JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Targeting the Tumor Vasculature with an Engineered Cystine-Knot Miniprotein Applying an Optimized Phage Display Discovery Platform

Die Dissertation zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Bonny Gaby Lui

geb. am 16.06.1990 in Offenbach am Main

Mainz, 2019

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung:

Abstract

Engineered cystine-knot miniproteins represent an auspicious alternative protein scaffold for therapeutic and diagnostic applications, allowing specific tailoring of several molecular properties in addition, to target affinity and specificity, in particular stability solubility and pharmacokinetic behavior of the binding protein. Furthermore, the simple architecture of those small peptidic molecules facilitates conventional chemical manufacturing and the construction of multi-functional fusion molecules. Cystine-knot miniproteins are characterized by a remarkable stability and an extraordinary tolerance to sequence variation enabling the construction of combinatorial libraries. The starting basis for this thesis was a previously established cystine-knot miniprotein technology platform that allows the development of new drug candidates for diagnostic and therapeutic applications. The platform comprised a randomized cystine-knot miniprotein phage library (MCopt 1.0) that was constructed on the basis of an open chain variant of the squash trypsin inhibitor MCoTI-II (*Momordica cochinchinensis*) and a phage display screening system for ligand selection against soluble target proteins.

The first objective of this thesis was the establishment of an efficient cell-based phage panning protocol and the development of a robust procedure to identify cystine-knot miniproteins that bind to plasma membrane proteins in their native conformation. In a proof of concept experiment a previously identified model ligand (MC-FA-010), which binds to FAP-α was utilized. MC-FA-010 phages were mixed at different ratios into the established MCopt 1.0 phage library and the mix was screened against FAP-α expressing eukaryotic cell line. Using an adapted screening protocol, MC-FA-010 phages could be specifically enriched after one, two or three successive screening rounds, depending on the initially applied mixing ratio. Additionally, a novel cell-based hit identification procedure consisting of high-throughput compatible protein expression and purification of individual clones combined with target binding analysis based on flow cytometry was established. The applicability of the whole cell-based panning and downstream hit identification procedure was confirmed by successful identification of high-affinity binders upon screening of the naïve MCopt 1.0 phage library against a FAP-α expressing eukaryotic cell line. The detection sensitivity was further enhanced with the development of a second expression system enabling biotin/streptavidin based tetramerization of proteins in order to gain higher binding strength as a result of the avidity effect.

The second objective of this thesis was the design and construction of novel cystine-knot miniprotein phage libraries to increase the repertoire and size of the library. Three rationally designed sub-libraries with scattered randomized positions in loop 1 (MCopt 2.0), loop 5 (MCopt 2.1) or both loops (MCopt 3.0) in the open chain sequence of MCoTI-II were successfully generated. Phage displaying of cystine-knot miniproteins through the minor coat protein pIII was obtained using a pJuFo phagemid system, comprising Fos-cystine-knot miniprotein and Jun-pIII´ gene fusions for independent protein folding. Library quality controls included the expression analysis on phage level via ELISA and western blotting indicating the functional surface presentation of cystine-knot miniproteins. MCopt 2.0 library clones showed good protein expression rates and overall proper folding capabilities, while MCopt 2.1 and MCopt 3.0 library clones performed in a marginal inferior manner. This suggested a moderate tolerance of MCoTI-II scaffold towards sequence randomization in loop 5 or in both adjacent loops 1 and 5. Moreover, also the phagemid backbone was optimized to circumvent an enrichment of deletion mutants (clones lacking the M13 coat protein and the cystine-knot miniprotein encoding sequence part). In the novel phagemid vector, pPDIII-1, repetitive sequence parts were completely eliminated and different genetic elements for tighter gene expression control were introduced. A

model selection experiment showed a clearly improved genetic stability of the pPDIII-1-MCopt 2.0 sub-library as compared to the original pJuFo vector.
The third objective of this thesis focused on the development of engineered cystine-knot miniproteins for *in vivo* targeting of the tumor vasculature. To this end, MC-FN-010, a cystine-knot miniprotein identified by screening of MCopt 1.0 und MCopt 2.0 libraries that specifically binds to fibronectin extra domain B (EDB) was used. Detailed characterization of MC-FN-010 via alanine scanning mutagenesis revealed its target binding relevant amino acids and resulted in defining a second derivative candidate (MC-FN-016). Both cystine-knot miniproteins featured high EDB specificity but relatively low affinities. However, chemical oligomerization of the ligands and site-directed fluorescence dye conjugation increased the binding strength enormously, while retaining its high specificity. The resulting trimeric constructs were analyzed *in vivo* in a U-87 MG based xenograft glioblastoma mouse model. For both variants, a strong accumulation in the tumor and overall low background signals could be detected via *in vivo* and *ex vivo* fluorescence measurement. These findings together emphasize the high potential of cystine-knot miniproteins as molecular scaffolds for tumor imaging technologies.

Zusammenfassung

Cystin-Knoten Miniproteine stellen ein vielversprechendes alternatives Proteingerüst für therapeutische und diagnostische Anwendungen dar, welches neben der Target Affinität und Spezifität, eine spezifische Anpassung verschiedener molekularer Eigenschaften ermöglicht. Insbesondere die Stabilität, Löslichkeit und das pharmakokinetische Verhalten dieser Bindeproteine sind von großer Bedeutung. Die einfache Architektur dieser kleinen peptidischen Moleküle erleichtert zudem die konventionelle chemische Herstellung und die Konstruktion multifunktionaler Fusionsmoleküle. Cystin-Knoten Miniproteine zeichnen sich durch eine bemerkenswerte Stabilität und außergewöhnliche Toleranz gegenüber Sequenzvariationen aus, die den Aufbau von kombinatorischen Bibliotheken ermöglicht. Ausgangspunkt für diese Arbeit war eine zuvor etablierte Cystin-Knoten Miniprotein Technologie Plattform, die eine Entwicklung neuer Wirkstoffkandidaten für diagnostische und therapeutische Applikationen ermöglicht. Die Plattform umfasste eine randomisierte Cystin-Knoten Miniprotein Phagen-Bibliothek (MCopt 1.0), die auf einer offenkettigen Variante des *Squash*-Trypsin-Inhibitors MCoTI-II (*Momordica cochinchinensis*) und eines Phagen-Display *Screening* Systems zur Liganden-Selektion gegen lösliche Zielproteine beruht.

Das erste Ziel dieser Arbeit war die Etablierung eines effizienten, zellbasierten Phagen *Panning* Protokolls und die Entwicklung eines robusten Verfahrens zur Identifizierung von Cystine-Knoten Miniproteinen, die an Plasmamembranproteinen in ihrer nativen Konformation binden. In einem *Proof of Concept* Experiment wurde ein zuvor identifizierter Modellligand (MC-FA-010) verwendet, der an FAP-α bindet. MC-FA-010 Phagen wurden in verschiedenen Verhältnissen in die etablierte MCopt 1.0 Phagenbibliothek gemischt und die Mischung wurde gegen ein auf einer eukaryotischen Zelllinie exprimiertes FAP-α selektiert. Abhängig vom initialen Mischungsverhältnis´ konnten die MC-FA-010 Phagen mit einem angepassten *Screening* Protokoll nach ein, zwei oder drei aufeinanderfolgenden Selektionsrunden spezifisch angereichert werden. Zusätzlich wurde ein neuartiger zellbasierter Hit Identifizierungsprozess bestehend aus einer hochdurchsatzfähigen Protein Expression und Reinigung von individuellen Klonen in Kombination mit einer durchflusszytometrischen Bindungsanalyse etabliert. Die Anwendbarkeit des gesamten zellbasierten *Panning* und des nachgeschalteten Hit Identifizierungsprozess´ wurde durch ein *Screening* der naiven MCopt 1.0 Phagen-Bibliothek gegen eine FAP-α exprimierenden eukaryotischen Zelllinie verifiziert. Zudem wurde die Detektionssensitivität mit der Entwicklung eines zweiten Expressionssystems verbessert, welches eine Biotin/Streptavidin basierte Tetramerisierung von Proteinen ermöglicht, um durch den Aviditätseffekt eine höhere Bindungsstärke zu erreichen.

Das zweite Ziel dieser Arbeit war das Design und die Konstruktion von neuen Cystin-Knoten Miniprotein Phagen-Bibliotheken, um das Repertoire und die Größe der Bibliothek zu erweitern. Drei rational konzipierte Bibliotheken mit verstreuten, randomisierten Positionen in *Loop* 1 (MCopt 2.0), *Loop* 5 (MCopt 2.1) oder in beiden *Loops* (MCopt 3.0) in der offenkettigen Sequenz Variante von MCoTI-II wurden erfolgreich generiert. Die Phagen-Präsentation von Cystin-Knoten Miniproteinen über das kleine Hüllprotein pIII wurde mittels eines pJuFo Phagemid Systems bewerkstelligt, das aus Fos-Cystin-Knoten Miniprotein und Jun-pIII´ Genfusionen besteht und eine unabhängige Proteinfaltung ermöglicht. Die Qualitätskontrolle der Bibliotheken beinhaltete die Expressionsanalyse auf Phagenebene mittels ELISA und Western Blot und deutete auf die funktionale Oberflächenpräsentation von Cystin-Knoten Miniproteinen hin. MCopt 2.0-Bibliotheksklone zeigten gute Proteinexpressionsraten und Faltungseigenschaften, während MCopt 2.1- und MCopt 3.0-Bibliotheksklone zu marginal schlechteren Ergebnisse führten. Dies wies auf eine moderate Toleranz

des MCoTI-II Gerüstes gegenüber Sequenz Randomisierung in *Loop* 5 oder in beiden benachbarten *Loops* 1 und 5 hin. Darüber hinaus wurde auch der Phagemid Vektor optimiert, um eine Anreicherung von Deletionsmutanten zu umgehen (Klone ohne M13 Hüllprotein und Cystin-Knoten Miniprotein Sequenzen). Im neuen Phagemid Vektor, pPDIII-1, wurden repetitive Sequenzeinheiten vollständig eliminiert und verschiedene genetische Elemente zur strengeren Kontrolle der Genexpression eingeführt. Ein Modell Selektionsexperiment zeigte eine deutlich verbesserte genetische Stabilität der pPDIII-1-MCopt 2.0 Bibliothek im Vergleich zum ursprünglichen pJuFo Vektor.

Das dritte Ziel dieser Arbeit konzentrierte sich auf die Entwicklung von Cystin-Knoten Miniproteinen für die *in vivo* Targetierung von Tumorgefäßen. Zu diesem Zweck wurde MC-FN-010 verwendet, ein Cystin-Knoten Miniprotein, das durch das *Screening* von MCopt 1.0 und MCopt 2.0 Bibliotheken identifiziert wurde und das spezifisch an die Extra Domäne B von humanem Fibronektin (EDB) bindet. Eine detaillierte Charakterisierung von MC-FN-010 mittels Alaninscan Mutagenese deckte die bindungsrelevanten Aminosäuren auf und führte zur Bestimmung eines zweiten Derivatkandidaten (MC-FN-016). Beide Cystin-Knoten Miniproteine wiesen eine hohe EDB Spezifität, aber eine relativ geringe Affinität auf. Die chemische Oligomerisierung der Liganden und die gerichtete Konjugation eines Fluoreszenzfarbstoffs erhöhten wiederum die Bindungsstärke deutlich, wobei die hohe Spezifität konstant blieb. Die daraus resultierenden trimeren Konstrukte wurden *in vivo* in einem U-87 MG basierten Xenograft Glioblastom Mausmodel analysiert. Für beide Varianten wurde eine starke Akkumulation im Tumor und insgesamt niedrige Hintergrundsignale mittels *in vivo* und *ex vivo* Fluoreszenzbestimmung detektiert. Diese Ergebnisse unterstreichen das hohe Potenzial der Cystin-Knoten Miniproteine als molekulare Grundgerüste für tumorbildgebende Technologien.

List of figures

List of tables

List of abbreviations

ADME	absorption, distribution, metabolism and excretion
AF680	Alexa Fluor 680
AgRP	Agouti-related protein
APC	allophycocyanin
BirA	biotin holoenzyme synthetase
CDRs	complementarity-determining regions
cFN	cellular FN
CFU	colony forming units
CTX	chlorotoxin
DMSO	dimethyl sulfoxide
DPBS	Dulbecco´s phosphate-buffered saline
DSG-3	Desmoglein-3
dsDNA	double-stranded DNA
dsRed	*Discosoma* red fluorescent protein
DTT	dithiotreitol
e.g.	exempli gratia
EC	extracellular cadherin
ECM	extracellular matrix
EDA	extra domain A
EDB	extra domain B
EDTA	ethylenediaminetetraacetic acid
EETI-II	*Ecballium elaterium* trypsin inhibitor II
ELISA	enzyme-linked immunosorbent assay
EMA	European Medicines Agency
EtOH	ethanol
FACS	fluorescence-activated cell sorting
FAP-α	fibroblast activation protein-α
FDA	Food and Drug Administration
FN	fibronectin
FSC	forward scatter
GFP	green fluorescent protein
GMP	good manufacturing practice
H6	hexahistidine
HC	heavy chain
HRP	horseradish peroxidase
i.v.	intravenously
IF	immunofluorescence
IIICS	type III connecting sequence
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilodalton
LC	light chain
mAbs	monoclonal antibodies
MCoTI-II	*Momordica cochinchinensis* trypsin inhibitor II
MeOH	methanol

MFI	mean fluorescence intensity
MOI	multiplicity of infection
MRI	magnetic resonance imaging
n.s.	not significant
NaCl	sodium chloride
OD	optical density
oMCoTI-II	open chain variant of *Momordica cochinchinensis* trypsin inhibitor II
PCR	polymerase chain reaction
PDB	protein data base
PEG	polyethylene glycol
PEI	polyethylenimine
PET	positron emission tomography
pFN	plasma FN
pIII	minor coat protein of filamentous phages
pIX	minor coat protein of filamentous phages
POI	protein of interest
PVDF	polyvinylidene difluoride transfer membrane
pVI	minor coat protein of filamentous phages
pVII	minor coat protein of filamentous phages
pVIII	major coat protein of filamentous phages
RCBS	Research Collaboratory for Structural Bioinformatics
RPC	reverse phase chromatography
rSAP	alkaline phosphatase
RU	response unit
s.c.	subcutaneously
SA	streptavidin
SAR	structure-activity relationship
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOE	splicing by overlap extension
SOTI	*Spinacia oleracea* trypsin inhibitor
SPECT	single photon emission computed tomography
SPR	surface plasmon resonance spectroscopy
SSC	side scatter
ssDNA	single-stranded DNA
TAE	tris/acetic acid/EDTA
TFA	trifluoracetic acid
U.S.	United States
v/v	volume per volume
w/v	weight per volume

Table of contents

1 Introduction

1.1 The process of protein drug discovery and development

The discovery and development of drugs represents an essential process to identify potential biopharmaceutical agents aiming for the treatment or diagnosis of diseases. Novel and improved medicines are constantly needed to provide best therapy options for affected patients, as existing medications may lose effectiveness due to probable alteration in disease progression[1]. In the last decades, many different drug molecules from chemical and biological classes have been developed and entered the market for cancer treatment[2–4]. But also successful approval of new molecular entities as imaging agents has been reported[5]. However, the route of drug development is an elaborate and long-term process, in which different successive steps take place as shown in Figure 1.1. These include disease-related target selection, screening and identification of drug molecules, preclinical *in vivo* studies, clinical trials and eventually regulatory approvals[6]. The preclinical procedure follows no ultimate guideline, hence the single steps and the experimental assays varies with dependency on pre-defined requirements and aims. This chapter introduces the general goals of each drug discovery phase and presents the target selection based on cancer as well as highlights representative experimental approaches suitable for the development of biological drug molecules.

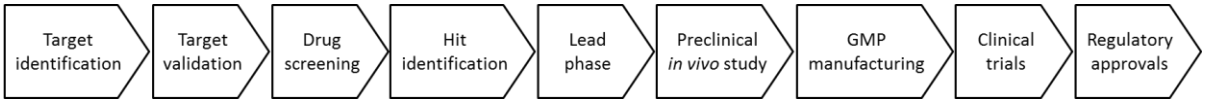


Figure 1.1 Overview of traditional protein drug discovery process. The drug discovery process typically begins with target identification and validation, followed by drug screening to select putative candidates. Those are further assessed within hit identification and optimized in the lead phase before starting *in vivo* studies. All these steps account to the preclinical phase. The most suitable drug candidates can be manufactured under good manufacturing practice (GMP) and enter clinical trials to receive regulatory approvals for entering the market. Adapted from [6].

A complete understanding of molecular mechanisms in disease pathogenesis is essential for directed development of novel therapeutic applications. The identification of disease-associated targets and their functional classification plays a crucial role for the development of ideally suited drugs[1]. Drug targets cover enzymes, receptors, metabolites, substrates, ion channels, transport proteins, DNA, RNA and ribosomes. With this, diverse therapeutic targets can be addressed to interfere with tumor growth and progression[7]. So far, many tumor-associated molecules have been described, mostly being overexpressed or selectively or aberrantly expressed in tumor cells[8]. However, selected target proteins are not necessarily druggable and they need to be validated. The purpose of this validation is to confirm experimentally the disease-related role in advance and additionally to ensure therapeutic benefit by addressing the molecule. The validation process relies on multiple approaches that involve different *in vitro* and *in vivo* assays[6]. Target expression can be determined on genetic or protein level, e.g. based on whole transcriptome sequencing, *in situ* hybridization techniques, quantitative PCR methods or by immunohistochemistry analysis on patient samples[9–11]. The clarification of target functionalities includes cellular experiments with RNA and protein modulation approaches. Furthermore, animal models using knockout or transgenic mice represent an important tool to mimic pathogen conditions in a complex biological system[9].

After a suitable target protein has been chosen, in the following step potential drug candidates are selected using diverse screening campaigns[6]. The selection is done either using classical techniques

like hybridoma technologies for antibody discovery, rational design approaches e.g. based on known target-specific ligands or by applying screening systems that enable the identification of novel target interaction ligands[12,13]. These include phage, yeast or ribosomal display technologies, all conceptually similar as they rely on the physical linkage of the presented protein of interest (POI) and its corresponding genetic information. With state-of-the-art recombinant DNA technologies large combinatorial libraries of peptides or proteins can be generated with up to billions of individual variants that are screened against the target molecule in a high-throughput manner[14–16]. Ribosomal display library sizes can reach up to 10^{14} members, but panning against whole cells is challenging due to weak ribosome-protein-mRNA complex stability[16]. In contrast, yeast display is mostly limited to a rather smaller library size, while allowing a panning procedure combined with fluorescence-activated cell sorting (FACS) to directly select target binding clones[15] that facilitate subsequent hit screening. Phage display represents a widely applied technology for drug discovery and enables target proteins to be used in soluble form or directly exposed on cell surfaces depending on their structural features[17,18]. A detailed overview of this system is described in chapter 1.2.

The hit identification process, which is closely connected to the screening section, aims to identify and confirm putative target interacting ligands using a limited number of pre-selected candidates[1,6]. The assessment of binding ability or biological activity is commonly conducted using biochemical and cellular high-throughput assays[19]. Flow cytometry screening for instance represents a widely applied system to gain multiparametric data sets[20]. Furthermore, in some cases a functional assay has to be developed in advance that fulfills the intended outcome robustness[6]. However, in hit screening approaches a high chance of detecting false positive or false negative hits exits, which is generally difficult to prevent. A subsequent second screen and deeper analysis of individual drug candidates is therefore usually conducted to confirm putative hits[19]. Further analysis, such as dose-response-curves, can be implemented as well[6]. Many drug molecules are being excluded in the course of the hit screening and identification, while the remaining candidates that passed all the testing enter the following drug discovery phase.

The lead phase describes the process of refining and optimizing the most promising drug candidates. A deeper characterization of the drug molecule facilitates directed fine-tuning. Important parameters concern the biochemical properties and the pharmacological profile, such as solubility, stability and absorption, distribution, metabolism and excretion (ADME). Whereas revealing the structure-activity relationship (SAR) is required to determine target binding sites and potential regions accessible for modifications. Hereby, starting points for improvements can be identified towards drug potency once having e.g. adverse biochemical properties or pharmacological profiles[6].

The final goal of the preclinical drug discovery process is to deliver most suitable drug candidates for clinical trials. Drug testing with *in vivo* animal models provides a characterization of the pharmacokinetic and pharmacodynamic performance in a whole biological environment to consider ADME. Furthermore, the drug efficacy is studied in *in vivo* disease models in order to analyze pharmacological response[6,21]. Once the preclinical studies are completed and the drug candidate passes all desired requirements the clinical trial process can be initiated.

Before entering the market the potential drug undergoes clinical testing. Clinical trials are separated into different stages aiming for an investigation of drug tolerability, dose, safety, side effects and efficacy in human cohorts. Furthermore, regulatory guidelines require preclinical data of *in vivo* and toxicology studies, manufacturing information as well as an outline of the clinical study protocol[22]. The first human study can be performed in phase 0, which was introduced to offer the opportunity of early drug testing with non-therapeutic purpose. This study is typically conducted with a small

cohort of healthy volunteers or patients and a very low single-dose application aiming to obtain first insights of drug performance[23]. In a phase I trial, 20 to 80 healthy volunteers are recruited for drug safety validation, including reporting of adverse effect, determination of dosing range and toxicity profile[22]. However, a phase I study is not necessarily performed with healthy volunteers, e.g. many anti-tumor drug molecules are directly tested in cancer patients[24]. The phase II study requires 100 to 300 patients to analyze drug efficacy and gain more safety information. The last pre-approval step, the phase III trial, with 1.000 to 3.000 patients is conducted to achieve ultimate drug safety and efficacy results. After accomplishment of clinical trials, a new drug application with all drug information and clinical data can be filed for evaluation and approval by the United States (U.S.) Food and Drug Administration (FDA). Once the FDA grants the approval, the drug molecule is commercially available in the U.S. and additional clinical studies are subjected to phase IV trials[22]. The drug evaluation for the European market is regulated by the European Medicines Agency (EMA). While the clinical trials follow roughly similar guidelines as described, the drug approval relies on four different pathways based on the drug class and the preference of the manufacturer[25].

1.2 Phage display technology

Since Smith introduced phage display by using filamentous bacteriophage fd for displaying peptides on the virion surface more than three decades ago[26], this technique gained strong value for drug discovery and development[17]. In the early 1990s, this methodology has been also demonstrated to enable functional presentation of complex folded proteins on phage surface[27–29]. The powerful phage display technique allows selection of polypeptides from large combinatorial libraries based on the genetic manipulation of filamentous phage surface proteins. Most commonly, the gene coding for the POI is either fused to the gene coding for the minor coat protein (pIII) or major coat protein (pVIII). The physical linkage of genotype and phenotype results in a surface presentation of the POI, while the phage contains the corresponding genetic information[26,30,31]. Both presentation systems differ with respect to the display valency of the presented protein with one to five copies in the case of pIII or the truncated form pIII´ and up to 2700 in pVIII[32,33]. In addition, the usage of the other coat proteins pVI, pVII or pXI for POI displaying is possible as well[34,35]. Figure 1.2 presents the structure of Ff filamentous bacteriophage and its coat proteins.

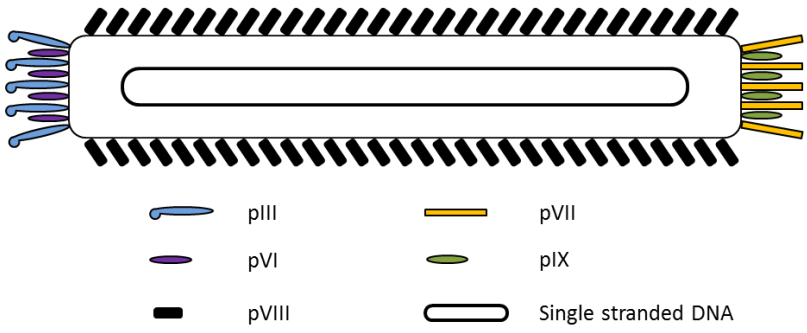


Figure 1.2 Structure of filamentous bacteriophage. An Ff bacteriophage consists of ssDNA encapsulated by major coat protein pVIII. The minor coat proteins pIII and pVI as well as pVII and pIX are located at each tip of the phage, respectively. Adapted from [36].

The Ff filamentous phage particles (M13, f1 and fd) infect *E. coli* that harbors the F episome coding for the F-pilus[37]. Bacterial infection is mediated by pIII coat protein interaction, which consists of N1-, N2- and C-domain. The N2-domain of the pIII coat protein binds to the tip of the bacterial F-pilus, while the N1-domain interacts with the periplasmic domain of the TolA protein. The C-domain is

involved in translocation of viral DNA into the cytoplasm of the *E. coli* cell. After translocation, the viral single-stranded DNA (ssDNA) is converted to double-stranded replicative form DNA (dsDNA) by the host replication machinery and serves as template to generate ssDNA through rolling circle replication and to express required phage proteins. Resulting ssDNA is covered by protein V dimer prior to packaging into coat proteins of phages and release from the cell through the bacterial membrane after assembly of the particle[32,38], Figure 1.3.

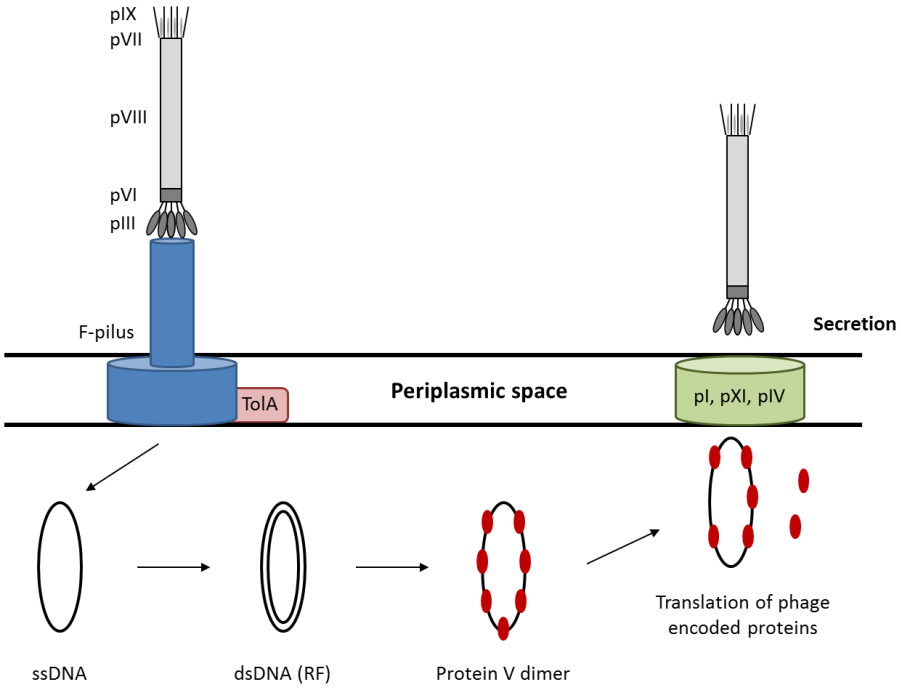


Figure 1.3 Life cycle of filamentous bacteriophage. Phage binds with pIII to tip of F-pilus on host cell and mediates transfer of viral ssDNA into cytoplasm. Host enzymes convert ssDNA to dsDNA for replication, protein expression and ssDNA synthesis. The single ssDNA is protected by dimers of protein V until phage assembling and export. Adapted from [39].

Phagemids are very popular vectors for creating phage display libraries that contain M13 and plasmid elements. A phagemid typically carries two replication origins to enable dsDNA as well as ssDNA synthesis and a gene coding for the POI fused to gene III/III´ or gene VIII protein, but lacking phage packaging relevant genes. Consequently, host cells harboring the phagemid cannot produce recombinant phages on their own, but they rather require a superinfection with helper phage particles (typically VCS M13 or M13K07) to provide missing gene elements[32,36].

A modified phagemid system, called pJuFo, was initially engineered by Crameri *et al.* for the construction and screening of cDNA libraries[40]. In this system, gene products are displayed through indirect fusion to phage coat proteins based on attached leucine zipper heterodimerization domains jun and fos. The expression cassettes of a gene coding for the POI and *fos* gene as well as a gene coding for the respective coat protein and *jun* gene allow an independent processing of the POI and coat protein to ensure functional displaying. Both constructs are secreted to the oxidative environment of the periplasmic space, in which an assembly and folding of complex proteins occurs[40–42].

Phage display is a long and well established method for screening combinatorial libraries that present large diversity of peptides or proteins with various size and complexity[43–45]. Conceptual design of phage libraries follows different rational and experimental strategies. Especially the displaying of antibody fragments showed high success in development of target specific antibodies based on

immune, naïve, semi synthetic or fully synthetic libraries[46–49]. Numerous reports describe the isolation of high affinity antibody fragments from phage display libraries directed against many individual therapeutic targets. Those molecules became candidates for engineering to accomplish clinical relevant therapeutics[50]. In the last decades, also non-immunoglobulin proteins have been functionally displayed on phages, including Affibodies, Anticalins, DARPins and Fynomers. The authors showed the capability of phage technology to screen large libraries consisting of a variety of protein scaffolds[51–54].

The success of this technology relies on many synergistic factors, such as the protein scaffold of choice, efficient protein displaying and library size[17]. An effective panning process is also required for ligand selection from a large phage library. Figure 1.4 shows a standard panning procedure of a phage library. A panning cycle begins with an incubation of the phage library with the target molecule that can e.g. be coated on tubes or beads. Removal of non-binding phages is done via washing procedure and target-binding phages are eluted for amplification in *E. coli*. In general, several successive panning cycles are performed in order to enrich target interacting proteins[32]. Consequently, even rare presented variants can be selected from a large library.

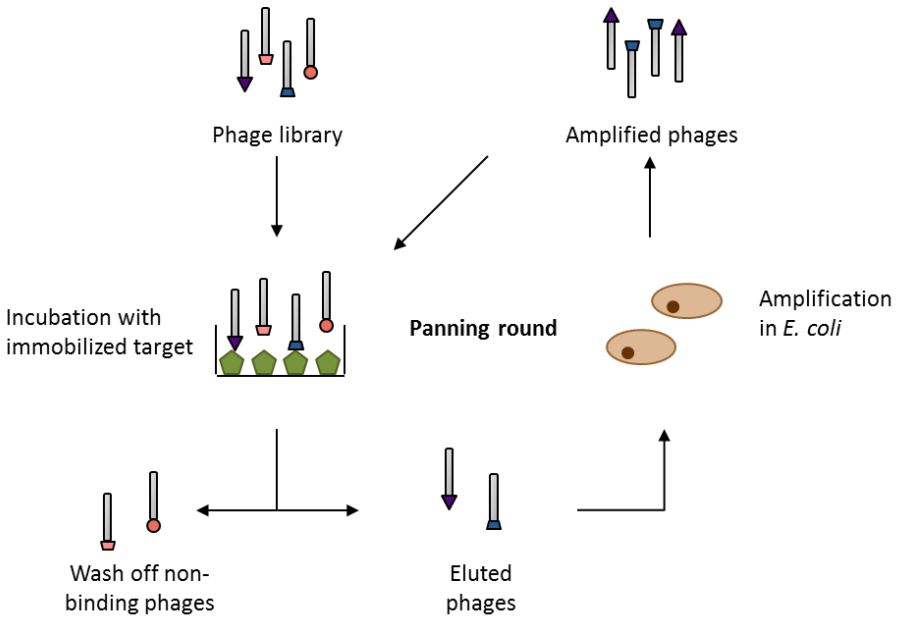


Figure 1.4 Schematic overview of a standard phage panning round. Phage library is incubated with target coated on beads or tubes. Non-binding phages are removed by washing, while target-interacting phages are eluted and amplified in *E. coli*. Adapted from [32].

Different panning opportunities of phage libraries are possible depending on the target type. The most common methods are based on the usage of target proteins in a soluble form, but this strategy is limited to proteins that are amenable to recombinant soluble production. A direct selection against whole cells on the other hand enables addressing of proteins in their native cellular environment, which is especially preferable for complex and membrane-bound proteins. Also *in vivo* phage library selection approaches have been described representing a further possibility especially for addressing organ- or tissue-specific targets in their natural environment[18,32].

In summary, phage display is a very powerful technology for selection of novel peptide or protein ligands against any desired target molecule and a key methodology in early phase drug discovery. An efficient and fast identification of promising candidates facilitates the development of innovative therapeutics.

1.3 Protein-based applications

The evolution of protein therapeutics highly accelerated with the first approval of a recombinant protein, which was human insulin for the treatment of diabetes mellitus in 1982[55]. Since then, proteins have become an important drug class for a broad range of clinical indications with more than hundred approved proteins[56]. Protein therapeutics are derived from a large variety of biological molecule types, including natural cytokines, growth factors, enzymes, soluble receptors, coagulation factors, monoclonal antibodies (mAbs) or engineered variants of these proteins[57]. Furthermore, they can be assigned into different categories based on their pharmacological activities. The groups are divided into enzymatic or regulatory activity, special targeting activity, vaccine purposes and diagnostic approaches[58].

Protein-based drugs gained a valuable source for targeted cancer therapy. In contrast to conventional surgery, chemotherapy and irradiation treatment, the high specificity and selectivity are advantageous to minimize numerous adverse effects mainly caused by undirected addressing of normal tissue[59].

1.3.1 Antibody-based therapeutics

mAbs represent a well-established and rapidly growing class of protein therapeutics[8]. By the end of 2014, 47 mAbs were approved for the treatment of various diseases[60], whereby the annual novel approvals increases constantly and numerous additional candidates are under clinical investigation[61,62]. Antibody-based therapeutics showed great success in clinical outcome of patients suffering from haematologic and solid cancers. Thereby, antibodies are engineered as anticancer and immunological drugs to specifically target tumor-associated molecules[63].

Antibodies, also termed as immunoglobulins, are naturally synthesized by B lymphocytes and as a part of the immune system, they play a major role in host defensing from bacterial or viral pathogens. The large ~150 kDa molecule consists of two identical heavy chains (HC) and light chains (LC) that are connected and stabilized via disulfide bonds. The complementarity-determining regions (CDRs) of the variable HC- and LC- domains form a highly specific antigen recognition site. While the constant fragment of an antibody mediates effector function, such as target neutralization or activation of cell-killing mechanism[64].

The remarkably inherit antibody properties are exploited for therapeutic applications. These especially involve the high target selectivity and specificity that allows precise binding to the tumor cells or compartments within the tumor microenvironment. Thereby, a direct or immune-mediated tumor killing can be evoked by alteration and modulation of various mechanisms. In addition, an antibody can serve as vehicle system to deliver conjugated drug compounds directly to the tumor site[8].

The development of therapeutically relevant antibodies relies on a wide range of engineering technologies that evolved over the last decades[65]. The advent was basically the discovery of the hybridoma technique allowing a continuously production of mAbs[66]. Chimerization and humanization strategies of rodent-derived mAbs were applied to improve their drug safety profile in terms of immunogenicity and tolerability. In chimeric mAbs the constant region is exchanged with human sequence, while humanized mAbs undergo complete replacement except for the murine CDRs[67]. The introduction of transgenic animals and displaying technologies further facilitated the development of fully human antibodies[68,69].

Despite the great success of mAb therapeutics, limitations emerged mainly related to their large molecular size and complex structural features. In general, the large size elicits poor penetration into solid tumors and therefore lowers the efficacy. The complex architecture and various post-translational modifications hamper controlled manufacturing process and results in high production costs[70,71]. Furthermore, the long circulation time of full-length antibody is adverse for molecular imaging, as a high-quality contrast resolution requires fast blood clearance[72].

To overcome the given limitations of antibodies, designing of antibody fragments, including Diabodies, Fab fragments and single-chain Fv fragments, with lowered molecular size while maintaining characteristic antigen binding properties, has been made. Despite the clinical approval of some Fab fragments, antibody fragments generally are prone to aggregate as a consequence of protein instability that can lead to increased immunogenicity potential[70,73].

To counteract the issue of large sized molecules and putative immunogenicity, a new class of non-immunoglobulin molecules emerged to fulfill the need. The next-generation binders evolved during the last decades with enhanced pharmacokinetic properties, efficacy and better safety profile over classical antibody[74], while being mostly simple structured, small, monomeric, soluble and stable[65].

1.3.2 Non-immunoglobulin protein scaffolds

Non-immunoglobulin protein scaffolds are a valuable alternative for the evolution of therapeutic and diagnostic agents, providing promising benefits apart from antibody-based therapeutics. These proteins are mainly divided into two categories, the domain-sized molecules and the constrained peptides. Affibodies, Affilins, Anticalins, DARPins, Fynomers, Kunitz domains with molecular weights between 6 and 20 kDa belong to the domain-sized category, while small-sized Avimers, bicyclic peptides and cystine-knot miniproteins are part of the constrained peptides category[75]. Figure 1.5 shows structural features of representative non-immunoglobulin protein scaffolds.

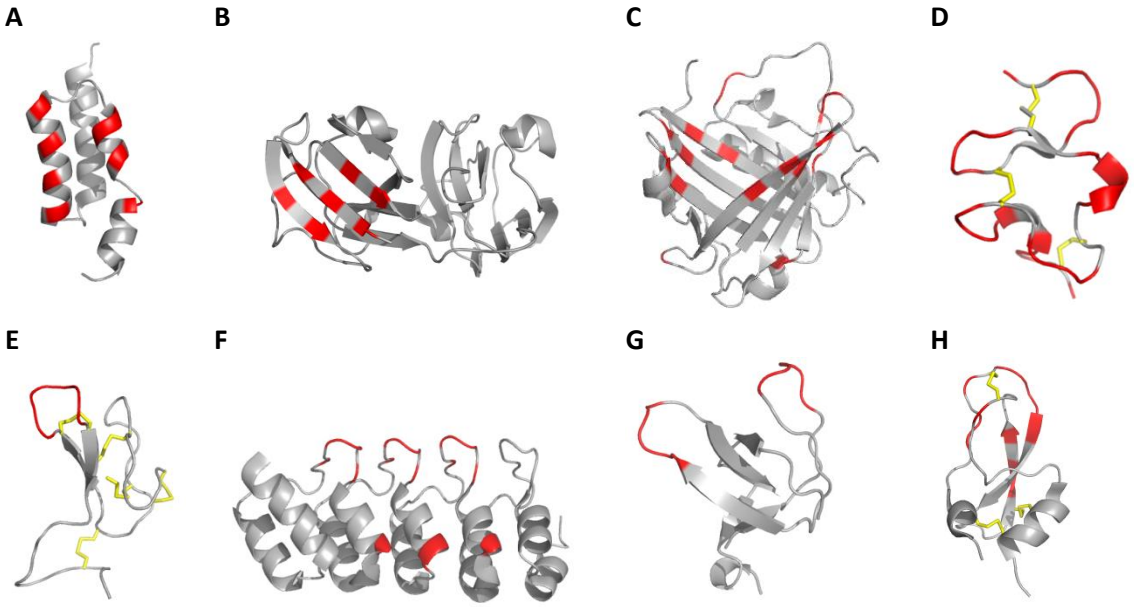


Figure 1.5 Three-dimensional structures of representative non-immunoglobulin protein scaffolds. **(A)** Affibody (PDB 2KZI) **(B)** Affilin (PDB 2JDF) **(C)** Anticalin (PDB 4GH7) **(D)** Avimer (PDB 1AJJ) **(E)** Cystine-knot miniprotein (PDB 1HYK) **(F)** DARPin (PDB 1SVX) **(G)** Fynomer (PDB 4AFS) **(H)** Kunitz domain (PDB 4BQD). Domain or loop accessibility for amino acid substitutions are presented in red and disulfide bridges are highlight as yellow sticks. Adapted from [75].

Typical characteristics of alternative protein scaffolds are their small size and high stability, leading to altered pharmacokinetic properties and target preference compared to mAbs. The compact size offers the opportunity to achieve good tissue penetration in therapeutic approaches. The generally simple architecture of non-immunoglobulin scaffolds exhibits limited protein complexity compared to antibody molecules, which are usually highly post-translationally modified. Therefore, the protein scaffolds are accessible for recombinant expression in bacterial hosts and in case of the smaller ones also for conventional solid phase chemical synthesis. This consequently leads to facile, controlled and cost-efficient manufacturing[74,76,77].

Although the alternative protein scaffolds have some general characteristics in common, their differences in sequence length, sequence composition and structural shape arise in altered pharmacokinetic behavior, such as half-time life, stability, immunogenicity, tissue penetration and bioavailability. Therefore in principle, the choice of the protein scaffold itself depends on the intended application[75,78].

Over the last decades, substantial progress has been made to generate tailored binding proteins using domain or loop specific engineering. Thereby, sequence and length modifications in the respective protein framework are created without having impact on the intrinsic structural fold. The derived molecules reveal binding properties, including specificity and affinity, comparable to traditional antibody behavior. Platforms for display of large combinatorial libraries represent the most common way in *de novo* discovery of target interacting ligands. So far, a large number of engineered protein scaffolds for diverse medical indications are currently under preclinical and clinical investigations[75,77].

The plasma kallikrein inhibitor Ecallantide (DX-88) is an approved molecule administered for the treatment of hereditary angioedema and originally derived from a Kunitz domain[79]. The approval of Ecallantide together with two further cystine-knot miniprotein-based drug molecules laid in principle the foundation of alternative protein scaffolds in terms of their clinical relevance[77]. With this, the cystine-knot miniprotein family gained particular attention for evolution of *in vivo* imaging technologies and targeted drug delivery purposes due to its beneficial pharmacological properties (chapter 1.3.3).

1.3.3 Cystine-knot miniproteins

The structural cystine-knot miniprotein (knottins) family comprises three main sub-groups, the inhibitor cystine-knot miniproteins, the cyclotides and the growth factor cystine knot proteins. These molecules are small peptides with typically 30 – 50 amino acids in size and a characteristic tertiary structure, which consists of intramolecular disulfide bonds connecting the antiparallel β-strands. The disulfide bonds within inhibitor cystine-knot miniproteins and cyclotides are formed between the first and fourth and between the second and fifth cysteine residues resulting in a macrocycle, through that a third disulfide bond between third and sixth cysteine residues passes (Figure 1.6). In growth factor knottin members the positions of cysteine residues connectivity differs[80]. The rigid knot-like structure mediates an extraordinary thermal, proteolytic and chemical stability of this knottin family[81]. The cyclotides subclass entails a further feature of the N-terminal to C-terminal cyclized backbone[82].

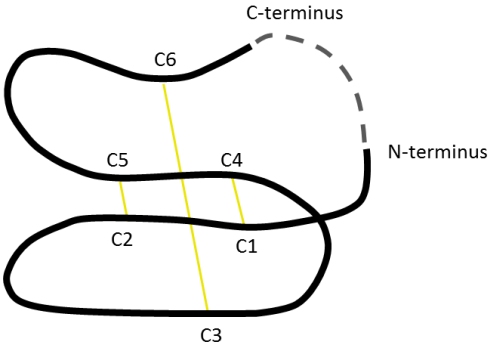


Figure 1.6 Schematic illustration of a cystine-knot miniprotein structure. Connective disulfide bonds are depicted in yellow. Dashed line in grey represents cyclization of C- and N-terminus present in the cyclotide family. Adapted from [83].

Cystine-knot miniproteins are found in a variety of different species and have diverse biological activities, including inhibitory, antimicrobial, cytotoxic or antifungal functions[80,83]. The functional loop regions of the constrained fold are important for bioactivity and are flexible with respect to sequence and length among various cystine-knot miniprotein types[84]. A large group of cystine-knot miniproteins have ion channel blocking activity, naturally occurring e.g. in the venom of spiders or marine cone snails[85]. The squash inhibitors represent a further group with protease inhibitory activities and prominent family members, such as cyclic *Momordica cochinchinensis* trypsin inhibitor II (MCoTI-II) and *Ecballium elaterium* trypsin inhibitor II (EETI-II) from the seeds of Cucurbitaceae[86,87]. Both proteins mediate their functional activity by a single loop located between the first and the second cysteine residues[88,89]. Figure 1.7 represents the three-dimensional structure and amino acid sequence of prominent cystine-knot miniproteins.

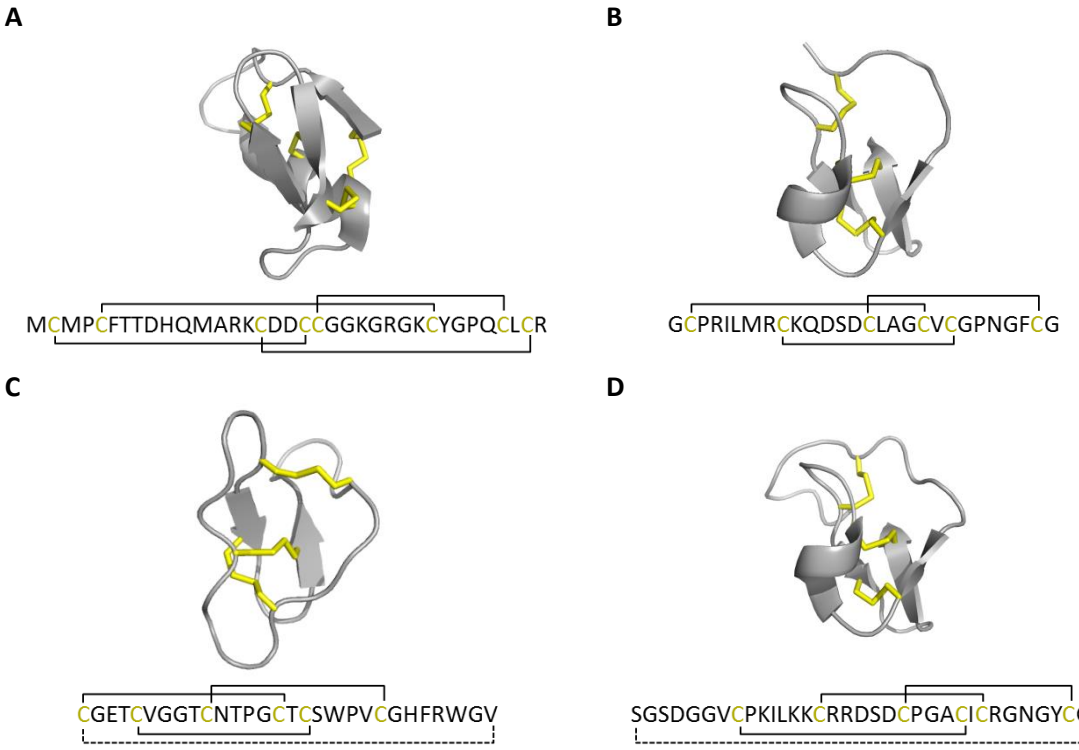


Figure 1.7 Three-dimensional structure and amino acid sequence of cystine-knot miniproteins. (A) Cholorotoxin (PDB 1CHL) **(B)** EETI-II (PDB 2IT7) **(C)** Kalata B1 (PDB 2LUR) **(D)** MCoTI-II (PDB 1HA9). Cystine-knot miniproteins are shown in grey and disulfide bridges are presented as yellow sticks. Dashed line represents cyclization of C- and N-terminus. The structures were obtained from RCBS PDB and illustrated using PyMol. Adapted from [83].

In recent years, several cystine-knot miniprotein members, including MCoTI-II, EETI-II, Agouti-related protein (AgRP) and *Spinacia oleracea* trypsin inhibitor (SOTI), have been engineered to specifically bind a variety of target proteins. Mostly, combinatorial libraries were created on the basis of the protein scaffold for yeast displaying approaches[90–93]. Despite the small size and limited randomization space, the exposed surface loops of cystine-knot miniproteins well tolerate sequence and length modifications without losing the structural characteristics.

The EETI-II framework underwent functional loop grafting in combination with directed evolution via yeast display to develop $\alpha_v\beta_3/\alpha_v\beta_5$ or $\alpha_v\beta_3/\alpha_v\beta_5/\alpha_5\beta_1$ integrin binding variants. These integrins are highly tumor-associated receptors expressed on the surface of various cancer cells[90]. Ever since, the selected 2.5D and 2.5F knottins were intensively investigated for the path to tumor imaging approaches. Both molecules thereby served as targeted moiety and were functionalized with various radionuclides or fluorescence dyes. A tumor-specific targeting was demonstrated with different peptide variants in a human glioblastoma xenograft mouse model[94–96]. While a ^{64}Cu-DOTA conjugated 2.5F knottin variant specifically accumulated in lung nodules of a transgenic mouse model[97]. Another compound, chlorotoxin (CTX) from scorpion venom, was initially discovered to target glioma-specific chloride ion channel[98]. This molecule was the starting scaffold for designing imaging compounds with specific accumulation in different tumor models and great progress as tumor diagnostic tool[99–101]. But also AgRP-based imaging agents for visualization of glioblastoma xenograft tumors[102] and an acyclic variant of MCoTI-II for A431 xenografts have been reported[91].

Apart from the intense preclinical development of many promising imaging candidates for the clinic, to date two knottin-derived drug molecules achieved regulatory approval[77]. Ziconotide, a ω-conotoxin MVIIa from *conus magnus* snail venom, was the first approved knottin-based protein therapeutic for the treatment of neuropathic pain. This drug is administered intrathecal and interacts with presynaptic N-type calcium channels, which leads to an inhibition of neurotransmission[103]. Linaclotide represents a further knottin drug molecule to treat patients suffering from irritable bowel syndrome with constipation. Linaclotide shares homology with the hormones guanylin and uroguanylin and acts as agonist of the guanylate cyclase C receptor[104].

Taken together, cystine-knot miniproteins exhibit beneficial pharmacological properties for tumor diagnostic and therapeutic approaches. Their simple structure allows convenient chemical manufacturing and facilitates site-directed compound conjugation[84]. Based on the high proteolytic and thermal stability an oral drug administration may be feasible with the potential to simplify regular medications[81]. Furthermore, the fast clearance of knottin-like compounds is advantageous for imaging purposes in order to achieve high tumor to background contrast[83]. The excellent performance of this binding molecule class as molecular scaffolds in preclinical studies and the FDA approval of two knottin-based drugs encourage their further engineering for cancer and other diseases.

1.4 Targeting fibronectin extra domain B

The extracellular matrix (ECM) is a complex system composed of different macromolecules creating a dynamic and well-defined network in the pericellular space. The matrix components contribute to structural integrity and regulate multiple cellular functions[105]. Fibronectin (FN) represents an essential multidomain glycoprotein of the ECM with a critical role in distinct physiological processes, including cell adhesion, migration, differentiation and growth[106]. FN is a large, dimeric protein paired by nearly two identical subunits through C-terminal disulfide bonds. Each monomer consists of type I, type II and type III domains that arrange to form a linear array. Type I and type II modules possess

45 – 60 amino acids and intramolecular disulfide bonds, whereas type III modules are around 90 amino acids in length without disulfide bonds[107,108]. FN typically contains 12 type I, 2 type II and 15-17 type III modules[109]. Additionally, extra domain A (EDA), extra domain B (EDB) and type III connecting sequence (IIICS), all type III modules of FN, can be inserted upon alternative splicing[110]. As a multifunctional protein FN exhibits multiple binding sites for various cellular and matrix proteins such as fibrin, collagen and heparin[111]. Figure 1.8 shows a schematic representation of FN with the three types of repeating modules and its splice-in domains.

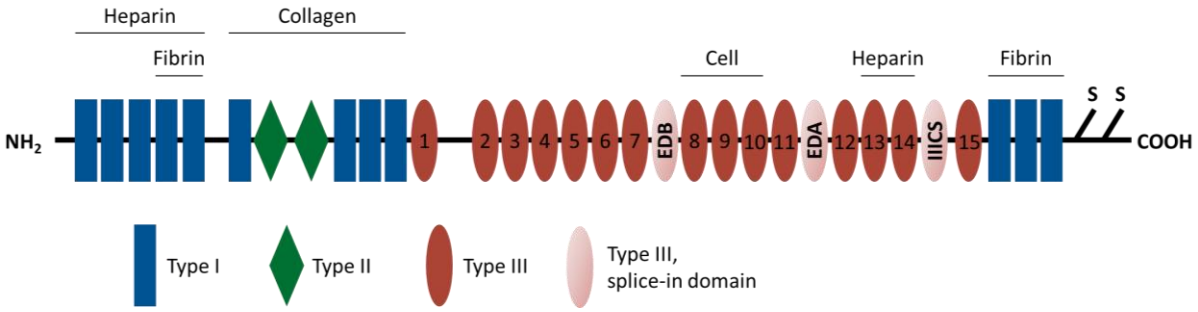


Figure 1.8 Structure of fibronectin. Fibronectin consists of three different repeating modules, type I (blue), type II (green) and type III (red) as well as spliced-in domains (rose). Binding sites for various molecules and two disulfide bonds are depicted. Adapted from [111].

FN naturally occurs in two distinct forms as plasma FN (pFN) and cellular FN (cFN)[112]. The pFN is expressed by hepatocytes and secreted as soluble protein into the blood[113]. pFN plays a role in hemostasis, in which it accumulates at the damaged tissue and incorporates in fibrin clots[114]. In pFN the splice-in domains EDA and EDB are completely absent, while the IIICS domain exists in one subunit[110,115]. In contrast, the insoluble cFN is produced by fibroblasts, endothelial cells and other cell types[109], and found in the ECM assembled to a fibrillar network[116]. FN is expressed as multiple alternatively spliced variants derived from a primary transcript[117,118]. The splice-in domains are inserted at three distinct positions of FN with EDB being located between type III domains 7 and 8, EDA between type III domains 11 and 12 and variable regions of IIICS between type III domains 14 and 15[119]. The oncofetal EDA and EDB domains arise from a single exon and can be included in the FN mRNA upon exon skipping[110,120].

The highly conserved EDB region consists of 91 amino acids and occurs with 100 % sequence identity among different species, including mice, rats, rabbits, dogs and humans[121]. EDB is usually absent in normal adult tissues, but involved in tissue remodeling and angiogenesis[121,122]. In addition, EDB has been found in many different cancer types with abundant expression pattern in the neovasculature of several aggressive solid human tumors such as head and neck, breast and glioma cancer[121,123–125]. Being a selective tumor marker EDB has been investigated as a target for cancer imaging and therapy in research and development as well as in several early clinical studies[126]. The most advanced EDB ligand has been isolated from an antibody phage display library in the lab of Dario Neri[127]. This L19 antibody exhibits specific tumor targeting properties and has been extensively investigated as vehicle for many diagnostic and therapeutic applications[128–131]. Despite the high clinical impact of the L19 antibody, several non-antibody ligands have been developed to specifically target EDB. These include binding ligands based on Fynomers, peptides, Anticalins and an ubiquitin-derived protein, which were all selected from combinatorial phage libraries with nanomolar or micromolar affinities[54,132–135]. The EDB binding Anticalins were used as diagnostic tracer to detect human gliomas[136]. Apart from protein-based molecules, two EDB recognizing peptides were generated and conjugated with fluorescence dyes for preclinical *in vivo* imaging studies. One report demonstrates the effective visualization of

human prostate tumor xenografts with a cyclic nonapeptide ZD2[132]. In another study, artificial high-affinity peptides (so called aptides) have been shown to specifically target human glioblastoma tumor xenografts[135]. The development of many EDB-specific molecules emphasizes this tumor-associated protein to be an ideal model for the generation of further non-antibody ligands[126].

1.5 Purpose of the thesis

In the last decades non-immunoglobulin protein scaffolds emerged as a promising alternative to complement antibodies and antibody derivatives for therapeutic and diagnostic applications. Engineered cystine-knot miniproteins represent an alternative protein scaffold with drug-like properties such as high target affinity and specificity, extraordinary stability, solubility and favorable pharmacokinetic characteristics. Additionally, these small peptidic molecules with a simple architecture facilitate a straight-forward chemical production and the construction of multi-functional fusion molecules. Thus far, our internal cystine-knot miniprotein technology platform consists of one combinatorial MCopt 1.0 phage library, which was constructed with sequence variation in the open chain variant of the squash trypsin inhibitor MCoTI-II originally found in gac fruit. The discovery of novel candidates is based on phage display screenings within the soluble screening pipeline using recombinant target proteins. But standard phage display protocols are not readily applicable for certain proteins, especially those with complex structural features.

Therefore, one aim of this thesis was the development of a cell-based screening approach for addressing target membrane proteins or other proteins that are not amenable to recombinant production in a soluble form. The phage screening procedure using target overexpressing cell lines should be efficient for cystine-knot miniprotein selection. Whereas the downstream hit identification process should be reliable in analysis of individual clones with respect to their target binding activity and specificity. Additionally, the library repertoire should be extended by the construction of novel cystine-knot miniprotein phage libraries based on sophisticated rational design strategy. The libraries should reveal good quality and functionality, including high clonal diversity and proper phage displaying of cystine-knot miniproteins. Another part of this work is focusing on the advancement of a cystine-knot miniprotein for targeting of EDB expressing tumors. The ligand should be further characterized and engineered to develop an EDB specific imaging probe, which should be finally evaluated in significant murine tumor model systems.

2 Materials and methods

2.1 Materials

2.1.1 Cell lines

CHO-K1	ATCC
Flp-In™-CHO	Thermo Fisher Scientific
U-87 MG	ATCC

2.1.2 Cell culture media

Dulbecco's Modified Eagle Medium (DMEM)/-F12 + GlutaMAX™	Gibco
Eagle's Minimum Essential Medium (EMEM)	ATCC
Ham's F-12 Nutrient Mix	Thermo Fisher Scientific

Cell line	**Cell culture media**	**Supplements**
CHO-K1	DMEM/-F12 + GlutaMAX™	10 % FBS (v/v)
CHO-K1-DSG-3, CHO-K1-dsRed-Mock, CHO-K1-FAP-α, CHO-K1-Mock	DMEM/-F12 + GlutaMAX™	10 % FBS (v/v), 200 µg/mL hygromycin-B
U-87 MG	EMEM	10 % FBS (v/v)
Flp-In™ wt	Ham's F-12 Nutrient Mix	10 % FBS, 2 mM L-glutamine, 100 µg/mL zeocin
Flp-In™-CHO-FAP-α	Ham's F-12 Nutrient Mix	10 % FBS (v/v), 2 mM L-glutamine, 200 µg/mL hygromycin-B

2.1.3 Microorganisms

***E. coli* strain**	**Genotype**	**Purpose**	**Manufacturer**
SHuffle® T7 Express	*fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (Spec^R, lacI^q) ΔtrxB sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10 --Tet^S) endA1 Δgor Δ(mcrC-mrr)114::IS10*	Expression of cystine-knot miniproteins	New England BioLabs
SS320 (MC1061F\')	[F\'*proAB+lacIqlacZ*ΔM15 *Tn10* (tetr)*] hsdR mcrB araD139 Δ(araABC-leu)7679ΔlacX74 galUgalK rpsL thi*	MCopt 2.0/2.1/3.0 phage library construction	Lucigen
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacI^{q}ZΔM15 Tn10 (Tet^r)]*)	Phagemid preparation, phage sub-library construction	Stratagene

2.1.4 Bacterial growth media

LB-agar was prepared by adding of 1.5 % agar-agar powder to LB-medium. LB-agar and cultivation medium was autoclaved at 121 °C at 1 bar for 20 min. Sterile filtered antibiotics and glucose were added to LB-agar after cooling below 60 °C. LB and SOC-medium were stored at 4 °C and supplements were added before usage.

LB-medium	1 % (w/v) bacto-tryptone, 0.5 % (w/v) yeast-extract, 0.5 % (w/v) NaCl in 1 L H_2O
$LB_{CarbGlu}$-medium	LB-medium + 100 µg/mL carbenicillin + 0.4 % (w/v) glucose
$LB_{CarbKanalPTG}$-medium	LB-medium + 100 µg/mL carbenicillin + 25 µg/mL kanamycin + 1 mM IPTG
$LB_{TetKana}$-medium	LB-medium + 12.5 µg/mL tetracycline hydrochloride + 25 µg/mL kanamycin
LB_{Tet}-medium	LB-medium + 12.5 µg/mL tetracycline hydrochloride
$LB_{CarbChlora}$-medium	LB-medium + 100 µg/mL carbenicillin + 10 µg/mL chloramphenicol + 0.4 % (w/v) glucose
SOC-medium	0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 0.05 % (w/v) NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ in 1 L H_2O; upon usage, add 20 mM glucose

2.1.5 Viral strain

VCSM13 bacteriophage	Stratagene

2.1.6 Vectors

Vector name	Purpose	Manufacturer/origin
pBirAcm	Generation of *E. coli* SHuffle® T7 Express pBirAcm cells	Avidity
pJuFoIII	Generation of MCopt 2.0/2.1/3.0 phage library	Prof. Mark Suter, SIAF, Davos, Switzerland
pPDIII-1	Generation of MCopt 2.0 phage sub-library	Derivative of pJuFoIII
pET32-LibEx	Expression of MCopt 1.0 cystine-knot miniprotein fusions	Derivative of pET-32a, Novagen
pET32-LibEx_GS	Expression of MCopt 2.0/2.1/3.0 cystine-knot miniprotein fusions	Derivative of pET-32a, Novagen
pET-32-AX	Expression of biotinylated cystine-knot miniprotein fusions	Derivative of pET-32a, Novagen
Piggy Bac transposase	Generation of stable cell line	Sanger institute
pPB transposon	Generation of stable CHO-K1-DSG-3 cell line and CHO-K1-dsRed-Mock cell line	Sanger institute

The pPB transposon vector was used to construct pPB transposon_DSG-3 and pPB transposon_dsRed vectors. The pPB transposon_DSG-3 carries gene segments coding for human Desmoglein-3 (DSG-3, Uniprot ID P32926, amino acids M1-I999) and green fluorescent protein (GFP). The pPB transposon_dsRed carries the coding gene for *Discosoma* red fluorescent protein (dsRed).

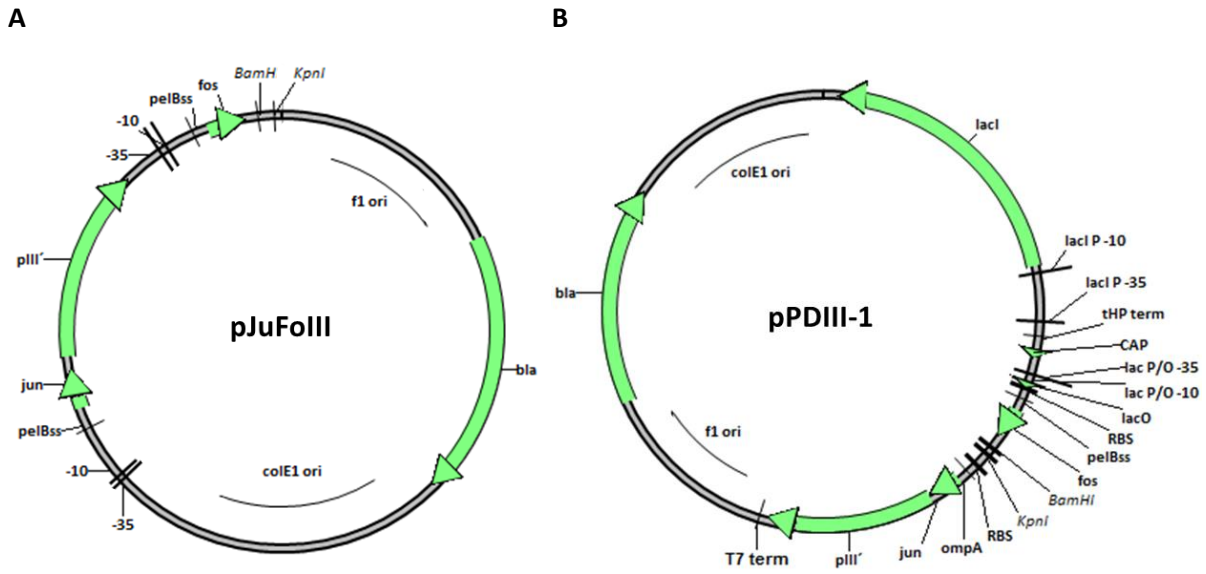


Figure 2.1 Phagemid vector maps. Genetic elements: colE1 ori: *E. coli* origin of replication; f1 ori: phage origin of replication; pelBss: pelB periplasmic leader sequence; ompA: ompA periplasmic leader sequence, bla: beta-lactamase; lacO: lac operator; -10/-35: lac promotor/operator region; fos/jun: leucine zipper domains; pIII´: gene encoding for a truncated recombinant minor coat protein of M13; RBS: ribosomal binding site; tHP term: tHP terminator sequence; T7 term: T7 terminator sequence. Arrows are shown in direction of transcription. DNA fragments encoding for MCopt 2.0/2.1/3.0 sequences were cloned using unique *Bam*HI and *Kpn*I restriction sites in frame with a pelBss and fos. **(A)** pJuFoIII phagemid vector. **(B)** pPDIII-1 phagemid vector.

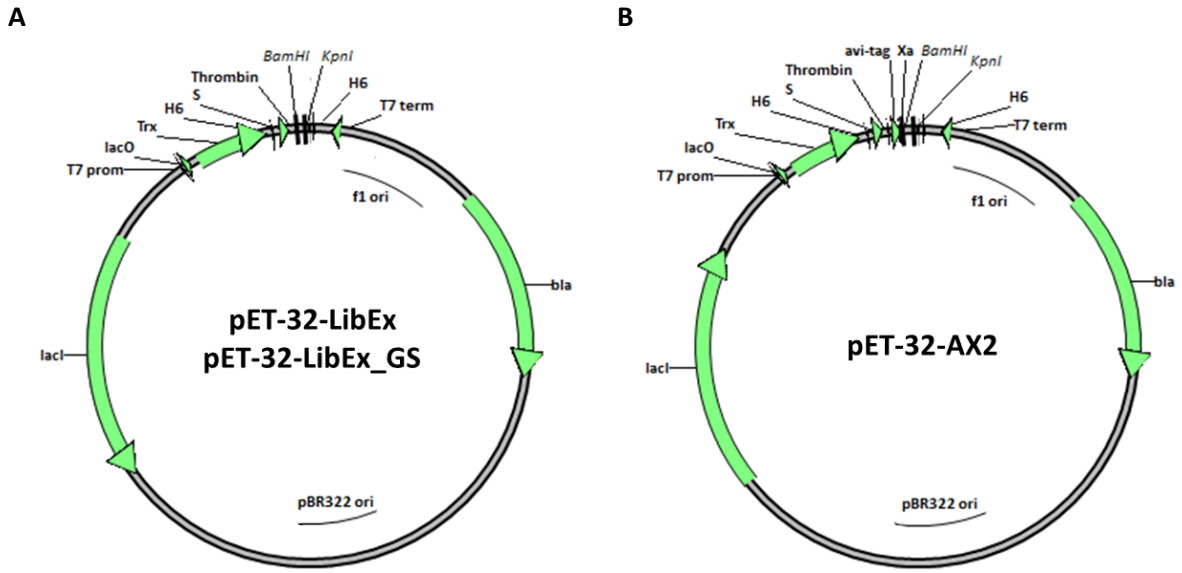


Figure 2.2 Expression vector maps. Genetic elements: f1 ori: phage origin of replication; bla: beta-lactamase; T7 prom: regulatory T7 promotor; lacO: lac operator; Trx: gene encoding for *E. coli* thioredoxin-A; H6: hexahistidine-tag; S: S-tag, peptide derived from RNase A, Thrombin: cleavage site for thrombin protease; avi-tag: specific peptide for biotinylation[137]; Xa: cleavage site for factor Xa protease; T7 term: T7 terminator sequence. Arrows are presented in direction of transcription. DNA fragments encoding for MCopt 1.0/2.0/2.1/3.0 sequences were cloned using unique *Bam*HI and *Kpn*I restriction sites. **(A)** pET-32-LibEx and pET-32-LibEx_GS vector. The pET-32-LibEx_GS vector carries six additional nucleotides encoding for the amino acids glycine and serine downstream of the thrombin cleavage site. **(B)** pET-32-AX2 vector.

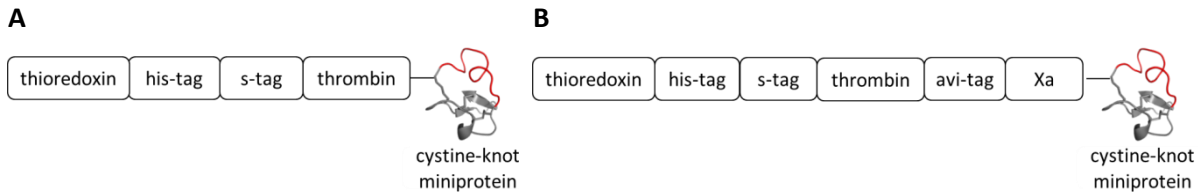


Figure 2.3 Schematic representation of the cystine-knot miniprotein fusion elements. (A) Protein expression via pET-32-LibEx or pET-32-LibEX_GS resulted in cystine-knot miniprotein fusion with thioredoxin-A (for efficient disulfide bond formation in the cytoplasm), His-tag (for rapid purification or detection by an antibody), S-tag (for detection by an antibody), thrombin cleavage site (for removal of fusion tag) and cystine-knot miniprotein in successive order. **(B)** Protein expression using pET-32-AX2 resulted in the similar cystine-knot miniprotein fusion as described before, but with an additional avi-tag sequence (for site-specific biotinylation using BirA biotin ligase) and a factor Xa protease cleavage site.

2.1.7 Antibodies and streptavidin conjugates

Alexa Fluor® 647 rat anti-mouse CD31, clone 390	BioLegend
APC mouse anti-his-tag, clone AD1.1.10	R&D Systems
APC streptavidin	eBioscience
Cy3 streptavidin	Rockland Immunochemicals
Cy5 goat anti-mouse IgG, 115-175-146	Dianova
HRP goat anti-mouse IgG, 554002	BD Bioscience
HRP goat anti-rabbit IgG, A6154-1ML	Sigma-Aldrich
HRP goat anti-S-tag, ab19324	Abcam
HRP rabbit anti-S-tag antibody, ab18589	Abcam
HRP streptavidin	Thermo Fisher Scientific
Mouse anti-human DSG-3, clone MM0245-4W25	Abcam
Mouse anti-human EDB+Fn, clone BC-1	Abcam
Mouse anti-human FAP-α, clone 1E5	Abnova
Rabbit anti-c-Fos, HPA018531	Sigma-Aldrich
Rabbit anti-M13, 27-9420-01	Sigma-Aldrich

2.1.8 Enzymes

Accutase®	Sigma-Aldrich
Alkaline Phosphatase (rSAP)	New England BioLabs
*Apa*I	New England BioLabs
*Bam*HI	New England BioLabs
Benzonase®	Merck Millipore
*Bgl*II	New England BioLabs
FastDigest *Apa*I	Thermo Fisher Scientific
FastDigest *Bam*HI	Thermo Fisher Scientific
FastDigest *Kpn*I	Thermo Fisher Scientific
*Kpn*I	New England BioLabs
Lysozyme	Merck Millipore
T4 DNA ligase	New England BioLabs
T4 DNA ligase	Thermo Fisher Scientific
Taq DNA polymerase	Thermo Fisher Scientific
Thrombin from human plasma	Sigma-Aldrich

Restriction enzymes and ligase for phage library generation as well as for vector preparation was used from New England BioLabs. All other cloning works were performed with enzymes obtained from Thermo Fishier Scientific. Enzymes were applied according to manufacturer´s instruction.

2.1.9 Recombinant proteins

Name	Protein description	Purpose
FN-B	Single Fibronectin EDB domain	Target protein
FN-6789	Fibronectin type III domains 6 – 9	Off-target protein
FN-67B89	Fibronectin EDB domain, flanked by surrounding type III domains	Target protein
Trx-MC-Myc-010	Cystine-knot miniprotein with grafted c-myc epitope in loop 1	Negative control

In this work, all recombinant proteins were produced in *E. coli* and purified in-house. Recombinant human fibronectin EDB either as a single domain (FN-B, Uniprot ID P02751, isoform 7, amino acid E1265-T1355) or flanked by its surrounding type III domains (FN-67B89) served as target proteins, while domains 6 – 9 without EDB (FN-6789) was used as control. All variants were expressed with a C-terminal hexahistidine (H6) tag and stored in aliquots in PBS supplemented with 5 % mannitol and 5 % trehalose at −20 °C. Biotinylated proteins were obtained via primary NH_2-groups using EZ-Link Sulfo-NHS-LC-Biotin according to manufacturer's instructions. Trx-MC-Myc-010 protein, an oMCoTI-II scaffold with grafted c-myc epitope in loop 1, was stored in PBS at −20 °C.

2.1.10 Synthetic DNA fragments

Gene fragments encoding for MCopt 2.0 or MCopt 2.1 library sequences were directly assembled by GeneArt™ (Thermo Fisher Scientific) using trinucleotide mixtures encoding for nineteen amino acids excluding cysteine at the variable positions. The MCopt 2.0 and MCopt 2.1 DNA fragments were amplified by the manufacturer and used for library construction or as a template for fragment amplification.
For generation of alanine scanning MC-FN-010 derivatives mutations were either introduced via PCR or the whole coding sequence was assembled via direct synthesis of GeneArt™ Strings™ fragments (Thermo Fisher Scientific).

2.1.11 Oligonucleotides

Name	Oligonucleotides sequence (5´-3´)	Purpose
BamHI-Lip-up	GCAGATCTGGGGGTGGTTCGGCCGGATCCGC	MCopt 1.0 insert amplification
Bsp120I-MC-lo	GCGCGGGCCCTCATTAACCGCAGTAACCGTTACCA	MCopt 1.0 insert amplification
M13 fwd long	TTGTAAAACGACGGCCAGTGAGC	MCopt 2.0 DNA fragment or MCopt 3.0 DNA fragment amplification
M13 rev long	GGAAACAGCTATGACCATGTTAATGCAGC	MCopt 2.1 DNA fragment or MCopt 3.0 DNA fragment amplification
M13 uni (-43) lo	AGGGTTTTCCCAGTCACGACGTT	Colony PCR
MC-FA-010-up	GGAAAGTGTCCTTATTCGAACTGG	Colony PCR
MCopt 2.0 rev	AATACATGCACCCGGGCAATCGCTATCGCGAC	MCopt 2.0 DNA fragment amplification
MCopt 2.1 fwd	GATAGCGATTGCCCGGGTGCATGT	MCopt 2.1 DNA fragment amplification

2.1.12 Kits

EZ-Link Sulfo-NHS-LC-Biotin	Thermo Fisher Scientific
High Pure PCR Product Purification Kit	Roche
High Pure Plasmid Isolation Kit	Roche
Plasmid DNA purification kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
Wizard® *Plus* SV Minipreps DNA Purification System	Promega Corporation
Wizard® SV Gel and PCR Clean-up System	Promega Corporation

2.1.13 Reagents and chemicals

10x Blocking solution, B6429-500ML	Sigma Aldrich
10x Fermentas fast digest buffer	Thermo Fisher Scientific
10x Taq buffer (NH_{4+})	Thermo Fisher Scientific
2-methylbutane	Carl Roth
50x Tris/Acetic Acid/EDTA (TAE)Puffer	Bio-Rad Laboratories
6x DNA loading dye	Thermo Fisher Scientific
Acetic acid, 100 %	Carl Roth
Acetonitrile	Carl Roth
Active cole	Carl Roth
Agarose low melting	Bio-Rad Laboratories
BactoTM Trypton	AppliChem
BactoTM Yeast extract	BD Biosciences
Betaine solution, 5 M, PCR reagent	Sigma Aldrich
Beta-Mercaptoethanol	Gibco
Bromophenol blue sodium salt	AppliChem
Carbenicillin disodium salt	Carl Roth

Chloramphenicol	Carl Roth
Complete EDTA free protease inhibitor cocktail	Roche
Coomassie Brilliant blue R-250	AppliChem
D(+)-glucose monohydrate	Merck Miilipore
D-biotin	Carl Roth
Difco™ Agar	BD Biosciences
Dimethyl sulfoxide (DMSO)	AppliChem
Disodium hydrogen phosphate	Carl Roth
Dithiotreitol (DTT)	Carl Roth
dNTP set (100 mM, each A,C,G,T)	Thermo Fisher Scientific
DPBS	Gibco
eFluor 450™	eBioscience
E-PAGE™ loading dye, 4x	Life Technologies
Ethanol >96 % (EtOH)	Carl Roth
Ethylenediaminetetraacetic acid (EDTA)	AppliChem
GelRed™ Nucleic Acid Gel Stain	Biotium
Glycerol	Carl Roth
Glycine	Carl Roth
Hoechst 33342	Thermo Fisher Scientific
Hydrochloric acid, 37 %	Carl Roth
Hydrochloric acid, 5 M	AppliChem
Imidazole	Carl Roth
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	AppliChem
Kanamycin sulfate	Carl Roth
Magnesium chloride hexahydrate	AppliChem
Magnesium sulfate heptahydrate	AppliChem
Mannitol	Carl Roth
Methanol, > 99.7 % (MeOH)	Carl Roth
$MgCl_2$ (25 mM)	Thermo Fisher Scientific
Milk powder	Carl Roth
Nuclease free, sterile water	Carl Roth
NuPAGE® MES SDS Running Buffer (20x)	Life Technologies
Polyethylene glycol (PEG)	AppliChem
Polyethylenimine (PEI)	Sigma-Aldrich
Potassium chloride	AppliChem
Sodium carbonate	AppliChem
Sodium chloride (NaCl)	Carl Roth
Sodium dodecyl sulfate (SDS), ultra-pure	Carl Roth
Tetracycline hydrochloride	Carl Roth
TMB substrate (3,3',5,5'-Tetramethylbenzidine)	Sigma-Aldrich
Trehalose	AppliChem
Triethylamine (TEA)	Sigma-Aldrich
Trifluoracetic acid (TFA)	AppliChem
Tris ultrapure	AppliChem

2.1.14 Standard DNA and protein marker

Gene Ruler™ 1 kbp DNA ladder	Thermo Fisher Scientific
Gene Ruler™ 50 bp DNA ladder	Thermo Fisher Scientific

Precision Plus Protein™ Dual Color Standard	Bio-Rad Laboratories
Precision Plus Protein™ Dual Xtra Prestained Standard	Bio-Rad Laboratories

2.1.15 Buffer

4x non-reducing SDS-loading buffer	250 mM Tris, 34 % (w/v) glycerol, 8.2 % (w/v) SDS, spatula bromophenol blue, pH 6.8
4x reducing SDS-loading buffer	250 mM Tris, 34 % (w/v) glycerol, 8.2 (w/v) % SDS, 5 % (w/v) β-Mercaptoethanol, spatula bromophenol blue, pH 6.8
Coating buffer	50 mM Na_2CO_3 pH 9.4
Coomassie destain solution	45 % (w/v) MeOH, 10 % (w/v) acetic acid
Coomassie stain solution	0.25 % (w/v) coomassie, 40 % (w/v) EtOH, 10 % (w/v) acetic acid
FACS buffer	DPBS, 5 mM EDTA, 5 % FBS, pH 7.5
Freezing buffer	FBS, 10 % DMSO
IMAC A buffer	20 mM Tris-HCl pH 8.0, 10 % glycerol, 500 mM NaCl, 10 mM imidazole
IMAC B buffer	20 mM Tris-HCl pH 8.0, 10 % glycerol, 500 mM NaCl, 500 mM imidazole
Lysis buffer	20 mM Tris pH 8.0, 2 mM $MgCl_2$, 20 mM NaCl, 0.1 mg/mL Lysozym, 15 U/mL Benzonase
PBS	14 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.5
PBS-T	14 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.1 % (w/v) Tween® 20pH 7.5
PEG/NaCl solution	25 % (w/v) polyethylene glycol, 2.6 M NaCl
TAE buffer	40 mM Tris, 20 mM acetic acid, 1 mM EDTA
TBS-T	50 mM Tris, 150 mM NaCl, 0.1 % (w/v) Tween® 20, pH 7.4
Thrombin cleavage buffer	20 mM Tris, 150 mM NaCl, 1.5 mM $CaCl_2$ and 5 % (w/v) glycerol, pH 8.45
Western Blot transfer buffer	25 mM Tris, 192 mM glycine, 20 % (w/v) MeOH

2.1.16 Consumables

24-deepwell plates, 10 mL	Thermo Fisher Scientific
48-well PCR plates	Thermo Fisher Scientific
6-well plates, 12-well plates, 24-well plates, 96-well plates	Thermo Fisher Scientific
96-deepwell plates, 2 mL	Carl Roth
96-well plates polystyrene and MaxiSorp™	Thermo Fisher Scientific
96-well round-bottom plates	Thermo Fisher Scientific
Adhesive PCR film	Thermo Fisher Scientific
Adhesive seals	Thermo Fisher Scientific
Amicon® Ultra-4 and -15 Centrifugal Filter Units (3 kDa NMWL, 10 kDa NMWL)	Merck Millipore
AMPure XP Beads	Beckman Coulter
Aspiration pipette 2 mL	Greiner bio-one
CASY cup	OLS OMNI Life Science

CASYton solution	OLS OMNI Life Science
Cell culture flask 175 cm^2, 75 cm^2, 25 cm^2	Greiner bio-one
Cell culture triple flask	Thermo Fisher Scientific
Cell scraper	Heathrow Scientific
Cell strainer 70 µm	Greiner bio-one
Combitip 25 mL, 50 mL	Eppendorf
Cover slip 75x25 mm	Carl Roth
Cryo tubes 1.8 mL	Sigma-Aldrich
Cuvette	BRAND
Dialysis hose (exclusion size, 10 kDa)	Carl Roth
Dialysis tubes	Carl Roth
Durapore membrane filter (0.22 µm)	Merck Millipore
Dynabeads® M-280 Streptavidin	Life Technologies
Electroporation cuvette 0.1 cm	Bio-Rad Laboratories
E-PAGE™ 96-well gels (6 %)	Life Technologies
FACS tubes, polystyrene	BD Biosciences
Filter Tips 10 µL, 20 µL, 100 µL, 200 µL, 300 µL, 1000 µL	Eppendorf
FluoroTrans® polyvinylidene difluoride transfer membrane(PVDF)	Pall Life Sciences
High Sensitivity DNA chip	Agilent Technologies
High speed falcon tubes 50 mL	Carl Roth
HisPur™ Ni-NTA spin columns	Thermo Fisher Scientific
HisTrap column 1 mL	GE Healthcare Life Science
Immunotube	Thermo Fisher Scientific
Lumi-Light Western Blotting Substrate	Sigma-Aldrich
MagicMedia™ (autoinduction medium)	Invitrogen
Microscope slides 76x26 mm	Carl Roth
Mounting medium	Dako
Nitrocellulose membrane (0.1 µm pore size)	GE Healthcare Life Science
NuPAGE® Novex 12 % Bis-Tris gels, 1.00 mm (10, 12 and 15 well)	Invitrogen
NuPAGE® Novex 4-12 % Bis-Tris gels, 1.00 mm (10, 12 and 15 well)	Invitrogen
PCR reaction tubes 0.2 mL, 8-tube-strip	Molecular BioProducts
Petri dish square, 245x245 mm	Sigma-Aldrich
Petri dishes 100x15 mm, 140x20 mm	Greiner bio-one
Pipette tips 10 µL, 200 µL, 1000 µL	Eppendorf
Reaction tubes 5 mL, 2 mL, 1.5 mL, 0.5 mL	Eppendorf
Reaction tubes 5 mL, 2 mL, 1.5 mL, 0.5 mL, protein LoBind	Eppendorf
Reagent reservoir 25 mL, 50 mL	Roth
Rod pipette 5 mL, 10 mL, 25 mL, 50 mL	Greiner bio-one
Serological pipettes 2 mL, 5 mL, 10 mL, 25 mL, 50 mL	Greiner bio-one
SOC medium	New England BioLabs
Streptividin sensor chip	GE Healthcare Life Science
SuperSignal™ West Dura Extended Duration Substrate	Thermo Fisher Scientific
Syringe 30 mL	BD Biosciences

Syringe filter 0.45 µm, sterile	Sartorius
Syringes 5 mL, 1 mL, 0.5 mL	BD Biosciences
Syringes, fixed needle, 0.3 mL, 0.5 mL, 1 mL	BD Biosciences
Tubes 15 mL, 50 mL	Greiner bio-one
Whatman paper	Carl Roth

2.1.17 Laboratory equipment

Agarose gel comb	Bio-Rad Laboratories
Agarose gel tray	Bio-Rad Laboratories
Agilent 1260 Infinity Quarternary LC system	Agilent Technologies
ÄKTA Pure	GE Healthcare Life Science
ÄKTAprime plus	GE Healthcare Life Science
Autoclave	SHP Steriltechnik AG
BD FACSCanto II	BD Biosciences
Biacore T-100 device	GE Healthcare Life Science
Bioanalyzer 2100	Agilent Technologies
Blotting Roller	Invitrogen
Branson Digital Sonifier 250	Branson
CASY TCC	OLS OMNI Life Science GmbH
Centrifugal flasks 1 L	Thermo Fisher Scientific
Centrifuge 4K15	Sigma
Centrifuge 5810R	Eppendorf
Centrifuge, SorvallTM LYNXTM Superspeed, Multitron pro	Thermo Fisher Scientific
ChemiDocMP	Bio-Rad Laboratories
CO_2 incubators HERAcell 150	Heraeus
Conical flask 250 mL	Duran
DynaMag™-2	Invitrogen
ELGA PURELAB Chorus 1	ELGA Veolia
E-PAGE™ System	Life Technologies
FACS Aria™	BD Biosciences
Impulse sealer Typ300	GEHO Pack Services
Incubation shaker Multitron Pro	INFORS HT
IVIS Spectrum System	PerkinElmer
Laminar flow bench HERAsafe 2020	Heraeus
LCMS Single Quad G6130B System	Heraeus
Microplate Reader Infinite M200 PRO	Tecan
MicroPulser Electroporation device	Bio-Rad Laboratories
Multichannel pipette	Eppendorf
Multistep pipette	Eppendorf
PCR thermocycler (3 blocks, 48-well microplates)	Biometra
peqSTAR 96X Universal (gradient PCR thermocycler)	Peqlab Brand
Pipette 10 µL, 200 µL, 1000 µL	VWR Eppendorf
Reciprocal shaker CAT RM5-80L	CAT
Reciprocal shaker MaxQ 2508 Modell 4363	Thermo Fisher Scientific
Resource RPC column 3 mL	GE Healthcare Life Science
Rotational vacuum concentrator, RVC 2-18-CD Plus	Christ

Sample loop, 5 mL	GE Healthcare
Scale	Mettler Toledo
Spectral photometer Ultrospec 2100	GE Healthcare
Spectrophotometer, Nanodrop2000c	Thermo Fisher Scientific
Superloop 10 mL	GE Healthcare
Table top centrifuge Fresco 21	Heraeus
Thermomixer 0.5 mL, 1.5 mL, 2 mL	Eppendorf
Trans-Blot® Turbo™ Blotting System	Bio-Rad Laboratories
TSKgelG3000SWXL	Tosoh Bioscience
UV transilluminator	NIPPON Genetics Europe
Vi-Cell XR	Beckman Coulter, Inc. Brea
Vortex-Genie	Scientific Industries
Water bath JULABO TW 8	JULABO
XCell SureLock Mini-Cell Electrophoresis chamber	Invitrogen
Zeiss Apotome microscope	Carl Zeiss

2.1.18 Software and database

Biacore T-100 Evaluation Software	GE Healthcare Life Science
ChemStation Software	Agilent Technologies
E-Editor 2.0	Thermo Fisher Scientific
FlowJo X	TreeStar
GraphPad Prism	GraphPad Software Inc.
ImageQuant™ TL 8.1	GE Healthcare
Living Image® software	PerkinElmer
PrimeView Evaluation Program	GE Healthcare Life Science
ProtParam	Bioinformatics Resource Portal
PyMol	Schrödinger
ZEN software	Carl Zeiss

2.2 Methods

2.2.1 Cell biological methods

2.2.1.1 Cell culture

Cell lines were handled under sterile conditions inside a laminar flow bench. All cell lines (chapter 2.1.1) were cultured under aseptic conditions in an incubator at 37 °C with 5 % CO_2 and >80 % humidity.
A tube containing cryopreserved cell bank was placed in a 37 °C bath to thaw cells until a small piece of ice was left. Cells were immediately transferred to 10 mL pre-warmed cell culture medium and centrifuged at 300 x g for 5 min. The supernatant was discarded and the cell pellet was re-suspended in the respective cell culture medium. Cell number and viability were determined as described in chapter 2.2.1.2 and seeded in a 75 cm^2 flask for cultivation.
For subcultivation of adherent cell lines medium was removed by aspiration, cells were washed with 1x Dulbecco´s phosphate-buffered saline (DPBS) and detached with 3 mL Accutase® solution for 3 – 5 min at 37 °C. Enzymatic reaction was stopped by adding 10 mL of cultivation medium and cell clusters were separated by pipetting up and down. Subsequently, cells were counted as described in chapter 2.2.1.2 and seeded with desired cell density (Table 2.1) into a new culture 175 cm^2 flask with respective amount of culture medium for further cultivation (chapter 2.1.2).

Table 2.1 Cultivation conditions of cell lines.

Parameters	CHO-K1	CHO-K1 transfectant	Flp-In™ wt	Flp-In™-CHO-FAP-α	U-87 MG
Cell type	adherent	adherent	adherent	adherent	adherent
Seeding density (subcultivation)	$1x10^4$ cells/cm^2	$1x10^4$ cells/cm^2	$1x10^4$ cells/cm^2	$1x10^4$ cells/cm^2	$2x10^4$ cells/cm^2
Subcultivation frequency	every 3rd day	every 3rd day	every 3rd day	every 3rd day	every 4th day

2.2.1.2 Cell counting

Cell number and viability were measured with automated cell counters Vi-Cell XR or CASY TCC.
The Vi-Cell XR cell optical counting system determines cell density and viability according to size, morphology, contrast via bright field and a trypan blue staining. For cell counting cell suspension was diluted 1:2 or 1:10 with DPBS in a total volume of 1 mL.
The CASY TCC automated cell counting device is based on electrical current exclusion and pulse area analysis. Cells travel in an electrolyte solution through a measuring pore into a capillary and are exposed to an electric field. Thereby, living cells generate greater pulse height compared to dead cells as they have intact membranes. For counting purposes, 10 µL cell suspensions was added to 10 mL CASYton solution in a CASY cup, which was then placed under the capillary for analysis.

2.2.1.3 Generation of stable cell lines

In this work stable cell lines were generated using the PiggyBac transposon system[138]. After cell harvesting (chapter 2.2.1.1), 2.4 x 10^5 CHO-K1 cells were seeded into a 6-well plate and grown overnight. A mixture of 8.4 µg PEI, 0.8 µg Piggy Bac transposase vector and 2 µg pPB transposon vector (pPB transposon_DSG-3 or pPB transposon_dsRed vector) was incubated at RT for 30 min and

added dropwise to the cells in transfection medium (DMEM/F12 + GlutaMAX™). Cells were then rested for around 8 h at 37 °C before changing transfection medium to growth medium. After three days of cultivation the medium was removed and replaced with fresh growth medium supplemented with 200 µg/mL hygromycin-B to induce the selection of successfully transfected cells.
A monoclonal cell from the pool of transfected cells was selected by FACS using the FACS Aria™ device. Single cells showing a high expression of dsRed were sorted into a 96-well, flat bottom plate. To generate a stable cell line expressing DSG-3, prior to sorting, cells were stained with an anti-DSG-3 antibody and anti-mouse IgG (chapter 2.1.7). Hereby, single cells with strong co-expression of GFP and DSG-3 were selected. After sorting, the plate was centrifuged at 300 x g for 5 min at RT to sediment the single cells. The plate was then incubated at 37 °C with 5 % CO_2 for cell outgrowth.
For cryopreservation, $1x10^7$ cells were re-suspended in 1 mL freezing buffer in a cryo tube and stored in an isopentane filled cryo freezing container at −80 °C overnight. Afterwards, the cryo tube was transferred to a liquid nitrogen tank (−196 °C) for long-term storage.

2.2.2 Molecular biological methods

2.2.2.1 DNA preparation

Isolation of plasmid DNA was performed according to manufacturer´s instruction using Wizard® *Plus* SV Minipreps DNA Purification System or plasmid DNA purification kit.

2.2.2.2 Restriction of vectors and DNA fragments

Vectors were digested with *Apa*I*, Bam*HI or *Kpn*I. Enzymes were applied with 10 U per 1 µg DNA and enzymatic reaction was conducted at 37 °C for 2 h. pET-32-LibEx vector was digested with *Apa*I and *Bam*HI. pET-32-LibEx_GS, pET-32-AX, pJuFoIII and pPDIII-1 vectors were restricted with *Bam*HI and *Kpn*I. Digested vectors were separated by gel electrophoresis (chapter 2.2.2.3) and DNA was subsequently purified via gel extraction (chapter 2.2.2.4). The 5´-ends of digested and purified vector backbone were then dephosphorylated with 1 U rSAP per 1 µg DNA at 37 °C for 1 h and 5 min at 65 °C for enzyme inactivation. As a final step, DNA was purified with the Wizard® SV Gel and PCR Clean-up system and eluted in an appropriate volume of H_2O. Prepared vectors were stored at −20 °C.
DNA fragments were restricted with FastDigest *Apa*I, FastDigest *Bam*HI or FastDigest *Kpn*I with 10 U per 1 µg DNA fragment. Incubation was performed at 37 °C for 1 h.

2.2.2.3 Agarose gel electrophoresis

DNA was separated according to its molecular size in an electric field within an agarose gel. For long DNA fragments gels with 1 % (w/v) agarose in TAE buffer were prepared and for short DNA fragments 2 % agarose gels were used. To this end, an appropriate amount of solid agarose was dissolved in TAE buffer by heating up inside a microwave at ~600 Watt. For detection of DNA fragments the agarose solution was supplemented with GelRed™ (Biotium, diluted 1:10000) and poured into a gel tray with a suitable well comb. 6x DNA loading dye was added to the DNA samples (final concentration 1x) and transferred together with an appropriate DNA ladder (50 bp DNA ladder or 1 kb DNA ladder) into the pockets of the gel. Depending on the gel size an optimal constant voltage of 100 V was adjusted and run for 40 – 90 min. Separated DNA bands were visualized with UV-light and documented with ChemiDoc MP. For preparative purposes, respective DNA bands were cut out of the gels on an UV transilluminator (chapter 2.2.2.4).

2.2.2.4 DNA purification from agarose gels

A piece of agarose gel containing the desired DNA fragment was cut with a sterile scalpel under UV light on a UV transilluminator. Isolation of DNA fragments out of an agarose gel was performed according to manufacturer´s protocol using QIAquick Gel Extraction Kit or Wizard® SV Gel and PCR Clean-up System.

2.2.2.5 Ligation of DNA fragments

DNA fragments were ligated enzymatically to the vector backbone using T4 DNA ligase. Ligation reactions were composed of 100 ng vector DNA and a fivefold molar excess of DNA fragments in a total volume of 20 µL, unless other described elsewhere. Enzymatic reactions were conducted at RT for 2 h, followed by a transformation of *E. coli* cells (chapter 2.2.2.6).

2.2.2.6 Transformation of *E. coli* SHuffle® T7 Express and E. *coli* SHuffle® T7 Express pBirAcm cells

The transformation was performed with competent *E. coli* SHuffle® T7 Express cells or *E. coli* SHuffle® T7 Express pBirAcm cells. The ligation sample was added to the bacterial cells and placed on ice for 30 min. The cells were then heat-shocked at 42 °C for 30 sec and incubated on ice for 5 min. For expansion of cells 950 µL pre-warmed SOC medium was added and the mixture was incubated at 37 °C for 1 h before plating them on a selective agar plate.

2.2.2.7 Generation of chemical competent *E. coli* SHuffle® T7 Express pBirAcm cells

E. coli SHuffle® T7 Express cells were transformed with pBirAcm plasmid according to the protocol as described in chapter 2.2.2.6. A single clone was grown in 5 mL $LB_{CarbChlora}$-medium culture at 37 °C and 200 rpm overnight. At the following day, a 50 mL culture containing LB-medium was inoculated with 500 µL of bacterial preculture. The culture was incubated at 37 °C at 150 rpm until an OD_{600} of 0.5 was reached. The cells were harvested by centrifugation at 4 °C and 4000 rpm for 10 min. Cell pellet was carefully taken up in 25 mL ice-cold 100 mM $CaCl_2$ and centrifuged (4 °C, 4000 rpm for 10 min). The supernatant was discarded, cell pellet re-suspended in 20 mL ice-cold 100 mM $CaCl_2$ and centrifuged again. As a final step of generating chemical competent cells, pellet was taken up again in 2 mL ice-cold 100 mM $CaCl_2$ and incubated for 30 min on ice. For long term storage at −80 °C, 300 µL glycerol was added to cell suspensions and aliquots of 200 µL were prepared.

2.2.2.8 Photometric determination of DNA concentration

DNA concentration and purity was determined by spectroscopic measurement at 260 nm absorption using NanoDrop2000c UV-Vis spectrophotometer. The spectrophotometer was blanked with 1.5 µL of the respective buffer and 1.5 µL of DNA solution was applied for quantification.

2.2.2.9 Sequencing

DNA sequencing was performed by a commercial provider (Eurofins Genomics, GATC or Microsynth AG). Plasmid DNA was prepared as described in chapter 2.2.2.1 and adjusted to a final concentration of 50 – 100 ng/µL in a volume of 15 µL. For sequencing, 2 µL of forward or reverse primer (10 pmol/µL) were added to purified plasmid DNA, respectively and send to Eurofins Genomics in a 1.5 mL reaction tube.

For high-throughput sequencing in 96-well plates, single *E. coli* clones were placed in 150 µL $LB_{CarbGlu}$-medium and incubated at 37 °C for 4 h. The plate together with a tube containing 200 µL of primer (10 pmol/µL) was shipped to Microsynth AG.
For high-throughput sequencing at GATC, a provided 96-well agar plate containing ampicillin was inoculated with 5 µL *E. coli* preculture from a 96-deepwell plate. The primer (10 pmol/µL) in a total volume of 100 µL was provided in a separate tube.

2.2.2.10 Colony polymerase chain reaction

Polymerase chain reaction (PCR) is a sensitive method to amplify DNA sequences with heat-stable polymerase enzyme and specific oligonucleotide primers. The reaction starts with a separation of dsDNA into two strands by a heat denaturation step. During the annealing step primers are able to bind their complementary DNA strand. This step is followed by the synthesis of the complementary DNA strand in 5´-3´direction via polymerase within the elongation step.
Colony PCR was performed with *E. coli* XL1-Blue cells to analyze enrichment of MC-FA-010 in spike-in selection experiments. *E. coli* cells were obtained from dilution plating of a cell-based panning round (chapter 2.2.5.1). Single clones were picked from selective agar plates with a sterile tip and dipped into a well of a 48-well PCR plate that contained 25 µL of PCR reaction mix (Table 2.2).

Table 2.2 Reagents and volumes for colony PCR.

Reagent	Volume (µL)
10x Taq buffer (NH_{4+})	2.5
$MgCl_2$ (25 mM)	1.5
dNTPs (10 mM)	0.5
MC-FA-010-up (10 µM)	0.5
M13 uni (-43) lo (10 µM)	0.5
Taq polymerase (1 U/µL)	0.25
H_2O	ad 25

A pJuFoVIII vector with a gene coding for MC-FA-010 was used as positive control, while pJuFoVIII vector and H_20 served as negative controls. PCR reaction program was performed as stated in Table 2.3 in a PCR thermocycler.

Table 2.3 PCR reaction program for colony PCR.

Duration	Temperature	Step	
2 min	94 °C	First denaturation	
30 sec	94 °C	Denaturation	30x
30 sec	53 °C	Annealing	
1 min	72 °C	Synthesis	
2 min	72 °C	Final elongation	
∞	4 °C	Hold	

After PCR reaction, 4 µL of 6x loading dye was added to each well. To analyze PCR products 10 µL of each reaction sample was electrophoretically separated on a 2 % agarose gel (chapter 2.2.2.3).

2.2.2.11 Gradient PCR

The optimal PCR condition for amplification of MCopt 2.0, MCopt 2.1 and MCopt 3.0 DNA fragments was assessed using gradient PCR and with addition of different supplements. The general composition of a PCR reaction is shown in Table 2.4 with addition of either 2.5 %, 5 %, 7.5 % or 10 % DMSO or 0.4 M, 0.8 M, 1.2 M or 1.6 M betaine. PCR reactions were performed in a gradient PCR cycler according to Table 2.5 with an annealing temperature gradient ranging from 53 °C to 64 °C.

Table 2.4 Reagents and volumes for gradient PCR. DMSO was added to a final concentration of 2.5 %, 5 %, 7.5 % or 10 % and betaine to 0.4 M, 0.8 M, 1.2 M or 1.6 M. Primer sets for amplification of MCopt 2.0, MCopt 2.1 and MCopt 3.0 DNA fragment pools are given in chapter 2.1.11.

Reagent	Volume (µL)
10x Taq buffer (NH_{4+})	2
$MgCl_2$ (25 mM)	1.2
dNTPs (10 mM)	0.4
Forward primer (10 µM)	0.4
Reverse primer (10 µM)	0.4
Taq polymerase (1 U/µL)	0.5
Supplement	DMSO or betaine
Template	50 ng DNA fragments
H_2O	ad 20

Table 2.5 PCR reaction program for gradient PCR. In a gradient PCR, annealing temperatures of 53 °C, 54.6 °C, 56.1 °C, 57.7 °C, 59.3 °C, 60.9 °C, 62.4 °C or 64 °C were applied.

Duration	Temperature	Step	
5 min	95 °C	First denaturation	
30 sec	95 °C	Denaturation	30x
30 sec	53 °C – 64 °C	Annealing (gradient)	
1 min	72 °C	Synthesis	
5 min	72 °C	Final elongation	
∞	4 °C	Hold	

For analysis of amplified PCR products, samples were mixed with 4 µL 6x loading dye and separated on a 2 % agarose gel (chapter 2.2.2.3). Based on the visual evaluation, single PCR samples were chosen for a control of fragment size and quality. Therefore, PCR products were purified via AMPure XP Beads according to manufacturer´s instruction and inspected on a High Sensitivity DNA chip using a Bioanalyzer 2100 chip-based capillary electrophoresis device, which was performed by colleagues from the Medical Genomics-NGS unit at TRON gGmbH.

2.2.2.12 Splicing by overlap extension PCR

The MCopt 3.0 DNA fragment pool for phage library construction was generated using splicing by overlap extension (SOE) PCR as depicted in Figure 2.4. Here, the MCopt 2.0 and MCopt 2.1 DNA fragments assemble through their complementary overlapping regions. By means of a PCR with a primer pair, which binds on each DNA fragment, both parts were genetically combined. This resulted in a novel amplified DNA fragment pool.

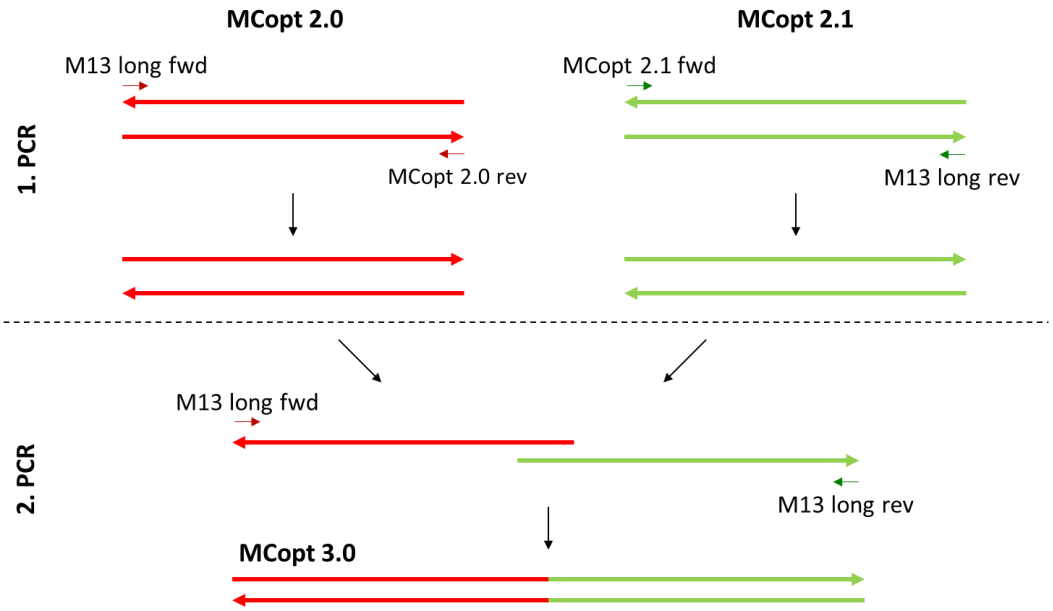


Figure 2.4 Experimental outline of MCopt 3.0 DNA fragment amplification via SOE PCR.

In a first PCR, fully synthetic MCopt 2.0 DNA fragments were amplified using M13 fwd long and MCopt 2.0 rev primers, while for synthetic MCopt 2.1 DNA fragments MCopt 2.1 fwd and M13 rev long primers were applied (chapter 2.1.11). The reaction mix as well as the applied PCR program is shown in Table 2.6 and Table 2.7. After fragment amplification in a PCR thermocycler, PCR products were gel-electrophoretically separated and respective amplicons were purified (chapter 2.2.2.3, chapter 2.2.2.4).

In the second PCR, the purified MCopt 2.0 and MCopt 2.1 DNA fragments served as templates for the construction of MCopt 3.0 DNA fragments with the primer pair M13 fwd long and M13 rev long (chapter 2.1.11). Resulting PCR products were again separated vial gel-electrophoresis and purified (chapter 2.2.2.3, chapter 2.2.2.4). Obtained MCopt 3.0 DNA fragments were used for phage library generation as described in chapter 2.2.2.14.

Table 2.6 PCR reagents and volumes for MCopt 2.0, MCopt 2.1 and MCopt 3.0 DNA fragment amplification.

	Volume (µL)		
Reagent	**MCopt 2.0 fragment**	**MCopt 2.1 fragment**	**MCopt 3.0 fragment**
10x Taq buffer (NH_{4+})	20	20	20
$MgCl_2$ (25 mM)	12	12	12
dNTPs (10 mM)	4	4	4
Forward primer (10 µM)	4	4	4
Reverse primer (10 µM)	4	4	4
Taq polymerase (1 U/µL)	5	5	5
Betaine (0.8 M)	0	32	0
DMSO (5 %)	0	0	10
DMSO (10 %)	20	0	0
Template	50 ng	50 ng	42 ng of each template
H_2O	ad 200	ad 200	ad 200

Table 2.7 PCR reaction program for MCopt 2.0, MCopt 2.1 and MCopt 3.0 DNA fragment amplification.

Duration	Temperature	Step	
5 min	95 °C	First denaturation	
30 sec	95 °C	Denaturation	30x
30 sec	53 °C *	Annealing	
1 min	72 °C	Synthesis	
5 min	72 °C	Final elongation	
∞	4 °C	Hold	

* 64 °C annealing temperature for MCopt 2.1 and MCopt 3.0 DNA fragment amplification.

2.2.2.13 Cloning of cystine-knot miniprotein sequence into expression vector

Plasmids of pJuFo-VIII-MCopt 1.0, pJuFoIII-MCopt 2.0/2.1/3.0 were prepared from 5 mL overnight *E. coli* culture (chapter 2.2.2.1). For cloning of MCopt 1.0 inserts into pET-32-LibEx vector, a *Bam*HI restriction site was introduced via PCR using BamHI-Lib-up primer, whereas Bsp120I-MC-lo served as reverse primer (chapter 2.1.11). A standard reaction mix is shown in Table 2.6 (similar to MCopt 2.1 DNA fragment pool amplification), but with 100 ng template DNA. The applied program is presented in Table 2.3. After amplification, PCR products were separated via gel electrophoresis (chapter 2.2.2.3), purified and eluted in 60 µL H_2O (chapter 2.2.2.4). Amplified MCopt 1.0 inserts were digested with *Apa*I and *Bam*HI in a volume of 200 µL at 37 °C for 1 h. Restricted MCopt 1.0 inserts were purified via Wizard® SV Gel and PCR Clean-up System for following ligation into prepared pET-32-LibEx vector (chapter 2.2.2.5).
For MCopt 2.0/2.1/3.0 insert cloning, 5 µg plasmid were directly digested with *Bam*HI and *Kpn*I in a total volume of 200 µL at 37 °C for 2.5 h (chapter 2.2.2.2). Inserts were purified after gel electrophoretic separation (chapter 2.2.2.3, chapter 2.2.2.4) and used for ligation into pET-32-LibEx_GS or pET-32-AX2 vector (chapter 2.2.2.5). Cloned expression vectors were introduced into *E. coli* SHuffle® T7 Express cells as described in chapter 2.2.2.6.

2.2.2.14 Generation of cystine-knot miniprotein phage libraries

Combinatorial phage libraries were constructed based on an open chain variant of the trypsin inhibitor II from *Momordica cochinchinensis* (oMCoTI-II). Three libraries differed with respect to the randomization scheme. MCopt 2.0 was randomized in loop 1, MCopt 2.1 in loop 5 and MCopt 3.0 in both loops with ten amino acids.
DNA fragments encoding for MCopt 2.0 or MCopt 2.1 sequences were directly used after manufacturing by GeneArt™ (chapter 2.1.10), while MCopt 3.0 DNA fragments were obtained using SOE-PCR (chapter 2.2.2.12).
In the first step, 5 µg of each DNA fragment pool were digested with *Bgl*I in a total volume of 200 µL for 4 h at 37 °C, gel-electrophoretically separated (chapter 2.2.2.3), purified (chapter 2.2.2.4), followed by a second digestion of remaining DNA with *Bam*HI and *Kpn*I in a total volume of 150 µL for 4 h at 37 °C and subsequent purification using Wizard® SV Gel and PCR Clean-up System. For ligation, 7200 ng pJuFoIII vector backbone was combined with a five-fold molar excess of purified PCR fragments and 30 µL T4 DNA ligase (40 U/µL) in a total volume of 480 µL and incubated at 16 °C overnight. Ligation product was purified using DNA purification kit and eluted in 60 µL H_2O. Subsequently, ligation product was introduced into electrocompetent *E. coli* SS320 cells. In total 12 electroporation reactions were performed with 5 µL ligation products and 50 µL *E. coli* cells respectively in a 0.1 cm cuvette using the program Ec1 with 1.8 kV of MicroPulser electroporation

device. Immediately after the electroporation pulse, 1 mL of pre-warmed SOC medium was added and the cuvette was rinsed three times with 1 mL SOC medium. Cell suspensions were combined and incubated at 37 °C for 1 h at 150 rpm. In order to calculate the library size a serial dilution with 4.8 µL of the pooled cell suspension was made and plated on small agar plates containing 100 µg/mL carbenicillin and 0.4 % (w/v) glucose. Remaining 4795.2 µL suspension was plated on large square agar selective plates. All plates were incubated overnight at 30 °C. At the following day, colonies were scrapped off the large plates by adding 4 mL LB-medium to each plate and stored in the presence of 10 % DMSO at −80 °C. Library size and diversity was calculated based on colony forming cells from the dilution plates and sequencing of individual clones.
In order to eliminate double transformed *E. coli* SS320 cells, infection of *E. coli* XL1-Blue cells with library phages was performed with a multiplicity of infection (MOI) of 0.1. To this end, an *E. coli* XL1-Blue culture (2.1 L) at an optical density (OD) of 0.5 was infected with 5 x 10^{10} produced MCopt 2.0, MCopt 2.1 or MCopt 3.0 library phages (chapter 2.2.5.4). Cells were rested for 30 min at 37 °C and agitated for 30 min at 37 °C with 120 rpm. Afterwards, cell suspensions were harvested by centrifugation at 4500 x g for 10 min. The pellet was taken up in 1.25 L $LB_{CarbGlu}$-medium for incubation at 37 °C with 120 rpm overnight. Bacterial cells were harvested by centrifugation (4500 x g for 10 min) and stored with 10 % DMSO in small aliquots at −80 °C for further analysis and phage production. Library size was calculated by plating serial dilution of bacterial suspension on respective agar plates.

2.2.2.15 Generation of MCopt 2.0 phage sub-library

A MCopt 2.0 phage sub-library was constructed in a second phagemid vector (pPDIII-1). Library generation was performed essentially as described in chapter 2.2.2.14, but in a smaller scale. After synthetic MCopt 2.0 DNA fragment pool restriction, ligation of 60 ng MCopt 2.0 DNA fragments and 400 ng restricted pPDIII-1 vector was conducted with 6 µL T4 DNA ligase in a total volume of 80 µL and incubated overnight at 16 °C. Thereafter, ligation product was purified via Wizard® SV Gel and PCR Clean-up system and eluted in 30 µL H_2O. Three electroporation reactions were performed each with 10 µL ligation product and 100 µL electrocompetent *E. coli* XL1-Blue competent cells in a 0.1 cm cuvette using MicroPulser electroporation device. Cells were re-suspended in 1 mL of pre-warmed SOC medium right after electrical pulse and the cuvette was rinsed three further times with 1 mL SOC medium. Cell suspension (12 mL) was rested at 37 °C for 1 h at 150 rpm, 1.2 µL was taken for dilution plating and remaining suspension was plated on agar plates containing 100 µg/mL carbenicillin and 0.4 % (w/v) glucose. After overnight incubation at 30 °C, colonies were scrapped off by dispensing 4 mL LB-medium to each plate and stored in the presence of 10 % DMSO at −80 °C. Colonies from the dilution plates were counted and individual clones were sequenced for library size as well as quality determination.

2.2.3 Microbiological methods

2.2.3.1 High-throughput production of cystine-knot miniprotein fusions

Recombinant protein production was carried out using *E. coli* SHuffle® T7 Express cells carrying either pET-32-LibEx or pET-32-LibEx_GS vector encoding for the respective cystine-knot miniprotein sequence. For a 96-well production of cystine-knot miniprotein fusions, single *E. coli* clones were expanded overnight at 37 °C in 1 mL $LB_{CarbGlu}$-medium in a 96-deepwell plate. The following day, 1 mL of MagicMedia™ (autoinduction medium) supplemented with 100 µL/mL carbenicillin were

inoculated with 10 µL of the respective bacterial preculture and incubated at 30 °C for 18 – 24 h at 200 rpm.
Cells were harvested by centrifugation at 4000 rpm for 30 min. Cell pellets were re-suspended in 250 µL lysis buffer and frozen at −80 °C for 1 h. The suspension was thawed and purified by heat-incubation at 80 °C for 10 min. After a final centrifugation for the removal of cell debris, the bacterial supernatants containing the cystine-knot miniprotein fusions were transferred to a fresh plate and stored at 4 °C until use.

2.2.3.2 *In vivo* biotinylation of cystine-knot miniprotein fusions in *E. coli*

In vivo biotinylation of cystine-knot miniprotein fusion was performed using chemically competent *E. coli* SHuffle® T7 Express pBirAcm cells bearing pET-32-AX2 vectors encoding for the respective cystine-knot miniprotein sequence. Single clones were grown in 1 mL $LB_{CarbChlora}$-medium overnight at 37 °C and 180 rpm in a 96-deepwell plate. At the next day, 10 µL of the bacterial suspension was transferred to a fresh 96-deepwell plate containing 1 mL MagicMedia™ (autoinduction medium with 100 µL/mL carbenicillin and 50 µM D-biotin) per well and incubated at 30 °C for 18 – 24 h at 200 rpm. Further processing was done as described in chapter 2.2.3.1.

2.2.3.3 Large-scale production of cystine-knot miniprotein fusion

Recombinant protein production was done using *E. coli* SHuffle® T7 Express cells carrying pET-32-LibEx or pET-32-LibEx_GS vectors encoding for respective cystine knot miniprotein sequence. For this, a 50 mL $LB_{CarbGlu}$-medium preculture was inoculated with a single clone from a selective agar plate and incubated at 37 °C overnight with 180 rpm. The day after, the main culture with 750 mL $LB_{CarbGlu}$-medium was inoculated to an OD_{600} of 0.01. The culture was incubated at 30 °C with 120 rpm until an OD_{600} of approximately 0.7 was reached. Protein production was then induced by adding 1 mM IPTG and the culture was incubated further at 25 °C with 120 rpm overnight. *E. coli* cells were harvested, re-suspended in 10 mL IMAC A buffer, lysed by sonification and heated to 80 °C for 10 min. After centrifugation of cell debris (15000 x g for 30 min, 4 °C), the supernatant was purified by immobilized metal affinity chromatography (IMAC) (chapter 2.2.4.1). The cystine-knot miniprotein fusion containing fractions were collected and dialyzed against thrombin cleavage buffer as described in chapter 2.2.4.2.
Cystine-knot miniprotein fusions were either directly used for enzyme-linked immunosorbent assay (ELISA) or flow cytometry assays or processed further in case that the untagged miniprotein was required, e.g. for surface plasmon resonance (SPR) spectroscopy. For this purpose cystine-knot miniprotein fusions were cleaved with thrombin protease (chapter 2.2.4.3).
Separation of cleaved protein fragments was performed by reverse phase chromatography (RPC) as described in chapter 2.2.4.4. Respective fractions containing cystine-knot miniprotein were lyophilized in a rotational vacuum concentrator as a final step. Amount of cystine-knot miniprotein was determined by weighing and the proteins were stored in lyophilized form at −20 °C. Proteins were resolved in H_2O or DPBS depending on the subsequent application. Cystine-knot miniprotein identity was verified by mass spectrometry with a LCMS Single Quad G6130B system using a standard electrospray ionization protocol and performed by colleagues from the Serodiscovery/Peptide unit at TRON gGmbH.

2.2.3.4 *E. coli* growth curves

A 50 mL LB-medium culture (supplemented with 100 µg/mL carbenicillin) was inoculated to an OD_{600} of 0.01 with *E. coli* preculture carrying respective vector and incubated at 37 °C and 180 rpm. To analyze bacterial density over time, the OD was determined every 30 min via absorption measurement at 600 nm (chapter 2.2.3.5). After the culture reached an OD_{600} of 0.5, induction of protein production was initiated by adding 50 µL 1 M IPTG. Bacterial growth was further observed every 30 min.

2.2.3.5 Determination of bacterial cell density

OD of bacterial culture was determined by absorption measurement at 600 nm using NanoDrop2000c UV-Vis spectrophotometer. An OD_{600} of 1 represents a concentration of 8×10^8 cells/mL. The spectrophotometer was blanked with 1 mL of respective bacterial growth medium in a cuvette and then 1 mL bacterial suspension was measured pure or diluted with medium.

2.2.4 Biochemical methods

2.2.4.1 Immobilized metal affinity chromatography

IMAC was used to purify cystine-knot miniprotein fusions from a heterogeneous protein solution. This method is based on the binding of His-tagged proteins to Ni^{2+}-NTA loaded column resin, whereas other proteins straight pass through. Retained proteins can subsequently be eluted upon increasing imidazole concentration.
Protein solution was filled into a 50 mL super loop and connected to the injection valve of an ÄKTAprime™ plus system. 1 mL HisTrap column was used and the program “His Tag Purification His Trap” was applied with a constant flow rate of 1 mL/min. The program started with a system equilibration with IMAC A as running buffer and followed by sample injection with a flow rate of 1 mL/min. Subsequently, the column was washed with 20-fold column volumes. Protein elution was done using a gradient from 0 − 100 % IMAC B buffer in 20-fold column volumes. Eluted proteins were collected in 1 mL fraction size and chromatograms were evaluated in PrimeView Evaluation Program. Protein containing fractions were analyzed on SDS-PAGE and pooled for further assays.

2.2.4.2 Dialysis

Protein solutions were filled into semipermeable dialysis tubes with an exclusion limit of approximately 14 kDa. Depending on the solution volumes, a ~1000-fold higher volume of thrombin cleavage buffer or PBS was used for dialysis at 4 °C overnight.

2.2.4.3 Thrombin cleavage of cystine-knot miniprotein fusion

Separation of cystine-knot miniproteins from its fusion protein was achieved using thrombin protease cleavage. Per 1 mg protein 0.5 U of thrombin was added to the purified cystine-knot miniprotein fusion solution and enzymatic reaction was incubated at 37 °C overnight.

2.2.4.4 Reverse phase chromatography

RPC was the method of choice to separate fusion protein from cystine-knot miniproteins after thrombin cleavage. For this chromatography technique, amino acid residues interact varyingly strong

with the hydrophobic column material, which usually contains alkyl chains covalently coupled to silica gel.
Agilent 1260 Infinity Quarternary LC system and a 3 mL Resource RPC column were used to isolate monomeric cystine-knot miniproteins. As running buffer 2 % acetonitrile in H_2O supplemented with 0.05 % TFA and as elution buffer 80 % acetonitrile in H_2O containing 0.05 % TFA was applied. The chromatography system was washed with H_2O and equilibrated with running buffer. Protein solution was injected in a 500 µL of 5 mL sample loop and a constant flow rate of 1 mL/min was set. Running buffer was substituted by running a gradient to 100 % elution buffer. Protein solution in a constant flow system passes absorption measurement at 220 nm and 280 nm, prior to collection of 250 µL fractions in a 96-deepwell plate. Fractions containing the protein of interests were identified in the chromatogram using ChemStation software.

2.2.4.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Proteins were separated according to their electrophoretic mobility and molecular mass by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Protein samples were loaded with 4x reducing SDS-loading buffer or 4x non-reducing SDS-loading buffer after denaturating via heating at 98 °C for 10 min. Depending on the protein size, protein samples and 5 µL protein standard (Precision Plus Protein™ Dual Color Standard or Precision Plus Protein™ Dual Xtra Prestained Standard) were transferred into pockets of NuPAGE Novex Bis-Tris 4-12 % 1 mm gels or NuPAGE Novex Bis-Tris 12 % 1 mm gels. Gels were run for 45 min with a constant voltage of 200 V. Subsequently, gels were stained in coomassie stain solution until the whole gel turned dark blue. Afterwards, gels were destained in coomassie destain solution to remove background signals and photographically documented with a ChemiDoc MP.

2.2.4.6 E-PAGE™ high-throughput gel system

The E-PAGE™ high-throughput gel system allows the analysis of 96 protein samples in parallel according to the classical SDS-PAGE principle. Protein samples were supplemented with 5 µL E-PAGE™ loading buffer and 2 µL DTT and heated at 70 °C for 10 min. The pockets of an E-PAGE™ 96-well gel (6 %) were loaded with 10 µL H_2O before adding 13.5 µL of prepared protein sample. As a protein reference Trx-MC-Myc-010 with an amount of 3 µg, 10 µg and 20 µg was additionally applied. At the end of each row 3 µL of a protein standard (Precision Plus Protein™ Dual Color Standard) was loaded. The gel was placed into the E-Holder™ platform and run for 16 min with a voltage according to manufacturer´s instruction. Proteins were visualized by coomassie staining as described in chapter 2.2.4.5. E-PAGE™ was documented photographically with a ChemiDoc MP and lanes sorted using the E-Editor program. The protein bands were densitometrically quantified using the ImageQuant™ TL 8.1 software and the concentration was determined by using a calibration standard resulting from known amount of Trx-MC-Myc-010.

2.2.4.7 Photometric determination of protein concentration

Spectrophotometric determination of protein concentration relies on the distinct absorbance of proteins at 280 nm. This is mainly mediated by the aromatic amino acids tyrosine and tryptophan as well as the disulfide bonds of the protein, which results in a molar absorption coefficient of the protein. The protein concentration can be calculated based on the molar absorption coefficient and according to Lambert-Beer law[139]:

$$A = \varepsilon \times d \times c$$

A = absorbance; ε = molar absorption coefficient (protein specific; L/mol x cm); d = utilized cuvette thickness (cm); c = protein concentration (mol/L)

Protein concentration and purity were determined by spectroscopic measurement at 280 nm and 260 nm using NanoDrop2000c UV-Vis spectrophotometer. The spectrophotometer was blanked with 1.5 µL of respective protein buffer and 1.5 µL of protein solution was applied.

2.2.4.8 Alanine scanning mutagenesis of selected MC-FN-010

In order to identify the residues within MC-FN-010 that contribute to EDB-binding an alanine scanning mutagenesis was performed. This method includes a systemic substitution of amino acids against alanine at defined sequence positions and subsequent binding analysis of the generated mutants. For generation of alanine scanning MC-FN-010 derivatives, respective DNA fragments (chapter 2.1.10) were cloned into pET-32-LibEx expression vector using unique *Bam*HI and *Kpn*I restriction sites and introduced into *E. coli* SHuffle®T7 Express competent cells (chapter 2.2.2.2, chapter 2.2.2.5, chapter 2.2.2.6). All mutations were verified by DNA sequencing (chapter 2.2.2.9). The alanine scan mutagenesis variants were expressed in 24-deepwell plates using 5 mL of selective autoinduction medium. Cystine-knot miniprotein fusion production and purification was performed as described in chapter 2.2.3.1 for the 96-well format, but included a further purification step of the supernatant using HisPur™ Ni-NTA spin columns. Binding ability and specificity of cystine-knot miniprotein fusions to target and off-target protein was carried out with an antibody-based ELISA assay as described in chapter 2.2.6.2.

2.2.4.9 Surface plasmon resonance spectroscopy

Binding kinetics of monomeric and trimeric cystine-knot miniprotein ligands to its target protein was determined using SPR in a Biacore T-100 device. For this, PBS-T was used as running buffer and the biotinylated FN-67B89 protein (200 – 300 µg/mL) was captured by binding to a flow cell of a SA sensor chip. To analyze monomeric ligands an immobilized target density of maximum 750 response units (RU) was applied and for trimeric variants a RU of maximum 400 was aimed for. Binding analysis of monomeric ligands was performed using a multi-cycle kinetic method with concentrations ranging from 50 – 4000 nM. A cycle started with an association period of 90 sec, followed by a dissociation period of 420 sec and a final regeneration step. Kinetic measurement was conducted applying a flow rate of 20 µL/min. Trimeric variants were analyzed under the same association and dissociation conditions, but using the single-cycle kinetic measurement mode in a constant flow of 30 µL/mL. In this case the analyte concentration was between 1.25 – 10 nM. Binding kinetics and steady state analysis were calculated using a global kinetic fit model (1:1 Langmuir, Biacore T-100 Evaluation Software).

2.2.5 Phage experimental methods

2.2.5.1 Cell-based panning

CHO-K1-Mock (negative cell line), CHO-K1-DSG-3 and CHO-K1-FAP-α (target expressing cell lines) cells were prepared as described in chapter 2.2.1.1. Phage library particles were diluted to 1×10^{13} or 2.4×10^{13} (first round) and 1×10^{12} or 2.4×10^{12} (second and third round) phages in 2 mL 2 % milk powder in PBS and 2 µL of phage suspension was removed for phage input titration (chapter 2.2.5.3).

For a panning round 1 x 10^7 CHO-K1-Mock cells were incubated with the actual phage suspension of 1998 µL for 2 h at 4 °C for negative selection of phages. Cells were discarded after centrifugation at 1500 x g for 1 min and remaining 1996 µL supernatant (2 µL removed for phage titration) was added to 1 x 10^7 CHO-K1-FAP-α or CHO-K1-DSG-3 cells for 2 h at 4 °C. Subsequently, unbound phages were washed off by rinsing five times with TBS-T and centrifugation at 1500 x g for 1 min. In order to elute bound phages, a pH-shift elution was performed by adding 200 µL 100 mM glycine (pH 2) to the washed cells. The suspension was incubated for 10 min at 300 rpm, centrifuged at 1500 x g for 1 min and the supernatant was transferred to a fresh tube containing 37.5 µL 1 M Tris/HCl (pH 9) for neutralization. 2.37 µL of elution solution was removed for phage titration (chapter 2.2.5.3).
Exponentially growing *E. coli* XL1-Blue cells at an OD_{600} of 0.5 (1762.5 µL) were infected with eluted phages. Therefore, *E. coli* cells were incubated with the phages for 30 min at 37 °C without agitation and further 30 min with agitation at 150 rpm. Afterwards, infected cells were plated onto two large agar plates (supplemented with 100 µg/mL carbenicillin and 0.4 % (w/v) glucose) and incubated at 37 °C overnight. At the following day, cells were scraped from the plates by dispensing 4 mL LB-medium. Resulting cell suspension was used for phage production as described in chapter 2.2.5.4. Produced phage particles were applied for subsequent panning round according to the same experimental procedure as described.

2.2.5.2 Panning against soluble target proteins

Panning against soluble FN-B protein was used as model system to compare screening outcome of pJuFoIII-MCopt 2.0 library and pPDIII-1-MCopt 2.0 sub-library. Both libraries were screened in three successive panning rounds, starting with a tube-based panning and two successive panning rounds with streptavidin-coated (SA) magnetic beads.
For the first round of tube-based panning, 1.2 mL 2 % milk powder in Na_2CO_3 (pH 9.4) were added in one immunotube (negative selection) and 1.2 mL Na_2CO_3 (pH 9.4) with 100 µg FN-B was added to a second immunotube for an incubation overnight at 4 °C and 30 rpm. Afterwards, the solutions were discarded and both immunotubes were blocked with 1.2 mL 2 % milk powder in TBS for 2 h at RT and 30 rpm. The uncoated immunotube was washed twice with TBS-T at 500 rpm for 3 min respectively and then 1 x 10^{13} phages in 1.2 mL 2 % milk powder in TBS was added to the tube. Incubation was conducted for 30 min at RT, while target coated immunotube was blocked for additional 30 min and washed twice with TBS-T. The phage solution from the negative selection was transferred to the target coated immunotube (1 µL removed for phage titration (chapter 2.2.5.3) and incubated for 1 h at RT and 30 rpm. Afterwards, the target coated immunotube was washed six times with TBS-T and two times with TBS for 3 min at 80 rpm, respectively. For phage elution, 950 µL 100 mM triethylamine (TEA) was added to the tube for 6 min at 80 rpm and afterwards transferred to a fresh tube containing 950 µL 1 M Tris/HCl (pH 7) for neutralization. A second elution was performed with 950 µL 100 mM glycine (pH 2) for 10 min at RT and 80 rpm. Subsequently, the supernatant was added to the first elution solution (2850 µL) and 2.85 µL was removed for phage titration (chapter 2.2.5.3). The elution solution was used to infect exponentially growing *E. coli* XL1-Blue cells (27.2 mL culture). The culture was incubated for 30 min at 37 °C without agitation and further 30 min with agitation at 150 rpm. Afterwards, infected cells were plated onto two large agar plates containing 100 µg/mL carbenicillin and 0.4 % (w/v) glucose and incubated at 37 °C overnight. At the next day, cells were scraped off the plates by adding 4 mL LB-medium. The cell suspension was then used for phage particle production (chapter 2.2.5.4) and served as input phage pool for the next panning round.

For the second or third panning round 2 x 50 µL Dynabeads® M-280 SA were transferred to a 2 mL tube each and washed with 1 mL TBS-T. To the first tube 100 µg biotinylated FN-B in 200 µL TBS was added, whereas 200 µL TBS without target (for negative selection of phages) was added to the second tube and beads were incubated on the rolling mixer for 20 min at 30 rpm. Tubes were placed into a DynaMag™-2 to recover beads and the beads were washed twice with TBS-T. For blocking, beads were taken up in 2 % milk powder in TBS for 1 h at 4 °C and 30 rpm. Target coated SA beads remained meanwhile in blocking solution, whereas 1 x 10^{13} (second and third round) phages were added to uncoated SA beads in 1 mL 2 % milk powder in TBS and incubated for 30 min at RT, 30 rpm (negative selection). Afterwards, blocking solution of target coated SA beads was removed, the beads were washed twice with TBS-T and re-suspended in phage supernatant (1 µL removed for phage titration chapter 2.2.5.3) resulting from the negative selection. After incubation for 1 h at RT and 30 rpm, unbound phages were removed by washing the beads six times with TBS-T and twice with TBS. Bound phages were eluted by adding 50 µL 100 mM TEA to the beads for 6 min at 2000 rpm and neutralized by transferring the supernatant to a fresh tube with 100 µL 1 M Tris/HCl (pH 7). A further elution was performed with 50 µL 100 mM glycine (pH 2) for 10 min at 2000 rpm. Both eluted phage solutions (200 µL) were combined and 2 µL of the solution were removed for phage titration (chapter 2.2.5.3), while remaining solution was used for infection of 1.8 mL *E. coli* XL1-Blue cell culture. Plating and phage rescue was done as described above.

2.2.5.3 Phage titration

Phage titration was performed to monitor phage titer within a panning round using dilution plating. During a cell-based panning phage suspension was removed for dilution plating before and after negative selection against CHO-K1-Mock cells as well as from the elution solution, which was referred to input, post-input and output (chapter 2.2.5.1). Within a panning against soluble target protein, phage suspension after negative selection in immunotube or on magnetic beads was used for input titration and suspensions of eluted phage solution for output titration (chapter 2.2.5.2). Serial dilution of phage suspension was performed with LB-medium. For infection, 10 µL of each phage dilution was added to 90 µL exponentially growing *E. coli* XL1-Blue cells. Cells were incubated at 37 °C for 30 min without agitation and thereafter 30 min with agitation at 150 rpm. Subsequently, cell suspensions were plated on agar plates containing 100 µg/mL carbenicillin and 0.4 % (w/v) glucose and incubated at 37 °C overnight. At the following day, colony forming units (CFU) were counted to determine the number of infectious phages.

2.2.5.4 Production of recombinant phages

For production of recombinant phages, 50 mL $LB_{CarbGlu}$-medium was inoculated with the respective *E. coli* suspension to an OD_{600} of 0.2 and incubated at 37 °C with 250 rpm until an OD_{600} of 0.5 was reached. Then 5 x 10^{11} VCSM13 helper phages were added for phage production from the phagemid vector and incubated at 37 °C for 30 min without agitation, followed by 30 min with agitation at 150 rpm. Cell suspensions were harvested by centrifugation at 4500 x g for 10 min at RT and supernatant was discarded. Cell pellet was re-suspended in 50 mL $LB_{CarbKanaIPTG}$-medium for induction of protein production and incubated at 30 °C overnight with 250 rpm.
The following day, cells were centrifuged at 4500 x g for 15 min, 4 °C. Phages were precipitated by adding 10 mL PEG/NaCl solution to 40 mL supernatant and incubation on ice for 30 min. After a centrifugation at 15000 x g for 20 min at 4 °C, the supernatant was removed and precipitated phages were dissolved in 800 µL 10 mM Tris/HCl (pH 8.0) for a following centrifugation at 15000 x g for

10 min at 4 °C. The solution was transferred to a fresh tube containing 400 µL PEG/NaCl for a second precipitation on ice for 20 min, followed by a further centrifugation (15000 x g for 10 min at 4 °C). Phage pellet was taken up in 800 µL 10 mM Tris/HCl (pH 8.0), heated at 65 °C for 15 min as a last purification step and finally centrifuged (15000 x g for 10 min at RT). The phage particle concentration of the supernatant was determined photometrically using the dual wavelength modus A_{269}-A_{320} on a spectral photometer. The concentration was calculated based on the nucleotide content and molar extinction coefficient of M13 phages[30].

$$\text{Phage particle/mL} = \frac{(A_{269} - A_{320})\ x\ (6x10^{16})}{6407\ (nucleotde\ of\ M13\ genome)}$$

Phage particle suspension was diluted 1/10 volume of 75% glycerol for storage at −80 °C.

2.2.6 Immunological methods

2.2.6.1 Whole-cell enzyme-linked immunosorbent assay

Functional display of MC-FA-010 on phage particles was determined with a whole-cell ELISA. CHO-K1-FAP-α and CHO-K1-Mock cells were harvested as described in chapter 2.2.1.1 and diluted to 5 x 10^5 cells/mL in cultivation medium. A volume of 200 µL was added to the wells of a 96-well plate and cultivated overnight under aseptic conditions. At the following day, the supernatant was gently removed and blocking of cells was performed with 300 µL 5 % milk powder in PBS for 1 h at RT. Subsequently the plate was washed three times with PBS-T and 1 x 10^{11} phage particles diluted in 100 µL PBS were added to each well. After incubation for 1 h at 4 °C the plate was washed again three times with PBS-T and horseradish peroxidase (HRP)-conjugated M13 antibody was applied diluted 1:1000 in PBS-T for 1 h at 4 °C. The plate was washed three times with PBS before adding TMB as a chromogenic substrate (100 µl per well). Color development was stopped with 50 µL of 0.2 M HCl per well after approximately 5 min and absorbance was measured at 450 nm using an Infinite M200 PRO microplate reader.

2.2.6.2 Specificity analysis of MC-FN-010 variants and quality control of recombinant FN-67B89 protein via ELISA

Binding ability and specificity of cystine-knot miniprotein fusion to target and off-target proteins was analyzed with an antibody-based detection method. Cavities of 96-well microtiter plates (MaxiSorp™) were coated with 1 µg FN-B, FN-67B89 or FN-6789 in a total volume of 100 µL coating buffer overnight at 4 °C. The cavities were washed three times with PBS-T, blocked with 1x casein diluted in PBS for 2 h at RT and washed again three times with PBS-T. Different concentrations of cystine-knot miniprotein fusion variants diluted in 100 µL PBS were applied to each well and incubated for 1 h at 4 °C. After a washing step (three times) with PBS-T, binding of respective cystine-knot miniprotein fusion was detected with an HRP-conjugated anti-S-tag antibody. Enzymatic reaction was detected with TMB used as a chromogenic substrate and was stopped with 0.2 M HCl after approximately 5 min. The measurement of absorbance at 450 nm was performed using an Infinite M200 PRO microplate reader.

To determine the quality of expressed and purified FN-67B89 protein, ELISA was used as technique to analyze protein-protein interactions in a 96-well format. The procedure was similar as described above, but using 100 µL of BC-1 antibody diluted 1:1000 in PBS-T as primary antibody and an HRP-conjugated anti-mouse antibody diluted 1:5000 in PBS as detection antibody.

2.2.6.3 Western blot analysis

Detection of biotin in cystine-knot miniprotein fusion was performed using western blot analysis. Protein samples were separated using gel electrophoresis (chapter 2.2.4.5). For blotting, gels were placed on a nitrocellulose membrane and covered with transfer buffer and Whatman paper stacks (equilibrated in transfer buffer) at each site. Protein transfer was performed in a Trans-Blot® Turbo™ Blotting System at constant 0.3 A and up to 25 V for 60 min. After blotting the membrane was blocked in 5 % milk powder in PBS-T for 1 h at RT on an orbital shaker. Membrane wash was performed with TBS-T for 5 min before placing the membrane in a HRP-conjugated SA solution, 1:5000 diluted in TBS-T, and incubated at 4 °C overnight. Subsequently, the membrane was washed twice quickly with TBS-T and once with TBS. For chemiluminescent detection of HRP reaction, SuperSignal™ West Dura Extended Duration Substrate with a 1:1 mixing ratio of Luminol/Enhancer Solution and Stable Peroxide Solution was distributed on the membrane. Protein signals were analyzed with a defined exposure time on ChemiDoc MP and documented photographically.

2.2.6.4 Analysis of Fos presentation on phages via ELISA and western blot

As a part of the phage library quality control, surface presentation of Fos-cystine-knot miniprotein complex was analyzed by ELISA. Plates were coated with 1 µg anti-M13 antibody per well in a total volume of 100 µL coating buffer and incubated at 37 °C for 2 h. Blocking of cavities was conducted with 300 µL of 1x casein in TBS at 4 °C overnight. Afterwards, washing with TBS was performed and 1 x 10^{12} phages diluted in 100 µL coating buffer were applied and incubated for 1 h at 37 °C. An anti-c-Fos antibody, diluted 1:1000 with 1x casein in TBS, was used as primary antibody and incubated for 1 h at RT. After two washing steps with TBS-T and another one with TBS, binding of primary antibody was detected with a HRP-conjugated anti-rabbit IgG antibody 1:5000 diluted with 1x casein in TBS. After washing twice with TBS-T and twice with TBS, assay was detected and completed as described in chapter 2.2.6.2.
The second quality control was a western blot analysis. For this purpose, 5 x 10^{11} phages were boiled in 15 µL 1x reducing SDS-loading buffer at 98 °C for 15 min and subjected to an SDS-PAGE according to chapter 2.2.4.5. After electrophoresis, gels were placed on a PVDF transfer membrane and further experimental process was performed essentially as described in chapter 2.2.6.3. Anti-c-Fos antibody was diluted 1:100 in 1 % milk powder in PBS and secondary HRP-conjugated anti-rabbit IgG antibody was diluted 1:5000. PBS-T was applied as washing solution and final washing was conducted with PBS. Lumi-Light western blotting substrate was applied as detection substrate.

2.2.6.5 Flow cytometry

Flow cytometry is a laser based technique to analyze intracellular proteins and extracellular surface proteins of single cells. Single cells pass different laser beams within a fluid stream system to allow the analysis of multiple physical parameters. Scattered and reflected light, detected by forward (FSC) and side scatter (SSC), correspond to cell size and granularity. Protein expression of cells can be detected directly or indirectly with fluorescence labeled antibodies, which emit light on their respective wavelength.
Flow cytometry was used to analyze binding ability of cystine-knot miniprotein fusions as a part of the hit identification procedure with roughly purified proteins in bacterial supernatant or purified proteins. Furthermore, this method was conducted to assess cell surface presentation of target proteins.

Binding analysis of cystine-knot miniprotein fusions was performed with bacterial supernatant (chapter 2.2.3.1) or fully purified proteins (chapter 2.2.3.3). Mammalian cells were harvested (chapter 2.2.1.1) by centrifugation at 300 x g for 5 min and washed with 300 µL FACS buffer. After centrifugation (300 x g for 5 min), cells were re-suspended in 10 µL bacterial supernatant diluted with 90 µL DPBS or in 100 µL purified cystine-knot miniprotein fusion diluted in DPBS and incubated at 4 °C for 1 h. The cells were washed three times with 300 µL FACS buffer, centrifuged again (300 x g for 5 min) and eventually taken up in 100 µL allophycocyanin (APC) conjugated anti-His-Tag antibody 1:100 diluted in DPBS for staining at 4 °C for 30 min. Subsequently, cells were washed twice with FACS buffer and once with DPBS. For staining of dead cells, 100 µL fixable dye eFluor 450™ 1:7500 diluted in DPBS was added, incubated at 4 °C for 20 min followed by two washing steps with FACS buffer. For flow cytometry analysis, cells were re-suspended in 200 µL FACS buffer. The measurement was performed on BD FACSCanto II and data was analyzed using FlowJo X.
Experimental procedure for binding analysis of biotinylated cystine-knot miniprotein fusion differed and is described as followed. For this purpose, biotinylated cystine-knot miniprotein was incubated with a five-fold molar excess of SA-APC in 100 µL DPBS protected from light at RT for 30 min. The mixture was added to the cells and placed at 4 °C for 1 h. After washing with FACS buffer and DPBS, staining of dead cells was performed as described above.
Analysis of surface presentation of human FAP-α and human DSG-3, was analyzed with anti-FAP-α antibody diluted 1:50 and anti-DSG-3 antibody diluted 1:100, respectively. Cy5-conjugated anti-mouse IgG diluted 1:100 was applied as detection antibody.

2.2.6.6 Immunofluorescence staining

Immunofluorescence (IF) staining is an imaging method to visualize biomolecule distribution on cells or tissue sections using antigen specific fluorescence labeled probes.
For IF staining cryopreserved tumor or brain pieces were cut in six micron thick sections, fixed in ice cold acetone for 5 min and air-dried. Slides were then blocked in PBS with 3 % BSA at RT for 5 min. For staining of EDB, 1 µg of the respective biotinylated cystine-knot miniprotein was incubated with 2.9 µg SA-Cy3 conjugate at RT for 30 min. The preformed complex was then added to the tumor sections and incubated for 30 min at 37 °C. Afterwards, the slides were washed three times with PBS containing 1 % BSA. CD31 staining was performed with an Alexa Fluor® 647 rat anti-mouse CD31 antibody diluted 1:100 in PBS with 1 % BSA for 30 min at 37 °C. After three washing steps in PBS, cell nuclei were stained with Hoechst 33342 diluted 1:5000 in PBS for 30 min at RT. Slides were washed again as described above and covered with coverslips in a thin layer of mounting medium. Images were captured with a Zeiss Apotome microscope and analyzed with ZEN software.

2.2.7 Peptide synthesis

Trimeric Alexa Fluor 680 (AF680) conjugated ligands as well as N-terminally biotinylated miniproteins were purchased from Pepscan. The obtained peptidic constructs, listed in Table 2.8, were stored in 100 µg aliquots at −20 °C. For the performed experiments all peptides were dissolved in 100 µL DPBS resulting in a concentration of 1 µg/µL. Their identities were verified by electrospray ionisation mass spectrometry (Pepscan). Purity was analyzed by analytical RPC by Pepscan. Additionally, trimeric ligands were analyzed via SDS-PAGE (chapter 2.2.4.5) and SPR (chapter 2.2.4.9) in order to characterize purity, target binding properties (binding to FN-67B89) and specificity (binding to FN-6789).

Table 2.8 Synthetic peptides with molecular weight and purity.

Peptidic construct	Molecular weight	Purity
AF680-(MC-FN-010)$_3$	13681.9 Da	90.5 %
AF680-(MC-FN-016)$_3$	13510.6 Da	90.5 %
AF680-(MC-FN-0115)$_3$	13129.0 Da	90.1 %
MC-FN-010-bio	4171.9 Da	95.5 %
MC-FN-0115-bio	4168.7 Da	94.9 %

2.2.8 Animal experimental methods

2.2.8.1 Tumor models

Mice were housed in the animal facility at BioNTech AG and all animal protocols were approved by the German Tierschutzkommision des Landesuntersuchungsamts Rheinland-Pfalz. Four weeks old Fox n1/nu mice ranging in weights between approximately 25 and 28 g were purchased from Janvier Labs. For xenograft mouse studies, 7 x 10^6 human U-87 MG glioblastoma cells were subcutaneously (s.c.) injected into the right flank of Fox n1/nu mice. Tumors were allowed to grow for approximately five weeks. Subcutaneous tumor size was determined using ellipsoid formula ($\frac{width\ x\ length^2}{2}$). All animals with tumor volumes between 100 – 1200 mm^3 were included in the studies and mice were randomly assigned to experimental cohorts.

2.2.8.2 *In vivo* and *ex vivo* imaging

Mice carrying a desired tumor size were included for analysis of biodistribution and tumor targeting of trimeric constructs. All trimeric constructs were injected intravenously (i.v.) via the retrobulbar venous plexus in a final volume of 100 µL DPBS buffer (3.34 nmol/mice). Mice groups (n=3 for each construct) were imaged in an IVIS Spectrum system using excitation range of 615 – 665 nm and monitoring emission signals at 695 – 770 nm. Imaging process was performed 1 h, 2 h or 6 h post-injection. At the end of the experiment, after euthanization of the mice, the tumor and specific organs were excised, imaged, weighed and cryo-conserved in a 2-methyl butane containing tubes, which were placed in liquid nitrogen. Fluorescence intensity of regions of interest was quantified using Living Image® software. Statistically significance was calculated based on triplicate data sets using two-way ANOVA analysis with corrected p values (Dunnett´s method) for comparison of different time and organ signals in GraphPad Prism 6. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, n.s. = not significant. All data are presented as mean ± SD.

3 Results

3.1 Development of customized cell-based drug discovery platform

The efficient discovery of lead compounds plays a crucial role for the further development of effective diagnostic and therapeutic products. As described in chapter 1.1, the screening of combinatorial molecular libraries represents a major pillar within the lead discovery field. Many selection strategies using surface display of molecules have been developed, which facilitate the *de novo* discovery of target interacting ligands. The phage display technology is the most widely applied method using target proteins either captured on beads, coated on immunotubes or exposed on the cell surface[32]. However, the usage of recombinant soluble proteins is partially limited due to poor expression, solubility or complex structure, which holds true especially for membrane proteins with multiple membrane spanning regions and small extracellular domains. Targeting whole cells could overcome these limitations and allow addressing surface proteins in their native conformation[140,141]. Even though many standard experimental protocols for phage panning have been described, they are not inevitably amenable for all combinatorial phage libraries. The panning as well as the following hit identification protocols rather needs to be verified and customized to the individual drug discovery platforms.

To this end, the following sections are focusing on the establishment of a cell-based panning protocol applicable for cystine-knot miniprotein phage libraries. Furthermore, an efficient high-throughput system has been developed for the analysis of selected ligands in conjugation with flow cytometry analysis. In addition, the hit identification process was optimized to allow simultaneous detection of target and off-target binding as well as to increase binding sensitivity. In order to proof the feasibility of binder identification by applying the established cell-based panning protocol and the downstream hit identification process, a naïve library was screened against a target overexpressing cell line.

3.1.1 Establishment of a cell-based phage panning protocol with a model system

The aim of this work package was the establishment of an efficient cell-based panning protocol suitable for cystine-knot miniprotein based phage selections. In a proof of concept experiment the enrichment of a known tumor specific cystine-knot miniprotein (MC-FA-010), presented through the major coat protein on phages, within a panning against FAP-α overexpressing eukaryotic cell line (CHO-K1-FAP-α) was analyzed. Therefore, different amounts of MC-FA-010 displaying phages were spiked into the naïve MCopt 1.0 phage library and screened against CHO-K1-FAP-α cells. Phage titer and MC-FA-010 enrichment was monitored by colony PCR before and after each panning round. This model system mimics real selection conditions and demonstrates the general applicability for panning of naïve complex phage libraries.

At first, as a part of a quality control of the materials needed for the selection experiment, target expression on CHO-K1-FAP-α cells was analyzed by flow cytometry (Figure 3.1 A) and western blot (Figure 3.1 B). Both assays showed a clear target expression in CHO-K1-FAP-α cells. A surface presentation of FAP-α was also observed, while the average of bulk population revealed a receptor density of less than 10^6 FAP-α molecules per cell. The amount was estimated based on the known receptor density of monoclonal CHO-K1-FAP-α cells.

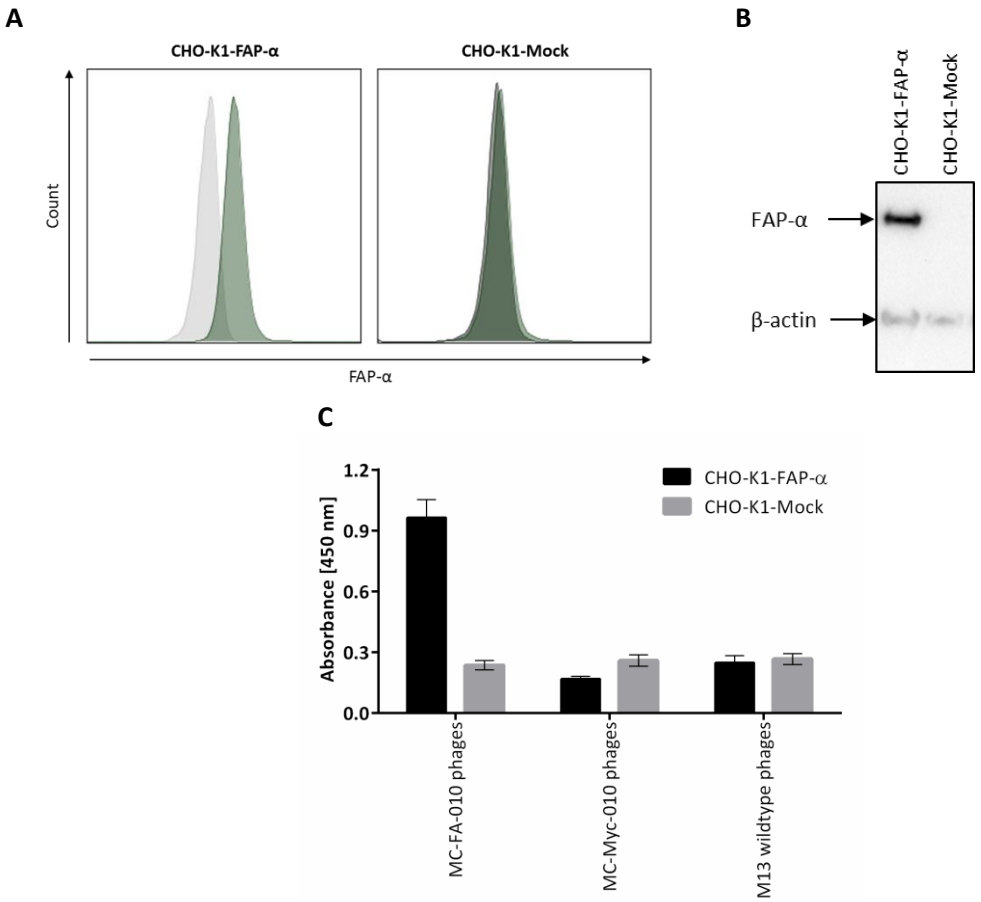


Figure 3.1 Quality control of CHO-K1-FAP-α cells and MC-FA-010 displaying phages. (A) FAP-α expression on CHO-K1-FAP-α and CHO-K1-Mock cells. Expression of FAP-α was determined by flow cytometry using an anti-human-FAP-α antibody and detected by a Cy5-conjugated anti-mouse antibody (green histograms). Grey histograms represent staining only with Cy5-conjugated anti-mouse antibody. Representative histograms are shown. **(B)** Analysis of FAP-α expression in CHO-K1-FAP-α and CHO-K1-Mock cells monitored by western blot. Target expression was analyzed with an anti-human-FAP-α antibody. Comparative loading amounts were measured using an anti-β-actin antibody. HRP-conjugated anti-mouse antibody was applied to detect primary antibody binding. **(C)** Binding analysis of MC-FA-010 presenting phages to CHO-K1-FAP-α via whole-cell ELISA. 1×10^{11} of each phage particles were applied to adherent CHO-K1-FAP-α or CHO-K1-Mock cells. A HRP-conjugated anti-M13 antibody was used to analyze binding. ELISA was performed in triplicates ± SD with MC-Myc-010 presenting phages and M13 phages serving as negative controls.

Additionally, the presentation of MC-FA-010 on phages was analyzed using whole-cell ELISA (Figure 3.1 C). MC-FA-010 presenting phages exhibited about three-fold higher binding signals to CHO-K1-FAP-α in comparison to CHO-K1-Mock cells. In contrast, the signals generated by the negative controls MC-Myc-010 presenting phages and M13 phages towards both cell lines were overall low. The described quality control experiments clearly indicated correct surface presentation of MC-FA-010 on phages and a considerably high expression of the target protein FAP-α on the surface of the generated cell line, both being a prerequisite for the intended cell-based panning protocol validation. For the selection experiments, MC-FA-010 presenting phages were spiked into the MCopt 1.0 phage library in three different dilutions ($1{:}10^3$, $1{:}10^6$, $1{:}10^9$). The phage pools were screened separately against CHO-K1-FAP-α cells after depletion with CHO-K1-Mock cells according to the experimental outline as depicted in Figure 3.2. Panning was performed with 1×10^{12} phage particles at each round.

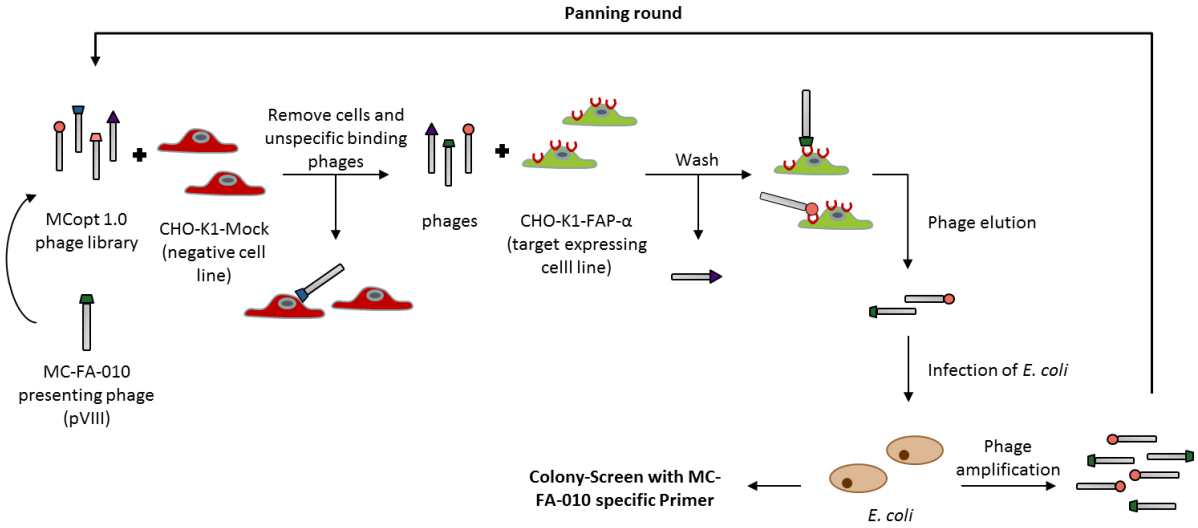


Figure 3.2 Experimental overview of a standard cell-based phage display panning procedure. MC-FA-010 presenting phages were spiked into the MCopt 1.0 library in different ratios. The pools were then incubated with the negative cell line to eliminate unspecific binding phages. After removal of cells, the remaining supernatants were added to target expressing cells. Subsequently, unbound phages were removed by washing steps, cell-bound phages were eluted via pH shift and amplified for the following panning round. Presence of MC-FA-010 presenting phages was analyzed by colony PCR with respective oligonucleotides.

Infectious phage input, post-input (after negative selection) and output titer was monitored within every panning round by counting CFU of infected *E. coli* XL1-Blue cells. The ratio of post-input to output generally decreased with increasing selection rounds, typically indicating an enrichment of a phage population (Figure 3.3 A).

A

Dilution	Panning round	Input	Post-input	Output	Post-input/output ratio
$1:10^3$	1	4.80×10^{11}	2.35×10^{11}	3.43×10^{5}	6.85×10^{5}
$1:10^6$	1	3.35×10^{11}	1.75×10^{11}	1.05×10^{6}	1.67×10^{5}
	2	7.95×10^{10}	5.36×10^{10}	1.90×10^{6}	2.82×10^{4}
$1:10^9$	1	1.70×10^{11}	3.53×10^{11}	3.43×10^{5}	1.03×10^{6}
	2	5.85×10^{10}	4.42×10^{10}	1.69×10^{5}	2.62×10^{5}
	3	1.40×10^{11}	1.43×10^{11}	5.95×10^{6}	2.40×10^{4}

B

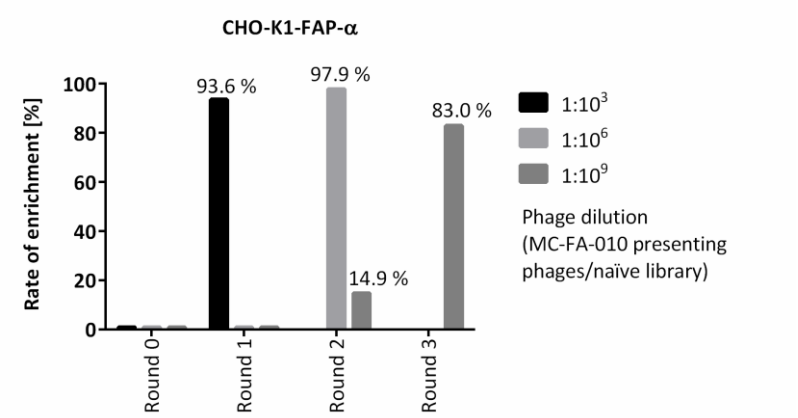


Figure 3.3 Phage titers and enrichment of spike-in experiment against CHO-K1-FAP-α cells. (A) Titers of phage populations throughout the successive panning rounds against CHO-K1-FAP-α cells. Titers were determined by counting CFU. **(B)** Enrichment of MC-FA-010 presenting phages through cell-based selection against CHO-K1-FAP-α cells. Three different amounts of MC-FA-010 presenting phages were spiked into the MCopt 1.0 phage library. Clones were analyzed after the respective selection round by colony PCR with MC-FA-010 specific oligonucleotides. Enrichment was calculated based on the ratio of MC-FA-010 carrying clones and the analyzed clones in total. Round 0 represents the starting pool before the selection process.

The experimental evaluation of successful MC-FA-010 enrichment was assayed by colony PCR analysis with 47 recovered clones after each panning round as shown in Figure 3.3 B. With the highest amount of MC-FA-010 presenting phages mixed into the library, 93.6 % of the analyzed clones were PCR-positive already after one panning round. In the experiment with the initial dilution factor of $1:10^6$, MC-FA-010 enrichment could be observed after second panning round with a proportion of 97.9 %. With the last example of $1:10^9$ phage dilution the specific enrichment could be detected after the second panning round with 14.9 % positive clones and reached 83 % in the third round. Taken together, depending on the initial dilution ratio, MC-FA-010 presenting phages could be enriched specifically after one, two or three successive panning rounds. At least for this particular model system this outcome thus demonstrates the high efficiency of the cell-based panning protocol for ligand selections based on our phage library.

In a control experiment, the two highest amounts of MC-FA-010 presenting phages were again added into the MCopt 1.0 library ($1:10^3$, $1:10^6$). However, this time the screening was performed only against CHO-K1-Mock cells and consequently no negative depletion was done in this experiment. The input to output ratio increased over consecutive panning rounds and suggesting no certain phage enrichment (Figure 3.4 A). The resulting clones after each phage panning cycle were again used for colony PCR in order to assay for MC-FA-010 existence. As expected, none of those analyzed clones

carried the MC-FA-010 insert, so that the calculated enrichment rate was 0 % in all cases (Figure 3.4 B). With this control experiment growth advantages of MC-FA-010 presenting phages could be excluded.

A

Dilution	Panning round	Input	Output	Input/output ratio
$1:10^3$	1	1.26×10^{11}	8.40×10^4	1.50×10^6
	2	1.04×10^{11}	2.80×10^4	3.71×10^6
$1:10^6$	1	2.24×10^{11}	8.60×10^4	2.60×10^6
	2	4.71×10^{11}	1.60×10^4	2.94×10^7

B

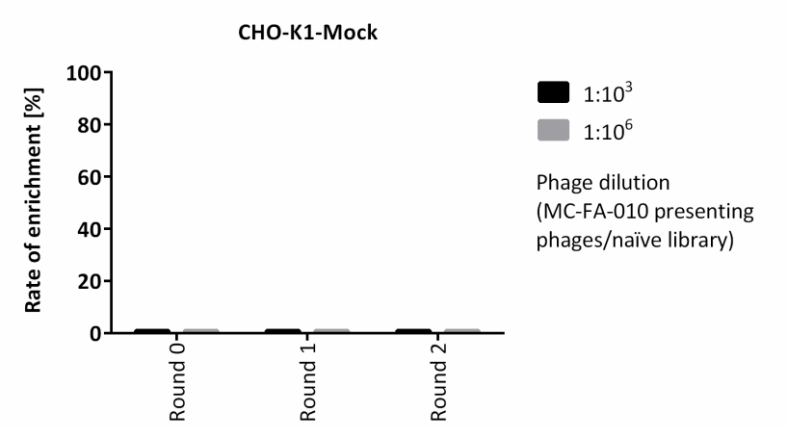


Figure 3.4 Phage titers and enrichment of a spike-in experiment against CHO-K1-Mock cells. (A) Phage titers throughout panning rounds of control experiment against CHO-K1-MOCK cells. Counting CFU was performed for titer determination. **(B)** Analysis of growth advantages of MC-FA-010 presenting phages in a spike-in model by cell-based selection against CHO-K1-Mock cells. MC-FA-010 presenting phages were added with two different proportions into MCopt 1.0 phage library for panning. The presence of MC-FA-010 after each round was analyzed via colony PCR with corresponding oligonucleotides.

3.1.2 Development of a flow cytometry-based hit identification process

Having shown the efficiency of the cell-based panning protocol to specifically enrich MC-FA-010 presenting phages, in the following a flow cytometry-based screening procedure has been developed to allow binding analysis of individual ligands in a high-throughput manner. The experimental overview of the whole hit identification procedure is depicted in Figure 3.5. For hit screening, selected cystine-knot miniproteins were expressed N-terminally fused to thioredoxin-A, a His-tag and an S-tag (Trx-cystine-knot miniprotein fusion) in a 96-well mini scale format. The expression rate of each clone was determined via E-PAGE™ analysis. The corresponding sequence was revealed and additionally binding of each variant to target and off-target protein was assayed via flow cytometry. Based on the derived signal-to-noise ratio of cellular binding and the expression value, a ranking value for each candidate was calculated as a measure for specific target interaction.

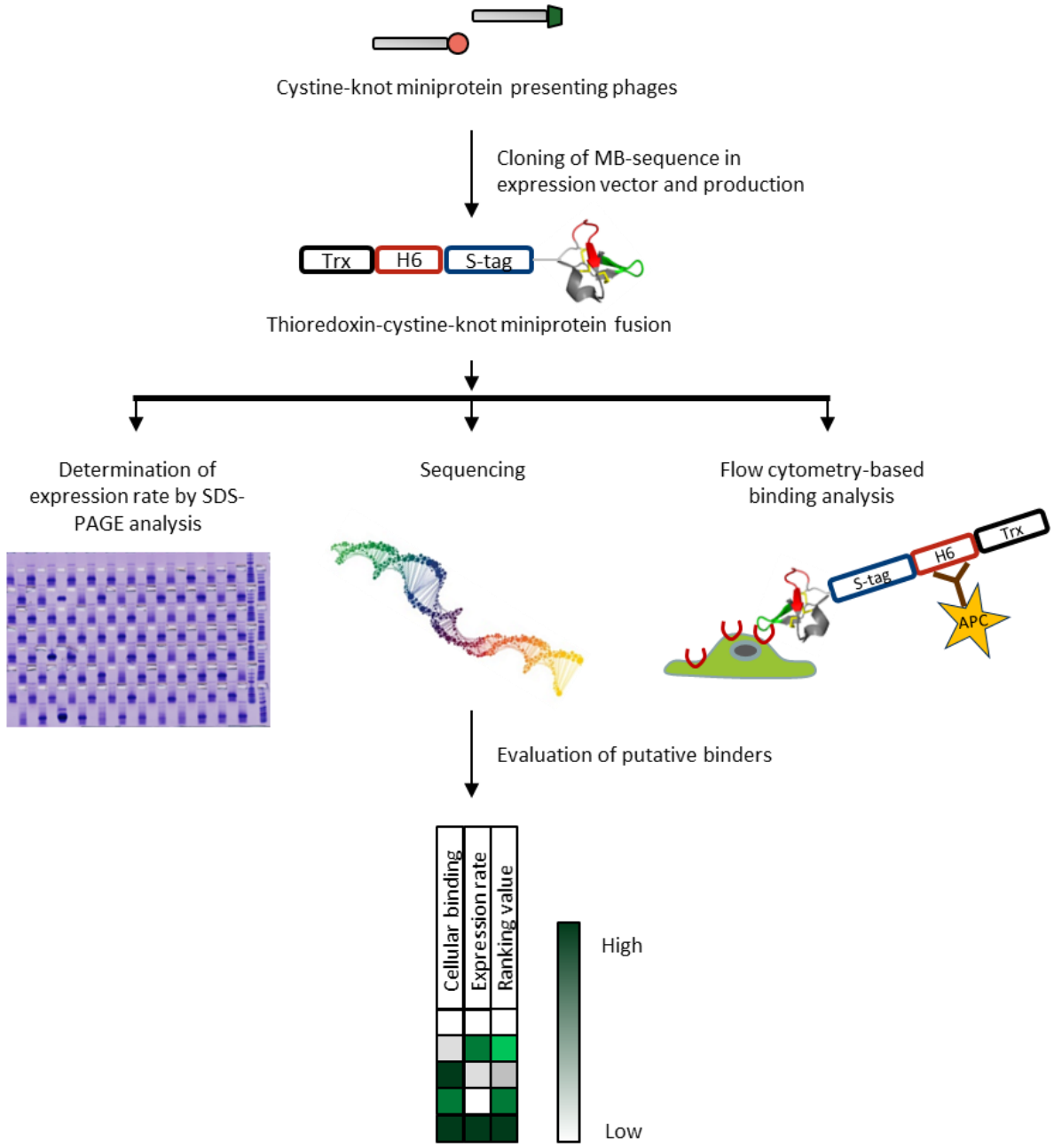


Figure 3.5 Experimental outline of downstream hit identification procedure. DNA sequences of selected cystine-knot miniproteins were cloned into an expression vector for 96-well production as Trx-cystine-knot miniprotein fusion. For each clone respective sequence, expression rate and flow cytometry-based target binding was determined. Cellular binding value [signal (target expressing cell)/signal (control cell)] was correlated to the expression rate resulting in a ranking value as a measure for specific target interaction.

The phage pool of the second selection round (initial dilution ratio of $1:10^9$) from the model selection against CHO-K1-FAP-α cells was used for establishment of the cellular binding analysis within the hit identification procedure. Enriched cystine-knot miniproteins were cloned as a pool into the pET-32-LibEx expression vector and introduced into SHuffle® T7 Express competent *E. coli* cells. In order to analyze the sequence composition of the resulting repertoire, 92 single clones were randomly chosen for sequencing. As shown in Figure 3.6, different cystine-knot miniprotein sequences were presented at 55.4 % and peptide sequences at 10.9 %. Furthermore, a small portion of frameshift sequences, vector re-ligands and sequence errors were observed. But importantly, enriched MC-FA-010 dominated the pool with a proportion of 23.9 %.

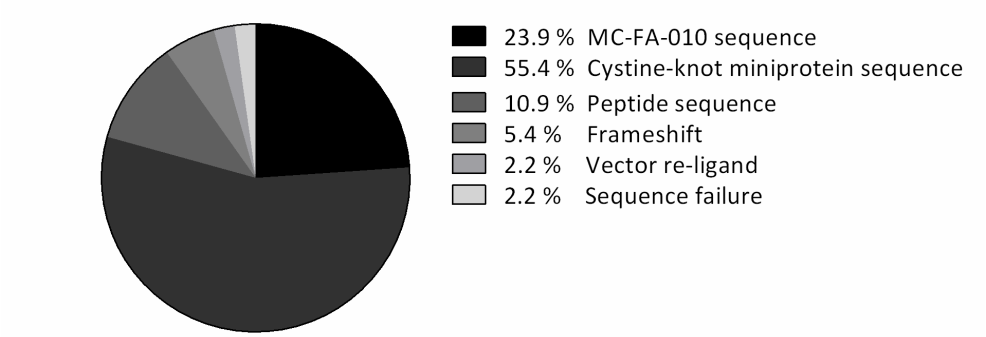


Figure 3.6 Sequence composition after cloning of the model screen pool into the pET-32-LibEx expression vector. In total 92 clones of the pool were randomly chosen for sequencing.

In the course of the hit identification procedure, individual clones were expressed as Trx-cystine-knot miniprotein fusion in the cytoplasm of the host cell. A simple heat-purification step of cell lysates was performed to recover thermotolerant Trx-cystine-knot miniprotein fusions from endogenous proteins yielding in general a purity of approximately 70 – 99 % (supplementary Figure 5.1). The supernatant was directly used for cellular binding analysis towards CHO-K1-FAP-α cells and CHO-K1-Mock cells in two separate 96-well plates, as both cell lines express intrinsic GFP. As presented in Figure 3.7, several constructs showed increased interaction with CHO-K1-FAP-α cells in comparison to CHO-K1-Mock cells. These included the known MC-FA-010 proteins, as depicted with asterisks. Importantly, no interaction of all analyzed constructs with CHO-K1-Mock cells could be observed allowing a specific discrimination of target and off-target binding. Furthermore, the low background signal indicated an efficient negative depletion during phage panning.

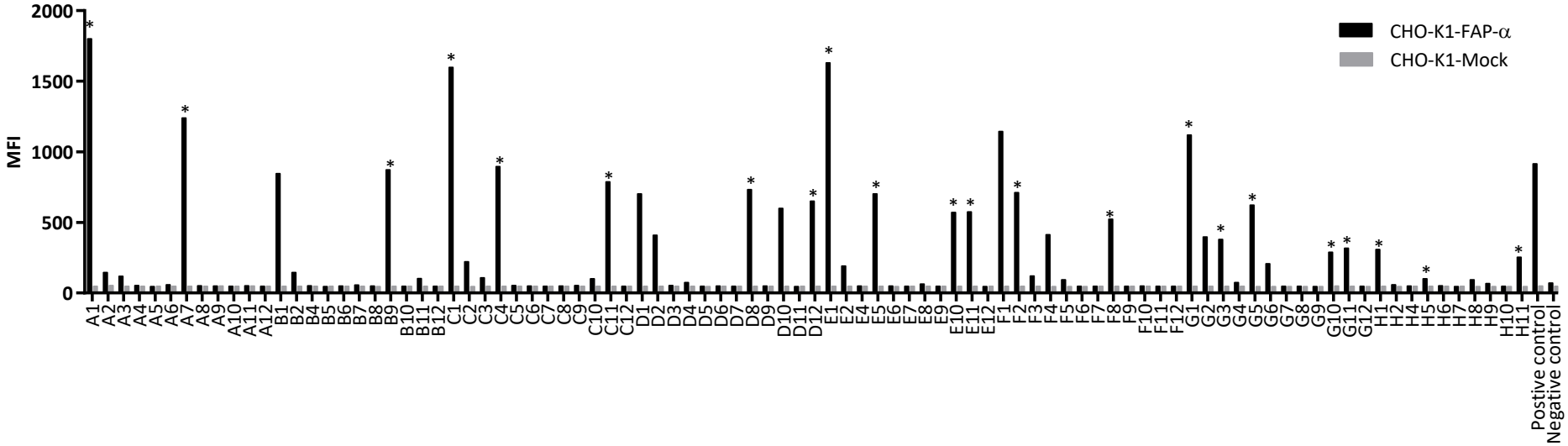


Figure 3.7 Cellular binding analysis of single variants on CHO-K1-FAP-α and CHO-K1-Mock cells. For flow cytometry measurements, 10 µL *E. coli* supernatant was applied to the cells and binding of Trx-cystine-knot miniprotein fusion was detected using an APC-conjugated anti-His antibody. The mean fluorescence intensity (MFI) of the APC signal was calculated for data evaluation. Trx-MC-FA-010 variants are marked with an asterisk. Purified Trx-MC-FA-10 (1 µM) served as positive control, while Trx-MC-Myc-010 (1 µM) was used as negative control.

In order to take the potential differential expression of the candidates within binding analysis into account, the amount of each Trx-cystine-knot miniprotein fusion was determined using E-PAGE™ analysis calculated based on a protein standard with a defined protein concentration. The expression value was included for generation of ranking values, which consequently represent a point of reference for identifying promising candidates. The expression and cellular binding values of individual variants are shown in the supplementary (Table 5.1). The FAP-α binding cystine-knot miniprotein (MC-FA-010) was observed 22 times in 92 analyzed variants. Remarkably, for all MC-FA-010 clones middle to high ranking values could be observed as summarized in Figure 3.8. In addition, two cystine-knot miniproteins (B01, D01) and two peptides (D10, F01) interacted with the CHO-K1 FAP-α cells and displayed high ranking values as well. But follow-up analysis of these clones was not performed. Altogether, these findings demonstrate the high efficiency of the downstream cell-based hit identification procedure to allow selection of target interacting cystine-knot miniproteins.

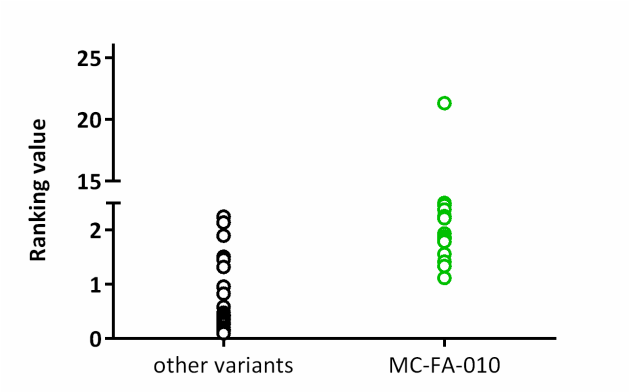


Figure 3.8 Ranking values of clones after CHO-K1-FAP-α model selection and hit identification procedure. MC-FA-010 variants are depicted in green apart from other variants in black. Cellular binding, expression and ranking value as well as corresponding protein sequence of each variant is depicted in the supplementary Table 5.1.

3.1.3 Panning and hit identification of naïve MCopt 1.0 library against CHO-K1-FAP-α cells

Based on the described model system, a clear enrichment of MC-FA-010 displaying phages after panning and its identification within downstream hit screening could be achieved, providing evidence for a general feasibility of the established protocol. The aim of the following experiment was to apply a completely naïve MCopt 1.0 library for analyzing the whole process performance under real conditions. For this purpose, MCopt 1.0 library was screened in three successive panning rounds against CHO-K1-FAP-α cells according to the experimental outline in Figure 3.2. The first panning round started with 1 x 10^{13} MCopt 1.0 library phages in order to yield a high library coverage of ~10000 fold. The two following rounds were conducted with 1 x 10^{12} amplified phage particles as input. The phage titers were again determined by CFU counting of infected *E. coli* XL1-Blue cells (Figure 3.9 A).

A

Panning round	Input	Post-input	Output	Post-input/output ratio
1	5.65×10^{11}	3.53×10^{11}	1.10×10^{6}	3.21×10^{5}
2	6.25×10^{10}	6.01×10^{10}	437	1.38×10^{8}
2	9.50×10^{10}	2.61×10^{10}	2.74×10^{3}	9.53×10^{6}
3	2.68×10^{11}	4.32×10^{11}	6.40×10^{4}	6.75×10^{6}

B

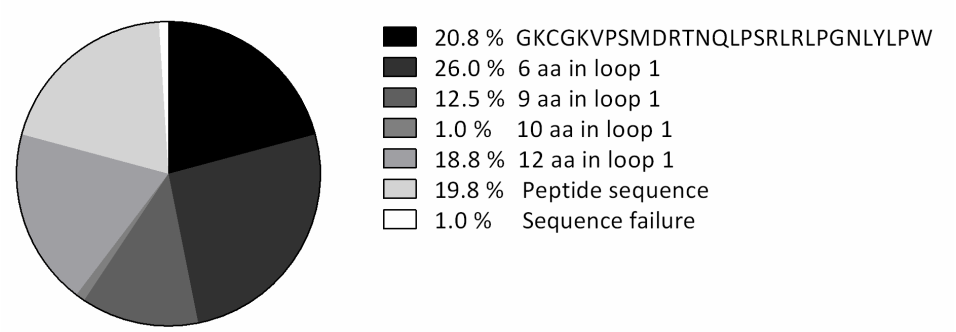


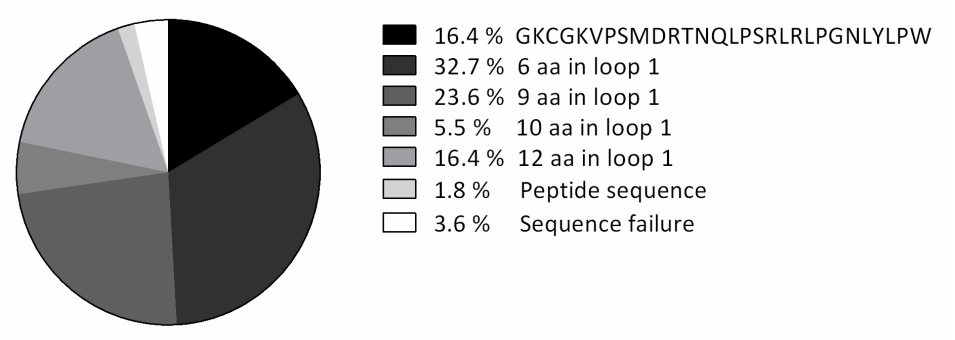
Figure 3.9 Phage titers and sequence composition of the screening of the naïve MCopt 1.0 library against CHO-K1-FAP-α cells. (A) Titers of phage populations throughout panning rounds of MCopt 1.0 library screening against CHO-K1-FAP-α cells. Input, post-input and output was determined within every panning round by counting CFU. The second panning round was performed twice. **(B)** Sequence composition of the screen pool after the third panning round against CHO-K1-FAP-α cells. In total 96 clones were randomly selected and sequenced.

At each selection round the input titers were much lower than intended, which indicated a portion of non-infectious phages in the pool. Nevertheless, in the first round an input of 5.65×10^{11} phages could be achieved yielding a 500-fold library coverage. The post-input titer was expected to be in general lower than the input upon negative depletion of phages, except in the third screening round. The second panning round was extracurricular performed twice, as the initial output with 437 clones was very low. Both obtained output phage pools were combined for the last panning round. The post-input to output ratio as a measure of enrichment rate increased unexpectedly over panning rounds. However, the progress of post-input to output ratio only served as rough indications of binder enrichment. For detailed clarification, 96 clones were sequenced after panning completion.

Figure 3.9 B shows the sequence distribution of clones derived after the third screening round. The proportion of cystine-knot miniproteins are arranged according to their amino acids length in loop 1 and the proportion of peptide sequences and sequence failures are given. One peptide sequence (GKCGKVPSMDRTNQLPSRLRLPGNLYLPW) dominated the screen pool with 20.8 %, meaning that it has been strongly enriched during selection. Furthermore, many single cystine-knot miniprotein sequences with different lengths of randomized amino acids in loop 1 could be observed. Sequences containing a loop length of six amino acids occurred to 26 %, nine amino acids to 12.5 % and twelve amino acids to 18.8 %. Unexpectedly, also a sequence comprising ten amino acids in loop 1 appeared, even though this loop length was not included in the library design.

In order to identify ligands specific for CHO-K1-FAP-α, a downstream hit identification procedure was performed as depicted in Figure 3.5. After cloning of the screen pool into the pET-32-LibEx expression vector for Trx-cystine-knot miniprotein production, single clones were sequenced again to monitor sequence distribution and to provide sequence information for hit identification.

A



B

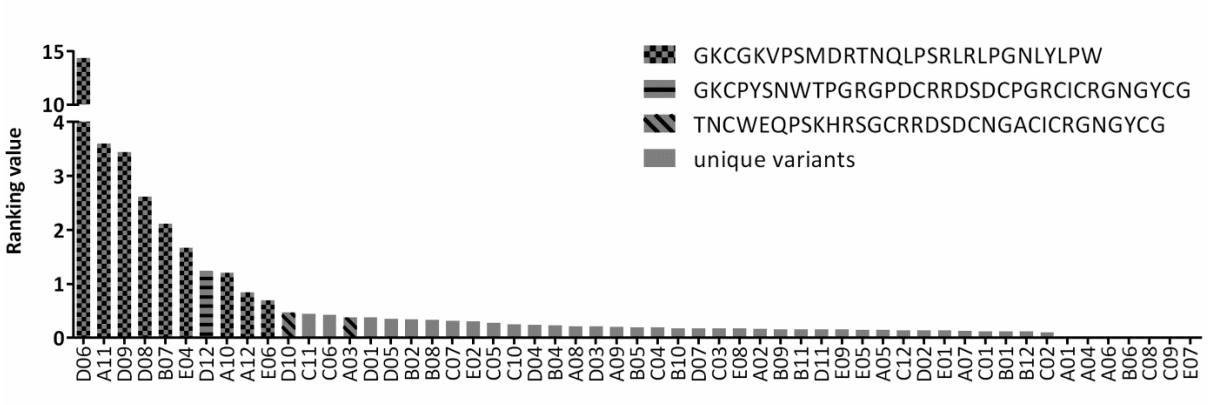


Figure 3.10 Hit identification of MCopt 1.0 library panning against CHO-K1-FAP-α cells. (A) Sequence composition after cloning of the screen pool into the pET-32-LibEx expression vector. For sequencing 55 clones were randomly selected. **(B)** Ranking value of hit identification. For each clone the cellular binding value was divided by the expression rate to calculate its ranking value. Identical clones are highlighted with sequence information. Expression value, cellular binding value, ranking value and sequence of each clone are listed in the supplementary Table 5.2.

Figure 3.10 A shows the sequence distribution of clones after cloning into the expression vector. Again, the sequences are arranged according to their amino acids length in loop 1 and the proportion of peptide sequences and sequence failures are depicted. As expected, the enriched peptide sequence (GKCGKVPSMDRTNQLPSRLRLPGNLYLPW) was highly represented after screen pool cloning with a proportion of 16.4 %, which is consistent to the findings on phage pool level (Figure 3.9 B). Apart from that, unique cystine-knot miniproteins with six and nine amino acids in loop 1, 32.7 % and 23.6 %, represented the highest proportion, followed by cystine-knot miniproteins carrying twelve amino acids (16.4 %) and ten amino acids in loop 1 (5.5 %) as well as a peptide sequence (1.8 %).

Individual Trx-cystine-knot miniprotein fusions were then produced in a high-throughput compatible manner combined with subsequent expression and cellular binding analysis. Binding ability of the Trx-cystine-knot miniprotein fusions was assayed via flow cytometry using the CHO-K1-FAP-α as well as the CHO-K1-Mock cells. As presented in Figure 3.10 B, the outcome of performed hit identification process resulted in ranking values, which were determined by dividing cellular binding value with the expression value for each clone. The enriched peptide sequence and furthermore a cystine-knot miniprotein showed strong FAP-α interaction, as the corresponding ranking value was relatively high (supplementary Table 5.2). Interestingly, the cystine-knot miniprotein sequence could be verified as MC-FA-010, the known FAP-α specific protein, which was previously identified via MCopt 1.0 screening against soluble FAP-α. Even though on phage pool level MC-FA-010 sequence was not detected and consequently no specific enrichment could be observed, the downstream hit screening

allowed its selection. These findings underline the ability of the cell-based phage panning and subsequent hit identification to identify novel target binding cystine-knot miniproteins.

3.1.4 Optimization and extension of hit identification process

High-throughput hit screening represents a fundamental step for identifying target binding ligands after phage display selection. The capacity of analyzing large numbers in parallel would be highly beneficial for sufficient coverage of many clones. For the in-house drug discovery platform, an efficient high-throughput compatible expression and purification of individual clones combined with flow cytometry has been demonstrated. The cellular assay relied on distributing target positive and negative cell lines into two separate plates for binding analysis, as both cell lines endogenously overexpress GFP due to stable transfection. With the aim to further increase the throughput, a second negative cell line was generated that overexpresses another fluorescence dye (dsRed). Target expressing cells and red fluorescent control cells could then be merged for simultaneous detection of target and off-target binding in one well of a microtiter plate.
For this purpose, stable CHO-K1 cells expressing dsRed were generated using the PiggyBac transposon system. A gene cassette carrying the dsRed expression gene as well as a hygromycin resistance gene for selection of transfected cells was integrated into the host cell genome. The transfection efficiency was monitored by flow cytometry and based on the percentage of dsRed positive CHO-K1 cells. As shown in Figure 3.11 A, five days after transfection 18.4 % of the cell population expressed dsRed. At day 18 dsRed positive cells reached 47.7 % expression and slightly increased to 55.3 % at day 29. This progress indicated a polyclonal cell population containing clones with different dsRed expression levels and an insufficient selection process. To isolate a monoclonal cell population, cells with strong dsRed expression were selected using FACS. Single cells were expanded for analysis of clonal diversity in terms of dsRed expression. Figure 3.11 B shows a flow cytometric analysis of monoclonal cells with high and homogeneous dsRed expression in comparison to the prior polyclonal population. The generated cells were termed CHO-K1-dsRed-Mock and represent a second negative cell line for the hit identification process.

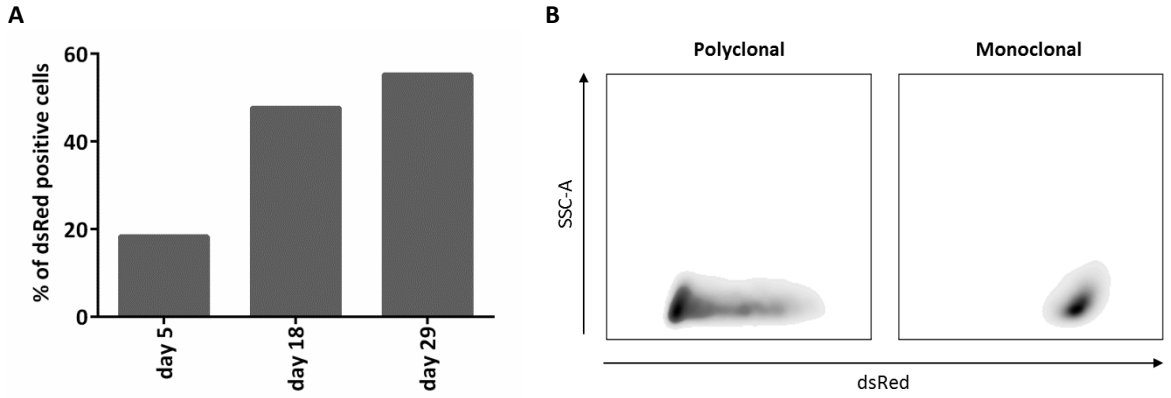


Figure 3.11 Generation of stable CHO-K1-dsRed-Mock cell line. (A) Transfection efficiency of dsRed expressing CHO-K1 cells determined by flow cytometry. Percentage of dsRed-expressing cells derived from total cell population and determined by referencing to CHO-K1 wild type cells. Day indicates analysis after transfection. **(B)** Representative density plot of polyclonal and monoclonal CHO-K1-dsRed-Mock population. Expression of dsRed was analyzed by flow cytometry before and after generation of monoclonal cell line via FACS.

The applicability of a second negative cell line to be used for off-target binding analysis in the hit identification procedure was investigated in combination with CHO-K1-FAP-α cells and based on the known cystine-knot miniprotein MC-FA-010 again. For the validation of the flow cytometry assay

equal cell numbers of both cell lines were mixed prior to incubation with purified Trx-MC-FA-010 miniprotein fusion or protein containing bacterial lysate. Bacterial lysate was applied in order to simulate cellular binding conditions like intended for downstream hit screening. Gating of the two cell lines was unproblematic due to the clear differentiation of GFP and dsRed fluorescence dyes (Figure 3.12 A). Both different presented versions of Trx-MC-FA-010 miniprotein fusion clearly interacted with CHO-K1-FAP-α cells, while no binding to CHO-K1-dsRed-Mock cells was observed (Figure 3.12 B). The proof of concept experiment demonstrates the ability to duplicate the throughput for hit screening as a result of combining CHO-K1-FAP-α and CHO-K1-dsRed-Mock cells allowing simultaneously target and off-target binding analysis of selected cystine-knot miniproteins within one 96-well plate.

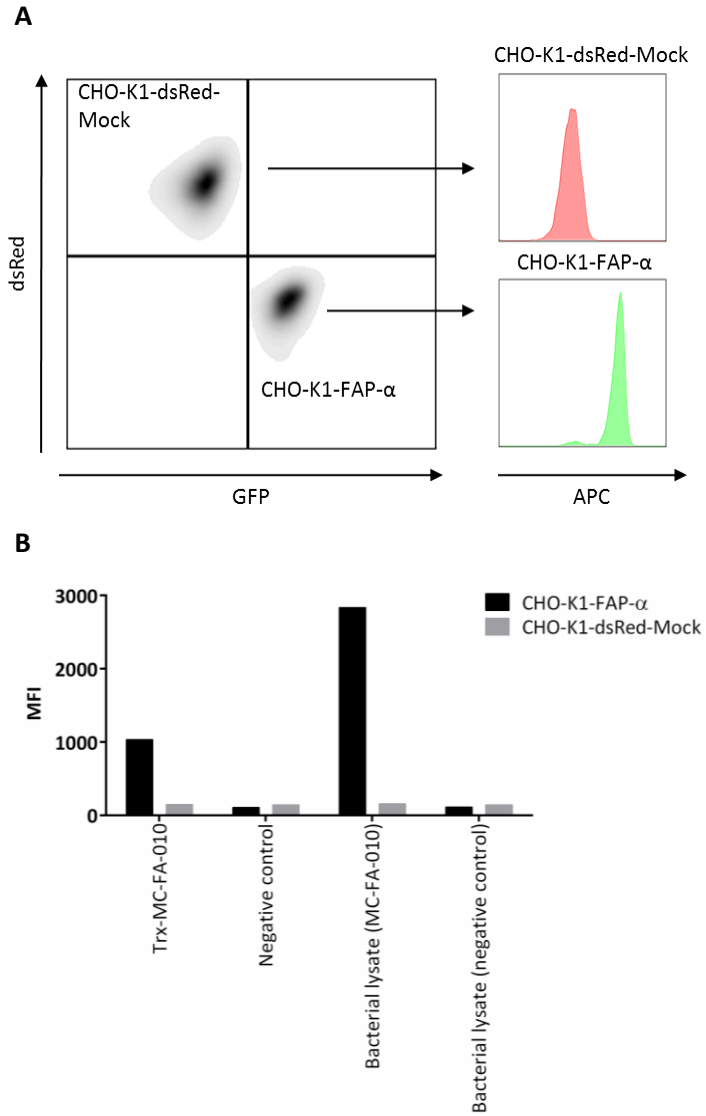


Figure 3.12 Simultaneous analysis of cellular target and off-target binding using flow cytometry. (A) Representative density plot of mixed CHO-K1-dsRed-Mock and CHO-K1-FAP-α cell population presenting the dsRed and GFP signal. The histograms show the APC signal of gated CHO-K1-dsRed-Mock and CHO-K1-FAP-α cells that represents protein binding to the respective cells. **(B)** CHO-K1-dsRed-Mock and CHO-K1-FAP-α cells were equally mixed for binding analysis of 1 µM purified Trx-cystine-knot miniprotein fusion or 10 µL bacterial lysate. Protein binding was detected with an APC-conjugated anti-His-tag antibody. Graph represents the MFI of APC signals. Experiment was repeated three times with similar results.

The extension of the cellular hit identification repertoire also included an optimization of flow cytometry-based assay sensitivity to lower the possibility to miss putative binders due to low affinity. A previous report from Kim and co-workers described a remarkably increase of knottin-like peptides target affinity by dimerization as a result of exploiting the avidity effect[142]. Tetramerization was the strategy of choice to gain increased target binding strength of cystine-knot miniproteins via SA complexation. While keeping high-throughput suitable manner in sight, a system was developed to allow production of Trx-cystine-knot miniprotein fusion directly linked to a biotin molecule in a 96-well scale. *In vivo* biotinylation in *E. coli* using the enzyme named biotin holoenzyme synthetase (BirA) has been shown to selectively attach biotin to lysine residue side chain within a specific recognition sequence termed AviTag™ (unique 15 amino acid peptide) attached to the protein of interest[143].
For this purpose, FAP-α binding cystine-knot miniprotein MC-FA-012, a derivative of parental MC-FA-010, and Flp-In™-CHO-FAP-α cells were used as model system. Corresponding sequence encoding for cystine-knot miniprotein was genetically inserted to the pET-32-AX2 vector to conduct expression with an AviTag™. After cloning, vectors were introduced into *E. coli* SHuffle® T7 Express pBirAcm cells. Subsequent protein expression was done in a 96-well plate in autoinduction medium with or without the addition of D-biotin to the culture. Cystine-knot miniprotein fusions were subsequently roughly purified via heat incubation. Western blot analysis in Figure 3.13 A displays biotinylation success of Trx-MC-FA-012 miniprotein fusion with correct size, in case that D-biotin was present during production. For analysis of stronger binding, biotinylated Trx-MC-FA-012 miniprotein fusion of bacterial lysate was mixed with APC-conjugated SA to pre-assemble a tetramerized complex and compared to monomeric Trx-MC-FA-012 miniprotein fusion of bacterial lysate in a flow cytometry set up (Figure 3.13 B). Obviously, the tetramerized format of the Trx-cystine knot miniprotein led to an approximately 30-fold higher binding signal than the monomeric version. Therefore, it can be concluded that the tetramerization strategy indeed exploited avidity effect and resulted in stronger target binding as well as improved assay sensitivity for future hit screenings. Furthermore, the high-throughput expression system for direct biotin conjugation to the proteins facilitates the implementation into the hit identification process.

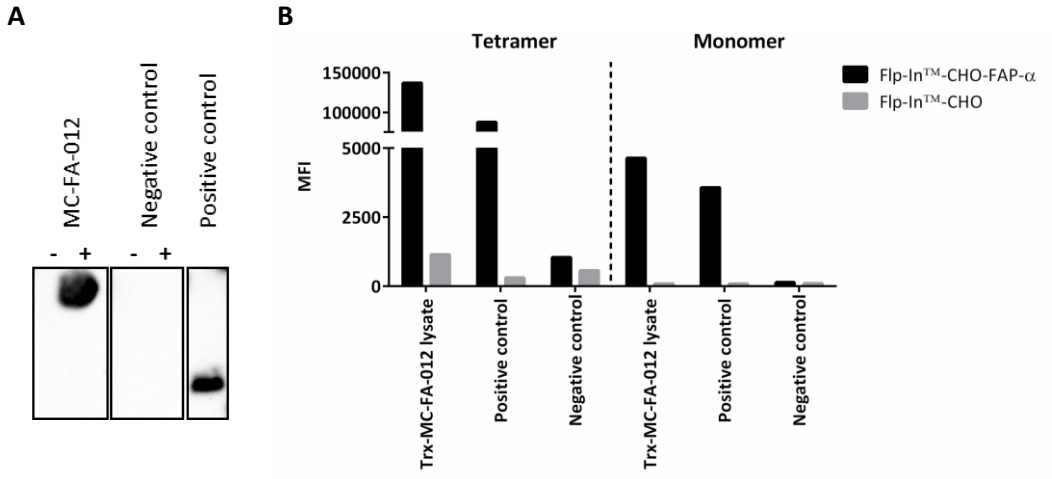


Figure 3.13 Analysis of biotinylated Trx-cystine-knot miniprotein fusions. (A) Successful *in vivo* biotinylation of Trx-cystine-knot miniprotein fusion. For western blot analysis 20 µL of protein containing bacterial lysate was applied and detected with a HRP-conjugated SA. The symbol "+" represents the presence and the symbol "–" indicates the absence of D-biotin during production. **(B)** Increased binding strength of tetramerized Trx-MC-FA-012 miniprotein fusion determined via flow cytometry. Tetramerization of protein was achieved by complexation with APC-conjugated SA. Monomeric and tetramerized Trx-MC-FA-012 fusion proteins were applied for binding analysis and binding was detected using APC-conjugated anti-His-tag antibody. The experiment was conducted with 4 µL bacterial lysate. Graph shows the MFI of APC signal.

The establishment of an efficient cell-based panning and downstream hit identification platform, which is suitable for selection of target-specific cystine-knot miniproteins, provides a novel opportunity to target membrane associated proteins in their natural conformation. A major progress of developing robust downstream methods has been made to enable high-throughput hit screening of selected cystine-knot miniproteins against target expressing cells and a control cell line. Furthermore, two different expression systems for cystine-knot miniprotein production in a 96-well scale allowed their application as monomeric or tetramerized protein directly from bacterial lysate for analysis of cellular binding capacity.

3.2 Novel cystine-knot miniprotein phage libraries

The open chain sequence of MCoTI-II was the basis for combinatorial phage library construction[144]. An already existing phage library (MCopt 1.0) comprised sequences with randomized amino acids in loop 1, scattered positions in the third loop and two variable residues in front of the first cysteine. In this library cystine-knot miniproteins are presented via the pVIII on the surface of filamentous M13 phages. This well characterized library represents the basis of our drug discovery platform for the development of novel cystine-knot miniprotein-based agents. In order to extent the internal library repertoire and to enhance ligand selection outcome, novel cystine-knot miniprotein phage libraries have been generated in this work applying an improved rational design strategy.
This chapter describes the design, construction and characterization of three novel cystine-knot miniprotein phage libraries. As a part of quality evaluation, one phage library was additionally screened against a highly tumor-related target protein (DSG-3) recombinantly presented on mammalian cells for cell-based screening and ligand identification. However, panning experiments revealed an increased occurrence of deletion mutants lacking a large segment comprising Jun-pIII´ and Fos-cystine-knot miniprotein gene fusions. This genetic instability of the pJuFo vector has been mentioned in a published doctoral thesis[145], but has besides not yet been described in the literature. At the end of this chapter a possible solution for this emerging issue is described that comprises a complete re-design of the phagemid vector and initial data showing feasibility of this new system.

3.2.1 Cystine-knot miniprotein phage library design

The generation of large and diverse libraries with correctly folded cystine-knot miniproteins as well as their phage surface presentation was based on the pJuFo phagemid vector[42]. Randomization of certain sequence positions of the oMCoTI-II scaffold was intended to gain high library diversity. The amino-terminal loop in the wildtype miniprotein represents the active site of its trypsin inhibitory function and has been shown to well tolerate engineering for target binding[88,91,92]. Therefore, one library (MCopt 2.0) was developed with amino acid exchanges and sequence length extension in loop 1. In a previous report, the authors showed a general possibility to modify EETI-II protein in sequence diversity and length variation in two loop positions without losing conformational properties and trypsin-binding ability. Both loop positions are not directly involved in trypsin interaction, but located in proximity to the active loop site[146]. Although EETI-II and MCoTI-II have sequence differences, both share high structural similarity and belong to the same cystine-knot miniprotein family[147]. Loop 3 in the EETI-II scaffold seemed to slightly tolerate randomization better than loop 2[146], which loop position is basically analog to loop 5 in MCoTI-II. Based on these observations, a second library (MCopt 2.1) was designed with loop 5 randomization and an additionally extended loop size in the oMCoTI-II scaffold. Furthermore, in another report an EETI-II protein was engineered with two separate integrin binding epitopes in the active loop and an adjacent loop 3 position. The mutated variants resulting from combinatorial yeast library selection showed proper folding to native functional knottin and high affinity binding capability to integrins[148]. Encouraged by their strategy, in the third library (MCopt 3.0) both adjacent loops 1 and 5 are randomized simultaneously in order to enlarge the potential binding surface of the oMCoTI-II protein and thus potentially increase its binding affinity and specificity. Figure 3.14 presents the three different library design strategies and the three-dimensional structure of the oMCoTI-II protein[149]. All three libraries comprise a completely randomized sequence of ten amino acid residues in loop 1, loop 5 or both loops, including the whole range of canonical nineteen amino acids, except cysteine. Apart from the randomized region the

wildtype scaffold framework was maintained in all libraries with the exception of MCopt 2.1, which contained lysine substitutions to alanine, isoleucine and arginine in loop 1. Furthermore, the basic scaffold comprises three modified amino acids in front of the first cysteine due to cloning reasons.

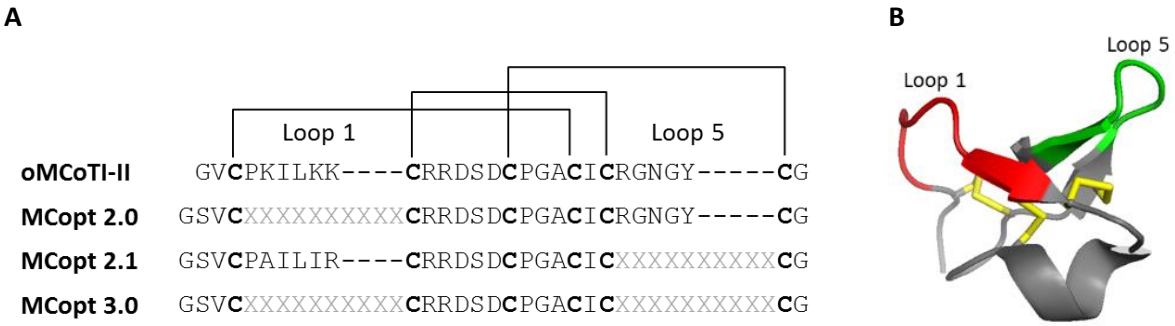


Figure 3.14 Library design strategy and three-dimensional structure of oMCoTI-II protein. (A) Schematic overview of the oMCoTI-II sequence framework and amino acid sequences of MCopt 2.0, MCopt 2.1 as well as MCopt 3.0 library. Bold letters represent amino acid cysteine and brackets indicate disulfide bonds between cysteines. Grey letters show randomized positions and X depicts a random amino acid, expect cysteine. **(B)** Three-dimensional structure of oMCoTI-II protein (PDB 2IT8)[149]. Surface exposed loop 1 is shown in red and loop 5 in green. Disulfide bonds are depicted as yellow sticks. The oMCoTI-II structure was obtained from RCBS PDB and illustrated using PyMol.

Phage libraries were constructed using a pJuFo phage display system to present cystine-knot miniproteins through a recombinant pIII´ and thus maintain a low presentation valency. More precisely, cystine-knot miniprotein sequences were genetically fused to *fos* gene and pIII´ to *jun* gene, respectively[42], allowing a simultaneous expression but autonomous folding of both fusions. After secretion of the fusions into the periplasmic space they assemble through the leucine zipper of heterodimeric Jun-Fos elements after individual processing. Subsequently, the recombinant coat protein fusions are incorporated into the phage envelope, leading to surface presentation of the encoded cystine-knot miniproteins.

3.2.2 Generation and characterization of three novel cystine-knot miniprotein phage libraries

For library generation, MCopt 2.0 and MCopt 2.1 DNA library fragments were directly synthesized with the intended randomized loop region. In contrast, MCopt 3.0 DNA library fragments were constructed based on the MCopt 2.0 and MCopt 2.1 DNA fragments using SOE PCR. As a start, the optimal conditions for amplifying MCopt 2.0 and MCopt 2.1 DNA fragments were analyzed to achieve high yields and likewise low amounts of side products. To this end, gradient PCR reactions were performed with annealing temperatures ranging from 53 °C to 64 °C in the presence of different betaine or DMSO concentrations. Figure 3.15 shows the analysis of yielded PCR products using agarose gel electrophoresis and two electropherograms of selected samples.

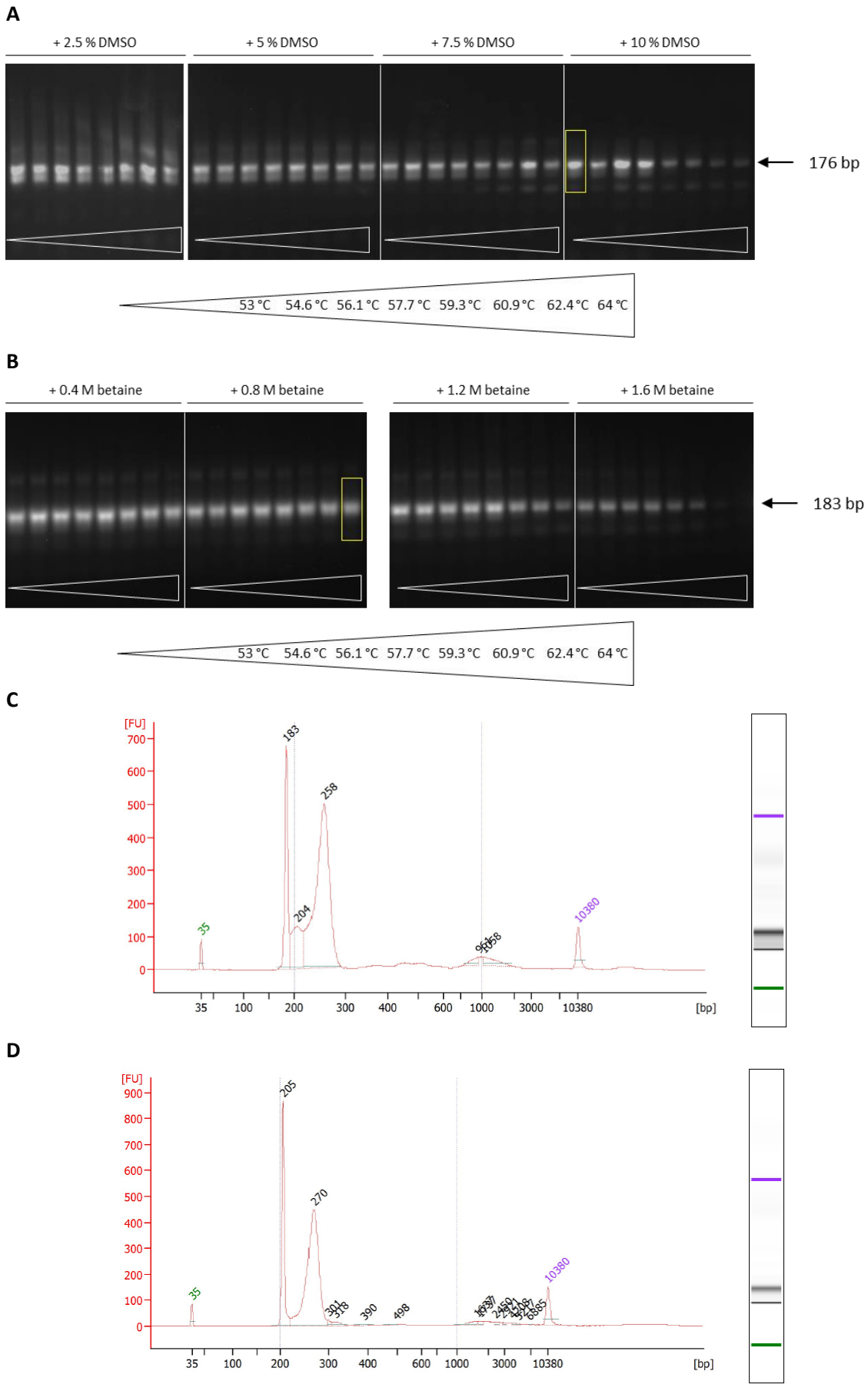


Figure 3.15 Determination of optimal PCR conditions for MCopt 2.0 and MCopt 2.1 DNA fragment pool amplification. Agarose gel electrophoretic analysis of MCopt 2.0 **(A)** and MCopt 2.1 **(B)** DNA fragment pools

amplified using different annealing temperatures. Synthetic MCopt 2.0 DNA fragment pool was amplified in the presence of 2.5 %, 5 %, 7.5 % or 10 % DMSO and synthetic MCopt 2.1 DNA fragment was amplified in the presence of 0.4 M, 0.8 M, 1.2 M or 1.6 M betaine at annealing temperatures of 53 °C, 54.6 °C, 56.1 °C, 57.7 °C, 59.3 °C, 60.9 °C, 62.4 °C and 64 °C (from left to right). Intended MCopt 2.0 DNA fragment size was 176 bp, while MCopt 2.1 DNA fragment had a length of 183 bp (depicted with an arrow). Yellow boxes indicate PCR products selected for detailed fragment size analysis. Bioanalyzer electropherograms of amplified MCopt 2.0 **(C)** and MCopt 2.1 **(D)** DNA fragment pools. Size distribution of PCR amplicons are shown in the electropherograms (left) and in the gel image (right). The left peak represents the lower marker with 35 bp and the right peak the upper marker with 10380 bp.

The addition of 2.5 % DMSO led to strong side products at all tested annealing temperatures during amplification of MCopt 2.0 DNA fragments, while the presence of 5 %, 7.5 % and 10 % DMSO resulted in less side products. The amplification of MCopt 2.1 DNA fragments with different concentrations of betaine showed in general a lower proportion of side products. Based on the visual appearance of separated PCR products on the agarose gel, a single sample of each library fragment pool was further assayed in size and quality via Bioanalyzer. The selected PCR condition for MCopt 2.0 DNA fragment amplification was 53 °C annealing temperature in the presence of 10 % DMSO and for MCopt 2.1 DNA fragment amplification with the addition of 0.8 M betaine and an annealing temperature of 64 °C.
MCopt 2.0 and MCopt 2.1 DNA amplicons were used as template in the SOE PCR to generate MCopt 3.0 DNA fragments. Again, different DMSO concentrations were added to the PCR samples and the annealing temperatures was varied in the gradient PCR. All PCR products were assessed on an agarose gel and one sample was applied on a Bioanalyzer (Figure 3.16).

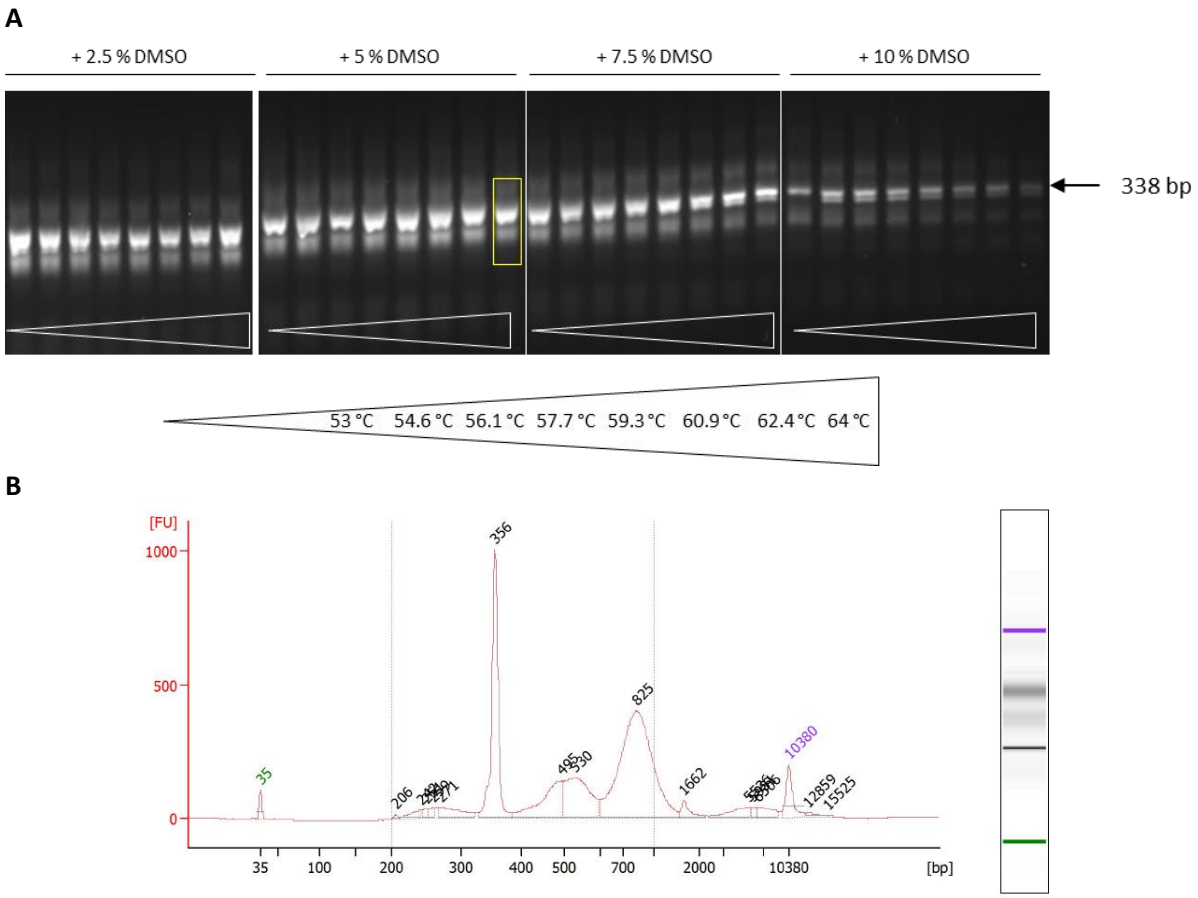


Figure 3.16 Determination of optimal PCR conditions for MCopt 3.0 DNA fragment pool amplification. (A) Agarose gel electrophoresis of MCopt 3.0 DNA fragments using SOE PCR. Amplified MCopt 2.0 and MCopt 2.1 DNA pool served as template to generate MCopt 3.0 DNA fragments. Gradient PCR was performed with annealing temperature of 53 °C, 54.6 °C, 56.1 °C, 57.7 °C, 59.3 °C, 60.9 °C, 62.4 °C and 64 °C (from left to right), in the presence of 2.5 %, 5 %, 7.5 % or 10 % DMSO. Yellow box indicates PCR sample chosen for detailed size analysis. **(B)** Bioanalyzer electropherogram of amplified MCopt 3.0 DNA fragments. PCR product was applied on a Bioanalyzer high sensitivity chip for analysis of size distribution. Results are depicted as electropherogram (left) and gel image format (right). The left peak shows the lower marker with 35 bp and the right peak the upper marker with 10380 bp.

The gelelectrophoretic separation indicates, besides the expected PCR amplicon at 338 bp, also many side products with different size distribution. The Bioanalyzer results show several peaks in the electropherogram, which indicated the presence of side products too and marginally lowered the product quality. Nevertheless, the selected condition for SOE PCR was the addition of 5 % DMSO and an annealing temperature of 64 °C.

After determining the optimal PCR conditions, a large amount of MCopt 3.0 DNA fragments was synthesized via SOE PCR. The synthetic MCopt 2.0 and MCopt 2.1 DNA fragments from GeneArt™ and the generated MCopt 3.0 DNA fragments were prepared for cloning. The restricted MCopt 2.0, MCopt 2.1 and MCopt 3.0 DNA fragments were cloned into the pJuFoIII phagemid vector and subsequently introduced into electrocompetent *E. coli* SS320 cells. Dilution plating of transformed cells was performed to count colony numbers and calculate the number of primary transformants (Figure 3.17 A). A primary library size of 1.06 x 10^9 cells for MCopt 2.0, 1.82 x 10^9 cells for MCopt 2.1 and 2.16 x 10^9 cells for MCopt 3.0 could be achieved.

A

Library	Number of primary transformants
MCopt 2.0	1.06×10^9
MCopt 2.1	1.82×10^9
MCopt 3.0	2.16×10^9

B

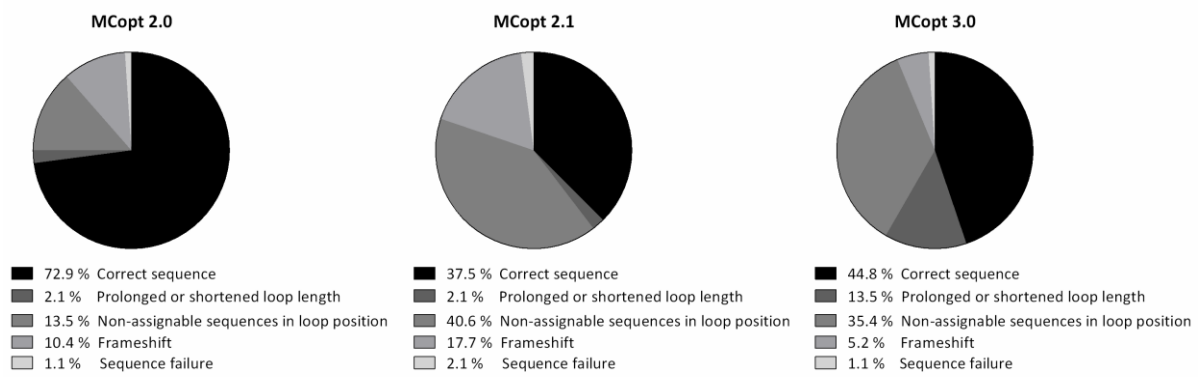


Figure 3.17 Primary MCopt 2.0, MCopt 2.1 and MCopt 3.0 library transformants and sequence composition after electroporation of *E. coli* SS320 cells. (A) Primary library size of MCopt 2.0, MCopt 2.1 and MCopt 3.0. *E. coli* SS320 colonies were counted after electroporation and dilution plating. **(B)** Sequence composition of the generated MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries. In total 96 clones were randomly selected for sequencing.

The quality of each library was determined by sequencing of 96 randomly chosen clones and evaluated as percentage of sequence composition (Figure 3.17 B). In general, sequences were assigned to five categories. In the MCopt 2.0 library, the largest proportion represented correct cystine-knot miniproteins sequences with 72.9 %, while correct MCopt 2.1 library sequences and MCopt 3.0 library sequences were detected to only 37.5 and 44.8 %, respectively. In all libraries, a low frequency of clones was observed with longer or shorter randomized loop sizes than originally designed, but their presence is not expected to influence library quality. Furthermore, a proportion between 5.2 % and 17.7 % of sequences contained frameshift mutations and < 2.1 % revealed sequence failure. Surprisingly, many clones showed non-assignable sequences in the randomized loop region, while the conserved scaffold framework was sequenced correctly. These clones were especially observed in MCopt 2.1 library at 40.6 % and in MCopt 3.0 at 35.4 %. Curiosity about their occurrence led to a deeper analysis of the respective clones, because a high frequency of incorrect clones decreases the library quality and might deteriorate ligand selection outcome.

Non-assignable sequences could be a sign of poor sequencing quality, in which a clear assignment of one nucleotide within the analyzed DNA strand was missing. As a consequence, translation into amino acid level led to an X at the corresponding position. Interestingly, these events predominantly occurred in randomized loop positions (Figure 3.17 B). In order to exclude general sequencing failure, several clones were repeatedly sequenced. Table 3.1 shows representative amino acid sequences of MCopt 2.1 library clones. The data shows that the non-assignable residues mainly appeared in loop 5 with comparative sequence pattern of the same clone. Such occurrence might be attributed to double or multiple transformed *E. coli* cells[150], which would have a severe negative impact on the library quality, since this would impede the fundamental linkage of genotype and phenotype of the generated recombinant phages.

Table 3.1 Amino acid sequences of MCopt 2.1 library clones after electroporation of *E. coli* SS320 cells. Sequencing of the same clone was performed independently. X in bold letters counts for a non-assignable residue.

Sequencing run	Clone number	Sequences
1	Clone 1	GSVCPAILIRCRRDSDCPGACIC**XXXXXXXX**R**X**CG
2		GSVCPAILIRCRRDSDCPGACIC**XXXXXXXXXX**CG
1	Clone 2	GSVCPAILIRCRRDSDCPGACIC**X**TWY**XXXXXX**CG
2		GSVCPAILIRCRRDSDCPGACIC**XXXXXXXX**TTCG
1	Clone 3	GSVCPAILIRCRRDSDCPGACICW**XXXXXXXX**LCG
2		GSVCPAILIRCRRDSDCPGACICWVES**XXX**D**XX**CG

Sblattero and colleagues described a method to expand diversity of conventional antibody phage libraries. This involved recombination of VH and VL genes located on two separate phagemid vectors in a single bacteria after infection at a high MOI[151]. In contrast, an infection at a low MOI lead to entering of only one phagemid per cell. In order to reduce the strong occurrence of clones carrying more than one phagemid vector and restore the genotype-phenoptype linkage, phage particles of the primary MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries were produced and *E. coli* XL1-Blue cells were infected using a low MOI. Library titers after infection showed cell numbers between 1.54 x 10^9 and 8.72 x 10^9 (Figure 3.18 A), which was nearly similar to electroporated *E. coli* SS320 and for MCopt 2.0 even higher. Consequently, infection approach caused no significant changes in library size.

A

Library	Number of infected cells
MCopt 2.0	8.72 x 10^9
MCopt 2.1	1.54 x 10^9
MCopt 3.0	2.22 x 10^9

B

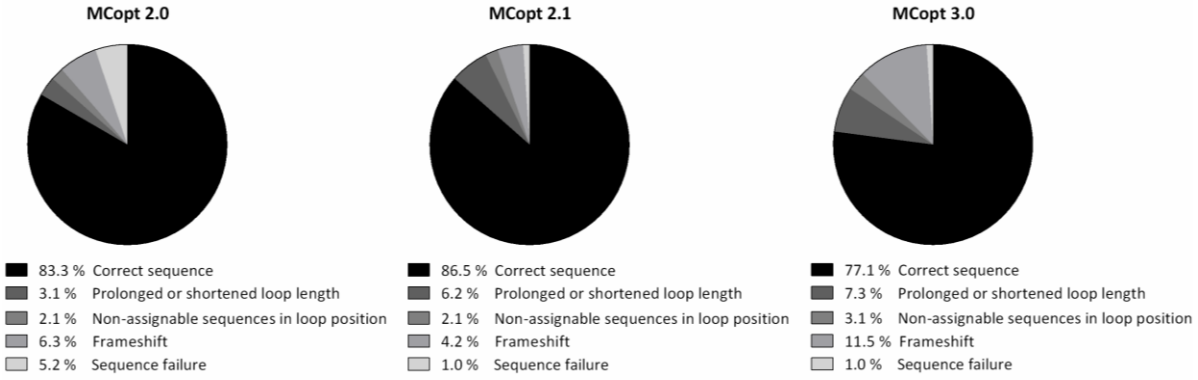


Figure 3.18 Properties of the final MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries after infection of *E. coli* XL1-Blue cells. (A) Library titers of MCopt 2.0, MCopt 2.1 and MCopt 3.0 after infection of *E. coli* XL1-Blue cells. Colony numbers of dilution plating was counted. **(B)** Sequence composition of MCopt 2.0, MCopt 2.1 and MCopt 3.0 library after infection of *E. coli* XL1-Blue cells with a low MOI. Sequencing of 96 randomly chosen clones was performed.

Sequencing of individual clones enabled the assessment of library status on molecular level. Figure 3.18 B presents the sequencing outcome of the final libraries after re-infection with a low MOI. A small proportion of sequences with prolonged or shortened loop length and frameshifts as well as clones with sequence failure still remained. But importantly, the frequency of correct cystine-knot

miniprotein sequences increased in all libraries and the non-assignable sequences in the loop positions likewise decreased dramatically. Particularly, for the MCopt 2.1 (86.5 %) and the MCopt 3.0 library (77.1 %) the percentage of correct sequences were much higher compared to the primary *E. coli* SS320 library (Figure 3.17 B). The reduction of clones carrying more than one phagemid vector should result in enormously improved library qualities.

Displaying of functional cystine-knot miniprotein on the phage surface represented another important quality control of the generated libraries. Based on our knowledge, no suitable antibody for the detection of oMCoTI-II exists but the fusion of the cystine-knot miniprotein to Fos enables an indirect detection using anti-Fos antibody. In a sandwich ELISA setting MCopt 2.0, MCopt 2.1 and MCopt 3.0 library phage pools showed higher signals than negative control phages (Figure 3.19 A). Furthermore, individual phages of each library were randomly selected for analysis (Figure 3.19 B). The majority of single phages revealed strong signals, even though their strength varied. These findings clearly proved a Fos presentation on the phage envelope and thus indirectly indicated the display of the fused cystine-knot miniprotein as well. The expression of Fos on library phages was additionally determined in a western blot analysis as depicted in Figure 3.19 C. The Fos-cystine-knot miniprotein complex alone was not observed. But in all novel libraries a band was detected at around 37 kDa, whose size could be referred to assembled Fos-cystine-knot miniprotein and Jun-pIII´ complex. Furthermore, a distinct band above the protein complex at around 48 kDa occurred, which appearance could not be clearly assigned to proteins with expected molecular size. This band appeared in all three libraries and may result due to additional protein interaction of the Fos-cystine-knot miniprotein and Jun-pIII´ complex.

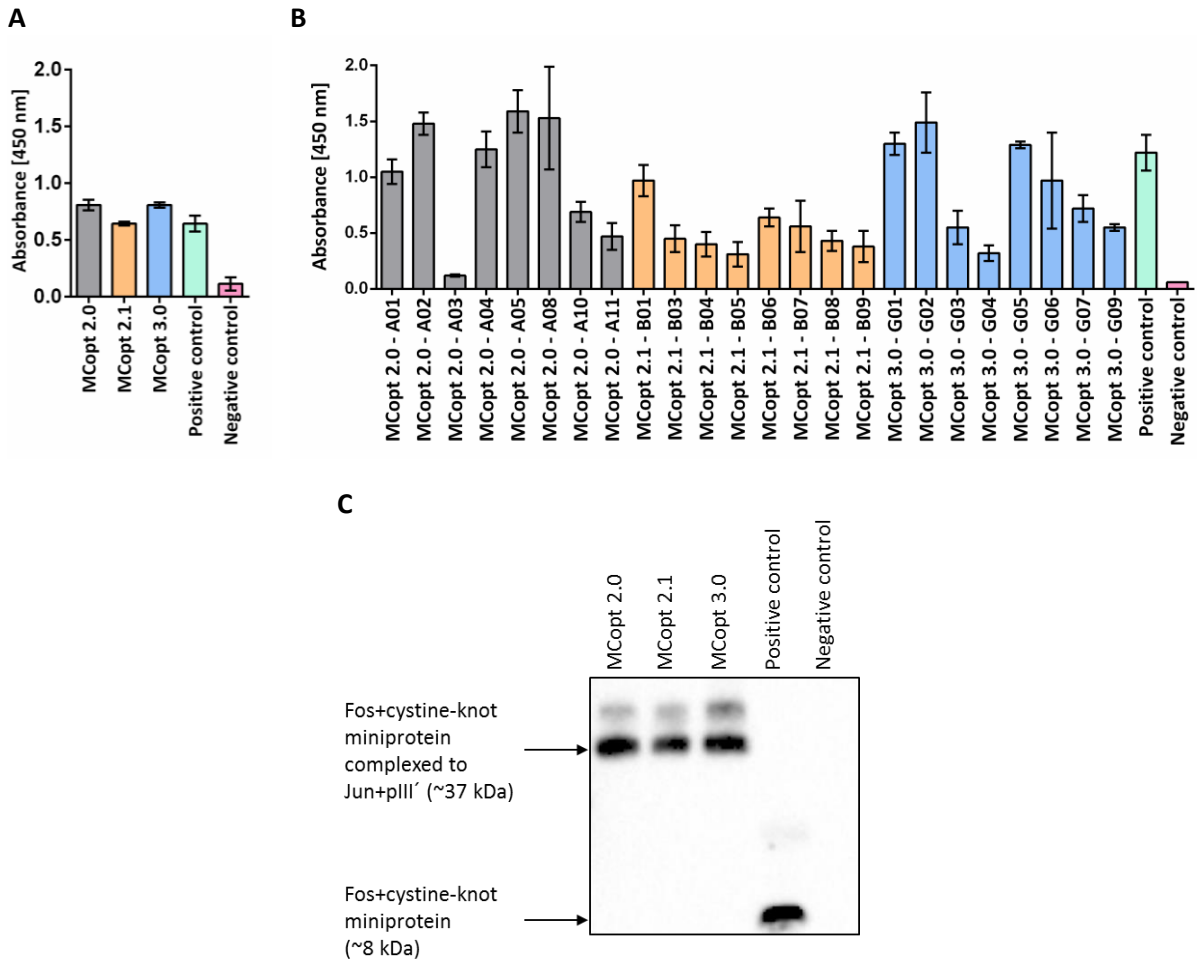


Figure 3.19 Analysis of Fos expression. Surface presentation of Fos-cystine-knot-miniproteins on library phage pools **(A)** and single phages **(B)**. Phages of MCopt 2.0, MCopt 2.1 and MCopt 3.0 library were applied at 1 x 10^{12} particles for ELISA. Binding of anti-Fos antibody to phages was detected using HRP-conjugated anti-rabbit antibody. ELISA was performed in triplicates ± SD. **(C)** Western Blot analysis of Fos expression on MCopt 2.0, MCopt 2.1 and MCopt 3.0 library phages. Western blot was performed with 5 x 10^{11} phage particles for binding analysis of anti-Fos antibody and detection with HRP-conjugated anti-rabbit antibody. For all analysis MCopt 1.0 library phages served as positive control and M13 wildtype phages as negative control (performed by Tanja Ziß).

3.2.3 Characterization of cystine-knot miniproteins on protein level

To determine whether loop randomization and length extension has an impact on the biophysical properties of cystine-knot miniproteins, a producibility and folding analysis was performed. For the producibility study a 96-well mini-scale format was applied, as this method is also part of the hit identification procedure within the in-house drug discovery platform. Furthermore, single cystine-knot miniproteins were expressed via large-scale production as well and purified to analyze their protein mass as a measure of proper folding.

The overall producibility of Trx-cystine-knot miniprotein fusions resulting from MCopt 2.0, MCopt 2.1, and MCopt 3.0 library was analyzed in a 96-well production for yield determination. For this purpose, cystine-knot miniprotein sequences of library pools were cloned into pET-32-LibEx_GS expression vector and subsequently introduced into SHuffle® T7 Express competent *E. coli* cells. Single clones of each library were randomly picked for expression in 1 mL volume. Bacterial supernatant after heat-treatment, containing roughly purified Trx-cystine-knot miniprotein fusions, was applied on an E-PAGE™ high-throughput gel system (supplementary Figure 5.2). The concentration of each clone was calculated based on known protein references. For evaluation the proportional distribution of yielded

concentrations were assigned into three categories of < 1 mg/mL, 1 – 1.49 mg/mL and > 1.5 mg/mL (Figure 3.20). MCopt 1.0 library clones were considered as reference and their protein expression mainly resulted in concentrations between 1 – 1.49 mg/mL with a proportion of 74.2 %. Trx-cystine-knot miniprotein fusions below 1 mg/mL and above 1.5 mg/mL were represented at 9.7 % and 16.1 %, respectively. The distribution of MCopt 2.0, MCopt 2.1, and MCopt 3.0 library clones differed. A high frequency between 49.5 % and 68.8 % had concentrations below 1 mg/mL, whereas nearly no or a very small number of clones showed yields above 1.5 mg/mL. In the range between 1 – 1.49 mg/mL were library clones at 31.2 % to 49.5 %. Especially the average yield of MCopt 2.1 and MCopt 3.0 library clones were weaker in comparison to MCopt 2.0 library clones.

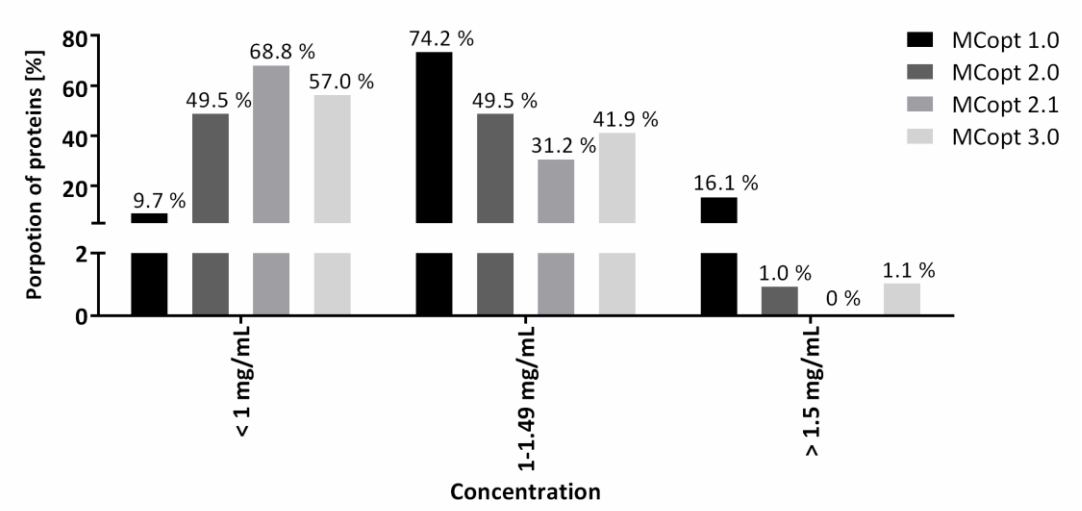


Figure 3.20 Proportional distribution of yielded Trx-cystine-knot miniprotein fusion concentrations after expression in 96-well format. Protein concentrations of MCopt 2.0, MCopt 2.1 and MCopt 3.0 library clones were determined densitometrically based on a protein reference (supplementary Figure 5.2). The data set of 93 analyzed clones was categorized in concentrations of < 1 mg/mL, 1 – 1.49 mg/mL and > 1.5 mg/mL.

In order to evaluate correct folding of the cystine-knot miniproteins derived from MCopt 1.0, MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries, individual clones were randomly picked for large-scale production. Expressed Trx-cystine-knot miniprotein fusions were purified via heat-treatment and IMAC. The fusion protein-tag was cleaved from the cystine-knot miniprotein part and separated using RPC. The purified cystine-knot miniprotein was analyzed in mass spectrometry. A detection of the expected cystine-knot miniprotein mass with all three disulfide bonds accounted as correctly folded protein (Table 3.2). All three analyzed MCopt 1.0 library clones revealed proper protein mass. Whereas in one out of five MCopt 2.0 library clones, one out of four MCopt 2.1 library clones and one out of three MCopt 3.0 library clones the correct mass was not observed.

Table 3.2 Mass spectrometry based folding analysis of individual MCopt 1.0, MCopt 2.0, MCopt 2.1, and MCopt 3.0 library clones. Expected mass of two-, three- and four-fold charged cystine-knot miniprotein (2 H^+, 3 H^+, 4 H^+) and the actual detected mass resulting from mass spectrometry. Protein sequences of all analyzed library clones are listed in the supplementary Table 5.3.

Library clone	Molecular mass of cystine-knot miniprotein	Expected cystine-knot miniprotein mass	Detected cystine-knot miniprotein mass
MCopt 1.0 clone 1	3270.51 Da	1636.26 (2 H^+) 1091.18 (3 H^+) 818.63 (4 H^+)	1090.90 (3 H^+)
MCopt 1.0 clone 2	3771.13 Da	1886.57 (2 H^+) 1258.05 (3 H^+) 943.79 (4 H^+)	1887.30 (2 H^+) 1257.90 (3 H^+) 943.70 (4 H^+)
MCopt 1.0 clone 3	4231.57 Da	2116.79 (2 H^+) 1441.53 (3 H^+) 1058.90 (4 H^+)	2116.7 (2 H^+) 1411.2 (3 H^+)
MCopt 2.0 clone 1	3784.22 Da	1893.12 (2 H^+) 1262.41 (3 H^+) 947.06 (4 H^+)	1893.10 (2 H^+) 1262.10 (3 H^+)
MCopt 2.0 clone 2	3866.28 Da	1934.15 (2 H^+) 1289.77 (3 H^+) 967.58 (4 H^+)	1934.2 (2 H^+) 1289.40 (3 H^+)
MCopt 2.0 clone 3	3767.26 Da	1884.61 (2 H^+) 1256.75 (3 H^+) 942.81 (4 H^+)	1884.6 (2 H^+) 1256.40 (3 H^+)
MCopt 2.0 clone 4	3822.19 Da	1912.10 (2 H^+) 1275.07 (3 H^+) 956.56 (4 H^+)	-
MCopt 2.0 clone 5	3585.96 Da	1793.99 (2 H^+) 1196.33 (3 H^+) 897.50 (4 H^+)	1196.20 (3 H^+)
MCopt 2.1 clone 1	3885.54 Da	1943.78 (2 H^+) 1296.19 (3 H^+) 972.39 (4 H^+)	-
MCopt 2.1 clone 2	3962.54 Da	1982.28 (2 H^+) 1321.86 (3 H^+) 991.64 (4 H^+)	1981.90 (2 H^+) 1321.60 (3 H^+) 991.50 (4 H^+)
MCopt 2.1 clone 3	3885.48 Da	1943.75 (2 H^+) 1296.17 (3 H^+) 972.38 (4 H^+)	1943.20 (2 H^+) 1295.80 (3 H^+)
MCopt 2.1 clone 4	4008.53 Da	2005.27 (2 H^+) 1337.18 (3 H^+) 1003.14 (4 H^+)	2005.6 (2 H^+) 1336.80 (3 H^+)
MCopt 3.0 clone 1	4428.98 Da	2215.50 (2 H^+) 1447.34 (3 H^+) 1108.25 (4 H^+)	2215.40 (2 H^+) 1477.30 (3 H^+) 1108.00 (4 H^+)
MCopt 3.0 clone 2	4409.02 Da	2205.52 (2 H^+) 1470.68 (3 H^+) 1103.26 (4 H^+)	-
MCopt 3.0 clone 3	4568.12 Da	2285.07 (2 H^+) 1523.71 (3 H^+) 1143.04 (4 H^+)	2284.70 (2 H^+) 1523.00 (3 H^+)

The overall performance of MCopt 2.0, MCopt 2.1, and MCopt 3.0 library clones was slightly weaker in comparison to clones of the established MCopt 1.0 library. Nevertheless, the 96-well expression analysis showed producibility of all cystine-knot miniprotein fusions, but MCopt 2.0 library clones with general better expression yields of all new libraries. The rough analysis of correct cystine-knot miniprotein folding furthermore enabled insights into protein level with no major differences in all libraries.

3.2.4 Panning of MCopt 2.0 library against a novel target-overexpressing cell line

The functionality of the MCopt 2.0 phage library was in the following analyzed in a selection experiment using a tumor-associated antigen as a target. DSG-3, a transmembrane adhesion protein of desmosomes[152], was defined as target protein for a cell-based panning approach. Initially, a stable CHO-K1 cell line that overexpresses DSG-3 was constructed with a PiggyBac transposon system. Transfected cells were selected through hygromycin resistance. FACS of transfected cells allowed an isolation of strong DSG-3 expressing clones to gain a monoclonal population, which was designated as CHO-K1-DSG-3. The expression of DSG-3 was determined by flow cytometry as shown in Figure 3.21.

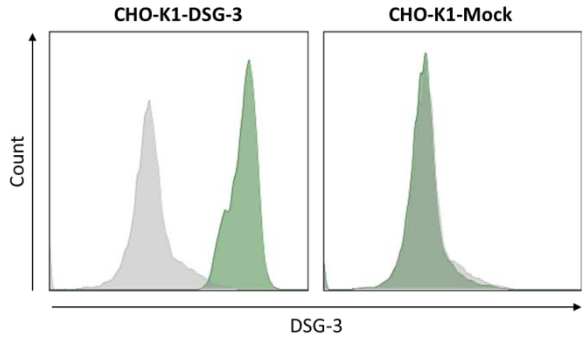


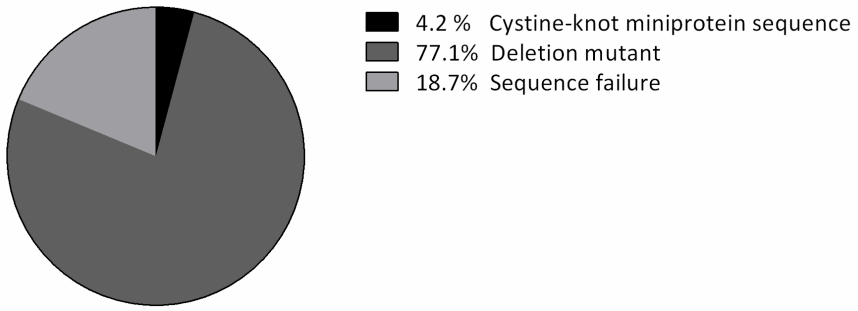
Figure 3.21 DSG-3 expression on CHO-K1-DSG-3 and CHO-K1-Mock cells. Analysis of DSG-3 expression was determined via flow cytometry with an anti-human-DSG-3 antibody and detected with a Cy5-conjugated anti-mouse antibody (green histograms). Grey histograms indicate staining with Cy5-conjugated anti-mouse antibody. Representative histograms are shown.

A Screening of the MCopt 2.0 library was performed in three successive rounds against CHO-K1-DSG-3 cells based on the experimental outline given in Figure 3.2. In the first panning round 2.4×10^{13} MCopt 2.0 library phages were applied and 2.4×10^{12} amplified phage particles in the second and third round. During each panning round phage titers were determined for input, post-input and output monitoring via CFU counting (Figure 3.22 A).

A

Panning round	Input	Post-input	Output	Post-input/output ratio
1	2.78×10^{12}	4.90×10^{12}	4.40×10^{4}	1.11×10^{8}
2	2.59×10^{12}	1.68×10^{12}	4.90×10^{5}	3.42×10^{6}
3	8.85×10^{11}	1.15×10^{12}	4.99×10^{6}	2.11×10^{5}

B



C

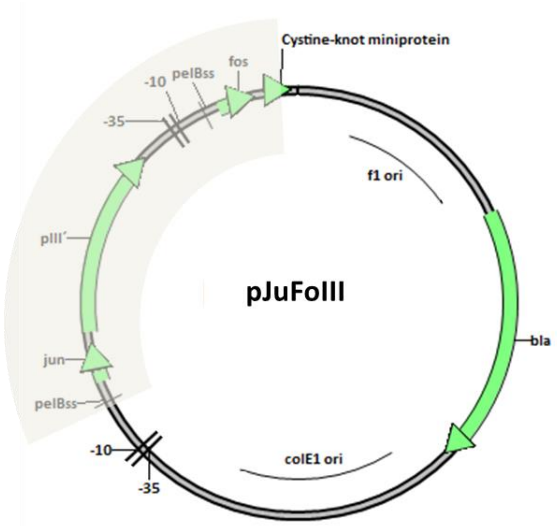


Figure 3.22 Phage titers and sequence compostion after MCopt 2.0 panning against CHO-K1-DSG-3 cells. **(A)** Phage titers of MCopt 2.0 library panning rounds against CHO-K1-DSG-3 cells. Input, post-input and output was determined based on CFU counting at each panning round. **(B)** Sequence composition resulting from the third panning round against CHO-K1-DSG-3 cells. Sequencing of 48 clones was performed. **(C)** Schematic overview of molecular gene elements in the observed deletion mutants. Grey area represents lacking gene segment of the deletion mutants. Genetic elements: colE1 ori: *E. coli* origin of replication; f1 ori: phage origin of replication; pelBss: pelB periplasmic leader sequence; bla: beta-lactamase; -10/-35: promotor elements; fos/jun: leucine zipper domains; pIII´: gene encoding for a truncated recombinant minor coat protein of M13.

The input titer at every panning round was slightly lower as originally intended. The post-input was unexpectedly higher in the first and third round and a possible indication of insufficient negative depletion. However, the post-input to output ratio decreased over the successive panning rounds, which suggested an enrichment of a specific phage population. But to gain precise insights on a molecular level, 48 output clones of the third panning round were sequenced.

Classification of sequence composition into three categories is depicted in Figure 3.22 B. Surprisingly, only 4.2 % of total clones revealed to be cystine-knot miniprotein sequences. A proportion of 18.7 % showed sequence failure for unknown reasons and a considerable high proportion of 77.1 % showed no cystine-knot miniprotein related sequences at all (deletion mutant). Sequence alignments of the

respective clones with the pJuFo phagemid vector revealed an absence of the complete Fos-cystine-knot miniprotein and Jun-pIII´ gene cassettes as well as all segments in between, which is schematically presented in Figure 3.22 C.
In order to gain an understanding of the reasons for this strong deletion mutant enrichment, the bacterial growth behavior of the *E. coli* XL1-Blue MCopt 2.0 library pool and the *E. coli* XL1-Blue deletion mutant clone (carrying the mutated pJuFo phagemid vector) was monitored over a period of 300 min (Figure 3.23). The deletion mutant grew considerably faster compared to the MCopt 2.0 library pool under the same cultivation conditions. The occurrence and enrichment of the deletion clones is highly problematic because selection of potential binders might be suppressed by the fast-growing mutants. The dominant incidence of deletion mutants was also observed in several other screening experiments using diverse target proteins and under different selection conditions.

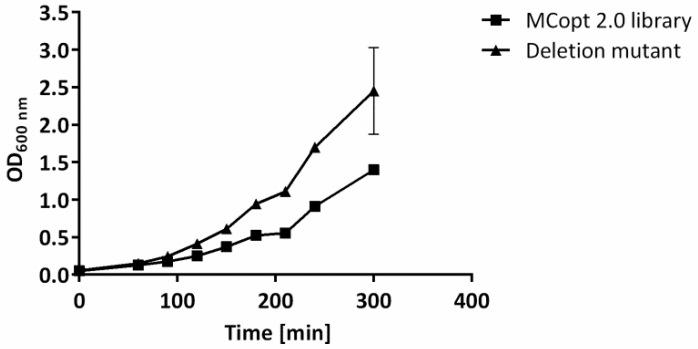
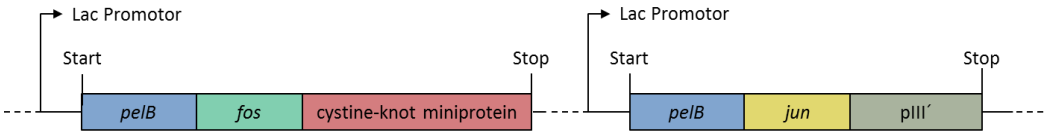


Figure 3.23 Growth comparison of *E. coli* XL1-Blue MCopt 2.0 library pool and *E. coli* XL1-Blue deletion mutant. Cells were cultured in LB-medium supplemented with 100 µg/mL carbenicillin and 0.4 % (w/v) glucose at 37 °C. OD of bacterial culture was measured at different time points. Experiment was performed in duplicates ± SD.

3.2.5 Optimization of MCopt 2.0 phage library

As shown in chapter 3.2.4, the dominant occurrence of deletion mutants negatively affected screening outcome and represented a major interference factor. Based on the experimental results an exact clarification of the deletion mutant origin remained to be challenging, since they could in principle either be already present in the MCopt 2.0 library repertoire from the start or might have developed via unknown recombination events during the panning procedure.
The pJuFoIII phagemid vector contains tandem repeats in the promotor and the pelB leader sequences of Fos-cystine-knot miniprotein and Jun-pIII´ gene cassettes (Figure 3.24 A), which may be prone to homologous rearrangement events[153]. In order to proof this hypothesis and to overcome deletion mutant development an alternative phagemid vector (pPDIII-1) was designed and constructed. In the new vector design, the expression cassettes were genetically remodeled aiming to avoid repetitive sequences, while retaining the same molecular pathways and protein features. As schematically depicted in Figure 3.24 B, the new pPDIII-1 vector has, in contrast to pJuFo, a monocistronic structure where both gene fusions are driven by a single lac promotor sequence only separated by a shine dalgarno sequence in front of the second part. This feature and the exchange of one of the two pelB leader sequence to an ompA leader completely removed repetitive sequence parts from the phagemid. Additionally, new genetic elements were introduced (the lacI repressor gene, a CAP binding site and additional tHP terminator sequence 3´ and 5 ´of the operon) that should result in a tighter gene expression control of the recombinant fusions and thus should minimize potential growth difference of individual clones.

A



B

Figure 3.24 Schematic overview of expression cassettes. (A) Expression cassette of pJuFoIII phagemid vector. **(B)** Expression cassette of pPDIII-1 phagemid vector. *pelB*: pelB periplasmic leader sequence; *ompA*: ompA periplasmic leader sequence; *fos*/*jun*: leucine zipper domains; pIII´: recombinant minor coat protein; CAP: catabolite activator protein binding region; start and stop represent ATG and TAA codon, respectively. The complete vector maps of pJuFoIII phagemid vector and pPDIII-1 phagemid vector is shown in chapter 2.1.6.

The benefit of the designed new phagemid vector in preventing deletion mutant appearance was analyzed by construction of a small MCopt 2.0 sub-library and a test panning against a soluble tumor-associated antigen. The MCopt 2.0 sub-library was generated by cloning synthetic MCopt 2.0 DNA library fragments into the pPDIII-1 phagemid vector and a following electroporation of electrocompetent *E. coli* XL1-Blue cells. The number of transformed cells was determined using dilution plating and counting CFU (Figure 3.25 A).The primary size of the pPDIII-1-MCopt 2.0 sub-library was 8.9 x 10^5 cells.

A small number of 12 clones were sequenced to enable insights into molecular library composition (Figure 3.25 B). The pPDIII-1-MCopt 2.0 sub-library revealed 83.3 % with correct and unique cystine-knot miniproteins sequences, while 16.7 % represented frameshift sequences.

A

Sub-library	Number of primary transformants
MCopt 2.0	8.9 x 10^5

B

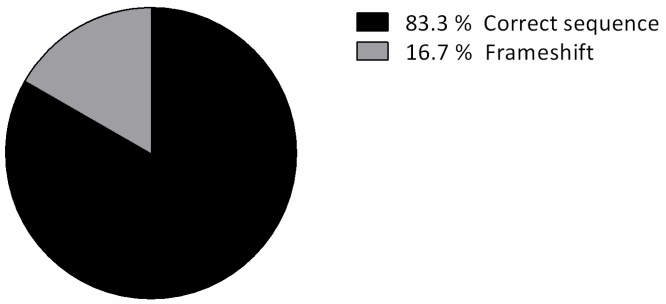


Figure 3.25 pPDIII-1-MCopt 2.0 sub-library size and sequence composition. (A) pPDIII-1-MCopt 2.0 sub-library titer. Colonies of electroporated *E. coli* XL1-Blue forming units were counted after dilution plating. **(B)** pPDIII-1-MCopt 2.0 sub-library sequences. Evaluation was based on 12 sequences.

Surface display of cystine-knot miniproteins on pPDIII-1-MCopt 2.0 sub-library phages was indirectly evaluated with an anti-Fos antibody in an ELISA-based set up (Figure 3.26 A). The pPDIII-1-MCopt 2.0 sub-library pool showed a stronger signal than the negative control, but a slightly weaker signal in comparison to the pJuFoIII-MCopt 2.0 library phages indicating a marginally lower surface density of the Fos-cystine-knot-miniprotein / Jun-pIII´ complex.

A previous experiment showed faster growth of the *E. coli* XL1-Blue deletion mutant in comparison to the *E. coli* XL1-Blue MCopt 2.0 library pool (Figure 3.23). Figure 3.26 B presents a similar experiment, in which no significant growth difference of *E. coli* pPDIII-1-MCopt 2.0 sub-library pool and *E. coli* XL1-Blue deletion mutant was observed, especially in the early time frame. *E. coli* XL1-Blue pJuFoIII-MCopt 2.0 library pool remained to have a delayed growth.

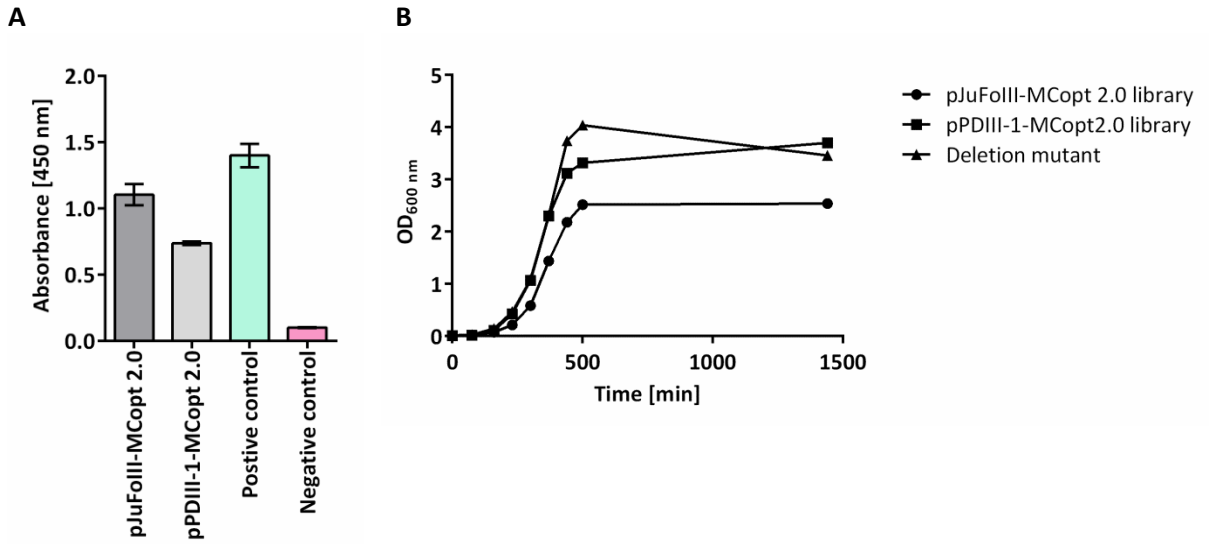


Figure 3.26 Phage surface presentation analysis and bacterial growth properties of the pPDIII-1-MCopt 2.0 sub-library. (A) Phage surface display analysis of Fos-cystine-knot-miniprotein on pJuFoIII-MCopt 2.0 and pPDIII-1-MCopt 2.0 sub-library pools. ELISA was performed with 1 x 10^{12} phage particles for binding analysis with anti-Fos antibody and subsequent detection with HRP-conjugated anti-rabbit antibody. Assay was conducted in duplicates ± SD. **(B)** Bacterial growth curves of *E. coli* XL1-Blue pJuFoIII-MCopt 2.0 library, pPDIII-1-MCopt 2.0 sub-library and the pJuFo derived deletion mutant. OD of bacterial culture were determined at different time points during cell cultivation in LB-medium supplemented with 100 µg/mL carbenicillin at 37 °C.

The novel phagemid vector was constructed with the aim to prevent the occurrence and the enrichment of deletion mutants during phage selection experiments. To experimentally evaluate an advantage of the new system, a comparative panning of the pJuFoIII-MCopt 2.0 library and the pPDIII-1-MCopt 2.0 sub-library against soluble FN-B target protein was performed for direct comparison of the screening outcome. Both libraries were screened in three successive panning rounds with 1 x 10^{13} applied phages at each round. Infectious phage input and output titers were determined during every panning round using CFU counting (Figure 3.27 A).

A

	Panning round	Input	Output	Input/output ratio
pJuFoIII-MCopt 2.0 library	1	8.40×10^{11}	2.97×10^{7}	2.80×10^{4}
	2	6.55×10^{11}	1.00×10^{4}	6.55×10^{7}
	3	2.55×10^{11}	1.00×10^{4}	2.55×10^{7}
pPDIII-1-MCopt 2.0 sub-library	1	1.64×10^{12}	1.17×10^{7}	1.40×10^{5}
	2	8.35×10^{11}	1.00×10^{4}	8.35×10^{7}
	3	8.10×10^{11}	1.00×10^{4}	8.10×10^{7}

B

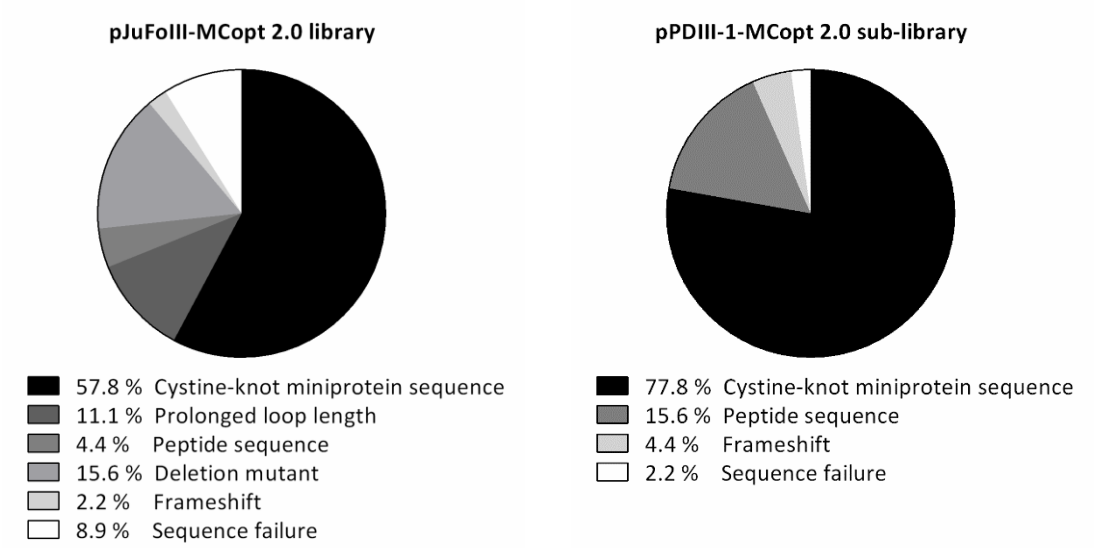


Figure 3.27 Analysis of phage titers and sequence composition after a comparative selection experiment against FN-B target protein. (A) Phage titers of screen pools resulting from pJuFoIII-MCopt 2.0 library and pPDIII-1-MCopt 2.0 sub-library panning against soluble FN-B target. Input and output was calculated within every panning round by determining CFU. **(B)** Composition of screen pool sequences after the third panning round of pJuFoIII-MCopt 2.0 library and pPDIII-1-MCopt 2.0 sub-library screening against soluble FN-B target. For sequence analysis 45 clones of each library screening approach were randomly chosen.

Input titers were obviously lower than 1×10^{13} applied phages, indicating a proportion of non-infectious phages. The output titer decreased throughout panning rounds in both library screenings. Sequencing of 45 clones after panning completion allowed monitoring of sequence compositions and particularly of the presence of deletion mutants (Figure 3.27 B). In both screening approaches different frequencies of correct cystine-knot miniprotein sequences, peptide sequences, frameshift sequences and clones with sequence failure were observed. The pJuFoIII-MCopt 2.0 library pool additionally revealed a small proportion of sequences with prolonged loop length. But importantly, deletion mutants were detected to a proportion of 15.6 % in the case of the pJuFoIII-MCopt 2.0 library screening, whereas no mutant was observed in the pPDIII-1-MCopt 2.0 sub-library pool at all.
In summary, the usage of the novel vector system seems to prevent either the formation of the deletion mutants or its enrichment during selection. The described selection experiment thus provides strong evidence for the successful remodeling of the pJuFo phagemid vector.

3.3 Targeting of fibronectin extra domain B (EDB)

This section describes the development of a cystine-knot miniprotein lead candidate targeting EDB for tumor imaging approaches. The target-specific cystine-knot miniprotein (MC-FN-010) was selected based on conventional phage display technology and initial *in vitro* target specificity was analyzed. The discovery of EDB binding cystine-knot miniproteins was performed by Nadja Salomon during her master thesis. Subsequently, in order to characterize MC-FN-010 more deeply, *in vitro* assays for revealing the target relevant binding site and affinity determination were applied. Furthermore, preclinical *in vivo* studies with two lead structures demonstrated their ability to detect EDB-positive tumors in a U-87 MG based xenograft glioblastoma mouse model.

The MCopt 1.0 library and MCopt 2.0 library were screened separately for cystine-knot miniprotein identification against single domain FN-B. After three selection rounds of each library, six different cystine-knot miniprotein sequences could be enriched. Interestingly, five out of six amplified sequences comprise a common R-I/V-R-(L) motif at the C-terminal end of loop 1. Only those five cystine-knot miniproteins showed binding capability to EDB presented in different formats with low off-target (FN-6789) interaction as tested in ELISA-based set ups. Furthermore, binding ability of tag-free proteins was determined by SPR analysis. Surprisingly, only one cystine-knot miniprotein (MC-FN-010) revealed strong binding towards FN-67B89 target protein across all five candidates in the tested concentration range from 50 – 1000 nM. Therefore, MC-FN-010 was chosen as lead candidate for further analysis and optical imaging probe development.

3.3.1 Dependency of EDB binding on the R-I/V-R-(L) motif and further sequence optimization of the selected cystine-knot miniprotein

The occurrence of the amino acid motif R-I/V-R-(L) in the enriched and specifically binding cystine-knot miniproteins suggested involvement of this sequence in EDB binding. We performed a systematic alanine scanning mutagenesis of MC-FN-010, in which every single amino acid in the variable region and the arginine residues in loop 2 and 5 were individually exchanged for alanine (consecutively numbered from MC-FN-011 to MC-FN-0114 as shown in supplementary Table 5.4). All constructs were tested for binding to single domain FN-B (Figure 3.28 A).

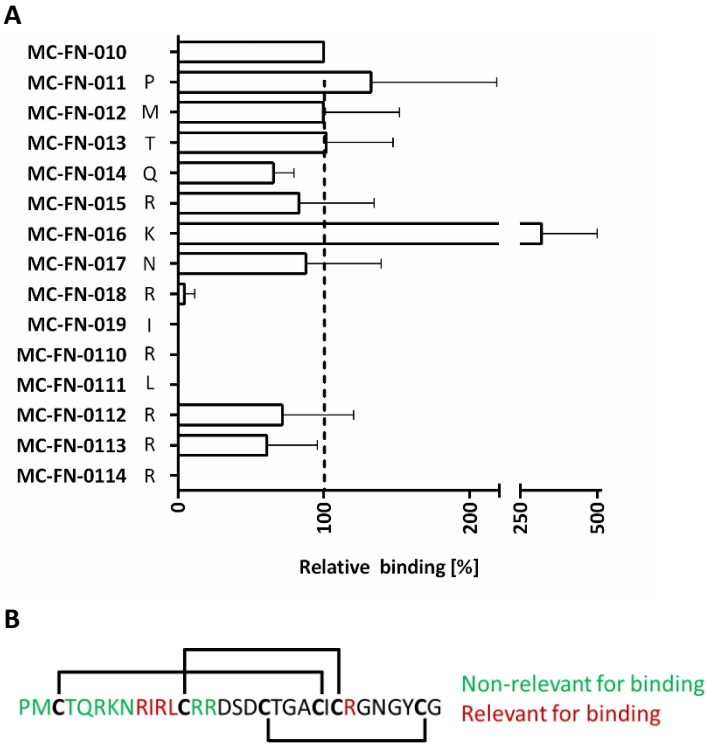


Figure 3.28 Dependency of EDB binding on the R-I/V-R-(L) motif. (A) Identification of functionally relevant amino acids in the MC-FN-010 sequence by alanine scanning mutagenesis. Trx-cystine-knot miniproteins (parental MC-FN-010 and alanine scan variants) were incubated with pre-coated human FN-B. Binding was detected with HRP-conjugated anti-s-tag-antibody. Relative binding of each variant was calculated by comparing the respective apparent binding constant to parental MC-FN-010. Error bars represent the standard derivation of three independent duplicate measurements ± SD. **(B)** EDB-binding relevant amino acid residues of parental MC-FN-010. Disulfide bonds (brackets) between cysteine residues (bold), residues relevant (red) or at best marginally relevant (green) for target-binding are indicated.

Five constructs with exchanges of single amino acids within the common motif positions had lower binding as compared to that of the respective original sequence. An additional alanine substitution in the fifth loop also led to reduced target interaction. In contrast, target binding was preserved in the seven constructs with alanine exchanges in the beginning of the sequence, indicating that those positions were not crucial for EDB binding. Together, these results confirm that the four amino acid residues in the first loop (RIRL), and the arginine residue in the fifth loop impact binding to EDB either directly or indirectly by affecting miniprotein conformation (Figure 3.28 B).

This analysis also identified mutagenized variants, which were comparable with regard to binding to the EDB domain. We leveraged these data for selection of a second development candidate, with which we could also address that the lysine residue in loop 1 of MC-FN-010 may interfere with the selective linkage to primary amines required for full exploitation as an imaging reagent. MC-FN-016 had a primary amine only at the N-terminus, and was found to be comparable to MC-FN-010 with regard to its specificity for binding to FN-67B89 (Figure 3.29). The affinities of tag-free MC-FN-010 and MC-FN-016 towards biotinylated FN-67B89 as tested by SPR analysis were in the single-digit micromolar range with fast off-rates (Table 3.3).

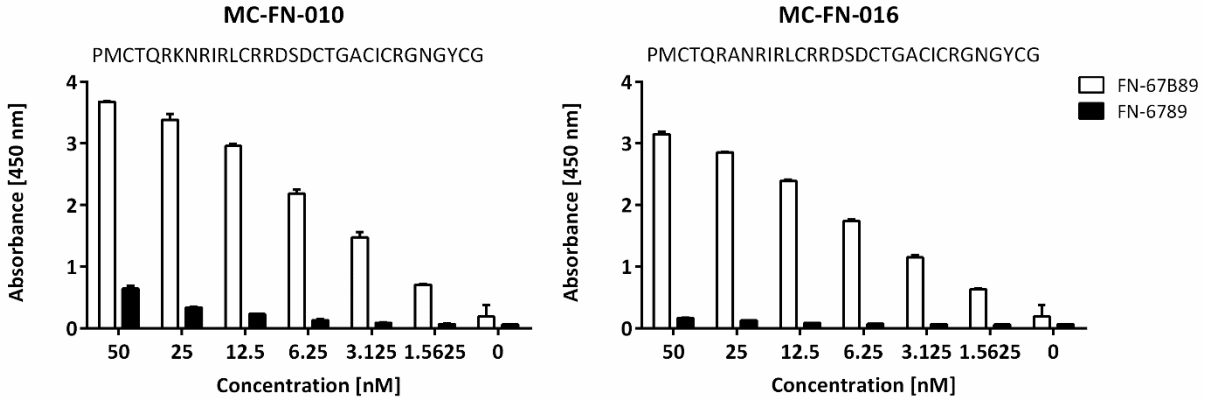


Figure 3.29 ELISA-based binding of MC-FN-010 and MC-FN-016 to human FN-67B89 and FN-6789. Trx-cystine-knot miniproteins were applied at different concentrations and binding was detected with HRP-conjugated anti-S-tag-antibody. Data shown as mean ± SD resulting from duplicates of a single performed ELISA using 1 µg coated FN-67B89 or FN-6789 per well.

Table 3.3 SPR binding analysis of cystine-knot miniproteins to EDB. Kinetic parameters of EDB-specific cystine-knot miniprotein variants resulting from SPR analysis. Biotinylated human FN-67B89 was immobilized to a SA sensor to determine affinity of MC-FN-010 and MC-FN-016 using two-fold serial dilutions starting from 4000 nM. Kinetic parameters were calculated using a 1:1 Langmuir fitting model applied to generated sensorgrams (supplementary Figure 5.4).

Name	Concentration (nM)	Kon (1/Ms)	Koff (1/s)	KD (µM)
MC-FN-010	50 – 4000	4.95×10^4	0.07831	1.58
MC-FN-016	62.5 – 4000	1.33×10^5	0.1873	1.41

3.3.2 Targeting the tumor vasculature on tumor tissue sections and in mice with selected cystine-knot miniproteins

Glioblastomas are known to express the fibronectin EDB isoform in their vasculature[154] and models derived from them are commonly used for *in vitro* and *in vivo* characterization of EDB-binding ligands[136,155]. We resorted to a model based on inoculating the human glioblastoma cell line U-87 MG subcutaneously into the flanks of Fox n1/nu mice to obtain xenograft tumors.

First, we tested the capacity of MC-FN-010 to stain tissue sections derived from these xenograft tumors. MC-FN-0115 was used as non-binding negative control (alanine substitutions in three positions; PMCTQRANRIAACRRDSDCTGACICRGNGYCG) (supplementary Figure 5.3). For IF imaging experiments, biotinylated MC-FN-010 and MC-FN-0115 were tetramerized with Cy3-labeled SA to increase avidity. Tetramerized MC-FN-010-bio bound almost selectively to areas around tumor vessels, as confirmed by co-localization with the vascular marker CD31, a surface protein ubiquitously expressed on endothelial cells (Figure 3.30). In addition, tetramerized MC-FN-010-bio was found to localize in perivascular areas. In contrast, MC-FN-010-bio did not stain normal mouse brain sections and U-87 MG tumor sections stained with the negative control construct MC-FN-0115 showed no fluorescent signal at all.

Next, we assessed the capacity of both MC-FN-010 and MC-FN-016 in the U-87 MG xenografted mouse model *in vivo* with regard to their capacity to visualize tumor vasculature. The cystine-knot miniproteins were chemically trimerized by oxime ligation and tagged with a near-infrared fluorescent dye, Alexa Fluor 680. All three trimeric constructs showed no critical aberration in size and purity as demonstrated by SDS-PAGE and RPC analysis (supplementary Figure 5.5). Oligomerization profoundly improved affinity of both EDB binding cystine-knot miniproteins (AF680-$(MC\text{-}FN\text{-}010)_3$ and AF680-

$(MC\text{-}FN\text{-}016)_3$) to a three-digit picomolar binding constant and remarkably slower off-rates (Table 3.4) compared to those of the monomeric variants (Table 3.3).

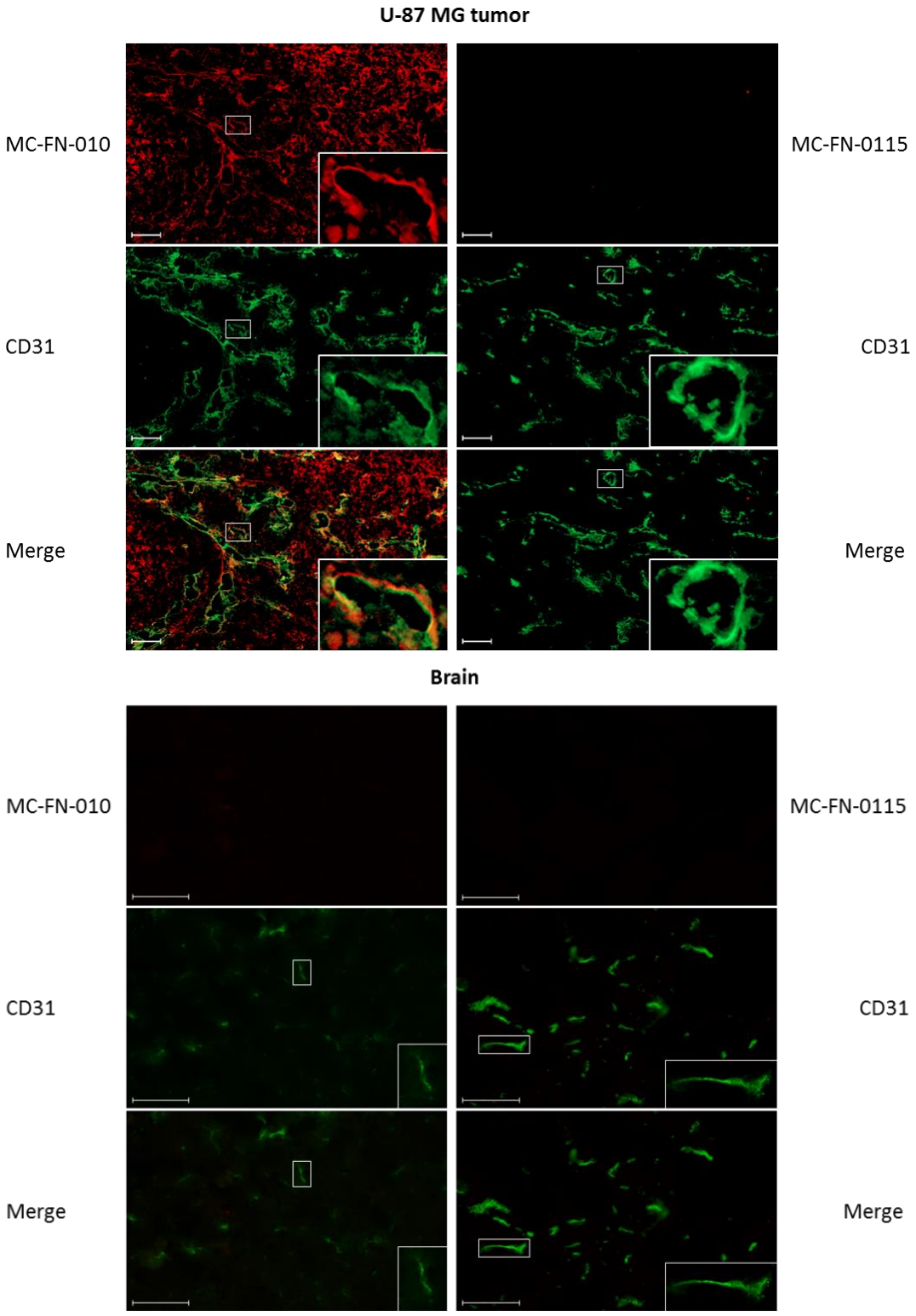


Figure 3.30 Specific binding of MC-FN-010 to tissue sections derived from the human U-87 MG glioblastoma cell line grown as mouse xenograft tumor. Representative IF staining of U-87 MG tumor tissues **(A)** and normal mouse brain **(B)** with EDB ligand MC-FN-010 and negative control MC-FN-0115. Tissue sections (6 µm) were stained with tetramerized cystine-knot miniprotein-biotin / SA-Cy3 complex (red) and an anti-CD31 antibody to visualize vasculature (green). Scale bars, 100 µm.

Table 3.4 SPR binding analysis of trimeric cystine-knot miniprotein variants to EDB. Kinetic parameters of AF680-(MC-FN-010)$_3$ and AF680-(MC-FN-016)$_3$ resulting from SPR with single-cycle kinetic analysis. Biotinylated human FN-67B89 was immobilized to SA chip to perform binding determination of AF680-(MC-FN-010)$_3$ and AF680-(MC-FN-016)$_3$ using two-fold serial dilution starting from 10 nM. Kinetic parameters were calculated using a 1:1 Langmuir fitting model applied to generated sensorgrams (supplementary Figure 5.6).

Name	Concentration (nM)	Kon (1/Ms)	Koff (1/s)	KD (pM)
AF680-(MC-FN-010)$_3$	1.25 – 10	8.04×10^6	3.47×10^{-3}	431
AF680-(MC-FN-016)$_3$	1.25 – 10	2.56×10^6	2.31×10^{-3}	902
AF680-(MC-FN-0115)$_3$	1.25 – 10	No binding to FN-67B89		

Whole-body imaging of mice and *ex vivo* measurement of resected organs was performed after retro-orbital i.v. injection of 3.34 nmol AF680-(MC-FN-010)$_3$, AF680-(MC-FN-016)$_3$ and the negative control AF680-(MC-FN-0115)$_3$. Both AF680-(MC-FN-010)$_3$ and AF680-(MC-FN-016)$_3$, but not the control variant strongly accumulated in the tumor within the first hours and remained for the whole recording period of 6 h. Signals for all trimeric constructs were detected in liver and kidney 1 h and 2 h post injection, but were of transient nature (Figure 3.31 A).

The kinetics of *ex vivo* measured fluorescence in the tumor, kidney and liver, normalized to their respective weight, were assessed for the two EDB-specific probes as compared to the AF680-(MC-FN-0115)$_3$ negative control (Figure 3.31 B). The parental AF680-(MC-FN-010)$_3$ as well as AF680-(MC-FN-016)$_3$ were found to target to the tumor and to remain significantly enriched over the recorded 6 h time period as compared to the triple alanine negative control. The results show that the identified EDB-specific cystine-knot miniproteins are suitable for use as imaging agents for tumor targeting.

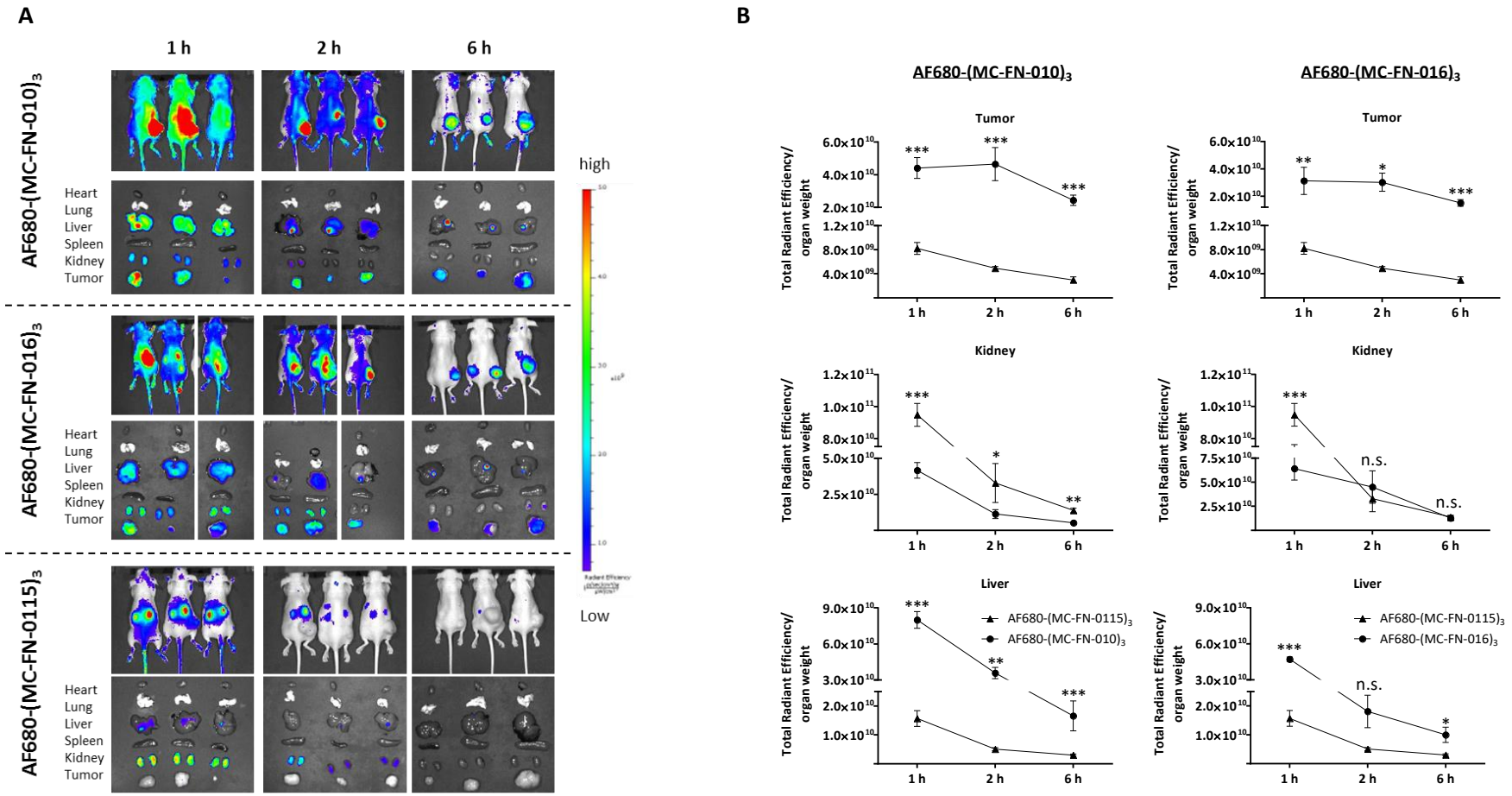


Figure 3.31 Targeting the tumor vasculature in mice with selected cystine-knot miniproteins. (A) *In vivo* and *ex vivo* imaging of U-87 MG bearing mice. Tumors derived from human U-87 MG cells injected s.c. in flanks of mice were imaged after i.v. application of 3.34 nmol AF680-(MC-FN-010)$_3$, AF680-(MC-FN-016)$_3$ (EDB binder) and control AF680-(MC-FN-0115)$_3$. Groups of three mice per group were stratified to each represent a range of tumor sizes. **(B)** Kinetics of fluorescence signals in different organs quantified and normalized to the respective organ / tumor weight. Data are shown as mean ± SD of triplicates from a single experiment. Statistical significance was calculated with two-way ANOVA (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, n.s. = not significant).

4 Discussion

This work comprises the process and technology extension of the internal cystine-knot miniprotein discovery platform adjusted for the early research and development phase. In addition, another part presents the generation of novel tumor imaging probes based on trimeric fibronectin EDB specific ligand and a conjugated near-infrared fluorescence dye. In the following sections the single parts of the work are summarized and discussed in detail.

4.1 An efficient cell-based panning approach for cystine-knot miniprotein selection

Phage display represents a powerful technology for the discovery of novel target-specific binding ligands. A very common and widely applied method is to use target proteins in their soluble form, which can e.g. be presented on immunotubes or captured on beads for panning[32]. This requires a recombinant production of target protein in suitable expression systems. In addition, the protein purification is also crucial to obtain pure products with ideally high yields[156,157]. The protein quality such as purity, solubility and natural conformation plays a critical role for a successful ligand identification using e.g. phage display[140,158]. Nonetheless, this methodology is not necessarily accessible for all proteins due to structural complexity, expression and purification difficulties or solubility issues. The flexibility of the phage display approach however allows not only panning against soluble targets, but also the usage of whole cells to provide proteins in their native conformational environment[18].

Many cell-based panning strategies have been described with differences in the detailed protocol and alteration of the panning conditions[18]. These include incubation conditions, the type of applied phages, cells as well as washing and elution conditions as briefly summarized in Table 4.1. The variety of available cell-based panning protocols and the possible modifications of single parameters basically demonstrates the high diversity of implementing this approach. Consequently, a universally applicable protocol for any phage library cannot be directly defined, but nevertheless the overall goal of effective ligand selection remains the same. However, most publications describe the ligand selection against whole cells from large antibody or peptide libraries, whereby primary cells, tumor cells or stably transfected cell lines were applied[159–162].

Table 4.1 Overview of cell-based panning parameters and its potential modifications.

Panning parameter	Alteration opportunity
Incubation conditions	Blocking agent, incubation buffer, duration, temperature
Phages	Library (type, presented POI), applied phage particles
Negative depletion	Cell line, duration, repetition
Target cells	Primary or immortalized cells, adherent or in suspension
Washing conditions	Buffer, duration, temperature, repetitions
Elution conditions	Buffer, duration, temperature, repetitions

Cain and colleagues screened a mutated C5a phage library against human C5aR-expressing CHO cell line and selected interacting ligands[163]. Another study described a similar panning strategy of a peptide phage library against human CCR5-expressing CHO cells, but the authors included an additional subtraction of phages towards non-transfected CHO cells[159]. An implementation of a counter-selection step should decrease the number of unspecific ligands from the phage population

and enable a target panning with a roughly pre-selected library[141]. In close relation to the published panning procedure from Wang and colleagues[159], the proof of concept experiment in this work demonstrates the feasibility to specifically enrich MC-FA-010 presenting phages by screening against CHO-K1-FAP-α cells (Figure 3.3 B). Based on the initial dilution factor of MC-FA-010 presenting phages to MCopt 1.0 library, the respective clones were detected after one, two or three successive panning rounds. But even more convincing was the successful identification of FAP-α binding MC-FA-010 out of the complete naïve MCopt 1.0 phage library by applying the protocol (Figure 3.10 B), which simulated real panning conditions. This outcome highlights the efficiency of the established cell-based panning protocol to select target interacting cystine-knot miniprotein from a combinatorial phage library.

Stable transfected cell lines represent a good basis to establish a cell-based protocol suitable for our cystine-knot miniprotein phage library, as the target protein is overexpressed on the cell surface and an ideal negative counterpart (non-transfected cell line) is easily accessible[141]. Thereby, the target amount plays a crucial role during phage panning, which depends on the protein density on the cell surface and the applied cell number. The protein surface density of stable cell lines relies on the transgene integration efficiency into the host genome mediated by PiggyBac transposon system and its consistent gene expression[164]. A bulk cell population consists of single clones with different target density, which consequently hampers a full control of antigen availability during cell-based panning. The CHO-K1-FAP-α cell line represents a heterogeneous FAP-α expressing pool, but nonetheless allowed specific cystine-knot miniprotein selection (chapter 3.1.1). Alternatively, a homogeneous CHO-K1-FAP-α cell population can be obtained by isolation and expansion of monoclonal cells based on the same strategy that was applied to generate the CHO-K1-dsRed-Mock cell line in this work (Figure 3.11). A further possibility to generate desired stable target expression cell line represents the Flp-FRT system, a yeast-derived Flp recombinase-mediated site-specific transgene integration. Zhou and colleagues constructed an engineered CHO cell line that harbors a single site-specific recombinase-recognition sequence, which allowed directed target gene integration and yielded cells with homologous protein expression[165]. However, the negative selection on the corresponding non-transfected cell line is also important to deplete ligands from the library that interact with endogenous receptors. A stringent depletion of irrelevant ligands can thus positively influence target selection outcome.

Specific gene silencing or knockdown using e.g. RNA interference or CRISPR/Cas system[166] is another approach to obtain matching target expressing and off-target cell line counterparts. Thereby, tumor cells with suppressed target gene expression could serve for negative depletion, whereas the naïve tumor cells are applied for positive selection that provides an unmodified presentation of the target protein in its natural frequency and environment.

The implementation of a cellular phage screening pipeline increases the power of the cystine-knot miniprotein discovery platform enormously. So far, the cell-based panning technology was adjusted to target overexpressing cell lines, but a direct usage of primary cells or above-mentioned tumor cell lines is also conceivable. The flexible application of various cell types provides on the one hand further selection opportunities and on the other hand requires the adaption panning procedures.

4.2 A robust cell-based hit identification approach in a high-throughput manner

The purpose of the hit identification process is to select individual clones with respect to target binding ability and specificity. A fast, robust and high-throughput approach is favorable in order to cover an analysis of all selected clones from previous phage panning. ELISA for example represents a

traditional methodology for assay protein interaction of putative candidates in a 96-well format[167]. This involves a specific antibody for binding detection via HRP or alkaline phosphatase, while the analyte can either be directly applied exposed on the phage surface after selection or on protein level itself. Moreover, the target can be provided in soluble form or expressed on the cell surface[168–171]. An ELISA-based assay is also applied within our platform for hit screening against soluble target proteins. We used Trx-cystine-knot miniprotein fusions directly after cloning, expression and purification via heat-incubation in a 96-well format for binding analysis. The cloning of selected cystine-knot miniprotein sequences from the enriched phage pool into the expression vector is a critical step, as the sequence distribution can be altered and sequences with lower frequencies might be lost during amplification. However, this approach was initially implemented to be able to screen cystine-knot miniproteins outside from the phage background in order to avoid potential avidity effects resulting from the pVIII display and probable stickiness of the phages. Although we showed the general feasibility to perform whole-cell ELISA with phages displaying cystine-knot miniproteins (Figure 3.1 C), the low sensitivity required high amounts of phages. As these high amounts could not be generated in a 96-well format the whole-cell ELISA system was not considered for further method development. We preferred an adaption of flow cytometry based analysis for cell-based hit screening in relation with the pre-existing high-throughput protein expression system. Flow cytometric analysis is advantageous for hit identification purposes, since a measurement generates multiparametric data of a single cell[172], while covering a large number of cells for a statistically relevant amount. This work presents a robust flow cytometry-based system for the analysis of enriched cystine-knot miniproteins. For assay development the phage pool of the second panning round against CHO-K1-FAP-α was used (the initial dilution ratio of $1{:}10^9$), which contains a moderate number of MC-FA-010 displaying phages (Figure 3.3 B). The cystine-knot miniprotein sequences from the screen pool were cloned into the expression vector and subsequently underwent routinely performed 96-well production procedure. For binding analysis of Trx-cystine-knot miniprotein fusion containing bacterial supernatant, CHO-K1-FAP-α and CHO-K1-Mock cells were assessed in two separate plates. Increased fluorescence signals in CHO-K1-FAP-α cell population was observed and resulted from Trx-MC-FA-010 miniprotein fusion (Figure 3.7). For data processing, interaction of each clone was evaluated as a cellular binding value by dividing the CHO-K1-FAP-α signal with the corresponding CHO-K1-Mock signal. At the end, the three major components of our downstream hit identification process were collected in a ranking list consisting of the cellular binding value, the protein expression rate and the sequence (Figure 3.5). This resulted in a generation of a ranking value for every clone as a measure for specific target interaction. Table 5.1 represents the ranking list of the performed hit identification procedure with middle to high ranking values from all Trx-MC-FA-010 variants as well as from two additional cystine-knot miniproteins and peptide clones. Consequently, our ranking system enables a differentiation of clones with target binding ability from those with no interaction. In addition, the developed downstream process was verified with the successful identification of Trx-MC-FA-010 and target interacting peptides resulting from the naïve MCopt 1.0 library panning against CHO-K1-FAP-α cells (Figure 3.10). Overall, these results demonstrate the applicability of the flow cytometry-based hit identification system for the efficient screening of target binding cystine-knot miniproteins in a 96-well scale format. The data evaluation scheme allows a correlation of obtained values for reliable identification of promising candidates. However, this system also harbors limitations, including the lack of a suitable threshold to be able to distinguish hit candidates from unspecific binders or the arising of false positive clones due to single sample measurements. Hit screening is thus supposed to

give first insights of putative candidates, while a more detailed analysis of single clones afterwards is imperative in any case.

The high-throughput nature of the system was further optimized by the generation of a red fluorescent mock cell line for applying simultaneous detection of target and off-target binding (Figure 4.1 A). The second negative cell line (CHO-K1-dsRed-Mock) was generated to endogenously express a red fluorescent dye allowing discrimination from the target expressing cell lines that carried the GFP reporter gene. Figure 3.12 shows the feasibility study with the mixed cell populations and the crucial FAP-α binding of Trx-MC-FA-010 miniprotein fusion to CHO-K1-FAP-α cells, in which no interaction towards CHO-K1-dsRed-Mock cells was observed. In particular, adequate functional assay sensitivity can be concluded, while the doubling of hit screening throughput additionally leads to an improved cost- and time-efficient analysis.

Furthermore, a 96-well expression system for the *in vivo* biotinylation of Trx-cystine-knot miniproteins in *E. coli* was implemented to improve the potential sensitivity of detection. The biotinylated protein can be tetramerized for target binding in order to increase putative binding strength by exploiting potential avidity effects. Figure 4.1 illustrates the binding detection principle of monomeric Trx-cystine-knot miniprotein fusion and tetramerized Trx-cystine-knot miniprotein fusion. This procedure increases the assay sensitivity extremely by means of signal enhancement, which was generated by tetramerized Trx-MC-FA-012 miniprotein fusion in comparison to monomeric version in a flow cytometric setup (Figure 3.13 B).

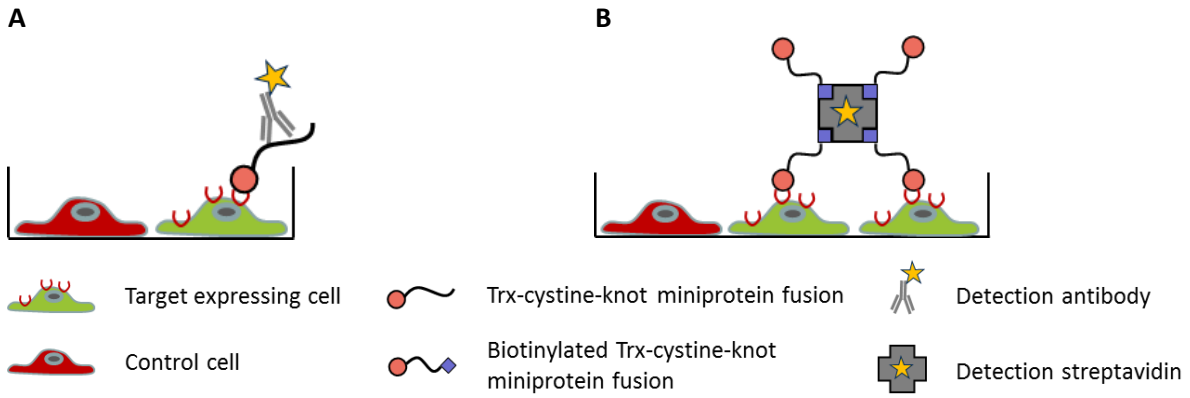


Figure 4.1 Schematic illustration of monomeric and tetramerized Trx-cystine-knot miniprotein fusion for cellular binding analysis via flow cytometry. (A) Target binding analysis of monomeric applied Trx-cystine-knot miniprotein via fluorescence-dye labeled anti-His-tag antibody. **(B)** Tetramerization of biotinylated Trx-cystine-knot miniprotein fusion using fluorescence-dye conjugated SA for target binding detection.

In sum, the presented approach allows the downstream hit screening of individual clones in a high-throughput manner. It consists of two 96-well production systems yielding ligand candidate as Trx-cystine-knot miniproteins either unmodified or in site-specifically biotinylated format. Both versions can be used for flow cytometry-based target interaction assessment using on the one hand monomeric Trx-cystine-knot miniprotein fusion and on the other hand tetramerized variants of biotinylated Trx-cystine-knot miniprotein fusions. Cellular data outcome can be easily correlated with the protein expression values for the generation of a ranking list that enables the identification of promising hit candidates. Our system is advantageous, as already a small volume of roughly purified proteins in the bacterial lysate can be directly assessed with sufficient sensitivity outcome in a primary hit screen. However, putative hits generally undergo secondary or further functional characterization, for which the developed high-throughput cellular binding assay would be suitable to analyze e.g.

different concentration range of fully purified candidates. For this purpose, 96-well His-tag purification system can be easily introduced to achieve purified Trx-cystine-knot miniprotein fusions. Yang and Nolan presented a high-throughput approach for screening and characterization of selected phage clones. The authors set up a multiplexed binding assay based on optically-encoded microsphere and flow cytometry. To assess target binding ability, the phages were captured on anti-M13 antibody-bearing microspheres, while the target protein was fluorescence labeled for the measurement[173]. A similar strategy could be considered for our platform as well. Binding of cystine-knot miniprotein presenting phage to target expressing cell could be directly assessed within panning using flow cytometry. In any case, a reliable screening protocol has to be developed and validated.

4.3 Improved cystine-knot miniprotein phage libraries

In this work the design, construction and characterization of three cystine-knot miniprotein phage libraries is reported that are based on previous obtained findings from the literature combined with sophisticated rational design strategies. These represent fundamental and crucial steps for efficient ligand selections within the internal drug discovery platform. As described in detail in chapter 3.2.1, the oMCoTI-II protein scaffold was the basis for the new libraries. The MCopt 2.0 library consists of ten randomized amino acids in the functional amino-terminal loop of the oMCoTI-II protein, while MCopt 2.1 library harbors ten amino acid sequence randomization in loop 5. A third library (MCopt 3.0) was constructed with ten randomized amino acids in both adjacent loop 1 and loop 5.
Crucial requisites of combinatorial libraries are high clonal diversity, properly folded and displayed proteins as well as correct amino acid sequence compositions[161,174]. In this work, three phage libraries were generated with reasonable primary transformation numbers of at least 1.06×10^9 clones (Figure 3.17 A). Sequencing was done to analyze library quality on a molecular level of all three libraries (Figure 3.17 B). In the MCopt 2.0 library a proportion of 72.9 % represented the expected cystine-knot miniprotein sequences. On the contrary, the frequency of correct sequences in MCopt 2.1 and MCopt 3.0 library was much lower, while a similar proportion of non-assignable sequences in the loop regions emerged. Those clones were shown to be most probably cells with double or multiple transformed phagemids (chapter 3.2.2), which consequently negatively affects the essential characteristic of a genotype and phenotype coupling within the phage display system[32]. We present an *E. coli* infection approach to overcome this problem according to a procedure described by Sblattero *et al.*, who established a scFv library with high clonal diversity using a Cre recombinase system[151]. We assumed that an infection of *E. coli* with the primary library phages (resulting from the multiple transformed cells) would restore the genotype-phenotype linkage when a low MOI is used. As a prerequisite each phage must contain only a single DNA species and the MOI must be sufficiently low to ensure that a single phage only infects an individual cell. The hypothesis was confirmed, as the occurrence of non-assignable sequences in the loop position in all libraries decreased enormously with a proportion of below 3.1 % (Figure 3.18 B). However, the reinfection procedure of primary phage particles hampers a reliable calculation of library diversity. Although an absolute number could not be calculated, the theoretical clonal diversity of all libraries is supposed to be below the determined number of infected cells (Figure 3.18 A). A further quality requisite of functional cystine-knot miniprotein presentation on phage particles was indirectly proven with an antibody directed against the assembled Fos-molecule (Figure 3.19 A, Figure 3.19 B). Therefore, the MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries accomplished quality requirements on molecular and phage level. For downstream hit identification procedure, the accessibility of Trx-cystine-knot miniproteins derived from MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries for recombinant expression in a mini-

scale setting was shown. The majority of MCopt 2.1 and MCopt 3.0 library clones yielded protein concentrations below 1 mg/mL, while the product yield of MCopt 2.0 library clones was equally distributed in the range of < 1 mg/mL and 1.49 mg/mL (Figure 3.20). Furthermore, an influence of the sequence exchanges and extension in specific loop positions on correct oMCoTI-II protein folding and functionality was assessed. Individual clones of MCopt 2.0, MCopt 2.1 and MCopt 3.0 libraries were recombinantly produced for mass determination. The formation of three characteristic disulfide bonds of oMCoTI-II framework was taken as a measure of tolerance to sequence and length modification. Most of the cystine-knot miniproteins had the expected mass, while MCopt 2.0 proteins revealed the highest frequency (Table 3.2). A general tolerance of loop 1 to substitutions can be concluded, while retaining constrained structural motif. In contrast, amino acid exchanges and length alteration in loop 5 or in dual-loop positions seemed only to be bear in a moderate manner.

The feasibility of modifying the inhibitory loop position in the oMCoTI-II sequence for randomized combinatory library construction has been previously demonstrated[92,175]. But also the in-house experience based on the MCopt 1.0 phage library emphasized these observations. Furthermore, Kimura and co-workers presented a loop-grafting approach in the related acyclic MCoTI-II sequence to a pharmacokinetically stabilized knottin with $\alpha_v\beta_6$ integrin binding ability[91]. However, loop 5 or dual-loop randomization of oMCoTI-II scaffolds has not been systemically investigated so far. In a previous report from Glotzbach and colleagues, sequence composition in loop 5 of oMCoTI-II protein has been substituted in a very moderate variation manner as a part of combinatorial library design. Yeast display screening resulted in selection of engineered knottin directed against human matriptase-1 with binding and inhibitory function[92].

The MCopt 3.0 library design with loop-selective randomization scheme was based on a conceptual strategy to increase the potential binding area, which might result in a higher target affinity. Moreover, cystine-knot miniproteins with dual-loop functionalities are a very interesting option to gain bispecific targeting. Successful ligand selection of MCopt 2.0 and MCopt 2.1 phage libraries against independent target molecules offers the opportunity to combine loop 1 and loop 5 subsequently via loop grafting for achieving dual-binding functionalities. Recently, Chan and co-workers grafted two anti-angiogenic epitopes in loop 1 and loop 5 or loop 5 and loop 6 of the cyclic MCoTI-II framework with inhibitory effects on HUVEC and tumor cell proliferation[176]. Based on their experimental results, the general concept of dual-targeting seems to be feasible and encouraged our rational designed approach.

In recent years, various combinatorial cystine-knot miniprotein-based libraries were constructed for tumor-associated antigen identifications. This mostly involved engineered knottin presentation via yeast display and high-throughput screening in combination with flow-cytometry. The evolution-directed strategy resulted in selection of several knottin ligands against target proteins[90,92,93,175]. But also functional displaying of EETI-II and kalata B1 derived knottin libraries on bacterial surface has been reported[177,178]. However, we present the functional phage surface presentation of cystine-knot miniprotein using the pJuFo phagemid system[42]. A low valency display through the recombinant pIII´ protein was intended to provide the possibility of high target affinity cystine-knot miniproteins selection, based on the hypothesis of avoiding strong avidity effects on phage level. The control of protein copy numbers alone might however not be sufficient enough, an inclusion of panning stringency might support the route to this purpose[179].

4.4 Challenges and hurdles of novel MCopt 2.0 phage library

The successful ligand selection from a combinatorial phage library relies on different parameters, including library quality, panning stringency or binding affinity[53]. After panning completion of MCopt 2.0 library against CHO-K1-DSG-3 cells a strong enrichment of deletion mutants lacking the whole Fos-cystine-knot miniprotein and Jun-pIII´ sequence part was observed (Figure 3.22). The deletion mutants showed growth advantages over MCopt 2.0 library pool (Figure 3.23) and therefore developed to be a critical issue that needed to be overcome. Their general incidence became challenging, since a lack of sufficient target specific phage population enrichment resulted in no novel cystine-knot miniprotein selection and furthermore deteriorated the library quality. Mueller reported about a similar observation of deletion in the Fos/Jun region, which predominantly occurred under low selection pressure. The author solved this problem by introduction of an additional antibiotic resistance marker gene[145]. Another report describes a genetic instability of repetitive gene segments, which eventually can cause insertion or deletion of DNA sequences[153]. We could show that the deletion mutants in our case most likely developed based on repetitive DNA sequences in the promotor and pelB leader segments. The construction of a pPDIII-1 phagemid vector, in which all tandem repeats were removed clearly prevents the occurrence of deletion mutants during panning procedure (Figure 3.27 B). This effect could have been achieved due to suppression of putative recombination processes of genetic elements. Or secondly, a deletion mutant outgrowth or enrichment was prevented, as pPDIII-1-MCopt 2.0 library clones revealed equal growth behavior (Figure 3.26 B). Although the pelB as well as the ompA leader sequence follow the sec pathway for translocation of proteins to the cell membrane and are routinely used for antibody phage display[180], the usage of ompA in pPDIII-1-MCopt 2.0 sub-library resulted in a weaker frequency of molecule presentation on the phage surface than pJuFoIII-MCopt 2.0 library (Figure 3.26 A). However, the monocistronic gene structure of expression cassettes within the pPDIII-1 vector might also contribute to that incidence (Figure 3.24 B). So far, however an interfering impact on library quality was not observed, as a display of Fos-cystine-knot miniprotein was clearly detected.

In addition to the observed genetic instability of the phagemid, other factors concerning the DSG-3 target protein structure or the target density on the cell surface could have negatively affected the selection. The single-pass transmembrane protein is composed of an intracellular domain and four extracellular cadherin (EC) domains[181], which results in a rather linear structure. The structural feature of the expressed DSG-3 ectodomains on CHO-K1 cells might thus not be presented ideally for tight cystine-knot miniprotein binding and selection using our applied cell-based phage screening approach. Moreover, the protein cell surface density could have an additional impact on accessibility during panning. Both hypotheses are possible, but however not explicitly proven. In contrast to the described cell-based selection against DSG-3, a soluble protein-based selection against single domain FN-B protein was successful, although FN also forms to a linear array[108] and shares overall structural similarity with DSG-3 (chapter 3.3). Consequently, a phage panning against a single EC domain of DSG-3 in a soluble form might lead to successful cystine-knot miniprotein identification as well. In our case it can be concluded that likewise multiple parameters influence the selection outcome enormously, as known in advance. Not only the library quality and panning strategy plays a critical role, but most probably also the target protein of choice and the target protein presentation.

4.5 EDB specific cystine-knot miniproteins as tumor imaging agents

Targeted tumor imaging with peptide or protein-based molecular probes may improve cancer diagnosis, staging and treatment monitoring[182]. The selective expression of the extra domain B containing isoform of FN in tumor-associated neovasculature[121,122], makes it an ideal target for ligand-based delivery of drug molecules. Derivatives of the EDB binding human L19 antibody have been widely used as molecular system to deliver radioisotopes or cytokines in cancer imaging and therapy[183–185]. Several L19-based approaches are currently under clinical investigation with promising early results[186,187]. This together with the fact that cystine-knot miniproteins entail a broad range of favorable characteristics encouraged us to generate EDB-binders based on this scaffold.

In this study, we present the development of novel cystine-knot miniprotein-based ligands with high specificity for fibronectin EDB and its preclinical evaluation as an *in vivo* imaging probe in a glioblastoma mouse model. The screening of two combinatorial libraries against an EDB-containing protein fragment resulted in strong enrichment of six clones, five of which shared the distinctive R-I/V-R-(L) motif. We characterized the identified hits and selected MC-FN-010 based on the robust affinity and specificity of its binding to EDB for further development. Mapping of the MC-FN-010 binding site through alanine scanning mutagenesis revealed that the amino acid residues R-I-R-L in the first loop and an arginine residue in the fifth loop are required for EDB binding activity of the peptide (Figure 3.28). One may speculate that the consensus sequence together with amino acids located in loop 2 (R-R-D-S-D) and loop 5 (R-G-N-G-Y) forms a patch of positively charged residues that mediate binding to EDB mainly via electrostatic interactions. This hypothesis is supported by the dominance of negative charges with only two positively charged residues being represented within the 91 amino acids of EDB[188].

The measured binding affinities of the parental MC-FN-010 and its derivative MC-FN-016 to FN-67B89 were in the single-digit micromolar range (Table 3.3). As a higher binding strength is may be favorable for efficient tumor uptake and retention *in vivo*, in particular for small peptides that are excreted rapidly via the kidneys[189,190], we oligomerized these ligands and conjugated the resulting trimers with a fluorescent dye. This resulted in apparent affinities in the picomolar range (~1500- to 3600-fold stronger as compared to the monomeric versions) (Table 3.4), comparable with the reported binding constants for the single-chain Fv L19 fragment[127].

The ability of the trimerized and labelled formats of MC-FN-010 and MC-FN-016 to bind to EDB-expressing tumors *in vivo* was investigated in an U-87 MG xenograft mouse model. Both molecules were selectively taken up and maintained within the tumor over the entire recorded period of 6 h in contrast to a triple alanine mutant that served as a negative control (Figure 3.31). Signals observed in liver and kidney decreased over time. The uptake and retention of oMCoTI-II-based peptides in the kidneys was reported in other studies and attributed to a high content of arginine residues[91].

Taken together, the strong tumor uptake of these trimeric constructs and their favorable biodistribution indicate the suitability for diagnostic imaging applications. These EDB ligands may provide a platform, which can support a diversity of *in vivo* cancer imaging and targeting approaches including conjugation of radioactive tracers, use for PET or SPECT nuclear imaging techniques and targeting of bioactive payloads.

4.6 Concluding remarks and future perspectives

Cystine-knot miniproteins emerged as a promising non-immunoglobulin class with high potential as molecular scaffold for a variety of different clinically relevant applications[84]. In particular, the ability

of various cystine-knot miniprotein derived formats for tumor imaging approaches has been extensively evaluated in preclinical studies[96,98,191,192]. One of the most advanced imaging agents is based on the CTX peptide, which has been conjugated to the radioisotope iodine-131 for clinical phase I dose study in human glioblastoma patients. The patients tolerated the single-dose administration, while the imaging compound specifically accumulated in the tumor[100]. Moreover, a modified version of CTX was conjugated to indocyanine green (BLZ-100) for intraoperative and diagnostic application in a preclinical setting[101,193]. Beside the strong advancement for tumor imaging technologies, therapeutic applications can also be taken into account. These include the exploitation of inhibiting functionalities or antagonistic activity after target binding[92,176,194,195] or by conjugation of bioactive payloads that are specifically delivered to the tumor site. The versatility of the cystine-knot miniprotein scaffold for targeted delivery have been successfully demonstrated by conjugated radioisotope or toxins and as antibody surrogates via fusion to IgG Fc portion[196–200].

So far, the internal phage screening pipeline enabled a reasonable *de novo* discovery of target interacting cystine-knot miniproteins by using soluble target proteins. This thesis reports about the establishment of an efficient phage panning procedure for cystine-knot miniprotein selection against target protein overexpressing mammalian cells. In addition, a robust hit identification process was developed for high-throughput screening of single clones towards target interaction ability. The flow cytometry-based measurement allows the analysis of monomeric or tetramerized Trx-cystine-knot miniprotein fusions by using two different detection systems. The implementation of a cell-based screening pipeline could potentially augment the power of our discovery platform and may offer novel opportunities to additionally address surface receptors in their native conformation, especially to overcome limited accessibility of target proteins for recombinant production. The second part of this work describes the extension of the cystine-knot miniprotein phage library repertoire. Novel phage libraries were constructed based on an improved rational design strategy and resulted in ten randomized amino acids in loop 1, loop 5 or dual loop positions of the oMCoTI-II protein scaffold. All phage libraries were successfully generated with reasonable quality. The last part of this thesis reported for the first time about the development of an engineered cystine-knot miniprotein for targeting of EDB expressing tumors, which was initially selected from a combinatorial phage library. This molecule underwent each step of our cystine-knot miniprotein technology platform and entered together with a second candidate a preclinical *in vivo* imaging testing. Their usage as imaging probes was evaluated in a U-87 MG based xenograft glioblastoma mouse model. All in all, the whole work demonstrates the availability of a powerful platform for the preclinical discovery and the development of cystine-knot miniprotein candidates, which were further engineered as protein scaffolds for tumor imaging technologies as presented in Figure 4.2. In addition, the results further emphasize the relevance of cystine-knot miniprotein framework as tumor diagnostic tool.

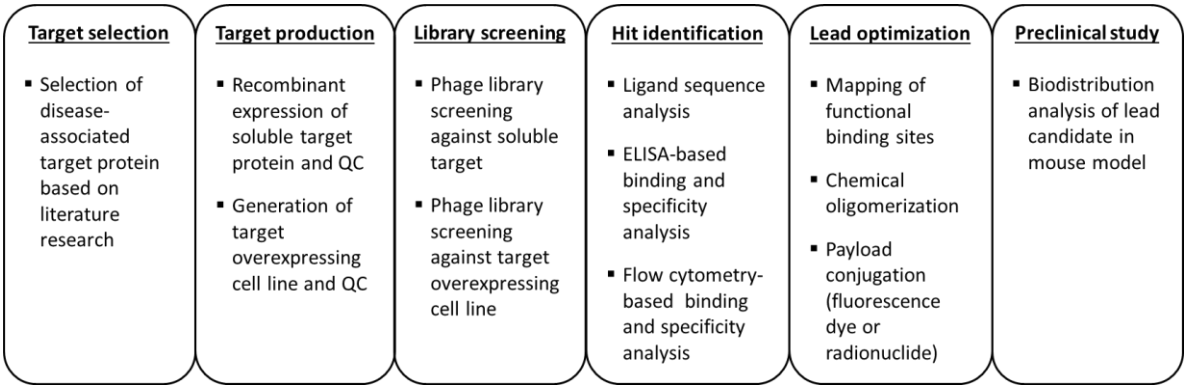


Figure 4.2 Overview of internal cystine-knot miniprotein discovery platform.

5 Supplementary

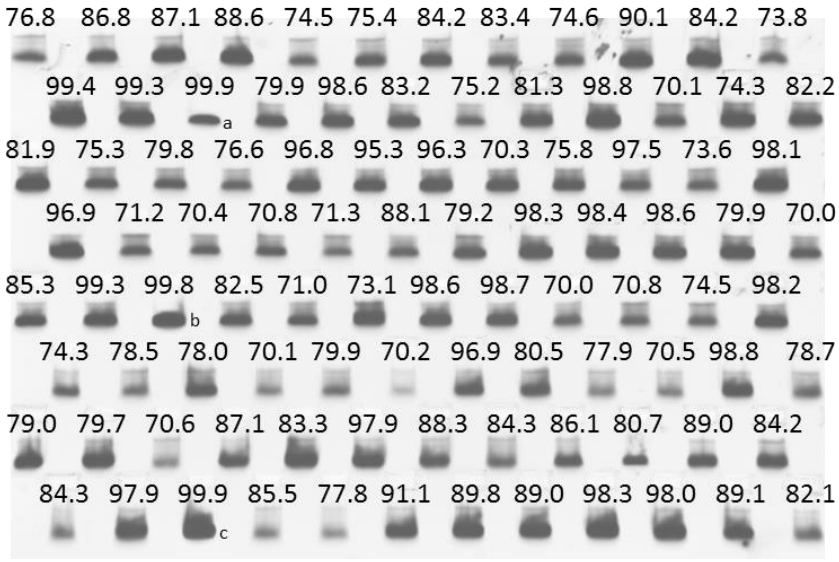


Figure 5.1 Purity of Trx-cystine-knot miniprotein fusion after heat-purification step. Representative E-PAGE™ analysis of expressed Trx-cystine-knot miniprotein fusions from MCopt 1.0 library in a 96-well scale. Analysis was performed with 10 µL of bacterial supernatant that contained roughly purified Trx-cystine-knot miniprotein. Purified Trx-MC-Myc-010 was applied as protein reference at 3 µg (a), 10 µg (b) and 20 µg (c). Percentage of protein purity was determined via densitometric analysis and shown in numbers above each band.

Table 5.1 Ranking list of hit identification procedure for spike-in screening against CHO-K1-FAP-α. Phage pool after second panning round of initial dilution ratio of $1:10^9$ was used. Cellular binding value [signal (CHO-K1-FAP-α)/signal (CHO-K1-Mock)], protein expression value and sequence are shown for each variant. Ranking value was generated by dividing cellular binding value with the corresponding expression value. Variants were sorted according to the highest ranking value. MC-FA-010 sequences are highlighted in bold letters.

Variant	Cellular binding value	Expression value	Ranking value	Sequence
H01	6.61	0.31	21.32	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
A01	38.40	4.90	7.84	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
A07	26.08	4.37	5.97	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
C01	33.71	5.90	5.71	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
E01	35.28	7.12	4.96	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
G01	23.86	4.97	4.80	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
C11	16.71	5.12	3.26	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
F01	24.22	7.72	3.14	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
C04	18.84	6.68	2.82	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
G05	13.52	5.40	2.50	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
B09	18.34	7.46	2.46	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
H05	2.12	0.89	2.38	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
D08	15.67	6.95	2.26	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
B01	17.79	7.90	2.25	KYCTDESHGCRRDSDCIAACICRGNGYCG
E05	15.16	6.82	2.22	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
D10	12.74	5.96	2.14	HQVPSRLRLSDGMHLPW
D12	13.94	7.18	1.94	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**

D01	15.14	7.96	1.90	DECTNLEKYKRPCRRDSDCNLACICRGNGYCG
F02	15.04	7.99	1.88	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
E11	12.29	6.56	1.87	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
G03	8.19	4.45	1.84	**GKCPYSNWTPGRGPNCRRDSDCPGRCICRGNGYCG**
E10	12.36	6.91	1.79	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
H11	5.50	3.52	1.56	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
G02	8.57	5.68	1.51	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
H08	2.04	1.40	1.46	NQCSRNTDSLSSRLRLPGSLYLPW
F08	11.30	7.94	1.42	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
G11	6.86	5.11	1.34	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
D02	8.63	6.56	1.32	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
G10	6.02	5.37	1.12	**GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG**
H09	1.47	1.52	0.96	frameshift
G06	4.47	5.38	0.83	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
F04	8.98	10.98	0.82	frameshift
C02	4.79	8.16	0.59	GECGANNDTKRTCRRDSDCPGGCICRGNGYCG
E02	4.13	7.18	0.58	DRCNQHQMSCRRDSDCPMACIRRGNGYCG
B02	3.03	6.34	0.48	GTCRKRIRSCRRDSDCQGACICRGNGYCG
C10	2.15	4.87	0.44	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
H07	0.98	2.23	0.44	sequence failure
H06	1.10	2.59	0.42	GGCKSQSPRKTDKWGCRRDSDCPLACICRGNGYCG
H02	1.25	2.97	0.42	RDCTEQADRRYNCRRDSDCPGACICRGNGYCG
A02	2.78	6.93	0.40	TDCEKEPGTDKERTPCRRDSDCEGACICRGNGYCG
G04	1.58	4.23	0.37	YKCKNPGRYCRRDSDCDMACICRGNGYCG
A03	2.50	7.16	0.35	PHCKPQGWPCRRDSDCPIACICRGNGYCG
C03	2.29	7.38	0.31	STCTSEPSPCRRDSDCPGSCICRGNGYCG
F03	2.54	8.49	0.30	HTCPREKENCRRDSDCERACICRGNGYCG
A09	1.02	3.51	0.29	PWCNSIYHWCRRDSDCPGACICRGNGYCG
H04	1.06	3.68	0.29	EYCKISSKNPGYCRRDSDCTRACICRGNGYCG
B11	2.17	7.57	0.29	NRCKENHMDTYQHQQVVPSRLRLPDGMHLPW
A11	1.03	3.72	0.28	QRCRGYTPSDHTCRRDSDCKMACICRGNGYCG
H10	1.02	3.77	0.27	HTCKGRYHYCRRDSDCPGACICRGNGYCG
C09	1.12	4.30	0.26	RPCNKWPSNYHYTKGCRRDSDCHWACICRGNGYCG
F05	2.02	7.97	0.25	IDCSTYPQQPGQQSGCRRDSDCPGDCICRGNGYCG
A12	1.00	4.68	0.21	QYCLDRGPHGSSPSHCRRDSDCNGACICRGNGYCG
D04	1.58	7.42	0.21	HWCTRYPHNCRRDSDCPVACICRGNGYCG
G08	1.00	5.03	0.20	VNCPSGRRGHRIGIWMPSRLRLPGSMHLPW
A06	1.20	6.13	0.20	QGCTRTHSPDHQCRRDSDCGYRCICRGNGYCG
B06	1.03	5.30	0.19	sequence failure
A08	1.06	5.46	0.19	DTCKSYGYYYDECRRDSDCSGACICRGNGYCG
B04	1.09	6.22	0.17	frameshift
E08	1.36	8.00	0.17	TDCPTNKSECRRDSDCPKACICRGNGYCG
C06	1.03	6.16	0.17	TKCYHKEQQDYTPPKCRRDSDCPWRCICRGNGYCG
C12	1.01	6.14	0.16	NSCIQSRNPCRRDSDCPGACICRGNGYCG
A04	1.10	6.83	0.16	EHCPQPYSRCRRDSDCPRACICRGNGYCG

A10	0.99	6.24	0.16	QYCLDRGPHGSSPSHCRRDSDCNGACICRGNGYCG
D09	1.04	6.64	0.16	TYCKRQHETCRRDSDCDMRCICRGNGYCG
B12	1.00	6.44	0.16	HKCEQPLDRCRRDSDCNGACICRGNGYCG
G09	1.00	6.41	0.16	vector re-ligand
D07	1.02	6.58	0.15	QPCASQRAYRNVCRRDSDCPIACICRGNGYCG
E12	0.96	6.35	0.15	KSCSTDEPKNEYCRRDSDCYGRCICRGNGYCG
C08	1.05	7.01	0.15	STCTNFRGGCRRDSDCPKQCICRGNGYCG
G12	1.02	6.86	0.15	PRCQTRNPYQQSESPCRRDSDCPGACICRGNGYCG
D03	1.14	7.63	0.15	TGCYSSNHYCRRDSDCPTNCICRGNGYCG
D05	1.03	6.91	0.15	YKCKPGEEQCRRDSDCPGACICRGNGYCG
B07	1.21	8.30	0.15	frameshift
C07	1.00	6.89	0.15	SYCTGNQLECRRDSDCPKHCICRGNGYCG
G07	1.03	7.12	0.14	TDCYRTKEQCRRDSDCQMACICRGNGYCG
B08	1.06	7.32	0.14	RPCKPNMSQCRRDSDCPGACICRGNGYCG
A05	0.97	6.74	0.14	GKCDKPSDYCRRDSDCPGECICRGNGYCG
D06	1.06	7.81	0.14	HYCPHQEDDCRRDSDCPGACICRGNGYCG
E06	1.05	7.76	0.14	TGCLQYYTEGPECRRDSDCPGACICRGNGYCG
E04	1.03	7.70	0.13	PPCGMSEPSCRRDSDCPVACICRGNGYCG
B10	1.02	7.63	0.13	QPCISYGTTCRRDSDCPHNCICRGNGYCG
E07	1.00	7.92	0.13	SYCGPAYYYCRRDSDCPGSCICRGNGYCG
B05	0.98	8.21	0.12	GECSTDESEHSDHYGCSRDSDCKMNCICRGNGYCG
F07	1.02	8.56	0.12	GHCIRRSSNCRRDSDCPKECICRGNGYCG
C05	1.12	9.39	0.12	RTLPGNPVEVPSRLRLPWSMHLPW
F10	1.04	8.87	0.12	PGCRTHMPYEHTRPECRRDSDCSRACICRGNGYCG
E09	1.03	8.88	0.12	QSCKNKSEQCRRDSDCGMACICRGNGYCG
F06	1.04	9.24	0.11	frameshift
F12	1.00	9.11	0.11	QYCKMSHRKERKCRRDSDCPGACICRGNGYCG
F09	1.01	9.51	0.11	PECNRKYTTCRRDSDCPIACICRGNGYCG
F11	1.01	9.86	0.10	GQCRIRGEDCRRDSDCQGACICRGNGYCG
D11	1.00	9.93	0.10	vector re-ligand

Table 5.2 Ranking list of hit identification procedure for MCopt 1.0 library screening against CHO-K1-FAP-α after third panning round. Cellular binding value [signal (CHO-K1-FAP-α)/signal (CHO-K1-Mock)], protein expression value and sequence are shown for each variant. Ranking value was generated by dividing cellular binding value with the corresponding expression value. Variants were sorted according to the highest ranking value. X represents sequencing error.

Variant	Cellular binding value	Expression value	Ranking value	Sequence
D06	9.93	0.69	14.39	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
A11	1.40	0.39	3.60	GKCGKVPSMDRTNQLPSRLRLPXXMYLPW
D09	4.89	1.42	3.44	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
D08	5.61	2.14	2.62	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
B07	5.86	2.77	2.12	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
E04	11.78	7.05	1.67	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
D12	3.18	2.57	1.24	GKCPYSNWTPGRGPDCRRDSDCPGRCICRGNGYCG
A10	4.45	3.69	1.21	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW

A12	1.73	2.05	0.84	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
E06	6.08	8.62	0.70	GKCGKVPSMDRTNQLPSRLRLPGNLYLPW
D10	1.15	2.46	0.47	TNCWEQPSKHRSGCRRDSDCNGACICRGNGYCG
C11	1.02	2.27	0.45	GQCRXXDQXXLGXCRRDSDCPVTCICRGNGYCG
C06	3.01	6.98	0.43	LTCMDGREQCRRDSDCSGACICRGNGYCG
A03	1.26	3.3	0.38	TNCWEQPSKHRSGCRRDSDCNGACICRGNGYCG
D01	1.56	4.15	0.38	KECYPPRQHGHECRRDSDCRGACICRGNGYCG
D05	2.31	6.68	0.35	KDCRNDYDGCRRDSDCQGACICRGNGYCG
B02	1.15	3.41	0.34	SGCHNKISYDYGCRRDSDCSYRCICRGNGYCG
B08	1.03	3.12	0.33	DKCNLQHEYGHNGGPCRRDSDCQWACICRGNGYCG
C07	1.04	3.27	0.32	KPCPTKMYGCRRDSDCPGACICRGNGYCG
E02	1.11	3.66	0.30	HPCNSKLNHCRRDSDCTWACICRGNGYCG
C05	1.12	3.97	0.28	STCYHPHYESHYAYPCRRDSDCDRACICRGNGYCG
C10	1.05	4.18	0.25	TGCQMEHSRHPENRSCRRDSDCTGGCICRGNGYCG
D04	1.15	4.86	0.24	TKCTGSTKNCRRDSDCPGACICRGNGYCG
B04	1.23	5.28	0.23	QQCPYNSTHGDYCRRDSDCPTACICRGNGYCG
A08	1.17	5.47	0.21	PECRYDPEYKPQCRRDSDCHGACICRGNGYCG
D03	1.29	6.25	0.21	QRCGKETEQCRRDSDCNTACICRGNGYCG
A09	0.99	4.96	0.20	PLCQHPKYQCRRDSDCPTACICRGNGYCG
B05	1.35	6.96	0.19	TNCSETPNLCRRDSDCYAECICRGNGYCG
C04	1.19	6.35	0.19	QHCNEDCRRDSDCPVKCICRGNGYCG
B10	1.08	5.86	0.18	HTCQPGRNQGYRCRRDSDCEGQCICRGNGYCG
D07	1.05	5.75	0.18	SHCTTATSPCRRDSDCTGQCICRGNGYCG
C03	1.13	6.25	0.18	SKCQQPPHQIQPCRRDSDCPMECICRGNGYCG
E08	1.13	6.24	0.18	TRCNTYPRSYTPNQICRRDSDCPMDCICRGNGYCG
A02	1.10	6.33	0.17	ERCPDKKSECRRDSDCPWACICRGNGYCG
B09	1.03	6.3	0.16	TSCTSYKAHCRRDSDCPGRCICRGNGYCG
B11	1.07	6.68	0.16	PGCGSLGPDCRRDSDCDKACICRGNGYCG
D11	1.02	6.43	0.16	GECQNNRTQEEGGRQCRRDSDCQGSCICRGNGYCG
E09	1.04	6.7	0.16	GHCNKTTRQCRRDSDCPIECICRGNGYCG
E05	1.12	7.35	0.15	IRCKYGSQQCRRDSDCPGHCICRGNGYCG
A05	1.28	8.79	0.15	TQCSDPRYWSQQCRRDSDCGGACICRGNGYCG
C12	1.02	7.09	0.14	HECSRSQPKCRRDSDCPLACICRGNGYCG
D02	1.29	9.29	0.14	EQCEHKRPERASDGDCRRDSDCPVACICRGNGYCG
E01	1.03	7.5	0.14	TGCTQKDSHPPNCRRDSDCPLACICRGNGYCG
A07	1.13	8.97	0.13	KSYNSNEYNCRRDSDCPAACICRGNGYCG
C01	0.99	7.98	0.12	SQCLDMPNQSETLPPCRRDSDCPRACICRGNGYCG
B01	1.02	8.49	0.12	HKCPEYGLTTRTCRRDSDCPAECICRGNGYCG
B12	1.05	8.97	0.12	MHCDDEPTTTKACRRDSDCEKNCICRGNGYCG
C02	1.08	11.28	0.10	EHCTMPEHQCRRDSDCGGQCICRGNGYCG
A01	0.93	0	0	KRCNRNDRWYYPCRRDSDCPGRCICRGNGYCG
A04	1.01	0	0	sequence failure
A06	0.97	0	0	GKCGKVPSMDRTNQLPSRLRLPWRLYLPW
B06	1.00	0	0	KYCGPPEYHCRRDSDCYRSCICRGNGYCG
C08	0.98	0	0	sequence failure
C09	0.99	0	0	RTCRERVEYTLQCRRDSDCTIDCICRGNGYCG
E07	1.09	0	0	KQCQKNGPQQSDYHQCRRDSDCDWACICRGNGYCG

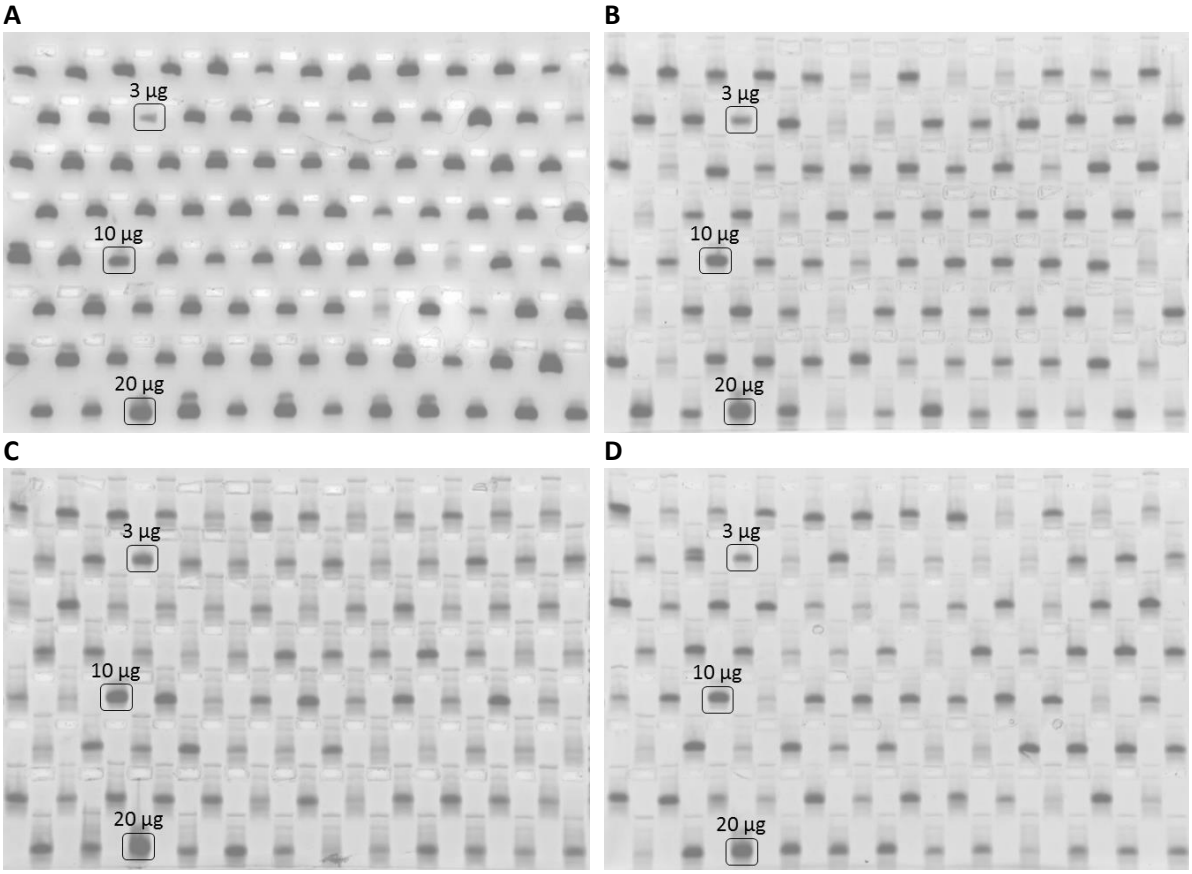


Figure 5.2 Analysis of high-throughput production of Trx-cystine-knot miniprotein fusions. E-PAGE™ analysis of produced Trx-cystine-knot miniprotein fusions from MCopt 1.0 library **(A)**, MCopt 2.0 library **(B)**, MCopt 2.1 library **(C)** and MCopt 3.0 library **(D)** in a 96-well format. Bacterial supernatant containing roughly purified Trx-cystine-knot miniprotein fusions were applied in a volume of 10 µL for E-PAGE™ analysis. Trx-MC-Myc-010 was used as protein reference at 3 µg, 10 µg and 20 µg (black borders).

Table 5.3 Amino acid sequences of MCopt 1.0, MCopt 2.0, MCopt 2.1 and MCopt 3.0 library clones of protein production and mass analysis.

Library clone	Amino acid sequence
MCopt 1.0 clone 1	GSAHDCPGDSPNCRRDSDCSVACICRGNGYCG
MCopt 1.0 clone 2	GSAMNCNTDPSRYGLCRRDSDCRGDCICRGNGYCG
MCopt 1.0 clone 3	GSAQVCYYGPEDSNYYPYCRRDSDCNLACICRGNGYCG
MCopt 2.0 clone 1	GSGSVCMSWAIDQETMCRRDSDCPGACICRGNGYCG
MCopt 2.0 clone 2	GSGSVCQWVSQKQYNLCRRDSDCPGACICRGNGYCG
MCopt 2.0 clone 3	GSGSVCTVWSWMGAEKCRRDSDCPGACICRGNGYCG
MCopt 2.0 clone 4	GSGSVCAQEVQTYPWQCRRDSDCPGACICRGNGYCG
MCopt 2.0 clone 5	GSGSVCTPRVALGQGDCRRDSDCPGACICRGNGYCG
MCopt 2.1 clone 1	GSGSVCPAILIRCRRDSDCPGACICGWKQKRPSLPCG
MCopt 2.1 clone 2	GSGSVCPAILIRCRRDSDCPGACICGRWSLDRRVECG
MCopt 2.1 clone 3	GSGSVCPAILIRCRRDSDCPGACICDIGWNSNLFMCG
MCopt 2.1 clone 4	GSGSVCPAILIRCRRDSDCPGACICHYDSEWRDVICG
MCopt 3.0 clone 1	GSGSVCHHPLIWSMPTCRRDSDCPGACICHNGPWFGEMECG
MCopt 3.0 clone 2	GSGSVCTPLAEFEQIRCRRDSDCPGACICDSSVDMMMPWCG
MCopt 3.0 clone 3	GSGSVCHPDIQQLYWMCRRDSDCPGACICWHAQEVVSFQCG

Table 5.4 Amino acid sequence of EDB ligand MC-FN-010 and its alanine scanning variants. Respective amino acids in the first, second or fifth loop of parental MC-FN-010 sequence were exchanged with alanine. Alanine substitutions are highlighted in bold letter.

Construct name	Amino acid sequence
MC-FN-010 (parental sequence)	PMCTQRKNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-011	**A**MCTQRKNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-012	P**A**CTQRKNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-013	PMC**A**QRKNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-014	PMCT**A**RKNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-015	PMCTQ**A**KNRIRLCRRDSDCTGACICRGNGYCG
MC-FN-016	PMCTQR**A**NRIRLCRRDSDCTGACICRGNGYCG
MC-FN-017	PMCTQRK**A**RIRLCRRDSDCTGACICRGNGYCG
MC-FN-018	PMCTQRKN**A**IRLCRRDSDCTGACICRGNGYCG
MC-FN-019	PMCTQRKNR**A**RLCRRDSDCTGACICRGNGYCG
MC-FN-0110	PMCTQRKNRI**A**LCRRDSDCTGACICRGNGYCG
MC-FN-0111	PMCTQRKNRIR**A**CRRDSDCTGACICRGNGYCG
MC-FN-0112	PMCTQRKNRIRLC**A**RDSDCTGACICRGNGYCG
MC-FN-0113	PMCTQRKNRIRLCR**A**DSDCTGACICRGNGYCG
MC-FN-0114	PMCTQRKNRIRLCRRDSDCTGACIC**A**GNGYCG

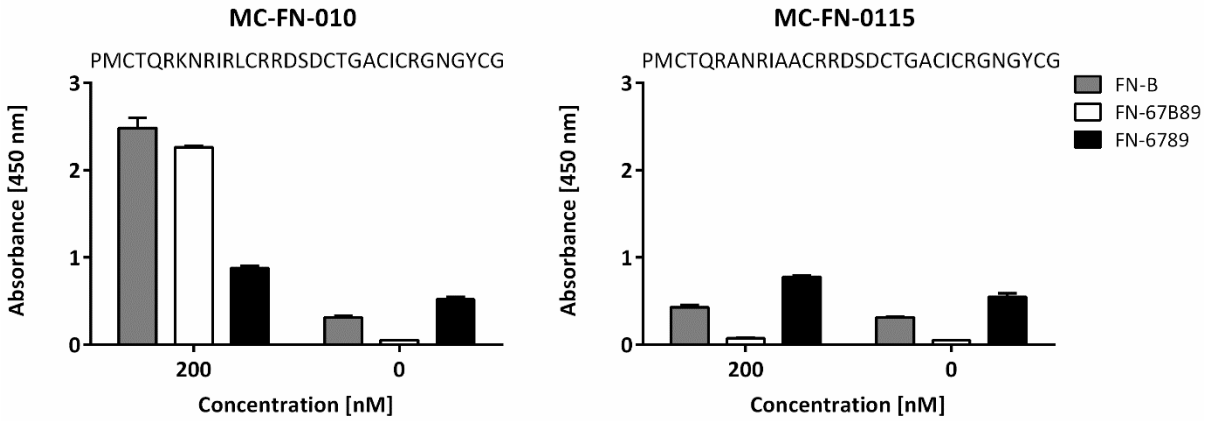


Figure 5.3 Binding analysis of MC-FN-010 and negative control MC-FN-0115. ELISA-based binding analysis of Trx-cystine-knot miniproteins with applied concentration at 200 nM to coated FN-B and FN-67B89 target proteins and to the control protein FN-6789 (coated at 1 µg/well). Binding was detected with 10 ng of HRP-conjugated anti-S-tag antibody. ELISA was performed in duplicates (± SD).

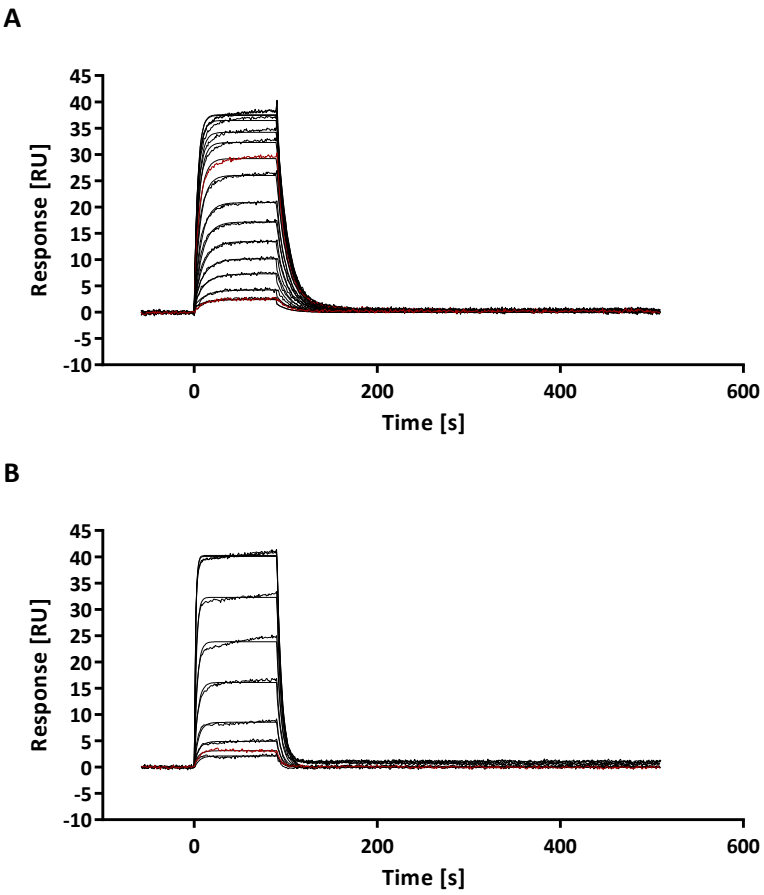


Figure 5.4 SPR analysis of MC-FN-010 and and MC-FN-016. Sensorgrams resulting from multi-cycle kinetic analysis of monomeric MC-FN-010 **(A)** and MC-FN-016 **(B)** binding to biotinylated FN-67B89 captured on SA sensor chip. Concentrations ranging from 50 nM to 4000 nM for MC-FN-010 and 62.5 nM to 4000 nM for MC-FN-016 (curves from bottom down to top) were applied. The curves were corrected to a blank SA control flow cell. Kinetic parameters obtained by applying a 1:1 Langmuir fitting model are shown in Table 3.3.

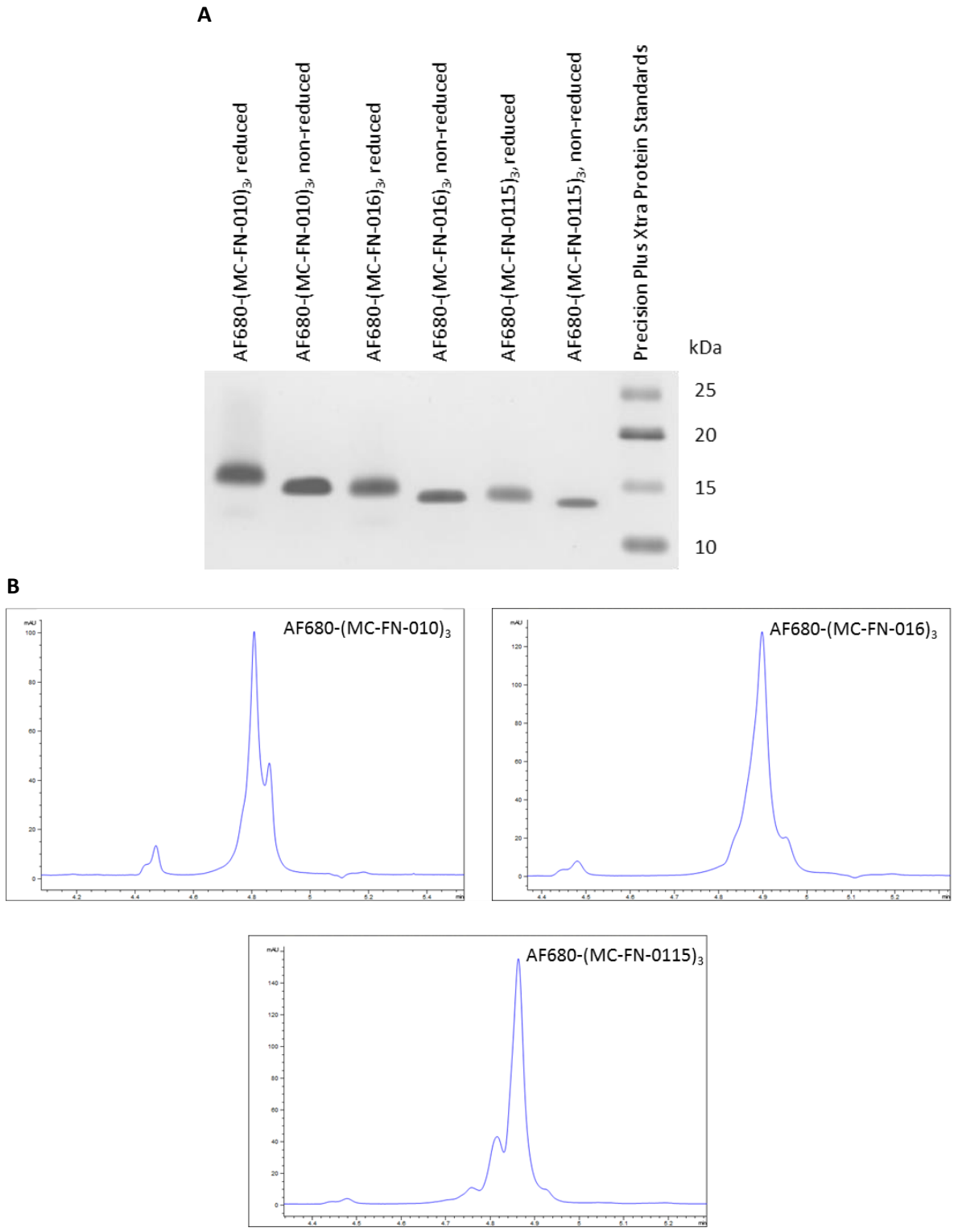


Figure 5.5 Quality control of AF680-(MC-FN-010)$_3$, AF680-(MC-FN-016)$_3$ and AF680-(MC-FN-0115)$_3$. **(A)** Analytical SDS-PAGE of AF680-(MC-FN-010)$_3$, AF680-(MC-FN-016)$_3$ and AF680-(MC-FN-0115)$_3$. Proteins were applied at 1 µg to SDS-PAGE under reducing (with β-mercaptoethanol) and non-reducing (without β-mercaptoethanol) conditions. AF680-(MC-FN-010)$_3$ has a molecular weight of 13681.9 Da, AF680-(MC-FN-016)$_3$ of 13510.6 Da and AF680-(MC-FN-0115)$_3$ of 13129 Da. Contrast and brightness has been altered in this image for better visualization. **(B)** RPC chromatograms of AF680-(MC-FN-010)$_3$, AF680-(MC-FN-016)$_3$ and AF680-(MC-FN-0115)$_3$. A total of 2 µg was applied to a HPLC column Poroshell 120 (EC-C18) in acetonitrile 2 % and 0.05 % TFA buffer at a flow rate of 2 mL/min. Proteins were analyzed spectroscopically via absorbance measurement at 280 nm.

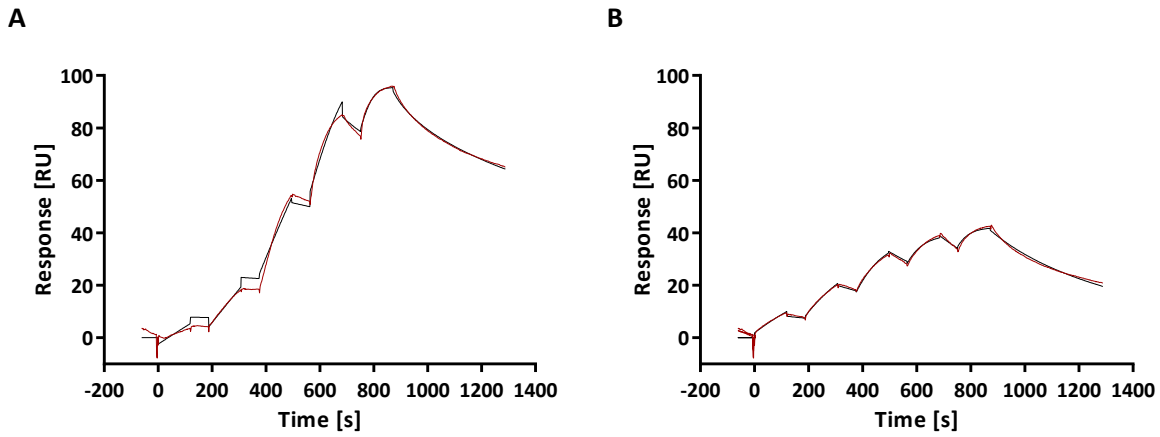


Figure 5.6 SPR analysis of AF680-(MC-FN-010)$_3$ and AF680-(MC-FN-016)$_3$. Sensorgrams arises from a single-cycle kinetic analysis of AF680-(MC-FN-010)$_3$ **(A)** and AF680-(MC-FN-016)$_3$ **(B)** binding to biotinylated FN-67B89 captured on SA sensor chip. Different concentrations of trimeric constructs from 1.25 – 10 nM were subsequently analyzed. The sensorgrams were referenced to a blank SA control flow cell. Kinetic parameters obtained by applying a 1:1 Langmuir fitting model are depicted in Table 3.4.

6 References

1. Sinha, S. & Vohora, D. in *Pharmaceutical Medicine and Translational Clinical Research* (Elsevier2018), pp. 19–32.
2. Hoelder, S., Clarke, P. A. & Workman, P. Discovery of small molecule cancer drugs: successes, challenges and opportunities. *Molecular oncology* **6,** 155–176; 10.1016/j.molonc.2012.02.004 (2012).
3. Buffery, D. The 2015 Oncology Drug Pipeline: Innovation Drives the Race to Cure Cancer. *American health & drug benefits* **8,** 216–222 (2015).
4. Kinch, M. S. & Griesenauer, R. H. 2017 in review: FDA approvals of new molecular entities. *Drug discovery today* **23,** 1469–1473; 10.1016/j.drudis.2018.05.011 (2018).
5. Kinch, M. S. & Woodard, P. K. Analysis of FDA-approved imaging agents. *Drug discovery today* **22,** 1077–1083; 10.1016/j.drudis.2017.03.006 (2017).
6. Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *British journal of pharmacology* **162,** 1239–1249; 10.1111/j.1476-5381.2010.01127.x (2011).
7. Imming, P., Sinning, C. & Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nature reviews. Drug discovery* **5,** 821–834; 10.1038/nrd2132 (2006).
8. Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nature reviews. Cancer* **12,** 278–287; 10.1038/nrc3236 (2012).
9. Lindsay, M. A. Target discovery. *Nature reviews. Drug discovery* **2,** 831–838; 10.1038/nrd1202 (2003).
10. Deepak, S. *et al.* Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes. *Current genomics* **8,** 234–251 (2007).
11. Jiang, Z. *et al.* Whole transcriptome analysis with sequencing: methods, challenges and potential solutions. *Cellular and molecular life sciences : CMLS* **72,** 3425–3439; 10.1007/s00018-015-1934-y (2015).
12. Chiu, M. L. & Gilliland, G. L. Engineering antibody therapeutics. *Current opinion in structural biology* **38,** 163–173; 10.1016/j.sbi.2016.07.012 (2016).
13. Lounnas, V. *et al.* Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Computational and structural biotechnology journal* **5,** e201302011; 10.5936/csbj.201302011 (2013).
14. Nixon, A. E., Sexton, D. J. & Ladner, R. C. Drugs derived from phage display: from candidate identification to clinical practice. *mAbs* **6,** 73–85; 10.4161/mabs.27240 (2014).
15. Cherf, G. M. & Cochran, J. R. Applications of Yeast Surface Display for Protein Engineering. *Methods in molecular biology (Clifton, N.J.)* **1319,** 155–175; 10.1007/978-1-4939-2748-7_8 (2015).
16. Groves, M. A. T. & Osbourn, J. K. Applications of ribosome display to antibody drug discovery. *Expert opinion on biological therapy* **5,** 125–135; 10.1517/14712598.5.1.125 (2005).

17. Omidfar, K. & Daneshpour, M. Advances in phage display technology for drug discovery. *Expert opinion on drug discovery* **10,** 651–669; 10.1517/17460441.2015.1037738 (2015).

18. Molek, P., Strukelj, B. & Bratkovic, T. Peptide phage display as a tool for drug discovery: targeting membrane receptors. *Molecules (Basel, Switzerland)* **16,** 857–887; 10.3390/molecules16010857 (2011).

19. Keseru, G. M. & Makara, G. M. Hit discovery and hit-to-lead approaches. *Drug discovery today* **11,** 741–748; 10.1016/j.drudis.2006.06.016 (2006).

20. Black, C. B., Duensing, T. D., Trinkle, L. S. & Dunlay, R. T. Cell-based screening using high-throughput flow cytometry. *Assay and drug development technologies* **9,** 13–20; 10.1089/adt.2010.0308 (2011).

21. Amore, B. M., Gibbs, J. P. & Emery, M. G. Application of in vivo animal models to characterize the pharmacokinetic and pharmacodynamic properties of drug candidates in discovery settings. *Combinatorial chemistry & high throughput screening* **13,** 207–218 (2010).

22. van Norman, G. A. Drugs, Devices, and the FDA. Part 1. *JACC: Basic to Translational Science* **1,** 170–179; 10.1016/j.jacbts.2016.03.002 (2016).

23. Kummar, S. *et al.* Phase 0 clinical trials: conceptions and misconceptions. *Cancer journal (Sudbury, Mass.)* **14,** 133–137; 10.1097/PPO.0b013e318172d6f3 (2008).

24. Salzberg, M. First-in-Human Phase 1 Studies in Oncology: The New Challenge for Investigative Sites. *Rambam Maimonides medical journal* **3,** e0007; 10.5041/RMMJ.1074 (2012).

25. van Norman, G. A. Drugs and Devices Comparison of European and U.S. Approval Processes. *JACC: Basic to Translational Science* **1,** 399–412; 10.1016/j.jacbts.2016.06.003 (2016).

26. Smith, G. P. Filamentous fusion phage. novel expression vectors that display cloned antigens on the virion surface. *Science* **228,** 1315–1317 (1985).

27. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348,** 552–554; 10.1038/348552a0 (1990).

28. Bass, S., Greene, R. & Wells, J. A. Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* **8,** 309–314; 10.1002/prot.340080405 (1990).

29. Lowman, H. B., Bass, S. H., Simpson, N. & Wells, J. A. Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* **30,** 10832–10838; 10.1021/bi00109a004 (1991).

30. Barbas, C. F., Burton, D. R., Scott, J. K. & Silverman, G. J. *Phage Display* (Cold Spring Harbor Laboratory Pr, 2001).

31. Pande, J., Szewczyk, M. M. & Grover, A. K. Phage display: concept, innovations, applications and future. *Biotechnology advances* **28,** 849–858; 10.1016/j.biotechadv.2010.07.004 (2010).

32. Azzazy, H. M. & Highsmith, W. Phage display technology. Clinical applications and recent innovations. *Clinical Biochemistry* **35,** 425–445; 10.1016/S0009-9120(02)00343-0 (2002).

33. Paschke, M. Phage display systems and their applications. *Applied microbiology and biotechnology* **70,** 2–11; 10.1007/s00253-005-0270-9 (2006).

34. Hufton, S. E. *et al.* Phage display of cDNA repertoires: the pVI display system and its applications for the selection of immunogenic ligands. *Journal of immunological methods* **231,** 39–51 (1999).

35. Løset, G. Å. & Sandlie, I. Next generation phage display by use of pVII and pIX as display scaffolds. *Methods (San Diego, Calif.)* **58,** 40–46; 10.1016/j.ymeth.2012.07.005 (2012).

36. Sidhu, S. S. Engineering M13 for phage display. *Biomolecular engineering* **18,** 57–63; 10.1016/S1389-0344(01)00087-9 (2001).

37. Karlsson, F., Borrebaeck, C. A. K., Nilsson, N. & Malmborg-Hager, A.-C. The mechanism of bacterial infection by filamentous phages involves molecular interactions between TolA and phage protein 3 domains. *Journal of Bacteriology* **185,** 2628–2634 (2003).

38. Rakonjac, J., Bennett, N. J., Spagnuolo, J., Gagic, D. & Russel, M. Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Current issues in molecular biology* **13,** 51–76 (2011).

39. Rahbarnia, L. *et al.* Evolution of phage display technology: from discovery to application. *Journal of drug targeting* **25,** 216–224; 10.1080/1061186X.2016.1258570 (2017).

40. Crameri, R. & Suter, M. Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene* **137,** 69–75 (1993).

41. Crameri, R., Jaussi, R., Menz, G. & Blaser, K. Display of expression products of cDNA libraries on phage surfaces. A versatile screening system for selective isolation of genes by specific gene-product/ligand interaction. *European Journal of Biochemistry* **226,** 53–58 (1994).

42. Weichel, M., Jaussi, R., Rhyner, C. & Crameri, R. Display of E. coli Alkaline Phosphatase pIII or pVIII Fusions on Phagemid Surfaces Reveals Monovalent Decoration with Active Molecules. *The open biochemistry journal* **2,** 38–43; 10.2174/1874091X00802010038 (2008).

43. O'Neil, K. T. & Hoess, R. H. Phage display: protein engineering by directed evolution. *Current opinion in structural biology* **5,** 443–449 (1995).

44. Sidhu, S. S., Lowman, H. B., Cunningham, B. C. & Wells, J. A. Phage display for selection of novel binding peptides. *Methods in enzymology* **328,** 333–363 (2000).

45. Zhao, A., Tohidkia, M. R., Siegel, D. L., Coukos, G. & Omidi, Y. Phage antibody display libraries: a powerful antibody discovery platform for immunotherapy. *Critical reviews in biotechnology* **36,** 276–289; 10.3109/07388551.2014.958978 (2014).

46. Kruif, J. de, Boel, E. & Logtenberg, T. Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions. *Journal of molecular biology* **248,** 97–105; 10.1006/jmbi.1995.0204 (1995).

47. Vaughan, T. J. *et al.* Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nature Biotechnology* **14,** 309–314; 10.1038/nbt0396-309 (1996).

48. Hoogenboom, H. R. *et al.* Antibody phage display technology and its applications. *Immunotechnology : an international journal of immunological engineering* **4,** 1–20 (1998).

49. Chassagne, S. *et al.* A high-affinity macaque antibody Fab with human-like framework regions obtained from a small phage display immune library. *Molecular immunology* **41,** 539–546; 10.1016/j.molimm.2004.03.040 (2004).

50. Frenzel, A., Schirrmann, T. & Hust, M. Phage display-derived human antibodies in clinical development and therapy. *mAbs* **8,** 1177–1194; 10.1080/19420862.2016.1212149 (2016).

51. Nord, K. *et al.* Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nature Biotechnology* **15,** 772–777; 10.1038/nbt0897-772 (1997).

52. Beste, G., Schmidt, F. S., Stibora, T. & Skerra, A. Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 1898–1903 (1999).

53. Steiner, D., Forrer, P. & Plückthun, A. Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *Journal of molecular biology* **382,** 1211–1227; 10.1016/j.jmb.2008.07.085 (2008).

54. Grabulovski, D., Kaspar, M. & Neri, D. A novel, non-immunogenic Fyn SH3-derived binding protein with tumor vascular targeting properties. *The Journal of biological chemistry* **282,** 3196–3204; 10.1074/jbc.M609211200 (2007).

55. Chance, R. E. & Frank, B. H. Research, development, production, and safety of biosynthetic human insulin. *Diabetes care* **16 Suppl 3,** 133–142 (1993).

56. Walsh, G. Biopharmaceutical benchmarks 2014. *Nature Biotechnology* **32,** 992–1000; 10.1038/nbt.3040 (2014).

57. Krejsa, C., Rogge, M. & Sadee, W. Protein therapeutics: new applications for pharmacogenetics. *Nature reviews. Drug discovery* **5,** 507–521; 10.1038/nrd2039 (2006).

58. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nature reviews. Drug discovery* **7,** 21–39; 10.1038/nrd2399 (2008).

59. Padma, V. V. An overview of targeted cancer therapy. *BioMedicine* **5,** 19; 10.7603/s40681-015-0019-4 (2015).

60. Ecker, D. M., Jones, S. D. & Levine, H. L. The therapeutic monoclonal antibody market. *mAbs* **7,** 9–14; 10.4161/19420862.2015.989042 (2015).

61. Reichert, J. M. Antibodies to watch in 2017. *mAbs* **9,** 167–181; 10.1080/19420862.2016.1269580 (2017).

62. Kaplon, H. & Reichert, J. M. Antibodies to watch in 2018. *mAbs* **10,** 183–203; 10.1080/19420862.2018.1415671 (2018).

63. Coulson, A., Levy, A. & Gossell-Williams, M. Monoclonal Antibodies in Cancer Therapy: Mechanisms, Successes and Limitations. *The West Indian medical journal* **63,** 650–654; 10.7727/wimj.2013.241 (2014).

64. Hoffman, W., Lakkis, F. G. & Chalasani, G. B Cells, Antibodies, and More. *Clinical journal of the American Society of Nephrology : CJASN* **11,** 137–154; 10.2215/CJN.09430915 (2016).

65. Carter, P. J. Introduction to current and future protein therapeutics: a protein engineering perspective. *Experimental cell research* **317,** 1261–1269; 10.1016/j.yexcr.2011.02.013 (2011).

66. Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256,** 495–497 (1975).

67. Carter, P. J. Potent antibody therapeutics by design. *Nature reviews. Immunology* **6,** 343–357; 10.1038/nri1837 (2006).

68. Lonberg, N. Fully human antibodies from transgenic mouse and phage display platforms. *Current Opinion in Immunology* **20,** 450–459; 10.1016/j.coi.2008.06.004 (2008).

69. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nature protocols* **1,** 755–768; 10.1038/nprot.2006.94 (2006).

70. Chames, P., van Regenmortel, M., Weiss, E. & Baty, D. Therapeutic antibodies: successes, limitations and hopes for the future. *British journal of pharmacology* **157,** 220–233; 10.1111/j.1476-5381.2009.00190.x (2009).

71. Beckman, R. A., Weiner, L. M. & Davis, H. M. Antibody constructs in cancer therapy: protein engineering strategies to improve exposure in solid tumors. *Cancer* **109,** 170–179; 10.1002/cncr.22402 (2007).

72. Romer, T., Leonhardt, H. & Rothbauer, U. Engineering antibodies and proteins for molecular in vivo imaging. *Current opinion in biotechnology* **22,** 882–887; 10.1016/j.copbio.2011.06.007 (2011).

73. Nelson, A. L. Antibody fragments: hope and hype. *mAbs* **2,** 77–83 (2010).

74. Wurch, T., Pierré, A. & Depil, S. Novel protein scaffolds as emerging therapeutic proteins: from discovery to clinical proof-of-concept. *Trends in Biotechnology* **30,** 575–582; 10.1016/j.tibtech.2012.07.006 (2012).

75. Vazquez-Lombardi, R. *et al.* Challenges and opportunities for non-antibody scaffold drugs. *Drug discovery today* **20,** 1271–1283; 10.1016/j.drudis.2015.09.004 (2015).

76. Kintzing, J. R., Filsinger Interrante, M. V. & Cochran, J. R. Emerging Strategies for Developing Next-Generation Protein Therapeutics for Cancer Treatment. *Trends in Pharmacological Sciences* **37,** 993–1008; 10.1016/j.tips.2016.10.005 (2016).

77. Simeon, R. & Chen, Z. In vitro-engineered non-antibody protein therapeutics. *Protein & cell* **9,** 3–14; 10.1007/s13238-017-0386-6 (2017).

78. Fiedler, M. & Skerra, A. in *Handbook of Therapeutic Antibodies,* edited by S. Dübel & J. M. Reichert (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2014), pp. 435–474.

79. Farkas, H. & Varga, L. Ecallantide is a novel treatment for attacks of hereditary angioedema due to C1 inhibitor deficiency. *Clinical, cosmetic and investigational dermatology* **4,** 61–68; 10.2147/CCID.S10322 (2011).

80. Kintzing, J. R. & Cochran, J. R. Engineered knottin peptides as diagnostics, therapeutics, and drug delivery vehicles. *Current opinion in chemical biology* **34,** 143–150; 10.1016/j.cbpa.2016.08.022 (2016).

81. Kolmar, H. Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. *FEBS J.* **275,** 2684–2690; 10.1111/j.1742-4658.2008.06440.x (2008).

82. Craik, D. J., Daly, N. L., Bond, T. & Waine, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *Journal of molecular biology* **294,** 1327–1336; 10.1006/jmbi.1999.3383 (1999).

83. Ackerman, S. E., Currier, N. V., Bergen, J. M. & Cochran, J. R. Cystine-knot peptides: emerging tools for cancer imaging and therapy. *Expert review of proteomics* **11,** 561–572; 10.1586/14789450.2014.932251 (2014).

84. Sarah J. Moore, Cheuk Lun Leung & Jennifer R. Cochran. Knottins: disulfide-bonded therapeutic and diagnostic peptides. *Drug discovery today. Technologies* **9,** e1-e70; 10.1016/j.ddtec.2011.07.003 (2012).

85. Craik, D. J., Daly, N. L. & Waine, C. The cystine knot motif in toxins and implications for drug design. *Toxicon : official journal of the International Society on Toxinology* **39,** 43–60 (2001).

86. Chiche, L. *et al.* Squash inhibitors: from structural motifs to macrocyclic knottins. *Current protein & peptide science* **5,** 341–349 (2004).

87. Favel, A. *et al.* Protease inhibitors from Ecballium elaterium seeds. *International journal of peptide and protein research* **33,** 202–208 (1989).

88. Daly, N. L. *et al.* Structural insights into the role of the cyclic backbone in a squash trypsin inhibitor. *The Journal of biological chemistry* **288,** 36141–36148; 10.1074/jbc.M113.528240 (2013).

89. Le, N. D. *et al.* Molecular recognition between serine proteases and new bioactive microproteins with a knotted structure. *Biochimie* **72,** 431–435 (1990).

90. Kimura, R. H., Levin, A. M., Cochran, F. V. & Cochran, J. R. Engineered cystine knot peptides that bind alphavbeta3, alphavbeta5, and alpha5beta1 integrins with low-nanomolar affinity. *Proteins* **77,** 359–369; 10.1002/prot.22441 (2009).

91. Kimura, R. H. *et al.* Pharmacokinetically stabilized cystine knot peptides that bind alpha-v-beta-6 integrin with single-digit nanomolar affinities for detection of pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **18,** 839–849; 10.1158/1078-0432.CCR-11-1116 (2012).

92. Glotzbach, B. *et al.* Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1. *PLoS ONE* **8,** e76956; 10.1371/journal.pone.0076956 (2013).

93. Silverman, A. P., Levin, A. M., Lahti, J. L. & Cochran, J. R. Engineered cystine-knot peptides that bind alpha(v)beta(3) integrin with antibody-like affinities. *Journal of molecular biology* **385,** 1064–1075; 10.1016/j.jmb.2008.11.004 (2009).

94. Kimura, R. H., Miao, Z., Cheng, Z., Gambhir, S. S. & Cochran, J. R. A dual-labeled knottin peptide for PET and near-infrared fluorescence imaging of integrin expression in living subjects. *Bioconjugate chemistry* **21,** 436–444; 10.1021/bc9003102 (2010).

95. Miao, Z. *et al.* An engineered knottin peptide labeled with 18F for PET imaging of integrin expression. *Bioconjugate chemistry* **20,** 2342–2347; 10.1021/bc900361g (2009).

96. Kimura, R. H., Cheng, Z., Gambhir, S. S. & Cochran, J. R. Engineered knottin peptides: a new class of agents for imaging integrin expression in living subjects. *Cancer Research* **69,** 2435–2442; 10.1158/0008-5472.CAN-08-2495 (2009).

97. Nielsen, C. H. *et al.* PET imaging of tumor neovascularization in a transgenic mouse model with a novel 64Cu-DOTA-knottin peptide. *Cancer Research* **70,** 9022–9030; 10.1158/0008-5472.CAN-10-1338 (2010).

98. Soroceanu, L., Gillespie, Y., Khazaeli, M. B. & Sontheimer, H. Use of chlorotoxin for targeting of primary brain tumors. *Cancer Research* **58,** 4871–4879 (1998).

99. Veiseh, M. *et al.* Tumor paint: a chlorotoxin:Cy5.5 bioconjugate for intraoperative visualization of cancer foci. *Cancer Research* **67,** 6882–6888; 10.1158/0008-5472.CAN-06-3948 (2007).

100. Mamelak, A. N. *et al.* Phase I single-dose study of intracavitary-administered iodine-131-TM-601 in adults with recurrent high-grade glioma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **24,** 3644–3650; 10.1200/JCO.2005.05.4569 (2006).

101. Fidel, J. *et al.* Preclinical Validation of the Utility of BLZ-100 in Providing Fluorescence Contrast for Imaging Spontaneous Solid Tumors. *Cancer Research* **75,** 4283–4291; 10.1158/0008-5472.CAN-15-0471 (2015).

102. Jiang, L. *et al.* Evaluation of a (64)Cu-labeled cystine-knot peptide based on agouti-related protein for PET of tumors expressing alphavbeta3 integrin. *J.Nucl.Med.* **51,** 251–258; 10.2967/jnumed.109.069831 (2010).

103. Pope, J. E. & Deer, T. R. Ziconotide: a clinical update and pharmacologic review. *Expert Opinion on Pharmacotherapy* **14,** 957–966; 10.1517/14656566.2013.784269 (2013).

104. Layer, P. & Stanghellini, V. Review article: Linaclotide for the management of irritable bowel syndrome with constipation. *Alimentary pharmacology & therapeutics* **39,** 371–384; 10.1111/apt.12604 (2014).

105. Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *Journal of cell science* **123,** 4195–4200; 10.1242/jcs.023820 (2010).

106. Ruoslahti, E. Fibronectin and its receptors. *Annual review of biochemistry* **57,** 375–413; 10.1146/annurev.bi.57.070188.002111 (1988).

107. Patel, R. S., Odermatt, E., Schwarzbauer, J. E. & Hynes, R. O. Organization of the fibronectin gene provides evidence for exon shuffling during evolution. *The EMBO journal* **6,** 2565–2572 (1987).

108. Proctor, R. A. Fibronectin: a brief overview of its structure, function, and physiology. *Reviews of infectious diseases* **9 Suppl 4,** S317-21 (1987).

109. Mao, Y. & Schwarzbauer, J. E. Fibronectin fibrillogenesis, a cell-mediated matrix assembly process. *Matrix biology : journal of the International Society for Matrix Biology* **24,** 389–399; 10.1016/j.matbio.2005.06.008 (2005).

110. Schwarzbauer, J. E., Patel, R. S., Fonda, D. & Hynes, R. O. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *The EMBO journal* **6,** 2573–2580 (1987).

111. Potts, J. R. & Campbell, I. D. Fibronectin structure and assembly. *Current Opinion in Cell Biology* **6,** 648–655 (1994).

112. Pankov, R. & Yamada, K. M. Fibronectin at a glance. *Journal of cell science* **115,** 3861–3863 (2002).

113. Tamkun, J. W. & Hynes, R. O. Plasma fibronectin is synthesized and secreted by hepatocytes. *The Journal of biological chemistry* **258,** 4641–4647 (1983).

114. To, W. S. & Midwood, K. S. Plasma and cellular fibronectin: distinct and independent functions during tissue repair. *Fibrogenesis & tissue repair* **4,** 21; 10.1186/1755-1536-4-21 (2011).

115. Wilson, C. L. & Schwarzbauer, J. E. The alternatively spliced V region contributes to the differential incorporation of plasma and cellular fibronectins into fibrin clots. *The Journal of cell biology* **119,** 923–933 (1992).

116. Wolanska, K. I. & Morgan, M. R. Fibronectin remodelling: cell-mediated regulation of the microenvironment. *Biochemical Society transactions* **43,** 122–128; 10.1042/BST20140313 (2015).

117. Paul, J. I., Schwarzbauer, J. E., Tamkun, J. W. & Hynes, R. O. Cell-type-specific fibronectin subunits generated by alternative splicing. *The Journal of biological chemistry* **261,** 12258–12265 (1986).

118. Schwarzbauer, J. E. Alternative splicing of fibronectin: three variants, three functions. *BioEssays : news and reviews in molecular, cellular and developmental biology* **13,** 527–533; 10.1002/bies.950131006 (1991).

119. White, E. S., Baralle, F. E. & Muro, A. F. New insights into form and function of fibronectin splice variants. *The Journal of pathology* **216,** 1–14; 10.1002/path.2388 (2008).

120. Gutman, A. & Kornblihtt, A. R. Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **84,** 7179–7182 (1987).

121. Neri, D. & Bicknell, R. Tumour vascular targeting. *Nature reviews. Cancer* **5,** 436–446; 10.1038/nrc1627 (2005).

122. Castellani, P. *et al.* The fibronectin isoform containing the ed-b oncofetal domain. A marker of angiogenesis. *Int. J. Cancer* **59,** 612–618; 10.1002/ijc.2910590507 (1994).

123. Birchler, M. T. *et al.* Expression of the extra domain B of fibronectin, a marker of angiogenesis, in head and neck tumors. *The Laryngoscope* **113,** 1231–1237 (2003).

124. Ohnishi, T. *et al.* Role of fibronectin-stimulated tumor cell migration in glioma invasion in vivo: clinical significance of fibronectin and fibronectin receptor expressed in human glioma tissues. *Clinical & Experimental Metastasis* **16,** 729–741 (1998).

125. Kaczmarek, J. *et al.* Distribution of oncofetal fibronectin isoforms in normal, hyperplastic and neoplastic human breast tissues. *International journal of cancer* **59,** 11–16 (1994).

126. Kumra, H. & Reinhardt, D. P. Fibronectin-targeted drug delivery in cancer. *Advanced Drug Delivery Reviews* **97,** 101–110; 10.1016/j.addr.2015.11.014 (2016).

127. Pini, A. *et al.* Design and use of a phage display library Human antibodies with subnanomolar affinity against a marker of angiogenesis eluted from a two-dimensional gel. *Journal of Biological Chemistry* **273,** 21769–21776 (1998).

128. Sauer, S. *et al.* Expression of the oncofetal ED-B-containing fibronectin isoform in hematologic tumors enables ED-B-targeted 131I-L19SIP radioimmunotherapy in Hodgkin lymphoma patients. *Blood* **113,** 2265–2274; 10.1182/blood-2008-06-160416 (2009).

129. Pasche, N. & Neri, D. Immunocytokines: a novel class of potent armed antibodies. *Drug discovery today* **17,** 583–590; 10.1016/j.drudis.2012.01.007 (2012).

130. Demartis, S., Tarli, L., Borsi, L., Zardi, L. & Neri, D. Selective targeting of tumour neovasculature by a radiohalogenated human antibody fragment specific for the ED-B domain of fibronectin. *European journal of nuclear medicine* **28,** 534–539 (2001).

131. Fabbrini, M. *et al.* Selective occlusion of tumor blood vessels by targeted delivery of an antibody-photosensitizer conjugate. *International journal of cancer. Journal international du cancer* **118,** 1805–1813; 10.1002/ijc.21412 (2006).

132. Han, Z. *et al.* EDB Fibronectin Specific Peptide for Prostate Cancer Targeting. *Bioconjugate chemistry* **26,** 830–838; 10.1021/acs.bioconjchem.5b00178 (2015).

133. Lorey, S. *et al.* Novel ubiquitin-derived high affinity binding proteins with tumor targeting properties. *J. Biol. Chem.* **289,** 8493–8507; 10.1074/jbc.M113.519884 (2014).

134. Gebauer, M., Schiefner, A., Matschiner, G. & Skerra, A. Combinatorial design of an Anticalin directed against the extra-domain b for the specific targeting of oncofetal fibronectin. *Journal of molecular biology* **425,** 780–802; 10.1016/j.jmb.2012.12.004 (2013).

135. Kim, S. *et al.* Bio-inspired design and potential biomedical applications of a novel class of high-affinity peptides. *Angewandte Chemie (International ed. in English)* **51,** 1890–1894; 10.1002/anie.201107894 (2012).

136. Albrecht, V. *et al.* Anticalins directed against the fibronectin extra domain B as diagnostic tracers for glioblastomas. *International journal of cancer* **138,** 1269–1280; 10.1002/ijc.29874 (2016).

137. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein science : a publication of the Protein Society* **8,** 921–929; 10.1110/ps.8.4.921 (1999).

138. Wu, S. C.-Y. *et al.* piggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 15008–15013; 10.1073/pnas.0606979103 (2006).

139. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein science : a publication of the Protein Society* **4,** 2411–2423; 10.1002/pro.5560041120 (1995).

140. Jones, M. L. *et al.* Targeting membrane proteins for antibody discovery using phage display. *Scientific reports* **6,** 26240; 10.1038/srep26240 (2016).

141. Alfaleh, M., Jones, M., Howard, C. & Mahler, S. Strategies for Selecting Membrane Protein-Specific Antibodies using Phage Display with Cell-Based Panning. *Antibodies* **6,** 10; 10.3390/antib6030010 (2017).

142. Kim, J. W., Cochran, F. V. & Cochran, J. R. A chemically cross-linked knottin dimer binds integrins with picomolar affinity and inhibits tumor cell migration and proliferation. *Journal of the American Chemical Society* **137,** 6–9; 10.1021/ja508416e (2015).

143. Kay, B. K., Thai, S. & Volgina, V. V. High-throughput biotinylation of proteins. *Methods in molecular biology (Clifton, N.J.)* **498,** 185–196; 10.1007/978-1-59745-196-3_13 (2009).

144. Avrutina, O. *et al.* Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II. *Biological chemistry* **386,** 1301–1306; 10.1515/BC.2005.148 (2005).

145. Mueller, M. Persistent bacterial infections: Identification of immunogenic structures of Borrelia burgdorferi sensu lato and Chlamydophila pneumoniae by phage surface display. *Dissertation* (2004).

146. Lahti, J. L., Silverman, A. P. & Cochran, J. R. Interrogating and predicting tolerated sequence diversity in protein folds: application to E. elaterium trypsin inhibitor-II cystine-knot miniprotein. *PLoS computational biology* **5,** e1000499; 10.1371/journal.pcbi.1000499 (2009).

147. Daly, N. L., Clark, R. J. & Craik, D. J. Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides. *The Journal of biological chemistry* **278,** 6314–6322; 10.1074/jbc.M210492200 (2003).

148. Kimura, R. H. *et al.* Functional mutation of multiple solvent-exposed loops in the Ecballium elaterium trypsin inhibitor-II cystine knot miniprotein. *PLoS ONE* **6,** e16112; 10.1371/journal.pone.0016112 (2011).

149. Heitz, A. *et al.* Knottin cyclization. impact on structure and dynamics. *BMC.Struct.Biol.* **8,** 54; 10.1186/1472-6807-8-54 (2008).

150. Goldsmith, M., Kiss, C., Bradbury, A. R. M. & Tawfik, D. S. Avoiding and controlling double transformation artifacts. *Protein engineering, design & selection : PEDS* **20,** 315–318; 10.1093/protein/gzm026 (2007).

151. Sblattero, D. & Bradbury, A. Exploiting recombination in single bacteria to make large phage antibody libraries. *Nature Biotechnology* **18,** 75–80; 10.1038/71958 (2000).

152. Brown, L. & Wan, H. Desmoglein 3: a help or a hindrance in cancer progression? *Cancers* **7,** 266–286; 10.3390/cancers7010266 (2015).

153. Bzymek, M. & Lovett, S. T. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **98,** 8319–8325; 10.1073/pnas.111008398 (2001).

154. Mariani, G. *et al.* Tumor targeting potential of the monoclonal antibody BC-1 against oncofetal fibronectin in nude mice bearing human tumor implants. *Cancer* **80,** 2378–2384 (1997).

155. Mohammadgholi, M. *et al.* Human Fibronectin Extra-Domain B (EDB)-Specific Aptide (APTEDB) Radiolabelling with Technetium-99m as a Potent Targeted Tumour-Imaging Agent. *Anti-cancer agents in medicinal chemistry*; 10.2174/1871520617666170918125020 (2017).

156. Andersen, D. C. & Krummen, L. Recombinant protein expression for therapeutic applications. *Current opinion in biotechnology* **13,** 117–123 (2002).

157. Wingfield, P. T. Overview of the purification of recombinant proteins. *Current protocols in protein science* **80,** 6.1.1-35; 10.1002/0471140864.ps0601s80 (2015).

158. Zhang, M.-Y. & Dimitrov, D. S. Sequential antigen panning for selection of broadly cross-reactive HIV-1-neutralizing human monoclonal antibodies. *Methods in molecular biology (Clifton, N.J.)* **562,** 143–154; 10.1007/978-1-60327-302-2_11 (2009).

159. Wang, F.-Y. *et al.* Selection of CC chemokine receptor 5-binding peptide from a phage display peptide library. *Bioscience, biotechnology, and biochemistry* **70,** 2035–2041; 10.1271/bbb.50654 (2006).

160. Shukla, G. S. & Krag, D. N. Phage display selection for cell-specific ligands: development of a screening procedure suitable for small tumor specimens. *Journal of drug targeting* **13,** 7–18; 10.1080/10611860400020464 (2005).

161. Moutel, S. *et al.* NaLi-H1: A universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *eLife* **5**; 10.7554/eLife.16228 (2016).

162. Wu, C.-H., Liu, I.-J., Lu, R.-M. & Wu, H.-C. Advancement and applications of peptide phage display technology in biomedical science. *Journal of biomedical science* **23,** 8; 10.1186/s12929-016-0223-x (2016).

163. Cain, S. A., Williams, D. M., Harris, V. & Monk, P. N. Selection of novel ligands from a whole-molecule randomly mutated C5a library. *Protein Engineering Design and Selection* **14,** 189–193; 10.1093/protein/14.3.189 (2001).

164. Zhao, S. *et al.* PiggyBac transposon vectors: the tools of the human gene encoding. *Translational lung cancer research* **5,** 120–125; 10.3978/j.issn.2218-6751.2016.01.05 (2016).

165. Zhou, H., Liu, Z.-G., Sun, Z.-W., Huang, Y. & Yu, W.-Y. Generation of stable cell lines by site-specific integration of transgenes into engineered Chinese hamster ovary strains using an FLP-FRT system. *Journal of Biotechnology* **147,** 122–129; 10.1016/j.jbiotec.2010.03.020 (2010).

166. Boettcher, M. & McManus, M. T. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular cell* **58,** 575–585; 10.1016/j.molcel.2015.04.028 (2015).

167. Cariuk, P., Gardener, M. J. & Vaughan, T. J. Evolution of biologics screening technologies. *Pharmaceuticals (Basel, Switzerland)* **6,** 681–688; 10.3390/ph6050681 (2013).

168. O'Kennedy, R., Byrne, M., O'Fagain, C. & Berns, G. A Review of Enzyme-Immunoassay and a Description of a Competitive Enzyme-Linked Immunosorbent Assay for the Detection of Immunoglobulin Concentrations. *Biochemical Education* **18,** 136–140; 10.1016/0307-4412(90)90219-E (1990).

169. Xu, M.-Y. *et al.* Production of a human single-chain variable fragment antibody against esophageal carcinoma. *World journal of gastroenterology* **10,** 2619–2623 (2004).

170. Hammers, C. M. & Stanley, J. R. Antibody phage display: technique and applications. *The Journal of investigative dermatology* **134,** 1–5; 10.1038/jid.2013.521 (2014).

171. Lequin, R. M. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). *Clinical chemistry* **51,** 2415–2418; 10.1373/clinchem.2005.051532 (2005).

172. Adan, A., Alizada, G., Kiraz, Y., Baran, Y. & Nalbant, A. Flow cytometry: basic principles and applications. *Critical reviews in biotechnology* **37,** 163–176; 10.3109/07388551.2015.1128876 (2017).

173. Yang, L. & Nolan, J. P. High-throughput screening and characterization of clones selected from phage display libraries. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **71,** 625–631; 10.1002/cyto.a.20417 (2007).

174. Bratkovic, T. Progress in phage display: evolution of the technique and its application. *Cellular and molecular life sciences : CMLS* **67,** 749–767 (2010).

175. Maaß, F. *et al.* Cystine-knot peptides targeting cancer-relevant human cytotoxic T lymphocyte-associated antigen 4 (CTLA-4). *J. Pept. Sci.*; 10.1002/psc.2782 (2015).

176. Chan, L. Y., Craik, D. J. & Daly, N. L. Dual-targeting anti-angiogenic cyclic peptides as potential drug leads for cancer therapy. *Scientific reports* **6,** 35347; 10.1038/srep35347 (2016).

177. Christmann, A., Walter, K., Wentzel, A., Krätzner, R. & Kolmar, H. The cystine knot of a squash-type protease inhibitor as a structural scaffold for Escherichia coli cell surface display of conformationally constrained peptides. *Protein engineering* **12,** 797–806 (1999).

178. Getz, J. A., Rice, J. J. & Daugherty, P. S. Protease-resistant peptide ligands from a knottin scaffold library. *ACS chemical biology* **6,** 837–844; 10.1021/cb200039s (2011).

179. Kabir, M. E., Krishnaswamy, S., Miyamoto, M., Furuichi, Y. & Komiyama, T. An improved phage-display panning method to produce an HM-1 killer toxin anti-idiotypic antibody. *BMC biotechnology* **9,** 99; 10.1186/1472-6750-9-99 (2009).

180. Thie, H., Schirrmann, T., Paschke, M., Dübel, S. & Hust, M. SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli. *New biotechnology* **25,** 49–54; 10.1016/j.nbt.2008.01.001 (2008).

181. Harrison, O. J. *et al.* Structural basis of adhesive binding by desmocollins and desmogleins. *Proceedings of the National Academy of Sciences of the United States of America* **113,** 7160–7165; 10.1073/pnas.1606272113 (2016).

182. Haque, A., Faizi, M. S. H., Rather, J. A. & Khan, M. S. Next generation NIR fluorophores for tumor imaging and fluorescence-guided surgery: A review. *Bioorganic & medicinal chemistry* **25,** 2017–2034; 10.1016/j.bmc.2017.02.061 (2017).

183. Johannsen, M. *et al.* The tumour-targeting human L19-IL2 immunocytokine: preclinical safety studies, phase I clinical trial in patients with solid tumours and expansion into patients with advanced renal cell carcinoma. *European journal of cancer (Oxford, England : 1990)* **46,** 2926–2935; 10.1016/j.ejca.2010.07.033 (2010).

184. Hemmerle, T. *et al.* The antibody-based targeted delivery of TNF in combination with doxorubicin eradicates sarcomas in mice and confers protective immunity. *British journal of cancer* **109,** 1206–1213; 10.1038/bjc.2013.421 (2013).

185. Tijink, B. M. *et al.* Radioimmunotherapy of head and neck cancer xenografts using 131I-labeled antibody L19-SIP for selective targeting of tumor vasculature. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* **47,** 1127–1135 (2006).

186. Danielli, R. *et al.* Intralesional administration of L19-IL2/L19-TNF in stage III or stage IVM1a melanoma patients: results of a phase II study. *Cancer immunology, immunotherapy : CII* **64,** 999–1009; 10.1007/s00262-015-1704-6 (2015).

187. Weide, B. *et al.* Intralesional treatment of stage III metastatic melanoma patients with L19-IL2 results in sustained clinical and systemic immunologic responses. *Cancer immunology research* **2,** 668–678; 10.1158/2326-6066.CIR-13-0206 (2014).

188. Fattorusso, R. *et al.* NMR structure of the human oncofoetal fibronectin ED-B domain, a specific marker for angiogenesis. *Structure* **7,** 381–390; 10.1016/S0969-2126(99)80051-3 (1999).

189. Schmidt, M. M. & Wittrup, K. D. A modeling analysis of the effects of molecular size and binding affinity on tumor targeting. *Mol.Cancer Ther.* **8,** 2861–2871; 10.1158/1535-7163.MCT-09-0195 (2009).

190. Zahnd, C. *et al.* Efficient tumor targeting with high-affinity designed ankyrin repeat proteins: effects of affinity and molecular size. *Cancer Research* **70,** 1595–1605; 10.1158/0008-5472.CAN-09-2724 (2010).

191. Moore, S. J., Leung, C. L., Norton, H. K. & Cochran, J. R. Engineering agatoxin, a cystine-knot peptide from spider venom, as a molecular probe for in vivo tumor imaging. *PloS one* **8,** e60498; 10.1371/journal.pone.0060498 (2013).

192. Zhu, X. *et al.* 99mTc-labeled cystine knot peptide targeting integrin αvβ6 for tumor SPECT imaging. *Molecular Pharmaceutics* **11,** 1208–1217; 10.1021/mp400683q (2014).

193. Butte, P. V. *et al.* Near-infrared imaging of brain tumors using the Tumor Paint BLZ-100 to achieve near-complete resection of brain tumors. *Neurosurgical focus* **36,** E1; 10.3171/2013.11.FOCUS13497 (2014).

194. Gunasekera, S. *et al.* Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. *Journal of medicinal chemistry* **51,** 7697–7704; 10.1021/jm800704e (2008).

195. Thongyoo, P., Bonomelli, C., Leatherbarrow, R. J. & Tate, E. W. Potent inhibitors of beta-tryptase and human leukocyte elastase based on the MCoTI-II scaffold. *Journal of medicinal chemistry* **52,** 6197–6200; 10.1021/jm901233u (2009).

196. Jiang, L. *et al.* Preliminary evaluation of (177)Lu-labeled knottin peptides for integrin receptor-targeted radionuclide therapy. *European journal of nuclear medicine and molecular imaging* **38,** 613–622; 10.1007/s00259-010-1684-x (2011).

197. Cox, N., Kintzing, J. R., Smith, M., Grant, G. A. & Cochran, J. R. Integrin-Targeting Knottin Peptide-Drug Conjugates Are Potent Inhibitors of Tumor Cell Proliferation. *Angewandte Chemie (International ed. in English)* **55,** 9894–9897; 10.1002/anie.201603488 (2016).

198. Currier, N. V. *et al.* Targeted Drug Delivery with an Integrin-Binding Knottin-Fc-MMAF Conjugate Produced by Cell-Free Protein Synthesis. *Molecular cancer therapeutics* **15,** 1291–1300; 10.1158/1535-7163.MCT-15-0881 (2016).

199. Moynihan, K. D. *et al.* Eradication of large established tumors in mice by combination immunotherapy that engages innate and adaptive immune responses. *Nature medicine* **22,** 1402–1410; 10.1038/nm.4200 (2016).

200. Kwan, B. H. *et al.* Integrin-targeted cancer immunotherapy elicits protective adaptive immune responses. *The Journal of Experimental Medicine* **214,** 1679–1690; 10.1084/jem.20160831 (2017).

Acknowledgement

Publication

Parts of this thesis (including some figures and text fragments) were submitted for publication:

LUI BG. Salomon N. Wüstehube-Lausch J. Daneschdar M. Schmoldt HU. Türeci Ö. Sahin U. **Targeting the tumor vasculature with engineered cystine-knot miniproteins** (submitted and under review).

Inventor within the frame of an international PCT (Patent Cooperation Treaty) patent application. Applicant: BioNTech AG; Investigator: LUI BG. Salomon N. Wüstehube-Lausch J. Daneschdar M. Schmoldt HU. Fiedler M. Sahin U. **Composition and methods for diagnosis and treatment of cancer.** International patent application number: PCT/EP2018/065205.

Bonny Gaby Lui

Curriculum vitae

Statutory declaration

I hereby declare that the submitted doctoral thesis „Targeting the Tumor Vasculature with an Engineered Cystine-Knot Miniprotein Applying an Optimized Phage Display Discovery Platform“ is, to the best of my knowledge and belief, in all parts my original work and that I have not received assistance from outside other than acknowledged. I have clearly indicated and referenced all used material and sources. This work has not been submitted, either substantially or in whole, for examination purposes at this or any other University before.

Mainz, …….

……
Bonny Gaby Lui