655	1980	0.03	0.01	-1.53	48	10.6	1	-6.9	62.3	103	53	2.0			0
KYKYQY	230	219	0.01	0.01	-1.99	0	6.1	1	-4.2	5.4	32	19	1.5			0
EFEFQF	73	246	0.00	0.00	-2.49	45	2.8	1	-24.5	4.3	405	111	2.6			0
CEFEFQF	117	91	0.01	0.00	-2.28	53	1.4	0	-13.7	13.1	367	494	2.6			1
CKFKFQFNMW	21517	80	0.96	0.21	-0.02	48	6.3	1	18.4	3.2	4107	2513	3.6			1
CKYKYQY	1083	2692	0.04	0.02	-1.39	49	3.1	1	1.5	3.6	1061	439	3.0			0
KWKWQW	387	383	0.01	0.01	-1.84	30	1.5	0	7.6	7.3	1545	1957	3.2			0
CKWKWQW	6760	136	0.25	0.05	-0.60	34	1.3	0	14.7	3.1	7511	924	3.9			1
KLKQLL	197	575	0.01	0.00	-2.13	32	1.1	0	-8.5	3.3	320	216	2.5			0
CKIKIQINMWQ	18623	25	0.70	0.14	-0.16	45	4.6	1	13.9	0.4	1881	680	3.3	89	2	1
RGDKIKIQI	417	1812	0.02	0.02	-1.81	40	5.6	1	-0.9	14.5	298	401	2.5			1
CKIKIQINM	30273	394	1.59	0.30	0.20	43	2.6	1	16.8	0.5	4228	485	3.6	75	3	1
CKIKIQI	19407	1919	1.02	0.23	0.01	53	3.9	1	12.8	3.5	2433	1343	3.4	54	1	1
CKIKQII	797	2633	0.04	0.02	-1.38	25	3.9	1	-3.8	6.4	75	27	1.9	55	1	1
CKIKQIINMWQ	23697	276	1.24	0.31	0.09	51	48.5	1	13.4	3.4	8317	3111	3.9	50	2	1
KIKIQIRGD	400	4304	0.02	0.01	-1.68	47	40.5	1	6.4	4.6	114	32	2.1			1
RGDKIKIQIC	19382	163	0.83	0.17	-0.08	41	19.8	1	16.5	3.8	7509	8351	3.9	80	4	1
RGDKIKIQINM	9554	2669	0.41	0.10	-0.38	46	9.1	1	16.4	1.9	57	32	1.8	61	13	1
RGDKIKIQINMWQ	12503	1540	0.76	0.15	-0.12	49	1.8	0	6.6	7.7	1554	1432	3.2	66	14	1
KFKFEFEF	577	1244	0.03	0.02	-1.55	41	2.7	1	-2.6	3.6	2186	1813	3.3			1
KIKQII	100	319	0.00	0.00	-2.31	39	1.0	0	-12.7	6.4	122	146	2.1	13	7	0
KIEIQINM	707	82	0.03	0.01	-1.46	43	1.0	0	-0.4	-1.0	275	215	2.4			1
CKIKIQIRGD	6660	181	0.33	0.16	-0.48	58	0.0	0	7.2	0.6	5777	938	3.8	91	1	1
CKIKQIIRGD	993	3045	0.05	0.01	-1.31	nd	1.4	0	13.2	3.4	3495	2551	3.5			0
RGDKIKIQINMC	21717	198	1.07	0.26	0.03	nd	1.4	0	12.2	3.7	1404	774	3.1	60	8	1
CQFQFQF	7570	3707	0.35	0.08	-0.45	35	13.0	1	-24.8	3.8	7929	4815	3.9			1
KIKIQII	400	955	0.02	0.01	-1.73	45	1.4	0	5.6	9.3	1039	219	3.0			1
KFKFQFFF	8657	93	0.40	0.18	-0.39	49	2.6	1	13.3	0.7	3039	608	3.5	98	1	1
KIKIKIQI	2073	3591	0.10	0.03	-1.01	40	2.0	1	7.0	6.6	999	801	3.0	80	1	1
KAKAKAQA	187	563	0.01	0.00	-2.06	40	1.2	0	-9.0	5.9	77	65	1.9			0
KLKLKLQL	2097	55	0.10	0.04	-1.01	42	1.6	0	8.4	8.8	43	2	1.6			1
KFKFKFQF	18697	862	0.80	0.16	-0.10	40	41.9	1	15.2	3.3	1036	442	3.0	94	5	1
KVKVQVV	213	2009	0.01	0.00	-2.04	45	1.3	0	-7.9	5.6	28	21	1.4			0

KLKLQLL	163	53	0.01	0.01	-2.16	40	0.8	0	-4.9	4.4	73	62	1.9			1
KFKFQFF	2850	135	0.12	0.04	-0.91	48	2.6	1	6.9	1.7	38	13	1.6			1
KIKIQIII	8260	702	0.35	0.08	-0.45	46	2.3	1	19.8	9.3	95	40	2.0	87	2	1
KVKVQVVV	820	1160	0.04	0.01	-1.46	46	0.9	0	0.1	10.5	163	192	2.2			1
HHHHKAKAKAKAYYYY	203	78	0.01	0.00	-1.88	44	1.4	0	7.0	1.4	241	83	2.4			0
KIKIQIC	9447	17	0.61	0.11	-0.22	53	4.1	1	19.9	1.6	981	305	3.0	52	11	1
RGDSKIKIQIC	10816	150	0.60	0.19	-0.22	39	13.8	1	15.1	3.9	2699	210	3.4	43	11	1
RGDSGGGGGGKIKIQIC	13453	3541	0.87	0.18	-0.06	41	42.0	1	16.7	0.7	10416	5402	4.0	55	25	1
ILKNLSRSRKIKIQIC	12740	1659	0.34	0.08	-0.47	38	2.7	1	18.0	4.2	2392	1866	3.4			1
KIKIKIKIWWWW	22347	1999	0.59	0.11	-0.23	38	2.3	1	15.2	3.0	12421	2307	4.1			0
KIKIKIKI	1237	1418	0.03	0.01	-1.49	40	3.3	1	5.5	13.4	155	38	2.2			0
CEFEFQFNMWQ	2517	411	0.14	0.08	-0.85	51	6.1	1	-34.2	0.4	186	44	2.3	99	0	1
CSISIQI	3087	326	0.06	0.01	-1.25	59	15.0	1	-25.9	1.8	10241	5508	4.0	79	3	1
CEIEIQINMWQ	12067	422	0.22	0.05	-0.66	50	3.7	1	-25.8	1.3	1570	1451	3.2	67	7	1
CEIEIQINM	3593	1935	0.07	0.01	-1.19	46	5.1	1	-15.3	8.4	58	75	1.8	43	9	1
CSISIQINM	10943	497	0.20	0.06	-0.70	56	5.9	1	-19.4	3.2	3291	1256	3.5	86	3	1
KAKAQANM	497	2645	0.01	0.01	-2.05	33	1.1	0	-2.6	4.3	43	25	1.6			0
CKAKAQA	1740	335	0.03	0.01	-1.50	35	1.2	0	-8.0	5.8	363	387	2.6	44	9	0
HHHHKIKIQINMYYYY	22409	406	1.11	0.62	0.04	47	12.2	1	7.8	0.6	5659	3747	3.8	98	0	1
KIKIKIKIWW	20557	1379	0.37	0.09	-0.43	49	3.5	1	15.4	3.1	9436	981	4.0	92	3	0
HHHHKIKIKIKIWWWW	20797	3303	0.38	0.07	-0.42	49	4.0	1	14.5	3.4	23757	9196	4.4	96	1	0
MKIKIQINM	15000	1065	0.27	0.05	-0.56	39	1.3	0	9.1	2.9	981	707	3.0			1
RGDCKIKIQINM	7203	1002	0.24	0.05	-0.63	43	1.8	0	11.3	5.5	1623	1122	3.2			1
MKIKIQINMWQ	12787	869	0.42	0.10	-0.38	39	4.1	1	10.8	2.9	1475	1141	3.2			1
HHHHKIKIKIKIYYYY	20713	2171	0.68	0.20	-0.17	53	9.9	1	13.2	3.8	12793	5972	4.1			0
RGDCKFKFQF	4637	4934	0.15	0.05	-0.82	47	6.8	1	13.9	1.5	1338	300	3.1	85	9	1
KIKIQIW	4280	1272	0.14	0.03	-0.85	49	2.0	1	12.5	4.2	145	168	2.2			1
KFKFQFW	4953	453	0.16	0.03	-0.79	35	3.5	1	6.5	2.8	122	56	2.1			1
CKFKFQFW	17862	357	0.88	0.18	-0.06	43	1.8	0	15.2	1.1	29211	17451	4.5	99	0	1
EIKIQINM	3607	2252	0.10	0.04	-1.00	48	3.8	1	-22.3	2.1	2816	2882	3.4	51	7	1
IKVAVKIKIQINM	17920	1317	0.50	0.09	-0.30	58	1.8	0	14.5	3.5	4848	1475	3.7			1
HHHHKAKAKAKAWWWW	3233	592	0.09	0.02	-1.05	40	3.7	1	17.4	8.6	659	233	2.8			0
CKIW	290	434	0.05	0.03	-1.29	46	1.0	0	-3.7	0.6	3	2	0.5	48	15	0
CKIKIQINMW	7253	159	1.29	0.24	0.11	43	1.4	0	18.3	3.7	4915	5114	3.7			1
ILKNLSRSRKIKIQINMWQ	1797	478	0.32	0.07	-0.49	39	3.3	1	19.2	3.8	307	189	2.5			1
EIEIQINM	203	221	0.04	0.01	-1.44	54	1.5	0	-23.3	1.1	134	85	2.1	78	4	1
KIKIQINMWWQ	17977	44	3.21	0.56	0.51	50	1.3	0	11.8	4.0	3893	1970	3.6			0
KIKIQINMWWWQ	15400	408	2.75	0.51	0.44	48	1.3	0	9.5	2.6	7596	2548	3.9			0
KFKQFFINMWQ	14953	937	0.92	0.26	-0.04	38	1.1	0	8.0	3.2	8996	4347	4.0			1
KIKIQIMWNQ	9147	3226	0.56	0.10	-0.25	41	1.2	0	3.0	4.0	1887	306	3.3			1
KIKIQIMQWN	11623	168	0.72	0.18	-0.15	40	1.3	0	9.4	10.5	2203	415	3.3			1
KIKIQINMWQRGD	197	2183	0.01	0.01	-1.92	48	1.5	0	1.7	6.4	2783	3327	3.4			0
EIEIQINMWQ	2137	85	0.13	0.03	-0.88	54	4.1	1	-19.1	1.3	1264	133	3.1	90	1	0
KIKIKIQINMWQ	20353	176	0.79	0.15	-0.10	40	1.6	0	13.1	2.9	1603	498	3.2		3	1

Table S6 Summary of experimental descriptors measured for the peptides designed via pattern analysis: Absolute Infectivity at 1.3 µM peptide concentration (Abs. Inf.) and the corresponding standard deviation from triplicate measurements (Std. Dev.), the infectivity relative to EF-C (Inf. rel. EF-C) and the logarithmic infectivity relative to EF-C (Log_{10} Inf. rel EF-C), percentual β–sheet content from FT-IR measurements (**Figure S38**), ThT-fluorescence relative to control (Rel. ThT) and categorical ThT fluorescence (ThT. Cat is 1, if Rel. ThT >2), zeta potential (Zeta-pot.) and the derived count rate of scattered light from zeta-potential measurements (Count Rate), and fibril formation (1 = fibrils present, 0 non-fibrillar structures) as evaluated qualitatively from TEM measurements (**Figure S25**).

Sequence	Abs. Inf. / RLU s^{-1}	Std. Dev. / RLU s^{-1}	Inf. rel. EF-C	Std. Dev.	Log_{10} Inf. rel. EF-C	ß-sheet / %	Rel. ThT	ThT Cat.	Zeta pot. / mV	Std. Dev. / mV	Count Rate / kcps	Std. Dev. / kcps	Log_{10} Count Rate	Fibril
CKIWIQIY	115280	27300	0.57	0.14	-0.24	31	6	1	4.5	0.1	1842	818	3.3	1
KIWIWIW	105670	2895	0.53	0.05	-0.28	33	9	1	11.3	0.3	6187	4991	3.8	1
KFQFNMW	36267	3869	0.18	0.02	-0.74	28	1	0	8.8	0.7	332	261	2.5	1
KIKIWIW	65420	6552	0.33	0.04	-0.49	40	1	0	17.2	0.4	608	46	2.8	1
CKIKIWIW	173537	22193	0.86	0.13	-0.06	39	2	1	19.1	0.3	3321	309	3.5	1
RVHLGLG	4140	1152	0.02	0.01	-1.69	30	1	0	-6.0	1.0	28	34	1.5	1
CRLSLSLS	63003	10341	0.31	0.06	-0.50	38	2	0	9.0	0.6	1697	1030	3.2	1
CRLRVGVS	8737	3852	0.04	0.02	-1.36	0	1	0	10.5	0.5	35	16	1.5	0
RVRVTVT	11983	2526	0.06	0.01	-1.22	40	1	0	-3.7	1.5	25	12	1.4	0
RINFWFW	9033	4840	0.04	0.02	-1.35	52	16	1	10.8	0.2	50996	8911	4.7	0
RFWFWFW	64303	19492	0.32	0.10	-0.49	31	7	1	13.6	0.1	30037	37815	4.5	0
RIRFHIW	2087	1777	0.01	0.01	-1.98	37	1	0	0.1	0.2	12	4	1.1	0
CRFRFHFH	106307	27039	0.53	0.14	-0.28	38	1	0	19.6	0.5	430	217	2.6	1
KAQCTLP	14133	5124	0.07	0.03	-1.15	0	1	0	-8.5	1.8	97	119	2.0	0
KCQCQCQ	230517	3686	1.15	0.10	0.06	33	15	0	11.8	0.3	15236	3247	4.2	1
KAKLTCT	13367	5300	0.07	0.03	-1.18	41	1	0	-2.3	1.6	17	15	1.2	0
KLKLTLT	7733	6934	0.04	0.03	-1.41	40	1	1	1.7	0.1	199	83	2.3	1
CKFKFPF	4987	4704	0.02	0.02	-1.61	0	1	0	-1.1	1.4	143	147	2.2	0

Table S7 Summary of experimental descriptors measured for the peptides found from data base screening: Absolute Infectivity at 1.3 µM peptide concentration (Abs. Inf.) and the corresponding standard deviation from triplicate measurements (Std. Dev.), the infectivity relative to EF-C (Inf. rel. EF-C) and the logarithmic infectivity relative to EF-C (Log_{10} Inf. rel EF-C), percentual β–sheet content from FT-IR measurements (**Figure S39**), ThT-fluorescence relative to control (Rel. ThT) and categorical ThT fluorescence (ThT. Cat is 1, if Rel. ThT >2), zeta potential (Zeta-pot.) and the derived count rate of scattered light from zeta-potential measurements (Count Rate), and fibril formation (1 = fibrils present, 0 non-fibrillar structures) as evaluated qualitatively from TEM measurements (**Figure S36**).

Sequence	Abs. Inf. / RLU s^{-1}	Std. Dev. / RLU s^{-1}	Inf. rel. EF-C	Std. Dev.	Log_{10} Inf. rel. EF-C	ß-sheet / %	Rel. ThT	ThT Cat.	Zeta pot. / mV	Std. Dev. / mV	Count Rate / kcps	Std. Dev. / kcps	Log_{10} Count Rate	Fibril
IKFLSVN	8601	3301	0.07	0.02	-1.16	48	1	0	5.9	0.7	137	43	2.1	1
VQIVYK	7994	1434	0.07	0.01	-1.17	45	2	0	-2.7	1.0	26	22	1.4	0
IIKIIK	9742	3518	0.08	0.03	-1.10	52	1	0	-7.0	1.0	4	2	0.6	0
GQTVTK	6998	2415	0.06	0.02	-1.25	29	1	0	-5.3	1.0	34	8	1.5	0
DFNKF	9988	2476	0.08	0.02	-1.07	53	1	0	-11.8	0.4	1065	1483	3.0	1
RCELARTLKR	13881	4241	0.09	0.03	-1.06	16	4	1	-0.9	0.6	78	40	1.9	0
ISKLEYSNFSVRY	92993	11998	0.54	0.09	-0.26	43	1	0	9.3	0.2	7807	2555	3.9	1
GYVIIK	13085	3603	0.10	0.02	-0.99	42	1	0	9.9	1.3	30	4	1.5	0
KLVFFAK	13708	3677	0.07	0.02	-1.15	42	1	0	3.2	0.3	14	9	1.2	1
RATVYV	7568	4530	0.04	0.03	-1.38	46	1	0	-5.2	1.6	42	41	1.6	1
RQGNINIVA	21525	7173	0.12	0.04	-0.93	50	5	1	10.2	0.8	36	29	1.6	1
KYFIL	11684	2396	0.07	0.02	-1.14	21	203	1	-7.2	0.5	18903	20938	4.3	0
NKGAII	10918	3461	0.06	0.02	-1.20	31	1	0	-9.5	2.5	68	82	1.8	0
IIIQGK	11434	2250	0.06	0.01	-1.21	44	1	0	-7.6	0.1	10	2	1.0	0
LASLIYR	12491	2722	0.06	0.02	-1.20	49	1	0	-5.0	2.6	105	123	2.0	0
IIIK	9273	3071	0.05	0.02	-1.29	48	1	0	-3.8	1.9	26	7	1.4	0
MKVIFLKDVKG	20404	3725	0.12	0.02	-0.93	53	1	0	9.6	0.8	137	109	2.1	1
LANWMCLAKW	76339	8813	0.40	0.06	-0.40	36	11	1	14.7	0.3	15062	4245	4.2	0
VHDCVNITIK	103498	13829	0.53	0.08	-0.27	51	428	1	9.3	0.3	12542	10750	4.1	1
IFQINS	12629	2531	0.06	0.01	-1.21	39	7	1	-11.0	3.8	36	20	1.6	1
NSGAITIG-NH_2	11611	2899	0.05	0.02	-1.26	32	4	1	-8.4	0.9	20	18	1.3	0
GAITIG-NH_2	8902	2091	0.04	0.01	-1.35	39	1	0	-5.4	0.9	30	31	1.5	0
VLYVGSKT	12423	2982	0.06	0.02	-1.24	48	1	0	-9.1	3.3	24	19	1.4	0
GNNQQNY	11825	4465	0.05	0.02	-1.28	31	2	1	-6.6	1.1	37	31	1.6	0

Table S8 Summary of aggregate analysis from widefield microscopy measurements (**Figure S32**). The Number of aggregated peptides, the median size area, average size area and the cumulative total size area of all particles (> 10 µm^2) in an area of 1330 µm × 1330 µm was determined via automated 3D object counting (ImageJ) in triplicate measurements (Std. Dev.).

Sequence	Abs. Inf. / RLU s^{-1}	Std. Dev. / RLU s^{-1}	Count Rate / kcps	Std. Dev. / kcps	Aggregate Number	Std. Dev. Aggregate Number	Median Area / µm²	Std. Dev. Median Area / µm²	Average Area / µm²	Std. Dev. Average Area / µm²	Total Area/ µm²	Std. Dev. Total Area/ µm²
CKFKQFFNM	15720	4394	27264	5488	704	90	15	1	31	44	21881	2095
CKFKFQFW	17862	357	29211	17451	517	80	17	1	28	35	14537	2313
QCKIKQIINMWQ	21312	70	2602	937	408	35	28	3	72	112	29314	1897
HHHHKIKIKIKIWWWW	20797	3303	23757	9196	446	122	25	4	50	70	22294	7605
QCKIKIQINMWQ	21221	3959	8271	7432	398	139	12	1	21	28	8454	4577
HHHHKIKIQINMYYYY	22409	406	5659	3747	234	44	28	2	97	217	22687	8518
CKIKIQI	19407	1919	2433	1343	317	36	16	2	59	196	18615	2924
CKFKFQFNM	28183	3701	2224	2193	170	33	12	0	20	27	3388	573
CKIKQIINMWQ	23697	276	8317	3111	195	28	92	33	183	263	35687	4610
HHHHKIKIKIKIYYYY	20713	2171	12793	5972	147	39	20	3	72	155	10553	5882
CKFKFQF	20067	2227	5303	4282	128	22	22	5	87	199	11107	3665
CSISIQI	3087	326	10241	5508	123	6	19	3	112	366	13871	3490
RGDSGGGGGGKIKIQIC	13453	3541	10416	5402	72	5	53	44	373	836	26844	7348
CKWKWQW	6760	136	7511	924	87	10	29	18	96	211	8346	4684
KIKIQIC	9447	17	981	305	82	6	13	1	23	31	1915	358
CQFQFQF	7570	3707	7929	4815	46	23	15	5	73	186	3337	308
KFKFQFNMWQ	20677	104	735	170	64	9	26	5	599	1583	38349	5028
CKIKQII	797	2633	75	27	76	12	36	5	111	231	8372	789
KIKQIINMWQ	9363	178	616	488	39	5	199	92	398	443	15530	2319
HIHIQIC	4621	1135	261	128	33	7	86	68	748	1646	24920	8795
CHLHLQL	2934	106	1155	381	36	10	36	18	330	1136	11885	4102
KWKWQW	387	383	1545	1957	28	25	41	21	112	221	3098	2408
KIKQIIN	323	1079	64	61	12	2	16	8	52	91	601	454
CKAKAQANMWQ	600	134	9	7	6	3	480	345	858	1579	5432	2557

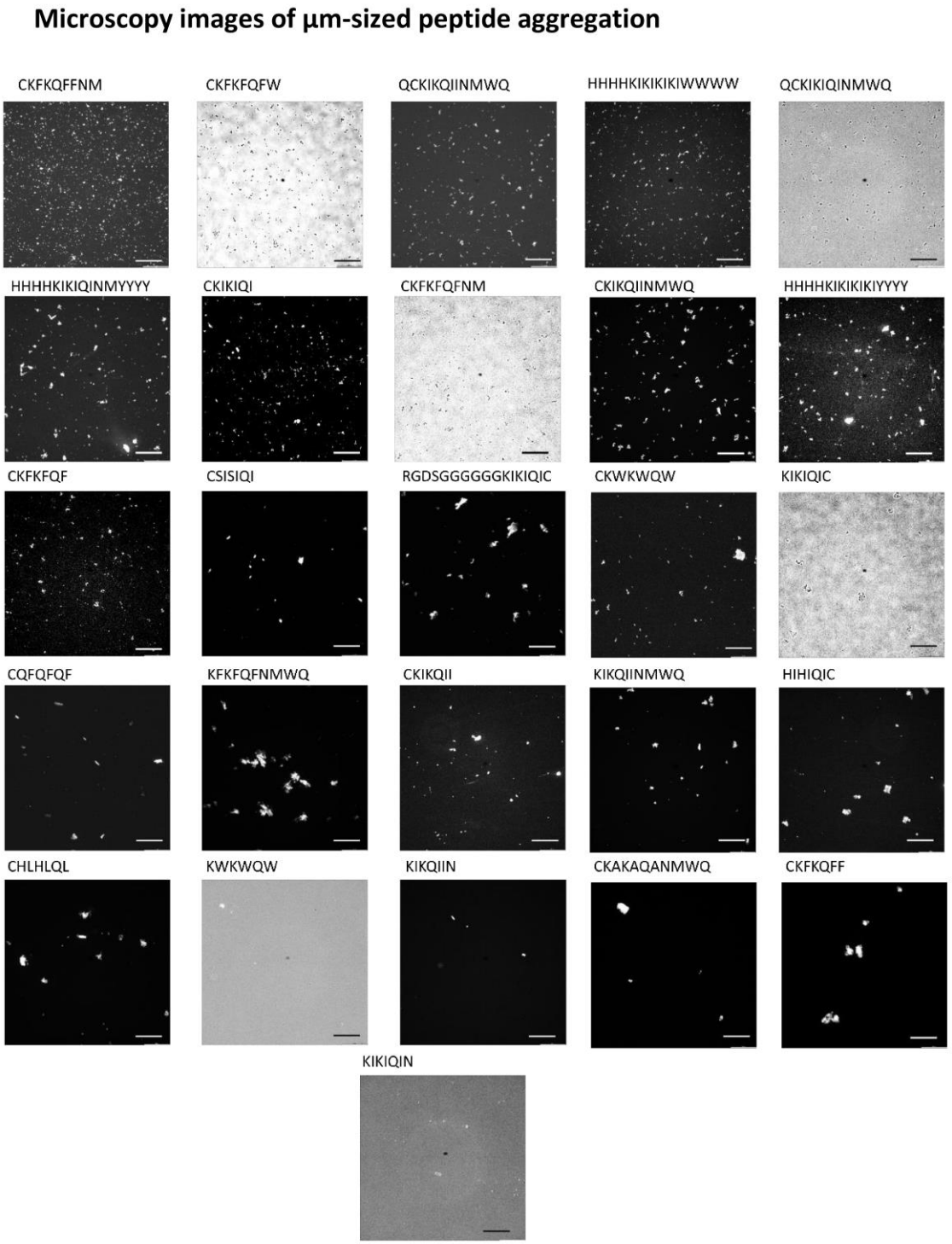


Figure S32 Representative selection of widefield microscopy images of peptides from the peptide library. Images with a dark background (white scale bar) are fluorescence microscopy images, and images with a bright background (black scale bar) are brightfield microscopy images. Peptides were diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Before the measurement they were diluted to 0.1 mg/mL with 50 µM ThT-solution in PBS, scale bar 200 µm. Microscopy measurements were conducted once with at least three microscopy images recorded for each peptide sample.

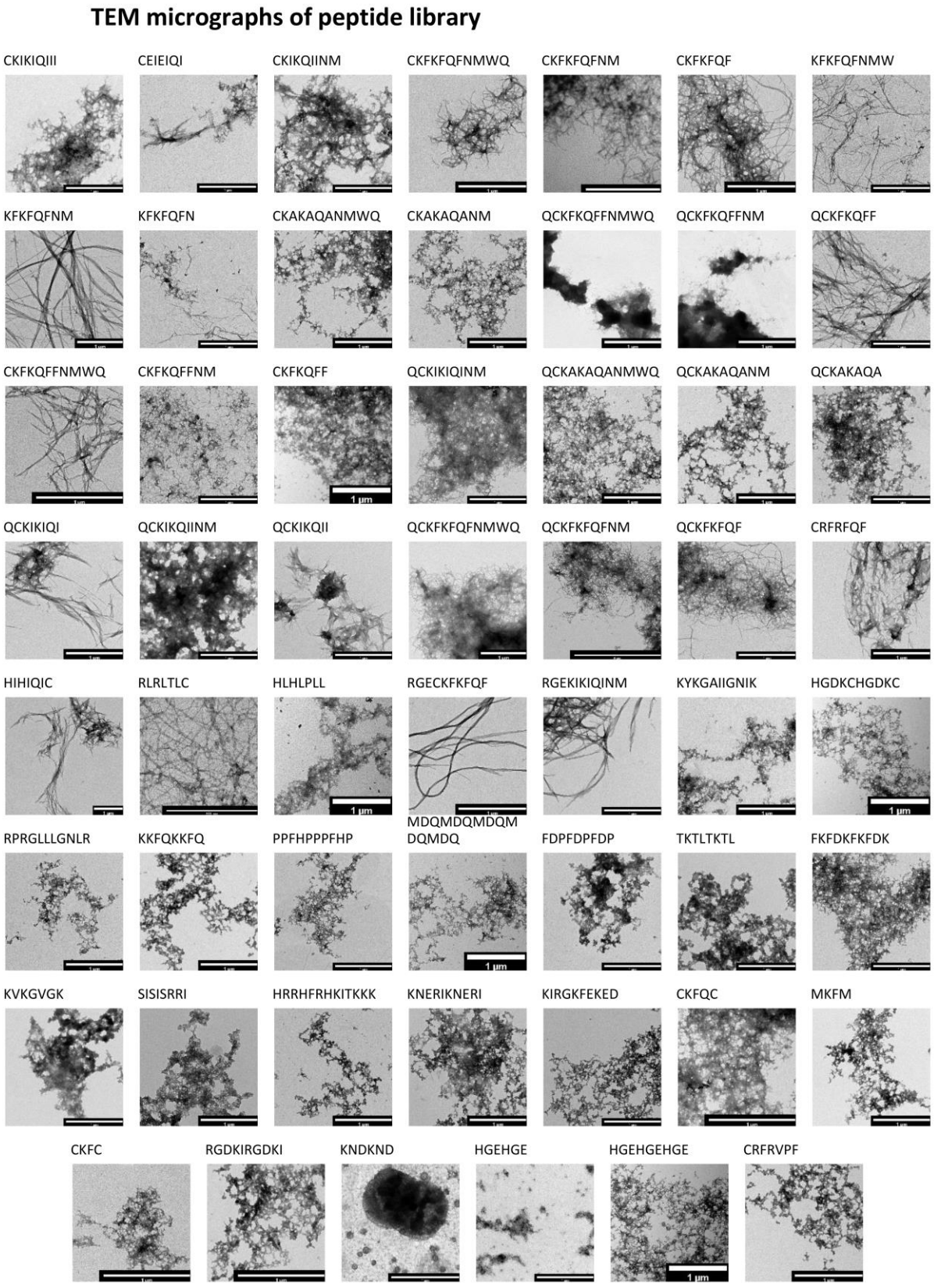


Figure S33 TEM micrographs of peptide library. Peptides were diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT, scale bar 1 µm. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

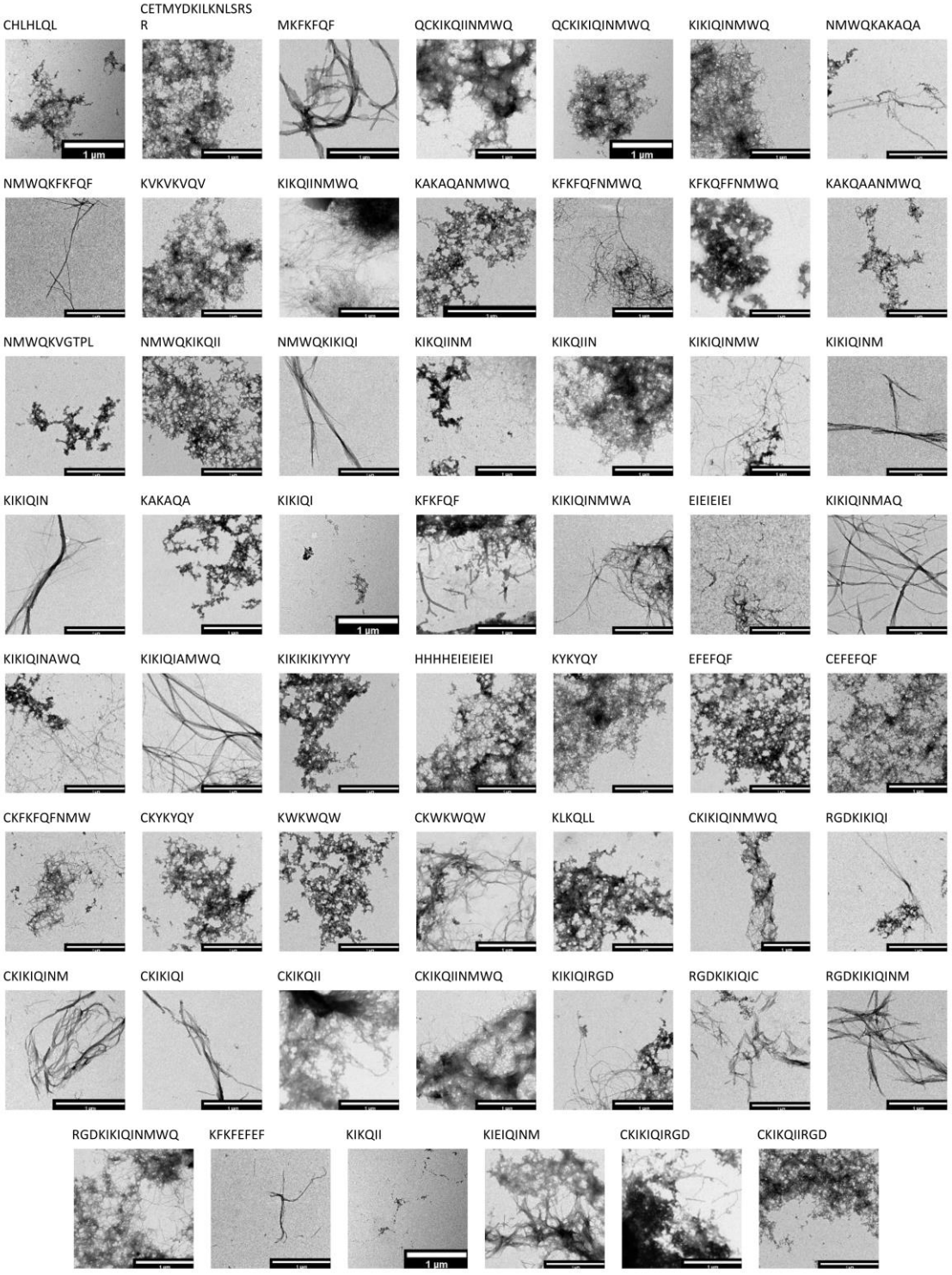


Figure S33 cont. TEM micrographs of peptide library. Peptides were diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT, scale bar 1 µm. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

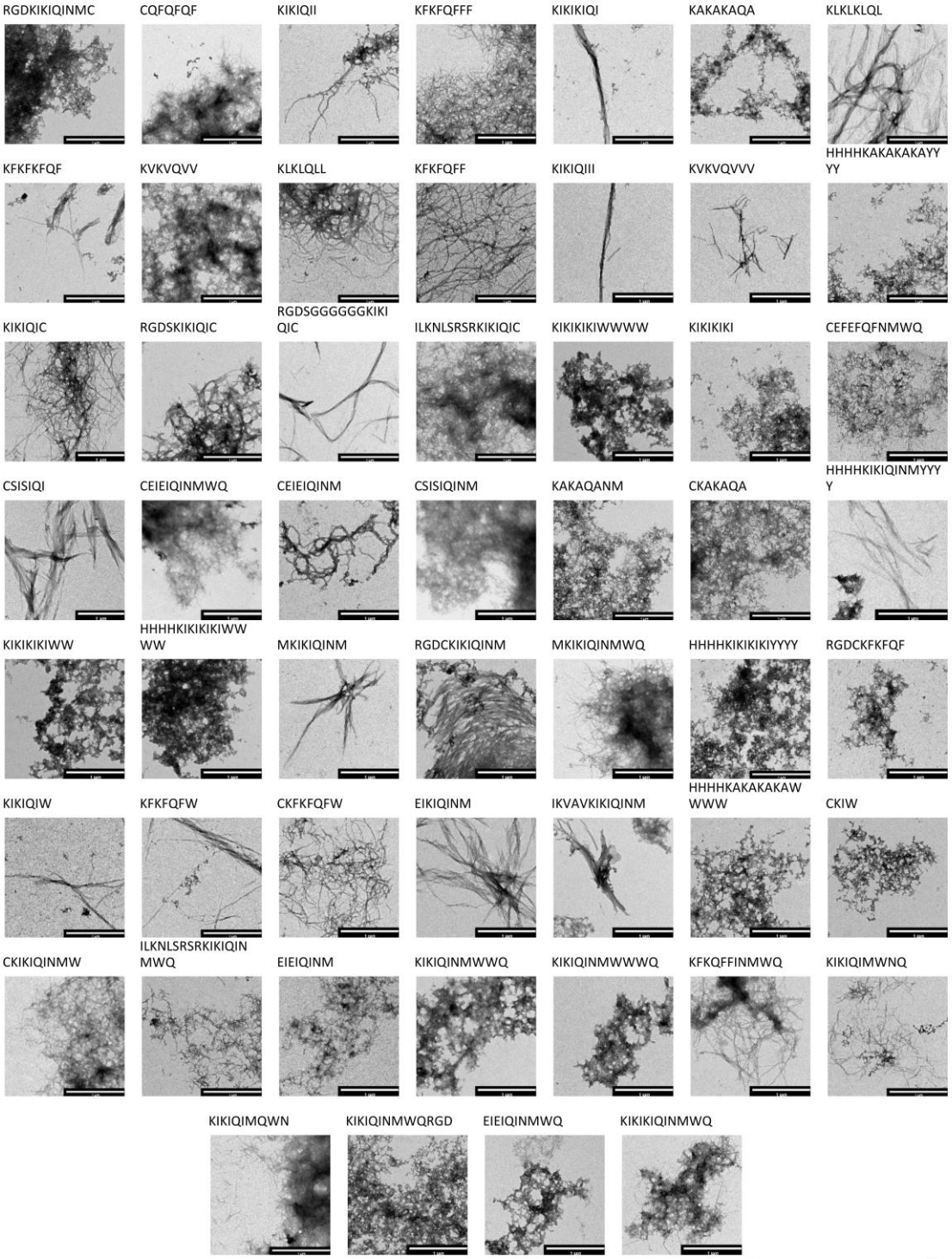


Figure S33 cont. TEM micrographs of peptide library. Peptides were diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT, scale bar 1 µm. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

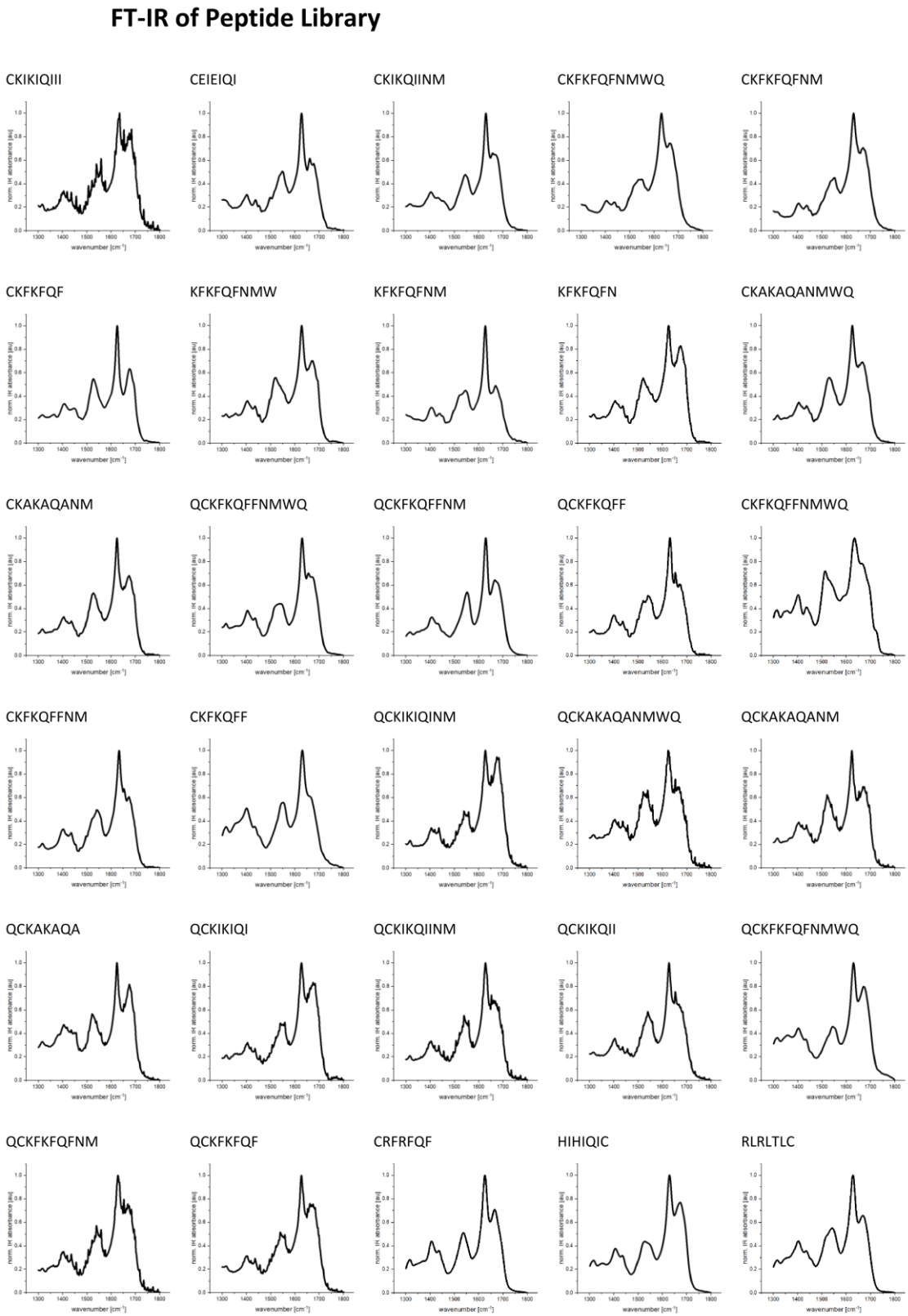


Figure S34 ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

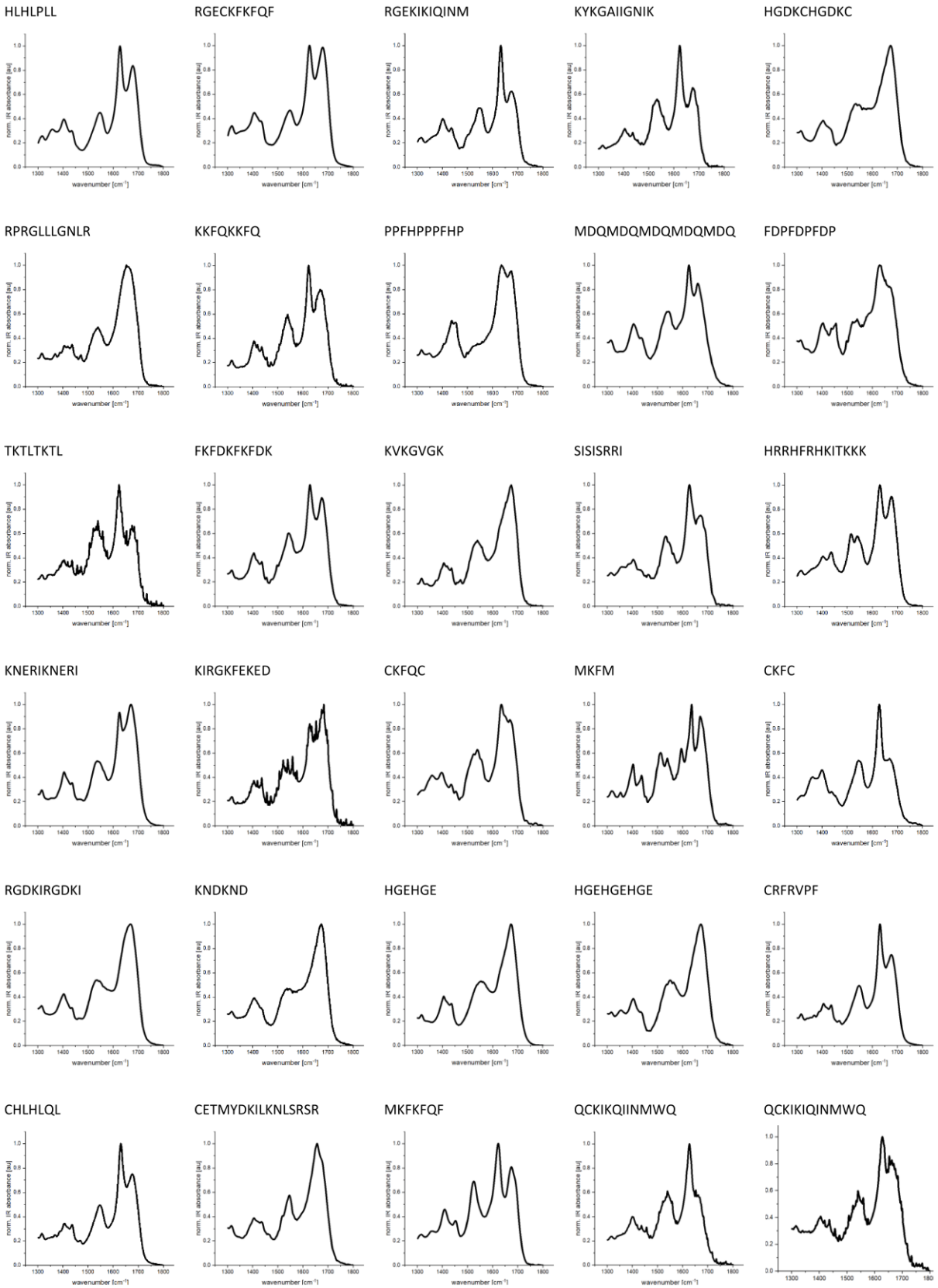


Figure S34 cont. ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

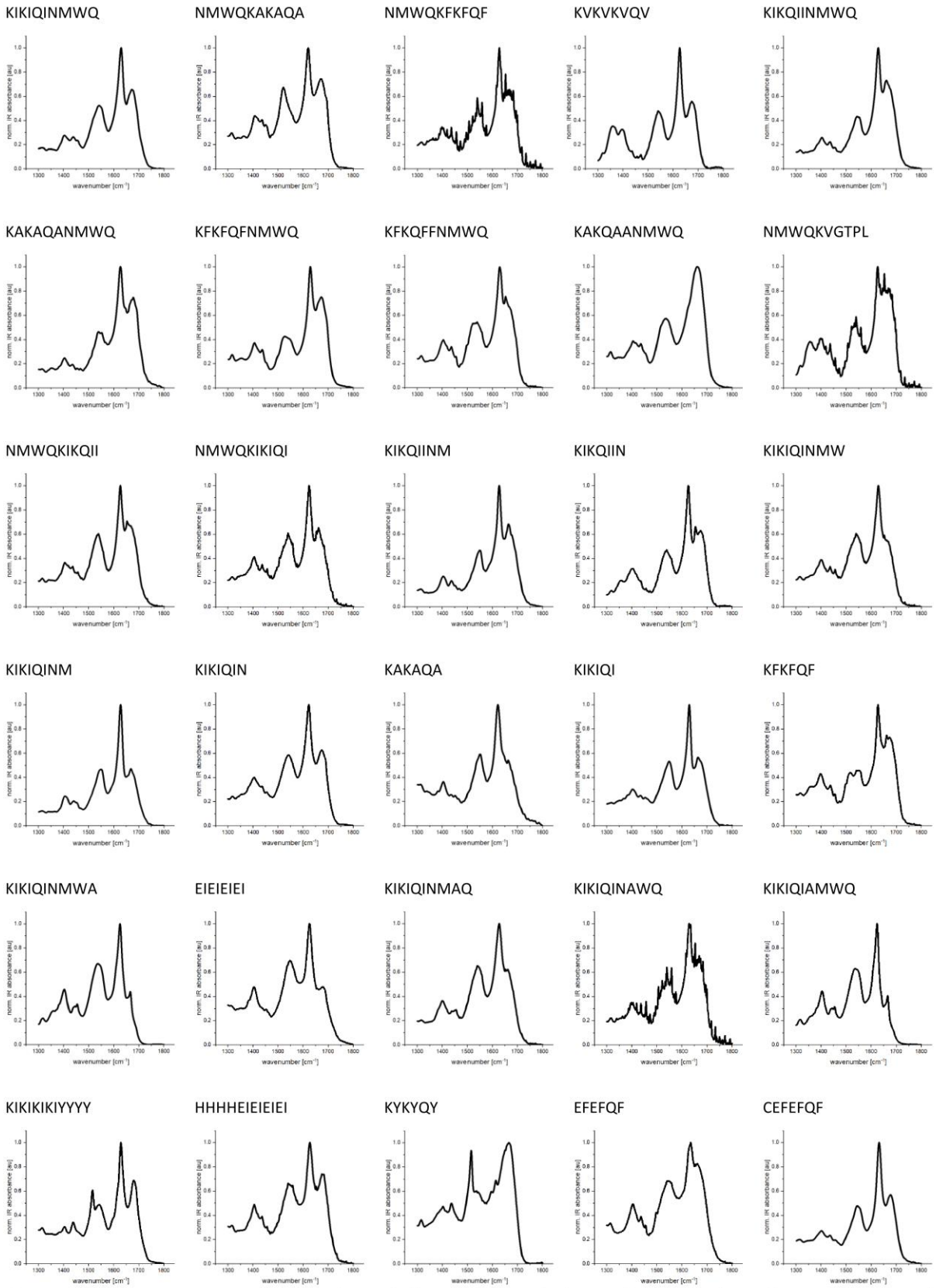


Figure S34 cont. ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

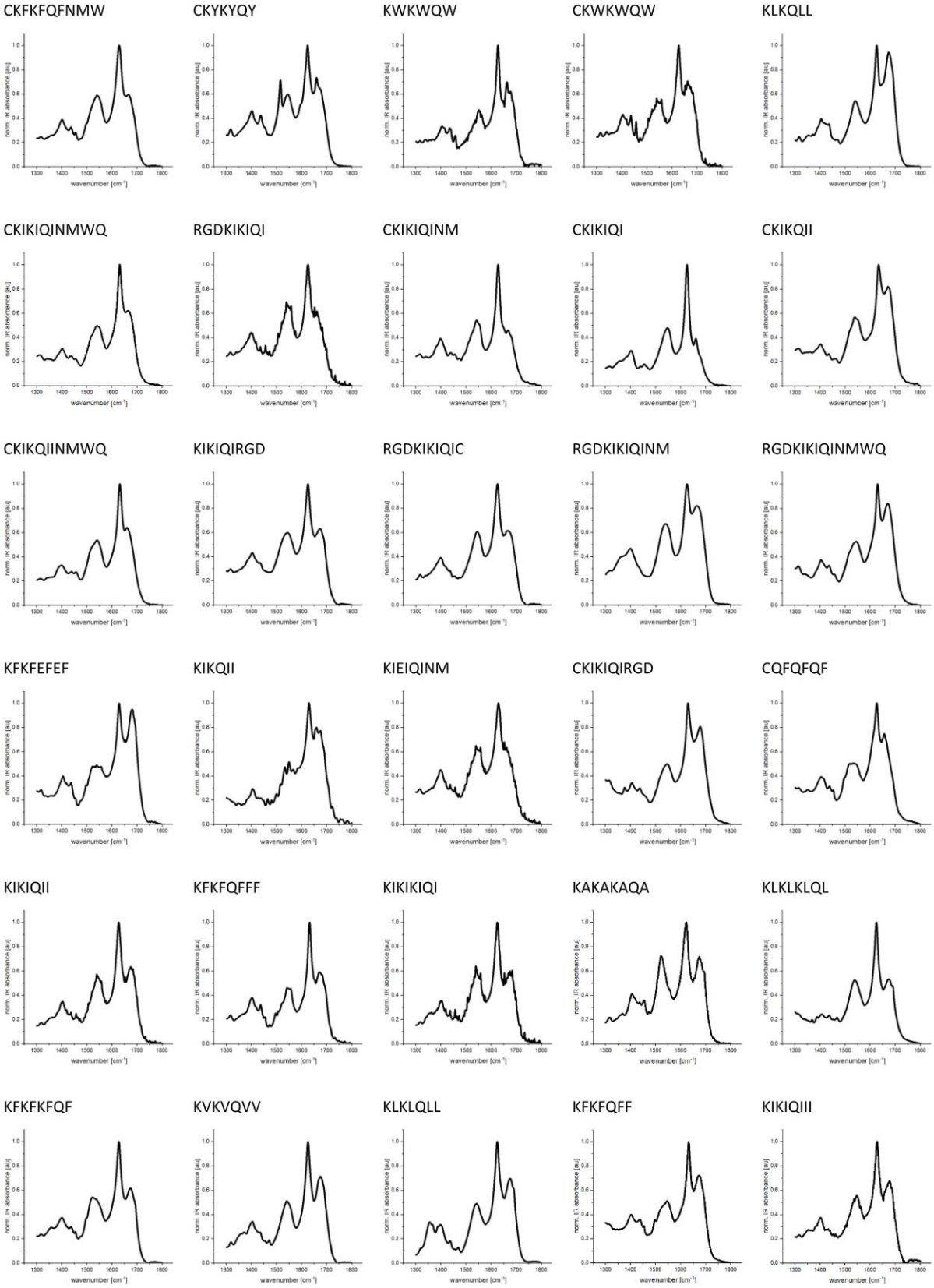


Figure S34 cont. ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

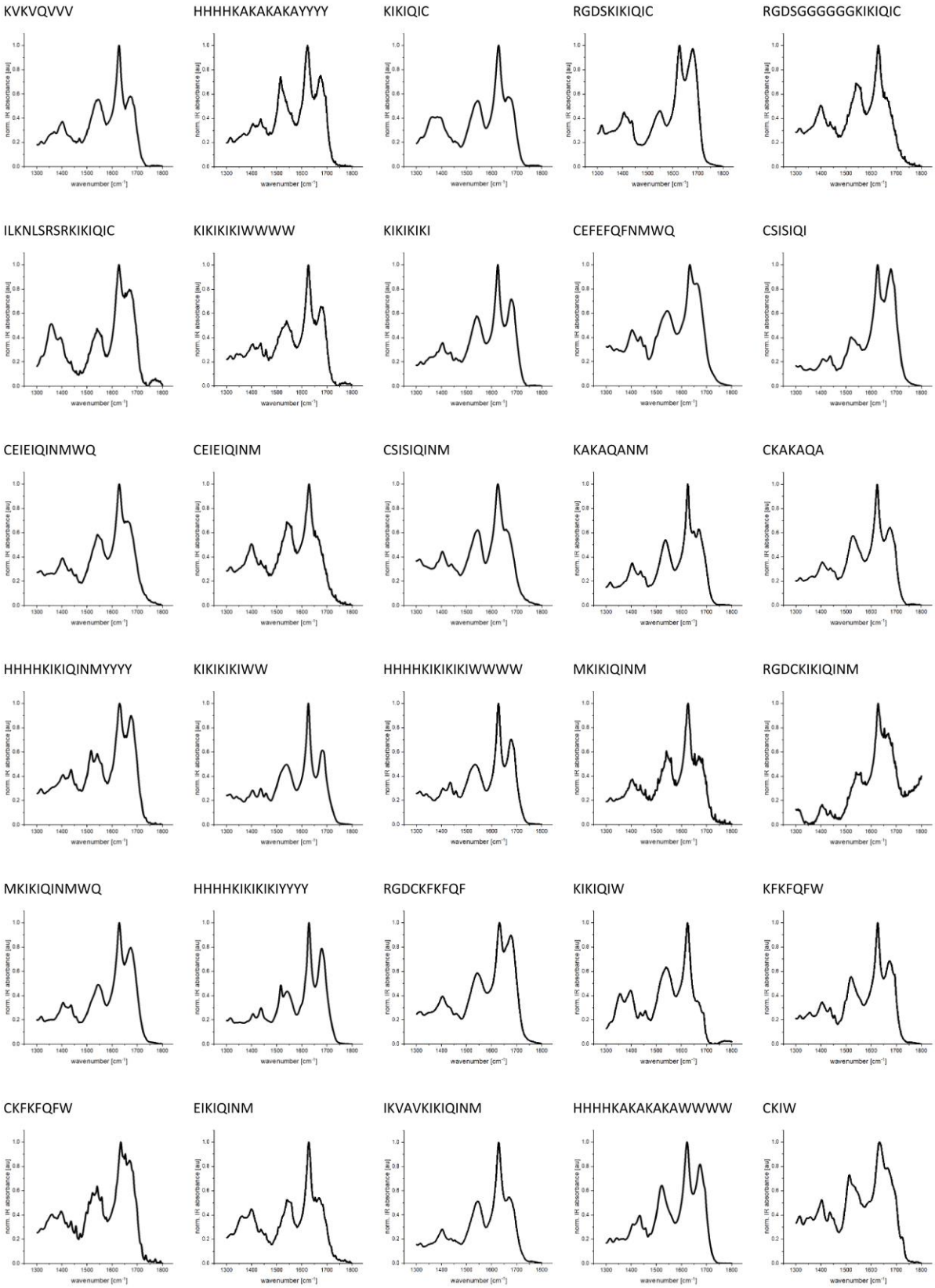


Figure S34 cont. ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

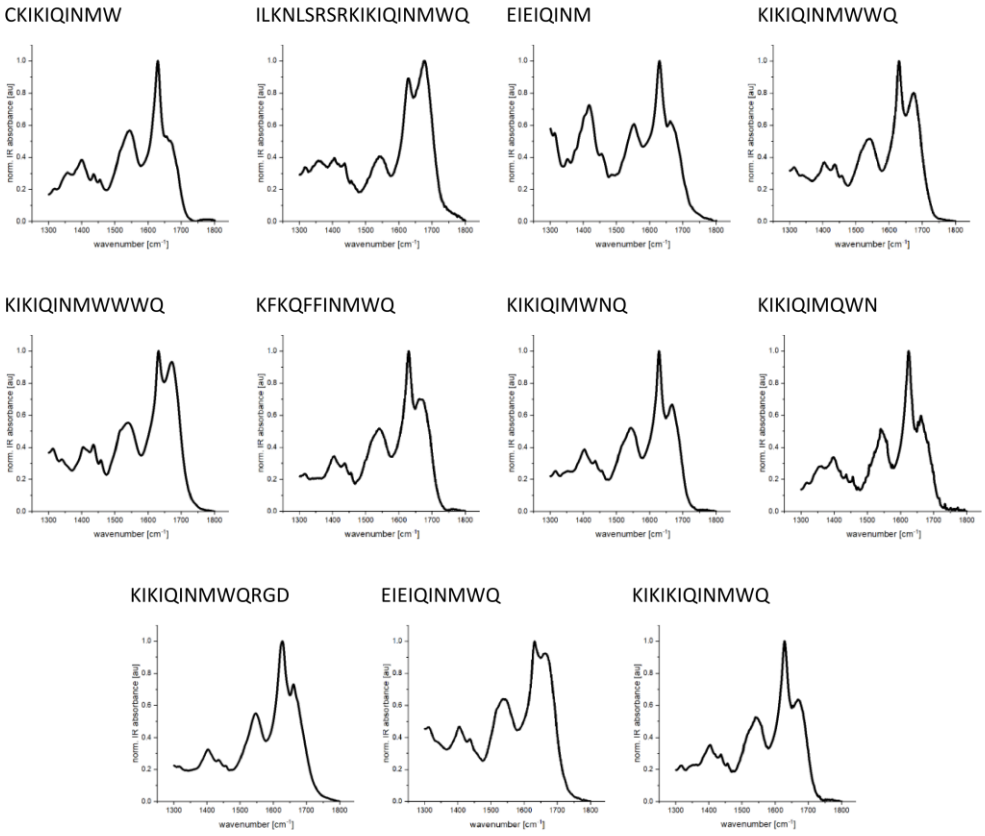


Figure S34 cont. ATR FT-IR-measurement of peptide library. Each graph shows infrared absorbance from 1300–1800 cm^{-1} normalized to the maximum of the amide I region. For the measurement peptides were lyophilized after diluted from 10 mg/mL in DMSO to 1 mg/mL in PBS and incubated for 1d at RT. Source data for Figure S34 is provided.

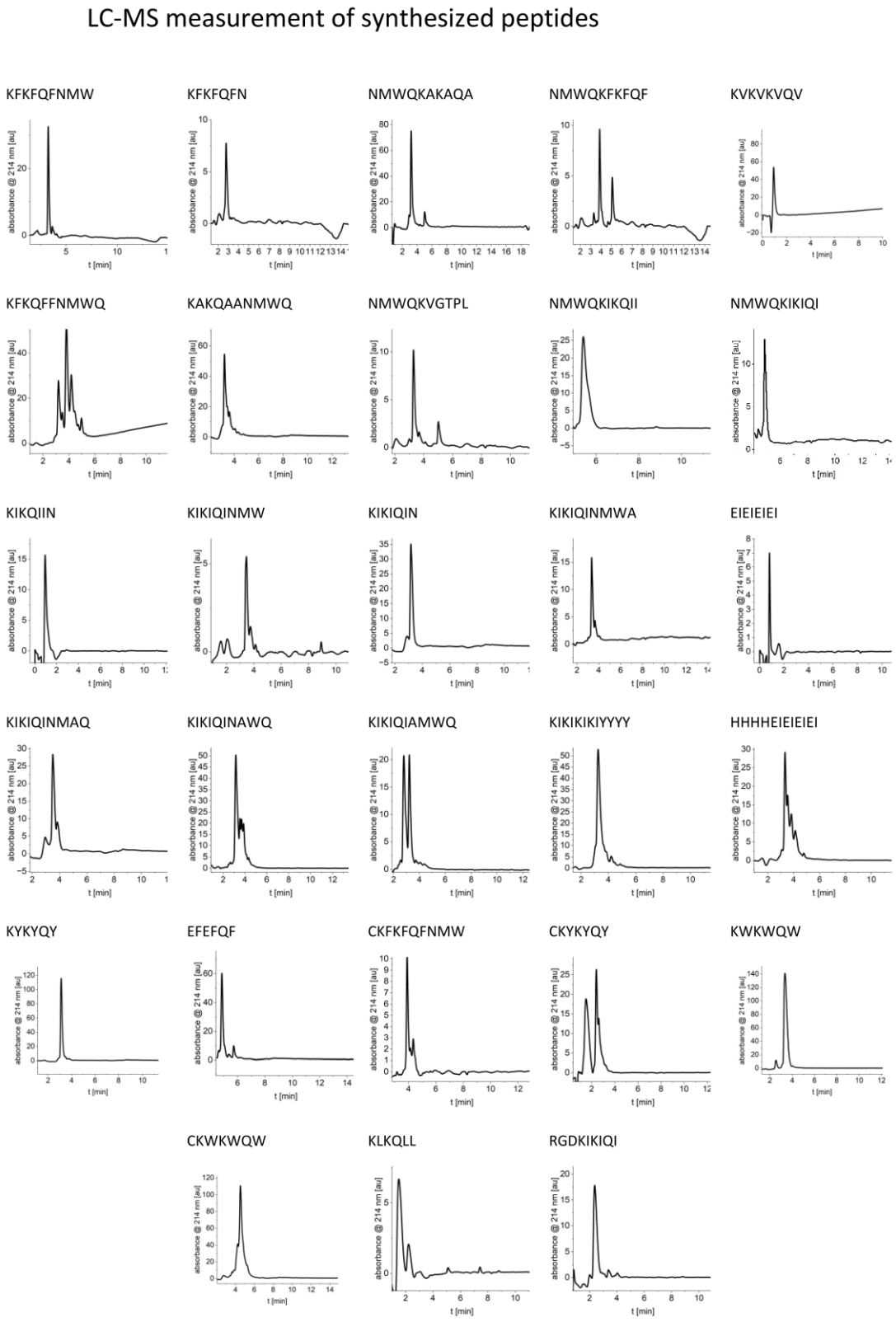


Figure S35 LC-MS chromatogram of peptides, which were not commercially obtained or reported before. Source data for Figure S35 is provided.

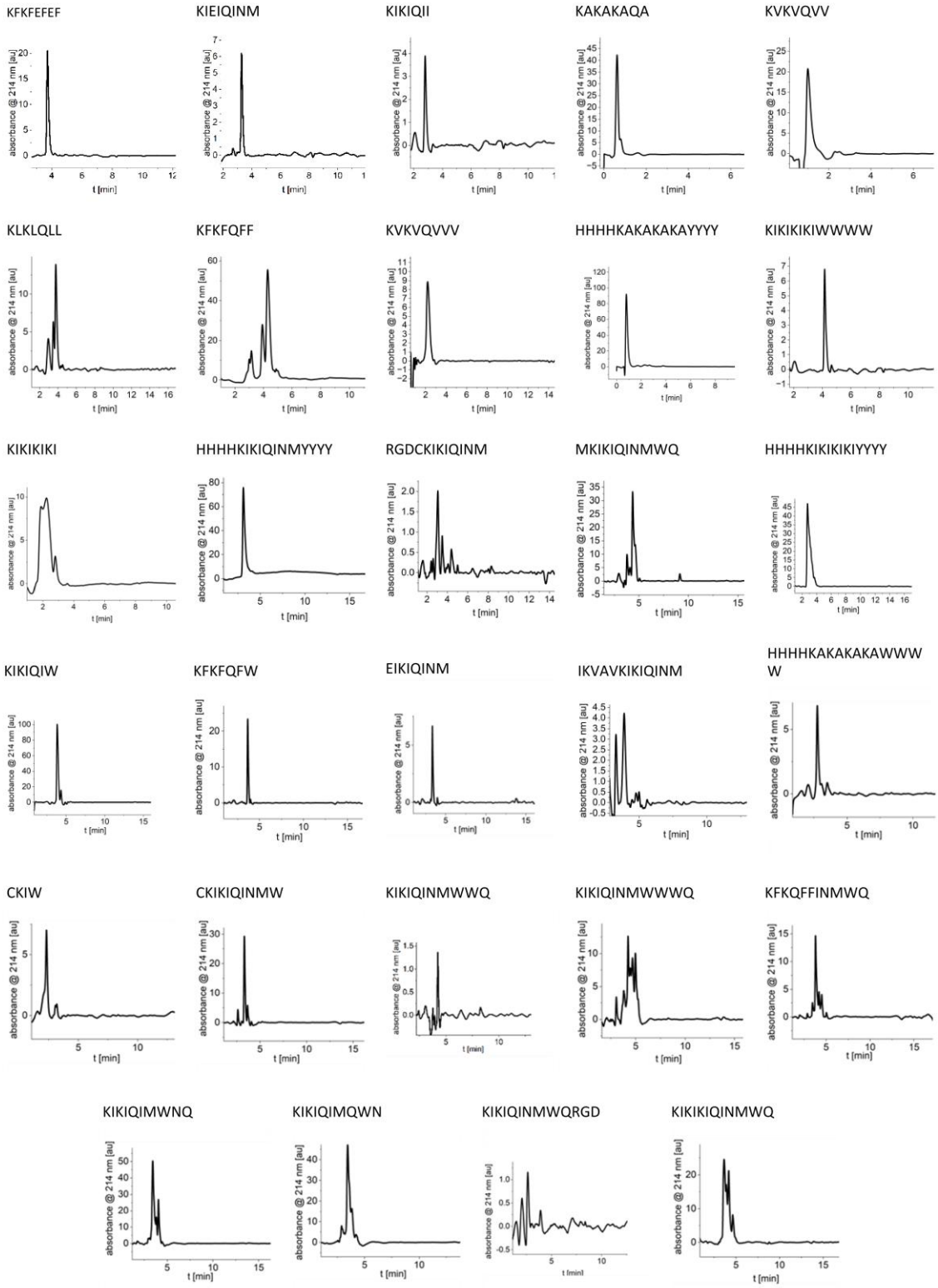


Figure S35 cont. LC-MS chromatogram of peptides, which were not commercially obtained or reported before. Source data for Figure S35 is provided.

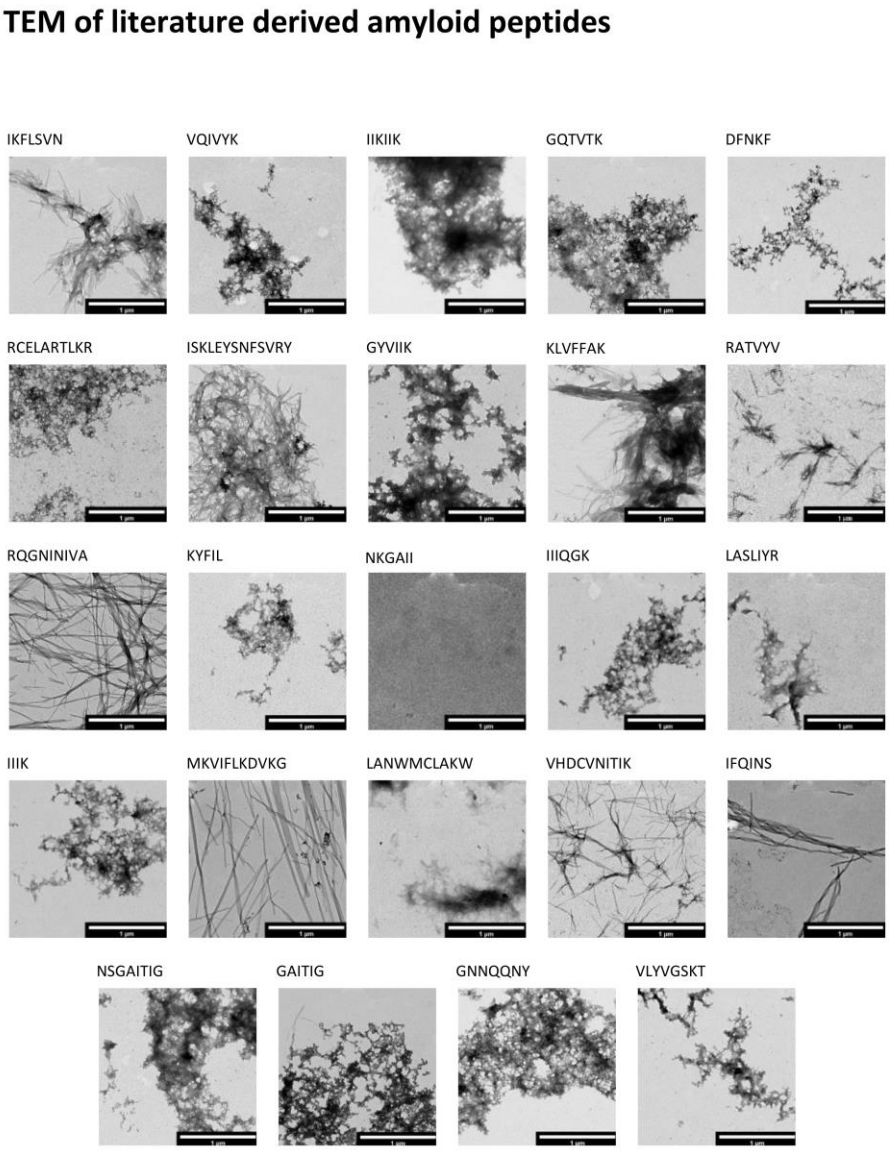


Figure S36 TEM images of literature derived peptides prepared according to standard procedure from 10 mg/mL DMSO diluted in PBS to 1mg/mL. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

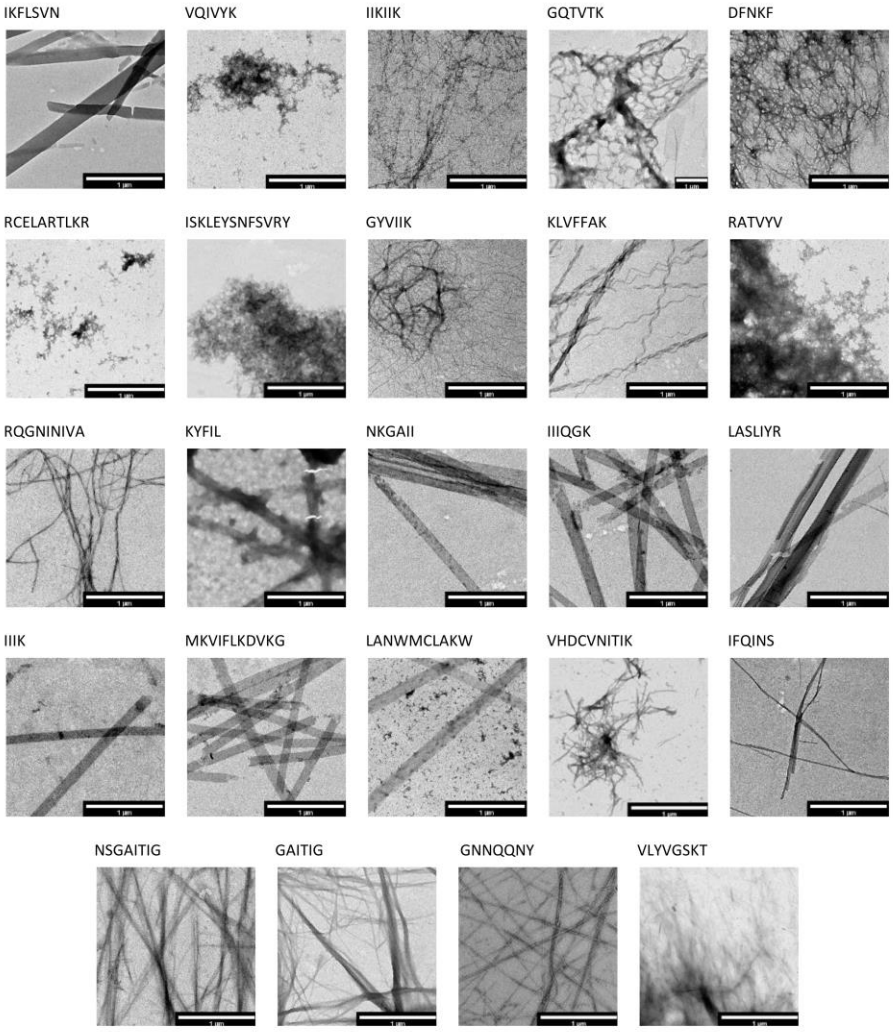


Figure S37 TEM images of literature derived peptides prepared according to literature protocol. TEM measurements were conducted once with at least three microscopy images recorded for each peptide sample.

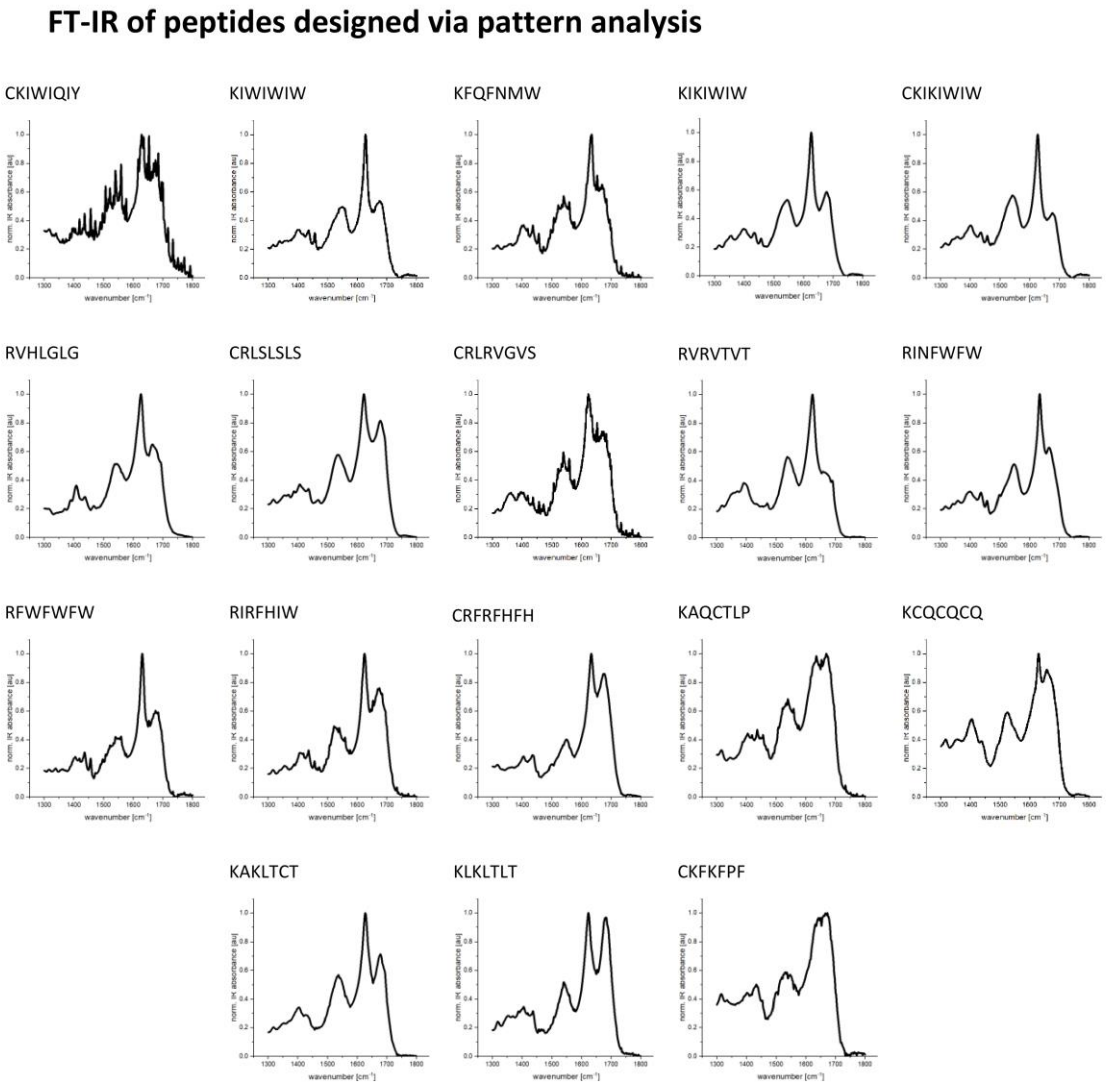


Figure S38 ATR FT-IR measurements of peptides designed via pattern analysis prepared according to standard procedure from 10 mg/mL DMSO diluted in PBS to 1 mg/mL. Source data for Figure S38 is provided.

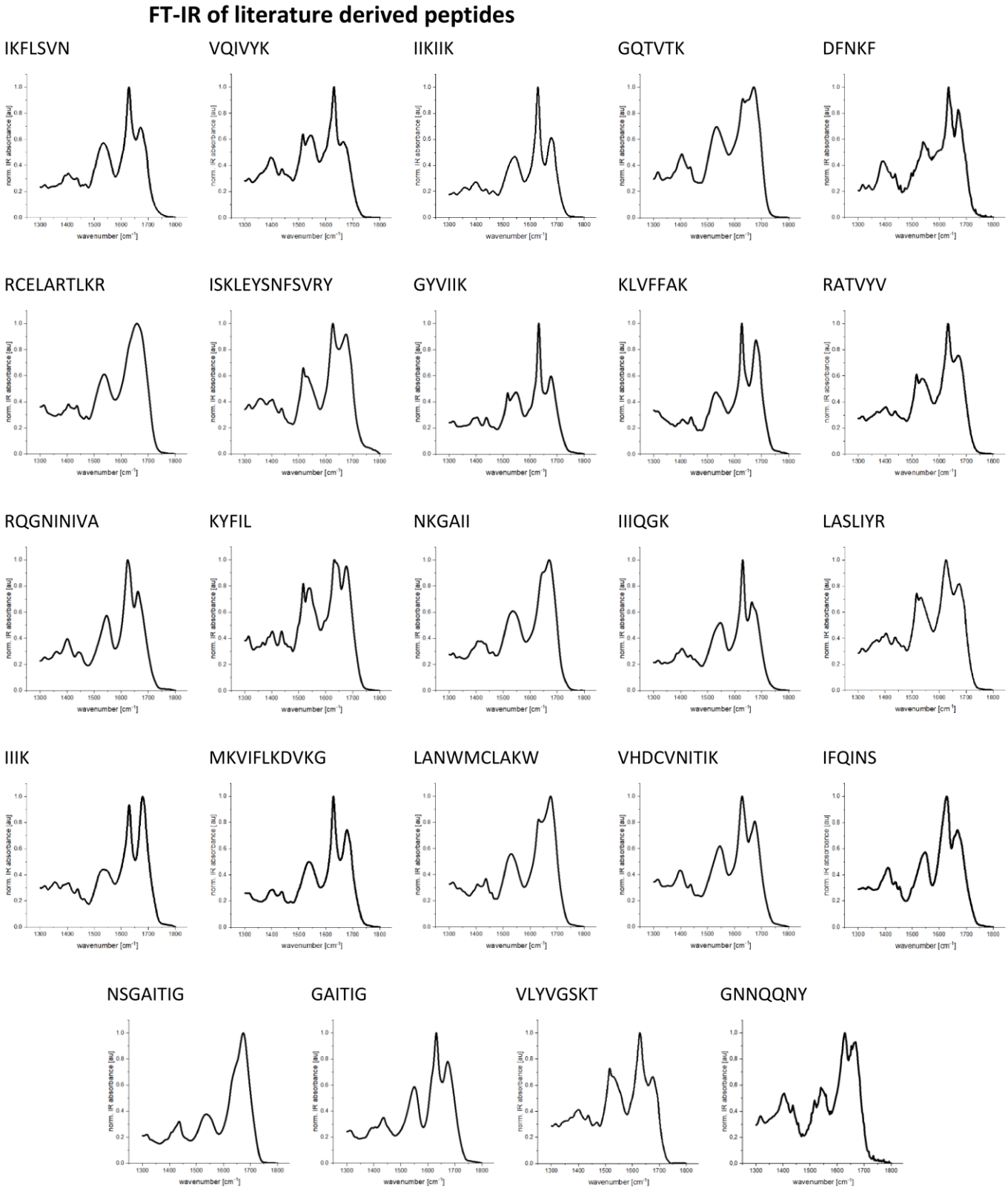


Figure S39 ATR FT-IR measurements of literature peptides prepared according to standard procedure from 10 mg/mL DMSO diluted in PBS to 1 mg/mL. Source data for Figure S39 is provided.

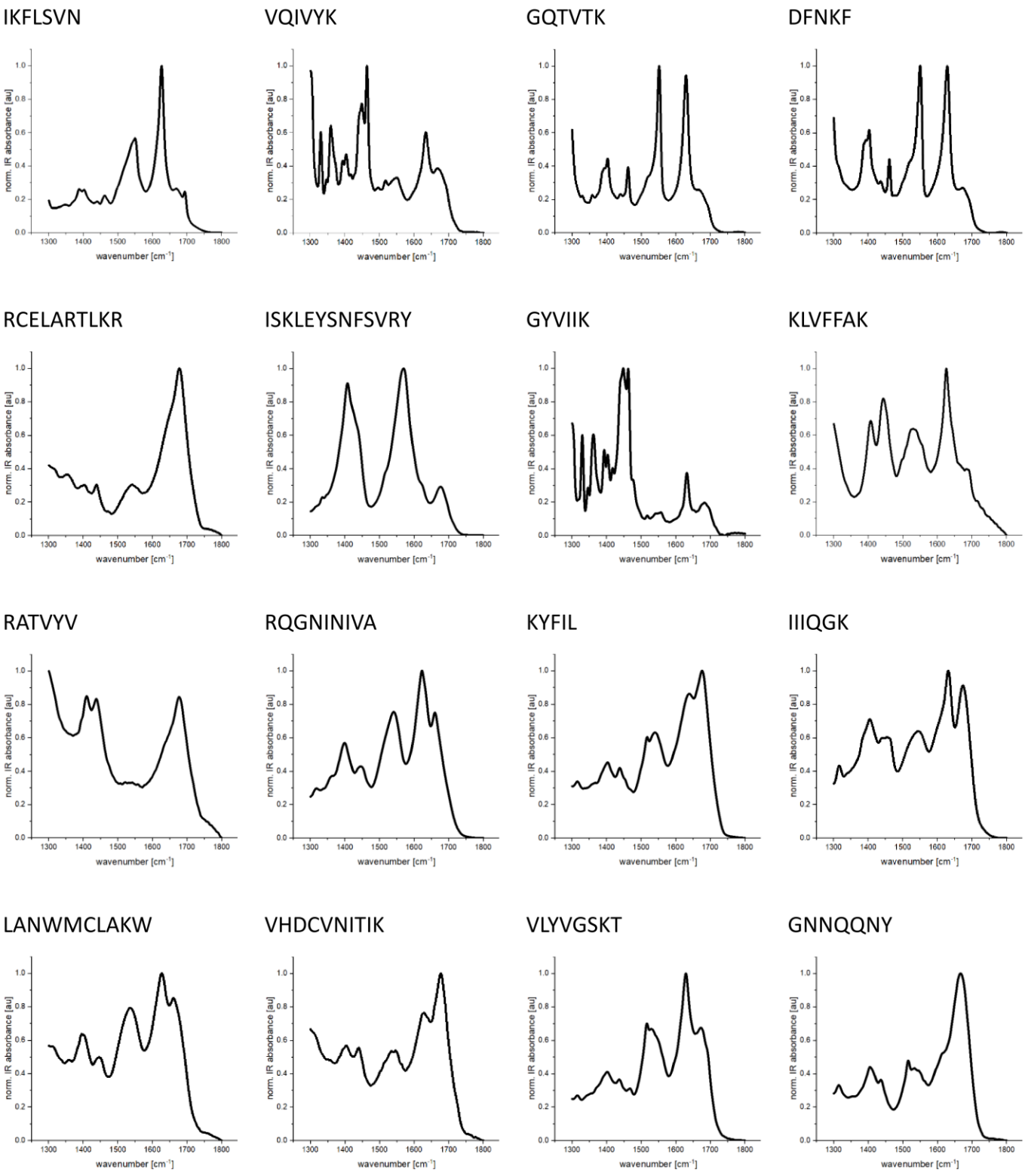


Figure S39 cont. ATR FT-IR measurements of literature peptides prepared according to literature protocol. Source data for Figure S39 is provided.

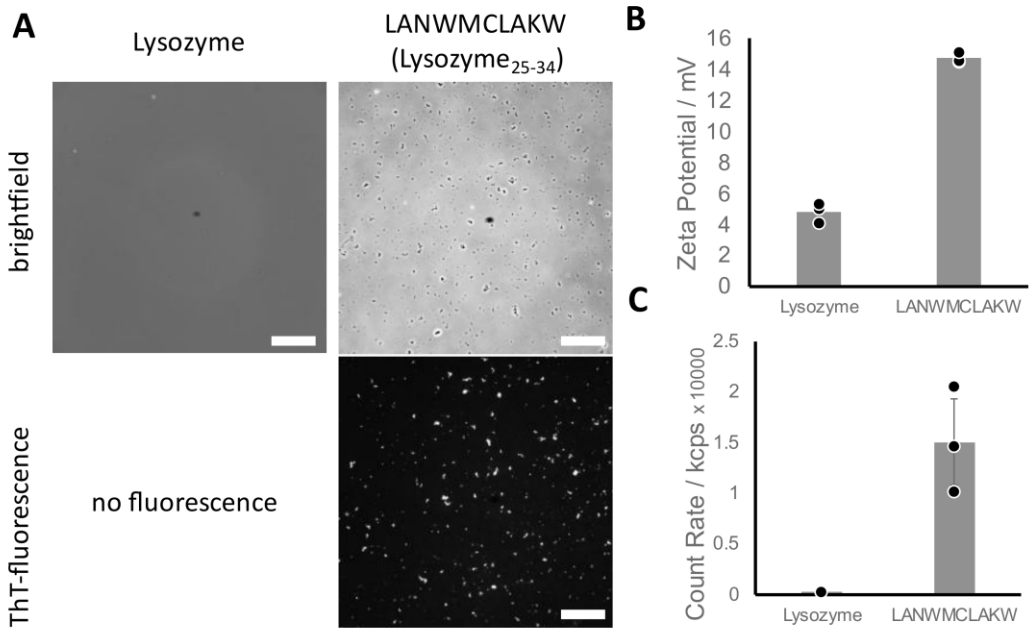


Figure S40 Comparison of physicochemical properties of the full protein lysozyme and a sequence fragment Lysozyme$_{25-34}$ LANWMCLAKW. **A** No µm-sized aggregation or ThT-fluorescence can be observed for lysozyme, while LANWMCLAKW is yielding clearly visible µm-sized aggregates in brightfield microscopy and ThT-fluorescent aggregates via fluorescence microscopy, scale bar 200 µm. Preparation of both samples at 809.7 µM (1 mg/mL LANWMCLAKW) in PBS (10 vol% DMSO), incubation for 1d at rt. Just before microscopy measurement samples were diluted to 80.9 µM with 50 µM ThT-dye solution in PBS. **B** Zeta-Potential for lysozyme and LANWMCLAKW and **C** derived Count Rate of scattered light for lysozyme and LANWMCLAKW from zeta-potential measurements. Error bars indicate Std. Dev. from triplicate measurements. Samples were incubated at 809.7 µM in PBS (10 vol% DMSO), for 1d at rt before diluting for zeta-potential measurement samples to 80.9 µM with 1 mM KCl. Mean values are displayed with error bars indicating standard deviation from triplicate measurements. Source data for Figure S40 B, C is provided.

Supplementary References

1. Yoshimura, Y. *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14446–14451 (2012).

2. Maurer-Stroh, S. *et al.* Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods* **7**, 237–242 (2010).

3. Adachi, M., So, M., Sakurai, K., Kardos, J. & Goto, Y. Supersaturation-limited and Unlimited Phase Transitions Compete to Produce the Pathway Complexity in Amyloid Fibrillation. *J. Biol. Chem.* **290**, 18134–18145 (2015).

4. Ke, P. C. *et al.* Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **49**, 5473–5509 (2020).

5. Biancalana, M., Makabe, K., Koide, A. & Koide, S. Molecular Mechanism of Thioflavin-T Binding to the Surface of β-Rich Peptide Self-Assemblies. *J. Mol. Biol.* **385**, 1052–1063 (2009).

6. Naiki, H., Higuchi, K., Hosokawa, M. & Takeda, T. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine T. *Anal. Biochem.* **177**, 244–249 (1989).

7. Sidhu, A., Vaneyck, J., Blum, C., Segers-Nolten, I. & Subramaniam, V. Polymorph-specific distribution of binding sites determines thioflavin-T fluorescence intensity in α-synuclein fibrils. *Amyloid* **25**, 189–196 (2018).

8. Arad, E., Green, H., Jelinek, R. & Rapaport, H. Revisiting thioflavin T (ThT) fluorescence as a marker of protein fibrillation – The prominent role of electrostatic interactions. *J. Colloid Interface Sci.* **573**, 87–95 (2020).

9. Namioka, S., Yoshida, N., Konno, H. & Makabe, K. Residue-Specific Binding Mechanisms of Thioflavin T to a Surface of Flat β-Sheets within a Peptide Self-Assembly Mimic. *Biochemistry* **59**, 2782–2787 (2020).

10. Zhao, R. *et al.* Measurement of amyloid formation by turbidity assay—seeing through the cloud. *Biophys. Rev.* **8**, 445–471 (2016).

11. Pignataro, M. F., Herrera, M. G. & Dodero, V. I. Evaluation of Peptide/Protein Self-Assembly and Aggregation by Spectroscopic Methods. *Molecules* **25**, 4854 (2020).

12. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8**, 409–427 (2016).

13. Malm, A. V. & Corbett, J. C. W. Improved Dynamic Light Scattering using an adaptive and statistically driven time resolved treatment of correlation data. *Sci. Rep.* **9**, 1–11 (2019).

14. Martinek, T. A. *et al.* Secondary Structure Dependent Self-Assembly of β-Peptides into Nanosized Fibrils and Membranes. *Angew. Chemie* **118**, 2456–2460 (2006).

15. Toksoz, S., Mammadov, R., Tekinay, A. B. & Guler, M. O. Electrostatic effects on nanofiber formation of self-assembling peptide amphiphiles. *J. Colloid Interface Sci.* **356**, 131–137 (2011).

16. Chen, Y., Gan, H. X. & Tong, Y. W. pH-controlled hierarchical self-assembly of peptide amphiphile. *Macromolecules* **48**, 2647–2653 (2015).

17. Fauchère, J.-L., Charton, M., Kier, L. B., Verloop, A. & Pliska, V. Amino acid side chain parameters for correlation studies in biology and pharmacology. *Int. J. Pept. Protein Res.* **32**, 269–278 (2009).

18. Osorio, D., Rondón-Villarreal, P. & Torres, R. Peptides: A Package for Data Mining of Antimicrobial Peptides. *R J.* **7**, 4 (2015).

19. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).

20. Ikai, A. Thermostability and Aliphatic Index of Globular Proteins. *J. Biochem.* **88**, 1895–1898 (1980).

21. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature* **299**, 371–374 (1982).

22. Boman, H. G. Antibacterial peptides: basic facts and emerging concepts. *J. Intern. Med.* **254**, 197–215 (2003).

23. Guruprasad, K., Reddy, B. V. B. & Pandit, M. W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng. Des. Sel.* **4**, 155–161 (1990).

24. Sheik, D. A., Dewhurst, S. & Yang, J. Natural Seminal Amyloids as Targets for Development of Synthetic Inhibitors of HIV Transmission. *Acc. Chem. Res.* **50**, 2159–2166 (2017).

25. Lee, Y.-H. & Ramamoorthy, A. Semen-derived amyloidogenic peptides-Key players of HIV infection. *Protein Sci.* **27**, 1151–1165 (2018).

26. Tan, S. *et al.* Peptides derived from HIV-1 gp120 co-receptor binding domain form amyloid fibrils and enhance HIV-1 infection. *FEBS Lett.* **588**, 1515–1522 (2014).

27. Hauber, I., Hohenberg, H., Holstermann, B., Hunstein, W. & Hauber, J. The main green tea polyphenol epigallocatechin-3-gallate counteracts semen-mediated enhancement of HIV infection. *Proc. Natl. Acad. Sci.* **106**, 9033–9038 (2009).

28. Viswanathan, G. *et al.* Inhibitory Effect of Naphthoquinone-Tryptophan Hybrid towards Aggregation of PAP f39 Semen Amyloid. *Molecules* **23**, 3279 (2018).

29. Schilling, C. *et al.* Sequence-Optimized Peptide Nanofibers as Growth Stimulators for Regeneration of Peripheral Neurons. *Adv. Funct. Mater.* **29**, 1809112 (2019).

30. Kaygisiz, K. *et al.* Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides. *Zenodo.org* doi:10.5281/ZENODO.8079727 (2023).

31. Mendes, A. C., Baran, E. T., Reis, R. L. & Azevedo, H. S. Self-assembly in nature: using the principles of nature to create complex nanobiomaterials. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **5**, 582–612 (2013).

32. Fujiwara, K., Toda, H. & Ikeguchi, M. Dependence of alpha-helical and beta-sheet amino acid propensities on the overall protein fold type. *BMC Struct. Biol.* **12**, 18 (2012).

33. Dai, B. *et al.* Tunable assembly of amyloid-forming peptides into nanosheets as a retrovirus carrier. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 2996–3001 (2015).

34. Medini, K. *et al.* Controlling gelation with sequence: Towards programmable peptide

52. Kurnellas, M. P., Adams, C. M., Sobel, R. A., Steinman, L. & Rothbard, J. B. Amyloid fibrils composed of hexameric peptides attenuate neuroinflammation. *Sci. Transl. Med.* **5**, (2013).

53. Louros, N. *et al.* WALTZ-DB 2.0: an updated database containing structural information of experimentally determined amyloid-forming peptides. *Nucleic Acids Res.* **48**, D389–D393 (2020).

54. Luisi, D. L., Wu, W.-J. & Raleigh, D. P. Conformational analysis of a set of peptides corresponding to the entire primary sequence of the N-terminal domain of the ribosomal protein L9: evidence for stable native-like secondary structure in the unfolded state. *J. Mol. Biol.* **287**, 395–407 (1999).

55. Sieste, S. *et al.* Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction. *Adv. Funct. Mater.* **31**, 2009382 (2021).

6.4 Unraveling Structure-Property Relationships in Self-Assembling Peptide Nanofibrils via Simulations

Authors: Aysenur Iscen*, Kübra Kaygisiz, Christopher V. Synatschke, Tanja Weil, Kurt Kremer*
* corresponding author
Manuscript to be submitted.
Author Contributions:
Aysenur Iscen conceptualized, designed and performed all computational work, and wrote the manuscript. Kübra Kaygisiz conceptualized and performed all experimental work, and revised the manuscript. Christopher Synatschke revised the manuscript. Tanja Weil conceived the idea, supervised, acquired funding for the project, and revised the manuscript. Kurt Kremer conceptualized, supervised, acquired funding for the project, and revised the manuscript.

Copyrights: The authors hold copyright for the following part that is reproduced from an unpublished manuscript to be submitted.

Unraveling structure-property relationships in self-assembling peptide nanofibrils via simulations

Aysenur Iscen,* Kübra Kaygisiz, Christopher V. Synatschke, Tanja Weil, and Kurt Kremer*

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

E-mail: iscena@mpip-mainz.mpg.de; kremer@mpip-mainz.mpg.de

Abstract

Assemblies of peptides and proteins through specific intermolecular interactions set the basis for macroscopic materials found in Nature. Many researchers have taken advantage of easily tunable hydrogen bonding interactions, which make highly stable β-sheets, to develop supramolecular peptide nanofibrils (PNFs), that can mimic biological structures, such as the extracellular matrix, in their ability to act as scaffolds, provide stability, transport materials, and signal cells. We previously reported a model system of PNFs, in which we systematically investigated material properties of biologically active peptides. Inspired by our earlier work, here we provide a microscopic view of the self-assembled PNFs and investigate how different modifications, such as sequence truncation, amino acid substitution and sequence re-ordering, affect the secondary structure and dynamic properties of peptides in supramolecular fibrils to determine the relationship between structure, morphology and biological activity. While a certain peptide-peptide aggregation propensity is required to form bioactive β-sheet rich structures, we show in contrast to previous studies that very high peptide aggregation propensity can cause instabilities when fibers are in close proximity to

other fibers in solution, which in turn decreases bioactivity. Also, the dynamic behavior of the individual PNFs is different than aggregated ones, which is why we propose that fiber-fiber interaction should be taken into account in simulations for structure-property-activity correlation. Our multi-scale simulation approach enables to trace changes in macroscopic aggregation behavior back to single amino acid modifications. Understanding these relationships can lead to efficient materials discoveries and help in studying naturally occurring amyloid fiber aggregation in neurodegenerative diseases.

Introduction

In Nature, formation of hierarchical, complex structures from the assembly of simple building blocks can be frequently observed in functional and pathogenic contexts. Cell cytoskeleton and extracellular matrix are two examples in which various protein assemblies form intricate networks inside and outside the cells, respectively, to ensure vital functions such as providing stability, motility, aid division and transport of materials.[1] Another example of self-assembly observed in biology, amyloid fibril formation, is a specific and ordered aggregation process that is observed in the extracellular spaces and is associated with several lethal diseases, such as Alzheimer's and Parkinson's.[2] In both cases, the characteristic function of the β-sheet rich protein fibrils is determined by their high rigidity, tensile strength, and enhanced resistance towards biodegradation. Inspired by the stability of amyloid fibrils many researchers in the field of materials science and supramolecular chemistry have designed amyloid-type biomaterials to act as scaffolds for applications in different fields, such as biomedicine,[3–5] soft robotics,[6] catalysis,[7] sensing[8] and many others.[9–12]

Amyloid fibrils as functional biomaterials

We and others previously reported a strong dependency of structural factors, such as β-sheet formation, on biological activity. For example, PNFs from self-assembling short peptide sequences ($\sim$ 10 amino acids) with varying amino acid sequences have been used in neuron

regeneration applications by directly injecting into tissues of living mice with facial whisker nerve injuries.[13] Peptide nanofibers attached to the injured nerve cells and improved recovery of the whisker movement, while peptides with higher β-sheet content performed better than those with less β-sheet content. Similarly, another study showed that self-assembling peptides are effective in promoting corticospinal axon regrowth after acute spinal cord injury in mice.[14]

Furthermore, PNFs with high β-sheet content cannot only interact efficiently with cells but also with viruses which have been exploited for gene-therapeutic applications, e.g by enhancing the retroviral gene transfer.[4,15,16] The affinity of amyloid-type peptides towards cells and viruses is attributed to the fact that positively charged amino acids (i.e. lysine) can interact with the negatively charged cell membranes and viral membranes, bringing them in close proximity and increasing the chances of gene transfer.

In other target applications, it was found that the lysine groups on the PNFs bind carbon dioxide molecules from the environment, which makes them efficient, biodegradable and affordable carbon capture materials.[16,17] While experiments have shown that the specific position of the lysine groups significantly affect the carbon capture capability of the fibers, the findings lack clear conclusions about why some fibers perform better than others even though they have similar lysine contents. Depending on the intended application for PNFs, there may be different mechanisms at play, making it more complex to make clear distinctions between structure and properties. This raises a need for strategies that blend advanced computational methods with experimental techniques to that provide more detailed insights.

Molecular simulations of amyloid fibrils

For many decades, researchers have tried to answer the question of why some of these proteins in our cells, which are normally soluble under physiological conditions, assemble into an ordered fibrillar structure with defined β-sheets that persist, accumulate and macroscopically aggregate to pathogenic amyloid "plaques."

Synthetic short self-assembling peptides allow an experimentally direct way to study the

impact of amino acid sequences on the formation of β-sheet rich structures.[18,19] While model peptides allow a systematic quantification of changes in structure properties, experimental methods lack the possibility to mechanistically unveil the assembly process due to the multiple scale dependencies.[20]

The self-assembly of amyloid-like peptide nanofibrils relies on intermolecular interactions, such as van der Waals, hydrogen bonding, π-stacking, electrostatic and hydrophobic interactions. Amino acids are commonly grouped into four categories, i.e. polar, nonpolar, acidic or basic, based on the structure of their side chains and small differences in structure affect their hydrophobicities[21] and propensities for making secondary structures.[22] These two factors then contribute to the resulting morphology of the supramolecular polymer. However, determining the morphology of resulting supramolecular polymer given a peptide a sequence, i.e. whether it is a micelle, fibril, or vesicle, is not straightforward. As shown in previous studies, different ordering of the same amino acid sequences produces different morphologies, i.e. polymorphism, when self-assembled in solution.[15,23]

Molecular simulations are a powerful tool to shed light on mechanisms that govern peptide self-assembly at the experimentally inaccessible time and length scales. Previously, the structure of self-assembled peptide nanofibers composed of specific peptide sequences experimentally shown to form β-sheet networks have been studied with molecular dynamics simulations using atomistic models.[24,25] In addition, coarse-grained models have also investigated the self-assembly mechanisms of peptides into nanofibers.[26–29] Because each of these studies considers only a few specific peptide sequences, it is difficult to make generalized conclusions about other peptide nanofibers. In order to explore the potential space of amino acid sequences for pentapeptide aggregation, a machine learning-based strategy was utilized where an AI expert conducted a search to identify pentapeptides based on their aggregation propensities calculated from coarse-grained molecular dynamics simulations and compared to sequences suggested by human experts in the field.[30] In this case, the scoring system for their model only relied on hydrophobicity and aggregation propensity of amino acids and did

not differentiate between different morphologies that peptides can self-assemble into, such as micelles, fibrils, vesicles, or sheets. Therefore, a systematic investigation into the effect of amino acid sequence on nanofibril morphology is still missing.

In this study, we apply molecular dynamics simulations to demonstrate how amino acid side chain hydrophobicity and position determine structural relationships between morphology and biological property of self-assembling PNFs on multiple length scales. To this end, we exploit a previously reported amyloid peptide library for neuron regeneration or retroviral gene transfer,[13,15,20] in which small sequence modifications modulated the β-sheet content, and resulted in materials with varying bioactivity, as shown in Figure 1. By utilizing coarse-grained Martini[31] model and atomistic CHARMM36 force fields[32] and comparing changes in the self-assembly rate, solvent accessible surface area, β-sheet content, fiber-fiber interactions, we find that having moderately hydrophobic N-terminal residues, e.g cysteine, enhances the self-assembly rate of peptides by weakening peptide-water interactions. In addition, retaining an amphiphilic order of hydrophobic-hydrophilic residues in the core region help increase the stability and β-sheet formation in self-assembled fibrils.We check the validity of our model and verify simulation results with experimental data on morphology and β-sheet content. We also investigate fiber-fiber interactions in solution to understand the surface properties of fibrils and their dynamic behavior.

With our multiscale computational approach, we identify which amino acid sequences are more likely to self-assemble into supramolecular fibrils with high β-sheet contents in solution. When peptides have low aggregation propensities, the self-assembled fibrils are less likely to aggregate with other fibers in solution, which makes the peptides more mobile in the supramolecular fibril. Our findings suggest that amino acid sequences that form fibrils with high β-sheet content and low mobility result in the highest biological activity for retroviral gene transfer. We envision that these microscopic insights from molecular design to macroscopic (bio)properties will not only lead to efficient and economical materials discoveries but also provide insights into aggregation mechanisms, which are commonly associated

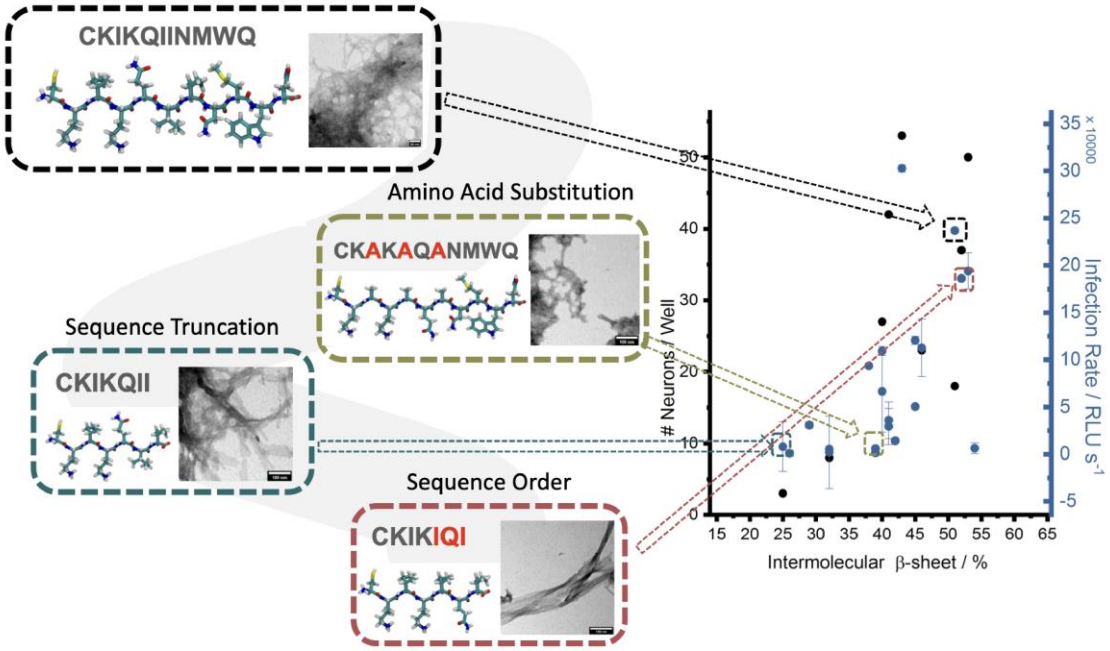


Figure 1: Relationship between intermolecular β-sheets and biological activity of PNFs in neuron regeneration (left axis, black dots) and retroviral transduction (right axis, blue dots) experiments. The effects of various chemical modifications of the peptide sequence such as amino acid substitution, sequence truncation and change in sequence order will be discussed in this study.

with several lethal diseases in naturally occurring amyloids.

Amyloid-like self-assembly of PNFs

In this study, we aim to investigate how specific amino acid mutations can lead to amyloid formation by analyzing structure-morphology-property relationships in short self-assembling peptides from a molecular simulation perspective. As an experimental readout for bioactivity, we selected infectivity enhancement. The increase in infection rates is a result of promoted delivery through the interaction of PNFs with virus and cell and can be quantified via a luminescence-based assay.[15] We apply infection rates relative to a reference PNF (Rel. Infectivity, see methods) to compare infectivity rates of different peptides with each other. As a starting point for the molecular simulation, we selected the peptide sequence CKIKQIINMWQ from our previous reports.[15,20] Experimental methods confirm that the peptide

sequence CKIKQIINMWQ forms long, thin fibrils with β-sheet structures in solution, but it is difficult to understand what these peptide assemblies look like from the available experimental data. In order to understand how each amino acid (placement along the chain, hydrophobicity, charge, etc.) contributes to the formation of β-sheet structures, first we performed simulations of the self-assembly process of peptides in solution.

One way of validating our methodology is to compare our results to a peptide sequence known to naturally form amyloid fibrils, such as the human prion protein (PrP). Misfolding of PrP and accumulation in the brain is responsible for an infectious fatal neurodegenerative disease called Prion disease in humans and many other animals.[33,34] Because simulating the misfolding and amyloid formation of the whole protein is a complicated problem and not in the scope of this study, we chose a short, β-sheet forming region of the PrP, VHDCVNITIK (residues 176-185), for our simulations. Furthermore, our experiments suggest VHDCVNITIK peptides form nanofibers with the same β-sheet content, but slightly different infection rate as CKIKQIINMWQ fibers, shown in Figure 2.

In Figure 2, we show the amino acid sequence for VHDCVNITIK (PrP) and the CKIKQIINMWQ peptide. VHDCVNITIK peptide consists of charged (acidic and/or basic) amino acids on both N- and C-termini of peptide chain and a core region with alternating polar and nonpolar amino acids. In comparison, CKIKQIINMWQ peptide only has lysine (basic) amino acids close to the N-terminus followed by a repeating pattern of polar-nonpolar-nonpolar amino acids in the core region and finally a polar amino acid on the C-terminus. We should note here, that even though we categorize the C-terminus Q in CKIKQIINMWQ peptide as polar, all peptides have a negatively-charged, acidic carboxyl group ($-CO_2^-$) on the C-terminus and a positively-charged amine group ($-NH_3^+$) on the N-terminus in neutral pH. As we will later show, placement and order of charged, polar and nonpolar residues play a significant role in the structure of the self-assembled fibers, which consequently affect their biological activity.

The self-assembly process and morphology of the structures that form highly depend on

simulation parameters, such as the force field, concentration, and time. For low concentration of peptides (0.04 peptides/nm^3), we observe an ordered amyloid-like fibril formation in under 5 µs for both VHDCVNITIK and CKIKQIINMWQ peptides, shown in Figure 2B-C. The alignment of the peptides is approximately perpendicular to the long axis of the fiber, along which the self-assembled structure is continuous over the periodic boundaries of the simulation box. The morphology of the fibers show twisting, where peptides are slightly disordered, but nevertheless, these structures are stable once they are formed (Movie X). At intermediate concentration (0.08 peptides/nm^3), we observe larger aggregated structures that are also fiber-like in morphology with peptides aligned more or less perpendicular to the long axis of the fiber (Fig. S3). The fibers are thicker and the twisting is more pronounced for these structures due to higher peptide concentration. Regarding the kinetics of self-assembly, both peptides self-assemble into a single fibrillar structure in less than 100 ns for low concentration and less than 5 ns at intermediate concentration (Fig. 2D). Having a more concentrated solution leads to the enhancement of hydrophobic contacts and speeds up the aggregation process. In terms of the different peptide sequences, the VHDCVNITIK peptide self-assembles into a fiber more quickly than CKIKQIINMWQ peptide at low concentration. This difference in self-assembly kinetics can be attributed to the arrangement of amino acids in the sequence. Both peptides are of similar length (10 vs 11 total residues per peptide chain, respectively) and contain similar ratios of nonpolar (5/10 vs 6/11) and polar (2/10 vs 3/11) amino acids. However, the interaction of the peptides with water in assembled fibers is different. The solvent accessible surface area (SASA) of the amino acids in the assembled fiber structures is similar when we consider the two ends of the peptides (N- and C-terminus) as they contain charged amino acids at both ends that interact with water and ions in solution. However, in the case of CKIKQIINMWQ, the charged lysine groups are located in 2nd and 4th position, where the latter position is close to the β-sheet forming core region of the peptide. The effect of these lysine residues is clearly visible with high SASA value, whereas VHDCVNITIK fiber becomes more hydrophobic closer to the

core region (Fig. 2E). The stronger hydrophobic interactions between the peptides increases the kinetics of self-assembly for VHDCVNITIK peptide compared to the CKIKQIINMWQ peptide. This result is in agreement with earlier work by Fu et al.[28], where increasing the hydrophobic interactions between amino acids lead to more fibrillar morphologies of self-assembled peptide amphiphiles.[28] The increases in hydrophobicity might also cause lateral attachment of fibrils to each other resulting in a more sheet-like morphology (Fig. 2G). Fiber-fiber interactions in solution will be discussed in more detail in the final section of the Results.

While coarse-grained simulations at low and intermediate concentrations give us information about how these peptides self-assemble into larger structures, they cannot provide information about the peptide secondary structure, e.g. β-sheets, α-helix or random coil, since the secondary structure in the coarse-grained model is fixed at the start of the simulation. Formation of β-sheet networks plays an important role in the stability of the fibrils once they assemble in solution. We can examine the secondary structure of peptides using atomistic models, but this is computationally expensive with the number of molecules used in our low and intermediate concentration simulations. Therefore, we performed additional coarse-grained simulations at a higher concentration (20 peptides/125 nm^3 or 0.16 peptides/nm^3) with fewer number of peptides and smaller simulation box. This concentration is carefully chosen so that the randomly distributed peptides self-assemble into a single fiber. We then converted the coarse-grained single fiber structures into atomistic model using a backmapping procedure (see Methods Section in SI). The snapshots of a short section of VHDCVNITIK and CKIKQIINMWQ fibers at high concentration are shown in Figure 3 with both coarse-grained and atomistic model. For both peptide sequences, the fiber structure remains stable after conversion to atomistic model and equilibration for 250 ns. On average, 16 % of VHDCVNITIK peptides form β-sheet structures, which is about twice the amount formed by CKIKQIINMWQ peptides (7.8 %) (Fig. 3E). It's important to note here that the β-sheet content from backmapped atomistic structures is very small compared to

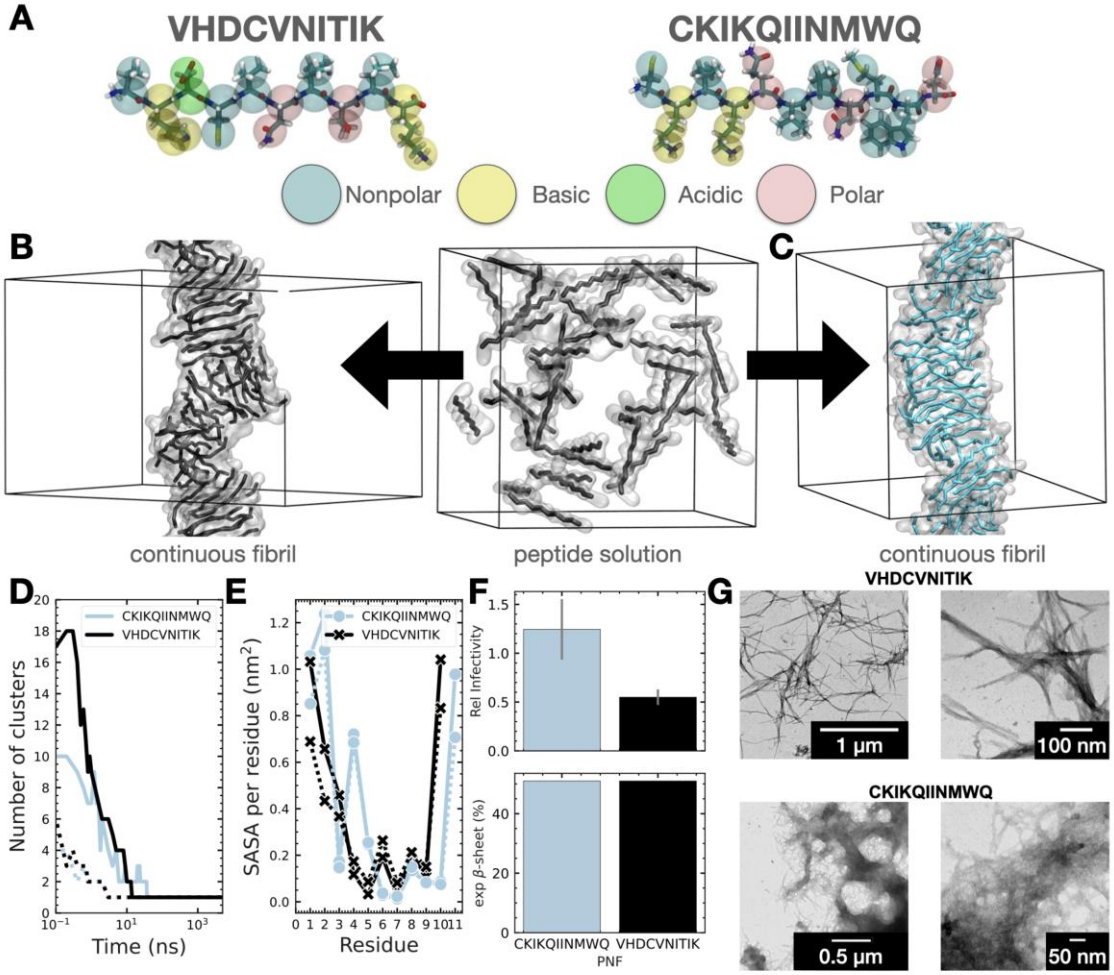


Figure 2: (A) The coarse-grained model used in simulations for the peptide sequences VHDCVNITIK and CKIKQIINMWQ. Here, the beads are colored according to amino acid side hydropathy and charge which is coarse-grained into four categories (nonpolar, basic, acidic and polar). After 5 µs, randomly distributed peptides in solution self-assemble to form fibrils that are extended over the periodic boundaries of the simulation box. Peptide backbones are colored black for (B) VHDCVNITIK and blue for (C) CKIKQIINMWQ. The rest of the peptide is shown in white. Water and ions are omitted for clarity. Comparisons between (D) self-assembly kinetics depicted by the change in number of clusters over simulation time and (E) average solvent accessible surface area (SASA) per residue from self-assembled fibrils from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration (shown in B and C), while dotted lines refer to simulations at intermediate peptide concentration (see Fig. S3 in SI). Experimental data comparing (F) relative infectivity at 1.3 µM concentrations and intermolecular β-sheet content of peptides. The error bars refer to standard deviation from different measurements. (G) TEM images for VHDCVNITIK and CKIKQIINMWQ fibers.

experiments, which is due to the missing secondary structure information when backmapping from Martini to CHARMM model. During backmapping, it is especially difficult to restore extended structures, which lead to β-sheet formation between peptides, because the backbone beads in the Martini model are placed at the center of mass of each residue instead of C_α position, and the backmapping procedure lacks the information to build the correct peptide planes that would result in extended regions in atomistic model.[35] Nevertheless, we use these backmapped atomistic structures to understand the trends in the experimental data and also to more accurately design the pre-formed fibers, which will be discussed later.

For the VHDCVNITIK peptide, there is an increase in the amount of β-sheets that each amino acid makes starting from the N-terminus that reaches a maximum with the 8^{th} residue (threonine, T). Amino acids close to the two ends of the peptide chain are exposed to the water and ions, and the strong interactions with solution prevent them from forming hydrogen bonds with other peptides that are necessary for forming intermolecular β-sheets structures. The number of β-sheets is strongly correlated with the solvent-accessible surface area of each amino acid in the assembled fiber (Fig. 3F and S4). This explains the low amount of β-sheets observed in CKIKQIINMWQ fiber, where instead of a very hydrophobic core region as in VHDCVNITIK fiber, the 4^{th} amino acid residue (lysine, K) is highly hydrophilic. This means that the lysine residues are on the surface of the fiber, where water and ions can penetrate and prevent peptide-peptide contacts. We should also mention that the SASA trends for each amino acid calculated from atomistic simulations at high concentration matches very well to the SASA calculated from coarse-grained simulations at low and intermediate concentrations, suggesting trends we observe in terms of how peptides self-assemble and interact with solvent are not dependent on peptide concentration or the computational model used.

Recently, a study by Álvarez et al.[14] showed that similar self-assembling peptides are effective in promoting corticospinal axon regrowth after acute spinal cord injury in mice. However, surprisingly the authors of this study found that the disruption of the β-sheet

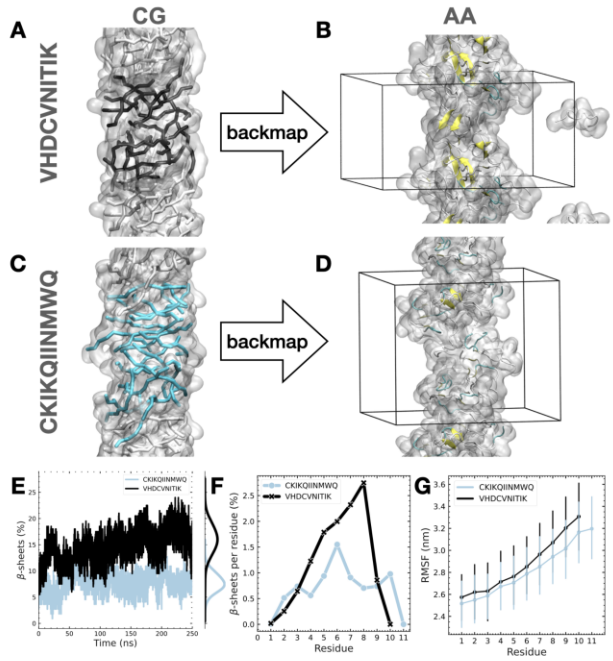


Figure 3: The self-assembled fibers at high peptide concentration using the coarse-grained (CG) model (A, C) and the corresponding structures after 250 ns of equilibration with the all atom (AA) model (B, D). In B and D, peptide backbones are drawn and colored according to the secondary structure, where β-sheet is yellow, random coil is white and turn is cyan. The total percentage of β-sheets over time (E), average β-sheets per residue (F) and average root mean square fluctutation per residue (RMSF) after the backmapping procedure is shown for two peptides. In E and F, percentage is calculated per total number of residues in the simulation box. For F and G, only the last 100 ns of the simulation is used for calculating the average values.

region increases the efficiency of nanofibers in promoting neuronal differentiation of neural progenitor cells in vitro. Authors speculate that the added motion of disordered protein domains that results from disruption of β-sheet structures provides greater capacity to signal efficiently in the biological environment. Their coarse-grained model and simulations provide useful information about the mobility of peptides in assembled structures, but the authors do not extend their computational model to provide any information about secondary structures, i.e. β-sheet formation, in peptide nanofibers. Our simulations show that while the mobility of the peptides, which we determine by calculation of root mean square fluctuation of each residue, in self-assembled fibers is comparable for both peptides, VHDCVNITIK also has a higher percentage of β-sheets (Fig. 3E,G). Interestingly, while both fibers start

with similar amounts of β-sheets structures after the backmapping procedure, VHDCVNITIK develops more β-sheets over 250 ns simulation time while CKIKQIINMWQ remains the same. This suggests that perhaps high mobility of peptides in the fibers does not necessarily mean loss of stability as suggested by Álvarez et al.[14], but can also contribute to structure formation.

In reality, the mechanism of amyloid aggregation is a very complicated one and simple coarse-grained models like Martini that do not allow for folding of the peptide, i.e. change in secondary structure, may not give the perfect amyloid structures isolated in experiments with Cryo-EM.[34] Furthermore, self-assembling simulations with fully atomistic models in explicit solvent would take a very long time if started from random peptides in solution. Therefore, we constructed an amyloid fiber from VHDCVNITIK and CKIKQIINMWQ peptides using an atomistic model and tested their stability and β-sheet formation. In Figure 4, we show the amount of β-sheet structures over the course of 300 ns for pre-formed fibers of CKIKQIINMWQ and VHDCVNITIK peptides. Both fibers remain stable and contain the same amount of β-sheet structures on average, which is closer to experimentally measured values in Fig. 2F. The pre-formed fibrils have higher β-sheet contents, because the peptides are packed together with minimal defect in the initial fiber structure and since it is fully atomistic, there is no backmapping procedure that may result in loss of structural information. While β-sheets in pre-formed VHDCVNITIK fiber is distributed in a very similar manner to the self-assembled fiber, where we observe the highest β-sheet content for residues 4-9, the trends observed for pre-formed CKIKQIINMWQ fiber is slightly different to its self-assembled counterpart. Pre-formed CKIKQIINMWQ fiber shows a steady increase in the number of β-sheets starting from the N-terminus and reaching a maximum at the 8^{th} residue (asparagine, N) and slowly loses structure towards the C-terminal of the peptide chain. With perfectly aligned peptides, the effect of water-peptide interaction at the core region of the fiber is eliminated and strong peptide-peptide interactions provide increased stability against losing structure. This suggests that over time CKIKQIINMWQ peptides

can form nanofibers with similar β-sheet content as those of VHDCVNITIK peptides, which might explain the similarities between β-sheet content measured in experiments.

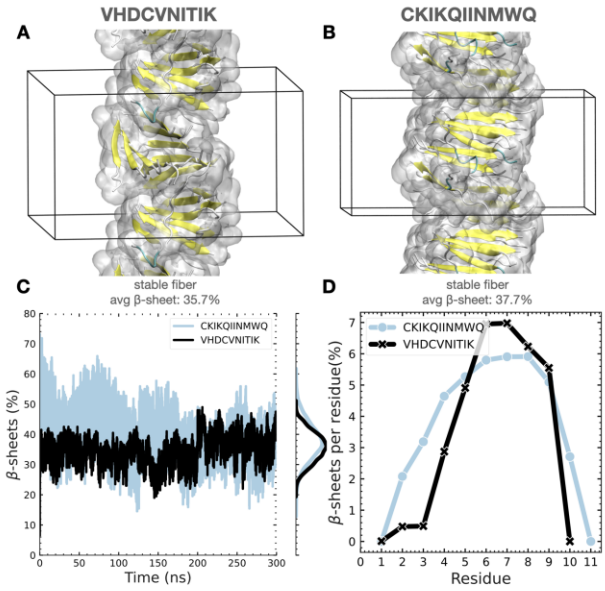


Figure 4: Snapshots of side view of (A) VHDCVNITIK and (B) CKIKQIINMWQ pre-formed fibers after 300 ns atomistic simulation. Peptide backbones are drawn and colored according to the secondary structure, where β-sheets are shown in yellow. (C) Total percentage of amino acid residues making β-sheets structures over the simulation and (D) average β-sheet amount per residue calculated from the last 100 ns of simulation. Both in (C) and (D), percentage is calculate with respect to the total number of amino acid residues.

Earlier work suggests that different parts of the CKIKQIINMWQ peptide sequence contribute to the desirable properties of PNFs, such as their ability to self-assemble efficiently in solution, form stable fibers through strong β-sheet structures and bind to cells for retroviral transduction. Experiments investigate the effects N-terminal cysteine, amphiphilic KIKQII, and C-terminal NMWQ regions have on structural properties, such as morphology and β-sheet content, and biological activity, such as transduction enhancement[15] and neuron regeneration.[13] In previous experimental work, through sequence modifications the role of certain regions of the peptide in its self-assembly behavior was evaluated. Now, we will use a computational approach to study the self-assembly kinetics and structure in more detail.

Role of N-terminal cysteine

According to experimental data, presence of the N-terminal cysteine (C) promotes formation of β-sheet structures and enhances transduction rates in PNFs.[15,20] However, it is not fully clear how cysteine does this. When cysteine is replaced with the more hydrophobic methionine amino acid that is very similar in structure, the enhancement in intermolecular β-sheets and transduction activity was not observed, suggesting that the influence of N-terminal C is not of a hydrophobic nature.[15]

Removing cysteine

We performed simulations comparing CKIKQIINMWQ peptide to KIKQIINMWQ, where only the N-terminal C is missing, shown in Figure 5. Qualitatively, the coarse-grained self-assembled structures are very similar in nature, with both sequences self-assembling into amyloid-like structures. At high concentration, we observe a fibrillar structure, which remains stable when backmapped to the atomistic model. However, only 4.54 % of peptides form β-sheet structures, compared to 7.8 % from the same method with CKIKQIINMWQ peptides. While the structures backmapped to atomistic from Martini have very small amount of β-sheets compared to experiments (Fig. 5C), the simulations of pre-formed fibrils are quantitatively in better agreement with experiments. From atomistic simulations of pre-formed fibrils, the amount of KIKQIINMWQ peptides forming β-sheets is 35.3 %, which is slightly less than that of CKIKQIINMWQ (37.8 %), shown in Fig. S7.

Compared to CKIKQIINMWQ, the number of peptides that form β-sheets decreases upon removal of N-terminal cysteine. The decrease in β-sheet formation is also observed in experiments for KIKQIINMWQ fibers suggesting that N-terminal C may play a role in structure formation and stability of the PNFs. This loss in stability is observed in low and intermediate concentration simulations (Fig. S5), where the structures resemble fragment of a fiber instead of a continuous one that we obtained earlier with CKIKQIINMWQ. At

low concentration, absence of the N-terminal C leads peptides to aggregate slower and the formed structure fluctuates between one or two aggregates.

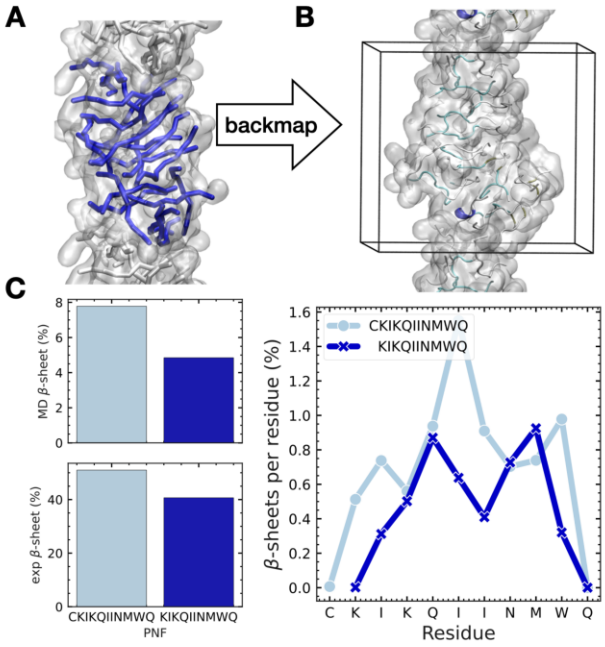


Figure 5: Self-assembled KIKQIINMWQ fiber at high peptide concentration from (A) coarse-grained and (B) backmapped atomistic simulations. (C) The change in the average β-sheet content with removal of N-terminal cysteine from experiments and MD simulations and the average percentage of β-sheets per residue.

The slowing down of self-assembly kinetics may be attributed to the role of N-terminal C in shielding the rest of the peptide from water and ions. The surface hydrophobicity of the KIKQIINMWQ fiber is very similar compared to the original sequence except for a significant increase in peptide-solvent interactions at the N-terminus (Fig. S5D). Because lysine has a charged side chain, as opposed to the hydrophobic side chain of cysteine, the first lysine residues in KIKQIINMWQ fiber are fully exposed to the solvent on the surface of the fiber (Fig. 6A-B), whereas the first cysteine residue in CKIKQIINMWQ fiber reduces the contact area of lysine (residue 2) with solvent (Fig. 6C-D). As a consequence, the amino acids close to the ends of the peptide interact more with solvent and are less likely to form β-sheets (Fig. 5C).

We conclude, that removing N-terminal cysteine reduces hydrophobicity and slows down

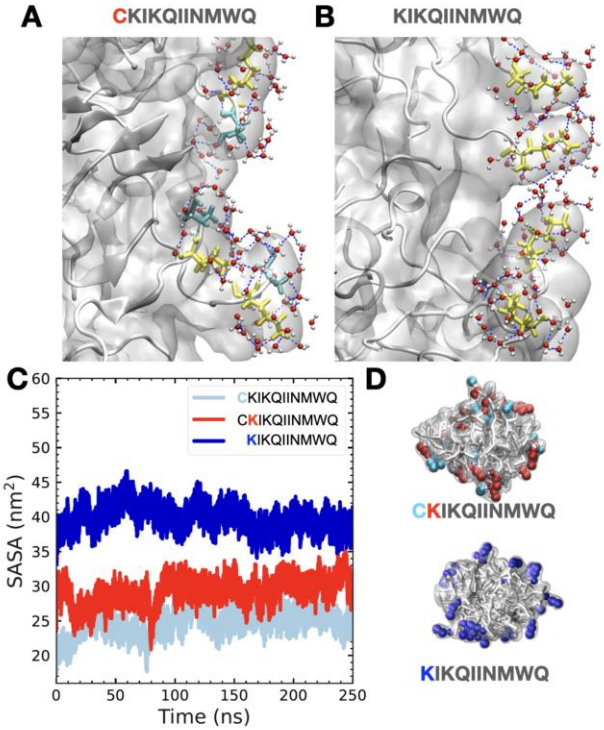


Figure 6: Interaction of water with the N-termini of self-assembled (A) CKIKQIINMWQ and (B) KIKQIINMWQ fibers. Here, cysteine is colored in cyan and lysine is yellow. Only water molecules within the first solvation shell of selected residues are shown. Hydrogen bonds are drawn with dashed black lines. (C) The solvent accessible surface area (SASA) for cysteine (residue 1, light blue), lysine (residue 2, red) in CKIKQIINMWQ and lysine (residue 1, dark blue) in KIKQIINMWQ is shown over simulation time. (D) Cross-sectional view of the self-assembled fibers and distribution of N-terminal residues.

self-assembly kinetics. Similar observations are made upon reduction of hydrophobicity of peptides substituted with less-hydrophobic amino acids such as alanine, e.g. CKIKIQINMWQ to CKAKAQANMWQ (Fig. S6-S7). However, if the N-terminus is too hydrophobic, e.g cysteine to methionine substitution, this may also lead to disruption of the β-sheet region during the self-assembly process (see section "Tuning surface hydrophobicity").

Switching cysteine position

Presence of N-terminal cysteine increases self-assembly kinetics and enhances retroviral transduction, but what about if the cysteine was located at the other end of the peptide chain, i.e. the C-terminus? Experiments show that switching the position of the N-terminal

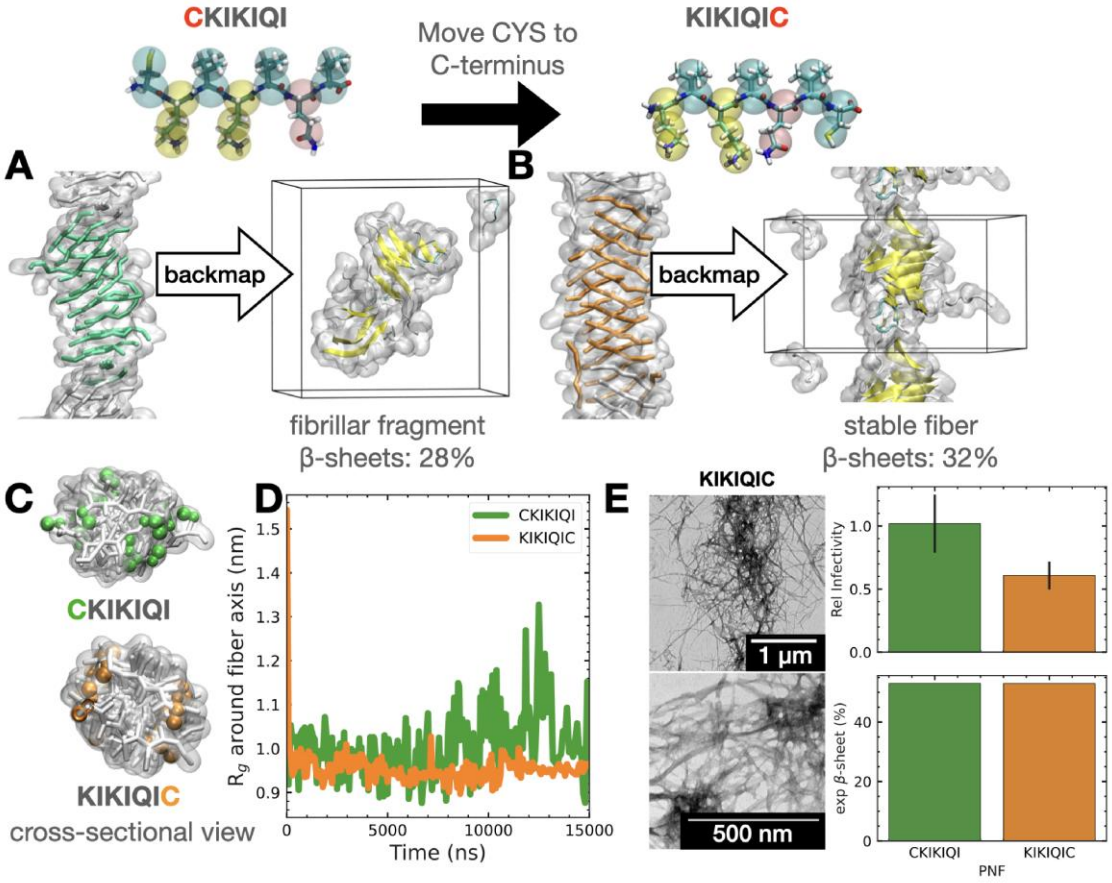


Figure 7: Snapshots of self-assembled fibers in solution from coarse-grained and atomistic simulations for (A) CKIKIQI and (B) KIKIQIC. (C) Cross-sectional view of the self-assembled fibers showing the distribution of cysteine residues. (D) The radius of gyration around the fiber axis from coarse-grained simulations. (E) Experimental TEM images of KIKIQIC, β-sheet content and infectivity relative to EF-C. For TEM images of CKIKIQI, see Fig. 8.

cysteine to the C-terminus leads to ribbon-like fiber morphologies with similar β-sheet content for short peptide sequences (CKIKIQI vs KIKIQIC), but the infection rate is lower when cysteine is located on the C-terminus, shown in Fig. 7. These findings were confirmed by our simulations, where we found very similar self-assembly behavior at low and intermediate concentrations (Fig. S8). The biggest difference in the two structures is how the two ends of the peptide chains interact with water. Without the cysteine, the N-terminus of KIKIQIC sequence is highly hydrophilic, as discussed in the previous section, and the C-terminal cysteine has less exposed surface area to solvent compared to the N-terminal cysteine (Fig.

S8D). We further investigate the effect cysteine-water interactions have on structure formation and stability of PNFs with simulations of self-assembled single fibers at high peptide concentration, shown in Figure 7. KIKIQIC peptides form flat, ribbon-like fibers, where the peptides are aligned so that the hydrophobic amino acids are buried and the lysines are facing the surface. What is different from CKIKIQI in this structure is that peptides form perfectly aligned dimers, where N-terminus of peptides on one side of the fiber directly face the C-terminus of the peptides on the other side, which are then stacked along the length of the fiber (Fig. 7A-C). Thus, the cross-sectional radius of the fiber is smaller with less fluctuations (Fig. 7D). We observed very similar results when we performed simulations for other sequences, such as KIKQIINMWQC and KIKQIIC (Fig. S9-S11). Overall, our results were consistent with KIKIQIC. We did not observe significant differences in properties such as morphology of self-assembled peptides, self-assembly kinetics or amount of β-sheet structures. However, C-terminal cysteine was always more hydrophobic and interacted less with solvent compared with N-terminal cysteine.

Whether or not cysteine arrangement on the fiber surface has a direct impact on biological activity can be better assessed by investigating interaction of the supramolecular fibers with other fibers in solution, which will be discussed in the section "Dynamic properties of PNFs and correlation to biological activity", and with cellular membranes, which is not the scope of this paper, but something we would like to explore in the future.

Role of C-terminal NMWQ

On the other end of the CKIKQIINMWQ peptide, the NMWQ sequence was initially thought to enhance transduction in PNFs as the removal of NMWQ from CKIKQIINMWQ peptide resulted in a drastic decrease in biological activity. However, after testing other sequences, authors concluded motif NMWQ was not essential for enhancing the retroviral transduction properties.[15] Our simulations show that when NMWQ sequence is removed from the

CKIKQIINMWQ peptide, CKIKQII peptides form disordered aggregates at low and intermediate concentrations instead of a fiber (Fig. S12). This is not surprising since the residues least exposed to solvent, i.e. most hydrophobic, in the original CKIKQIINMWQ sequence are IINMW, which are now either missing or on the C-terminus interacting with solvent. A fibrillar structure forms at high concentration (Fig. S13, Movie X), but this structure is not stable and dissociates completely when backmapped to atomistic representation. The loss in stability can be expected for any backmapping procedure from coarse-grained to atomistic model simply because coarse-grained models lack certain amount of information that is necessary to correctly rebuild the atomistic structures. However, in this case, we should note that the CKIKQII fiber that forms in the coarse-grained simulations also loses stability during the 15 µs simulation (Movie X), while the CKIKQIINMWQ fiber remains stable throughout the simulation time. Even when we start from a pre-formed fiber, CKIKQII loses stability and β-sheet structures very rapidly within about 100 ns, shown in Figure 8A. Nevertheless, this does not mean that all short peptide sequences perform poorly. A single change in the order of amino acids, i.e. CKIKQII to CKIKIQI, enhances the stability and β-sheet propensity of peptides. There is a decrease in amount of β-sheets when NMWQ is removed (Fig. 8B), but CKIKIQI still has a fibrillar morphology with comparable β-sheet composition to the original sequence CKIKQIINMWQ. In experiments, there is a decrease in infectivity and β-sheet formation for CKIKQIINMWQ to CKIKQII transformation, which is not true for CKIKIQI (Fig. 8C-D). TEM images also show considerable differences in the morphologies of CKIKIQI, where thicker, more-defined fibers are present. The effect of amphiphilic pattern in PNF structure is discussed in the next section in more detail.

Effect of amphiphilic pattern

So far, we investigated the role of the amino acids at two ends of the peptide sequences, which are important for surface properties of the assembled nanofibers and their interaction with the

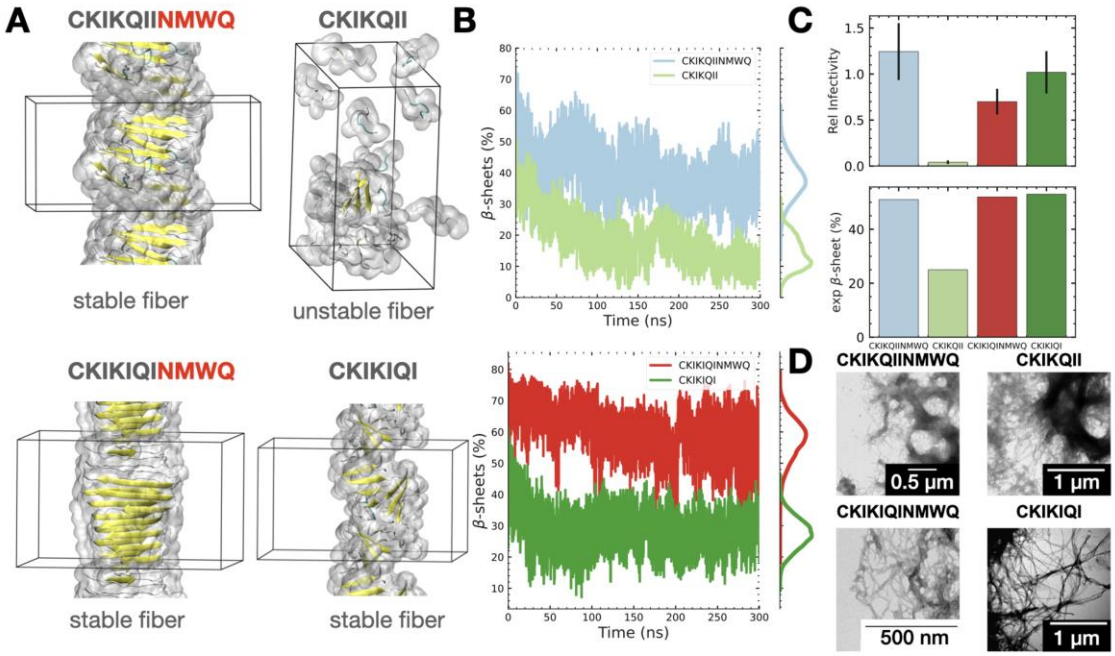


Figure 8: (A) Snapshots of pre-formed fibers after 300 ns of atomistic simulations and (B) the corresponding percentage of total β-sheets, showing the effect of NMWQ on structure formation. Experimental data on (C) infectivity relative to EF-C and β-sheet content and (D) TEM images for the chosen peptide sequences.

cell membranes. Furthermore, previous experiments have shown that secondary structure, which is a direct consequence of amyloid-like packing of the peptides, is highly correlated with the infection rates. In order to optimize the internal packing of the peptides, Sieste et al.[15] proposed an intrinsic switch of 5^{th} and 6^{th} residues (CKIKQII to CKIKIQI), which produces an alternating hydrophilic-hydrophobic pattern of amino acids. Coarse-grained simulations at low and intermediate concentrations did not show a significant difference in self-assembly kinetics or the solvent accessible surface area, but the peptides that form the aggregates in solution were more ordered in an amyloid-like packing for CKIKIQI (Fig. S14). Furthermore, when self-assembled fiber at the high concentration was backmapped to atomistic representation, the periodic structure was disconnected but still retained internal order, shown in Figure 7A, with high percentage of β-sheets (28.6 %) instead of complete dissociation as in the case of CKIKQII. We also observed the same effect for the longer peptide sequence (Fig. S15-S16), but since presence of NMWQ already increases stability,

effects of CKIKQIINMWQ to CKIKIQINMWQ transition is not as remarkable as in the short peptide sequences that lack the NMWQ fragment, in agreement with Sieste et al.[15]. The difference between the pre-formed fibers, previously shown in Figure 8 also follows the same trends, where the intrinsic -QII- to -IQI- switch in long sequences slightly increases β-sheet content in both experiments and simulations, but a significant increase is observed for the short peptide sequences only.

The importance of a single switch in the amino acid sequence in terms of the fiber properties is best explained through its interaction with solvent. Figure 9 shows the spatial distribution of the 5^{th} residue, where the switch happens, and its interaction with water in the self-assembled fiber structures. The 5^{th} residue in the peptide sequence is important since it is preceded by a pattern of -basic-hydrophobic-basic- (-KIK-) amino acids, which already determines the orientation of the peptide chain with respect to the fiber surface so that the basic lysine residues can interact with water, ions and/or negatively-charged cellular membranes. Glutamine (Q) has a polar side chain that can form hydrogen bonds with water, whereas isoleucine (I)'s hydrophobic side chain minimizes its interaction with water. This is shown in Figure 9, where Q residues are distributed closer to the fiber surface and I residues are located at the core region, even though they are both the 5^{th} residue in the peptide chain. The number of water around these residues is shown qualitatively in Fig. 9A-B and quantitatively in SASA calculation in Fig. 9E. Following the -KIK- pattern with a hydrophobic amino acid ensures similar side chains to face the same side of the peptide. In other words, having a hydrophobic amino acid near the core of the peptide allows for better packing and increases peptide-peptide interactions that lead to formation of β-sheets.

Tuning surface hydrophobicity

So far we have shown that stability and formation of β-sheet structures in supramolecular nanofibers can be enhanced by increasing the peptide-peptide interaction and minimizing

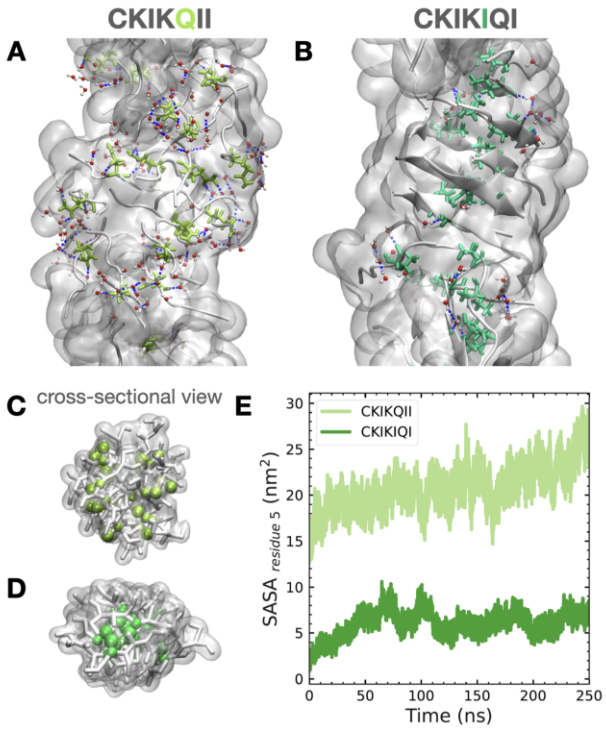


Figure 9: The distribution of 5th residue (green) in self-assembled fibers for (A) CKIKQII and (B) CKIKIQI peptides from high concentration atomistic simulations. Because for CKIKQII, the fiber structure is not very stable (Movie X), here we used a snapshot from the trajectory when the fibrillar morphology was observed. Water molecules within the first solvation shell of 5th residue is shown with hydrogen bonds in blue. The cross-sectional snapshots of the structures are shown in (C) and (D), respectively. (E) The solvent accesible surface area (SASA) for 5th residue from each structure calculated using the atomistic simulations.

the peptide-water interaction. This can be done by modification of the ends of the peptide chains via presence of a hydrophobic amino acid, such as cysteine, or optimization of internal amphiphilic pattern. However, it is not clear to what extent the degree of hydrophobicity plays a role in the self-assembly process and the observed biological properties. In this section, we explore how various substitutions of amino acids can tune the properties of peptide nanofibers.

Cysteine to methionine substitution

While trying to assess the role of N-terminal cysteine in enhancing infection rates, experiments where cysteine was replaced with more hydrophobic methionine revealed no enhance-

ment in transduction efficiency,[15] shown in Figure 10. We investigated the effect of hydrophobicity by replacing the cysteine with methionine in CKIKIQI peptide. Overall, the morphology of the aggregated peptides and self-assembly process is very similar for CKIKIQI and MKIKIQI (Fig. S17). This is not surprising, since the parameters used for methionine and cysteine are identical in the coarse-grained model with the exception of a larger bond length between side chain and backbone bead, since methionine side chain is bulkier compared to cysteine. At high concentration, MKIKIQI peptides assemble to form a fiber, shown in Figure 10, that is less ordered compared to CKIKIQI (Fig. 7A). While some of the peptides are aligned perpendicular to the long axis of the fiber, others are oriented in different directions, which lead to more random distribution of the methionine over the cross-sectional area of the fiber. While cysteine is more concentrated on the surface of CKIKIQI fiber, some of the methionine is also buried close to the core of the MKIKIQI fiber (Fig. 10C). This has direct consequences in terms of the peptide-peptide interaction at the atomistic level. The bulky side chain of methionine interrupts the hydrogen bonding interactions between peptides that lead to β-sheets. Therefore, we observe a considerable decrease in the number of β-sheets in the core region where CKIKIQI fiber shows the highest amount (Fig. 10D).

If we had better alignment of MKIKIQI peptides, would the increase in hydrophobicity of the methionine with respect to cysteine increase the stability of the fiber in solution? Our pre-formed fiber simulations, where peptides are pre-aligned in solution, reveal that substitution of cysteine with methionine increases the amount of β-sheet structures that form at the core region of the fiber (Fig. S20A-D). This shows that the more hydrophobic methionine amino acid is better at maintaining the stability of the fibrillar morphology by weakening the interaction of the peptides with water and ions, but because of this increased hydrophobicity it can also prefer to be buried near the core of the peptide and lead to misalignment of peptide chains during the self-assembly process, which eventually prevents formation of β-sheet structures. Thus, an intricate balance between hydrophobicity and position of the amino acid in the sequence is needed for optimizing properties of PNFs.

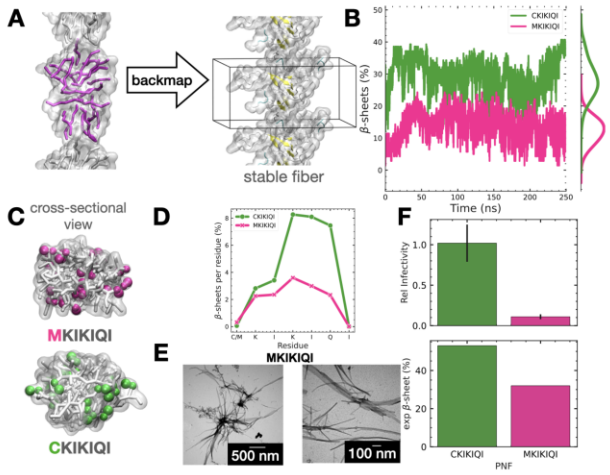


Figure 10: The self-assembled MKIKIQI fiber from coarse-grained and atomistic simulations are shown in (A), where both structures show misalignment of peptides around the periodic boundaries of the simulation box. For comparison to CKIKIQI, see Fig. 7A. The total percentage of β-sheets from are shown in (B). The distribution of the first residue (methionine in pink and cysteine in green) over the cross-section of the fiber is shown in (C). Here, we can see methione residues in the core region disrupting the β-sheet network over the high β-sheet-propensity region (D). Experimental (E) TEM images for MKIKIQI and (F) calculated properties.

Isoleucine to alanine/phenylalanine substitution

Furthermore, we discovered that having an internal amphiphilic pattern of hydrophobic and hydrophilic amino acids optimized the packing of the peptides when they assemble to form a fiber in solution. Sieste et al.[15] designed different peptide nanofibers by tuning the hydrophobicity of internal hydrophobic residues. When isoleucine residues were replaced with alanine, which has a smaller and less hydrophobic side chain, they observed no enhancement in transduction. Our simulations revealed that CKAKAQANMWQ peptides assembled into aggregates with peptides not as ordered as in CKIKIQINMWQ aggregates, shown in Figure 11. CKIKIQINMWQ structures resemble a fibrillar fragment at low concentration and a twisted fibril at intermediate concentration, where the peptides are stacked mostly perpendicular to the long axis of the fiber. CKAKAQANMWQ, on the other hand, self-assembles to form aggregates that are not continuous over the periodic box in any direction, and the peptide is more randomly packed together. Interestingly, the self-assembly kinetics calculated by

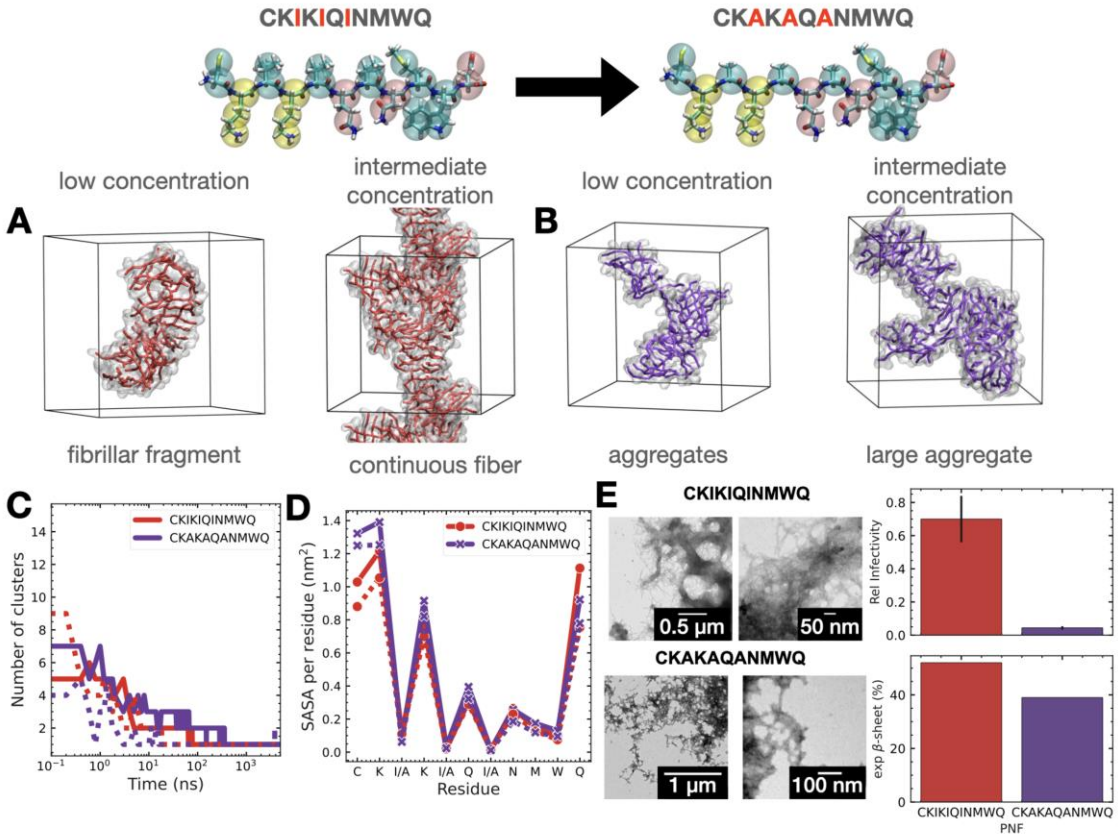


Figure 11: Comparison of self-assembled structures from low and intermediate concentrations of (A)CKIKIQINMWQ and (B) CKAKAQANMWQ peptides. Although the (C) self-assembly kinetics and (D) solvent accessible surface area (SASA) are very similar for both sequences, the morphology of the resulting structures in (A-B) simulations and (E) experiments, as well as experimental β-sheet content and infectivity relative to EF-C are different.

number of clusters over time are very similar for both peptides, but the polar residues close to the core of the fiber (i.e K and Q) have slightly more surface area accessible to the solvent when they are found around alanine instead of isoleucine (Fig. 11C-D). This small decrease in hydrophobicity near the core region of the peptide when isoleucine is replaced by alanine results in more peptide-water interactions that may lead to the less ordered aggregated structures (Fig. 11E).

At high concentration, the self-assembled CKAKAQANMWQ aggregate shows some β-sheet formation (Fig. S18), but the fibrillar structure did not remain stable in atomistic simulations. Moreover, while pre-formed CKIKIQINMWQ fiber revealed remarkable sta-

bility and a large amount of β-sheet structures, CKAKAQANMWQ peptides were more disordered and we observed a decrease in β-sheets over time as well as formation of 3_{10}-helix structures (Fig. S20E-G). This is in agreement with experimentally determined propensity of amino acids to form helix structures, in which alanine has the highest propensity compared to all other amino acids.[36]

On the other hand, replacing alanine with a more hydrophobic amino acid, such as phenylalanine, results in formation of fibrils even without the N-terminal cysteine present, shown in Figure 12. KFKFQFNMWQ peptide sequence forms continuous fibrils for all concentrations tested in our simulations. This observation agrees well with the TEM images (Fig. 12E), where KAKAQANMWQ self-assembles into aggregates while KFKFQFNMWQ produces long, thin fibrils. The self-assembly time for the KFKFQFNMWQ peptides is faster for both concentrations (Fig. 12C) and the resulting structure has a lower solvent accessible surface area, meaning the peptides are packed closely together in amyloid-like fashion. While the self-assembled high concentration single fiber shows only 11.4 % β-sheet formation when backmapped to atomistic representation (Fig. S19), pre-formed KFKFQFNMWQ fiber remains very stable in solution and 54.9 % of peptides form β-sheet structures (Fig. S20H-J). Regardless of which model is used, there is an increase in the β-sheet content for the alanine to phenylalanine substitution.

Dynamic properties of PNFs and correlation to biological activity

So far, we investigated the effect peptide primary structure has on morphology by simulating how peptides self-assemble in solution, their secondary structure formation and stability in solution. However, in all of our simulations, when the peptides formed fibers, we only observe a single fiber in solution. This is due to the finite size of simulation boxes. The TEM images obtained in experiments usually show bundles of fibers in close proximity. If

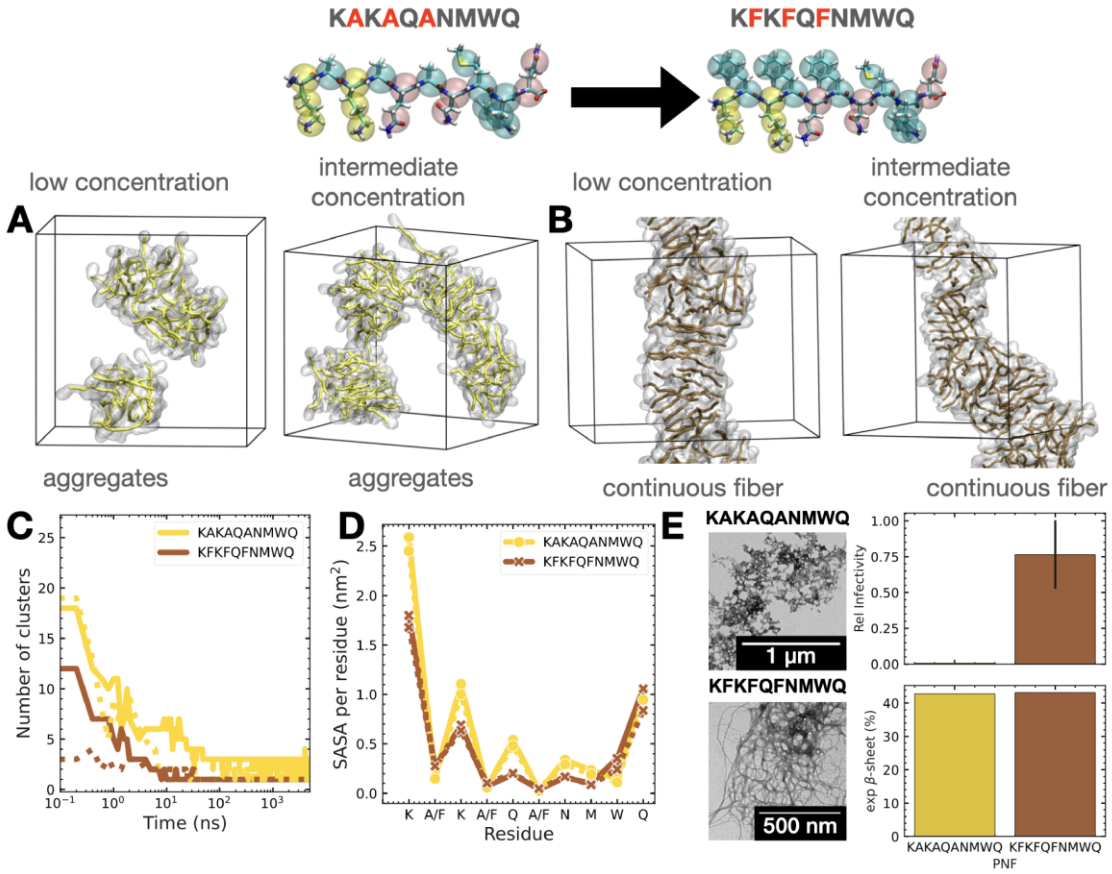


Figure 12: Comparison of self-assembled structures from low and intermediate concentrations of (A)KAKAQANMWQ and (B) KFKFQFNMWQ peptides. There are significant differences in (C) self-assembly kinetics and (D) solvent accessible surface area (SASA), where the increase in hydrophobicity in phenylalanine hinders peptide-water interactions of hydrophilic residues (K and Q) in close proximity. (E) The TEM images and experimentally measured properties show significant enhancement of properties with KFKFQFNMWQ.

multiple fibers interact and form superstructures, then the surface properties of fibers will be different than those of single fibers in solution.[20] Therefore, the interaction of fibers is important for understanding not only how the fibers interact with the cell membranes, but more importantly how higher order assemblies are formed since the self-assembly process continues after formation of single fibers and leads to hierarchical structures.

To study the interaction between fibers, we performed simulations by placing them in close proximity with each other allowing peptides to interact. These simulations were performed with the coarse-grained model since the size of the system and simulation time are

too large for atomistic simulations. With these simulations, we expect to gain information about the dynamical properties of these supramolecular polymers, such as whether and how the peptides in separate fibers interact with each other, diffuse in and out of fibers, rearrange and form new structures, etc. These dynamical features of supramolecular polymers are still mostly unclear and experimental methods lack the ability to investigate such events.

Some of the presented peptide sequences form dispersed single fibers, whereas others tend to aggregate in solution to µm-sized particles, as determined via combinatorial experimental TEM, light scattering, and widefield microscopy measurement, which are covering multiple assembly scales.[20] In our simulations, we used self-assembled fibers at high peptide concentration and placed four in very close proximity shown in the initial configuration (Fig. S1), so that not only the peptides in each supramolecular fiber can interact through "intra-fiber" interactions, but also they can sense the presence of peptides in other fibers through "inter-fiber" interactions. After 10 µs simulation for each peptide sequence, we observed three different behaviors: I) fibers remain stable and isolated in solution II) fibers remain stable and interact with each other III) fibers lose stability, re-mix and form new aggregates and/or fibers in solution. Because not all the peptides have the same number of amino acids, it is difficult to compare the structures based on a property such as surface area. Therefore, we use distance between center of mass of fibers in solution, normalized by the diameter of a single fiber for each peptide sequence, to distinguish the different aggregation behavior of fibers, shown in Figure 13. Rapidly decreasing fiber-fiber distance shows which fibers lose stability and re-mix in solution (III), while a slight decrease in distance refers to those fibers that remain stable and interact to form aggregated fiber bundles (II). On the other hand, an increase in fiber-fiber distance means the fibers remain stable and isolated in solution (I). Some examples of these three behaviors are shown in Figure 13A, where we normalize the fiber-fiber distance by the fiber diameter for each peptide sequence. Here, we notice that the peptide sequences that dissociate and re-mix are those lacking the C-terminal NMWQ fragment, which helps to improve their stability in solution, with the exception of the KIKIQIC

sequence that remains stable and isolated in solution.

Furthermore, we investigate whether there is a direct relationship between the aggregation of fibrils in solution and the ability of peptides to self-assemble into a fibril. In our coarse-grained simulations, the peptides are randomly dispersed in the simulation box at the beginning of the simulation (Fig. 2). This means the total surface area of peptides exposed to solvent is high. Over time, peptides self-assemble into supramolecular structures through favorable intermolecular interactions. This reduces their solvent accessible surface area (SASA). We calculate an aggregation propensity (AP) score as the ratio of SASA($t = 0$) to SASA(t). Larger values of AP score indicate a greater tendency of peptides to self-assemble. Our simulations reveal that this tendency for peptides to self-assemble is also directly linked to the fiber-fiber distance when there are multiple interacting fibers in solution, as shown in Figure 13B. Those peptide sequences that have high AP scores, such as VHDCVNITIK, also result in their self-assembled fibers to have more favorable interaction and aggregate in solution. Smaller peptides lacking the NMWQ fragment, with the exception of KIKIQIC, have aggregation propensities that are so high that individual fibers are not able to remain stable in solution when in close distance to other fibers and remix to form new fibrils. However, our data also suggests that similar AP scores of peptides in solution can result in very different aggregation behaviors when fibers are in contact with each other, depending on the hydrophobicity and interaction of amino acids on their surface. Thus, peptide AP score by itself is not a very reliable determining factor and simulations with multiple fibers in solution should be performed to understand formation of hierarchical structures.

In Figure 14, we show some examples of fibers that remain stable solution with varying degrees of interaction. As seen from the cross-sectional view of snapshots, CKIKQIINMWQ fibers do not aggregate in solution, but there is still some fiber-fiber interaction that keeps the fibers in close proximity. When the internal amphiphilic order of amino acids is optimized (CKIKIQINMWQ), there is a slight decrease in the fiber-fiber distance as the fibers

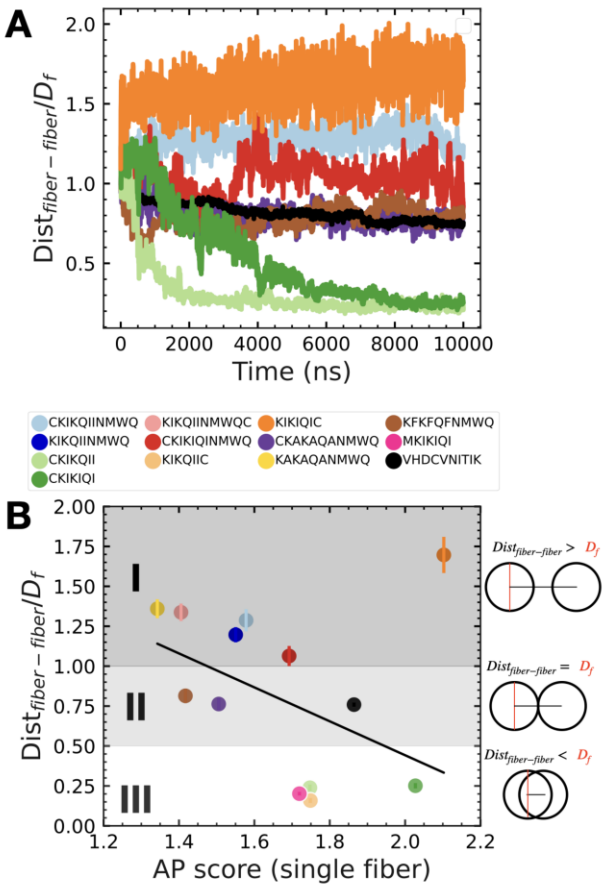


Figure 13: (A) Average fiber-fiber distance over time from coarse-grained simulations of four interacting fibers normalized by fiber diameter. A large value indicates minimal or no inter-fiber interaction, while a value close to 0 indicates remixing of peptides to form new structures. (B) Relationship between peptide aggregation propensity (AP) score from single fiber self-assembly simulations and fiber-fiber distance from multi-fiber simulations. Three regions are shown for no fiber-fiber aggregation (I), moderate fiber-fiber aggregation (II), large fiber-fiber aggregation where fibers lose stability and remix to form new fibers (III).

start to interact more with each other. The interaction between fibers happens through the displacement of water between fiber surfaces. Despite the fact that the fiber surface is decorated with charged amino acids, such as N and C-terminal residues and lysine residues, the fibers can still form aggregated structures through favorable electrostatic and hydrophobic interactions. This structure closely resembles cross β-spine structures found in Amyloid-like fibrils,[37] where a stack of steric zipper structures forms the dry interface between the β-sheets and different fibrils in solution interact through the wet interface.

When the hydrophobic isoleucine residues are replaced with alanine, CKAKAQANMWQ fibers also remain stable and aggregate in the solution. However, when investigated closely, we observe some mixing of peptides within the aggregated fiber bundle (purple and green peptides in orange fiber). This observation further confirms our previous discussion on how replacing isoleucine decreases the stability of the PNFs and is more prone to form aggregates in solution rather than fibers.

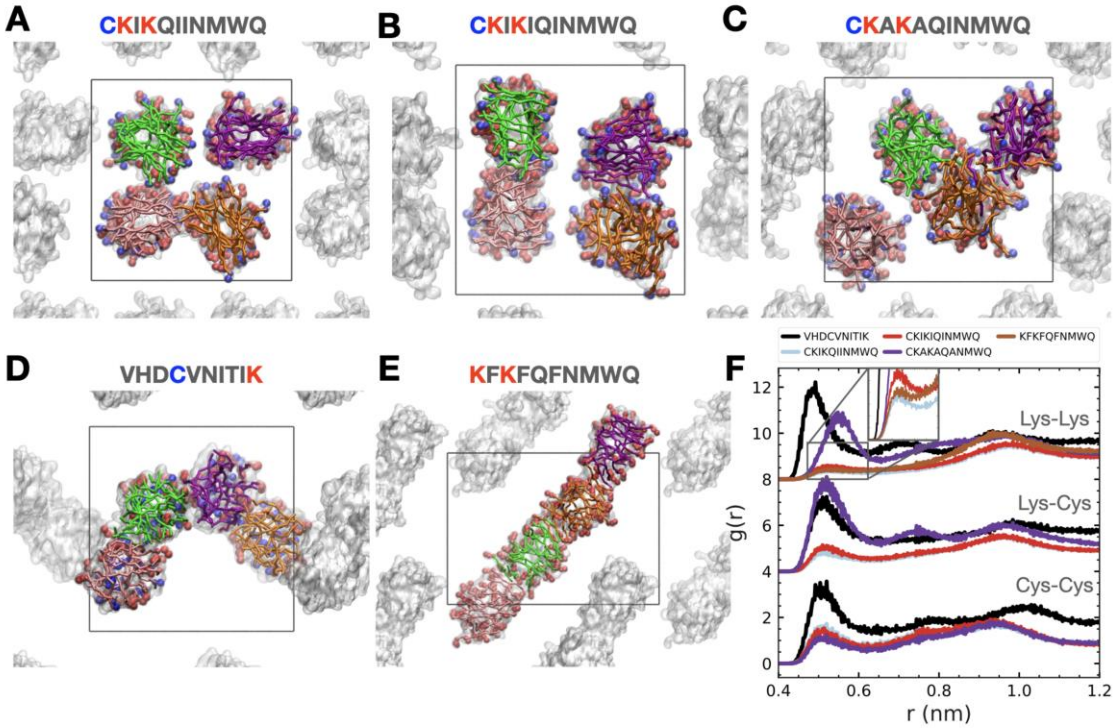


Figure 14: Cross-sectional views from 10 µs of simulations of interacting fibers that remain stable in solution for (A) CKIKQIINMWQ, (B) CKIKIQINMWQ, (C) CKAKAQANMWQ, (D) VHDCVNITIK and (E) KFKFQFNMWQ. Here, the backbones of peptides that belong to the same fiber in the initial structure (see Fig. S1) are colored with the same color. Cysteine residues are colored in red, lysine residues are blue. Size of the simulation box is shown in black to show the interaction over periodic images. The side chain bead of the peptide are colored in white. Water and ions are omitted for clarity. (F) The intermolecular pair distribution functions, g(r), between cysteine and lysine residues are shown. The g(r) for lysine-cysteine and lysine-lysine are shifted for easier viewing.

A very favorable inter-fiber interaction exists for VHDCVNITIK sequence, which is part of human prion protein. In this sequence, cysteine is not on the N-terminal and therefore it is more hydrophobic. From the cross-sectional view, we can see that most of cysteine is

buried in the core of the fiber. Furthermore, the single lysine residue is on the C-terminus of the peptide chain, meaning it has a positively-charged side chain and negatively-charged (-COO^-) backbone. This results in more favorable interactions between fiber surfaces and draws the fibers closer in solution (Fig 14D). Another example of a peptide sequence that displays a strong inter-fiber interaction is KFKFQFNMWQ. Here, the resulting aggregated fiber structure is different compared to VHDCVNITIK, where the fibers align in a way that maximizes their contact while maintaining lysine-rich surfaces on each side. Simulations performed by Sasselli et al.[29] suggested that inter-fiber interactions are enhanced by presence of hydrophilic groups on surfaces and are weakened when they are replaced by hydrophobic ones. However, our simulations show otherwise where removing N-terminal cysteine creates a more hydrophilic nanofibril surface, which prevents aggregation of peptides in solution (Fig. S22). Furthermore, switching the position of N-terminal cysteine to the C-terminus of the peptide increases intra-fiber cysteine-lysine interactions, neutralizing the surface charge. This in turn results in weakening of inter-fiber interactions and increases the distance between fibers in solution (Fig. S23,15).

Recent experiments[38] have shown evidence of disulfide bridge formation between the N-terminal cysteine residues in self-assembled fibers. Because disulfide bridge is a chemical bond that forms between the sulfur atoms of two cysteine residues, we are not able to observe it in our coarse-grained simulations. However, for this reaction to happen the side chains of cysteine residues must be very close to each other in solution. We can quantify the distribution of cysteine and positively-charged lysine residues by calculating intermolecular pair distribution functions, g(r), shown in Figure 14F. Since the VHDCVNITIK fibers are closest in solution, the magnitude of first peak for lysine-lysine, cysteine-lysine and cysteine-cysteine distributions are very large compared to other structures in Figure 14. CKAKAQANMWQ also aggregates in solution, but while lysine-lysine contacts are strong, cysteine-cysteine interaction is the weakest among the compared structures. This means disulfide bridge formation is very unlikely in CKAKAQANMWQ fibers. After VHDCVNI-

TIK, CKIKQIINMWQ shows the most favorable cysteine-cysteine interaction followed by CKIKIQINMWQ. Lysine-lysine interactions follow an opposite trend for these structures, with more closely distributed lysine residues in CKIKIQINMWQ. What is interesting is that even though the KFKFQFNMWQ fibers are highly aggregated in solution, the lysine-lysine contacts are small due their orientation with respect to each other.

We previously discussed how the lack of NMWQ sequence in the short peptides can result in loss of stability in solution. This is also true in presence of other fibers in solution, shown in Figure 15, where for both CKIKQII and CKIKIQI, the peptides in different fibers can interact and exchange to form aggregated structures. We should note that even though the initial fiber structures lose their stability, peptides from different fibers interact to form new stable fibers in solution. If we examine the newly formed fibers more closely, we see that CKIKQII forms three structures that are of different sizes, isolated in solution (Fig. 15A,D). The CKIKQII fibers do not aggregate further in the course of 10 µs of simulation. On the other hand, CKIKIQI re-assembles to form an interacting network of fibers in solution, where the supramolecular structure resembles a cross-linked network of fibers rather than isolated fibers in solution (Fig. 15B,E). Furthermore, by moving the N-terminal cysteine to C-terminus, the loss of stability in CKIKIQI is not observed in KIKIQIC, shown in Figure 15C,F. Two of the KIKIQIC fibers aggregate to form a single long fiber, while the rest remain stable and isolated in solution. The aggregation of the fibers in solution can be quantified by the number of clusters that are present (Fig. 15G), where we use a cutoff distance of 0.5 nm for peptides to be considered in the same cluster according to first peak in g(r) in Fig. 15H. The tendency to aggregate is highest for CKIKIQI compared to other sequences, which is also observed in light scattering measurements that indicate the highest microscopic aggregation for CKIKIQI (Fig. 15I).

Although both CKIKQII and KIKIQIC aggregate into three isolated fibers in solution, their morphologies are very different. CKIKQII fibers differ in size with less ordered peptides, while KIKIQIC fibers are all similar and exhibit amyloid-like arrangement with pep-

tides stacked perpendicularly with respect to the length of the fiber. KIKIQIC fibers owe their stability to the perfectly aligned peptides where the N-terminal lysine (+ charge) interact directly with C-terminal cysteine (- charge), visible in the very large magnitude of lysine-cysteine peak in Figure 15H. However, there is very little surface to surface inter-fiber interaction, which is also observed in experiments (Fig. 15I). Interestingly, although KIKIQIC has a high β-sheet content and is highly stable, its infection rates are not as high as CKIKIQI in experiments. Our simulations confirm our previous experimental observation that microscopic aggregation is another factor that is important for enhancing efficiency of PNFs in retroviral transfer.[20] Therefore, not only β-sheet formation but also fiber-fiber interactions in PNFs may be a strong indication of their ability to interact with cell membranes and thus improve biological activity.

As briefly discussed in the section "Amyloid-like self-assembly of PNFs", the mobility of peptides after they self-assemble to form fibers in solution may be different depending on the peptide-peptide and peptide-solvent interactions. This, in turn, can potentially have consequences on the stability of fibers, their interaction with surroundings, retroviral enhancement properties, etc. We quantified the mobility of peptides by calculating the root mean square fluctuation (RMSF) of each residue from our coarse-grained simulations. For all peptides, RMSF of amino acids in a single fiber in solution (Fig. 16A) is smaller and varies less along the peptide chain compared to the RMSF of amino acids when fibers are interacting in solution (Fig. 16B). Interestingly, for peptides with favorable fiber-fiber interaction, we observe an increase in the motion of the residues closer to the C-terminus, whereas the N-terminal amino acids in the peptide sequence are more stationary. This result along with qualitative examination of simulation trajectories, show that even after the fibers have reached an equilibrium structure, i.e. the fiber-fiber distance does not change, there is considerable motion in the peptides. This result shows that the dynamic behavior of these supramolecular fibers are different when individual fibers aggregate in solution to form "bundles" as opposed to fibers in dilute solutions surrounded by solvent. Furthermore,

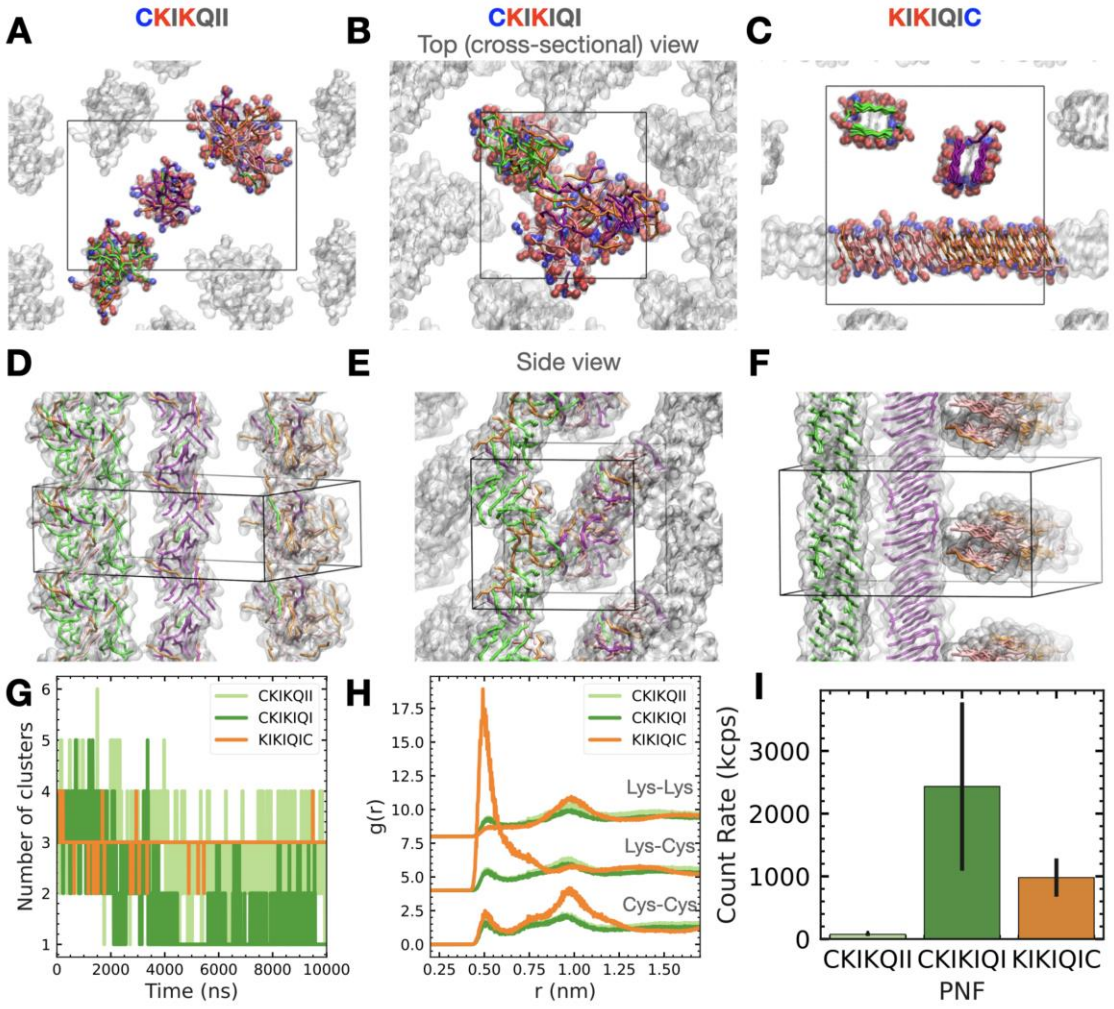


Figure 15: The (A, B, C) cross-sectional and (D, E, F) side views of CKIKQII, CKIKIQI and KIKIQIC fibers, respectively, in solution. The peptide backbones that belong to the same fiber in the initial structure are colored with the same color. Cysteine residues are colored in red, lysine residues are blue. Size of the simulation box is shown in black to show the interaction over periodic images. The side chain bead of the peptide are colored in white. Water and ions are omitted for clarity. (G) The number of clusters (i.e. fibers) over the simulation time, showing remixing and formation of new fibrillar structures. (H) The intermolecular pair distribution functions, g(r), between lysine and cysteine groups. (I) Count rate of scattered light from zeta-potential measurements, indicating microscopic aggregation.

large RMSF values are observed for fibers that do not aggregate in solution (e.g. CKIKQI-INMWQ) or those that lose stability and remix to form new fibers (e.g. CKIKQII), but small RMSF values are found for those structures where fiber aggregation is the strongest (e.g. VHCVNITIK and KFKFQFNMWQ). Thus, when fibers aggregate to form hierarchical

assemblies, the mobility of peptides that make up the supramolecular fiber is reduced.

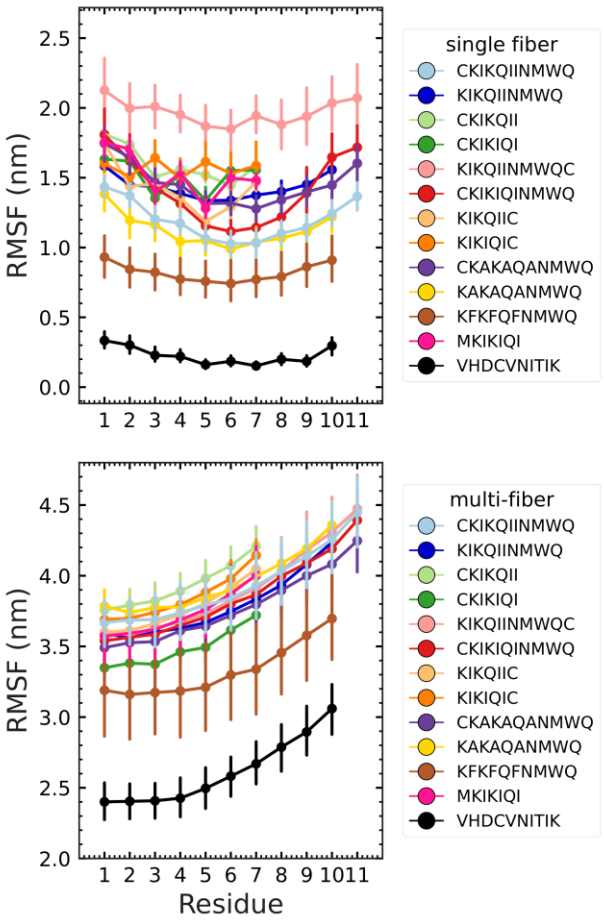


Figure 16: Average room mean square fluctuation (RMSF) for each residue calculated from coarse-grained single fiber (top) and multi-fiber (bottom) simulations. For single fiber the last 4 µs of 15 µs simulation, and for the multi-fiber system the last 2 µs of 10 µs simulation is used for averaging. Error bars are calculated from the standard deviation between different peptides in the system.

But how does this dynamic behavior affect the stability of fibers? While the aggregated fibers remain stable in solution in our coarse-grained simulations with multiple interacting fibers, we do not observe any change in fiber morphologies over the 10 µs of simulation time. However, since these simulations are performed with coarse-grained models, we cannot make direct connections to their secondary structure or β-sheet content. Furthermore, we computed the RMSF from atomistic model of a single fiber in solution, both self-assembled and pre-formed, and the corresponding β-sheet formation in these structures (Fig. S27).

We find that there is no correlation (r^2 = 0.0055 and 0.035 for self-assembled and pre-formed single fibers, respectively) between number of β-sheets and RMSF of peptides. It should be noted that this is only true for single fibers in solution. We do not know whether peptide mobility in aggregated fibers have different consequences on their β-sheet content. Larger scale atomistic simulations of aggregated fibers would need to be performed, which is something we would like to pursue in future studies.

Discussion

Our simulations and experiments show that the primary structure (amino acid sequence) of a peptide is an important factor in determining the morphology of the self-assembled structures, whether it's a long amyloid-like fibril or small aggregates in solution. Although experimental and simulation conditions, such as peptide concentration or the type of model used, can lead to differences in the observed structures, we can learn from the trends in our data to derive structure-property relationships. Figure 17 show how properties correlate with the infectivity rates for the subset of peptide sequences we have chosen for our simulations. In the time scale of experiments, which are much longer than our simulations, we expect to find a variety of structures. Therefore, in this plot the percentage of β-sheets calculated from MD simulations is an average of self-assembled and pre-formed atomistic simulations. Because both experimental and MD β-sheet amounts depend on conditions, such as peptide concentration, here we compare the trends in our data instead of the absolute quantity. Overall, there is remarkable agreement between the simulations and experiments, where (1) the presence of N-terminal cysteine increases β-sheet formation for isoleucine containing peptides (KIKQIINMWQ vs CKIKQIINMWQ), but not for alanine containing peptides (KAKAQANMWQ vs CKAKAQANMWQ); (2) switching cysteine position from N-terminal to C-terminal has no effect on the β-sheet amount (CKIKIQI vs KIKIQIC); (3) removing the C-terminal NMWQ eliminates formation of β-sheets in the original sequence (CKIKQIIN-

MWQ vs CKIKQII); (4) optimization of the amphiphilic pattern induces β-sheet formation even for short sequences (CKIKQII vs CKIKIQI); (5) replacing the N-terminal cysteine with more hydrophobic methionine causes loss of β-sheet rich structure (CKIKIQI vs MKIKIQI) and finally (6) increasing the hydrophobicity of internal residues by replacing alanine with phenylalanine greatly improves the β-sheet formation (KAKAQANMWQ vs KFKFQFNMWQ).

While all of the peptides self-assemble into supramolecular fibrils if the peptide concentration is high enough, this alone does not guarantee that they will remain stable when in contact with other fibers or cell membranes in solution. Peptides with high aggregation propensity have favorable peptide-peptide interactions that trigger self-assembly in solution and also induce aggregation of fibers. Formation of β-sheets is one of the factors that make these supramolecular fibrils stable. However, if the peptide-peptide interactions are too strong, fibers in solution lose stability and remix to form new fibers. Therefore, the dynamic behaviour of peptides in self-assembled fibrils become important.

If we investigate the differences in mobility with respect to the specific changes in the peptide sequence already discussed in this study, we can make several observations. Considering there are inter-fiber interactions in solution (Fig. 16B), firstly, removing the N-terminal cysteine leads to a slight decrease in the mobility (CKIKQIINMWQ vs KIKQIINMWQ). CKIKQIINMWQ fibers have a closer fiber-fiber distance on average compared to KIKQIINMWQ (Fig. S22), but both of these fibers do not aggregate to form bundles in solution. Therefore, the increased fiber-fiber interaction in CKIKQIINMWQ might be the reason for the increased mobility of peptides. However, for the alanine counterpart (CKAKAQANMWQ vs KAKAQANMWQ), we see the opposite effect, because presence of N-terminal cysteine leads to aggregation of fibers in CKAKAQANMWQ. In contrast, the aggregated fibers have lower mobility in solution. Removing the C-terminal NMWQ sequence doesn't have any effect for CKIKQIINMWQ vs CKIKQII, but we observe a decrease in the mobility for the optimized amphililic order (CKIKIQINMWQ vs CKIKIQI). Changing the amphiphilic

order of amino acids decreases the mobility of peptides (CKIKQIINMWQ vs CKIKIQINMWQ and CKIKQII vs CKIKIQI), since better packing of the peptides create a more stable structure. Replacing the N-terminal cysteine with methionine increases mobility (CKIKIQI vs MKIKIQI). We also observe an increase in mobility for isoleucine to alanine substitution (CKIKIQINMWQ vs CKAKAQANMWQ), while alanine to phenylalanine substitution decreases the mobility significantly (KAKAQANMWQ vs KFKFQFNMWQ).

While dynamic nature of these supramolecular peptide nanofibers is one of the properties that sets these materials apart from conventional polymers, so far it is not clear how this dynamic property affects their function. Thus, we also investigated whether the peptide mobility has an effect on the biological activity of the PNFs. As previously discussed, the number of β-sheets from both experiments and simulations is positively correlated with infection rates in cells. We also find that there is a negative correlation between RMSF of peptides and infectivity in multi-fiber simulations where there is fiber-fiber interaction, shown in Figure 17(top panel). Our results suggest aggregation of fibrillar structures reduces the mobility of peptide chains in the supramolecular fibril and formation of hierarchical structures through fiber-fiber aggregation correlate positively with bioactivity. It is unclear what role the density of fibrils and their aggregation play in the mechanism of retroviral transduction, but is of utmost interest for future studies. However, this is not true if we consider the RMSF from atomistic and coarse-grained single-fiber simulations (Figure 17, bottom panel). These results confirm that while single fiber simulations in solution are beneficial in providing structural properties, such as β-sheet formation, fiber-fiber interactions provide important information about the dynamical behavior of peptides and should be taken into account when structure-morphology-property relationships are investigated.

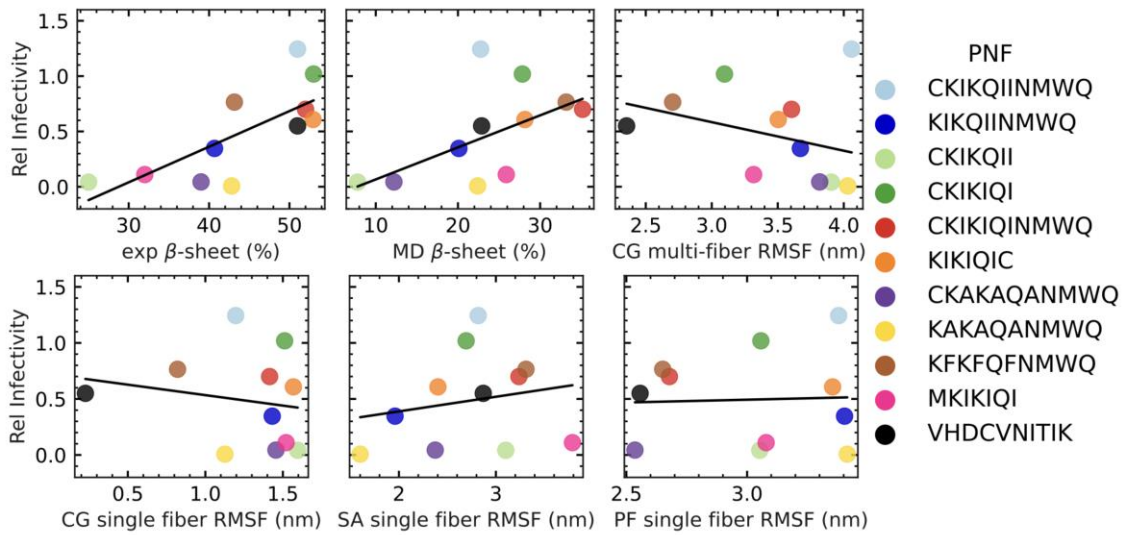


Figure 17: Correlation between different measured properties and relative infectivity. While we observe correlation for the data in the top row, there is no direct link between properties shown in the bottom row and infectivity. Here, SA and PF single fiber simulations refer to self-assembled (backmapped) and pre-formed atomistic simulations of single fibers in solution.

Conclusions

Understanding structure-property relationships of supramolecular peptide nanofibrils is important for identifying amino acid sequences to facilitate the discovery of functional materials with optimized properties. We previously reported a library of PNFs and their properties for enhancing retroviral gene transfer. In this work, we chose a subset of peptide sequences from our previous work in order to determine how specific changes in the primary structure, such as sequence truncation, reordering and amino acid substitutions, affect the morphology of self-assembled supramolecular structures, their stability, secondary structure formation and aggregation in solution. Using a multiscale modeling approach, we combined coarse-grained and atomistic simulations to make connections between microscopic properties and biological activity of PNFs.

Our results have shown that sequence truncation directly affects the stability of the supramolecular fibrils in solution. Presence of C-terminal cysteine increase the hydrophobicity of PNF surface by the shielding the β-sheet forming region of the peptide from water

and ions. Truncating the sequence by removing C-terminal cysteine reduces the number of β-sheets. Furthermore, when the N-terminal NMWQ motif is missing, less stable fibers with very low amount of β-sheets form and the stability of fibrils can be recovered only with reordering of the internal amino acids to make an amphiphilic pattern. Amino acid substitutions with more hydrophobic residues improves the properties of fibrils when the substitution is done on the internal residues. However, making the C-terminus more hydrophobic causes the disruption of the β-sheet forming region.

We not only investigate the self-assembly of peptides to form fibrils, but also the dynamic behavior of peptides after they have formed supramolecular fibrils. While high aggregation propensity of peptides leads to fast self-assembly in solution, it can also be detrimental to the stability of fibrils when in close proximity of other fibrils in solution. On the other hand, low aggregation propensity means no aggregated hierarchical structures, which are highly correlated with increased biological activity. Those peptide sequences with moderate aggregation propensity self-assemble to form fibrils and also show stable, aggregated fibrils in solution. Because these PNFs are inspired by Amyloid fibrils, we validated our model by comparing the self-assembly of our chosen peptide sequences to a short, β-sheet forming region of human prion protein, which is one of the naturally-occuring Amyloid-forming proteins that accumulate in the brain leading to a fatal, neurodegenerative disease. Thus, we have shown that our multiscale approach is applicable for studying the properties of both synthetic and naturally-occuring Amyloid-like peptide nanofibrils.

Acknowledgement

This study was funded by the Deutsche Forschungs-gemeinschaft (DFG, German Research Foundation)—Projektnummer316249678—SFB 1279 (A05, C01) and Projectnumber 441734479.

Supporting Information Available

A detailed description of the methods and additional simulation data can be found in the supplementary information.

References

(1) Mouw, J. K.; Ou, G.; Weaver, V. M. Extracellular matrix assembly: a multiscale deconstruction. *Nature Reviews Molecular Cell Biology* **2014**, *15*, 771–785.

(2) Sunde, M.; Blake, C. C. F. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly Reviews of Biophysics* **1998**, *31*, 1–39.

(3) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. *Science* **2001**, *294*, 1684–1688.

(4) Yolamanova, M. et al. Peptide nanofibrils boost retroviral gene transfer and provide a rapid means for concentrating viruses. *Nature Nanotechnology* **2013**, *8*, 130–136.

(5) Yan, X.; He, Q.; Wang, K.; Duan, L.; Cui, Y.; Li, J. Transition of Cationic Dipeptide Nanotubes into Vesicles and Oligonucleotide Delivery. *Angewandte Chemie International Edition* **2007**, *46*, 2431–2434.

(6) Li, C.; Iscen, A.; Sai, H.; Sato, K.; Sather, N. A.; Chin, S. M.; Álvarez, Z.; Palmer, L. C.; Schatz, G. C.; Stupp, S. I. Supramolecular–covalent hybrid polymers for light-activated mechanical actuation. *Nature Materials* **2020**, *19*, 900–909.

(7) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stöhr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V. Short peptides self-assemble to produce catalytic amyloids. *Nature Chemistry* **2014**, *6*, 303–309.

(8) Yemini, M.; Reches, M.; Rishpon, J.; Gazit, E. Novel Electrochemical Biosensing Platform Using Self-Assembled Peptide Nanotubes. *Nano Letters* **2005**, *5*, 183–186, PMID: 15792436.

(9) Kholkin, A.; Amdursky, N.; Bdikin, I.; Gazit, E.; Rosenman, G. Strong Piezoelectricity in Bioinspired Peptide Nanotubes. *ACS Nano* **2010**, *4*, 610–614, PMID: 20131852.

(10) Ulijn, R. V.; Smith, A. M. Designing peptide based nanomaterials. *Chem. Soc. Rev.* **2008**, *37*, 664–675.

(11) Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J. Biomimetic peptide self-assembly for functional materials. *Nature Reviews Chemistry* **2020**, *4*, 615–634.

(12) Cherny, I.; Gazit, E. Amyloids: Not Only Pathological Agents but Also Ordered Nanomaterials. *Angewandte Chemie International Edition* **2008**, *47*, 4062–4069.

(13) Schilling, C.; Mack, T.; Lickfett, S.; Sieste, S.; Ruggeri, F. S.; Sneideris, T.; Dutta, A.; Bereau, T.; Naraghi, R.; Sinske, D.; Knowles, T. P. J.; Synatschke, C. V.; Weil, T.; Knöll, B. Sequence-Optimized Peptide Nanofibers as Growth Stimulators for Regeneration of Peripheral Neurons. *Advanced Functional Materials* **2019**, *29*, 1809112.

(14) Álvarez, Z.; Kolberg-Edelbrock, A. N.; Sasselli, I. R.; Ortega, J. A.; Qiu, R.; Syrgiannis, Z.; Mirau, P. A.; Chen, F.; Chin, S. M.; Weigand, S.; Kiskinis, E.; Stupp, S. I. Bioactive scaffolds with enhanced supramolecular motion promote recovery from spinal cord injury. *Science* **2021**, *374*, 848–856.

(15) Sieste, S.; Mack, T.; Lump, E.; Hayn, M.; Schütz, D.; Röcker, A.; Meier, C.; Kaygisiz, K.; Kirchhoff, F.; Knowles, T. P. J.; Ruggeri, F. S.; Synatschke, C. V.; Münch, J.; Weil, T. Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction. *Advanced Functional Materials* **2021**, *31*, 2009382.

(16) Li, D.; Furukawa, H.; Deng, H.; Liu, C.; Yaghi, O. M.; Eisenberg, D. S. Designed amyloid fibers as materials for selective carbon dioxide capture. *Proceedings of the National Academy of Sciences* **2014**, *111*, 191–196.

(17) Li, D.; Jones, E. M.; Sawaya, M. R.; Furukawa, H.; Luo, F.; Ivanova, M.; Sievers, S. A.; Wang, W.; Yaghi, O. M.; Liu, C.; Eisenberg, D. S. Structure-Based Design of Functional Amyloid Materials. *Journal of the American Chemical Society* **2014**, *136*, 18044–18051.

(18) Wagoner, V. A.; Cheon, M.; Chang, I.; Hall, C. K. Impact of sequence on the molecular assembly of short amyloid peptides. *Proteins: Structure, Function, and Bioinformatics* **2014**, *82*, 1469–1483.

(19) Betush, R. J.; Urban, J. M.; Nilsson, B. L. Balancing hydrophobicity and sequence pattern to influence self-assembly of amphipathic peptides. *Peptide Science* **2018**, *110*, e23099.

(20) Kaygisiz, K.; Rauch-Wirth, L.; Dutta, A.; Yu, X.; Nagata, Y.; Bereau, T.; Münch, J.; Synatschke, C. V.; Weil, T. Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides. *ChemRxiv* **2023**, This content is a preprint and has not been peer-reviewed.

(21) White, S. H.; Wimley, W. C. Hydrophobic interactions of peptides with membrane interfaces. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1998**, *1376*, 339–352.

(22) Fujiwara, K.; Toda, H.; Ikeguchi, M. Dependence of α-helical and β-sheet amino acid propensities on the overall protein fold type. *BMC Structural Biology* **2012**, *12*, 18.

(23) Cui, H.; Cheetham, A. G.; Pashuck, E. T.; Stupp, S. I. Amino Acid Sequence in Constitutionally Isomeric Tetrapeptide Amphiphiles Dictates Architecture of One-Dimensional Nanostructures. *Journal of the American Chemical Society* **2014**, *136*, 12461–12468, PMID: 25144245.

(24) Lee, O.-S.; Stupp, S. I.; Schatz, G. C. Atomistic Molecular Dynamics Simulations of Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers. *Journal of the American Chemical Society* **2011**, *133*, 3677–3683, PMID: 21341770.

(25) Iscen, A.; Schatz, G. C. Peptide amphiphile self-assembly(a). *Europhysics Letters* **2017**, *119*, 38002.

(26) Lee, O.-S.; Cho, V.; Schatz, G. C. Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics. *Nano Letters* **2012**, *12*, 4907–4913, PMID: 22924639.

(27) Kang, M.; Cui, H.; Loverde, S. M. Coarse-grained molecular dynamics studies of the structure and stability of peptide-based drug amphiphile filaments. *Soft Matter* **2017**, *13*, 7721–7730.

(28) Fu, I. W.; Markegard, C. B.; Chu, B. K.; Nguyen, H. D. Role of Hydrophobicity on Self-Assembly by Peptide Amphiphiles via Molecular Dynamics Simulations. *Langmuir* **2014**, *30*, 7745–7754, PMID: 24915982.

(29) Sasselli, I. R.; Syrgiannis, Z.; Sather, N. A.; Palmer, L. C.; Stupp, S. I. Modeling Interactions within and between Peptide Amphiphile Supramolecular Filaments. *The Journal of Physical Chemistry B* **2022**, *126*, 650–659, PMID: 35029997.

(30) Batra, R.; Loeffler, T. D.; Chan, H.; Srinivasan, S.; Cui, H.; Korendovych, I. V.; Nanda, V.; Palmer, L. C.; Solomon, L. A.; Fry, H. C.; Sankaranarayanan, S. K. R. S. Machine learning overcomes human bias in the discovery of self-assembling peptides. *Nature Chemistry* **2022**, *14*, 1427–1435.

(31) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation* **2013**, *9*, 687–697, PMID: 26589065.

(32) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. J. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ^1 and χ^2 Dihedral Angles. *Journal of Chemical Theory and Computation* **2012**, *8*, 3257–3273, PMID: 23341755.

(33) Prusiner, S. B. Prions. *Proceedings of the National Academy of Sciences* **1998**, *95*, 13363–13383.

(34) Wang, L.-Q.; Zhao, K.; Yuan, H.-Y.; Wang, Q.; Guan, Z.; Tao, J.; Li, X.-N.; Sun, Y.; Yi, C.-W.; Chen, J., et al. Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nature Structural & Molecular Biology* **2020**, *27*, 598–602.

(35) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *Journal of Chemical Theory and Computation* **2014**, *10*, 676–690, PMID: 26580045.

(36) Nick Pace, C.; Martin Scholtz, J. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophysical Journal* **1998**, *75*, 422–427.

(37) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-β spine of amyloid-like fibrils. *Nature* **2005**, *435*, 773–778.

(38) Kaygisiz, K.; Ender, A. M.; Gačanin, J.; Kaczmarek, L. A.; Koutsouras, D. A.; Nalakath, A. N.; Winterwerber, P.; Mayer, F. J.; Räder, H.-J.; Marszalek, T.; Blom, P. W. M.; Synatschke, C. V.; Weil, T. Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning. *Macromolecular Bioscience n/a*, 2200294.

Supporting Information

Supporting Information: Unraveling structure-property relationships in self-assembling peptide nanofibrils via simulations

Aysenur Iscen, Kübra Kaygisiz, Christopher V. Synatschke, Tanja Weil, and Kurt Kremer*

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

E-mail: kremer@mpip-mainz.mpg.de

Computational Methods

Because self-assembly may take a long time (ns to µs), we performed coarse-grained (CG) simulations starting from randomly distributed peptides in solutions using three different concentrations: low concentration (40 peptides/1000 nm^3 or 0.04 peptides/nm^3), intermediate concentration (80 peptides/1000 nm^3 or 0.08 peptides/nm^3) and high concentration (20 peptides/125 nm^3 or 0.16 peptides/nm^3). The high concentration simulations were designed with fewer peptides in a smaller simulation box in order to force formation of a single fiber instead of aggregated structures.

After coarse-grained self-assembly simulations, we converted or "backmapped" our coarse-grained structures to all-atom (AA) model to gain additional information that can only be obtained from atomistic models, such as peptide secondary structure (β-sheet, helix, random coil), hydrogen bonding between peptides, hydrogen bonding with water, water configurations at the peptide surface, etc. When structures obtained from coarse-grained simulations

are backmapped, sometimes the packing of molecules may not be perfect either due to low precision of interactions between molecules or time constraints. Therefore, we also performed atomistic simulations of perfectly aligned, pre-formed peptide nanofibrils for comparison. This latter approach is useful in order to give us some ideas about ideal packing situations, which may exist in experiments, and is a good way to interpret the results of coarse-grained simulations where peptides are allowed to self-assemble in solution without any preconceived notion about what the structures may look like.

Coarse-grained simulations

We performed molecular dynamics (MD) simulations of self-assembly of peptides. For each peptide sequence, we used Packmol[S1] to randomly place peptides in a box for three peptide concentrations:

1. Low concentration: 40 peptides in 10 nm x 10 nm x 10 nm initial box size (0.04 peptides/nm^3, 0.066 M)

2. Intermediate concentration: 80 peptides in 10 nm x 10 nm x 10 nm initial box size (0.08 peptides/nm^3, 0.13 M)

3. High concentration: 20 peptides in 5 nm x 5 nm x 5 nm initial box size (0.16 peptides/nm^3, 0.26 M)

The coarse-grained peptides were modeled with the MARTINI force field version 2.2.[S2] Each simulation box was solvated using Polarizable MARTINI water model.[S3] We also added 0.138 M NaCl to mimic experiments and additional counterions (Cl^-) to balance the total charge in the system. All coarse-grained simulations were performed with Gromacs 2019.4[S4,S5] software according to the following protocol. We first minimized for 10,000 steps using steepest descent method. Then, we ran a constant volume (NVT) simulation for 1 ns using 10 fs time step at 298 K in order to equilibrate the temperature. In the second step of the equilibration, we ran a constant pressure (NPT) simulation with 20 fs time

step for 2 µs at 298 K and 1 bar pressure to equilibrate the volume of the box. Following the equilibration, we performed a production run for 5 µs using NPT ensemble with 20 fs time step for the low and intermediate concentration simulations and 15 µs for the high concentration simulations. We used the v-rescale coupling method[S6] for temperature and Parrinello-Rahman barostat[S7,S8] for pressure coupling. The compressibility for the pressure coupling was 3 x 10^{-4} bar^{-1}. Bond constraints were handled using the LINear Constraint Solver (LINCS) algorithm.[S9] The Verlet cutoff scheme[S10] was used for neighbor searching. Long-range electrostatics were treated with the reaction field method. A cut-off of 1.1 nm was used for evaluation of all non-bonded interactions. For electrostatic interactions, we used the group method with dielectric constant 2.5, which is appropriate for polarizable Martini water model. The neighbor list was updated every 10 steps. Atomic coordinates were saved every 500 ps for the trajectory analysis. Periodic boundary conditions were set in x, y and z directions. All snapshots from simulations were rendered using the Visual Molecular Dynamics (VMD) software.[S11]

We also performed simulations, where four self-assembled fibers from high concentration CG simulations were placed in a box at a close distance as shown in Fig. S1. The initial box size along the length of the fiber (z) was taken directly from the high concentration self-assembly simulations ($\sim$ 5 nm) and the box size for the other two dimensions (x, y) were set to 10 nm. There was a total of 80 peptides (20 x 4) in each box. The simulation boxes were solvated and ionized as before. The same equilibration protocol was used for these simulations, but the total simulation time was 10 µs.

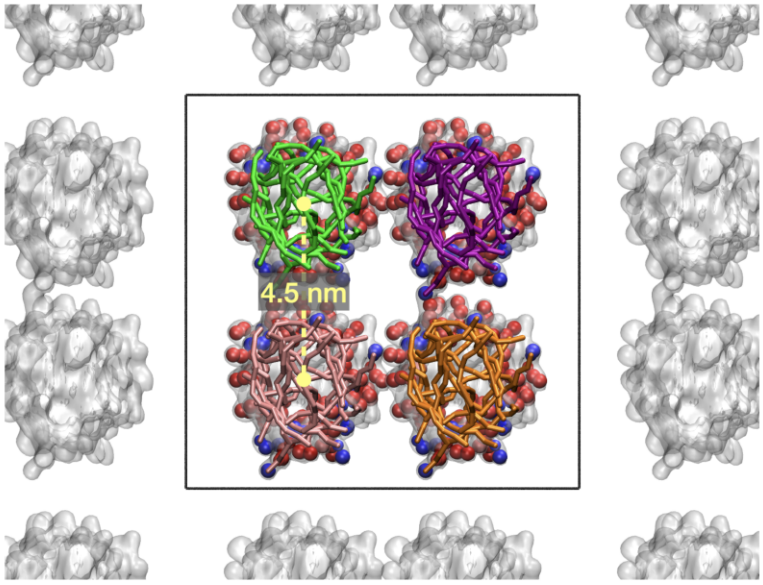


Figure S1: The cross-sectional view of initial configuration of four peptides placed with fiber-fiber center of mass distance of 4.5 nm. Peptides that belong to the same fiber have backbones colored green, purple, pink and orange. Cysteine residues are colored in blue, lysine is red. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity.

Backmapping of CG to AA model

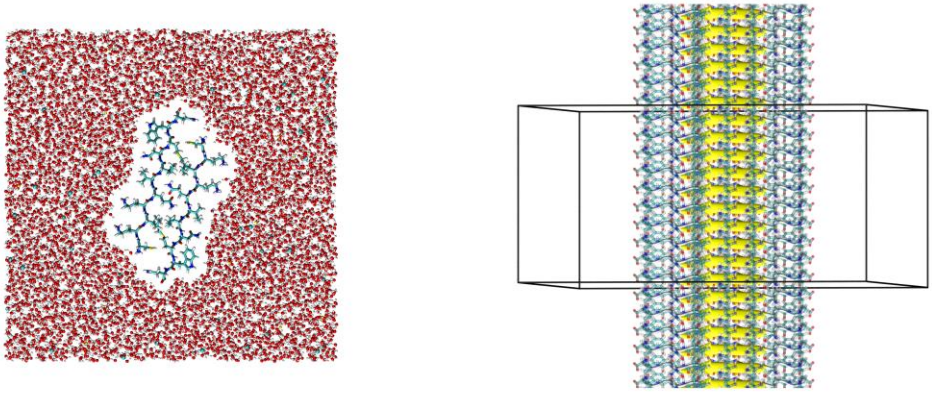
Due to limitations on system size and timescales associated with atomistic simulations, we only backmapped our high concentration CG simulations to atomistic model. This is useful for our analysis because due to high concentrations of peptides, we observe a single, self-assembled fiber for all peptide sequences at this concentration with the CG model. However, even at the CG level, some of these fibers are more stable than others. By converting our CG fibers to atomistic representation, we gain information on stability of the fibers in solution and secondary structure formation. For the backmapping procedure, we used our equilibrated single fibers (after 15 µs) as initial coordinates and converted them into full atomistic resolution according to the CHARMM36[S12] force field. We minimized the backmapped peptides using two steps of minimization: 1) 500 steps with steepest descent method where nonbonded interactions between peptides are excluded and 2) 500 steps with steepest descent method without any exclusions. After the short minimization, we solvated

and ionized the atomistic fiber with TIP3P water[S13] and 0.138 M NaCl as before. After this step, following simulation protocol was used for equilibration of the atomistic structures. We first minimized the system for 1000 steps, followed by a short NVT equilibration at 298 K for 200 ps, where the peptide was position constrained. Following this, we performed an NPT equilibration at 298 K and 1 bar using isotropic pressure coupling for 100 ns. An addition production NPT simulation was performed for 150 ns at the same temperature and pressure, but using semiisotropic pressure coupling to allow for box size along the z direction to change independently from x and y. For atomistic simulations, the time step was 2 fs. Long-range electrostatics were determined with the smooth particle mesh Ewald (PME)[S14] method using cubic interpolation and Fourier grid spacing of 0.16 nm. Van der Waals interactions were smoothly switched to zero between 1.0 and 1.2 nm. A cut-off of 1.2 nm was used for evaluation of all non-bonded interactions. Atomic coordinates were saved every 100 ps for the trajectory analysis. The compressibility for the pressure coupling was 4.5 x 10^{-5} bar^{-1}. Only hydrogen bonds were constrained using the LINCS algorithm.[S9] Secondary structure analysis was performed using the STRIDE algorthym[S15] included in VMD.[?]

Atomistic simulations of pre-formed fibers

In order to investigate the properties of perfectly aligned fibers, we performed MD simulations of pre-formed fibers. Since we know from experiments that these peptides self-assemble into Amyloid-like fibers with high β-sheet content, our initial structure was designed in order to optimize intermolecular hydrogen bonding between peptides by stacking 10 peptides along the length of the fiber with a peptide-peptide center of mass distance of 0.45 nm. This allows for a close packing of peptides that allows for making intermolecular hydrogen bonding. From our previous CG self-assembly simulations, we observed that almost all peptides assembled to form dimer structures, where the hydrophobic residues of two peptides face each other, while hydrophilic residues remain exposed to the surface of the fiber. Therefore, we placed

another 10 peptides stacked in the same across from the first layer to form a dimer making sure that the charged lysine residues remain on the surface of the fiber. This pre-formed fiber structure was modeled with CHARMM36 force field[S12] and solvated with TIP3P water[S13] and 0.138 M NaCl as described in the previous section. The same MD settings were used for the equilibration and production runs. As discussed in the section before, the box volume was equilibrated firs with isotropic coupling, followed by semiisotropic coupling using the Parrinello-Rahman barostat.[S7,S8] The total simulation run was 300 ns with a 2 fs time step.



(a) Top view (b) Side view

Figure S2: The initial (a) cross-sectional top and (b) side views of the pre-formed fibers. Along the cross-section of the fiber (top view), two peptide chains antiparallel to each other are placed so that their lysine residues are on the surface of the fiber, while hydrophobic isoleucine residues are buried in the core. Along the length of the fiber (side view), ten peptides are placed with an initial center of mass distance of 4.5 nm. There are a total of 20 peptides in this system. Simulation box is shown in black. Peptides only interact with their periodic image along the length of the fiber in order to simulate an infinitely long fiber. Water and ions are omitted for clarity.

A summary of all simulations performed in this study is shown below in Table S1. We tested a total of 13 peptide sequences and performed 6 different simulations for each peptide sequence. Four of these simulations were coarse-grained simulations using the Martini model, and the other two were atomistic simulations using the CHARMM36 force field. The total simulation time for a single peptide sequence was 35,000 ns (35 µs) and 550 ns for coarse-grained and atomistic simulations, respectively. In total, we ran 455 µs of coarse-grained simulations and 7150 ns of atomistic simulations.

Table S1: Summary of simulation details for different types of coarse-grained and atomistic simulations performed in this study.[1]

Peptide sequences	Force field	Type of simulation	Number of peptides	Number of water	Number of ions	Total number of beads/atoms	Simulation time (ns)
CKIKQIINMWQ, CKIKIQINMWQ, KIKQIINMWQ, KIKQIINMWQC, CKIKQII, CKIKIQI, MKIKIQI, KIKQIIC, KIKIQIC, CKAKAQANMWQ, KAKAQANMWQ, KFKFQFNMWQ, VHDCVNITIK	Coarse-grained Martini	Self-assembly, low concentration (0.066 M)	40	6992 (PW model)	83 Na, 163 Cl	22302	5,000
		Self-assembly, intermediate concentration (0.13 M)	80	2706 (PW model)	42 Na, 202 Cl	21968	5,000
		Self-assembly, high concentration (0.26 M)	20	628 (PW model)	10 Na, 50 Cl	2484	15,000
		Interacting fibers (4 self-assembled, high concentration fibers)	80	2563 (PW model)	42 Na, 202 Cl	10093	10,000
	All-atom CHARMM36	Self-assembled, high concentration backmapped from Martini	20	7138 (TIP3P model)	21 Na, 61 Cl	25576	250
		Pre-formed fiber	20	7433 (TIP3P model)	21 Na, 61 Cl	26461	300

[1]The number of peptides for each simulation is same for all 13 peptide sequences studied. However, due to differences in the peptide sequences, the number of water and ions (total number of bead/atoms) slightly differ. The numbers given in the table refer to simulations with CKIKQIINMWQ peptide sequence.

Experimental Methods

Supporting Figures for "Amyloid-like self-assembly of PNFs"

One of the factors that significantly affect the morphology of the self-assembled structures is the peptide concentration both in experiments and simulations. Because we use periodic conditions in our simulations, depending on the peptide concentration, we can observe different morphologies such as aggregates or continuous fibrillar structures. Therefore, we consider three different concentrations as discussed in the Methods section. Both VHDCVNITIK and CKIKQIINMWQ self-assemble to form fibers in low (0.04 peptides/nm^3) concentration, shown in Fig. 2 in main text. In the next step, we keep the box size constant and double the number of peptides to investigate whether similar morphologies are available at intermediate (0.08 peptides/nm^3) concentration. As seen in the figure below, the peptides still self-assemble into fibers, however the fiber twists to accommodate the larger number of peptides.

We also consider a high peptide concentration (0.16 peptides/nm^3), but with a smaller box size and less number of peptides. This is done in order to investigate how a small section of the PNF looks like when it self-assembles. The morphologies for CKIKQIINMWQ and VHDCVNITIK are shown in Fig. 3 in the main text. Below we show how the solvent accessible surface area changes along the length of the peptide chain calculated from atomistic model after backmapping and 250 ns of simulation. Although the concentration is different and the model is no longer coarse-grained, the SASA trends are very similar to those for low concentration simulations.

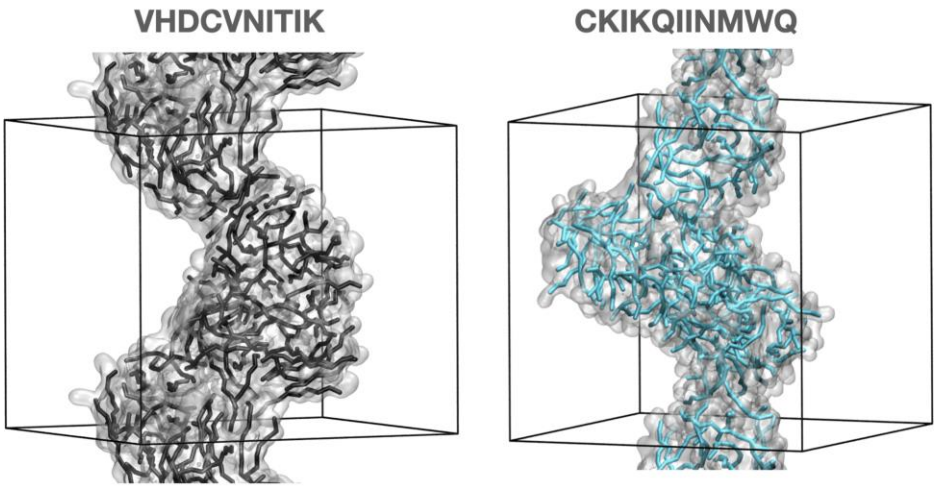


Figure S3: Snapshots from intermediate concentration (0.08 peptides/nm^3) coarse-grained self-assembly simulations after 5 µs. Only peptide backbones are colored for easier view of peptide assembly. The rest of the peptides are shown in white. Water and ions are omitted for clarity.

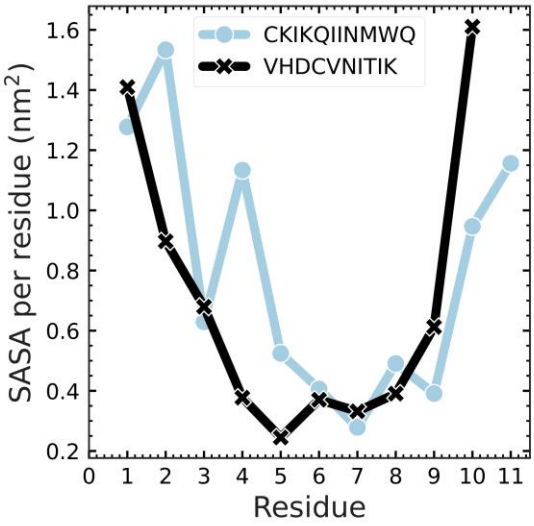


Figure S4: Average solvent accessible surface area (SASA) per residue calculated from self-assembled atomistic fiber simulations.

Supporting Figures for "Role of N-terminal cystine"

The structures from self-assembly simulations at low and intermediate concentrations are shown in the Figure below for KIKQIINMWQ peptide sequence. This peptide self-assembles

to form a fiber-like segment at both concentrations, which is not continuous over any of the periodic boundaries of the simulation box. Surface hydrophobicity is very similar except for the N-terminus, where the cysteine amino acid is missing, making this fiber less hydrophobic. This enhanced interaction of the N-terminus with water prevents the peptide to form a continuous fiber for these two concentrations.

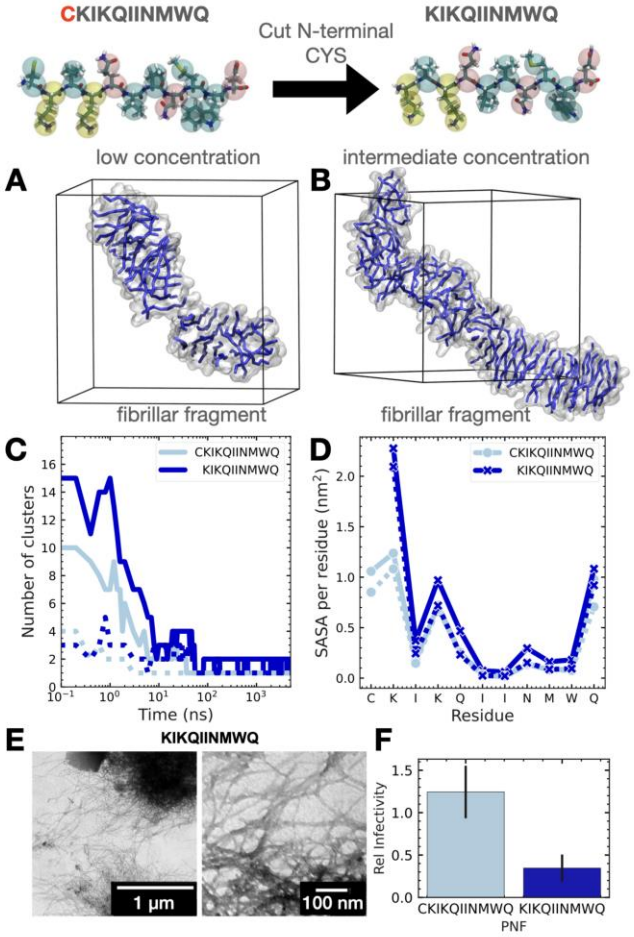


Figure S5: Snapshots from (A) low and (B) intermediate concentration coarse-grained self-assembly simulations after 5 µs for KIKQIINMWQ. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (C) Number of clusters during self-assembly as a function of simulation time. (D) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration. (E) TEM images for KIKQIINMWQ fibers. (F) Experimental infecivity relative to EF-C comparing CKIKQIINMWQ to KIKQIINMWQ.

Role of the N-terminal cysteine is further investigated by comparing CKAKAQANMWQ and KAKAQANMWQ peptide sequences in the figure below. Self-assembled structures for CKAKAQANMWQ are shown in Fig. 10 in the main text. Overall effect of removing N-terminal cysteine on hydrophobicity of the fiber is similar to the results discussed for KIKQIINMWQ.

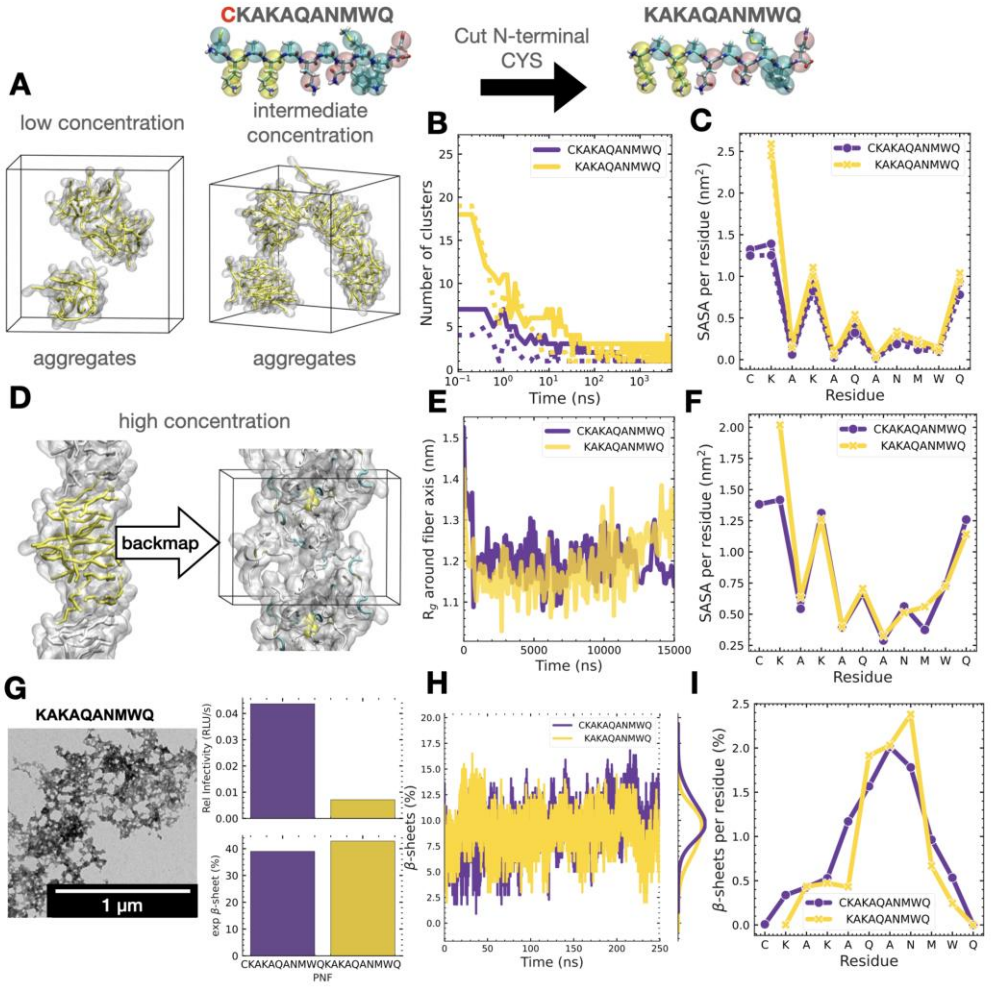


Figure S6: (A) Snapshots from low and intermediate concentration coarse-grained self-assembly simulations after 5 µs for KAKAQANMWQ. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (B) Number of clusters during self-assembly as a function of simulation time. (C) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration. (D) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping. (E) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (F) Average solvent accessible surface area (SASA) per residue calculated from the self-assembled atomistic fiber simulation. (G) TEM images for KAKAQANMWQ aggregates and experimental infecivity relative to EF-C and β-sheet content comparing CKAKAQANMWQ and KAKAQANMWQ. (H) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (I) Average percentage of β-sheets per residue.

In the simulations with pre-formed fibers, the results for alanine contaning peptide sequences differ from isoleucine containing peptides, shown in the figure below. Here KAKAQANMWQ peptide shows a high amount of β-sheets structures compared to CKAKAQANMWQ. However, careful investigation of the snapshots from these simulations indicate that KAKAQANMWQ fiber starts to dissociate into small aggregates. Therefore, we expect that these fiber structures for alanine containing peptides is not very stable.

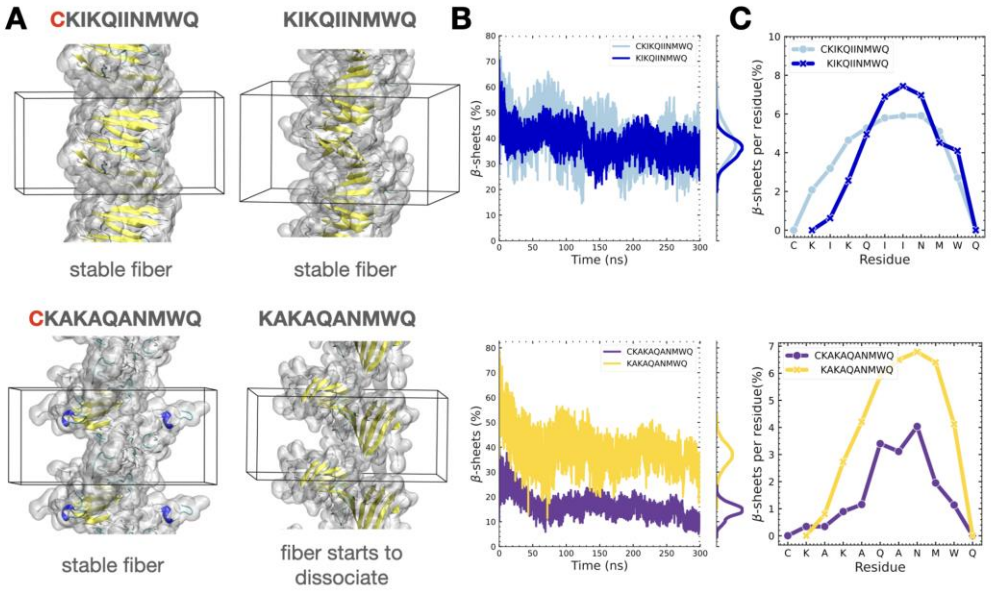


Figure S7: Effect of N-terminal cysteine from pre-formed fiber simulations. (A) Snapshots of side views of CKIKQIINMWQ, KIKQIINMWQ, CKAKAQANMWQ and KAKAQANMWQ pre-formed fibers after 300 ns atomistic simulation. Peptide backbones are drawn and colored according to the secondary structure, where β-sheets are shown in yellow and 3_{10}-helix structures are in dark blue. (B) Total percentage of amino acid residues making β-sheets structures over the simulation and (C) average β-sheet amount per residue calculated from the last 100 ns of simulation. Both in (B) and (C), percentage is calculated with respect to the total number of amino acid residues.

Here, we investigate the effect of switching cysteine from the N-terminus to C-terminus of the peptide. Only when the amphiphilic order is optimized, we observe fiber formation at intermediate and high concentrations for KIKIQIC. For snapshots of CKIKIQI, see Fig. S14.

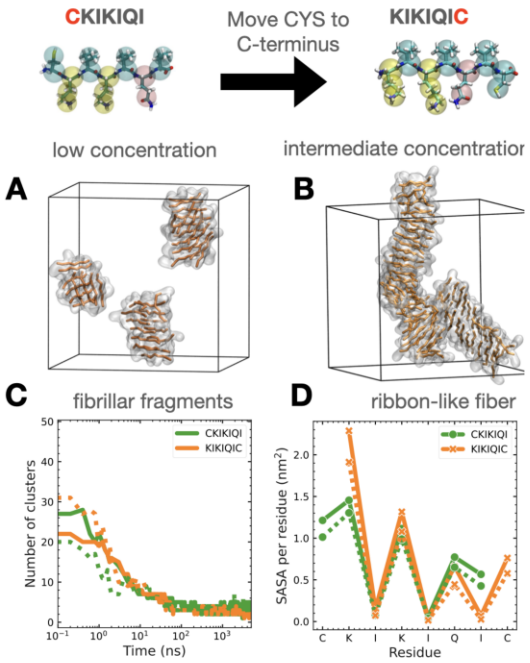


Figure S8: Snapshots from (A) low and (B) intermediate concentration coarse-grained self-assembly simulations after 5 µs for KIKIQIC. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (C) Number of clusters during self-assembly as a function of simulation time. (D) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration.

When we switch the position of the cysteine from N-terminus to C-terminus, the fibrillar morphology of the CKIKQIINMWQ disappears and instead we observe a segment of the fiber that is non-interacting with its periodic image. Only at high concentration, we can form a fiber that remains stable after being backmapped to atomistic representation. The surface hydrophobicity differs at the two termini of the peptide as expected, but the overall β-sheet percentage is similar.

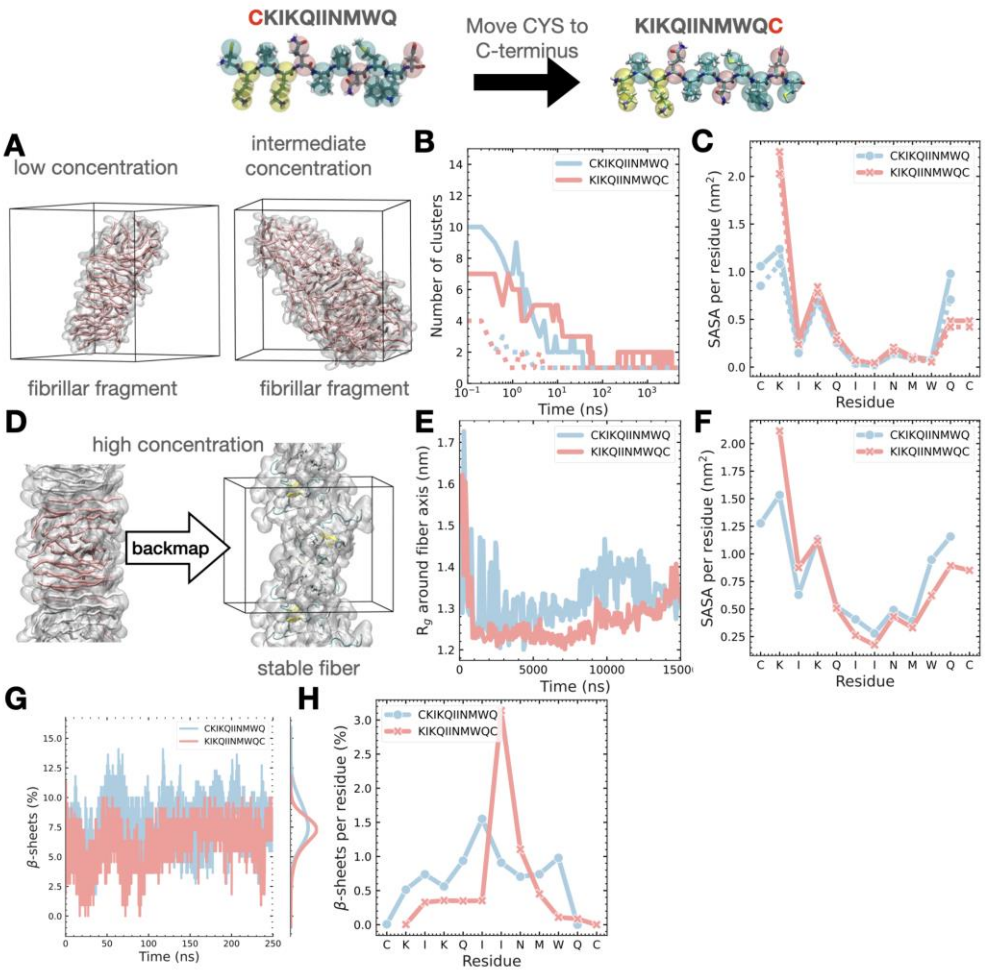


Figure S9: (A) Snapshots from low and intermediate concentration coarse-grained self-assembly simulations after 5 µs for KIKQIINMWQC. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (B) Number of clusters during self-assembly as a function of simulation time. (C) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration. (D) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping. (E) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (F) Average solvent accessible surface area (SASA) per residue calculated from the self-assembled atomistic fiber simulation. (G) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (H) Average percentage of β-sheets per residue.

For the short peptide sequences, we don't observe fibrillar structures for any of the tested concentrations. See Fig. S12 and S13 for CKIKQII.

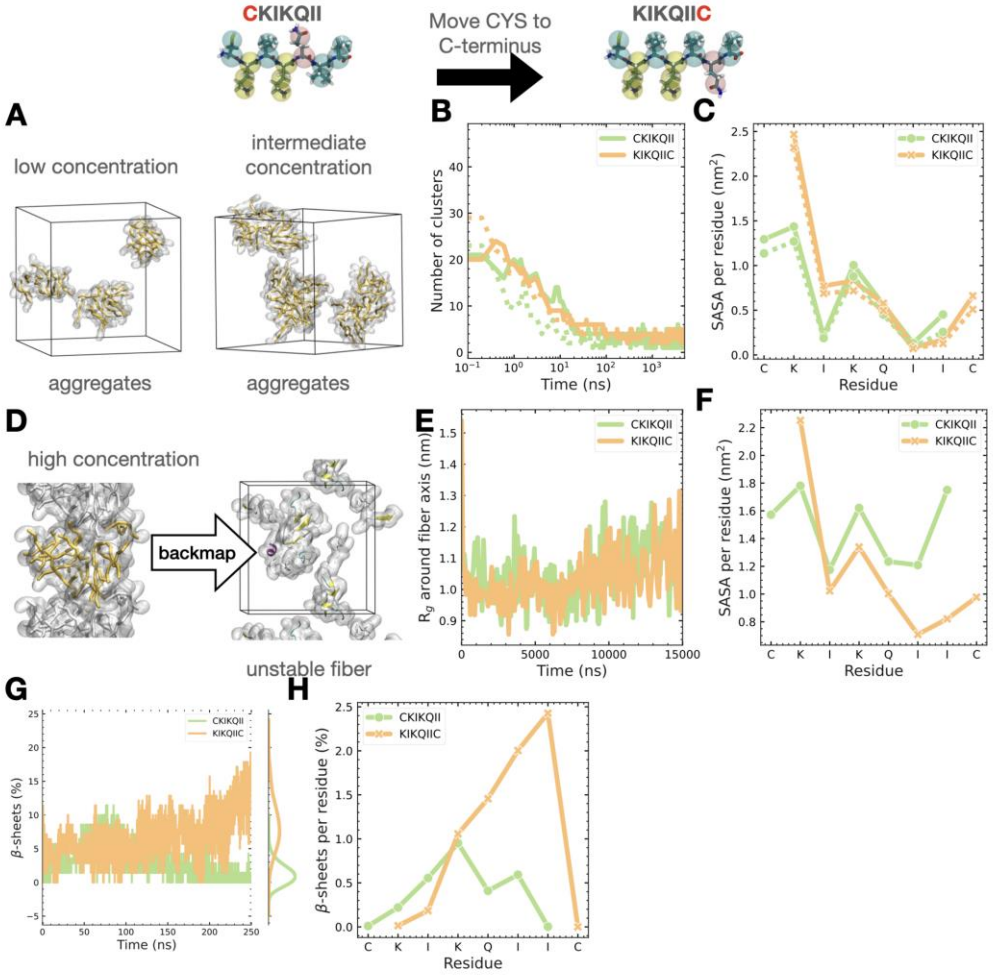


Figure S10: (A) Snapshots from low and intermediate concentration coarse-grained self-assembly simulations after 5 µs for KIKQIIC. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (B) Number of clusters during self-assembly as a function of simulation time. (C) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration. (D) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping. (E) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (F) Average solvent accessible surface area (SASA) per residue calculated from the self-assembled atomistic fiber simulation. (G) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (H) Average percentage of β-sheets per residue.

The effect of switching cysteine position is shown below for the pre-formed fiber simulations. Here, we see that the total β-sheet content is similar, but location of β-sheets in the

peptide sequence shifts from the N-termius to C-terminus. Thus, once again showing that there is a correlation between surface hydrophobicity and β-sheet formation.

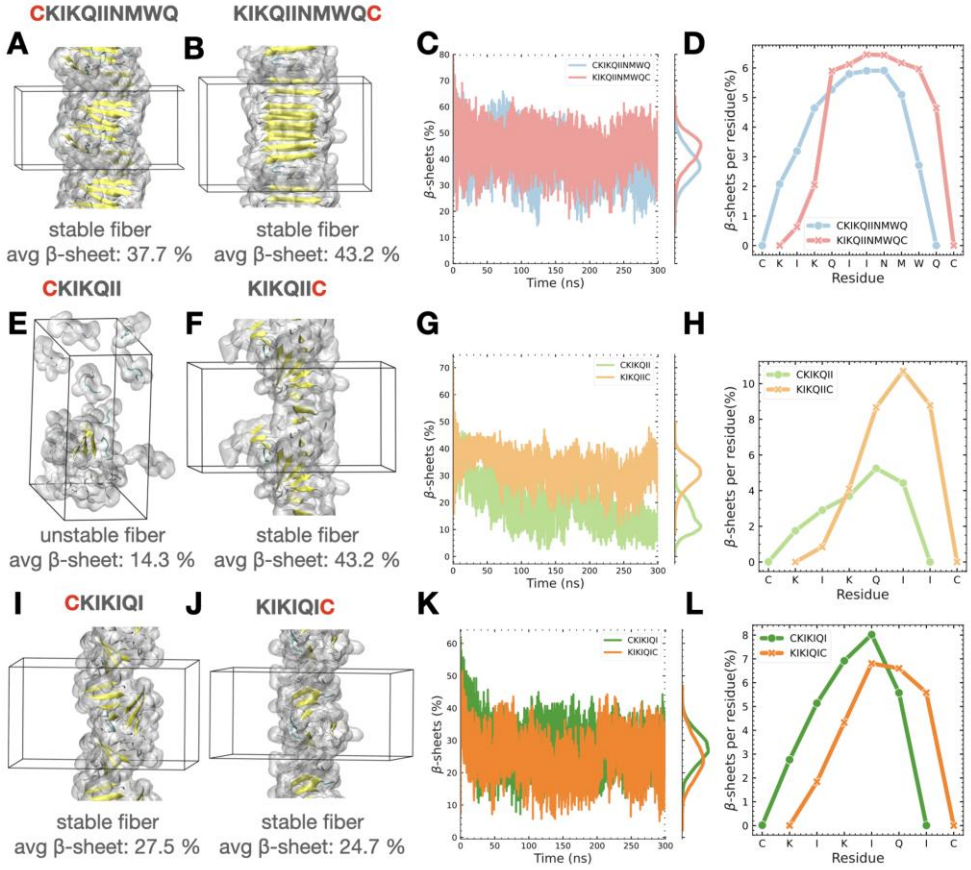


Figure S11: Effect of switching N-terminal cysteine to C-terminus from pre-formed fiber simulations. Snapshots of side views of (A) CKIKQIINMWQ, (B) KIKQIINMWQC, (E) CKIKQII, (F) KIKQIIC, (I) CKIKIQI and (J) KIKIQIC pre-formed fibers after 300 ns atomistic simulation. Peptide backbones are drawn and colored according to the secondary structure, where β-sheets are shown in yellow. (C, G, K) Total percentage of amino acid residues making β-sheets structures over the simulation and (D, H, L) average β-sheet amount per residue calculated from the last 100 ns of simulation. The percentage of β-sheets is calculated with respect to the total number of amino acid residues.

Supporting Figures for "Role of C-terminal NMWQ"

Lack of NMWQ amino acids in the original peptide (CKIKQIINMWQ) results in formation of disordered aggregates at low and intermediate concentrations. The self-assembly of the

peptides into the aggregated structures is also slower. After the peptides assemble, the resulting structures have more solvent accessible surface area since they are aggregates in solution instead of a continuous fiber. These aggregates are also visible in the TEM images from experiments.

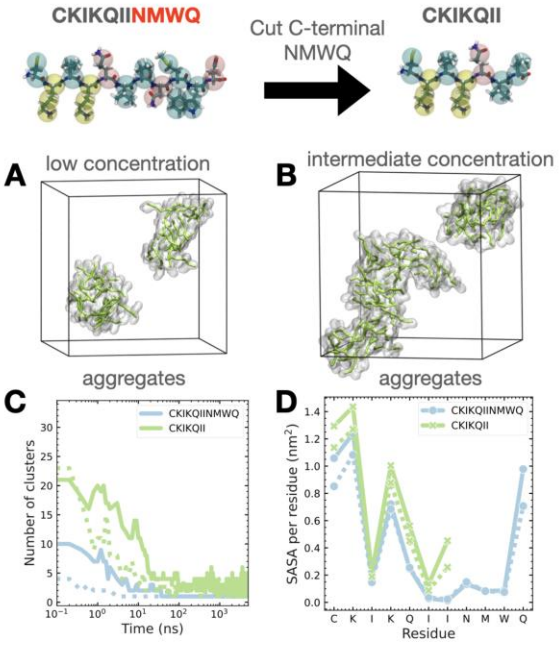


Figure S12: Snapshots from (A) low and (B) intermediate concentration coarse-grained self-assembly simulations after 5 µs for CKIKQII. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (C) Number of clusters during self-assembly as a function of simulation time. (D) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration.

At high concentration, the CKIKQII peptides are able to form a fiber, but this structure is not very stable both in coarse-grained system (Movie X) and atomistic system.

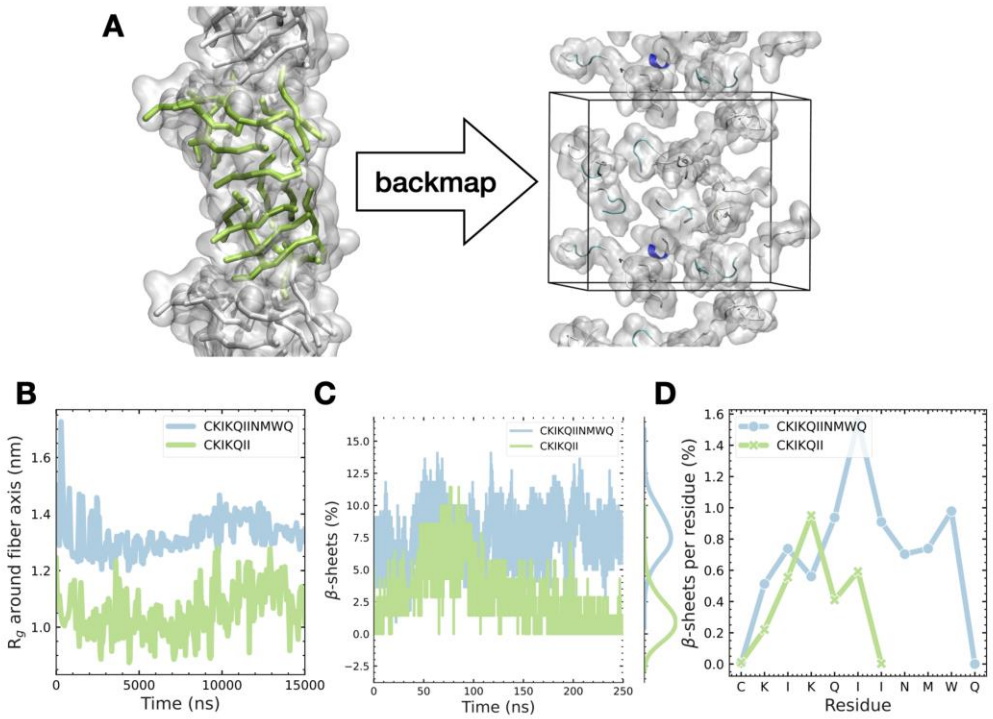


Figure S13: (A) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping for CKIKQII. (B) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (C) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (D) Average percentage of β-sheets per residue.

Supporting Figures for "Effect of amphiphilic pattern"

Changing the internal order of the amino acids from CKIKQII to CKIKIQI results small segments that are Amyloid-like in the way they are aligned, instead of disordered aggregates shown in Fig. S12. A fiber forms at high concentration, which loses its periodicity after 250 ns of atomistic simulation. However, this small segment of fiber still retains a high percentage of β-sheets compared to other peptide sequences. These β-sheets correspond to the core region of the peptide sequence where the intrinsic switch takes place.

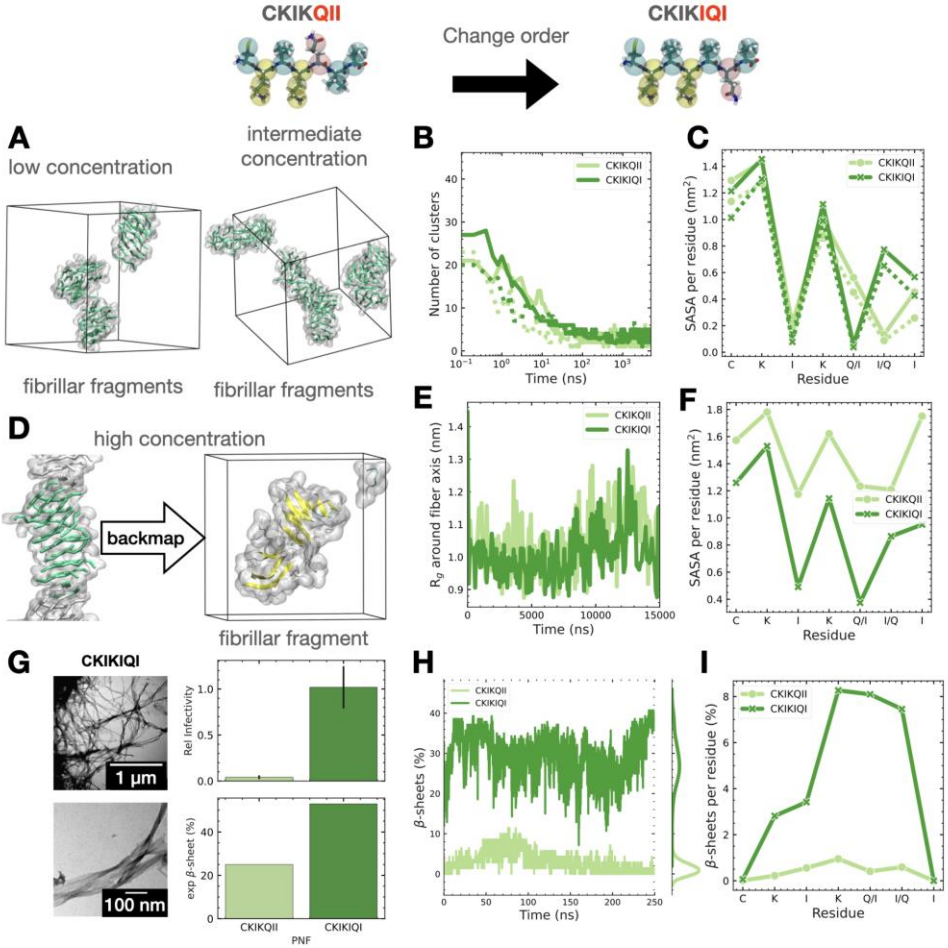


Figure S14: (A) Snapshots from low and intermediate concentration coarse-grained self-assembly simulations after 5 µs for CKIKIQI. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (B) Number of clusters during self-assembly as a function of simulation time. (C) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration. (D) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping. (E) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (F) Average solvent accessible surface area (SASA) per residue calculated from the self-assembled atomistic fiber simulation. (G) TEM images for CKIKIQI aggregates and experimental infecivity relative to EF-C and β-sheet content comparing CKIKQII and CKIKIQI. (H) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (I) Average percentage of β-sheets per residue.

For the original sequence, optimizing the amphiphilic pattern has the effect shown below. Overall morphology and surface hydrophobicity is similar. The self-assembly kinetics for the

low and intermediate concentrations is between those of CKIKQIINMWQ. In experiments, we did not observe a significant difference in the infection rates.

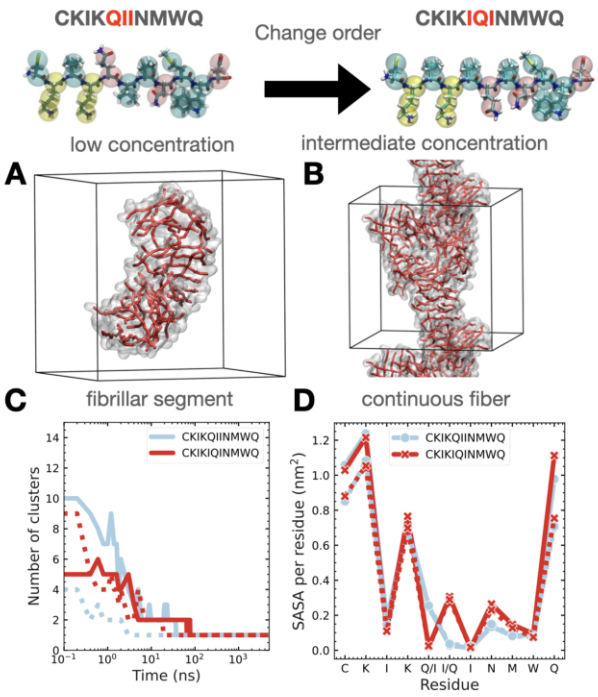


Figure S15: Snapshots from (A) low and (B) intermediate concentration coarse-grained self-assembly simulations after 5 µs for CKIKIQINMWQ. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (C) Number of clusters during self-assembly as a function of simulation time. (D) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration.

At high concentration, we observe a slight increase in the total number of β-sheets. This change corresponds to the core region of the peptide sequence, where the number of β-sheets increases continuously up to residue 5, which is where the intrinsic switch happens, plateaus and then starts to decrease after residue 7 toward the C-terminus of the peptide chain.

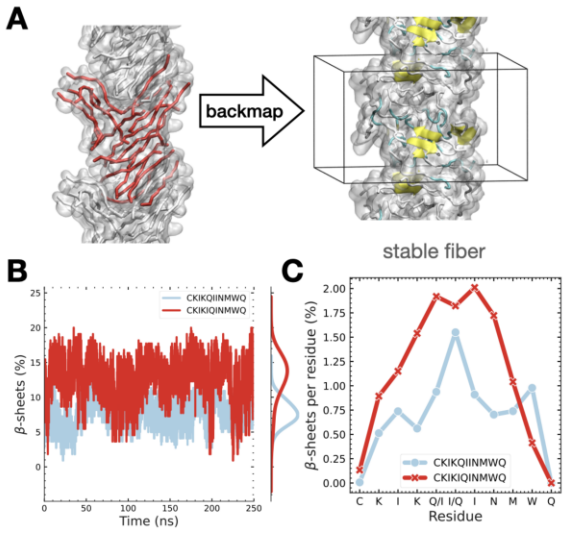


Figure S16: (A) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping for CKIKIQINMWQ. (B) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (C) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (D) Average percentage of β-sheets per residue.

Supporting Figures for "Tuning surface hydrophobicity"

In this section, we explored the effect of hydrophobocity by replacing specific amino acids. Below, we show the effect of replacing N-terminal cysteine with methionine. Cysteine and methionione are both sulfur containing amino acids, but because methionine has a larger alkyl side chain, it is more hydrophobic. When it comes to our coarse-grained model, both cysteine and methionine residues are represented with the same bead types, meaning their interaction with the other parts of the peptide and solvent is the same. This is due to the simplification of the coarse-graining procedure. However, since methionine has a bulkier side chain, the bond length between the backbone and side chain bead for methionine is larger compared to the cysteine. As expected, we see no differences between MKIKIQI and CKIKIQI self-assembly in solution for the low and intermediate concentrations, shown below. For high concentration and atomistic structure, refer to the main text.

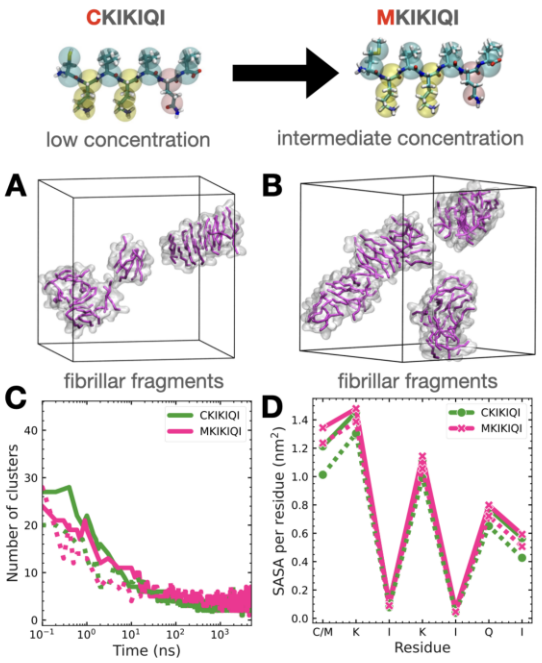


Figure S17: Snapshots from (A) low and (B) intermediate concentration coarse-grained self-assembly simulations after 5 µs for MKIKIQI. Only peptide backbones are colored for easier view of peptide assembly. The side chains are shown in white. Water and ions are omitted for clarity. (C) Number of clusters during self-assembly as a function of simulation time. (D) Average solvent accessible surface area (SASA) per residue calculated from the last 2 µs of simulation. The solid lines refer to simulations at low peptide concentration, while dotted lines refer to simulations at intermediate peptide concentration.

We also replaced the isoleucine in CKIKIQINMWQ peptide with alanine, which is less hydrophobic and has a smaller side chain. The self-assembled single fiber at high peptide concentration is shown below. Here, we can see that after backmapping to atomistic model, there is a loss of stability in CKAKAQANMWQ, as well as β-sheet formation. For the effect of replacing alanine with more hydrophobic phenylalanine, see the main text.

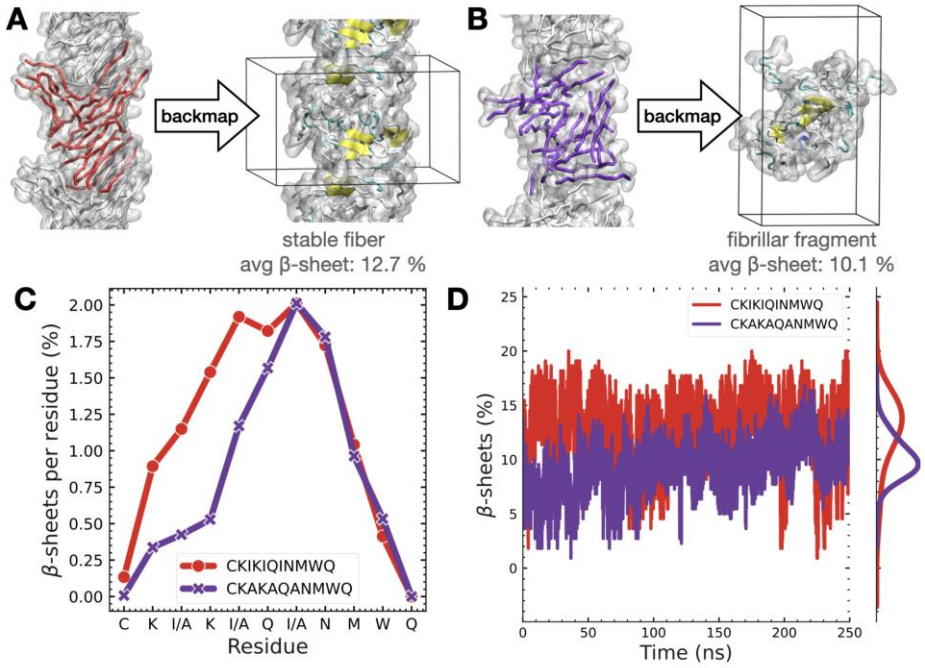


Figure S18: Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping for (A) CKIKIQINMWQ and (B) CKAKAQANMWQ. (C) Average percentage of β-sheets per residue. (D) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations.

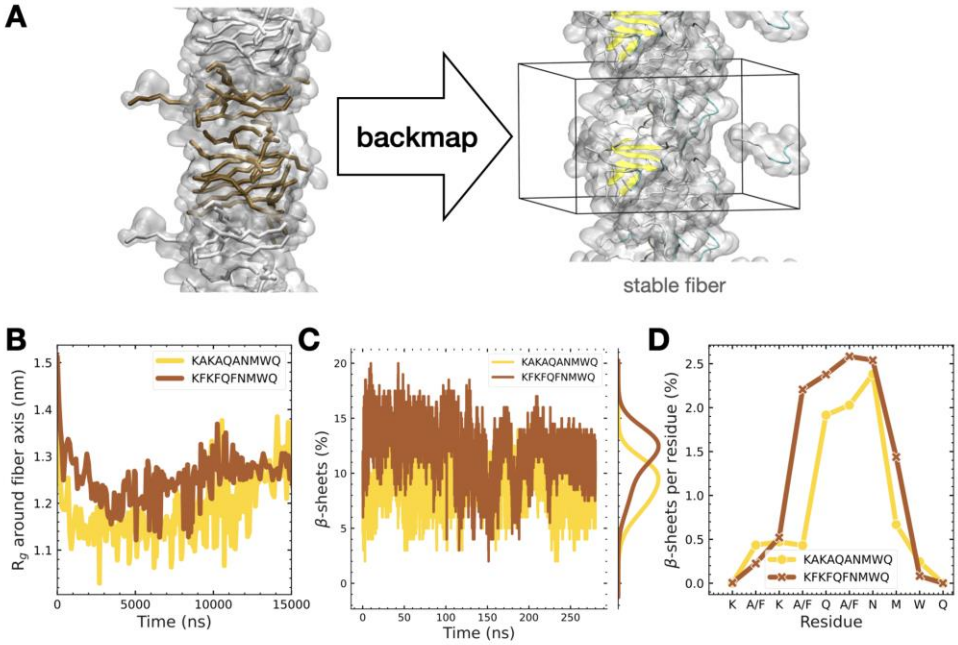


Figure S19: (A) Snapshots from after 15 µs of high concentration coarse-grained self-assembly simulation and after 250 ns atomistic simulation after backmapping for KFKFQFNMWQ. (B) The average radius of gyration around the fiber axis from high concentration self-assembly simulation. (C) Calculated percentage of β-sheets from self-assembled atomistic fiber simulations. (D) Average percentage of β-sheets per residue.

Finally, in the figure below, we show that effect of hydrophobicity in the secondary structure and stability of single pre-formed fibers. Overall, for the pre-formed fibers, we discovered that the peptide sequence with the more hydrophobic amino acid lead to a structure with larger amount of β-sheets.

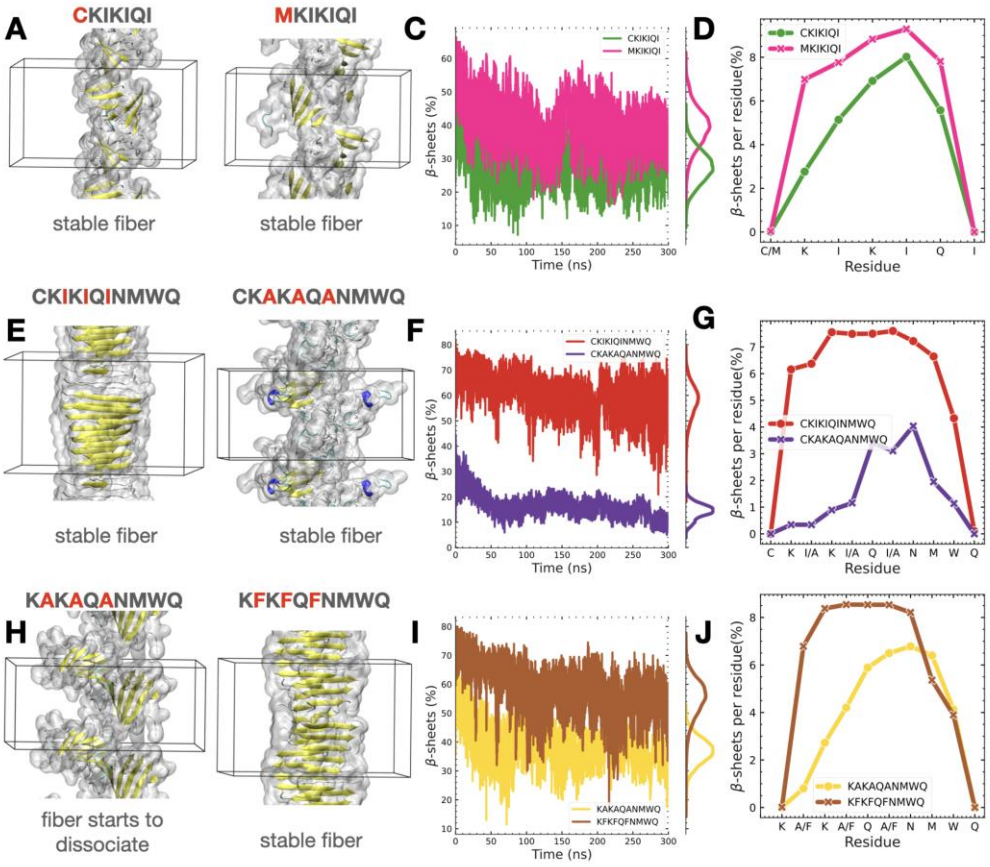


Figure S20: Effect of hydrophobicity from pre-formed fiber simulations. Snapshots of side views of (A) CKIKIQI, MKIKIQI, (E) CKIKIQINMWQ, CKAKAQANMWQ (H) KAKAQANMWQ and KFKFQFNMWQ pre-formed fibers after 300 ns atomistic simulation. Peptide backbones are drawn and colored according to the secondary structure, where β-sheets are shown in yellow and 3_{10}-helix structures are in dark blue. (C, F, I) Total percentage of amino acid residues making β-sheets structures over the simulation and (D, G, J) average β-sheet amount per residue calculated from the last 100 ns of simulation. All β-sheet percentages are calculate with respect to the total number of amino acid residues.

Supporting Figures for "Dynamic properties of PNFs"

In order to see whether fibers in solution interact with one another and form larger structures in solution, we placed four fibers in close proximity from each other. The cross-sectional view of the initial structure for CKIKQIINMWQ is shown in Figure S1. For each sequence, we used the coordinates of a stable, self-assembled single fiber from high concentration coarse-grained simulations as initial configurations.

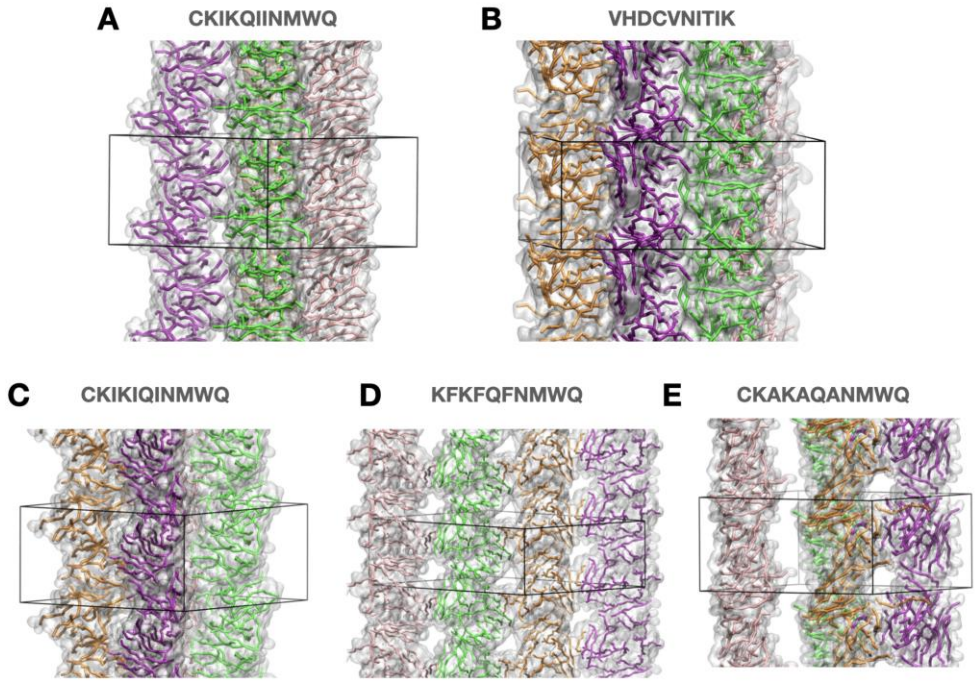


Figure S21: Snapshots of the side view of (A) CKIKQIINMWQ, (B) VHDCVNITIK, (C) CKIKIQINMWQ, (D) KFKFQFNMWQ and (E) CKAKAQANMWQ after 10 µs of coarse-grained multi-fiber simulations. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity.

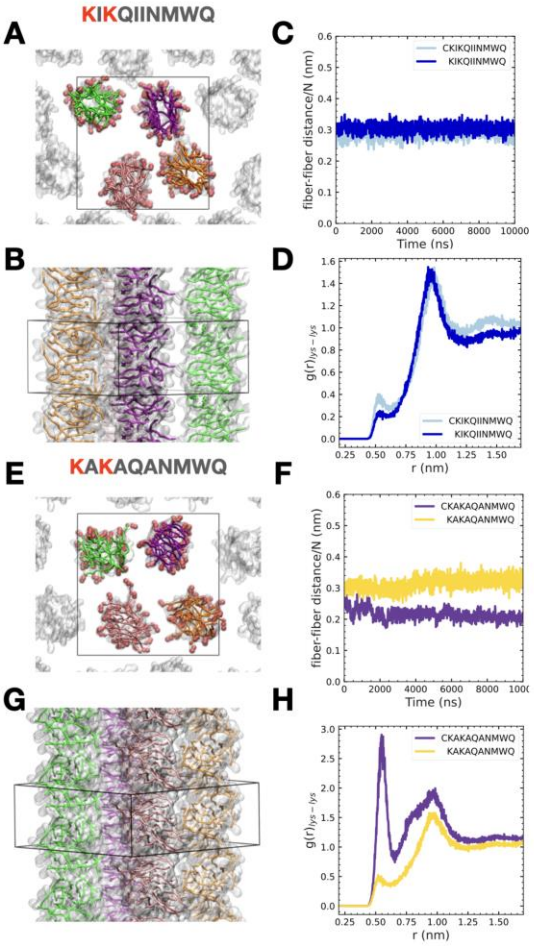


Figure S22: Effect of removing N-terminal cysteine on fiber-fiber interactions. The (A,E) cross-sectional and (B,G) side views of KIKQIINMWQ and KAKAQANMWQ fibers, respectively, in solution. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity. (C, F) The fiber-fiber distance between the initial structures over the simulation time. (D, H) The intermolecular pair distribution functions, g(r), between lysine groups.

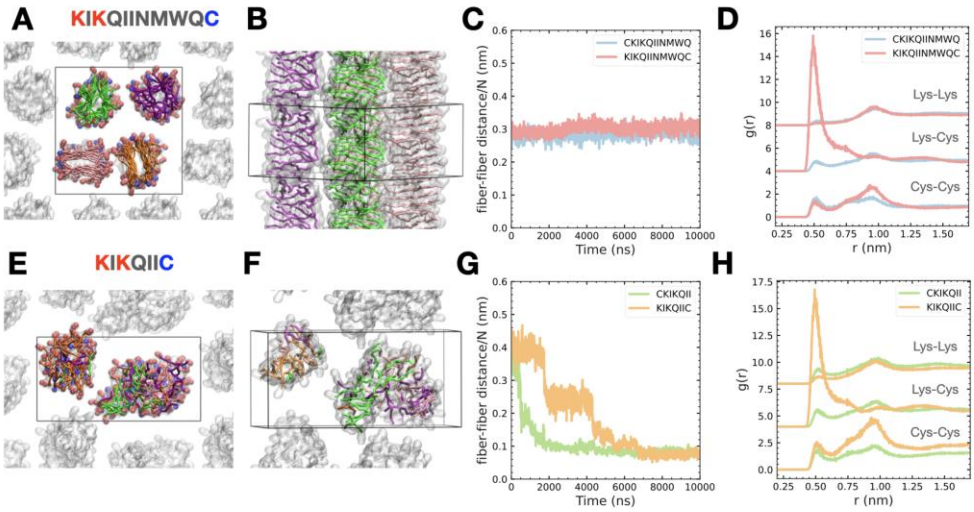


Figure S23: Effect of switching N-terminal cysteine to C-terminus on fiber-fiber interactions. The (A,E) cross-sectional and (B,F) side views of KIKQIINMWQC and KIKQIIC fibers, respectively, in solution. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity. (C, G) The fiber-fiber distance between the initial structures over the simulation time. (D, H) The intermolecular pair distribution functions, g(r), between lysine and cysteine groups.

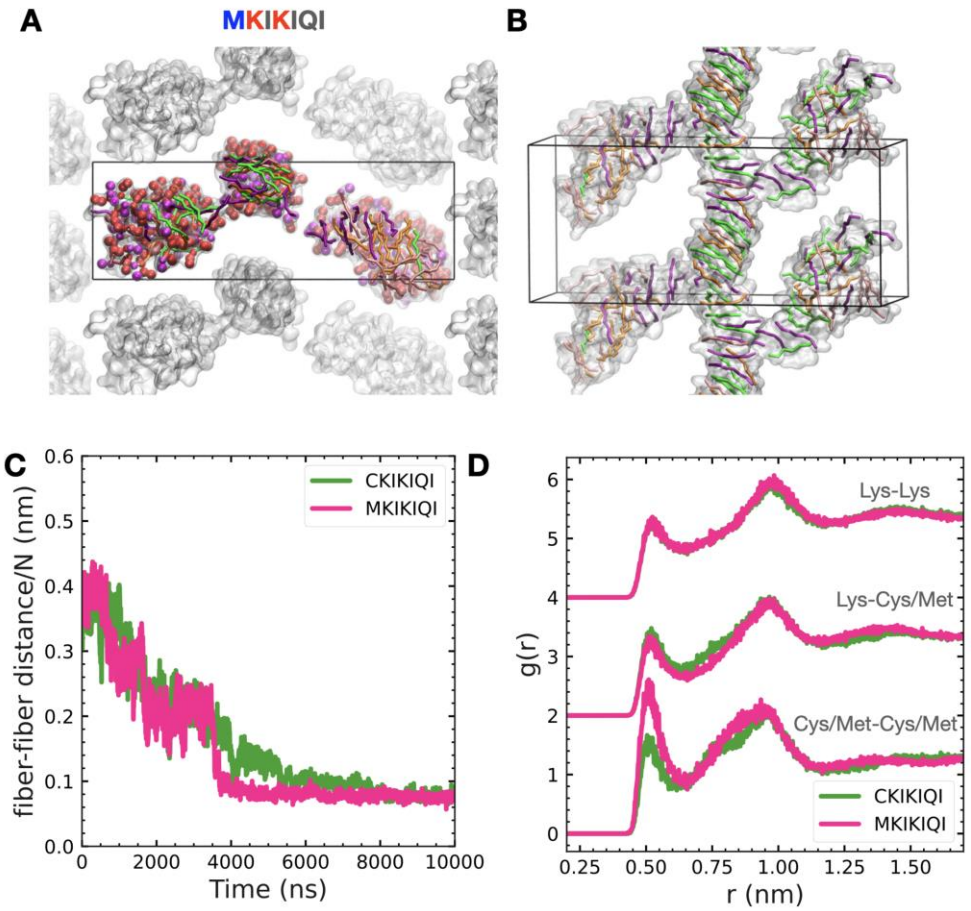


Figure S24: Effect of cysteine to methionine mutation on fiber-fiber interactions. The (A) cross-sectional and (B) side views of MKIKIQI fibers, respectively, in solution. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity. (C) The fiber-fiber distance between the initial structures over the simulation time. (D) The intermolecular pair distribution functions, g(r), between lysine and cysteine/methionine groups.

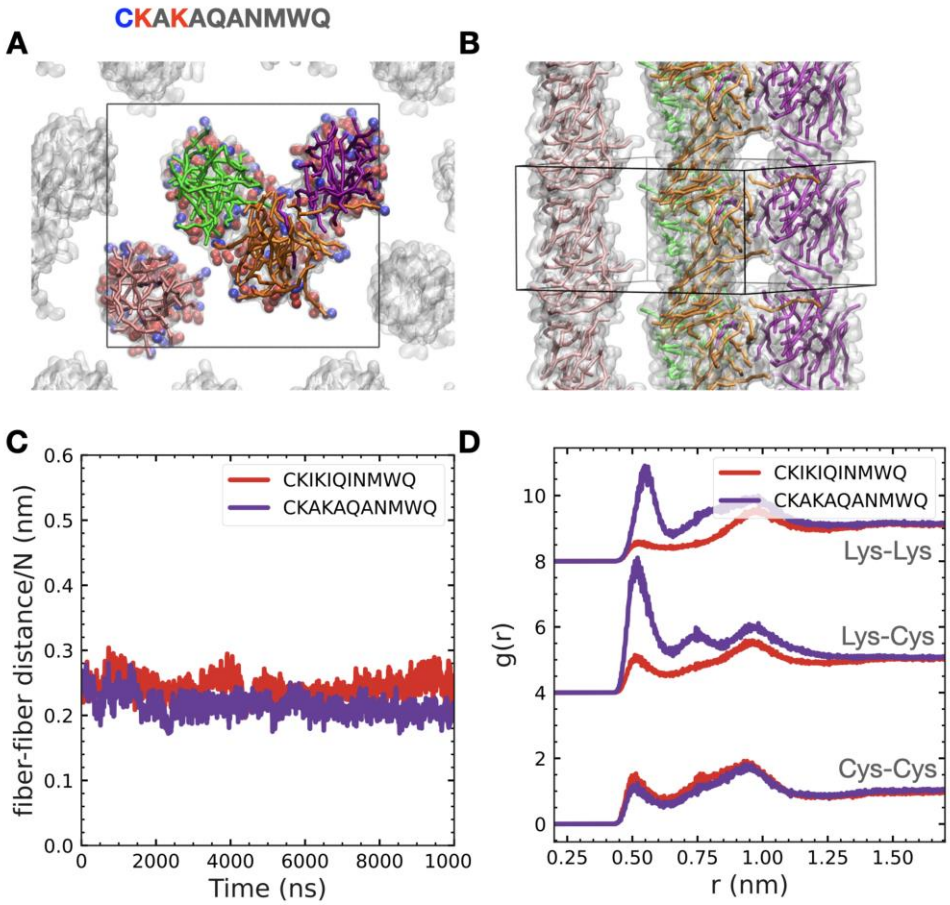


Figure S25: Effect of isoleucine to alanine mutation on fiber-fiber interactions. The (A) cross-sectional and (B) side views of CKAKAQANMWQ fibers, respectively, in solution. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity. (C) The fiber-fiber distance between the initial structures over the simulation time. (D) The intermolecular pair distribution functions, g(r), between lysine and cysteine groups.

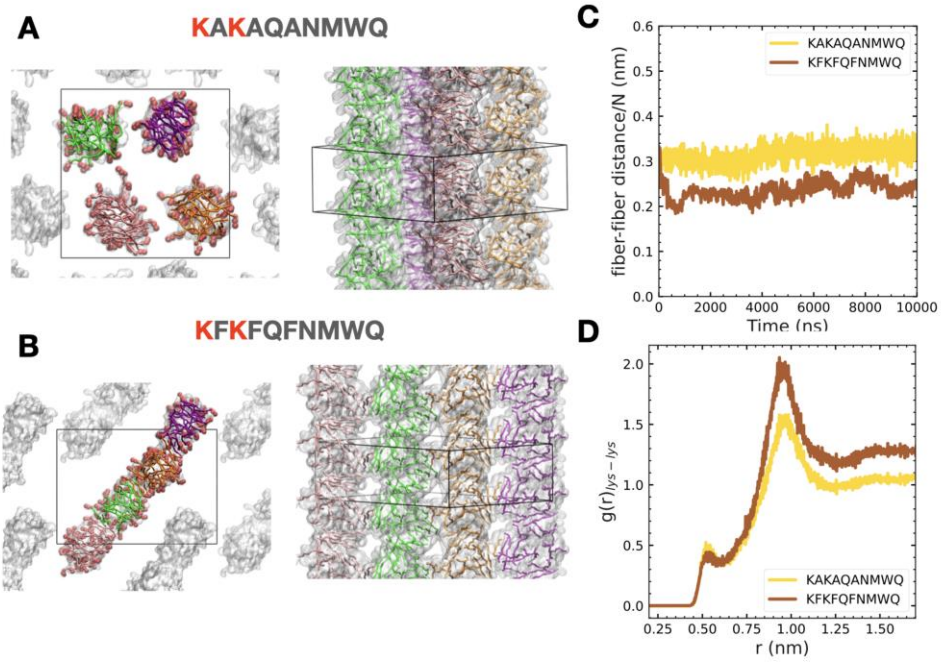


Figure S26: Effect of alanine to phenylalanine mutation on fiber-fiber interactions. The cross-sectional and side views of (A) KAKAQANMWQ and (B) KFKFQFNMWQ fibers, respectively, in solution. Peptides that belong to the same fiber at time = 0 have same color backbones. The side chains of peptides are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity. (C) The fiber-fiber distance between the initial structures over the simulation time. (D) The intermolecular pair distribution functions, g(r), between lysine groups.

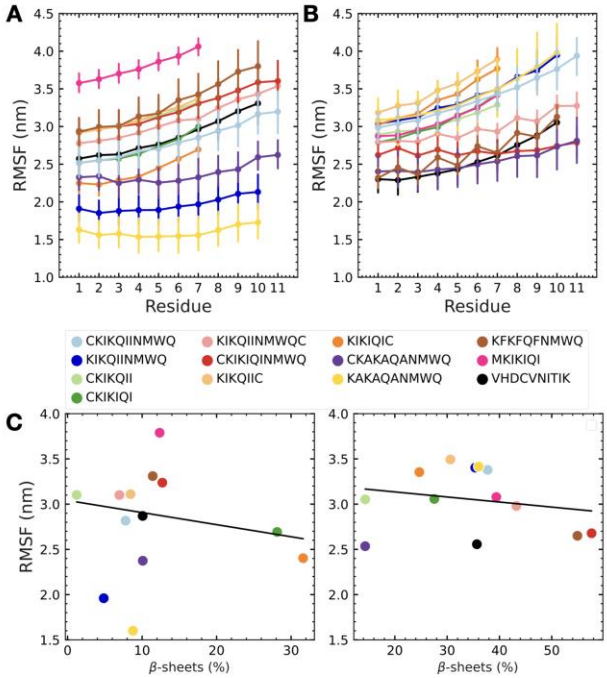


Figure S27: Average room mean square fluctuation (RMSF) for each residue calculated from atomistic (A) self-assembled and (B) pre-formed single fiber simulations. For both, the 100 ns of simulation is used for averaging. (C) Average RMSF for each peptide chain is plotted against average percentage of β-sheets calculated from self-assembled (left) and pre-formed (right) single fiber simulations, respectively.

References

(S1) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **2009**, *30*, 2157–2164.

(S2) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation* **2013**, *9*, 687–697, PMID: 26589065.

(S3) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable Water

Model for the Coarse-Grained MARTINI Force Field. *PLOS Computational Biology* **2010**, *6*, 1–17.

(S4) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **1995**, *91*, 43–56.

(S5) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(S6) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(S7) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(S8) Nosé, S.; Klein, M. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055–1076.

(S9) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

(S10) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Molecular Simulation* **1991**, *6*, 121–142.

(S11) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(S12) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. J. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting

Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ^1 and χ^2 Dihedral Angles. *Journal of Chemical Theory and Computation* **2012**, *8*, 3257–3273, PMID: 23341755.

(S13) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(S14) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(S15) Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function, and Bioinformatics* **1995**, *23*, 566–579.

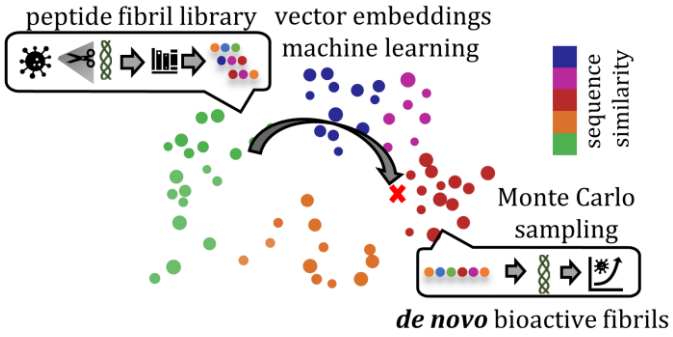
6.5 Inverse Design of Viral Infectivity-Enhancing Peptide Fibrils from Continuous Protein-Vector Embeddings

Authors: Kübra Kaygisiz, Arghya Dutta, Lena Rauch-Wirth, Christopher V. Synatschke, Jan Münch, Tristan Bereau* and Tanja Weil*
* corresponding author
Published in: Biomaterials Science, **2023,** 11 (15), 5251–5261.
DOI: 10.1039/D3BM00412K
Date of Publication: 15th of June 2023

Author Contributions:
Kübra Kaygisiz conceptualized, designed, carried out the experiments, analyzed data, performed the formal analysis related to peptide design, synthesis and characterization, wrote python scripts for amino acids analysis, performed and analyzed aggregation prediction, and wrote the original draft. Arghya Dutta wrote python scripts, designed, analyzed data, and performed the formal analysis related to vector embeddings and Monte Carlo sampling and revised the manuscript. Lena Rauch-Wirth designed and carried out the experiments related to virus transduction, analyzed data, performed the formal analysis related to virus transduction, and revised the manuscript. Christopher V. Synatschke supervised and revised the manuscript. Tristan Bereau conceived the idea and conceptualized the study, analyzed data and provided supervision related to vector embeddings and Monte Carlo sampling, provided the funding resources, and revised the manuscript. Tanja Weil conceptualized, supervised, provided the funding resources, and revised the manuscript.

Copyrights: The following part is published in Biomater. Sci., **2023,** 11 (15), 5251–5261, https://doi.org/10.1039/D3BM00412K and licensed under CC BY 3.0 (https://creativecommons.org/licenses/by/3.0/). Further distribution of the material is possible if appropriate credit, a link provided to the license, and indications of changes are made.

Graphical abstract:



We apply machine learning to find infectivity-enhancing fibrillar peptides. Using a regression model on the continuous vector representations of a peptide library and employing Monte Carlo Sampling, we predict *de novo* sequences with a 25% hit rate.

Biomaterials Science

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Biomater. Sci.*, 2023, **11**, 5251

Inverse design of viral infectivity-enhancing peptide fibrils from continuous protein-vector embeddings†

Kübra Kaygisiz, [a] Arghya Dutta, [‡b] Lena Rauch-Wirth, [c] Christopher V. Synatschke, [a] Jan Münch, [c] Tristan Bereau [*§b] and Tanja Weil [*a]

Amyloid-like nanofibers from self-assembling peptides can promote viral gene transfer for therapeutic applications. Traditionally, new sequences are discovered either from screening large libraries or by creating derivatives of known active peptides. However, the discovery of *de novo* peptides, which are sequence-wise not related to any known active peptides, is limited by the difficulty to rationally predict structure–activity relationships because their activities typically have multi-scale and multi-parameter dependencies. Here, we used a small library of 163 peptides as a training set to predict *de novo* sequences for viral infectivity enhancement using a machine learning (ML) approach based on natural language processing. Specifically, we trained an ML model using continuous vector representations of the peptides, which were previously shown to retain relevant information embedded in the sequences. We used the trained ML model to sample the sequence space of peptides with 6 amino acids to identify promising candidates. These 6-mers were then further screened for charge and aggregation propensity. The resulting 16 new 6-mers were tested and found to be active with a 25% hit rate. Strikingly, these *de novo* sequences are the shortest active peptides for infectivity enhancement reported so far and show no sequence relation to the training set. Moreover, by screening the sequence space, we discovered the first hydrophobic peptide fibrils with a moderately negative surface charge that can enhance infectivity. Hence, this ML strategy is a time- and cost-efficient way for expanding the sequence space of short functional self-assembling peptides exemplified for therapeutic viral gene delivery.

Received 8th March 2023,
Accepted 14th June 2023

DOI: 10.1039/d3bm00412k

rsc.li/biomaterials-science

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

1 Introduction

New peptides to increase viral transduction: a multi-parameter challenge

Short self-assembling peptides have attracted much interest as functional materials in recent years since they can be designed in a precise and cost-efficient way.[1] However, a small change in sequence can drastically change physicochemical properties on multiple length scales, which makes it challenging to rationally design *de novo* self-assembling peptide sequences with a desired bioactivity.[2]

For example, the emerging field of gene therapy requires efficient transduction of target cells by viral vectors that deliver the therapeutic gene.[3] In this regard, self-assembled peptide fibrils that increase the colocalization of viral vectors and cellular membranes and thereby enhance gene delivery are promising candidates for new or optimized gene-therapeutic applications.[4–6] However, the discovery of self-assembling peptides as enhancers of viral transduction *via* screening methods is challenging because of the complexity in predicting sequences that show the required physicochemical properties such as sequence amphiphilicity, charge, and assembly, for biological activity.[7,8]

Traditionally, new peptides with certain desired properties, *e.g.*, viral transduction/infection enhancement, have been found mainly by serendipity during screening processes,[4,9] or by nature-inspired rational design.[7,10,11] A common strategy is

[a]*Department Synthesis of Macromolecules, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. E-mail: weil@mpip-mainz.mpg.de*
[b]*Polymer Theory, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany. E-mail: bereau@thphys.uni-heidelberg.de*
[c]*Institute of Molecular Virology, Ulm University Medical Center, Meyerhofstraße 1, 89081 Ulm, Germany*

†Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3bm00412k

‡Present address: Institute of Biochemistry II, Faculty of Medicine, Goethe University, Theodor–Stern–Kai 7, 60590 Frankfurt, Germany.
§Present address: Institute for Theoretical Physics, Heidelberg University, Philosophenweg 19, 69120 Heidelberg, Germany.

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

finding recurring motifs in known active peptides to generate new sequences.[12,13] Although changing one amino acid at a time to screen derivatives of a known active compound is often a direct and efficient way to find active peptides with similar structure and study property–activity relationship, it is not a tractable way to discover new sequences. The sequence space of peptides is huge—there are 20^6 = 64 million peptides if we only consider all possible 6-mers composed of 20 canonical amino acids. Further, the size of the sequence space increases exponentially with the number of residues. Consequently, exploring the peptide space to discover new peptides by creating randomly generated peptides or derivatives of known structures and studying them in experiments quickly becomes unfeasible. In the quest to discover new peptides with certain target bioactivity, computational methods have been established for fast and inexpensive prescreening of peptide sequences.[14]

Navigating through sequence space *via* machine learning

Machine learning (ML) can be applied to bridge the limited experimental dataset with the vast compound space by analyzing the sequence–activity relationships to reverse engineer novel non-intuitive compounds.[15–17] ML models are often used to study structure–activity relationships since they can mitigate the curse of dimensionality: peptides with similar values of a target property can be far apart in sequence space but close in some higher-dimensional underlying feature space.[18] Various supervised and unsupervised ML algorithms have been used, for example, to predict liquid–liquid phase separating protein sequences[19] and to detect antimicrobial[20] or cell-penentrating[21] peptide sequences. Recently, an ML approach, in combination with Monte Carlo tree search and molecular dynamics simulations, was reported for predicting unexpected *de novo* β-sheet rich self-assembling peptides.[22] While all these reports successfully identified either bioactive or self-assembling peptides separately, they did not predict bioactive self-assembling peptides.

Inverse design of sequences *via* continuous vector embeddings

One way of training an ML model that can infer structure–activity relationships of bioactive self-assembling peptides is to use the minimal available information as input to the model—the peptide's sequence. Continuous vector representation is the method of choice to represent complex sequence order and composition of proteins and peptides for training ML algorithms. For instance, the word embedding model Word2Vec,[23,24] that was originally developed as a natural language processing tool, can be applied to extract structural concepts to encode latent material property information across various datasets.[25]

In this direction, Asgari and Mofrad recently proposed a method that can convert any protein sequence into a unique, dense, 100-dimensional numerical vector, termed ProtVec.[26] They used a method from natural language processing that employs an artificial neural network that, while attempting to determine the context in which a word is most likely to occur in a sentence, generates a continuous distributed representation of the word; further, the ProtVecs were found to accurately capture the physicochemical properties of proteins.

Here, we report a three-step approach to explore the sequence space of bioactive self-assembling peptides and discover *de novo* sequences with high bioactivity by only using sequence and activity information. First, we trained a LASSO (Least Absolute Shrinkage and Selection Operator) regression model using ProtVec representations of peptides from a relatively small library of 163 sequences with known activity values. LASSO aims at identifying a minimal subset of parameters relevant to the prediction, thereby enhancing explainability. Then, we utilized the trained model to systematically sample the sequence space of 6-mer peptides using a Monte Carlo approach.[27] Finally, we screened the sequences with highest predicted activities based on their charge and tendency to aggregate. The search yielded 16 new peptides, which were tested experimentally and found to be active with a 25% hit rate or a 50% hit rate if further predictive parameters like aggregation are included. Strikingly, the newly created peptides are very different in sequence from the ones comprising the training set and shorter than any previously reported sequence for infectivity enhancement. Taken together, our method offers a fast and computationally inexpensive way of predicting and screening potentially bioactive, self-assembling peptides for any desired target bioactivity; consequently, it can accelerate peptide screening in the early stages of research.

2 Results

To gain insight into sequence elements and predict *de novo* sequences, which enhance retroviral transduction, we exploited an already reported peptide library for viral transduction enhancement consisting of 163 sequences. This library was based on the self-assembling 12-mer peptide EF-C (QCKIKQIINMWQ) corresponding to residues 417–428 of the HIV envelope protein gp120 and was created by systematic peptide sequence alterations (Fig. 1A).[4,7,8] The rational design of the library results in similar sequences (Fig. 1B), which show a wide range of activity (Fig. 1C). We decided to use this peptide library as a training set to demonstrate the usefulness of ML approaches because this data set represents a common situation in the early stage of research, where one wants to explore the peptide sequence space starting from a relatively small library consisting of derivatives of a known active compound.

Exploration of sequence space

The ML approach we used is summarized in Fig. 1D. The self-assembling bioactive peptides that were used to train the ML model were previously reported by us.[8] As described in the introduction, we represented the peptide sequences as dense, 100-dimensional numerical vectors, termed as ProtVecs, using a continuous distributed representation proposed by Asgari and Mofrad.[26] One of the advantages in using continuous dis-

 This journal is © The Royal Society of Chemistry 2023

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

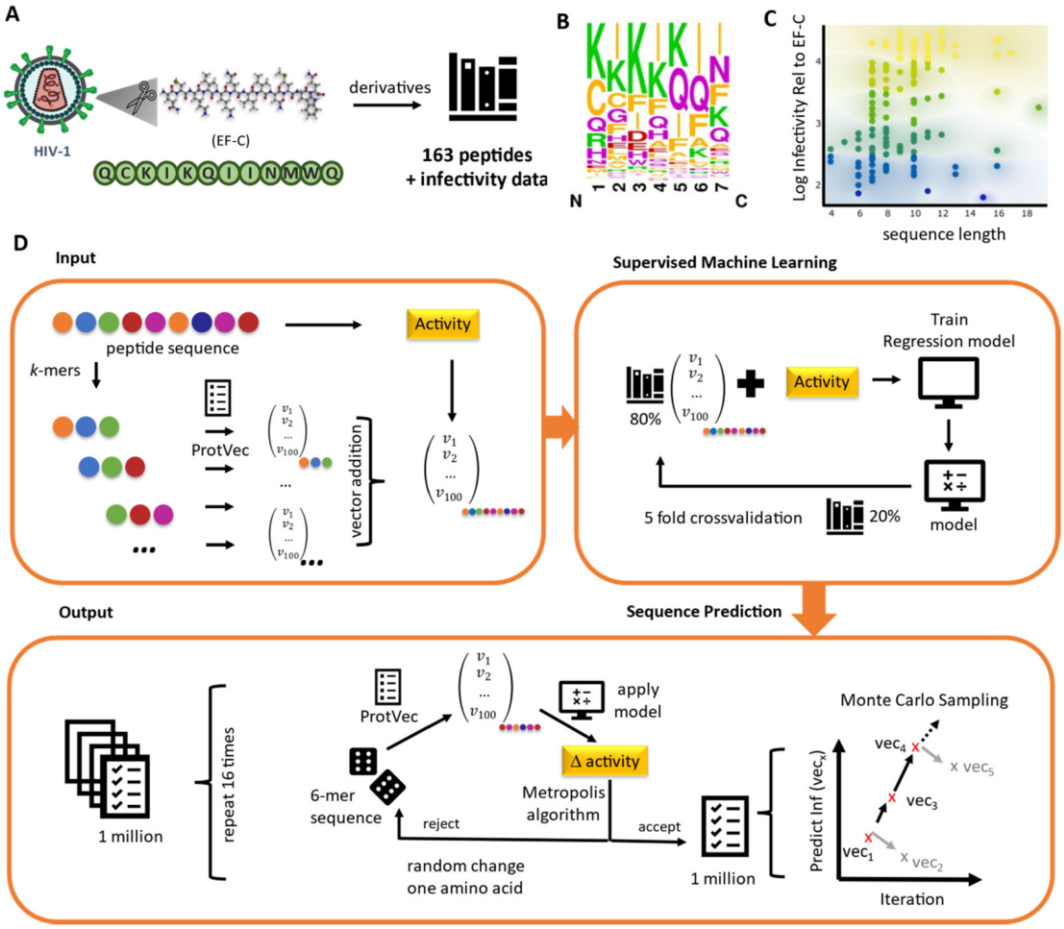


Fig. 1 Schematic overview of the workflow. (A) A peptide library[8] was created by systematic sequence derivations of the infectivity-enhancing peptide EF-C. This database contains 163 peptide sequences and their respective biological activity data (*i.e.* enhancement of HIV-1 infection, "infectivity". (B) The sequence logo plot summarizes the amino acid frequency in the library and shows that most sequences are composed of positive charged (K, lysine) and hydrophobic uncharged (F, phenylalanine or I, isoleucine) amino acids in alternating order. (C) The diagram shows the length distribution of infectivity enhancement in log scale relative to the reference peptide EF-C. Despite the sequence similarity, a wide range of activities can be found within the library. (D) Flowchart summarizing the machine-learning approach in this study. As input data, the EF-C based peptide library (A) was used. Each peptide sequence in the library was represented as a 100-dimensional vector using a continuous distributed representation, ProtVec.[26] The vector and respective activity were used in a supervised LASSO regression model to compute a linear relationship between the activity and relevant components of the 100-d ProtVec. The trained model was then used in a Monte Carlo search where 1 million 6-mers were generated, and the 6-mers that were predicted to perform better than EF-C were retained based on a Metropolis criterion (see Methods). By repeating the Monte Carlo step 16 times, each time starting with a random 6-mer, a broad portion of the 6-mer sequence space was covered.

tributed representation is that it captures underlying bioinformatic properties and information on distance relationships of neighboring amino acids and uniformly casts any arbitrary-sized peptide to a unique numerical vector with a fixed size (*e.g.*, 100-d ProtVecs), which can be readily used in numerical computations.[14,26] We constructed ProtVecs for training set peptide sequences by dividing the sequences into 3-grams (3 consecutive amino acids) and summing up the ProtVecs representing those 3-grams[26] by linear vector addition. In that way, 100-d vectors are created for each sequence, which can be assigned to an experimentally determined activity. The 100-d vectors in combination with the bioactivity data were then used to train a LASSO linear regression ML model (Fig. S1†).[28] The model (ProtVec LASSO, ESI eqn (1)†) was validated *via* 5-fold cross-validation and found to correlate well with the bioactivity (Pearson correlation coefficient R = 0.84, Fig. 2A). Though a RIDGE regression model offers an alternative with a similar Pearson correlation (R = 0.83, Fig. S1†), we used a LASSO model since, by minimizing the number of non-zero regression coefficients, it provides a simpler model with fewer parameters.

The strength of this ML model is that it can be applied on any other sequences made of canonical amino acids, which can be represented in a 100-d vector. With the trained ML model, we decided to screen for bioactive 6-mers. We chose peptides with only 6 amino acids because they would be shorter than any infectivity enhancing peptide from our library or the literature.

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

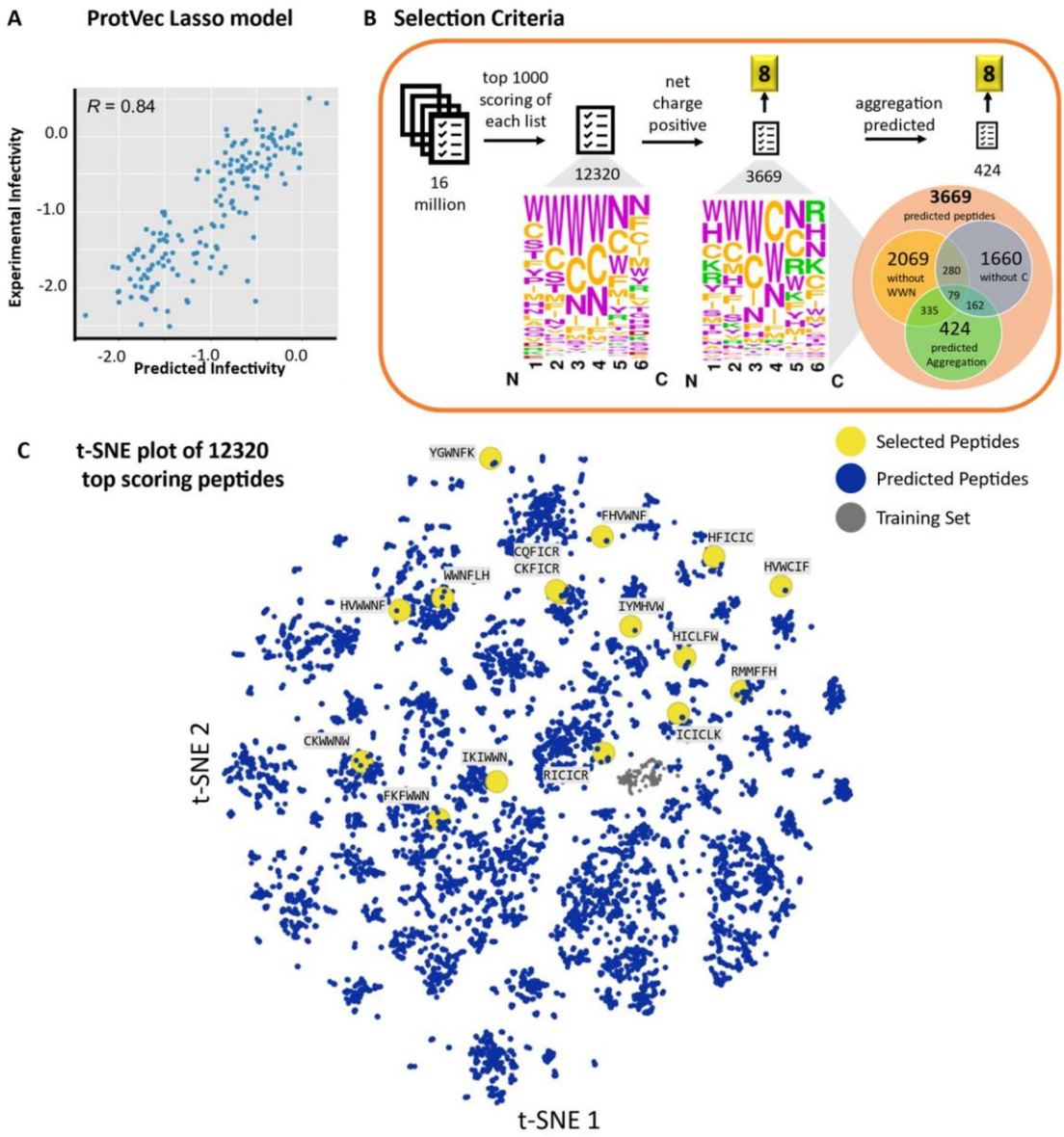


Fig. 2 (A) Performance of supervised linear regression model trained from EF-C based library. The experimentally measured data correlate highly with the predicted infectivity enhancement (Pearson R = 0.84). (B) Scheme showing the selection criteria for narrowing down putatively infectivity-enhancing peptides for experimental evaluation. From each independent prediction Monte Carlo Sampling run the best predicted 1000 sequences were selected. From this subset 3669 sequences showed a positive net charge and were further considered as promising candidates. 8 sequences predicted for aggregation and 8 sequences not predicted for aggregation were selected for experimental testing. The letter size in sequence logo plots is visualizing the amino acid frequency at the corresponding positions in the predicted 6-mer sequences. A detailed listing of amino acids abundancy for the 3669 peptides is shown in Fig. S4.† The Venn diagram is summarizing the main composition motifs of the 3669 peptides. (C) t-SNE dimension reduction plot of the predicted sequences (12 320, blue), the selected subset (16, large yellow symbols with sequence information) and training set (grey). The data points are projected from a 100-d space to a 2-d space with t-SNE.

To sample the sequence space in a time and computational wise cost-efficient way we applied a Monte Carlo model, so that not all possible 6-mer sequences (20^6 = 64 million) must be calculated. Starting from 16 initial 6-mers with randomly chosen residues, we executed 16 independent MC runs which were continued for 1 million steps, generating a peptide at each step. The generated peptides were retained based on a Metropolis criterion (Fig. 2B, see Methods; full lists of the retained peptides, along with code, data, and the trained ML models, can be found at **https://gitlab.com/arghyadutta/seq-to-infect**).

From each of the 16 MC runs, only 1000 sequences with the largest predicted infectivity values were kept, yielding 12 320 sequences after removing duplicates (Table S3†). The predicted 6-mer sequences were represented in a 2-d t-distributed stochastic neighbor embedding (t-SNE)[29] dimensionality

 This journal is © The Royal Society of Chemistry 2023

View Article Online
Biomaterials Science Paper

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

reduction plot from 100-d vector space to check their semantic varieties (Fig. 2C). The t-SNE algorithm attempts to cluster sequences that are semantically close to each other regarding their amino acid composition. As expected, the training set peptides, which have similar amino acid compositions by design, formed a cluster that is distinctly separate from the widely distributed clusters of the generated sequences.

Criteria for selecting sequences

To further narrow down the list of potentially interesting sequences for experimental evaluation, we applied two selection criteria based on previously found evidence for infectivity-enhancing self-assembling peptides: charge and propensity for aggregation. Aggregation is important for bioactivity because without the amyloid structure, the unique properties in cell-adhesion[30] and retroviral transduction enhancement[8] are lost. To find the peptides, that are prone to aggregate and form fibrillar structure, we applied open source protein-aggregation tools Tango,[31] APPNN,[32] Waltz,[33] PATH,[34] Aggrescan[35] and PASTA 2.0.[36] Since these tools are developed for polypeptides and proteins, we first evaluated the reliability of each tool to predict aggregation of self-assembling short peptides with the accuracy and receiver operating characteristic (ROC) value (Fig. S2†) by using our training set (Table S1†).[8] APPNN,[32] PATH,[34] and Aggrescan[35] performed best (Fig. S2†) and were used to select aggregation prone peptides.

Out of the 12 320 predicted sequences for infectivity enhancement, 3669 peptides have a net positive charge; 424 of these 3669 peptides were predicted for aggregation by at least two of Aggrescan, PATH, and APPNN (for detailed analysis see ESI Section 2, Table S4†).

Most of these peptides contain the motif "WWN" (1600 of 3669) or the amino acid Cysteine (2009 of 3669) as visualized in the Venn diagram and in the sequence logo plot (Fig. 2B, and Fig. S4†). Interestingly, the motif "WWN" does not appear in any of the training set peptides, whereas Cysteine was shown previously by us to contribute positively to infectivity enhancement.[7,8] For experimental evaluation, peptides with a large variation in sequences were selected from different clusters in the t-SNE map in order to cover a large sequence space (Fig. 2C). Other than predicted infectivity, hydrophobicity, and aggregation propensity (Fig. S3†), we considered *N*-gram similarity scores (Fig. S5, and Table S5†) to ensure diverse selection of sequences. Finally, from the total 16 peptides selected for experimental evaluation, 8 of the peptides were predicted for aggregation, and 8 sequences were not predicted for aggregation as a control group (Fig. S3†). All these peptides strongly differ sequence wise from the training set, as visualized *via* sequence plot and *N*-gram similarity scores (Fig. 2B, and Fig. S5†).

Prediction yields highly active *de novo* peptides

16 short 6-mer peptides predicted for high biological activity (enhanced HIV-1 infection rates) were selected for experimental evaluation of their physicochemical properties and infectivity enhancement. All the tested peptides are biocompatible as tested *via* a cell-viability assay (Fig. S8†).

4 of the 16 peptides show remarkable infectivity enhancement above 10% relative to EF-C (Fig. 3A, and Fig. S7†). It is important to recognize the inherent difference in sequence length when comparing the infectivity enhancement of the newly found sequences with EF-C. The four infectivity enhancing 6-mer peptides have roughly half the molecular weight of the 12-mer EF-C. Therefore, these peptides exhibit approximately twice the infectivity enhancing efficiency in terms of mass concentration when compared to EF-C. Consequently, a direct comparison needs to consider this significant difference of sequence length. Further, these 4 peptides were predicted for aggregation, resulting in a hit rate of 50% based on the selected 8 aggregation prone peptides (Fig. S3†) or a hit rate of 25% relative to the entire selection. Interestingly, among these peptides only one peptide (ICICLK) shows a positive zeta-potential. The other 3 peptides (HVWCIF, HICLFW, HFICIC) form fibrils, colocalize with cell-membranes (Fig. 3B–E) and show infectivity enhancement despite their moderately negative zeta-potentials. We wondered whether these hit peptides show a different mode of action and applied a property–activity correlation model, which was developed with the training set (ESI section 4†).[8] The newly created peptides fit in the model well (R = 0.72, Fig. 3F), which indicates an interaction mode comparable to the training set: the peptide fibrils associate with viruses and colocalize them with cellular membranes, which facilitates the uptake and increase infection rate.[4]

3 Discussion

ML is an emerging approach to discover new active compounds from a diverse chemical space.[37] The main benefit of a ML-based approach is that it can extract underlying structure–activity relationships and use this information to predict new structures with the desired activity. With the increasing complexity of systems, it is becoming more difficult to rationally design structures from scratch, as experienced often for self-assembling structures that span nm to μm length scales and differ strongly in their biological activities. In this regard, ML is a promising approach to discover bioactive self-assembling peptides, which are difficult to rationally design *de novo*. Here, we show that a ML approach, relying on continuous vector representation for proteins, can be applied to design and predict short 6-mer peptides with diverse sequence and physicochemical properties.

Our training set contains 163 peptide sequences based on derivatives of an active compound. Small training sets are common in early stages of research, but they are rarely considered for ML approaches that aim to predict new sequence spaces since training the model is difficult.[20,27,38–42] As shown here, new peptide sequence spaces can be discovered *via* a computational approach that combines ML and MC with further screening, while still using a small training set with a wide variation in bioactivity.

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

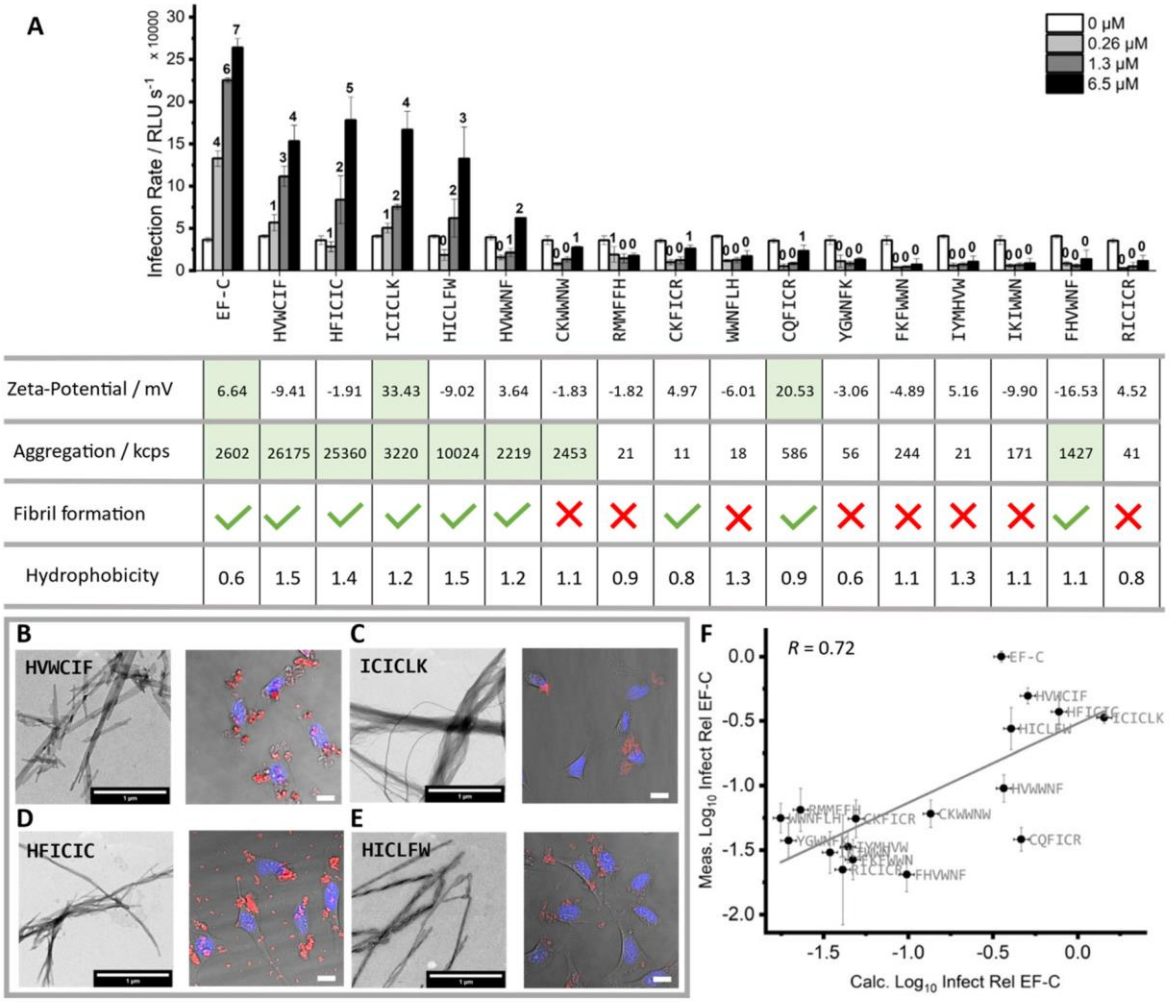


Fig. 3 (A) Summary of infectivity enhancement and physicochemical properties of *de novo* predicted peptides (Table S2†). Peptides are incubated in PBS at 1 mg mL^{-1} for 1 d at RT before characterization. Absolute HIV-1 infection rates in the presence of EF-C (QCKIKQIINMWQ) and peptides from ML-prediction at 6.5 µM, 1.3 µM, 0.26 µM and 0 µM (virus only infection). The n-fold infection rates relative to virus only control is shown for each column. The aggregation into µm-sized colloids was determined by light scattering count rate during zeta-potential measurements. The molecular aggregation into nm-sized fibrils was determined by transmission electron microscopy (TEM, Fig. S6†). Hydrophobicity was calculated according to Fauchere hydropathy scale.[50] (B–E) TEM showing fibril morphology (scale bar 1 µm) and confocal fluorescence microscopy showing cell-fibril-colocalization (scale bar 20 µm) of hit peptides (B) HVWCIF, (C) ICICLK, (D) HFICIC, (E) HICLFW. For TEM measurements the peptides were stained with 4 wt% uranylacetat during preparation. For confocal fluorescence microscopy the fibrils were stained with Proteostat and diluted to 20 µg mL^{-1} before adding to HeLa cells (40 000/1 cm^2) which nucleus was stained with Hoechst 33342. (F) Model describing property–activity correlation for the training set, see ESI Section 4,† applied on newly found peptides (Pearson correlation coefficient R = 0.72).[8] The plot shows the infectivity enhancement of peptides relative to the infectivity enhancement of EF-C in a logarithmic scale. The calculated value is determined by ESI eqn (3).†

Interestingly, most of the predicted peptides are rich in hydrophobic amino acids cysteine and tryptophan, the latter mainly from the sequence motif "WWN". Sequences which contain these hydrophobic amino acids enhance infectivity if they form fibrillar structures that are influenced by cysteine's capability to form disulfide bonds (ESI chapter 9, Fig. S10A–D†). Notably, in a previous study on the training set, we discovered a higher prevalence of cysteine in active peptides.[7,8] However, it was found that cysteine was not essential for activity.[7] In contrast to the *de novo* peptides found in this study, such as ICICLK, the non-essential role of cysteine in the training set can be attributed to the strong self-assembly tendency of peptides with amphiphilic sequence patterns. These patterns stabilize the structure even after the disulfide bonds are broken (Fig. S10F†). Therefore, our machine learning approach successfully extracted the importance of cysteine for peptide fibril formation and incorporated this information in the newly found sequences. We showed here that fibrillar structure formation can be predicted with an accuracy of ~75% through the combination of open-source aggregation prediction tools Aggrescan,[35] APPNN,[32] and PATH.[34] We found that while these algorithms were developed for polypeptides

 This journal is © The Royal Society of Chemistry 2023

View Article Online
Biomaterials Science Paper

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

and proteins, they also perform well for short self-assembling peptides.

The infectivity enhancement of peptide fibrils occurs due to improved colocalization of viruses with cell-membrane.[4] The main driving force for this interaction is believed to be electrostatic interactions;[7,9,11,43] where positively charged fibrils sequester negatively charged virions which in turn bind to the negatively charged cellular membrane. As virion attachment to the cell membrane is the major rate limiting step during viral entry, increased numbers of virions at the cell surface result in higher cell entry and infection rates. However, we here found that fibrils with a moderately negative zeta-potential can also increase infectivity. These kinds of fibrils were not included in the training set and not reported before. We hypothesize that oversimplification of the fibril–cell-membrane interaction by reducing it to solely electrostatic interactions can be misleading since the cell interaction of fibrils is regulated by an intricate balance between charge and hydrophobicity. For example, the peptide CQFICR (Fig. 3A) forms fibrils and has a positive zeta-potential but does not enhance infectivity. The low hydrophobicity for CQFICR (0.9) results in less aggregated fibrils and decreases hydrophobic cell-membrane interaction. Another example is demonstrated with the hydrophobic peptide FHVWNF (Fig. 3A), which forms aggregating fibrils with a negative zeta-potential but does not enhance infectivity due to the contribution of the strong negative zeta-potential, as supported by our property–activity model (Fig. 3F). A further example are fibrils derived from the immunoglobulin light chain that have a net negative surface charge and retain virion-binding activity but lack cell-binding and viral transduction enhancing properties.[43] More recently it has also been shown that cellular protrusions actively engage EF-C fibril/virion complexes, suggesting that not only electrostatic interactions may account for bioactivity.[44] It is important to note that meaningful comparisons of zeta-potential can only be made among peptides that either form aggregates or do not form aggregates. This is because the size of colloidal particles has an influence on the measured zeta-potential.[45]

Our method demonstrates that moderately negatively charged peptide fibrils can be active if the hydrophobicity and aggregation features are both strongly pronounced. Hydrophobic amino acids such as tryptophan, phenylalanine, and cysteine can facilitate these desired properties; the continuous vector representation of peptides successfully extracts this underlying information by processing sequence and activity information of the training set without the requirement to assume a predetermined set of relevant descriptors as often done in traditional prediction approaches.[46] The predicted sequences show a higher hydrophobicity, on average, than reported for the training set (Fig. S11†).

Taken together, our method offers a promising tool to yield diverse peptide structures, which cannot be created rationally from derivatives of active compounds or by using conventional approaches such as sequence–pattern analysis.[8,47]

Finally, all these newly found active peptides are the shortest infectivity-enhancing peptides known to us and not found in any protein databases, which makes them truly *de novo*.

4 Conclusions

Discovering new supramolecular nanostructures with a desired bioactivity is challenging due to complex multi-parameter and multi-scale dependencies. In this work, we report an inverse design approach by applying a computational pipeline that includes using a continuous vector representation-based Machine Learning model, a Monte Carlo sampling, and a final screening based on charge and aggregation propensity. A small peptide library, based on derivatives of an active peptide, was utilized to predict short 6-mer self-assembling peptides with prospects for their application as transduction enhancers in basic research and the clinics.

The strength of a continuous vector representation-based approach is that it can encode sequence and physicochemical information of a peptide into a numerical vector which can then be used to train an ML model. Monte Carlo sampling, using the trained ML model, enables us to screen a large sequence space in a time- and cost-efficient way and yield *de novo* active peptide sequences, which are structure- and property-wise very different from the training set. We envision that our data-driven method will substantially accelerate the early stages of research by screening large sequence spaces and predicting *de novo* peptides starting from a small dataset, which are unexpected by human experience and rational design.

5 Materials and methods

Materials for peptide characterization

Dimethylsulfoxide (DMSO, ACS reagent, ≥99.9%) was purchased from Honeywell, Riedel-de Haën®. Uranyl acetate was purchased from Merck. PBS was purchased from Sigma Aldrich. Proteostat® was purchased from Enzo Life Sciences. All chemicals were used as received.

Creation of peptide library and synthesis of new peptides

The data on infectivity of a peptide library based on the gp120 derivative enhancing factor C (EF-C) was reported previously by us.[8] The peptides were synthesised according to standard Fmoc solid phase peptide synthesis and were commercially obtained Phtd Peptides industrial Co. limited with purity of ≥95% as determined by liquid chromatography mass spectrometry.

Preparation of peptide fibrils

The peptides were dissolved in DMSO (c = 10 mg mL^{-1}, stored at 4 °C before usage) and added to PBS buffer to a final incubation concentration of 650 μM or 1 mg mL^{-1} as indicated in the text. Other concentrations described in the text were achieved by further dilution of the peptides.

Physicochemical characterization of peptide fibrils

TEM. 5 μL of peptide fibrils (1 mg mL^{-1}) were placed on copper grids which were placed on copper grids coated with carbon and formvar layer (300 mesh, Plano GmbH). After

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

10 min incubation time, the grids were staining with 4% uranyl acetate solution for 2.5 min and washed with water. Measurements were performed on a Jeol 1400 electron microscope with 120 kV acceleration voltage.

ATR FT-IR. To determine the β-sheet content of peptide fibrils ATR FT-IR spectroscopy measurements were conducted by lyophilizing 200 μL of the respective 1 mg mL^{-1} peptide solution. All spectra were recorded on a Bruker Tensor 27 spectrometer with a diamond crystal as ATR element (PIKE Miracle™, spectral resolution 2 cm^{-1}). Every sample was measured with 64 scans and processed with OriginLab software according to a previous report.[7]

ThT-assay. 4 μL of 1 mg mL^{-1} peptide fibrils solution is added to a 20 μL, 50 μM ThT-solution in PBS. The mixture is incubated for 15 min at RT. For reference PBS containing 10% DMSO (4 μL) was added instead of peptide fibril solution. The samples were placed in black UV Star® 384 microliter well-plates (Greiner bio-one). Fluorescence spectra were recorded on an Infinite® M1000 PRO microplate reader (Tecan) at e_{em} = 488 nm upon excitation at e_{ex} = 440 nm with 10 nm bandwidths and multiple reads per well (3 × 3). Measurements were repeated in triplicates and averaged with standard deviation. A peptide fibril was considered as ThT-active if the fluorescence intensity was at least twice as strong compared to the control.

Zeta-potential. For the determination of the surface charge of aggregated peptides zeta-potential measurements were conducted with a Zetasizer Nano ZS, Malvern Instruments. Unless stated otherwise the peptides were dissolved from DMSO (10 mg mL^{-1}) in PBS (pH 7.4) to concentration of 1 mg mL^{-1} and incubated for 1d at RT. Just before the measurement the 60 μL of the peptides were further diluted in 600 μL KCl (aq, 1 mM) in a 1 mL disposable folded capillary cells (DTS-1060, Zetasizer Nano series, Malvern). The zeta-potential of the peptides was derived from the electrophoretic mobility based on the Smoluchowski formula. All measurements were averaged in triplicates and reproduced at least once for each sample.

Aggregate analysis. The derived count rate of scattered light at 633 nm, 173° from the zeta-potential measurement was used as information on the light scattering intensity and turbidity of the sample as an indicator for microscopic aggregation.[8,48]

Virus-peptide interaction

Human wild-type HeLa cell line were obtained from abcam, ab260075, LOT: GR3292155-1. TZM-bl cell line and R5-tropic HIV-1 stock plasmid (pBRNL4.39-92TH014) were obtained from the National Institutes of Health AIDS Research and Reference Reagent Program as reported earlier.[49]

R5-tropic HIV-1 stocks and HIV-1 infection assays were prepared analogous to a previous report.[7] Briefly, the effect of peptide fibrils (final concentration on cells 6.5, 1.3, 0.26, 0 μM) on HIV-1 infection was studied *via* a luminescence assay for detection of β-galactosidase, which is expressed upon HIV-1 infection of TZM-bl cells. The HIV-1 infection assay was conducted in three technical replicates and reproduced at least once. Note, that *n*-fold infectivity enhancement of peptides relative to virus only infection rates are strongly dependant of initial virus concentration. To compare independent measurements with each other EF-C (QCKIKQIINMWQ) was always used as a reference peptide. EF-C is the original sequence on which the training set is based and was applied previously by us to quantify infection rates.[4,7,8]

Cell viability was determined after addition of peptides to TZM-bl cells *via* the CellTiter-Glo assay. To this end, 10 000 cells were seeded and on the next day serial diluted peptides were added. After 3 days the supernatant was removed and 100 μL CellTiter-Glo Reagent 1 : 1 diluted in PBS was added. After 10 min 50 μL was transferred to white microplate and luminescence was recorded by Orion microplate luminometer.

Confocal laser scanning microscopy studies were performed for the visualization of the cell-peptide interaction. HeLa cells were seeded one day prior to conducting the assay (40 000 per well) in an 8-well IBIDI slide. 4 μL of the preformed peptide fibrils (1 mg mL^{-1}) were diluted with 4 μL Proteostat (Enzo Life Science, 1 μL stock in 999 μL PBS) and further diluted with medium to receive a final peptide concentration of 20 μg mL^{-1}. The nucleus of the HeLa cells was stained with Hoechst 33342 (NucBlue™, Thermo Fisher Scientific). The peptide solution mixture was transferred to the HeLa cells and incubated for 30 min at 37 °C before washing three times with PBS. The interaction of fibril clusters with cells was monitored after 30 min incubation time on a Stellaris 8 confocal laser scanning microscope (Leica) equipped with a 20× air objective and laser excitation wavelength of 405 nm (Hoechst) and 561 nm (Proteostat).

Calculation of hydrophobicity and net charge

Net charge. The net charge of a peptide was calculated at pH 7.4 and the EMBOSS pKscale.

Hydrophobicity. The calculation of the hydrophobicity is based on experimental based on water–octanol partition coefficients of amino acids according to Fauchere[50] hydropathy scale. The calculation of the net charge and hydrophobicity were conducted *via* the "peptides" package in R.[51]

Details of computational pipeline

Machine learning. A peptide library consisting of 163 sequences derived from EF-C (QCKIKQIINMWQ), which was previously reported by us (Table S1†),[8] was used as the training set for the ML model. To this end, each peptide of the training set was represented as 100-dimensional numerical vectors (ProtVecs) by summing up ProtVecs of their constituent 3-grams (comprising 3 consecutive amino acid) as outlined by Asgari and Mofrad.[26] Scikit-learn was used to train and cross-validate LASSO and RIDGE regression models that connected the ProtVecs to the experimentally determined infectivity of the peptides.

Metropolis criterion. To sample the vast sequence space of the 6-mers, we used a MC method as described in the Results section. At each MC step, we generated a new peptide sequence by randomly changing one amino acid of the old peptide from the previous step and computed the infectivity of

 This journal is © The Royal Society of Chemistry 2023

View Article Online
Biomaterials Science Paper

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

the old and new peptides using the trained ProtVec LASSO model (ESI eqn (1)†). We denote the infectivity of an old and a new peptide sequence I_0 and I_n, respectively. Following the standard Metropolis protocol, we accept the Monte Carlo move under the acceptance criterion

$$\mathrm{acc}(0 \rightarrow n) = \begin{cases} 1, I_n - I_0 \geq 0 \\ \exp\left(\frac{I_n - I_o}{t}\right), I_n - I_0 < 0 \end{cases}$$

where t denotes the reduced temperature (set to 1 throughout). The Metropolis criterion is evaluated by comparing the second term to a uniformly distributed random number in the interval $[0, 1)$.

All code and data used for ML and MC analysis are openly available at **https://gitlab.com/arghyadutta/seq-to-infect**.

Data visualization

Sequence plot. To visualize the amino acid frequency in the peptide library the sequence logo generator tool WebLogo[52] was applied. The sequence plot shows the amino acid frequency in the first seven positions starting from the N-terminus of the library or the first 6 positions starting from the N-terminus of the ML-generated peptides. All sequences of the library shorter than 7 AA (14 out of 163 peptides) were excluded for this visualization. The color code represents amino acid side chain hydrophobicity and charge categorization according to Kyte–Doolittle hydropathy scale. Green represents hydrophilic positive charge, yellow hydrophobic no charge, purple hydrophilic no charge and red hydrophilic negative charge.

t-SNE plot. For the visualization of sequence similarities in the library and the predicted peptides dimensionality reduction approach t-SNE (t-distributed stochastic neighbor embedding) was applied. To this end, the sequences were vectorized by linear vector addition of 3-mer ProtVec vector embeddings. The 100-d vector embeddings were applied as features in t-SNE, which was operated in the data analysis software Orange3 (V. 3.26.0)[53] with the following conditions: Perplexity 30, exaggeration 1, PCA components 20, normalized data. The parameters were selected empirically after testing several conditions for displaying distinct clusters.

Author contributions

K.K. designed, carried out the experiments, analyzed data, performed the formal analysis, and wrote the original draft. A.D. designed and carried out the ML prediction, analyzed data, performed the formal analysis, and revised the manuscript. L. R.W. designed and carried out the experiments related to virus transduction, analyzed data, performed the formal analysis, and revised the manuscript. T.B. conceived the idea, conceptualized, and analyzed data. C.V.S., J.M, T.B. and T.W supervised, provided the funding resources, and revised the manuscript. The final manuscript was approved by all authors.

Conflicts of interest

J.M. and L.R. hold patents for application of peptide fibrils for retroviral transduction enhancement.

Acknowledgements

The authors would like to thank Prof. Dr. Pascal Friederich for valuable discussions. Further, the authors acknowledge open-source ProtVec vector embedding script,[26,54] WebLogo,[52] Orange3[53] NumPy,[55] Pandas,[56] and Scikit-learn[57] and the freely available aggregation prediction tools Tango,[31] APPNN,[32] Waltz,[33] Path,[34] Aggrescan[35] and PASTA 2.0[36] that were used in this work.

The study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number 316249678—SFB 1279 (A02, A03, A05, C01). T. B. acknowledges support from the Emmy Noether program of the Deutsche Forschungsgemeinschaft (DFG). A. D. acknowledges support by BiGmax, the Max Planck Society's Research Network on Big-Data-Driven Materials Science. Open Access funding provided by the Max Planck Society.

References

1 N. J. Sinha, M. G. Langenstein, D. J. Pochan, C. J. Kloxin and J. G. Saven, *Chem. Rev.*, 2021, **121**, 13915–13935.
2 A. L. Boyle and D. N. Woolfson, *Chem. Soc. Rev.*, 2011, **40**, 4295.
3 L. Naldini, *Nature*, 2015, **526**, 351–360.
4 M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständker, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff and J. Münch, *Nat. Nanotechnol.*, 2013, **8**, 130–136.
5 C. Meier, T. Weil, F. Kirchhoff and J. Münch, *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.*, 2014, **6**, 438–451.
6 K. Kaygisiz and C. V. Synatschke, *Biomater. Sci.*, 2020, **8**, 6113–6156.
7 S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, K. Kaygisiz, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch and T. Weil, *Adv. Funct. Mater.*, 2021, **31**, 2009382.
8 K. Kaygisiz, L. Rauch-Wirth, A. Dutta, X. Yu, Y. Nagata, T. Bereau, J. Münch, C. V. Synatschke and T. Weil, *ChemRxiv*, 2023, DOI: **10.26434/chemrxiv-2023-hfqxb**.
9 J. Münch, E. Rücker, L. Ständker, K. Adermann, C. Goffinet, M. Schindler, S. Wildum, R. Chinnadurai, D. Rajan, A. Specht, G. Giménez-Gallego, P. C. Sánchez,

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

D. M. Fowler, A. Koulov, J. W. Kelly, W. Mothes, J. C. Grivel, L. Margolis, O. T. Keppler, W. G. Forssmann and F. Kirchhoff, *Cell*, 2007, **131**, 1059–1071.
10 B. Dai, D. Li, W. Xi, F. Luo, X. Zhang, M. Zou, M. Cao, J. Hu, W. Wang, G. Wei, Y. Zhang and C. Liua, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 2996–3001.
11 S. Kirti, K. Patel, S. Das, P. Shrimali, S. Samanta, R. Kumar, D. Chatterjee, D. Ghosh, A. Kumar, P. Tayalia and S. K. Maji, *ACS Biomater. Sci. Eng.*, 2019, **5**, 126–138.
12 P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn and T. Tuttle, *Nat. Chem.*, 2015, **7**, 30–37.
13 C. Wu and J. E. Shea, *Curr. Opin. Struct. Biol.*, 2011, **21**, 209–220.
14 C. Wu, R. Gao, Y. Zhang and Y. De Marinis, *BMC Bioinf.*, 2019, **20**, 456.
15 F. Strieth-Kalthoff, F. Sandfort, M. H. S. Segler and F. Glorius, *Chem. Soc. Rev.*, 2020, **49**, 6154–6168.
16 F. Sandfort, F. Strieth-Kalthoff, M. Kühnemund, C. Beecks and F. Glorius, *Chem*, 2020, **6**, 1379–1390.
17 R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams and A. Aspuru-Guzik, *ACS Cent. Sci.*, 2018, **4**, 268–276.
18 D. Ofer, N. Brandes and M. Linial, *Comput. Struct. Biotechnol. J.*, 2021, **19**, 1750–1758.
19 K. L. Saar, A. S. Morgunov, R. Qi, W. E. Arter, G. Krainer, A. A. Lee and T. P. J. Knowles, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2019053118.
20 E. Y. Lee, B. M. Fulan, G. C. L. Wong and A. L. Ferguson, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 13588–13593.
21 E. M. López-Vidal, C. K. Schissel, S. Mohapatra, K. Bellovoda, C.-L. Wu, J. A. Wood, A. B. Malmberg, A. Loas, R. Gómez-Bombarelli and B. L. Pentelute, *JACS Au*, 2021, **1**, 2009–2020.
22 R. Batra, T. D. Loeffler, H. Chan, S. Srinivasan, H. Cui, I. V. Korendovych, V. Nanda, L. C. Palmer, L. A. Solomon, H. C. Fry and S. K. R. S. Sankaranarayanan, *Nat. Chem.*, 2022, **14**, 1427–1435.
23 T. Mikolov, I. Sutskever, K. Chen, G. Corrado and J. Dean, *Adv. Neural Inf. Process. Syst.*, 2013, 3111–3119.
24 R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu and P. Kuksa, *J. Mach. Learn. Res.*, 2011, **12**, 2493–2537.
25 V. Tshitoyan, J. Dagdelen, L. Weston, A. Dunn, Z. Rong, O. Kononova, K. A. Persson, G. Ceder and A. Jain, *Nature*, 2019, **571**, 95–98.
26 E. Asgari and M. R. K. Mofrad, *PLoS One*, 2015, **10**, e0141287.
27 C. Hoffmann, R. Menichetti, K. H. Kanekal and T. Bereau, *Phys. Rev. E*, 2019, **100**, 033302.
28 R. Tibshirani, *J. R. Stat. Soc. Ser. B*, 1996, **58**, 267–288.
29 L. Van Der Maaten and G. Hinton, *J. Mach. Learn. Res.*, 2008, **9**, 2579–2605.
30 K. Kaygisiz, A. M. Ender, J. Gačanin, L. A. Kaczmarek, D. A. Koutsouras, A. N. Nalakath, P. Winterwerber, F. J. Mayer, H. Räder, T. Marszalek, P. W. M. Blom, C. V. Synatschke and T. Weil, *Macromol. Biosci.*, 2023, **23**, 2200294.
31 A.-M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz and L. Serrano, *Nat. Biotechnol.*, 2004, **22**, 1302–1306.
32 C. Família, S. R. Dennison, A. Quintas and D. A. Phoenix, *PLoS One*, 2015, **10**, e0134679.
33 J. Beerten, J. Van Durme, R. Gallardo, E. Capriotti, L. Serpell, F. Rousseau and J. Schymkowitz, *Bioinformatics*, 2015, **31**, 1698–1700.
34 J. W. Wojciechowski and M. Kotulska, *Sci. Rep.*, 2020, **10**, 7721.
35 O. Conchillo-Solé, N. S. de Groot, F. X. Avilés, J. Vendrell, X. Daura and S. Ventura, *BMC Bioinf.*, 2007, **8**, 65.
36 I. Walsh, F. Seno, S. C. E. Tosatto and A. Trovato, *Nucleic Acids Res.*, 2014, **42**, W301–W307.
37 S. Giguère, F. Laviolette, M. Marchand, D. Tremblay, S. Moineau, X. Liang, É. Biron and J. Corbeil, *PLoS Comput. Biol.*, 2015, **11**, e1004074.
38 K. H. Kanekal and T. Bereau, *J. Chem. Phys.*, 2019, **151**, 164106.
39 A. Dutta, J. Vreeken, L. M. Ghiringhelli and T. Bereau, *J. Chem. Phys.*, 2021, **154**, 244114.
40 A. Capecchi and J.-L. Reymond, *Med. Drug Discovery*, 2021, **9**, 100081.
41 C. Rauer and T. Bereau, *J. Chem. Phys.*, 2020, **153**, 014101.
42 H. Altae-Tran, B. Ramsundar, A. S. Pappu and V. Pande, *ACS Cent. Sci.*, 2017, **3**, 283–293.
43 D. Schütz, C. Read, R. Groß, A. Röcker, S. Rode, K. Annamalai, M. Fändrich and J. Münch, *ACS Omega*, 2021, **6**, 7731–7738.
44 D. Schütz, S. Rode, C. Read, J. A. Müller, B. Glocker, K. M. J. Sparrer, O. T. Fackler, P. Walther and J. Münch, *Adv. Funct. Mater.*, 2021, **31**, 2104814.
45 Y. Chen, H. X. Gan and Y. W. Tong, *Macromolecules*, 2015, **48**, 2647–2653.
46 V. Samsoninkova, N. L. Venkatareddy, W. Wagermaier, A. Dallmann and H. G. Börner, *Soft Matter*, 2018, **14**, 1992–1995.
47 C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil and B. Knöll, *Adv. Funct. Mater.*, 2019, **29**, 1809112.
48 Y. Yoshimura, Y. Lin, H. Yagi, Y. H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki and Y. Goto, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14446–14451.
49 A. Papkalla, J. Münch, C. Otto and F. Kirchhoff, *J. Virol.*, 2002, **76**, 8455–8459.
50 J.-L. Fauchère, M. Charton, L. B. Kier, A. Verloop and V. Pliska, *Int. J. Pept. Protein Res.*, 2009, **32**, 269–278.
51 D. Osorio, P. Rondón-Villarreal and R. Torres, *R J.*, 2015, **7**, 4.
52 G. E. Crooks, G. Hon, J.-M. Chandonia and S. E. Brenner, *Genome Res.*, 2004, **14**, 1188–1190.

 This journal is © The Royal Society of Chemistry 2023

53 J. Demšar, A. Erjavec, T. Hočevar, M. Milutinovič, M. Možina, M. Toplak, L. Umek, J. Zbontar and B. Zupan, *J. Mach. Learn. Res.*, 2013, **14**, 2349–2353.
54 Replication Data for: Continuous Distributed Representation of Biological Sequences for Deep Proteomics and Genomics - Harvard Dataverse, **https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/JMFHTN**, (accessed 3 February 2023).
55 C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke and T. E. Oliphant, *Nature*, 2020, **585**, 357–362.
56 W. Mckinney, in Proc. of the 9th Python in Science Conf. (SCIPY 2010), 2010, pp. 56–61.
57 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel and B. Thirion, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

Open Access Article. Published on 15 June 2023. Downloaded on 7/31/2023 6:41:07 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Supporting Information

Electronic Supplementary Material (ESI) for Biomaterials Science.
This journal is © The Royal Society of Chemistry 2023

Supplementary Information for

Inverse design of viral infectivity-enhancing peptide fibrils from continuous protein-vector embeddings

Kübra Kaygisiz, Arghya Dutta, Lena Rauch-Wirth, Christopher V. Synatschke, Jan Münch, Tristan Bereau*, Tanja Weil*

K. Kaygisiz, Dr. C. V. Synatschke, Prof. Dr. T. Weil
Department Synthesis of Macromolecules
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
Email: weil@mpip-mainz.mpg.de, synatschke@mpip-mainz.mpg.de

Dr. A. Dutta, Dr. T. Bereau
Polymer Theory
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany

Dr. A. Dutta,
Institute of Biochemistry II, Faculty of Medicine,
Goethe University,
Theodor–Stern–Kai 7, 60590 Frankfurt, Germany (*present address*)

Dr. T. Bereau
Institute for Theoretical Physics, Heidelberg University,
Philosophenweg 19, 69120 Heidelberg, Germany (*present address*)
Email: bereau@thphys.uni-heidelberg.de

L. Rauch-Wirth, Prof. Dr. J. Münch
Institute of Molecular Virology
Ulm University Medical Center
Meyerhofstraße 1, 89081 Ulm, Germany
Email: jan.muench@uni-ulm.de

Corresponding authors: Tristan Bereau (bereau@thphys.uni-heidelberg.de) and Tanja Weil (weil@mpip-mainz.mpg.de)

Contents

1. Regression Model: Infectivity Prediction 3
2. Aggregation Prediction 4
3. N-gram similarity for predicted peptides with training set 8
4. Evaluation of De Novo Peptide Activity with Property–Activity Model[4] 9
5. TEM micrographs 10
6. Infectivity Data 11
7. Cell-Viability 11
8. FT-IR Data 12
9. Impact of Disulfide Bond Formation on Self-Assembly 13
10. Amino acid Composition Analysis 15
11. Training Set 16
12. Predicted Peptides Characterization 19
13. References 21

1. Regression Model: Infectivity Prediction

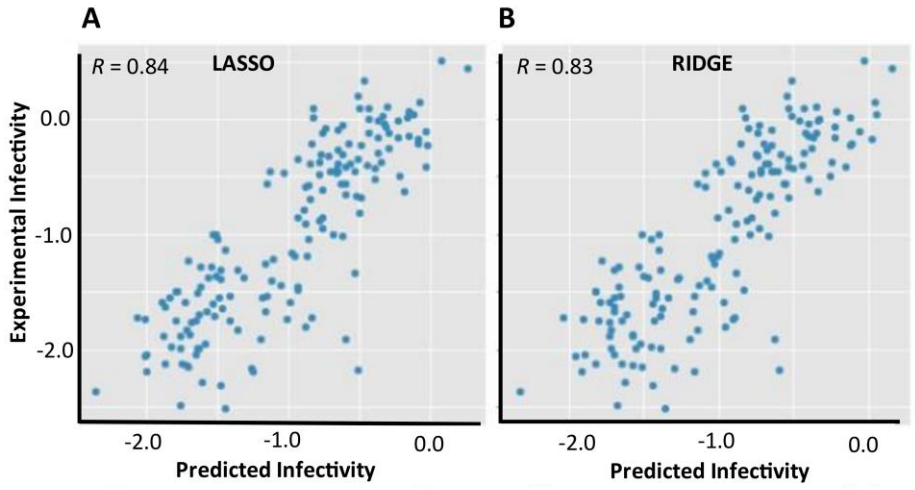


Figure S1 A LASSO and **B** RIDGE linear regression models were trained via a 5-fold cross validation.

LASSO and RIDGE regression models were trained on the training set peptides, represented as 100-d numerical vectors using continuous vector representations; the models perform similarly (**Figure S1**). Since LASSO regression applies regularization by minimizing the number of non-zero coefficients, the resulting model contains only the relevant parameters. This results in a simpler model with fewer parameters. For example, for our model only 21 vectors have a non-zero coefficient (**Eqn. 1**). Interestingly, while RIDGE regression equation contains all the vector components, it offers slightly poorer correlation as shown in (**Figure S1**).

(**Eqn. 1**) Log Infect. Rel EF-C = -2.004 + (-0.496) * vec40 + (-0.467) * vec4 + (-0.374) * vec30 + (-0.276) * vec68 + (-0.239) * vec8 + (-0.212) * vec20 + (0.179) * vec55 + (-0.17) * vec88 + (0.153) * vec16 + (0.131) * vec59 + (0.129) * vec57 + (0.099) * vec100 + (-0.095) * vec43 + (0.06) * vec22 + (0.057) * vec46 + (-0.05) * vec99 + (-0.03) * vec71 + (0.024) * vec60 + (-0.009) * vec61 + (-0.009) * vec54 + (-0.002) * vec52

Mean squared error = 0.179
Mean absolute error = 0.328
R^2= 0.695
Pearson R = 0.837

All code and data used in ML training are openly available at https://gitlab.com/arghyadutta/seq-to-infect.

2. Aggregation Prediction

Aggregation was found as a necessary property for infectivity enhancement of peptides previously by us and others.[1–4] Therefore, we applied the open accessible protein-aggregation tools Tango,[5] APPNN,[6] Waltz,[7] PATH,[8] Aggrescan[9] and PASTA 2.0[10] to preselect promising *de novo* created peptides.

Aggrescan is based on statistical analysis of the aggregation-propensity value for each amino acid residue in the sequence and a subsequent aggregation prediction by hot spot regions, identified from the peptide aggregation profile. Here, we consider a sequence as amyloidogenic if there is at least one predicted hotspot.

Waltz applies statistical analysis of a sequence and was originally developed by position specific matrix for 1089 short 6-mer peptides sequences, which were experimentally determined for fibril formation.[7] Here, we considered a sequence as amyloidogenic if at least one amyloidogenic region was detected upon entry of following parameters: threshold custom 0-100 and pH 7.0.

Tango is designed to predict aggregating regions in unfolded polypeptide chains. statistical mechanics algorithm. The method is benchmarked against experimentally observed 179 peptides.[5] Here, we applied following input parameters to determine the β-sheet aggregation tendency (aggregation parameter): pH 7.4, 298 K, ionic strength 0.1724. We select a threshold above 5.0% over 5 residues to identify hotspots for aggregation as suggested by the authors to determine amyloidogenic sequences.[5]

PATH is a structure-based method for predicting amyloidogenicity by threading and machine learning. Here, we considered a peptide as aggregating if at least one amyloidogenic region was calculated.

PASTA 2.0 is based on energetic functions which were determined experimentally from protein structures interactions potential and H-bond formation between all non-consecutive residues for parallel and anti-parallel β-pairing. A sequence is considered amyloidogenic if the pasta energy for the lowest predicted pairing is lower or equal to the threshold stated by the authors (-4.0).[10] The parameters for the prediction was threshold custom, top pairing energy 20, energy threshold -2 PEU, large scale true, protein-protein analysis: false.

APPNN applies a neural network machine learning approach based on the analysis of seven physicochemical and biochemical features such as β-sheet frequency, hydrophobic moment, helix termination parameters or isoelectric point. A sequence was considered amyloidogenic if at least one of these six amino acid windows was classified amyloidogenic.

Except for Waltz, these prediction tools were developed based on a polypeptide and protein aggregation and not on short self-assembling peptides. To find the best performing tool for our self-assembling peptide library, we applied the experimental data on self-assembly by electron microscopy[4] (**Table S1**) as a dataset to evaluate the accuracy and reliability of each tool for self-assembly with the accuracy and receiver operating characteristic (ROC) value. The accuracy was calculated from the confusion matrix according to **Eqn.2**.

(**Eqn.2**) $$accuracy = \frac{\sum true\ positive + \sum true\ negative}{\sum total\ population}$$

The ROC value for the prediction (**Figure S2**) was calculated with 10-fold stratified cross-validation and the experimental fibril formation as target value and a logistic regression learner and LASSO regularization model (17 strength) with the data-mining software orange3.[11]

The prediction tools Aggrescan, APPNN and PATH performed best with an accuracy of 76%, 69% and 69%, respectively. Even though these aggregation prediction tools are trained on polypeptides and proteins, the reported accuracy for these tools match well to our self-assembling peptide library composed of short peptides. Noteworthy, combining Aggrescan, APPNN and PATH increase the performance of aggregation further (**Figure S2**).

Therefore, we applied Aggrescan, APPNN and PATH to predict aggregation propensity of the *de novo* predicted 3669 sequences. A sequence was considered aggregating, if at least two of Aggrescan, APPNN or PATH were positive. By applying this method 424/3669 peptides were predicted for aggregation by at least two of these tools.

As shown in **Figure S3** the aggregation tools performed with comparable accuracy for the selected 16 peptides as determined by Aggrescan 75% for the training set (**Figure S3**C) and 63% for the *de novo* predicted peptides (**Figure S3**D).

A

Aggregation prediction on training set

Tool	Accuracy	ROC
Aggrescan	76%	0.680
PATH	69%	0.696
APPNN	69%	0.619
PASTA 2.0	50%	0.512
Tango	43 %	0.437
Waltz	39%	0.487

B

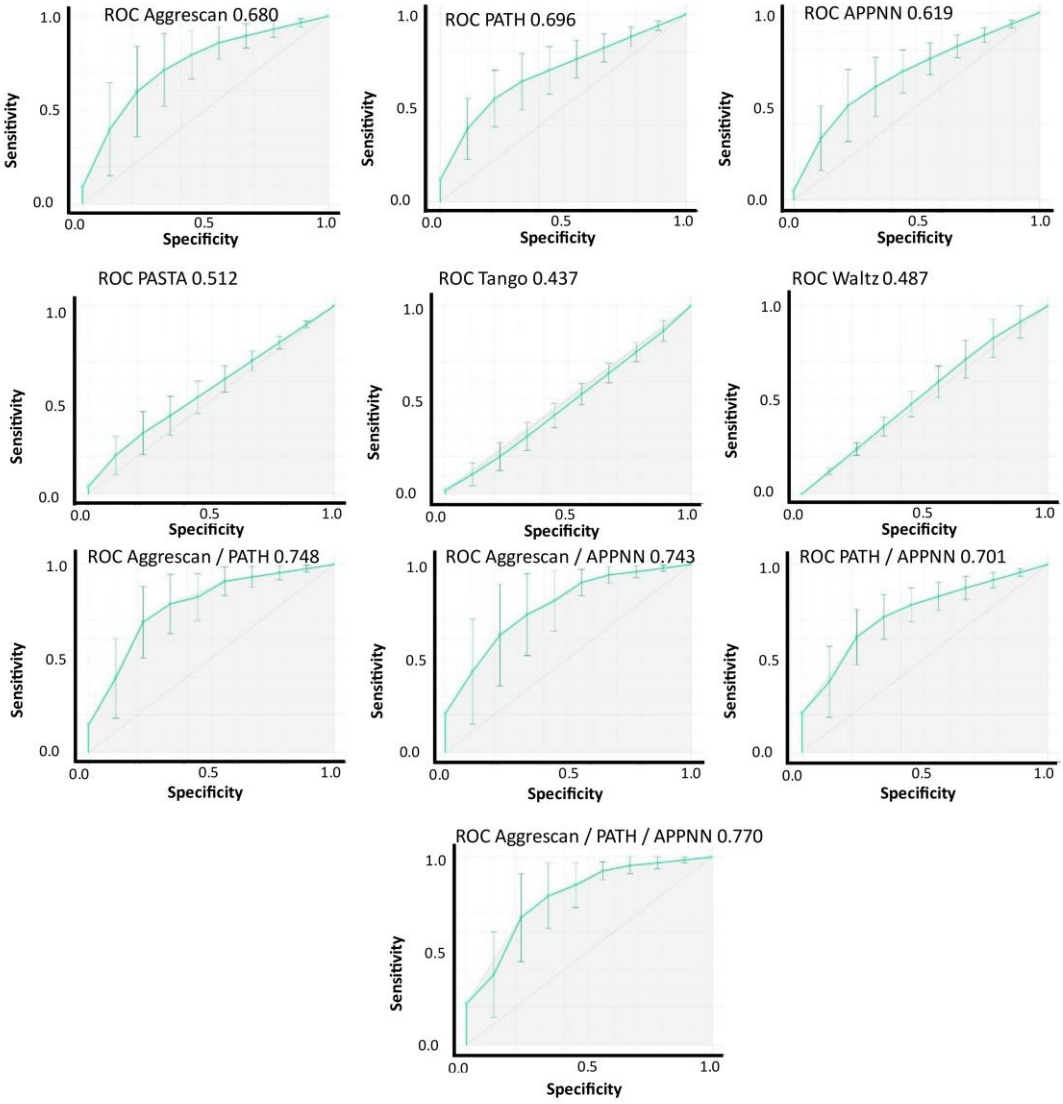
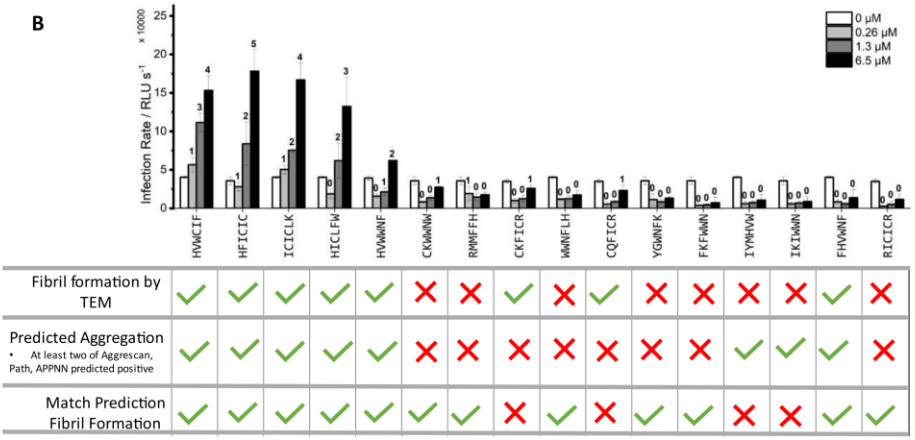


Figure S2 A Evaluation of the protein-aggregation tools Tango,[5] APPNN,[6] Waltz,[7] Path,[8] Aggrescan[9] and PASTA 2.0[10] with the training set (EF-C based library, **Table S1**). **B** ROC value for the prediction calculated with 10-fold stratified cross-validation and the experimental fibril formation as target value and a logistic regression learner and Lasso regularization model with 17 strength.

A

	Aggrescan	Path	APPNN	PASTA 2.0	Tango	Waltz	Predicted Infecticity Rel EF-C	Calculated hydrophobicity	Selection based on
HVWCIF	1	1	1	1	1	1	1.10	1.45	Highest predicted infectivity and aggregation without WWN motif
HFICIC	1	1	1	1	0	0	0.76	1.43	High N-gram similarity (0.42) with training set (HIHIQIC)
ICICLK	1	1	1	1	0	0	0.72	1.23	Cysteine rich sequence predicted for aggregation
HICLFW	1	1	1	1	0	1	0.69	1.53	Highest hydrophobicity and predicted for aggregation
HVWWNF	1	1	1	1	0	0	2.29	1.17	Highest predicted infectivity and aggregation
CKWWNW	1	0	0	0	0	0	0.86	1.12	High N-gram similarity (0.36) with training set (CKWKWQW)
RMMFFH	1	0	0	0	0	0	0.70	0.86	Low N-gram similarity (Avg 0.92, highest 0.66) with training set, not predicted for aggregation
CKFICR	1	0	0	0	0	0	0.82	0.78	High N-gram similarity (0.33) with training set (CKFQC)
WWNFLH	0	0	1	0	0	0	2.10	1.25	3rd highest infectivity prediction
CQFICR	1	0	0	0	0	0	0.96	0.91	High N-gram similarity (0.50) with training set (CQFQFQF)
YGWNFK	0	0	0	0	0	0	1.01	0.57	Low N-gram similarity (Avg 0.92, highest 0.72) with training set, not predicted for aggregation
FKFWWN	1	1	0	0	0	0	0.90	1.08	High N-gram similarity (0.56) with training set (KFKFQFNM)
IYMHVW	1	1	1	1	0	0	0.93	1.26	Low N-gram similarity (Avg 0.93, highest 0.78) with training set, predicted for aggregation
IKIWWN	1	1	1	1	0	0	0.90	1.09	High N-gram similarity (0.58) with training set (KIKIKIKIWWWW)
FHVWNF	1	1	1	1	0	0	1.09	1.10	3rd highest predicted infectivity and aggregation without WWN motif
RICICR	1	0	0	1	0	0	0.82	0.77	Moderate N-gram similarity (Avg 0.88, highest 0.66) with training set, not predicted for aggregation

B



C

Prediction accuracy **75%**
(at least two of Aggrescan, Path, APPNN predicted positive)

	Pred.	0	1	sum
actual	0	6	2	8
	1	2	6	8
sum		8	8	16

D

Prediction accuracy aggrescan only **63%**

	Pred.	0	1	sum
actual	0	2	6	8
	1	0	8	8
sum		2	14	16

Figure S3 Aggregation prediction tools applied on 16 de novo created peptides. **A** Summary of aggregation prediction results, predicted infectivity according to ProtVec *LASSO* model **Eqn. 1**, calculated hydrophobicity and comments on selection criteria. **B** Comparison of experimental and predicted aggregation. Experimental aggregation was determined by TEM fibril formation. 8 peptides were predicted for aggregation and 8 peptides were not predicted for aggregation by at least two of the tools Aggrescan, APPNN and PATH. **C** The accuracy of aggregation prediction by applying at least two prediction tools is determined 75 %. **D** Aggregation prediction accuracy for Aggrescan only calculated by confusion matrix for predicted peptides is 63 %.

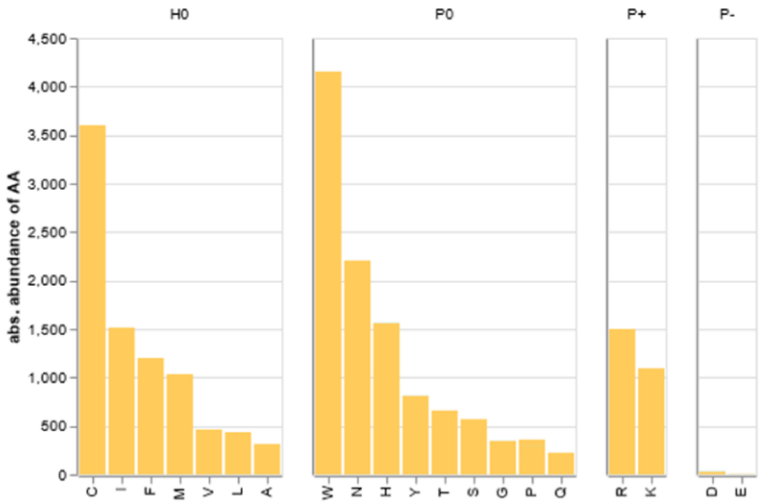


Figure S4 Absolute abundance of amino acids in net charge positive peptides (total 3669) predicted for infectivity enhancement. Cysteine (C) and Tryptophan (W) are the most prevalent amino acids.

3. N-gram similarity for predicted peptides with training set

The N-gram sequence similarity of the net charge positive peptides (total 3669) predicted for infectivity enhancement with the training set was calculated to ensure a diverse selection of peptides semantically close and far away from the training set. The N-gram similarity factor quantifies the similarity of two strings and returns 0 for the same sequence and 1 for sequences

Selected sequence	Average N-gram similarity with training set	Highest similarity	Corresponding sequence for highest similarity in the training set
HVWCIF	0.91	0.64	HIHIQIC
HFICIC	0.85	0.43	HIHIQIC
ICICLK	0.88	0.69	KIKIKIKI
HICLFW	0.88	0.64	HLHLPLL
HVWWNF	0.93	0.75	HGEHGE
CKWWNW	0.81	0.36	CKWKWQW
RMMFFH	0.93	0.67	MKFM
CKFICR	0.78	0.33	CKFQC
WWNFLH	0.94	0.80	NMWQKFKFQF
CQFICR	0.85	0.50	CKFQC
YGWNFK	0.93	0.73	KYKGAIIGNIK
FKFWWN	0.83	0.50	KFKFQFN
IYMHVW	0.93	0.79	KIKIQIW
IKIWWN	0.79	0.50	KIKIQIN
FHVWNF	0.92	0.75	KFKFQFNM
RICICR	0.88	0.67	KIKIQI

Figure S5 Overview of N-gram similarity values between the selected sequences and the training set. Average N-gram similarity describes the mean N-gram value between one selected sequence with every sequence of the training set. The highest similarity value shows the lowest value (highest similarity) for each selected sequence and the corresponding sequences from the training set. Values are colored gradually from red (0) – blue (1).

which have nothing in common. We applied the algorithm by Kondrak[12] for 2-grams with the python script shown in https://github.com/luozhouyang/python-string-similarity.git

In **Table S5** a matrix of all 3669 peptides N-gram similarity values with the training is listed. As shown in **Figure S5** the N-gram similarity values of the selected 16 peptides cover a wide range between 0.33 to 0.93 to quantify the diversity of selected sequences.

4. Evaluation of De Novo Peptide Activity with Property–Activity Model[4]

Three of the newly predicted peptides show unexpected activity despite negative Zeta-potential. To test whether these peptides follow a different mode of action a property activity model determined for the training set was applied on the *de novo* created peptides. The model established by multivariate analysis is shown in **Eqn. 3**.[4]

(**Eqn. 3**) Log Infect Rel Ef-C = -2.33462 + 0.02128*(Zeta-potential)+ 0.29879*(Log Count Rate) +0.27355*(fibril formation) + 0.26241 (Hydrophobicity) + 0.10744 (ThT-activity) + 0.00356 (β—sheet).

The peptides found from machine-learning are matching the model with a Pearson correlation coefficient of $R = 0.75$, which is comparable to the Pearson correlation coefficient of $R = 0.82$ found for the training set.[4] Noteworthy, the peptides which show at first glance unexpected behavior can be explained well with this model. According to this model the peptides HVWCIF, HFICIC, HICLFW are active due to their extraordinary high hydrophobicity and successive aggregation which outweighs the contributions of the moderately negative zeta-potential.

5. TEM micrographs

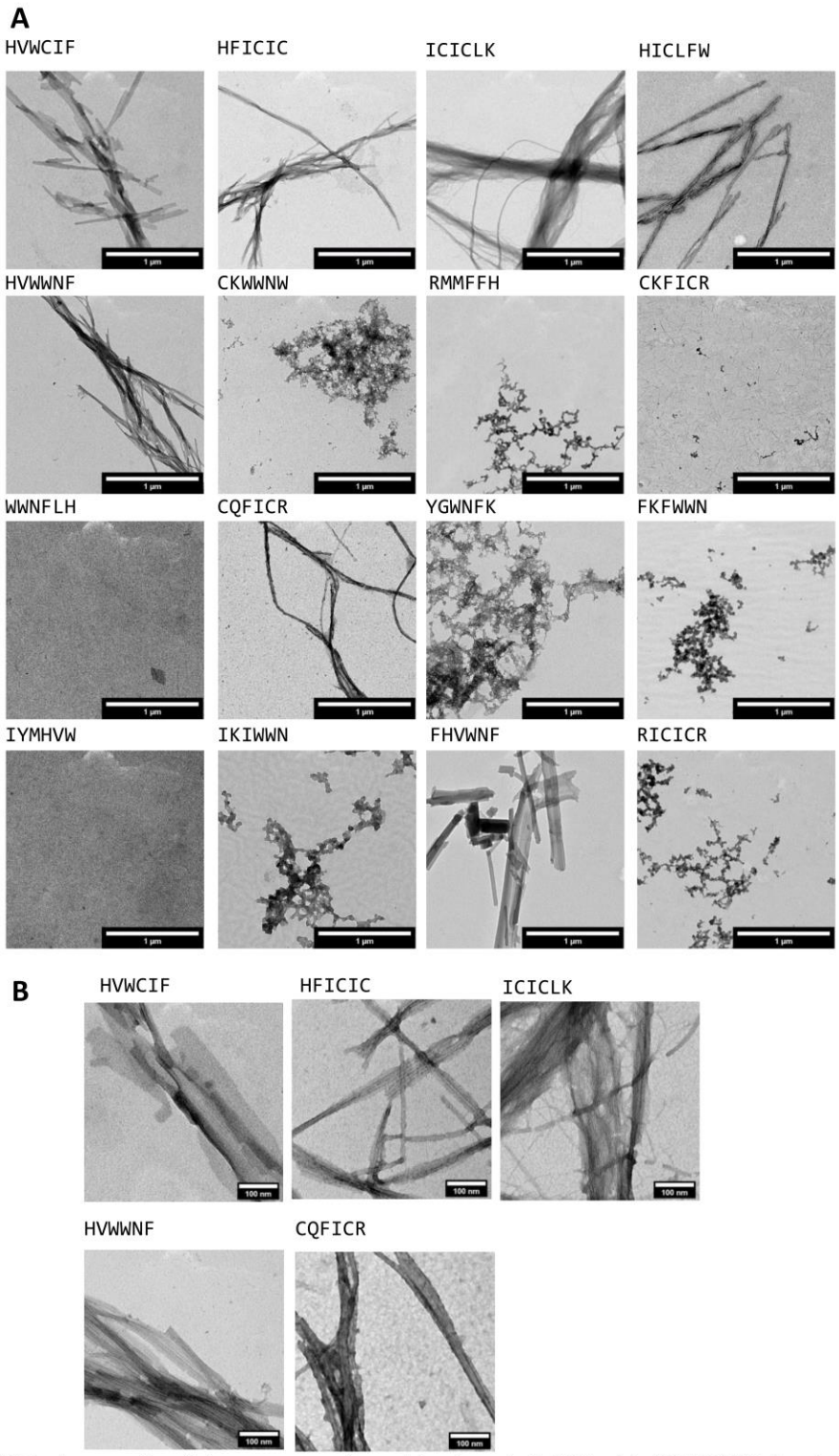


Figure S6 A TEM micrographs of *de novo* created peptides,1 mg/mL, PBS with 10% DMSO, incubated for 1d at RT, scale bar 1 µm. **B** Enlarged view on selected peptide fibrils of **A**, scale bar 100 nm.

6. Infectivity Data

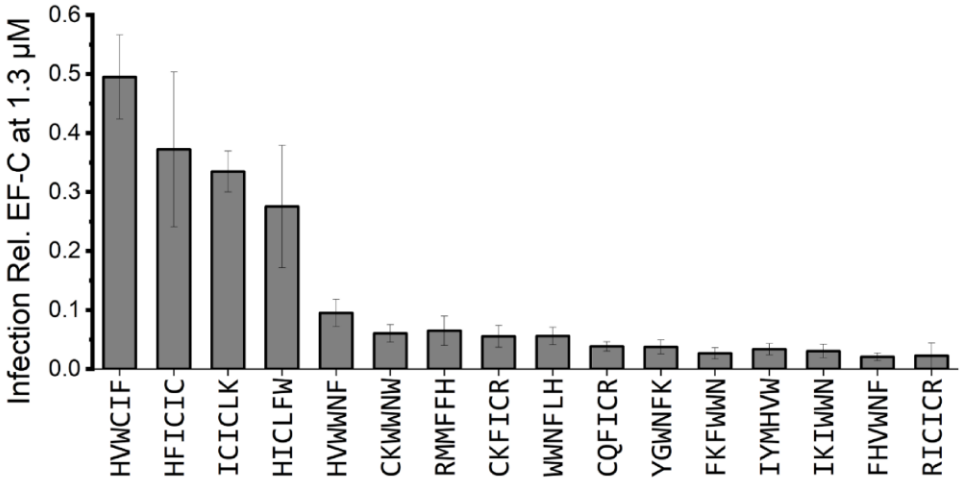


Figure S7 HIV-1 infection rates relative to EF-C (QCKIKQIINMWQ) at 1.3 µM concentration shown for the peptides from ML-prediction at 1.3 µM.

7. Cell-Viability

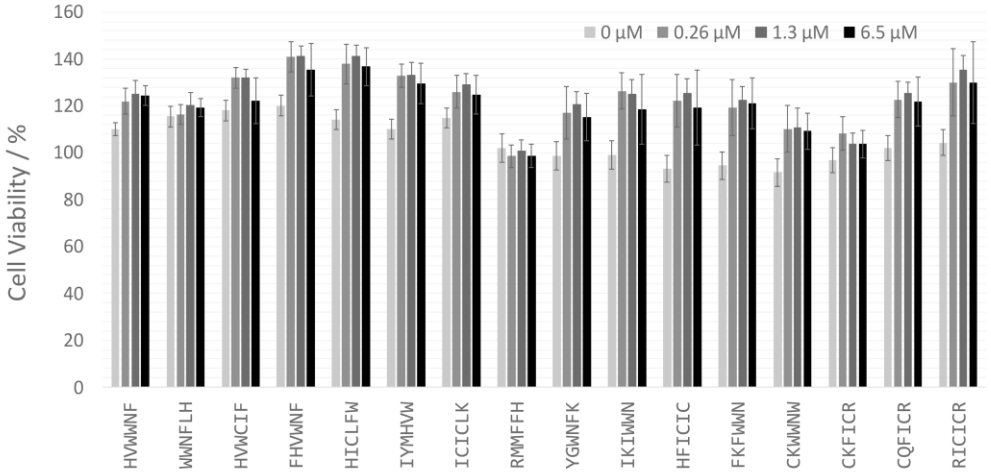


Figure S8 Cell viability normalized to 100% metabolic activity as determined by CellTiterGlo Assay of the 16 predicted peptides. Cell viability is maintained for all peptides at all tested concentrations (0.26 µM, 1.3 µM and 6.5 µM).

8. FT-IR Data

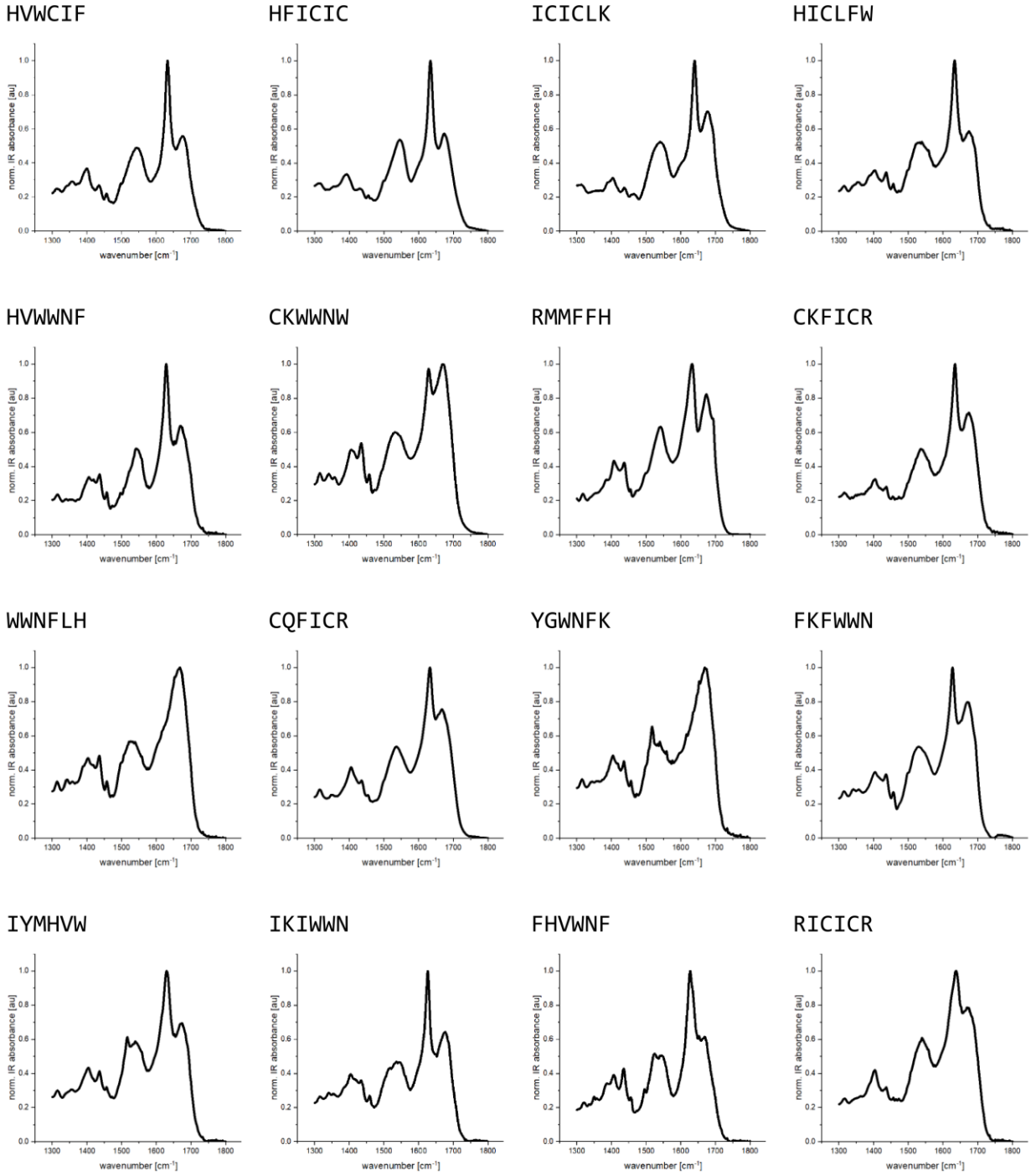


Figure S9 ATR FT-IR spectra of *de novo* created peptides.

9. Impact of Disulfide Bond Formation on Self-Assembly

Thiol groups from the side chain of cysteine can undergo disulfide bond formation with other thiol groups, which is known to influence self-assembly properties.[13]

To study the impact of disulfide bond formation of cysteine rich short peptides, we applied tris(2-carboxyethyl)phosphine (TCEP)[14] in 10 molar equivalents excess to break disulfide bonds, exemplarily studied for the peptide ICICLK. Transmission electron microscopy was performed to evaluate nanoscopic morphology, and brightfield microscopy was performed to evaluate microscopically large aggregation (Leica DMi8, 10x air objective). Surface charge and microscopic aggregation were evaluated via zeta-potential measurements.

Breaking disulfide bonds drastically changes the peptide assembly properties of ICICLK (**Figure S10**). Without disulfide bonds, no fibril formation (**Figure S10 A**) and no microscopic aggregation (**Figure S10 B, D**) can be observed, which also results in reduced surface charge (**Figure S10 C**).

Interestingly, for the original peptide EF-C in the training set, the addition of TCEP is has no visible influence on fibril formation and aggregation (**Figure S10 F**). This is likely due to the stabilizing effect of the alternating amphiphilic sequence pattern found in EF-C, which was identified earlier by us to drive assembly also without the presence of cysteine.[2,4]

Thus, we conclude that disulfide bond formation is a critical feature for self-assembly of the newly identified 6-mer peptides.

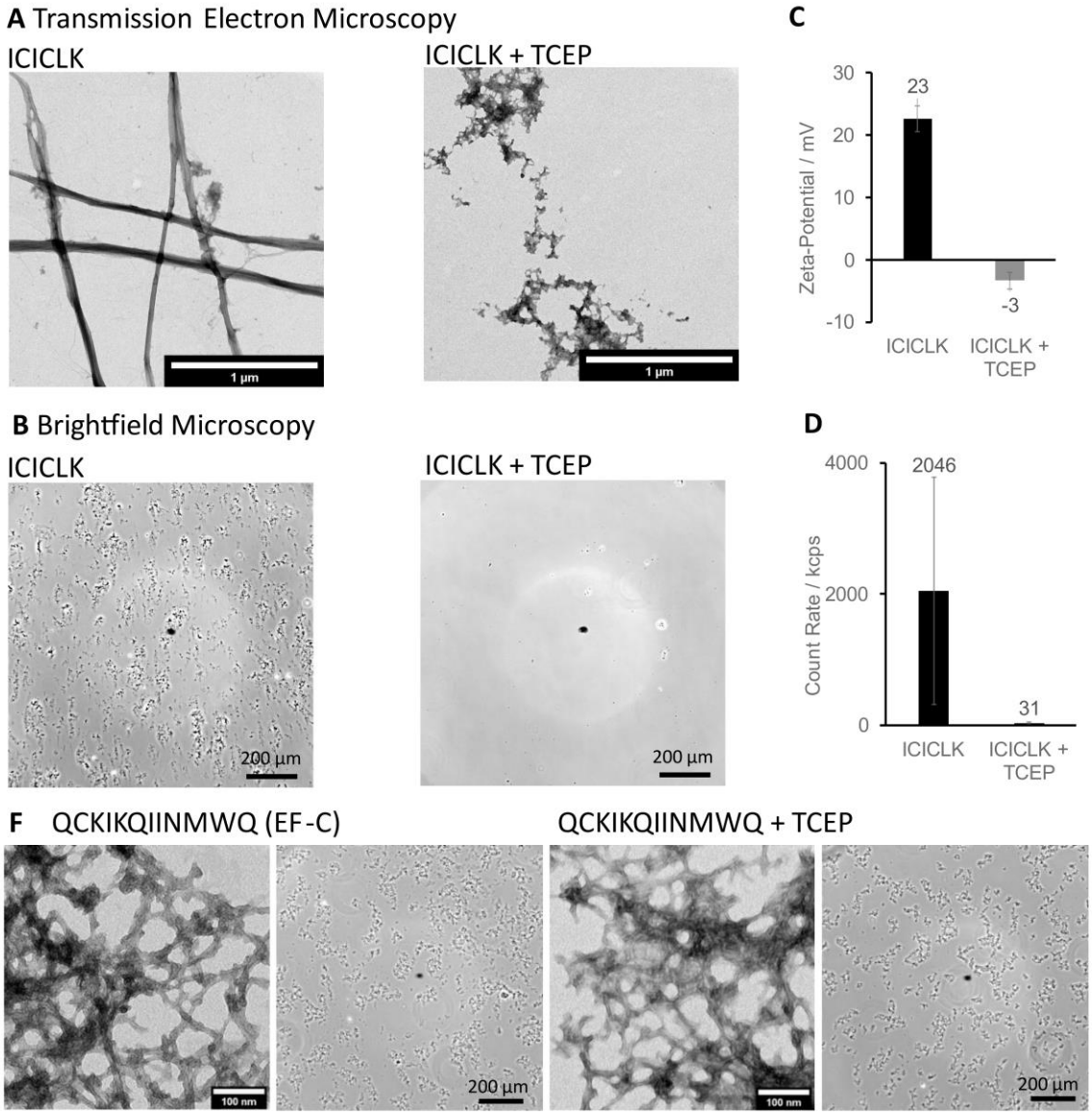


Figure S10 Effects of breaking disulfide bond on peptide self-assembly, aggregation, and surface charge of cysteine-rich peptide ICICLK. The peptide was incubated at room temperature for 1 day without and with 10 molar equivalents excess Tris(2-carboxyethyl) phosphine (TCEP). Then, the peptide was diluted from 10 mg/mL DMSO to 1 mg/mL in phosphate buffered saline, pH 7.4 (ICICLK) or in TCEP in phosphate buffered saline, pH 7.4. **A** Transmission electron microscopy micrographs depicting fibrillar (ICICLK) and non-fibrillar (ICICLK + TCEP) structures, scale bar 1 µm. **B** Brightfield microscopy images of peptide samples without (visible aggregates) and with TCEP (no visible aggregation), scale bar 200 µm. **C** Zeta-potential measurements and **D** derived count rate of scattered light of peptide samples without and with TCEP. **F** Transmission electron microscopy micrographs, scale bar 100 nm and brightfield microscopy images, scale bar 200 µm depicting fibrillar EF-C (QCKIKQIINMWQ) structures without (left) and with TCEP (right).

10. Amino acid Composition Analysis

To explore potentially common amino acid compositions between highly active peptides in the training set and the newly discovered active 6-mer peptides, we conducted a simplified, coarse-grained analysis. This analysis calculates the percentage of charged, hydrophobic, and hydrogen-bonding amino acids in peptides using the Hopp–Woods amino acid classification (**Figure S11 A, Code S1**).[15] Additionally, we coarse-grained the peptide activity into three thresholds: "high" (infectivity relative to EF-C > 0.7), "medium" (infectivity relative to EF-C > 0.1), and "low" (infectivity relative to EF-C < 0.1) active sequences.

The analysis of the training set revealed that peptides categorized as "high" and "medium" active contained a higher proportion of hydrophobic amino acids, while "low" activity peptides displayed a greater prevalence of charged amino acids (**Figure S11 B**). Remarkably, this same trend was observed in the *de novo* predicted peptides. The four active peptides, HVWCIF, HFICIC, ICICLK, and HICLFW, displayed a significantly higher content of hydrophobic amino acids compared to charged or hydrogen-bonding classified amino acids, which were predominantly found in non-active sequences (**Figure S11 C**).

It is worth noting that traditional prediction methods often rely on a predetermined set of descriptors. In contrast, the vector embedding approach employed in our study allows for the identification of underlying descriptors without any such assumptions. Therefore, a data-driven approach utilizing vector embeddings provides the flexibility to uncover latent descriptors that may have been not considered previously.

A Amino acid composition analysis

Coarse Grain

Amino Acids according to Hopp-Woods

Charged amino acids: E, D, R, K, H
Hydrogen bonding amino acids: S, T, E, D, K, R, H, N, Q, Y
Hydrophobic amino acids: W, V, I, L, F, M, C, A, G

Activity

High: Infection relative to EF-C > 0.7
Medium: Infection relative to EF-C > 0.1
Low: Infection relative to EF-C < 0.1

B Amino acid composition analysis for the **training set**

Activity category	Charged / %	Hydrogen bonding / %	Hydrophobic / %
High	24	47	53
Medium	25	48	52
Low	34	52	47

C Amino acid composition analysis for the **predicted peptides**

Sequence	Charged / %	Hydrogen bonding / %	Hydrophobic / %
HVWCIF	17	17	83
HFICIC	17	17	83
ICICLK	17	17	83
HICLFW	17	17	83
HVWWNF	17	33	67
CKWWNW	17	33	67
RMMFFH	33	33	67
CKFICR	33	33	67
WWNFLH	17	33	67
CQFICR	17	33	67
YGWNFK	17	50	50
FKFWWN	17	33	67
IYMHVW	17	33	67
IKIWWN	17	33	67
FHVWNF	17	33	67
RICICR	33	33	67

Figure S11 Comparison of the amino acid composition of the training set and the de novo predicted peptides. **A** To quantify the amino acid distribution, the amino acids and the activity were coarse-grained. The amino-acid compositions of a peptide were calculated by counting the number of each Hopp–Woods type amino acids (charged, hydrogen-bonding, or hydrophobic) in it and normalizing each count by the peptide's length (**Code S1**). **B** The peptides of the training set were categorized in high, medium, and low active sequences. High active peptides have on average a higher percentage of hydrophobic amino acids. **C** The active *de novo* predicted sequences (bold) have a higher percentage of hydrophobic amino acids compared to inactive sequences.

11. Training Set

Table S1 Binary representation of aggregation prediction results by protein-aggregation tools Tango,[5] APPNN,[6] Waltz,[7] PATH,[8] Aggrescan[9] and PASTA 2.0[10] for the training set. The experimental evaluation for fibril formation by TEM was reported previously by us.[4] Fibril formation is indicated by 1, no fibril formation is indicated by 0. The Log Infectivity enhancement relative to EF-C (QCKIKQIINMWQ) at 1.3 μM is retrieved from our previous report.[4]

sequence	Log Infect Rel. EF-C at 1.3 μM	Fibril formation by TEM	Tango	APPNN	Waltz	Path	Aggrescan	PASTA 2.0
CKIKIQIII	-0.01	1.0	0	1	1	1	1	1
CEIEIQI	-1.55	1.0	0	1	0	0	0	0
CKIKQIINM	-0.30	1.0	0	1	0	0	1	1
CKFKFQFNMWQ	0.06	1.0	0	1	0	0	1	0
CKFKFQFNM	0.09	1.0	0	0	0	0	1	0
CKFKFQF	0.09	1.0	0	0	0	0	1	0
KFKFQFNMW	-0.35	1.0	0	1	0	0	1	0
KFKFQFNM	-0.33	1.0	0	0	0	0	1	0
KFKFQFN	-1.74	1.0	0	0	0	0	1	0
CKAKAQANMWQ	-1.36	0.0	0	1	0	0	0	0
CKAKAQANM	-1.50	0.0	0	0	0	0	0	0
QCKFKQFFNMWQ	-0.61	1.0	1	1	0	0	1	0
QCKFKQFFNM	-0.37	1.0	0	0	0	0	1	0
QCKFKQFF	-1.91	1.0	0	0	0	0	0	0
CKFKQFFNMWQ	-0.21	1.0	1	1	0	0	1	0
CKFKQFFNM	-0.39	1.0	0	0	0	0	1	0
CKFKQFF	-2.17	0.0	0	0	0	0	0	0
QCKIKIQINM	-0.40	1.0	0	1	0	1	1	0
QCKAKAQANMWQ	-1.14	0.0	0	1	0	0	0	0
QCKAKAQANM	-1.83	0.0	0	0	0	0	0	0
QCKAKAQA	-1.60	0.0	0	0	0	0	0	0
QCKIKIQI	-0.95	1.0	0	0	0	0	1	0
QCKIKQIINM	-0.46	1.0	0	0	0	0	1	1
QCKIKQII	-2.19	1.0	0	1	0	0	0	0
QCKFKFQFNMWQ	-0.21	1.0	0	0	0	0	1	0
QCKFKFQFNM	-0.29	1.0	0	1	0	0	1	0
QCKFKFQF	-0.12	1.0	0	0	0	0	1	0
CRFRFQF	-0.46	1.0	0	0	0	0	0	0
HIHIQIC	-0.57	1.0	0	0	0	1	1	1
RLRLTLC	-1.29	1.0	0	1	0	1	1	0
HLHLPLL	-2.00	0.0	0	0	0	0	0	0
RGECKFKFQF	-0.48	1.0	0	0	0	0	1	0
RGEKIKIQINM	-0.61	1.0	0	1	0	1	1	0
KYKGAIIGNIK	-2.51	0.0	1	1	0	0	1	1
HGDKCHGDKC	-1.99	0.0	0	0	0	0	0	0
RPRGLLLGNLR	-1.32	0.0	0	1	0	0	1	0
KKFQKKFQ	-1.54	0.0	0	0	0	0	0	0
PPFHPPPFHP	-1.75	0.0	0	0	0	0	0	0
MDQMDQMDQMDQMDQ	-2.37	0.0	0	0	0	0	0	0
FDPFDPFDP	-1.63	0.0	0	0	0	0	0	0
TKTLTKTL	-1.67	0.0	0	0	0	0	0	0
FKFDKFKFDK	-1.33	0.0	0	0	0	0	0	0
KVKGVGK	-1.59	0.0	0	0	0	0	0	0
SISISRRI	-1.55	0.0	0	0	0	0	0	0
HRRHFRHKITKKK	-1.87	0.0	0	0	0	0	0	0
KNERIKNERI	-1.89	0.0	0	0	0	0	0	0
KIRGKFEKED	-1.24	0.0	0	0	0	0	0	0
CKFQC	-1.64	0.0	0	0	0	0	0	0
MKFM	-1.76	0.0	0	0	0	0	0	0
CKFC	-1.96	0.0	0	0	0	0	0	0
RGDKIRGDKI	-2.13	0.0	0	0	0	0	0	0

KNDKND	-1.98	0.0	0	0	0	0	0	0
HGEHGE	-2.19	0.0	0	0	0	0	0	0
HGEHGEHGE	-1.72	0.0	0	0	0	0	0	0
CRFRVPF	-1.70	0.0	0	0	0	0	0	0
CHLHLQL	-1.01	0.0	0	1	0	0	0	0
CETMYDKILKNLSRSR	-1.72	0.0	0	0	0	0	0	0
MKFKFQF	-1.20	1.0	0	0	0	0	1	0
QCKIKQIINMWQ	0.00	1.0	1	1	0	0	1	1
QCKIKIQINMWQ	0.02	1.0	0	1	0	1	1	1
KIKIQINMWQ	-0.53	1.0	0	1	0	1	1	1
NMWQKAKAQA	-1.38	1.0	0	0	0	0	0	0
NMWQKFKFQF	-0.57	1.0	0	0	0	0	1	0
KVKVKVQV	-1.76	0.0	0	0	0	0	1	0
KIKQIINMWQ	-0.46	1.0	1	1	0	0	1	1
KAKAQANMWQ	-2.15	0.0	0	1	0	0	0	0
KFKFQFNMWQ	-0.12	1.0	0	1	0	0	1	0
KFKQFFNMWQ	-1.05	0.0	1	1	0	0	1	0
KAKQAANMWQ	-1.51	0.0	0	1	0	0	0	0
NMWQKVGTPL	-1.61	0.0	0	0	0	0	0	0
NMWQKIKQII	-1.60	0.0	0	0	0	0	0	0
NMWQKIKIQI	-0.69	1.0	0	0	0	0	1	0
KIKQIINM	-0.35	1.0	0	1	0	0	1	1
KIKQIIN	-1.40	0.0	0	1	0	0	1	0
KIKIQINMW	0.33	1.0	0	1	0	1	1	1
KIKIQINM	-0.44	1.0	0	1	0	1	1	0
KIKIQIN	-0.58	1.0	0	1	0	1	1	0
KAKAQA	-1.74	0.0	0	0	0	0	0	0
KIKIQI	-1.17	0.0	0	0	0	0	1	0
KFKFQF	-1.21	0.0	0	0	0	0	1	0
KIKIQINMWA	-0.16	1.0	0	1	0	1	1	1
EIEIEIEI	-1.75	1.0	0	1	0	0	0	0
KIKIQINMAQ	-0.67	1.0	0	1	0	1	1	0
KIKIQINAWQ	-0.27	1.0	0	1	0	1	1	0
KIKIQIAMWQ	-0.41	1.0	1	1	0	1	1	0
KIKIKIKIYYYY	0.14	0.0	1	1	1	1	1	1
HHHHEIEIEIEI	-1.53	0.0	0	1	0	0	0	0
KYKYQY	-1.99	0.0	0	0	0	0	0	0
EFEFQF	-2.49	0.0	0	0	0	0	0	0
CEFEFQF	-2.28	1.0	0	0	0	0	0	0
CKFKFQFNMW	-0.02	1.0	0	1	0	0	1	0
CKYKYQY	-1.39	0.0	0	0	0	0	0	0
KWKWQW	-1.84	0.0	0	0	0	0	0	0
CKWKWQW	-0.60	1.0	0	0	0	0	0	0
KLKQLL	-2.13	0.0	0	0	0	0	0	0
CKIKIQINMWQ	-0.16	1.0	0	1	0	1	1	1
RGDKIKIQI	-1.81	1.0	0	0	0	0	1	0
CKIKIQINM	0.20	1.0	0	1	0	1	1	0
CKIKIQI	0.01	1.0	0	0	0	0	1	0
CKIKQII	-1.38	1.0	0	0	0	0	0	0
CKIKQIINMWQ	0.09	1.0	1	1	0	0	1	1
KIKIQIRGD	-1.68	1.0	0	0	0	1	1	0
RGDKIKIQIC	-0.08	1.0	0	1	0	1	1	0
RGDKIKIQINM	-0.38	1.0	0	1	0	1	1	0
RGDKIKIQINMWQ	-0.12	1.0	0	1	0	1	1	1
KFKFEFEF	-1.55	1.0	0	0	0	0	0	0
KIKQII	-2.31	0.0	0	0	0	0	0	0
KIEIQINM	-1.46	1.0	0	1	0	1	0	0
CKIKIQIRGD	-0.48	1.0	0	0	0	1	1	0
CKIKQIIRGD	-1.31	0.0	0	0	0	0	1	0
RGDKIKIQINMC	0.03	1.0	0	1	0	1	1	0
CQFQFQF	-0.45	1.0	0	1	0	0	0	0
KIKIQII	-1.73	1.0	0	1	0	1	1	1
KFKFQFFF	-0.39	1.0	0	0	1	1	1	0
KIKIKIQI	-1.01	1.0	0	0	0	1	1	0
KAKAKAQA	-2.06	0.0	0	0	0	0	0	0
KLKLKLQL	-1.01	1.0	0	0	0	0	0	0
KFKFKFQF	-0.10	1.0	0	0	0	0	1	0
KVKVQVV	-2.04	0.0	0	1	0	0	1	1
KLKLQLL	-2.16	1.0	0	0	0	1	0	0

KFKFQFF	-0.91	1.0	0	0	0	1	1	0
KIKIQIII	-0.45	1.0	0	1	1	1	1	1
KVKVQVVV	-1.46	1.0	0	1	0	0	1	1
HHHHKAKAKAKAYYYY	-1.88	0.0	0	0	1	0	1	0
KIKIQIC	-0.22	1.0	0	1	0	1	1	0
RGDSKIKIQIC	-0.22	1.0	0	1	0	1	1	0
RGDSGGGGGGKIKIQIC	-0.06	1.0	0	1	0	1	1	0
ILKNLSRSRKIKIQIC	-0.47	1.0	0	1	0	1	1	0
KIKIKIKIWWWW	-0.23	0.0	0	1	1	1	1	1
KIKIKIKI	-1.49	0.0	0	0	0	1	1	0
CEFEFQFNMWQ	-0.85	1.0	0	1	0	1	0	0
CSISIQI	-1.25	1.0	0	1	0	0	1	0
CEIEIQINMWQ	-0.66	1.0	0	1	0	1	0	0
CEIEIQINM	-1.19	1.0	0	1	0	1	0	0
CSISIQINM	-0.70	1.0	1	1	0	1	1	0
KAKAQANM	-2.05	0.0	0	0	0	0	0	0
CKAKAQA	-1.50	0.0	0	0	0	0	0	0
HHHHKIKIQINMYYYY	0.04	1.0	1	1	1	1	1	1
KIKIKIKIWW	-0.43	0.0	0	0	0	1	1	1
HHHHKIKIKIKIWWWW	-0.42	0.0	0	1	1	1	1	1
MKIKIQINM	-0.56	1.0	0	1	0	1	1	0
RGDCKIKIQINM	-0.63	1.0	0	1	0	1	1	0
MKIKIQINMWQ	-0.38	1.0	0	1	0	1	1	1
HHHHKIKIKIKIYYYY	-0.17	0.0	1	1	1	1	1	1
RGDCKFKFQF	-0.82	1.0	0	0	0	0	1	0
KIKIQIW	-0.85	1.0	0	1	0	1	1	1
KFKFQFW	-0.79	1.0	0	0	0	1	1	0
CKFKFQFW	-0.06	1.0	0	0	0	1	1	0
EIKIQINM	-1.00	1.0	0	1	0	1	0	0
IKVAVKIKIQINM	-0.30	1.0	0	1	0	1	1	1
HHHHKAKAKAKAWWWW	-1.05	0.0	0	0	1	0	1	0
CKIW	-1.29	0.0	0	0	0	0	0	0
CKIKIQINMW	0.11	1.0	0	1	0	1	1	1
ILKNLSRSRKIKIQINMWQ	-0.49	1.0	0	1	0	1	1	1
EIEIQINM	-1.44	1.0	0	1	0	1	0	0
KIKIQINMWWQ	0.51	0.0	1	1	0	1	1	1
KIKIQINMWWWQ	0.44	0.0	1	1	1	1	1	1
KFKQFFINMWQ	-0.04	1.0	1	1	1	1	1	1
KIKIQIMWNQ	-0.25	1.0	1	1	0	1	1	1
KIKIQIMQWN	-0.15	1.0	0	1	0	1	1	0
KIKIQINMWQRGD	-1.92	0.0	0	1	0	1	1	1
EIEIQINMWQ	-0.88	0.0	0	1	0	1	0	0
KIKIKIQINMWQ	-0.10	1.0	0	1	0	1	1	1

12. Predicted Peptides Characterization

Table S2 Physicochemical characterization of the selected *de novo* peptides. Shown are the absolute infection rates (Abs. Infect) at 6.5, 1.3, 0.26 and 0 µM peptide concentration. The logarithmic infection relative to EF-C (Log Infect Rel EF-C) refers to 1.3 µM concentration. Fibril formation is determined by transmission electron microscopy (**Figure S6**) and β-sheet content is determined by ATR-FT-IR spectroscopy (**Figure S9**). Standard deviations (Std Dev) are determined by triplicate measurements.

	Abs Infect [RLU/s] at 6.5 µM	Std Dev.	Abs Infect [RLU/s] at 1.3 µM	Std Dev.	Abs Infect [RLU/s] at 0.26 µM	Std Dev.	Abs Infect [RLU/s] at 0 µM	Std Dev.	Infect. Rel-to EFC 1.3 µM	Std Dev.	Log Infect Rel to EFC	Std Dev.	Zeta-Pot.	Std Dev.	Count Rate / kcps	Count Rate Std	Fibril	Hydro-phobicity	ThT active	β sheet [%]
HVWCIF	2E+05	18915	111453	12016	56677	9587	40407	1946	0.55	0.08	-0.26	0.06	-9.41	0.40	26175	7913	1	1.46	1	44
HFICIC	2E+05	27604	83800	28416	28200	5615	35653	5134	0.42	0.15	-0.38	0.15	-1.91	0.22	25360	7713	1	1.43	1	54
ICICLK	2E+05	21794	75363	2873	50423	5753	40407	1946	0.37	0.04	-0.43	0.04	33.43	1.20	3220	1783	1	1.23	0	38
HICLFW	1E+05	37786	61983	22561	18680	6151	40407	1946	0.31	0.12	-0.51	0.16	-9.02	0.67	10024	3381	1	1.54	1	44
HVWWNF	62177	790.5	21423	4777	15823	2689	39163	2823	0.11	0.03	-0.97	0.10	3.64	0.23	2219	758	1	1.17	1	37
CKWWNW	27287	1426	13610	3071	8167	1749	35653	5134	0.07	0.02	-1.17	0.10	-1.83	0.82	2453	2899	0	1.12	1	26
RMMFFH	17607	3346	14603	5405	19163	8998	35653	5134	0.07	0.03	-1.14	0.16	-1.82	0.40	21	9	0	0.86	0	33
CKFICR	25873	4155	12463	3991	9960	3079	35190	2493	0.06	0.02	-1.21	0.14	4.97	0.65	11	11	1	0.78	0	37
WWNFLH	17167	6199	12610	3089	11870	1369	40407	1946	0.06	0.02	-1.20	0.11	-6.01	0.70	18	3	0	1.25	0	2
CQFICR	23253	6640	8617	1592	5157	2902	35190	2493	0.04	0.01	-1.37	0.09	20.53	0.60	586	136	1	0.91	1	33
YGWNFK	13143	2312	8433	2530	11340	7237	35653	5134	0.04	0.01	-1.38	0.14	-3.06	2.50	56	25	0	0.57	0	6
FKFWWN	7317	6784	4607	1344	3757	983	35653	5134	0.02	0.01	-1.64	0.13	-4.89	0.26	244	159	0	1.08	0	32
IYMHVW	10627	6870	7553	2037	5907	2901	40407	1946	0.04	0.01	-1.43	0.12	5.16	1.42	21	5	0	1.27	0	40
IKIWWN	8810	6012	6837	2474	6010	2049	35653	5134	0.03	0.01	-1.47	0.16	-9.90	1.26	171	57	0	1.09	0	37
FHVWNF	13743	10451	6007	2070	8390	2290	40407	1946	0.03	0.01	-1.52	0.15	-16.53	0.68	1427	931	1	1.10	0	49
RICICR	11550	6543	5030	4903	2860	163	35190	2493	0.03	0.02	-1.60	0.42	4.52	0.28	41	18	0	0.78	0	47

Table S3 contains information on top 12320 sequences from Monte Carlo ProtVec LASSO model screening with information on predicted infectivity, hydrophobicity, and net charge and is openly available at the following data repository DOI: 10.5281/zenodo.7708290

Table S4 contains information on top 3669 peptides with a net positive charge with information on aggregation prediction results from Aggrescan, APPNN, and PATH and is openly available at the following data repository DOI: 10.5281/zenodo.7708290

Table S5 contains information on N-gram similarity matrix composed of top 3669 peptides and 163 peptides from the training set and is openly available at the following data repository DOI: 10.5281/zenodo.7708290

Code S1 is a python script for calculating the amino acids composition of charged, hydrogen bonding, and hydrophobic amino acids in a peptide sequence library according to Hopp–Woods classification.[15] The code and the corresponding training set of coarse-grained peptides are openly available at the following data repository DOI: 10.5281/zenodo.8004720

13. References

1 M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständker, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff and J. Münch, *Nat. Nanotechnol.*, 2013, **8**, 130–136.

2 S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, K. Kaygisiz, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch and T. Weil, *Adv. Funct. Mater.*, 2021, **31**, 2009382.

3 S. Kirti, K. Patel, S. Das, P. Shrimali, S. Samanta, R. Kumar, D. Chatterjee, D. Ghosh, A. Kumar, P. Tayalia and S. K. Maji, *ACS Biomater. Sci. Eng.*, 2019, **5**, 126–138.

4 K. Kaygisiz, L. Rauch-Wirth, A. Dutta, X. Yu, Y. Nagata, T. Bereau, J. Münch, C. V. Synatschke and T. Weil, *ChemRxiv*, 2023, https://doi.org/10.26434/chemrxiv-2023-hfqxb.

5 A.-M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz and L. Serrano, *Nat. Biotechnol.*, 2004, **22**, 1302–1306.

6 C. Família, S. R. Dennison, A. Quintas and D. A. Phoenix, *PLoS One*, 2015, **10**, e0134679.

7 J. Beerten, J. Van Durme, R. Gallardo, E. Capriotti, L. Serpell, F. Rousseau and J. Schymkowitz, *Bioinformatics*, 2015, **31**, 1698–1700.

8 J. W. Wojciechowski and M. Kotulska, *Sci. Rep.*, 2020, **10**, 7721.

9 O. Conchillo-Solé, N. S. de Groot, F. X. Avilés, J. Vendrell, X. Daura and S. Ventura, *BMC Bioinformatics*, 2007, **8**, 65.

10 I. Walsh, F. Seno, S. C. E. Tosatto and A. Trovato, *Nucleic Acids Res.*, 2014, **42**, W301–W307.

11 J. Demšar, A. Erjavec, T. Hočevar, M. Milutinovič, M. Možina, M. Toplak, L. Umek, J. Zbontar and B. Zupan, *J. Mach. Learn. Res.*, 2013, **14**, 2349–2353.

12 G. Kondrak, *SPIRE*, 2005, **3772**, 115–126.

13 D. Komáromy, M. C. A. Stuart, G. Monreal Santiago, M. Tezcan, V. V. Krasnikov and S. Otto, *J. Am. Chem. Soc.*, 2017, **139**, 6234–6241.

14 J. A. Burns, J. C. Butler, J. Moran and G. M. Whitesides, *J. Org. Chem.*, 1991, **56**, 2648–2650.

15 T. P. Hopp and K. R. Woods, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 3824–8.

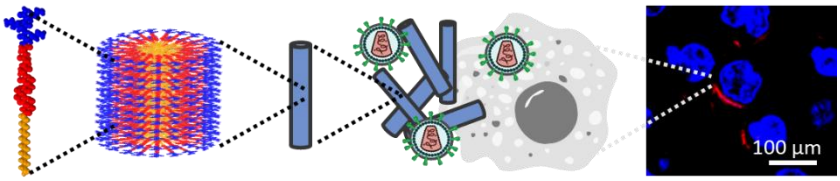
6.6 Peptide Amphiphiles as Biodegradable Adjuvants for Efficient Retroviral Gene Delivery

Authors: Kübra Kaygisiz, Lena Rauch-Wirth, Aysenur Iscen, Jan Hartenfels, Kurt Kremer, Jan Münch*, Christopher V. Synatschke*, Tanja Weil*
* corresponding author
Manuscript in revision.

Author Contributions:
Kübra Kaygisiz conceptualized the study, conceived the idea and designs for biodegradability, carried out the experiments related to peptide design, synthesis and characterization and *in vitro* peptide degradability, analyzed all data, performed the formal analysis, and wrote the original draft. Lena Rauch-Wirth designed, and carried out the experiments related to virus transduction and cell-viability, analyzed data related to virus transduction and cell-viability, performed the formal analysis, and revised the manuscript. Aysenur Iscen performed and analyzed data related to molecular dynamics simulations, and revised the manuscript. Jan Hartenfels analyzed data and supported peptide synthesis and characterization of P1-P3 and PA1-PA3. Kurt Kremer supervised, provided the funding resources, and revised the manuscript. Jan Münch supervised experiments related to virus transduction assays, provided the funding resources, and revised the manuscript. Christopher V. Synatschke initiated the project, supervised and revised the manuscript. Tanja Weil conceptualized, supervised, provided the funding resources, and revised the manuscript.

Copyrights: The authors hold copyright for the following part that is reproduced from an unpublished manuscript in revision.

Graphical abstract:



Peptide Amphiphiles: Biodegradable Viral Delivery Agents

We report peptide amphiphiles as enhancers of retroviral gene delivery through improved cell-virus interaction and biodegradability. The peptide amphiphiles exhibit comparable efficacy to existing transduction enhancers and hold promise for potential use in *ex vivo* gene therapy.

Peptide amphiphiles as biodegradable adjuvants for efficient retroviral gene delivery

Kübra Kaygisiz[1], Lena Rauch-Wirth[2], Aysenur Iscen[3], Jan Hartenfels[1], Kurt Kremer[3], Jan Münch[2]*, Christopher V. Synatschke[1]*, Tanja Weil[1]*

1 Department Synthesis of Macromolecules, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

2 Institute of Molecular Virology, Ulm University Medical Center, Meyerhofstraße 1, 89081 Ulm, Germany

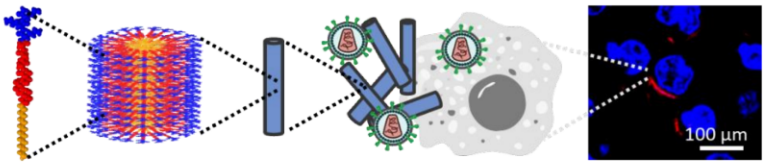
3 Polymer Theory Department, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

Correspondence to be addressed to: Jan.Muench@uni-ulm.de, synatschke@mpip-mainz.mpg.de; weil@mpip-mainz.mpg.de

Keywords: self-assembly, nanofibrils, peptide amphiphiles, retroviral transduction enhancer, gene delivery, biodegradability.

Abstract:

Retroviral gene delivery is the key technique for *in vitro* and *ex vivo* gene therapy. However, inefficient virion-cell attachment resulting in low gene transduction efficacy remains a major challenge in clinical applications. Adjuvants for *ex vivo* therapy settings need to increase transduction efficiency while being easily removed or degraded post-transduction to prevent the risk of venous embolism after infusing the transduced cells back to the bloodstream of patients, yet no such peptide system have been reported thus far. In this study, we introduce peptide amphiphiles (PAs) with a hydrophobic fatty acid and a hydrophilic peptide moiety that reveal enhanced viral transduction efficiency. The PAs form β-sheet-rich fibrils that assemble into positively charged aggregates, promoting virus adhesion to the cell membrane. The block-type amphiphilic sequence arrangement in the PAs ensures efficient cell-virus interaction and biodegradability. Good biodegradability is observed for fibrils forming small aggregates and we show via molecular dynamics simulations that the fibril-fibril interactions of PAs are governed by fibril surface hydrophobicity. Our findings establish PAs as additives in retroviral gene transfer, rivalling commercially available transduction enhancers in efficiency and degradability with promising translational options in clinical gene therapy applications.



Peptide Amphiphiles: Biodegradable Viral Delivery Agents

Graphical abstract We report peptide amphiphiles as enhancers of retroviral gene delivery through improved cell-virus interaction and biodegradability. The peptide amphiphiles exhibit comparable efficacy to existing transduction enhancers and hold promise for potential use in *ex vivo* gene therapy.

Introduction

Non-pathogenic viral particles offer promising perspectives as vectors for the delivery of genetic material into cells in the context of gene therapy[1,2] and vaccines.[3,4] Lentiviruses and γ-retroviruses are the vectors of choice in fundamental research as well as most clinical trials that are currently underway.[5–7] These vectors are superior compared to other viral vectors as they do not replicate and induce only very low inflammatory responses, which is essential for safe, stable and long-lasting gene expression.[8–10] For instance, chimeric antigen receptor (CAR) T cell therapy is one of the most promising applications of gene therapy approved for patient use.[11] For the generation of CAR T cells, the T cells of the patient are isolated, activated and genetically modified *ex vivo* to express the CAR. After infusion of CAR T cells back into the patient, the CAR T cells recognize specific tumor surface antigens by their CAR receptor, which mediates the elimination of tumor cells.[12] An indispensable but challenging step in the engineering process of CAR T cells is the efficient introduction of the CAR gene via retroviral transduction.[13] However, low concentrations of the viral vectors must be used to avoid side effects such as cytotoxicity and immunogenic reactions. Therefore, efficient virion-cell attachment is essential for gene transduction, which remains a major challenge in retroviral gene delivery.[14–17] In *ex vivo* clinical applications, it is of utmost importance to prevent exposure of possibly harmful substances to the patient.[18] Therefore, any aggregated adjuvants should be removed or degraded after successful transduction.

To overcome these challenges, transduction-enhancing additives, including polymers, lipids, peptides and others, have been reported.[19] Self-assembling peptides which form nanofibrils with a positive net charge and a high β-sheet secondary structure, so called β-amyloids, have been identified previously as efficient transduction enhancers.[20,21] These peptides are characterized by the presence of amphiphilic patterns composed of positively charged and lipophilic amino acids. In aqueous media, these peptides form very stable peptide aggregates. Several studies have confirmed that their positive net charge and microscopic aggregation is important for binding to negatively charged viral particles.[20–23] Subsequent attachment of the amyloid-virus complexes to the cellular membrane leads to increased transduction rates.[23] However, due to their intrinsic high stability, the *in vitro* biodegradation and removal of these often micrometer-sized amyloid-like fibrils is challenging. Therefore, introducing these fibrils along with the transduced cells into a patient's bloodstream may lead to undesired side effects, such as venous embolism. This poses significant challenges for their use in *ex vivo* therapy settings.[24] Thus, successful peptide-based adjuvants for clinical applications have to increase transduction efficiency and be easily removed or degraded after successful transduction. To date, no self-assembling peptide has been reported that show these features.

Peptide amphiphiles (PAs) constitute a class of chimeric molecules that consist of a hydrophobic tail and a charged, hydrophilic peptide domain.[25,26] While different types of hydrophobic tails have been explored, fatty acid residues are most prevalent. The hydrophobic domain triggers self-assembly and structure formation in water, which is then stabilized by the charged amino acid domains. In addition, certain amino acid sequences can give rise to ordered β-sheet structures, which results in fibrillar nanostructures with high aspect ratios. Pioneering works by Stupp and coworkers developed and applied this material class for various medicinal applications.[27–32] In recent decades, the variety of peptide sequences employed in PAs has been greatly expanded,[33] as well as their applications ranging from neural tissue regeneration[34] to drug delivery,[35,36] antimicrobial systems,[37] peptide vaccines,[38,39] immune-therapeutic applications[40,41] and non-viral gene delivery.[42–44] For instance, by attaching a fatty acid residue to antibacterial[37,45] and antiviral peptides,[46–51] their potency was greatly improved. Recent works have shown that PAs are internalized via lipid-raft mediated endocytosis by a variety of cell lines without affecting cell viability.[52–54] Interestingly, the morphology of PA-based

nanostructures can be destroyed via enzymatic cleavage of the fatty acid from the peptide, which renders them inherently biodegradable.[55–57] These unique features make PAs very attractive for *ex vivo* gene delivery applications. However, to the best of our knowledge, PAs have not been applied as an additive in the context of retroviral transduction.

Herein, for the first time, we propose PAs as highly efficient and biodegradable retroviral transduction enhancers. In our PA design, we disconnect the aggregate forming and membrane-binding features of amyloid peptides. The PA peptide sequence comprises positively charged amino acids that support the interactions with the viral and cellular membrane required for high transduction efficiency of the retroviral particles, whereas self-assembly is mainly driven by the palmitic acid residues, which should facilitate the formation of smaller aggregates with high *in vitro* degradability. Peptides prone to high aggregation are compared with peptides that reveal lower tendency for aggregate formation in view of their *in vitro* degradability. We report on PA1 (C_{16}-VVVAAAKKK-NH_2) and new peptide amphiphiles such as eicosapentaenoic acid-VVVAAAKKK-NH_2, C_{16}-*thioester*-CVVVAAAKKK-NH_2, C_{16}-*ester*-SVVVAAAKKK-NH_2, C_{16}-ALAAGKK-NH_2, and C_{16}-AKAVGK-NH_2 which demonstrate a combination of *in vitro* degradability and high transduction efficacy. In fact, their effectiveness as transduction enhancers is comparable to the gold standard transduction enhancers RetroNectin, Vectofusin-1 or Protransduzin (enhancing factor C, EF-C).[23,58–60]

Results and Discussion:

Peptide Fibril Design

Cross β-sheet rich amyloid forming peptide transduction enhancers based on the enhancing factor C (EF-C) are usually characterized by alternating amphiphilic patterns composed of hydrophilic (lysine)

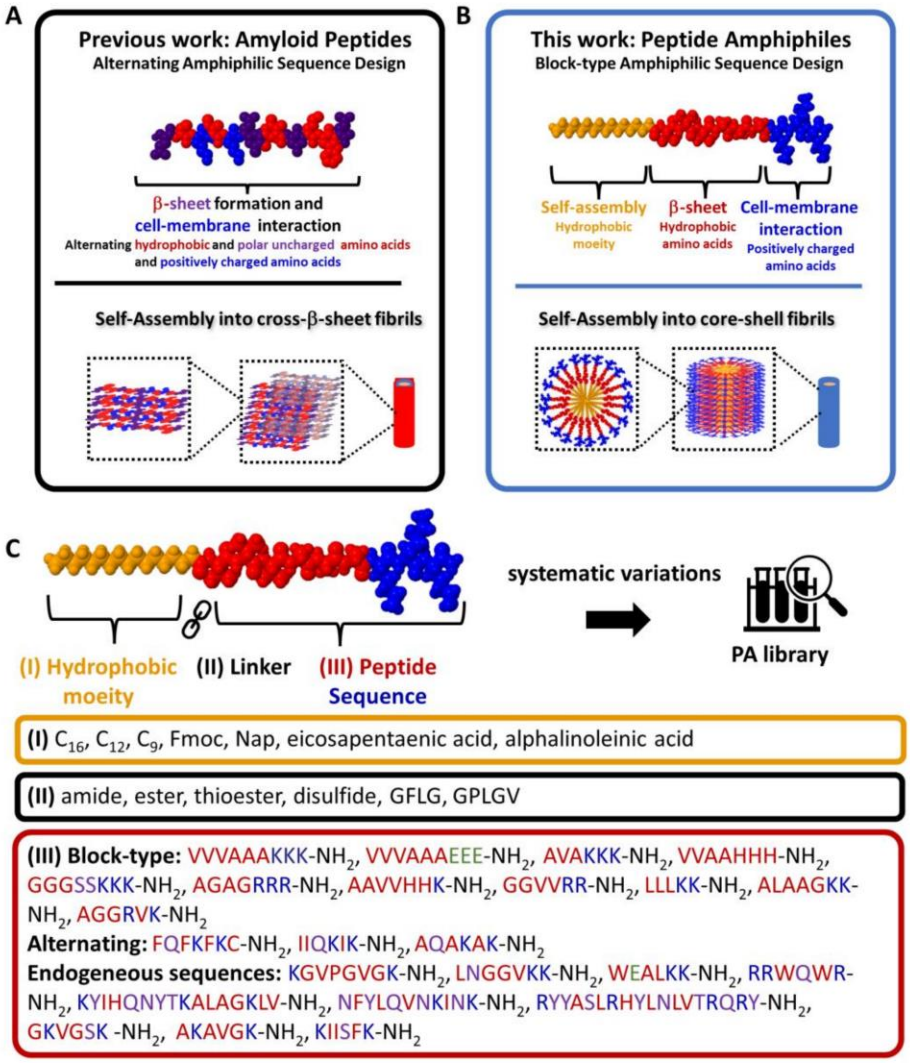


Figure 1 A Schematic illustration showing the sequence design of amyloid peptides (EF-C, QCKIKQIINMWQ) previously reported for infectivity enhancement[23] and **B** peptide amphiphiles (PA1, C_{16}-VVVAAAKKK-NH_2) reported in this study. Amyloid peptides have an alternating amphiphilic sequence that leads to self-assembly into rigid, stable cross-β sheet-rich fibrils. In contrast, PAs self-assemble based on hydrophobic fatty acids, resulting in core-shell fibril structures. The β-sheet secondary order is determined by non-polar amino acids (red), while the cationic, polar residues (blue) are oriented outwards, making them interactive with virus particles and cell membranes. **C** Peptide amphiphile library created by systematic variations of the original sequence PA1 to study the effects of the (I) hydrophobic moiety, (II) the linkage chemistry and (III) the peptide sequence on infectivity enhancement and degradability. The peptide sequences were designed either according to a block-type amphiphilic scheme, alternating amphiphilic, or derived from endogenous peptides not following any design scheme. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

and aliphatic (isoleucine or valine) or aromatic (phenylalanine, tyrosine, tryptophan) amino acids (**Figure 1A**).[21] In order to design PA based nanofibrils that reveal (I) high viral transduction efficiency and (II) high *in vitro* degradability, the amphiphilic patterns were changed to a block-type arrangement as depicted in **Figure 1B**.

The sequence design for block-type PAs is derived from the non-self-assembling peptides VVVAAAKKK-NH_2 (P1) and VVVAAAEEE-NH_2 (P3) which contain β-sheet-prone hydrophobic amino acids sequence "VVVAAA" as well as positively "KKK" and negatively "EEE" charged residues at physiologic pH, respectively. Combining fatty acid chains with these peptides afford the conjugates C_{16}-VVVAAAKKK-NH_2 (PA1) and C_{16}-VVVAAAEEE-NH_2 (PA3). These PAs form nanofibers that have been extensively studied by Stupp and coworkers for a variety of applications.[32,61,62] For comparison, we designed the self-assembling peptide FQFKFKC-NH_2 (P2) with an alternating amphiphilic pattern based on transduction enhancers previously reported by us.[20] Both FQFKFKC-NH_2 and C_{16}-FQFKFKC-NH_2 (PA2) readily form nanofibrils (**Figure 2A**).

To cover a wide variety of different PA designs and study its influence on aggregate formation and degradation, we designed a library inspired by PA1 and PA2 consisting of 37 distinct peptide amphiphiles by varying the (I) hydrophobic moiety, (II) the linkage chemistry and (III) the peptide sequence (**Figure 1C**, **SI chapter 7**). Peptide synthesis was conducted via standard solid-phase chemistry using the Fmoc protection strategy (**see methods**). PAs were accomplished by linking a palmitic acid (C_{16}) to the *N*-terminus according to a protocol described previously.[62,63] The purity of the conjugates (> 90%) was determined by LC-MS or HPLC and the expected *m/z* values were determined by MALDI-TOF for each peptide (**Figure S1**). After confirming the successful synthesis of the peptides, they were characterized by various methods to assess aggregation and fibril formation as discussed below (zeta-potential, Thioflavin T (ThT) fluorescence, attenuated total reflection infrared (ATR-IR) spectroscopy, and transmission electron microscopy (TEM)) (**Figure 2**, **SI chapter 7**, **see methods**).

Coupling Peptides to Palmitic Acid Enables Fibrillar Assembly

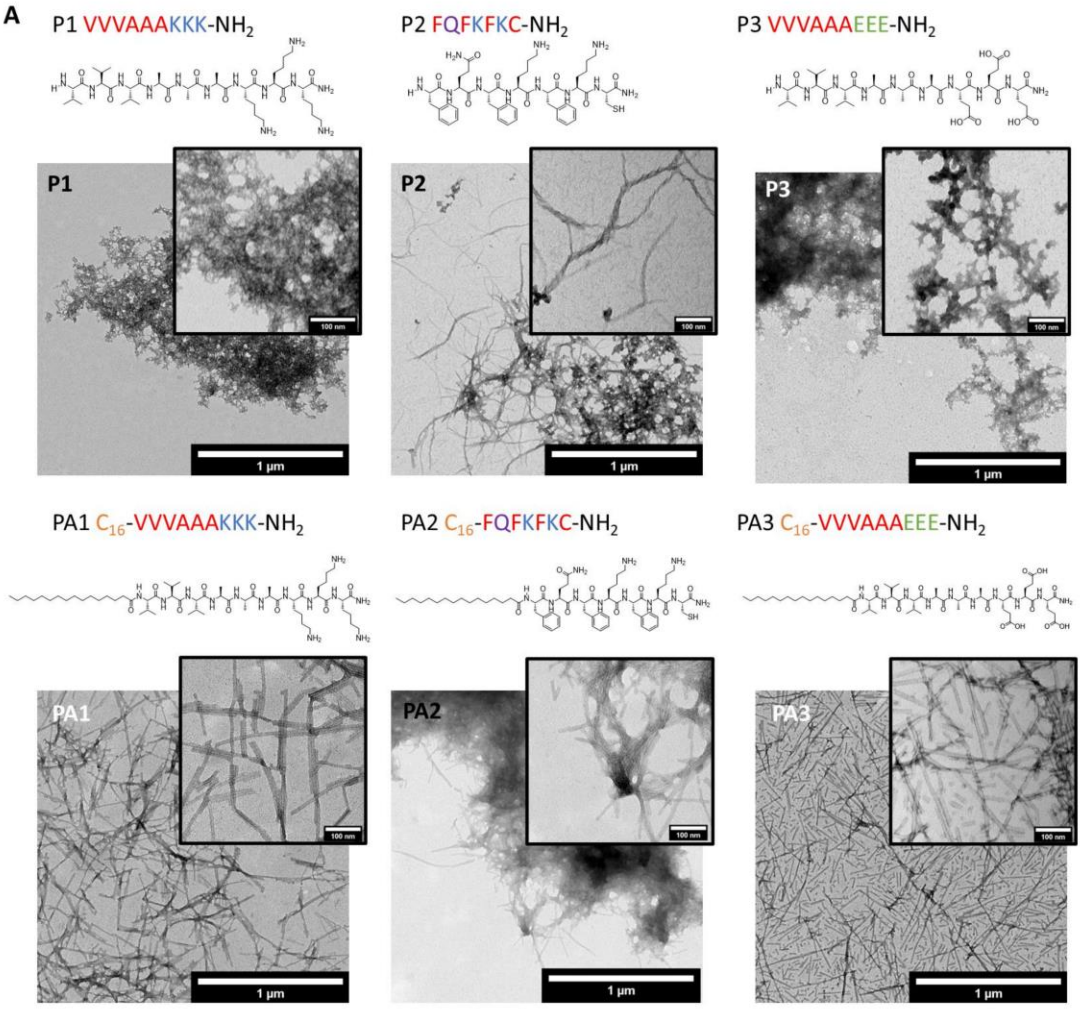
In the next sections, we will focus on the representatively selected examples PA1-PA3 and the parent peptides P1-P3, to illustrate their assembly, activity, and biodegradability, as we present only a portion of the PAs library for the sake of clarity and brevity. Detailed results and further discussion of the entire library can be found in **SI chapter 7**.

The block-type amphiphilic peptides P1 and P3 without the palmitic tail formed non-fibrillar amorphous aggregates, whereas the alternating amphiphilic sequence P2 assembled into twisted fibrils upon dilution from dimethyl sulfoxide (DMSO) stock solution in phosphate-buffered saline (PBS), according to TEM images (**Figure 2A**). The fatty acid conjugates PA1 and PA2 assembled into distinct fibrils in water (**Figure S2C-D**), whereas in PBS, micrometer-sized clusters of fibrils were formed (**Figure S3B-C**). Multivalent ions such as phosphate ions are known to interact with basic amino acid residues like lysine causing aggregation, probably through salt bridge formation between negative charged phosphate ions and positively charge lysine residues.[64,65] As expected, substituting lysine residues with negatively charged glutamic acid residues in PA3 results in distinct fibrils (**Figure 2A**), which did not bundle or adopt µm-sized aggregates in PBS (**Figure S3D**).

Beside morphologic properties, we have previously identified positive zeta-potential as an important factor for highly active transduction enhancers.[20] The zeta-potential characterizes the electrokinetic potential at the slipping plane of colloidal particles and can be measured for self-assembling peptides if they form fibrils or amorphous aggregates. The peptide sequences of PA1 and PA2 contain lysine residues that are positively charged at pH 7.4. PA3 however, provides negatively charged glutamic

acids. As expected, PA1 and PA2 displayed a positive zeta-potential while PA3 showed a negative zeta-potential (**Figure 2B**). Compared to PA1–PA3, the parent peptides P1–P3 displayed a less positive zeta-potential despite having the same peptide sequence. This observation underlines that the zeta-potential is not only dependent on the charged groups within the peptide sequence, but may also be influenced by the size and morphology of the assembled structure, which was also previously shown for other PAs.[66,67]

To gain information on the secondary structure of the assembled peptides, ATR-IR measurements were conducted with a focus on the amide I (1600–1700 cm^{-1}) and amide II (1500–1600 cm^{-1}) regions.[68,69] Interestingly, β-sheet structural elements were identified for all peptides P1–P3 and PA1–PA3 with a main peak around 1630 cm^{-1}, which is assigned to parallel β-sheets (see **SI chapter 3**). A quantitative analysis of the amide I region revealed that PA1–PA3 show a 9–15 % higher β-sheet content compared to their respective parent peptides P1–P3 (**Figure 2B**, **Figure S4**). The secondary structure of the PAs is further confirmed in solution via circular dichroism (CD) spectroscopy (**Figure 2C**) and ThT fluorescence (**Figure 2B**, **Figure S5**). Interestingly, P1–P3 show CD spectra with characteristics of a random coil structure and no β-sheet structure, which can be traced back to the preparation in pure water lacking counter-ions, which are likely facilitating intermolecular interaction via charge screening of charged residues “K” and “E” (**SI chapter 2**). By introducing palmitic acid in PA1–PA3, a predominant β-sheet structure is observed (**Figure 2C**). The CD spectra of PA1 and PA3 are in good agreement with literature reports.[62,70] The fluorescence of ThT upon excitation at 488 nm can only be detected if the molecular rotation of the dye is hindered, that is upon binding to highly ordered structures such as amyloid or β-sheet containing fibrils. As expected, ThT fluorescence could be observed for β-sheet fibril-forming PA1–PA3 but not for P1–P3 (**Figure 2C** and **Figure S5**). Therefore, only PA1 and PA3 form fibrils with high β-sheet content in PBS, which was not observed for the parent peptides P1 and P3 without the palmitic acid moiety. In our PA library we found that 21 of 23 tested PAs with newly designed peptide sequences assembled into β-sheet rich fibrils and conclude that attaching palmitic acid to the *N*-terminus of non-assembling peptides is a reliable way to create self-assembling PAs (**Figure S19, Figure S27**). Thus, the introduction of the fatty acid residue drives the assembly of the peptides into ordered β -sheet rich nanostructures as reported,[71] which is required for boosting retroviral gene transduction.



B

Code	Sequence	Zeta-Potential [mV]	ThT	β-sheet [%]
PA1	C_{16}–VVVAAAKKK–NH_2	+22.2 ± 0.7	+	57
PA2	C_{16}–FQFKFKC–NH_2	+19.2 ± 3.4	+	59
PA3	C_{16}–VVVAAAEEE–NH_2	−29.8 ± 2.7	+	61
P1	VVVAAAKKK–NH_2	−2.3 ± 0.4	–	48
P2	FQFKFKC–NH_2	+10.1 ± 3.0	–	46
P3	VVVAAAEEE–NH_2	−21.2 ± 1.5	–	46

C

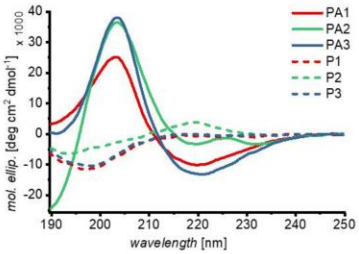


Figure 2 Characterization of PA1–PA3 and P1–P3. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids. **A** Representative TEM micrographs of P1–P3 and PA1–PA3 reveal the morphologies of peptide assemblies (1 mg mL^{-1} in PBS). Formation of fibrils are observed for PA1–PA3 and P2, whereas P1 and P3 assemble in amorphous aggregates (scale bars 1 µm, inset: 100 nm). **B** Tabular summary of physicochemical features of PA1–PA3 and P1–P3 (1mg mL^{-1}, PBS) including zeta-potential, ThT fluorescence (+ = ThT active, − = no ThT fluorescence observed). The β-sheet contents were determined via a quantitative analysis by peak integration of parallel β-sheet at 1630 cm^{-1} in the amide I region of the 2nd derivative of ATR-IR spectra (**Figure S4**). **C** CD spectra display β-sheet secondary order for PA1–PA3 and random coil conformation for P1 and P3 (all peptides 0.1 mg mL^{-1} in pure water).

Sequence Hydrophilicity and Arrangement are Important for Aggregation

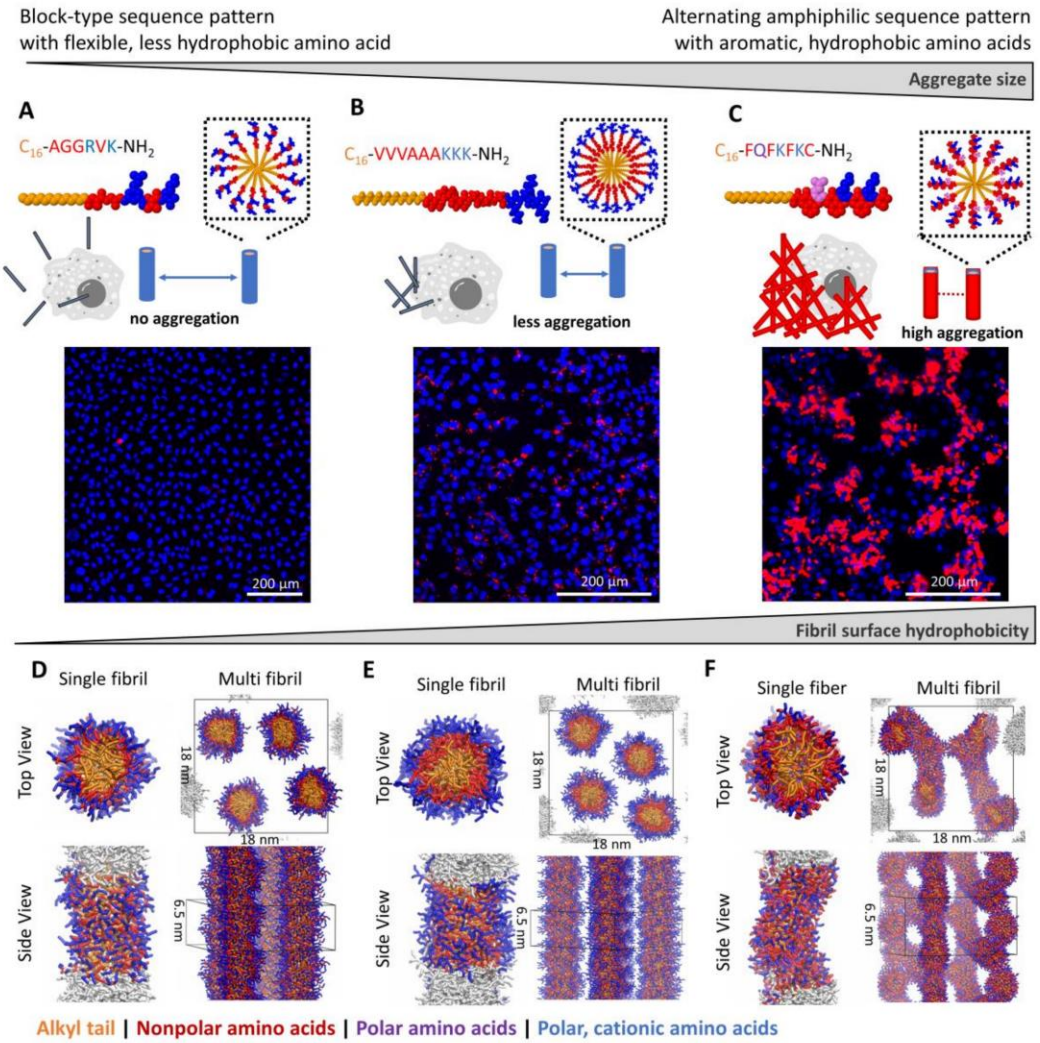


Figure 3 Schematic representation of three selected PAs **A** C_{16}-AGGRVK-NH_2, **B** PA1, and **C** PA2, that show distinct aggregation properties. Laser scanning microscopy images of Proteostat stained PAs (red) and HeLa cells (nucleus in blue) reveal formation of µm-sized PA clusters for PA1 and PA2. The PAs (20 µg mL^{-1}) were added to cells for 30 min, washed and samples were analyzed. PA1 and PA2 revealed association with cellular membranes while C_{16}-AGGRVK-NH_2 did not associate with cells, scale bar 200 µm. **D – F** Coarse grained molecular dynamics simulations. Snapshots show cross-sectional (top) and side views of selected PAs self-assembled into single fibers after 15 µs from randomly dispersed PAs in solution in a simulation box with a size 13 nm × 13 nm along the fibril cross-section (top view) and 6.5 nm along the lateral direction (side view). The snapshots for the interaction of multiple fibers are shown after self-assembly following an additional 10 µs simulation time in a box with a size 18 nm × 18 nm along the fibril cross-section (top view) and 6.5 nm along the lateral direction (side view). Color codes represent alkyl chains in yellow, nonpolar amino acids in red, polar amino acids in purple and polar, cationic amino acids in blue. Water and ions in the simulation box are omitted in the snapshots for clarity. For the single fibril system, the molecules in the simulation box are shown in color, while the periodic image is shown in white. The simulation box is shown in black for the multi fibril system. **D** Simulation of C_{16}-AGGRVK-NH_2 and **E** PA1 reveal stable fibril formation with high-density of polar groups (blue) on the fibril surface that prevents fibril-fibril interaction. **F** In contrast, PA2 displays hydrophobic amino acids (red) on the fibrils surface that is likely the main driving force for multi-fibril networks formation.

In general, amyloid fibrils with high β-sheet content are usually degraded slowly by cells.[72] However, PAs can be internalized by cells[73,74] which may be facilitated by interactions of the fatty acids with cellular membranes.[52,75] As assemblies of PAs are likely degraded by cleavage of the fatty acid residue in cells,[52,53,55,76] we hypothesized that the difference in biodegradability can be traced back to (I) the size of the initial aggregates and (II) the different self-assembling properties and biodegradability of the parent peptides.

Our observations reveal three different aggregation types in the PA library: fibers that do not aggregate or attach to cells (*e.g.* C_{16}-AGGRVK-NH_2, **Figure 3A**), fibers that form small µm-sized aggregates that attach to cells (*e.g.* PA1 **Figure 3B**), and fibers that form large µm-sized aggregates that attach to cells (*e.g.* PA2 **Figure 3C**).

The aggregation of PAs is a complex process that occurs at various length scales, from molecular self-assembly dynamics to fiber-fiber interactions and micrometer-sized colloid formation. This makes it challenging to identify specific molecular design rules for PAs as a driving force behind microscopic aggregation. However, findings from our systematic PA library (**SI chapter 7.3**) suggest that fibrils that do not aggregate share a common feature: they consist of a high number of less hydrophobic, flexible (A, G) and polar amino acids (R, K, E) close to the fatty acid (**Figure 3A**). This could increase dynamicity and solvation of the self-assembling PAs, which results in less aggregating fibers. PAs with a block-type arrangement of hydrophobic (V, I) and hydrophilic, charged (K, R) amino acids, as found in PA1, form small-sized aggregates (**Figure 3B**). In contrast, strongly aggregating fibrils consist of high numbers of hydrophobic aromatic amino acids and an alternating amphiphilic sequence order as found in PA2 (**Figure 3C**).

From these observations, we hypothesized that surface hydrophobicity of the assembled fibrils could be a key feature that determines their stability and aggregation. To evaluate this hypothesis, we performed coarse-grained (CG) molecular dynamics simulations on three peptide amphiphiles that show distinct aggregation behavior: C_{16}-AGGRVK-NH_2, PA1, and PA2 (**Figure 3D – F**).

In the first step, we studied the self-assembly of PAs in solution from randomly dispersed molecules to fibrils using CG Martini model (**SI chapter 13**).[77,78] After attaining stable equilibrium structures in under 15 µs, we further investigated fibril-fibril interactions by placing fibrils in close proximity in solution to gain insight into the different degrees of aggregation observed in experiments. Although the Martini model has limitations, such as determining the evolution of peptide secondary structure, in this case, we are interested in not the secondary structure of peptide amphiphiles but in the aggregation behavior to make connections with experiments.

In line with experimental findings, our simulations show differences in the arrangement of amino acid residues in the core and on the surface of the fibrils depending on the PA sequence. For example, while the block-type PA1 forms distinct domains such as the hydrophobic core and the hydrophilic surface of the fibrils (**Figure 3E**), PA2's alternating sequence allows for less homogeneous distribution of the amino acids on the surface. While the fibrils are stabilized by the hydrophobic tail, we observe the presence of not only hydrophilic but also hydrophobic (phenylalanine and cysteine) residues surface-exposed and in contact with solvent molecules (**Figure 3F**). This, in turn, allows for a more attractive interaction between fibrils in solution, which aggregate to form connecting networks of "supramolecular cross-links" between the fibrils. In contrast, PA1 and C_{16}-AGGRVK-NH_2 (**Figure 3D**) show very little aggregation in solution. This is qualitatively reflected in fiber-fiber distance (**Figure S34**) and solvent accessible surface area (SASA, **Figure S35A**) in **SI chapter 13.2**.

The aggregation behavior also affects the mobility of the supramolecular components measured by root mean square fluctuations (RMSF) of residues, where increased fibril-fibril aggregation in PA2 results in less mobile fibrils (**Figure S36**). Other non-aggregating fibrils, such as PA1 and C_{16}-AGGRVK-NH_2 also remain stable but show larger movements in response to interactions with solvent and other fibrils in solution (**SI chapter 13.3**). Noteworthy the RMSF changes for PA1 depending on whether single (**Figure S36A)** or multiple fibers (**Figure S36B)** are considered, which indicates that it is important to not only model the self-assembled fiber but also the inter-fiber interactions for evaluation of fibril properties.

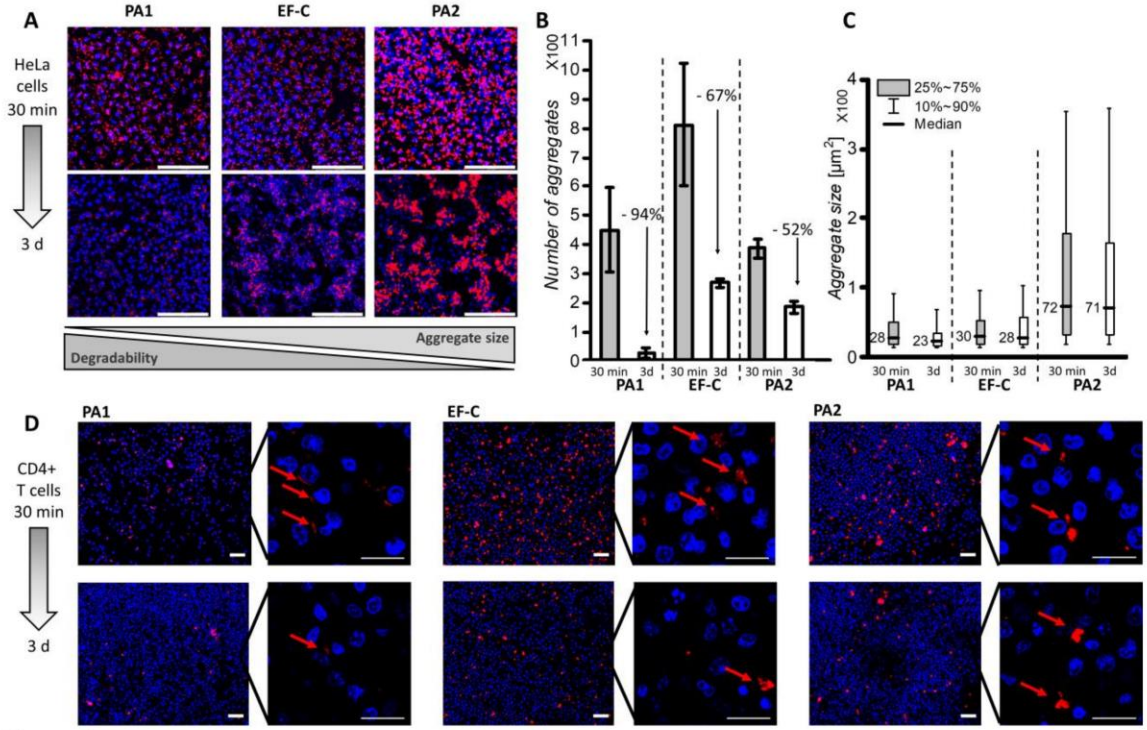


Figure 4 Degradability of µm-sized fibrillar aggregates formed by EF-C, PA1 and PA2 after 30 min and 3 d in HeLa and CD4+ T cell culture. **A** Representative laser scanning microscopy images of Proteostat-labeled EF-C, PA1 and PA2 (red) incubated with HeLa cells (nucleus in blue). In general, PA1 reveals less and smaller µm-sized aggregates than PA2 and EF-C. After 3 days in cell-culture, almost all small aggregates of PA1 are degraded, whereas aggregates formed by PA2 stay almost unchanged, scale bar 200 µm. **B** Bar plot showing number of aggregates > 10 μm^2 in an area of 581.25 µm × 581.25 µm in cell culture after 30 min and 3 d incubation (4 µg peptide added to 40,000 cells), error bars indicate standard deviations from six technical replicates. **C** Box plot showing size distribution of aggregates in an area of 581.25 µm × 581.25 µm in cell culture after 30 min and 3 d incubation (4 µg peptide added to 40,000 cells), deviations indicate three technical replicates. **D** Representative laser scanning microscopy images of Proteostat-labeled EF-C, PA1 and PA2 (red) incubated with CD4+ T cells (nucleus in blue). PA1 reveals less and smaller µm-sized aggregates that attach as thin layers to cell membrane compared to PA2 and EF-C (indicated by red arrows). After 3 days in cell culture, almost all small aggregates of PA1 are degraded by CD4+ T cells (1 µg peptide added to 250,000 cells), scale bar 20 µm.

Peptides with an alternating amphiphilic sequence order and high content of hydrophobic amino acids are commonly found in amyloidogenic peptides.[21] Our findings suggest that these properties can also promote intermolecular and fiber-fiber interactions of PAs, leading to increased colloidal aggregation.

PAs – Forming Small Aggregates are Degraded Efficiently

To assess how the aggregate size affects the biodegradability of the PAs, the number of aggregates > 10 μm^2 were evaluated via confocal microscopy in the presence of HeLa cells after 30 min and 3 d of incubation at 37°C (**SI chapter 7**). This time interval was selected, because the corresponding infectivity assays are always performed for at least 3 d. Most of the peptide nanostructures such as PA1 and PA2 investigated in this work as well as other commercially available peptides for transduction efficiency enhancement such as EF-C,[23] Vectofusin-1[79] form µm-sized aggregates (**Figure 4A**, **Figure S6**). While aggregate formation can be an obstacle in clinical applications[24] it does not appear to affect cell viability (**Figure 5D**, **Figure S7, Figure S25**). To avoid possible complications upon re-administration of e.g. transduced T-cells, the reduction of aggregates (> 10 μm^2) is highly desired for clinical applications.[80,81] We found that 16 of the 35 investigated PAs could be degraded by cells (**SI chapter 7, Figure S12**). For example, PA1 forms less and smaller aggregates compared to EF-C and PA2, which could be well degraded (94%) over the course of 3 days in the presence of HeLa cells (**Figure 4A, B, C, Figure S8**) as well as CD4+ T cells (**Figure 4D, Figure S9**). In contrast, PA2 forms larger µm-sized aggregates, which remain nearly unchanged in size after 3 days incubation and are not degraded as efficiently (52%) in presence of HeLa cells (**Figure 4B**, **Figure S8**) and CD4+ T cells (**Figure 4D**, **Figure S9**). The efficient cellular biodegradation for PA1 is further emphasized in control experiments without cells where PA1, PA2 and EF-C are not significantly changing in size and number over the course of 3 days (**Figure S10**).

In contrast to P1, P2 forms fibrils (**Figure 2A**). Thus, after cleavage of the fatty acid, fibril formation and microscopic aggregation of P2 could hinder degradation as observed here. To examine how peptide sequence biodegradability impacts overall PA degradability, we designed pa1, which is the D-amino acid equivalent of PA1 (**Figure S11**). All physicochemical properties of pa1 remain similar to those of PA1. However, fibrillar aggregates of pa1 cannot be degraded by HeLa cells within a 3-day timeframe. Thus, the reduced biodegradability for pa1 can be attributed directly to sequence degradability, specifically the inability of cells to degrade D-amino acids.

We conclude that PA1 forms smaller aggregates and it is degraded more quickly *in vitro*, which distinguishes them from amyloid-like peptide fibrils, e.g. EF-C or PAs consisting of alternating amphiphilic amyloid-peptide sequences such as PA2, which show in general higher *in vitro* stability (**Figure 4**, **Figure S7**).[82,83] PA1 is much better degraded compared to PA2 because of non-assembling parent peptide sequence and the small size of aggregated fibrils.

Positively Charged PAs are Potent Enhancers of γ-Retroviral Transduction

Finally, we evaluated the capacity of the PA library to enhance HIV-1 infection of TZM-bl cells, which is a well-established and frequently applied model to quantifying HIV-1 infectivity.[84] EF-C (QCKIKQIINMWQ), which has been reported earlier by us,[20] served as a reference. EF-C is a highly potent enhancer on par with other commercially available transduction enhancers such as RetroNectin and Vectofusin-1 serving as gold standards for transduction enhancers.[23,79]

Out of the 35 PAs in our library, 21 showed a significant increase in infectivity by at least 2-fold compared to the virus alone at a concentration of 1.3 µM (**Figure S12**). This confirms the effectiveness of adding fatty acids to the N-terminus to transform the inactive parent peptide sequences **(Figure S23)** into ones that enhance infectivity (**Figure S19**). Inactive PAs lack the physicochemical properties which were reported earlier by us as important requirements for infectivity enhancement.[21] They either cannot self-assemble into fibrils (*e.g.* C_{16}-KGVPGVGK-NH_2, alpha linolenic acid-VVVAAAKKK-NH_2,

C_{16}-WEALKK-NH_2), cannot form µm-sized aggregates (*e.g* Nap-VVVAAAKKK-NH_2, C_{16}-GGGSSKKK-NH_2, C_{16}-AGAGRRR-NH_2, C_{16}-GGVVRR-NH_2, C_{16}-RRWQWR-NH_2, C_{16}-GKVGSK -NH_2, **SI chapter 7**) or have a negative zeta-potential (PA3, **Figure 5A**). In accordance with the above demonstrated aggregate analysis, PAs which cannot form fibrillar aggregates contain less hydrophobic amino acids (G, A), more flexible hydrophobic moieties (alpha linolenic acid, naphthalene) or charged amino acids (R, K, E) close to the hydrophobic moiety.

PA1 and PA2 are amongst the most active PAs and show a strong infectivity enhancement of HIV-1 in the range of EF-C (**Figure 5A**). Our data indicate a similar mechanism for the activity as reported previously for EF-C.[23] The positively charged fibrillar aggregates of PA1 and PA2 interact with negatively cellular membranes of HeLa cells as shown by confocal fluorescence microscopy measurements depicted in **Figure 4A** and **Figure S6**. Importantly, the parent peptides P1 and P3 without fatty-acid residues are inactive (**Figure 5A**) and did not interact with the cellular membrane (**Figure S6**). These findings demonstrate that even inactive peptides can be rendered active when a fatty acid residue is introduced and induces self-assembly into fibrils with high β-sheet content. P2 with the alternating amphiphilic sequence design can assemble into moderately positive charged fibrils (**Figure 2A, B**) and thus enhances infectivity even without the fatty acid residue (**Figure 5A**). The addition of a fatty acid moiety to peptides can promote the formation of β-sheet rich fibrils, which in turn increases their transduction enhancing activity, given that the overall charge is positive.

To investigate the potential use for gene therapy applications we evaluated the effect of P1–P3 and PA1–PA3 on the transduction of Jurkat cells and clinically relevant primary T cells. First, we tested Jurkat cells, a T-cell line frequently applied for studying leukemia,[85] with a γ-retroviral vector pseudotyped with the glycoprotein derived from gibbon ape leukemia virus (GALV-RV). Without a PA additive, almost no transduced GFP+ cells were detected (**Figure 5B**). However, with an increasing concentration of PA1 and PA2 strong transduction enhancement similar to EF-C was achieved. Next, PA1 and PA2 were tested in a more clinically relevant transduction setting. To this end, primary CD4+ T cells were isolated, activated and transduced with GALV-RV in presence of different transduction enhancers. Strikingly, PA1 and PA2 showed transduction efficacies in the range of commercially available gold-standard enhancers RetroNectin and Vectofusin-1 (**Figure 5C**), which were utilized as

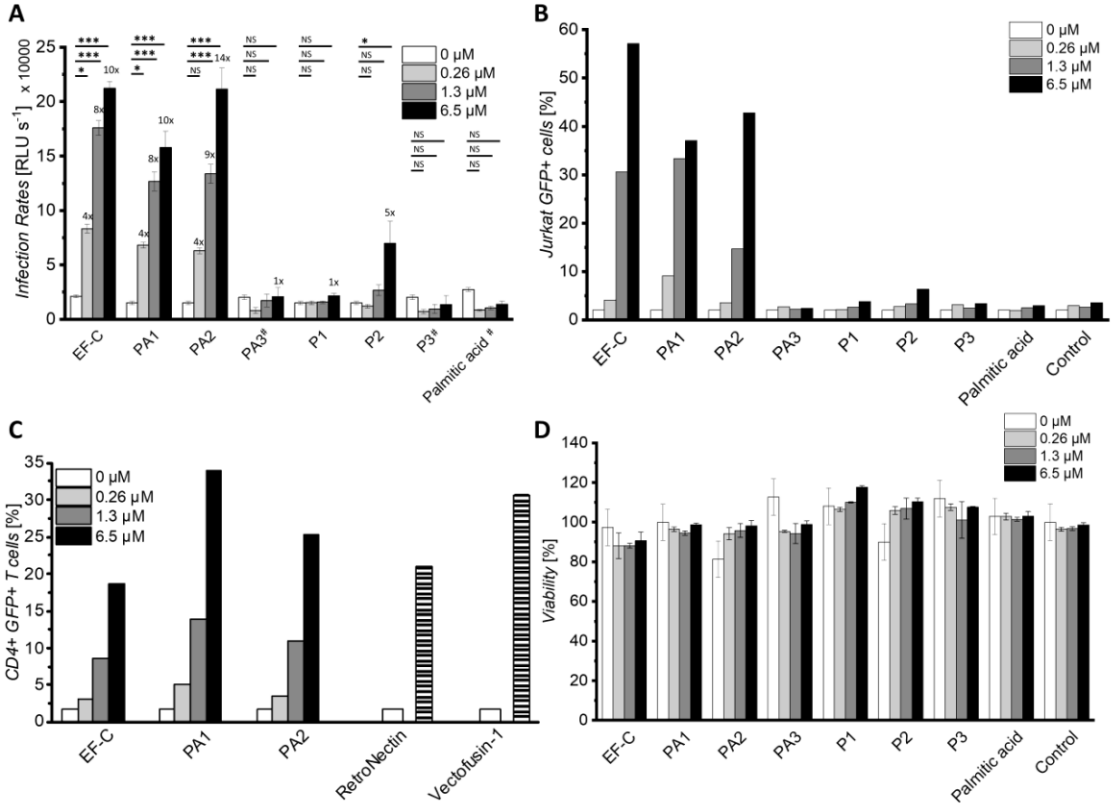


Figure 5 Peptide amphiphiles PA1 and PA2 are potent enhancers of HIV-1 infection and γ-retroviral transduction**. A** Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of peptides PA1–PA3 and P1–P3 (0.26 µM, 1.3 µM, 6.5 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. Two repetitions of triplicate infections were used for samples marked with (#). P-values for statistical significance were determined by two-way ANOVA test with Dunnett's multiple comparison test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant, RLU s^{-1} are relative light units per second. **B** Percentage of green fluorescent protein (GFP) positive Jurkat cells at indicated concentrations of peptides three days after transduction with a gibbon ape leukemia virus pseudotyped γ-retroviral vector (GALV-RV). Experiment has been performed once. **C** Percentage of GFP positive CD4+ T cells after transduction with GALV-RV. Comparing the transduction efficiency enhancement of EF-C, PA1 and PA2 (0.26 µM, 1.3 µM, 6.5 µM, for EF-C 6.5 µM equals 10 µg mL^{-1}) with benchmark retroviral transduction enhancers RetroNectin (20 µg mL^{-1}, 0.3 µM) and Vectofusin-1 (10 µg mL^{-1}, 3.3 µM) as recommended by the manufacturers. Experiment has been performed once. **D** CellTiter-Glo cell viability assay of Jurkat cells treated with increasing concentrations of the peptides of biological triplicate with standard deviation.

recommended by the manufacturers, *i.e.* time-consuming coating and washing steps in the case of RetroNectin. Thus, PAs are powerful transduction enhancers with similar transduction enhancing efficacies as commercially available systems. Furthermore, PA1 and PA2 did not reduce the viability of Jurkat cells (**Figure 5D**), or CD4+ T cells (**Figure S7**), which is important in terms of CAR-T cell therapy.

In summary, PAs that have block-type sequence design with hydrophobic and cationic charged amino acids, *i.e.* in PA1 form moderately sized aggregates that are attaching to cellular membranes, deliver viruses and are biodegradable (**Figure 6A**). Replacing cationic amino acids with anionic amino acids, *i.e.* in PA3 prevent cellular attachment and are not suitable for retroviral gene delivery (**Figure 6C**). Changing the sequence design from block-type to an alternating amphiphilic sequence design with aromatic, hydrophobic amino acids results in PAs that form large aggregates (> 50 µm^2, *i.e.* PA2), which can increase virus uptake but that cannot be fast and efficiently degraded by the cells (**Figure 6B**). While a certain degree of aggregation is required **(SI chapter 7)** to efficiently interact with cells and viruses, the size of the aggregates is crucial to facilitate biodegradability.

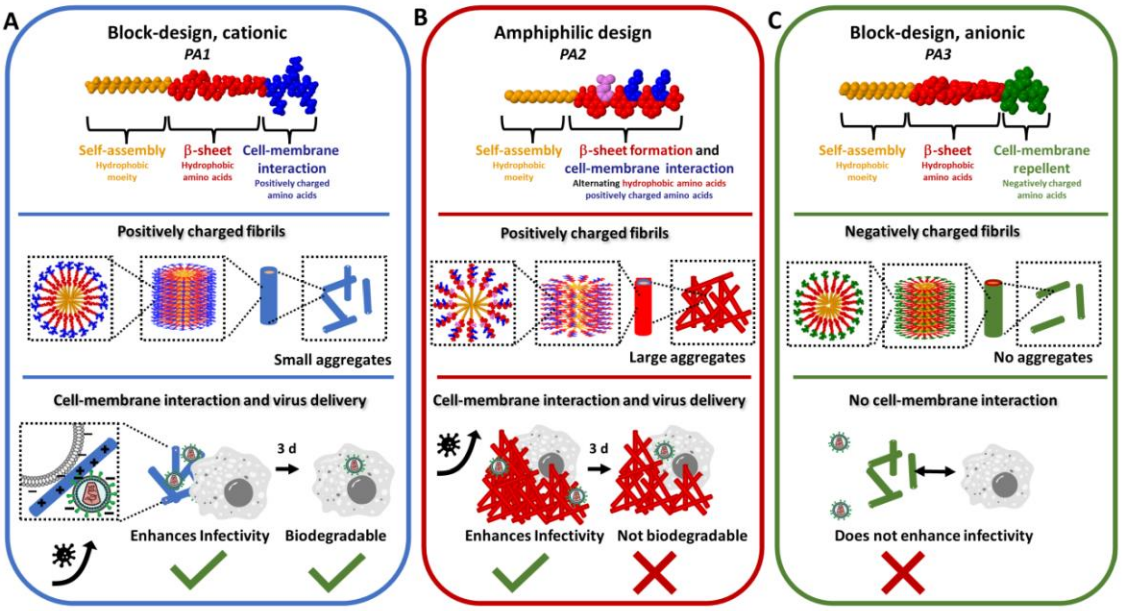


Figure 6 Schematic illustration of main findings in this study. **A** PAs with block-type design of hydrophobic and cationic amino acids (e.g. PA1, C_{16}-VVVAAAKKK-NH_2) form moderately sized aggregates that can attach to cellular membranes, deliver viruses and are biodegradable in the timeframe of 3 d. **B** PAs with an alternating amphiphilic sequence design with hydrophobic amino acids (e.g. PA2, C_{16}-FQFKFKC-NH_2) form large aggregates (> 50 µm^2) that can deliver viruses but cannot be efficiently biodegraded. **C** PAs containing anionic amino acids (e.g. PA3, C_{16}-VVVAAAEEE-NH_2) hinder cellular attachment and are not suitable for retroviral gene delivery.

Conclusion

In this study, we present peptide amphiphiles (PAs) as biodegradable additives for promoting retroviral transduction. The conjugation of fatty acid residues can render inactive peptide sequences active as demonstrated by evaluating 35 different PAs in a virus infection assay representing a promising approach to discover new PAs for enhancing retroviral transduction on a par with established materials like RetroNectin, Vectofusin-1 and EF-C. Active PAs showed the characteristic ability to form β-sheet rich fibrous aggregates with a positive zeta-potential making them highly potent transduction enhancers. The arrangement and hydrophilicity of amino acids in the peptide sequence plays a crucial role in controlling the aggregate morphologies and biodegradability of the peptide amphiphiles (PAs). We could demonstrate via molecular dynamics simulations that fibril-fibril aggregation of PAs could be traced back to the fibril surface hydrophobicity and self-assembly dynamics, with fibrils consisting of more mobile PAs causing less aggregation. Specifically, block-type amphiphilic peptide motifs afford PAs forming small aggregates that can be efficiently degraded in the presence of cells, while alternating amphiphilic peptide sequences yield PAs with high tendency to form large aggregates that are not degraded by cells within the relevant time of the transduction experiment.

In summary, we have introduced PAs with a block-type sequence arrangement as a novel class of transduction enhancers. These PAs share key physicochemical traits with previously reported amyloid-like self-assembling peptides, but they reveal excellent biocompatibility and degradability. We envision that this class of transduction enhancers could outperform conventional amyloid-like peptides in future *ex vivo* clinical gene transfer applications.

Experimental Section

Materials

For peptide synthesis, Rink Amide AM resin LL (100-200 mesh), OxymaPure®, (2-(1*H*-benzotriazol-1yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Val-OH, Fmoc-Ala-OH and Fmoc-Glu(OtBu)-OH were purchased from Novabiochem®, Merck. Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Piperidine, Trifluoroacetic acid, *N,N*-Di*iso*propylethylamine and *N,N*'-Di*iso*propylcarbodiimide were purchased from Carl Roth GmbH + Co. KG. Palmitic acid and Tri*iso*propylsilane were purchased from TCI Tokyo Chemical Industry. Peptide-synthesis-grade solvents (Dichloromethane, Dimethylformamide) were acquired from Acros Organics. Solvents were purchased in HPLC grade purity or higher from Sigma-Aldrich (Diethyl ether, Dimethylsulfoxide (DMSO)), Fisher Scientific (Dichloromethane, Dimethylformamide, Acetonitrile). Peptide EF-C (>95% purity) were purchased from Zhengzhou Phtd Peptides industrial Co. Vectofusin-1 was purchased from Miltenyi Biotech and RetroNectin was purchased from Takara Bio. Proteostat® was purchased from Enzo Life Sciences. Thioflavin-T, Uranyl acetate and Dulbecco's Phosphate Buffered Saline without Magnesium and Calcium (DPBS) were purchased from Sigma-Aldrich. Dulbecco's Modified Eagle Medium (DMEM), Roswell Park Memorial Institute (RPMI) 1640 Medium, fetal calf serum (FCS), Dynabeads™ Human T-Activator CD3/CD28 were purchased from Gibco (Germany) was purchased from Sigma-Aldrich. Penicillin-Streptomycin and L-Glutamine (200 mM) were purchased from PAN-Biotech. Human interleukin-2 (IL-2), interleukin-7 (IL-7) and interleukin-15 (IL-15) were purchased from Miltenyi Biotech. RosetteSep™ Human CD4+ Enrichment Cocktail was purchased from STEMCELL Technologies.

Instruments

Peptides were synthesized with an automated microwave peptide synthesizer (CEM, Liberty BlueTM). High pressure liquid chromatography (HPLC) on preparative scale was conducted with a Shimadzu system equipped with the following modules DGU-20A5R, LC-20AP, CBM-20A, SPD-M20A, SIL-10AP, FRC-18A with a reverse phase column Phenomenex Gemini® (5 µm NX-C18, 110 Å, 150 x 30 mm). Analytical scale HPLC was performed using a Shimadzu system with modules DGU-20A5R, LC-20AT, CBM-20A, SPD-M20A, SIL-10ACHT, CTO-20AC, column Zorbax XDB-C18, 9.4 x 250 mm, 5 µm pore size. LC-MS was conducted on a Shimadzu LC-2020 Single Quadrupole MS instrument with the modules LC-20AD, SIL-20ACHT, SPD-20A, CTO-using a Kinetex EVO and C18 100 Å LC 50 × 2.1 mm column with 2.6 µm pore size. MALDI-TOF-MS was measured on a Bruker rapifleX MALDI-TOF/TOF and a Waters MALDI Synapt G2-SI instrument. TEM measurements were performed on a JEOL 1400 instrument with 120 kV acceleration voltage equipped with a CCD camera. Fluorescence intensity was measured using a Spark 20 M microplate reader by the company Tecan Group, Ltd. Zeta-potential measurements were conducted with a Zetasizer Nano ZS (Malvern Instruments) with 1 mL disposable folded capillary cells. FT-IR spectra of solid samples using a Bruker Tensor II spectrometer equipped with a diamond crystal as ATR element with a spectral resolution of 2 cm–1, each spectrum was an average of 64 scans. Circular dichroism measurements were conducted on a JASCO 1500 instrument. Fluorescence microscopy measurements were recorded on a Leica DMi8 microscope with 10x air objective and equipped with a Leica MC170 HD camera. Confocal laser scanning microscopy was performed on a Stellaris 8 confocal laser scanning microscope (Leica) equipped with a 20x air objective or on a Zeiss LSM 710 confocal microscope equipped with a 20x and 63x objective. Luminescence measurement for the evaluation of infection efficiency was conducted on Orion microplate luminometer (Berthold). Flow cytometry was conducted on CytoFLEX LX Flow Cytometer (Beckman Coulter) and flow cytometry data were analyzed using FlowJo (BD Bioscience).

Synthesis and Purification of Peptides

Peptides (PA1-3 and P1-3) were synthesized using an automated microwave peptide synthesizer (CEM, Liberty BlueTM) from C to N-terminus according to fluorenylmethyloxycarbonyl (Fmoc) solid phase peptide synthesis (SPPS) strategy by Merrifield using Rink Amide AM Resin LL (100-200 mesh, Novabiochem). The beads were swollen in DMF while shaking the reaction vessel for 1 h. The Fmoc-protecting group was cleaved off by a piperidine solution (20 % in DMF) at 155 W, 75 °C for 15 s and at 30 W, 90 °C for 50 s. The beads were washed three times with DMF before the addition of the Fmoc-protected amino-acid (2M, 5 equiv. relative to the resin loading capacity) in DMF, DIC (5 equiv.) in DMF and Oxyma Pure® (10 equiv) in DMF. The coupling was conducted by microwaving at 170 W, 75 °C for 15 s and at 30 W, 90 °C for 110 s, the solution was removed, and the beads were washed with DMF. This procedure was repeated for each amino acid. Coupling of the fatty acid chain was conducted according to a previous reported literature procedure[62,63] by adding palmitic acid (4 molar equiv), DIPEA (6 molar equiv), HBTU (4 molar equiv) in DMF/DCM (5/1, v/v) at rt and shaking for 2 h. Cleavage from the resin was carried out in TFA/water/tri*iso*propylsilane (95/2.5/2.5, v/v/v) for 2 h at room temperature. This solution was precipitated in cold diethyl ether (40 mL) and centrifuged 3 times at 4 °C, 4000 rpm for 15 min. The precipitate was dissolved in water with 0.1 % TFA for peptides with lysine residue and with 0.1 % NH_4OH for peptides with glutamic acid residues yielding PAs as colourless powders (PA1 34%, PA2 25%, PA3 46% yield). The purification was conducted via HPLC (reverse phase column Phenomenex Gemini® 5 µm NX-C18, 110 Å, 150 x 30 mm) using a gradient of water and acetonitrile containing 0.1 % TFA for peptides with lysine residues or 0.1 % NH_4OH for peptides with glutamic acid residues as the mobile phase at a flow rate of 25 mL min^{-1}. Chromatography was

monitored with an UV absorption detector at 190 and 254 nm. The peptides were identified via MALDI-TOF-MS and purity was confirmed via LC-MS or analytical HPLC as described in an previous report.[86] The data was processed with Shimadzu LabSolutions and OriginLab. The PA and peptide library (SI chapter 7-12) were commercially obtained from GenScript Inc. > 95 % purity except of C_{16}-*ester*-SVVVAAAKKK-NH_2, C_{16}-*disulfide*-CVVVAAAKKK-NH_2, C_{16}-*thioester*-CVVVAAAKKK-NH_2, and Eicosapentaenoic acid – VVVAAAKKK-NH_2 which were purchased from Bachem AG > 87% purity.

Fibril Formation

The peptides were predissolved in DMSO to a concentration of 10 mg mL^{-1}. The fibril formation was induced by further diluting the solution to a final incubation concentration of 650 µM or 1 mg mL^{-1} in PBS as indicated in the text and incubated for 1 d at RT. Other concentrations described in the text were achieved by further dilution of preformed fibrils.

Characterization of Fibrils

Transmission electron microscopy (TEM) measurements, Thioflavin (ThT), zeta-potential measurements for fibril characterization, ATR-IR spectra acquisition and the calculation of β–sheet were prepared and conducted analogous to a previous report.[20,86] The morphology of the peptide assemblies was evaluated via TEM to determine the self-assembly at the nanoscale into fibrillar structures or amorphous aggregates. Amorphous aggregates lack a fibrillar structure and have less internal order.[87–89] Fluorescence microscopy of dye-stained peptides was applied to distinguish between microscopically aggregating or distinct fibrils. Fibrils were defined as distinct, if no visible structures could be identified in fluorescence and brightfield microscopy. A peptide fibril was considered as ThT-active if the fluorescence intensity was at least twice as strong compared to the control (PBS containing 10% DMSO). The derived count rate of the zeta-potential measurement was used as information on the light scattering intensity and turbidity of the sample as a further indicator for microscopic aggregation.[90]
The experimental procedure is briefly described below.
TEM: 5 µL of fibrils (1 mg mL^{-1}) were placed on 300 mesh, Formvar layered copper grids. After 10 min incubation time, the grids were staining with 4 % uranyl acetate solution for 2.5 min. After three washing steps with dd H_2O excess solvent was removed by filter paper. Measurements were performed on a Jeol 1400 electron microscope with 120 kV acceleration voltage.
ThT-fluorescence: 20 µL 50 µM of ThT solution in PBS was pipetted in a black 384 well-plate and 4 µL of the nanofibril solution 1 mg mL^{-1} were added. For reference PBS with 10% DMSO (4 µL) instead of fibril solution was added. Subsequently, fluorescence emission was recorded λem= 488 nm upon excitation at λex = 440 nm with 10 nm bandwidth and multiple reads per well (3 × 3) using a Spark 20 M microplate reader (Tecan Group, Ltd). Data processing was performed with Origin software. A peptide fibril was considered as ThT-active if the fluorescence intensity was at least twice as strong compared to the control (PBS containing 10% DMSO).
Proteostat fluorescence: 1 µL of a Proteostat stock solution (Enzo Life Science) were diluted in 100 µL PBS. 5 µL of the Proteostat solution was added to 10 µL of the peptide fibrils (1 mg mL^{-1}) in a 384 well plate. Fluorescence intensity was determined at λem= 600 nm upon excitation at λex = 550 nm with 10 nm bandwidth and multiple reads per well (3 × 3) using a Spark 20 M microplate reader (Tecan Group, Ltd). Data processing was performed with Origin software.

Zeta-potential: 60 µL of the peptides (1 mg mL^{-1}) were diluted in 600 µL, 1 mM KCl solution and placed in a disposable folded capillary cuvette. All zeta-potential measurements (Zetasizer, Malvern Instruments) were conducted in triplicates for each sample.

ATR-FT-IR: ATR-IR spectra acquisition and the calculation of β–sheet content were conducted according to a previous report.[20] Briefly, FT-IR spectra of solid samples were recorded after lyophilization of fibril solutions using a Bruker Tensor II spectrometer equipped with a diamond crystal as ATR element with a spectral resolution of 2 cm–1, each spectrum was an average of 64 scans. The data was processed with Origin software. β–sheet amount was determined by calculating the integral of the peak at 1630 cm^{-1} (β-sheet) relative to the integral between 1650-1700 cm^{-1} (native structure) from the 2nd derivative of the FT-IR spectra.

CD: Circular dichroism measurements were performed on a JASCO J-1500 instrument in a 1 mm path length quartz cuvette (HellmaAnalytics) from 190 – 250 nm with peptides directly dissolved in MilliQ water to 1 mg mL^{-1} and subsequently pH adjusted to pH 7.4 with 0.1 M NaOH (aq.) or HCl (aq.) and incubated for 1 d to induce assembly. CD measurements were conducted in pure water to avoid aggregate formation which occurs in presence of phosphate ions from PB buffer and disturbs the measurement. It was previously shown that CD measurements in PB or MilliQ yield comparable results for other PA.[91] The samples were incubated for 1 d and diluted to 0.1 mg mL^{-1} before measurement. Three accumulations were averaged for each measurement with a wavelength step size of 2 nm and scanning speed of 20 nm min^{-1}. For each measurement, the corresponding background measurements were conducted and subtracted from the original spectrum.

Fluorescence microscopy: 1 µL Fibrils (1 mg mL^{-1}) were stained with 9 µL, 50 µM ThT. 10 µL of the stained fibrils (0.1 mg mL^{-1}) were placed on a microscopy glass. Fluorescence microscopy measurements were performed on a Leica DMi8 microscope with 10x air objective and equipped with a Leica MC170 HD camera with a filter setting λ_{em} = 527/30 nm upon excitation at λ_{ex} = 480/40 nm) and processed with the Leica LASX software.

Cell culture

TZM-bl and HeLa cells were cultured in Dulbecco´s Modified Eagle Medium (DMEM) supplemented with 10 % (v/v) inactivated fetal calf serum (FCS), L-glutamine (2 mM), penicillin (100 units mL^{-1}) and streptomycin (100 µg mL^{-1}). Jurkat cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 Medium supplemented with 10 % (v/v) fetal calf serum (FCS), L-glutamine (2 mM), penicillin (100 units mL^{-1}) and streptomycin (100 µg mL^{-1}). CD4+ T cells were cultured in supplemented RPMI medium in presence of interleukin-2 (IL-2) or interleukin-7 (IL-7) and interleukin-15 (IL-15) as described below.

Virus—Peptide Fibril—Cell Interaction

Infectious HIV-1 stocks, GALV-pseudotyped γ-retroviral vector (GALV-RV) were prepared analogous to a previous report and briefly described in the following sections.[20]

Virus stocks: For generating virus stocks 800,000 HEK293T cells were seeded. Next day, transfection using different plasmids and TransIT-LT1 Transfection Reagent (Mirus Bio, Cat: Mir 2305) according to manufacturer's protocol was performed. Two days after transfection virus stocks were harvested, centrifuged (3 min at 1300 rpm) and supernatants were stored at – 80 °C. Infectious HIV-1 stocks were prepared by transfection of HEK293T cells using pBRNL4.39-92TH14 (5 µg), a plasmid containing a CCR5 tropic molecular HIV-1 clone.[92] GALV retroviral vector (GALV-RV) was produced by cotransfection of HEK293T cells with following plasmids: GFP expressing Murine leukemia virus (MLV) vector E200

pcmE26-gfp (2.3 µg), E848 pCsGPpA-ed (1.9 µg) and glycoprotein derived from Gibbon ape Leukemia Virus (GALV, 0.77 µg).[20]

HIV-1 Infection Assay: The effect of peptide fibrils on HIV-1 infection was studied via a luminescence assay for β-galactosidase, which was reported by TZM-bl cells upon infection. 10,000 TZM-bl cells were seeded in 180 µl supplemented DMEM one day prior infection. The cells were inoculated with 20 µl PNF solution 1:1 (v/v) mixed with virus (1:500 dilution) after an incubation time of 10 min. The HIV-1 infection assay was conducted in three technical replicates and repeated in three independent experiments unless stated otherwise. The assay was conducted as described before.[20]

γ-Retroviral Infection Assay of Jurkat cells: 50,000 Jurkat cells were seeded in 180 µl supplemented RPMI. The cells were inoculated with 20 µl PNF solution 1:1 (v/v) mixed with virus (1:3.2 dilution) after an incubation time of 10 min. The effect of peptide fibrils on retroviral transduction of Jurkat cells was analyzed 3 days post infection by determining the percentage of GFP + cells by flow cytometry. All assays were conducted as described before.[20]

γ-Retroviral transduction Assay CD4+ T cells: Human CD4+ T cells were isolated from a Buffy coat using RosetteSep™ Human CD4+ Enrichment Cocktail (STEMCELL Technologies) according to manufacturer's protocol. Isolated T cells were activated using Dynabeads™ Human T-Activator CD3/CD28 (GibcoTM) according to manufacturer's protocol in presence of 10 ng mL^{-1} IL-2 (Miltenyi Biotec) for three days. After 3 days the magnetic Dynabeads were removed using a magnet and T cells were cultured in presence of 9 ng mL^{-1} IL-7 (Miltenyi Biotec) and 10 ng mL^{-1} IL-15 (Miltenyi Biotec). After 1 day 5×10^5 T cells were transduced with GALV-γ-RV in absence and presence of PA1, PA2, EF-C (0.26, 1.3 and 6.5 µM) and Vectofusin®-1 (10 µg mL^{-1}; Miltenyi Biotec), or RetroNectin (20 µg mL^{-1}; Takara Bio). The concentration of the transduction enhancers corresponds to the concentration on cells and in the case for RetroNectin it corresponds to the concentration that was used to precoat the 24-well plate. The 24-well plate was coated with RetroNectin and incubated overnight at 4 °C. Wells were blocked with PBS containing 2 % BSA for 30 min and washed twice with PBS. Virus was centrifuged onto coated wells at 2,000 g at 32 °C for 2 h. Viral supernatant was removed and CD4+ T cells were added and subsequently spin-infected (300 g for 10 min) with virus again. 1 day after transduction the medium was exchanged. After 3 days GFP+ T cells were analyzed by flow cytometry. Viability was determined using LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit (ThermoFisher Scientific) according to manufacturer's protocol. For the investigation of the transduction enhancing effect and viability of peptides on CD4+ T cells the gating strategy shown in Figure S8A was applied.

Cell Viability with TZM-bl and Jurkat cells: The cell viability after addition of peptides to TZM-bl cells and Jurkat was studied via the CellTiter-Glo assay. To this end, 20,000 cells TZM-bl or 50,000 Jurkat cells were incubated with serial diluted peptides. After 3 days the supernatant was removed and 100 µL CellTiter-Glo Reagent 1:1 diluted in PBS was added. After 10 min 50 µL was transferred to white microplate and luminescence was recorded by Orion microplate luminometer.

Confocal laser scanning microscopy and degradation study with PA1-3 and HeLa cells: HeLa cells were seeded one day prior to conducting the assay (40,000/well) in an 8-well IBIDI slide. 4 µL of the preformed peptide fibrils (1 mg mL^{-1}) were diluted with 4 µL Proteostat (Enzo Life Science, 1 µL stock in 999 µL PBS) and incubated for 10 min and further diluted with medium to receive a final peptide concentration of 20 µg mL^{-1}. The nucleus of the HeLa cells was stained with Hoechst 33342 (NucBlue™, Thermo Fisher Scientific). The peptide solution was transferred to the HeLa cells and incubated for 30 min at 37°C before washing three times with PBS. The interaction of fibril clusters with cells was monitored after 30 min incubation time on a Stellaris 8 confocal laser scanning microscope (Leica) equipped with a 20x air objective and laser excitation wavelengths of 405 nm (Hoechst) and 561 nm

(Proteostat). For the evaluation of aggregate degradation, the samples were further incubated at 37°C for 3 d and analyzed without any additional washing step. The evaluation of aggregate size and number was conducted by analyzing Proteostat stained aggregates via the 3D objects counter function of ImageJ 1.53f51 for 3 individual measurements (area 581.25 µm × 581.25 µm). The size distribution shows aggregates sizes larger than the threshold 10 μm^2 in an area of 581.25 µm × 581.25 µm and repeated in three technical triplicates. The number of aggregates larger than 10 μm^2 were counted via the 3D objects counter function of ImageJ 1.53f51. The mean number and standard deviations were determined by averaging at least three technical triplicates.

Confocal laser scanning microscopy and degradation study with PA1-3 and CD4+ T cells: The degradation of the peptide assemblies on CD4+ T cells was determined by analyzing fluorescence microscopy images. To this end, 4 µL of the preassembled peptide fibrils EF-C, PA1 and PA2 (1 mg mL^{-1}, 1 d in PBS) were mixed with 4 µL of Proteostat (Enzo Life Science, 1 µL stock in 99 µL PBS) and incubated for 15 min. 2 µL of the stained fibrils were added to 198 µL preactivated and nucleus stained (Hoechst 33342) CD4+ T cells (200 µL, $0.25x10^6$) resulting in a final 5 µg mL^{-1}concentration of the peptide. After incubation for 30 min at 37°C the cells were washed with 3x with PBS and incubated in medium (supplemented RPMI containing IL-15 and 7). Microscopy was conducted on a Zeiss LSM 710 confocal microscope with 20x and 63x objective and laser excitation wavelengths of 405 nm (Hoechst) and 561 nm (Proteostat). The measurements were performed after 30 min and 3 d of incubation at 37°C without any washing step in between. Control experiments to analyze changes of fibril aggregation were conducted analogously without the presence of cells. The evaluation of aggregate size and number was conducted by analyzing Proteostat stained aggregates via the 3D objects counter function of ImageJ 1.53f51 for 5 individual measurements (area 708.5 µm × 708.5 µm) of three individually repeated experiments.

Confocal laser scanning microscopy and degradation study with PA library and HeLa cells: Confocal laser scanning microscopy studies were performed for the visualization of the cell-PA interaction. HeLa cells were seeded one day prior to conducting the assay (20,000/well) in an 8-well IBIDI slide. 2 µL of the preformed peptide fibrils (1 mg mL^{-1}) were diluted with 2 µL Proteostat (Enzo Life Science, 1 µL stock in 99 µL PBS) and incubated for 10 min and further diluted with medium to receive a final peptide concentration of 10 µg mL^{-1}. The nucleus of the HeLa cells was stained with Hoechst 33342 (NucBlue™, Thermo Fisher Scientific). The peptide solution mixture was transferred to the HeLa cells and incubated for 30 min at 37°C before washing three times with PBS. The interaction of fibril clusters with cells was monitored after 30 min incubation time on a Stellaris 8 confocal laser scanning microscope (Leica) equipped with a 20x air objective and laser excitation wavelengths of 405 nm (Hoechst), and 561 nm (Proteostat). For the evaluation of aggregate degradation, the samples were further incubated at 37°C for 3 d and analyzed without any additional washing step. The evaluation of aggregate size and number was conducted by analyzing Proteostat stained aggregates via the 3D objects counter function of ImageJ 1.53f51 for 3 individual measurements (area 1008 µm × 1008 µm). The size distribution shows aggregates sizes larger than the threshold 10 μm^2 in an area of 1008 µm × 1008 µm. The number of aggregates were determined by counting aggregates larger than the threshold 10 μm^2 in an area of 1008 µm × 1008 µm and repeated in three technical triplicates.

Control experiments degradation analysis: To check changes of fibril aggregation without cells fluorescence microscopy measurements (Leica DMi8 microscope, 10x air objective, Leica MC170 HD camera, λem = 527/30 nm, λex = 480/40 nm) were performed analogously to the cell-experiments with PA1-3. Briefly, 4 µL of the preformed peptide fibrils (1 mg mL^{-1}) were diluted with 4 µL Proteostat (Enzo Life Science, 1 µL stock in 999 µL PBS) and incubated for 10 min and further diluted with medium to receive a final peptide concentration of 20 µg mL^{-1}. The fibril aggregates were monitored after 30

min incubation at 37°C and for the evaluation of aggregate degradation, the samples were further incubated at 37°C for 3 d and analyzed without any additional washing step. The evaluation of aggregate size and number was conducted by analyzing Proteostat stained aggregates via the 3D objects counter function of ImageJ 1.53f51 for 3 individual measurements (area 1331.2 µm × 1331.2 µm). The evaluation of aggregate size and number was conducted by analyzing Proteostat stained aggregates via the 3D objects counter function of ImageJ 1.53f51 for at least three individual measurements (area 1331.2 µm × 1331.2 µm) of three individually repeated experiments.

Statistics: Statistical analysis were performed using GraphPad Prism version 9.5.1. P-values for statistical significance were determined by two-way ANOVA with Dunnett's multiple comparison test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, ** $p \leq 0.001$, NS not significant.

Conflicts of Interest

J.M. and L.R. are inventors of granted and filed patents applying peptide nanofibrils as promoters of retroviral gene transfer. T.W., J.M., C.V.S., J.H., L.R. and K.K., filed a patent application for the usage of peptide amphiphiles as degradable enhancers.

Author Contributions

K.Kaygisiz conceived the idea, conceptualized, designed, and carried out the experiments, analyzed data, performed the formal analysis, and wrote the original draft. L.R.W. designed, and carried out the experiments related to virus transduction and cell-viability, analyzed data, performed the formal analysis, and revised the manuscript. A.I. performed molecular dynamics simulation, analyzed data and revised the manuscript. J.H. analyzed data and carried out experiments related to peptide synthesis. K.Kremer, J.M., C.V.S and T.W conceptualized and supervised, provided the funding resources, and revised the manuscript. The final manuscript was approved by all authors.

Acknowledgements

The study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 316249678—SFB 1279 (A02, A03, A05, C01). The authors thank Prof. Samuel Stupp for inspiring discussions on the properties of PAs that sparked the idea for this project.

Supporting Information

Supporting Information is available from online or from the authors.

References

[1] K. A. High, M. G. Roncarolo, *N. Engl. J. Med.* **2019**, *381*, 455.

[2] F. Goker, L. Larsson, M. Del Fabbro, F. Asa'ad, *Int. J. Mol. Sci.* **2019**, *20*, 3551.

[3] M.-W. Ku, P. Charneau, L. Majlessi, *Expert Rev. Vaccines* **2021**, *20*, 1571.

[4] V. F. I. Van Tendeloo, C. Van Broeckhoven, Z. N. Berneman, *Leukemia* **2001**, *15*, 545.

[5] O. W. Merten, S. Charrier, N. Laroudie, S. Fauchille, C. Dugué, C. Jenny, M. Audit, M. A. Zanta-Boussif, H. Chautard, M. Radrizzani, G. Vallanti, L. Naldini, P. Noguiez-Hellin, A. Galy, *Hum. Gene Ther.* **2011**, *22*, 343.

[6] A. V. Joglekar, S. Sandoval, *Hum. Gene Ther. Methods* **2017**, *28*, 291.

[7] M. A. Morgan, M. Galla, M. Grez, B. Fehse, A. Schambach, *Gene Ther.* **2021**, *28*, 494.

[8] S. Froelich, A. Tai, P. Wang, *Immunopharmacol. Immunotoxicol.* **2010**, *32*, 208.

[9] E. Martínez-Molina, C. Chocarro-Wrona, D. Martínez-Moreno, J. A. Marchal, H. Boulaiz, *Pharmaceutics* **2020**, *12*, 1051.

[10] D. B. Kohn, *Curr. Opin. Biotechnol.* **2019**, *60*, 39.

[11] G. Hucks, S. R. Rheingold, *Blood Cancer J.* **2019**, *9*, 10.

[12] R. C. Sterner, R. M. Sterner, *Blood Cancer J.* **2021**, *11*, 69.

[13] S. P. Kerkar, L. Sanchez-Perez, S. Yang, Z. A. Borman, P. Muranski, Y. Ji, D. Chinnasamy, A. D. M. Kaiser, C. S. Hinrichs, C. A. Klebanoff, C. D. Scott, L. Gattinoni, R. A. Morgan, S. A. Rosenberg, N. P. Restifo, *J. Immunother.* **2011**, *34*, 343.

[14] D. L. Puhl, A. R. D'Amato, R. J. Gilbert, *Brain Res. Bull.* **2019**, *150*, 216.

[15] A. S. Chuck, M. F. Clarke, B. O. Palsson, *Hum. Gene Ther.* **1996**, *7*, 1527.

[16] J. Park, S. Inwood, S. Kruthiventi, J. Jenkins, J. Shiloach, M. Betenbaugh, *Curr. Opin. Chem. Eng.* **2018**, *22*, 128.

[17] M. V. Ferreira, E. T. Cabral, A. S. Coroadinha, *Biotechnol. J.* **2021**, *16*, 2000017.

[18] M. Perez, A. Maiguy-Foinard, C. Barthélémy, B. Décaudin, P. Odou, *Pharm. Technol. Hosp. Pharm.* **2016**, *1*, 91.

[19] K. Kaygisiz, C. V. Synatschke, *Biomater. Sci.* **2020**, *8*, 6113.

[20] S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, K. Kaygisiz, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, *31*, 2009382.

[21] K. Kaygisiz, L. Rauch-Wirth, A. Dutta, X. Yu, Y. Nagata, T. Bereau, J. Münch, C. V. Synatschke, T. Weil, *Nat. Commun., accepted* **2023**.

[22] S. Kirti, K. Patel, S. Das, P. Shrimali, S. Samanta, R. Kumar, D. Chatterjee, D. Ghosh, A. Kumar, P. Tayalia, S. K. Maji, *ACS Biomater. Sci. Eng.* **2019**, *5*, 126.

[23] M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständker, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff, J. Münch, *Nat. Nanotechnol.* **2013**, *8*, 130.

[24] J. Wong, A. Brugger, A. Khare, M. Chaubal, P. Papadopoulos, B. Rabinow, J. Kipp, J. Ning, *Adv. Drug Deliv. Rev.* **2008**, *60*, 939.

[25] A. Dehsorkhi, V. Castelletto, I. W. Hamley, *J. Pept. Sci.* **2014**, *20*, 453.

[26] S. Zhang, *Nat. Biotechnol.* **2003**, *21*, 1171.

[27] J. D. Hartgerink, E. Beniash, S. I. Stupp, *Science* **2001**, *294*, 1684.

[28] J. D. Hartgerink, E. Beniash, S. I. Stupp, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5133.

[29] G. A. Silva, C. Czeisler, K. L. Niece, E. Beniash, D. A. Harrington, J. A. Kessler, S. I. Stupp, *Science* **2004**, *303*, 1352.

[30] R. N. Shah, N. A. Shah, M. M. D. R. Lim, C. Hsieh, G. Nuber, S. I. Stupp, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 3293.

[31] E. Sleep, B. D. Cosgrove, M. T. McClendon, A. T. Preslar, C. H. Chen, M. H. Sangji, C. M. R. Pérez, R. D. Haynes, T. J. Meade, H. M. Blau, S. I. Stupp, *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E7919.

[32] M. J. Webber, J. Tongers, M.-A. Renault, J. G. Roncalli, D. W. Losordo, S. I. Stupp, *Acta Biomater.* **2010**, *6*, 3.

[33] P. Palladino, V. Castelletto, A. Dehsorkhi, D. Stetsenko, I. W. Hamley, *Langmuir* **2012**, *28*, 12209.

[34] Z. Álvarez, A. N. Kolberg-Edelbrock, I. R. Sasselli, J. A. Ortega, R. Qiu, Z. Syrgiannis, P. A. Mirau, F. Chen, S. M. Chin, S. Weigand, E. Kiskinis, S. I. Stupp, *Science* **2021**, *374*, 848.

[35] C. Zhao, H. Chen, F. Wang, X. Zhang, *Colloids Surfaces B Biointerfaces* **2021**, *208*, 112040.

[36] J. Boekhoven, R. H. Zha, F. Tantakitti, E. Zhuang, R. Zandi, C. J. Newcomb, S. I. Stupp, *RSC Adv.* **2015**, *5*, 8753.

[37] N. Rodrigues de Almeida, Y. Han, J. Perez, S. Kirkpatrick, Y. Wang, M. C. Sheridan, *ACS Appl. Mater. Interfaces* **2019**, *11*, 2790.

[38] Y.-G. Kim, Y. Lee, J.-W. Jung, H.-E. Jin, *J. Nanosci. Nanotechnol.* **2020**, *20*, 5329.

[39] C. L. O'Neill, P. C. Shrimali, Z. P. Clapacs, M. A. Files, J. S. Rudra, *Acta Biomater.* **2021**, *133*, 153.

[40] Y. Li, L. L. Lock, Y. Wang, S.-H. Ou, D. Stern, A. Schön, E. Freire, X. Xu, S. Ghose, Z. J. Li, H. Cui, *Biomaterials* **2018**, *178*, 448.

[41] Y. Li, L. L. Lock, J. Mills, B. S. Ou, M. Morrow, D. Stern, H. Wang, C. F. Anderson, X. Xu, S. Ghose, Z. J. Li, H. Cui, *Nano Lett.* **2020**, *20*, 6957.

[42] N. Yi, B. Oh, H. A. Kim, M. Lee, *J. Drug Target.* **2014**, *22*, 156.

[43] H. Wang, Z. Feng, B. Xu, *Theranostics* **2019**, *9*, 3213.

[44] C. Zhang, T. Zhang, S. Jin, X. Xue, X. Yang, N. Gong, J. Zhang, P. C. Wang, J. H. Tian, J. Xing, X. J. Liang, *ACS Appl. Mater. Interfaces* **2017**, *9*, 4425.

[45] M. V. Húmpola, M. C. Rey, N. M. Carballeira, A. C. Simonetta, G. G. Tonarelli, *J. Pept. Sci.* **2017**, *23*, 45.

[46] M. J. Gómara, R. Pons, C. Herrera, P. Ziprin, I. Haro, *Bioconjug. Chem.* **2021**, *32*, 1999.

[47] A. O. Matemu, K. Nakamura, H. Kayahara, H. Murasawa, S. Katayama, S. Nakamura, *J. Food Sci.* **2011**, *76*, M299.

[48] C. Hüttl, C. Hettrich, R. Miller, B.-R. Paulke, P. Henklein, H. Rawel, F. F. Bier, *BMC Biotechnol.* **2013**, *13*, 51.

[49] Y.-G. Kim, Y. Lee, J. H. Kim, S.-Y. Chang, J.-W. Jung, W.-J. Chung, H.-E. Jin, *Nanomaterials* **2020**, *10*, 2342.

[50] X. Ding, X. Zhang, H. Chong, Y. Zhu, H. Wei, X. Wu, J. He, X. Wang, Y. He, *J. Virol.* **2017**, *91*, 831.

[51] G. D'Errico, A. M. D'Ursi, D. Marsh, *Biochemistry* **2008**, *47*, 5317.

[52] D. Missirlis, H. Khant, M. Tirrell, *Biochemistry* **2009**, *48*, 3304.

[53] D. Missirlis, T. Teesalu, M. Black, M. Tirrell, *PLoS One* **2013**, *8*, e54611.

[54] M. R. Schnorenberg, J. A. Bellairs, R. Samaeekia, H. Acar, M. V. Tirrell, J. L. LaBelle, *Materials* **2019**, *12*, 2567.

[55] A. Dehsorkhi, I. W. Hamley, J. Seitsonen, J. Ruokolainen, *Langmuir* **2013**, *29*, 6665.

[56] M. Mazza, A. Patel, R. Pons, C. Bussy, K. Kostarelos, *Faraday Discuss.* **2013**, *166*, 181.

[57] H.-W. Jun, V. Yuwono, S. E. Paramonov, J. D. Hartgerink, *Adv. Mater.* **2005**, *17*, 2612.

[58] H. Hanenberg, X. L. Xiao, D. Dilloo, K. Hashino, I. Kato, D. A. Williams, *Nat. Med.* **1996**, *2*, 876.

[59] D. Fenard, D. Ingrao, A. Seye, J. Buisset, S. Genries, S. Martin, A. Kichler, A. Galy, *Mol. Ther. Nucleic. Acids* **2013**, *2*, e90.

[60] T. Tsukamoto, S. Okada, *J. Virol. Methods* **2017**, *248*, 234.

[61] J. H. Ortony, B. Qiao, C. J. Newcomb, T. J. Keller, L. C. Palmer, E. Deiss-Yehiely, M. Olvera de la Cruz, S. Han, S. I. Stupp, *J. Am. Chem. Soc.* **2017**, *139*, 8915.

[62] E. T. Pashuck, H. Cui, S. I. Stupp, *J. Am. Chem. Soc.* **2010**, *132*, 6041.

[63] S. E. Paramonov, H. W. Jun, J. D. Hartgerink, *J. Am. Chem. Soc.* **2006**, *128*, 7291.

[64] E. Strandberg, F. Schweigardt, P. Wadhwani, J. Bürck, J. Reichert, H. L. P. Cravo, L. Burger, A. S. Ulrich, *Sci. Rep.* **2020**, *10*, 12300.

[65] W. C. Fowler, C. Deng, G. M. Griffen, T. Teodoro, A. Z. Guo, M. Zaiden, M. Gottlieb, J. J. De Pablo, M. V. Tirrell, *J. Am. Chem. Soc.* **2021**, *143*, 4440.

[66] Y. Chen, H. X. Gan, Y. W. Tong, *Macromolecules* **2015**, *48*, 2647.

[67] S. Toksoz, R. Mammadov, A. B. Tekinay, M. O. Guler, *J. Colloid Interface Sci.* **2011**, *356*, 131.

[68] M. Jackson, H. H. Mantsch, *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95.

[69] A. Adochitei, G. Drochioiu, *Rev. Roum. Chim.* **2011**, *56*, 783.

[70] S. Sur, F. Tantakitti, J. B. Matson, S. I. Stupp, *Biomater. Sci.* **2015**, *3*, 520.

[71] M. P. Hendricks, K. Sato, L. C. Palmer, S. I. Stupp, *Acc. Chem. Res.* **2017**, *50*, 2440.

[72] L. L. Lock, C. D. Reyes, P. Zhang, H. Cui, *J. Am. Chem. Soc.* **2016**, *138*, 3533.

[73] R. M. Peitzsch, S. McLaughlin, *Biochemistry* **1993**, *32*, 10436.

[74] B. F. Lin, D. Missirlis, D. V. Krogstad, M. Tirrell, *Biochemistry* **2012**, *51*, 4658.

[75] P. Zhang, L. L. Lock, A. G. Cheetham, H. Cui, *Mol. Pharm.* **2014**, *11*, 964.

[76] H. Acar, R. Samaeekia, M. R. Schnorenberg, D. K. Sasmal, J. Huang, M. V. Tirrell, J. L. LaBelle, *Bioconjug. Chem.* **2017**, *28*, 2316.

[77] C. Li, A. Iscen, H. Sai, K. Sato, N. A. Sather, S. M. Chin, Z. Álvarez, L. C. Palmer, G. C. Schatz, S. I. Stupp, *Nat. Mater.* **2020**, *19*, 900.

[78] A. Iscen, G. C. Schatz, *J. Phys. Chem. B* **2019**, *123*, 7006.

[79] L. S. Vermeer, L. Hamon, A. Schirer, M. Schoup, J. Cosette, S. Majdoul, D. Pastré, D. Stockholm, N. Holic, P. Hellwig, A. Galy, D. Fenard, B. Bechinger, *Acta Biomater.* **2017**, *64*, 259.

[80] L.-H. Han, S. Yu, T. Wang, A. W. Behn, F. Yang, *Adv. Funct. Mater.* **2013**, *23*, 346.

[81] K. D. Ratanji, J. P. Derrick, R. J. Dearman, I. Kimber, *J. Immunotoxicol.* **2014**, *11*, 99.

[82] J. Schönfelder, P. B. Pfeiffer, T. Pradhan, J. Bijzet, B. P. C. Hazenberg, S. O. Schönland, U. Hegenbart, B. Reif, C. Haupt, M. Fändrich, *Amyloid* **2021**, *28*, 243.

[83] O. S. Makin, E. Atkins, P. Sikorski, J. Johansson, L. C. Serpell, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 315.

[84] M. Trotard, N. Tsopoulidis, N. Tibroni, J. Willemsen, M. Binder, A. Ruggieri, O. T. Fackler, *J. Virol.* **2016**, *90*, 2064.

[85] R. T. Abraham, A. Weiss, *Nat. Rev. Immunol.* **2004**, *4*, 301.

[86] A. M. Ender, K. Kaygisiz, H.-J. Räder, F. J. Mayer, C. V. Synatschke, T. Weil, *ACS Biomater. Sci. Eng.* **2021**, *7*, 4798.

[87] M. Nilsson, *Methods* **2004**, *34*, 151.

[88] R. Khurana, J. R. Gillespie, A. Talapatra, L. J. Minert, C. Ionescu-Zanetti, I. Millett, A. L.

Fink, *Biochemistry* **2001**, *40*, 3525.

[89] V. A. Wagoner, M. Cheon, I. Chang, C. K. Hall, *Proteins Struct. Funct. Bioinforma.* **2014**, *82*, 1469.

[90] Y. Yoshimura, Y. Lin, H. Yagi, Y. H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki, Y. Goto, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14446.

[91] D. Missirlis, M. Farine, M. Kastantin, B. Ananthanarayanan, T. Neumann, M. Tirrell, *Bioconjug. Chem.* **2010**, *21*, 465.

[92] A. Papkalla, J. Münch, C. Otto, F. Kirchhoff, *J. Virol.* **2002**, *76*, 8455.

Supporting Information

Supporting Information for

Peptide amphiphiles as biodegradable adjuvants for efficient retroviral gene delivery

Kübra Kaygisiz[1], Lena Rauch-Wirth[2], Aysenur Iscen[3], Jan Hartenfels[1], Kurt Kremer[3], Jan Münch[2]*, Christopher V. Synatschke[1]*, Tanja Weil[1]*

1 Department Synthesis of Macromolecules, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

2 Institute of Molecular Virology, Ulm University Medical Center, Meyerhofstraße 1, 89081 Ulm, Germany

3 Polymer Theory Department, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

Correspondence to be addressed to: Jan.Muench@uni-ulm.de, synatschke@mpip-mainz.mpg.de; weil@mpip-mainz.mpg.de

Content

1 Purity and Characterization of PA1-3 and P1-3 3
2 Aggregation of PA1-3 and P2 fibrils in ultrapure water and phosphate buffer 4
3 ATR FT-IR analysis of PA1-PA3, P1-P3 5
4 ThT fluorescence of PA1-3 and P1-3 7
5 Cell-PA interaction 8
6 D-Amino acids substitution of PA1 13
7 Effects of PA modifications on properties and activity 14
7.1 Effects of hydrophobic moiety exchange on properties and activity 15
7.2 Effects of linker exchange on properties and activity 19
7.3 Effects of sequence exchange on properties and activity 23
8 Cell-Viability of PA library 31
9 Proteostat Assay of PA library 32
10 TEM of PA library 33
11 ATR-IR of PA library 35
12 Chemical structures of PA library 37
13 Computational Methods 41
14 References 47

1 Purity and Characterization of PA1-3 and P1-3

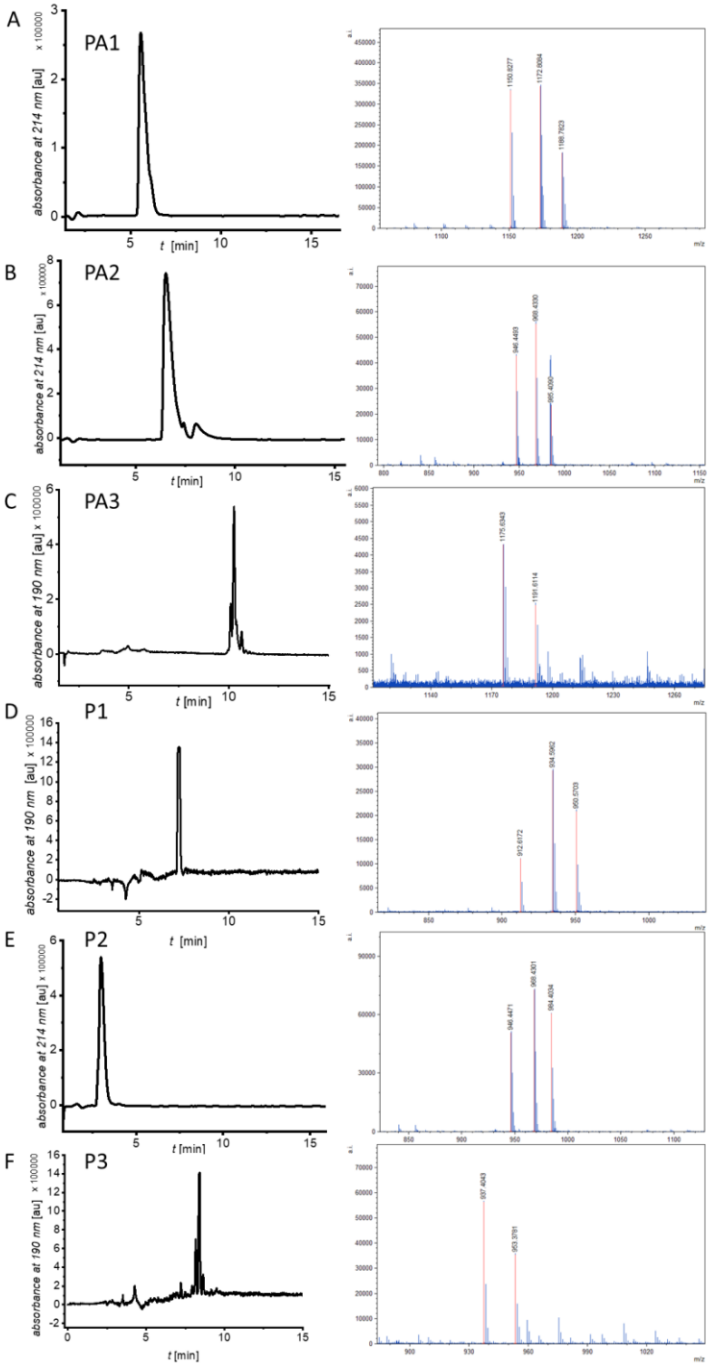


Figure S1 Mass determination and characterization of PA1-3 and P1-3. High performance liquid chromatography traces (left) and mass spectrometry (right). **A** PA1, M = $C_{58}H_{111}N_{13}O_{10}$, LC-MS (m/z) calc. for $[M+H]^+$ 1150.87; found: 1150.95, purity >95%. MALDI-ToF MS (m/z) calc. for $[M+H]^+$, 1150.8650; found: 1150.8277. **B** PA2, M = $C_{63}H_{97}N_{11}O_9S$, LC-MS (m/z) calc. for $[M+H]^+$ 1184.73; found: 1184.62, purity >90%. MALDI-ToF MS (m/z) calc. for $[M+H]^+$, 1184.7264; found: 1184.6454. **C** PA3, M = $C_{58}H_{111}N_{13}O_{10}$, HPLC purity >90%. MALDI-ToF MS (m/z) calc. for $[M+Na]^+$, 1150.6903; found: 1175.6343. **D** P1, M = $C_{42}H_{81}N_{13}O_9$, HPLC purity >95% and MALDI-ToF MS (m/z) calc. for $[M+H]^+$, 912.6353; found: 912.6172. **E** P2, M = $C_{47}H_{67}N_{11}O_8S$, LC-MS (m/z) calc. for $[M+H]^+$ 946.50; found: 946.59, purity >95%. MALDI-ToF MS (m/z) calc. for $[M+H]^+$, 946.4973; found: 946.4471. **F** P3, M = $C_{39}H_{66}N_{10}O_{15}$, HPLC purity >70%. MALDI-ToF MS (m/z) calc. for $[M+Na]^+$, 937.4607; found: 937.4043.

2 Aggregation of PA1-3 and P2 fibrils in ultrapure water and phosphate buffer

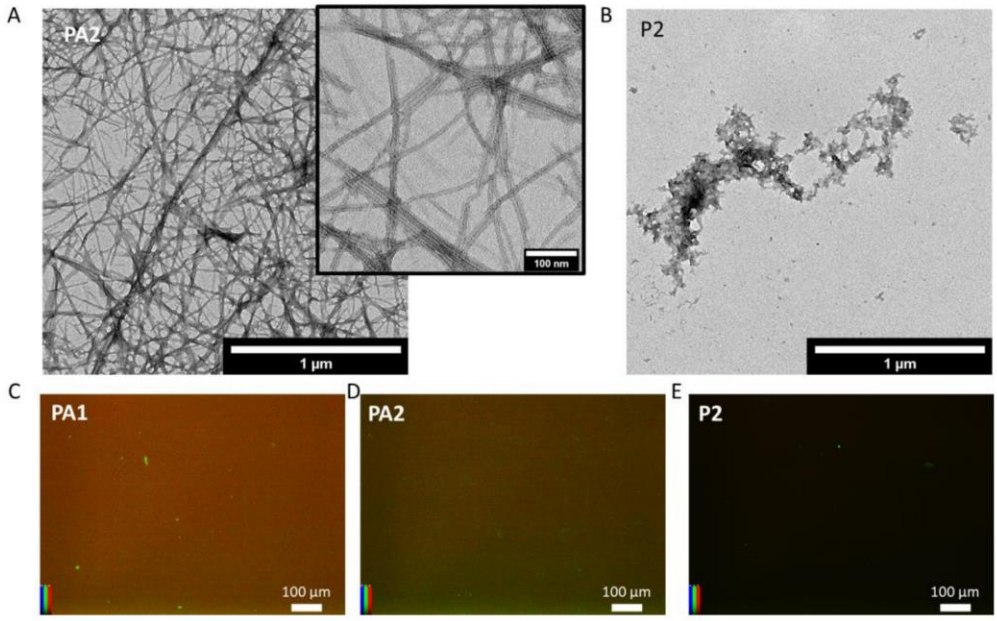


Figure S2 Characterization of PA2 and P2 incubated 1 d in pure MilliQ water. **A** TEM image of PA2, scale bar 1 µm, detail 100 nm. **B** TEM image of P2, scale bar 1 µm. **C, D, E** Fluorescence microscopy images of PA1, PA2 and P2, respectively. Samples were diluted from 1 mg mL^{-1} to 0.1 mg mL^{-1} with ThT-solution in MilliQ, final concentration 45 µM, contrast setting 0-150 for all images, scale bar 100 µm.

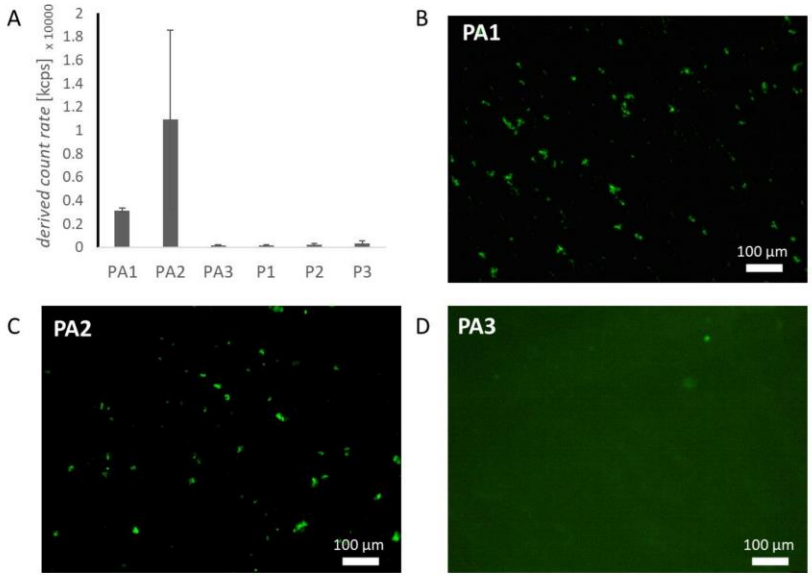


Figure S3 Formation of µm-sized aggregates of PA1 and PA2 in PBS. **A** Scattering intensity of light (633 nm) during zeta-potential measurements for PA1-3 and P1-3. **B, C, D** Fluorescence microscopy images of PA1, PA2 and PA3, respectively. Samples were diluted from 1 mg mL^{-1} to 0.1 mg mL^{-1} with ThT-solution in PBS, final concentration 45 µM, scale bar 100 µm.

3 ATR FT-IR analysis of PA1-PA3, P1-P3

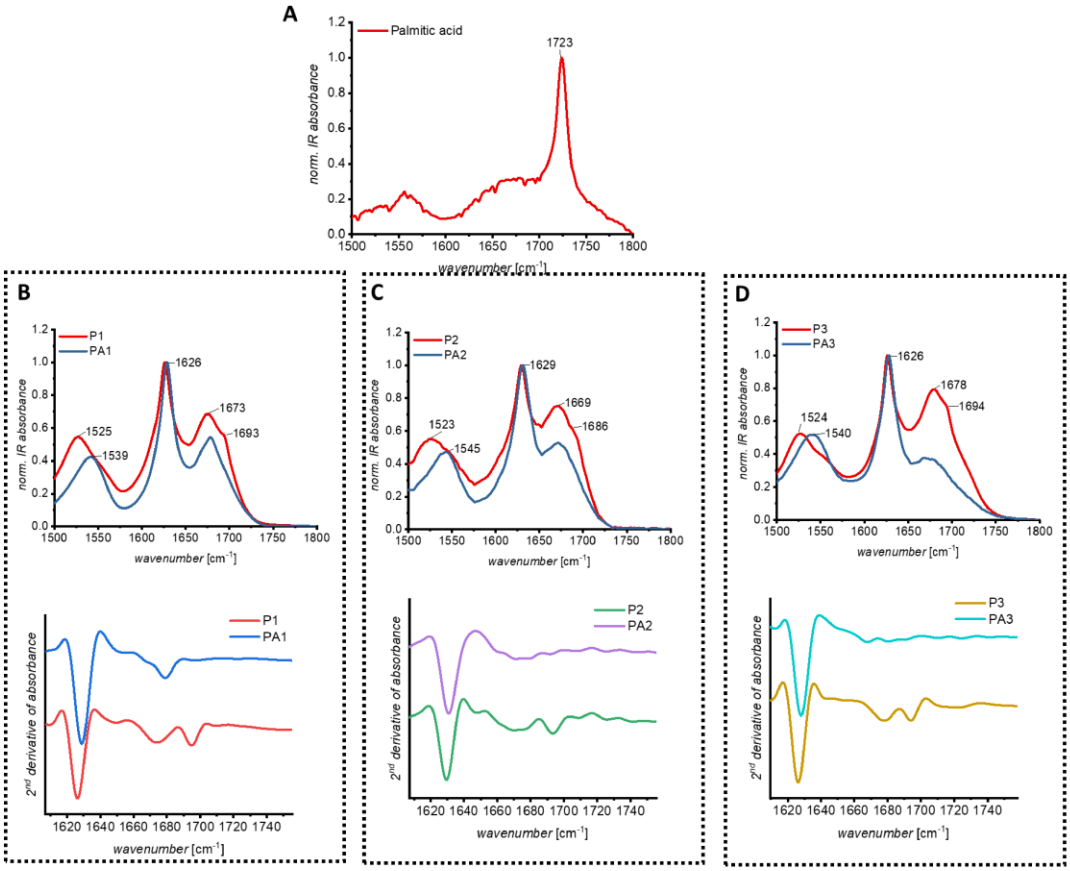


Figure S4 ATR-IR measurements. **A** ATR-IR spectrum of palmitic acid. **B-D Top** ATR-IR spectra of the amide I and II regions of PA1–PA3 and P1–P3 (1 mg mL^{-1} in PBS lyophilized). Chemically similar peptides are displayed in one diagram: PA1 and P1 (**B**), PA2 and P2 (**C**), PA3 and P3 (**D**). **B-D Bottom** 2nd derivation of IR absorbance of PA1 (C_{16}-VVVAAAKKK-NH_2) and P1 (VVVAAAKKK-NH_2) **(B),** 2nd derivation of IR absorbance of PA2 (C_{16}-FQFKFKC -NH_2) (**C**), P2 (FQFKFKC-NH_2) and 2nd derivation of IR absorbance of PA3 (C_{16}-VVVAAAEEE-NH_2) and P3 (VVVAAAEEE-NH_2) (**D**).

All peptides P1–P3 and PA1–PA3 show a main peak around 1630 cm^{-1} which is assigned to parallel β-sheets (**Figure S4B-D**). Another intense, broad peak in the amide I region was found at approx. 1660-80 cm^{-1} indicating β-turn structural elements.[1,2] A small peak at 1694 cm^{-1} can be assigned to antiparallel β-sheets and was found only for P1–P3, but not in the PAs.[3,4] The amide II peak maximum for P1–P3 is located at approx. 1520 cm^{-1}, whereas for PA1–PA3, it occurs at approx. 1540 cm^{-1}. The changes in the amide I and amide II region indicate the transformation of amorphous aggregates (P1, P3) to fibrillar assemblies (PA1, PA3) in accordance with TEM measurements.[5] The decrease of the peak at 1694 cm^{-1} and the shift of the amide II band to higher wavenumbers for PA1–PA3 can be assigned to an increase of parallel β-sheet structures and decrease of antiparallel β-sheet. Similar peak shifts are reported for observations regarding the transition of amorphous aggregates to fibrillar assemblies of Aβ(1–40)[6–8] and lysozyme[9] for time-dependent assembly processes. Further, amorphous aggregates show less β-sheet content compared to fibrillar assemblies.[6,10] The relative amount of β-sheet secondary structural elements was quantified by integration of the IR absorption bands in the amide I region.[11] Since amino acid side chains can also contribute to the IR absorption in the amide I region, each peptide was only compared to its fatty acid conjugate. The palmitic acid

residue does not contribute to the IR absorbance in the amide I region (**Figure S4**). A quantitative analysis of the amide I region was executed by integration of the 2nd derivative peak area of the β-sheet (1630 cm^{-1}) and non-ordered or other structural elements (1640–1690 cm^{-1}, **Figure S4**).[11,12]

4 ThT fluorescence of PA1-3 and P1-3

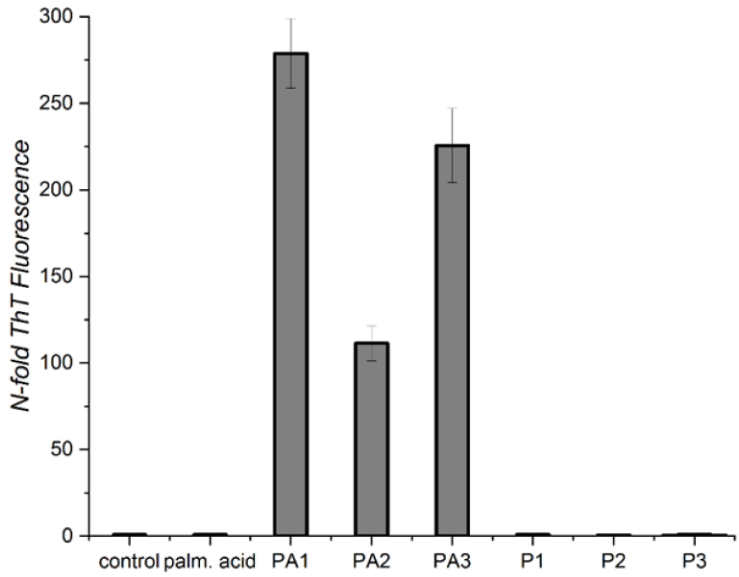


Figure S5 N-fold Thioflavin-T fluorescence intensity at 488 nm excitation for palmitic acid, PA1-3 and P1-3 relative to PBS (control).

5 Cell-PA interaction

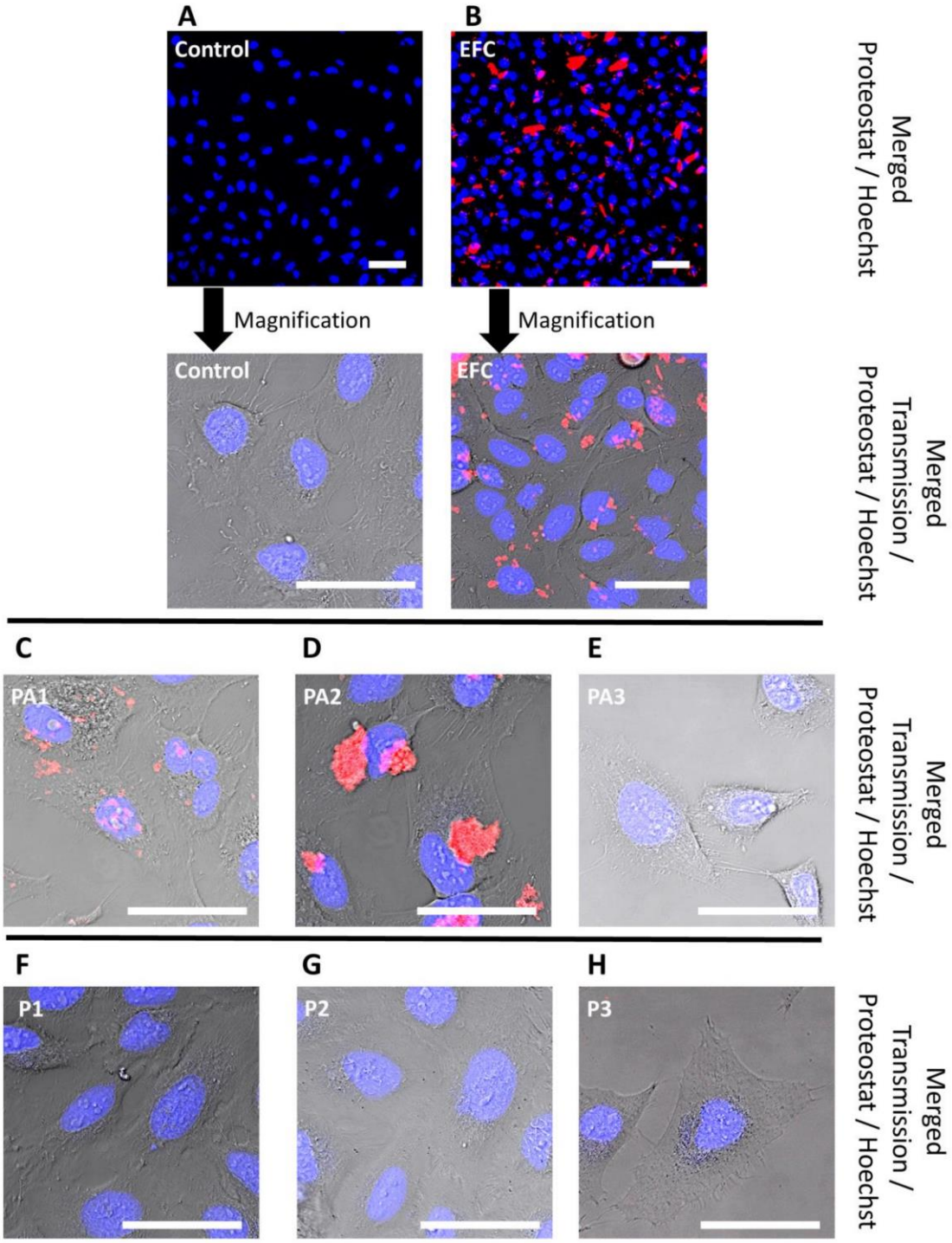


Figure S6 Laser scanning microscopy micrographs of **A** cells only (control), and Proteostat-stained (red) peptides **B** EF-C (QCKIKQIINMWQ), **C** PA1, **D** PA2, **E** PA3, **F** P1, **G** P2, **H** P3 added to HeLa cells (nucleus in blue). The peptides (20 µg mL^{-1}) were added to cells for 30 min, washed and analyzed. EF-C, PA1, PA2 revealed association with cellular membranes while PA3, P1-3 does not associate with cells, scale bar 50 µm.

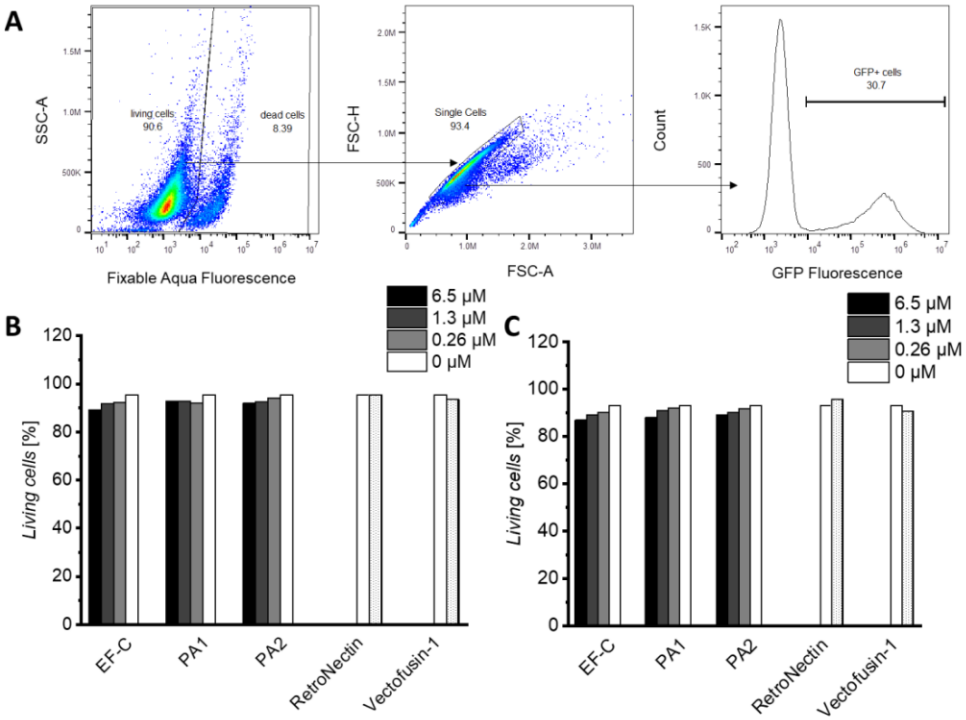


Figure S7 Viability of CD4+ T cells. **A** Gating strategy for analysis of living GFP+ cells. Cells were stained with LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit. Dead cells were initially eliminated by gating on viability dye-negative cells (SSC-A vs. Fixable Aqua Fluorescence) and afterwards for singlets (FSC-H vs. FSC-A). GFP+ cells were measured using the FITC channel. **B** Viability of CD4+ T cells after transduction with GALV-RV and treatment with different concentrations of transduction enhancing peptides EF-C, PA1 and PA2 (0, 0.26, 1.3, 6.5 µM) and 10 µg mL^{-1} (Vectofusin-1), 20 µg mL^{-1} (RetroNectin) after 3 days. **C** Control experiments with non-transduced CD4+ T cells to determine viability after treatment with transduction enhancers for 3 days.

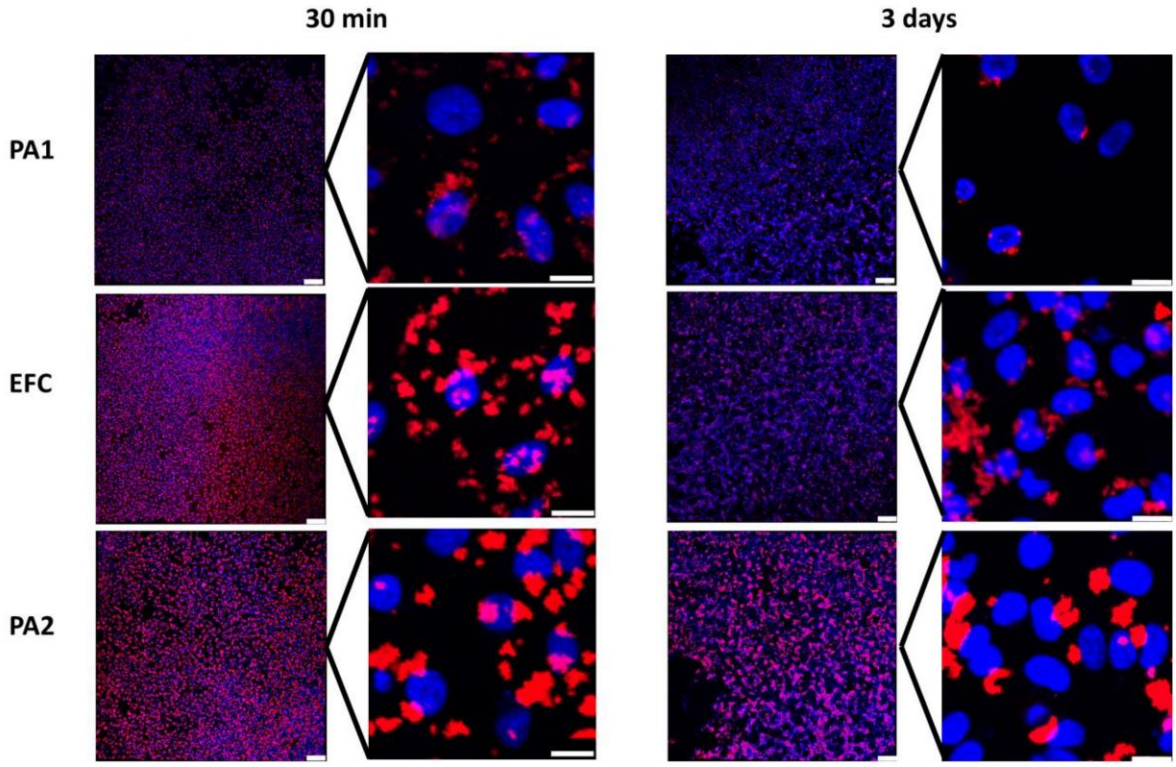


Figure S8 Degradability of µm-sized fibrillar aggregates formed by PA1 and PA2 after 30 min and 3 d in cell culture shown via laser scanning microscopy images of Proteostat-labeled EF-C, PA1 and PA2 (red) incubated with HeLa cells (nucleus in blue). In general, PA1 reveals smaller µm-sized aggregates than PA2. After 3 days in cell culture, almost all small aggregates of PA1 are degraded, whereas aggregates formed by PA2 stay almost unchanged, scale bar 200 µm for overview (left) and 20 µm for detail (right).

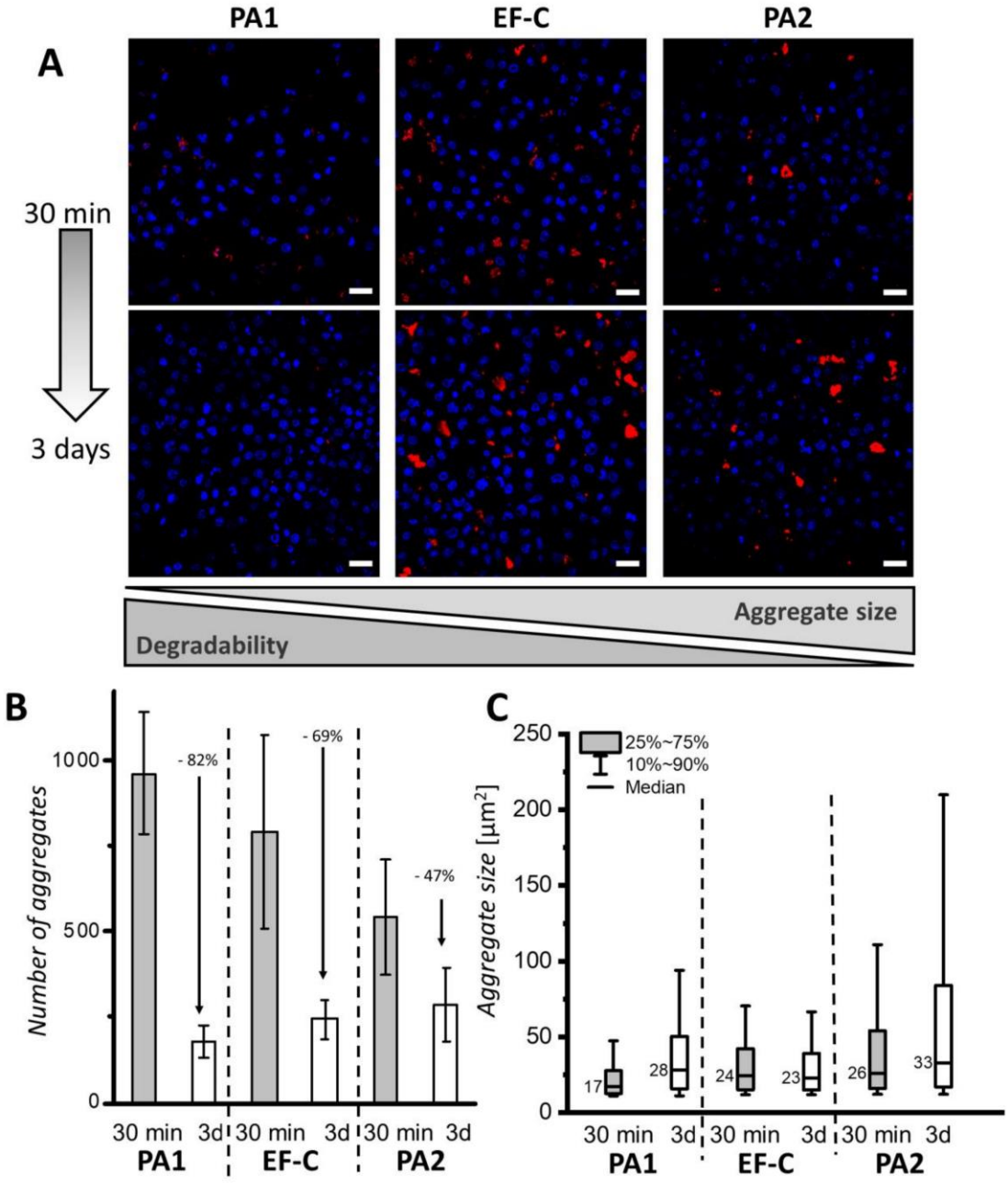


Figure S9 Degradability of µm-sized fibrillar aggregates formed by EF-C, PA1 and PA2 after 30 min and 3 d in CD4+ T cell culture. **A** Representative laser scanning microscopy images of Proteostat-labeled EF-C, PA1 and PA2 (red) incubated with CD4+ T cells (nucleus in blue). In general, PA1 reveals smaller µm-sized aggregates than PA2 and EF-C. After 3 days, almost all small aggregates of PA1 are degraded, scale bar 20 µm. **B** Bar plot showing number of aggregates > 10 µm^2 in an area of 708.5 µm × 708.5 µm in cell culture after 30 min and 3 d incubation (1 µg peptide added to 250,000 cells), error bars indicate standard deviations from five technical replicates. **C** Box plot showing size distribution and in an area of 3543 µm × 3543 µm in cell culture after 30 min and 3 d incubation (1 µg peptide added to 250,000 cells).

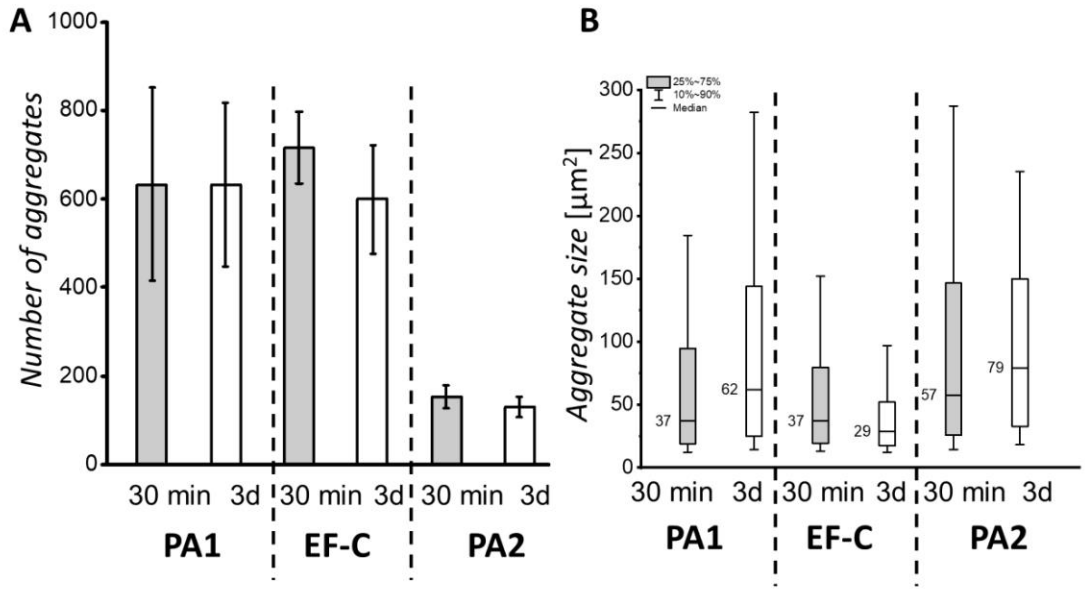


Figure S10 Stability test of Proteostat stained fibrils without the presence of cells. **A** Bar plot showing number of aggregates > 10 µm^2 in an area of 1331.2 µm × 1331.2 µm at 0.1 mg mL^{-1} incubated at 37°C, 250 rpm in PBS after 30 min and 3 d incubation, error bars indicate standard deviations from five technical replicates. **B** Box plot show size of peptide fibrils > 10µm^2 in 3991.65 µm × 3991.65 µm at 0.1 mg mL^{-1} incubated at 37°C, 250 rpm in PBS. Aggregate analysis was conducted via fluorescence microscopy by cumulating triplicate measurements. There is no significant change in aggregate size and number over the course of 3 days. Under these conditions PA2 forms less aggregates, which are larger in size compared to PA1 and EF-C.

6 D-Amino acids substitution of PA1

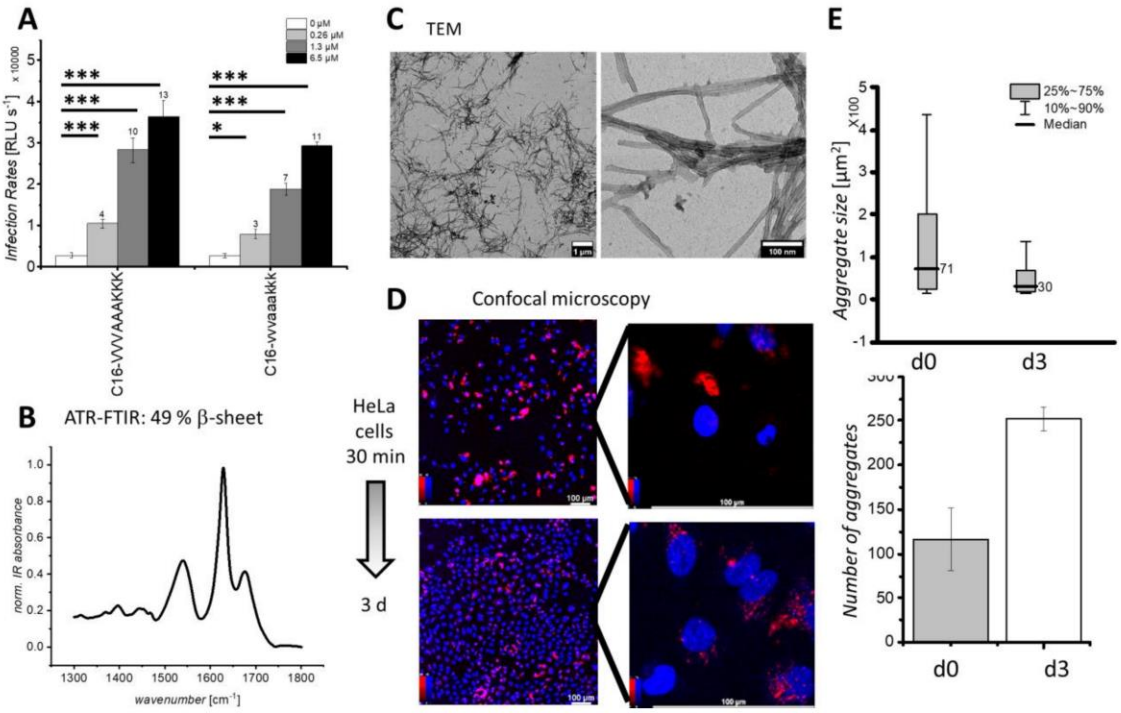
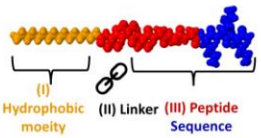


Figure S11 A Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of PA1 and pa1, the D-amino acid pendant of PA1 (6.5 µM, 1.3 µM, 0.26 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. P-values for statistical significance were determined by two-way ANOVA test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant. RLU s^{-1} are relative light units per second. **B** The β-sheet content of pa1 is 49% and is calculated from the FT-IR spectra. **C** pa1 assembling into fibrils as determined by TEM, scale bars 1 µm (left) and 100 nm (right). **D** Representative confocal laser scanning microscopy images of Proteostat-labeled pa1 (red) incubated with HeLa cells (nucleus in blue), scale bar 100 µm. **E** Bar plot showing number of aggregates > 10 μm^2 and box plot showing size distribution in an area of 1008 µm × 1008 µm in cell culture after 30 min and 3 d incubation (2 µg peptide added to 20,000 cells), deviations indicate three technical replicates. In general, pa1 reveals larger µm-sized aggregates than PA1. After 3 days in cell-culture, aggregates of pa1 are incompletely degraded into smaller aggregates (>10 μm^2), which causes an increase of number of aggregates.

C_{16}-vvvaaakkk-NH_2 (pa1) is analogous to PA1 but contains only D-amino acids, which cannot be degraded from cells (**Figure S11**). As expected C_{16}-vvvaaakkk-NH_2 is not degraded and forms large aggregates (71 μm^2) which are further decreasing in size (71 μm^2 at d0 → 30 μm^2 at d3) but not in number (increase in 116%), which indicates that aggregates are not degraded but rather breaking up to smaller parts.

7 Effects of PA modifications on properties and activity

To investigate the effects of modification in the PA sequence, we created various peptide amphiphiles derived from PA1 by varying the (I) hydrophobic moiety, (II) the linkage chemistry and (III) the peptide sequence (**Figure S12, SI chapter 7.1 – 7.3**). The color code for sequence representation is yellow for the hydrophobic moiety, blue for cationic amino acids, red for non-polar amino acids, purple for polar uncharged amino acids, and green for anionic amino acids. To be concise, only the summarized information will be discussed, all original data for the peptides are summarized in the (**Figure S26 - Figure S32**). The tested peptides were biocompatible and not cytotoxic at the tested concentrations on TZM-bl cells (**Figure S25**).



Original sequence		N-fold infectivity at 1.3 µM	Degradability / %
PA1	C_{16} VVVAAAKKK-NH_2	10±3	94%

(I)

Hydrophobic moeity	Replace C16	N-fold infectivity at 1.3 µM	Degradability / %
Fatty acid length	C12	6±1	-32%
	C9	2±0	78%
Aromatic residues	Fmoc	5±1	28%
	Nap	1±1	ND
Unsaturated fatty acids	Eicosapentaene acid	12±1	99%
	Alphalinoleinic acid	1±0	ND

(II)

Linker	Replace Amide:	N-fold infectivity at 1.3 µM	Degradability / %
Functional group	Ester	17±5	31%
	Thioester	11±4	87%
	Disulfide	4±2	ND
Enzyme responsive sequence	GFLG (Cathespin B)	11±4	4%
	GPLGV (MMP)	8±3	28%

(III)

Peptide sequence	Replace VVVAAAKKK-NH_2	N-fold infectivity at 1.3 µM	Degradability /%
PA1 inspired	VVVAAAEEE-NH_2	1±0	ND
	AVAKKK-NH_2	2±0	94%
	VVAAHHH-NH_2	2±0	26%
	GGGSSKKK-NH_2	2±1	ND
	AGAGRRR-NH_2	1±0	ND
	AAVVHHK-NH_2	6±3	-123%
	GGVVRR-NH_2	1±1	ND
β-sheet prone sequences	FQFKFKC-NH_2	9±1	53%
	IIQKIK-NH_2	4±1	28%
	AQAKAK-NH_2	4±1	-83%
	LLLKK-NH_2	3±0	27%
	KGVPGVGK-NH_2	1±0	ND
Endogeneous sequences	AGGRVK-NH_2	3±1	ND
	LNGGVKK-NH_2	6±1	48%
	ALAAGKK-NH_2	5±1	87%
	WEALKK-NH_2	1±0	ND
	RRWQWR-NH_2	0±0	ND
	KYIHQNYTKALAGKLV-NH_2	5±2	-110%
	NFYLQVNKINK-NH_2	5±1	-123%
	RYYASLRHYLNLVTRQRY-NH_2	5±1	43%
	GKVGSK -NH_2	2±1	ND
	AKAVGK-NH_2	6±3	89%
	KIISFK-NH_2	2±1	-84%

Figure S12 Summarized overview of peptide amphiphile library in which systematic variations the original sequence PA1 (C_{16}-VVVAAAKKK-NH_2) were conducted to study the effects of the (I) hydrophobic moiety, (II) the linkage chemistry and (III) the peptide sequence on infectivity enhancement and degradability (see **SI chapter 7.1 – 7.3**).

In the following section a compound is considered infectivity enhancing if n-fold infectivity at 1.3 µM is larger than 2. A compound is considered to form µm-sized aggregates if visible particles were observed in widefield microscopy (brightfield or ThT-staining filter, objective 10x) or confocal microscopy (stained with proteostat, objective 20 x) and samples showed a derived count rate larger than 1000 kcps in Zeta-Potential measurements. A compound is considered ThT active if a ThT-fluorescence larger than 2 relative to a control samples (PBS with 10% DMSO) is observed.

For the degradability only the PAs which could be stained with Proteostat (**Figure S26**) and were visible in microscopy (**Figure S15**, **Figure S18**, **Figure S21**, and **Figure S22**) were studied in detail. To determine

whether a PA is degraded or not, Proteostat stained PAs which are added to Hela cells (2 µg peptide added to 20.000 cells) and studied via confocal laser scanning microscopy. Microscopy is conducted at day 0 (d0, after 30 min of incubation) and at day 3 (d3, after 72 h of incubation). The number of aggregates >10 μm^2 is counted via 3D object counter in ImageJ in an area of 1008 µm × 1008 µm.

A PA is considered as degradable if the number and the size of aggregates are reduced after 3 days. The efficiency is determined by the percentual degradation (aggregate number at d3 / aggregate number at d0). Three technical replicates were evaluated for each sample. A large deviation in number and size of aggregates represents a nonhomogeneous distribution and polymorphism of aggregates.

In the following, we will distinguish between three aggregation types for fibrillar structures: Fibrils which were not aggregating (below resolution limit of the microscope, approx. 650 nm), moderately aggregating (< 50 μm^2) and aggregating to large particles (> 50 μm^2).

7.1 Effects of hydrophobic moiety exchange on properties and activity

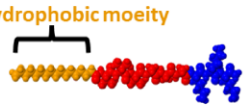
To investigate the effects of varying hydrophobic moieties on assembly, aggregation and activity, we attached different lengths of fatty acids, various unsaturated fatty acids and different size of aromatic groups to the sequence VVVAAAKKK-NH_2 (**Figure S29**).

Combining the VVVAAAKKK-NH_2 peptide sequence with different hydrophobic residues can maintain infectivity enhancement, if the resulting structures are able to form fibrils, aggregate into µm-sized particles and show a positive zeta-potential (**Figure S13**). The hydrophobic residues naphthalene (Nap), pelargonic acid (C_9) and alpha linolenic acid are not able to maintain activity because they either don't form fibrils (alpha linolenic acid) or do not aggregate (Nap, C_9, alpha linolenic acid) or show almost neutral zeta-potential (nap, alpha linolenic acid). We find a minimum alkyl chain length of lauric acid (C_{12}) necessary to observe efficient infectivity enhancement. However, modifications with C_{12} result in large aggregate formation, which even increases the number of aggregates after 3 days, because the aggregates are partially degraded – most likely due to their initial size. C_9 modified PA is reduced in number of aggregates, but increases in size after 3 days in presence of HeLa cells, which indicates that beside degradation also aggregation into larger clusters takes place.

Combining VVVAAAKKK-NH_2 with long, polyunsaturated fatty acid residues like eicosapentaenoic acids increases infectivity enhancement strongly. Eicosapentaenoic acid-VVVAAAKKK-NH_2 forms spherical structures composed of fibrillar structures (**Figure S27**), which aggregate into a high number of equally distributed small aggregates with average size of 6-7 μm^2. The infectivity enhancement of eicosapentaenoic acid-VVVAAAKKK-NH_2 is similar to EF-C or PA1, which makes it a clinically interesting degradable PA. Interestingly, alpha linolenic acid is a fatty acid, which has 2 less unsaturated sites than eicosapentaenoic acid, and is neither forming fibrils, nor it is aggregating or active (**Figure S13**). In contrast the active PA Fmoc-VVVAAAKKK is forming fibrils leading to very large µm-sized aggregates, which cannot be degraded by cells (**Figure S14**).

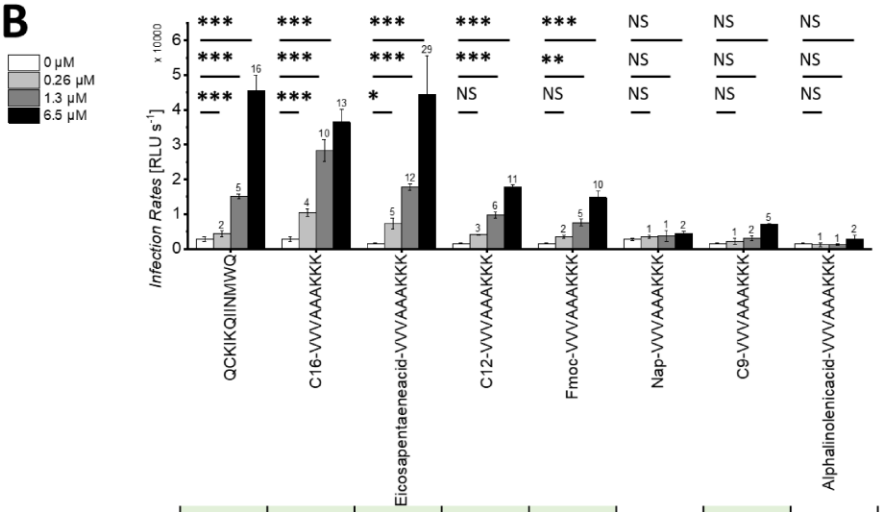
We therefore conclude that the capacity to form fibrils, which can aggregate can be traced back to the flexibility of the hydrophobic moiety, *i.e.* the increased flexibility for the less unsaturated alpha linolenic acid is limiting self-assembly and aggregation.

A



Sequence	Infectivity enhancing n-fold at 1.3 µM	Fibril Formation	Degradability / % (rel. to initial particles > 10 µm²)
C_{16}-VVVAAAKKK-NH_2	10±3	Yes	94%
C_{12}-VVVAAAKKK-NH_2	6±1	Yes	-32%
C_9-VVVAAAKKK-NH_2	2±0	Yes	78%
alpha linolenic acid-VVVAAAKKK-NH_2	1±0	No	ND
eicosapentaene acid -VVVAAAKKK-NH_2	12±1	Yes	99%
Fmoc-VVVAAAKKK-NH_2	5±1	Yes	28%
Nap-VVVAAAKKK-NH_2	1±1	Yes	ND

B



Zeta-Potential / mV	6.6	22.2	16.9	19.4	15.2	1.2	13.5	4.4
Count Rate / kcps	2602.3	3138.4	6069.8	3558.1	509.5	9.3	77.2	255.7
N-fold ThT-fluorescence	16.2	278.7	61.9	185.7	269.9	1.3	37.1	1.8
Fibril Formation	1	1	1	1	1	1	1	0
β-sheet / %	52	57	51	49	52	46	49	37

Figure S13 A Summary of sequence, infectivity enhancement at 1.3 µM concentration relative to virus only (n-fold), assembly into fibrils determined by TEM, **Figure S27** and degradability determined via confocal microscopy aggregate analysis of particles > 10 µm². Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids. **B** Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of peptides EF-C and newly designed PAs (6.5 µM, 1.3 µM, 0.26 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. P-values for statistical significance were determined by two-way ANOVA test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant. RLU s^{-1} are relative light units per second. The table summarizes the respective physicochemical properties of the PAs. PAs assembling into fibrils as determined by TEM (**Figure S27**) are indicated with "1". The β-sheet content is calculated from the FT-IR spectra (**Figure S28**). Properties which were previously found to favor infectivity are colored green.

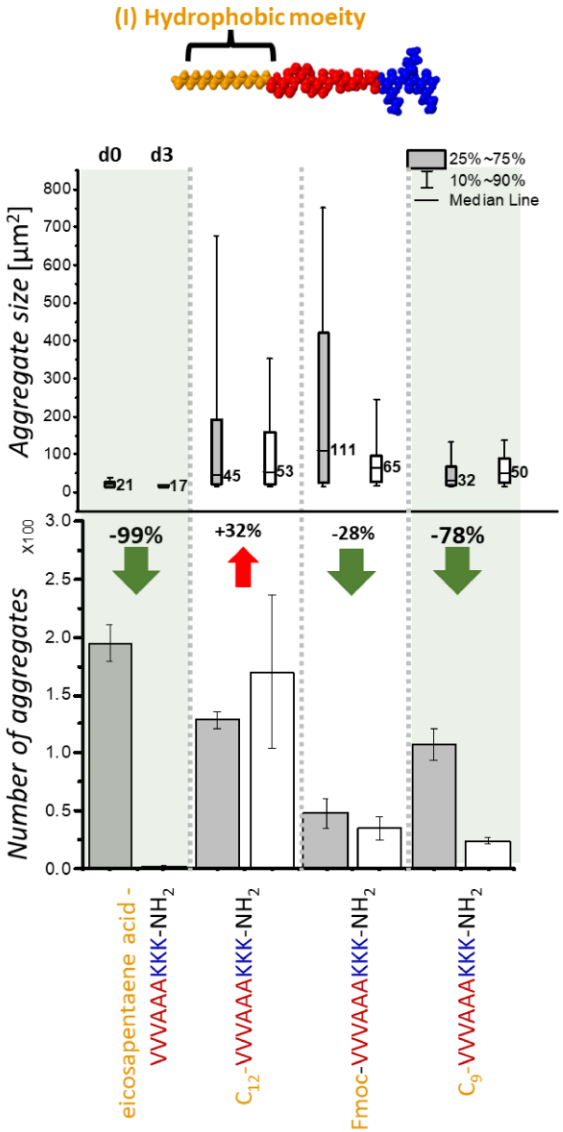


Figure S14 Degradability of PAs investigated via the aggregate size and number of aggregates larger than 10 µm² in an area of 1008 µm × 1008 µm. Error bars in aggregate number depict technical triplicate measurements. The percentual degradability was determined by number of aggregates after 3 days (d3, white boxes and bars) relative to the initial number of aggregates after 30 min incubation (d0 black boxes and bars). Box plots show aggregate size distribution with 10-90% shown as whiskers, 25-75% as boxes and the median line. The arrows give a qualitative evaluation of size reduce (green) or increase (red), number reduce (green) or number increase (red). Green highlighted section are pointing out best performing candidates for degradability.

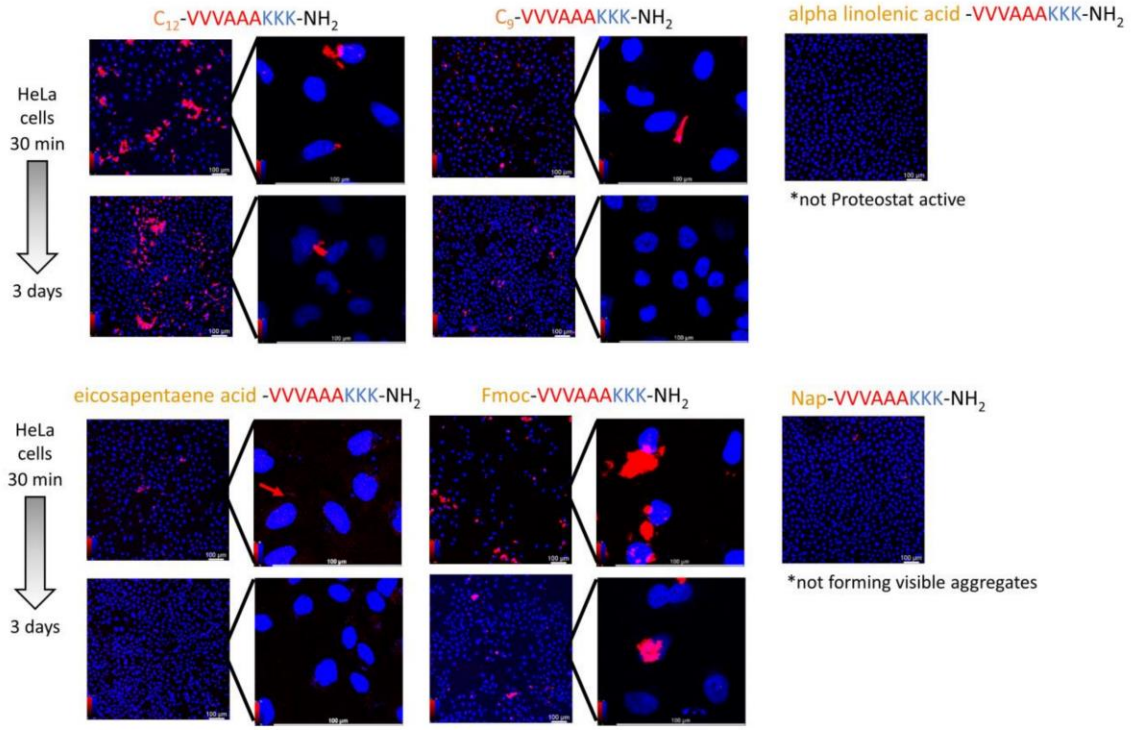


Figure S15 Representative confocal laser scanning microscopy images of Proteostat-labeled PAs (red) incubated with HeLa cells (nucleus in blue) after 30 minutes of incubation (top rows) and 3 days of incubation (bottom rows), scale bar 100 µm.

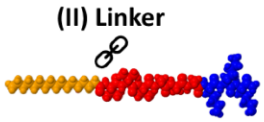
7.2 Effects of linker exchange on properties and activity

The linker is the functional group connecting the hydrophobic moiety with the peptide sequence. In the original sequence the linker is an amide bond. To investigate the effects of the linker on infectivity enhancement and degradability, we tested ester, thioester, disulfide and enzyme degradable peptide sequences GPLGV[13] (matrix metalloproteinase sensitive) and GFLG[14] (cathepsin B sensitive) (**Figure S30**) to investigate potential chemical and enzymatic degradation pathways.

Changing the linker does not significantly affect the infectivity enhancement for all tested functional groups except in the case of disulfide. The infectivity enhancement of the disulfide linker-containing PA is reduced approx. 60 % (n-fold at 1.3 µM reduced from 10 to 4). This decline in activity can be traced back to the missing fibril aggregation of the disulfide linked PA (C_{16}-*disulfide*-CVVVAAAKKK-NH_2) as is evidenced by low light scattering count rate (**Figure S16**). C_{16}-*disulfide*-CVVVAAAKKK-NH_2 can be stained with the imaging agent Proteostat (**Figure S26**), but not observed via microscopy (**Figure S18**), because it is not forming visible aggregates.

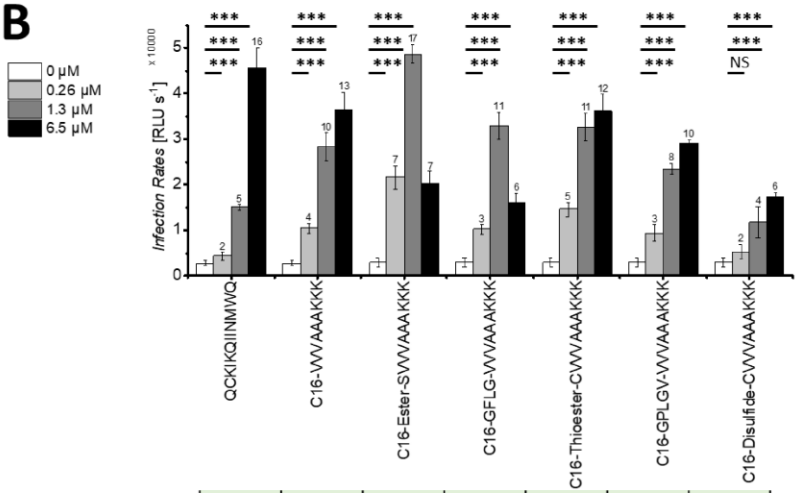
All investigated linker-modified PAs are degraded in the presence of cells (**Figure S17**). The best performing linker exchange is observed for the thioester modification, which shows a degradation of 87% and small, well distributed aggregates (**Figure S18**). The ester modified PA is also forming small aggregates and degraded well, but not as efficient as the thioester modified or PA1. The PAs with enzymatically degradable peptide sequences (MMP and cathepsin B) did not perform better than other linkers in terms of degradability (**Figure S17**). This may be because the peptide sequences are not sterically not accessible to the relevant enzymes.

A



Sequence	Infectivity enhancing n-fold at 1.3 µM	Fibril Formation	Degradability / % (rel. to initial particles > 10 µm²)
C_{16}-VVVAAAKKK-NH_2	10±3	Yes	94%
C_{16}-*ester*-SVVVAAAKKK-NH_2	17±5	Yes	31%
C_{16}-*disulfide*-CVVVAAAKKK-NH_2	4±2	Yes	ND
C_{16}-*thioester*-CVVVAAAKKK-NH_2	11±4	Yes	**87%**
C_{16} -GPLGV-VVVAAAKKK-NH_2	8±3	Yes	28%
C_{16} –GFLG- VVVAAAKKK-NH_2	11±4	Yes	4%

B



Zeta-Potential / mV	6.6	22.2	20.5	18.8	18.3	17.3	19.8
Count Rate / kcps	2602.3	3138.4	3217.3	5544.1	2058.5	5058.7	926.1
N-fold ThT-fluorescence	16.2	278.7	277.1	340.8	166.1	395.5	21.4
Fibril Formation	1	1	1	1	1	1	1
β-sheet / %	52	57	43	57	50	64	57

Figure S16 A Summary of sequence, infectivity enhancement at 1.3 µM concentration relative to virus only (n-fold), assembly into fibrils determined by TEM (**Figure S27**), and degradability determined via confocal microscopy aggregate analysis of particles > 10 µm². Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids. **B** Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of peptides EF-C and newly designed PAs (6.5 µM, 1.3 µM, 0.26 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. P-values for statistical significance were determined by two-way ANOVA test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant. RLU s^{-1} are relative light units per second. The table summarizes the respective physicochemical properties of the PAs. PAs assembling into fibrils as determined by TEM (**Figure S27**) are indicated with "1". The β-sheet content is calculated from the FT-IR spectra (**Figure S28**). Properties which were previously found to favor infectivity are colored green.

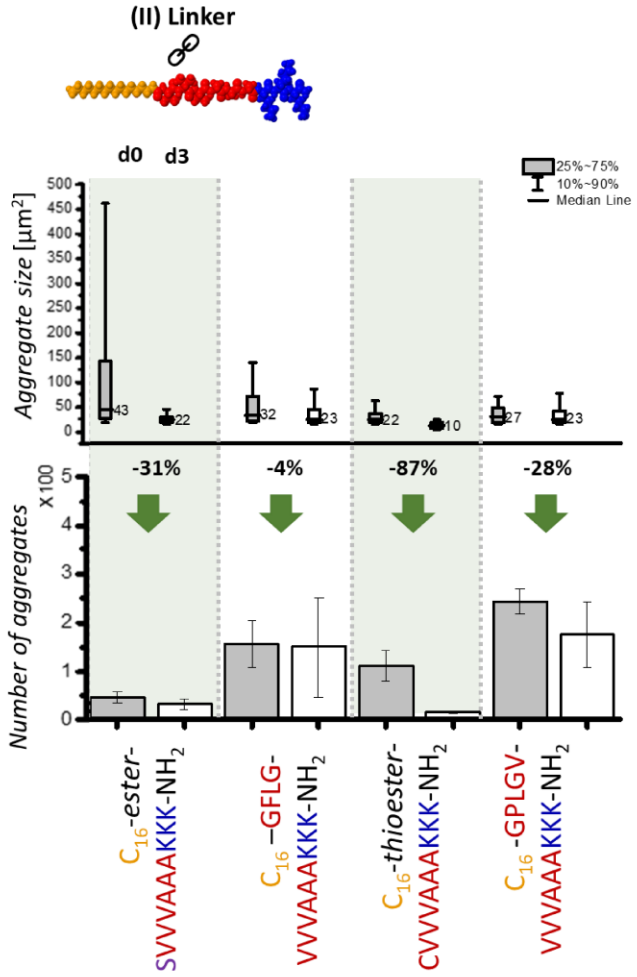


Figure S17 Degradability of PAs investigated via the aggregate size and number of aggregates larger than 10 µm² in an area of 1008 µm × 1008 µm. Error bars in aggregate number depict technical triplicate measurements. The percentual degradability was determined by number of aggregates after 3 days (d3, white boxes and bars) relative to the initial number of aggregates after 30 min incubation (d0 black boxes and bars). Box plots show aggregate size distribution with 10-90% shown as whiskers, 25-75% as boxes and the median line. The arrows give a qualitative evaluation of size reduce (green) or increase (red), number reduce (green) or number increase (red). Yellow circle means no great change from initial state. Green highlighted section is pointing out best performing candidates for degradability.

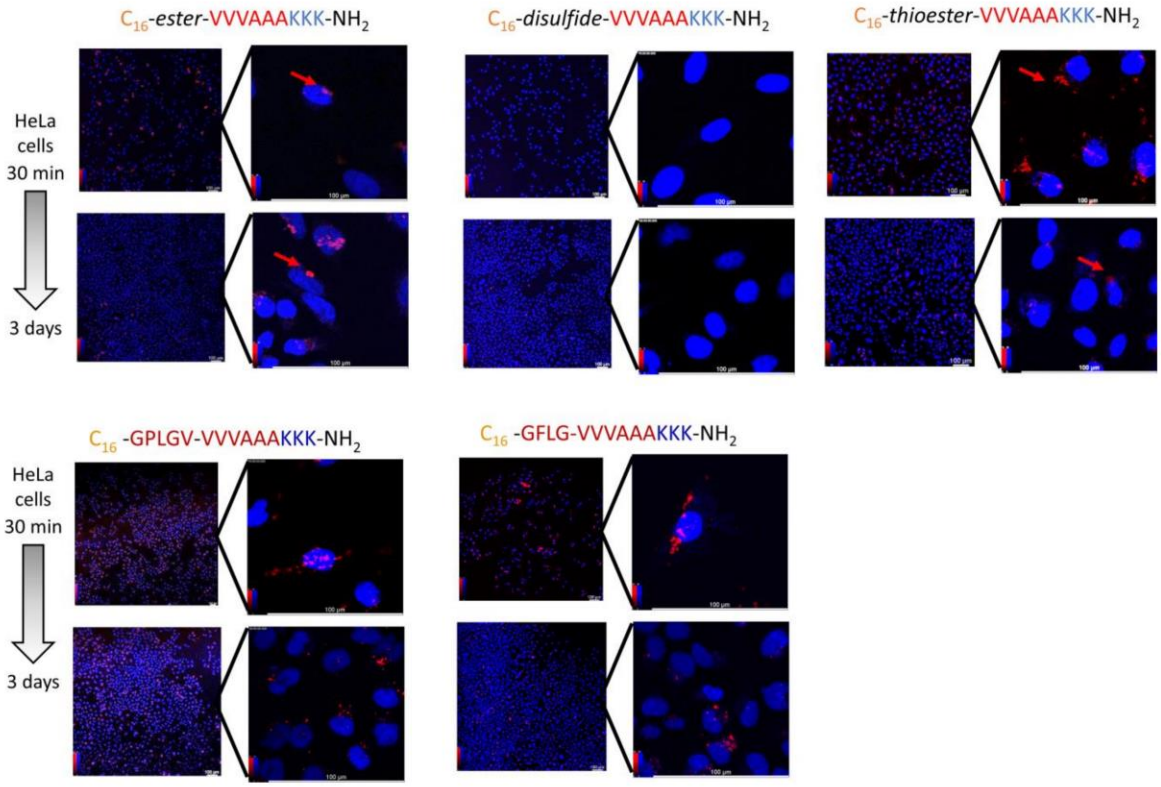


Figure S18 Representative confocal laser scanning microscopy images of Proteostat-labeled PAs (red) incubated with HeLa cells (nucleus in blue) after 30 minutes of incubation (top rows) and 3 days of incubation (bottom rows), scale bar 100 µm.

7.3 Effects of sequence exchange on properties and activity

Next, we evaluated if attaching a hydrophobic moiety to non-assembling peptides is a reliable way to achieve assembly into fibrils and infectivity enhancement, and what design rules for the sequence are important to obtain infectivity enhancing PAs (**Figure S19** – **Figure S22**). To this end, we attached palmitic acid to peptides (**Figure S31**, **Figure S32**) which are

(I) inspired from the original sequence design of PA1 (hydrophobic amino acids are close to palmitic acid residue, polar amino acids are mostly found at the C-terminus)

(II) known to have β-sheet prone sequences but are not assembling into fibrils without fatty acid at the investigated conditions. We selected peptide sequences which follow the design rule of PA1 (hydrophobic amino acids are close to palmitic acid residue, polar amino acids are mostly found at the C-terminus).
or

(III) endogenous peptide sequences which are found in β-sheet rich regions of proteins but are sequence-wise non-related to PA1.

The respective peptide sequences without palmitic acid were analyzed analogously (**Figure S23**, **Figure S24**).

We found that attaching palmitic acid to the *N*-terminus of non-assembling peptides is a very reliable way to create peptide amphiphiles which form fibrils (90%, 19 of 21 tested compounds assembled into fibrils, **Figure S19**, **Figure S27**). The only peptide sequence which are not forming fibrils after functionalization with palmitic acid are C_{16}-KGVPGVGK-NH_2 and C_{16}-WEALKK-NH_2. Both sequences include charged amino acids (K or E) close to the hydrophobic moiety, which might disturb self-assembly. Moreover, all the PAs except of C_{16}-KGVPGVGK showed a positive Zeta-potential, while the corresponding peptides without fatty acid moiety showed negative or almost neutral (< 5 mV) zeta-potential. Attaching palmitic acid to net positive charged peptide sequences is therefore a reliable way to obtain nanostructures with positive Zeta-potential.

Strikingly, 11 of the 21 newly created PAs with different sequences show infectivity enhancement (n fold > 2 at 1.3 µM) with similar or less activity compared to PA1. No significant difference regarding to design principles (I-III) could be found, but all the non-active (n-fold at 1.3 µM < 2) and weakly active (n-fold at 1.3 µM 2–3) PAs had in common that they are not forming µm-sized aggregates and forming too few/less aggregates, respectively (**Figure S21** - **Figure S22**: C_{16}-WEALKK-NH_2, C_{16}-AGGRVK-NH_2, C_{16}-AVAKKK-NH_2, C_{16}-RRWQWR-NH_2, C_{16}-KGVPGVGK-NH_2, C_{16}-GGGSSKKK-NH_2, C_{16}-GGVVRR-NH_2, C_{16}-AGAGRRR-NH_2).

PAs that form no or too small aggregates have in common that they contain a high number of flexible amino acids (G, A) and polar amino acids (R, K, E) close to the fatty acid. This could increase dynamicity and solvation of the self-assembling PAs, which result in less aggregating fibers. Further, sterically demanding amino acids (W) close the fatty acid moiety and kinked amino acids (P) in the sequence hinder efficient assembly, which can increase dynamicity of the fibers and result in less aggregation.

On the other side having few, large aggregates (approximately less than 100 aggregates per 1008 µm × 1008 µm with more than 10% larger than 200 µm) reduces infectivity enhancement and

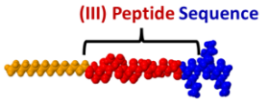
degradability. PAs that are forming these large aggregates as observed via confocal microscopy aggregate size analysis (**Figure S21**- **Figure S22**) are C16-AQAKAK, C_{16}-NFYLQVNKINK-NH_2, C_{16}-KYIHQNYTKALAGKLV-NH_2, C_{16}-KIISFK-NH_2, C_{16}-AAVVHHK-NH_2, and C_{16}-VVAAHHH-NH_2 form few but very large aggregates.

Most of the PAs that form very large aggregates have in common, that their sequences are composed of alternating, amphiphilic amino acids, comparable to PA2, that also forms large aggregates that cannot degraded well (**Figure 4**). C_{16}-VVAAHHH-NH_2 does not contain any charged amino acid at pH 7.4 and C_{16}-AAVVHHK-NH contains only one cationic charged residue. Aggregation is likely caused due to the hydrophobic, aromatic character of histidine (H) at the C-terminus, which decreases solubility and favors fiber-fiber aggregation.

Further, most of the PAs that form very large aggregates are degraded incompletely as evidenced by decrease of aggregate size but increase in aggregate number for C_{16}-AAVVHHK-NH_2, C_{16}-NFYLQVNKINK-NH_2, C_{16}-KYIHQNYTKALAGKLV-NH_2, C_{16}-KIISFK-NH_2, C16-AQAKAK. This observation could indicate a fragmentation of large aggregates that occurs simultaneously to the degradation. The opposite case (increase in size, decrease in number) indicates an aggregation of small particles to larger particles, and can be observed for C_{16}-VVAAHHH-NH_2 and C_{16}-RYYASLRHYLNLVTRQRY-NH_2. Both cases are not desired in *in vitro* applications, thus these PAs are not well suited as degradable additives to enhance viral infectivity.

The best performing candidates in terms of infectivity and degradability are C_{16}-AKAVGK-NH_2 (infect *n* fold = 6, degradability 89%), then C_{16}-LNGGVKK-NH_2 (infect *n* fold = 6, degradability 48%) and finally C_{16}-ALAAGKK-NH_2 (infect *n* fold = 5, degradability 87%).

Noteworthy the peptides NFYLQVNKINK-NH_2, RYYASLRHYLNLVTRQRY-NH_2 and KGVPGVGK-NH_2 are forming fibrils even without palmitic acid. With the exception of NFYLQVNKINK-NH_2 these peptides were found to be non-aggregating and inactive (**Figure S23**, **Figure S24**). The sole active sequence (NFYLQVNKINK-NH_2) also showed the required physicochemical properties but no good degradability (**Figure S24**).



Sequence	Infectivity enhancing n-fold at 1.3 µM	Fibril Formation	Degradability / % (rel. to initial particles > 10 µm²)	Sequence Design Information
C_{16}-VVVAAAKKK-NH_2	10±3	Yes	94%	Original Sequence
C_{16}-AVAKKK-NH_2	2±0	Yes	94%	inspired **from original sequence**
C_{16}-VVAAHHH-NH_2	2±0	Yes	26%	inspired **from original sequence**
C_{16}-GGGSSKKK-NH_2	2±1	Yes	ND	inspired **from original sequence**
C_{16}-AGAGRRR-NH_2	1±0	Yes	ND	inspired **from original sequence**
C_{16}-AAVVHHK-NH_2	6±3	Yes	-123%	inspired **from original sequence**
C_{16}-GGVVRR-NH_2	1±1	Yes	ND	inspired **from original sequence**
C_{16}-IIQKIK-NH_2	4±1	Yes	28%	**non-assembling**, β-sheet prone sequence
C_{16}-AQAKAK-NH_2	4±1	Yes	-83%	**non-assembling**, β-sheet prone sequence
C_{16}-LLLKK-NH_2	3±0	Yes	27%	peptide **not assembling** by its own, but related derivatives (β–sheet prone)
C_{16}-KGVPGVGK-NH_2	1±0	No	ND	peptide **not assembling** by its own, but related derivatives (β-sheet prone)
C_{16}-AGGRVK-NH_2	3±1	Yes	ND	**Endogeneous** Inspired from AGGRV (138-142 T-cell receptor on myeloid cells)
C_{16}-LNGGVKK-NH_2	6±1	Yes	48%	**Endogeneous** Inspired from 116-121 EGFR factor
C_{16}-ALAAGKK-NH_2	5±1	Yes	87%	**Endogeneous** Reverse seq of KGAALA: 105-110 Linoleate 9S-lipoxygenase
C_{16}-WEALKK-NH_2	1±0	No	ND	**Endogeneous** Reverse sequence of KKLAEW: 155-160 Protein phosphatase 1D
C_{16}-RRWQWR-NH_2	0±0	Yes	ND	**Endogeneous** Derived from cell-penetrating peptide of bovine lactoferrin
C_{16}-KYIHQNYTKALAGKLV-NH_2	5±2	Yes	-110%	**Endogeneous** 532 - 548 sequence from procollagen I
C_{16}-NFYLQVNKINK-NH_2	5±1	Yes	-123%	**Endogeneous** 282- 293 sequence from lumican
C_{16}-RYYASLRHYLNLVTRQRY-NH_2	5±1	Yes	43%	**Endogeneous** 46 - 63 PYY gut hormone truncated
C_{16}-GKVGSK -NH_2	2±1	Yes	ND	**Endogeneous** Serumalbumin derived
C_{16}-AKAVGK-NH_2	6±3	Yes	89%	**Endogeneous** GAPDH derived
C_{16}-KIISFK-NH_2	2±1	Yes	-84%	**Endogeneous** IL-18 derived

Figure S19 Summary of sequence, infectivity enhancement at 1.3 µM concentration relative to virus only (n-fold), assembly into fibrils determined by TEM (**Figure S27**), and degradability determined via confocal microscopy aggregate analysis of particles > 10 µm². Additionally, sequence design information describes the rationally behind the sequence selection. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

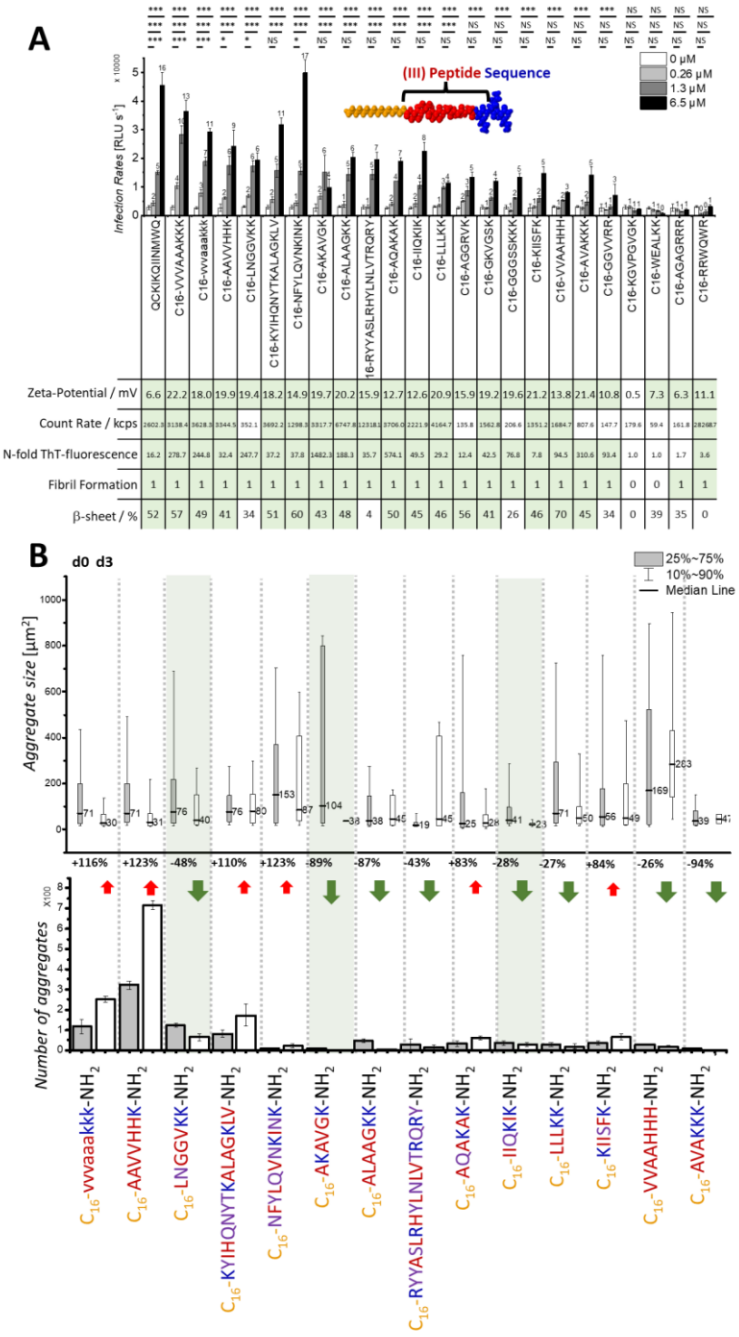


Figure S20 A Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of peptides EF-C and newly designed PAs (6.5 µM, 1.3 µM, 0.26 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. P-values for statistical significance were determined by two-way ANOVA test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant.RLU s^{-1} are relative light units per second. The table summarizes the respective physicochemical properties of the PAs. PAs assembling into fibrils as determined by TEM (**Figure S27**) are indicated with "1". The β-sheet content is calculated from the FT-IR spectra (**Figure S28**). Properties which were previously found to favor infectivity are colored green. **B** Degradability of PAs investigated via the aggregate size and number of aggregates larger than 10 μm^2 in an area of 1008 µm × 1008 µm. Error bars in aggregate number depict technical triplicate measurements. The percentual degradability was determined by number of aggregates after 3 days (d3, white boxes and bars) relative to the initial number of aggregates after 30 min incubation (d0 black boxes and bars). Box plots show aggregate size distribution with 10-90% shown as whiskers, 25-75% as boxes and the median line. The arrows give a qualitative evaluation of size reduce (green) or increase (red), number reduce (green) or number increase (red). Green highlighted section is pointing out best performing candidates for degradability.

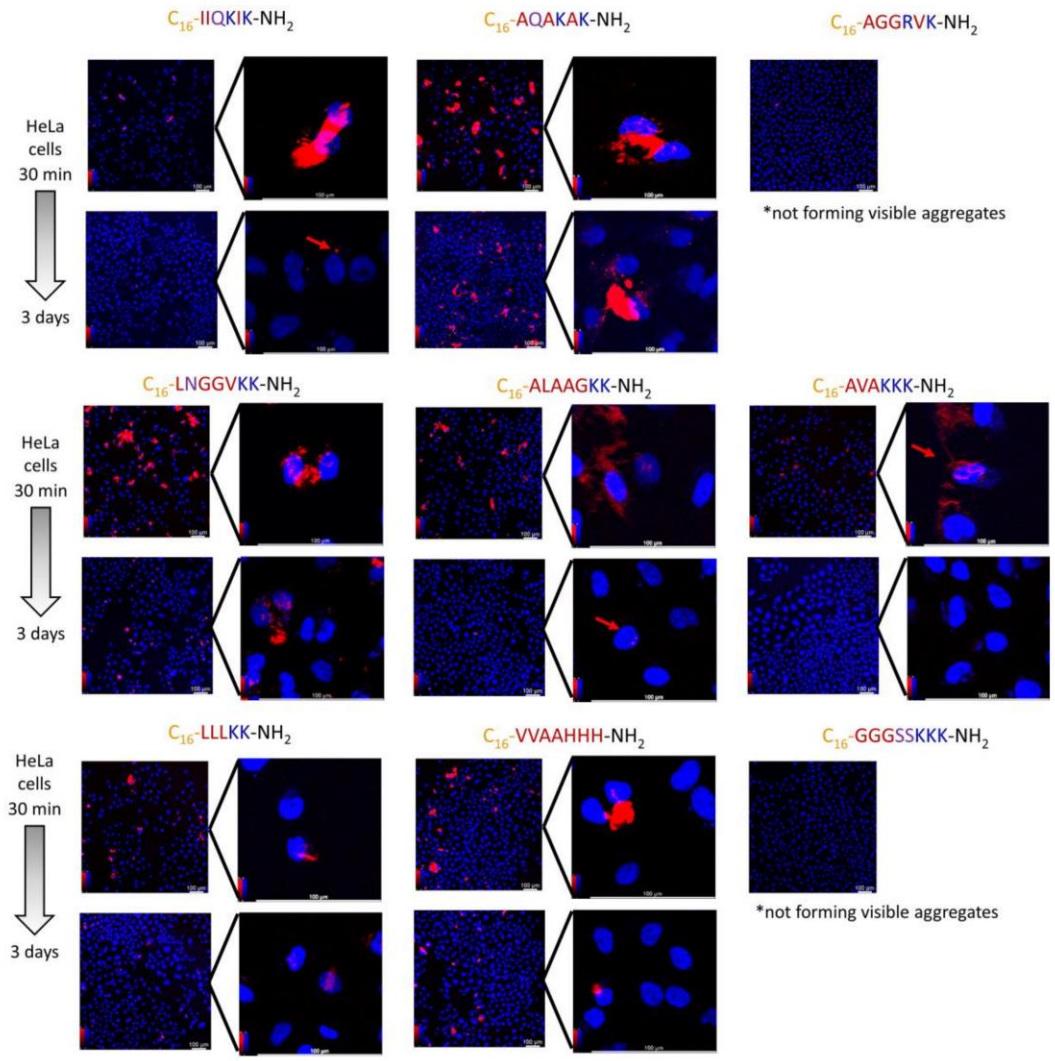


Figure S21 Representative confocal laser scanning microscopy images of Proteostat-labeled PAs (red) incubated with HeLa cells (nucleus in blue) after 30 minutes of incubation (top rows) and 3 days of incubation (bottom rows), scale bar 100 µm.

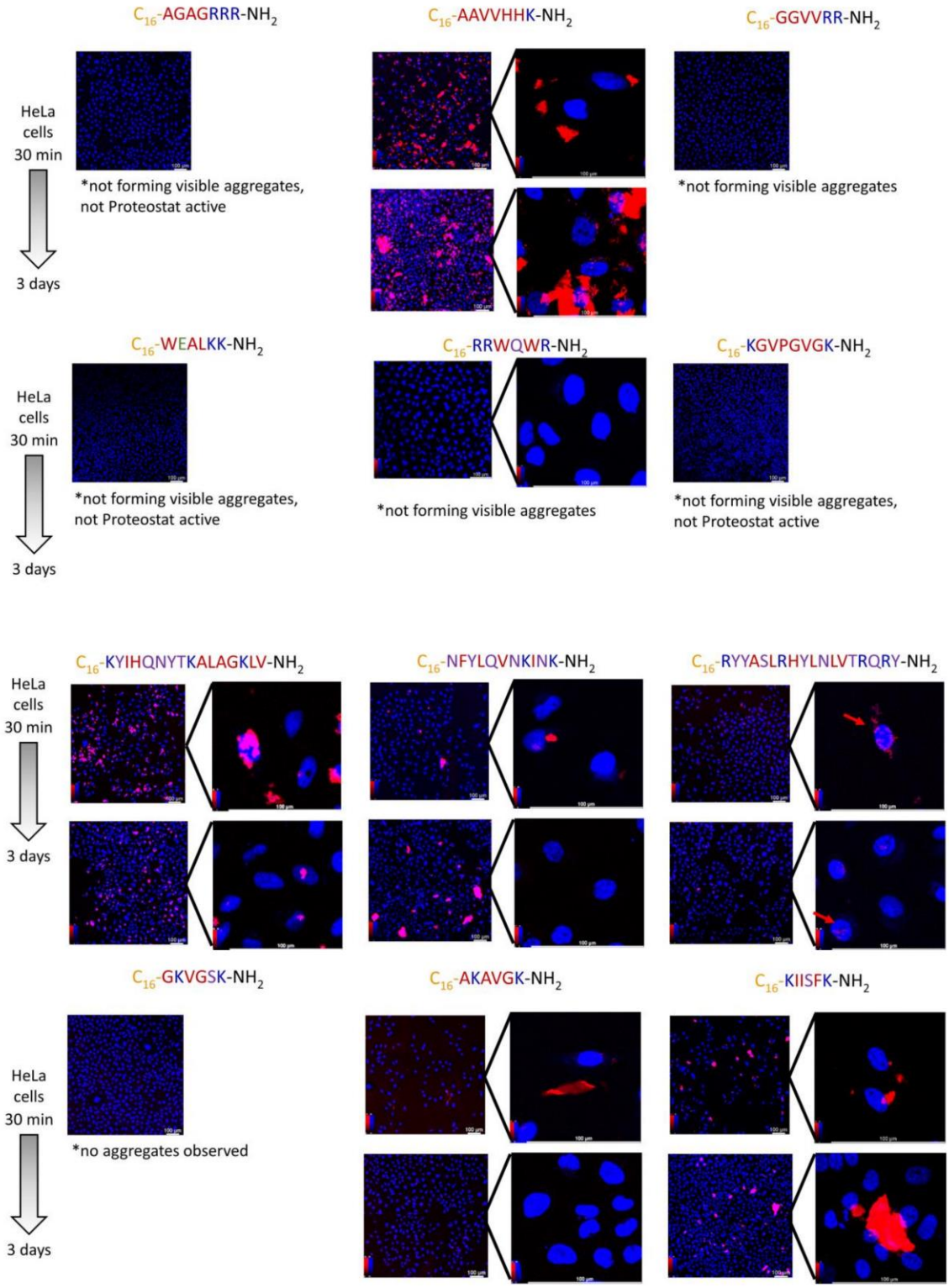
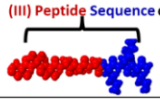


Figure S22 Representative confocal laser scanning microscopy images of Proteostat-labeled PAs (red) incubated with HeLa cells (nucleus in blue) after 30 minutes of incubation (top rows) and 3 days of incubation (bottom rows), scale bar 100 µm.

(III) Peptide Sequence control without fatty acid



Sequence	Infectivity enhancing n-fold at 1.3 µM	Fibril Formation	Degradability / % (rel. to initial particles > 10 µm²)	Sequence Design Information
C_{16}-VVVAAAKKK-NH_2	10±3	Yes	94%	Original Sequence
VVVAAAKKK-NH_2	0.5±0.3	No	ND	
AVAKKK-NH_2	1.2±0.2	No	ND	inspired **from original sequence**
VVAAHHH-NH_2	0.7±0.1	No	ND	inspired **from original sequence**
GGGSSKKK-NH_2	1.0±0.3	No	ND	inspired **from original sequence**
AGAGRRR-NH_2	0.3±0.1	No	ND	inspired **from original sequence**
AAVVHHK-NH_2	0.8±0.3	No	ND	inspired **from original sequence**
GGVVRR-NH_2	0.8±0.6	No	ND	inspired **from original sequence**
IIQKIK-NH_2	1.2±0.4	No	ND	**non-assembling**, β-sheet prone sequence
AQAKAK-NH_2	0.6±0.2	No	ND	**non-assembling**, β-sheet prone sequence
LLLKK-NH_2	0.8±0.3	No	ND	peptide **not assembling** by its own, but related derivatives (β-sheet prone)
KGVPGVGK-NH_2	0.9±0.1	No	ND	peptide **not assembling** by its own, but related derivatives (β-sheet prone)
AGGRVK-NH_2	0.7±0.2	No	ND	**Endogeneous** Inspired from AGGRV (138-142 T-cell receptor on myeloid cells)
LNGGVKK-NH_2	0.8±0.2	No	ND	**Endogeneous** Inspired from 116-121 EGFR factor
ALAAGKK-NH_2	1.0±0.3	No	ND	**Endogeneous** Reverse seq of KGAALA: 105-110 Linoleate 9S-lipoxygenase
WEALKK-NH_2	0.8±0.2	No	ND	**Endogeneous** Reverse sequence of KKLAEW: 155-160 Protein phosphatase 1D
RRWQWR-NH_2	0.8±0.3	No	ND	**Endogeneous** Derived from cell-penetrating peptide of bovine lactoferrin
KYIHQNYTKALAGKLV-NH_2	1.2±0.5	No	ND	**Endogeneous** 532 - 548 sequence from procollagen I
NFYLQVNKINK-NH_2	7.9±1.7	Yes	10%	**Endogeneous** 282- 293 sequence from lumican
RYYASLRHYLNLVTRQRY-NH_2	1.9±1.0	Yes	ND	**Endogeneous** 46 - 63 PYY gut hormone truncated
GKVGSK -NH_2	0.8±0.3	No	ND	**Endogeneous** Serumalbumin derived
AKAVGK-NH_2	0.9±0.3	No	ND	**Endogeneous** GAPDH derived
KIISFK-NH_2	0.9±0.3	No	ND	**Endogeneous** IL-18 derived

Figure S23 Summary of sequence, infectivity enhancement at 1.3 µM concentration relative to virus only (n-fold), assembly into fibrils determined by TEM (**Figure S27**), and degradability determined via confocal microscopy aggregate analysis of particles > 10 µm^2. Additionally, sequence design information describes the rationally behind the sequence selection. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

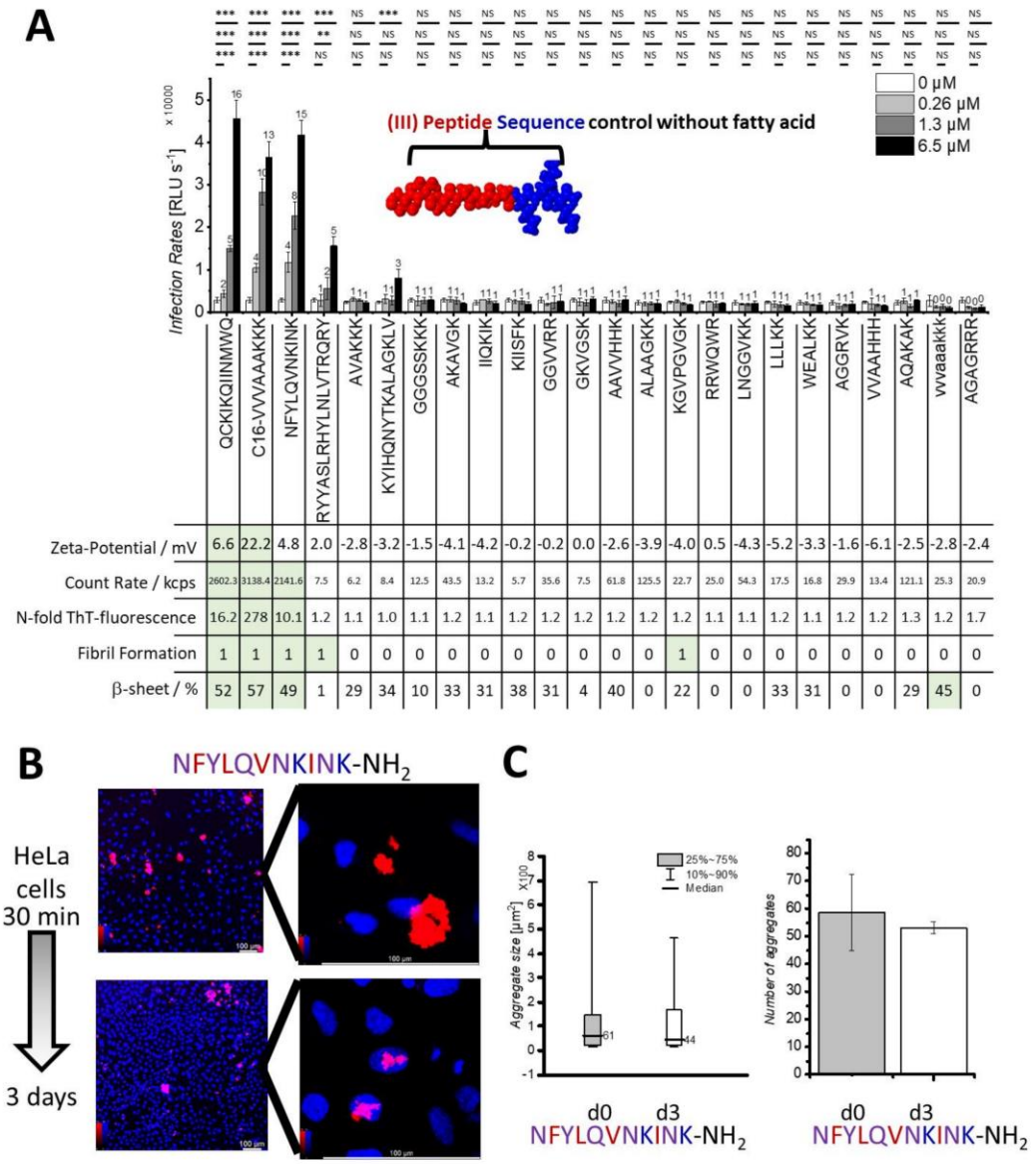


Figure S24 A Infectivity assay showing HIV-1 infection rates of TZM-bl cells observed in the presence of increasing concentrations of peptides EF-C and newly designed PAs (6.5 µM, 1.3 µM, 0.26 µM). The baseline (0 µM) shows infection rate without the addition of peptides. Values above each bar represent the n-fold enhancement compared to 0 µM. Mean values and standard deviations are derived from three repetitions of triplicate infections. P-values for statistical significance were determined by two-way ANOVA test and are indicated with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, NS not significant. RLU s^{-1} are relative light units per second. The table summarizes the respective physicochemical properties of the PAs. PAs assembling into fibrils as determined by TEM (**Figure S27**) are indicated with "1". The β-sheet content is calculated from the FT-IR spectra (**Figure S28**). Properties which were previously found to favor infectivity are colored green. **B** Representative confocal laser scanning microscopy images of Proteostat-labeled PA (red) incubated with HeLa cells (nucleus in blue) after 30 minutes of incubation (top row) and 3 days of incubation (bottom row), scale bar 100 µm. **C** Degradability of Pat46 investigated via the aggregate size and number of aggregates larger than 10 μm^2 in an area of 1008 µm × 1008 µm. Error bars in aggregate number depict technical triplicate measurements. The percentual degradability was determined by number of aggregates after 3 days (d3, white boxes and bars) relative to the initial number of aggregates after 30 min incubation (d0 black boxes and bars). Box plots show aggregate size distribution with 10-90% shown as whiskers, 25-75% as boxes and the median line.

8 Cell-Viability of PA library

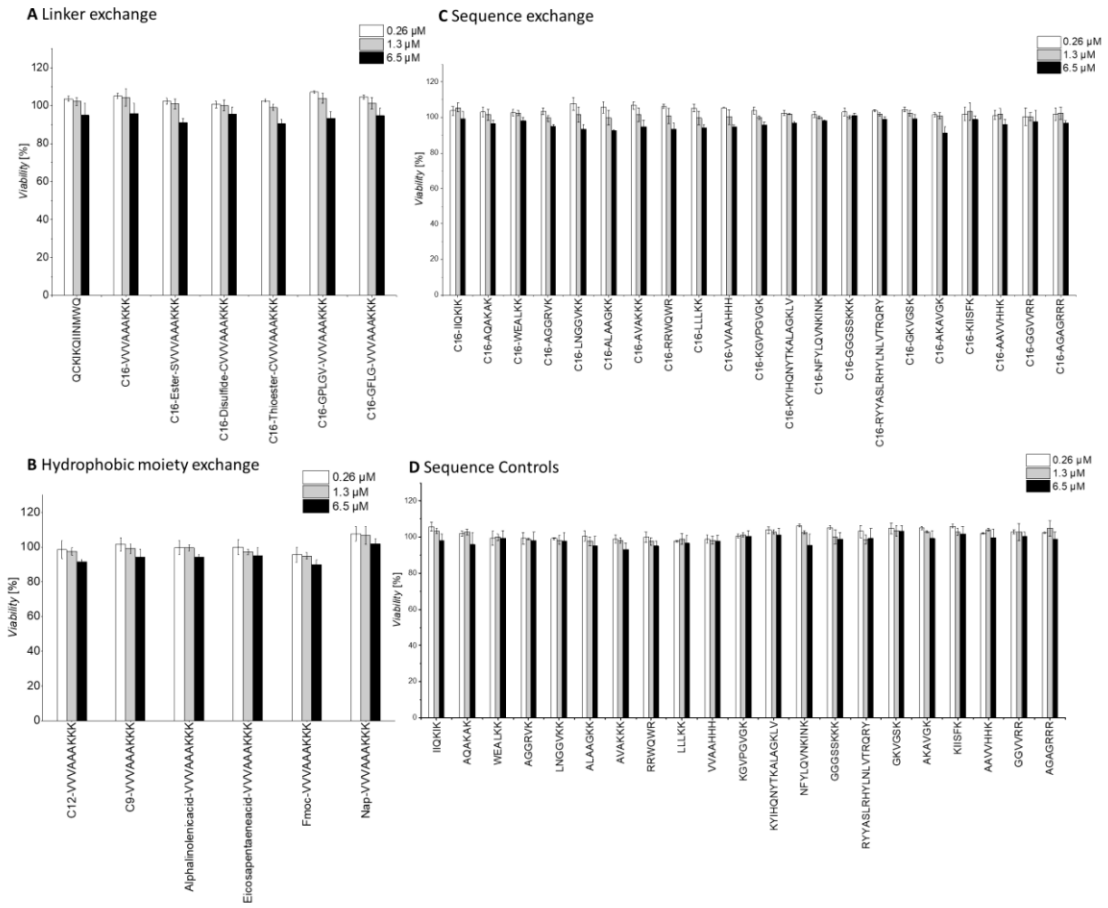


Figure S25 CellTiter-Glo cell viability assay of TZM-bl cells treated with increasing concentrations (0.26, 1.3 and 6.5 µM) of the peptides of biological triplicate with standard deviation.

9 Proteostat Assay of PA library

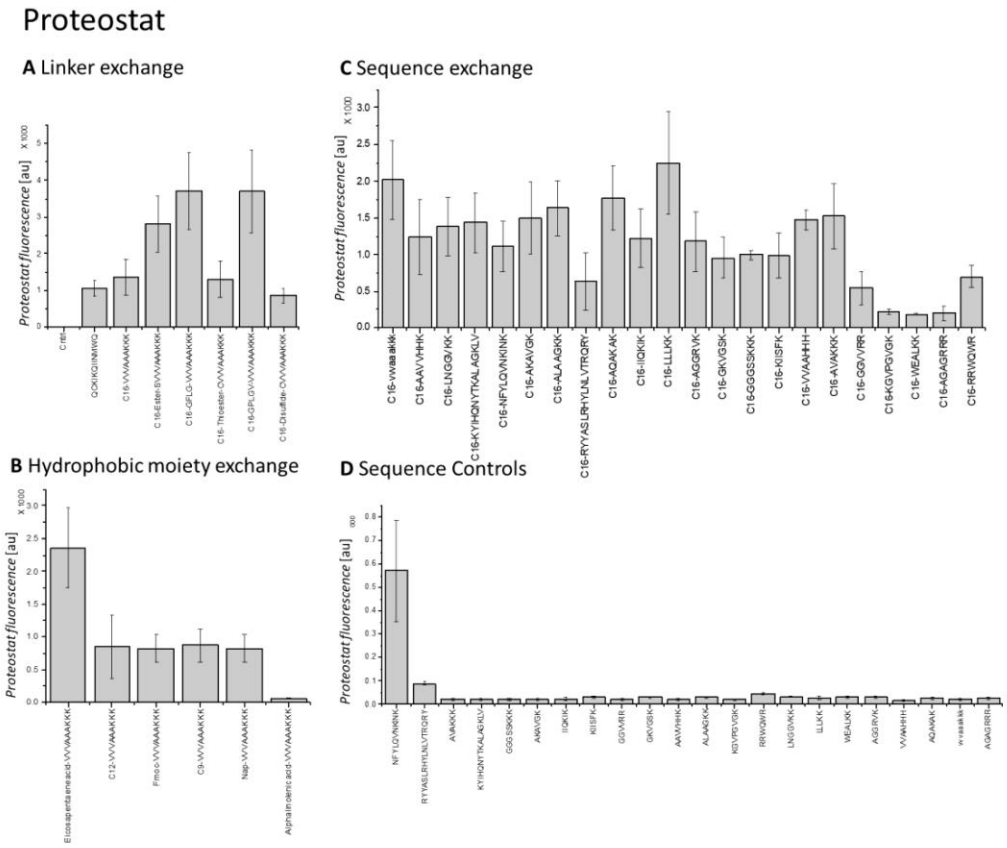


Figure S26 Proteostat Assay on PAs and peptides. **A** Linker exchange includes original sequence PA1 and variants with different linkers. For reference also EF-C and Control (PBS with 10% DMSO) are shown. **B** Exchange of hydrophobic moiety, **C** Exchange of peptide sequence **D** Controls showing peptide sequences without fatty acid residues.

10 TEM of PA library

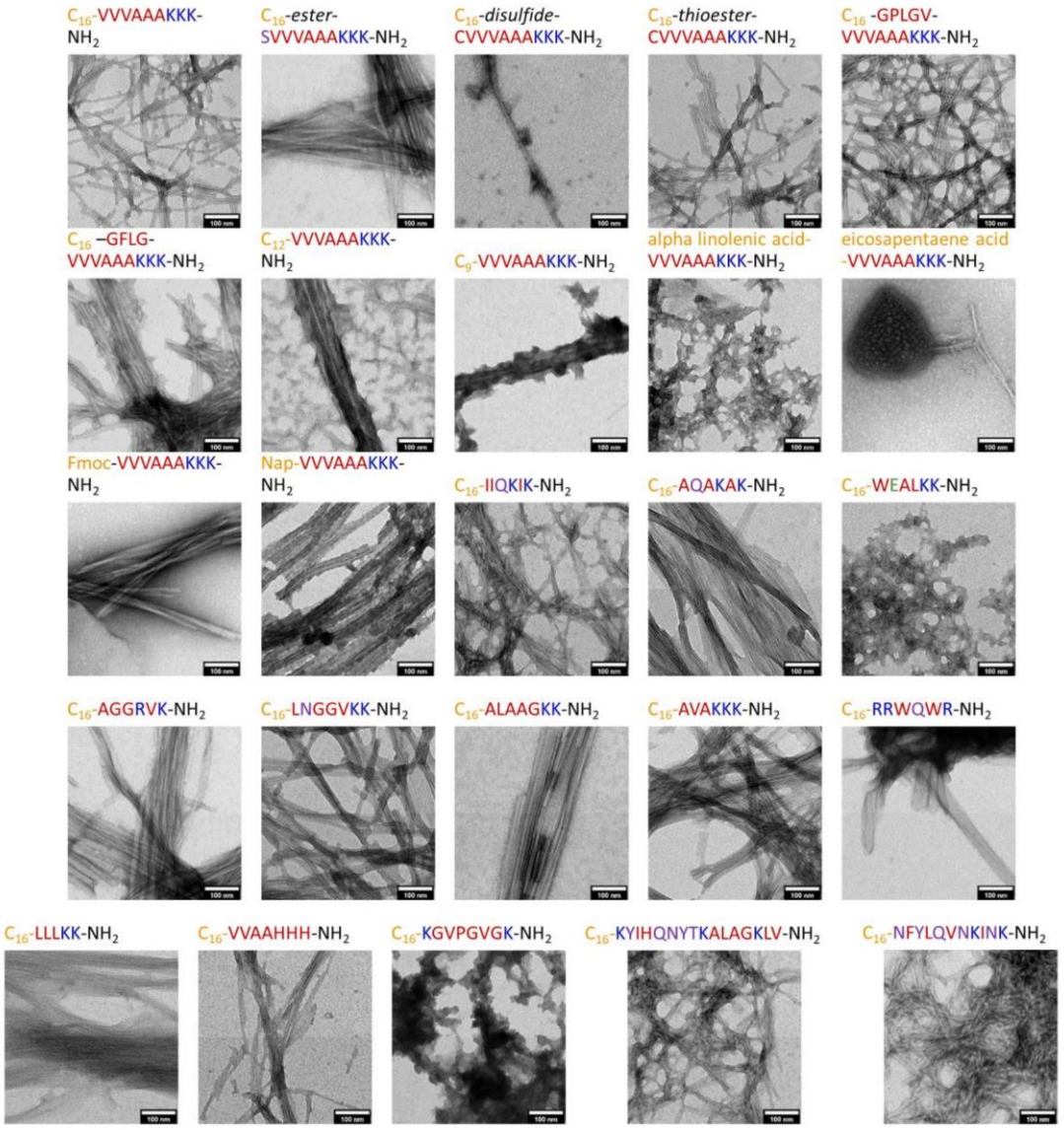


Figure S27 TEM micrographs of the PA and peptide library, scale bar 100 nm. The compounds were diluted from DMSO (10 mg mL^{-1}) in PBS (1 mg mL^{-1}) and incubated for 1d at rt. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

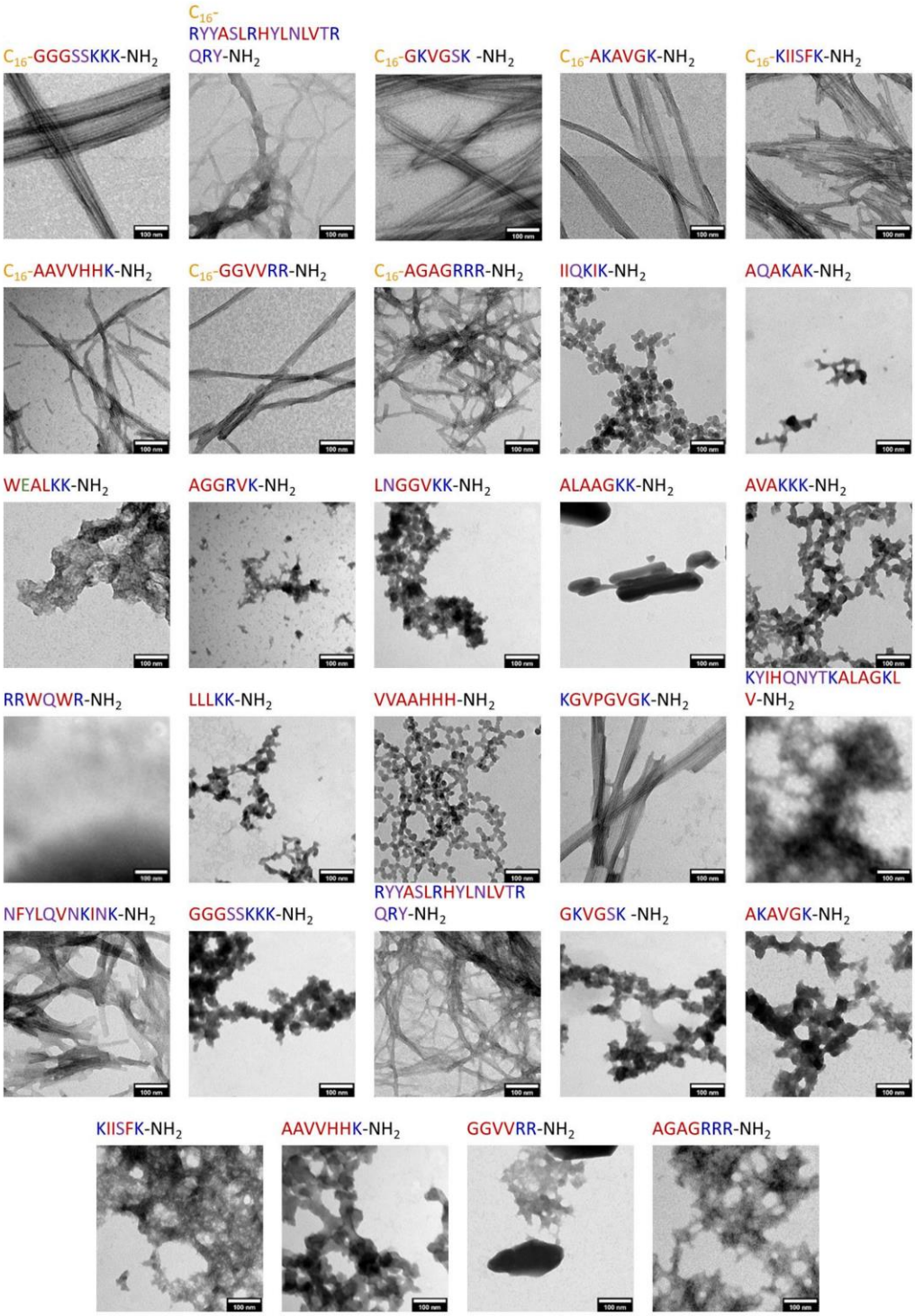


Figure S27 cont. TEM micrographs of the PA and peptide library, scale bar 100 nm. The compounds were diluted from DMSO (10 mg mL^{-1}) in PBS (1 mg mL^{-1}) and incubated for 1d at rt. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

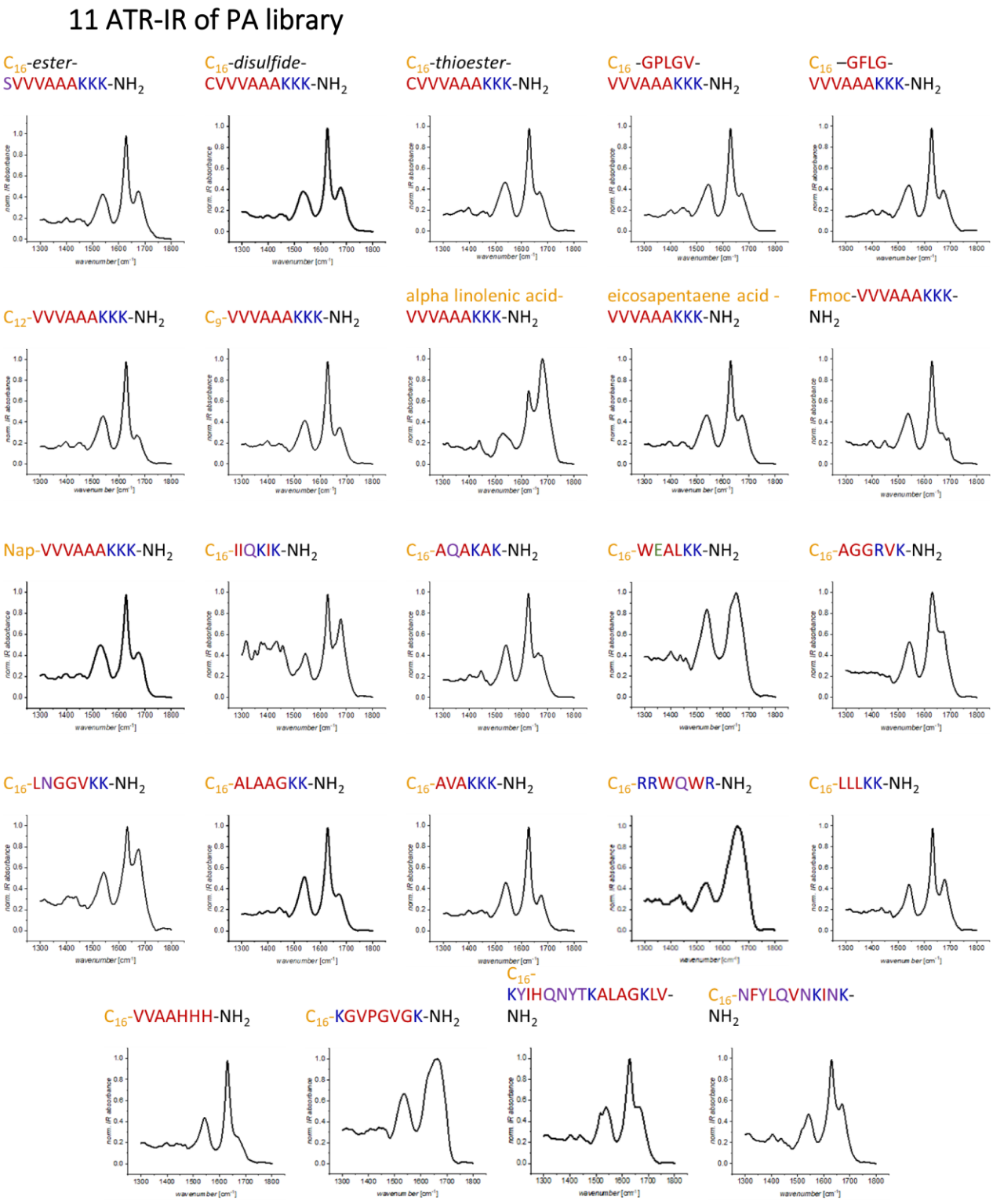


Figure S28 ATR-IR spectra of the PA and peptide library. The compounds were diluted from DMSO (10 mg mL^{-1}) in PBS (1 mg mL^{-1}) and incubated for 1d at rt before lyophilization. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

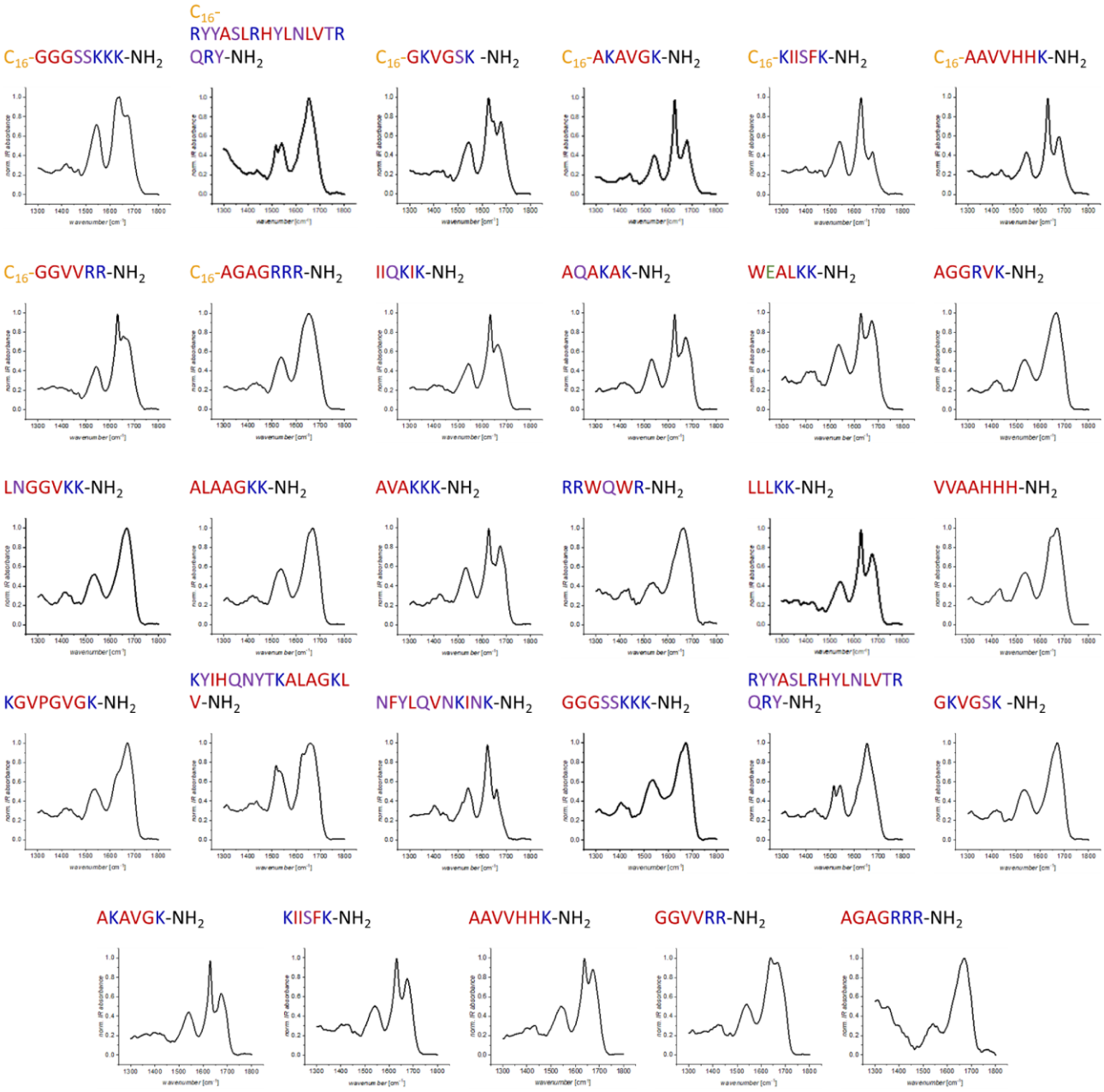


Figure S28 cont. ATR-IR spectra of the PA and peptide library. The compounds were diluted from DMSO (10 mg mL^{-1}) in PBS (1 mg mL^{-1}) and incubated for 1d at rt before lyophilization. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

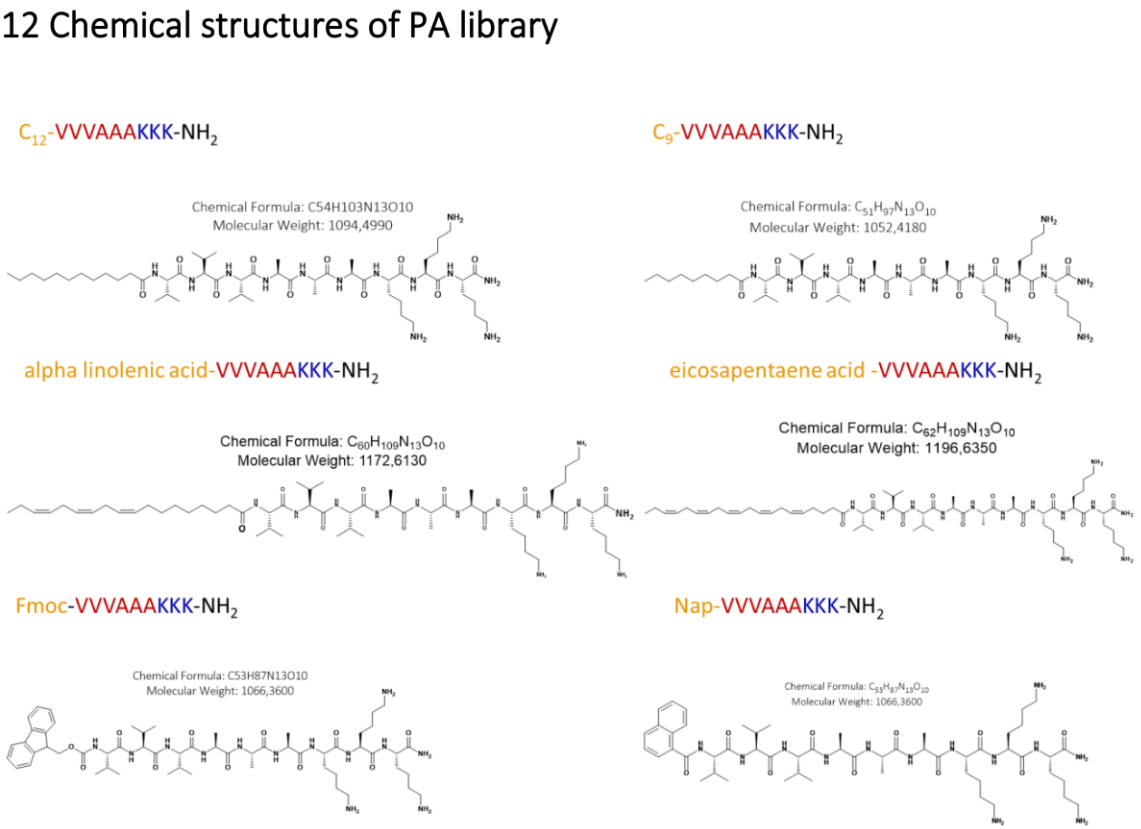


Figure S29 Chemical structure of the PA library with exchange of hydrophobic moiety. The PAs and their respective peptides without hydrophobic moiety were commercially obtained. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

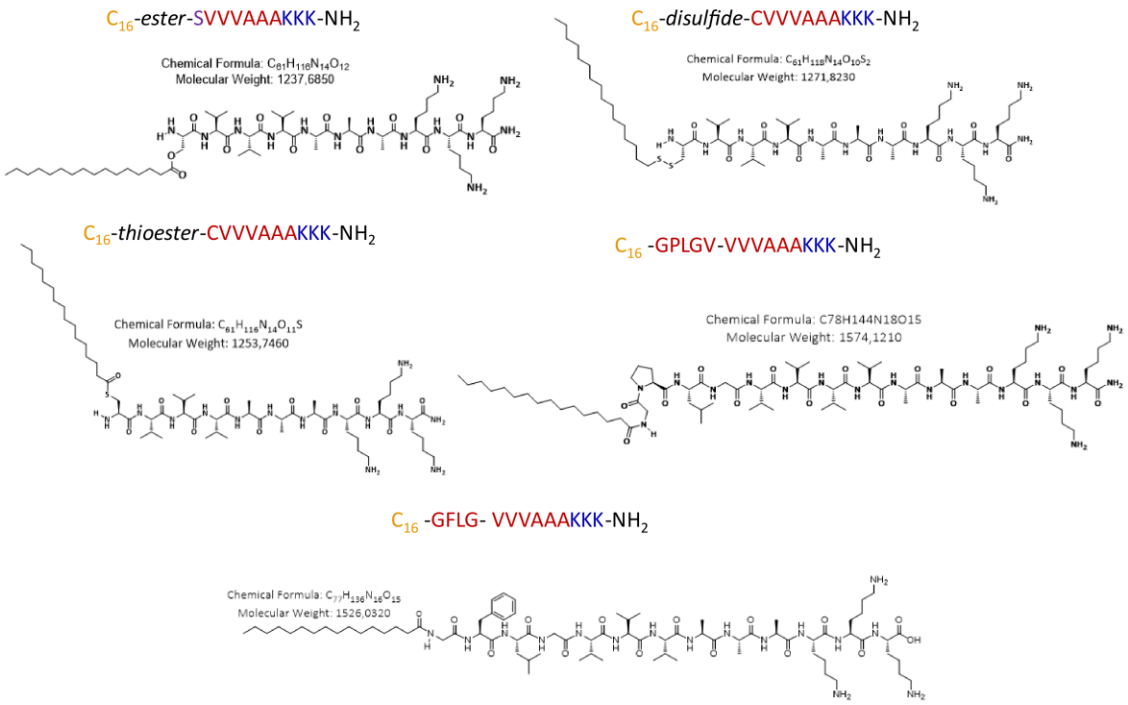


Figure S30 Chemical structure of the PA library with exchange of linker. The PAs and their respective peptides without hydrophobic moiety were commercially obtained. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

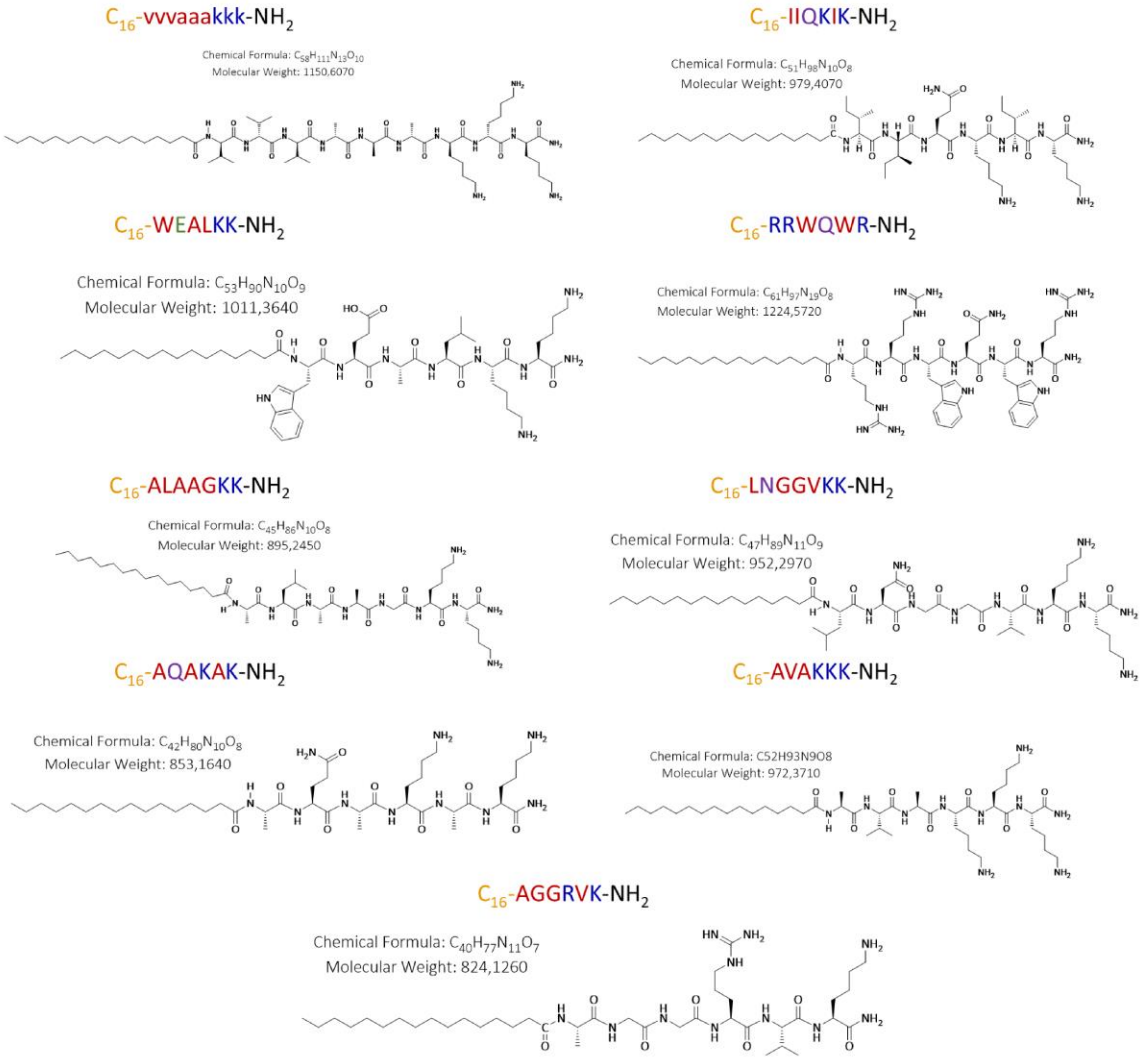


Figure S31 Chemical structure of the PA library with exchange of peptide sequence. The PAs and their respective peptides without hydrophobic moiety were commercially obtained. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

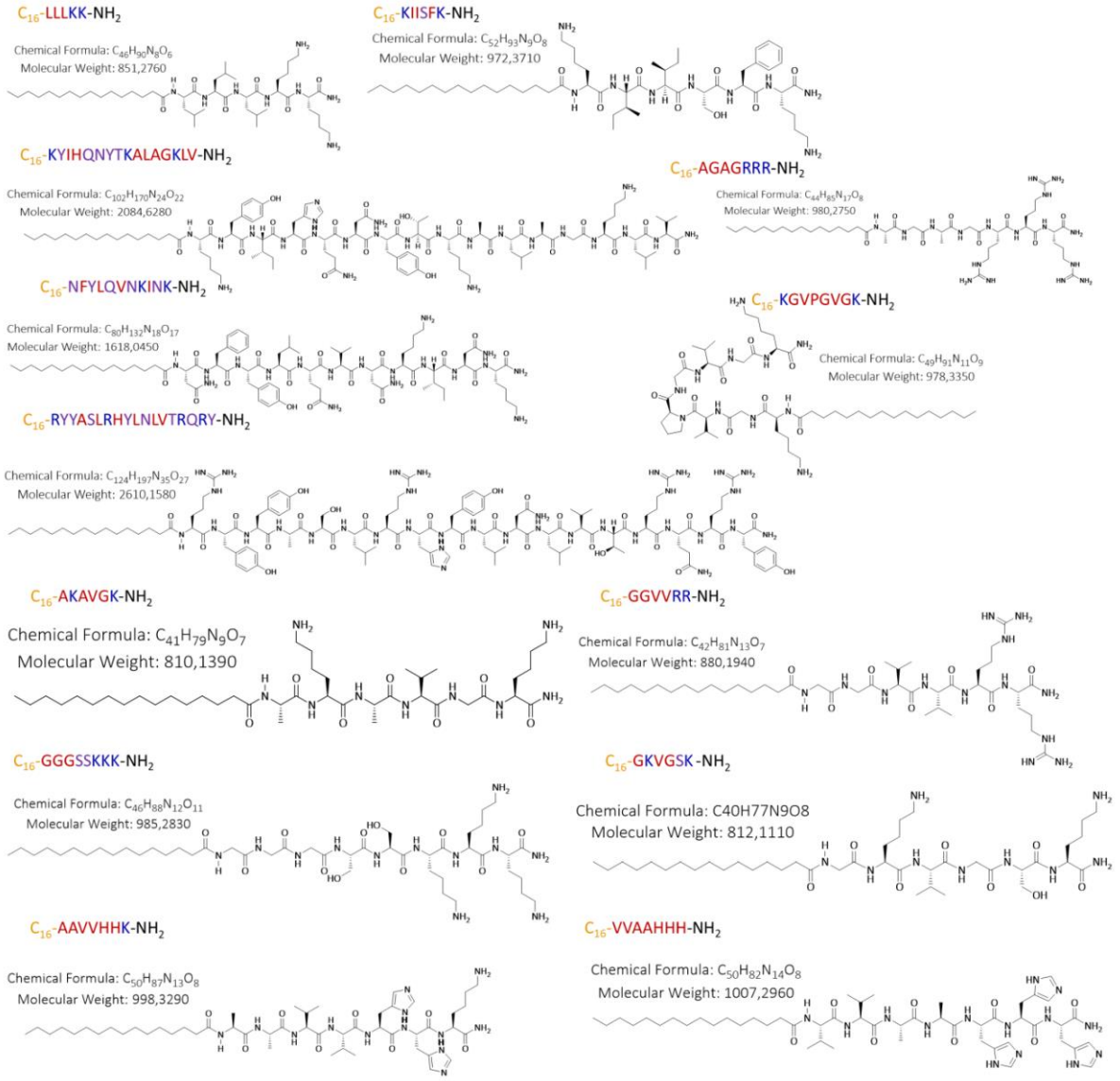


Figure S32 Chemical structure of the PA library with exchange of peptide sequence. The PAs and their respective peptides without hydrophobic moiety were commercially obtained. Color code for sequence representation: Yellow hydrophobic moiety, blue cationic amino acids, red non-polar amino acids, purple polar amino acids, green anionic amino acids.

13 Computational Modeling

13.1 Methods

In order to study the self-assembly of peptides into supramolecular nanofibrils, we performed molecular dynamics (MD) simulations using the MARTINI coarse-grained model version 2.2.[15] We opted to use the MARTINI force field here because we are interested in the microscopic aggregation of the peptides in solution in order to make connections with the even larger scale experiments. However, it is important to mention that the MARTINI force field does not provide information about how the secondary structure of the peptides change over time and we do not use our simulations to determine secondary structures. Therefore, when evaluating fibril-fibril interactions, we assume that the structure of the peptides that form the fibril remain the same. We simulated three different peptide amphiphiles (PAs) that show different aggregation behaviors based on our experimental data: C_{16}-VVVAAAKKK-NH_2 (PA1), C_{16}-FQFKFKC-NH_2 (PA2) and C_{16}-AGGRVK-NH_2. For each of the PAs, the C_{16} tail was represented by 4 beads based on earlier work.[16,17] The rest of the peptide sequence was modeled using standard MARTINI bead types for amino acids for pH of 7.4. The amidated C-terminal of the PAs was modeled using a polar, uncharged standard backbone bead (P5). Using the Packmol[18] software, we prepared the initial configuration of our simulations by randomly distributing 126 PAs in a box with initial size 13 nm × 13 nm × 6.5 nm, which makes a concentration of 0.11 peptides/nm^3 or 0.18 M. This high concentration allows us to accelerate the diffusion of peptides and study the self-assembly process at a reasonable computational cost. Although this concentration is very high compared to the experimental conditions, we argue that there will be similar high concentration regions in experiments before the self-assembly process takes place. Each system was solvated using Polarizable MARTINI water model[19] and neutralized with counterions (Cl^-) in order to account for the total charge on PA molecules. We also added 0.138 M NaCl to mimic experimental conditions. The exact number of water and ions differ slightly for each PA depending on the size and net charge of the PA molecule. For the system with PA1, there are a total of 126 PAs, 6133 water, 476 Cl^- and 98 Na^+ molecules. For PA2, there are 126 PAs, 6288 water, 350 Cl^- and 98 Na^+ molecules. For C_{16}-AGGRVK-NH_2, there are 126 PAs, 6689 water, 350 Cl^- and 98 Na^+ molecules.

We performed all simulations with Gromacs 2019.4[20,21] software according to the following protocol: (1) 10,000 steps minimization using steepest descent method, (2) 1 ns NVT (constant number of molecules, volume and temperature) simulation using 10 fs time step at 298 K, (3) 15 µs NPT (constant number of molecules, pressure and temperature) with 20 fs time step at 298 K and 1 bar pressure. We used the v-rescale coupling method[22] for temperature and Parrinello-Rahman barostat[23,24] for pressure coupling. The compressibility for the pressure coupling was 3 x 10^{-4} bar^{-1}. Bond constraints were handled using the LINear Constraint Solver (LINCS) algorithm.[25] The Verlet cutoff scheme[26] was used for neighbor searching. Long-range electrostatics were treated with the reaction field method. A cut-off of 1.1 nm was used for evaluation of all non-bonded interactions. For electrostatic interactions, we used the group method with dielectric constant 2.5, which is appropriate for polarizable Martini water model. The neighbor list was updated every 10 steps. Atomic coordinates were saved every 500 ps for the trajectory analysis. Periodic boundary conditions were set in x, y and z directions. All snapshots from simulations were rendered using the Visual Molecular Dynamics (VMD) software.[27]

After 15 µs of self-assembly simulations, we extracted the coordinates of the equilibrated PA fibrils to set up a larger simulation with 4 fibrils to study the aggregation behavior of the fibrils in solution. To

this end, we replicated the equilibrated self-assembled PA for each sequence 4 times and placed them in proximity from each other as shown in **Figure S33**. The initial distance between center of mass of fibrils in the box was about 2 times their diameter. At this distance, the peptides from different fibrils can interact with each other and aggregate or be repelled away based on the nature of their interaction. The initial box size along the length of the fibril (z) was taken directly from the self-assembly simulations (~6.5 nm) and the box size for the other two dimensions (x, y) were set to 18 nm. There was a total of 504 PAs (126 × 4) in each box. The simulation boxes were solvated and ionized as before. The same equilibration protocol was used for these simulations, but the total simulation time was 10 µs. For PA1, there are 13266 water, 1736 Cl^- and 224 Na^+ molecules. For PA2, there are 19366 water, 1301 Cl^- and 293 Na^+ molecules. For C_{16}-AGGRVK-NH_2, there are 11486 water, 1188 Cl^- and 180 Na^+ molecules.

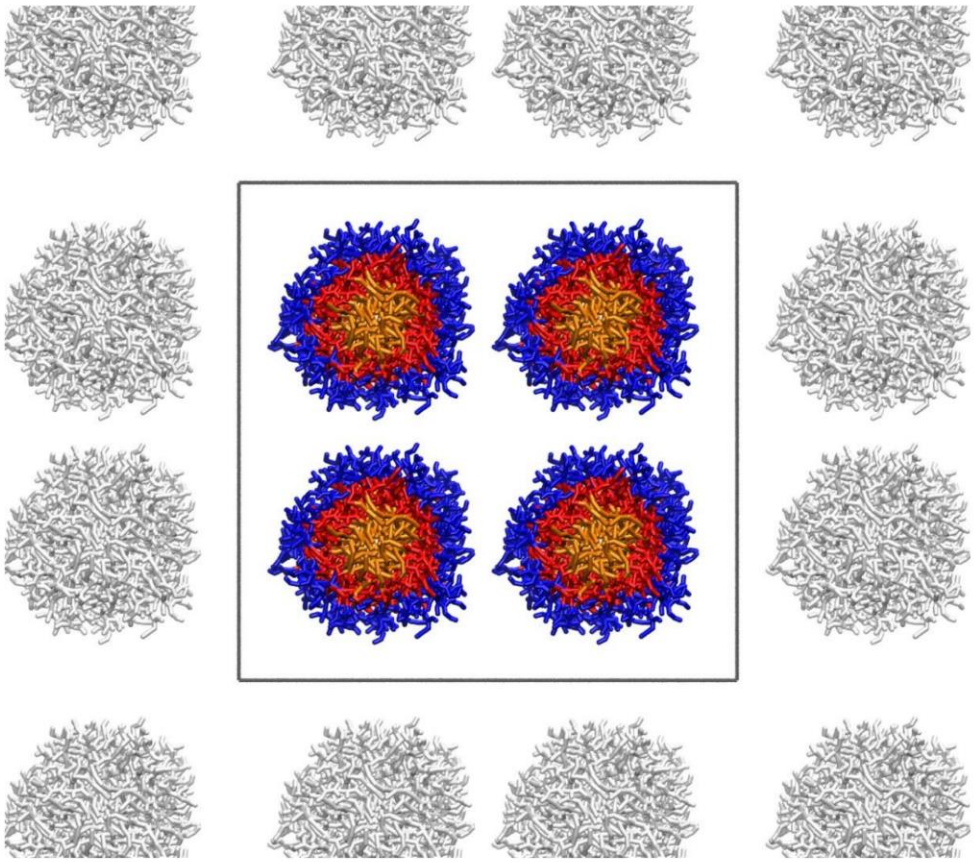


Figure S33 The cross-sectional (top) view of initial configuration of four self-assembled PA fibers for C_{16}-VVVAAAKKK-NH_2 (PA1). The PAs are colored according to the type of residues: C_{16} tail in orange, nonpolar amino acids in red, polar amino acids in purple (none in this PA sequence), and basic amino acids in blue. The periodic images in x and y are shown in white. The simulation box boundaries are drawn in black. Water and ions are omitted for clarity.

To quantitatively represent the aggregation, the distance between the center of mass of fibrils $Dist_{fiber-fiber}$ 1 to 4 over time was measured. Because each PA sequence has a different number of amino acids, the distance was normalized by the diameter of a single fiber D_f to ensure comparability in changes of relative distance. In the initial state (time=0), the $Dist_{fiber-fiber}/D_f$ is > 1.0 because the fibrils are not in contact with each other. If two fibrils are in contact at their surfaces the $Dist_{fiber-fiber}/D_f$ is equal 1.0. If the fibrils aggregate and there is mixing of monomers between the fibrils, the value of $Dist_{fiber-fiber}/D_f$ will be less than 1.

In **Figure S34** the changes of fibril-fibril distance upon multi-fibril simulation are shown and clearly shows the aggregation of PA2 whereas the fibril-fibril distance of PA1 and C_{16}-AGGRVK-NH_2 stay unchanged.

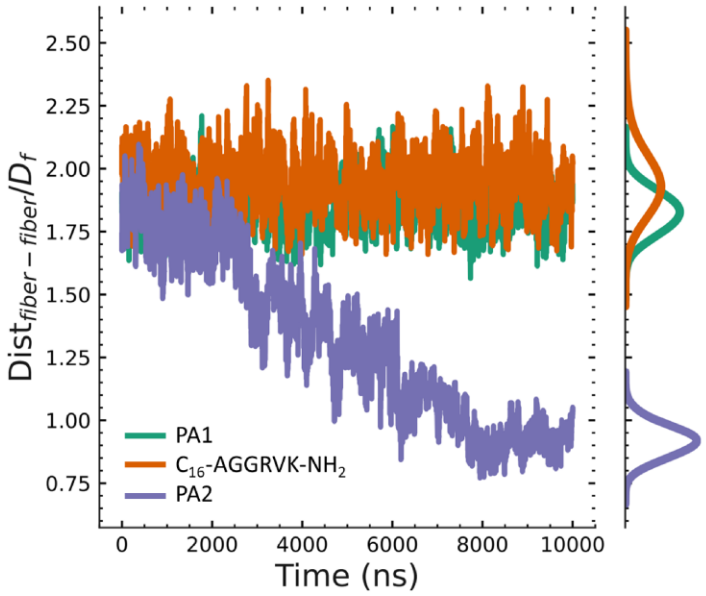


Figure S34 The distance was averaged over 4 fibers, which equals 6 distances.

13.2 Comments on Solvent Accessible Surface Area (SASA)

By calculating the total SASA (normalized by the length of the PA chain) for each simulation, as well as SASA per residue we can learn about the total solvent accessible solvent area of a fibril and how each amino acid in the sequence contributes to this solvent accessibility.

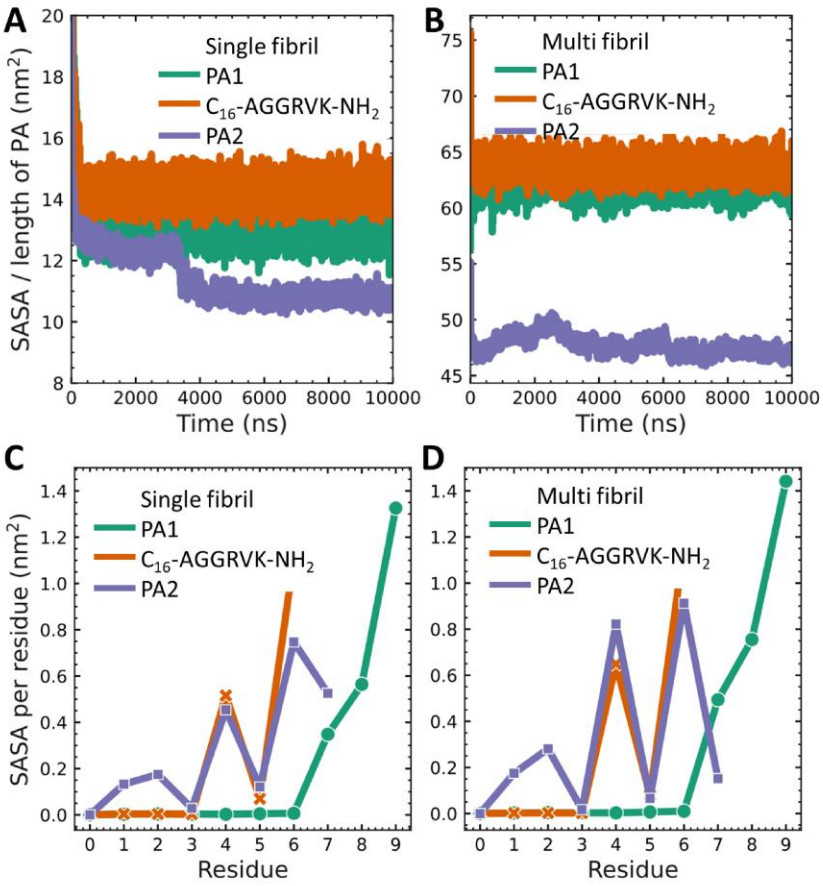


Figure S35 Quantitative evaluation of the solvent accessible surface area (SASA) of single and multiple-fibril simulations. **A** Time-dependent single fibril SASA normalized to the amino acid length of the PA. **B** Time-dependent multi fibril SASA normalized to the amino acid length of the PA**. C** SASA per amino acid residue of the single fibril and **D** of the multi-fibril simulation.

The total SASA shows that both single fibril and multi-fibril simulations have the same trend (**Figure S35A, B**), and the total SASA for PA2 is low compared to the other PAs in multi fibril simulations. This shows the aggregation of the fibrils in PA2.

The SASA per residue shows for example in the case of PA1 that the PA1 is even more hydrophobic than PA2 until the polar, charged lysine residues "KKK "residues that point towards the fibril surface and increase the SASA drastically (**Figure S35C**). For PA2 and C_{16}-AGGRVK-NH_2, the fluctuation in the SASA observed for residues 3-6, is due to the polar and nonpolar amino acids (**Figure S35C**). For both C_{16}-AGGRVK-NH_2 and PA2 (C_{16}-FQFKFKC-NH_2), the residues 4 and 6 are polar and cause an increase of SASA. However, for C_{16}-AGGRVK-NH_2, the polar residue 6 (lysine) is the last residue in the sequence

that results in a polar fibril surface comparable to PA1 (**Figure S35C**). In contrast, for PA2, the SASA decreases for the final nonpolar residue (cysteine), which renders the fibril surface hydrophobic (**Figure S35C**). The contributions of the final amino acids to the increase/decrease of fibril surface hydrophobicity is even more pronounced in multi-fibril simulations (**Figure S35D**).

13.3 Comments on Root Mean Square Fluctuation (RMSF)

The root mean square fluctuation (RMSF) analysis provides quantitative information on how fibrils respond to water vs. presence of other fibrils nearby. This can be evaluated by comparing single and multi-fibril RMSF with each other.

In single fibril simulations, the average RMSF of PA1 and PA2 is similarly low in contrast to the high RMSF of single fibril C_{16}-AGGRVK-NH_2 (**Figure S36A, B**). However, there is a larger deviation for RMSF of PA2 (**Figure S36A**), which means that some monomers are more dynamic (high RMSF) compared to others (low RMSF). The RMSF for PA1 is lower for all residues compared to PA2, which indicates a more stable assembly in solution when there is only a single fibril (**Figure S36A**). However, this observation inverts once multi-fibril interactions are considered (**Figure S36C, D**).

In the multi-fibril scenario, the fibrils already aggregated and reached equilibrium. Now, PA1 and C_{16}-AGGRVK-NH_2 show comparable high RMSF and PA2 a low average RMSF (**Figure S36C, D**) that indicates that the fibrils a PA1 and C_{16}-AGGRVK-NH_2 are less stable and more dynamic than PA2. The high distribution of RMSF for multi-fibril aggregated PA2 indicate that for some parts of the fibril the monomers are more mobile within the aggregated structure than others, but because the aggregated fibril is more stable than not aggregating fibrils the average RMSF for all residues is low.

In other words, fibrils from PA2 aggregate to form a network, in which monomers are still free to move around, but more restricted compared to monomers from PA1 and C_{16}-AGGRVK-NH_2 because PA2 monomers display more of favorable interactions such as low SASA that makes them aggregate. Finite size effects that result from the size of the simulation box play an important role in the network formation. However, we compare different PAs using the same simulation size and only PA2 results in formation of networks, while the other two PA fibrils remain isolated in solution.

For PA1 and C_{16}-AGGRVK-NH_2, the RMSF increases in presence of other fibrils, because there is some interaction between fibrils that is not leading to aggregation, but it is also making them more mobile.

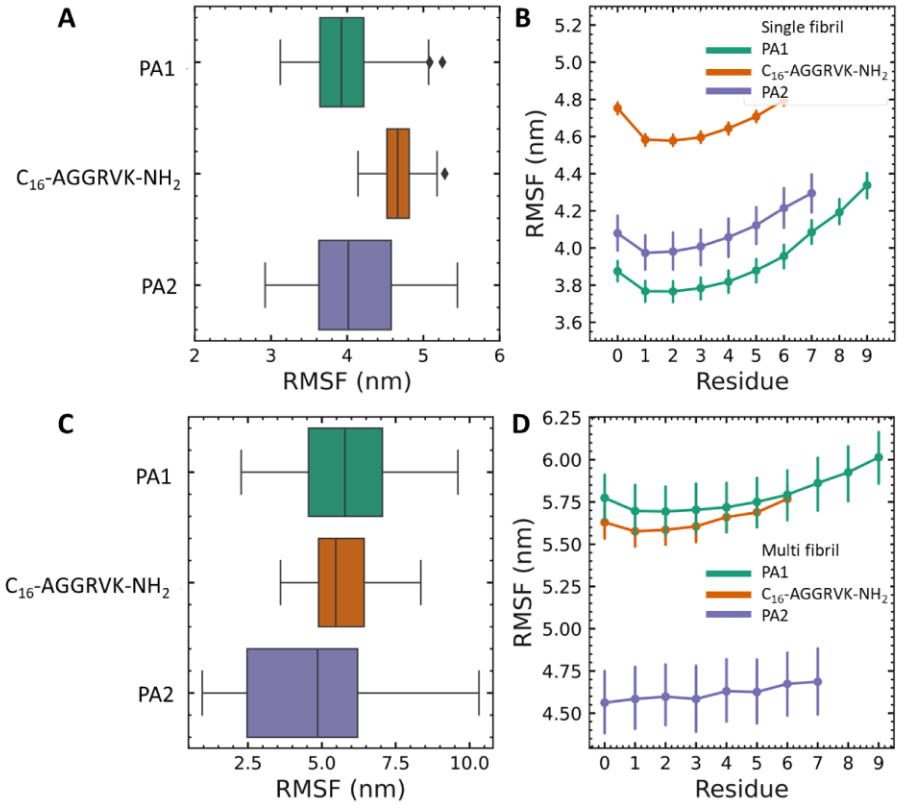


Figure S36 Root mean square fluctuation (RMSF) analysis of single and multi fibril PA simulations. **A** Box plot showing the distribution of RMSF in a single fibril. Solid line represents the median, the boxes show 50% and the whiskers 100% of the data. **B** Mean RMSF per residue in single fibrils, error bars represent standard deviation. **C** Box plot showing the distribution of RMSF in multi fibril simulations. Solid line represents the median, the boxes show 50% and the whiskers 100% of the data. **D** Mean RMSF per residue in multi fibril simulations, error bars represent standard deviation.

14 References

[1] E. Goormaghtigh, V. Cabiaux, J. Ruysschaert, *Eur. J. Biochem.* **1990**, *193*, 409.

[2] D. M. Byler, H. Susi, *Biopolymers* **1986**, *25*, 469.

[3] M. Jackson, H. H. Mantsch, *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95.

[4] A. Adochitei, G. Drochioiu, *Rev. Roum. Chim.* **2011**, *56*, 783.

[5] V. A. Wagoner, M. Cheon, I. Chang, C. K. Hall, *Proteins Struct. Funct. Bioinforma.* **2014**, *82*, 1469.

[6] R. Sarroukh, E. Cerf, S. Derclaye, Y. F. Dufrêne, E. Goormaghtigh, J. M. Ruysschaert, V. Raussens, *Cell. Mol. Life Sci.* **2011**, *68*, 1429.

[7] E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclayes, Y. F. Dufrênes, V. Narayanaswami, E. Goormaghtigh, J. M. Ruysschaert, V. Raussens, *Biochem. J.* **2009**, *421*, 415.

[8] R. Sarroukh, E. Goormaghtigh, J.-M. Ruysschaert, V. Raussens, *Biochim. Biophys. Acta - Biomembr.* **2013**, *1828*, 2328.

[9] A. Gustot, V. Raussens, M. Dehousse, M. Dumoulin, C. E. Bryant, J. M. Ruysschaert, C. Lonez, *Cell. Mol. Life Sci.* **2013**, *70*, 2999.

[10] M. Adachi, M. So, K. Sakurai, J. Kardos, Y. Goto, *J. Biol. Chem.* **2015**, *290*, 18134.

[11] S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, K. Kaygisiz, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, *31*, 2009382.

[12] C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil, B. Knöll, *Adv. Funct. Mater.* **2019**, *29*, 1809112.

[13] K. Ryu, G. J. Lee, J.-Y. Choi, T. Kim, T.-I. Kim, *Adv. Mater. Sci. Eng.* **2015**, *2015*, 1.

[14] R. Shankar, A. Samykutty, C. Riggin, S. Kannan, U. Wenzel, R. Kolhatkar, *Mol. Pharm.* **2013**, *10*, 3776.

[15] D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, S. J. Marrink, *J. Chem. Theory Comput.* **2013**, *9*, 687.

[16] O.-S. Lee, V. Cho, G. C. Schatz, *Nano Lett.* **2012**, *12*, 4907.

[17] C. Li, A. Iscen, H. Sai, K. Sato, N. A. Sather, S. M. Chin, Z. Álvarez, L. C. Palmer, G. C. Schatz, S. I. Stupp, *Nat. Mater.* **2020**, *19*, 900.

[18] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comput. Chem.* **2009**, *30*, 2157.

[19] S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta, S. J. Marrink, *PLoS Comput. Biol.* **2010**, *6*, e1000810.

[20] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Comput. Phys. Commun.* **1995**, *91*, 43.

[21] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19.

[22] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*.

[23] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182.

[24] S. Nosé, M. L. Klein, *Mol. Phys.* **1983**, *50*, 1055.

[25] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463.

[26] H. Grubmüller, H. Heller, A. Windemuth, K. Schulten, *Mol. Simul.* **1991**, *6*, 121.

[27] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.

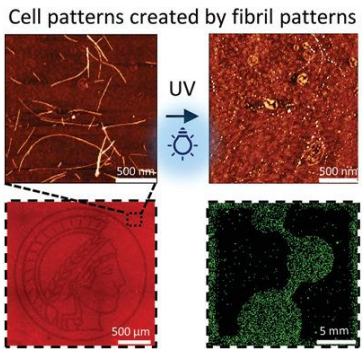
6.7 Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning

Authors: Kübra Kaygisiz[#], Adriana M. Ender[#], Jasmina Gačanin, L. Alix Kaczmarek, Dimitrios A. Koutsouras, Abin N. Nalakath, Pia Winterwerber, Franz J. Mayer, Hans-Joachim Räder, Tomasz Marszalek, Paul W. M. Blom, Christopher V. Synatschke*, Tanja Weil*
[#] shared first authorship * corresponding author
Published in: Macromolecular Bioscience, **2023,** 23, 2200294.
DOI: 10.1002/mabi.202200294
Date of Publication: 25th of October 2022

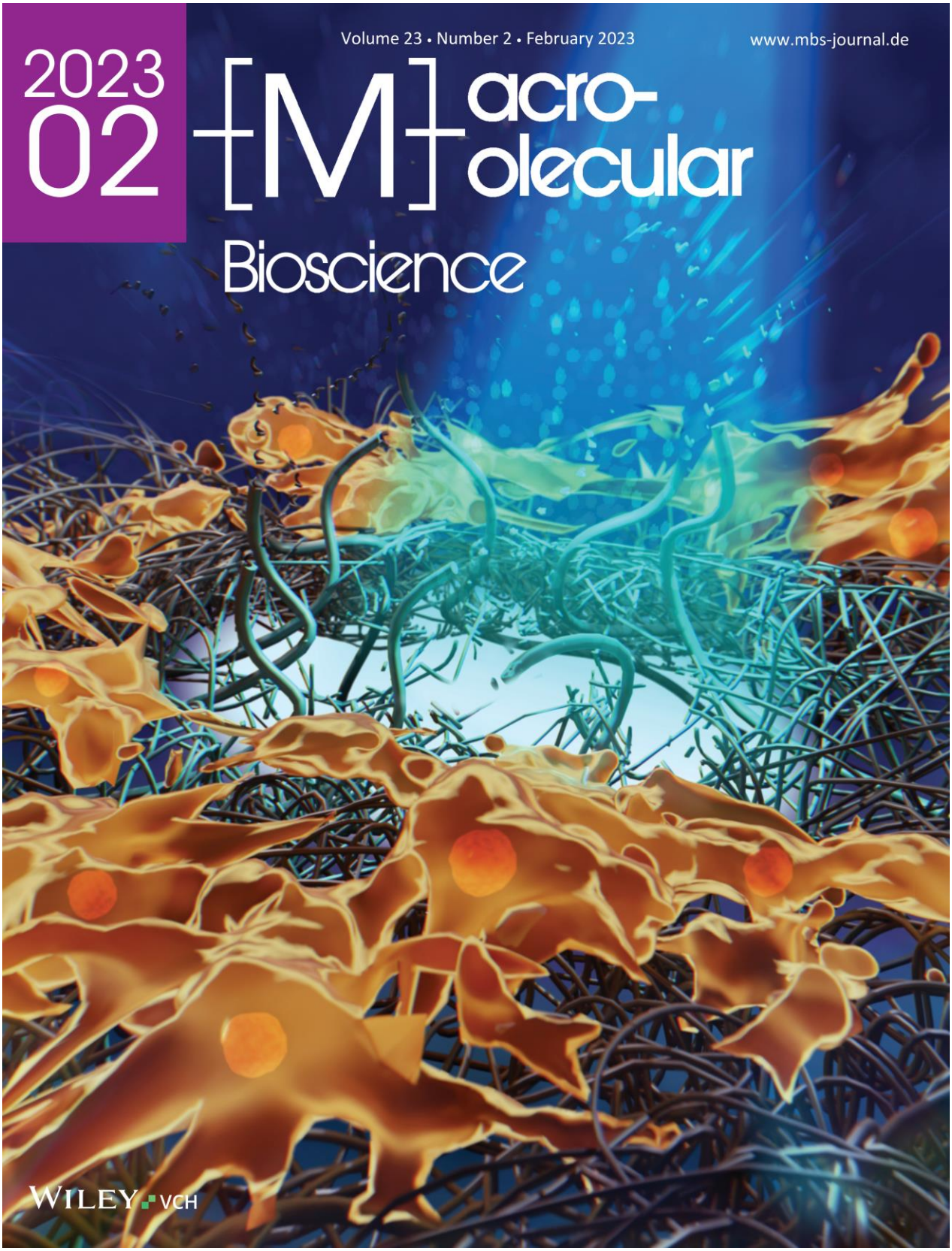
Author Contributions:
Kübra Kaygisiz conceptualized, designed, carried out the experiments, analyzed data, performed the formal analysis related to linker synthesis and characterization and peptide design, synthesis and characterization. She prepared parts of fibril and agarose coatings, performed and analyzed parts of the irradiation studies and cell-patterning assays, and wrote the original draft. Adriana M. Ender conceptualized, designed, carried out the experiments related to peptide characterization, cell-viability assays and wrote the original draft. Kübra Kaygisiz and Adriana M. Ender contributed equally to this work. Jasmina Gačanin conducted parts of the cell-assays, analyzed data and revised the manuscript. L. Alix Kaczmarek performed parts of peptide synthesis and characterization. Dimitrios A. Koutsouras performed parts of photo-patterning, supervised by Paul W. M. Blom. Abin N. Nalakath performed parts of the agarose coatings, supervised by Tomasz Marszalek. Pia Winterwerber performed and analyzed parts of the AFM measurements. Franz J. Mayer and Hans-Joachim Räder performed and analyzed MALDI-ToF MS Imaging. Tomasz Marszalek and Paul W. M. Blom provided the funding, and revised the manuscript. Christopher V. Synatschke conceived the idea of cell-adhesive peptide coating, supervised and revised the manuscript. Tanja Weil conceived the idea, conceptualized, supervised, provided the funding, and revised the manuscript.

Copyrights: The following part is published in *Macromol. Biosci.* **2023**, 23, 2200294. https://doi.org/10.1002/mabi.202200294 and licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/). Further distribution of the material is possible if appropriate credit, a link provided to the license, and indications of changes are made. The cover is reproduced from *Macromol. Biosci.* **2023**, 23, 2200294 with permission of 2023 Wiley.

Graphical abstract:



Photoreactive self-assembling peptide sequences can form cell-adhesive fibrillar structures. The fibril morphology can locally be destroyed through an external trigger, that is UV-light irradiation, resulting in immediate loss of β-sheet structures, nanofibril morphology, and cell adhesiveness of the peptides. Thus, coatings of the photocleavable self-assembling fibrils can be applied for controlled cell attachment via a photopatterning strategy.


2023
02
Volume 23 • Number 2 • February 2023
www.mbs-journal.de
[M]acromolecular
Bioscience
WILEY VCH

Check for updates

RESEARCH ARTICLE

Macromolecular Bioscience
www.mbs-journal.de

Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning

Kübra Kaygisiz, Adriana M. Ender, Jasmina Gačanin, L. Alix Kaczmarek, Dimitrios A. Koutsouras, Abin N. Nalakath, Pia Winterwerber, Franz J. Mayer, Hans-Joachim Räder, Tomasz Marszalek, Paul W. M. Blom, Christopher V. Synatschke, and Tanja Weil**

Amyloid-like fibrils are a special class of self-assembling peptides that emerge as a promising nanomaterial with rich bioactivity for applications such as cell adhesion and growth. Unlike the extracellular matrix, the intrinsically stable amyloid-like fibrils do not respond nor adapt to stimuli of their natural environment. Here, a self-assembling motif (CKFKFQF), in which a photosensitive *o*-nitrobenzyl linker (PCL) is inserted, is designed. This peptide (CKFK-PCL-FQF) assembles into amyloid-like fibrils comparable to the unsubstituted CKFKFQF and reveals a strong response to UV-light. After UV irradiation, the secondary structure of the fibrils, fibril morphology, and bioactivity are lost. Thus, coating surfaces with the pre-formed fibrils and exposing them to UV-light through a photomask generate well-defined areas with patterns of intact and destroyed fibrillar morphology. The unexposed, fibril-coated surface areas retain their ability to support cell adhesion in culture, in contrast to the light-exposed regions, where the cell-supportive fibril morphology is destroyed. Consequently, the photoresponsive peptide nanofibrils provide a facile and efficient way of cell patterning, exemplarily demonstrated for A549, Chinese Hamster Ovary, and Raw Dual type cells. This study introduces photoresponsive amyloid-like fibrils as adaptive functional materials to precisely arrange cells on surfaces.

1. Introduction

Nature creates function through complex self-organization processes. In the past few decades, great progress has been made in mimicking these ordered nanostructures with synthetic biomaterials.[1] To mimic natural cell distribution and study cell migration and communication, the spatial control of cell-attachment (cell patterning) is of great interest.[2–4] The most frequently used approaches for cell-patterning focus on controlling the stiffness and texture of substrates or the position of cell attractive molecules on inorganic substrates or biomaterials.[2,5–8] However, these methods usually lack spatiotemporal control of the morphology of the cellular environment, which plays an essential part in natural matrices that are constantly reshaped by various stimuli.[9] For instance, directed cell alignment is controlled by patterning of fibrillar structures from self-assembling peptides (SAPs).[10] These fibrillar morphologies have crucial functions in living cells; and are thus, highly prevalent in the extracellular matrix to enhance material–cell interactions to support proliferation and cell infiltration at the mesoscopic length scale.[11–13]

Amyloid fibrils are a subclass of SAPs and nature's intrinsically bioactive nanofibrils. They are characterized by their high cross β-sheet content, mechanical stiffness, stability toward enzymatic degradation, and strong adhesion to various substrates.[14] The best known are amyloid structures in misfolded proteins that lead to insoluble plaques associated with several neurodegenerative diseases.[15–17] However, naturally occurring amyloids fulfill various purposes in organisms as scaffold materials.[18,19] For example, they are involved in the formation of bacteria biofilms[20,21] or as scaffolds for hormone storage[22] and melanin polymerization.[23] Recently, the potential of amyloid nanostructures serving in beneficial roles has been exploited in 3D cell culture, gene delivery, tissue, and neuronal regeneration, as well as in drug delivery.[24–26] This is due to the cell adhesive properties that were found for some artificial amyloid fibrils. However, because of the inherent stability of the amyloid fibrils, they cannot readily be quantitatively disassembled in a short time,[27] which would be a prerequisite to study morphology dependent

K. Kaygisiz, A. M. Ender, J. Gačanin, L. A. Kaczmarek, P. Winterwerber, F. J. Mayer, H.-J. Räder, C. V. Synatschke, T. Weil
Department Synthesis of Macromolecules
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
E-mail: synatschke@mpip-mainz.mpg.de; tweil@mpip-mainz.mpg.de
D. A. Koutsouras, A. N. Nalakath, T. Marszalek, P. W. M. Blom
Department of Molecular Electronics
Max Planck Institute for Polymer Research
Ackermannweg 10, 55128 Mainz, Germany
T. Marszalek
Department of Molecular Physics
Faculty of Chemistry
Lodz University of Technology
Zeromskiego 116, Lodz 90-924, Poland

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/mabi.202200294

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/mabi.202200294

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

cell-adhesive properties, for example, for the application as adaptive scaffold biomaterials. Synthetic amyloid-like short SAPs that spontaneously form β-sheet-rich nanofibrils through noncovalent interactions[25,28] are not related to amyloids found in neurodegenerative diseases and exhibit several distinctively different features. They can be chemically customized to design next generation functional biomaterials with controllable stability and bioactivity.[17,29,30]

To gain control over the formation of amyloid structures and thereby their bioactivities, several groups have established strategies to disassemble already formed fibrils. Changing environmental parameters such as peptide concentration and pH can lead to disassembly as these parameters affect the equilibrium of fibril formation.[31] Furthermore, additives such as ruthenium(II)-complexes,[32] fullerenes,[33] water–ethanol mixtures,[34] amphiphilic polyphenylene dendrimers,[35] PEG-chains,[36] carbon nanotubes,[37] and molecular chaperons[38,39] have been applied to disassemble amyloid fibrils. However, these methods cannot be applied for controlled spatiotemporal modification after fibrils are embedded in complex environments. Instead, stimulus-sensitive reactive groups have been incorporated in the peptide sequence to initiate morphological changes of fibrils.[40] For example, azobenzene and hydrazone groups that undergo a *cis–trans* isomerization after irradiation[41–43] or depsipeptides that reveal a pH-induced structural transition from a kinked to a linear structure[44,45] have been proposed to control peptide assembly and disassembly. In these instances, highest spatiotemporal control can only be achieved with light. Since the first pioneering works on photoresponsive manipulation of the cell environment,[46–49] a variety of spatial cell-patterning techniques have been applied.[50] The overwhelming amount of reports modulates cell attachment by introducing or removing bioactive molecules[51–54] or locally changing stiffness of the polymeric (bio)material by crosslinking or bond scission.[55–57] Photoresponsive approaches for modulating structural changes in self-assembly of peptides are rarely reported[58] and not applied for cell-patterning. *o*-Nitrobenzyl ester is a well-established photoresponsive molecule and can be combined with peptides to attach bioactive molecules.[11,59] For example, using this photocleavable linker (PCL) attached to amyloid-like fibrils, we previously created functional cell-gradients by photocleavage of a bioactive moiety from the fibrillar backbone.[60] It is important to note that in these published reports, the fibrillar backbone of the nanofibrils remained intact and also cell-adherent to some extent, limiting the generation of unique cell patterns.

In this work, we solve this limitation by introducing the PCL cleavage motif into the backbone of an amphiphilic SAP sequence. UV-irradiation fully destroys the amyloid-like fibrillar morphology, resulting in a complete loss of bioactivity. Consequently, cell-patterning is achieved by coating the fibrils on surfaces and irradiating them with UV-light through a photomask. Our strategy demonstrates a direct connection between amyloid morphology and bioactivity, which is exploited to study cell attachment and migration in artificial cell-matrices. We envision that the spatiotemporal control over structural integrity gives access to tailored and customizable 4D cell-culture.

2. Results and Discussion

2.1. Design, Synthesis, and Characterization of the Photoresponsive Self-Assembling Peptide

The peptide sequence CKFKFQF assembles into cell-adhesive nanofibrils in aqueous media (Figure S1, Supporting Information) and was used as scaffold for the design of photocleavable nanofibrils that provide the capacity to disintegrate upon irradiation. CKFKFQF forms biocompatible amyloid-like structures, which were previously shown to enhance neuronal outgrowth in vitro.[61] We hypothesized that the bioactivity can be traced back to the fibril morphology and speculated that changes in sequence length or amphiphilicity will affect the assembly and bioactivity for short sequences such as CKFKFQF. Furthermore, due to its amphiphilic sequence composed of hydrophobic, non-charged phenylalanine (F) and hydrophilic, cationic Lysine (K) amino acids, this peptide represents an ideal starting point to study the effect on assembly by cleaving the hydrophilic charged peptide part from the lipophilic part. In our concept, a PCL was implemented between the third (F) and fourth (K) amino acid of this sequence, resulting in CKFK-PCL-FQF 1. Even after integration of the PCL linker, this sequence formed amyloid-like nanofibrils serving as bioactive scaffold and stimulating cell adhesion and growth. UV irradiation leads to a molecular scission of PCL, resulting in the disintegration of the nanofibrils and the formation of fragments 2 and 3 (**Figure 1**A), which cannot form fibrils (Figure S1, Supporting Information). In addition, the non-photoreactive control peptide CKFK-NCL-FQF (4, Scheme S1, Supporting Information) was synthesized and characterized (Figures S8 and S9, Supporting Information).

Peptide synthesis of 1 and 4 was conducted via solid phase peptide synthesis (Scheme S1, Supporting Information) using Wang resin. The PCL as well as NCL linkers were synthesized and conjugated to the solid phase bound peptides as described previously.[60] Purification of the peptides was accomplished applying reversed-phase high performance liquid chromatography (rp-HPLC). Successful synthesis and purity (>95%) were validated by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-ToF-MS) and liquid chromatography mass spectrometry (LC-MS), respectively (Figure 1B; Figure S2, Supporting Information). Next, the photocleavage kinetics were determined by time-dependent measurements, in which aliquots of 1 (1 mg mL^{-1} in water) were analyzed by HPLC starting with a non-irradiated sample and irradiation times of up to 20 min (Figure 1C). Using LC-MS measurements (Figure S3, Supporting Information), the signal with the retention time of 10.2 min was assigned to the intact peptide showing a mass-to-charge ratio (m/z) of 1285.70, matching the protonated $[1+H]^+$ adduct. After 10 min UV treatment at 365 nm, less than 10% of the intact peptide was detected (Figure 1D). An additional kinetic study was performed, by incubating 1 in dimethylsulfoxide (DMSO). After 1 min of UV-irradiation, complete photocleavage of the peptide CKFK-PCL-FQF was observed (Figure S4, Supporting Information), indicating that the cleavage kinetics strongly depend on the selected solvent.[62] In contrast, UV-irradiation of the non-cleavable control peptide CKFK-NCL-FQF did not alter its structure nor its elution time after 10 min of

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macro-olecular Bioscience
www.mbs-journal.de

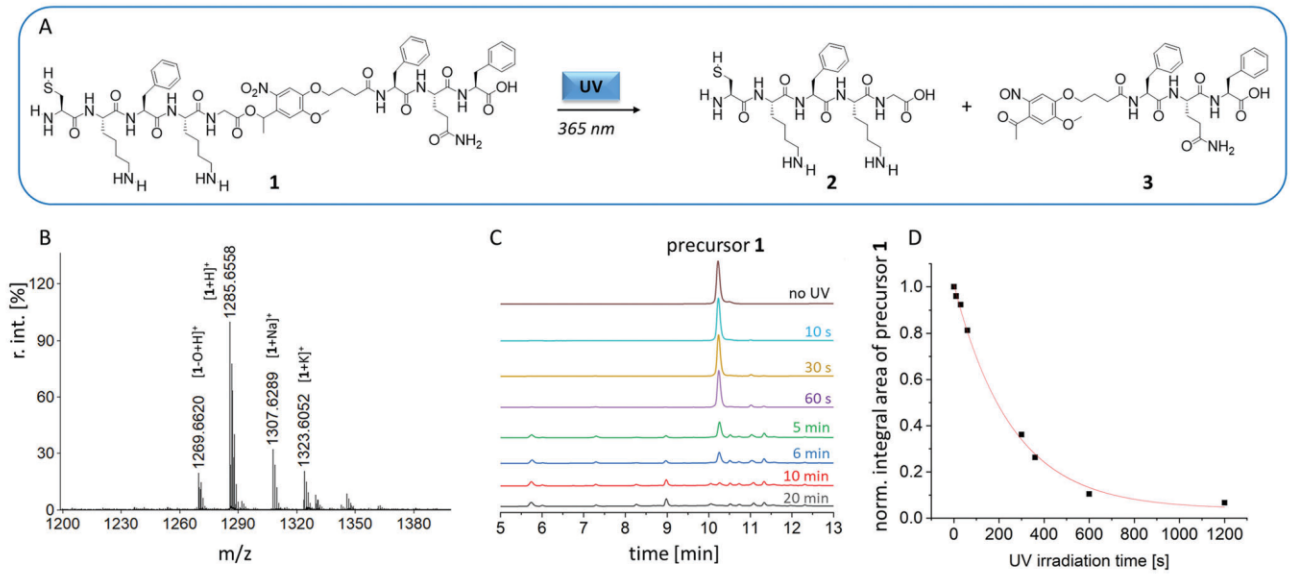


Figure 1. A) Scheme of molecular cleavage of the designed peptide CKFK-PCL-FQF 1 under UV-irradiation. B) MALDI-ToF-MS spectra of the purified peptide CKFK-PCL-FQF, 1, MS calcd. for $[1+H]^+$, 1285.59 g mol $^{-1}$; found m/z 1285.66, calcd for $[1+Na]^+$, 1307.57 g mol^{-1}; found 1307.63 m/z, calcd. for $[1\text{-}O+H]^+$, 1269.60 g mol^{-1} found m/z 1269.66. The loss of oxygen ascribed with the mass 1269.66 $[1\text{-}O+H]^+$ is caused by high laser power of the MALDI-ToF-MS measurement in HCCA matrix.[75] C) HPLC spectra of CKFK-PCL-FQF peptide after different UV-irradiation times in aqueous solution. After 10 min irradiation, the precursor molecule 1 is completely cleaved. Several decomposition products of 1 are detected via LC-MS (Figure S3, Supporting Information). Absorbance detected at 254 nm. D) First-order kinetics of the integral area from LC traces of the precursor molecule 1 at t_{Ret} 10.2 min normalized to 1 without irradiation (no UV).

irradiation in water (Figures S10–S12, Supporting Information) or DMSO (Figure S1, Supporting Information).

2.2. Secondary Structure and Morphology of Amyloid-Like Fibrils Are Altered After Irradiation

We designed CKFK-PCL-FQF to form amyloid-like fibrils in water at pH 7.4 with similar morphology (**Figure 2**E) as the original bioactive peptide sequence CKFKFQF (Figure S1, Supporting Information).[63] The only difference in the structure is the PCL linker, which serves as cleavage site for controlled linker degradation (Figure 2A). After UV irradiation of the CKFK-PCL-FQF nanofibrils, the β-sheet secondary structure was lost, as evidenced by the decrease of the characteristic cotton peaks for β-sheet structure[64] elements at 196 and 222 nm in circular dichroism (CD) measurements (Figure 2B).

The photoinduced cleavage and disintegration of the β-sheet structures were further supported by attenuated total reflectance fourier transform infrared (ATR-FTIR) spectroscopy measurements of lyophilized peptides (Figure 2C). Here, the non-irradiated fibrils showed absorbance maxima in the amide I region at 1630 and 1664 cm^{-1} corresponding to β-sheet and β-turn structures, respectively, as well as absorbance in the amide II region at ≈1520 cm^{-1}, corresponding to β-sheet structures.[65] Upon irradiation of the fibrils in solution, the peak at 1630 cm^{-1} in the amide I region disappeared, whereas the peak at 1664 cm^{-1} increased and the amide II band at 1520 cm^{-1} decreased compared to the non-irradiated sample. These experimental results indicated loss of β-sheet structures and increase of non-ordered and α-helical structural elements.[44,66]

To elucidate whether the β-sheet structures of the fibrils dispersed in aqueous media, the Thioflavin-T (ThT) assay was conducted, which indicates the presence of amyloid structures by an increase in fluorescence intensity of the ThT molecule after binding to ordered β-sheets structures.[67] In this assay, fluorescence intensity of ThT decreased after irradiation of CKFK-PCL-FQF fibrils (Figure 2D), revealing the reduction of β-sheet-rich amyloid-like structures. Interestingly, the ThT fluorescence further decreased, when incubating the irradiated sample in solution for 72 h. This was in contrast to the non-irradiated sample, which showed no decrease in ThT-fluorescence after 72 h. The time-dependent decrease in ThT-fluorescence indicates a hysteretic degradation of irradiated fibrils. Transmission electron microscopy (TEM) images that were recorded directly after irradiation for 20 min (Figure 2E) revealed fragments of broken fibrils, which further transformed to irregular aggregates during the selected observation time of 72 h (Figure S5, Supporting Information). Fragments of broken fibrils may retain the ability to bind ThT to a certain extent whereas the irregular aggregates cannot. Obviously, the supramolecular structures were destabilized over a prolonged period of time whereas the intermolecular changes altering the packing of the peptide monomers within the nanofibrils already occurred during the irradiation process.

Liquid-mode atomic force microscopy (liquid-AFM) measurements of in situ irradiated nanofibrils allowed us to directly study the fragmentation of the nanofibril during irradiation. A clear transition from elongated nanofibrils into spherical aggregates (Figure 2F) was observed. The degradation of the fibrillar morphology was also supported by turbidity measurements of the fibril solution, which increased directly after

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

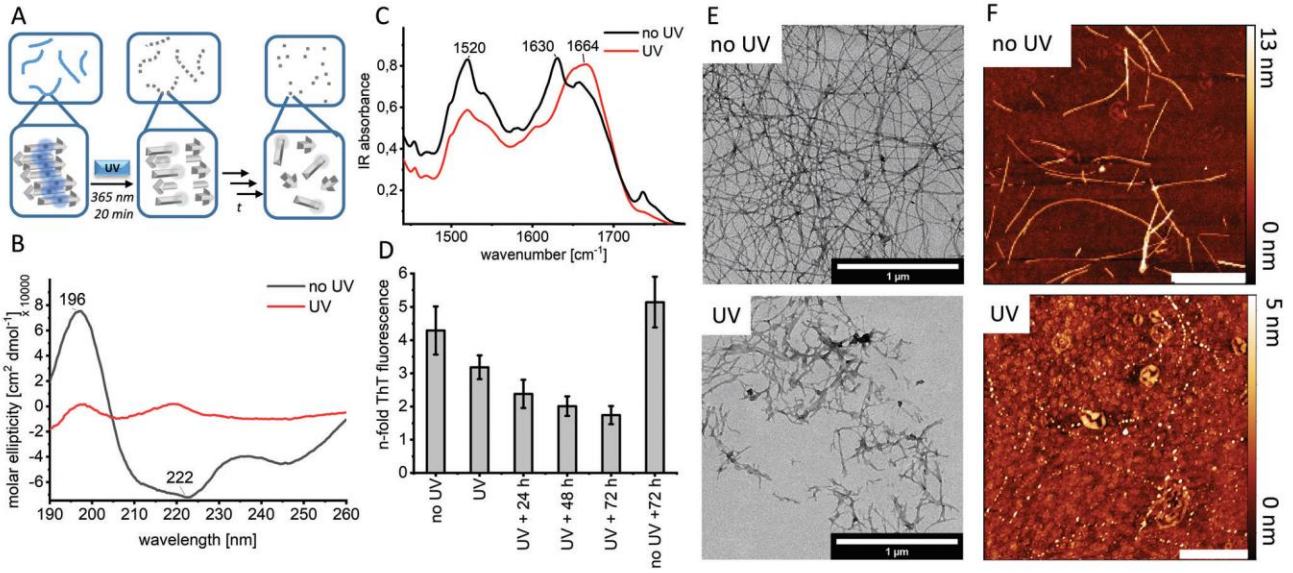


Figure 2. Secondary structure and morphology of CKFK-PCL-FQF fibrils before and after UV irradiation for 20 min. A) Schematic illustration of photocleavage of the peptide sequence upon irradiation, which destabilizes the amyloid-like fibrils. B) CD spectra of non-irradiated versus irradiated fibrils reveal loss of secondary structure after irradiation. C) ATR-FTIR spectra of non-irradiated and irradiated fibril samples indicate loss of β-sheet structures in the amide I and II regions. D) Thioflavin-T Assay of non-irradiated fibrils as well as fibrils irradiated for 20 min at 365 nm and incubated at room temperature for 24, 48, and 72 h. The bar plot shows the n-fold fluorescence relative to ThT fluorescence in water at pH 7.4, excitation 440/10 nm, and emission 488/10 nm. E) TEM micrographs of fibrils before (top) and after irradiation (bottom) displaying morphological changes, scalebar: 1 µm. F) Liquid AFM measurements of the fibrils (top) after in situ irradiation for 20 min (bottom) showing destabilization of fibrils into fragments, scalebar: 500 nm.

irradiation and further over the following three days. We interpreted the increased turbidity by the formation of a large number of aggregates of the newly formed peptide fragments, indicated by enhanced light scattering of the dispersed sample (Figure S6, Supporting Information). The degradation of the nanofibrils did not result in a complete disassembly and solvation of the peptide as shown by microscopy measurements (Figure S5, Supporting Information) and by a conversion assay which quantifies aggregated monomers (Figure S7, Supporting Information).[63] We traced the aggregated structures back to the peptide fragment 3, which is highly lipophilic due to the phenylalanine side chain residues and the *N*-terminal nitrosobenzyl group. Noteworthy, these aggregated peptide fragments did not have β-sheet secondary order as shown by CD and FTIR (Figure 2B,C).

In contrast, the integrity of the control peptide CKFK-NCL-FQF was not affected by photoirradiation (TEM, Figure S11, Supporting Information; AFM, Figure S12, Supporting Information; and ThT-assay, Figure S13, Supporting Information), confirming that the observed structural changes of CKFK-PCL-FQF were due to the photoinduced bond cleavage of the PCL. Photoirradiation of the CKFK-PCL-FQF nanofibrils therefore resulted in immediate loss of the β-sheet structures because of the scission of the peptide backbone in a cationic, hydrophilic (CKFK), and lipophilic (FQF) part. The resulting destabilization of the intermolecular interactions between peptide monomers further led to fragmentation of the fibrillar assembly into aggregates with no β-sheet secondary order.

2.3. Photopatterning of Fibril-Coated Surfaces

Next, we investigated whether photolithography could be used to create defined patterns of intact and degraded amyloid-like nanofibrils to create distinct cell-adhesive regions on substrates (**Figure 3**A). Initially, optical microscope glass slides were covered with a thin agarose layer[60] using a dip coating technique. The agarose layer of ≈50 nm thickness (Figure S14, Supporting Information) shows a cell-repellent character and is applied to suppress nonspecific cell adhesion.[61,68–70] Subsequently, the slides were coated with an amyloid-like fibril layer via drop casting. As agarose forms gels by physical crosslinks via hydrogen bonds, these interactions may also stabilize the attachment of the peptide fibrils coating and agarose layer.[71–73] A photolithography mask was used to cover parts of the peptide film, leaving the rest of it exposed to UV-light. Thereby, a pattern of choice can be imbued on the substrate that was previously coated with CKFK-PCL-FQF.

We applied different photomasks with different resolution for the photopatterning of peptide nanofibrils (Figure S15, Supporting Information). We established mass spectrometry imaging (MALDI-MSI) for characterizing the identity of the formed fragments on the surface. MALDI-MSI directly determines the chemical species on the surface and reveals the distribution of the fragmented peptide **2** with a spatial resolution up to 100 µm (Figure S17, Supporting Information). Moreover, photopatterning was visually demonstrated by the Proteostat assay that only stains intact amyloid structures, which can then be detected by fluorescence microscopy. The surface areas that were directly exposed to

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

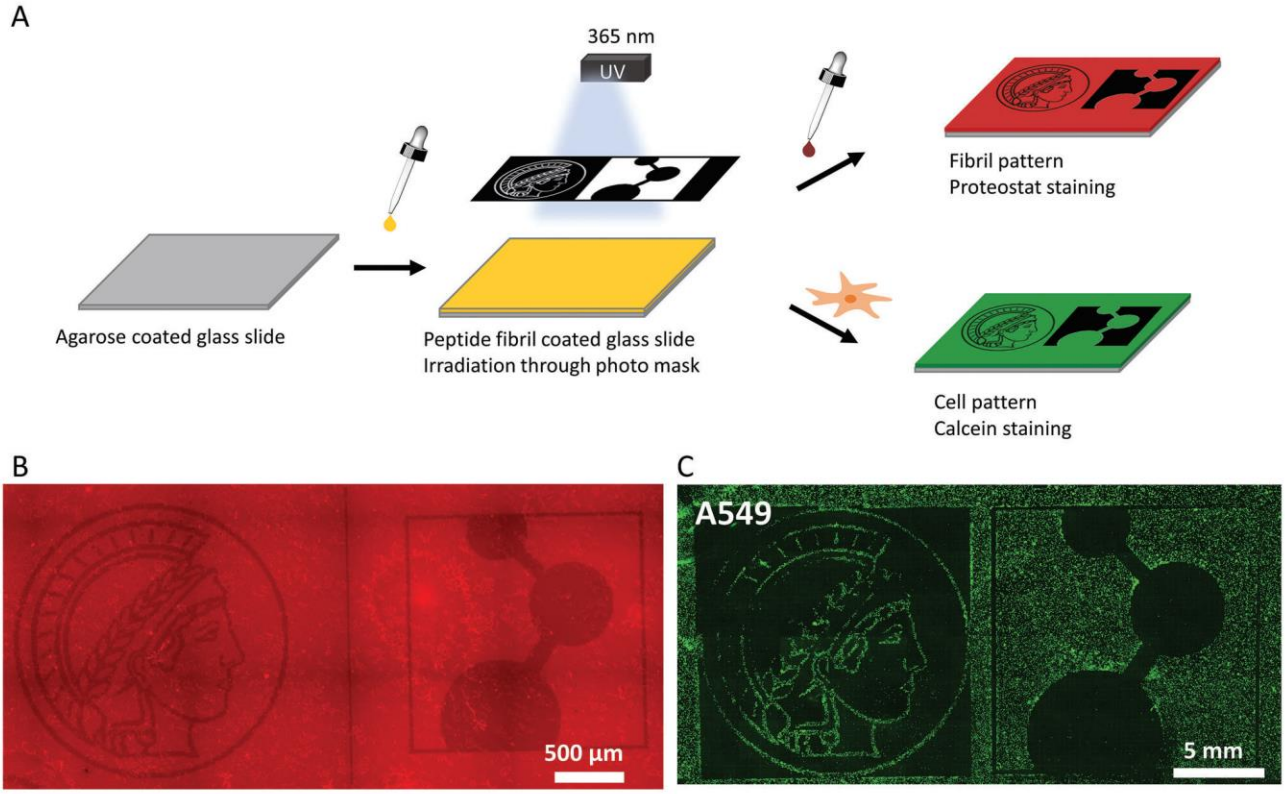


Figure 3. Patterning of glass slides coated with photocleavable amyloid-like fibrils. A) Schematic representation of the workflow. The peptide layer is applied on an agarose pre-treated glass slide and exposed to UV-light through a photomask. The fibril morphology is destroyed where UV-light can reach the surface, which is visualized via fluorescence microscopy of Proteostat-stained samples. Before seeding the cells, the photopatterned surface is stored for 5 h in water to allow morphology change of the irradiated parts. Cells seeded for 15 h on the patterns only attach to the non-irradiated regions, where fibrils remain intact. B) Fluorescence microscope image of Proteostat-stained photopatterned peptide layer depicting the head of Minerva and MPIP logo (scalebar: 500 µm). C) Fluorescence microscopy image of A549 cells seeded on photopatterned peptide layer. The UV exposure alters the ability of the amyloid-like fibrils to sustain cell culturing and results in cell patterning (scalebar: 5 mm).

UV-light reveal less fluorescence and appear darker compared to the non-exposed regions. Via this photopatterning technique, resolutions down to 10 µm could be achieved (Figure 3B; Figure S16, Supporting Information), which is in the size-range of a single A549 cell and thereby suitable for accurate cell patterning.

2.4. Spatially Controlled Cell Pattern

We have reported previously that bioactive amyloid-like nanofibrils, which bind to cellular membranes[63] and support neuronal outgrowth,[61] require a positive net charge, fibrillar morphology, as well as high content of β-sheet structures. As demonstrated in the previous sections, the intact CKFK-PCL-FQF peptide can provide these features in regard to their morphology (Figure 2) and charge (Figure S22, Supporting Information) but loses them upon irradiation. We therefore aimed to create spatial cell patterns by photoinduced degradation of these fibril properties and applied A549, Chinese Hamster Ovary (CHO), and Raw-Dual type cells as model cell-lines. The cell distribution was analyzed by fluorescence microscopy and attached viable cells were stained in green with calcein (Figure S18, Supporting Information). The CellTiter-Glo assay confirmed the non-toxic character of irradiated and non-irradiated fibril regions (Figure S19, Supporting Information). A549 cells (Figure 3C), CHO, and Raw-Dual cells (Figure S20, Supporting Information) did not attach to irradiated areas whereas non-irradiated areas with intact fibril-coating showed robust cell attachment which was due to differences in morphology and secondary structure of irradiated and non-irradiated fibril areas. Control experiments without peptide coating and with the non-photoresponsive CKFK-NCL-FQF fibrils did not show any cell attachment and cell patterning, respectively (Figure S21, Supporting Information). To determine the resolution limit for cell-patterning, we applied a photomask with elements of different size down to 10 µm (Figure S15B, Supporting Information **Figure 4**A). Cell patterns could be clearly identified for fibril patterned elements larger than 100 µm (Figure 4B-i,ii). Motifs smaller than 40 µm (Figure 4B-iii,C-iii) could not be resolved. In summary, cell patterning was achieved by photopatterning of photocleavable peptides on surfaces. Intact fibrillar morphology induced cellular attachment whereas photocleavage of irradiated fibril surface areas resulted in complete loss of their ability to support cell adhesion.

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

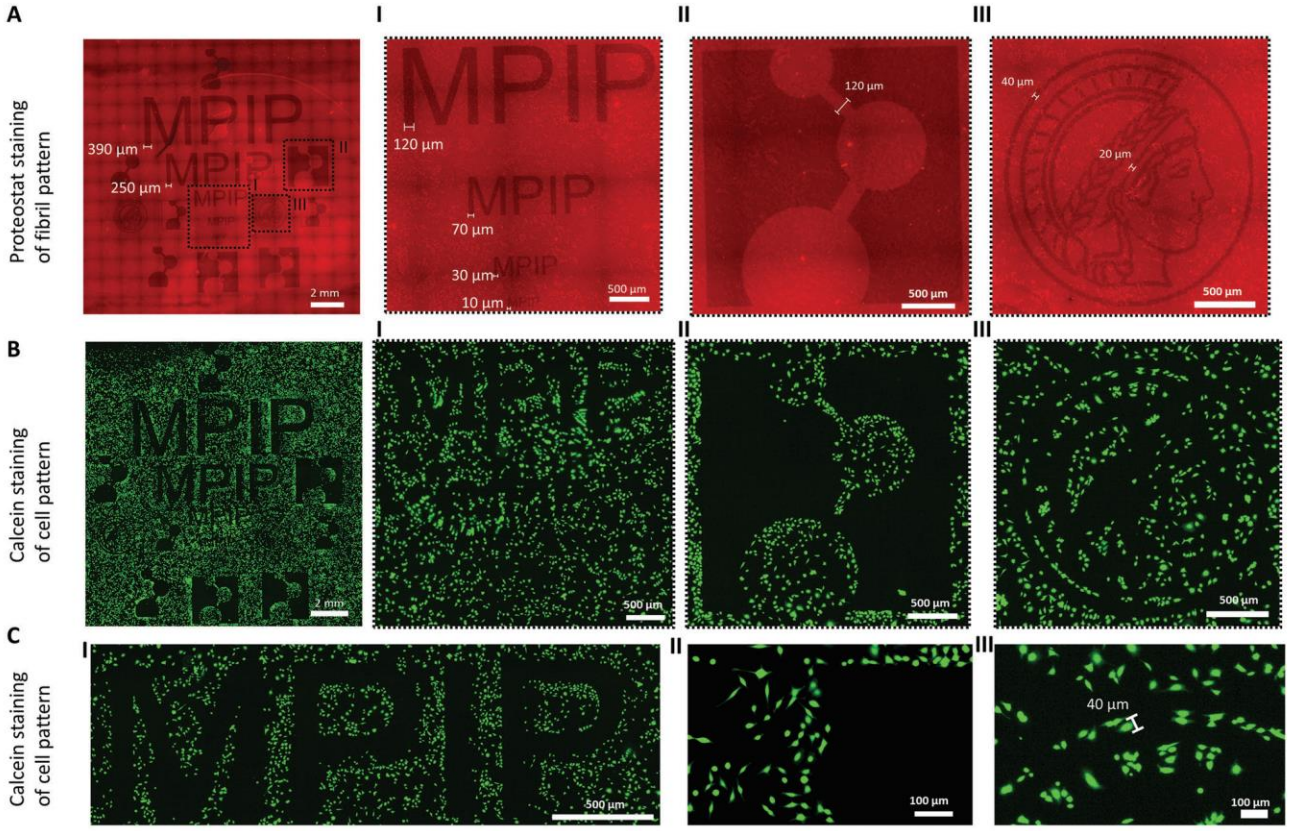


Figure 4. Cell pattern resolution limit shown for A549 cells, which were incubated on a photopatterned peptide fibril surface. A) The photomask (Figure S15B, Supporting Information) enables the creation of motifs of various sizes as shown via Proteostat staining (scalebar: 2 mm). A-i–iii) shows regions of interest depicting the smallest resolvable motif, which is 10 µm as part of the letter "M" (scalebar: 500 µm). B) Cells seeded on this photopatterned surface (B, scalebar: 2 mm) show that cell pattern can be clearly observed for motifs with a size larger than 100 µm (B-i-iii, scalebar: 500 µm) but not for motifs smaller than 40 µm (B-iii, scalebar: 500 µm). C) Enlarged details of B-i-iii) (C-i; scalebar: 500 µm), (C-ii,iii; scalebar: 100 µm) show cell alignment at the borders of the patterning motifs. Distance bars show width of indicated patterning elements.

3. Conclusion

We have successfully demonstrated the synthesis of a photoreactive SAP sequence that forms cell-adhesive fibrillar surface coatings supporting spatially controlled cell attachment. Positioning the PCL at a critical position within the self-assembling sequence resulted, after the irradiation process, in the formation of a positively charged and a lipophilic fragment that were no longer able to form ordered, cell-adhesive nanofibrils. Fibril morphology was locally destroyed through UV-light irradiation as an external trigger resulting in immediate loss of β-sheet structures, nanofibril morphology, and cell adhesiveness of the SAPs. Using photomasks, cell-adhesive and cell-repulsive areas were fabricated with resolutions of up to 10 µm precision. We envision that this method could be applied to create UV-treatable peptide layers on various substrates and various cell-types to realize a versatile cell patterning technique and establish amyloid-like fibrils as adaptive functional materials in biologically relevant model environments. Embedding these photoresponsive peptide fibrils into a 3D cell-matrix for spatiotemporal control over structural integrity may provide access to tailored and customizable 4D cell culture materials. More broadly, these findings are important beyond biomaterials applications by providing a direct correlation between morphology and cell-interactions and for understanding bioactivity of amyloid fibrils in general.

4. Experimental Section

Materials: For peptide synthesis, OxymaPure, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Boc-Gly-OH, and Fmoc-Phe-Wang resin, and (2-(1*H*-benzotriazol-1yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Novabiochem, Merck. Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, piperidine, trifluoroacetic acid (TFA), *N,N*-di*iso*propylethylamine (DIPEA), and *N,N″*-diisopropylcarbodiimide were purchased from Carl Roth GmbH + Co. KG. Peptide-synthesis-grade solvents (dichloromethane and dimethylformamide) were acquired from Acros Organics. Solvents used for synthesis were purchased in HPLC grade purity or higher from Sigma–Aldrich (diethyl ether, DMSO, ethanol), Fisher Scientific (dichloromethane, chloroform, hexane, dimethylformamide, acetone, acetonitrile, ethyl acetate, and tetrahydrofuran), or VWR Chemicals (methanol). Proteostat was purchased from Enzo Life Sciences. Thioflavin-T was purchased from Sigma–Aldrich. Dulbecco's modified eagle medium (DMEM), F-12K Nutrient Mixture Kaighns's Modification (Nut Mix (1X)), heat-inactivated (30 min at 56 °C) fetal bovine serum (FBS), MEM non-essential amino acid solution, penicillin (10.000 U per mL), and streptomycin (10 000 µg mL^{-1}) were purchased from Gibco (Darmstadt, Germany). Normocin and Zeocin were purchased from InvivoGen.

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

Methods: *Linker-Synthesis*: The photocleavable (PCL, 4-(4-(1-(([(9*H*-fluoren-9-ylmethoxy)carbony]glycyl)oxy)ethyl)-2-methoxy-5-nitrophenoxy) butanoic acid) and non-photocleavable (NCL, 4-(4-(2-(2-([(9*H*-fluoren-9-ylmethoxy)carbonyl]amino)acetamido)ethyl) phenoxy)butanoic acid) linkers were synthesized according to literature.[11,60]

Solid-Phase Peptide Synthesis and Characterization of CKFK-PCL-FQF 1 and CKFK-NCL-FQF 4: Peptides were synthesized using an automated microwave peptide synthesizer (CEM, Liberty BlueTM) from *C* to *N*-terminus according to fluorenylmethyloxycarbonyl (Fmoc) solid phase peptide synthesis strategy by Merrifield using the Fmoc-L-Phe-Wang resin (Scheme S1, Supporting Information). The coupling reaction of the PCL[60] (5 equiv.) or NCL[60] (5 equiv.) to the peptide sequence FQF on solid phase was performed manually in DMF with HBTU (5 equiv.) and DIPEA (10 equiv.) for 24 h at room temperature. Successive coupling of the following amino acids (CKFK) was carried out using an automated microwave peptide synthesizer as described in detail in Section S1.1, Supporting Information. Cleavage from the resin was performed by shaking in a solution of TFA containing 2.5% water and 2.5% tri*iso*propylsilane for 2 h at room temperature. This solution was added to cold diethyl ether (40 mL) to afford a white precipitate and afterward centrifuged three times at 4 °C, 4000 rpm for 15 min. The precipitate was dissolved in water with 0.1% TFA and purified via high performance liquid chromatography (HPLC) (Shimadzu modules DGU-20A5R, LC-20AP, CBM-20A, SPD-M20A, SIL-10AP, and FRC-18A with Phenomenex Gemini 5 µm NX-C18, 110 Å, 150 × 30 mm) using a gradient of water and acetonitrile containing 0.1% TFA as the mobile phase at a flow rate of 25 mL min^{-1}. Chromatography was monitored with a UV absorption detector at 214 and 254 nm. After lyophilization, a white solid (yield 10%) was obtained. The peptides were identified via matrix assisted laser desorption/ionization time of flight (MALDI-ToF) mass spectrometry and purity (>95%) was confirmed by LC-MS. MALDI-ToF-MS was conducted on a Bruker rapifleX MALDI-TOF/TOF and a Waters MALDI Synapt G2-SI instrument. The samples were mixed with a saturated α-cyano-4-hydroxycinnamic acid (HCCA) solution in water/ACN 1:1 + 0.1% TFA before measurement and applied to the target via dried droplet method.[74] The data were evaluated with mMass software. LC-MS was conducted on a Shimadzu LC-2020 Single Quadrupole MS instrument with the modules LC-20AD, SIL-20ACHT, SPD-20A, and CTO-using a Kinetex EVO and C18 100 Å LC 50 × 2.1 mm column with 2.6 µm pore size. A gradient of acetonitrile/water mixture with 0.1% formic acid (5% ACN to 95% ACN over 20 min) was used as eluent. Samples were dissolved either in methanol or MilliQ water to a concentration of 0.01 mg mL^{-1} and monitored at 214 nm absorption. The data were processed with LabSolutions and Origin.

Photocleavage Kinetics in Solution: The photocleavage kinetics in solution were conducted analogous to a previous report.[60] Briefly, the peptides were predissolved in DMSO to a concentration of 10 mg mL^{-1} and subsequently diluted to a concentration of 1 mg mL^{-1} with pure MilliQ water. Irradiation was conducted using an LED by Opulent Americas (Starboard Luminus SST-10-UV-A130, $\lambda = 365$ nm) operating at a current of 1 A and a radiant flux of 875 mW with varying exposure times ranging from 0 to 20 min. 50 µL of the samples was injected per run in a Shimadzu system (modules DGU-20A5R, LC-20AT, CBM-20A, SPD-M20A, SIL-10ACHT, and CTO-20AC) with an Agilent ZORBAX Eclipse XDB-C18, 80 Å, 5 µm, 9.4 × 250 mm column at a flow rate of 4 mL min^{-1} and MilliQ water and ACN with 0.1% TFA as a gradual eluent from 5% to 80% ACN in 15 min.

Nanofibril-Formation: The peptides were predissolved in DMSO to a concentration of 10 mg mL^{-1}. The solution was immediately diluted to 1 mg mL^{-1} using MilliQ water. The pH was adjusted to 7.4 ± 0.4 using 0.1 M NaOH and 0.1 M HCl solutions and checked with a pH electrode. To complete fibril formation, the pH-adjusted peptide solution was incubated 24 h at room temperature.

Amyloid-Like Fibril Characterization From Solution: Transmission electron microscopy (TEM) measurements, ThioflavinT (ThT) measurements, and attenuated total reflection Fourier transform infrared (ATR-FTIR) spectra for nanofibril characterization were prepared and conducted analogous to a previous report.[60] Monomer to fibril conversion was determined as previously reported.[63] CD measurements were conducted on a JASCO 1500 instrument in a 1 mm quartz cuvette (HellmaAnalytics) from 190 to 260 nm with preformed fibrils diluted to 0.1 mg mL^{-1}. For the characterization of irradiated samples, the preformed fibrils were irradiated in solution with an LED by Opulent Americas (Starboard Luminus SST-10-UV-A130) operated at 365 nm at a current of 1 A and a radiant flux of 875 mW. Unless indicated otherwise, the characterization was conducted immediately after irradiation. The characterization of irradiated fibrils was conducted analogous to non-irradiated samples.

Precoating of Microscope Slides With Agarose: Glass slides were precleaned with acetone and *iso*propanol. An aqueous agarose solution (1 wt%) was cooked 1 h in order to obtain a fully dissolved and transparent solution. The slides were then dipped in the hot (90 °C) agarose solution for 3 min and thereafter slowly (800 µm s^{-1}) lifted up from the solution with an automated stage. The slides were air-dried before further usage.

Fabrication of Nanofibril-Coated Surfaces: Surface coatings of SAPs were achieved by homogenous distribution and incubation of 300 µL, 0.25 mg mL^{-1} preformed fibrils on agarose-coated microscopy slides (75 by 26 mm) overnight. Excess solvent was removed by flipping and the slides were dried. After photo-patterning, the slides were stored in MilliQ for 5 h before conducting cell studies.

Irradiation of Surface With Photomask: In order to create patterns of irradiated photoresponsive SAP, a photolithography mask (SELBA) was applied. UV exposure was realized with an LED by Opulent Americas (Starboard Luminus SST-10-UV-A130, 20 min, $\lambda = 365$ nm, 1 A, 875 mW) for the patterning of motifs (Figure S15A, Supporting Information) or with a SÜSS MicroTec MA6 mask aligner (10 s, $\lambda = 365$ nm, 9 mW cm^{-2} 350 W Hg) for the photopatterning of motif (Figure S15B, Supporting Information).

Atomic Force Microscopy: Atomic force microscopy (AFM) was conducted in liquid state with a JPK NW III atomic force microscope for compound 1 or with a Bruker Dimension FastScan BioTM atomic force microscope for compound 4. AFM probes with a nominal force constant of 2 N m^{-1} or 0.25 N m^{-1} were used, respectively and operated in tapping mode with a resonance frequency of 70 kHz (OLTESPA-R3, Bruker). Samples were irradiated in situ by using an LED by Opulent Americas (Starboard Luminus SST-10-UV-A130) operating at a current of 1 A, 875 mW at 365 nm for 20 min while being measured. Images were processed with Gwyddion 2.59.

Pattern Analysis by Proteostat Assay: Proteostat-staining was applied as an amyloid-sensitive dye to detect intact fibrils after UV treatment of 1. A Proteostat Kit by Enzo Lifesciences was used for analysis. 1 µL of the assay buffer was diluted in 99 µL MilliQ water. 0.1 µL of the Proteostat stock solution was added and thoroughly mixed and 500 µL of the solution was placed on the sample to cover the peptide-coated area as well as a part of the non-coated area for reference. After incubation in the dark for 15 min, the solution was removed and samples were dried. The measurement was conducted on a Leica Thunder DMI8 microscope coupled to a Leica DFC9000 GTC VSC-12365 camera with HC PL FLUOTAR 10×/0.32 DRY objective. The fluorescence emission was detected at 550 nm upon excitation at 475 nm. Images were processed with Leica Application Suite X (LAS X).

Cell Culture: A549 human wild-type cells, a human alveolar basal epithelial carcinoma cell line (obtained from Abcam), were cultivated in DMEM supplemented with 10% heat-inactivated (30 min at 56 °C) FBS, 1% MEM non-essential amino acid solution as well as 1% penicillin (10 000 U per mL) and 1% streptomycin (10 000 µg mL^{-1}) at 37 °C under a humidified atmosphere with 5% CO_2. Chinese Hamster Ovary (CHO) type cells, an epithelial cell line derived from hamster ovaries was cultivated in F-12K Nut Mix supplemented with 10% heat-inactivated (30 min at 56 °C) FBS, 1% penicillin (10 000 U per mL), and 1% streptomycin (10 000 µg per mL). The RAW-Dual reporter cell line (264.7) was purchased from InvivoGen (San Diego, CA, USA) and cultured in DMEM-GlutaMAXTM medium, which was supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin, 0.02% normocin, and 0.01% zeocin at 37 °C with 5% CO2 saturation. Cells were reseeded at least twice weekly. For splitting of A549 and CHO cells, the cells were washed with DPBS, trypsinated, and resuspended in supplemented culture medium. Splitting of Raw-Dual cells was conducted by scratching cells and resuspending in supplemented culture medium.

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

For the cell patterning, A549 and CHO cells were washed with PBS, trypsinated, centrifuged (4 min, RT, 500 rpm), and resuspended in supplemented medium, and the Raw-Dual cells were scratched, centrifuged (4 min, RT, 500 rpm), and resuspended prior to seeding on the respective surfaces with a density of 5×10^4 cells per petri dish (8 cm diameter) in 16 mL medium. This was followed by overnight (15 h) incubation in the respective supplemented culture medium at 37 °C under a humidified atmosphere with 5% CO_2. The cells were then treated with calcein-AM staining. To this end, the cells were washed three times with PBS. The cell-culture medium was then replaced by 15–16 mL supplemented medium containing 1 µM calcein-AM solution (prepared from 10 mg mL^{-1} solution in DMSO) and the cells were incubated at 37 °C under a humidified atmosphere with 5% CO_2 for 30 min. Live imaging was performed subsequently on a Leica Thunder DMI8 microscope coupled to a Leica DFC9000 GTC VSC-12365 camera with a N PLAN 5×/0.12 DRY objective. The fluorescence emission was detected using a 519 nm filter upon excitation at 475 nm. Images were processed with Leica Application Suite X (LAS X). All cell-culture experiments were repeated at least in triplicates.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

K.K. and A.M.E. contributed equally to this work. The authors thank Uwe Rietzler, Qi Lu for his experimental support, Pierpaolo Moscariello for his comments, and Judith Stickdorn and Alina Heck for cell culture donation and advice. Support by the IMB Microscopy Core Facility is gratefully acknowledged. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 316249678—SFB 1279 (A05, C01) and Projektnummer 441734479, the Federal Ministry of Education and Research of Germany (BMBF) in the framework of ProMatLeben Polymere InGel-NxG (FKZ: 13XP5086F) as well as Kosmogel (FKZ: 13XP5148). T.M. acknowledges the Foundation for Polish Science financed by the European Union under the European Regional Development Fund (POIR.04.04.00-00-3ED8/17). This work was supported by the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC).

Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

Keywords

amyloid degradation, controlled cell attachment, patterning, self-assembling peptide nanofibers, stimuli-responsive biomaterials

Received: July 19, 2022
Revised: October 14, 2022
Published online: November 10, 2022

[1] T. Aida, E. W. Meijer, S. I. Stupp, *Science* **2012**, *335*, 813.
[2] Y. Li, Y. Xiao, C. Liu, *Chem. Rev.* **2017**, *117*, 4376.
[3] J. d'Alessandro, A. Barbier–Chebbah, V. Cellerin, O. Benichou, R. M. Mège, R. Voituriez, B. Ladoux, *Nat. Commun.* **2021**, *12*, 4118.
[4] D. B. Brückner, N. Arlt, A. Fink, P. Ronceray, J. O. Rädler, C. P. Broedersz, *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2016602118.
[5] P. Premnath, A. Tavangar, B. Tan, K. Venkatakrishnan, *Exp. Cell Res.* **2015**, *337*, 44.
[6] M. Ventre, C. F. Natale, C. Rianna, P. A. Netti, *J. R. Soc., Interface* **2014**, *11*, 20140687.
[7] C. M. Kelleher, J. P. Vacanti, *J. R. Soc., Interface* **2010**, *7*, S717.
[8] Y. Lin, M. M. Mazo, S. C. Skaalure, M. R. Thomas, S. R. Schultz, M. M. Stevens, *Chem. Sci.* **2019**, *10*, 1158.
[9] C. T. Mierke, *Front. Phys.* **2021**, *9*, 619.
[10] E. C. Wu, S. Zhang, C. A. E. Hauser, *Adv. Funct. Mater.* **2012**, *22*, 456.
[11] S. Sur, J. B. Matson, M. J. Webber, C. J. Newcomb, S. I. Stupp, *ACS Nano* **2012**, *6*, 10776.
[12] I. Wheeldon, A. Farhadi, A. G. Bick, E. Jabbari, A. Khademhosseini, *Nanotechnology* **2011**, *22*, 212001.
[13] G. Ye, F. Bao, X. Zhang, Z. Song, Y. Liao, Y. Fei, V. Bunpetch, B. C. Heng, W. Shen, H. Liu, J. Zhou, H. Ouyang, *Nanomedicine* **2020**, *15*, 1995.
[14] J. Adamcik, F. S. Ruggeri, J. T. Berryman, A. Zhang, T. P. J. Knowles, R. Mezzenga, *Adv. Sci.* **2021**, *8*, 2002182.
[15] G. G. Glenner, C. W. Wong, *Biochem. Biophys. Res. Commun.* **1984**, *122*, 1131.
[16] A. V. Maltsev, S. Bystryak, O. V. Galzitskaya, *Ageing Res. Rev.* **2011**, *10*, 440.
[17] I. Cherny, E. Gazit, *Angew. Chem., Int. Ed.* **2008**, *47*, 4062.
[18] V. A. Iconomidou, G. Vriend, S. J. Hamodrakas, *FEBS Lett.* **2000**, *479*, 141.
[19] J. E. Podrabsky, J. F. Carpenter, S. C. Hand, *Am. J. Physiol.: Regul., Integr. Comp. Physiol.* **2001**, *280*, R123.
[20] D. Romero, C. Aguilar, R. Losick, R. Kolter, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2230.
[21] A. Taglialegna, I. Lasa, J. Valle, *J. Bacteriol.* **2016**, *198*, 2579.
[22] S. K. Maji, M. H. Perrin, M. R. Sawaya, S. Jessberger, K. Vadodaria, R. A. Rissman, P. S. Singru, K. P. R. Nilsson, R. Simon, D. Schubert, D. Eisenberg, J. Rivier, P. Sawchenko, W. Vale, R. Riek, *Science* **2009**, *325*, 328.
[23] B. Watt, G. van Niel, G. Raposo, M. S. Marks, *Pigment Cell Melanoma Res.* **2013**, *26*, 300.
[24] K. Sato, M. P. Hendricks, L. C. Palmer, S. I. Stupp, *Chem. Soc. Rev.* **2018**, *47*, 7539.
[25] A. K. Das, P. K. Gavel, *Soft Matter* **2020**, *16*, 10065.
[26] N. P. Reynolds, *Biointerphases* **2019**, *14*, 040801.
[27] E. Chuang, A. M. Hori, C. D. Hesketh, J. Shorter, *J. Cell Sci.* **2018**, *131*, jcs189928.
[28] T. Lee, S. Yoo, K. Lee, L. Hwang, Ryou, *Int. J. Mol. Sci.* **2019**, *20*, 5850.
[29] N. Balasco, C. Diaferia, G. Morelli, L. Vitagliano, A. Accardo, *Front. Bioeng. Biotechnol.* **2021**, *9*, 130.
[30] M. Jackson, E. Hewitt, *Biomolecules* **2017**, *7*, 71.
[31] N. Nespovitaya, J. Gath, K. Barylyuk, C. Seuring, B. H. Meier, R. Riek, *J. Am. Chem. Soc.* **2016**, *138*, 846.
[32] G. Son, B. Il Lee, Y. J. Chung, C. B. Park, *Acta Biomater.* **2018**, *67*, 147.
[33] K. Siposova, V. I. Petrenko, O. I. Ivankov, A. Musatov, L. A. Bulavin, M. V. Avdeev, O. A. Kyzyma, *ACS Appl. Mater. Interfaces* **2020**, *12*, 32410.
[34] S. Jordens, J. Adamcik, I. Amar-Yuli, R. Mezzenga, *Biomacromolecules* **2011**, *12*, 187.
[35] S. Xiang, J. Wagner, T. Lückerath, K. Müllen, D. Y. W. Ng, J. Hedrich, T. Weil, *Adv. Healthcare Mater.* **2022**, *11*, 2101854.
[36] P. A. Rühs, J. Adamcik, S. Bolisetty, A. Sánchez-Ferrer, R. Mezzenga, *Soft Matter* **2011**, *7*, 3571.
[37] D. Lin, J. Lei, S. Li, X. Zhou, G. Wei, X. Yang, *J. Phys. Chem. B* **2020**, *124*, 3459.

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Macromolecular Bioscience
www.mbs-journal.de

[38] P. Arosio, T. C. T. Michaels, S. Linse, C. Månsson, C. Emanuelsson, J. Presto, J. Johansson, M. Vendruscolo, C. M. Dobson, T. P. J. Knowles, *Nat. Commun.* **2016**, *7*, 10948.
[39] A. Franco, P. Gracia, A. Colom, J. D. Camino, J. Á. Fernández-Higuero, N. Orozco, A. Dulebo, L. Saiz, N. Cremades, J. M. G. Vilar, A. Prado, A. Muga, *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2105548118.
[40] S. Chagri, D. Y. W. Ng, T. Weil, *Nat. Rev. Chem.* **2022**, *6*, 320.
[41] V. Peddie, A. D. Abell, *J. Photochem. Photobiol., C* **2019**, *40*, 1.
[42] A. A. Deeg, T. E. Schrader, S. Kempter, J. Pfizer, L. Moroder, W. Zinth, *ChemPhysChem* **2011**, *12*, 559.
[43] K. Nakamura, W. Tanaka, K. Sada, R. Kubota, T. Aoyama, K. Urayama, I. Hamachi, *J. Am. Chem. Soc.* **2021**, *143*, 19532.
[44] M. Pieszka, A. M. Sobota, J. Gačanin, T. Weil, D. Y. W. Ng, *ChemBioChem* **2019**, *20*, 1376.
[45] J. Gačanin, J. Hedrich, S. Sieste, G. Glaßer, I. Lieberwirth, C. Schilling, S. Fischer, H. Barth, B. Knöll, C. V. Synatschke, T. Weil, *Adv. Mater.* **2019**, *31*, 1805044.
[46] Y. Luo, M. S. Shoichet, *Nat. Mater.* **2004**, *3*, 249.
[47] T. Matsuda, T. Sugawara, *J. Biomed. Mater. Res.* **1995**, *29*, 749.
[48] D. Ryan, B. A. Parviz, V. Linder, V. Semetey, S. K. Sia, J. Su, M. Mrksich, G. M. Whitesides, *Langmuir* **2004**, *20*, 9080.
[49] S. Zhang, L. Yan, M. Altman, M. Lässle, H. Nugent, F. Frankel, D. A. Lauffenburger, G. M. Whitesides, A. Rich, *Biomaterials* **1999**, *20*, 1213.
[50] F. M. Yavitt, B. E. Kirkpatrick, M. R. Blatchley, K. S. Anseth, *Biomater. Sci. Eng.* **2022**, acsbiomaterials.1c01450.
[51] C. A. DeForest, K. S. Anseth, *Angew. Chem., Int. Ed.* **2012**, *51*, 1816.
[52] R. J. Wade, E. J. Bassin, W. M. Gramlich, J. A. Burdick, *Adv. Mater.* **2015**, *27*, 1356.
[53] D. S. Hernandez, E. T. Ritschdorff, S. K. Seidlits, C. E. Schmidt, J. B. Shear, *J. Mater. Chem. B* **2016**, *4*, 1818.
[54] I. Batalov, K. R. Stevens, C. A. DeForest, *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2014194118.
[55] B. D. Fairbanks, S. P. Singh, C. N. Bowman, K. S. Anseth, *Macromolecules* **2011**, *44*, 2444.
[56] C. Yang, F. W. DelRio, H. Ma, A. R. Killaars, L. P. Basta, K. A. Kyburz, K. S. Anseth, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E4439.
[57] A. M. Kloxin, A. M. Kasko, C. N. Salinas, K. S. Anseth, *Science* **2009**, *324*, 59.
[58] J. B. Matson, Y. Navon, R. Bitton, S. I. Stupp, *ACS Macro Lett.* **2015**, *4*, 43.
[59] S. V. Wegner, O. I. Sentürk, J. P. Spatz, *Sci. Rep.* **2016**, *5*, 18309.
[60] A. M. Ender, K. Kaygisiz, H.-J. Räder, F. J. Mayer, C. V. Synatschke, T. Weil, *ACS Biomater. Sci. Eng.* **2021**, *7*, 4798.
[61] C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil, B. Knöll, *Adv. Funct. Mater.* **2019**, *29*, 1809112.
[62] M. E. Lee, E. Gungor, A. M. Armani, *Macromolecules* **2015**, *48*, 8746.
[63] S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, K. Kaygisiz, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, *31*, 2009382.
[64] N. J. Greenfield, *Nat. Protoc.* **2007**, *1*, 2876.
[65] A. Adochitei, G. Drochioiu, *Rev. Roum. Chim.* **2011**, *56*, 783.
[66] M. Jackson, H. H. Mantsch, *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95.
[67] C. Xue, T. Y. Lin, D. Chang, Z. Guo, *R. Soc. Open Sci.* **2017**, *4*, 160696.
[68] J. Jeong, Y. Lee, Y. Yoo, M. K. Lee, *Colloids Surf., B* **2018**, *162*, 306.
[69] W. Y. Seow, K. Kandasamy, G. S. L. Peh, J. S. Mehta, W. Sun, *ACS Biomater. Sci. Eng.* **2019**, *5*, 4067.
[70] F. Topuz, A. Nadernezhad, O. S. Caliskan, Y. Z. Menceloglu, B. Koc, *Carbohydr. Polym.* **2018**, *201*, 105.
[71] R. Armisen, *Hydrobiologia* **1991**, *221*, 157.
[72] S. Arnott, A. Fulmer, W. E. Scott, I. C. M. Dea, R. Moorhouse, D. A. Rees, *J. Mol. Biol.* **1974**, *90*, 269.
[73] K. Nakamura, R. Kubota, I. Hamachi, *ChemRxiv* **2022**, 10.26434/chemrxiv
[74] A. A. Patil, C.-K. Chiang, C.-H. Wen, W.-P. Peng, *Anal. Chim. Acta* **2018**, *1031*, 128.
[75] X. Zhan, D. M. Desiderio, *Int. J. Mass Spectrom.* **2009**, *287*, 77.

© 2022 The Authors. Macromolecular Bioscience published by Wiley-VCH GmbH

16165195, 2023, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/mabi.202200294 by MPI 355 Polymer Research, Wiley Online Library on [31/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supporting Information

WILEY-VCH

Supporting Information

Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning

Kübra Kaygisiz[#a], Adriana Maria Ender[#a], Jasmina Gačanin[a], Lara Alix Kaczmarek[a], Dimitrios A. Koutsouras[b], Abin N. Nalakath[b], Pia Winterwerber[a], Franz J. Mayer[a], Hans-Joachim Räder[a], Tomasz Marszalek[b, c], Paul W. M. Blom[b], Christopher V. Synatschke[a] and Tanja Weil[a]**

# denotes equal contribution

a. Kübra Kaygisiz, Adriana Maria Ender, Dr. Jasmina Gačanin, Lara Alix Kaczmarek, Pia Winterwerber, Dr. Franz J. Mayer, Dr. Hans-Joachim Räder, Dr. Christopher V. Synatschke and Prof. Dr. Tanja Weil

Department Synthesis of Macromolecules, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

E-mail: synatschke@mpip-mainz.mpg.de, weil@mpip-mainz.mpg.de

Dr. Dimitrios A. Koutsouras, Abin N. Nalakath, Dr. Tomasz Marszalek, Prof. Dr. Paul W. M. Blom

b. Department of Molecular Electronics, Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

Dr. Tomasz Marszalek

c. Department of Molecular Physics, Faculty of Chemistry, Lodz University of Technology Zeromskiego 116, 90-924 Lodz, Poland

WILEY-VCH

Table of Content

1. Methods ... 3
1.1. Microwave assisted solid phase peptide synthesis ... 3
1.2. Turbidity measurements ... 6
1.3. MALDI-MSI ... 6
1.4. Cell Viability Analysis (CellTiter-Glo Assay) ... 6
1.5. Zeta Potential ... 7
2. Additional Characterization CKFK-*PCL*-FQF **1** ... 8
2.1. Comparison of fibril morphology via TEM of CKFKFQF, CKFK-*PCL*-FQF **1** and CKFK-*NCL*-FQF **4** after irradiation of monomers in DMSO ... 8
2.2. LC-MS ... 9
2.3. Photocleavage kinetics of CKFK-*PCL*-FQF **1** in water ... 10
2.4. Photocleavage kinetics of CKFK-*PCL*-FQF **1** in DMSO ... 11
2.5. TEM ... 11
2.6. Turbidity of CKFK-*PCL*-FQF **1** before and after irradiation by DLS ... 12
2.7. Conversion of monomers to aggregated monomers before and after irradiation of CKFK-*PCL*-FQF **1** ... 13
3. Additional Characterization CKFK-*NCL*-FQF **4** ... 14
3.1. LC-MS ... 14
3.2. MALDI—ToF-MS ... 15
3.3. HPLC study of CKFK-*NCL*-FQF ... 16
3.4. TEM of CKFK-*NCL*-FQF fibrils before and after irradiation ... 16
3.5. AFM of *in situ* irradiated CKFK-*NCL*-FQF **4** ... 17
3.6. ThT-Assay of CKFK-*NCL*-FQF **4** ... 17
4. Peptide Coating and Patterning ... 18
4.1. Photomasks for patterning peptide fibrils ... 18
4.2. Proteostat staining of photopatterned surface ... 19
4.3. MALDI MSI of photopatterned peptide layer ... 20
4.4. Cell-Viability of A549 on photopatterned surface ... 21
4.5. Cell Titer Glo Assay of CKFK-*PCL*-FQF fibrils ... 22
4.6 Control patterning experiments with non-responsive peptide fibrils CKFK-*NCL*-FQF and with agarose coated slide without peptide fibrils. ... 24
5. Zeta Potential ... 25
6. References ... 26

1. Methods

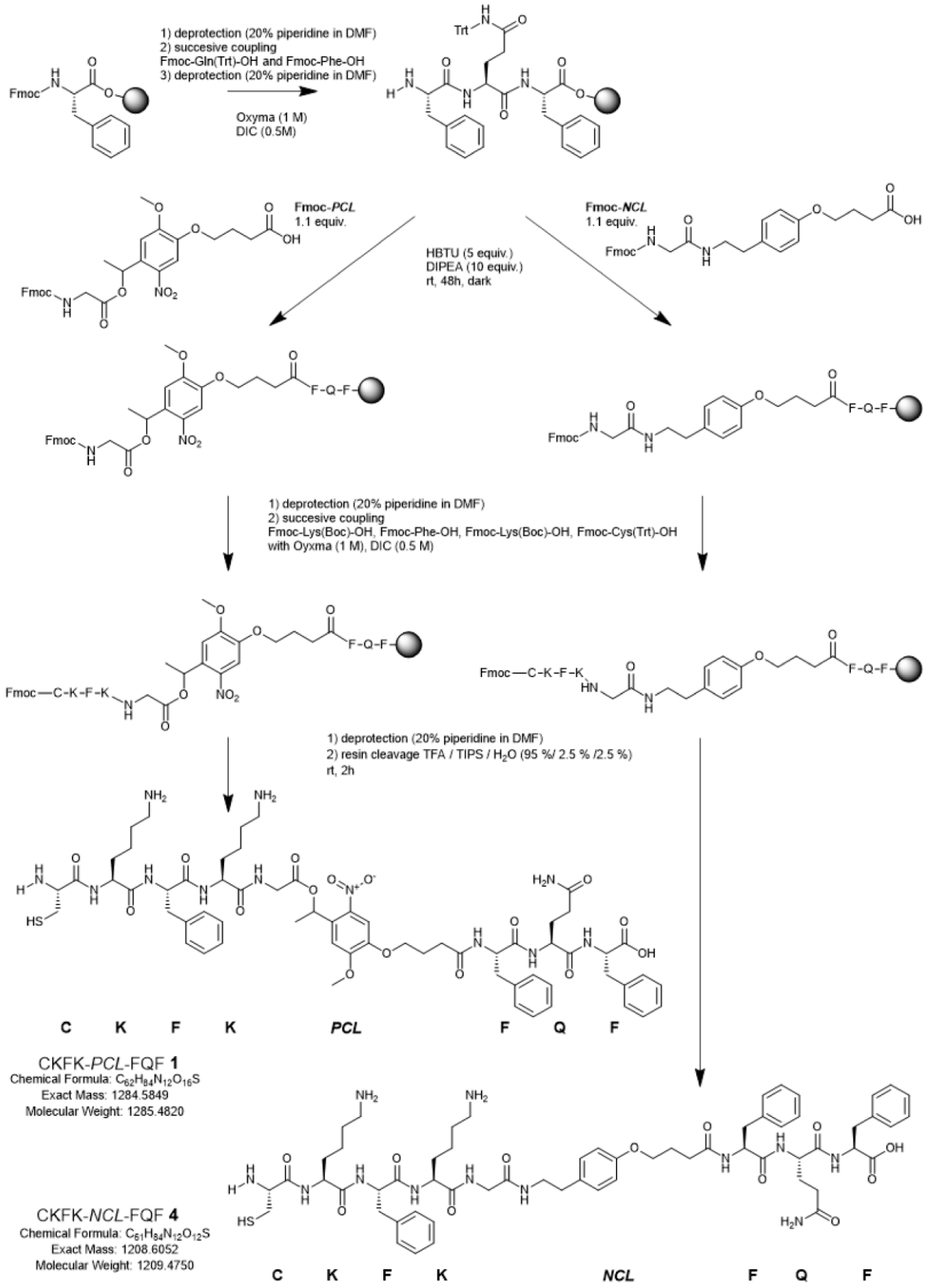
1.1. Microwave assisted solid phase peptide synthesis

Peptides were synthesized using an automated microwave peptide synthesizer (CEM, Liberty Blue™) from C to N terminus according to fluorenylmethyloxycarbonyl (Fmoc) solid phase peptide synthesis (SPPS) strategy by Merrifield (Scheme S1). *N,N'*-Di*iso*propylcarbodiimide (0.5 M in DMF) and ethyl cyano(hydroxyimino)acetate (1.0 M in DMF) were used as activator and base, a piperidine solution (20 vol% in *N,N'*-dimethylformamide (DMF)) was used for deprotection and DMF as the main washing agent. The preloaded polystyrene (PS) based Fmoc-L-Phe-Wang resin (Novabiochem®, loading: 0.68 mmol/g, 100–200 mesh) and side-chain functionality protected L-amino acids were used for coupling. Depending on the reaction scale, reagents were prepared as suggested by CEM in the Liberty Blue™ User Guide.

For 0.1 mmol synthesis scale, Fmoc-L-Phe-Wang resin (0.1 mmol, 0.65 mmol/g) was swelled in DMF for an hour before synthesis. In the automated peptide synthesizer (Liberty Blue, CEM), another short swelling in 10 mL DMF for 5 min was carried out. Fmoc-deprotection was carried out with 3 mL of a 20% piperidine in DMF solution while heating to 75 °C (155 W) for 15 s and 90 °C (30 W) for 50 s. Subsequently, the resin was washed with DMF twice with 2 mL and once with 3 mL for 5 min each. Coupling of the amino acids was conducted using 2.5 mL of a 0.2 M solution of the respective amino acid in DMF, 1 mL of a 0.5 M DIC in DMF solution and 0.5 mL of a 1.0 M Oxyma in DMF solution, followed by heating to 75 °C (170 W) for 15 s and 90 °C (30 W) for 110 s and flushing with 2 mL DMF. Final deprotection was carried out using 20% piperidine in DMF as previously described. The synthesized linkers were coupled to the FQF sequence on the resin in a separate reaction vessel. Successive coupling of CKFK to the resin was carried out using the same cycle as previously described. Cleavage from the resin was carried out by shaking in a solution of TFA (95%), TIPS (2.5%) and MilliQ water (2.5%) for 2 h at room temperature. Subsequently, the solution was transferred into ice-cold

diethyl ether for precipitation and the mixture was centrifuged three times at 4 °C, 4000 rpm for 15 min. The remaining liquid was disposed, and the pellet stored in the freezer.

WILEY-VCH



Scheme S1 Reaction scheme and molecular structures of the photoresponsive peptide CKFK-*PCL*-FQF **1** and non-photoresponsive control CKFK-*NCL*-FQF **4**. Synthesis of Fmoc-*PCL* and Fmoc-*NCL* was conducted as described previously.[1]

1.2. Turbidity measurements

Turbidity measurements were conducted with a Zetasizer Nano S instrument (Malvern Panalytic) in a micro cuvette (ZEN0040) with 70 µL of the undiluted nanofibril sample at room temperature. The derived count rate of scattered light provided information on turbidity of the sample before and after irradiation.

1.3. MALDI-MSI

For the matrix assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) measurements the peptide layer was coated on an agarose pre-treated indium tin oxide (ITO) slide. The peptide layer was photopatterned and spray-coated with a solution of MALDI matrix α-cyano-4-hydroxycinnamic acid (HCCA, 10 mg/mL in 70% ACN, 30% H2O, 0.2% TFA, Sigma-Aldrich, Germany) analogous to a previous report.[1] Acquisition of the MALDI MSI spectra were carried on a MALDI-TOF mass spectrometer (rapifleX TOF/TOF, Bruker, Bremen, Germany) in the reflection mode using the software Compass 2.0 (Bruker GmbH, Bremen, Germany) and FlexImaging 5.0, (Bruker GmbH Bremen, Germany). An area of 60 mm × 3.5 mm was scanned in steps of 10 µm using a laser profile M5. (59) A pixel size of 10 µm × 10 µm, 26% laser intensity, 1000 shots/pixel and a laser pulse repetition rate of 10 kHz was used. Distribution of peak intensities were visualized with the software Scils Lab Pro 2019 (Bruker, Germany) with a color gradient of ion signal intensities (2) normalized to total ion current (TIC).

1.4. Cell Viability Analysis (CellTiter-Glo Assay)

A549 cell viability in presence of peptide nanofibrils was tested using a CellTiter-Glo Assay (Promega G7571). A549 cells were seeded with a density of 18000 cells/mL in 96-well plates before treatment and incubated for 24 h. 5 µL of the prepared samples were added to 45 µL cell

culture medium to a final concentration of 0.1 mg/mL. 1 μM staurosporine was added as a negative control, pure cell culture medium served as positive control. The plates were incubated for 24 h at 37 °C. The wells were washed with DPBS, and the CellTiter-Glo Assay was performed according to manufacturer's instructions. Luminescence was detected using a GloMax Multi 96-well plate reader (Promega). All samples were measured in sextuplicates.

1.5. Zeta Potential

To reveal surface charges, the electrophoretic mobility of fibrils was measured and converted to ζ-potentials. Here, 60 μL of preformed fibrils (10 mg/mL DMSO solution diluted in PBS to obtain a 1 mg/mL concentration and subsequently incubated for 24 h) were diluted in 600 μL of 1 mM KCl solution. The measurement was conducted in 1 mL disposable DTS1070 folded capillary cells (Zetasizer Nano series, Malvern Panalytical) on a Zetasizer Nano ZS (Malvern Panalytical) at 25°C. The ζ-potential values were converted via Zetasizer Nano Software (Malvern Panalytical).

2. Additional Characterization CKFK-*PCL*-FQF **1**

2.1. Comparison of fibril morphology via TEM of CKFKFQF, CKFK-*PCL*-FQF **1** and CKFK-*NCL*-FQF **4** after irradiation of monomers in DMSO

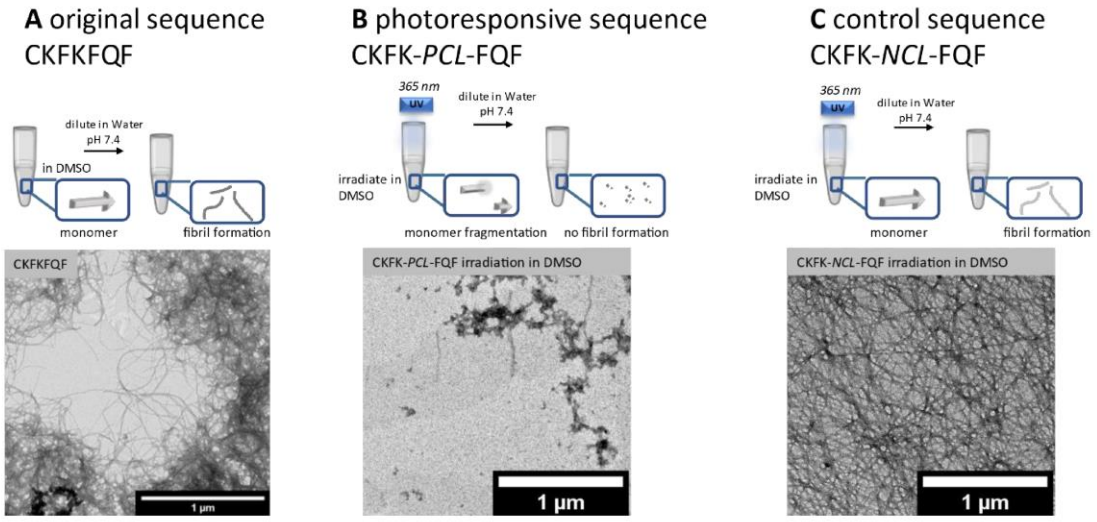


Figure S1 TEM micrographs of CKFKFQF,[2] CKFK-*PCL*-FQF **1** and CKFK-NCL-FQF **4**. **A** Fibril morphology of the original sequence CKFKFQF[2] diluted from DMSO (10 mg/mL) in water (1 mg/mL) at pH 7.4). **B** UV irradiation of monomeric CKFK-*PCL*-FQF in DMSO (10 mg/mL) for 10 min and subsequent incubation in water (1 mg/mL) at pH 7.4. **C** UV irradiation of monomeric CKFK-*NCL*-FQF in DMSO (10 mg/mL) for 10 min and subsequent incubation in water (1 mg/mL) at pH 7.4, scalebar 1µm.

2.2. LC-MS

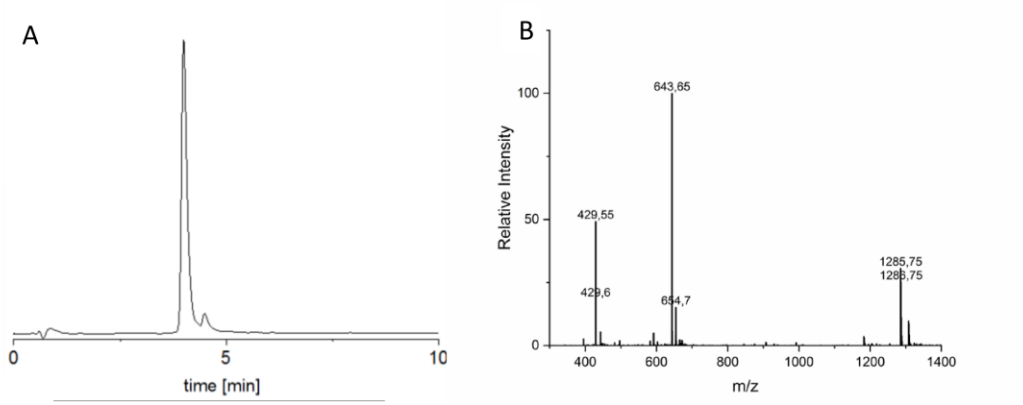
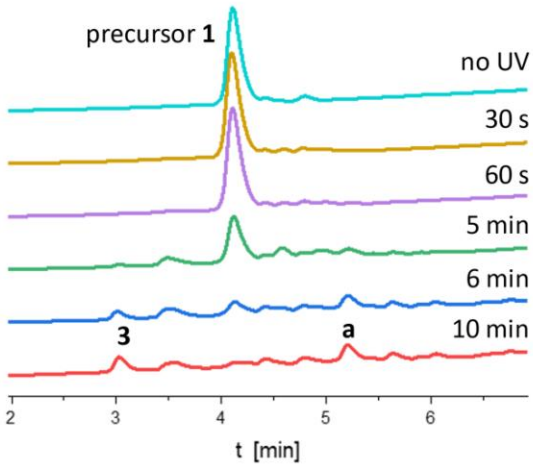


Figure S2 A LC-MS chromatogram (detector: 214 nm) of CKFK-*PCL*-FQF **1 B** MS (ESI+) m/z calc. for CKFK-*PCL*-FQF $[\mathbf{1}+3H]^{3+,}$ 429.19 g/mol; found m/z 429.55, $[\mathbf{1}+2H]^{2+}$, 643.29 g/mol; found m/z 643.65 and $[\mathbf{1}+H]^{+}$, 1285.58 g/mol; found m/z 1285.75.

2.3. Photocleavage kinetics of CKFK-*PCL*-FQF **1** in water



assignment	found	calculated
[**3**+Na]$^+$	725.3	725.3
a	728.3	728.3

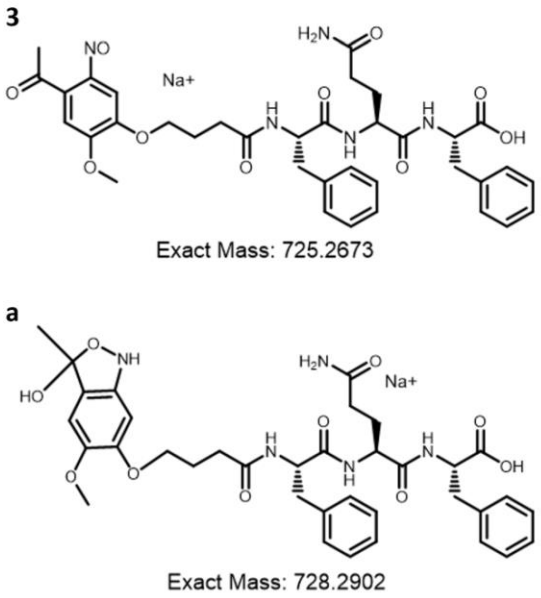


Figure S3 LC-MS traces of CKFK-PCL-FQF before UV and after UV exposure for 30 s, 60 s, 5 min, 6 min and 10 min in water. Absorbance detected at 254 nm. Photocleavage of **1** t_{Ret} 4.1 min result in several decomposition products because the photocleavage reaction is driven by radical reaction. Two main peaks which can be assigned to the proposed main fragment **3** t_{Ret} 3.0 min can be found with another by-product **a** t_{Ret} 5.2 min caused by thiol-side reactions as detected and assigned via ESI-MS.[3] **a** is a ring-chain tautomer of **3**, the ratio between the main fragment **3** and **a** is 49:51.

2.4. Photocleavage kinetics of CKFK-*PCL*-FQF **1** in DMSO

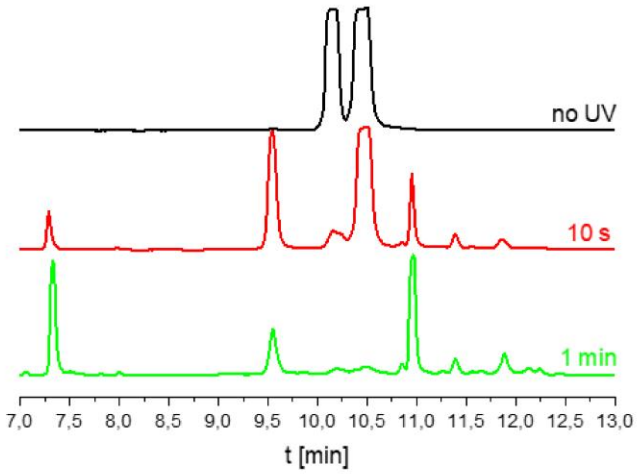


Figure S4 HPLC traces of CKFK-*PCL*-FQF **1** before UV and after UV exposure for 10 s and 60 s after 24 h incubation time in pure DMSO. Absorbance detected at 190 nm. After 60 s of irradiation, the precursor molecule is completely photocleaved. The two peaks visible in the non-irradiated sample can be assigned to CKFKF-*PCL*-FQF in its monomeric and dimerized form. Incubation in pure DMSO for 24 h results in disulfide bond formation as identified by mass spectrometry of the respective peaks (data not shown).

2.5. TEM

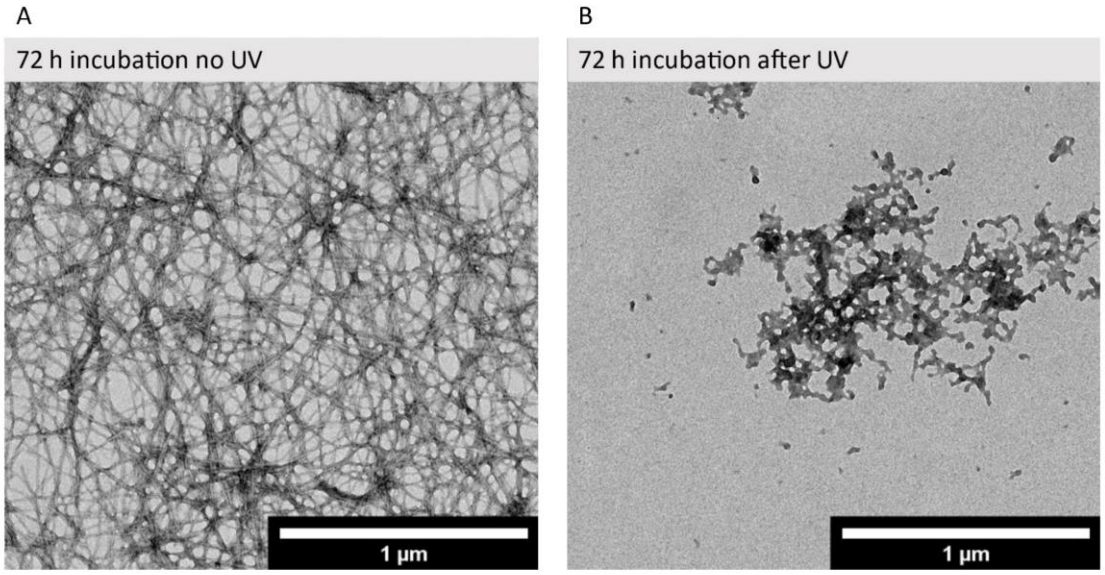


Figure S5 TEM micrographs of CKFK-PCL-FQF **1** after 72 h of incubation in water at pH 7.4. **A** without UV irradiation and **B** with UV irradiation of the peptide fibrils in solution for 20 min.

2.6. Turbidity of CKFK-*PCL*-FQF **1** before and after irradiation by DLS

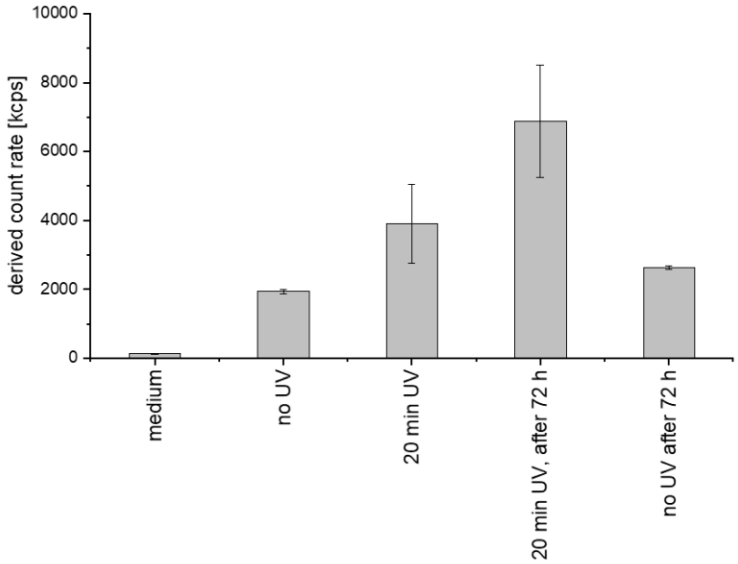


Figure S6 DLS count rate of scattered light (600 nm) of CKFK-PCL-FQF **1** solution 1 mg/mL before and after irradiation at 365 nm and after incubation at rt for 72 h.

2.7. Conversion of monomers to aggregated monomers before and after irradiation of CKFK-*PCL*-FQF **1**

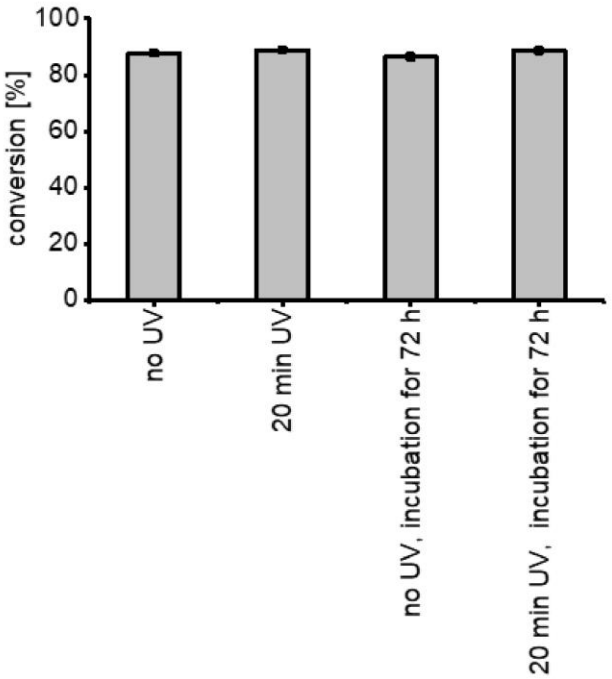


Figure S7 Conversion rate assay. Irradiated and non-irradiated CKFK-PCL-FQF **1** solution at 1 mg/mL before and after incubation at rt for 72 h.

3. Additional Characterization CKFK-*NCL*-FQF **4**

3.1. LC-MS

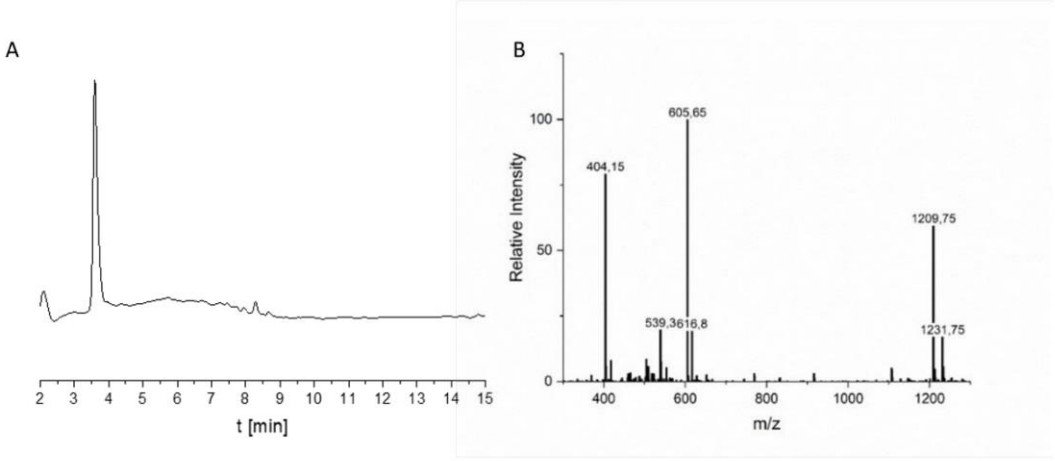


Figure S8 A LC-MS chromatogram (detector: 214 nm) of CKFK-*NCL*-FQF **4**. **B** MS (ESI+) m/z calc. for [**4**+3H]$^{3+}$, 403.87 g/mol; found m/z 404.15, [**4**+2H]$^{2+}$, 605.31 g/mol; found m/z 605.65, [**4**+H]$^{+}$, 1209.61 g/mol; found m/z 1209.75 and [**4**+Na]$^{+}$, 1231.61 g/mol; found m/z 1231.75.

3.2. MALDI—ToF-MS

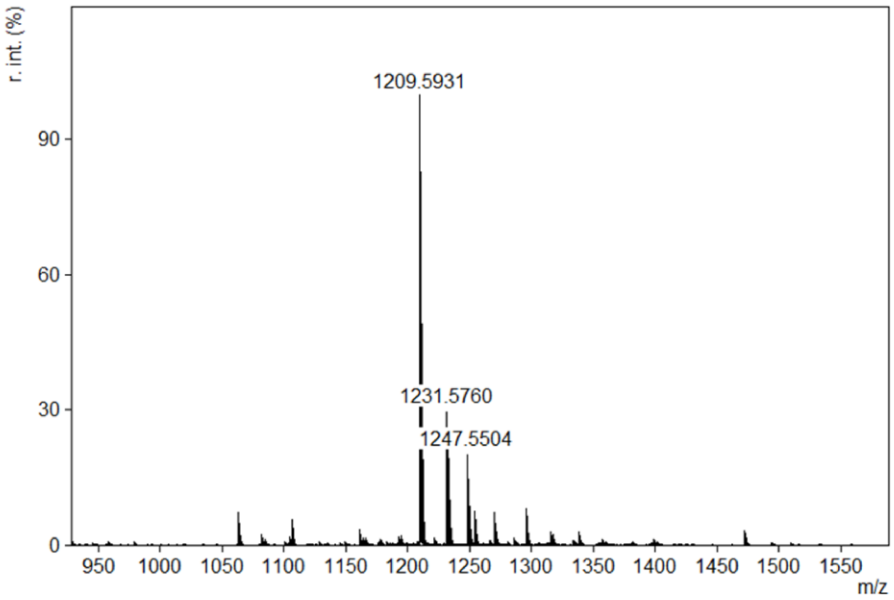


Figure S9 MALDI-ToF-MS of CKFK-*NCL*-FQF **4** with m/z calc. for [**4**+H]$^+$, 1209.61 g/mol; found m/z 1209.59, [**4**+Na]$^+$, 1231.61 g/mol; found m/z 1231.57 and [**4**+K]$^+$, 1247.61 g/mol; found m/z 1247.55.

3.3. HPLC study of CKFK-*NCL*-FQF

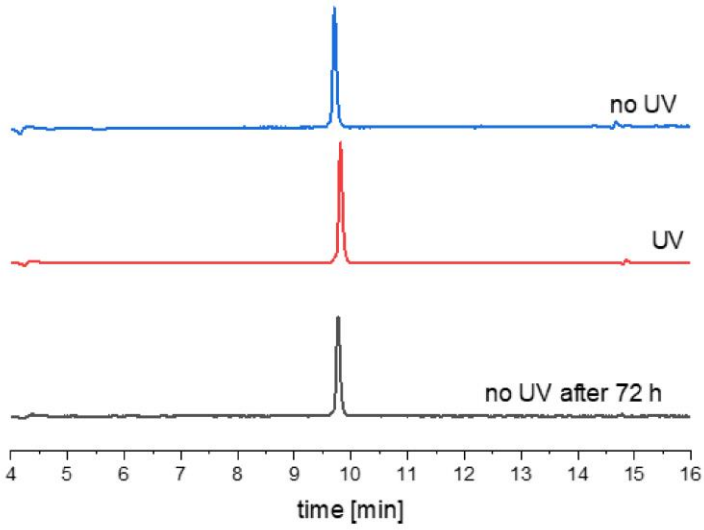


Figure S10 HPLC traces of CKFK-*NCL*-FQF **4** fibrils after no UV (24 h and 72 h incubation time) and UV exposure for 10 min after 24 h incubation time. Absorbance detected at 254 nm.

3.4. TEM of CKFK-*NCL*-FQF fibrils before and after irradiation

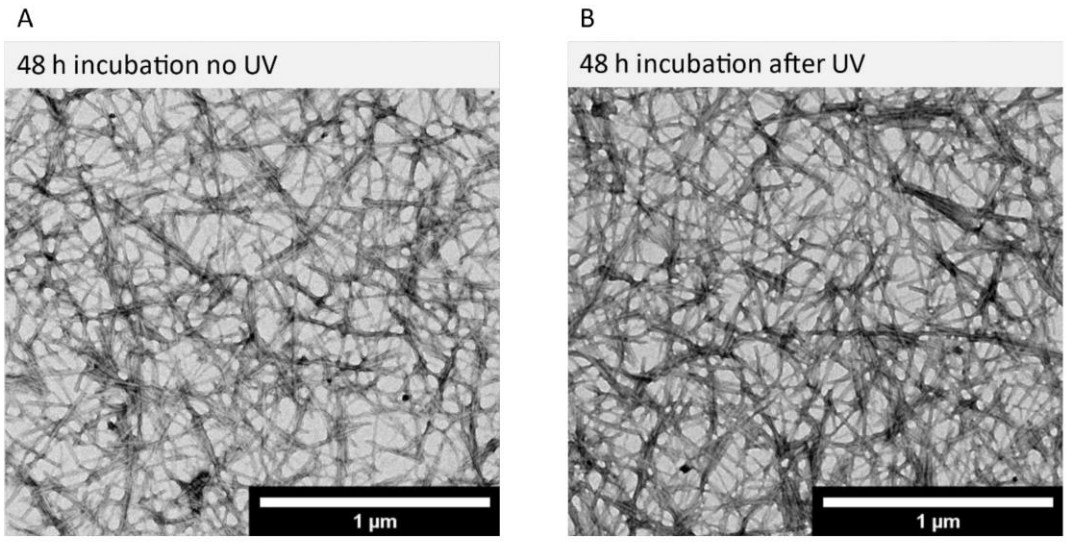


Figure S11 TEM micrographs of CKFK-*NCL*-FQF **4** after 48 h of incubation in water at pH 7.4. **A** without UV irradiation and **B** with UV irradiation of the peptide solution for 20 min.

3.5. AFM of *in situ* irradiated CKFK-*NCL*-FQF **4**

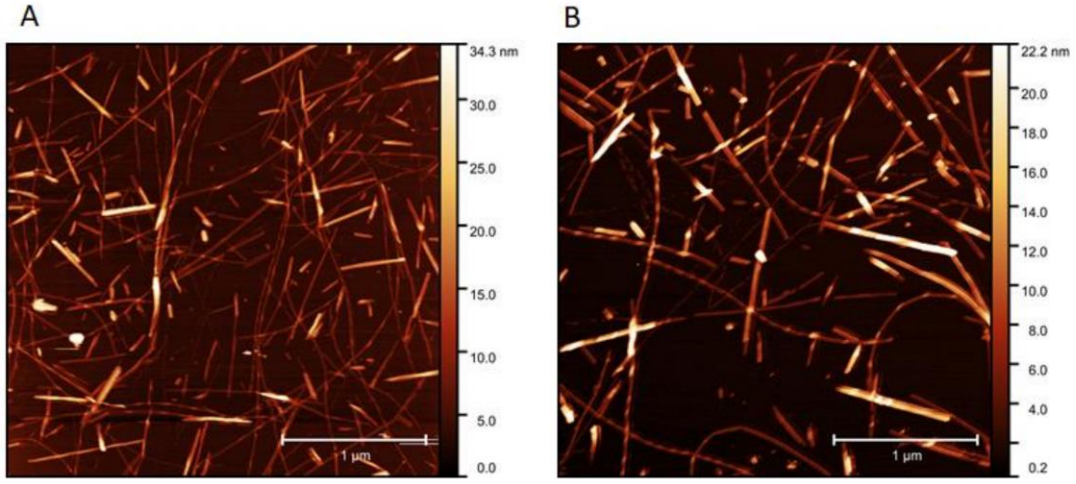


Figure S12 Liquid AFM measurement of CKFK-*NCL*-FQF **4** after of incubation in water at pH 7.4. **A** without UV irradiation and **B** with *in situ* UV irradiation of the peptide solution for 20 min.

3.6. ThT-Assay of CKFK-*NCL*-FQF **4**

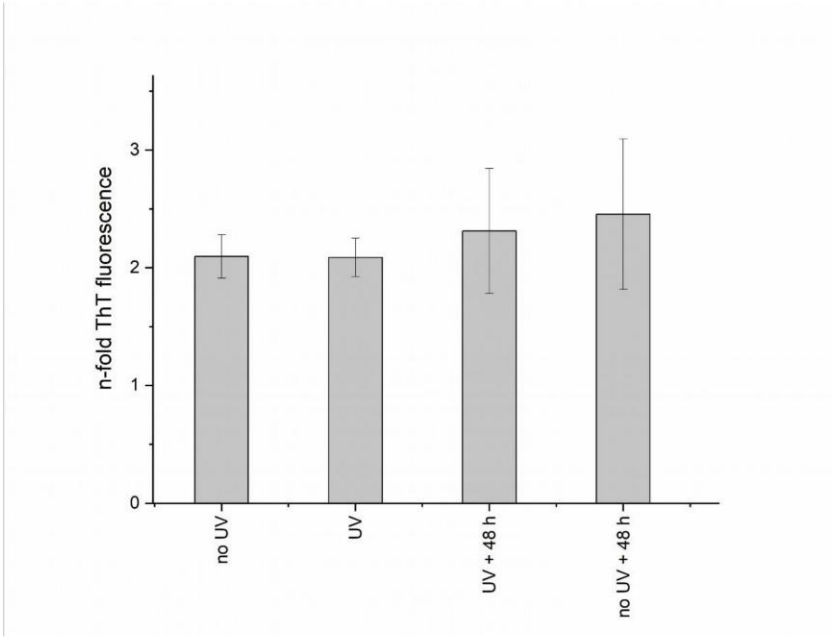


Figure S13 N-fold ThT fluorescence for CKFKF-*NCL*-FQF nanofibrils before and after UV irradiation for 20 min and after 48 h incubation before and after UV irradiation for 20 min.

4. Peptide Coating and Patterning

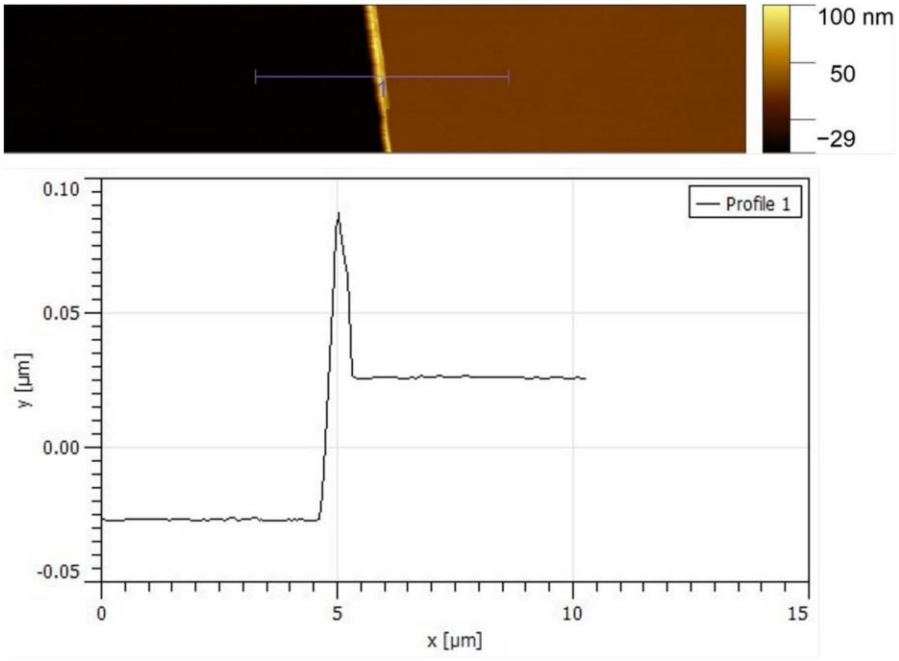


Figure S14 Atomic force microscopy measurement to determine agarose-coating thickness. The height profile was measured after scratching the surface with a pipette tip. The measured thickness of the agarose coating is approximately 50 nm which was found to be the minimum thickness to observe a cell repellent character.

4.1. Photomasks for patterning peptide fibrils

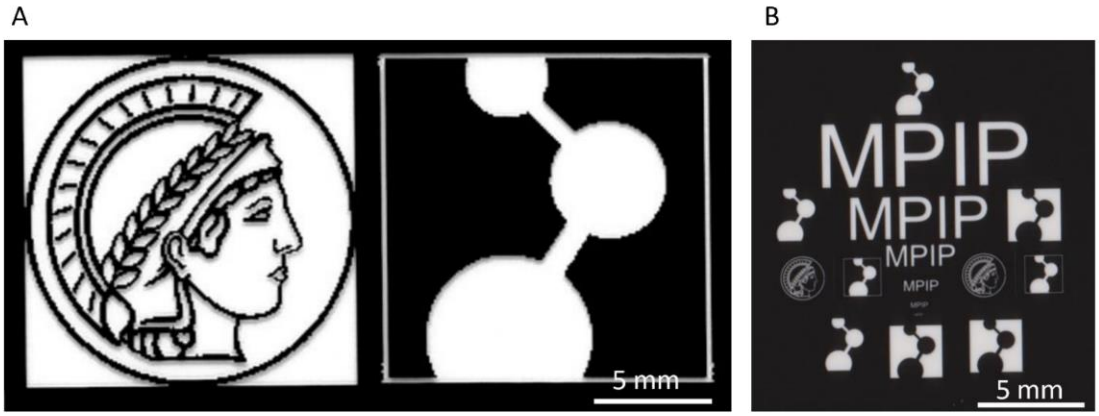


Figure S15 Overview of photomasks for patterning of peptide fibrils and cells. **A** Photomasks for patterning cells, scalebar 5 mm. **B** Photomask with smallest elements line thickness 10 - 40 µm, scalebar 5 mm.

4.2. Proteostat staining of photopatterned surface

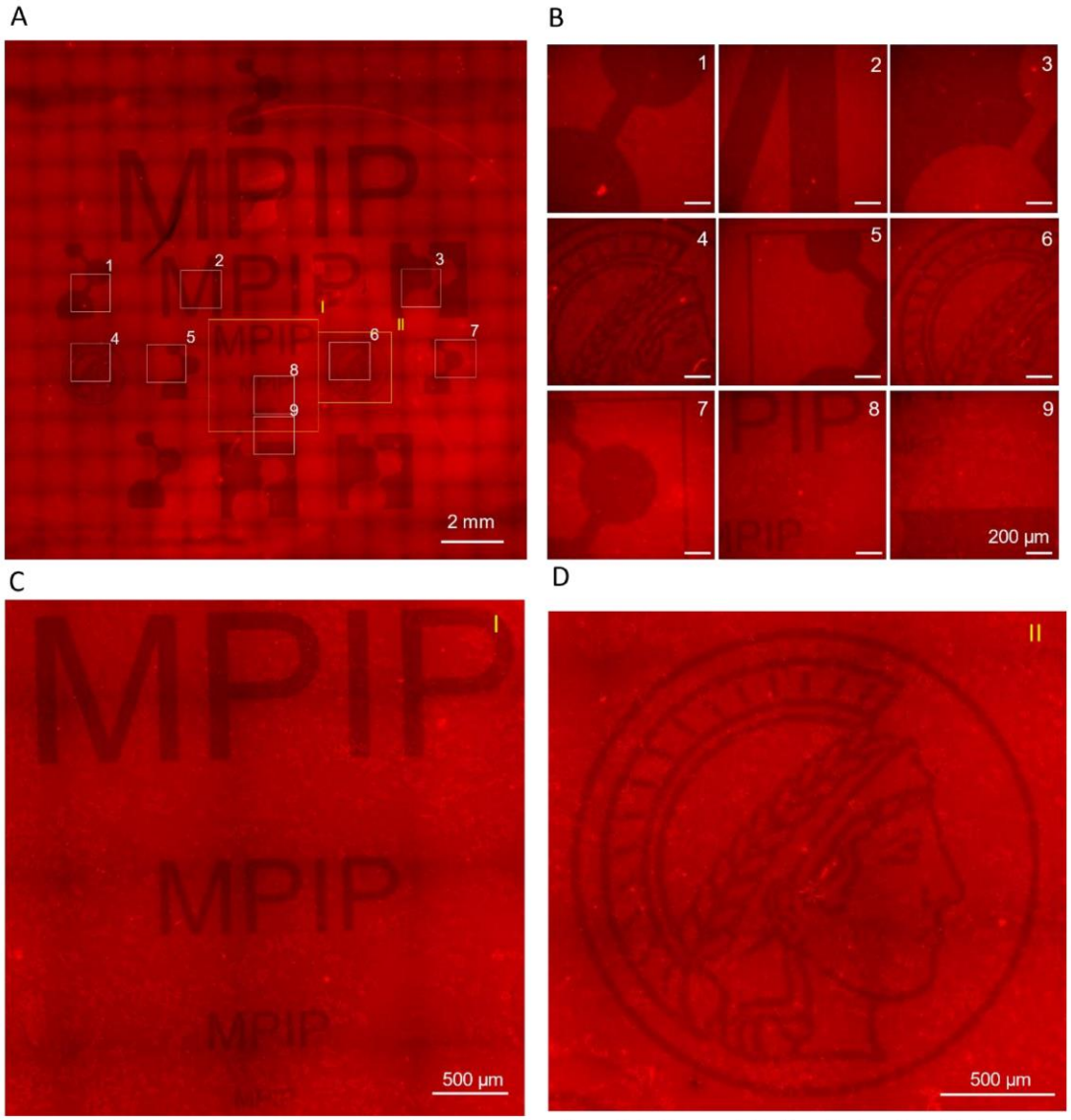


Figure S16 Fluorescence microscopy of photopatterned peptide layer of **1** after Proteostat-staining. Detailed overview of photopatterned surface: **A** tile scan over irradiated area (scale bar 2 mm), **B** ROI 1-9: single images of specific areas (scale bar 200 µm), **C** Region of interest (ROI) I: „MPIP" tile scan (scale bar 500 µm) (MPIP: Max Planck Institute for Polymer Research), **D** ROI II: Head of Minerva, a trademark of the Max Planck Society tile scan (scale bar 500 µm).

4.3. MALDI MSI of photopatterned peptide layer

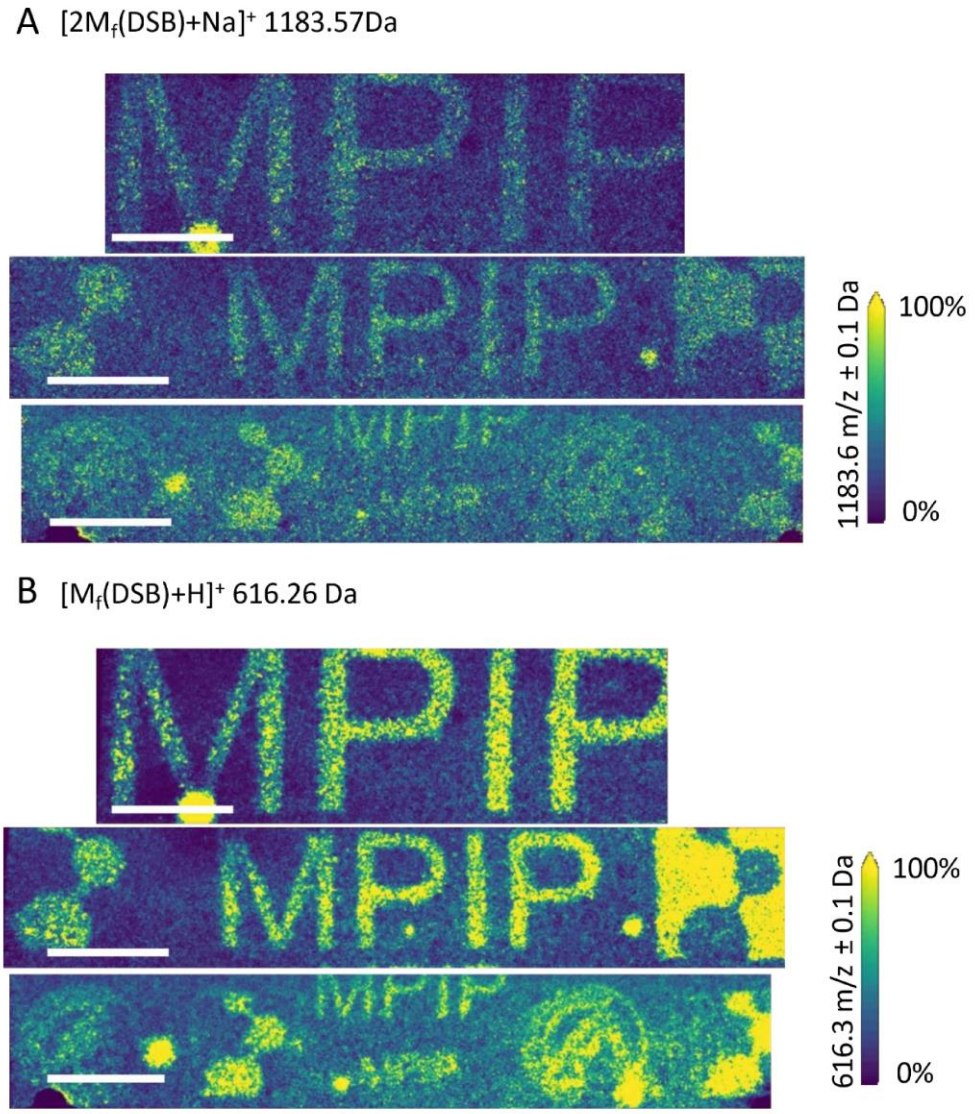


Figure S17 MALDI-MSI of photopatterned peptide layer of **1**. **A** MALDI-MSI results visualize the photopatterning via distribution of the mass intensity of the disulfide bonded peptide (DSB)[4] of the fragment (M_f(DSB): **2**+S) with sodium counter ion [2M_f (DSB) Na^+] = 1183.57 Da and **B** with the fragment peptide of [M_f (DSB) + H]$^+$= 616.26 Da (In-Source-Decay fragment of the disulfide bonded peptide of 2M_f (DSB)[4] Scalebars 2 mm.

4.4. Cell-Viability of A549 on photopatterned surface

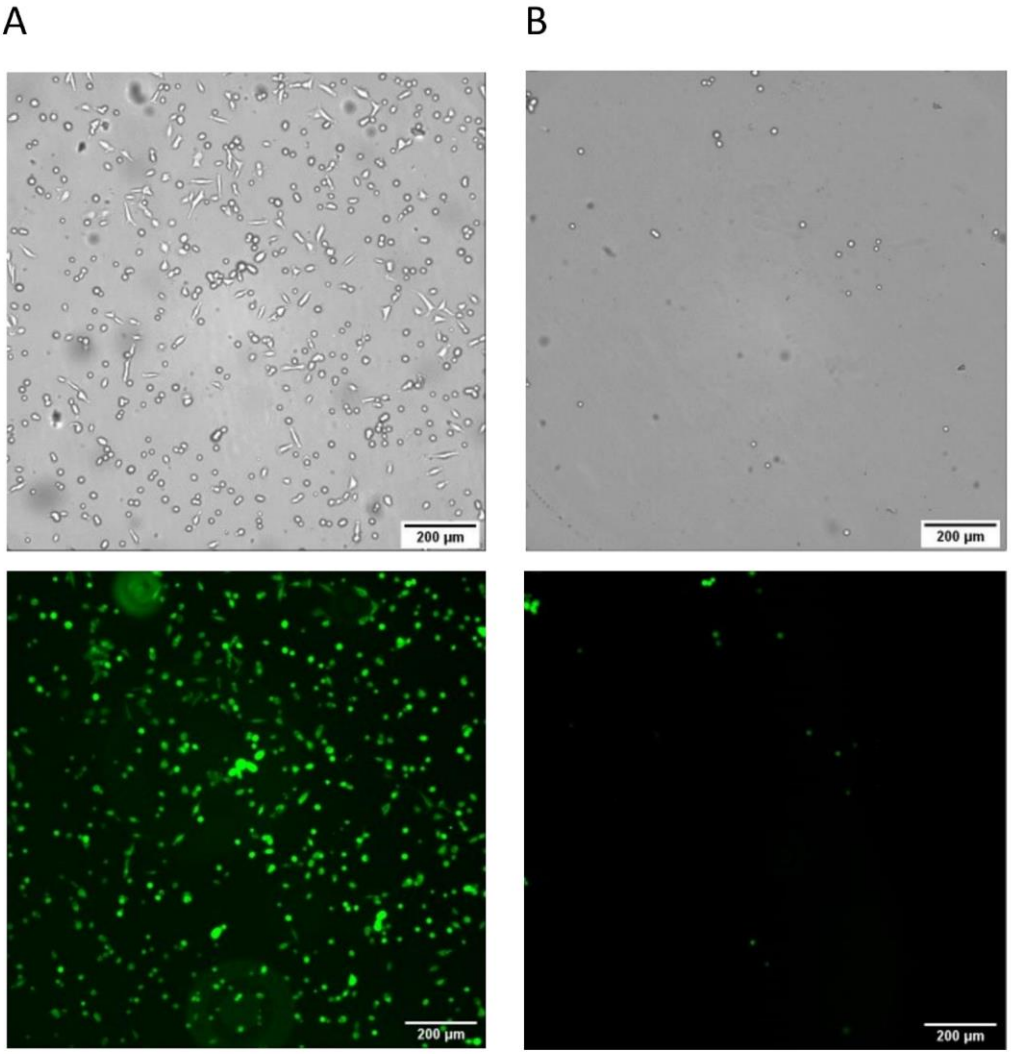


Figure S18 Imaging of calcein staining of A549 cells on CKFK-*PCL*-FQF peptide coated agarose slides **A** before irradiation with UV and **B** after UV-irradiation for 20 min. Top brightfield, bottom fluorescence image. Cells were seeded after irradiation. Scalebar 200 µm.

4.5. Cell Titer Glo Assay of CKFK-*PCL*-FQF fibrils

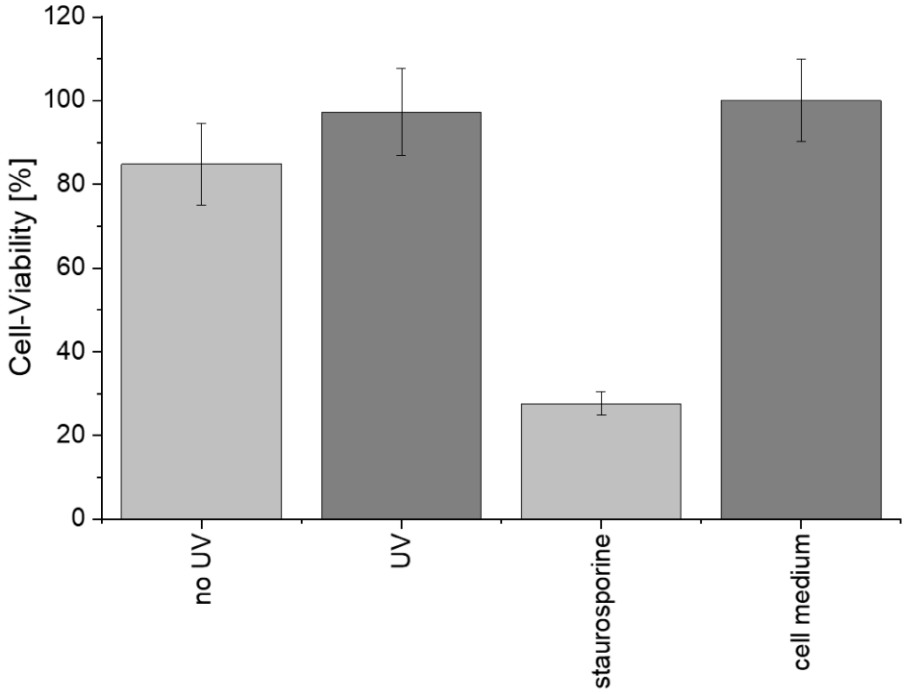


Figure S19 CellTiter-Glo Assay of preformed fibrils from CKFK-*PCL*-FQF **1** with and without UV irradiation for 10 minutes. 1 μM staurosporine was added for a negative control, pure cell culture medium served as positive control.

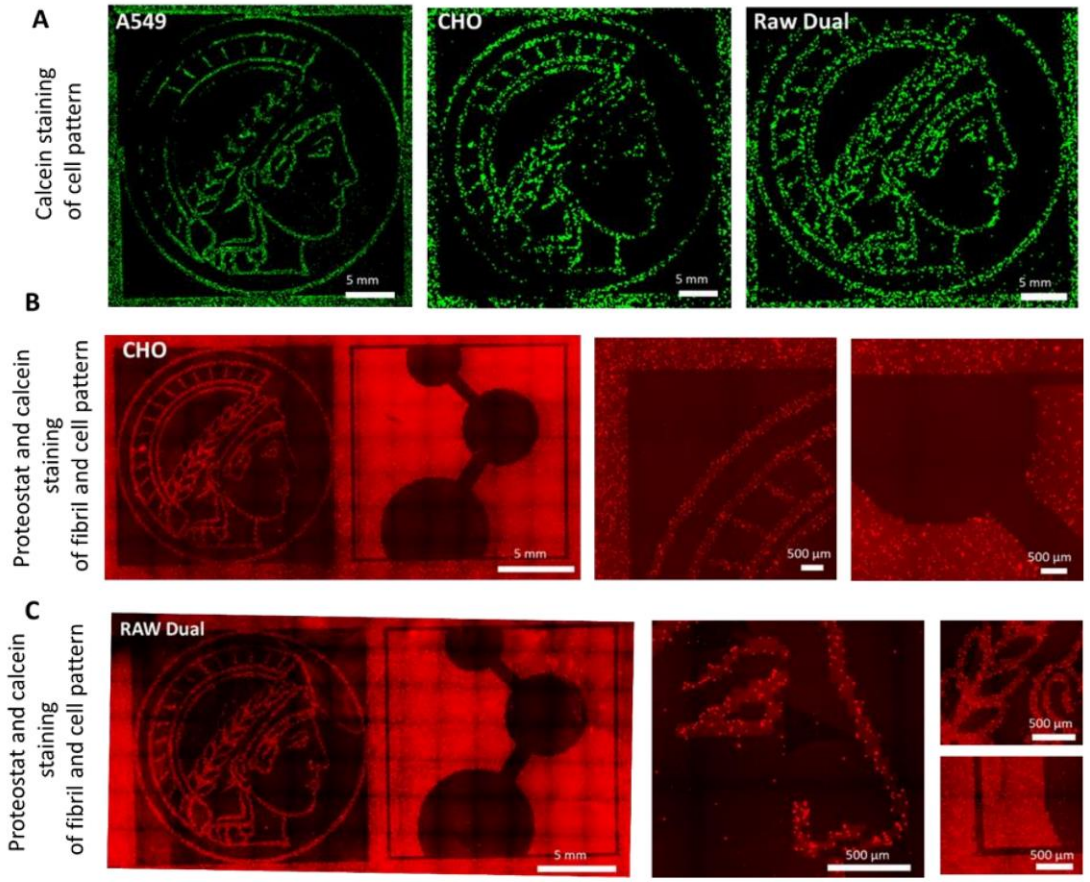


Figure S20 Cell patterning with three different cell types on surfaces irradiated with photo mask shown in Figure S15A. **A** Fluorescence microscopy measurement of calcein stained A549, CHO and Raw Dual cells display the head of Minerva, a trademark of the Max Planck Society, scalebar 5 mm. **B** Simultaneous calcein and Proteostat staining of CHO cells and **C** Raw Dual cells show that cells attach to intact peptide fibrils, which appear brighter upon Proteostat staining, whereas non-fibrillar parts, which are less fluorescent, cannot maintain cell attachment, scalebar left 5 mm, right 500 µm.

4.6 Control patterning experiments with non-responsive peptide fibrils CKFK-*NCL*-FQF and with agarose coated slide without peptide fibrils.

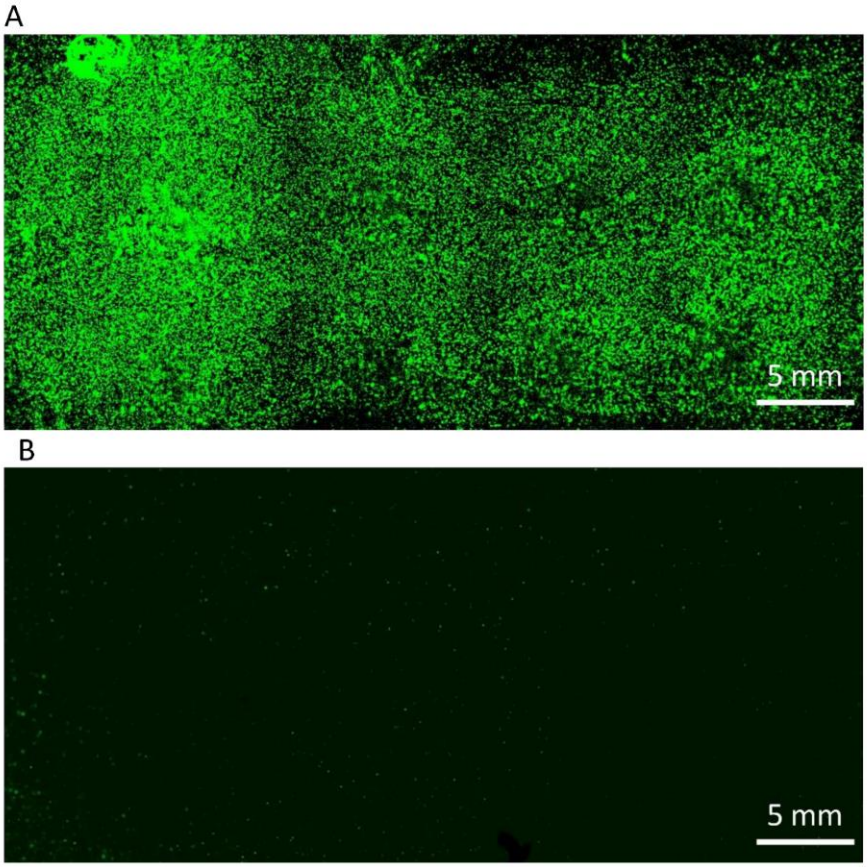


Figure S21 Control patterning study with A549 cells seeded on non-responsive CKFKF-*NCL*-FQF **4** fibrils and agarose treated slide without peptide fibrils, which were irradiated with MPI-P logo motif Figure S15A. **A** As expected, no cell patterning can be observed for CKFKF-*NCL*-FQF fibrils and **B** cells cannot attach well to agarose coated slide if peptide fibrils are not applied (scalebar 5 mm).

5. Zeta Potential

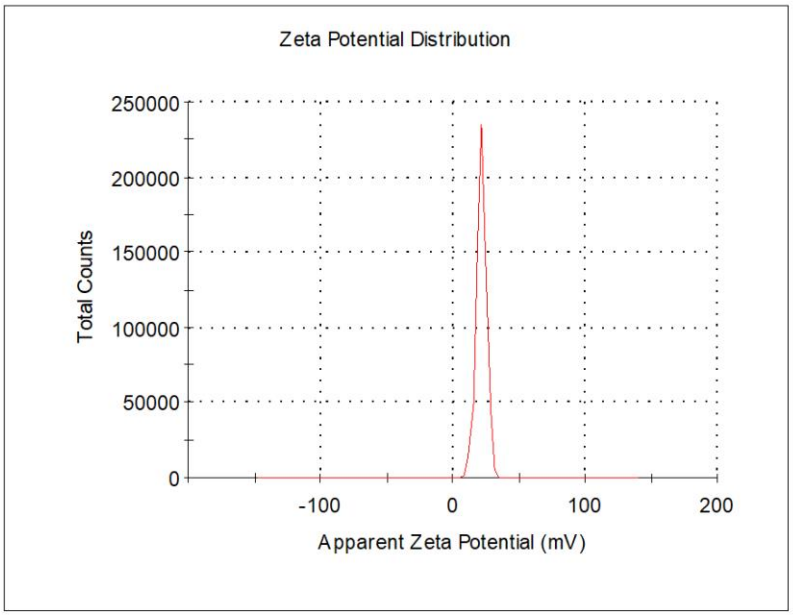


Figure S22 ζ-Potential distribution of fibrils from CKFK-*PCL*-FQF (**1**) peptide. The calculated mean value (21.0 ± 0.8 mV) for the ζ-potential was achieved by six independent measurements à 20 runs. ζ-potential of CKFKFQF is 22.3± 1.4 mV.[2]

6. References

[1] A. M. Ender, K. Kaygisiz, H.-J. Räder, F. J. Mayer, C. V. Synatschke, T. Weil, *ACS Biomater. Sci. Eng.* **2021**, *7*, 4798.

[2] C. Schilling, T. Mack, S. Lickfett, S. Sieste, F. S. Ruggeri, T. Sneideris, A. Dutta, T. Bereau, R. Naraghi, D. Sinske, T. P. J. Knowles, C. V. Synatschke, T. Weil, B. Knöll, *Adv. Funct. Mater.* **2019**, *29*, 1809112.

[3] A. Barth, J. E. T. Corrie, M. J. Gradwell, Y. Maeda, W. Mäntele, T. Meier, D. R. Trentham, *J. Am. Chem. Soc.* **1997**, *119*, 4149.

[4] A. Resemann, L. Liu-Shin, G. Tremintin, A. Malhotra, A. Fung, F. Wang, G. Ratnaswamy, D. Suckau, *MAbs* **2018**, *10*, 1200.

7. APPENDIX

Scientific Contributions to Peer-Reviewed Journals and Conferences

Peer-Reviewed Publications

8 **K. Kaygisiz**, L. Rauch-Wirth, A. Dutta, T. Bereau, X. Yu, Y. Nagata, J. Münch, C. V. Synatschke, T. Weil, *Nat. Commun.*, **2023**, *accepted*, DOI: 10.1038/s41467-023-40663-6, *Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides.*

7 L.-R. Olari, R. Bauer, M. Gil Miró, V. Vogel, L. Cortez Rayas, R. Groß, A. Gilg, R. Klevesath, A. A. Rodríguez Alfonso, **K. Kaygisiz**, U. Rupp, P. Pant, J. Mieres-Pérez, L. Steppe, R. Schäffer, L. Rauch-Wirth, C. Conzelmann, J. A. Müller, F. Zech, F. Gerbl, J. Bleher, N. Preising, L. Ständker, S. Wiese, D. R. Thal, C. Haupt, H. R. A. Jonker, M. Wagner, E. Sanchez-Garcia, T. Weil, S. Stenger, M. Fändrich, J. von Einem, C. Read, P. Walther, F. Kirchhoff, B. Spellerberg, J. Münch, *Cell. Mol. Life Sci.* **2023**, 80, 151, DOI: 10.1007/s00018-023-04795-8, *The C-terminal 32-mer fragment of hemoglobin alpha is an amyloidogenic peptide with antimicrobial properties*

6 **K. Kaygisiz**, A. Dutta, L. Rauch-Wirth, C. V. Synatschke, J. Münch, T. Bereau, T. Weil, *Biomater. Sci.* **2023,** 11 (15), 5251–5261. DOI: 10.1039/D3BM00412K, *Inverse Design of Viral Infectivity Enhancing Peptides Fibrils from Continuous Protein-Vector Embeddings*.

5 **K. Kaygisiz,** A. M. Ender, J. Gačanin, L. A. Kaczmarek, D. A. Koutsouras, A. N. Nalakath, P. Winterwerber, F. J. Mayer, H. J. Räder, T. Marszalek, P. W. M. Blom, C. V. Synatschke, T. Weil, *Macromol. Biosci.* **2022**, 2200294, DOI: 10.1002/mabi.202200294, *Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning*. *Selected as the Front Cover*

4 A. M. Ender, **K. Kaygisiz**, H. J. Räder, F. J. Mayer, C. V. Synatschke, T. Weil, *ACS Biomater. Sci. Eng.* **2021**, *7*, 4798–4808, DOI: 10.1021/acsbiomaterials.1c00889, *Cell-instructive surface gradients of photo-responsive amyloid-like fibrils*.
Selected as the Inside Cover

3 S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, **K. Kaygisiz**, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, 2009382, DOI: 10.1002/adfm.202009382, *Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction.*

2 **K. Kaygisiz**, C. V. Synatschke, *Biomat. Sci.* **2020**, *8*, 6113–6156, DOI: 10.1039/D0BM01367F, *Materials promoting viral gene delivery*.
Selected as the Back Cover

1 T. Ikai, R. Ishidate, K. Inoue, **K. Kaygisiz**, K. Maeda, E. Yashima, *Macromolecules*, **2020**, *53*, 973–981, DOI: 10.1021/acs.macromol.9b02727, *Chiral/Achiral Copolymers of Biphenylylacetylenes Bearing Various Substituents: Chiral Amplification through Copolymerization, Followed by Enhancement/Inversion and Memory of the Macromolecular Helicity.*

Conferences Contributions

2023 **Peptide Self-Assembly Conference**, Manchester, poster presentation
2023 **NFDI4Chem Chemistry Data Days**, Mainz, poster presentation
2023 **CRC1279 Students Retreat**, Lindenberg Allgäu, poster presentation
2023 **GRC Self-Assembly and Supramolecular Chemistry**, Les Diablerets, poster presentation
2022 **Merck Curious Conference**, Darmstadt, oral presentation (*invited*)
2022 **Novo Nordisk Future Chemist Symposium**, Kopenhagen, oral presentation (*invited*)
2022 **GRC Bioinspired Materials**, Les Diablerets, poster presentation
2022 **Bordeaux Polymer Conference**, Bordeaux, poster presentation.
2022 **Suprachem,** Mainz, poster presentation.
2022 **Macromolecular Colloquium Freiburg**, online, poster flash presentation.
2021 **ACS Spring Meeting,** online, poster presentation.
2021 **Suprabio21,** online, poster presentation.
2021 **CRC1279 Symposium,** Ulm, oral presentation.
2020 **Macromolecular Colloquium Freiburg**, online, poster presentation.

List of Thesis Publications

To be Submitted

2023 A. Iscen, **K. Kaygisiz**, C. V. Synatschke, T. Weil, K. Kremer, *Unraveling structure-property relationships in self-assembling peptide nanofibrils via simulations.*

Submitted

2023 **K. Kaygisiz,** L. Rauch-Wirth, A. Iscen, J. Hartenfels, K. Kremer, J. Münch, C. V. Synatschke, T. Weil, *In Revision, Peptide amphiphiles as biodegradable adjuvants for efficient retroviral gene delivery*

Accepted

5 **K. Kaygisiz**, L. Rauch-Wirth, A. Dutta, T. Bereau, X. Yu, Y. Nagata, J. Münch,C. V. Synatschke, T. Weil, *Nat. Commun.*, **2023**, *accepted*, DOI: 10.1038/s41467-023-40663-6, *Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides.*

4 **K. Kaygisiz**, A. Dutta, L. Rauch-Wirth, C. V. Synatschke, J. Münch, T. Bereau, T. Weil, *Biomater. Sci.* **2023,** 11 (15), 5251–5261. DOI: 10.1039/D3BM00412K, *Inverse Design of Viral Infectivity Enhancing Peptides Fibrils from Continuous ProteinVector Embeddings*.

3 **K. Kaygisiz,** A. M. Ender, J. Gačanin, L. A. Kaczmarek, D. A. Koutsouras, A. N. Nalakath, P. Winterwerber, F. J. Mayer, H. J. Räder, T. Marszalek, P. W. M. Blom, C. V. Synatschke, T. Weil, *Macromol. Biosci.* **2022**, 2200294, DOI: 10.1002/mabi.202200294, *Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning*.
Selected as the Front Cover

2 S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, **K. Kaygisiz**, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, 2009382, DOI: 10.1002/adfm.2020093, *Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction.*

1 **K. Kaygisiz**, C. V. Synatschke, *Biomat. Sci.* **2020**, *8*, 6113–6156, DOI: 10.1039/D0BM01367F, *Materials promoting viral gene delivery*.
Selected as the Back Cover

Declaration of Originality in Academic Work

I, Kübra Kaygisiz, hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg-University Mainz to insure standard scientific practice.

Mainz, 24.08.2023 [redacted]

Place, Date Signature

Curriculum Vitae

August 2023

CV – Kübra Kaygisiz

ORCiD 0000-0003-0077-7867

Google Scholar https://scholar.google.com/citations?hl=de&user=x5p83yYAAAAJ

Research Interests Self-Assembling Peptides, Supramolecular Chemistry, Peptide Materials Design, Data-Mining, Biomaterials, Viral Gene Delivery, Cell-Material Interactions.

Education

02/2020 – 08/2023	**PhD in Chemistry, Max Planck Institute for Polymer Research, Mainz, Germany.** Topic: „Self-Assembling Peptide Nanostructures for Controlled Cell-Material Interactions" Supervisor: [redacted].
10/2017 – 09/2019	**M.Sc., Chemistry, Johannes Gutenberg University Mainz, Germany (Grade: 1.7)** Thesis „Photo Patterned Peptide Materials." Grade 1.0, Supervisor: [redacted], Max Planck Institute for Polymer Research.
10/2015 – 03/2017	**Japanese Studies, Johannes Gutenberg University Mainz, Germany (Certificate)**
10/2014 – 09/2017	**B. Sc., Chemistry**, **Johannes Gutenberg University Mainz, Germany (Grade: 2.0)** Thesis „Redoxswitchable Collagen Mimetic Peptides." Grade 1.0, Supervisor: [redacted], Johannes Gutenberg University Mainz.

Research Stays

08/2022– 10/2022	**Research Stay, Kyoto University, Japan** Topic: "Dynamics and Patterning of Photo-Responsive Multicomponent Peptide Fibrils via Super Resolution CLSM" Supervisor: [redacted].
04/2018 – 09/2018	**Research Internship, Nagoya University, Japan** Topic: "Synthesis of Acetyloxy Type Poly(biphenylylacetylene)s and Investigation of their Chiral Amplification." Supervisor: [redacted].
08/2016 – 10/2016	**Research Internship, Johannes Gutenberg University Mainz** Establishing standard procedure for controlled calcite nucleation and growth using the ammonia diffusion method. Supervisor: [redacted].

Publications and Patents

In Preparation

2023 A. Iscen, **K. Kaygisiz**, C. V. Synatschke, T. Weil, K. Kremer, *Unraveling structure-property relationships in self-assembling peptide nanofibrils via simulations.*

Submitted

2023 L. Rauch-Wirth, A. Renner, **K. Kaygisiz**, L. Zimmermann, A. A. Rodríguez Alfonso, D. Schütz, S. Wiese, L. Ständker, T. Weil, D. Schmiedel and J. Münch, *In Revision, Optimized peptide nanofibrils as efficient transduction enhancers for in vitro and ex vivo gene transfer*

2023 **K. Kaygisiz,** L. Rauch-Wirth, A. Iscen, J. Hartenfels, K. Kremer, J. Münch, C. V. Synatschke, T. Weil, *In Revision, Peptide amphiphiles as biodegradable adjuvants for efficient retroviral gene delivery*

2023 **K. Kaygisiz**, L. Rauch-Wirth, J. Hartenfels, J. Münch, C. V. Synatschke, T. Weil *European Patent Number EP23171438.7 Self-assembling peptide amphiphiles as retroviral transduction enhancers*

Accepted

8 **K. Kaygisiz**, L. Rauch-Wirth, A. Dutta, T. Bereau, X. Yu, Y. Nagata, J. Münch, C. V. Synatschke, T. Weil, *Nat. Commun.*, **2023**, *accepted*, DOI: 10.1038/s41467-023-40663-6, *Data-Mining Unveils Structure-Property-Activity Correlation of Viral Infectivity Enhancing Self-Assembling Peptides.*

7 L.-R. Olari, R. Bauer, M. Gil Miró, V. Vogel, L. Cortez Rayas, R. Groß, A. Gilg, R. Klevesath, A. A. Rodríguez Alfonso, **K. Kaygisiz**, U. Rupp, P. Pant, J. Mieres-Pérez, L. Steppe, R. Schäffer, L. Rauch-Wirth, C. Conzelmann, J. A. Müller, F. Zech, F. Gerbl, J. Bleher, N. Preising, L. Ständker, S. Wiese, D. R. Thal, C. Haupt, H. R. A. Jonker, M. Wagner, E. Sanchez-Garcia, T. Weil, S. Stenger, M. Fändrich, J. von Einem, C. Read, P. Walther, F. Kirchhoff, B. Spellerberg, J. Münch, *Cell. Mol. Life Sci.* **2023**, 80, 151, DOI: 10.1007/s00018-023-04795-8, *The C-terminal 32-mer fragment of hemoglobin alpha is an amyloidogenic peptide with antimicrobial properties*

6 **K. Kaygisiz**, A. Dutta, L. Rauch-Wirth, C. V. Synatschke, J. Münch, T. Bereau, T. Weil, *Biomater. Sci.* **2023,** 11 (15), 5251–5261. DOI: 10.1039/D3BM00412K, *Inverse Design of Viral Infectivity Enhancing Peptides Fibrils from Continuous ProteinVector Embeddings*.

5 **K. Kaygisiz,** A. M. Ender, J. Gačanin, L. A. Kaczmarek, D. A. Koutsouras, A. N. Nalakath, P. Winterwerber, F. J. Mayer, H. J. Räder, T. Marszalek, P. W. M. Blom, C. V. Synatschke, T. Weil, *Macromol. Biosci.* **2022**, 2200294, DOI: 10.1002/mabi.202200294, *Photoinduced Amyloid Fibril Degradation for Controlled Cell Patterning*. *Selected as the Front Cover.*

4 A. M. Ender, **K. Kaygisiz**, H. J. Räder, F. J. Mayer, C. V. Synatschke, T. Weil, *ACS Biomater. Sci. Eng.* **2021**, *7*, 4798–4808, DOI: 10.1021/acsbiomaterials.1c00889, *Cell-instructive surface gradients of photo-responsive amyloid-like fibrils*. *Selected as the Inside Cover.*

3 S. Sieste, T. Mack, E. Lump, M. Hayn, D. Schütz, A. Röcker, C. Meier, **K. Kaygisiz**, F. Kirchhoff, T. P. J. Knowles, F. S. Ruggeri, C. V. Synatschke, J. Münch, T. Weil, *Adv. Funct. Mater.* **2021**, 2009382, DOI: 10.1002/adfm.2020093, *Supramolecular Peptide Nanofibrils with Optimized Sequences and Molecular Structures for Efficient Retroviral Transduction.*

2 **K. Kaygisiz**, C. V. Synatschke, *Biomat. Sci.* **2020**, *8*, 6113–6156, DOI: 10.1039/D0BM01367F, *Materials promoting viral gene delivery*.
Selected as the Back Cover

1 T. Ikai, R. Ishidate, K. Inoue, **K. Kaygisiz**, K. Maeda, E. Yashima, *Macromolecules*, **2020**, *53*, 973–981, DOI: 10.1021/acs.macromol.9b02727, *Chiral/Achiral Copolymers of Biphenylylacetylenes Bearing Various Substituents: Chiral Amplification through Copolymerization, Followed by Enhancement/Inversion and Memory of the Macromolecular Helicity.*

Awards and Fellowships

2023 **Travel Award** for the Peptide Self-Assembly conference in Manchester awarded by GDCh.
2023 **Poster Prize** at the Peptide Self-Assembly conference in Manchester.
2023 **Writing Fellowship** awarded from the Max Planck Foundation.
2023 **Poster Prize** at the NFDI4Chem symposium in Mainz.
2023 **Poster Prize** at the CRC1279 symposium in Lindenberg.
2022 **Research Scholarship** of Japanese Society for the Promotion of Science (JSPS).
2022 **Examined Project Manager** in Business Chemistry awarded from GDCh.
2021 **Poster Prize** at the MPIP symposium in Mainz.
2018 **Research Scholarship** of Materials Science in Mainz Graduate School for research stay in Japan.
2014-2019 **Study Scholarship** of the German National Academic Foundation Avicenna Studienwerk e.V.
2014 **International Chemistry Olympiad**, awarded for the 3rd national selection round
2013 **Jugend forscht**, ranked 3rd on federal state level in chemistry.

Conference Contributions

2023 **Peptide Self-Assembly Conference**, Manchester, poster presentation.
2023 **NFDI4Chem Chemistry Data Days**, Mainz, poster presentation.
2023 **CRC1279 Students Retreat**, Lindenberg Allgäu, poster presentation.
2023 **GRC Self-Assembly and Supramolecular Chemistry**, Les Diablerets, poster presentation.
2022 **Merck Curious Conference**, Darmstadt, oral presentation (*invited*).
2022 **Novo Nordisk Future Chemist Symposium**, Kopenhagen, oral presentation (*invited*).
2022 **GRC Bioinspired Materials**, Les Diablerets, poster presentation.
2022 **Bordeaux Polymer Conference**, Bordeaux, poster presentation.

2022 **Suprachem,** Mainz, poster presentation.
2022 **Macromolecular Colloquium Freiburg**, online, poster flash presentation.
2021 **ACS Spring Meeting,** online, poster presentation.
2021 **Suprabio21,** online, poster presentation.
2021 **CRC1279 Symposium,** Ulm, oral presentation.
2020 **Macromolecular Colloquium Freiburg**, online, poster presentation.

Teaching and Supervision

Lab courses	02/2021 – 03/2021 (Pharmaceutical Chemistry), 07/2021 – 08/2021 (Pharmaceutical Chemistry), 01/2022 – 03/2022 (Pharmaceutical Chemistry)
Master Students	[redacted] (02/2022 – 06/2022)
Bachelor Students	[redacted] (05/2021 – 09/2021)
Research Interns	[redacted] (10/2021 – 01/2022), [redacted] (10/2021 – 11/2021), [redacted] (03/2022 – 04/2022)

Voluntary Activities and Service

2022 **Conference Organization**, Friends of the Chemistry Olympiad 3-day conference in Mainz with 50 participants.
2017 – 2019 **Board Member** of Bildungsnetzwerk Mainz e.V. (Association for Education in Mainz): organization of events for public relations, leisure activities for children and young adults.
2016 – 2019 **Group leader** in Bildungsnetzwerk Mainz e.V. (Association for Education in Mainz): Supervising 10 volunteering tutors for middle school children.
2016 – 2019 **Member of advisory board** of Friends of the Chemistry Olympiad: main organization and supervision of two annual seminars (3-5 days) for 30 middle and 15 high school students.
2016 – 2017 **Head of Communication** in Deutsches Jungforscher Netzwerk juFORUM e.V. (German Association of Young Scientists): Board member, organization of annual conference with more than 100 participants and 3 workshops, responsible for public relations including homepage maintenance, collaborations with other associations and monthly newsletters.
2015 – 2019 **Volunteering Tutor** in Bildungsnetzwerk Mainz e.V. (Association for Education in Mainz) for refugee middle school students on a weekly basis.

ACKNOWLEDGEMENTS

My PhD time has been a memorable experience and several people have contributed to the completion of this thesis in numerous ways. They have not only enriched my scientific knowledge, interests, and ideas, but also shared enjoyable moments. Their support has truly made my time at the Max Planck Institute for Polymer Research immensely enriching.

First and foremost, I thank my PhD advisor, [redacted]. I am grateful for her invaluable suggestions, which allowed me to explore the fascinating field of bioactive peptide fibrils. Her advice, insightful perspectives, and unwavering support gave me much freedom in my research while having access to cutting-edge resources.

I would also like to extend my thanks to the members of my thesis committee [redacted], and [redacted]for their constructive feedback and evaluation of my thesis. I would like to acknowledge the support provided by [redacted]. Thanks for connecting me with internal and external collaborators and arranging meetings to discuss with visiting researchers.

I am grateful to [redacted] for providing me with the opportunity to attend international conferences, in particular, the Gordon Research Conferences on “Bioinspired Materials” and “Supramolecular Chemistry and Self-Assembly”. The insights gained from these experiences have broadened my horizon and enriched my understanding of the field. Furthermore, I thank [redacted] for giving me the opportunity to work within the CRC1279 and for initiating many collaborations within.

It was a great pleasure and a very rewarding experience to work together with the groups of [redacted], [redacted], and [redacted]. I was very fortunate to make valuable connections with great minds in peptide discovery from the biologist’s perspective, attend the weekly CRC lectures, monthly student meetings and annual retreats. I thank the CRC PhD students committee for always setting up the remote access for me to participate virtually in Ulm.

This highly interdisciplinary and collaborative thesis would not have been possible without the strong interactions between all of the cooperation partners. Their expertise and feedback have immensely shaped the outcome of this thesis, and I am grateful for their valuable input:

First, I would like to deeply thank [redacted] for the wonderful collaboration on all virus transduction assays. Her expertise and experience in virus-related questions has been an important source of learning for me. Thank you for the nice collaboration throughout our PhD journey.
I thank [redacted] for his invaluable input and proof-reading in the introductory parts regarding computational methods, which has helped me to refine sections of this thesis. Our research collaboration has served as a constant inspiration for this thesis. I thank [redacted] who made the technical aspects of machine-learning possible and sparking my interest coding.
Thanks to [redacted] for the molecular dynamics simulations, that are substantial contributions in chapters 3.4 and 6.4 of this thesis, and the nice discussions we had on peptide fibril properties from the very beginning until the end of my PhD. Your contribution significantly enhanced the mechanistic, molecular understandings and provided important insights into

peptide self-assembly. I would further extend my gratitude to [redacted], who supported the simulations with discussions and provided the resources.
I thank [redacted] for her dedication and motivation to conduct all the intricate SFG measurements that have been vital for understanding peptide fibrils in more detail. Thanks to [redacted] and [redacted] for providing the resources and supervision for SFG measurements.
I want to thank all the previous coworkers of the Weil group and Münch group who made possible that I could smoothly connect to previous research findings. In particular, I thank [redacted], [redacted], and [redacted] for setting a strong fundament on which I could build on my research. Furthermore, I thank [redacted] for supporting cell-culture experiments.

I feel very lucky to have been given the chance to supervise and work with motivated students. My special thanks go out to [redacted], whose contribution in his bachelor and master thesis on functionalization peptide fibrils with antiviral peptide motifs significantly advanced our understanding of the bioactivity of these peptides. [redacted] greatly assisted in advancing the peptide amphiphile synthesis and his careful evaluation provided important insights in solubility and aggregation of different peptides sequences.

During my Ph.D. time, I had the privilege of spending several months at Kyoto University and work within the group of [redacted]. I thank [redacted] for hosting my visit and [redacted] for teaching me super resolution microscopy and fluorescence lifetime imaging methods. This research stay has been a truly memorable experience that fostered personal and academic growth.

I owe gratitude to all researchers whose patient mentorship and teaching introduced me to various biophysical characterization techniques. These methods and techniques have been instrumental in advancing my research:
I thank [redacted] for teaching me how to perform TEM tomography for 3D visualization of fibril–cell interactions, [redacted] for teaching me TEM for fibril morphology analysis, and [redacted] for conducting all Cryo-EM measurements. I thank [redacted] and [redacted] for giving me an introduction into AFM and [redacted] teaching me liquid AFM methods. I acknowledge the access to several microscopes provided by the Institute of Molecular Biology (IMB) in Mainz. Thanks to [redacted] and [redacted] from the IMB for teaching me confocal fluorescence microscopy and live-cell imaging with widefield microscopy. I thank [redacted] for giving me an introduction to lifetime and live-cell imaging with confocal fluorescence microscopy and for inspiring scientific discussions on confocal fluorescence methods and selection of right dyes and experimental setups.
Thanks to [redacted] for providing access to DLS measurements and discussing measurement methods. Thanks to the group of [redacted], who helped me out with the particle size analyzer.
I thank the mass department, in particular [redacted] and [redacted] for the great teamwork during the MALDI-Imaging and [redacted] and [redacted], for measuring MALDI-ToF of my samples. I thank further [redacted] and [redacted] for maintaining the HPLC systems and supporting the measurements of my many peptide samples. Thanks to the BioCore for the support regarding questions on cell culture. I thank [redacted] for introducing and teaching me cell culture and bacteria culture, the NMR

department, in particular, [redacted] for supporting with NMR kinetic measurements and discussions on spectra evaluation, and [redacted] for her continuous support in manuscript submission and communications to the editorial offices.

I acknowledge funding from the CRC1279 that enabled my research and several conference travels such as to the GRCs in 2022 and 2023, and the Bordeaux Polymer Conference 2022. I thank JSPS for awarding me a research fellowship, enabling a research stay in Japan. Furthermore, I am grateful to receive a travel award from the GDCh to participate at the Peptide Self-Assembly Conference 2023 in Manchester. Finally, I feel happy to receive a writing fellowship to finalize my thesis and spent many memorable days at the Grundner residence of Max Planck, supported by the Max Planck Foundation.

On a more personal note, I thank the Weil group for enjoyable moments be it the weekly volleyball sessions or over dinner after work. Special thanks to [redacted] and [redacted] for the all the nice conversations and support throughout my PhD journey!

To the former and current office members [redacted], [redacted], [redacted], [redacted], [redacted], [redacted], [redacted], [redacted], [redacted] and [redacted]: you have been great colleagues and friends! I thank you all.

To all the individuals mentioned above and to those whose contributions may not have been explicitly stated, I offer my deepest gratitude.

Finally, my sincere gratitude goes to my family and friends. Your continuous support and encouragement have been invaluable to me.