# Synthesis, purification and testing of point-modified messenger RNA

Dissertation zur Erlangung des Grades

“Doktor der Naturwissenschaften“

im Promotionsfach Pharmazie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

der Johannes Gutenberg-Universität Mainz

Jasmin Hertler

geboren am 08.02.1992

in Bad Soden am Taunus

Mainz, Januar 2022

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

Dekanin: [redacted]

1. Berichterstatter: [redacted]

2. Berichterstatter: [redacted]

Datum der mündlichen Prüfung: 02.03.2022

D77 (Dissertation Mainz)

Die vorgelegte Dissertation wurde am Institut für Pharmazeutische und Biomedizinische Wissenschaften der Johannes Gutenberg-Universität in Mainz zur Erlangung des Grades “Doktor der Naturwissenschaften“ unter der Betreuung von [redacted] angefertigt.

1. Berichterstatter

[redacted]

Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

Institut für Pharmazeutische und Biomedizinische Wissenschaften

Johannes Gutenberg-Universität Mainz

2. Berichterstatter

[redacted]

Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

Institut für Pharmazeutische und Biomedizinische Wissenschaften

Johannes Gutenberg-Universität Mainz

Hiermit versichere ich eidesstattlich:

1. Ich habe die jetzt als Dissertation vorgelegte Arbeit selbst angefertigt und alle benutzten Hilfsmittel (Literatur, Apparaturen, Material) in der Arbeit angegeben.

2. Ich habe oder hatte die jetzt als Dissertation vorgelegte Arbeit nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht.

3. Ich hatte weder die jetzt als Dissertation vorgelegte Arbeit noch Teile davon bei einer anderen Fakultät bzw. einem anderen Fachbereich als Dissertation eingereicht.

Ort, Datum

Jasmin Hertler

*„There will come a time when you believe everything is finished. That will be the beginning"*

Louis L'Amour

# Acknowledgements

# Contents

# Abstract

Ribonucleic acids (RNA) are involved in a huge variety of important cellular processes. They have been a subject of scientific research for decades now, and the huge variety of non-canonical nucleosides, namely modifications, that have been discovered, add another layer to the regulatory capabilities of the various RNA species. However, there are gaps in the knowledge regarding some aspects of point-modified messenger RNAs (mRNAs), as mechanistic studies have been performed mainly with short mRNA surrogates. This PhD thesis aims to develop a suitable synthesis method of point-modified mRNA to investigate the influence of modifications on translation efficiency.

The mRNA of the enhanced green fluorescent protein with coupled internal ribosomal entry site (IRES-eGFP) served as a model and was first divided retrosynthetically into several fragments, with the middle fragment intended to introduce a modification into the mRNA with point accuracy. Subsequently, the mRNA was reconstituted in a 3-way-one-pot splint ligation and purified via the newly developed method of real-time gel elution. The splint ligation approach was adapted from shorter RNA and adjusted to long RNA species. The purification method was investigated with respect to its practicability and safety against RNA damage, with particular emphasis on gel particles on the one hand and oxidative damage due to irradiation during elution on the other. The obtained point-modified mRNAs were analyzed with focus on the number and position of the respective modification. After confirmed correct incorporation of the desired modification and reconstitution of the full-length mRNA, the obtained modified mRNAs were translated in vitro and the protein yield was measured against the protein yield of an unmodified reference mRNA. In particular, 2'-O-methylations introduced into the Kozak sequence in and around the start codon revealed remarkable results for cap-independent IRES-driven translation, which had not been reported before.

# Zusammenfassung

Ribonukleinsäuren (RNA) sind in eine große Anzahl wichtiger zellulärer Prozesse involviert. Seit Jahrzehnten werden sie wissenschaftlich erforscht und die riesige Vielfalt nicht-kanonischer Nukeloside, d.h. Modifikationen, die entdeckt wurden, ergänzen die Regulationsmöglichkeiten der verschiedenen RNA Spezies um eine weitere Ebene. Es gibt jedoch Wissenslücken in Bezug auf einige Aspekte punktmodifizierter Boten-RNAs (mRNAs), da mechanistische Studien hauptsächlich mit kurzen mRNA-Surrogaten durchgeführt wurden. Diese Promotionsarbeit zielt deshalb auf dieses Gebiet ab und diente dazu eine geeignete Synthesemethode von Punkt-modifizierten mRNA zu entwickeln um den Einfluss von Modifikationen auf die Translation hin zu untersuchen. Die mRNA des verbesserten grün fluoreszierenden Proteins mit gekoppelter interner ribosomalen Eintrittsstelle (IRES-eGFP) wurde als Model verwendet und es erfolgte zunächst die retrosynthetische Aufspaltung in mehrere Fragmente, wobei das mittlere Fragment dazu diente, eine Modifikation punktgenau in die mRNA einzubringen. Anschließend wurde die mRNA in einer „3-way-one pot“ Splint Ligation rekonstituiert und über die neu entwickelte „real-time“ Gel Elution aufgereinigt. Dabei wurde ein Ansatz der Splint Ligation von kürzeren RNA übernommen und auf lange RNA Spezies angepasst. Die verwendete Aufreinigung wurde hinsichtlich ihrer Praktikabilität und Sicherheit gegenüber RNA Schädigungen untersucht, dabei wurde besonders Wert auf noch vorhandele Gel Partikel einerseits und oxidative Schädigungen durch die Bestrahlung während der Elution andererseits gelegt. Die erhaltenen Punkt-modifizierten mRNAs wurden hinsichtlich der Anzahl und Position der jeweiligen Modifikation hin analysiert. Nach bestätigtem, korrektem Einbau der gewünschten Modifikation und Rekonstitution der volllängen mRNA wurden die erhaltenen modifizierten mRNAs *in vitro* translatiert und die Proteinausbeute mit der einer unmodifizierten Referenz mRNA verglichen. Dabei zeigten sich besonders bei 2‘-O-methylierungen die in die Kozak Sequenz in und um das Startcodon eingebracht wurden erstaunliche Ergebnisse für die Cap-unabhängige IRES-getriebene Translation, die so bisher noch nicht berichtete worden waren.

# List of figures

# List of tables

# Abbreviations

| | |
|---|---|
| 8OxoG | 8-oxoguanosine |
| A | adenosine |
| *A. thaliana* | *Arabidopsis thaliana* |
| aa-tRNA | aminoacyl-tRNA |
| ABCF | ATP-binding cassette sub-family F |
| ALKBH5 | alpha-ketoglutarate-dependent hydroxylase homolog 5 |
| $A_m$ | 2'-O-methyladenosine |
| AMP | adenosine monophosphate |
| APS | ammonium persulfate |
| ATP | adenosine triphosphate |
| BSA | bovine serum albumin |
| C | cytidine |
| cDNA | complementary DNA |
| CDS | coding sequence |
| CF | cleavage factor |
| circRNA | circular RNA |
| CPSF | cleavage and polyadenylation specific factor |
| CrPV | cricklet paralysis virus |
| CSTF | cleavage stimulation factor |
| Cy5 | cyanine 5-succinimidylester |
| D | dihydrouridine |
| DAP5 | death-associated protein 5 |
| DCP2 | mRNA-decapping enzyme 2 |
| DEPC | diethyl pyrocarbonate |
| dMRM | dynamic multiple reaction monitoring |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| dNTP | deoxynucleoside triphosphate |

| | |
|---|---|
| dsRNA | double-stranded RNA |
| DTT | dithiothreitol |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylenediaminetetraacetic acid |
| EF-TU | elongation factor thermo unstable |
| eGFP | enhanced green fluorescent protein |
| eIF | eukaryotic translation initiation factor |
| EMCV | encephalomyocarditis virus |
| ESI | electrospray ionization |
| EUTP | 5-ethynyluridine triphosphate |
| FBS | fetal bovine serum |
| FTO | fat mass and obesity-associated protein |
| FTSJ3 | pre-rRNA 2'-O-ribose RNA methyltransferase 3 |
| G | guanosine |
| GFP | green fluorescent protein |
| Gly | glycine |
| $G_m$ | 2'-O-methylguanosine |
| GTP | guanosine triphosphate |
| HEK | human embryonic kidney |
| HILIC | hydrophilic interaction chromatography |
| HIV | human immunodeficiency viruses |
| HNRNP | heterogeneous nuclear ribonucleoprotein |
| hnRNP | heterogeneous ribonucleoprotein particle |
| HPLC | high performance liquid chromatography |
| I | inosine |
| IF | translation initiation factor |
| IFN | interferone |
| IRES | internal ribosome entry site |
| ITP | inosine triphosphate |
| IVT | *in vitro* transcript / transcription |
| KanR | kanamycin resistance |

| | |
|---|---|
| LC-MS/MS | liquid chromatography – Mass spectrometry |
| LMG | low melting/gelling |
| Lys | lysine |
| $m^1A$ | 1-methyladenosine |
| $m^1\Psi$ | $N^1$-methylpseudouridine |
| $m^3C$ | 3-methylcytidine |
| $m^5C$ | 5-methylcytidine |
| $m^5CTP$ | 5-methylcytidine triphosphate |
| $m^6A$ | $N^6$-methyladenosine |
| $m^6A_m$ | $N^6$,2'-O-dimethyladenosine |
| $m^6AMP$ | $N^6$-methyladenosine monophosphate |
| $m^6ATP$ | $N^6$-methyladenosine triphosphate |
| $m^7G$ | 7-methylguanosine |
| MDA5 | melanoma differentiation-associated protein 5 |
| Met | methionine |
| METTL | methyltransferase like |
| mRNA | messenger RNA |
| ncRNA | non-coding RNA |
| NDP | nucleoside diphosphate |
| NLS | neutral loss scan |
| $N_m$ | 2'-O-methylated nucleoside |
| NTP | nucleoside triphosphate |
| nts | nucleotides |
| ORF | open reading frame |
| ori | origin of replication |
| PABP | poly(A) binding protein |
| PAGE | polyacrylamide gel electrophoresis |
| PAP | poly(A) polymerase |
| pDNA | plasmid DNA |
| pol | polymerase |
| prM-E | premembrane and envelope protein |

| | |
|---|---|
| PRR | pattern recognition receptor |
| PS-DVB | polystyrene-divinylbenzene |
| Psi / Ψ | pseudouridine |
| ΨTP | pseudouridine triphosphate |
| PTB | polypyrimidine tract-binding protein |
| PUS | pseudouridine synthase |
| R.S. | restriction Site |
| Rat1 | 5' - 3' exoribonuclease 2 |
| Rbm15 | RNA-binding motif protein 15 |
| RIG-I | retinoic acid inducible gene I |
| RNA | ribonucleic acid |
| RNAP | RNA polymerase |
| RNase | ribonuclease |
| RNA-seq | RNA sequencing |
| ROS | reactive oxygen species |
| RP | reversed phase |
| RRL | rabbit reticulocyte lysate |
| rRNA | ribosomal RNA |
| *S. cerevisiae* | *Saccharomyces cerevisiae* |
| $s^2U$ | 2-thiouridine |
| SAM | S-adenosyl methionine |
| SDS | sodium dodecyl sulfate |
| Ser | serine |
| SILIS | stable isotope labeled internal standard |
| siRNA | small interfering RNA |
| snoRNA | small nucleolar RNA |
| snoRNP | small nucleolar ribonucleoprotein |
| SRSF | serine/arginine-rich splicing factor |
| ssRNA | single stranded RNA |
| T | thymidine |
| TBE | tris/borate/ethylenediaminetetraacetic acid |

| | |
|---|---|
| TEAA | triethylamoniumaccetate |
| TEMED | tetramethylethylenediamine |
| TFII | transcription factor for polymerase II |
| TLC | thin-layer chromatography |
| TLR | toll like receptor |
| TRBP | transactivation response element RNA-binding protein |
| tRNA | transfer RNA |
| Tyr | tyrosine |
| U | uridine |
| $U_m$ | 2'-O-methyluridine |
| UTR | untranslated region |
| UV | ultraviolet |
| VIMRA | vir like $m^6A$ methyltransferase associated |
| WTAP | wilms tumor 1-associated protein |
| Xrn1 | 5' - 3' exoribonuclease 1 |
| YBX1 | Y-Box binding protein 1 |
| YTHDC | YTH domain-containing proteins |
| YTHDF | YTH domain family members |

# 1 Introduction

## 1.1 It is an RNA world

The sequence of genetic information is reflected through three major classes of biopolymers, namely deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and proteins. This central dogma of molecular biology [1,2] describes a flow of information comparable to a one-way street, once information reaches the protein level, whereas information can be transferred between nucleic acids (DNA and RNA) [2]. Around 1953, the structure of DNA as a double helix was discovered [3], and later the structure and function of RNA was elucidated. While DNA occurs in principle as a double-stranded structure that forms the well-known double helix, RNA does not necessarily consist of two complementary strands. The RNA strand forms a heteropolymer composed of four canonical nucleobases linked by a sugar-phosphate backbone. These four nucleobases, namely adenine (A), guanine (G), cytosine (C) and uracil (U), can be chemically divided into purine nucleobases (adenine and guanine) and pyrimidine nucleobases (cytosine and uracil). In addition, the pyrimidine nucleobase thymine (T) replaces uracil in the context of DNA sequences. Figure 1.1 illustrates the difference between nucleosides (nitrogen-containing base attached to the ribose moiety via an N-glycosidic bond) and nucleotides (nucleoside with phosphate groups) and between the nucleosides uridine and thymidine. In addition, the overall structure of RNA includes possible double-stranded regions mediated by hydrogen bonds between base pairs A-U and G-C [4] depicted in the lower part of figure 1.1. After many years of research, several RNA species are known that combine these basic structural elements, but can also have complex secondary and tertiary structures, making them suitable for a variety of different functions, from the basics of protein

biosynthesis [4] and beyond. The latter include low abundance RNA species such as long non-coding RNA (long ncRNA), small nuclear RNA (snoRNA), small interfering RNA (siRNA) or circular RNA (circRNA) (reviewed in [5,6]). The best known and most abundant RNA classes are ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA). While only mRNAs belong to the coding RNA, tRNA and rRNA belong to the non-coding RNA class but are still essential for protein biosynthesis. Transfer RNAs, when interacting directly with mRNAs by base pairing, provide the respective amino acids for peptide formation, and ribosomal RNAs, together with ribosomal proteins, form the ribosome and thus the environment for protein biosynthesis [4]. As mentioned earlier, mRNAs belong to the coding class of RNA because they are direct transcripts of genomic information. They serve as templates for protein production by transporting genomic information from the nucleus in eukaryotes to the cytosol (where protein biosynthesis occurs) [7].

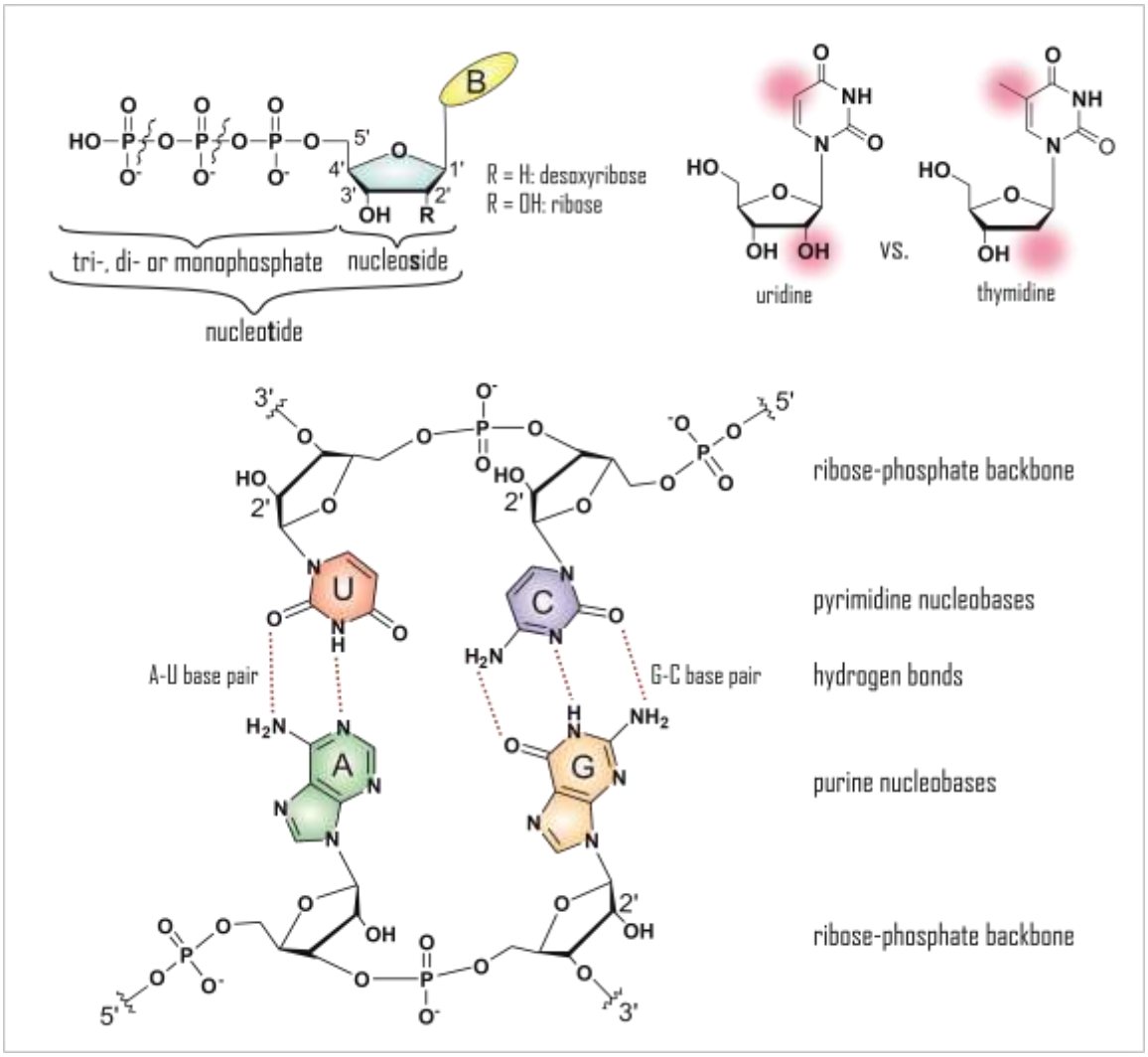


**Figure 1.1 Chemical structure of nucleosides, nucleotides and RNA base pairing.**
Chemical structure of an exemplary nucleoside triphosphate (upper left) and structural differences between uridine and thymidine (upper right), namely the 2'-hydroxyl group and the methyl group at position 5, were depicted in the upper part of the scheme. The lower part illustrates the covalent binding of the ribose-phosphate backbone and the possible hydrogen bonds between the canonical base pairs adenosine (A) – uridine (U) and guanosine (G) – cytidine (C).

## 1.2 Lifecycle of eukaryotic mRNA

As mentioned above, mRNAs are transcribed from genomic DNA in the nucleus of eukaryotes. This transcription is carried out by one of three different polymerases (pol I, pol II, and pol III). Polymerase II has been shown to be the most important polymerase because it is the only polymerase that can support maturation of nascent mRNA during transcription [8,9]. Important for the initiation of mRNA synthesis in many eukaryotic genes and the first described promoter element is the “TATAA” sequence element on the DNA, namely the “TATA box” [10], which is recognized by the transcription factor TFIID and initiates the transcription process of polymerase II [11-13]. Other transcription factors stabilize the transcription complex [11], are essential for selection of the required DNA site (TFIIB) or unwind the DNA double helix (subunit of TFIIF) allowing transcription to begin a few nucleotides downstream of the “TATA box” [14,15]. These maturation steps are crucial for mRNA export and include packaging into RNA/protein complexes (heterogeneous ribonucleoprotein particles, hnRNPs). Other important elements for regulating mRNA transcription and maturation are so-called cis-acting elements, i.e. sequence elements on nucleic acids in general that are recognized by trans-acting factors to regulate gene expression (such as the TATA box mentioned above). Trans-acting factors include proteins that bind to DNA to mediate transcription, but also proteins that bind to mRNA to mediate mRNA processing or translation (reviewed in [16-19]). Lei and colleagues proposed the idea of coupled transcription, processing, and packaging of mRNA [20], which changed the understanding of the two processes (transcription and maturation) from the idea of only post-transcriptional maturation to partial co-transcriptional maturation.

In addition to the described packaging of mRNA, the pathway from nascent to mature mRNA involves other very important steps, some of which are exemplified below. Maturation of pre-mRNA in eukaryotic systems involves processing of the 5' end (namely, attachment of a cap structure) as well as splicing and processing of the 3' end (attachment of a poly(A) tail) [21-23].

The cap structure is critical for canonical cap-dependent translation in eukaryotic cells and consists of a 7-methylguanosine ($m^7G$) bound to the mRNA via a 5' - 5' triphosphate linker (known as $m^7GpppN$, cap 0) [24].

Additional methylation of the first and/or second transcribed nucleoside by 2'-O-methyltransferase using S-adenosylmethionine (SAM) as a methyl donor [25] yields cap 01 ($m^7GpppN_mpN$), cap 02 ($m^7GpppNpN_m$), or cap 012 ($m^7GpppN_mpN_m$), respectively (cf. figure 1.2A) [26]. Prevention of degradation from the 3' end of the mRNA is achieved by a poly(A) sequence (poly(A) tail) that is also added to the nascent mRNA (cf. figure 1.2A) [22]. Processing of the 3' end begins with a cleavage step. Then, polyadenylation is initiated by recognition of the polyadenylation signal sequence “AAUAAA” by the cleavage and polyadenylation specificity factor (CPSF), followed by recruitment of the cleavage stimulation factor (CSTF) and cleavage factors I and II (CFI and CFII, respectively) [27-29]. After this cleavage step, poly(A) polymerase (PAP) adds the poly(A) tail to the mRNA. Both mRNA capping and polyadenylation are important for mRNA stability [30] and contribute to mRNA nuclear export [31] and translation mechanism [32,33] in eucaryotic cells.

Another important processing step includes shortening of the pre-mRNA (known as splicing). Since the pre-mRNA sequence contains coding and non-coding parts (exons and introns, respectively), the splicing mechanism removes the introns and joins the exons [34]. This also results in the possibility of “alternative splicing,” i.e. alternative inclusion (or exclusion) of certain exons. This increases the density of information from the same DNA segment to different mature mRNAs and thus to different possible polypeptide sequences [35,36].

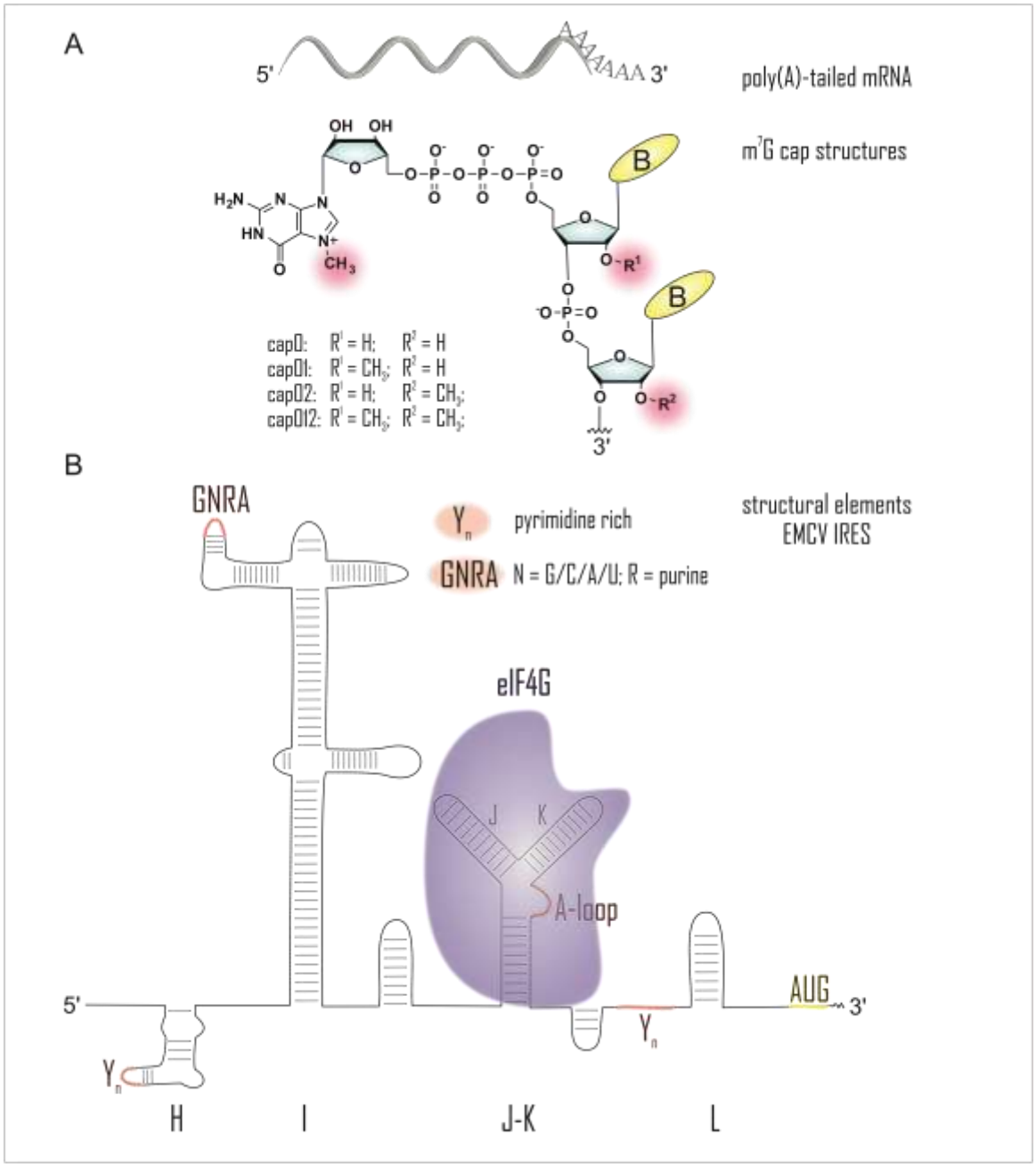


**Figure 1.2 Structural elements of mRNAs.**

**A** The poly(A) tail on the 3' end of mRNA as well as the $m^7G$ cap structure are depicted. The cap structure can be methylated on the two adjacent 2'-O positions ($R^1$ and $R^2$, respectively). The methylation on position 7 of guanine is highlighted. **B** Illustrated are some important structural elements of the EMCV IRES, namely the Regions H, I, J-K and L with pyrimidine rich regions $Y_n$ acting as protein binding site as well as the GNRA domain (N= G/C/A/U and R = purine). The J-K loop acts as binding region for eukaryotic initiation factor 4G and includes an A-loop. The start codon is depicted in yellow (AUG).

## 1.3 Translation in eukaryotes

After the mRNA is transported into the cytoplasm, the complexes of actively translating ribosomes are formed to direct the translation of the mRNAs into polypeptides and proteins. This complex formation involves different translation initiation factors, which differ between prokaryotic (IF) and eukaryotic (eIF) translation machinery. The canonical pathway of translation initiation involves the cap structure mentioned above to recruit eIF4F, a complex of different proteins involved in translation initiation. In addition to eIF4G as the scaffold protein and the RNA helicase eIF4A, eIF4F includes the cap-binding factor eIF4E, which interacts with the methyl group of the $m^7G$ cap structure. After this activation step of the mRNA, the 48S initiation complex is formed by binding of this mRNA complex to the 43S preinitiation complex. The 43S preinitiation complex, in turn, was formed by binding of the 40S ribosomal subunit to other initiation factors (e.g. eIF3) and the ternary complex eIF2-GTP-$tRNA^{Met}$. Once the 48S initiation complex is formed, it is able to “scan” the mRNA until it matches the start codon AUG. Marilyn Kozak first reported that not only the start codon, but also the surrounding sequence context is necessary to be recognized by the 48S complex in eukaryotes. This sequence was named after her (“Kozak sequence”) and describes the optimal sequence context 3 nucleotides (nts) upstream and 4 nts downstream of the “A” of the start codon (positions -3 to +4, respectively), namely RCCAUGG (R = A or G) [37]. Meanwhile, the poly(A)-binding protein (PABP) interacts with the poly(A) tail and is recruited by the aforementioned scaffold protein eIFF4, forming a “closed loop” structure of the mRNA (cf. figure 1.3) [38]. To form the final 80S ribosomal complex, some initiation factors are released while the 60S ribosomal subunit is recruited (reviewed in [39]).

### 1.3.1 Cap-independent translation

In contrast to the cap-dependent translation pathway mentioned above, it has been shown that up to 20 % of mammalian mRNAs can be translated in the absence of eIF4E (cap-binding protein) using other proteins (namely DAP5 and eIF3d) but still in a cap-dependent manner [40,41]. Translation initiation via a cap-independent mechanism is possible by recruitment of ribosomal subunits through an internal ribosome entry site (IRES) instead of the cap structure. These secondary structural elements of mRNA were first found in the viral genome, but later it was shown that several percent of eukaryotic mRNAs carry IRES elements and are translated mainly in the context of cellular stress [42-45] and also by using DAP5 [46,47]. One of the best-studied internal ribosome entry sites is the encephalomyocarditis virus IRES (EMCV IRES) of the picornavirus family, which is known to result in robust protein translation [48]. The authentic initiation site of EMCV IRES is the 11$^{th}$ AUG, counted from the 5' end of the RNA, and in rarely observed cases initiation by other AUG codons within the sequence was possible, depending on structural mutations of the IRES [49]. The sequence of the IRES includes several structural domains upstream of the initiation site, referred to as H, I, J-K, and L. Domains H and L serve as binding domains for various proteins with pyrimidine-rich motifs, e.g. polypyrimidine tract-binding protein (PTB) or eIF4B [50-52]. A GNRA motif with variable N-position and a purine base at position R in domain I enhances the structural integrity and function of the IRES motif [53-55]. The J-K domain comprises an oligo(A) loop and is recognized by the central domain of eIF4G, allowing subsequent binding of eIF4A to the formed complex (cf. figure 1.2B) [56,57]. This binding is an essential step for the assembly of the 48S complex and thus for the initiation of translation (cf. figure 1.3A) [58,59]. With the exception of eIF4E, virtually all other initiation factors are required for EMCV IRES-mediated translation [56,57,60], but this may be different for IRES motifs from other species such as Cricket paralysis virus IRES, which appears to be independent of all initiation factors (eIFs) [61].

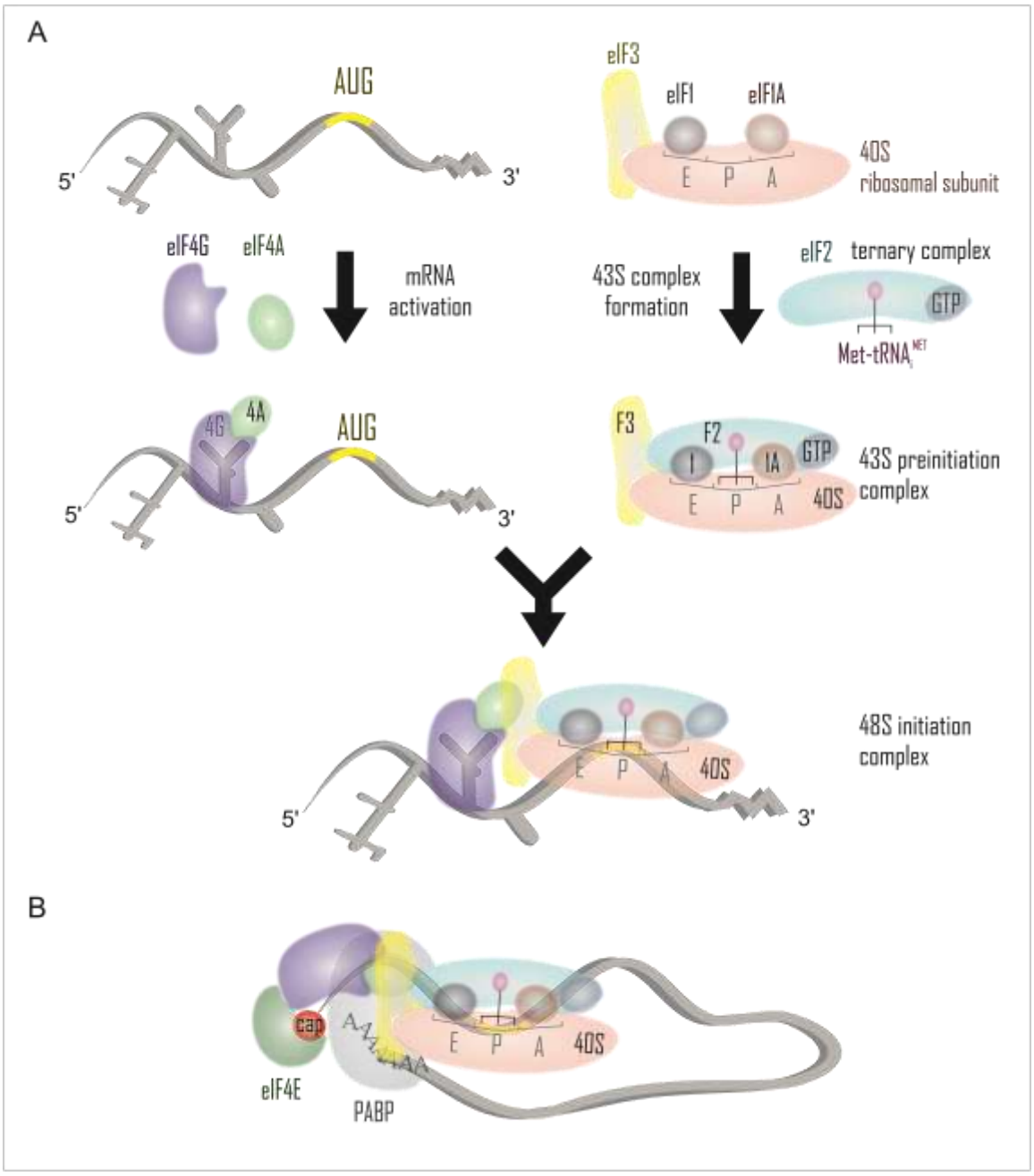


**Figure 1.3 Translation initiation in cap-independent and cap-dependent manner.**

**A** On the left side, mRNA activation is illustrated. The initiation factors eIF4G and eIF4A (purple and green) are recruited by structural elements of the IRES moiety. The formation of the 43S preinitiation complex is depicted on the right. The 40S ribosomal subunit is associated with eIF3, eIF1 and eIF1A (yellow, grey and light brown, respectively). The ternary complex (eIF2, light blue with GTP and the initiator $tRNA_i^{Met}$) is recruited. Subsequently, the activated mRNA and the 43S preinitiation complex together form the 48S initiation complex. **B** The closed loop formation of cap-dependent translation initiation is depicted. Here, PABP (poly(A)-binding protein, grey) and eIF4E (dark green) bind to the poly(A) tail and the cap structure, respectively.

## 1.4 Integrity of (messenger) RNA

As a consequence of the 2'-hydroxyl function, RNA molecules are more susceptible to degradation under alkaline conditions than DNA molecules (cf. figure 1.4). Hydrolysis of single-stranded RNA is observed at neutral pH and only slightly elevated temperature [62,63]. In addition, RNA is generally less associated with proteins and therefore more accessible to exogenous damage. The same is true for the protective duplex character of DNA, which RNA lacks. Another important aspect is enzymatic degradation by nucleases that act internally or can degrade the RNA molecule from either the 5' or 3' end [64]. These ribonucleases (RNases) can originate either from intracellular (e.g. RNase L) or exogenous sources (e.g. such as RNase A or synthetic RNases) (reviewed in [65]).

Not only is the (complete) degradation of RNA molecules a risk when working with these polymers, but other sources that damage RNA can have a significant impact on RNA function. Such RNA damage can affect base-pairing properties, which in the case of mRNAs lead to defects in protein synthesis and thus possibly to aberrant proteins or ribosome stalling.

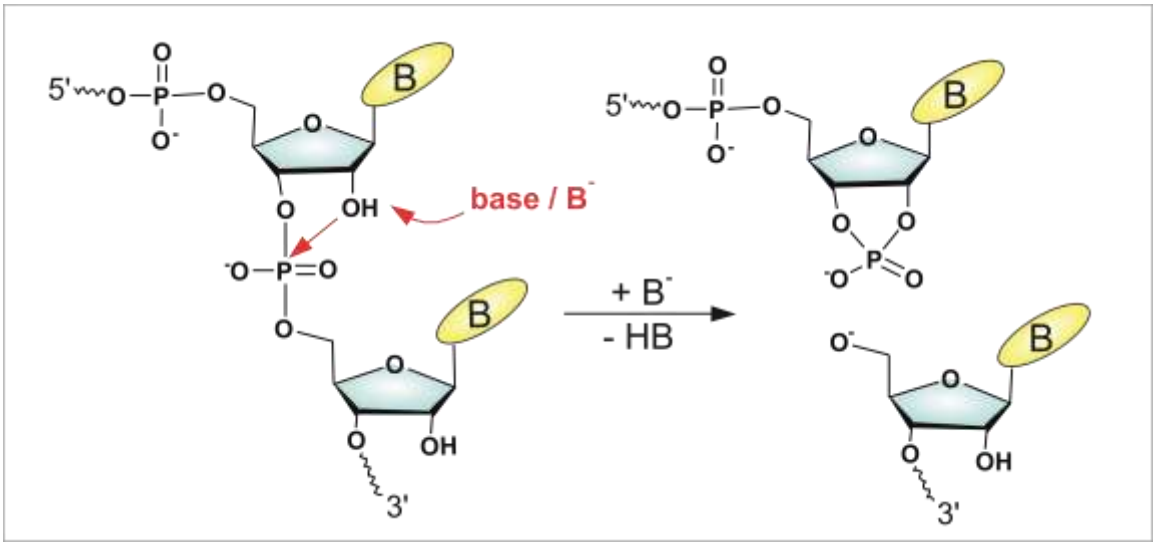


**Figure 1.4 Alkaline hydrolysis of RNA on the 2′-hydroxyl function.**
The ribose is deprotonated at the 2'-hydroxyl function under alkaline conditions and the subsequent attack of oxygen at the 3'-phosphate group leads to RNA strand break.

An important class of reagents responsible for the multiple modifications of RNA are reactive oxygen species (ROS) such as superoxide anion ($O_2^-$), which is produced by some cytosolic and membrane-bound proteins during the normal life cycle of a cell [66] and is essential for various aspects of the cell cycle such as autophagy, stress response, or signal transduction [67]. However, if the fragile balance is disturbed, this can lead to cell damage and even cell death [68]. The superoxide anion is extremely damaging due to its high reactivity and generates even more reactive oxygen species such as perhydroxyl radical ($HO_2^-$), hydrogen peroxide ($H_2O_2$), or hydroxyl radical ($^{\bullet}OH$) during reactions [69]. While a normal level of ROS can be removed in different ways depending on the organism (reviewed in [70,71]), a misbalance between ROS removal and production or exposure appears to lead to the formation of the hydroxyl radical [72,73]. In addition to hydrogen peroxide, reactions involving hydroxyl radicals are thought to cause most of the oxidative damage observed in nucleic acids [74].

Not only direct exposure or endogenous formation of reactive oxygen species leads to oxidative damage. Exposure to exogenous factors such as compounds from air pollution, environmental factors, X-rays, or ultraviolet (UV) irradiation also lead to ROS formation and thus oxidative damage to nucleic acids in general [75-80]. With respect to UV irradiation, it is important to consider the three main classes into which the UV spectrum is divided. The spectrum for UV light ranges from 400 nm down to 100 nm with low energy (long wavelength) UV-A (400 - 315 nm), medium energy UV-B (315 - 280 nm), and high energy UV-C light (280 - 100 nm), with the largest damage detectable from UV-C [78,80], but also lesions caused by UV-A irradiation are reported [79,81,82].

The most important non-canonical nucleosides on nucleic acids caused by reactive oxygen species originate from the reactions of guanines with hydroxyl radicals, followed by oxidation or reaction between guanine and singlet oxygen ($^1O_2$) [83], leading, for example, to 8-oxoguanosine (8OxoG) [79,81] or hydrates of cytidine and uridine [83] (see figure 1.5). 8OxoG, although the most common oxidative damage product induced in RNA, is also one of the most dangerous because it disrupts RNA-RNA interactions and therefore leads to a change in the decoding mechanism during

translation [84-86]. Other observed changes induced by ROS are ribose changes, base excision, cyclobutene dimers or strand breaks [80,81,83,87].

In addition to this oxidative damage, other widespread damage occurs from endogenous or exogenous alkylating agents (reviewed in [88]). The best known endogenous alkylating agent is probably S-adenosylmethionine (SAM), a universal methyl donor with a relatively high cellular concentration but weak nonenzymatic reactivity with other biomolecules [89]. Nevertheless, SAM is capable of forming various adducts *in vitro* after reaction with nucleic acids [90]. From exogenous sources, halocarbons, such as chloromethane, are the most common [91,92]. Alkylative adducts on a nitrogen such as N1 (e.g. 1-methyladenosine, $m^1A$) or N3 (e.g. $N^3$-methylcytosine, $m^3C$) alter the Watson-Crick surface and thus prevent canonical base pairing. Interestingly, 7-methylguanosine is the most common modification found after alkylation stress and can lead to depurination of nucleosides, namely abasic sites [93], thus likely adversely affecting RNA function.

## 1.5 Modified nucleosides in RNA

As mentioned in the previous chapters, a variety of non-canonical nucleoside structures can be found in nucleic acids, which are either modifications that have been specifically installed by enzymes (e.g. $m^7G$ cap structure) or originated from damage of the nucleic acids (e.g. 8OxoG). The number of non-canonical nucleosides is growing virtually daily and to date includes over 170 structures [94]. This adds another layer of information to the primary nucleotide sequence, making RNA more suitable for its complex and diverse function [95]. Modifications on RNAs are found in all three phylogenetic domains (archaea, bacteria, and eukarya) and in both coding RNA (i.e. mRNA) and noncoding RNA (reviewed in [96,97]). These modifications can range from simple methylation to acetylation or oxidation to hypermodifications with amino acids, sugar structures, or even more complex structures [95,98]. As diverse as the modifications themselves are, so are their functions, including structural and metabolic stabilization of RNA, modulation and control of decoding during translation,

and a variety of specific functions resulting from a direct interaction between the modification and proteins (the so-called “reader” proteins).

### 1.5.1 rRNA and tRNA modifications

In yeast eukaryotic ribosomal RNA (rRNA), nearly 2 % of nucleosides are modified [99], although the chemical diversity of these modifications is limited. The vast majority consist of pseudouridine and 2'-O-methylation [97,99], which are mostly introduced by small nucleolar ribonucleoproteins (snoRNPs) [100-102]. Although deletion of entire subsets of modifications results in measurable changes in ribosome biogenesis and function, deletion of individual modifications at functional sites such as the decoding site, peptidyl transfer center, or intersubunit bridge has no measurable effects on phenotype or cell growth [99,103].

In contrast to the low diversity of modifications found in rRNA, the other common non-coding RNA species, i.e. tRNA, has the highest abundance and diversity of RNA modifications. Examination of over 500 tRNAs from all three phylogenetic domains revealed a median of eight modified nucleosides per tRNA molecule, and the number of modifications was generally higher in eukaryotes compared to bacteria [95,104,105]. A large number of modifications are related to the anticodon loop and thus to the proper decoding of the degenerate genetic code. Positions one and two of the codon are reliably decoded with respect to canonical base pairing, whereas the situation at position three is more complex because of the discrepancy between the large number of codons to be read and the small number of anticodons [106]. A proposal to resolve this discrepancy was made by Francis Crick, who introduced the “wobble hypothesis” and thus the idea of noncanonical base pairing at position three [107], the very position with a large variety of nucleoside modifications, including residues such as the so-called “wobble uridines” [108,109]. Another highly modified position associated with the anticodon loop is tRNA position 37, which plays an important role in structuring the anticodon loop, promoting accurate translation [110] or enhancing tRNA selection [111]. In addition, Schaefer *et al.* have demonstrated that position 38 in the anticodon loop protects tRNAs from the endonuclease angiogenin and thus from cleavage when this position is methylated

[112]. Simpler modifications such as pseudouridine (Ψ), dihydrouridine (D), inosine (I), or methylations are located elsewhere in the tRNA and are known to be important for the local or overall structure of the tRNA molecule or may promote either stability or degradation [113]. A single 1-methyladenosine ($m^1A$) at position 9 of the mitochondrial $tRNA^{Lys}$ has been shown to be responsible for the classical cloverleaf structure of the tRNA molecule [114], and is critical at position 58 of the initiator $tRNA^{Met}$ for tRNA stability [115,116].

Overall, none of the modifications appeared to be essential, whereas their combined absence led to rapid decay of the tRNA by the 5'-exonucleases Xrn1 and Rat1 [117].

### 1.5.2 mRNA modifications

Although the variance of modifications found on mRNA is lower compared to tRNA and rRNA, and mRNAs are modified only in substochiometric amounts [118], since the 1970s and especially in the last decades, the modification sites found on mRNA have increased [119-123]. The modifications found are isomerization (e.g. Ψ [124,125]) or methylations either at the base (e.g. $m^1A$ [126], $N^6$-methyladenosine, $m^6A$ [119,127] or 5-methylcytidine, $m^5C$ [128-130]) or at the ribose moiety (i.e. 2'-O-methylations [131-133]). In addition to some general remarks, the modifications essential for the underlying work (namely $m^6A$, Ψ, and 2'-O-methylations) are discussed separately in greater detail later, and the following discussion focuses on the remaining modifications found on the mRNA. The aforementioned 5'-cap structure is one of the most modified sites of mRNA, containing the $m^7G$ and one to two 2'-O-methylations at the adjacent nucleosides [121,134,135]. In general, it is known that ribose methylations on RNA are important to distinguish between “self” and “non-self” RNA structures [136-141] and thus prevent the activation of the immune system through the activation of retinoic acid-inducible gene I (RIG-I), melanoma differentiation-associated protein 5 (MDA 5) or type 1 interferon effectors (type 1 IFN) [24,142-146]. In addition, double methylation can occur at the first transcribed nucleoside, resulting in $N^6$,2'-O-dimethyladenosine ($m^6A_m$) [147-149], which increases transcript stability by enhancing resistance to decapping by mRNA decapping enzyme 2 (DCP2) [135]. Reportedly, the abundance of $m^5C$ in DNA is up to 4 %, which is why it is referred to

as the “fifth base” in DNA, and it has been shown to be involved in the regulation of gene expression [150]. However, the reported abundance in RNA ranges from 0.02 to 0.1 % of all cytidines, and the exact number of sites in eukaryotic mRNAs varies dramatically in the literature depending on methodology and data analysis [151-154]. The functions of the $m^5C$ sites found are as controversial in the recent literature as the exact sites [155]. Several studies revealed an overlap of the found $m^5C$ sites with binding sites of proteins within mRNA, e.g. with the export adaptor protein ALYREF, which has been shown to be a so-called “reader” protein of $m^5C$ and to support nuclear export [156]. Another “reader” protein of $m^5C$ in the cytosol was identified as Y box binding protein 1 (YBX1), which is responsible for increasing the stability of mRNA [157,158]. In addition to the controversy surrounding $m^5C$ in mRNA, another controversial modification is $m^1A$ [159] with possible sites in the 5' untranslated region (5' UTR) [126] as well as in GC-rich regions [126,160]. Since the methylation is located in the Watson-Crick base pairing region [161], it is likely to modify the secondary and tertiary structure of the mRNA [126].

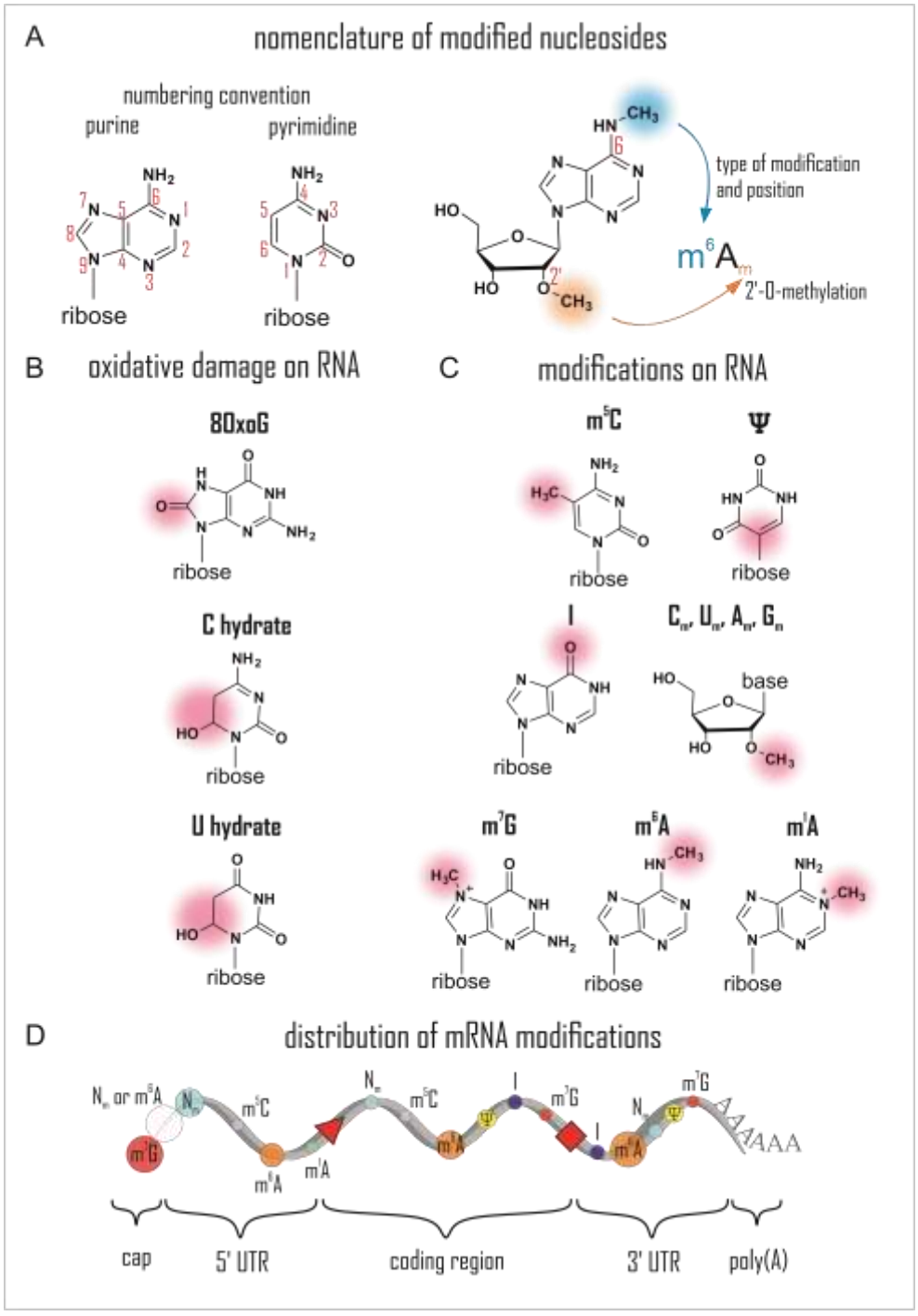


**Figure 1.5 Nomenclature and structure of RNA modifications and their distribution in mRNA.**
**A** The nomenclature of modified ribonucleosides is shown with the numbering convention of the base [162] and with an example of a doubly methylated adenosine (i.e. $m^6A_m$). **B** The chemical structures of some nucleosides found after oxidative damage are illustrated with highlighted structural changes compared to their canonical base: 8-oxoguanosine (8OxoG) and hydrates of cytidine and uridine **C** The chemical structures of some modifications are depicted: ribose methylation of all four nucleosides ($C_m$, $U_m$, $A_m$, $G_m$), pseudouridine (Ψ), 5-methylcytidine ($m^5C$), 7-methylguanosine ($m^7G$), $N^6$-methyladenosine ($m^6A$), 1-methyladenosine ($m^1A$), and inosine (I). **D** The reported modifications and their distribution in the different regions of mRNA are illustrated (adapted from [133]). The size of the circles correlates with the abundance of the corresponding nucleoside. The red triangle and the red square illustrate the start and stop codon, respectively.

### 1.5.3 $m^6A$ in mRNA

In the 1970s, $m^6A$ was the first modification found on mRNA [119] and since then has held a special position in the study of modified nucleosides on mRNA, especially in recent years [122,123]. Therefore, it may not be surprising that there is a wide variety of transcriptome-wide detection methods targeting $m^6A$ (reviewed in [133]). The amounts of $m^6A$ found in mammalian mRNAs vary from 0.1 - 0.4 % [121] or one to three $m^6A$ sites per mRNA for certain transcripts [120,163]. A specific consensus sequence for $m^6A$ sites, the DRACH motif (with D = G/A/U, R = G/A, H = A/U/C) [123,164], has been reported, but not all consensus sequences found appeared to be methylated [165], so the exact mechanism of targeting methylation sites is still obscure. In contrast, many studies have shown that the mechanism of methylation occurs via a complex containing methyltransferases 3 and 14 (METTL3 and METTL14, respectively), referred to as the “writer” complex. Within the complex, METTL14 supports the activation of METTL3 through allosteric interaction [119,122,166] and is important in RNA binding [167-169], while METTL3 transfers the methyl group of the methyl donor SAM to the target mRNA [166]. This process occurs co-transcriptionally [133,170] and involves several proteins, e.g. Wilms tumor 1 associated protein (WTAP) as a structural component of the METTL3-METTL14 complex, which supports stabilization in the methylation complex and interacts with splicing factors as well as other important proteins involved in RNA processing [164,171-173]. Deletion of another protein, VIRMA (former known as KIAA1429), resulted in dramatically reduced levels of $m^6A$ and thus is also thought to play an essential role in the methyltransferase complex [164], which controls regioselected methylation in the 3' UTR in close proximity to the stop codon [174]. Genomic depletion of METTL3 and inactivation of METTL14 in mouse embryonic stem cells resulted in clearly reduced poly(A)-mRNA levels after preparation [175] and allowed linkage of a variety of cellular functions to $m^6A$ [175,176]. RNA-binding motif protein 15 (RBM15) was shown to bind in close proximity to the target site of $m^6A$ methylation and is therefore thought to be involved in the recruitment of the methyltransferase complex to the target mRNA [177]. Depletion of this protein in mice

revealed important functions in vascular integrity, hematopoiesis, and B-cell proliferation [178,179].

Not only the writers of A to $m^6A$ are intensively studied, but also the second class of proteins associated with $m^6A$, the so-called “erasers” (demethylases), namely the two Fe(II)- and 2-oxoglutarate-dependent oxygenases FTO and ALKBH5. While ALKBH5 unambiguously demethylates $m^6A$, as it was shown that disrupting the balance of methylation and demethylation of $m^6A$ by knockout of ALKBH5 and thus the demethylation process affected sperm development and fertility in mice [180], substrate specificity for FTO has not been conclusively demonstrated. Even though FTO has been shown to promote $m^6A$ demethylation *in vitro* [181] and Zhao *et al.* demonstrated higher levels of $m^6A$ in mRNA *in vivo* when FTO was knocked down [182], according to several studies, it seems more likely that FTO targets $m^6A_m$ more efficiently than $m^6A$ [135,183].

In addition to the “writers” and “erasers”, proteins from the YTH domain family in particular are known to be “readers” of $m^6A$, i.e. they bind selectively to $m^6A$ sites and are therefore an important class of proteins in mediating the effects of $m^6A$ on mRNA [123,184-186]. While YTHDF1-3 are localized in the cytosol [187], YTHDC1 is located in the nucleus and interacts directly with the splicing factor SRSF3 while blocking SRSF10 to promote alternative splicing [185,188,189]. In addition, YTHDC1 plays a role in nuclear export of target mRNA [190]. Transcriptome-wide mapping revealed an enrichment of $m^6A$ sites along the 3' UTR near the stop codon and lower levels of methylation sites in the main coding region and the 5' UTR [122,123] of transcripts that exhibited functions in protein expression and mRNA stability mediated either by “reader” proteins or by alteration of local secondary structures [187,191]. This alternative mRNA folding, termed “$m^6A$ switch”, allows binding of the heterogeneous nuclear ribonucleoprotein C (hnRNPC) and thus alternative splicing [192-194]. The amount of $m^6A$ in the 5' UTR increases upon heat shock stress, possibly by a mechanism involving translocation of YTHDF2 to the nucleus to prevent demethylation of FTO and thus maintain the methylation level of the 5' UTR [186]. Cap-independent translation of such 5' UTR methylated mRNAs has been described by direct recruitment of eIF3 by $m^6A$ or by mediation of other proteins (e.g. ABCF1)

[195,196]. When $m^6A$ is localized in the 3' UTR, closed loop formation and thus cap-independent protein translation has been described. In the proposed mechanism, YTHDF1 actively promotes translation by interacting with several eukaryotic initiation factors (mainly eIF3) [197]. A possible explanation for this mechanism might be the enhanced protein translation of selective transcripts in response to stress stimuli [186,198]. For $m^6A$ located at different codon positions in the coding region of an mRNA, translational elongation was negatively affected in an *in vitro* translation assay, with the effect being greatest for the first codon position [199]. In another study, inhibition of the ribosome decoding process and tRNA accommodation was observed [200].

Given the involvement of $m^6A$ in many important biological processes, it is not surprising that $m^6A$ is associated with many diseases, including various types of cancer (reviewed in [194,201]), Alzheimer's disease [202], metabolic diseases [194], autoimmune diseases [203], HIV infection [204], hypertension, or cardiovascular disease [205]. It is likely that further intricacies of this important mRNA modification will be uncovered as part of ongoing and future studies to understand the role of $m^6A$ in more diseases.

### 1.5.4 Pseudouridine in mRNA

Pseudouridine is an isomer of uridine in which the bond between the ribose and the base moiety is changed from the canonical N-glycosidic bond to a C-C bond [206,207]. Because of its early discovery in the 1950s and its high abundance, 0.2 % to 0.6 % of uridines in eukaryotic mRNA, pseudouridine is sometimes referred to as the “fifth RNA nucleoside” [208-212]. Although the structure of pseudouridine differs from uridine due to isomerism, Watson-Crick base pairing is not affected, so pairing with A persists [213,214]. Several studies from several groups show a wide distribution of pseudouridine in different mRNA regions, including 5' UTR, 3' UTR and the coding sequence [125,212,215-217]. The exact role of pseudouridine within mRNA has not been fully elucidated, but increased pseudouridine levels after heat shock stress suggest a specific role in the stress response [125] and increased stability of the mRNA-tRNA interaction [218], as well as increased mRNA stability (reviewed in [219])

by contributing to increased phosphodiester backbone rigidity and base stacking [213,220]. In addition, Ψ is thought to expand the genetic code through its ability to alter protein synthesis by converting stop codons to sense codons [218,221]. Another study suggests a possible role in the formation of the spliceosome complex. Here, polypyrimidine-rich sequences are required at the 3' splice site and replacement of uridine with pseudouridine resulted in splicing defects in *Xenopus* oocytes [222].

Although recognition of target mRNAs for pseudouridylation is not clear, a consensus sequence has been proposed that guides a site-specific RNA pseudouridine synthase (PUS7) to its target site [125]. This mechanism involves snoRNAs and, in addition to PUS7, other enzymatic proteins that isomerize U into Ψ depending on the organism [223,224]. The observations on the effects of Ψ on protein synthesis are not consistent across studies. While a single Ψ resulted in decreased protein synthesis in a bacterial *in vitro* translation system and in human cells [199,225] other studies showed only a moderate decrease in protein content [226] or even an increase in protein synthesis when all uridines were replaced by Ψ [227]. Following these findings and the wide distribution of Ψ in eukaryotic mRNAs, further studies are required to elucidate the mechanism behind the altered protein synthesis. In that context, Kristin Koutmou's group proposed that Ψ alters the discrimination between near/noncognate and cognate aa-tRNAs during translation [225]. In addition, Karikó and coworkers demonstrated that incorporation of pseudouridine into mRNA lowers the immunogenicity of these *in vitro* transcripts by suppressing Toll-like receptor 7 (TLR7) signaling [228-230] and another pseudouridine modification ($N^1$ methylpseudouridine) is successfully used in therapeutic application (please refer to 1.9).

### 1.5.5 2'-O-methylation in mRNA

Methylation of the first and/or second transcribed nucleotide at the ribose moiety to reconstitute the different cap structures (cf. figure 1.2) has been reported to be necessary for efficient gene expression and to increases mRNA stability by distinguishing “self-“ and “non-self” RNA species [96,142], as mentioned before. In addition, numerous ribose methylated sites ($N_m$) were described in the coding region

of transcripts, with enrichment in codons encoding specific amino acids [131]. Elliot and coworkers proposed a methylation mechanism for mRNAs that is comparable to a mechanism found for noncoding RNA species (e.g. rRNA, tRNA, and miRNA) [231]. The proposed mechanism involves ribonucleoprotein complexes containing small nuclear C/D-box RNA (snoRNA) that direct methyltransferase to the target RNA [232,233]. Moreover, the biological role of ribose methylation in the coding sequence of mRNA transcripts remains unclear. However, in tRNA, a single $G_m$ residue has an antagonistic effect on TLR7 activation [139-141] by acting as a suppressor of RNA-induced pattern recognition receptor (PRR) activation [234], thereby suppressing immune stimulation, and similar effects have been reported using oligoribonucleotides with single 2'-O-methylations [228]. In addition, Ringeard *et al.* revealed a mechanism by which HIV-1 recruits a complex in human cells consisting of the RNA-binding protein TRBP and the 2'-O-methyltransferase FTSJ3 so that the viral RNA is 2'-O-methylated to evade recognition by the innate immune system [235].

Synthesis of mRNAs containing ribose methylations provides some evidence for this group of RNA modifications show repressed protein expression depending on codon context in *in vitro* translation systems [199,236] as well as in human cell line experiments [226]. Although ribose methylation does not directly interact with the Watson-Crick interface, the Puglisi group suggested a disruption of codon reading during cognate tRNA selection. 2'-O-methylation sterically interferes with the interaction of monitoring bases of the ribosome with codon-anticodon helices and thus with accommodation to A-site tRNA [236].

## 1.6 Detection methods for modifications in RNA

For a long time, paper-based or thin-layer chromatography (TLC) of nucleic acids was the main tool for identification of modifications [237,238], although large amounts of material were needed for detection. The discovery of the absorption of ultraviolet light by nucleic acids in the 1940s reduced the amount of material required to a few micrograms [239], and thereafter the introduction of radiolabeling with $^{32}P$ *in vitro* by T7 RNA polymerase [240] or *in vivo* labeling [241,242] by growth in $^{32}P$-containing medium enabled detection of minute amounts of nucleosides. Another breakthrough was the possibility of post-labeling RNA molecules at their 5' ends after the discovery of T4 polynucleotide kinase [243]. The first reported modification found in mRNA was pseudouridine (see 1.5.4), which was discovered by TLC around 1950 [209], as also 2'-O-methylations [244]. At the beginning of TLC analysis, modifications were identified by comparison with synthetic standards for assignment without considering sequence information [245].

In order to distinguish between random [246] and specific hydrolysis, including RNA cleavage by ribonucleases [247] or deoxyribozymes [248,249], optimized TLC analysis was used for sequence-dependent identification of modifications [212,250]. In addition, TLC analysis enabled the detection of $m^6A$ in various mRNA species (e.g. mRNA of *Arabidopsis thaliana* or in sporulating cells of *S. cerevisiae*) [173,251] by the group of Rupert Fray and was recently successfully applied to quantify relative $m^6A$ levels in *A. thaliana* to identify factors required for $m^6A$ mRNA methylation [252]. However, TLC analysis is very time-consuming, so automated methods have increasingly become the focus of RNA research.

### 1.6.1 High performance liquid chromatography coupled to mass spectrometry

The combination of analysis of changes in molecular mass, fragmentation pattern (by mass spectrometry), and retention time from liquid chromatography allows the identification and quantification of modifications from a mixture obtained by enzymatic digestion of RNA molecules into nucleosides. Early high performance liquid chromatography (HPLC) approaches, including the reversed-phase (RP) method, enabled separation of canonical and non-canonical nucleosides around 1980 [253], and improvements led to HPLC coupled to mass spectrometry (MS) [254]. The commonly used RP-HPLC is not always suitable for hydrophilic or polar nucleosides, but allowed coupling of HPLC with electrospray ionization mass spectrometry (ESI-MS), and as an alternative, further developments led to hydrophilic interaction liquid chromatography (HILIC) [255]. A variety of methods for nucleoside detection open the possibility of precise identification and quantification of the composition of RNA molecules, although sequence information is lost during digestion to nucleoside level [254]. Specific hydrolysis by RNase T1 [256] or other endonucleases [257,258] results in different RNA fragments in which sequence information is retained. However, this application is limited to abundant RNA species and low abundant species like mRNA cannot be analyzed in this way [259]. Further extensions of the field lead to LC-MS-based sequencing approaches [260] and automated analysis of data [261].

### 1.6.2 Next-generation sequencing

Until now the different sequencing techniques are divided into 1st, 2nd and 3rd generation. The First-generation techniques are based on polyacrylamide gel electrophoresis (PAGE) and involve either the synthesis of new nucleic acid strands (namely Sanger sequencing, sequencing by synthesis) [262] or, in contrast, nucleobase-specific cleavage, referred to as sequencing by chain scission [263]. Following the rapid adaptation of the DNA-based approach to RNA briefly after their respective publications, both methods allowed the initial study of RNA sequences [264,265]. Next (or 2nd) generation sequencing has evolved extensively over the last two decades and, in general, can be distinguished in methods utilizing sequencing by ligation [266,267] or sequencing by synthesis [268,269], and both approaches are classified as short-read sequencing. In contrast, the third generation has already been presented, focusing on long-read sequencing including real-time sequencing [270] enabled by newly developed techniques such as the so-called nanopore sequencing [271,272], which has already resulted in new techniques for $m^6A$ detection in native RNA sequences [273]. However, one of the most commonly used methods is next generation Illumina sequencing based on sequencing by synthesis [269]. Here, sample preparation is called library preparation and involves several steps from purification of target DNA to addition of sequencing adapters (short oligonucleotides). For RNA sequencing, a reverse transcription step must be performed to generate complementary DNA (cDNA) from the target RNA. The resulting template is then loaded into the so-called flow cell, where it hybridizes with anchored short synthetic oligonucleotides. In a final step, fluorescently labeled dNTPs are added for single nucleotide extension, which finally enables image acquisition. This leading technology is used in a variety of research areas, but for detection of modifications, library preparation must be adapted to the particular modification. Modified nucleosides can cause errors in reverse transcription of RNA sequences, such as stopping or blocking the enzyme or misincorporation [274]. An early link between modified nucleic acids and sequencing was an observed misincorporation attributed to various modifications involved in Watson-Crick base pairing [275-277]. For modifications that do not alter the Watson-Crick edge, such as 2'-O-methylations,

additional complementary steps must be performed during library preparation. For the detection of 2'-O-methylations, in the so-called RiboMethSeq alkaline hydrolysis and phosphodiester bond cleavage is added to the library preparation [100,278]. Methylation of the ribose at the 2'-O-position protects the modified position from random nucleolytic cleavage thus alkaline treatment of modified RNA resulting in fragments starting or ending at the +1-nucleotide relative to the 2'-O-methylation. This leads to the phenomenon that fragments starting or ending 3' before ribosemethylation are strongly underrepresented or even completely excluded. This protection is taken into account in the subsequent computational analysis of the data, resulting in a specific cleavage profile that allows the precise mapping of 2'-O-methylation [278,279].

## 1.7 Synthetic mRNA design

The manipulation of synthetic mRNAs began in the 1980s with the *in vitro* transcription of mRNA using the phage RNA polymerase SP6. Although the activity was validated by injection into the cytoplasm of frog cells or by *in vitro* translation with wheat germ extract, the major obstacle of the fragility of the synthesized mRNA initially led to a research standstill. The advantages of RNA over DNA, such as the lack of genomic integration and protein expression in the cytoplasm in one step, draw attention to the improvement of synthetic mRNA as a novel therapeutic tool. Therefore, it is important to understand the biological significance of the different modifications.

### 1.7.1 mRNA synthesis via *in vitro* transcription and solid phase synthesis

As described above, exogenous *in vitro* transcribed (IVT) mRNA can stimulate the immune system by activating PRRs and downstream activation of TLRs in endosomes. While TLR3 recognizes more dsRNA (double stranded RNA), TLR7 recognizes ssRNA (single stranded RNA) and thus IVT mRNA [280-282]. This stimulation of the immune system results in alteration of protein expression and suppression by various mechanisms [283-286], but the immunogenicity of synthesized RNA is significantly reduced by the use of modified nucleosides (namley Ψ, $m^6A$, $m^5C$, $s^2U$, and $N_m$) [228].

For potential use in therapeutic applications, such as vaccines, the effects of the modifications on the mechanism of translation decoding and immunostimulatory activity need to be investigated. One of the first modifications studied was pseudouridine, which was either partially [287-289] or completely [230] incorporated into mRNA. The strategies used so far to incorporate modifications into mRNAs can be divided into three main groups (illustrated in figure 1.6). The incorporation of modified nucleoside triphosphates (NTPs) during *in vitro* transcription is accomplished by replacing one of the four major NTPs with the modified NTP, resulting in a fully modified mRNA. In another strategy, a certain percentage of the modification is introduced into the NTP mixture, resulting in random incorporation into the synthesized mRNA (shown in figure 1.6A) [228]. Sophisticated strategies such as insertion of $m^6A$ at the first transcribed nucleotide using $N^6$-adenosine monophosphate ($m^6AMP$) [195] or the use of artificial mRNA sequences that allow insertion of only a single modification (i.e. inosine) using only ITP instead of GTP [290] are limited in their application. In addition, *in vitro* transcription of modified mRNAs is limited to those NTPs that are accepted by the polymerase used (e.g. T7 RNA polymerase), which excludes, for example, ribose methylated modifications in this synthesis strategy [291].

Another widely used synthesis strategy is chemical synthesis via a solid support material, known as solid phase synthesis. The RNA molecule is synthesized step by step in single reactions, where the phosphoramidites of the respective nucleotides are covalently bound to the growing RNA chain (cf. figure 1.6B). Although site-specific synthesis of modified mRNAs is possible [228], the length of the resulting RNA is limited to about 40 nts, making this method suitable only for short modified RNAs [292,293].

### 1.7.2 mRNA synthesis via splint ligation

For longer, site-specifically modified mRNA synthesis, the method of choice described in the literature is RNA ligation (figure 1.6C). Two consecutive RNA fragments are joined by specific ligation enzymes (i.e. ligases) to form a longer RNA. The ligases are mostly derived from bacteriophage T4 and are either specialized for ligation of single-

stranded nucleic acids, such as T4 RNA ligase 1 [294], or they more efficiently ligate a nick in double-stranded regions, allowing either DNA, RNA:RNA or RNA:DNA hybrids [295]. While T4 RNA ligase 1 performed well in techniques such as pCp labeling [296], for double-stranded regions only T4 DNA ligase and T4 RNA ligase 2 are suitable. These efficiently link the 3'-hydroxyl group to a 5'-monophosphate group of two consecutive nucleotides when hybridized to a so-called splint oligonucleotide [295,297-299]. T4 RNA ligase-2-like ligases occur in all three phylogenetic domains and require adenosine triphosphate (ATP) and a divalent cationic cofactor (e.g. $Mg^{2+}$ or $Mn^{2+}$) in addition to the reactive 3'-hydroxy group [295]. The ligation reaction consists of three nucleotidyl transfer steps [300-303]. In the first step, ATP is covalently bound to the ligase with the release of pyrophosphate, forming a ligase (lysyl-N)-AMP intermediate. Subsequently, AMP is transferred to the 5'-phosphate group of the donor RNA, forming an RNA adenylate intermediate [304]. Finally, the 3'-hydroxyl group of the acceptor RNA attacks the AppRNA, resulting in the release of AMP and the formation of a new phosphodiester bond. In addition to the creation of a double-stranded structure required for T4 RNA ligase 2, the use of a splint ensures proper alignment of RNA fragments and prevents incorrect ligations, such as ligation of the same fragment type to each other. The use of splints with overhangs between 8 and 35 nts on each site of the nick has been reported in the literature [199,299,305], with longer splints favoring the ligation reaction [299,306]. The individual RNA fragments can be synthesized either by solid phase synthesis or by *in vitro* transcription, allowing for fragments of flexible length depending on the synthesis method chosen. Although this method was originally used to investigate the importance of the 2'-hydroxyl group at the 3'-splice site for the splicing process [298], it has been used extensively by Erlacher and coworkers to produce point-modified (i.e. site-specifically modified) mRNA [199,226,305,307].

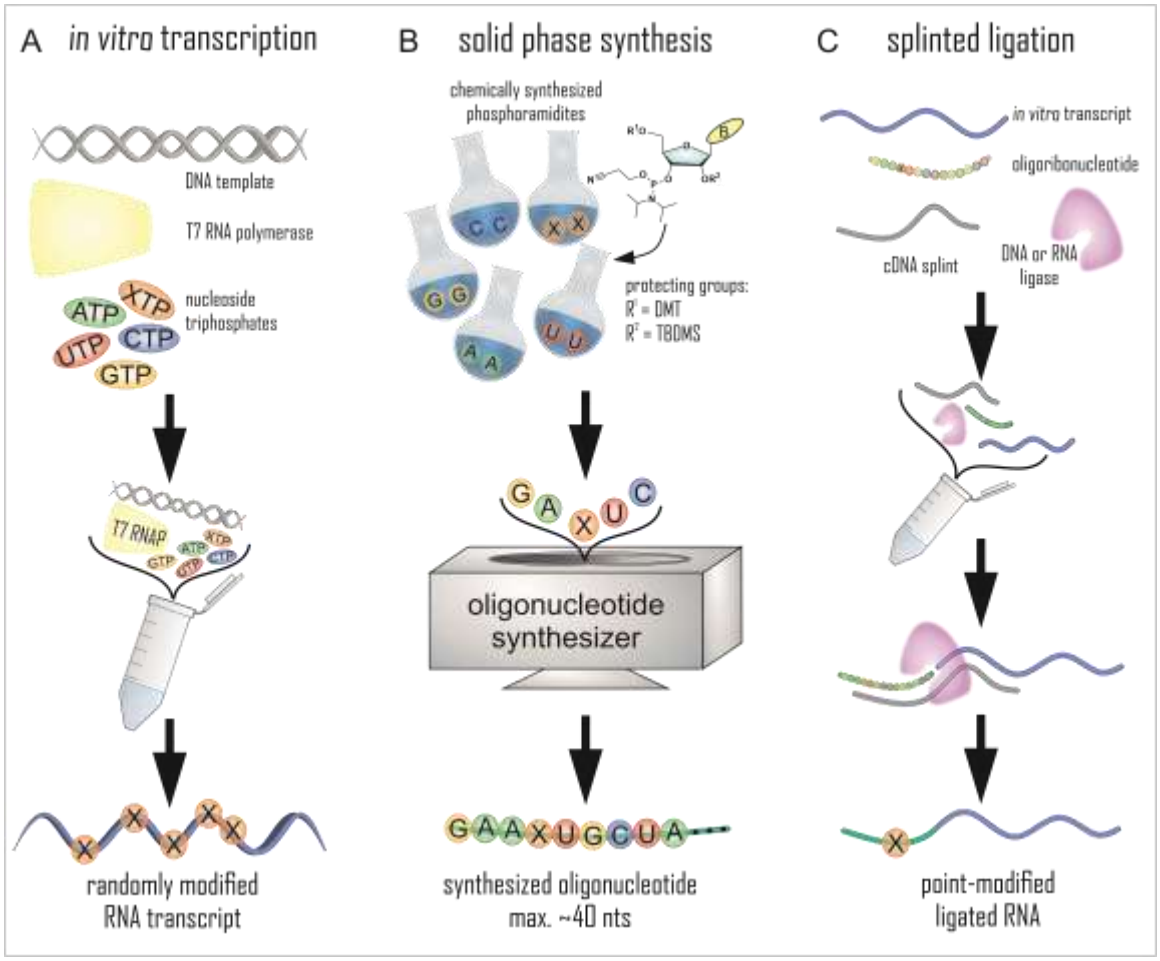


**Figure 1.6 Synthesis strategies of modified RNA.**

**A** *In vitro* transcription of modified RNA involves the DNA template, an RNA polymerase (T7 RNA polymerase in this case), and the nucleoside triphosphates (ATP, CTP, UTP, GTP) including the modified XTP. The resulting RNA is randomly modified. **B** For solid phase synthesis of modified oligoribonucleotides, the chemically synthesized phosphoramidites with respective base (B) are coupled to protecting groups for the 5'- and 2'-hydroxyl positions ($R^1$ = dimethoxytrityl, DMT and $R^2$ = *tert*-butyldimethylsilyl ether, TBDMS) and subsequent solid phase synthesis yields the oligoribonucleotide with specific sequence. **C** For splint ligation, the RNA fragments (*in vitro* transcribed or oligoribonucleotides) are hybridized to a DNA splint and enzymatic ligation is performed by either DNA or RNA ligase resulting in a point-modified RNA.

## 1.8 Green fluorescent protein (GFP) as reporter mRNA sequence

Synthetic mRNAs are applied to different research questions and thus the requirements for the mRNA depend on the particular experimental setup. While the study of the effects of modifications in a specific codon context on translation elongation was performed with a short artificial oligoribonucleotide [200,236], other publications aimed at a more global study of modifications on functional protein translation [226,230,308]. Commonly used reporter mRNAs for this purpose are luciferases (firefly or *renilla*) [230] or mRNAs encoding the green fluorescent protein family [226,308]. Development of bioluminescence of luciferases requires stoichiometric amounts of ATP and ambient oxygen as cofactors [309,310]. In contrast, only ambient oxygen is required for maturation of the green fluorescent protein [311,312] and the chromophore is formed after folding into the three-dimensional tertiary structure of the protein [313,314]. Furthermore, the barrel of beta-sheets around the fluorophore prevents loss of fluorescence under mild conditions [315-319], and assembly of the fragments can occur post-transcriptionally [320]. The chromophore of GFP is formed from the amino acid residues serine65 (Ser65), tyrosine66 (Tyr66), and glycine67 (Gly67), resulting in a *p*-hydroxybenzylidene imidazolinone [321,322]. The reaction mechanism of the chromophore formation is depicted in figure 1.7. In a first reaction step, the amide of Gly67 reacts with the carbonyl residue at position 65 by nucleophilic attack, followed by dehydration. Subsequently, ambient oxygen dehydrates the $\alpha,\beta$-bond of the residue at position 66, causing the aromatic group to conjugate with the imidazolinone ring [321,322]. In addition to the original mRNA sequence encoding the green fluorescent protein from the jellyfish *Aequorea Victoria*, first described by Osamu Shimomura in 1962 [323], an mRNA sequence concomitant with an altered amino acid sequence were found, resulting in an improved version of the protein, namely the enhanced green fluorescent protein (eGFP) [324-326]. The point mutation of serine at position 65 to threonine (S65T) has been reported to result in enhanced photostability of protein fluorescence due to a longer wavelength of excitation and

emission, and formation of the fluorophore from the nascent polypeptide chain occurs up to four times faster than for the wild type protein [324,327]. Although the formation of the chromophore in wild type GFP is considered to be fully elucidated, the mechanism behind the improved folding efficiency and fluorescence for the eGFP forms, such as S65T, remains unclear [326].

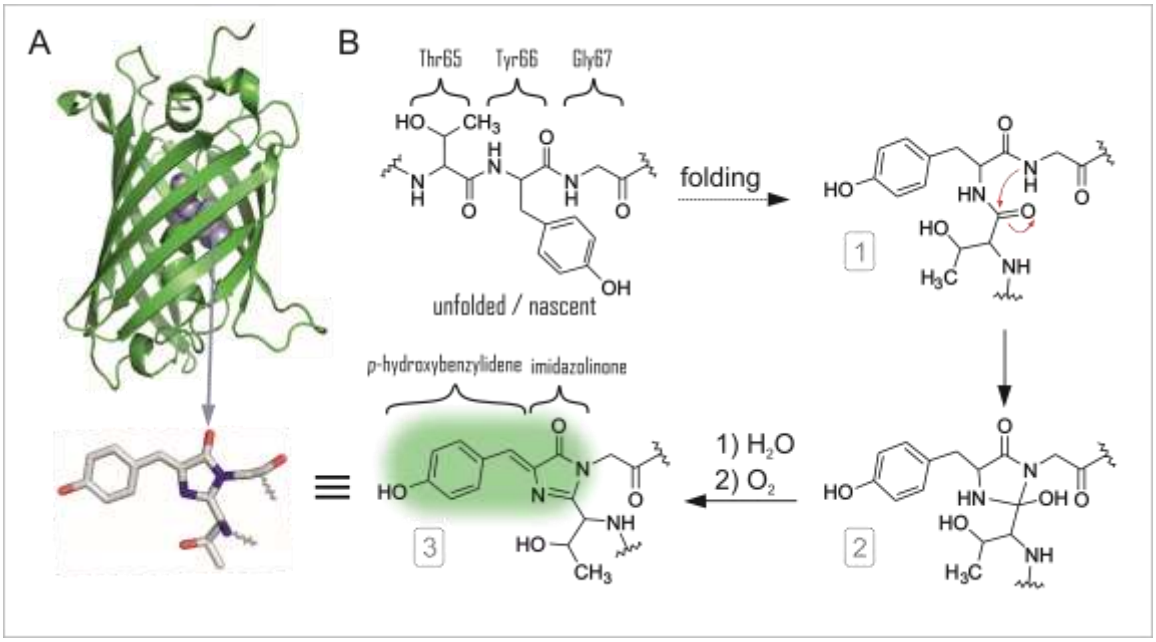


**Figure 1.7 eGFP and its chromophore formation.**

**A** The beta-barrel structure of eGFP is depicted with the chromophore inside the barrel (blue balls) and in greater detail the chromophore structure is illustrated in the lower part (nitrogen atoms blue and oxygen atoms red). **B** After post-transcriptionally folding of the nascent protein the cyclisation between Thr65, Tyr66 and Gly67 is enabled by close proximity of the residues. (1) Amide from Gly67 reacts with carbonyl residue at position 65 in a nucleophilic attack. (2) Subsequently, the imidazolinone ring gets dehydrated and ambient oxygen causes dehydratization of $\alpha,\beta$-bond of the residue at position 66 to form the *p*-hydroxybenzylidene imidazolinone function (3) [328].

## 1.9 Application of modified mRNA

In recent decades, the idea of nucleic acid-encoded therapeutic applications has become the focus of research. Indeed, transient expression of mRNA has both weaknesses and advantages over DNA-based approaches. Whereas a serious disadvantage of DNA therapeutics was genomic integration and thus persistent but irreversible translation [329], the propensity of long RNAs [330] to degradation and thus low protein yield had to be circumvented. As diverse as the degradation possibilities for the mRNA are, the formulation also has to be adapted. Degradation can already take place in the sample storage container or be promoted by the application and thus the circulation in the body and finally in and from the cell. Major advances in the introduction of modifications to maintain mRNA stability (see 1.5) combined with achievements in delivery, such as the use of liposomes or similar nanoparticle formulations [331-333], have enabled several clinical trials of therapeutic IVT mRNA [331,334-337]. Both passive and active immunization are reported, with active immunization becoming the focus of research. Pardi *et al.* used mRNA encoding the light and heavy chains of the anti-HIV-1 antibody VRC01 to generate immunity to intravenous HIV infection in humanized mice [338]. This mRNA was modified with $N^1$-methylpseudouridine ($m^1\Psi$), which has been reported to perform better than Ψ in cell-based experiments and in mice [285,339,340]. In active immunization, an antigen of the target pathogen must be identified. This sequence is then cloned into a DNA template and transcribed into mRNA *in vitro* before the vaccine is administered. An active immunization approach for a HIV vaccination has been studied in mice and rhesus macaques [341]. While unmodified mRNA stimulates the intrinsic immune system by activating TLR7 and TLR8 and thus can act as an intrinsic adjuvant [342,343], and certain modification in mRNA impede immune stimulation [228]. Depending on the approach, PRR signaling may be desired or better circumvented [230,280,344,345]. In this context, HIV genomic RNA has been reported to contain ribose methylations that impede cellular immune responses, providing a clue to this important area of research [235]. Several studies have demonstrated

mRNA as a plausible alternative vaccine vector against influenza [346,347], cancer [348-350], and other infectious diseases [351-353].

In the fight against pandemic and epidemic viruses such as Zika virus, modified mRNA vaccines have begun to gain acceptance in research. In the case of Zika virus, the development of modified mRNA encoding the pre-membrane and envelope [354] of Zika virus glycoproteins resulted in immunity in a mouse model [354], and immunization was maintained in mice and macaques for a period of 3 to 5 months [352,354].

To date, the success of mRNA vaccines for therapeutic applications has culminated in the fight against the severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) pandemic, where two mRNA-based vaccines have achieved not only clinical trials but also regulatory approval [355]. Consistent with the previously described importance of modifications in mRNA, it is worth noting that both mRNA-based vaccines encoding the trimerized receptor-binding domain spike glycoprotein of SARS-CoV-2 are quantitatively modified with $m^1\Psi$.

# 2 Motivation and Objectives

To date, more than 170 modified nucleosides are known in all kingdoms of life and in a variety of RNA species. These modifications are involved in versatile biological processes, such as RNA folding, recognition by the innate immune system and regulation of translation fidelity. As introduced previously, the effects of modified nucleosides in mRNA on translation is an intensively studied area, but a straightforward synthesis routine for point-modified mRNA is lacking. The insertion of modifications into mRNAs in a biological context is reported to be site-specific and regulated by enzymes. In contrast, the most widely used method for synthesizing modified mRNA to study its effect on translation is the random insertion of modified nucleoside triphosphates during *in vitro* transcription. Therefore, this work aimed to develop a method for the synthesis of point-modified mRNAs with naturally occurring modifications. The starting point was a previous study on the synthesis of short RNA species using splinted ligation [299]. This method had to be transferred to long RNAs and optimized. Special attention must be dedicated to the development of a suitable purification method, which becomes more difficult as the length of the particular RNA increases. Since long RNA is not exceptionally stable, this ought to be reflected in the purification method developed so that the integrity of the RNA is preserved. An important decision must be made regarding the translational mechanism under study, as there are both cap-dependent and cap-independent translational pathways in eukaryotes associated to modifications. In a next step, positions for potential modifications have to be identified. Interesting positions where such modifications can be placed and investigated mainly concern the start and stop codon of the mRNA, but in principle also the coding region of the mRNA. Finally, a read-out of the translated protein served to give a first insight into the influence of modifications on translation. For this purpose, different assay systems such as Western blotting or radioactive labeling might be utilized, and probably several will need to be tested to find the appropriate one.

# 3 Results and Discussion

## 3.1 Retrosynthesis of mRNA sequences

In protein biosynthesis, mRNA is a very important factor to regulate the protein expression. Apart from the very well documented impact of different structural elements of the mRNA (untranslated regions, poly(A)-tail or cap structure) on mRNA processing, translation, or immunogenicity [145,356-361] the modification grade of the mRNA also plays an important role. However, the described effects on translation regulation were observed for highly modified *in vitro* transcribed mRNAs [285,340], shorter singularly modified RNAs such as non-sense sequences [218,225,290] or short peptide sequences [307], or for natural RNA species [362]. To overcome these limitations, this work aimed to develop a method to synthesize and purify long, point-modified mRNAs allowing subsequent translation without further processing like capping or poly(A)-tailing. These mRNAs were only modified in a single position and were therefore termed “point-modified” mRNAs. Already single modifications were able to tune translation negatively in the context of short RNA sequences [200,236,307] and were therefore a very exciting approach for long mRNAs. Furthermore, single modifications in other RNA species were shown to reduce immunogenicity of these RNA species by suppression of TLR7 activation [363] and therefore the questions arose if these two effects of point modifications might be combined. The synthesis of such mRNAs was enabled by using a splint ligation approach. Hereby, a complementary DNA oligonucleotide was used to hybridize two (or more) successive RNA fragments in proper orientation [294]. Furthermore, the 5' phosphate group of one RNA and the 3' hydroxyl group of the adjacent RNA were brought into spatial proximity to enable proper ligation between these two ends. In

previous approaches, two RNA fragments were bridged by a cDNA to generate longer mRNAs [218,307] or tRNAs [299] (so-called 2-way splint ligation), or more than two RNA fragments (up to five) were fused to generate smaller RNA species [364].

The underlying construct was an eGFP mRNA with an internal ribosome entry site (IRES) (for complete sequence please refer to table 6.1 in the appendix). Other reporter genes such as *renilla* luciferase requires cofactors to develop its bioluminescence [309]. In contrast, the green fluorescent protein does not require any co-factors or enzymes to develop its fluorescence, since the chromophore is formed by the three-dimensional tertiary structure during maturation of the protein [313,314].

The fluorescence of GFP is very robust and stable under different conditions including SDS gel electrophoresis of the mature protein [316-319]. Therefore, in-gel detection of the fluorescence signal directly from SDS polyacrylamide gel after electrophoresis of the translated protein was possible. The translation was performed *in vitro* in rabbit reticulocyte lysate [365-367]. Furthermore, the used internal ribosome entry site from encephalomyocarditis virus (EMCV) was shown to ensure a stable *in vitro* translation of the luciferase reporter [366] which was reported to be more efficient than other IRESs or capped mRNAs [48]. The use of an IRES overcame the challenge of incomplete cap synthesis downstream of the mRNA synthesis

## 3.2 Investigation of splint ligation approach

### 3.2.1 Development of building blocks (with [redacted])

Starting from the final mRNA sequence, the development of the building blocks [368] faced some limitations. In order to allow synthesis of a point-modified mRNA, a 3-way-one-pot splint ligation was developed. Within this approach, the final mRNA sequence consisted of three different fragments enabling insertion of one single modification in a specific position of the mRNA by applying a modified middle fragment (cf. figure 3.1A). Long, point-modified mRNAs are not accessible via *in vitro* transcription, since modified nucleosides are inserted as triphosphates in different percentages (up to 100 %) instead of the canonical nucleotide (e.g. PsiTP/UTP,

$m^6$ATP/ATP) during transcription reaction. Besides the random incorporation of these modifications, the second limitation of this procedure is the restriction of the T7 RNA polymerase used by default for *in vitro* transcriptions. This enzyme tolerates a lot of naturally occurring modifications as nucleoside triphosphates, such as $N^6$-methyladenosine triphosphate ($m^6$ATP), 5-methylcytidine triphosphate ($m^5$CTP), pseudouridine triphosphate (ΨTP) [228] or functional groups used for click reactions like 5-ethynyluridine triphosphate (EUTP) or azides [369,370] but is not compatible with some other (naturally occurring) modifications such as 2'-O-methylations [291]. With a splint ligation approach, these limitations were circumvented. Although this work was aimed only at the incorporation of known modifications, the method can in principle be applied to all modifications available as phosphoramidites and which are incorporated during solid phase synthesis of the modifiable RNA fragment.

Following this idea, the mRNA was synthesized from two long RNA fragments (fragment #1 and #3) synthesized by *in vitro* transcription and one shorter RNA (fragment #2) synthesized by chemical synthesis and commercially available as oligoribonucleotide with inserted modifications. The solid phase synthesis of the modified oligonucleotide limits its length to a few dozen nucleotides [293] which is also the limitation of this synthesis method in general. Therefore, exclusive synthesis of long, point-modified mRNAs was not possible. The length of the designed IRES-eGFP mRNA included a total of more than 1300 nucleotides (nts), separated into 624 nts from 5' UTR including the EMCV IRES structure and 735 nts for the ORF including a 15 nts short 3' UTR. Fragments #1 and #3 were synthesized by *in vitro* transcription with T7 RNA polymerase since their length comprised several hundred nucleotides.

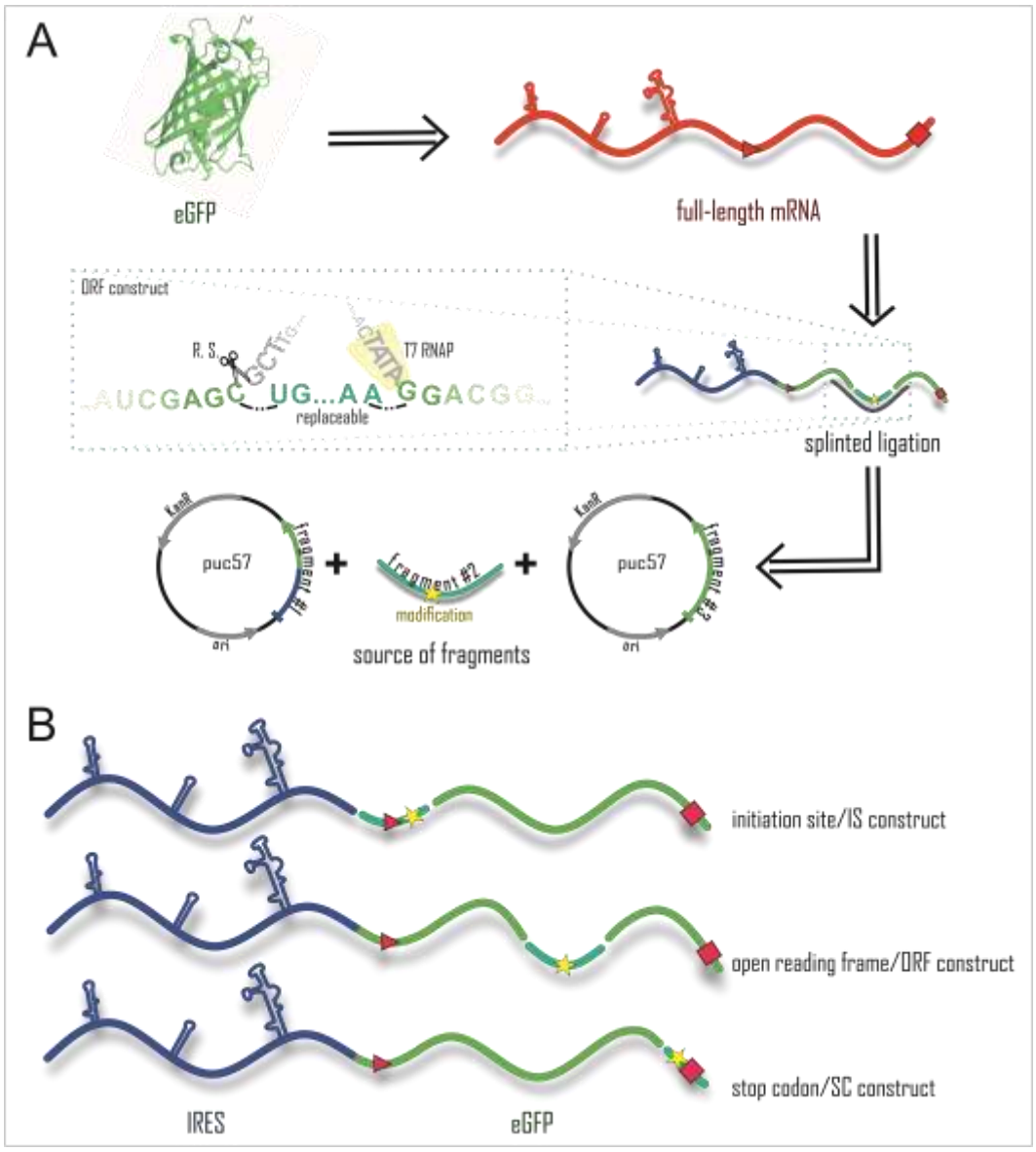


**Figure 3.1 Development of building blocks and their arrangement in the constructs.**

**A** Retrosynthetic view of the development, starting from protein down to the plasmid DNA level. Beta-barrel structure of eGFP synthesized by *in vitro* translation of the full-length mRNA (red). The mRNA contained a highly structured 5' internal ribosomal entry site (IRES) and was built from three fragments by splinted ligation. The red triangle and the red square symbolized the start- and stop codon of the eGFP sequence, respectively. The long RNA fragments for splinted mRNA ligation were synthesized from plasmid DNA visualized as circles with indicated origin of replication (ori), antibiotic resistance gene (KanR) and the sequence of interest, colored blue-green or green for fragment #1 (incl. IRES) and fragment #3, respectively. The point-modified short RNA fragment was synthesized by solid phase synthesis (fragment #2, turquoise). The inlet depicte a schematic zoom into the ligation site of the ORF construct with the solid phase synthesized replaceable fragment #2 in turquoise. Upstream (5') of this: fragment #1 with indicated restriction site from restriction enzyme (R.S. and scissor) of the plasmid DNA (grey). Downstream (3') of the replaceable fragment #2: the *in vitro* transcription of fragment #3 by T7 RNA polymerase (T7 RNAP, yellow trapeze) is depicted with its binding site (TATA box on the plasmid sequence in grey). The dashed lines indicate the two nicks to be ligation on RNA level. **B** Schematic sequences of the three different constructs with varied position of the replaceable fragment. The IRES sequence indicated in blue with secondary structure. Replaceable fragment in turquoise with modification depicted as yellow star and eGFP sequence in green. Start and stop codons indicated by red triangle and red square, respectively. The development of the constructs was done in collaboration with [redacted].

The aim was to obtain building blocks in which the plasmid DNA (pDNA) of the corresponding sequence was linearized and the RNA fragments were synthesized by T7 RNA polymerase and limited by falling off the enzyme at the end of the sequence. Therefore, linearization of the pDNA was performed with type II restriction enzymes, half of whose recognition sites were already present in the mRNA sequence, while the second half was incorporated into the pDNA sequence outside the subsequent RNA fragment. This was then followed by the commercial oligoribonucleotide and the second transcription product starting with at least a “GG”-motif. This was set as a prerequisite since the T7-RNA polymerase works most efficient with the +1 to +6 initially transcribed sequence “GGGAGA” [371,372]. With these limitations the sequence was inspected for half recognition sites of type II restriction enzymes with a downstream “GG” sequence in a distance of approximately 20 nts from the restriction site. The inlet in figure 3.1A detailed the ligation site of the ORF construct and indicated the different prerequisites made for a ligation. Upstream fragment #1 included the restriction site (R.S., “AGCGCT”) on the pDNA level and pDNA of downstream fragment #3 contained the T7 RNA polymerase promoter upstream of the initially transcribed sequence “GGA…“. Apart from inspecting the sequence for appropriate ligation sites, it was also important to consider possible purification methods for the final ligation product afterwards. In theory, the separation of a final ligation product differing in length significantly from the unligated fragments seemed more plausible than ligation of RNA fragments of only a few dozen nucleotides length, due to size separation methods like chromatography or gel electrophoresis and were therefore preferred.

At the end, three potential study objects named “initiation site (IS) construct”, “open reading frame (ORF) construct” and “stop codon (SC) construct” (figure 3.1B) were created. The names indicated the modifiable regions where fragment #2 was located. In the case of the SC construct, the ligation was a 2-way-one-pot ligation with an *in vitro* transcribed fragment #1 and solid phase synthesized fragment #2. Due to its incompatibility with the above purification requirements, it was not an ideal subject for study. Nevertheless, the idea was to investigate the ligation potential first on a higher number of study objects and focus afterwards on the most promising

candidates to optimize the approach and test the translatability *in vitro*. Both the initiation site construct and open reading frame construct showed a fragment pattern with fragments differing significantly in length from the reconstituted full-length mRNA and were therefore, in theory, more accessible for purification

### 3.2.2 Integrity of ligation products

A typical ligation reaction contained equimolar amounts of each fragment and the cDNA splint, hybridized to the RNA fragments prior to overnight incubation at 16 °C with T4 RNA ligase 2. It is reported in literature that lower incubation temperature and longer incubation time better preserve the integrity of the full-length product applying T4 DNA ligase [299], assuming that the benefits of longer incubation at lower temperatures are comparable between T4 DNA ligase and T4 RNA ligase 2, this concept was adapted for ligation with T4 RNA ligase 2. The enzymatic ligation of two successive RNA fragments was shown previously to be highly dependent on the reaction conditions including RNA sequence [364] and splint length [299] (please also refer to 3.2.3) and was applied earlier for smaller RNA species like tRNA [299] or other RNAs [294,364]. In an initial assay, the integrity of the potential ligation product was tested by hybridizing a fluorescent cDNA to the ligation product. For these experiments, ligation conditions were first transferred from unpublished data for the reaction. A cDNA splint with 35 nts overhang from each nick (total length of around 80 nts) and incubation temperature of 16 °C were used in an overnight reaction.

After separation of the hybridized probes on an agarose gel, the SC construct did not show a fluorescent signal corresponding to the full-length product, although the integrity of the individual RNA fragments was ensured. The SC construct was therefore excluded from further studies (data not shown). On the left site of figure 3.2, the exemplary results of the hybridization experiment are shown for the IS construct. The positive and negative controls of the IS construct (lanes 1 and 6-10 respectively) implicated that the hybridization only took place if at least two out of three fragments were successfully ligated. However, for the SC construct, with only two fragments, none of the ligation controls showed a fluorescence signal although the GelRed™ stained gel showed the presence of the RNA. In section 3.3.1, among others, the

results of HPLC experiments for SC construct were discussed and supported the results of the hybridization test (*vide infra*).

To validate the results from the hybridization of the fluorescent cDNA to the RNA, in a second step the ligation reaction of the IS and the ORF construct were performed including a cyanine 5 (Cy5)-labeled fragment #2. Internal labeling of the ligation product ensured that potential secondary structures and thus steric reasons did not interfere with the labeling procedure itself, which is otherwise a source of error in labeling by hybridization. Another opportunity while using internal labeling of the RNA was quantification of the ligation efficiency by quantifying the intensity of the Cy5 signals in the gel. The Cy5-labeled experiments supported the results from the hybridization tests: ligation only took place if the middle fragment #2 was present in both constructs and reasonable amounts of partially ligated fragment #2 with either fragment #1 or fragment #3 were seen in lanes 13 and 15 for IS construct and lanes 20 and 22 for ORF construct on the right site of figure 3.2.

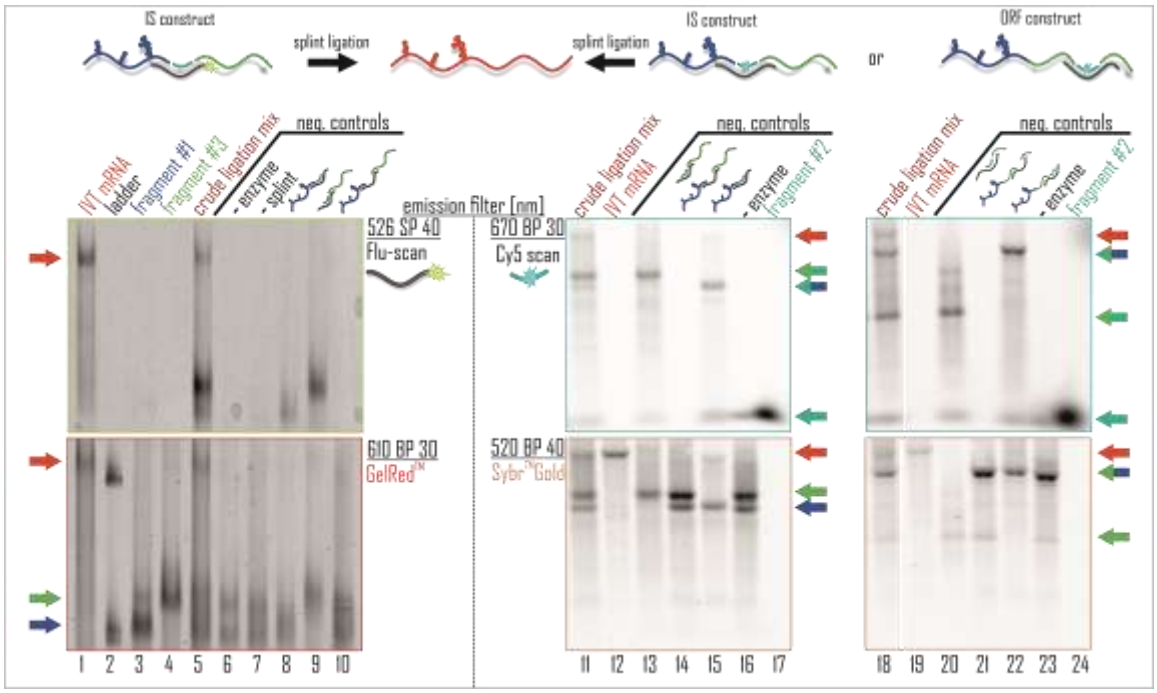


**Figure 3.2 Integrity of ligation product.**

The crude ligation mix were tested for integrity of the ligation product by either hybridizing to a fluorescent complementary cDNA (left side) or by using a Cy5-labeled replaceable/fragment #2 (right side). Above the different constructs and labelling possibilities are illustrated. The used emission filter is indicated and the arrows marked the respective RNA construct. **On the left site** [373] the fluorescence Scan (top) and a scan after staining the 2 % agarose gel with GelRed™ (bottom) are shown. The used controls were either fragment #1 or fragment #3 alone (lanes 3 and 4), an *in vitro* transcribed full-length mRNA (lane 1) or, as negative controls, lanes 6 – 10: ligation without enzyme or splint or without one of the respective fragments. A total amount of 750 ng RNA per lane was loaded. **On the right site** (lanes 11 – 24) either the Cy5-scans (top) or the scan of in-gel SYBR™Gold stained (bottom) 1 % agarose gel are shown (IS (left) and ORF (right) building block, respectively). Used controls: a full-length, not labeled, *in vitro* transcript (lanes 12 and 19), labeled fragment #2 (17 and 24) or the ligation reaction without enzyme (16 and 23) or without one of the respective fragments (lanes 13 – 15 and 20 – 22). 400 ng RNA was loaded per lane.

### 3.2.3 Ligation efficiency is dependent on ligation parameter

By staining the gel in a conventional way, an even staining of all RNA fragments was not be guaranteed due to the staining procedure itself and the interaction of the staining reagent with the RNA molecules. In-gel staining with SYBR™Gold implicated a ligation efficiency for the IS construct of about 30 %, whereas calculation of the ligation efficiency from the same gel with the Cy5 channel resulted in a ligation efficiency of only about 12 %. For detailed analysis, the Cy5 calculated ligation efficiencies were used, due to the above circumstances with stained RNA. Utilizing these intensities, only one molecule per mRNA will be guaranteed, making over- or under-representation of individual RNA species less likely.

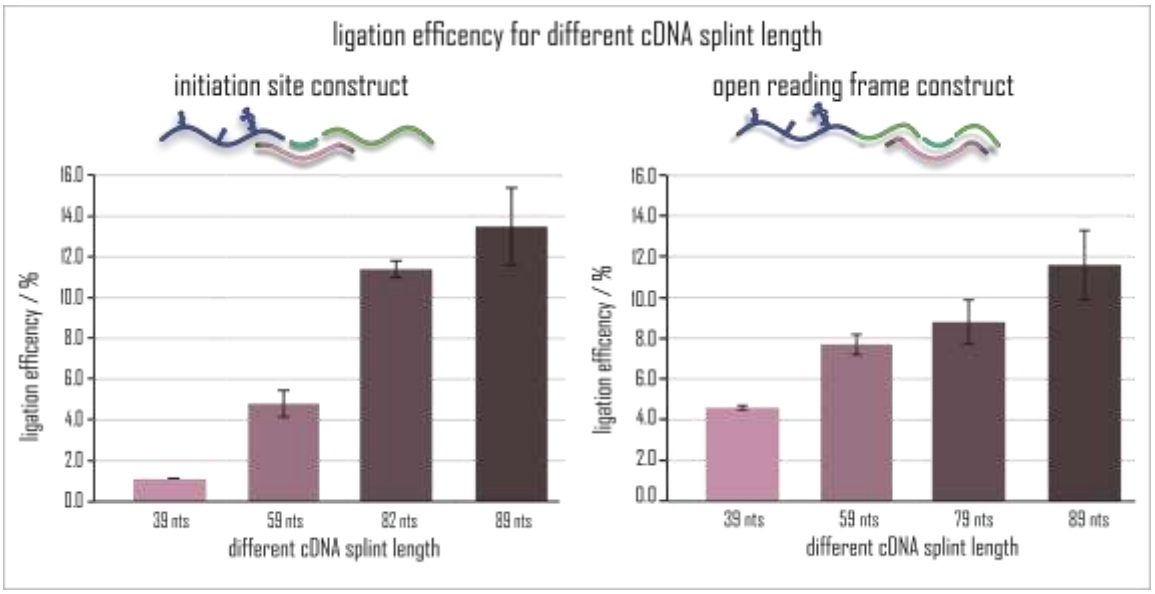


**Figure 3.3 Ligation efficiencies for different cDNA splint length.**
The ligation efficiency was calculated from the Cy5 signal in the gel after ligation with cDNA with a length of 39, 59, 82/79 or 89 nts for the initiation site (left) or the open reading frame (right) construct, respectively.

To further optimize the splinted ligation of the constructs different setups were tested to find optimal ligation conditions. Kurschat *et al.* used 2005 [299] cDNA with an "overhang" of 15 to 35 nts on each site of the nick. The same size range was used by Hoernes *et al.* 2016 and 2018 (8 to 20 nts respectively) [199,305]. In fact, a Cy5-labeled RNA with splint cDNAs with total length' of 39, 59, 82 and 89 nts for the initiation site construct and 39, 59, 79 and 89 nts for the open reading frame construct were used. This led to an overhang of 10 to 35 nts on each site (and additional 19 nts to span

fragment #2). The ligation efficiency for both building blocks were highest with the longest splint (cf. figure 3.3) but not substantially higher than with the 82 nts and 79 nts long splint for the IS construct and the ORF construct, respectively. In another optimization test, the conditions of the hybridization of the RNA fragments for the IS construct to the splint DNA were varried. Before adding the ligase, the reaction mixture was heated for 4 minutes to 75 °C to denature secondary structures of the RNA. Subsequently, the reaction was uncontrolled cooled to room temperature over several minutes. This cooling step was variied in three different ways to investigate a potential impact on hybridization efficiency of the splint to the RNA fragments. The experiments for the controlled cooling over 30 or 60 minutes were performed inside a thermocycler with cooling steps from 75 °C to 22 °C (assumed as room temperature) over indicated time. In addition to this slow cooling, another experimental setup tested cooling the heated samples on ice to produce a rapid temperature gradient. In figure 3.4A the results of these experiments are shown in a relative comparison towards the uncontrolled cooling. However, none of these different hybridization conditions resulted in a remarkable change in ligation efficiency compared to the uncontrolled reference.

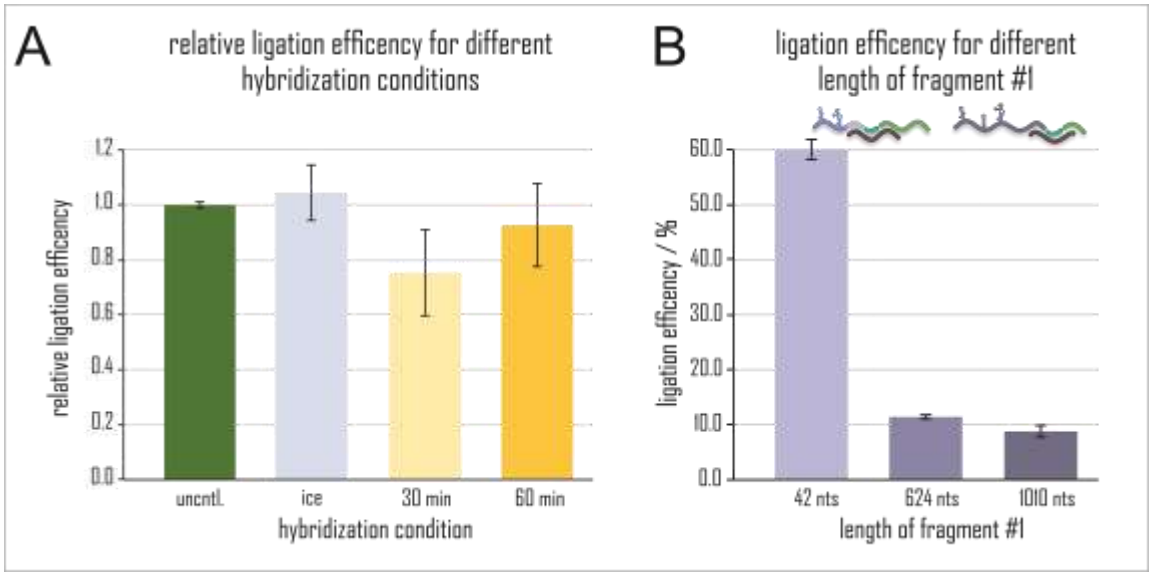


**Figure 3.4 Influence of hybridization conditions and fragment length on ligation efficiency.**
**A** Different cooling conditions after the denaturing step of the ligation reaction were tested and ligation efficacy relative to the uncontrolled cooling is shown. The ligation reaction was either cooled on ice or slowly cooled over 30 or 60 minutes to room temperature. **B** The ligation efficiency was calculated for ligations with different length of fragment #1. A 42 nts long fragment #1 was ligated to fragment #2 and #3 of the IS construct compared to ORF construct with 1010 nts long fragment #1. Experiments were performed in triplicates (± SD).

In the same experiment, it was tested whether hybridization, and thus subsequent ligation, differed when hybridization was split into two steps. In a first step the hybridization was performed only with the splint and the replaceable fragment #2 present. In a second step fragments #1 and #3 were added (either heated to denature secondary structures prior to mixing or not). In both cases no ligation was observed afterwards (data not shown). As mentioned in 3.2.2 the stop codon construct did not hybridize with fluorescent cDNA, and was thereofre probabyl not sucessfully ligated. Nevertheless, the question arose, if length of fragment #1 (and maybe the secondary structures of this fragment) had an impact on ligation efficiency. This was reinforced by observation that the ORF construct tended to have slightly lower ligation efficacy compared to the IS construct.

In another setup, replaceable fragment #2 and fragment #3 from the IS construct were used in combination with a short oligoribonucleotide labeled with a 5' fluorescein tag to act as fragment #1. Raplacing fragment #1 removed the highly structured IRES region of the construct and changed the 624 nts long RNA fragment into a 42 nts long fragment. This increased the ligation efficiency calculated by Cy5-labeled raplaceable (fragment #2) substainally compared to building blocks with longer and higher structured fragment #1. In further experiments, comparable ligation efficencies were observed for replacing fragment #1 against 40 nts or 48 nts long RNA sequences (cf. figure 6.1). In all cases, the presence of the very short RNA fragment #1 was confirmed by fluorescein labeling of these fragments and subsequent detection of the fluorescent signals (cf. figure 6.2).

However, based on this experiment, it was not possible to precisely elaborate whether the correlation exists in the length of fragment #1, in the amount of secondary structures or in both. Since the ligation site of the ORF construct was located several hundred nucleotides downstream of the IRES sequence, it is not reasonable to argue with the highly structured IRES motif alone to find a rationale for the low ligation yield. The three tested short fragments #1 all had a similar length. Therefore, further experiments with fragments with a greater variety of lengths needed to be performed to make this hypothesis more precise.

Overall, the presented synthesis route enabled the point-modified synthesis of long mRNAs in one reaction step. The idea of point-modified mRNAs synthesized by a ligation approach was alredy establisehd by Hoernes *et al.* in couple of different constructs [199,226,305,307] lacking the posibility of a one pot reaction for ligation of more then two succesive fragments, whereas a one-pot ligation avoided the need to apply yield-critical purification steps more than once. This was the advantage of the 3-way-one-pot ligation, including saving time by bypassing time-consuming steps wherever possible.

## 3.3 Purification of synthesized long mRNAs

Investigating a purification procedure for long RNA species, such as mRNA, was a challenging step in the underlying work. In particular, the purifications were aimed at separating long RNAs from smaller RNAs from a mixture, such as that resulting from a ligation reaction described above. A standard purification method for mRNAs from cells or tissues is the enrichment of the poly(A) fraction of a total RNA mixture by hybridization of magnetic poly(dT) beads, which was also adapted by Hoernes *et al.* for purification of point-modified mRNA after 2-way splint ligation [226,307]. The developed ligation approach described in 3.2.1 resulted in an mRNA without a poly(A) tail, and adapting the hybridization approach with a cDNA oligonucleotide hybridizing to the full-length ligation product, comparable to the splint used for ligation, was also not promising because impurities of the ligation reaction were mostly unligated fragments which were capable to hybridize in the same way to a potential cDNA. Therefore, this possibility was ruled out from the outset, and the focus was on three main purification ideas to be tested with the IS construct: Purification by high-performance liquid chromatography, excision from an agarose gel after electrophoresis, or gel elution in real time. An overview of the purification methods is provided in figure 3.5.

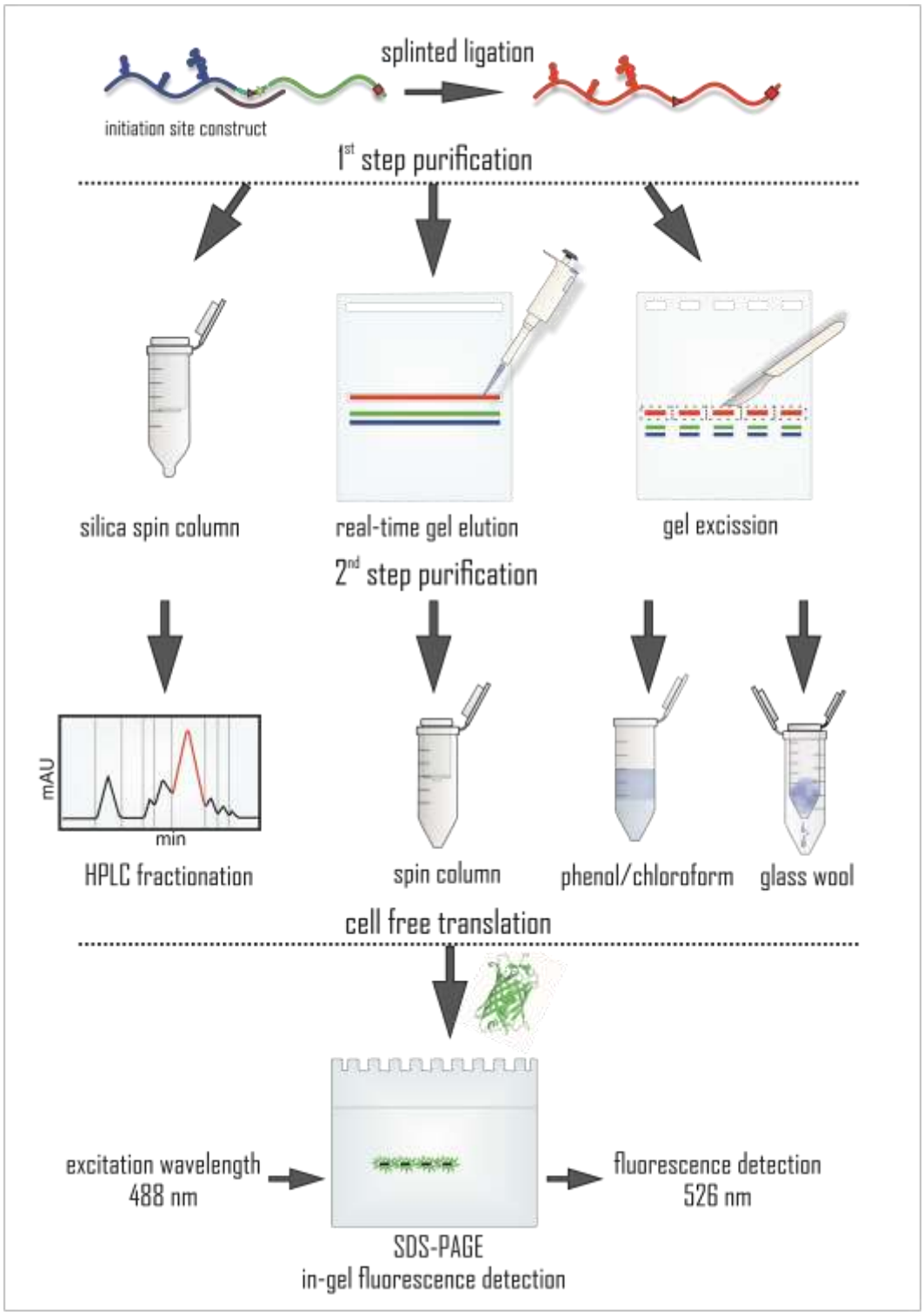


**Figure 3.5 Method overview of the tested purification methods.**
After a splinted ligation, the crude reaction mixture was purified. Left: after pre-cleaning with silica spin columns, the reaction mixture was subjected to HPLC fractionation. Middle: The crude ligation mix was fractionated by real-time gel elution from a conventional 1 % agarose gel by eluting the respective band from the collection pocket. Subsequently, the respective fraction was filtered via solid phase filter with a pore size of 0.2 nm. Right: The crude ligation mix was separated by a conventional or a low melting/gelling 1 % agarose gel and respective RNA was excised from the gel prior to extraction of the RNA either by phenol-chloroform extraction (low melting/gelling agarose) or by glass wool filtration (conventional agarose). Subsequent subjection of the purified full-length RNA to cell free translation and separation on SDS-PAGE allowed in-gel detection of eGFP fluorescence.

### 3.3.1 Purification via RP-HPLC (in collaboration with [redacted])

A promising method for purification of long RNA species is high performance liquid chromatography with a reversed phase column. This method was shown to purify modified mRNA after *in vitro* transcription and moreover to reduce immunogenicity of the mRNA solution [374]. Karikó *et al.* used preparative HPLC with a PS-DVB [375] column resin and triethylammonium acetate (TEAA) buffer as ion pair reagent [376]. Adapting this method to an Agilent HPLC system using a YMC-Triart reversed-phase C18 column with a hybrid silica column material also suitable for TEAA as an ion-pair reagent was challenging and was performed in close collaboration with [redacted].

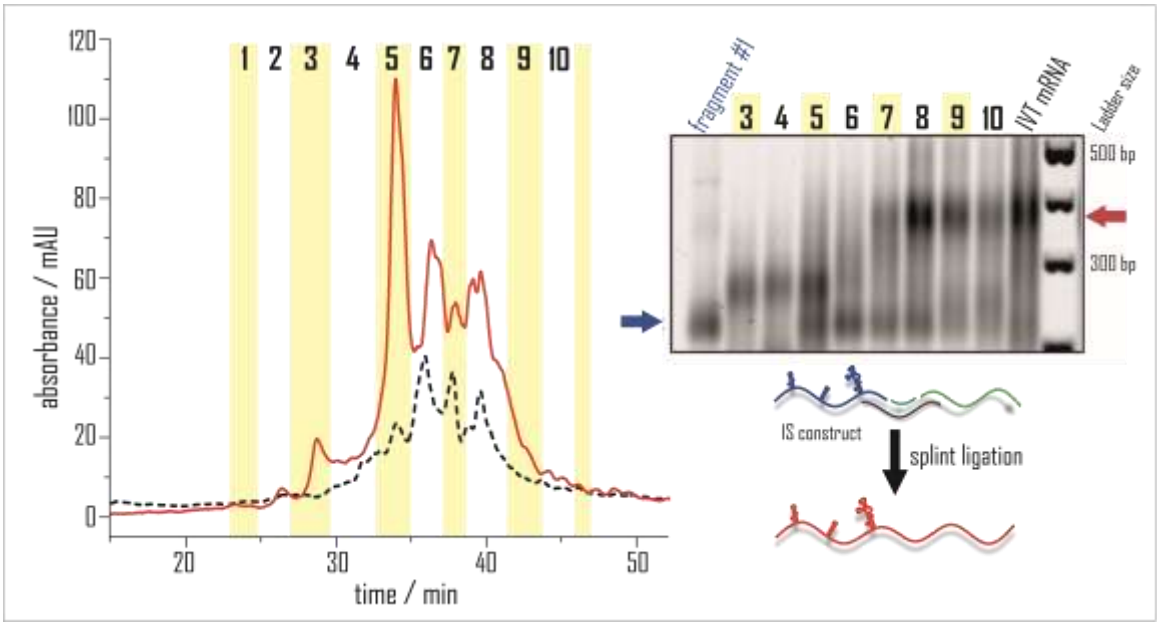


**Figure 3.6 Chromatogram of an HPLC run of ligation mix from IS construct and subsequent fraction analysis.**

A total amount of 30 µg crude ligation mix (red line) and the corresponding *in vitro* transcribed full-length mRNA (6 µg, dashed black line) were subjected to HPLC analysis. Ten small fractions were collected manually from the crud ligation mix, indicated by the yellow and white bars. Fractions three to ten were subsequently analyzed with a 2 % agarose gel, stained with SYBR™Gold (right side). The arrows indicated the expected positions of unligated fragment #1 (blue) and full-length product (red). 100 bp plus DNA Ladder from Thermo Fisher Scientific was applied next to the samples.

A full-length *in vitro* transcript, assumed to consist of only a single RNA species, was separated by RP-HPLC to elucidate the retention time of the full-length product. The chromatographic profile of UV absorbance is shown on the left site of figure 3.6 as a dashed black line next to the chromatographic profile of a crude ligation mixture applied on the system (solid red line). It is noticeable that the sample of the IVT transcript gave a heterogeneous chromatographic profile, although the sample was supposed to consist of ony a single RNA species. The crude ligation mixture, which was composed of both non-ligated fragments and the full-length ligation product, also gave a heterogeneous chromatographic profile. Fractionation of the crude ligation mixture resulted in an elution profile of the individual RNA fragments without sharp separation between them, as illustrated in the agarose gel image on the right site of figure 3.6. One had to note that no ligation product was observable for SC construct, neither in the UV chromatogram nor on the picture of the gel electrophoresis of the taken fractions (cf. figure 6.3).

Besides this, little literature is available for a reliable purification method for long RNA species without complex and/or expensive laboratory equipment. The use of electrophoretic techniques for the purification of nucleic acids is widely reported in the literature [199,377-380], but their potential for purifying long RNA species is limited. Since extraction of full-length ligation product from a denaturing polyacrylamide gel by passive elution after separation of crude ligation mix was only feasible by use of low percentage PAGE and not very reliable in a set of several purification tests, the use of agarose gels seemed more promising. The sufficient separation between the full-length product and the unligated or partially ligated fragments was already observed during the analysis of the building blocks (cf. figure 3.2).

### 3.3.2 Purification via agarose gel electrophoresis

There are numerous methods in the literature for extracting DNA molecules from agarose [381-384]. The applicability for larger RNA molecules was tested by adaption of some purification approaches from literature to the constructs used in this work. Since the 1970s, a lot of purification methods taking advantage of agarose gels were reported for DNA, involving different chemicals like hexadecyltrimethylammonium bromide in 1-butanol [385] to separate the nucleic acids from the agarose phase or to denature the agarose chemically by addition of “chaotrophic” anions like perchlorates [386] or iodide [387] to destroy the structure of the agarose and isolate the nucleic acids. However, these methods used rather harsh chemicals and are liable to demage the RNA. Another possibility described in the literature consists in the direct recovery of the nucleic acids from the solid agarose piece. Following this approach, the recovery of large RNA molecules by centrifugation through self-made glass wool filters either directly [388,389] or with a freezing step in liquid nitrogen prior to centrifugation through self-made glass wool filters [390,391] were tested (cf. figure 6.4). By using low melting/gelling agarose the addition of chemicals to destroy the integrity of the agarose piece, with the nucleic acids inside, became obsolete and allowed to perform the melting at temperatures around 65 °C. Since DNA molecules were more stable than RNA molecules, due to the lack of the 2'-hydroxyl group, heating them inside agarose for a couple of minutes seemed not detrimental for stability, while applicability to RNA had yet to be tested. The liquid agarose allowed extraction of the desired nucleic acids using, for example, chloroform [392]. Besides the low melting/gelling (LMG) agarose approach, another possibility to liquefy conventional agarose is the enzymatic digestion of the solid agarose particles by an enzyme called agarase, prior to phenol-chloroform extraction [392]. However, the enzymatic approach resulted in lower yields compared to heating low melting/gelling agarose and was more expensive due to the addition of the enzyme. Several of the cited methods were tested and all of them worked in general but had different disadvantages in detail. None of them showed satisfying results referring to integrity, purity and amounts of recoverd RNA. Centrifugation of the solid agarose pieces resulted in only very small yields for larger nucleic acids [391] and this was not

increased substantially in the underlying experiments by including additonal steps in the procedure like freezing the agarose beforehand.

Regarding the extraction of the RNA from liquid agarose by phenol-chloroform, only slightly higher yields were obtained, but resulted in a bigger pellet after precipitation of the RNA. This was a known impurity of these methods due to a fair amount of agarose inside the nucleic acid pellet [392] which impeded the subsequent LC-MS/MS analysis of the purified samples. For this reason other purification methods were tested to circumvent these agarose impurities or at least to reduce them to the lowest amount possible. All tested purification methods implicated that a certain amount of agarose was unavoidable, but was dependent on the ratio of gel slice volume to RNA. Figure 3.7 illustrates the example of extraction from LMG agarose in lane 1 compared to the crude ligation mix (lane 2). Based on this comparison, it can be assumed that the separation of the ligation product from most of the unliganded fragments was successful. Although this was also observed with the other agarose gel extraction methods, given that the ligation reaction resulted in only a small amount of full-length product, the above techniques did not seem to be the methods of choice to purify the full-length ligation product.

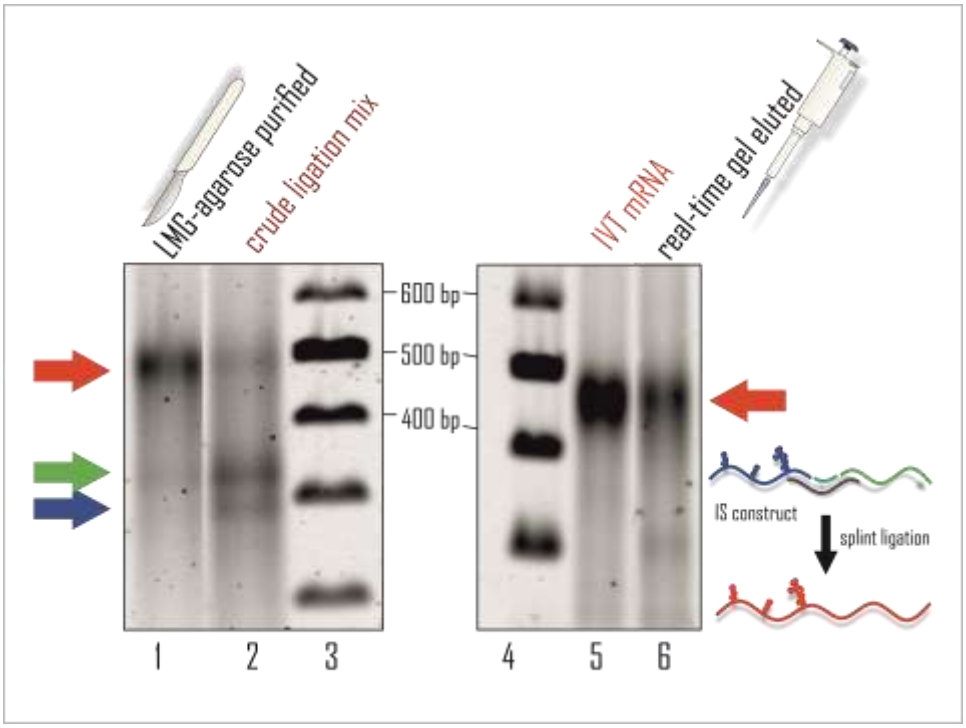


**Figure 3.7 Purification control after low melting/gelling (LMG) purification or real-time gel elution.** 100 ng of purified ligation mixture after LMG agarose purification (lane 1) or real-time gel elution (lane 5) were applied next to crude ligation mix (lane 2) or *in vitro* transcribed full-length RNA (lane 6) to a 1 % agarose gel stained with SYBR™Gold. Size marker: GeneRuler 100 bp plus DNA ladder (lane 3 and 4). The expected position of each RNA species (fragments or full-length RNA) is indicated by the colored arrows (color code from the ligation reaction depiction on the right).

### 3.3.3 Purification via real-time gel elution

Another possibilty to release nucleic acid from solid gel pieces consisted in electroelution. This is enabled by use of a dialysis bag inside a dialysis chamber [393-398] or onto a dialysis membrane [399] after a first step of gel electrophoresis and excision of the desired nucleic acid band. Another possibility was the elution directyl onto a cellulose membrane with a pore size of 45 µm [392]. The aforementioned methods are only feasible using equipment that is no longer commercially available and for which there is very little relevant literature [400], or are technically challenging due to the correct positioning of the cellulose membrane in the agarose gel after the nucleic acids have already been applied and separated [392] or the need to place a dialysis membrane very precisely during the casting of the gel [399]. In addition, all of the above methods had the disadvantage of requiring an initial purifying gel from which the RNA is then merely extracted. Moreover, elutions from PAGE slices showed the limitation in purifcation of small amounts of RNA in contrast to large volumes due to the elution process [401-403], which probably also applies for agarose preparations. Nevertheless, based on the idea of electroelution another purification method was developed called real-time gel elution to recover the desired nucleic acids directly from the agarose gel by live visualization of the separation process and without the need of a prior separation gel. The investigated system was inspired from Thermo Fisher Scientifics E-Gel™ CloneWell™ [404] but with a more flexible setup. For a detailed protocol please refer to section 5.3.2, but in general a conventional agarose gel chamber was placed onto a blue ligth transilluminator (i.e. DarkReader™ transilluminator) to visuallize in-gel stained nucleic acids (cf. figure 3.8), by a fluorescent staining dye (here SYBR™Gold), during the migation process. To recover the RNA from the gel, an additional comb was placed in the second half of the gel to create a “trough” (elution pocket) for pipetting out the desired RNA species as soon as they reentered the distal site. In the experiments presented here, this second pocket was created by either preparation of a conventional comb from the used gel chamber with adhesive tape to combine the predefined pockets to a bigger one or by direct use of an inhouse comb printed with a 3D printer (examplary pictures in figure 6.5).

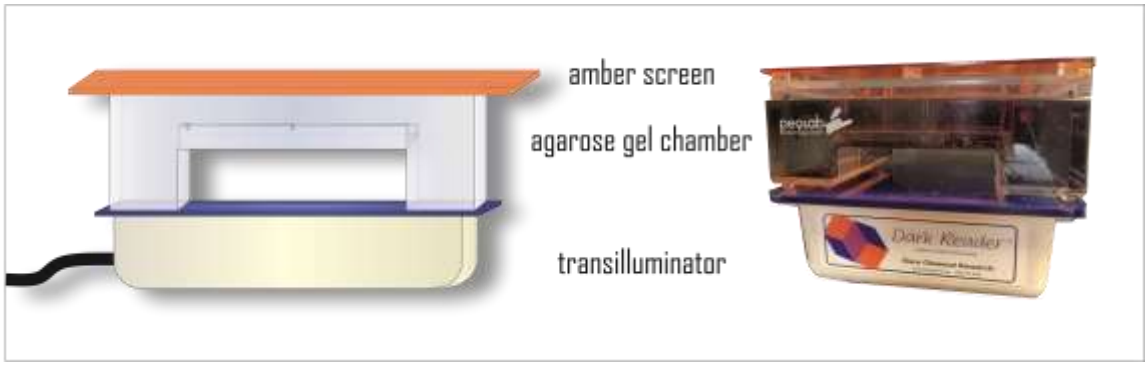


**Figure 3.8 Schematic setup (left) and photograph (right) of the real-time gel elution.**
The agarose gel chamber was placed on top of the transilluminator which used visible blue light instead of UV light as excitation source. To visualize the RNA during running of the gel an amber screen was needed and the agarose gel was prestained with SYBR™Gold. Background from photograph removed.

Direct and continuous visualization allowed determination of the exact time period in which the desired RNA migrated into the elution trough, and thus stepwise pipetting to recover the RNA was possible. This was put into practice by restarting electrophoresis after the initial recovery to allow more target RNA to enter the trough. This process was repeated for one to two minutes untill none of the target RNA was visible. A link/QR code is provided in figure 6.6 that leads to a video showing an examplary real-time gel elution of a ligation sample. The plastic housing of the agarose gel chamber prevented irradiation of the gel with UV light of wavelength 254/365 nm (mercury UV transilluminator) and impeded sensitive detection of RNA. In contrast, the setup used with a DarkReader™ transilluminator (from Clare Chemical Research) emitted blue light and was a suitable illumination source. To recover the purest possible RNA, the separation between the target RNA and the nearest migrating impurity (the longest not full-length ligation product) was a crucial step. In figure 3.7 purified full-length IS construct is shown after phenol-chloroform extraction of low melting/gelling agarose (left, lane 1) or after real-time gel elution (right, lane 6) next to the crude ligation mix (lane 2) or an *in vitro* transcript (lane 5) respectively. Both purification methods showed residual, not fully ligated, by-products, indicated by the arrows. Since the real-time gel elution setup allowed flexible positioning of the elution trough, the separation quality was able to be modulated by changing the gel length for migration. Another possibility was the percentage of agarose used to cast the gel, as

it was assumed that a higher agarose concentration led, in principle, to a better separation. However, increasing the concentration of agarose up to 2 % led to better separation in exchange for a lower recovery rate. For this reason, agarose concentrations between 0.8 to 1 % were used in all cases and only the position of the elution trough was adjusted. Additional to the improvement of the separation process, the elution trough was rinsed several times with 1x running buffer (TBE, tris/borat/EDTA) prior to the target RNA migrating into the pocket to flush residual impurities out of the pocket, thus ensuring the purest possible environment for the target RNA species. Under these conditions, after several dozen independent experiments, an elution rate of 15 - 20 % was obtained for the ligation product from the crude ligation mixture. Nevertheless, it was necessary to consider the substantial loss, which resulted in an overall yield of about 5 % due to the combination of the low ligation efficiency and the elution rate. However, there seemed to be a direct correlation between the elution rate and the length of the target RNA, since the recovery rate was increased up to 50 % for an RNA with only ~650 nts length in contrast to the full-length ligation product with >1300 nts. In addition, the input amount of RNA for the respective recovery experiment seemed not to be detrimental (cf. figure 6.7). This theory was also supported by the fact that purification of 28S rRNA subunit (>4700 nts long) from a mouse liver total RNA mixture yielded less than 10 % of the expected amount of rRNA [405] while the yield of the 18S subunit (>1800 nts) was comparable to the yield gained from full-length ligation product. One explanation for these observations might be the general tendency of long RNAs to be more strongly retained by the gel matrix than smaller RNAs [406,407] resulting in broader bands and a larger smear pattern along the entire length of the migrated gel pathway. Although these results appeared to be consistent, they were only the result of a series of individual experiments, and replicates of the experiments must be performed for validation.

## 3.4 Quality control of ligation products

Although the integrity of the full-length ligation product was ensured after real-time gel elution (cf. figure 3.7, lane 5), this was not the only criterion for good quality purification. As already mentioned, RNA molecules are more prone to degradation and therefore needed to be handled with caution compared to DNA molecules. All experiments were performed routinely with ultrapure water (Milli-Q$^{TM}$) to prevent degradation of the sample from ribonucleases present in the distilled water. Another method described in the literature for inactivating ribonucleases is to treat the water with DEPC (diethyl pyrocarbonate), which reacts irreversibly with amino acids and is thus able to destroy the active centre of the ribonucleases [408-410]). Since the performance of ultrapure water was shown to be as good as DEPC treated water [411], there was no need for this time-consuming treatment for solvents like the running buffer that were used during the procedure of real-time gel elution.

### 3.4.1 Agarose particle tracking

It has been previously mentioned that purification from agarose gels is associated with contamination by agarose particles in the sample [392]. Several purification methods described in section 3.3 as well as real-time gel elution ended up with rather large amounts of agarose that were visually detectable by precipitated pellets, which were generally larger than expected for pure RNA (please see figure 6.8 for an exemplary picture). For further investigation of these impurities within an RNA preparation, a series of samples were analyzed after purification by a device called “Nanotracker”. The instrument was routinely used to track trajectories of small particles, e.g. liposomes, by light scattering via a microscope visualized by a digital video camera [412]. This instrument allowed visual inspection of the samples with the additional possibility of automatically estimating their sizes based on diffusion coefficients. However, it is considered reliable only for approximately round-shaped nanoscale particles, but not for particles of undefined shape, possibly due to the low diffusion in the aqueous environment (data not shown).

For analysis, samples were filled up to a volume of one milliliter, injected into the sample reservoir, and visually inspected. To minimize personal bias, 20 to 30 randomly selected camera positions were analyzed by counting the number of particles per image and calculating their two-dimensional area. A more detailed description of the procedure was provided in section 5.3.2.

All samples subjected to the various agarose purification methods, including glass wool filtering of normal agarose, phenol-chloroform extraction of LMG agarose and real-time gel elution were subsequently analyzed (left site of figure 3.9). The samples were examined with regard to their total amount of agarose particles and the particles were categorized into small, medium and large based on their two dimensional area. Pictures for different particles sizes are shown in the middle of figure 3.9. For all samples an expected agarose accumulation was observed but the amount of agarose particles in the phenol-chloroform extracted sample differed noticeably from the samples after glass wool filtration and real-time gel elution, respectively. Interestingly, the distribution of particle sizes was slightly different for the real-time gel eluted sample. Here, demonstrating a tendency towards smaller particles compared to the sample filtered with glass wool, which contained the overall similar amount of particles.

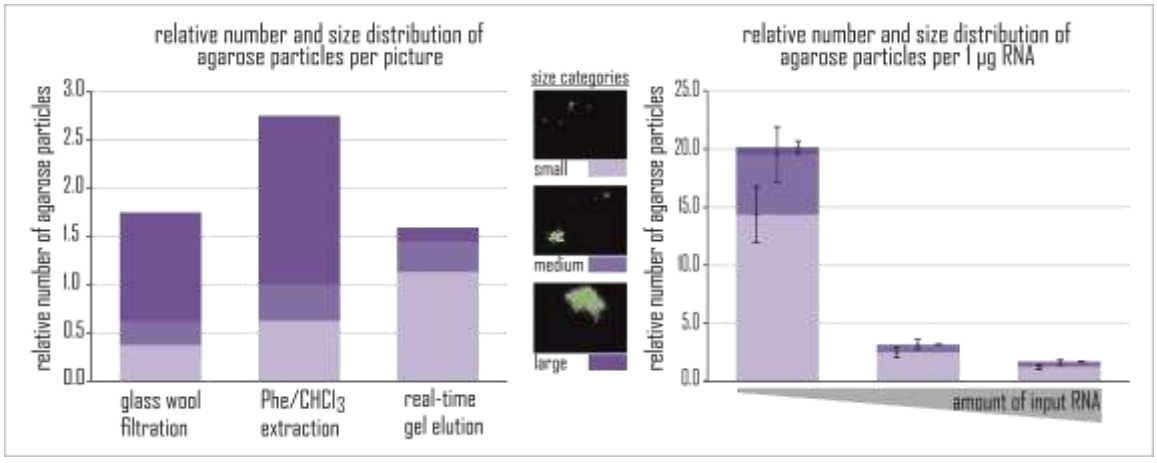


**Figure 3.9 Agarose particle tracking after different treatments.**

The particle tracking was performed by using the Nanoparticle tracking analyte and the respective software (Malvern Panalytical). The relative number of particles per inspected picture as well as the size of these particles are depicted for three different purification methods (glass wool filtration, phenol-chloroform (Phe/$CHCl_3$) extraction and real-time gel elution) on the left side. The different size categories are exemplary shown in the middle of the figure as well as the color legend for the different particle sizes (light to dark purple). On the right site the impact of increasing amounts of input RNA on the number of agarose particles per 1 µg real-time gel eluted RNA is seen for input amounts of 1.3, 4.5 and 16 µg, respectively. Standard deviations resulted from tree independent experiments.

The experiments showed an advantage of the real-time gel elution method over the other methods tested in terms of absolute number and size of agarose particles, but the question of the best RNA amount for one purification circle remained. To this end, the particle-to-RNA ratio was further examined to clarify a possible correlation between a variable amount of RNA input material and the volume recovered from the elution trough. Three samples with increasing amounts of RNA were purified by real-time gel elution and the number of agarose particles was considered relative to 1 µg of RNA. On the right site of figure 3.9, the relative number of agarose particles are shown in combination with their size distribution. Regarding the total amount of agarose particles, a clearly visible decrease was observed between the lowest input quantity and the medium input quantity. The elution volume for real-time gel elution was less dependent on the amount of RNA, but the size of the elution trough was pivotal, and thus the results supported the theory that the elution volume, rather than the amount of RNA, was the decisive factor for sample recovery.

The yield of recovered RNA in this setup was up to 50 %. While the short length of the used *in vitro* transcript of about 620 nts seemed to be a plausible explanation for the yields (cf. section 3.3.3), a dependency on the amount of input RNA was not observed in experiments with increasing amounts of input material (cf. figure 6.7).

Even though a high number of particles was detected in the sample with the lowest input quantity, these particles did not negatively affect translation efficiency (please read further in section 3.6). Nevertheless, handling, concentration determination, and downstream analysis such as LC-MS/MS were easier when the amount of particles in the sample was lower and the liquid was thus less viscous. Therefore, a purification method that results in as few agarose particles as possible shall always be preferred.

### 3.4.2 Detection of specific modifications after irradiation-induced damage

In addition to the characterization of contamination by agarose particles, another source of potential “contamination” in the form of chemical damage at the nucleoside level caused by the irradiation of the sample with blue light during real-time gel elution was investigated. It was known from the literature that UV irradiation may cause photoproducts such as 8-oxoguanosine [79,81] or hydrates of cytidine and uridine [81,83]. Although the RNA molecules tested were mainly artificial constructs, Estevez *et al.* have shown that even naturally modified RNAs can be damaged by only UV-A irradiation [79], indicating that it is essential to examine the harmfulness of the method studied.

Therefore, induced damage of *in vitro* transcripts was studied after irradiation with either UV light with a wavelength of 254 nm (UV-C) or with blue light with a wavelength between 400 and 500 nm, namely with the DarkReader™ transilluminator used for real-time gel elution. Irradiation with UV-C light (conventional mercury UV lamp, 254 nm/365 nm) was used as a positive control for conditions known from literature to cause irradiation damage in nucleic acids even after seconds [78]. A sample stored at room temperature in the dark for the same period as the treated samples served as a negative control. In the underlying approach, mRNA samples were treated for two hours, which corresponds to an exposure time four times longer than that required for real-time gel elution. After the treatment, samples were digested to nucleoside level and analyzed via a neutral loss scan (NLS) via LC-MS/MS [98,413]. Therefore, the nucleosides were ionized and filtered in quadrupole 1 of the triple quadrupole mass spectrometer for a predefined mass range. After this first filtering step, the ionized nucleosides were fragmented, in the collision cell (quadrupole 2) at the N-glycosidic bond, leading to a neutral loss of 132 Da. Finally, the transfer of nucleosides for detection was allowed only after a filtering step at quadrupole 3, which was focused on this mass transfer and to the exclusion of all other mass transfers. In the used setup this neutral loss of 132 Da was the ribose moiety of the nucleosides.

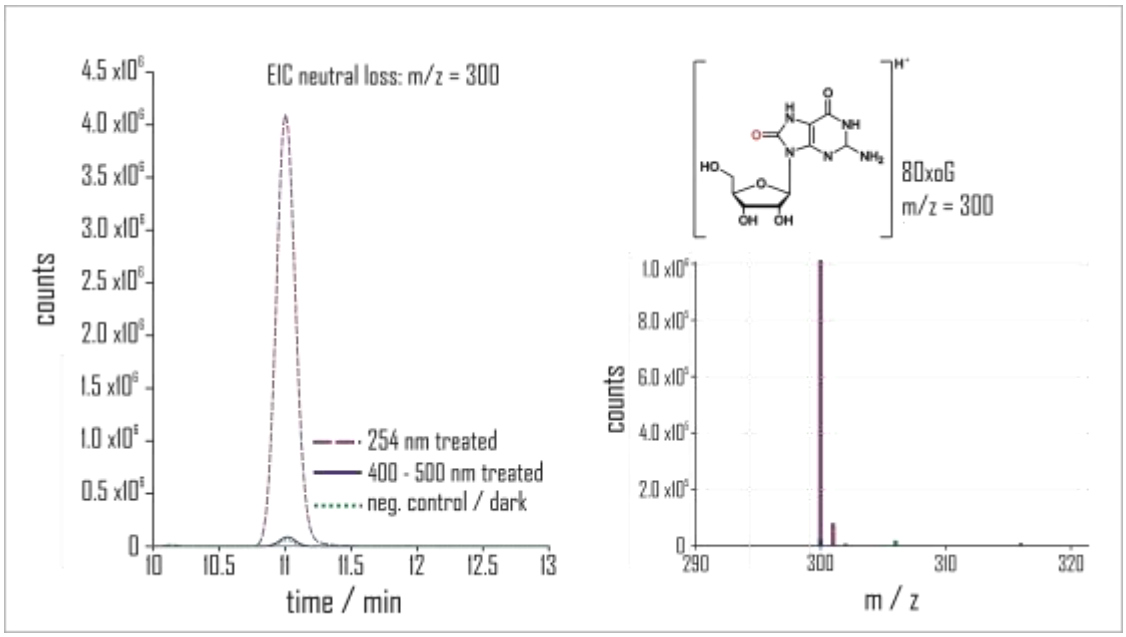


**Figure 3.10 Detection of oxidative damage products via LC-MS/MS.**
The left site illustrates the extracted ion chromatograms of mass-to-charge ratio of 300 (corresponding to 8-oxoguanosine, structure on the right) from three differently treated *in vitro* transcribed mRNAs after two hours of irradiation with UV-C (254 nm, purple dashed line), blue light (400 - 500 nm, blue line) from the DarkReader™ or incubation in the dark (neg. control, green dashed line). The right site showed the respective mass spectrums of the compound that was eluted after 11 minutes.

Based on the mass-to-charge ratios of the presumed photoproducts, the analysis of extracted ion chromatograms with exactly these mass-to-charge ratios was of greater interest. The left site of figure 3.10 shows the signal of a substance eluting after 11 minutes, detected in the extracted ion chromatogram corresponding to a mass-to-charge ratio of 300 Da, which is equal to the mass-to-charge ratio of 8-oxoguanosine (positively charged ion in the upper right part of figure 3.10). After irradiation of the sample with 254 nm (dashed purple line), a strong signal was detected, in contrast to only a very small signal after treatment with the DarkReader™ (blue line), which was comparable to the signal of the negative control (kept in the dark at room temperature for two hours, dashed green line). The mass-to-charge ratio was verified using the mass spectrums of the substance eluted after 11 minutes, shown on the lower right site of figure 3.10. In addition to the mass-to-charge ratio, 8-oxoguanosine was compared to an authentic standard to verify the results (data not shown).

Besides the results found for 8-oxoguanosine, the analysis did not reveal any other significant changes in the appearance of other RNA damage products known from literature [78,81,83], e.g. U-hydrate, between the three samples.

All in all, the artificial irradiation process with blue light, equal to the exposure during the real-time gel elution, did not provoke the formation of oxidative damage on nucleoside level. Consistent with the findings described in literature, the use of an excitation light source providing low energy blue light is milder compared to the often-used conventional UV-C light [78,81]. Kladwang and colleagues observed a high extent of oxidative damage after irradiation with 254 nm, but substantially less damage at already slightly longer wavelengths of 302 nm and finally no detectable damage after irradiation with 366 nm for 100 seconds. Since 366 nm wavelength already belong to the spectrum of UV-A light, they consider only a very inefficiently creation of photoproducts on nucleic acids after UV-A exposure [78]. Compared to the experiments performed here with even higher wavelengths, this indicates that the results were reliable and that no oxidative damage was to be expected from real-time gel elution.

### 3.4.3 Stability of $m^1A$ under applied reaction conditions

Of the various modifications considered for point modification, 1-methyladenosine ($m^1A$) was known to be susceptible under alkaline conditions. Among the reaction conditions to which the modified RNA was exposed, the TBE buffer with a pH of about 8.3 was supposed to be the most alkaline condition. In combination with a slightly elevated temperature during real-time gel elution, the question arose whether these conditions favoured a mechanism known from the literature called Dimroth rearrangement, in which $m^1A$ is chemically converted into $m^6A$ [414]. To investigate whether $m^1A$ was converted to $m^6A$ to a significant extent, which might affect downstream analysis of mRNA translation, two short oligoribonucleotides containing $m^1A$ were utilized. Even though the most alkaline condition of the workflow was gel elution in 1x TBE buffer, the conditions for phosphorylation of the 5' end of the oligoribonucleotide required for subsequent ligation, the ligation condition itself, and digestion of the splint DNA after ligation were

also tested. After treatment of the oligoribonucleotides with the respective reaction conditions, the samples were digested to the nucleoside level and analyzed by LC-MS/MS. The LC-MS/MS analysis was performed in the dynamic multiple reaction monitoring (dMRM) [415] mode. In contrast to the above-mentioned neutral loss scan, the dMRM only detected predefined nucleosides. For this purpose, the respective retention times and mass transitions were defined prior to analysis, which, together with a comparison to authentic standards, enables sensitive detection and quantifications [416]. During the analysis, the nucleosides were selected in quadrupole 1 for predefined precursor ions and after fragmentation at the N-glycosidic bond in the collision cell only the predefined product ions were allowed to pass quadrupole 3 for transmission to the detector.

The analysis was performed in a relative manner, i.e. comparing relative intensities of $m^1A$ and $m^6A$ normalized to the signal obtained from the UV absorption of cytidine in the respective samples. None of the tested conditions shifted the $m^1A$ to $m^6A$ ratio measured in the untreated samples towards higher amounts of $m^6A$. It must be noted, that the oligoribonucleotides used for these experiments, had already been stored for some years at -20 °C which might be a possible explanation for the ratio of almost 20 % $m^6A$ observed in the untreated control. However, under harsh alkaline conditions (5 minutes at 96 °C in bicarbonate buffer pH 9.2) the rearrangement of $m^1A$ into $m^6A$ was nearly complete (cf. figure 3.11 and figure 6.9) and confirmed the results reported by Macon & Wolfenden around 1960 [414].

The results of the performed experiments revealed that under the conditions during the ligation approach, no significant rearrangement of $m^1A$ into $m^6A$ was observed. This opened the possibility of synthesizing point-modified mRNAs with $m^1A$ modification. From this point on, however, the focus shifted to the use of other modifications, in particular the variants of 2'-O-methylation (2'-O-modivariants), since, on the one hand, the detection of these modifications was easily possible by LC-MS/MS and sequencing (i.e. RiboMethSeq) methods and, on the other hand, they showed interesting immunostimulatory properties (cf. section 1.5.5) in addition to their possible effects on translation.

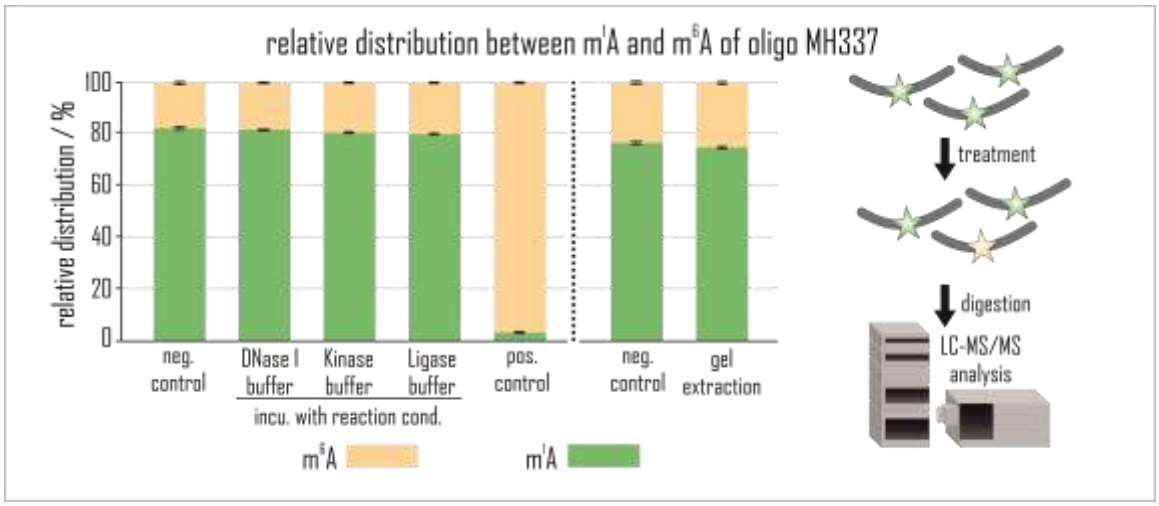


**Figure 3.11 Stability of $m^1A$ in oligoribonucleotide MH337 after different treatments.**
The distribution between $m^1A$ and $m^6A$ was analyzed from nucleosides via LC-MS/MS after exposure of the oligoribonucleotide MH337 to different reaction conditions and buffers and subsequent digestion. The negative controls were analyzed without further treatment and the positive control was treated with bicarbonate buffer pH 9.2 for 5 min at 96 °C. The conditions for the treated samples were as follows: DNase I buffer pH 7.5 for 30 min at 37 °C, buffer from T4 polynucleotide kinase pH 7.6 for 1 h at 37 °C, T4 RNA ligase 2 buffer pH 7.5 for 24 h at 16 °C. Extraction from agarose gel was performed independently as described in 5.3.4 incl. negative control (as above). All samples were analyzed in technical triplicates.

## 3.5 Quality control after including modified fragment #2

In the above sections, the integrity and quality of the unmodified ligation samples were verified. In the next step, the ligation of the IS construct was performed using a set of modified fragments #2 to introduce a modification around or in the start codon. A list of all modified oligoribonucleotides used, is listed in table 5.2. Besides a standard gel electrophoresis ensuring a successful purification, a few modivariants were selectively checked for proper incorporation of the modification after ligation via LC-MS/MS and via a sequencing approach.

Regarding the LC-MS/MS analysis, the measurement was performed in the dMRM mode including a combination of external and internal calibration which allowed an absolute quantification of modifications per mRNA. Samples were mixed in specific amounts with $^{13}C$ labeled nucleosides from *S. cerevisiae* total RNA for internal calibration after digestion to nucleoside level [416]. A detailed description of the procedure for the absolute quantification of modified nucleosides within the full-

length ligated mRNA was provided in section 5.3.4. The analysis of the samples containing fragment #2 with 2'-O-methylation ($A_m$, $U_m$ or $G_m$, respectively) revealed modification levels between 0.7 and 1 modification per mRNA and no quantifiable signal for an unmodified reference subjected to analysis was measured (figure 3.12). Due to the calculation steps and inaccuracies in this process, a modification level of 0.7 to 1 was not surprising and fragment #2 including the respective modification was assumed to be correctly incorporated in the majority of mRNAs. However, for a modivariant with pseudouridine instead of uridine inside the start codon (AΨG, Lig963), the analysis revealed an incorporation of roughly 0.55 pseudouridine molecules per mRNA. This modification level was quite low, although the purity and integrity of these samples were validated in the same way as for the other samples prior to LC-MS/MS (gel electrophoresis, data not shown). Further analysis is required, since the investigation here was only performed from a single ligation experiment. However, it had to be noticed, that this also applied for the analysis of the 2'-O-methylated modivariants and the $m^6A$ modivariant mentioned hereafter. Nevertheless, other 2'-O-modivariants were analyzed by sequencing, confirming the overall presence of a point modification (*vide infra*).

Besides the analysis of the 2'-O-methylated and pseudouridine modivariants with a reasonable result for the modification level, the analysis of a $m^6A$ modivariant revealed a modification status close to zero (cf. figure 3.12). In further analysis, it was shown that most commercial oligoribonucleotides purportedly synthesized with a ribose-$m^6A$ phosphoramidite in fact contained a deoxyribose $N^6$-methyladenosine at the position in question (figure 6.10). Therefore, these modivariants were excluded from further studies.

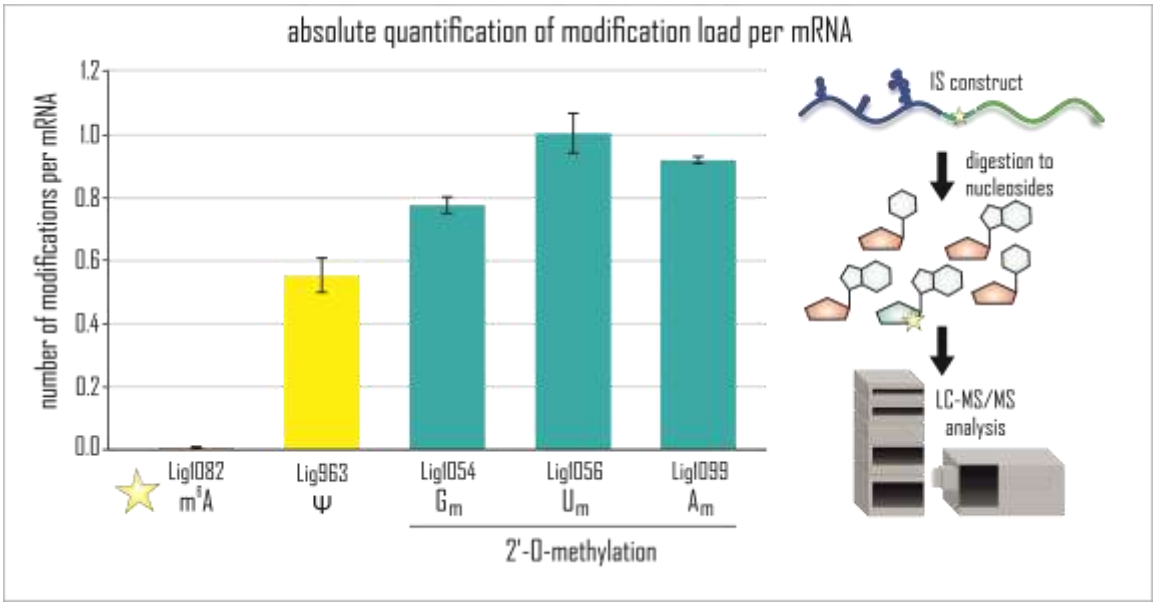


**Figure 3.12 Absolute quantification of modification levels of some IRES-eGFP ligated mRNAs by LC-MS/MS.**

After ligation using point-modified fragment #2, some modivariants were digested to nucleoside level and were analyzed using LC-MS/MS in dMRM mode. The analyzed modivariants contained either an $m^6A$ (Lig1082), a pseudouridine (Lig963) or a 2'-O-methylation (Lig1054: $G_m$, Lig1056: $U_m$ or Lig1099 $A_m$, respectively). All modivariants were named according to the used oligoribonucleotide and the position inserted modification is indicated by the star. For absolute quantification $^{13}C$, stable isotope-labeled nucleosides from *S. cerevisiae* were used as an internal standard and mixed in specific amounts with the analyzed sample. The experiment was performed in technical duplicates.

Not only the absolute number of modifications per mRNA, but also the exact position of some 2'-O-methylations inside the sequence was validated using RiboMethSeq, which allowed specific detection of ribose methylations in RNA [278,279]. In this approach, detection of 2'-O-methylation is based on a valuation of the chemical protection of this modification against random nucleolytic cleavage at the phosphodiester bond induced under alkaline conditions during sample processing (cf. 1.6.2). The normalized cleavage profile of three exemplary samples next to an unmodified control are illustrated in figure 6.11. The so-called cleavage profiles of the respective positions were evaluated as normalized cleavage counts and confirmed point modifications within all analyzed modivariants. The processing of the samples prior to RiboMethSeq and sequencing itself were performed by the group of [redacted] and analysis of the sequencing results was either performed by [redacted] or Florian Pichot.

## 3.6 Functional *in vitro* translation assay (performed by/with [redacted])

As outlined in section 3.1, the reporter of choice in the underlying work was eGFP. This optimized derivative of the original version of GFP gained higher fluorescence intensity and therefore a more sensitive read-out of the signal was possible [324,325]. One did not only have to consider the underlying reporter system, but also the various assays used for translational testing bear different limitations for various experimental setups. To test the functionality of protein expression in a first approach different *in vitro* (cell-free) translation systems were accessible. These commercially available translation assays were supposed to mimic the physiological cellular environment of the underlying system to the best possible opportunity [199,417-419] and circumvent some challenging steps from expression tests in a cell-based system such as transfection. From the above-mentioned aspects, all reporter systems in this work expressed for the enhanced derivative of GFP and were tested in a cell-free translation system based on eucaryotic translation. Therefore, the nuclease-treated rabbit reticulocyte lysate [48,365] was used, showing a stable translation of EMCV IRES fused mRNAs [420]. Subsequent readout of protein fluorescence was performed by in-gel fluorescence detection from a conventional SDS-PAGE [315]. The protein structure, crucial for its fluorescence, was found to be stable even under the conditions used during an SDS-PAGE if some requirements were fulfilled [315,317]. Therefore, heating in denaturing conditions, routinely performed prior to sample loading, had to be omitted for in-gel readout of protein fluorescence. This simple but straightforward assay to examine the integrity of the expressed protein allowed direct detection of the fluorescence activity, eliminating the need for further time-consuming analytical procedures. The development of this in-gel detection assays was performed by [redacted]. In a first validation test an *in vitro* transcript of the underlying reporter system was purified by either commercial clean-up kit (MEGAClear™ Transcription Clean-Up Kit) after *in vitro* transcription or by real-time gel elution to verify the functionality of the mRNA after the developed purification procedure. Indeed, both purification methods were shown to produce intact eGFP protein

visualized by its fluorescence. Of note, in the absence of mRNA no band corresponding to the fluorescent protein was observed (cf. figure 6.12A). At this point, it must be mentioned that the migration of native eGFP, with a theoretical molecular weight of ~27 kDa, under the gel electrophoresis conditions applied, did not match the protein ladder used, which is standardized for denatured proteins. However, this behavior was well described in literature and was another indication of the preserved folding of the protein [315,318,319]. Further optimizations were performed by [redacted] investigating optimal mRNA concentration and reaction time. The underlying experiments revealed a plateau of measured fluorescence intensity after 90 minutes of incubation, indicating a robust read-out at this timepoint with respect to the maturation time of eGFP needed for fluorescence development (cf. figure 6.12B) [313,314,326]. Even though a concentration of 36 nM mRNA was recommended by the manufacturer of the RRL *in vitro* translation system, a concentration of 20 nM was more suitable for the underlying application, because higher amounts of mRNA were leading to a saturation in protein synthesis. Therefore, a concentration in the dynamic range of the saturation curve was used for a sensitive relation to normal (cf. figure 6.12C).

In context of the study on agarose particle to mRNA ratio in purification of preparations with increasing amounts of starting material (cf. section 3.4.1), the samples were also examined with respect to differences in translation efficiency. In figure 3.13A, the results relative to a reference sample purified by commercial silica spin columns are shown. No inhibitory effect on the protein translation of agarose particles was observed, when comparing the sample with the highest agarose contamination (lowest amount of starting material) with the other two samples. Based on these results, it can be surmised that contamination with agarose particles had no measurable effect on translation efficiency. The compatibility of this translation system with point-modified mRNAs purified by real-time gel elution was demonstrated and allowed the use of this system to study the translation efficiency of a variety of modivariants. In addition, figure 3.13B shows the results of parallel detection of fluorescence and $^{35}$S-labeled methionine for three different modivariants. In contrast to the fluorescence detection, which was only able to detect fully reconstituted eGFP,

a detection with radioactive labeled protein was in principle able to detect abortive translation products. However, only full-length protein was detected with the radiolabeled translation assay. The robustness of the in-gel detection method was demonstrated by comparing these two methods with no detectable differences in relative translation efficiency. Although both methods resulted in the same pattern of translation efficiency for the investigated modification sites, it had to be considered that experiments involving work with radioactive substances must always be carried out with greater caution and involved an additional, albeit small, risk of damage to health. In addition to this disadvantage, the readout of the radioactive signal was coupled to the development of a phosphor imaging screen, which was a more time-consuming step than readout of the fluorescence in-gel subsequently of the SDS gel electrophoresis.

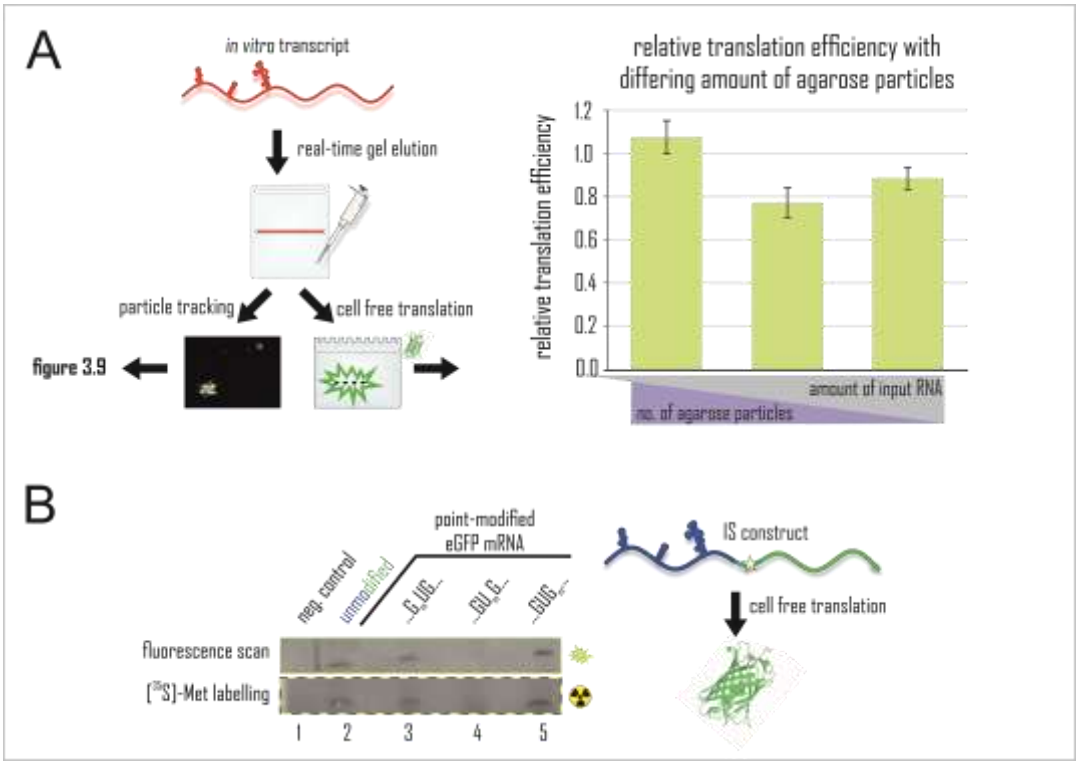


**Figure 3.13 Validation of purification and in-gel detection assay.**

**A** Depicted by the graphic on the left site: *in vitro* transcribed mRNA (red) was purified from increasing amounts (1.3, 4.5 and 16 µg, respectively) by real-time gel elution and afterwards subjected to agarose particle tracking (cf. figure 3.9) or cell free translation. The translation efficiency relative to a silica spin column (agarose particle free) purified mRNAis shown on the right. Results from three independent experiments (±SD). **B** Validation of in-gel fluorescence detection by side-by-side detection of fluorescence (fluorescence scan, upper gel picture) and development of a phosphor imaging screen ($^{35}$S-labeled eGFP, lower picture) after protein translation of unmodified, ligated eGFP mRNA (lane 2) or 2'-O-methylated eGFP mRNA on one of the three codon positions of codon two (lane 3, 4 and 5 respectively). The negative control was translated without addition of mRNA sample. Cell free translation was performed by [redacted].

## 3.7 Modified nucleosides and their impact in protein expression

Translation assays were mainly performed by [redacted], unless otherwise stated in the respective paragraph.

As already highlighted in section 1.6, the necessity to understand the possible impact some modification might have on protein biosynthesis arose from a vast increase of detection methods for these modifications in the last years [421,422]. Although the modivariants with point-modified adenosines were omitted from the presented work, it was worth noticing that especially $m^6A$ was one of the most intense studied modifications in mRNA. Previous studies found a promoting effect on translation when $m^6A$ was located in the 5' untranslated region (5' UTR) of the mRNA [195], while the translation was obstructed by disrupting tRNA selection when $m^6A$ was located in the coding sequence of the mRNA [200]. The latter also applies to 2'-O-methylations, as reported by the same working group [236]. Nevertheless, most of the investigations of effects of 2'-O-methylations in mRNA focused on the higher stability when used in combination with a 5' cap structure [24,25,423] and little literature exists investigating the translational effects of 2'-O-methylation in longer mRNAs [231]. However, the versatile literature on immunogenicity of 2'-O-methylations in tRNA [138-141] and of generally modified mRNAs [230,289,424] gave the impression that there was still much to discover on the end of 2'-O-methylated, long mRNAs. Also worth mentioning is the fact that the combination of an IRES sequence to initiate translation in the context of a point-modified mRNA had not yet been investigated.

In the following set of experiments the modivariants of point-modified pseudouridine (Ψ) and 2'-O-methylation (2'-O-Me), mentioned in section 3.5, were examined towards their translation efficacy by applying the investigated assay system presented under 3.6. It had to be noted that the principal ability of translation of the ORF construct was also confirmed, applying an unmodified ligation product for the translation assay (data not shown). Since the focus was set to the IS construct all modifications were inserted into the sequence around the start codon of the eGFP mRNA and with that the possibility to investigate the effect of a single modification

placed inside the start codon was given. The underlying sequence allowed introduction of pseudouridine at two different positions inside the sequence with one of them being the uridine of the start codon itself (AΨG). As shown in figure 3.14 this point-modified start codon almost completely abolished translation normalized to the translation rate of the unmodified mRNA. A smaller but still drastic effect was seen when inserting the Ψ at position two of the subsequent codon one, which led to approximately 75 % reduction in protein translation. These results fit to similar findings from a study published by Hoernes *et al.* 2015 using a procaryotic *in vitro* translation system and a 2-way-splint-ligation reconstituting a point-modified mRNA leading to a short peptide sequence [199]. Here, a single pseudouridine also led to an inhibitory effect in protein translation. Another important observation on the influence of a single pseudouridine modification in the context of artificial short oligoribonucleotide was made by Koutmou and coworkers. From kinetic observations of the peptide formation, they concluded that the reduction in protein expression was the consequence of reduced peptide bond formation in combination with decreased hydrolysis of guanosine-5'-triphosphate by the translation elongation factor EF-Tu leading to a reduced elongation rate [225].

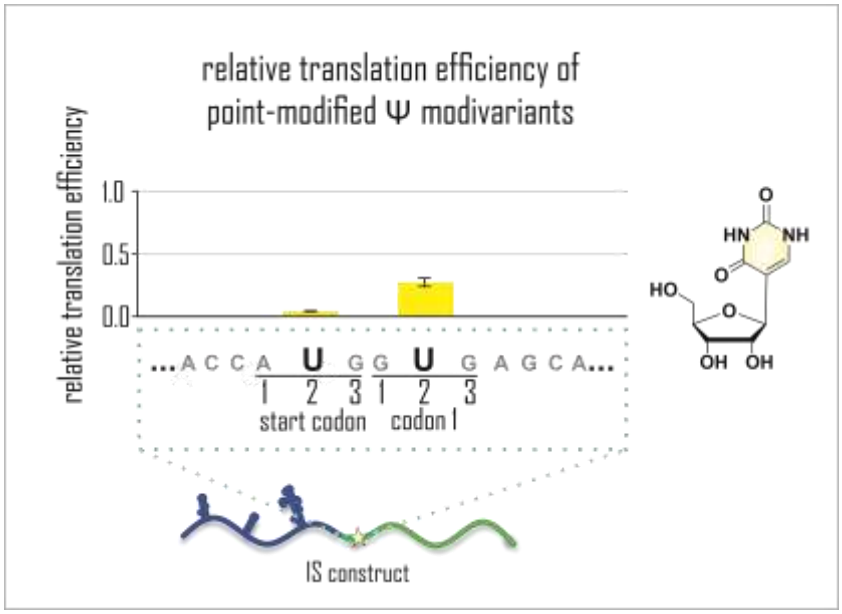


**Figure 3.14 Translation efficiency of pseudouridine modivariants.**
Two pseudouridine (structure on the right) modivariants were synthesized by splinted ligation and the translation efficiency relative to an unmodified control was investigated by in-gel fluorescence detection. The modification site was indicated by the sequence below the diagram (second position of start codon or codon 1 respectively). Results from three independent experiments (±SD). Cell free translation was performed with [redacted].

Similar to the smaller setup of pseudouridine modivariants the results of a larger cassette of 2'-O-methylated modivariants are shown in figure 3.15A. The ribosemethylation was introduced in the start codon as well as the two successive codons (codon one and codon two) of the IRES-eGFP mRNA. Although the impact of point modifications within the codon context was of great interest and will be discussed afterwards, there was opportunity to investigate point modifications within the context of the Kozak sequence as well. The Kozak sequence is a motif (RCCAUGG) discovered in eukaryotes including the start codon and was found to be responsible for protein translation initiation [425,426]. As already described in section 1.3, position between +4 (i.e. G) and position -3 (R = A or G) had a strong effect on translation initiation and were therefore well conserved [37]. Based on this knowledge, one modivariant was created with a ribose methylation at position +4 (R = $A_m$) besides the modivariants methylated in the coding sequence (including position -3, $G_m$). Interestingly, the separate modification of both positions did not show any drastic effect on normalized efficiency of translation (cf. figure 3.15). Since the Kozak sequence was crucial for eukaryotic translation initiation, this led to the assumption that the scanning mechanism of translation initiation factors was not disturbed by the ribosemethylation at position +4 or -3 ($A_m$CCAUGG and ACCAUGG$_m$ respectively).

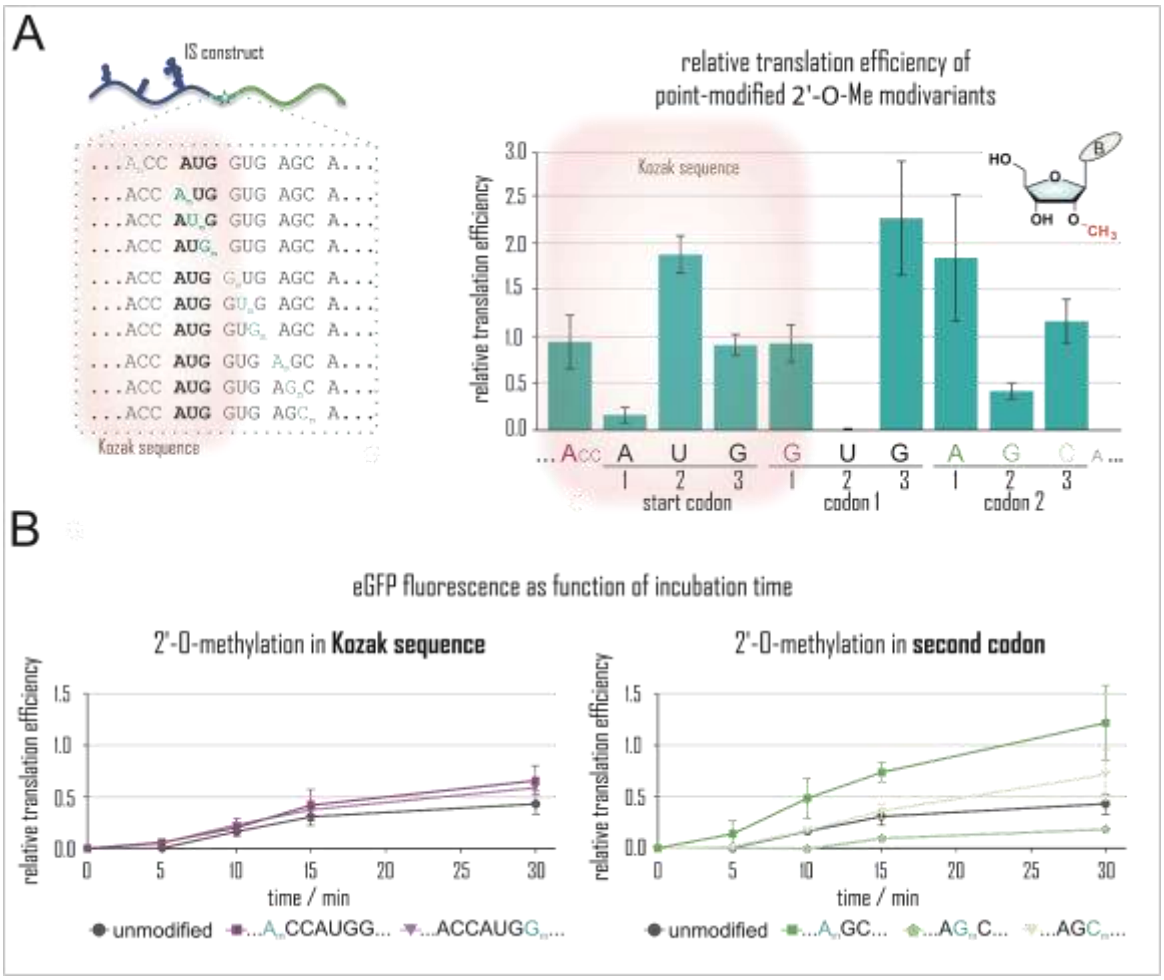


**Figure 3.15 Translation efficiency and kinetic study of 2'-O-methylated modivariants.**

**A** Each modification site of an IRES-eGFP modified mRNA was depicted by the sequences depicted on the left. The right site illustrates the results of in-gel detection of the 2'-O-methylated modivariants relative to an unmodified control after 90 minutes of incubation in RRL system. Aliquots from the translation reaction were separated on 10 % SDS-PAGE. The sequence context with respective modified position was indicated by the sequence below the bar plot, where each bar represented a single point-modified modivariant. The Kozak sequence was depicted by the pink box. **B** The fluorescence of eGFP was followed as a function of time for the first 30 minutes of translation reaction. After indicated time points, an aliquot of the translation reaction was taken, aliquots were separated on 10 % SDS-PAGE and fluorescence intensity was measured in-gel relative to unmodified control after full incubation time (90 minutes). The left site show development of fluorescence as function of time for unmodified control (grey dots), 2'-O-methylation of first position of Kozak sequence ($A_m$CCAUGG, dark purple square) and 2'-O-methylation of last position of Kozak sequence (ACCAUGG$_m$, light purple triangle). The right site show development of fluorescence as function of time for unmodified control (grey dots), 2'-O-methylation of first position of second codon ($A_m$GC, dark green square), 2'-O-methylation of second position of second codon (AG$_m$C, middle green pentagon) and 2'-O-methylation of third position of second codon (AGC$_m$, light green triangle). Results from three independent experiments (±SD). Cell free translation for kinetics was performed together with [redacted]. Analysis was performed by [redacted].

In a next step, the impact of ribosemethylation on each position of three successive codons (start codon and codon one and two) was analyzed whereas one had to consider that the start codon was belonging to the Kozak sequence as well. For position three of the start codon, the observed effect was comparable to the low effect observed for position +4 and -3 but for $A_m$ placed in the start codon ($A_mUG$) a strong inhibitory effect on protein translation was observed leading to only little fluorescent protein (figure 3.15 position 1, start codon). Even more striking was the analysis of the modivariant with $AU_mG$ start codon, which showed an almost twofold higher translation efficiency than the unmodified ligated reference mRNA. Such an effect of modification placed in second codon position had not been reported so far, especially not for the start codon.

After this closeup of the start codon and the surrounding sequence, in a next step the results were investigated from a more general point of view. Therefore, the focus was changed from investigating a single position towards the comparison of the different positions of codon triplets within the coding sequence (codon one and codon two respectively). Comparing these two codons it was noticeable that protein translation was reduced drastically (codon two) if not completely (codon one) by 2'-O-methylation of position two of both codons. One might assume that this also occurred with other codons than the two studied. Even though these experiments were not performed in the present work, a 2'-O-modivariant of the ORF construct might provide more insight. However, the presented results reflected findings from literature claiming that position two of a codon was most sensitive towards translation inhibition either on short model mRNA in procaryotic translation system [236] or after two-way-splint ligation of a longer reporter system in procaryotic *in vitro* translation system [199] or after transfection into an derivative of human embryonic kidney 293 cells (HEK293T) [226].

Choi *et al.* assumed that 2'-O-methylation disrupted translation elongation due to steric interruption of codon-anticodon recognition [236] and therefore the rejection rate of cognate tRNA during proofreading was higher compared to unmodified control. This higher rejection of cognate tRNA thus led to translational stalling. Whereby this effect was strongest on the middle position of the codon triplet. One had to recognize here that the above-mentioned studies were performed with cap-dependent

translation and in the underlying work it was possible to confirm these findings for IRES mediated cap-independent translation. Through the artificial mRNA sequence Choi *et al.* were able to study the stalling effect on position two in different codon combinations providing insights in effects like base pairing. They reported that G-C base pairing in codon-anticodon context reduced stalling duration and therefore led to a higher translation compared to other base pairs. However, this was not observed on codon one ($GU_mG$) where translation was completely abolished despite the presence of G-C base pairs on position one and two of the codon. This led to the assumption that effect of 2'-O-methylations here were only dependent on the codon position and at least for the investigated codons the presence of G-C base pairing was not of importance. Again, the findings discussed regarding G-C base pairing were observed in procaryotic translation system [236], which may explain the discrepant results.

Besides codon position two, codon position one and three did not exhibit a consistent pattern in translation efficiency. Although the first codon position of codon one is part of the Kozak sequence and therefore already discussed above, it was worth noting that compared to this modivariant ($G_mUG$ AGC), a 2'-O-methylation on position one of codon two (GUG $A_mGC$) showed higher protein translation. Noteworthy, the highest protein translation was found in the "$GUG_m$ AGC"-variant which also represented an increase in protein translation compared to the unmodified reference. This was somehow surprising since 2'-O-methylations on the third codon position, called wobble position, had shown the least effects in other studies [226,236].

To summarize the above-mentioned findings the underlying work validated the previously described inhibitory effect of 2'-O-methylations on codon position two and was able to show that this effect was independent from the mRNA sequence (i.e. short artificial sequence [236], reporter mRNA construct [226] or the presented eGFP assay system) and the used translation system (eukaryotic cells [226], procaryotic systems [236] or eukaryotic *in vitro* system, i.e. rabbit reticulocyte lysate).

The readout of eGFP fluorescence was not only dependent on protein production but on maturation of the protein as well, as mentioned previously in section 3.6. Therefore, it was of interest to study the kinetic of protein production in combination with protein maturation since fluorescence was only revealed if, on the one hand, the protein sequence was fully translated and, on the other hand, protein folding had fully proceeded. Here, focus was placed on early time points up to 30 minutes after start of translation. Figures 3.15B and 6.13 show the development of fluorescence signal starting already from 5 minutes incubation time. The time dependency of fluorescence was clearly seen comparing 2'-O-methylations on position one and two of codon two (cf. figure 3.15B, right site). The higher the protein translation, the higher the measured fluorescent signal and therefore the signal obtained from 2'-O-methylated position two indicated a weak protein synthesis and/or protein maturation since it was only measured after 10 minutes whereas a clear signal from modivariants methylated at position one was already obtained after 5 minutes. Considering that all samples were treated in the same way and that maturation of eGFP is reported to be quite robust [315], it is reasonable to assume that the delayed fluorescence measurement is due to slower protein synthesis, probably related to translational elongation [236].

However, a definitive conclusion as to why translation was affected in this particular system in the manner reported here cannot be drawn with the data collected, especially for the modivariant with respect to the start codon. Here, the observed protein synthesis revealed a remarkably different from findings previously reported [199,226,236].

# 4 Conclusion and Perspectives

Within this work the necessity of having a synthesis strategy for point-modified long mRNA at hand was addressed. This need arose from the question of the influence of mRNA modifications on protein synthesis when they are present in the coding region as well as in the non-coding region, both of which have been increasingly observed and thus this has become an urgent question [131,156,184]. The method developed is based on a 3-way-one-pot ligation moderated by a DNA splint and subsequent purification via real-time gel elution. For this purpose, the sequence of an eGFP mRNA with an EMCV IRES motif (total length >1300 nts) was selected as representative model mRNA. First, a systematic approach was developed to flexibly link individual RNA building blocks into a functional mRNA via 3-way-one-pot ligation, leading to the design of three initial constructs (namely IS, ORF and SC construct). Subsequent analysis of the ligation products confirmed the integrity of the full-length mRNA. Purification of the samples was achieved by real-time gel elution from an agarose gel, and the gel particles remaining in the sample were analyzed for their amount and size. After successful purification, the ligated mRNAs were characterized by LC-MS/MS and RiboMethSeq with respect to their modification load and possible oxidative damage after irradiation during real-time gel elution, and ligation precision was validated. Briefly, the same was true for the ligation of long mRNAs as for the ligation of short RNA species [299,364]: a cDNA splint of the maximal possible length is advantageous and hybridization is a crucial step for correct and successful ligation. However, ligation efficiency was dependent on the vicinity of the ligation site and, for reasons that are not yet elucidated, not all designed constructs resulted in a reconstituted mRNA (i.e. a ligation product). Therefore, it cannot be assumed that every designed construct leads to a successfully ligated product and a test ligation of new constructs is recommended before further experiments are planned.

Another obstacle was the use of so-called run off transcripts for fragment #1, as the T7 RNA polymerase is known to add some extra nucleotides to the 3' end of its transcript [427]. The effects of these additional nucleotides within the RNA sequence were dependent on the building block but were undesirable in all cases. Generally, these transcripts are excluded by hybridizing the RNA sequence to the DNA splint, but some might still occur. With reference to this, a first perspective experiment was performed with T4 DNA ligase, which is also specialized for double stranded nucleic acids, such as T4 RNA ligase 2 [295]. This single experiment did not reveal a remarkable difference regarding the protein yield for these two ligation approaches. However, for accurate statements in this respect, it is inevitable to repeat this experiment. In summary, reconstitutions of long mRNAs were successfully performed and ligation of point-modified mRNAs yielded the same amounts as the unmodified mRNA. The functionality of the in-gel approach investigated was validated using $^{35}$S-methionine labeling (performed with [redacted]), and point-modified mRNAs were examined for their effects on protein yield.

Several other studies have anticipated reduced, if not abolished, protein translation upon 2'-O-methylation at the second codon position [199,231], based on the idea that protein biosynthesis is affected by modifications in a position- and codon-dependent manner. The two codons of the coding sequence examined in this study showed a comparable effect, in contrast to the high increase in protein yield observed with a modified position two of the start codon. Furthermore, no evident pattern was observed for positions one and three of the codons studied. However, one had to note that the observed effects were only slightly inhibitory and, in some cases, even stimulatory, which is a remarkable observation for 2'-O-methylations and considering the literature to date, the first time observed for this type of modification. Previously reported stimulatory effects of modifications (namely $m^6A$) likely result from recognition of $m^6A$ by its reader proteins and downstream recruitment of eIF3 [197]. This may be different for 2'-O methylations, as no reader proteins are known to date, and a particular focus is on the tremendous effect of 2'-O methylation in the context of the start codon. Decoding of the start codon is subject to a special mechanism (involving specialized $tRNA_i^{Met}$) and is different from translation of the remaining RNA

sequence. Therefore, further investigation of the mechanism behind enhanced protein translation may provide insight into the potentially divergent interactions between ribosomes and mRNA in the case of point-modified mRNA.

A generalized conclusion from these results and a comparison with the literature is only possible to a limited extent, since a variety of publications indicates that the effect of modifications on translation can vary strongly depending on the translation system and sequence context [218,225,428,429]. Therefore, further studies with reconstituted mRNAs are essential to elaborate a possible dependence of the results on the underlying experimental approach. In addition, the effect of a point modification has only been tested for the vicinity of the start codon and shortly thereafter, so the effects of ribose methylation in the UTRs as well as in the rest of the coding sequence are still largely unknown. A reconstituted mRNA with potential modification site within the coding sequence has already been installed by the ORF construct and is waiting to be tested. In addition to the start codon, the stop codon might represent another interesting codon environment that failed to be covered by the SC construct, but with a slightly different ligation approach may lead to successful ligation.

Not only the modulation of the modification site, but also the variation of the translation approach might be of interest. The *in vitro* translation system utilized may be replaced or supplemented by transfection into cells or tissues. However, the limited yield after ligation allows such experiments only to a very limited extent so far.

Another possible approach might be to omit the IRES motif, since this does not correspond to the canonical translation context and most other publications operate with cap structures. Although this work targeted only point-modified mRNAs, it is in principle possible to introduce more than one modification to investigate possible synergistic effects of modifications not only on translation tuning but also with respect to their immunogenicity.

# 5 Materials and Methods

## 5.1 Materials

### 5.1.1 Instruments

| | |
|---|---|
| **Balances** | |
| Mettler Toledo Excellence Plus | Mettler Toledo (Gießen, Germany) |
| Sartorius Cubis Analytical Balance | Sartorius (Goettingen, Germany) |
| Mettler Toledo PM460 | Mettler Toledo (Gießen, Germany) |
| **Mixing & incubators** | |
| Thermoshaker Plus | Eppendorf (Hamburg, Germany) |
| VWR Digital Heatblock | VWR International (Radnor, USA) |
| GFL Schüttelinkubator 3032 | Gesellschaft für Labortechnik (Burgwedel, Germany) |
| BIOER ThermoCell | BIOER (Hangzhou, China) |
| IKA RH basic 2 | IKA-Werke GmbH & CO. KG (Staufen, Germany) |
| Heareus BB15 | Thermo Fisher Scientific (Waltham, USA) |
| Vortex Mixer (7-2020) | neoLab (Heidelberg, Germany) |
| **Centrifuges** | |
| Eppendorf Centrifuge 5810 | Eppendorf (Hamburg, Germany) |
| Sprout mini-centrifuge | Biozym (Hessisch Oldendorf, Germany) |
| 1 - 15 PK Sigma | Sigma (Osterode am Harz, Germany) |
| Eppendorf Centrifuge 5424 R | Eppendorf (Hamburg, Germany) |
| Eppendorf Centrifuge 5430 R | Eppendorf (Hamburg, Germany) |
| **HPLC & columns** | |
| Agilent 1100 HPLC system | Agilent (Böblingen, Germany) |
| Agilent 1260 Infinity LC | Agilent (Santa Clara, USA) |
| YMC-Triart C18 HPLC column (3 µm particle size, 120 Å pore size, 150 mm length, 3mm inner diameter) | YMC Europe GmbH (Dinslaken, Germany) |

| | |
|---|---|
| Synergy Fusion RP C18 column (4 µm particle size, 80 Å pore size, 250 mm length, 2 mm inner diameter) | Phenomenex (Aschaffenburg, Germany) |
| **Mass spectrometry** | |
| Agilent 6460 triple quadrupole | Agilent Böblingen (Germany) |
| Agilent 6470 TQ | Agilent Böblingen (Germany) |
| Peak scientific GENIUS XE NITROGEN | Peak Scientific Instruments GmbH (Düren, Germany) |
| NiGen LCMS 40-1 nitrogen generator | Claind (Tremezzina, Italy) |
| **Miscellaneous** | |
| Aagarose gel chamber | PeqLab (Erlangen, Germany) |
| DarkReader™ Blue Light Transilluminator | Clare Chemical Research (USA) |
| LSG-400-20 NA vertical chamber | C.B.S. Scientific (San Diego, USA) |
| Consort EV232 power supply | Consort (Turnhout, Belgium) |
| Model 250/2.5 power supply | BioRad (München, Germany) |
| Typhoon TRIO+ | GE Healthcare (Chicago, Illinois, USA) |
| Shaker (DOS-10L) | neoLab (Heidelberg, Germany) |
| Mini-PROTEAN® Tetra Vertical electrophoresis Cell | Bio-Rad (Feldkirchen, Germany) |
| Malvern NanoSight LM10 | Malvern Panalytical (Germany). |
| FiveEasy™ FE20 pH meter | Mettler Toledo (Gießen, Germany) |
| Micropipettes (Pipetting Discovery comfort 2, 10, 20, 100, 200 and 1000 µL) | Abimed (Langen, Germany) |
| Pipette boy (Integra) | VWR (Darmstadt, Germany) |
| Spectrophotometer NanoDrop ND 2000 | PeqLab (Erlangen, Germany) |
| Phosphorimager | GE Healthcare |
| Ultrapure water purification system Milli-Q | Millipore (Schwalbach, Germany) |
| CAMAG™ UV Lamp 4 | CAMAG (Muttenz, Switzerland) |
| Tri-CarbTM 4810TR | PerkinElmer (Waltham, USA) |
| GelAir™ Drying System | BioRad (München, Germany) |

### 5.1.2 Chemicals and consumables

| Gel electrophoresis and analysis | |
|---|---|
| GeneRuler 100 bp plus DNA Ladder | Thermo Fisher Scientific (Frankfurt, Germany) |
| Bromphenol blue | Merck (Darmstadt, Germany) |
| Formamide | Carl Roth (Karlsruhe, Germany) |
| Ammonium persulfate | Carl Roth (Karlsruhe, Germany) |
| TEMED | Carl Roth (Karlsruhe, Germany |
| Rotiphorese sequencing gel diluent | Carl Roth (Karlsruhe, Germany |
| Rotiphorese (10x) TBE buffer | Carl Roth (Karlsruhe, Germany) |
| Rotiphorese Gel 40 % acrylamide mix (19:1) | Carl Roth (Karlsruhe, Germany) |
| Rotiphorese sequencing gel concentrate | Carl Roth (Karlsruhe, Germany) |
| Rotiphorese® Gel B | Carl Roth (Karlsruhe, Germany) |
| Rotiphorese® Gel B (30 % acrylamide solution) | Carl Roth (Karlsruhe, Germany) |
| Rotiphorese 10x SDS-PAGE | Carl Roth (Karlsruhe, Germany) |
| Agarose | Biozym (Germany) |
| **Staining solutions** | |
| SYBR™Gold nucleic acid gel stain 10000x | Thermo Fisher Scientific (Frankfurt, Germany) |
| GelRed™ 3x | Biotium (Hayward, USA) |
| **Protein analytics** | |
| EasyTag™ L-[35S]-Methionine | PerkinElmer (Waltham, USA) |
| Tris-HCl | Carl Roth (Karlsruhe, Germany) |
| 30 % Acrylamide/Bis Solution, 29:1 | Bio-Rad (Feldkirchen, Germany) |
| Rotiphorese sequencing gel buffer concentrate | Carl Roth (Karlsruhe, Germany) |
| TWEEN 20 | Sigma Aldrich (Steinheim, Germany) |
| SDS | Roth (Karlsruhe, Germany) |
| Tris-HCl | Carl Roth (Karlsruhe, Germany) |
| **RNA/DNA isolation & EtOH precipitation and chemicals** | |
| Ammonium acetate | Merck (Darmstadt, Germany) |
| Glycogen | Thermo Fisher Scientific (Frankfurt, Germany) |
| Ethanol >99.5 % | Carl Roth (Karlsruhe, Germany) |
| Roti-aqua-phenol (RNA) | Carl Roth (Karlsruhe, Germany) |
| Roti-phenol (DNA) | Carl Roth (Karlsruhe, Germany) |
| Chloroform HPLC grade | Sigma-Aldrich (Steinheim, Germany) |

| RNA *in vitro* synthesis | |
|---|---|
| Dithiotreitol (DTT) | Thermo Fisher Scientific (Waltham, USA) |
| ATP (≥90 %, lyophilized) | Carl Roth (Karlsruhe, Germany) |
| GTP (1g, ≥90 %, lyophilized) | Carl Roth (Karlsruhe, Germany) |
| UTP (100 mg, ≥90 %, lyophilized) | Carl Roth (Karlsruhe, Germany) |
| CTP (1g, ≥98 %, lyophilized ) | Carl Roth (Karlsruhe, Germany) |
| Guanosine 5′ -monophosphate | Sigma-Aldrich (Steinheim, Germany) |
| Bovine serum albumin (BSA) | Thermo Fisher Scientific (Waltham, USA) |
| Triton X-100 | Sigma Aldrich (Steinheim, Germany) |
| **HPLC-(MS/MS) analysis** | |
| Triethylammonium acetate HPLC grade | Sigma Aldrich (Steinheim, Germany) |
| Ammonium acetate LCMS grade | Sigma Aldrich (Steinheim, Germany) |
| Acetonitrile LCMS grade | Honeywell (Morris Plains, USA) |
| Acetonitrile HPLC grade | Honeywell (Morris Plains, USA) |
| Acetic acid LCMS grade | Sigma Aldrich (Steinheim, Germany |
| **Disposables** | |
| Nanosep™ MF Centrifugal Devices (0.45 µm and 0.2 µm) | Pall (New York, USA) |
| Eppendorf tubes (1.5, 2.0 mL) | Roth (Karlsruhe, Germany) |
| Falcon® tubes (15, 50 mL) | CellStar (Frickenhausen, Germany) |
| Semi-micro cuvette, 3 ml | Sarstedt (Nümbrecht, Germany) |
| Disposable serological pipettes | Sarstedt (Nümbrecht, Germany) |
| Rotilabo-syringe filters (0.22 µm) | Carl Roth (Karlsruhe, Germany) |
| Pipette tips (10, 20, 100, 200, 1000 µL) | Carl Roth (Karlsruhe, Germany) |

### 5.1.3 Enzymes and kits

| | |
|---|---|
| T4 RNA ligase II (10 U/µL) | New England Biolabs (Frankfurt am Main, Germany |
| DNase I (50 U/µL) | Thermo Fisher Scientific (Frankfurt am Main, Germany |
| T7 RNA Polymerase | self-prepared in our lab (AK Helm) |
| T7 RNA Polymerase (20 U/µL) | Thermo Fisher Scientific (Frankfurt, Germany) |
| T4 Polynucleotide kinase (10 U/µL) | Thermo Fisher Scientific (Waltham, USA) |
| RNasin® Ribonuclease Inhibitor | Promega (Walldorf, Germany) |
| Cfr42I (SacII) (10 U/µL) | Thermo Fisher Scientific (Frankfurt, Germany) |
| BamHI (10 U/µL) | New England Biolabs (Frankfurt am Main, Germany) |
| AfeI (10 U/µL) | New England Biolabs (Frankfurt am Main, Germany) |
| BsrGI (10 U/µL) | New England Biolabs (Frankfurt am Main, Germany) |
| Nuclease P1 from *Penicillium citrinum* (lyophilized) | Sigma Aldrich (Steinheim, Germany) |
| snake venom phosphodiesterase from *C. adamanteus* (lyophilized) | Worthington (USA) |
| FastAP thermosensitive alkaline phosphatase (1 U/µL) | Thermo Fisher Scientific (Frankfurt, Germany) |
| Pentostatin >95 % | Sigma Aldrich (Steinheim, Germany) |
| InSolution Tetrahydrouridine | Merck (Darmstadt, Germany) |
| MEGAClear™ Kit | Thermo Fisher Scientific (Frankfurt, Germany) |
| Rabbit Reticulocyte Lysate System, Nuclease Treated | Promega (Walldorf, Germany) |

### 5.1.4 Buffers and solutions

All buffers and media were prepared with MilliQ-water.

| Buffer/solution | Composition |
|---|---|
| 5x Straßbourg (SB) buffer (transcription buffer) | 40 mM Tris-HCl (pH 8.1), 1 mM spermidine, 5 mM DTT, 0.01 % Triton X-100 |
| PAGE loading buffer, denaturing | 1x TBE (from 10x) and 90 % v/v formamide in water |
| PAGE loading buffer, denaturing, blue | 1x TBE (from 10x), 90 % v/v formamide, 0.1 % xylene cyanol, and 0.1 % bromphenol blue in water |
| 10 x PBS | 1.37 M NaCl, 27 mM KCl, 17 mM $KH_2PO_4$ and 100 mM $Na_2HPO_4$ (pH 6.8) |
| TEAA buffer (1 M stock) | 1 mol trimethylamine, 1 mol acetic acid in 1 L Milli-Q water, pH 7.0 |
| HPLC solvent A | 1:10 dilution of stock buffer 1 M TEAA in Milli-Q water (pH 7.0) |
| Denaturing PAGE mixtures for the separation of nucleic acids | For 100 mL 4 % denat. PAGE: 16 mL of gel concentrate (25 % solution (19:1)), 74 mL gel diluent and 10 mL sequencing gel buffer. For polymerization 400 µL 10 % APS ammonium persulfate and 50 µL Tetramethylethylendiamin (TEMED). The gel was casted using 1 mm spacers and one 1 mm comb. After polymerization, lower spacer and the comb were removed, pockets were rigorously rinsed with 1x TBE, before sample loading. |
| 10 % SDS-PAGE | **10 % resolving gel**: 3.3 mL Acrylamide/bis (30 %, Bio-Rad), 100 µL 10 % SDS, 3.75 mL, 1.5 M Tris HCl (Adjust to pH 8.8 with 6 N HCl), 2.85 mL distilled water, 10 µL TEMED and 60 µL of 10 % APS. |

| | |
|---|---|
| | **4.5 % Stacking gel**: 750 µL Acrylamide/bis (30 %), 650 µL 0.5 Tris HCl (Adjust to pH 6.8 with 6 N HCl), 3.56 mL distilled water, 50 µL of 10 % SDS, 50 µL TEMED and 50 µL of 10 % APS. |
| Bicarbonate buffer pH 9.2 | 1.05 g Sodium bicarbonate and 9.274 g sodium carbonate in 1 L water |

### 5.1.5 Software

| | |
|---|---|
| The PyMOL Molecular Graphics System, Version 2.0 | Schrödinger LLC (Braunschweig, Germay) |
| CorelDRAW V5 2020 | Corel Corporation (Ottawa, Canada) |
| IrfanView | created by Irfan Skiljan |
| Microsoft Office (Excel, Powerpoint, Access) | Microsoft (Redmond, USA) |
| Microsoft Endnote V9 | Microsoft (Redmond, USA) |
| ImagJ V5 | Created by Wayne Rasband (Madison, USA) |
| Typhoon scanner software | GE Healthcare (Buckinghamshire, UK) |
| Ape V5, | distributed through the Comprehensive R Archive Network [430] |
| ChemBioDraw Ultra 12.0 | CambridgeSoft/PerkinElmer (USA) |
| GIMP | Created by the GIMP-Team |
| MassHunter software B.05.00 | Agilent (Böblingen, Germany) |

## 5.2 Oligonucleotides and vectors

**Table 5.1 Unmodified oligoribonucleotides for ligation.**

| Name | Details | Sequence (5' to 3') | supplier |
|---|---|---|---|
| MH 905 | IS construct, fragment #2 | CCACAACCAUGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 819 | IS construct, fragment #1 | AAUUAAGAAUUCGGCUUUUACUGAUAGGUAUCGAGAUCGA | IBA (Göttigen, Germany) |
| MH 828 | IS construct, fragment #1 | GGGAGGUGUGUUAGCACACGAUUCAUAAUCAGCUACCCUCCC | IBA (Göttigen, Germany) |

| | | | |
|---|---|---|---|
| MH 829 | IS construct, fragment #1 | GGGCUCCUGUAAUUGGCGUAUGUAA CCCAGGCACCAAACACCCCCCUU | IBA (Göttigen, Germany) |
| MH 1095 | ORF construct, fragment #2 | UGAAGGGCAUCGACUUCAA | IBA (Göttigen, Germany) |
| MH 859 | SC construct, fragment #2 | AAGUAAAGCGGCCGCGGAUCCCC | Biomers (Ulm, Germany) |

**Table 5.2 Modified oligoribonucleotides for ligation (fragment #2).**

| Name | Details | Sequence (from 5’ to 3’) | supplier |
|---|---|---|---|
| MH 1052 | IS construct | CCACAACC**$A_m$**UGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1053 | IS construct | CCACAACCA**$U_m$**GGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1054 | IS construct | CCACAACCAU**$G_m$**GUGAGCAA | IBA (Göttigen, Germany) |
| MH 1055 | IS construct | CCACAACCAUG**$G_m$**UGAGCAA | IBA (Göttigen, Germany) |
| MH 1056 | IS construct | CCACAACCAUGG**$U_m$**GAGCAA | IBA (Göttigen, Germany) |
| MH 1057 | IS construct | CCACAACCAUGGU**$G_m$**AGCAA | IBA (Göttigen, Germany) |
| MH 1058 | IS construct | CCACAACCAUGGUG**$A_m$**GCAA | IBA (Göttigen, Germany) |
| MH 1059 | IS construct | CCACAACCAUGGUGA**$G_m$**CAA | IBA (Göttigen, Germany) |
| MH 1060 | IS construct | CCACAACCAUGGUGAG**$C_m$**AA | IBA (Göttigen, Germany) |
| MH 1061 | IS construct | CCACA**$m^6A$**CCAUGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1062 | IS construct | CCACAACCAUGGUGAGC**$m^6A$**A | IBA (Göttigen, Germany) |
| MH 1099 | IS construct | CCACA**$A_m$**CCAUGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 961 | IS construct | CCACAACCAUGGUG**$m^6A$**GCAA | IBA (Göttigen, Germany) |

| | | | |
|---|---|---|---|
| MH 962 | IS construct | CCACAACC**m$^6$A**UGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1082 | IS construct | CCAC**m$^6$A**ACCAUGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1083 | IS construct | CC**m$^6$A**CAACCAUGGUGAGCAA | IBA (Göttigen, Germany) |
| MH 960 | IS construct | CCACAACCAUGG**Ψ**GAGCAA | IBA (Göttigen, Germany) |
| MH 963 | IS construct | CCACAACCA**Ψ**GGUGAGCAA | IBA (Göttigen, Germany) |
| MH 1229 | IS construct | CCACAACCA(**dT-Cy5**)GGUGAGCAA | Biomers (Ulm, Germany) |
| MH 1228 | ORF construct | UGAAGGGCA(**dT-Cy5**)CGACUUCAA | Biomers (Ulm, Germany) |

**Table 5.3 Modified oligoribonucleotides for m$^1$A stability test.**

| Name | Details | Sequence (from 5' to 3') | supplier |
|---|---|---|---|
| MH 337 | Oligo used for m$^1$A stability test | AA**m$^1$A**GCUAACUUAGC | Gift from Prof. Dr. Piet Herdewijn |
| MH 340 | Oligo used for m$^1$A stability test | ACUUUUAA**m$^1$A**GGAUA | Gift from Prof. Dr. Piet Herdewijn |

**Table 5.4 Oligonucleotides applied as cDNA.**

| Name | Details | Sequence (from 5’ to 3’) | supplier |
|---|---|---|---|
| MH 1205 | 39 nts splint cDNA for ORF construct | GCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCG | Integrated DNA technologies (Coralville, USA) |
| MH 1206 | 59 nts splint cDNA for ORF construct | CCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGG | Integrated DNA technologies (Coralville, USA) |
| MH 1207 | 89 nts splint cDNA for ORF construct | ACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAAC | Integrated DNA technologies (Coralville, USA) |

| | | | |
|---|---|---|---|
| MH 1100 | standard splint cDNA for ORF construct | AGCTTGTGCCCCAGGATGTTGCCG TCCTCCTTGAAGTCGATGCCCTTC AGCTCGATGCGGTTCACCAGGGTG TCGCCCT | IBA (Göttigen, Germany) |
| MH 1182 | 39 nts splint cDNA for IS construct | CTCCTCGCCCTTGCTCACCATGGT TGTGGGATCCAAGCT | Integrated DNA technologies (Coralville, USA) |
| MH 1183 | 59 nts splint cDNA for IS construct | CGGTGAACAGCTCCTCGCCCTTGC TCACCATGGTTGTGGGATCCAAGC TTATCATCGTG | Integrated DNA technologies (Coralville, USA) |
| MH 1184 | 89 nts splint cDNA for IS construct | GGATGGGCACCACCCCGGTGAACA GCTCCTCGCCCTTGCTCACCATGG TTGTGGGATCCAAGCTTATCATCG TGTTTTTCAAAGGAAAA | Integrated DNA technologies (Coralville, USA) |
| MH 1168 | standard splint cDNA for IS construct | TGGGCACCACCCCGGTGAACAGCT CCTCGCCCTTGCTCACCATGGTTG TGGGATCCAAGCTTATCATCGTGT TTTTCAAAGG | Integrated DNA technologies (Coralville, USA) |
| MH 1217 | splint cDNA for Ligation with MH819 | GGCACCACCCCGGTGAACAGCTCC TCGCCCTTGCTCACCATGGTTGTG GTCGATCTCGATACCTATCAGTAA AAGCCGA | Integrated DNA technologies (Coralville, USA) |
| MH 1218 | splint cDNA for Ligation with MH828 | GGCACCACCCCGGTGAACAGCTCC TCGCCCTTGCTCACCATGGTTGTG GGGGAGGGTAGCTGATTATGAATC GTGTGCT | Integrated DNA technologies (Coralville, USA) |
| MH 1219 | splint cDNA for Ligation with MH829 | GGCACCACCCCGGTGAACAGCTCC TCGCCCTTGCTCACCATGGTTGTG GAAGGGGGGTGTTTGGTGCCTGG GTTACATA | Integrated DNA technologies (Coralville, USA) |
| MH 861 | splint cDNA for SC construct | TTTTTTTTTTTTTTTTTTAGCTCGG TACCCGGGGATCCGCGGCCGCTTT ACTTGTACAGCTCGTCCATGCCGA GAGTGATCCC | IBA (Göttigen, Germany) |
| MH 957 | 3'-fluorescein tagged cDNA for IS construct | GCTCCTCGCCCTTGCTCACCATGG TTGTGGCCATATTATC | IBA (Göttigen, Germany) |
| MH 958 | 3'-fluorescein tagged cDNA for SC construct | GCGGTGGCGGCCGCTTTACTTGTA CAGCTCGTCCATGC | IBA (Göttigen, Germany) |

**Table 5.5 pDNA vectors.**

| Name | Details | Resistance gene | Supplier |
|---|---|---|---|
| IRESeGFP5'-1-S1 _pUC57 | Encoding fragment #1 of IS construct | Ampicillin | GeneScript (Piscataway, USA) |
| IRESeGFP5'-3 _pUC57 | Encoding fragment #3 of IS construct | Ampicillin | GeneScript (Piscataway, USA) |
| IRESeGFP3'-1 _pUC57 | Encoding fragment #1 of SC construct | Ampicillin | GeneScript (Piscataway, USA) |
| IRESeGFPint2-1 _pUC57 | Encoding fragment #1 of ORF construct | Ampicillin | GeneScript (Piscataway, USA) |
| IRESeGFPint2-3_pUC57 | Encoding fragment #3 of ORF construct | Ampicillin | GeneScript (Piscataway, USA) |
| IRES-EMCV-S1 _pUC57 | Encoding full-length EMCV IRES-eGFP mRNA | Ampicillin | GeneScript (Piscataway, USA) |
| IRESeGFPpBSII_pUC 57 | Encoding for full-length EMCV IRES-eGFP mRNA | Ampicillin | GeneScript (Piscataway, USA) |

## 5.3 Molecular biology methods

### 5.3.1 From DNA to RNA

**Plasmid preparation and pDNA linearization**

The sequence of the respective RNA fragment was commercially available, integrated into circular plasmid pUC57 vector (GenScript, USA) and a T7-promoter. The pDNA was linearized by the respective restriction enzyme (cf. table 5.6) with enzyme concentrations applied from the manufacturer protocol and all reactions were incubated overnight at 37 °C. Subsequently, the linearized pDNA was isolated by phenol-chloroform extraction. The reaction mixture was firstly extracted with chloroform and phenol (TE-buffered, Carl Roth) at a ratio of 2:1:1 (v:v:v). This step was repeated a second time, followed by extraction with the same volume of chloroform and then once with the same volume of diethyl ether to remove residual traces of phenol. The two phases from each extraction step were well homogenized and subsequently separated by centrifugation. The DNA was recovered by ammonium acetate/ethanol precipitation (cf. section 5.3.2).

**Table 5.6 Restriction enzymes for plasmid linearization.**

| pDNA | Restriction enzyme |
|---|---|
| IRESeGFP3‘-1 _pUC57 | BsrGI (New England Biolabs) |
| IRESeGFP5‘-1_S1 _pUC57 | BamHI (New England Biolabs) |
| IRESeGFP5‘-3 _pUC57 | SacII (Thermo Fisher Scientific) |
| IRESeGFPint2-1 _pUC57 | AfeI (New England Biolabs) |
| IRESeGFPint2-3_pUC57 | SacII (Thermo Fisher Scientific) |

***In vitro* transcription of RNA and purification**

The *in vitro* transcription was performed using a T7 RNA polymerase previously in-house expressed from an addgene plasmid pQE 80L Kan or commercially available from Thermo Fisher Scientific. The linearized pDNA was incubated in 4 mM Tris-HCl (pH 8.1), 1 mM spermidine, 10 mM DTT, and 0.01 % Triton X-100, 30 mM $MgCl_2$ and 2.5 µg/mL bovine serum albumin (Thermo Fisher Scientific). NTPs (GTP, ATP, CTP, UTP) were added to a final concentration of 5 mM. A typical transcription was conducted utilizing 10 µg of DNA in a total volume of 200 µL.

For transcripts that required a 5' phosphate for downstream ligation, the final concentration of GTP was reduced to 2 mM and GMP was added at a final

concentration of 16 mM [431]. The DNA template was digested by adding 2 U of DNase I (Thermo Fisher Scientific, Germany) per 1 µg of DNA and subsequent incubation at 37 °C for 30 min. Transcription products were purified using the MEGAClear™ kit (Thermo Fisher Scientific, Germany). Sequences for commercially available oligoribonucleotides are listed in section 5.2 and *in vitro* transcribed sequences are listed in the appendix in table 6.1.

### 5.3.2 Purification, concentration and quality control of nucleic acids

**Ethanol precipitation**

The precipitation of nucleic acids was carried out by addition of 1/10 volume 5 M ammonium acetate (Merck-Millipore, Germany), 1 µL glycogen (5 mg/mL, Thermo Fisher Scientific, Germany) and 2 vol. 100 % ethanol (Carl Roth, Germany) followed by incubation at -80 °C for 1 hour or at -20 °C overnight. The DNA pellet was obtained by centrifugation at 12 000 g at -4 °C for 45 min, washed with 75 % ethanol, and centrifuged again at 12 000 g at -4 °C for 15 min. The resulting DNA pellet was resolved in ultrapure water.

**Analysis (concentration and purity determination)**

The respective DNA or RNA samples were mixed with denaturing loading dye and applied to an agarose gel with a concentration of 1 - 2 % agarose in 1x TBE buffer. The gel was either pre-stained with SYBR™Gold, which was diluted 1:10 000 when the gel was casted, or post-stained with 1x GelRed™ (from 3x Concentrate; BioTrend, Germany) for 20 minutes. The run of the gel was stopped when bromophenol blue had migrated almost to the end of the gel. Results were visualized on a Typhoon 9600 (GE Healthcare, United Kingdom). The following settings were made:

**Table 5.7 Typhoon settings.**

| Dye | Laser excitation [nm] | Emission filter [nm] |
|---|---|---|
| GelRed™ | 532 | 610 BP 30 |
| SYBR™Gold | 488/532 | 520 BP |
| Cy5 | 633 | 670 BP 30 |
| Fluorescein | 532 | 526 SP 40 |

The concentration of the samples was determined spectrophotometrically with a NanoDrop 2000 spectrophotometer (PeqLab, Germany).

**Glass wool filtration**

The glass wool filters were adapted according to previous publications for DNA extraction from agarose gels [388,389]. The bottom of a 500 µL reaction tube was punctured with a needle and packed with glass wool (Sigma-Aldrich, Germany). Subsequently, this "filter"-tube was placed into a 1.5 mL reaction tube (collection tube). An exemplary picture is depicted in figure 6.4.

**Low melting/gelling agarose**

A 1 % agarose gel was cast using low melting/gelling temperature agarose (Sigma-Aldrich, Germany). Samples were mixed 1:1 (v:v) with denaturing loading dye and subsequently loaded into the wells. RNA was visualized either by pre-staining the gel with 1x SYBR™Gold (Thermo Fisher Scientific) or by post-staining with 1 x GelRed™ (from 3x concentrate; BioTrend, Germany) for 20 minutes. Scanning was carried out using a Typhoon 9600 (GE Healthcare, United Kingdom). The desired bands were excised, and extraction of the ligation products was proceeded as described in the literature with some slight adjustments [432]. For all steps ultrapure water was utilized. The combined aqueous phases were extracted once with saturated phenol and thereafter once with chloroform.

**Real-time gel elution**

For real-time gel elution, the approach of the E-Gel™ CloneWell™ (Thermo Fisher Scientific) [404] was adapted with the DarkReader™ blue light transilluminator (Clare Chemical Research, USA) and a conventional agarose gel chamber (PeqLab, Germany). 12-well combs were utilized, with 10 wells being connected with adhesive tape. A second comb prepared in the same way was inserted into the second half of the gel chamber. 1x SYBR™Gold (Thermo Fisher Scientific, Germany) was added during the casting of the 1 % agarose gel (Biozym, Germany). The ligation reaction (mixed 2:1 with denaturing loading dye containing 90 % formamide (Roth, Germany) and 10 % 10x TBE) was added to the upper trough. The gel was run at 80 V for 15 -30 minutes and thereafter at 150 V. All fractions of interest were collected with a 1000 µl pipette. The collection trough was rinsed with 1x TBE buffer between fractions. To recover RNA from the eluted fractions, ammonium acetate/ethanol precipitation was performed by adding 1/10 volume of 5 M ammonium acetate (Merck-Millipore, Germany), 1 µL glycogen (5 mg/mL, Thermo Fisher Scientific, Germany), and two volumes 100 % ethanol (Carl Roth, Germany). Samples were incubated at -80 °C for 1 hour or at -20 °C overnight. The RNA pellet was collected by centrifugation at 12,000 g at -4 °C for 45 min, washed with 75 % ethanol, and centrifuged again at

12,000 g at -4 °C for 15 min. The final RNA pellet was dissolved in ultrapure water. Subsequently, all samples were filtered through 0.2 µm solid phase filters (Nanosep™ centrifugal device, Pall, USA).

**Agarose particle tracking**

A purified ligation sample was adjusted to a total volume of 1 mL, and subsequently injected into the sample chamber from a Malvern NanoSight LM10 (Malvern Panalytical, Germany). Scattering of a 532 nm laser allowed detection of the particles in suspension. This was visualized by a microscope with 20x magnification on which an sCMOS camera was mounted. Because of the uneven distribution of size and shape, the number of agarose particles was evaluated by visual inspection rather than by the automated Nanoparticle tracking analysis (NTA) software. Therefore, 20 positions of the camera were chosen at random, and the sizes of detected particles were measured using the software GIMP (freeware, The GIMP-Team). Particles were sorted into one of three categories, namely big, medium or small. As limits to separate the categories, the areas of particle profiles were used, such that particles with an area between 1 µm$^2$ and about 17 µm$^2$ in the above measurements were classified medium.

To test the degree of agarose particles in relation to RNA amount, a dilution series with samples containing 16, 4.5, and 1.3 µg RNA respectively was applied to real-time gel elution with typical recovery rates between 20 – 25 %. Afterwards, 1 µg of recovered RNA from the respective input amounts was pooled and analyzed. The agarose particle content corresponding to 1 µg was then compared among the samples.

**UV treatment**

7 µg of an *in vitro* transcript of full-length mRNA were diluted in 100 µL ultrapure water and were transferred to a cuvette. Afterwards, the samples were radiated for 2 hour either by 254 nm, blue light from the DarkReader™ or was kept in the dark at room temperature as negative control. Prior to digestion for LC-MS analysis the samples were precipitated.

### 5.3.3 Gel electrophoresis

**Polyacrylamide gel electrophoresis (PAGE) and gel elution**

For 4 % denaturing polyacrylamide gel solution the gel was casted in glass frames of different sizes. RNA sample (at least 100 ng) was dissolved in 2x denaturation loading buffer. In addition, a DNA ladder and a blue marker (bromophenol blue and xylene cyanol) were also applied to follow the migration process. The electrophoresis was running in 1x TBE buffer at 12 W for 3 hours in small glass frame (20 x 13 cm) and 7 hours and 10 minutes in big size glass frame (20 x 20 cm), while cooling with a fan. After migration, the gels were carefully transferred into a staining solution containing SYBR™Gold or GelRed™. After 15 - 20 minutes of shaking incubation, the gels were scanned using a Typhoon TRIO+ (GE Healthcare) (table 5.7). For gel extraction, GelRed™ solution was used for staining and band visualization. The picture resulting from the gel scanned was printed and placed under the gel to guide the excision of the corresponding band. After excision, the gel piece was crushed into smaller pieces, and transferred into a 1.5 mL Eppendorf® tube to incubate in 250 µL of 0.5 M $NH_4OAc$ (pH 5.0) for elution. After overnight incubation at 20 °C/750 rpm, the mixture was purified via a Nanosep™ spin filter with 0.45 µm membrane. The separation between the elution buffer from the gel residues by centrifugation over 15 minutes at 12 000 rpm and subsequent ethanol precipitation was performed.

**Agarose gel electrophoresis**

For the separation of DNA or RNA samples, 1 - 2 % agarose gels were prepared and pre-stained with SYBR™Gold nucleic acid gel stain or post stained with GelRed™. The staining solution was mixed in 1x TBE (roughly 1/10 of the prepared gel solution) and added to the final mixture after heating. A quantity of 75 ng - 1 µg of RNA or DNA samples were mixed with respective loading dye and after loading the samples, electrophoresis was carried out in 1x TBE buffer at 120 Volts for 90 minutes. If not pre-stained, gels were stained afterwards with either SYBR™Gold or GelRed™. Nucleic acid bands were visualized on a Typhoon TRIO+ (GE Healthcare) according to the settings illustrated in 5.7.

For calculation of ligation efficiencies, the Cy5-scan of the respective gel was analyzed by ImageJ and measured intensities of the RNA bands were normalized to the background and subsequent calculations were performed relative between ligation product and partially ligated RNA.

### 5.3.4 Liquid chromatography (LC) and mass spectrometry (MS)

**RNA analysis and purification via HPLC**

For RNA analytics and elution, the samples were applied to a YMC-Triart C18 (150 x 3 mm and 3 µm particle size) column from YMC Europe GmbH, Germany, combined with an Agilent 1100 HPLC series coupled with a DAD (diode array detector). The ligation samples were prepurified by MEGAClear™ kit (Thermo Fisher Scientific, Germany) and an amount of 30 - 40 µg RNA mixture was injected. The chromatography was realized by using 0.1 M triethylammonium acetate (pH 7.0, solvent A) and HPLC grade acetonitrile (solvent B) under denaturing conditions at 65 °C, with a flow rate of 0.4 mL/min. The elution gradient is shown in the table below.

**Table 5.8 HPLC gradient applied for ligated linear RNA purification.**

| Time / min | Solvent A / % | Solvent B / % |
|---|---|---|
| **5** | 90 | 10 |
| **6** | 87.5 | 12,5 |
| **36** | 86.7 | 13,3 |
| **41** | 86.7 | 13,3 |
| **51** | 86.2 | 13,8 |
| **57** | 86.2 | 13,8 |
| **65** | 0 | 100 |

**Absolute quantification via LC-MS/MS (dMRM mode)**

For absolute LC-MS/MS analysis in dMRM mode, 200 ng of each RNA sample was incubated overnight in 25 mM ammonium acetate (pH 7.5, Sigma-Aldrich, Germany) at 37 °C with 0.6 U nuclease P1 from *P. citrinum* (Sigma-Aldrich, Germany), 0.2 U snake venom phosphodiesterase from *C. adamanteus* (Worthington, USA), 2 U FastAP (Thermo Fisher Scientific, Germany), 200 ng pentostatin (Sigma-Aldrich, Germany), and 500 ng tetrahydrouridine (Merck-Millipore, Germany). For quantification, 50 ng of the digested sample was mixed together with 50 ng of internal standard ($^{13}$C-stable isotope-labeled nucleosides from *S. cerevisiae*, SILIS). The LC-MS/MS instrument that was used consisted of an Agilent 1260 series HPLC equipped with a diode array detector (DAD) and an Agilent 6460 triple quadrupole mass spectrometer coupled with electrospray ion source (ESI). The flow rate was 0.35 mL/min on a Synergi Fusion column (4 µM particle size, 80 Å pore size, 250 x 2.0 mm from Phenomenex). Solvents consisted of 5 mM ammonium acetate buffer (pH 5.3; solvent A) and LC-MS grade acetonitrile (solvent B; Honeywell). Elution started with 100 % solvent A, followed by a linear gradient to 8 % solvent B after 10 minutes and 40 % solvent B after 20 min. ESI parameters for quantification were as follows: Gas temperature 350 °C, gas flow

8 L/min, nebulizer pressure 50 psi, sheath gas temperature 350 °C, sheath gas flow 12 L/min, capillary voltage 3000 V. The mass spectrometer was operated in positive ion mode using Agilent MassHunter software. A combination of external and internal calibration was used for quantification as described previously [416].

**Relative quantification via LC-MS/MS for $m^1A$ stability test (dMRM mode)**

Oligoribonucleotide MH337 and MH340 were subjected for respective treatment (cf. section 3.4.3 and figure 6.9) and subsequently 200 ng of the samples were digested to nucleoside level.

Samples for relative quantification were prepared in the same manner and measured on the same instrument with the same settings. Thereafter, the resulted signal areas were normalized to the UV signal area of a main nucleoside as indicated in the results section or calculated relative to each other.

**Analysis in NLS mode**

For LC-MS analysis, after UV treatment of the sample, 5.6 µg of the respective RNA sample was incubated overnight in 25 mM ammonium acetate (pH 7.5, Sigma-Aldrich, Germany) at 37 °C with 0.6 U nuclease P1 from *P. citrinum* (Sigma-Aldrich, Germany), 0. 2 U snake venom phosphodiesterase from *C. adamanteus* (Worthington, USA) and 2 U FastAP (Thermo Fisher Scientific, Germany), 10 U benzonase (Sigma-Aldrich, Germany), 200 ng pentostatin (Sigma-Aldrich, Germany), and 500 ng tetrahydrouridine (Merck-Millipore, Germany). Subsequently, 5 µg of digested sample was administered to the system. Column, solvent and flow rate did not differ from above. ESI parameters were: Gas temperature 300 °C, gas flow 7 L/min, nebulizer pressure 60 psi, sheath gas temperature 400 °C, sheath gas flow 12 L/min, capillary voltage 3000 V. The mass spectrometer was operated in positive ion mode using Agilent MassHunter software in neutral loss scan mode with a loss of 132 Da [413].

### 5.3.5 *In vitro* translation and protein analytics

**_In vitro_ translation in Nuclease-treated Rabbit Reticulocyte Lysates**

Rabbit reticulocyte lysates, nuclease treated [367] (Promega) was programmed with 1 pmol mRNA in a final volume of 50 µL. Subsequently, the reaction was incubated for 90 minutes at 30 °C following the manufacturer's instructions. 2 µL aliquots of the reaction were mixed with non-reducing loading SDS-PAGE buffer (without $\beta$ - mercaptoethanol) and loaded without pre-heating on an 10 % SDS-polyacrylamide gel electrophoresis and run first for 30 minutes at 65 V and afterwards voltage was increased to 120 V. Protein activity was directly assessed by in-gel fluorescence detection. The gel was placed in Typhoon TRIO+ scanner (GE Healthcare, United Kingdom) and fluorescence was detected by blue laser settings (488/532 nm excitation wave length and 526 nm short pass emission filter). Quantification of eGFP signal intensity was analyzed by ImageJ normalized to the background.

For radioactive labeling, the reaction was carried as previously described and supplemented by an amino acid mixture containing all the amino acids except methionine. To that, 20 mCi [$^{35}$S]-methionine (PerkinElmer, Germany) was added. Gel was dried and subsequent translation products were visualized by Phosphor imaging.

# 6 Appendix

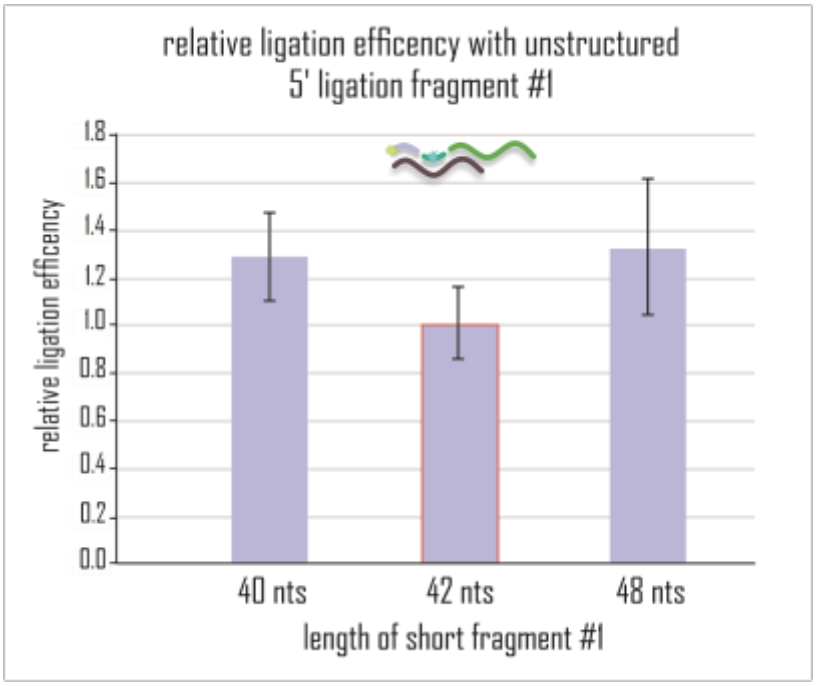


**Figure 6.1 Ligation efficiency with differing length of fragment #1.**
The ligation efficiency was calculated for ligations with different length of fragment #1 (40, 42 or 48 nts respectively) and ligation efficiency was calculated from Cy5 signals on 2 % agarose gel relative to ligation efficiency calculated for experiment with 42 nts long fragment #1 (orange box). Experiments were performed in triplicates (±SD).

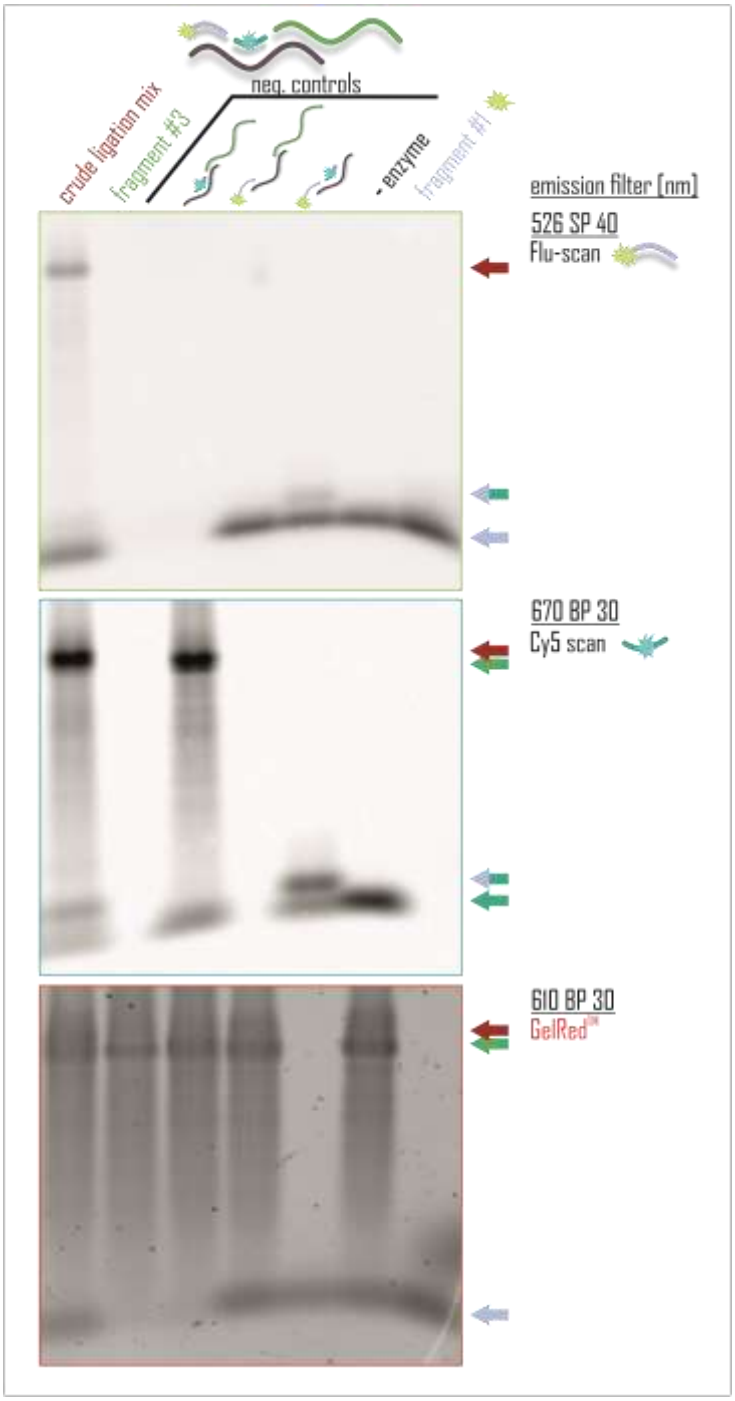


**Figure 6.2 Agarose gel picture of labeled ligation with short fragment #1.**

The crude ligation mix of ligation with short/unstructured fragment #1 were tested for integrity of the ligation product by applying fluorescein label to fragment #1 and Cy5 label to fragment#2 and analyzing the crude ligation together with some negative controls (without one of the respective fragments or without enzyme, right-handed from the black line) on 2 % agarose gel. Fragment #3 (green) and fragment #1 (light blue) served as size reference. The used emission filter is indicated, and the arrows marked the respective RNA construct with color indicating the involved fragments. From top to bottom: Fluorescein scan, Cy5 scan and scan after GelRed™ staining. A total amount of 750 ng RNA per lane was loaded.

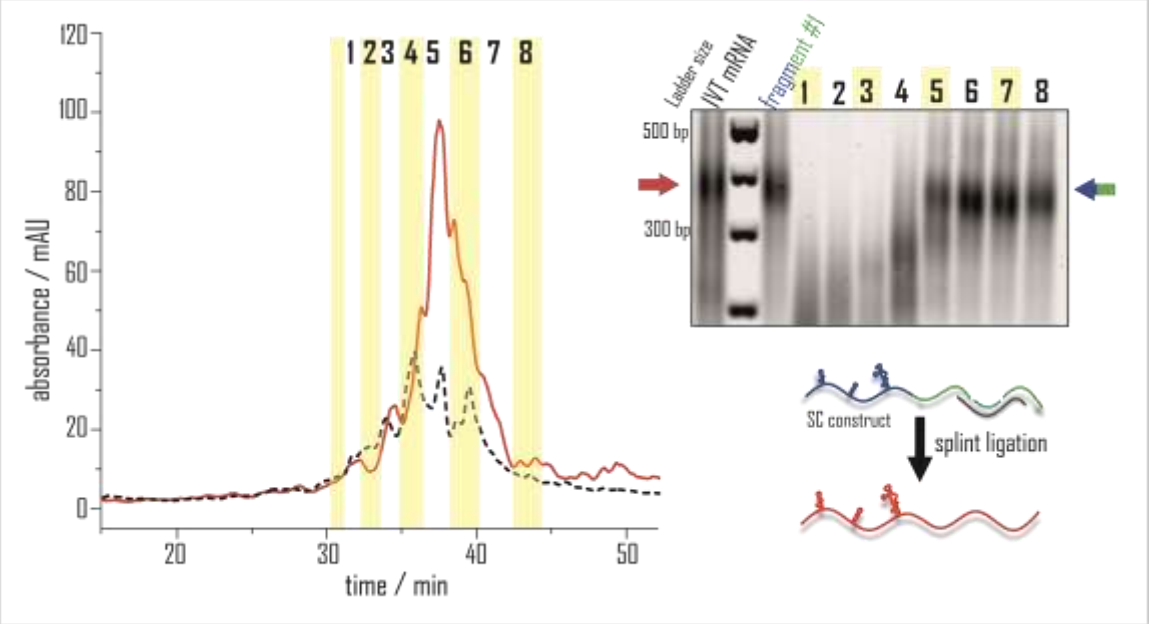


**Figure 6.3 Chromatogram of an HPLC run of ligation mixture from SC construct and subsequent fraction analysis.**

A total amount of 30 µg crude ligation mix (red line) and the corresponding *in vitro* transcribed full-length mRNA (6 µg, dashed black line) were subjected to HPLC analysis. Eight small fractions were collected manually from the crud ligation mix, indicated by the yellow and white bars. All fractions were subsequently analyzed with a 2 % agarose gel, stained with SYBR™Gold (right side). The arrows indicated the expected positions of unligated fragment #1 (blue/green) and full-length product (red). 100 bp plus DNA Ladder from Thermo Fisher Scientific was applied next to the samples.

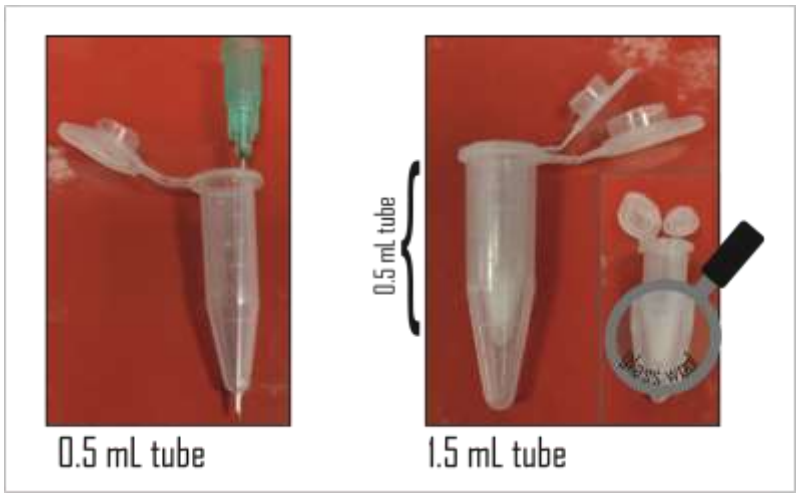


**Figure 6.4 Exemplary picture of self-made glass wool filter.**
On the left site the inner 0.5 mL tube is depicted with a hole in the bottom. Subsequently, the lower part was filled with glass wool and the tube was placed into a 1.5 mL tube (right site).

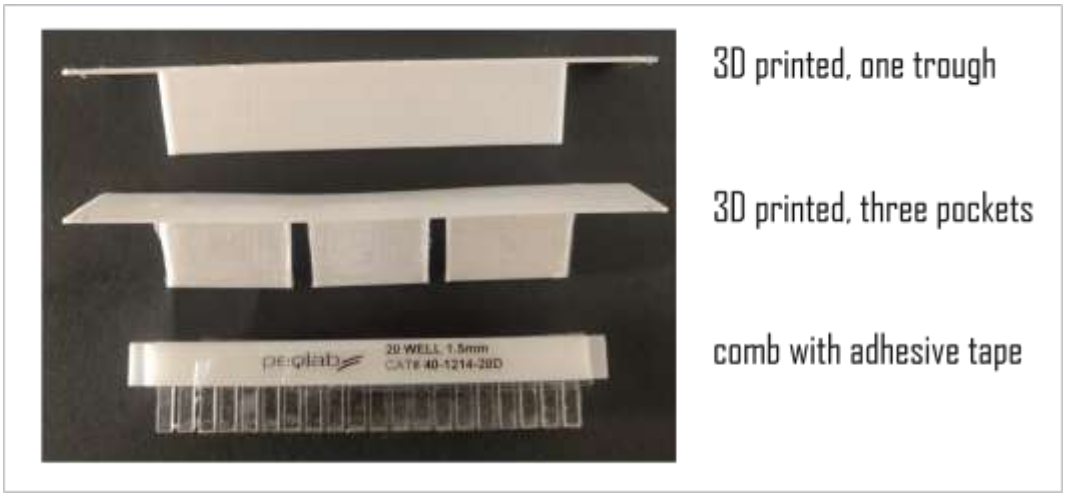


**Figure 6.5 Photographs of three different combs created for real-time gel elution.**
From Top to bottom: 3D printed comb with one trough, 3D printed comp with three large pockets and a conventional comb in which the standard pockets were joined with adhesive tape to form a trough.


**Figure 6.6 QR code for exemplary real-time gel elution.**
Scanning of the QR code with a mobile device or the link provided below the code will lead to an example elution of a ligation in real-time.

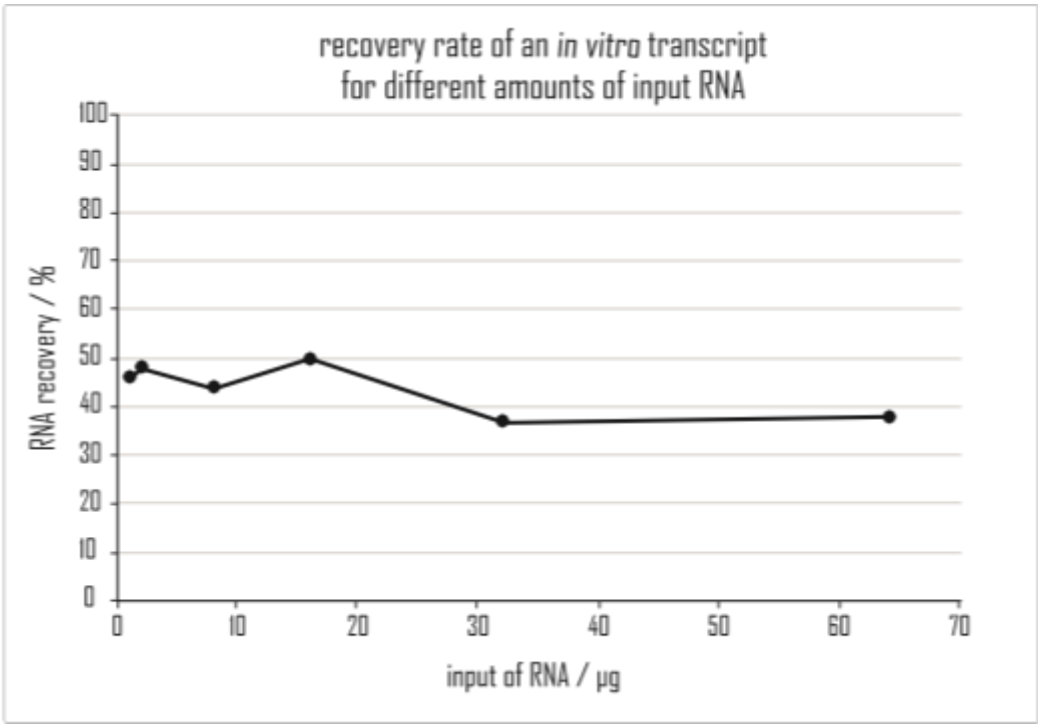


**Figure 6.7 Recovery rate after real-time gel elution for different amounts of input RNA.**
Different amounts of *in vitro* transcribed mRNA (around 620 nts long) were applied to real-time gel elution and subsequently analyzed regarding the recovery rate. The experiment was performed by Julie Wijns.


**Figure 6.8 Exemplary picture of pellet size.**

The photograph depicted the difference in size of pellets before and after filtration through 0.2 nm Nanosep™ filters.

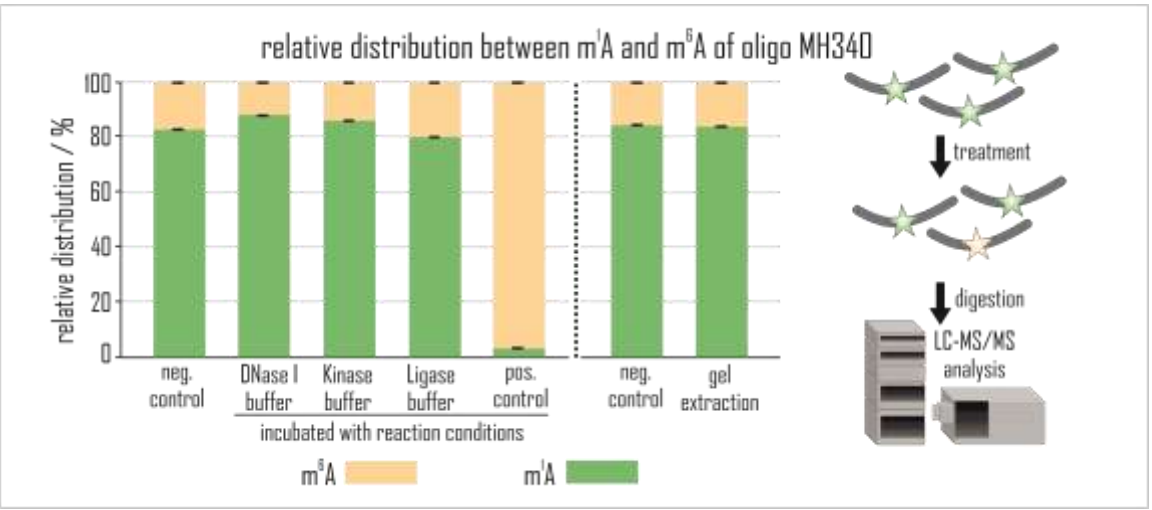


**Figure 6.9 Relative distribution between $m^1A$ and $m^6A$ of oligoribonucleotide MH340 after different treatments.**

The distribution between $m^1A$ and $m^6A$ was analyzed from nucleosides via LC-MS/MS after exposure of the oligoribonucleotide MH340 to different reaction conditions and buffers and subsequent digestion. The negative controls were analyzed without further treatment and the positive control was treated with bicarbonate buffer pH 9.2 for 5 min at 96 °C. The conditions for the treated samples were as follows: SB buffer for *in vitro* transcription pH 8.1, 5 h at 37 °C, DNase I buffer pH 7.5 for 30 min at 37 °C, buffer from T4 polynucleotide kinase pH 7.6 for 1 h at 37 °C, T4 RNA ligase 2 buffer pH 7.5 for 24 h at 16 °C. Extraction from agarose gel was performed independently as described in 5.3.2 incl. negative control (as above). All samples were analyzed in technical triplicates.

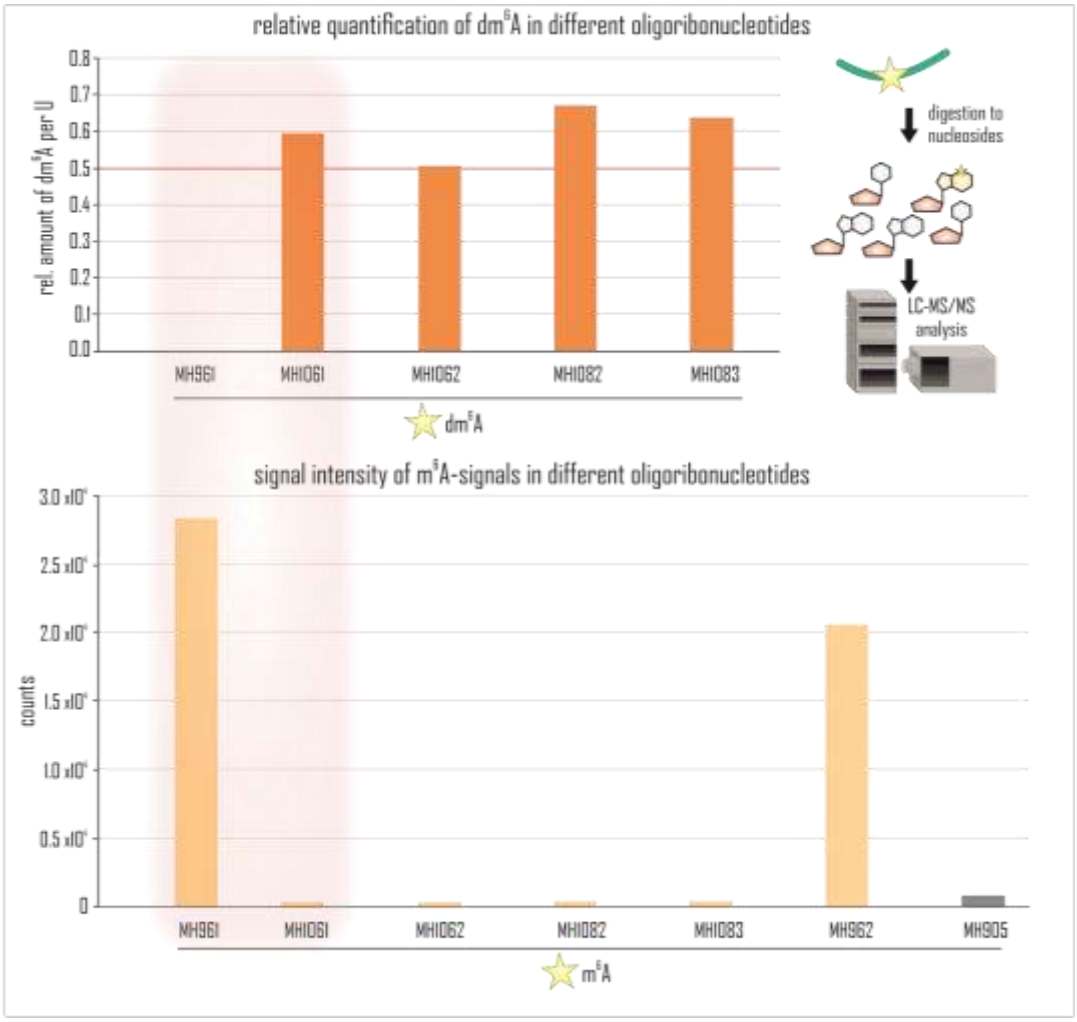


**Figure 6.10 Abundance of dm$^6$A and m$^6$A in oligoribonucleotides were measured by LC-MS/MS.**
The oligoribonucleotides function as fragment #2 and predicted to contain m$^6$A were analyzed by LC-MS/MS. In the upper bar diagram, the relative amount of dm$^6$A per U was calculated from the dMRM measurement for oligoribonucleotide MH961, MH1061, MH1062, MH1082 and MH1083. In contrast to this the abundance (as counts) of m$^6$A in the oligoribonucleotides is depicted in the lower bar diagram for the same oligoribonucleotides and additional for MH962 and unmodified MH905 (negative control).

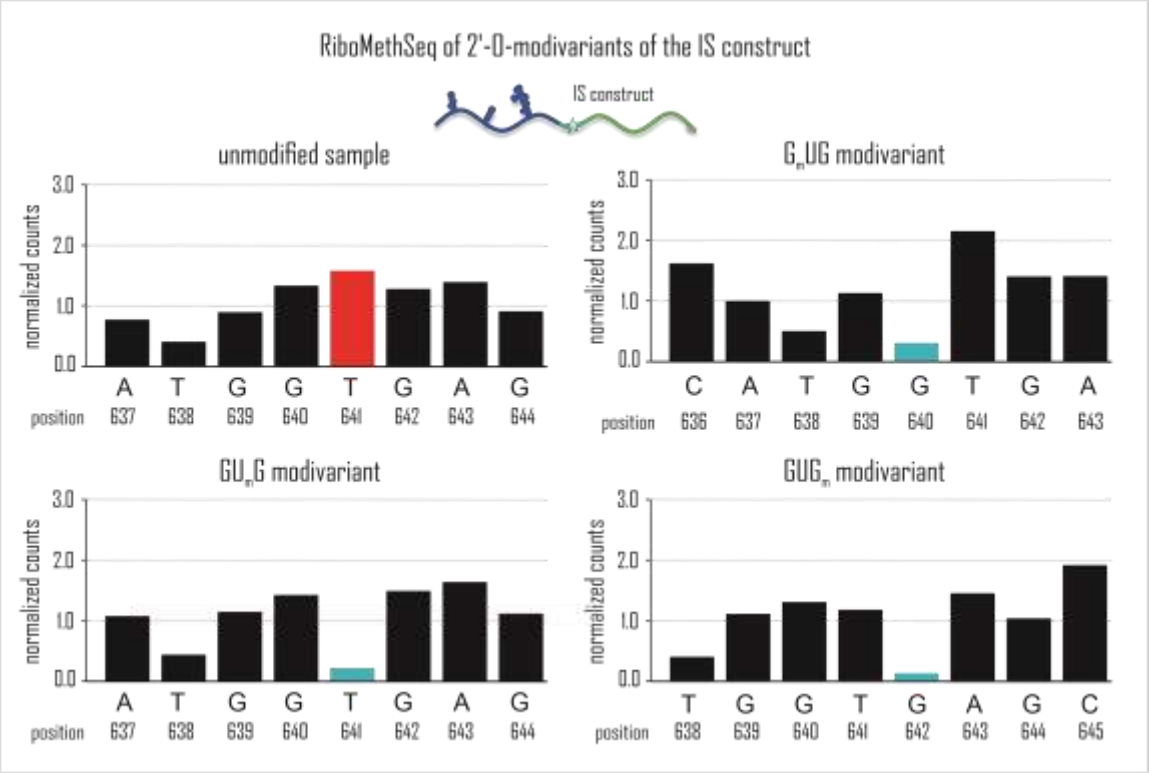


**Figure 6.11 RiboMethSeq data of three modivariants.**

RiboMethSeq protection profiles for unmodified (upper left) and three 2'-O-methylated constructs. Normalized RiboMethSeq cleavage profiles for the position 640, 641 and 642 (2'-O-methylated) are depicted with respective sequence under the bar plot. An unmodified reference is shown in the upper left with calculated cleavage profile for (unmodified) position 641. Position of modification is illustrated by the turquoise bar. Analysis was performed by [redacted].

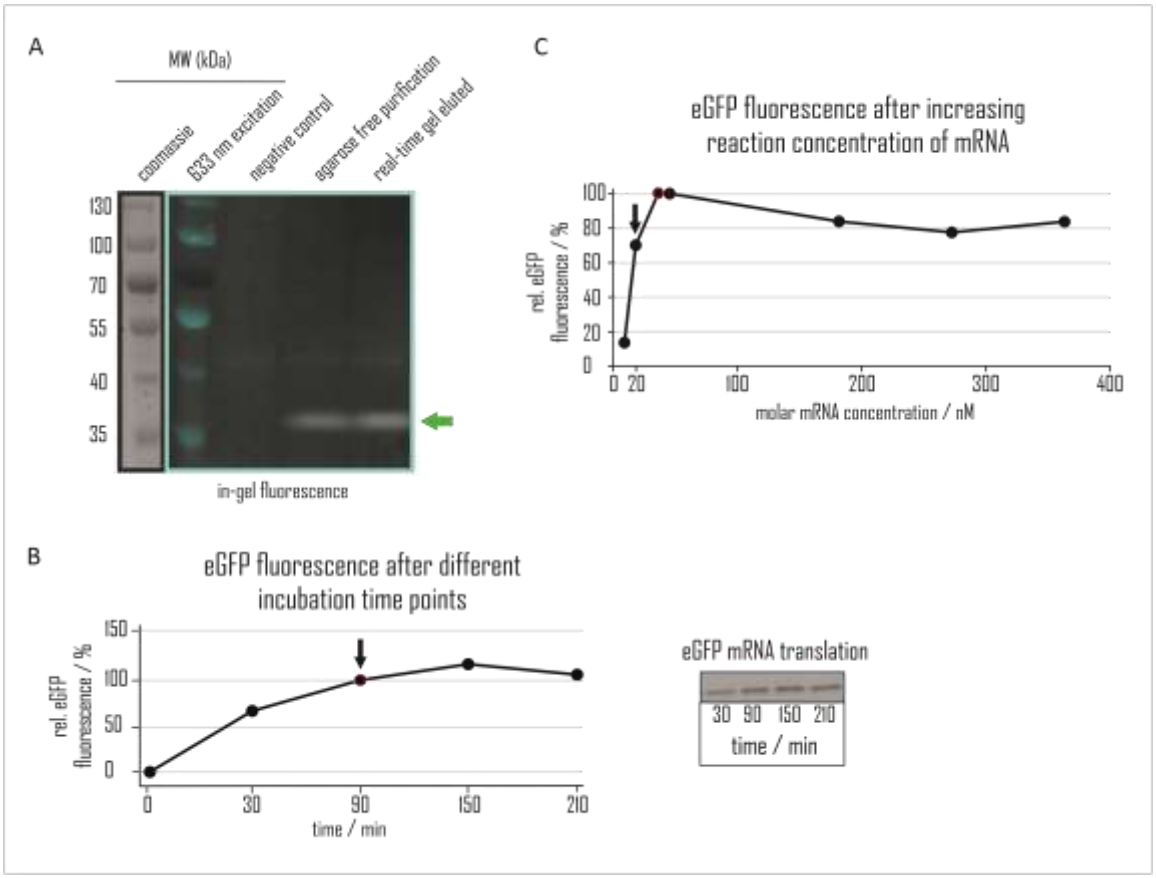


**Figure 6.12 Functionality of RRL *in vitro* translation assay.**

**A** IRES-eGFP mRNA purified either by silica-column or by real-time elution was *in vitro* translated in nuclease treated RRL. Subsequently, aliquots from the reaction mixture were loaded on 10 % SDS-PAGE. The eGFP fluorescence was measured in-gel by fluorescence scanning, using a blue laser (wavelength settings: 488 nm excitation/ 522 nm emission filter). On the left side, the gel was stained with Coomassie blue (Coomassie), then scanned using red laser 633 nm to reveal the weight size ladder bands. Negative control: Translation reaction mixture without mRNA. **B** IRES-eGFP mRNA was translated in nuclease treated RRL and protein synthesis was assessed at the indicated time points after incubation by aliquots analyzed on 10 % SDS-PAGE as described in **A**. **C** Translation reaction was performed with increasing concentrations of mRNA (10, 20, 36, 44, 180, 270 and 360 nM) and aliquots were subsequently analyzed on 10 % SDS-PAGE, in-gel detected as described in **A** and normalized to 36 nM (concentration recommended by the manufacturer, indicated with red circle). Black arrow depicted concentration used for the experiments.

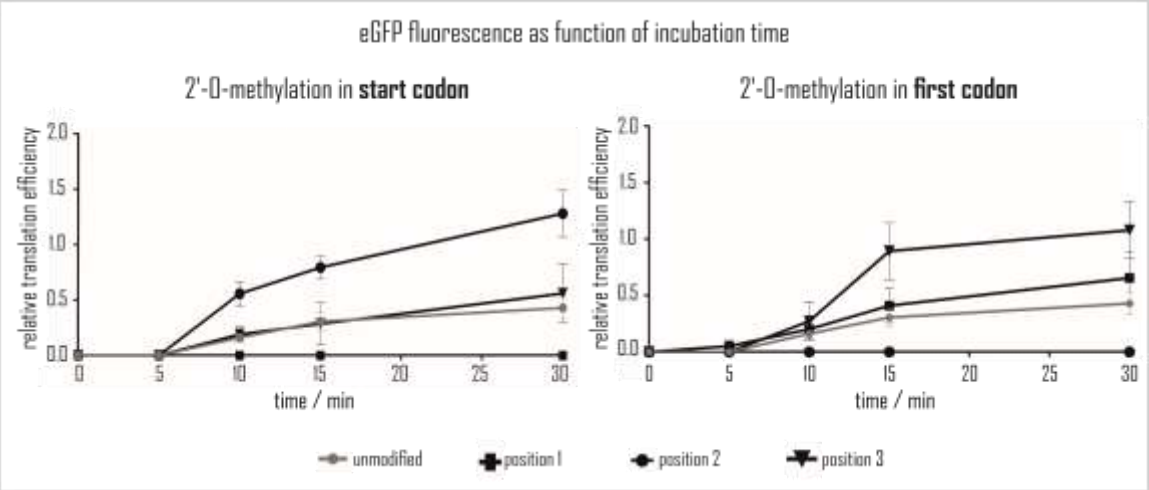


**Figure 6.13 Kinetic study of 2′-O-methylated modivariants of start codon and first transcribed codon.** The fluorescence of eGFP was followed as a function of time for the first 30 minutes of translation reaction. After indicated time points, an aliquot of the translation reaction was taken, aliquots were separated on 10 % SDS-PAGE and fluorescence intensity was measured in-gel relative to unmodified control after full incubation time (90 minutes). The left site showed development of fluorescence as function of time for unmodified control (grey dots), 2'-O-methylation of first ($A_mUG$, square), second ($AU_mG$, dots) and third ($AUG_m$, triangle) position of start codon. The right site showed development of fluorescence as function of time for unmodified control (grey dots), 2'-O-methylation of first ($G_mUG$, square), second ($GU_mG$, dot) and third ($GUG_m$, triangle) position of first transcribed codon. Results from three independent experiments (±SD). Cell free translation for kinetics were performed together with [redacted]. Analysis was performed by [redacted].

**Table 6.1 RNA sequences for all fragments and full-length mRNA.**

| Name | Sequence |
|---|---|
| IRES-eGFP S1 mRNA full length | GGGCGAAUUGGGUACCGGGCCCCCCCUCGAGGUCAUCGAAUUC<br>CGCCCCUCUCCCUCCCCCCCCCCCUAACGUUACUGGCCGAAGCC<br>GCUUGGAAUAAGGCCGGUGUGCGUUUGUCUAUAUGUUAUUUU<br>CCACCAUAUUGCCGUCUUUUGGCAAUGUGAGGGCCCGGAAACC<br>UGGCCCUGUCUUCUUGACGAGCAUUCCUAGGGGUCUUUCCCC<br>UCUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAA<br>GCAGUUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAG<br>CGACCCUUUGCAGGCAGCGGAACCCCCCACCUGGCGACAGGUG<br>CCUCUGCGGCCAAAAGCCACGUGUAUAAGAUACACCUGCAAAG<br>GCGGCACAACCCCAGUGCCACGUUGUGAGUUGGAUAGUUGUGG<br>AAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCU<br>GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUG<br>GGGCCUCGGUGCACAUGCUUUACAUGUGUUUAGUCGAGGUUA<br>AAAAAACGUCUAGGCCCCCCGAACCACGGGGACGUGGUUUUCC<br>UUUGAAAAACACGAUGAUAAGCUUGGAUCCCACAACCAUGGUG<br>AGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGG<br>UCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUC<br>CGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUG<br>AAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUGGCCCA<br>CCCUCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCG<br>CUACCCCGACCACAUGAAGCAGCACGACUUCUUCAAGUCCGCCA<br>UGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUCAAGGA<br>CGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGC<br>GACACCCUGGUGAACCGCAUCGAGCUGAAGGGCAUCGACUUCA<br>AGGAGGACGGCAACAUCCUGGGGCACAAGCUGGAGUACAACUA<br>CAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACG<br>GCAUCAAGGUGAACUUCAAGAUCCGCCACAACAUCGAGGACGG<br>CAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCG<br>GCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCAC<br>CCAGUCCGCCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACA<br>UGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG<br>GCAUGGACGAGCUGUACAAGUAAAGCGGCCGCCACCGC |
| IRES-eGFP mRNA full length | GGCGAAUUGGGUACCGGGCCCCCCCUCGAGGUCAUCGAAUUCC<br>GCCCCUCUCCCUCCCCCCCCCCCUAACGUUACUGGCCGAAGCCG<br>CUUGGAAUAAGGCCGGUGUGCGUUUGUCUAUAUGUUAUUUUC<br>CACCAUAUUGCCGUCUUUUGGCAAUGUGAGGGCCCGGAAACCU<br>GGCCCUGUCUUCUUGACGAGCAUUCCUAGGGGUCUUUCCCCU<br>CUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAAG<br>CAGUUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAGC<br>GACCCUUUGCAGGCAGCGGAACCCCCCACCUGGCGACAGGUGC<br>CUCUGCGGCCAAAAGCCACGUGUAUAAGAUACACCUGCAAAGG<br>CGGCACAACCCCAGUGCCACGUUGUGAGUUGGAUAGUUGUGGA<br>AAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCUG<br>AAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUGG<br>GGCCUCGGUGCACAUGCUUUACAUGUGUUUAGUCGAGGUUAAA<br>AAAACGUCUAGGCCCCCCGAACCACGGGGACGUGGUUUUCCUU<br>UGAAAAACACGAUGAUAAUAUGGCCACAACCAUGGUGAGCAAGG<br>GCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGC<br>UGGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGA<br>GGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUC<br>AUCUGCACCACCGGCAAGCUGCCCGUGCCCUGGCCCACCCUCG<br>UGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCC<br>CGACCACAUGAAGCAGCACGACUUCUUCAAGUCCGCCAUGCCC<br>GAAGGCUACGUCCAGGAGCGCACCAUCUUCUUCAAGGACGACG<br>GCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACAC<br>CCUGGUGAACCGCAUCGAGCUGGUGAACCGCAUCGAGCUGAAG<br>GAGGACGGCAACAUCCUGGGGCACAAGCUGGAGUACAACUACA |

| | |
|---|---|
| | ACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGC<br>AUCAAGGUGAACUUCAAGAUCCGCCACAACAUCGAGGACGGCAG<br>CGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGC<br>GACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCC<br>AGUCCGCCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUG<br>GUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCA<br>UGGACGAGCUGUACAAGUAAAGCGGCCGCCACCGCGG |
| Fragment #1 of IS construct | GGGCGAAUUGGGUACCGGGCCCCCCCUCGAGGUCAUCGAAUUC<br>CGCCCCUCUCCCUCCCCCCCCCCUAACGUUACUGGCCGAAGCC<br>GCUUGGAAUAAGGCCGGUGUGCGUUUGUCUAUAUGUUAUUUU<br>CCACCAUAUUGCCGUCUUUUGGCAAUGUGAGGGCCCGGAAACC<br>UGGCCCUGUCUUCUUGACGAGCAUUCCUAGGGGUCUUUCCCC<br>UCUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAA<br>GCAGUUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAG<br>CGACCCUUUGCAGGCAGCGGAACCCCCCACCUGGCGACAGGUG<br>CCUCUGCGGCCAAAAGCCACGUGUAUAAGAUACACCUGCAAAG<br>GCGGCACAACCCCAGUGCCACGUUGUGAGUUGGAUAGUUGUGG<br>AAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCU<br>GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUG<br>GGGCCUCGGUGCACAUGCUUUACAUGUGUUUAGUCGAGGUUA<br>AAAAAACGUCUAGGCCCCCCGAACCACGGGGACGUGGUUUUCC<br>UUUGAAAAACACGAUGAUAAGCUUGGAUC |
| Fragment #3 of IS construct | GGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGA<br>GCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGC<br>GAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGU<br>UCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUGGCCCACCCU<br>CGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUAC<br>CCCGACCACAUGAAGCAGCACGACUUCUUCAAGUCCGCCAUGC<br>CCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUCAAGGACGA<br>CGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGAC<br>ACCCUGGUGAACCGCAUCGAGCUGAAGGGCAUCGACUUCAAGG<br>AGGACGGCAACAUCCUGGGGCACAAGCUGGAGUACAACUACAAC<br>AGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGCAU<br>CAAGGUGAACUUCAAGAUCCGCCACAACAUCGAGGACGGCAGC<br>GUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGA<br>CGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAG<br>UCCGCCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGG<br>UCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAU<br>GGACGAGCUGUACAAGUAAAGCGGCCGCCACCGC |
| Fragment #1 of SC construct | UGGAGCUCCAGCUUUUGUUUACGACUCACUAUAGGGCGAAUUG<br>GGUACCGGGCCCCCCCUCGAGGUCAUCGAAUUCCGCCCCUCUC<br>CCUCCCCCCCCCCCUAACGUUACUGGCCGAAGCCGCUUGGAAUA<br>AGGCCGGUGUGCGUUUGUCUAUAUGUUAUUUUCCACCAUAUU<br>GCCGUCUUUUGGCAAUGUGAGGGCCCGGAAACCUGGCCCUGUC<br>UUCUUGACGAGCAUUCCUAGGGGUCUUUCCCCUCUCGCCAAAG<br>GAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAAGCAGUUCCUCU<br>GGAAGCUUCUUGAAGACAAACAACGUCUGUAGCGACCCUUUGC<br>AGGCAGCGGAACCCCCCACCUGGCGACAGGUGCCUCUGCGGCC<br>AAAAGCCACGUGUAUAAGAUACACCUGCAAAGGCGGCACAACCC<br>CAGUGCCACGUUGUGAGUUGGAUAGUUGUGGAAAGAGUCAAAU<br>GGCUCUCCUCAAGCGUAUUCAACAAGGGGCUGAAGGAUGCCCA<br>GAAGGUACCCCAUUGUAUGGGAUCUGAUCUGGGGCCUCGGUG<br>CACAUGCUUUACAUGUGUUUAGUCGAGGUUAAAAAAACGUCUA<br>GGCCCCCCGAACCACGGGGACGUGGUUUUCCUUUGAAAAACAC<br>GAUGAUAAUAUGGCCACAACCAUGGUGAGCAAGGGCGAGGAGC<br>UGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCG<br>ACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGCGAGGG<br>CGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUCUGCACC<br>ACCGGCAAGCUGCCCGUGCCCUGGCCCACCCUCGUGACCACCC<br>UGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAU |

| | |
|---|---|
| | GAAGCAGCACGACUUCUUCAAGUCCGCCAUGCCCGAAGGCUAC<br>GUCCAGGAGCGCACCAUCUUCUUCAAGGACGACGGCAACUACA<br>AGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA<br>CCGCAUCGAGCUGAAGGGCAUCGACUUCAAGGAGGACGGCAAC<br>AUCCUGGGGCACAAGCUGGAGUACAACUACAACAGCCACAACGU<br>CUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAAC<br>UUCAAGAUCCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCG<br>CCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUG<br>CUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCUGA<br>GCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGA<br>GUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUG<br>UAC |
| Fragment #1 of ORF construct | GGGCGAAUUGGGUACCGGGCCCCCCCUCGAGGUCAUCGAAUUC<br>CGCCCCUCUCCCUCCCCCCCCCCUAACGUUACUGGCCGAAGCC<br>GCUUGGAAUAAGGCCGGUGUGCGUUUGUCUAUAUGUUAUUUU<br>CCACCAUAUUGCCGUCUUUUGGCAAUGUGAGGGCCCGGAAACC<br>UGGCCCUGUCUUCUUGACGAGCAUUCCUAGGGGUCUUUCCCC<br>UCUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAA<br>GCAGUUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAG<br>CGACCCUUUGCAGGCAGCGGAACCCCCCACCUGGCGACAGGUG<br>CCUCUGCGGCCAAAAGCCACGUGUAUAAGAUACACCUGCAAAG<br>GCGGCACAACCCCAGUGCCACGUUGUGAGUUGGAUAGUUGUGG<br>AAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCU<br>GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUG<br>GGGCCUCGGUGCACAUGCUUUACAUGUGUUUAGUCGAGGUUA<br>AAAAAACGUCUAGGCCCCCCGAACCACGGGGACGUGGUUUUCC<br>UUUGAAAAACACGAUGAUAAGCUUGGAUCCCACAACCAUGGUG<br>AGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGG<br>UCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUC<br>CGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUG<br>AAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUGGCCCA<br>CCCUCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCG<br>CUACCCCGACCACAUGAAGCAGCACGACUUCUUCAAGUCCGCCA<br>UGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUCAAGGA<br>CGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGC<br>GACACCCUGGUGAACCGCAUCGAGC |
| Fragment #3 of ORF construct | GGAGGACGGCAACAUCCUGGGGCACAAGCUGGAGUACAACUAC<br>AACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGG<br>CAUCAAGGUGAACUUCAAGAUCCGCCACAACAUCGAGGACGGCA<br>GCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGG<br>CGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACC<br>CAGUCCGCCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAU<br>GGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGG<br>CAUGGACGAGCUGUACAAGUAAAGCGGCCGCCACCGC |

# Bibliography


[1] Crick, F.H. (1958) On protein synthesis. *Symp Soc Exp Biol*, 12, 138-163.

[2] Crick, F. (1970) Central dogma of molecular biology. *Nature*, 227, 561-563.

[3] Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, 737-738.

[4] Clancy, S. (2008) Chemical Structure of RNA. *Nature Education*, 7.

[5] Choudhuri, S. (2010) Small noncoding RNAs: biogenesis, function, and emerging significance in toxicology. *J Biochem Mol Toxicol*, 24, 195-216.

[6] Guttman, M. and Rinn, J.L. (2012) Modular regulatory principles of large non-coding RNAs. *Nature*, 482, 339-346.

[7] Fuchs, R.T., Grundy, F.J. and Henkin, T.M. (2006) The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat Struct Mol Biol*, 13, 226-233.

[8] Akiyama, T. (1978) [Qualitative analysis of urinary fibrin/fibrinogen degradation products in renal transplant rejection (author's transl)]. *Nihon Hinyokika Gakkai Zasshi*, 69, 1391-1403.

[9] Hirose, Y. and Manley, J.L. (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev*, 14, 1415-1429.

[10] Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S. (1978) The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. *Cold Spring Harb Symp Quant Biol*, 42 Pt 2, 1047-1051.

[11] Roeder, R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci*, 21, 327-335.

[12] Struhl, K. (1995) Yeast transcriptional regulatory mechanisms. *Annu Rev Genet*, 29, 651-674.

[13] Zawel, L. and Reinberg, D. (1993) Initiation of transcription by RNA polymerase II: a multi-step process. *Prog Nucleic Acid Res Mol Biol*, 44, 67-108.

[14] Wang, W., Carey, M. and Gralla, J.D. (1992) Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science*, 255, 450-453.

[15] Deng, W. and Roberts, S.G. (2007) TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma*, 116, 417-429.

[16] Keene, J.D. (2007) Biological clocks and the coordination theory of RNA operons and regulons. *Cold Spring Harb Symp Quant Biol*, 72, 157-165.

[17] Moroy, T. and Heyd, F. (2007) The impact of alternative splicing in vivo: mouse models show the way. *RNA*, 13, 1155-1171.

[18] Lee, J.E. and Cooper, T.A. (2009) Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans*, 37, 1281-1286.

[19] Fabian, M.R., Sonenberg, N. and Filipowicz, W. (2010) Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem*, 79, 351-379.

[20] Lei, E.P., Krebber, H. and Silver, P.A. (2001) Messenger RNAs are recruited for nuclear export during transcription. *Genes Dev*, 15, 1771-1782.

[21] Mortillaro, M.J., Blencowe, B.J., Wei, X., Nakayasu, H., Du, L., Warren, S.L., Sharp, P.A. and Berezney, R. (1996) A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 8253-8257.

[22] McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, 385, 357-361.

[23] Du, L. and Warren, S.L. (1997) A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. *J Cell Biol*, 136, 5-18.

[24] Ramanathan, A., Robb, G.B. and Chan, S.H. (2016) mRNA capping: biological functions and applications. *Nucleic Acids Res*, 44, 7511-7526.

[25] Lockless, S.W., Cheng, H.T., Hodel, A.E., Quiocho, F.A. and Gershon, P.D. (1998) Recognition of capped RNA substrates by VP39, the vaccinia virus-encoded mRNA cap-specific 2'-O-methyltransferase. *Biochemistry*, 37, 8564-8574.

[26] Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mittra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J. and Bujnicki, J.M. (2011) 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res*, 39, 4756-4768.

[27] Sadofsky, M. and Alwine, J.C. (1984) Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. *Mol Cell Biol*, 4, 1460-1468.

[28] Takagaki, Y. and Manley, J.L. (1997) RNA recognition by the human polyadenylation factor CstF. *Mol Cell Biol*, 17, 3907-3914.

[29] Yang, Q., Gilmartin, G.M. and Doublie, S. (2010) Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 10062-10067.

[30] Gallie, D.R. (1991) The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev*, 5, 2108-2116.

[31] Drummond, D.R., Armstrong, J. and Colman, A. (1985) The effect of capping and polyadenylation on the stability, movement and translation of synthetic messenger RNAs in Xenopus oocytes. *Nucleic Acids Res*, 13, 7375-7394.

[32] Munroe, D. and Jacobson, A. (1990) mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol Cell Biol*, 10, 3441-3455.

[33] Gray, N.K., Coller, J.M., Dickson, K.S. and Wickens, M. (2000) Multiple portions of poly(A)-binding protein stimulate translation in vivo. *EMBO J*, 19, 4723-4733.

[34] Chow, L.T., Gelinas, R.E., Broker, T.R. and Roberts, R.J. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12, 1-8.

[35] Berget, S.M., Moore, C. and Sharp, P.A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74, 3171-3175.

[36] Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*, 72, 291-336.

[37] Kozak, M. (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res*, 9, 5233-5252.

[38] Tomek, W. and Wollenhaupt, K. (2012) The "closed loop model" in controlling mRNA translation during development. *Anim Reprod Sci*, 134, 2-8.

[39] Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*, 11, 113-127.

[40] de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B. and Schneider, R.J. (2018) A widespread alternate form of cap-dependent mRNA translation initiation. *Nat Commun*, 9, 3068.

[41] Lee, S.H. and McCormick, F. (2006) p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. *EMBO J*, 25, 4008-4019.

[42] Martinez-Salas, E., Ramos, R., Lafuente, E. and Lopez de Quinto, S. (2001) Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *J Gen Virol*, 82, 973-984.

[43] Holcik, M., Sonenberg, N. and Korneluk, R.G. (2000) Internal ribosome initiation of translation and the control of cell death. *Trends Genet*, 16, 469-473.

[44] Hellen, C.U. and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev*, 15, 1593-1612.

[45] Sharma, D.K., Bressler, K., Patel, H., Balasingam, N. and Thakor, N. (2016) Role of Eukaryotic Initiation Factors during Cellular Stress and Cancer Progression. *J Nucleic Acids*, 2016, 8235121.

[46] Weingarten-Gabbay, S., Khan, D., Liberman, N., Yoffe, Y., Bialik, S., Das, S., Oren, M. and Kimchi, A. (2014) The translation initiation factor DAP5 promotes IRES-driven translation of p53 mRNA. *Oncogene*, 33, 611-618.

[47] Hundsdoerfer, P., Thoma, C. and Hentze, M.W. (2005) Eukaryotic translation initiation factor 4GI and p97 promote cellular internal ribosome entry sequence-driven translation. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 13421-13426.

[48] Borman, A.M., Bailly, J.L., Girard, M. and Kean, K.M. (1995) Picornavirus internal ribosome entry segments: comparison of translation efficiency and the requirements for optimal internal initiation of translation in vitro. *Nucleic Acids Res*, 23, 3656-3663.

[49] Kaminski, A., Howell, M.T. and Jackson, R.J. (1990) Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J*, 9, 3753-3759.

[50] Grover, R., Ray, P.S. and Das, S. (2008) Polypyrimidine tract binding protein regulates IRES-mediated translation of p53 isoforms. *Cell Cycle*, 7, 2189-2198.

[51] Lopez de Quinto, S., Lafuente, E. and Martinez-Salas, E. (2001) IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. *RNA*, 7, 1213-1226.

[52] Ochs, K., Rust, R.C. and Niepmann, M. (1999) Translation initiation factor eIF4B interacts with a picornavirus internal ribosome entry site in both 48S and 80S initiation complexes independently of initiator AUG location. *J Virol*, 73, 7505-7514.

[53] Lopez de Quinto, S. and Martinez-Salas, E. (1997) Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal initiation of translation. *J Virol*, 71, 4171-4175.

[54] Thapar, R., Denmon, A.P. and Nikonowicz, E.P. (2014) Recognition modes of RNA tetraloops and tetraloop-like motifs by RNA-binding proteins. *Wiley Interdiscip Rev RNA*, 5, 49-67.

[55] Fernandez-Miragall, O., Ramos, R., Ramajo, J. and Martinez-Salas, E. (2006) Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation. *RNA*, 12, 223-234.

[56] Pestova, T.V., Shatsky, I.N. and Hellen, C.U. (1996) Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol*, 16, 6870-6878.

[57] Kolupaeva, V.G., Pestova, T.V., Hellen, C.U. and Shatsky, I.N. (1998) Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *The Journal of biological chemistry*, 273, 18599-18604.

[58] Kolupaeva, V.G., Lomakin, I.B., Pestova, T.V. and Hellen, C.U. (2003) Eukaryotic initiation factors 4G and 4A mediate conformational changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. *Mol Cell Biol*, 23, 687-698.

[59] Lomakin, I.B., Hellen, C.U. and Pestova, T.V. (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Mol Cell Biol*, 20, 6019-6029.

[60] Pestova, T.V., Hellen, C.U. and Shatsky, I.N. (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol*, 16, 6859-6869.

[61] Jan, E. and Sarnow, P. (2002) Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J Mol Biol*, 324, 889-902.

[62] Eigner, J., Boedtker, H. and Michaels, G. (1961) The thermal degradation of nucleic acids. *Biochim Biophys Acta*, 51, 165-168.

[63] Tenhunen, J. (1989) Hydrolysis of single-stranded RNA in aqueous solutions--effect on quantitative hybridizations. *Mol Cell Probes*, 3, 391-396.

[64] Houseley, J. and Tollervey, D. (2009) The many pathways of RNA degradation. *Cell*, 136, 763-776.

[65] Ilinskaya, O.N. and Mahmud, R.S. (2014) Ribonucleases as antiviral agents. *Mol Biol*, 48, 615-623.

[66] Sohal, R.S. and Weindruch, R. (1996) Oxidative stress, caloric restriction, and aging. *Science*, 273, 59-63.

[67] Schieber, M. and Chandel, N.S. (2014) ROS function in redox signaling and oxidative stress. *Curr Biol*, 24, R453-462.

[68] Fimognari, C. (2015) Role of Oxidative RNA Damage in Chronic-Degenerative Diseases. *Oxid Med Cell Longev*, 2015, 358713.

[69] Lu, J.M., Lin, P.H., Yao, Q. and Chen, C. (2010) Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems. *J Cell Mol Med*, 14, 840-860.

[70] Imlay, J.A. (2008) Cellular defenses against superoxide and hydrogen peroxide. *Annu Rev Biochem*, 77, 755-776.

[71] Fukai, T. and Ushio-Fukai, M. (2011) Superoxide dismutases: role in redox signaling, vascular function, and diseases. *Antioxid Redox Signal*, 15, 1583-1606.

[72] Loschen, G., Azzi, A., Richter, C. and Flohe, L. (1974) Superoxide radicals as precursors of mitochondrial hydrogen peroxide. *FEBS Lett*, 42, 68-72.

[73] Kehrer, J.P. (2000) The Haber-Weiss reaction and mechanisms of toxicity. *Toxicology*, 149, 43-50.

[74] Imlay, J.A., Chin, S.M. and Linn, S. (1988) Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. *Science*, 240, 640-642.

[75] Cadet, J. and Wagner, J.R. (2013) DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb Perspect Biol*, 5.

[76] Zaira Leni, L.K., Marianne Geiser. (2020) Air pollution causing oxidative stress. *Current Opinion in Toxicology*, 20-21, 1-8.

[77] Maher, B.A., Gonzalez-Maciel, A., Reynoso-Robles, R., Torres-Jardon, R. and Calderon-Garciduenas, L. (2020) Iron-rich air pollution nanoparticles: An unrecognised environmental risk factor for myocardial mitochondrial dysfunction and cardiac oxidative stress. *Environ Res*, 188, 109816.

[78] Kladwang, W., Hum, J. and Das, R. (2012) Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci Rep*, 2, 517.

[79] Estevez, M., Valesyan, S., Jora, M., Limbach, P.A. and Addepalli, B. (2021) Oxidative Damage to RNA is Altered by the Presence of Interacting Proteins or Modified Nucleosides. *Front Mol Biosci*, 8, 697149.

[80] Singer, B. (1971) Chemical modification of viral ribonucleic acid. IX. The effect of ultraviolet irradiation on TMV-RNA and other polynucleotides. *Virology*, 45, 101-107.

[81] Sun, C., Limbach, P.A. and Addepalli, B. (2020) Characterization of UVA-Induced Alterations to Transfer RNA Sequences. *Biomolecules*, 10.

[82] Yagura, T., Schuch, A.P., Garcia, C.C.M., Rocha, C.R.R., Moreno, N.C., Angeli, J.P.F., Mendes, D., Severino, D., Bianchini Sanchez, A., Di Mascio, P. *et al.* (2017) Direct participation of DNA in the formation of singlet oxygen and base damage under UVA irradiation. *Free Radic Biol Med*, 108, 86-93.

[83] Wurtmann, E.J. and Wolin, S.L. (2009) RNA under attack: cellular handling of RNA damage. *Crit Rev Biochem Mol Biol*, 44, 34-49.

[84] Shan, X., Tashiro, H. and Lin, C.L. (2003) The identification and characterization of oxidized RNAs in Alzheimer's disease. *J Neurosci*, 23, 4913-4921.

[85] Shan, X., Chang, Y. and Lin, C.L. (2007) Messenger RNA oxidation is an early event preceding cell death and causes reduced protein expression. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 21, 2753-2764.

[86] Tanaka, M., Chock, P.B. and Stadtman, E.R. (2007) Oxidized messenger RNA induces translation errors. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 66–71.

[87] Rhee, Y., Valentine, M.R. and Termini, J. (1995) Oxidative base damage in RNA detected by reverse transcriptase. *Nucleic Acids Res*, 23, 3275-3282.

[88] Drablos, F., Feyzi, E., Aas, P.A., Vaagbo, C.B., Kavli, B., Bratlie, M.S., Pena-Diaz, J., Otterlei, M., Slupphaug, G. and Krokan, H.E. (2004) Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)*, 3, 1389-1407.

[89] Hoffman, D.R., Marion, D.W., Cornatzer, W.E. and Duerre, J.A. (1980) S-Adenosylmethionine and S-adenosylhomocystein metabolism in isolated rat liver. Effects of L-methionine, L-homocystein, and adenosine. *The Journal of biological chemistry*, 255, 10822-10827.

[90] Rydberg, B. and Lindahl, T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J*, 1, 211-216.

[91] Hamilton, J.T., McRoberts, W.C., Keppler, F., Kalin, R.M. and Harper, D.B. (2003) Chloride methylation by plant pectin: an efficient environmentally significant process. *Science*, 301, 206-209.

[92] McCulloch, A., Aucott, M.L., Benkovitz, C.M., Graedel, T.E., Kleiman, G., Midgley, P.M. and Li, Y.F. (1999) Global emissions of hydrogen chloride and chloromethane from coal combustion, incineration and industrial activities: Reactive Chlorine Emissions Inventory. *J Geophys Res-Atmos*, 104, 8391-8403.

[93] Rusyn, I., Asakura, S., Li, Y., Kosyk, O., Koc, H., Nakamura, J., Upton, P.B. and Swenberg, J.A. (2005) Effects of ethylene oxide and ethylene inhalation on DNA adducts, apurinic/apyrimidinic sites and expression of base excision DNA repair genes in rat brain, spleen, and liver. *DNA Repair (Amst)*, 4, 1099-1110.

[94] Boccaletto, P., Machnicka, M.A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A. *et al.* (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res*, 46, D303-D307.

[95] Carell, T., Brandmayr, C., Hienzsch, A., Muller, M., Pearson, D., Reiter, V., Thoma, I., Thumbs, P. and Wagner, M. (2012) Structure and function of noncanonical nucleobases. *Angew Chem Int Ed Engl*, 51, 7110-7131.

[96] Motorin, Y. and Helm, M. (2011) RNA nucleotide methylation. *Wiley Interdiscip Rev RNA*, 2, 611-631.

[97] Roundtree, I.A., Evans, M.E., Pan, T. and He, C. (2017) Dynamic RNA Modifications in Gene Expression Regulation. *Cell*, 169, 1187-1200.

[98] Dal Magro, C., Keller, P., Kotter, A., Werner, S., Duarte, V., Marchand, V., Ignarski, M., Freiwald, A., Muller, R.U., Dieterich, C. *et al.* (2018) A Vastly Increased Chemical Variety of RNA Modifications Containing a Thioacetal Structure. *Angew Chem Int Ed Engl*, 57, 7893-7897.

[99] Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.D., Lafontaine, D.L.J. and Bohnsack, M.T. (2017) Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA biology*, 14, 1138-1152.

[100] Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J. and Nielsen, H. (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew Chem Int Ed Engl*, 54, 451-455.

[101] Lestrade, L. and Weber, M.J. (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res*, 34, D158-162.

[102] Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., Nakayama, H. and Isobe, T. (2016) The complete chemical structure of Saccharomyces cerevisiae rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Res*, 44, 8951-8961.

[103] Decatur, W.A. and Fournier, M.J. (2002) rRNA modifications and ribosome function. *Trends Biochem Sci*, 27, 344-351.

[104] Sprinzl, M. and Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*, 33, D139-140.

[105] Phizicky, E.M. and Alfonzo, J.D. (2010) Do all modifications benefit all tRNAs? *FEBS Lett*, 584, 265-271.

[106] Lagerkvist, U. (1978) "Two out of three": an alternative method for codon reading. *Proceedings of the National Academy of Sciences of the United States of America*, 75, 1759-1762.

[107] Gabant, G., Auxilien, S., Tuszynska, I., Locard, M., Gajda, M.J., Chaussinand, G., Fernandez, B., Dedieu, A., Grosjean, H., Golinelli-Pimpaneau, B. *et al.* (2006) THUMP from archaeal tRNA:m22G10 methyltransferase, a genuine autonomously folding domain. *Nucleic Acids Res*, 34, 2483-2494.

[108] Bjork, G.R., Huang, B., Persson, O.P. and Bystrom, A.S. (2007) A conserved modified wobble nucleoside (mcm5s2U) in lysyl-tRNA is required for viability in yeast. *RNA*, 13, 1245-1255.

[109] Schaffrath, R. and Leidel, S.A. (2017) Wobble uridine modifications-a reason to live, a reason to die?! *RNA biology*, 14, 1209-1222.

[110] Agris, P.F., Vendeix, F.A. and Graham, W.D. (2007) tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol*, 366, 1-13.

[111] Bjork, G.R., Durand, J.M., Hagervall, T.G., Leipuviene, R., Lundgren, H.K., Nilsson, K., Chen, P., Qian, Q. and Urbonavicius, J. (1999) Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Lett*, 452, 47-51.

[112] Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M. and Lyko, F. (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev*, 24, 1590-1595.

[113] Motorin, Y. and Helm, M. (2010) tRNA stabilization by modified nucleotides. *Biochemistry*, 49, 4934-4944.

[114] Helm, M., Giege, R. and Florentz, C. (1999) A Watson-Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNALys. *Biochemistry*, 38, 13338-13346.

[115] Anderson, J., Phan, L., Cuesta, R., Carlson, B.A., Pak, M., Asano, K., Bjork, G.R., Tamame, M. and Hinnebusch, A.G. (1998) The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev*, 12, 3650-3662.

[116] Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G. and Anderson, J. (2004) Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. *Genes Dev*, 18, 1227-1240.

[117] Chernyakov, I., Whipple, J.M., Kotelawala, L., Grayhack, E.J. and Phizicky, E.M. (2008) Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5'-3' exonucleases Rat1 and Xrn1. *Genes Dev*, 22, 1369-1380.

[118] Frye, M., Jaffrey, S.R., Pan, T., Rechavi, G. and Suzuki, T. (2016) RNA modifications: what have we learned and where are we headed? *Nat Rev Genet*, 17, 365-372.

[119] Desrosiers, R., Friderici, K. and Rottman, F. (1974) Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, 71, 3971-3975.

[120] Perry, R.P., Kelley, D.E., Friderici, K. and Rottman, F. (1975) The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell*, 4, 387-394.

[121] Wei, C.M., Gershowitz, A. and Moss, B. (1975) Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell*, 4, 379-386.

[122] Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E. and Jaffrey, S.R. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*, 149, 1635-1646.

[123] Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M. *et al.* (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, 485, 201-206.

[124] Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M. and Gilbert, W.V. (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*, 515, 143-146.

[125] Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., Leon-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S. *et al.* (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell*, 159, 148-162.

[126] Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M.S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W.C. *et al.* (2016) The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature*, 530, 441-446.

[127] Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N. and Rechavi, G. (2013) Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. *Nat Protoc*, 8, 176-189.

[128] Edelheit, S., Schwartz, S., Mumbach, M.R., Wurtzel, O. and Sorek, R. (2013) Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS Genet*, 9, e1003602.

[129] Schaefer, M., Pollex, T., Hanna, K. and Lyko, F. (2009) RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res*, 37, e12.

[130] Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M. and Preiss, T. (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res*, 40, 5023-5033.

[131] Dai, Q., Moshitch-Moshkovitz, S., Han, D., Kol, N., Amariglio, N., Rechavi, G., Dominissini, D. and He, C. (2017) Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat Methods*, 14, 695-698.

[132] Frye, M., Harada, B.T., Behm, M. and He, C. (2018) RNA modifications modulate gene expression during development. *Science*, 361, 1346-1349.

[133] Zaccara, S., Ries, R.J. and Jaffrey, S.R. (2019) Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol*, 20, 608-624.

[134] Adams, J.M. and Cory, S. (1975) Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature*, 255, 28-33.

[135] Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.J., Chen, Q. *et al.* (2017) Reversible methylation of m(6)Am in the 5' cap controls mRNA stability. *Nature*, 541, 371-375.

[136] Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H. *et al.* (2010) 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature*, 468, 452-456.

[137] Freund, I., Eigenbrod, T., Helm, M. and Dalpke, A.H. (2019) RNA Modifications Modulate Activation of Innate Toll-Like Receptors. *Genes (Basel)*, 10.

[138] Keller, P., Freund, I., Marchand, V., Bec, G., Huang, R., Motorin, Y., Eigenbrod, T., Dalpke, A. and Helm, M. (2018) Double methylation of tRNA-U54 to 2'-O-methylthymidine (Tm) synergistically decreases immune response by Toll-like receptor 7. *Nucleic Acids Res*, 46, 9764-9775.

[139] Eigenbrod, T., Keller, P., Kaiser, S., Rimbach, K., Dalpke, A.H. and Helm, M. (2015) Recognition of Specified RNA Modifications by the Innate Immune System. *Methods Enzymol*, 560, 73-89.

[140] Gehrig, S., Eberle, M.E., Botschen, F., Rimbach, K., Eberle, F., Eigenbrod, T., Kaiser, S., Holmes, W.M., Erdmann, V.A., Sprinzl, M. *et al.* (2012) Identification of modifications in microbial, native tRNA that suppress immunostimulatory activity. *J Exp Med*, 209, 225-233.

[141] Jockel, S., Nees, G., Sommer, R., Zhao, Y., Cherkasov, D., Hori, H., Ehm, G., Schnare, M., Nain, M., Kaufmann, A. *et al.* (2012) The 2'-O-methylation status of a single guanosine controls transfer RNA-mediated Toll-like receptor 7 activation or inhibition. *J Exp Med*, 209, 235-241.

[142] Zust, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S. *et al.* (2011) Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol*, 12, 137-143.

[143] Srinivasan, M. and Lahiri, D.K. (2015) Significance of NF-kappaB as a pivotal therapeutic target in the neurodegenerative pathologies of Alzheimer's disease and multiple sclerosis. *Expert Opin Ther Targets*, 19, 471-487.

[144] Buratowski, S. (2009) Progression through the RNA polymerase II CTD cycle. *Mol Cell*, 36, 541-546.

[145] Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S. and Marcotrigiano, J. (2016) Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 596-601.

[146] Langberg, S.R. and Moss, B. (1981) Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from HeLa cells. *The Journal of biological chemistry*, 256, 10054-10060.

[147] Sun, H., Zhang, M., Li, K., Bai, D. and Yi, C. (2019) Cap-specific, terminal N(6)-methylation by a mammalian m(6)Am methyltransferase. *Cell Res*, 29, 80-82.

[148] Sikorski, P.J., Warminski, M., Kubacka, D., Ratajczak, T., Nowis, D., Kowalska, J. and Jemielity, J. (2020) The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5' cap modulates protein expression in living cells. *Nucleic Acids Res*, 48, 1607-1626.

[149] Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E. and Jaffrey, S.R. (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*, 12, 767-772.

[150] Breiling, A. and Lyko, F. (2015) Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin*, 8, 24.

[151] Legrand, C., Tuorto, F., Hartmann, M., Liebers, R., Jacob, D., Helm, M. and Lyko, F. (2017) Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. *Genome Res*, 27, 1589-1596.

[152] Amort, T., Rieder, D., Wille, A., Khokhlova-Cubberley, D., Riml, C., Trixl, L., Jia, X.Y., Micura, R. and Lusser, A. (2017) Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol*, 18, 1.

[153] Khoddami, V., Yerra, A., Mosbruger, T.L., Fleming, A.M., Burrows, C.J. and Cairns, B.R. (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 6784-6789.

[154] Huber, S.M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E.A. and Balasubramanian, S. (2015) Formation and abundance of 5-hydroxymethylcytosine in RNA. *Chembiochem : a European journal of chemical biology*, 16, 752-755.

[155] Grozhik, A.V. and Jaffrey, S.R. (2018) Distinguishing RNA modifications from noise in epitranscriptome maps. *Nat Chem Biol*, 14, 215-225.

[156] Yang, X., Yang, Y., Sun, B.F., Chen, Y.S., Xu, J.W., Lai, W.Y., Li, A., Wang, X., Bhattarai, D.P., Xiao, W. *et al.* (2017) 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res*, 27, 606-625.

[157] Chen, X., Li, A., Sun, B.F., Yang, Y., Han, Y.N., Yuan, X., Chen, R.X., Wei, W.S., Liu, Y., Gao, C.C. *et al.* (2019) 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat Cell Biol*, 21, 978-990.

[158] Yang, Y., Wang, L., Han, X., Yang, W.L., Zhang, M., Ma, H.L., Sun, B.F., Li, A., Xia, J., Chen, J. *et al.* (2019) RNA 5-Methylcytosine Facilitates the Maternal-to-Zygotic Transition by Preventing Maternal mRNA Decay. *Mol Cell*, 75, 1188-1202 e1111.

[159] Grozhik, A.V., Olarerin-George, A.O., Sindelar, M., Li, X., Gross, S.S. and Jaffrey, S.R. (2019) Antibody cross-reactivity accounts for widespread appearance of m(1)A in 5'UTRs. *Nat Commun*, 10, 5126.

[160] Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S. and Yi, C. (2016) Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat Chem Biol*, 12, 311-316.

[161] Hauenschild, R., Tserovski, L., Schmid, K., Thuring, K., Winz, M.L., Sharma, S., Entian, K.D., Wacheul, L., Lafontaine, D.L., Anderson, J. *et al.* (2015) The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. *Nucleic Acids Res*, 43, 9950-9964.

[162] Koller, A.N., Bozilovic, J., Engels, J.W. and Gohlke, H. (2010) Aromatic N versus aromatic F: bioisosterism discovered in RNA base pairing interactions leads to a novel class of universal base analogs. *Nucleic Acids Res*, 38, 3133-3146.

[163] Wei, C.M., Gershowitz, A. and Moss, B. (1976) 5'-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. *Biochemistry*, 15, 397-401.

[164] Schwartz, S., Mumbach, M.R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G.G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D. *et al.* (2014) Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep*, 8, 284-296.

[165] Wei, C.M. and Moss, B. (1977) Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. *Biochemistry*, 16, 1672-1676.

[166] Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G. and Rottman, F.M. (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA*, 3, 1233-1247.

[167] Sledz, P. and Jinek, M. (2016) Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife*, 5.

[168] Wang, P., Doxtader, K.A. and Nam, Y. (2016) Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol Cell*, 63, 306-317.

[169] Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C. *et al.* (2016) Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature*, 534, 575-578.

[170] Slobodin, B., Han, R., Calderone, V., Vrielink, J., Loayza-Puch, F., Elkon, R. and Agami, R. (2017) Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell*, 169, 326-337 e312.

[171] Horiuchi, K., Kawamura, T., Iwanari, H., Ohashi, R., Naito, M., Kodama, T. and Hamakubo, T. (2013) Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *The Journal of biological chemistry*, 288, 33292-33302.

[172] Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S. *et al.* (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res*, 24, 177-189.

[173] Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M. and Fray, R.G. (2008) MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell*, 20, 1278-1288.

[174] Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H. *et al.* (2018) VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov*, 4, 10.

[175] Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S. *et al.* (2015) Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science*, 347, 1002-1006.

[176] Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K. *et al.* (2014) m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell*, 15, 707-719.

[177] Patil, D.P., Chen, C.K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M. and Jaffrey, S.R. (2016) m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature*, 537, 369-373.

[178] Raffel, G.D., Mercher, T., Shigematsu, H., Williams, I.R., Cullen, D.E., Akashi, K., Bernard, O.A. and Gilliland, D.G. (2007) Ott1(Rbm15) has pleiotropic roles in hematopoietic development. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 6001-6006.

[179] Niu, C., Zhang, J., Breslin, P., Onciu, M., Ma, Z. and Morris, S.W. (2009) c-Myc is a target of RNA-binding motif protein 15 in the regulation of adult hematopoietic stem cell and megakaryocyte development. *Blood*, 114, 2087-2096.

[180] Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H. *et al.* (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell*, 49, 18-29.

[181] Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G. *et al.* (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol*, 7, 885-887.

[182] Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J. *et al.* (2014) FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res*, 24, 1403-1419.

[183] Garcia-Campos, M.A., Edelheit, S., Toth, U., Safra, M., Shachar, R., Viukov, S., Winkler, R., Nir, R., Lasman, L., Brandis, A. *et al.* (2019) Deciphering the "m(6)A Code" via Antibody-Independent Quantitative Profiling. *Cell*, 178, 731-747 e716.

[184] Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G. *et al.* (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, 505, 117-120.

[185] Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H.Y., Li, A., Ping, X.L., Lai, W.Y. *et al.* (2016) Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell*, 61, 507-519.

[186] Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R. and Qian, S.B. (2015) Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature*, 526, 591-594.

[187] Fu, Y., Dominissini, D., Rechavi, G. and He, C. (2014) Gene expression regulation mediated through reversible m(6)A RNA methylation. *Nat Rev Genet*, 15, 293-306.

[188] Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M. *et al.* (2016) m(6)A modulates neuronal functions and sex determination in Drosophila. *Nature*, 540, 242-247.

[189] Carroll, S.M., Narayan, P. and Rottman, F.M. (1990) N6-methyladenosine residues in an intron-specific region of prolactin pre-mRNA. *Mol Cell Biol*, 10, 4456-4465.

[190] Roundtree, I.A., Luo, G.Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P. *et al.* (2017) YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife*, 6.

[191] Ji, P., Wang, X., Xie, N. and Li, Y. (2018) N6-Methyladenosine in RNA and DNA: An Epitranscriptomic and Epigenetic Player Implicated in Determination of Stem Cell Fate. *Stem Cells Int*, 2018, 3256524.

[192] Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M. and Pan, T. (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*, 518, 560-564.

[193] Kierzek, E. and Kierzek, R. (2003) The synthesis of oligoribonucleotides containing N6-alkyladenosines and 2-methylthio-N6-alkyladenosines via post-synthetic modification of precursor oligomers. *Nucleic Acids Res*, 31, 4461-4471.

[194] Batista, P.J. (2017) The RNA Modification N(6)-methyladenosine and Its Implications in Human Disease. *Genomics Proteomics Bioinformatics*, 15, 154-163.

[195] Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B. and Jaffrey, S.R. (2015) 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell*, 163, 999-1010.

[196] Coots, R.A., Liu, X.M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X. and Qian, S.B. (2017) m(6)A Facilitates eIF4F-Independent mRNA Translation. *Mol Cell*, 68, 504-514 e507.

[197] Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H. and He, C. (2015) N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*, 161, 1388-1399.

[198] Zhou, J., Wan, J., Shu, X.E., Mao, Y., Liu, X.M., Yuan, X., Zhang, X., Hess, M.E., Bruning, J.C. and Qian, S.B. (2018) N(6)-Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. *Mol Cell*, 69, 636-647 e637.

[199] Hoernes, T.P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., Lindner, H., Huttenhofer, A. and Erlacher, M.D. (2016) Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Res*, 44, 852-862.

[200] Choi, J., Ieong, K.W., Demirci, H., Chen, J., Petrov, A., Prabhakar, A., O'Leary, S.E., Dominissini, D., Rechavi, G., Soltis, S.M. *et al.* (2016) N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat Struct Mol Biol*, 23, 110-115.

[201] Chen, X.Y., Zhang, J. and Zhu, J.S. (2019) The role of m(6)A RNA methylation in human cancer. *Mol Cancer*, 18, 103.

[202] Han, M., Liu, Z., Xu, Y., Liu, X., Wang, D., Li, F., Wang, Y. and Bi, J. (2020) Abnormality of m6A mRNA Methylation Is Involved in Alzheimer's Disease. *Front Neurosci*, 14, 98.

[203] Paramasivam, A., Priyadharsini, J.V. and Raghunandhakumar, S. (2020) Implications of m6A modification in autoimmune disorders. *Cell Mol Immunol*, 17, 550-551.

[204] Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G.M., Bansal, V., Wang, Y., Mason, C.E. and Rana, T.M. (2016) Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol*, 1, 16011.

[205] Paramasivam, A., Vijayashree Priyadharsini, J. and Raghunandhakumar, S. (2020) N6-adenosine methylation (m6A): a promising new molecular target in hypertension and cardiovascular diseases. *Hypertens Res*, 43, 153-154.

[206] Yu, C.T. and Allen, F.W. (1959) Studies on an isomer of uridine isolated from ribonucleic acids. *Biochim Biophys Acta*, 32, 393-406.

[207] Cohn, W.E. (1959) 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids. *Biochim Biophys Acta*, 32, 569-571.

[208] Cohn, W.E. and Volkin, E. (1951) Nucleoside-5'-Phosphates from Ribonucleic Acid. *Nature*, 167, 483-484.

[209] Davis, F.F. and Allen, F.W. (1957) Ribonucleic acids from yeast which contain a fifth nucleotide. *The Journal of biological chemistry*, 227, 907-915.

[210] Li, X., Ma, S. and Yi, C. (2016) Pseudouridine: the fifth RNA nucleotide with renewed interests. *Curr Opin Chem Biol*, 33, 108-116.

[211] Tardu, M., Jones, J.D., Kennedy, R.T., Lin, Q. and Koutmou, K.S. (2019) Identification and Quantification of Modified Nucleosides in Saccharomyces cerevisiae mRNAs. *ACS Chem Biol*, 14, 1403-1409.

[212] Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F. and Yi, C. (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat Chem Biol*, 11, 592-597.

[213] Hudson, G.A., Bloomingdale, R.J. and Znosko, B.M. (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *RNA*, 19, 1474-1482.

[214] Yu, Y.T. and Meier, U.T. (2014) RNA-guided isomerization of uridine to pseudouridine--pseudouridylation. *RNA biology*, 11, 1483-1494.

[215] Lovejoy, A.F., Riordan, D.P. and Brown, P.O. (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. *PLoS One*, 9, e110799.

[216] Charette, M. and Gray, M.W. (2000) Pseudouridine in RNA: what, where, how, and why. *IUBMB Life*, 49, 341-351.

[217] Gilbert, W.V., Bell, T.A. and Schaening, C. (2016) Messenger RNA modifications: Form, distribution, and function. *Science*, 352, 1408-1412.

[218] Karijolich, J. and Yu, Y.T. (2011) Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature*, 474, 395-398.

[219] Ge, J. and Yu, Y.T. (2013) RNA pseudouridylation: new insights into an old modification. *Trends Biochem Sci*, 38, 210-218.

[220] Zhou, K.I., Clark, W.C., Pan, D.W., Eckwahl, M.J., Dai, Q. and Pan, T. (2018) Pseudouridines have context-dependent mutation and stop rates in high-throughput sequencing. *RNA biology*, 15, 892-900.

[221] Fernandez, I.S., Ng, C.L., Kelley, A.C., Wu, G., Yu, Y.T. and Ramakrishnan, V. (2013) Unusual base pairing during the decoding of a stop codon by the ribosome. *Nature*, 500, 107-110.

[222] Chen, C., Zhao, X., Kierzek, R. and Yu, Y.T. (2010) A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing in vivo. *Mol Cell Biol*, 30, 4108-4119.

[223] Penzo, M., Guerrieri, A.N., Zacchini, F., Trere, D. and Montanaro, L. (2017) RNA Pseudouridylation in Physiology and Medicine: For Better and for Worse. *Genes (Basel)*, 8.

[224] Rintala-Dempsey, A.C. and Kothe, U. (2017) Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression? *RNA biology*, 14, 1185-1196.

[225] Eyler, D.E., Franco, M.K., Batool, Z., Wu, M.Z., Dubuke, M.L., Dobosz-Bartoszek, M., Jones, J.D., Polikanov, Y.S., Roy, B. and Koutmou, K.S. (2019) Pseudouridinylation of mRNA coding sequences alters translation. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 23068-23074.

[226] Hoernes, T.P., Heimdorfer, D., Kostner, D., Faserl, K., Nussbaumer, F., Plangger, R., Kreutz, C., Lindner, H. and Erlacher, M.D. (2019) Eukaryotic Translation Elongation is Modulated by Single Natural Nucleotide Derivatives in the Coding Sequences of mRNAs. *Genes (Basel)*, 10.

[227] Anderson, B.R., Muramatsu, H., Nallagatla, S.R., Bevilacqua, P.C., Sansing, L.H., Weissman, D. and Kariko, K. (2010) Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Res*, 38, 5884-5892.

[228] Kariko, K., Buckstein, M., Ni, H. and Weissman, D. (2005) Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity*, 23, 165-175.

[229] Kariko, K. and Weissman, D. (2007) Naturally occurring nucleoside modifications suppress the immunostimulatory activity of RNA: implication for therapeutic RNA development. *Curr Opin Drug Discov Devel*, 10, 523-532.

[230] Kariko, K., Muramatsu, H., Welsh, F.A., Ludwig, J., Kato, H., Akira, S. and Weissman, D. (2008) Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol Ther*, 16, 1833-1840.

[231] Elliott, B.A., Ho, H.T., Ranganathan, S.V., Vangaveti, S., Ilkayeva, O., Abou Assi, H., Choi, A.K., Agris, P.F. and Holley, C.L. (2019) Modification of messenger RNA by 2'-O-methylation regulates gene expression in vivo. *Nat Commun*, 10, 3401.

[232] Shubina, M.Y., Musinova, Y.R. and Sheval, E.V. (2016) Nucleolar Methyltransferase Fibrillarin: Evolution of Structure and Functions. *Biochemistry (Mosc)*, 81, 941-950.

[233] Somme, J., Van Laer, B., Roovers, M., Steyaert, J., Versees, W. and Droogmans, L. (2014) Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates. *RNA*, 20, 1257-1271.

[234] Hyde, J.L. and Diamond, M.S. (2015) Innate immune restriction and antagonism of viral RNA lacking 2-O methylation. *Virology*, 479-480, 66-74.

[235] Ringeard, M., Marchand, V., Decroly, E., Motorin, Y. and Bennasser, Y. (2019) FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature*, 565, 500-504.

[236] Choi, J., Indrisiunaite, G., DeMirci, H., Ieong, K.W., Wang, J., Petrov, A., Prabhakar, A., Rechavi, G., Dominissini, D., He, C. *et al.* (2018) 2′ -O-methylation in mRNA disrupts tRNA decoding during translation elongationn. *Nat Struct Mol Biol*, 25, 208-216.

[237] Consden, R., Gordon, A.H. and Martin, A.J. (1944) Qualitative analysis of proteins: a partition chromatographic method using paper. *Biochem J*, 38, 224-232.

[238] Keith, G. (1995) Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie*, 77, 142-144.

[239] Markham, R. and Smith, J.D. (1949) Chromatographic studies of nucleic acids; a technique for the identification and estimation of purine and pyrimidine bases, nucleosides and related substances. *Biochem J*, 45, 294-298.

[240] Zeevi, M. and Daniel, V. (1976) Aminoacylation and nucleoside modification of in vitro synthesised transfer RNA. *Nature*, 260, 72-74.

[241] Rottman, F.M., Desrosiers, R.C. and Friderici, K. (1976) Nucleotide methylation patterns in eukaryotic mRNA. *Prog Nucleic Acid Res Mol Biol*, 19, 21-38.

[242] Horowitz, S., Horowitz, A., Nilsen, T.W., Munns, T.W. and Rottman, F.M. (1984) Mapping of N6-methyladenosine residues in bovine prolactin mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 81, 5667-5671.

[243] Szekely, M. and Sanger, F. (1969) Use of polynucleotide kinase in fingerprinting non-radioactive nucleic acids. *J Mol Biol*, 43, 607-617.

[244] Maden, B.E. (1986) Identification of the locations of the methyl groups in 18 S ribosomal RNA from Xenopus laevis and man. *J Mol Biol*, 189, 681-699.

[245] Grosjean, H., Keith, G. and Droogmans, L. (2004) Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. *Methods Mol Biol*, 265, 357-391.

[246] Stanley, J. and Vassilenko, S. (1978) A different approach to RNA sequencing. *Nature*, 274, 87-89.

[247] Yu, Y.T., Shu, M.D. and Steitz, J.A. (1997) A new method for detecting sites of 2'-O-methylation in RNA molecules. *RNA*, 3, 324-331.

[248] Pyle, A.M., Chu, V.T., Jankowsky, E. and Boudvillain, M. (2000) Using DNAzymes to cut, process, and map RNA molecules for structural studies or modification. *Methods Enzymol*, 317, 140-146.

[249] Hengesbach, M., Meusburger, M., Lyko, F. and Helm, M. (2008) Use of DNAzymes for site-specific analysis of ribonucleotide modifications. *RNA*, 14, 180-187.

[250] Liu, N. and Pan, T. (2015) Probing RNA Modification Status at Single-Nucleotide Resolution in Total RNA. *Methods Enzymol*, 560, 149-159.

[251] Bodi, Z., Button, J.D., Grierson, D. and Fray, R.G. (2010) Yeast targets for mRNA methylation. *Nucleic Acids Res*, 38, 5327-5335.

[252] Ruzicka, K., Zhang, M., Campilho, A., Bodi, Z., Kashif, M., Saleh, M., Eeckhout, D., El-Showk, S., Li, H., Zhong, S. *et al.* (2017) Identification of factors required for m(6) A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol*, 215, 157-172.

[253] Gehrke, C.W., Kuo, K.C. and Zumwalt, R.W. (1980) Chromatography of nucleosides. *J Chromatogr*, 188, 129-147.

[254] Edmonds, C.G., Vestal, M.L. and McCloskey, J.A. (1985) Thermospray liquid chromatography-mass spectrometry of nucleosides and of enzymatic hydrolysates of nucleic acids. *Nucleic Acids Res*, 13, 8197-8206.

[255] Sakaguchi, Y., Miyauchi, K., Kang, B.I. and Suzuki, T. (2015) Nucleoside Analysis by Hydrophilic Interaction Liquid Chromatography Coupled with Mass Spectrometry. *Methods Enzymol*, 560, 19-28.

[256] Kowalak, J.A., Pomerantz, S.C., Crain, P.F. and McCloskey, J.A. (1993) A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. *Nucleic Acids Res*, 21, 4577-4585.

[257] Hossain, M. and Limbach, P.A. (2009) Multiple endonucleases improve MALDI-MS signature digestion product detection of bacterial transfer RNAs. *Anal Bioanal Chem*, 394, 1125-1135.

[258] Ross, R.L., Cao, X. and Limbach, P.A. (2017) Mapping Post-Transcriptional Modifications onto Transfer Ribonucleic Acid Sequences by Liquid Chromatography Tandem Mass Spectrometry. *Biomolecules*, 7.

[259] Wetzel, C. and Limbach, P.A. (2016) Mass spectrometry of modified RNAs: recent developments. *Analyst*, 141, 16-23.

[260] Zhang, N., Shi, S., Jia, T.Z., Ziegler, A., Yoo, B., Yuan, X., Li, W. and Zhang, S. (2019) A general LC-MS-based RNA sequencing method for direct analysis of multiple-base modifications in RNA mixtures. *Nucleic Acids Res*, 47, e125.

[261] Wein, S., Andrews, B., Sachsenberg, T., Santos-Rosa, H., Kohlbacher, O., Kouzarides, T., Garcia, B.A. and Weisser, H. (2020) A computational platform for high-throughput analysis of RNA sequences and modifications by mass spectrometry. *Nat Commun*, 11, 926.

[262] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74, 5463-5467.

[263] Maxam, A.M. and Gilbert, W. (1992) A new method for sequencing DNA. 1977. *Biotechnology*, 24, 99-103.

[264] Hamlyn, P.H., Browniee, G.G., Cheng, C.C., Gait, M.J. and Milstein, C. (1978) Complete sequence of constant and 3' noncoding regions of an immunoglobulin mRNA using the dideoxynucleotide method of RNA sequencing. *Cell*, 15, 1067-1075.

[265] Peattie, D.A. (1979) Direct chemical method for sequencing RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76, 1760-1764.

[266] Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G. *et al.* (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 327, 78-81.

[267] Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J.A., Costa, G., McKernan, K. *et al.* (2008) A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome Res*, 18, 1051-1063.

[268] Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437, 376-380.

[269] Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R. *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456, 53-59.

[270] Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B. *et al.* (2009) Real-time DNA sequencing from single polymerase molecules. *Science*, 323, 133-138.

[271] Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X. *et al.* (2008) The potential and challenges of nanopore sequencing. *Nat Biotechnol*, 26, 1146-1153.

[272] Clarke, J., Wu, H.C., Jayasinghe, L., Patel, A., Reid, S. and Bayley, H. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol*, 4, 265-270.

[273] Liu, H., Begik, O., Lucas, M.C., Ramirez, J.M., Mason, C.E., Wiener, D., Schwartz, S., Mattick, J.S., Smith, M.A. and Novoa, E.M. (2019) Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nat Commun*, 10, 4079.

[274] Motorin, Y., Muller, S., Behm-Ansmant, I. and Branlant, C. (2007) Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol*, 425, 21-53.

[275] Dunin-Horkawicz, S., Czerwoniec, A., Gajda, M.J., Feder, M., Grosjean, H. and Bujnicki, J.M. (2006) MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res*, 34, D145-149.

[276] Ebhardt, H.A., Tsang, H.H., Dai, D.C., Liu, Y., Bostan, B. and Fahlman, R.P. (2009) Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res*, 37, 2461-2470.

[277] Iida, K., Jin, H. and Zhu, J.K. (2009) Bioinformatics analysis suggests base modifications of tRNAs and miRNAs in Arabidopsis thaliana. *BMC genomics*, 10, 155.

[278] Marchand, V., Blanloeil-Oillo, F., Helm, M. and Motorin, Y. (2016) Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res*, 44, e135.

[279] Erales, J., Marchand, V., Panthu, B., Gillot, S., Belin, S., Ghayad, S.E., Garcia, M., Laforets, F., Marcel, V., Baudin-Baillieu, A. *et al.* (2017) Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 12934-12939.

[280] Diebold, S.S., Kaisho, T., Hemmi, H., Akira, S. and Reis e Sousa, C. (2004) Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science*, 303, 1529-1531.

[281] Heutinck, K.M., Kassies, J., Florquin, S., ten Berge, I.J., Hamann, J. and Rowshani, A.T. (2012) SerpinB9 expression in human renal tubular epithelial cells is induced by triggering of the viral dsRNA sensors TLR3, MDA5 and RIG-I. *Nephrol Dial Transplant*, 27, 2746-2754.

[282] Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H. and Bauer, S. (2004) Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science*, 303, 1526-1529.

[283] Ehrenfeld, E. and Hunt, T. (1971) Double-stranded poliovirus RNA inhibits initiation of protein synthesis by reticulocyte lysates. *Proceedings of the National Academy of Sciences of the United States of America*, 68, 1075-1078.

[284] Balachandran, S., Roberts, P.C., Brown, L.E., Truong, H., Pattnaik, A.K., Archer, D.R. and Barber, G.N. (2000) Essential role for the dsRNA-dependent protein kinase PKR in innate immunity to viral infection. *Immunity*, 13, 129-141.

[285] Svitkin, Y.V., Cheng, Y.M., Chakraborty, T., Presnyak, V., John, M. and Sonenberg, N. (2017) N1-methyl-pseudouridine in mRNA enhances translation through eIF2alpha-dependent and independent mechanisms by increasing ribosome density. *Nucleic Acids Res*, 45, 6023-6036.

[286] Nallagatla, S.R., Hwang, J., Toroney, R., Zheng, X., Cameron, C.E. and Bevilacqua, P.C. (2007) 5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops. *Science*, 318, 1455-1458.

[287] Kariko, K., Muramatsu, H., Keller, J.M. and Weissman, D. (2012) Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. *Mol Ther*, 20, 948-953.

[288] Yoshioka, N., Gros, E., Li, H.R., Kumar, S., Deacon, D.C., Maron, C., Muotri, A.R., Chi, N.C., Fu, X.D., Yu, B.D. *et al.* (2013) Efficient generation of human iPSCs by a synthetic self-replicative RNA. *Cell Stem Cell*, 13, 246-254.

[289] Kormann, M.S., Hasenpusch, G., Aneja, M.K., Nica, G., Flemmer, A.W., Herber-Jonat, S., Huppmann, M., Mays, L.E., Illenyi, M., Schams, A. *et al.* (2011) Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat Biotechnol*, 29, 154-157.

[290] Licht, K., Hartl, M., Amman, F., Anrather, D., Janisiw, M.P. and Jantsch, M.F. (2019) Inosine induces context-dependent recoding and translational stalling. *Nucleic Acids Res*, 47, 3-14.

[291] Padilla, R. and Sousa, R. (2002) A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. *Nucleic Acids Res*, 30, e138.

[292] McGregor, A., Rao, M.V., Duckworth, G., Stockley, P.G. and Connolly, B.A. (1996) Preparation of oligoribonucleotides containing 4-thiouridine using Fpmp chemistry. Photo-crosslinking to RNA binding proteins using 350 nm irradiation. *Nucleic Acids Res*, 24, 3173-3180.

[293] Pitsch, S. and Weiss, P.A. (2002) Chemical synthesis of RNA sequences with 2'-O-[(triisopropylsilyl)oxy]methyl-protected ribonucleoside phosphoramidites. *Curr Protoc Nucleic Acid Chem*, Chapter 3, Unit 3 8.

[294] Stark, M.R., Pleiss, J.A., Deras, M., Scaringe, S.A. and Rader, S.D. (2006) An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs. *RNA*, 12, 2014-2019.

[295] Nandakumar, J., Ho, C.K., Lima, C.D. and Shuman, S. (2004) RNA substrate specificity and structure-guided mutational analysis of bacteriophage T4 RNA ligase 2. *The Journal of biological chemistry*, 279, 31337-31347.

[296] England, T.E., Bruce, A.G. and Uhlenbeck, O.C. (1980) Specific labeling of 3' termini of RNA with T4 RNA ligase. *Methods Enzymol*, 65, 65-74.

[297] Wyatt, J.R., Sontheimer, E.J. and Steitz, J.A. (1992) Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev*, 6, 2542-2553.

[298] Moore, M.J. and Sharp, P.A. (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science*, 256, 992-997.

[299] Kurschat, W.C., Muller, J., Wombacher, R. and Helm, M. (2005) Optimizing splinted ligation of highly structured small RNAs. *RNA*, 11, 1909-1914.

[300] Uhlenbeck, O.C. (1980) T4 Rna Ligase. *Abstr Pap Am Chem S*, 180, 159-Biol.

[301] Silber, R., Malathi, V.G. and Hurwitz, J. (1972) Purification and properties of bacteriophage T4-induced RNA ligase. *Proceedings of the National Academy of Sciences of the United States of America*, 69, 3009-3013.

[302] Cranston, J.W., Silber, R., Malathi, V.G. and Hurwitz, J. (1974) Studies on ribonucleic acid ligase. Characterization of an adenosine triphosphate-inorganic pyrophosphate exchange reaction and demonstration of an enzyme-adenylate complex with T4 bacteriophage-induced enzyme. *The Journal of biological chemistry*, 249, 7447-7456.

[303] Sugino, A., Snoper, T.J. and Cozzarelli, N.R. (1977) Bacteriophage T4 RNA ligase. Reaction intermediates and interaction of substrates. *The Journal of biological chemistry*, 252, 1732-1738.

[304] Silverman, S.K. (2004) Practical and general synthesis of 5'-adenylated RNA (5'-AppRNA). *RNA*, 10, 731-746.

[305] Hoernes, T.P., Faserl, K., Juen, M.A., Kremser, J., Gasser, C., Fuchs, E., Shi, X., Siewert, A., Lindner, H., Kreutz, C. *et al.* (2018) Translation of non-standard codon nucleotides reveals minimal requirements for codon-anticodon interactions. *Nat Commun*, 9, 4865.

[306] Stephens, O.M., Yi-Brunozzi, H.Y. and Beal, P.A. (2000) Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry*, 39, 12243-12251.

[307] Hoernes, T.P., Clementi, N., Juen, M.A., Shi, X., Faserl, K., Willi, J., Gasser, C., Kreutz, C., Joseph, S., Lindner, H. *et al.* (2018) Atomic mutagenesis of stop codon nucleotides reveals the chemical prerequisites for release factor-mediated peptide release. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E382-E389.

[308] Zhang, D., Zhou, C.Y., Busby, K.N., Alexander, S.C. and Devaraj, N.K. (2018) Light-Activated Control of Translation by Enzymatic Covalent mRNA Labeling. *Angew Chem Int Ed Engl*, 57, 2822-2826.

[309] Anderson, J.M., Charbonneau, H. and Cormier, M.J. (1974) Mechanism of calcium induction of Renilla bioluminescence. Involvement of a calcium-triggered luciferin binding protein. *Biochemistry*, 13, 1195-1200.

[310] Marques, S.M. and Esteves da Silva, J.C. (2009) Firefly bioluminescence: a mechanistic approach of luciferase catalyzed reactions. *IUBMB Life*, 61, 6-17.

[311] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. (1994) Green fluorescent protein as a marker for gene expression. *Science*, 263, 802-805.

[312] Nienhaus, K. and Nienhaus, G.U. (2021) Fluorescent proteins of the EosFP clade: intriguing marker tools with multiple photoactivation modes for advanced microscopy. *RSC Chem Biol*, 2, 796-814.

[313] Reid, B.G. and Flynn, G.C. (1997) Chromophore formation in green fluorescent protein. *Biochemistry*, 36, 6786-6791.

[314] Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A. and Tsien, R.Y. (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci*, 20, 448-455.

[315] Inouye, S. and Tsuji, F.I. (1994) Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett*, 341, 277-280.

[316] Bokman, S.H. and Ward, W.W. (1981) Renaturation of Aequorea gree-fluorescent protein. *Biochem Biophys Res Commun*, 101, 1372-1380.

[317] Sorenson, A.E., Askin, S.P. and Schaeffer, P.M. (2015) In-gel detection of biotin–protein conjugates with a green fluorescent streptavidin probe. *Analytical Methods*, 7, 2087-2092.

[318] Aoki, T., Takahashi, Y., Koch, K.S., Leffert, H.L. and Watabe, H. (1996) Construction of a fusion protein between protein A and green fluorescent protein and its application to western blotting. *FEBS Lett*, 384, 193-197.

[319] Geertsma, E.R., Groeneveld, M., Slotboom, D.J. and Poolman, B. (2008) Quality control of overexpressed membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 5722-5727.

[320] Cabantous, S., Terwilliger, T.C. and Waldo, G.S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol*, 23, 102-107.

[321] Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. and Ward, W.W. (1993) Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry*, 32, 1212-1218.

[322] Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. and Cormier, M.J. (1992) Primary structure of the Aequorea victoria green-fluorescent protein. *Gene*, 111, 229-233.

[323] Shimomura, O., Johnson, F.H. and Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol*, 59, 223-239.

[324] Heim, R., Cubitt, A.B. and Tsien, R.Y. (1995) Improved green fluorescence. *Nature*, 373, 663-664.

[325] Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173, 33-38.

[326] Tsien, R.Y. (1998) The green fluorescent protein. *Annu Rev Biochem*, 67, 509-544.

[327] Hernandez, J.M., Morales, S.E. and Lopez, M.A. (1977) [Effect of pimozide on the prolactin and somatotropin responses to L-dopa, and on the somatotropin response to hypoglycemia (author's transl)]. *Rev Invest Clin*, 29, 201-207.

[328] Vasanthi, R.L., Arulvasu, C., Kumar, P. and Srinivasan, P. (2021) Ingestion of microplastics and its potential for causing structural alterations and oxidative stress in Indian green mussel Perna viridis- A multiple biomarker approach. *Chemosphere*, 283, 130979.

[329] Hasson, S.S.A.A., Al-Busaidi, J.K.Z. and Sallam, T.A. (2015) The past, current and future trends in DNA vaccine immunisations. *Asian Pacific Journal of Tropical Biomedicine*, 5, 344-353.

[330] Wolff, J.A., Ludtke, J.J., Acsadi, G., Williams, P. and Jani, A. (1992) Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. *Hum Mol Genet*, 1, 363-369.

[331] Conry, R.M., LoBuglio, A.F., Wright, M., Sumerel, L., Pike, M.J., Johanning, F., Benjamin, R., Lu, D. and Curiel, D.T. (1995) Characterization of a messenger RNA polynucleotide vaccine vector. *Cancer Res*, 55, 1397-1400.

[332] Kranz, L.M., Diken, M., Haas, H., Kreiter, S., Loquai, C., Reuter, K.C., Meng, M., Fritz, D., Vascotto, F., Hefesha, H. *et al.* (2016) Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. *Nature*, 534, 396-401.

[333] Reichmuth, A.M., Oberli, M.A., Jaklenec, A., Langer, R. and Blankschtein, D. (2016) mRNA vaccine delivery using lipid nanoparticles. *Ther Deliv*, 7, 319-334.

[334] Martinon, F., Krishnan, S., Lenzen, G., Magne, R., Gomard, E., Guillet, J.G., Levy, J.P. and Meulien, P. (1993) Induction of virus-specific cytotoxic T lymphocytes in vivo by liposome-entrapped mRNA. *Eur J Immunol*, 23, 1719-1722.

[335] Van Nuffel, A.M., Benteyn, D., Wilgenhof, S., Pierret, L., Corthals, J., Heirman, C., van der Bruggen, P., Coulie, P.G., Neyns, B., Thielemans, K. *et al.* (2012) Dendritic cells loaded with mRNA encoding full-length tumor antigens prime CD4+ and CD8+ T cells in melanoma patients. *Mol Ther*, 20, 1063-1074.

[336] Kyte, J.A., Mu, L., Aamdal, S., Kvalheim, G., Dueland, S., Hauser, M., Gullestad, H.P., Ryder, T., Lislerud, K., Hammerstad, H. *et al.* (2006) Phase I/II trial of melanoma therapy with dendritic cells transfected with autologous tumor-mRNA. *Cancer Gene Ther*, 13, 905-918.

[337] Su, Z., Dannull, J., Yang, B.K., Dahm, P., Coleman, D., Yancey, D., Sichi, S., Niedzwiecki, D., Boczkowski, D., Gilboa, E. *et al.* (2005) Telomerase mRNA-transfected dendritic cells stimulate antigen-specific CD8+ and CD4+ T cell responses in patients with metastatic prostate cancer. *J Immunol*, 174, 3798-3807.

[338] Pardi, N., Secreto, A.J., Shan, X., Debonera, F., Glover, J., Yi, Y., Muramatsu, H., Ni, H., Mui, B.L., Tam, Y.K. *et al.* (2017) Administration of nucleoside-modified mRNA encoding broadly neutralizing antibody protects humanized mice from HIV-1 challenge. *Nat Commun*, 8, 14630.

[339] Parr, C.J.C., Wada, S., Kotake, K., Kameda, S., Matsuura, S., Sakashita, S., Park, S., Sugiyama, H., Kuang, Y. and Saito, H. (2020) N 1-Methylpseudouridine substitution enhances the performance of synthetic mRNA switches in cells. *Nucleic Acids Res*, 48, e35.

[340] Andries, O., Mc Cafferty, S., De Smedt, S.C., Weiss, R., Sanders, N.N. and Kitada, T. (2015) N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J Control Release*, 217, 337-344.

[341] Pardi, N., LaBranche, C.C., Ferrari, G., Cain, D.W., Tombacz, I., Parks, R.J., Muramatsu, H., Mui, B.L., Tam, Y.K., Kariko, K. *et al.* (2019) Characterization of HIV-1 Nucleoside-Modified mRNA Vaccines in Rabbits and Rhesus Macaques. *Mol Ther Nucleic Acids*, 15, 36-47.

[342] Huijgens, P.C., Ossenkoppele, G.J., van der Lelie, J., Thomas, L.L., Wijngaarden, M.J. and Reijneke, R.M. (1988) Ifosfamide and VP-16213 combination chemotherapy combined with ablative chemotherapy and autologous marrow transplantation as salvage treatment for malignant lymphoma. *Eur J Cancer Clin Oncol*, 24, 483-486.

[343] Bourquin, C., Schmidt, L., Hornung, V., Wurzenberger, C., Anz, D., Sandholzer, N., Schreiber, S., Voelkl, A., Hartmann, G. and Endres, S. (2007) Immunostimulatory RNA oligonucleotides trigger an antigen-specific cytotoxic T-cell and IgG2a response. *Blood*, 109, 2953-2960.

[344] Schlee, M., Roth, A., Hornung, V., Hagmann, C.A., Wimmenauer, V., Barchet, W., Coch, C., Janke, M., Mihailovic, A., Wardle, G. *et al.* (2009) Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity*, 31, 25-34.

[345] Kallen, K.J., Heidenreich, R., Schnee, M., Petsch, B., Schlake, T., Thess, A., Baumhof, P., Scheel, B., Koch, S.D. and Fotin-Mleczek, M. (2013) A novel, disruptive vaccination technology: self-adjuvanted RNActive((R)) vaccines. *Hum Vaccin Immunother*, 9, 2263-2276.

[346] Freyn, A.W., Ramos da Silva, J., Rosado, V.C., Bliss, C.M., Pine, M., Mui, B.L., Tam, Y.K., Madden, T.D., de Souza Ferreira, L.C., Weissman, D. *et al.* (2020) A Multi-Targeting, Nucleoside-Modified mRNA Influenza Virus Vaccine Provides Broad Protection in Mice. *Mol Ther*.

[347] Feldman, R.A., Fuhr, R., Smolenov, I., Mick Ribeiro, A., Panther, L., Watson, M., Senn, J.J., Smith, M., Almarsson, Pujar, H.S. *et al.* (2019) mRNA vaccines against H10N8 and H7N9 influenza viruses of pandemic potential are immunogenic and well tolerated in healthy adults in phase 1 randomized clinical trials. *Vaccine*, 37, 3326-3334.

[348] Stadler, C.R., Bahr-Mahmud, H., Celik, L., Hebich, B., Roth, A.S., Roth, R.P., Kariko, K., Tureci, O. and Sahin, U. (2017) Elimination of large tumors in mice by mRNA-encoded bispecific antibodies. *Nat Med*, 23, 815-817.

[349] Okumura, K., Nakase, M., Inui, M., Nakamura, S., Watanabe, Y. and Tagawa, T. (2008) Bax mRNA therapy using cationic liposomes for human malignant melanoma. *J Gene Med*, 10, 910-917.

[350] Machovich, R. (1985) Blood coagulation-fibrinolytic system and endothelial cells. *Acta Biochim Biophys Acad Sci Hung*, 20, 135-153.

[351] Balmayor, E.R., Geiger, J.P., Aneja, M.K., Berezhanskyy, T., Utzinger, M., Mykhaylyk, O., Rudolph, C. and Plank, C. (2016) Chemically modified RNA induces osteogenesis of stem cells and human tissue explants as well as accelerates bone healing in rats. *Biomaterials*, 87, 131-146.

[352] Pardi, N., Hogan, M.J., Pelc, R.S., Muramatsu, H., Andersen, H., DeMaso, C.R., Dowd, K.A., Sutherland, L.L., Scearce, R.M., Parks, R. *et al.* (2017) Zika virus protection by a single low-dose nucleoside-modified mRNA vaccination. *Nature*, 543, 248-251.

[353] Espeseth, A.S., Cejas, P.J., Citron, M.P., Wang, D., DiStefano, D.J., Callahan, C., Donnell, G.O., Galli, J.D., Swoyer, R., Touch, S. *et al.* (2020) Modified mRNA/lipid nanoparticle-based vaccines expressing respiratory syncytial virus F protein variants are immunogenic and protective in rodent models of RSV infection. *NPJ Vaccines*, 5, 16.

[354] Richner, J.M., Himansu, S., Dowd, K.A., Butler, S.L., Salazar, V., Fox, J.M., Julander, J.G., Tang, W.W., Shresta, S., Pierson, T.C. *et al.* (2017) Modified mRNA Vaccines Protect against Zika Virus Infection. *Cell*, 168, 1114-1125 e1110.

[355] Walsh, E.E., Frenck, R., Falsey, A.R., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Neuzil, K., Mulligan, M.J., Bailey, R. *et al.* (2020) RNA-Based COVID-19 Vaccine BNT162b2 Selected for a Pivotal Efficacy Study. *medRxiv*.

[356] Sahin, U., Kariko, K. and Tureci, O. (2014) mRNA-based therapeutics--developing a new class of drugs. *Nat Rev Drug Discov*, 13, 759-780.

[357] Kariko, K., Kuo, A. and Barnathan, E. (1999) Overexpression of urokinase receptor in mammalian cells following administration of the in vitro transcribed encoding mRNA. *Gene Ther*, 6, 1092-1100.

[358] Leppek, K., Das, R. and Barna, M. (2018) Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol*, 19, 158-174.

[359] Schlake, T., Thess, A., Fotin-Mleczek, M. and Kallen, K.J. (2012) Developing mRNA-vaccine technologies. *RNA biology*, 9, 1319-1330.

[360] Park, J.E., Yi, H., Kim, Y., Chang, H. and Kim, V.N. (2016) Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Mol Cell*, 62, 462-471.

[361] Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E. and Mattaj, I.W. (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, 78, 657-668.

[362] Arango, D., Sturgill, D., Alhusaini, N., Dillman, A.A., Sweet, T.J., Hanson, G., Hosogane, M., Sinclair, W.R., Nanan, K.K., Mandler, M.D. *et al.* (2018) Acetylation of Cytidine in mRNA Promotes Translation Efficiency. *Cell*, 175, 1872-1886 e1824.

[363] Rimbach, K., Kaiser, S., Helm, M., Dalpke, A.H. and Eigenbrod, T. (2015) 2'-O-Methylation within Bacterial RNA Acts as Suppressor of TLR7/TLR8 Activation in Human Innate Immune Cells. *J Innate Immun*, 7, 482-493.

[364] Hengesbach, M., Kobitski, A., Voigts-Hoffmann, F., Frauer, C., Nienhaus, G.U. and Helm, M. (2008) RNA intramolecular dynamics by single-molecule FRET. *Curr Protoc Nucleic Acid Chem*, Chapter 11, Unit 11 12.

[365] Elroy-Stein, O., Fuerst, T.R. and Moss, B. (1989) Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 6126-6130.

[366] Beckler, G.S., Thompson, D. and Van Oosbree, T. (1995) In vitro translation using rabbit reticulocyte lysate. *Methods Mol Biol*, 37, 215-232.

[367] Jackson, R.J. and Hunt, T. (1983) Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol*, 96, 50-74.

[368] Holen, T., Amarzguioui, M., Babaie, E. and Prydz, H. (2003) Similar behaviour of single-strand and double-strand siRNAs suggests they act through a common RNAi pathway. *Nucleic Acids Res*, 31, 2401-2407.

[369] Hellmuth, I., Freund, I., Schloder, J., Seidu-Larry, S., Thuring, K., Slama, K., Langhanki, J., Kaloyanova, S., Eigenbrod, T., Krumb, M. *et al.* (2017) Bioconjugation of Small Molecules to RNA Impedes Its Recognition by Toll-Like Receptor 7. *Front Immunol*, 8, 312.

[370] Sawant, A.A., Tanpure, A.A., Mukherjee, P.P., Athavale, S., Kelkar, A., Galande, S. and Srivatsan, S.G. (2016) A versatile toolbox for posttranscriptional chemical labeling and imaging of RNA. *Nucleic Acids Res*, 44, e16.

[371] Sousa, R. (2001) T 7 RNA polymerase. *Uirusu*, 51, 81-94.

[372] Sousa, R. and Mukherjee, S. (2003) T7 RNA polymerase. *Prog Nucleic Acid Res Mol Biol*, 73, 1-41.

[373] Negishi, E.-i. and Baba, S. (1976) Novel stereoselective alkenyl–aryl coupling via nickel-catalysed reaction of alkenylanes with aryl halides. *J. Chem. Soc., Chem. Commun.*, 596b-597b.

[374] Kariko, K., Muramatsu, H., Ludwig, J. and Weissman, D. (2011) Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Res*, 39, e142.

[375] Zheng, B. and West, L.M. (2009) Estimating the Lipophilicity of Natural Products using a Polymeric Reversed Phase HPLC Method. *J Liq Chromatogr Relat Technol*, 33, 118-132.

[376] Dickman, M.J. and Hornby, D.P. (2006) Enrichment and analysis of RNA centered on ion pair reverse phase methodology. *RNA*, 12, 691-696.

[377] Fremont, P., Dionne, F.T. and Rogers, P.A. (1986) Recovery of biologically functional messenger RNA from agarose gels by passive elution. *Anal Biochem*, 156, 508-514.

[378] Flett, F. and Interthal, H. (2013) Separation of DNA oligonucleotides using denaturing urea PAGE. *Methods Mol Biol*, 1054, 173-185.

[379] Baronti, L., Karlsson, H., Marusic, M. and Petzold, K. (2018) A guide to large-scale RNA sample preparation. *Anal Bioanal Chem*, 410, 3239-3252.

[380] Abe, N., Matsumoto, K., Nishihara, M., Nakano, Y., Shibata, A., Maruyama, H., Shuto, S., Matsuda, A., Yoshida, M., Ito, Y. *et al.* (2015) Rolling Circle Translation of Circular RNA in Living Human Cells. *Sci Rep*, 5, 16435.

[381] Green, M.R. and Sambrook, J. (2020) Recovery of DNA from Low-Melting-Temperature Agarose Gels: Organic Extraction. *Cold Spring Harb Protoc*, 2020, pdb prot100461.

[382] Schmitt, J.J. and Cohen, B.N. (1983) Quantitative isolation of DNA restriction fragments from low-melting agarose by Elutip-d affinity chromatography. *Anal Biochem*, 133, 462-464.

[383] Wieslander, L. (1979) A simple method to recover intact high molecular weight RNA and DNA after electrophoretic separation in low gelling temperature agarose gels. *Anal Biochem*, 98, 305-309.

[384] Sambrook, J. and Russell, D.W. (2006) Recovery of DNA from Low-melting-temperature Agarose Gels: Enzymatic Digestion with Agarase. *CSH Protoc*, 2006.

[385] Langridge, J., Langridge, P. and Bergquist, P.L. (1980) Extraction of nucleic acids from agarose gels. *Anal Biochem*, 103, 264-271.

[386] Chen, C.W. and Thomas, C.A., Jr. (1980) Recovery of DNA segments from agarose gels. *Anal Biochem*, 101, 339-341.

[387] Vogelstein, B. and Gillespie, D. (1979) Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences of the United States of America*, 76, 615-619.

[388] Watanabe, M. (1999) Rapid and inexpensive recovery method of DNA fragments from agarose and polyacrylamide gels by a cotton-wool column tube. *Nucleic Acids Symp Ser*, 101-102.

[389] Sun, Y., Sriramajayam, K., Luo, D. and Liao, D.J. (2012) A quick, cost-free method of purification of DNA fragments from agarose gel. *J Cancer*, 3, 93-95.

[390] Thuring, R.W., Sanders, J.P. and Borst, P. (1975) A freeze-squeeze method for recovering long DNA from agarose gels. *Anal Biochem*, 66, 213-220.

[391] Tautz, D. and Renz, M. (1983) An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. *Anal Biochem*, 132, 14-19.

[392] Moore, D., Chory, J. and Ribaudo, R.K. (2001) Isolation and purification of large DNA restriction fragments from agarose gels. *Curr Protoc Immunol*, Chapter 10, Unit 10 15.

[393] Sambrook, J. and Russell, D.W. (2006) Recovery of DNA from agarose and polyacrylamide gels: electroelution into dialysis bags. *CSH Protoc*, 2006.

[394] Sheng, Y. and Bowser, M.T. (2014) Isolating single stranded DNA using a microfluidic dialysis device. *Analyst*, 139, 215-224.

[395] Gilbert, J.R. and Vance, J.M. (2001) Isolation of genomic DNA from mammalian cells. *Curr Protoc Hum Genet*, Appendix 3, Appendix 3B.

[396] Krerowicz, S.J., Hernandez-Ortiz, J.P. and Schwartz, D.C. (2019) A simple dialysis device for large DNA molecules. *Biotechniques*, 66, 93-95.

[397] Zarzosa-Alvarez, A.L., Sandoval-Cabrera, A., Torres-Huerta, A.L. and Bermudez-Cruz, R.M. (2010) Electroeluting DNA fragments. *J Vis Exp*.

[398] Lopez-Gomollon, S. and Nicolas, F.E. (2013) Purification of DNA Oligos by denaturing polyacrylamide gel electrophoresis (PAGE). *Methods Enzymol*, 529, 65-83.

[399] Fadouloglou, V.E. (2013) Electroelution of nucleic acids from polyacrylamide gels: a custom-made, agarose-based electroeluter. *Anal Biochem*, 437, 49-51.

[400] Jeffs, L.B., Palmer, L.R., Ambegia, E.G., Giesbrecht, C., Ewanick, S. and MacLachlan, I. (2005) A scalable, extrusion-free method for efficient liposomal encapsulation of plasmid DNA. *Pharm Res*, 22, 362-372.

[401] Helm, M., Brule, H., Degoul, F., Cepanec, C., Leroux, J.P., Giege, R. and Florentz, C. (1998) The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res*, 26, 1636-1643.

[402] Degoul, F., Brule, H., Cepanec, C., Helm, M., Marsac, C., Leroux, J., Giege, R. and Florentz, C. (1998) Isoleucylation properties of native human mitochondrial tRNAIle and tRNAIle transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the tRNAIle gene. *Hum Mol Genet*, 7, 347-354.

[403] Petrov, A., Wu, T., Puglisi, E.V. and Puglisi, J.D. (2013) RNA purification by preparative polyacrylamide gel electrophoresis. *Methods Enzymol*, 530, 315-330.

[404] Gibson, J.F., Kelso, S. and Skevington, J.H. (2010) Band-cutting no more: A method for the isolation and purification of target PCR bands from multiplex PCR products using new technology. *Mol Phylogenet Evol*, 56, 1126-1128.

[405] Richter, F., Plehn, J.E., Bessler, L., Hertler, J., Jorg, M., Cirzi, C., Tuorto, F., Friedland, K. and Helm, M. (2021) RNA marker modifications reveal the necessity for rigorous preparation protocols to avoid artifacts in epitranscriptomic analysis. *Nucleic Acids Res*.

[406] Kryndushkin, D.S., Alexandrov, I.M., Ter-Avanesyan, M.D. and Kushnirov, V.V. (2003) Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *The Journal of biological chemistry*, 278, 49636-49643.

[407] Sambrook, J. and Russel, D. (2001) Molecular cloning - A laboratory manual, 3rd edition. *Science*, 292, 446-446.

[408] Fedorcsak, I. and Ehrenberg, L. (1966) Effects of diethyl pyrocarbonate and methyl methanesulfonate on nucleic acids and nucleases. *Acta Chem Scand*, 20, 107-112.

[409] Gauthier, D. and Ven Murthy, M.R. (1987) Efficacy of RNase inhibitors during brain polysome isolation. *Neurochem Res*, 12, 335-339.

[410] Narumi, R., Yamamoto, T., Inoue, A. and Arata, T. (2012) Substrate-induced conformational changes in sarcoplasmic reticulum Ca2+-ATPase probed by surface modification using diethylpyrocarbonate with mass spectrometry. *FEBS Lett*, 586, 3172-3178.

[411] Huang, Y.H., Leblanc, P., Apostolou, V., Stewart, B. and Moreland, R.B. (1995) Comparison of Milli-Q PF Plus water to DEPC-treated water in the preparation and analysis of RNA. *Biotechniques*, 19, 656-661.

[412] Kestens, V., Bozatzidis, V., De Temmerman, P.J., Ramaye, Y. and Roebben, G. (2017) Validation of a particle tracking analysis method for the size determination of nano- and microparticles. *J Nanopart Res*, 19, 271.

[413] Kellner, S., Neumann, J., Rosenkranz, D., Lebedeva, S., Ketting, R.F., Zischler, H., Schneider, D. and Helm, M. (2014) Profiling of RNA modifications by multiplexed stable isotope labelling. *Chem Commun (Camb)*, 50, 3516-3518.

[414] Macon, J.B. and Wolfenden, R. (1968) 1-Methyladenosine. Dimroth rearrangement and reversible reduction. *Biochemistry*, 7, 3453-3458.

[415] Su, D., Chan, C.T., Gu, C., Lim, K.S., Chionh, Y.H., McBee, M.E., Russell, B.S., Babu, I.R., Begley, T.J. and Dedon, P.C. (2014) Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nat Protoc*, 9, 828-841.

[416] Kellner, S., Ochel, A., Thuring, K., Spenkuch, F., Neumann, J., Sharma, S., Entian, K.D., Schneider, D. and Helm, M. (2014) Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. *Nucleic Acids Res*, 42, e142.

[417] Castagnetti, S., Hentze, M.W., Ephrussi, A. and Gebauer, F. (2000) Control of oskar mRNA translation by Bruno in a novel cell-free system from Drosophila ovaries. *Development*, 127, 1063-1068.

[418] Michel, Y.M., Poncet, D., Piron, M., Kean, K.M. and Borman, A.M. (2000) Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for poly(A)-mediated stimulation of translation initiation. *The Journal of biological chemistry*, 275, 32268-32276.

[419] Hillebrecht, J.R. and Chong, S. (2008) A comparative study of protein synthesis in in vitro systems: from the prokaryotic reconstituted to the eukaryotic extract-based. *BMC Biotechnol*, 8, 58.

[420] Anastasina, M., Terenin, I., Butcher, S.J. and Kainov, D.E. (2014) A technique to increase protein yield in a rabbit reticulocyte lysate translation system. *Biotechniques*, 56, 36-39.

[421] Nachtergaele, S. and He, C. (2018) Chemical Modifications in the Life of an mRNA Transcript. *Annu Rev Genet*, 52, 349-372.

[422] Helm, M. and Motorin, Y. (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat Rev Genet*, 18, 275-291.

[423] Picard-Jean, F., Brand, C., Tremblay-Letourneau, M., Allaire, A., Beaudoin, M.C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J. *et al.* (2018) 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. *PLoS One*, 13, e0193804.

[424] Durbin, A.F., Wang, C., Marcotrigiano, J. and Gehrke, L. (2016) RNAs Containing Modified Nucleotides Fail To Trigger RIG-I Conformational Changes for Innate Immune Signaling. *mBio*, 7.

[425] Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44, 283-292.

[426] Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*, 15, 8125-8148.

[427] Gholamalipour, Y., Karunanayake Mudiyanselage, A. and Martin, C.T. (2018) 3' end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character-RNA-Seq analyses. *Nucleic Acids Res*, 46, 9253-9263.

[428] You, C., Dai, X. and Wang, Y. (2017) Position-dependent effects of regioisomeric methylated adenine and guanine ribonucleosides on translation. *Nucleic Acids Res*, 45, 9059-9067.

[429] Li, B., Luo, X. and Dong, Y. (2016) Effects of Chemically Modified Messenger RNA on Protein Expression. *Bioconjugate chemistry*, 27, 849-853.

[430] Paradis, E. and Schliep, K. (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526-528.

[431] Martin, C.T. and Coleman, J.E. (1989) T7 RNA polymerase does not interact with the 5'-phosphate of the initiating nucleotide. *Biochemistry*, 28, 2760-2762.

[432] Moore, D., Dowhan, D., Chory, J. and Ribaudo, R.K. (2002) Isolation and purification of large DNA restriction fragments from agarose gels. *Curr Protoc Mol Biol*, Chapter 2, Unit 2 6.

# List of Publications

(1) Marchand, V., Ayadi, L., Ernst, F.G.M., **Hertler, J.**, Bourguignon-Igel, V., Galvanin, A., Kotter, A., Helm, M., Lafontaine, D.L.J. & Motorine, I. AlkAniline-Seq: Profiling of m7G and m3C RNA Modifications at Single Nucleotide Resolution. Angew Chem Int Ed Engl. 57(51), 16785-16790 (2018).

(2) Slama, K., Galliot, A., Weichmann, F., **Hertler, J.**, Feederle, R., Meister, G., & Helm, M. Determination of enrichment factors for modified RNA in MeRIP experiments. Methods 156, 102–109 (2019).

(3) Richter, F., Plehn, J.E., Bessler, L., **Hertler, J.**, Jörg, M., Cirzi, C., Tuorto, F., Friedland, K. & Helm, M. RNA marker modifications reveal the necessity for rigorous preparation protocols to avoid artifacts in epitranscriptomic analysis. Nucleic Acid Res. 1150 (2021) Online ahead of print.

In revision:

(4) **Hertler, J.***, Slama, K.*, Marchand V., Motorin Y., & Helm M. Synthesis of point-modified mRNA.

*These authors contributed equally to this work.

# Curriculum Vitae