

Dissertation zur Erlangung des Doktorgrades der
Fakultät für Chemie und Pharmazie der
Ludwig-Maximilians-Universität München

**Untersuchungen zu den Interaktionspartnern
modifizierter Nukleoside
in *messenger* RNA**

Simon Geiger

aus
Dachau

2017

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München,

.....

(Simon Geiger)

Dissertation eingereicht am: 03.02.2017

1. Gutachter: Prof. Dr. Thomas Carell

2. Gutachter: Prof. Dr. Rasmus Linser

Mündliche Prüfung am: 17.03.2017

Inhaltsverzeichnis

Abstract	I
Zusammenfassung	VI
1. Einleitung	1
1.1 Grundlagen	1
1.2 Die chemische Diversität von Ribonukleinsäuren	3
1.3 Posttranskriptionelle Modifikationen von messenger RNA	7
1.3.1 Pseudouridin.....	7
1.3.2 2'-O-Methylmodifikationen.....	9
1.3.3 5-Methylcytidin.....	11
1.3.4 N ⁶ -Methyladenosin.....	12
1.3.5 N ¹ -Methyladenosin.....	16
1.3.6 Inosin.....	17
1.3.7 Die Isopentenylmodifikationen N ⁶ -Isopentenyladenosin und N ⁶ -Isopentenyl-2-thiomethyladenosin.....	18
2. Aufgabenstellung	35
3. Synthese der modifizierten Ribonukleoside und ihrer Phosphoramidite	37
3.1 Synthese der Isopentenylmodifikationen i⁶A und ms²i⁶A	37
3.1.1 Synthese des N ⁶ -Isopentenyladenosin Phosphoramidits.....	37
3.1.2 Synthese des 2-Thiomethyl-N ⁶ -isopentenyladenosin Phosphoramidits.....	40
3.2 Synthese der N⁶-methylierten Adenosine m⁶A und m⁶₂A	42
3.2.1 Synthese des N ⁶ -Methyladenosin Phosphoramidits.....	42
3.2.2 Synthese des N ⁶ ,N ⁶ -Dimethyladenosin Phosphoramidits.....	44
4. Synthese der Oligoribonukleotide	45
5. Interaktomstudien zu den Modifikationen m⁶A, m⁶₂A, i⁶A, ms²i⁶A, Ψ, Am, Gm, Cm und Um	54
5.1 Allgemeine Methode des SILAC Experimentes	54
5.2 Interaktomstudien zu m⁶A	56
5.2.1 SILAC-Experimente.....	56
5.2.2 Validierungsexperimente zu m ⁶ A.....	67
5.2.3 Zusammenfassung.....	70
5.3 Analyse des Interaktoms von N⁶,N⁶-Dimethyladenosin	75
5.4 Analyse des Interaktoms von Pseudouridin	77
5.5 Interaktomanalyse der 2'-O-Methylmodifikationen	81
5.6 Studien zur Funktion von ms²i⁶A	90
5.6.1 Interaktomanalyse.....	90
5.6.2 Validierungsexperimente.....	97
5.6.3 Lokalisierungsstudien zu ms ² i ⁶ A Bindern.....	109
5.6.4 Charakterisierung der Cdk5Rap1-Knockout Zelllinie.....	110
5.6.5 Addendum zum ms ² i ⁶ A Interaktom.....	118
6. Experimenteller Teil	120
6.1 Biochemische und mikrobiologische Methoden	120

6.1.1	Material und allgemeine Methoden	120
6.1.2	Transformation von chemisch kompetenten <i>E. coli</i> Zellen	124
6.1.3	Bakterien-Übernachtskultur	124
6.1.4	Plasmidpräparation	124
6.1.5	Konzentrationsbestimmung von Nukleinsäuren	125
6.1.6	Gateway-Klonierung.....	126
6.1.7	DNA Restriktion.....	129
6.1.8	DNA-Agarosegelelektrophorese.....	129
6.1.9	Sequenzierung.....	129
6.1.10	Konzentrationsbestimmung von Proteinen mittels Bradford Assay.....	130
6.1.11	Zellkultur	130
6.1.12	Interaktionsassays	131
6.1.13	Proteomweite <i>Label-Free Quantification</i> (LFQ).....	138
6.1.14	Konfokalmikroskopische Lokalisationsstudien.....	140
6.2	Oligonukleotidsynthese.....	142
6.2.1	Festphasensynthese	142
6.2.2	Entschützung, Aufreinigung und Charakterisierung	143
6.3	Chemische Synthesen.....	144
6.3.1	Material und allgemeine Methoden	144
6.3.2	Durchführung der Synthesen	145
7.	Anhang	III
7.1	Abkürzungen	III
7.2	Syntheszyklen für die Oligonucleotidsynthese.....	VII
8.	Literaturverzeichnis.....	XIII

Abstract

The central dogma of molecular biology describes the transfer of genetic information through the three essential layers DNA, messenger RNA and protein. Regulation of gene expression takes place during each individual step and is considered epigenetic when it includes inheritable modulation of genetic information by reversible chemical modification of the information carriers. The existence of this form of regulation is established on the levels of DNA and protein. In stark contrast, epigenetic regulation on the level of messenger RNA is an underdeveloped field of research that lately experienced a strong increase in scientific interest and number of publications.

Recent work addressing the biological function and sequence context of N^6 -methyladenosine, N^1 -methyladenosine and pseudouridine hints towards actively modulated mRNA modification patterns that represent a formerly unknown network of epigenetic regulation. In addition to a detailed analysis of the sequence context of such nucleoside derivatives, the knowledge of their protein interaction partners is an essential step towards a deeper understanding of their biological function.

In this work, interactome studies for a number of RNA modifications were performed, to better understand the described regulatory networks. For this purpose N^6 -methyladenosine (m^6A), N^6,N^6 -dimethyladenosine (m^6_2A), pseudouridine (Ψ), N^6 -isopentenyl-2-methylthioadenosine (ms^2i^6A), N^6 -isopentenyladenosine and the four 2'-*O*-methylated nucleosides Am, Gm, Cm and Um (**Figure 1**) were incorporated into oligoribonucleotides, using solid phase synthesis. In the case of i^6A , ms^2i^6A , m^6A and m^6_2A , this required synthesis of the phosphoramidites **15**, **22**, **29** and **31**.

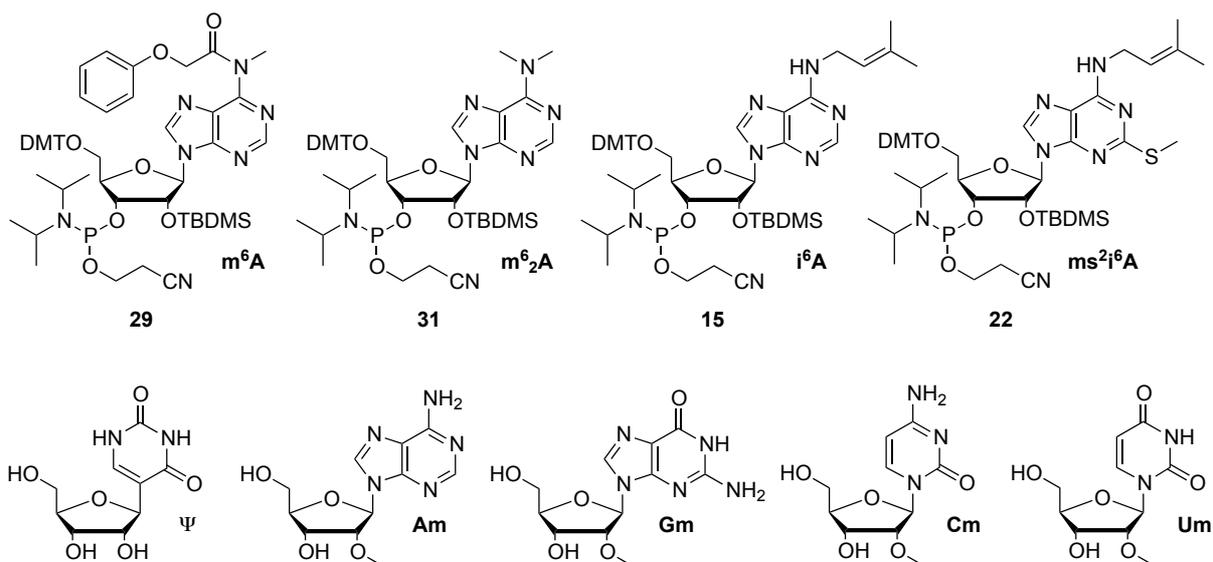


Figure 1: Structures of the phosphoramidite building blocks **15**, **22**, **29** and **31** that were synthesized for interactome studies in the course of this thesis. Solid phase monomers for the other investigated modifications were commercially available.

In a cooperative effort with the group of *Michiel Vermeulen*, the synthesized strands were used for interactome studies utilizing *stable isotope labeling with amino acids in cell culture* (SILAC). In these studies, groups of proteins that specifically interact with the modifications of interest (*reader* proteins), but also networks of *repelled* proteins that show preferential binding to the unmodified control strands, could be identified.

For m⁶A, undoubtedly the “epicenter” of mRNA epigenetics, a number of formerly unknown *readers* were observed. Among them are FMR1, FXR1 and FXR2, proteins that are functionally related to the fragile-X-syndrome, a major cause of inheritable mental retardation. The characteristic phenotype results from a loss in RNA-binding activity of FMR1 and the newly discovered connection between the RNA modification m⁶A and the binding behavior of the central proteins illuminates a highly interesting novel mechanistic detail of these critical RNA-protein interactions.

In addition to these *readers*, the negative interactome of repelled proteins was characterized for the first time. It shows a comparably high degree of variability, depending on the sequence context of the modification and is a rich source for future studies. For the G3BP1 and G3BP2 pair of repelled proteins, a selective interaction with unmodified sequences was validated. Furthermore, their stabilizing effect on bound transcripts was demonstrated which shows that they functionally antagonize m⁶A-binders of the YTH-family that destabilize mRNA. **Figure 2** is a short graphical summary of the findings regarding this modification.

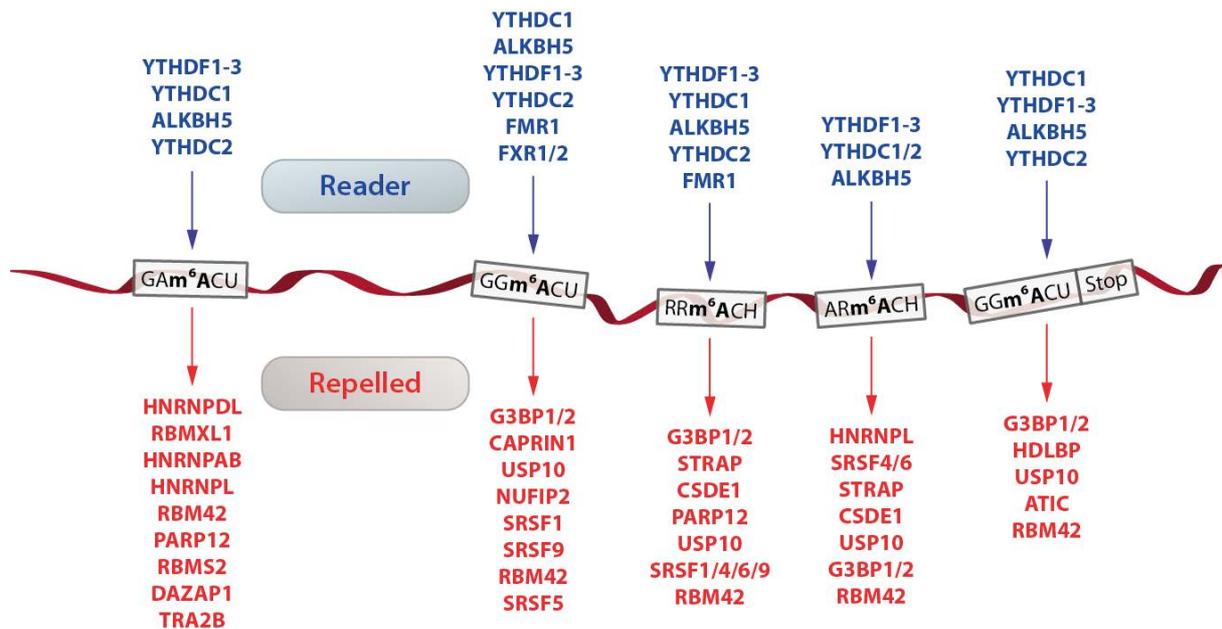


Figure 2: Selection of the most relevant members of the m⁶A interactome, sorted by affinity. Blue: proteins that were found to bind modified strands with higher affinity (*readers*). Red: Proteins that bind control strands with higher affinity (*repelled*).

The group of ribose-methylated nucleosides (Am, Cm, Um and Gm) was found to interact with a highly complex set of proteins that appears to be specific for the individual nucleobases. An involvement of the Nm group in an interferon response to viral infections has been published previously and the results of the investigations presented herein suggest that Um is the modification that mediates this involvement. Aside from this, the repulsion of nucleolytically active complexes and interactions with splicing regulators appear to be features of the group. The fact that Gm strongly repels the YTHDF family of proteins implies an antagonism between m⁶A and Gm that warrants further investigation, since this would enable a previously unknown mechanism of fine-tuning the m⁶A interactome. The results are summarized in **figure 3**.

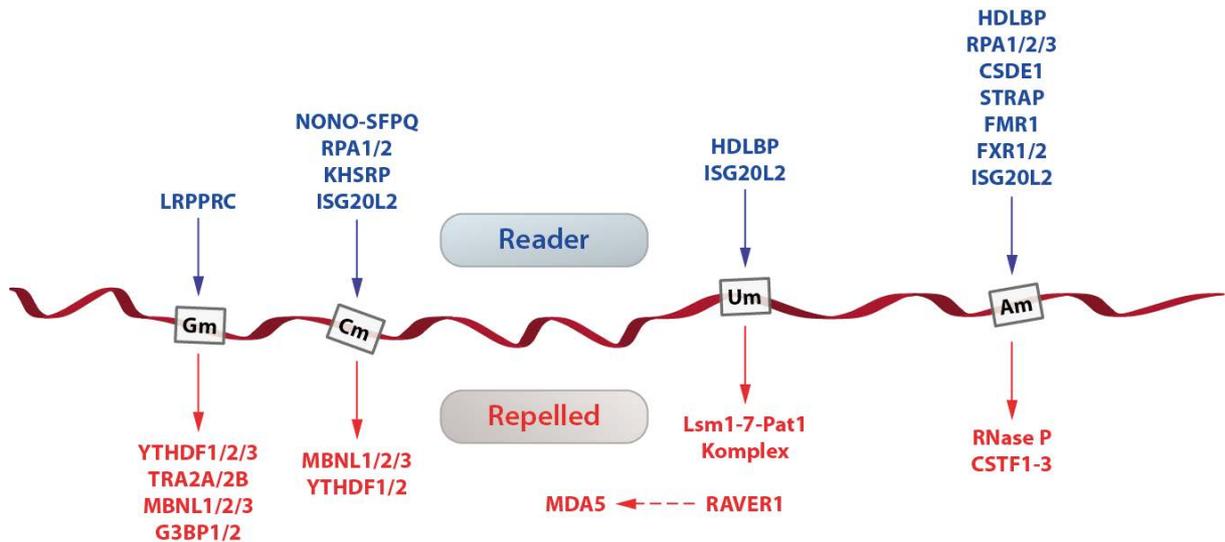


Figure 3: Selection of the most relevant interactome members for the 2'-O-methylated modifications Gm, Cm, Um and Am, sorted by affinity. Blue: proteins that were found to bind modified strands with higher affinity (*readers*). Red: Proteins that bind control strands with higher affinity (*repelled*).

The mRNA hypermodification ms^2i^6A was another key focus of this thesis, being an unusually complex mRNA modification. To further investigate the biological function of the modification, interactome studies were conducted and a knockout cell line of the biosynthetic enzyme, CDK5RAP1, was characterized. The analysis, combining deep transcriptomics with an analysis of the whole cellular proteome (**figure 4**), shows that the absence of ms^2i^6A leads to a general defect in mitochondrial protein biosynthesis, which is in accordance with published data. In addition to this, a pronounced lack of ribosomal proteins was apparent that hints towards a formerly unknown connection between the modification and their biosynthesis.

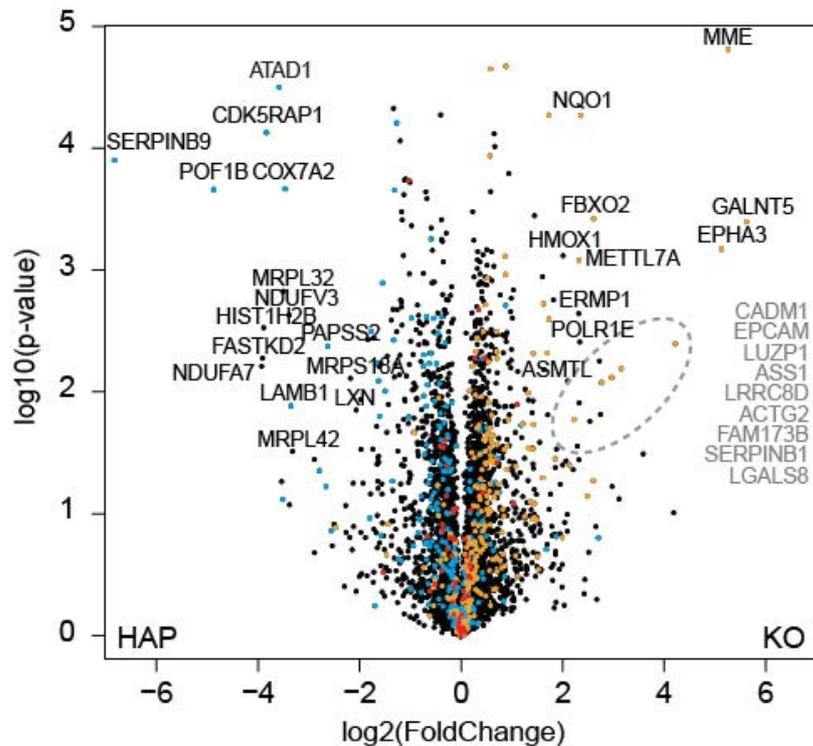


Figure 4: Whole proteome analysis of HAP1 wild type versus knockout cells. Significant proteins are labeled (significance threshold: fold change $> \pm 2$; $p < 0.05$). Proteins are color coded blue, if they are overrepresented in the wild type in transcriptomic data and orange if they are underrepresented. Red-labeled proteins appear as significant hits in SILAC studies using a ms^2i^6A -modified ORN as bait.

Preliminary studies for the remaining three modifications i^6A , m^6_2A and Ψ presented herein represent a good foundation for subsequent investigations.

In conclusion, the data presented in this dissertation analyses the complex interactome of a number of mRNA modifications in an unprecedented depth and quality, among them all four 2'-*O*-methylated nucleosides. Especially for the adenosine derivative m^6A a highly detailed network of readers and repelled proteins in relation to its sequence context could be obtained, which will serve as a valuable resource for future studies. Investigations on the biological role of ms^2i^6A revealed its interactome and include an in-depth characterization of a knockout cell line lacking this hypermodified mRNA modification.

Zusammenfassung

Das zentrale Dogma der Molekularbiologie beschreibt die Weitergabe von genetischer Information über die drei wesentlichen Ebenen DNA, *messenger* RNA und Protein. Als epigenetische Regulation wird eine vererbare Modulierung der genetischen Information durch reversible chemische Derivatisierung der Informationsträger, über die Abfolge der kanonischen Basen oder Aminosäuren hinaus, bezeichnet. Diese Form der Regulation ist auf den Stufen der DNA und der Proteine bekannt und teilweise verstanden, chemische Modifikationen der mRNA sind im Vergleich hierzu dagegen nur sehr unvollständig erforscht.

Aktuelle Arbeiten zu den RNA-Modifikationen N^6 -Methyladenosin, N^1 -Methyladenosin und Pseudouridin deuten darauf hin, dass diese einer aktiven Steuerung unterliegende Modifikationsmuster in mRNA ausbilden. Dies stellt eine bislang nicht bekannte Ebene der Epigenetik, bzw. epigenetischen Regulation dar. Neben Sequenzinformationen stellt die Kenntnis der spezifisch mit solchen Nukleotidderivaten interagierenden Proteine einen unabdingbaren Beitrag zum Verständnis ihrer Funktion dar.

In der hier vorgelegten Arbeit wurden Interaktionsstudien zu mehreren RNA-Modifikationen durchgeführt, um dieses Verständnis zu vertiefen. N^6 -Methyladenosin (m^6A), N^6,N^6 -Dimethyladenosin (m^6_2A), Pseudouridin (Ψ), N^6 -Isopentenyl-2-thiomethyladenosin (ms^2i^6A), N^6 -Isopentenyladenosin (i^6A) und die vier 2'-*O*-Methylmodifikationen Am, Gm, Cm und Um wurden zu diesem Zweck mittels Festphasensynthese in Oligonukleotide inkorporiert (**Abbildung 1**). Die hierfür nötigen Phosphoramidite wurden, im Falle von i^6A , ms^2i^6A , m^6A und m^6_2A synthetisch dargestellt (**15**, **22**, **29** und **31**).

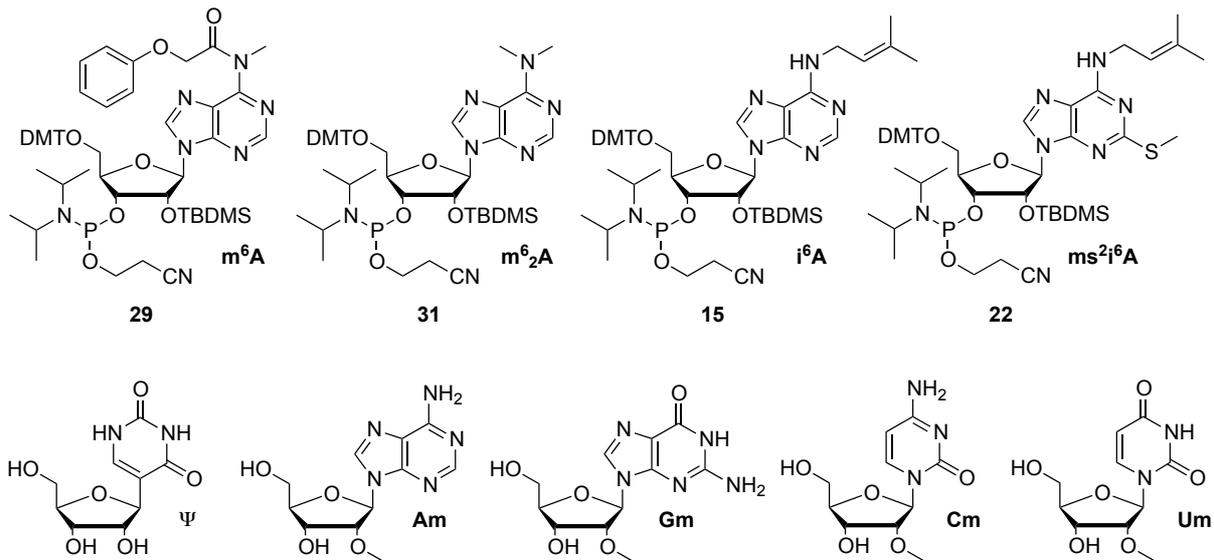


Abbildung 1: Strukturen der für Interaktomstudien dargestellten Phosphoramidite **15**, **22**, **29** und **31**. Festphasenmono-mere der weiteren untersuchten und hier abgebildeten Modifikationen waren kommerziell erhältlich.

In Kooperation mit der Gruppe von *Michiel Vermeulen* wurden die so erhaltenen RNA-Stränge mit Hilfe von *Stable isotope labeling with amino acids in cell culture* (SILAC) in Interaktomstudien eingesetzt. Auf diese Weise konnten sowohl Proteine identifiziert werden, die spezifisch mit den betrachteten Modifikationen interagieren (*Reader*), als auch solche, die bevorzugt an unmodifizierte Sequenzen binden, also abgestoßen werden.

Für das bislang im Zentrum des wissenschaftlichen Interesses stehende m⁶A konnten so neue *Reader* identifiziert werden, darunter die RNA-bindenden Proteine FMR1, FXR1 und FXR2, die mit dem Fragilen-X-Syndrom in funktionellem Zusammenhang stehen. Die Erkrankung resultiert aus dem Verlust der RNA-bindenden Funktion des Proteins FMR1 und stellt eine der häufigsten Ursachen für genetisch bedingte geistige Retardierung dar. Eine Steuerung der Interaktion zwischen diesen Proteinen und ihrem RNA-Substrat in Abhängigkeit von dessen Modifikationsgrad stellt einen hochinteressanten, bislang unbekanntem Zusammenhang zwischen m⁶A und der Erkrankung her.

Neben diesen *Readern* konnte zum ersten Mal das negative Interaktom aus abgestoßenen Proteinen charakterisiert werden. Es zeigt ein hohes Maß an Abhängigkeit vom Sequenzkontext der Modifikation und enthält eine Vielzahl von Anhaltspunkten für weitere Untersuchungen. So konnte für das Paar RNA-stabilisierender Proteine G3BP1 und G3BP2 eine selektive Interaktion mit unmodifizierten Oligonukleotiden belegt werden. Hieraus kann ein funktioneller Antagonismus zwischen diesen Proteinen und den bekannten, mRNA-destabilisierenden m⁶A-*Readern* der YTH-Familie abgeleitet werden, welcher ebenfalls

bislang nicht bekannt ist. Eine kurze grafische Zusammenfassung der Ergebnisse der m⁶A-Interaktomstudien ist in **Abbildung 2** dargestellt.

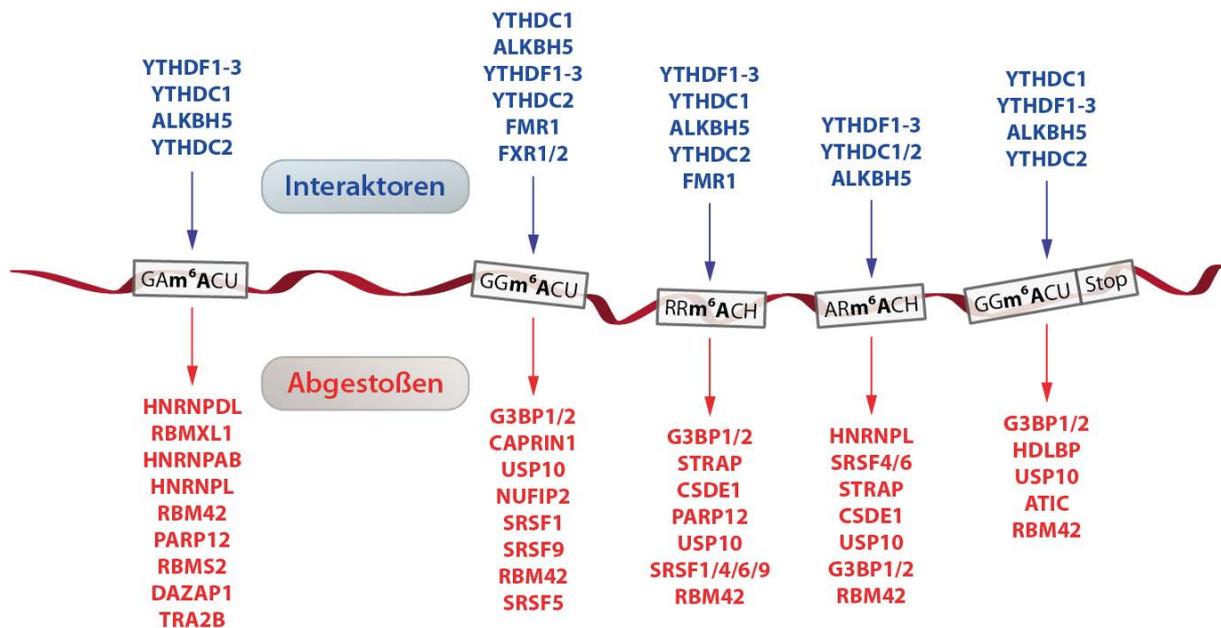


Abbildung 2: Auswahl einiger bedeutender, in SILAC-Studien identifizierten Interaktoren (selektiven Binder) und abgestoßenen Proteine, d.h. der Proteine, die kanonisches Adenosin mit höherer Affinität binden als m⁶A. Die Proteine sind absteigend nach ihrer Affinität geordnet, soweit möglich.

Auch für die Gruppe der ribosemethylierten Nucleoside (Am, Cm, Um, Gm) wurde ein komplexes, weitgehend individuelles Interaktom gefunden. Die Resultate liefern erste Hinweise auf die zellulären Funktionen dieser Gruppe von Modifikationen. Sie deuten darauf hin, dass der sich in aktuellen Studien andeutende Zusammenhang zwischen 2'-O-Methylierung und einer möglichen Interferonantwort auf eine virale Infektion durch Um vermittelt wird. Darüber hinaus zeigen sie, dass mehrere nukleolytisch wirksame Komplexe abgereichert werden und insbesondere Gm die YTHDF-Familie deutlich abstößt, was einen interessanten Hinweis auf einen möglichen m⁶A-Gm-Antagonismus darstellt. Die Ergebnisse sind in **Abbildung 3** dargestellt.

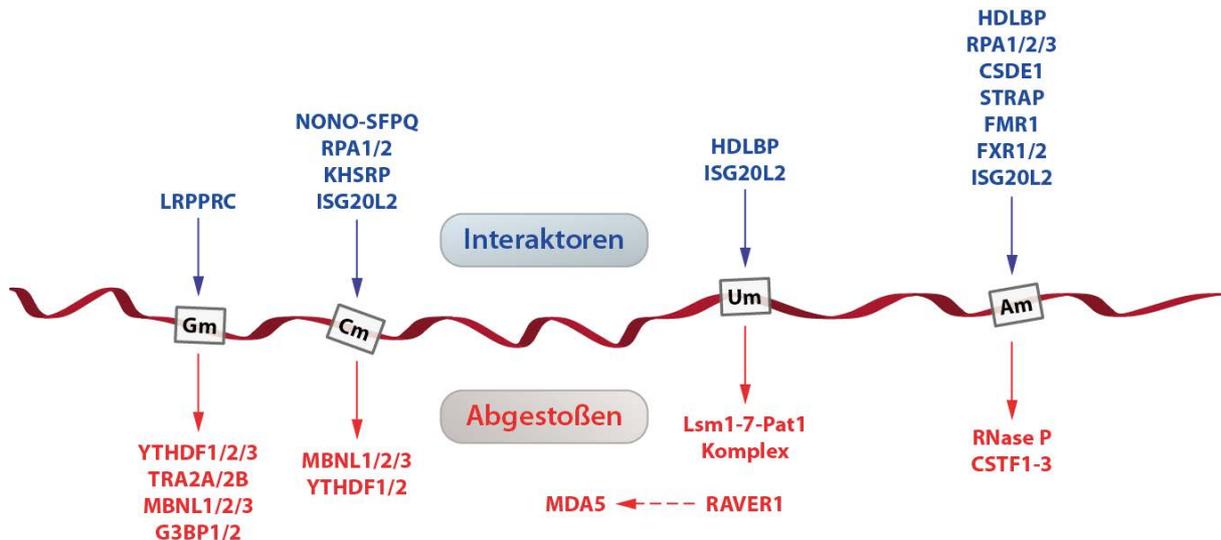


Abbildung 3: Auswahl einiger bedeutender, in SILAC-Studien identifizierten Interaktoren (selektiven Binder) und abgestoßenen Proteine für Gm, Cm, Um und Am. Die Proteine sind absteigend nach ihrer Affinität geordnet, soweit möglich.

Einen weiteren Schwerpunkt dieser Arbeit bildete die Untersuchung der Modifikation ms^2i^6A , welche eine strukturell ungewöhnlich komplexe Modifikation der mRNA darstellt. Um mehr über die Funktion der Modifikation zu erfahren, wurden Interaktomstudien durchgeführt und eine *Knockout*-Zelllinie charakterisiert, welche nicht dazu in der Lage ist ms^2i^6A aus seinem biosynthetischen Vorläufer i^6A herzustellen. Die Analysen auf Transkriptom- und Proteomebene (**Abbildung 4**) zeigten, dass die Abwesenheit des thiomethylierten Adenosinderivates zu einer deutlich messbaren Reduktion der mitochondrialen Proteinsynthese führt, was sich im Einklang mit literaturbekannten Studien befindet. Darüber hinaus zeigte sich aber auch ein schwerer Defekt in der Synthese ribosomaler Proteine, was auf eine bislang nicht bekannt Rolle der Modifikation in diesem Prozess hindeutet.

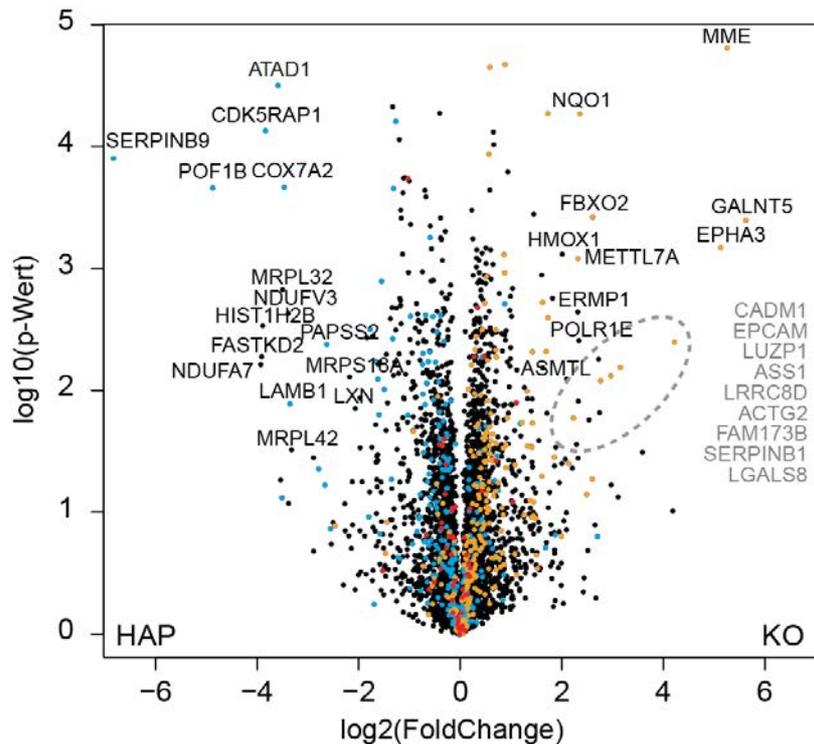


Abbildung 4: Proteomanalyse der HAP1-Wildtyp und Cdk5Rap1-Knockoutzellen. Signifikante Proteine sind mit Namen bezeichnet (Signifikanzgrenzen: FoldChange $> \pm 2$; $p < 0.05$). Proteine sind farblich hervorgehoben, wenn sie in der Transkriptomanalyse im Wildtyp überrepräsentiert (blau), bzw. unterrepräsentiert (orange) sind. Proteine, die im SILAC Interaktom eines ms^2i^6A -modifizierten Baits signifikant an- oder abgereichert wurden, sind rot markiert (Abbildung: Markus Müller).

Auch für i^6A , m^6_2A und Ψ konnten Gruppen interagierender Proteine identifiziert werden, die Daten bilden eine gute Grundlage für weiterführende Studien.

Zusammengenommen bieten die in dieser Dissertation erarbeiteten Daten einen in diesem Umfang und dieser Tiefe bislang nicht verfügbaren Einblick in den komplexen Raum aus Proteinen, die mit Modifikationen der mRNA interagieren. Insbesondere für das Adenosinderivat m^6A konnte ein umfassendes Bild seines Interaktoms in Abhängigkeit vom Sequenzkontext und in mehreren Zelltypen erstellt werden. Weiterhin konnte eine detailreiche Phänotypisierung einer ms^2i^6A -hypomodifizierten Zelllinie erreicht werden, die einen bedeutenden Schritt hin zum Verständnis der Rolle dieser mRNA-Hypermodifikation darstellt.

1. Einleitung

1.1 Grundlagen

Das Verständnis der Desoxyribonukleinsäure (DNA) als Speicher der genetischen Information hat sich seit den frühen Erkenntnissen zu ihrer chemischen Zusammensetzung und molekularen Struktur erheblich weiterentwickelt. Nachdem bereits 1918 durch Levene¹ eine korrekte Anordnung der einzelnen Bausteine in einer kettenartigen Struktur vorgeschlagen wurde, konnte in einer Reihe von darauf folgenden Arbeiten gezeigt werden, dass DNA für die Weitergabe und Speicherung der genetischen Information verantwortlich ist (Avery *et al.* 1944; Hershey und Chase 1952)^{2,3}, ihre Nukleotide jeweils paarweise angeordnet sind (Chargaff 1950)⁴ und dass sie über eine doppelhelikale Struktur verfügt (Watson und Crick, Wilkins *et al.*)^{5,6}. Basierend auf dieser Struktur wurde bereits 1953 ein erster Replikationsmechanismus postuliert⁷, der schnell experimentell untermauert werden konnte⁸.

Diese grundlegenden Kenntnisse ermöglichten die Formulierung erster und im Wesentlichen bis heute gültiger Modelle für die uni- und bidirektionale Weitergabe sequenzieller Information⁹ und des „zentralen Dogmas der Molekularbiologie“¹⁰. Die Vorstellung eines statischen Genoms das, abgesehen von schädlichen Umwelteinflüssen, keinen chemischen Veränderungen unterliegt war aber unzureichend um phänotypische Vielfalt unterschiedlicher Zelltypen desselben Organismus und Teile der Regulationsmechanismen in der Genexpression zu erklären. Sie musste daher erweitert werden. Eine neue Steuerungsebene, die zusätzlich zu der reinen Sequenzinformation eine vererbte Modifikation genetischer Regulationsmuster ermöglichte wurde in der Epigenetik identifiziert. Als epigenetisch gelten Veränderungen auf Nukleinsäure-, bzw. Proteinebene, welche eine phänotypische Modulation bewirken, die in meiotischen oder mitotischen Prozessen an die folgende Generation weitergegeben werden, ohne in der DNA-Sequenz selbst codiert zu sein.^{11,12}

Auf der Ebene der DNA ist heute klar, dass zusätzlich zu den kanonischen Basen Adenin, Guanin, Cytosin und Thymin weitere Derivate der Nukleinbasen existieren, die als chemisch stabile epigenetische Marker einer dynamischen Regulation unterliegen. Mit 5-Methylcytosin (m⁵dC), 5-Hydroxymethylcytosin^{13,14} (hm⁵dC), 5-Formylcytosin¹⁵ (f⁵dC), 5-Carboxycytosin^{16,17} (ca⁵dC) und 5-Hydroxymethyluracil¹⁸ (hm⁵dU) sind bislang fünf

Modifikationen bekannt, die über Methylierung und sequenzielle Oxidation aus Cytidin und Thymin entstehen (**Abbildung 5**).

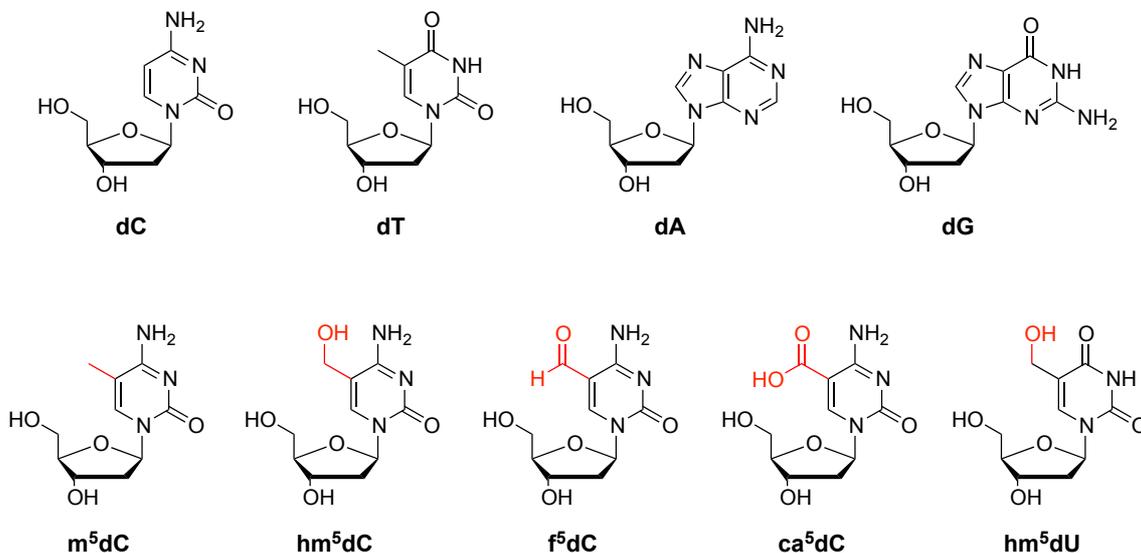


Abbildung 5: Kanonische DNA Basen und bekannte epigenetische Derivate.

Innerhalb dieser Gruppe ist vor allem m^5dC in seiner Wirkung erforscht,¹⁹⁻²¹ die Funktionen der anderen sind weniger klar und zum Teil ist unsicher, ob sie eine definierte Funktion besitzen, die über eine als Intermediate eines oxidativen Demethylierungsprozesses hinaus geht.^{22,23}

Weiterhin ist bekannt, dass auf der Proteinebene mit Histonmodifikationen ebenfalls vererbare Veränderungen mit regulatorischer Wirkung existieren, die zum Teil mit epigenetischen DNA-Modifikationen in einem funktionellem Zusammenhang stehen.²⁴ Die Histonproteine bilden mit der von ihnen gebundenen DNA die Nukleosomen und damit die grundlegenden Organisationseinheiten der genomischen DNA.²⁵ Modulierungen ihrer chemischen Struktur haben einen erheblichen Einfluss, u. A. auf die Packungsdichte des Chromatins und auf die transkriptionelle Aktivität der gebundenen DNA-Abschnitte.²⁶⁻²⁹ Der chemische Raum, den ihre Markierungen abdecken, ist erheblich größer und komplexer als die bislang bekannte Varianz in der DNA-Epigenetik. Er reicht von einfachen Mono-, Di- und Trimethylierungen über Acylierungen (Acetyl-, Propionyl-, Butyryl-, etc.) und ADP-Ribosylierung bis hin zur Anknüpfung kleiner Proteine (Ubiquitin, SUMO), wie in **Abbildung 6** dargestellt ist.³⁰

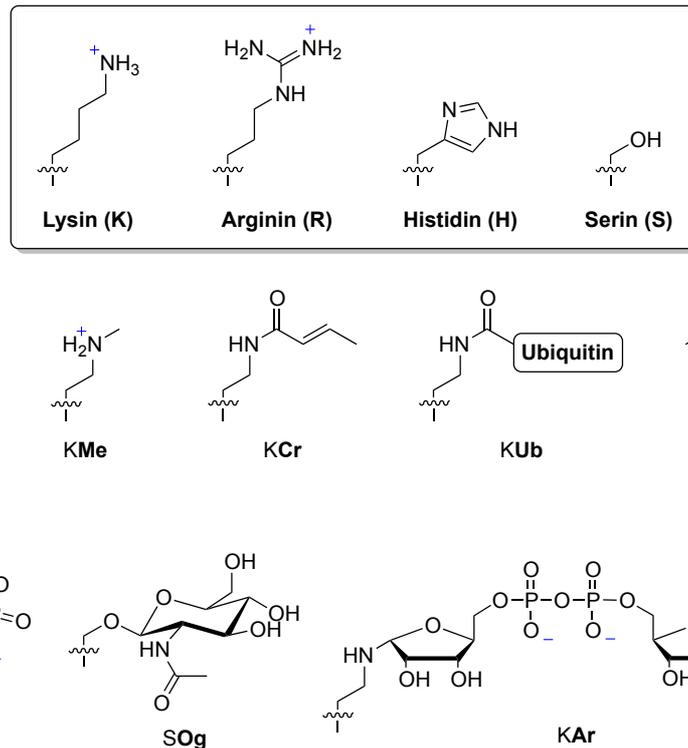


Abbildung 6: Beispiele für Aminosäureseitenketten und ihre möglichen posttranslationalen Modifikationen in Histonproteinen (PTMs). Ac = Acetylierung; Me = Methylierung; Cr = Crotonylierung; Ub = Ubiquitinierung; Ph = Phosphorylierung; Og = *O*-GlcNAcylierung; Ar = ADP-Ribosylierung.

Mit Blick auf den grundlegenden Ablauf der Genexpression lässt sich also sagen, dass sowohl auf der Ebene der DNA als auch der Proteine bekannte Regulationsmechanismen durch dynamisch modulierbare chemische Derivatisierung existieren. Sie bilden zum einen eine Brücke zwischen genomischer Homogenität und phänotypischer Heterogenität höher entwickelter Organismen. Zum anderen stellen sie eine Verbindung von *Nature* und *Nurture* dar, also zwischen weitgehend unveränderbarer genetisch codierter Information bzw. Determinierung und kurzfristig wirksamen Umwelteinflüssen und Anpassungsmechanismen.

Im Vergleich hierzu stellt die RNA-Epigenetik ein junges Feld der Forschung dar³¹ und es ist bislang vergleichsweise wenig über analoge Steuerungsmechanismen bekannt. Die chemische Diversität der Ribonukleinsäuren ist aber erheblich größer als die der Desoxyribonukleinsäure und aktuelle Forschung³²⁻³⁴ konnte zeigen, dass einige RNA Modifikationen in *messenger* RNA einer gezielten Regulation unterliegen.

1.2 Die chemische Diversität von Ribonukleinsäuren

Die strukturelle Vielfalt der bekannten Ribonukleotide ist erheblich größer als die der Desoxyribonukleotide. Derivate aller kanonischen RNA-Basen bilden eine Gruppe von mehr

als 140 verschiedenen Modifikationen mit zum Teil hoher chemischer Komplexität,³⁵⁻³⁷ wobei nahezu alle Positionen des Purin-, bzw. Pyrimidingerüsts, sowie die Riboseeinheit verändert sein können (**Abbildung 7**).

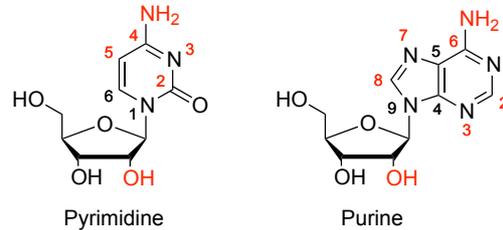


Abbildung 7: Cytidin und Adenosin mit Nomenklatur der Positionen des Purin- und Pyrimidingerüsts. Positionen, die chemischen Modifikationen unterliegen sind in rot hervorgehoben.

Methylierungen machen hierbei einen großen Teil der bekannten Modifikationen aus, aber auch komplexere Alkylierungen (z. B. Isopentenyl-), Acetylierung, Thiomethylierung bis hin zu Isomerisierung (Pseudouridin, Ψ) oder Erweiterungen des Grundgerüsts (Wybutosin, γ W) wurden entdeckt (**Abbildung 8**).

Sowohl die Anzahl unterschiedlicher Modifikationen als auch die Abundanz sind ungleichmäßig über die unterschiedlichen RNA-Spezies verteilt. *Transfer* RNA (tRNA) stellt die in beiderlei Hinsicht höchstmodifizierte Spezies dar. Es wird geschätzt, dass 10 bis 15% aller Nukleoside in tRNA Modifikationen tragen, was einer Rate von etwa acht Modifikationen pro tRNA-Molekül entspricht.³⁸ Einzelne Positionen sind hierbei nahezu quantitativ verändert, während andere nur kanonische Basen enthalten. Hochmodifizierte Positionen sind die *Wobble*-Position 34 und die direkt neben dem Anticodon befindliche Position 37. Vereinfacht betrachtet können die unterschiedlichen Modifikationen in tRNA nach ihrer Funktion in drei Gruppen eingeteilt werden.³⁹ Modifikationen des Anticodonloops beeinflussen die Translation,^{40,41} zum Beispiel über eine Stabilisierung der Codon-Anticodon-Wechselwirkung.⁴² Ihre Abwesenheit kann zu Wachstumsdefekten⁴³⁻⁴⁵ bis hin zur Letalität^{46,47} führen.

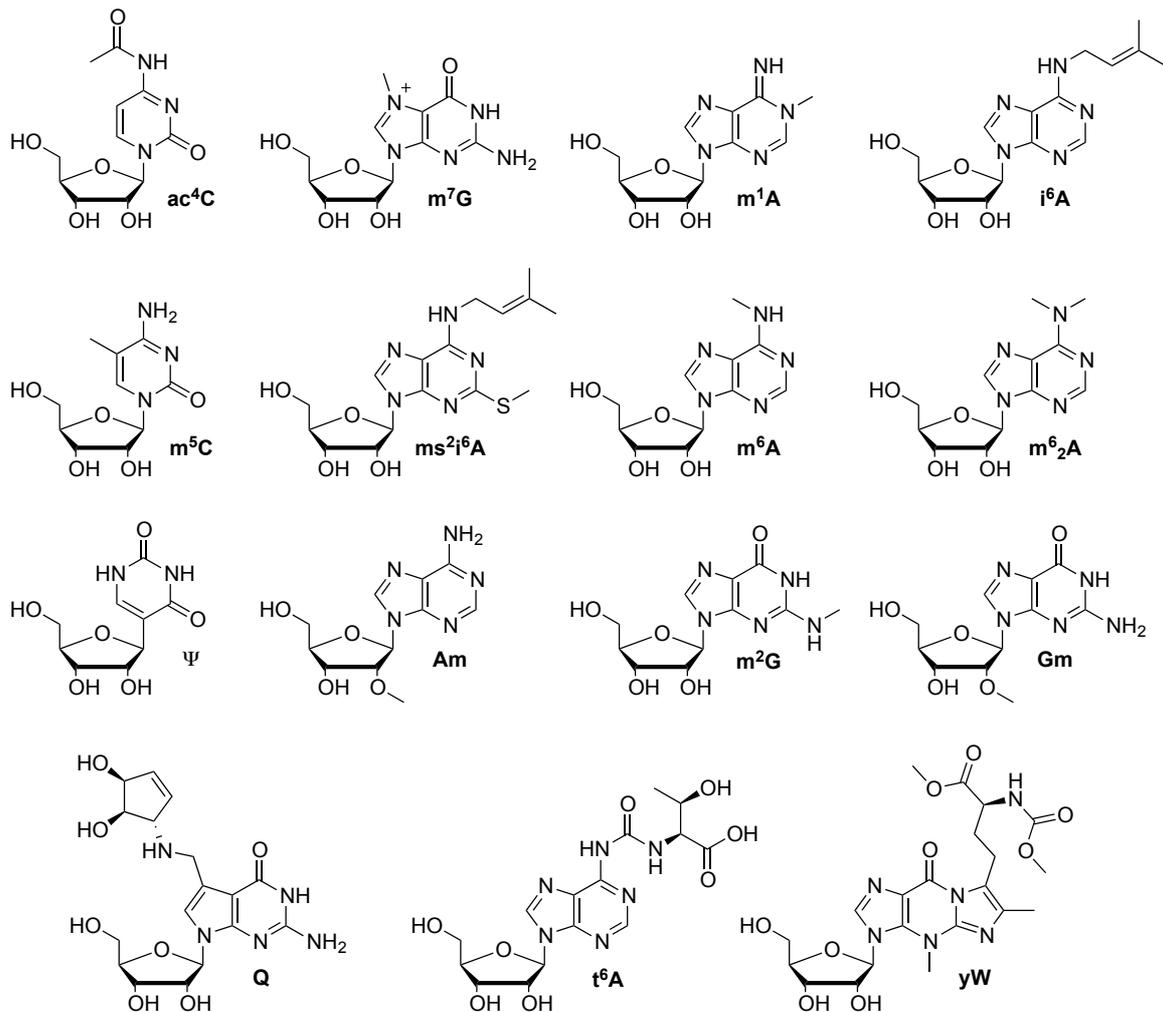


Abbildung 8: Strukturen für Beispiele einiger bekannter RNA-Modifikationen.

Modifikationen des tRNA-Hauptkörpers dienen häufig der korrekten Faltung und Stabilität.⁴⁸⁻
⁵¹ Die dritte Gruppe bilden über die gesamte Sequenz verteilte Modifikationen, die der Erkennung durch tRNA Ligasen dienen bzw. die Identität der tRNA festlegen und damit wichtig für eine korrekte Acylierung sind.⁵²⁻⁵⁵ Einige dieser Modifikationen sind konstitutiv, aber aktuelle Studien in Hefe konnten zeigen, dass sich Modifikationsmuster als Reaktion auf externe Stimuli, wie Toxine und oxidativen Stress ändern.⁵⁶⁻⁵⁸ Dies legt eine regulative Rolle dieser Modifikationen nahe. Darüber hinaus schwankt der Gehalt einiger Modifikationen stark in unterschiedlichen Geweben (Schwein) und in humanen Zellkulturlinien,^{59,60} was zumindest auf eine statische Relation mit dem Phänotyp hinweist.

Auch ribosomale RNA (rRNA) weist Modifikationen auf, ihre strukturelle Vielfalt ist aber erheblich geringer als die der tRNA. Humane rRNA trägt 91 Pseudouridine (Ψ), 105 2'-O-Methylierungen und zehn methylierte Basen.^{34,61} Die modifizierten Nukleoside befinden sich überwiegend in räumlicher Nähe zu den funktionellen Zentren des Ribosoms, wie dem

Peptidyl-Transferasezentrum (PTC).⁶² Es wird angenommen, dass ihre Rolle vor allem in der Stabilisierung von rRNA-Sekundärstrukturen besteht sowie in der Beeinflussung von Interaktionen zwischen rRNA und ribosomalen Proteinen.^{63,64} Sie sind essentiell für eine korrekte Funktion des Ribosoms⁶², scheinen aber statisch zu sein. Eine dynamische Regulation ist bislang rein spekulativ.⁶⁵ So fanden *Krogh et al.* in einer aktuellen Studie eine variabel modifizierte Subpopulation humaner rRNA Loci.⁶⁶ Sie leiten hieraus eine mögliche regulative Funktion der identifizierten Modifikationen ab. Im Kontext ribosomaler RNA Modifikationen ist auch interessant, dass zahlreiche bakterielle rRNA Methyltransferasen (z. B. *cfr*,⁶⁷ *ermC*,⁶⁸ *tsr*⁶⁹) für Resistenzen gegen Antibiotika verantwortlich sind. Die durch sie eingeführten Methylmodifikationen stören die Bindung der betroffenen Antibiotika an das bakterielle Ribosom und die verstärkte Aktivität der Methyltransferase stellt damit durchaus eine Anpassung an veränderte Umweltbedingungen dar.

Über modifizierte Nukleoside in anderen nicht-codierenden RNA Spezies ist bislang wenig bekannt. *Long non-coding RNAs* (lncRNAs) sind in diverse Regulationsprozesse des Chromatin *Remodellings*, die Transkription und die Prozessierung von mRNA involviert.⁷⁰ Einige von ihnen tragen mehrere *N*⁶-Methyladenosine (m⁶A),⁷¹⁻⁷³ die Funktion der Modifikation in dieser Spezies ist aber unklar.

Small nuclear RNAs (snRNAs) stellen essentielle Komponenten der Untereinheiten des Spliceosoms dar. Sie verfügen über vielfältige Modifikationen, darunter konstitutive Methylguanosine am 5'-Terminus (7-Methylguanodin (m⁷G) und 2,2,7-Trimethylguanodin (m₃G))⁷⁴, Pseudouridine⁷⁵⁻⁷⁸ und Methylierungen der Ribose, sowie der Base.^{79,80} Es ist anzunehmen, dass viele dieser Modifikationen zur Stabilisierung von Sekundärstrukturen dienen. Einige Pseudouridine scheinen aber einer dynamischen Regulation zu unterliegen.^{78,81} Weitere Klassen ncRNAs stellen *micro RNAs* (miRNAs), *piwi-interacting RNAs* (piRNAs) und *small interfering RNAs* (siRNAs) dar, die über verwandte Mechanismen wirken, aber unterschiedliche zelluläre Prozesse beeinflussen. Bislang sind lediglich nicht-basenspezifische 2'-*O*-Methylierungen und Inosin in diesen kleinen RNAs bekannt, diese Modifikationen scheinen allerdings eine zentrale Rolle zu spielen. In *Arabidopsis thaliana* tragen sowohl siRNAs, als auch miRNAs eine 3'-terminale 2'-*O*-Methylmodifikation^{82,83}, während piRNAs in Pflanzen nicht vorhanden sind.⁸⁴ Die Methylierung schützt die betroffenen Transkripte vor Polyuridylierung und erhöht ihre Stabilität.^{83,85} Dagegen werden in Tieren zwar piRNAs und siRNAs, nicht aber miRNAs 3'-terminal 2'-*O*-methyliert.⁸⁶⁻⁸⁸ Diese Klasse der RNA

Modifikationen scheint also eine wichtige, allerdings bislang noch weitgehend unverstandene Rolle in siRNAs zu spielen.

Adenosindeaminasen, deren Substrate RNAs sind (ADARs) wandeln Adenosin in Inosin um. Dieses *A-to-I editing* beeinflusst die Reifung und Funktion von miRNAs und siRNAs. Die genauen Mechanismen sind aber ebenfalls noch nicht abschließend geklärt.^{89,90}

Auch *small nucleolar RNAs* (snoRNAs), die als *guide RNAs* im Komplex mit Proteinen die Editierung ihrer Zieltranskripte steuern, wurden durch RNA-Sequenzierung als Träger von Modifikationen identifiziert. Bislang wurde m⁶A in muriner snoRNA und Ψ in Hefe-, sowie humaner snoRNA gefunden.^{75,78,91}

Abschließend ist festzustellen, dass alle erwähnten RNA Spezies in variablem Ausmaß modifiziert sind, wobei vor allem Ψ und 2'-O-Methylierungen zu den Modifikationen gehören, welche in nahezu allen RNAs von Bedeutung sind. Die nicht-kanonischen Basen sind verantwortlich für stabile Sekundärstrukturen, ermöglichen ungewöhnliche Basenpaarungen, sind essentiell in der Qualitätskontrolle, während oder nach Reifungsschritten und dienen der Festlegung der Identität von Transkripten.

1.3 Posttranskriptionelle Modifikationen von *messenger* RNA

Bei der Betrachtung der posttranskriptionellen Derivatisierung von RNA zur dynamischen Regulation zellulärer Prozesse stellt *messenger* RNA die interessanteste Spezies dar. Modifikationen der rRNA und tRNA resultieren notwendigerweise in globalen Effekten, in mRNA ermöglichen sie dagegen eine transkriptspezifische Kontrolle. Die Forschung in diesem Bereich hat in jüngster Zeit einen erheblichen Aufschwung erfahren und der Stand des Wissens soll im Folgenden kurz zusammengefasst werden.

1.3.1 Pseudouridin

Pseudouridin (Ψ), „die fünfte Base“ wurde als erste RNA-Modifikation 1957 von *Davis* und *Allen* entdeckt.⁹² Sie stellt eine der häufigsten Modifikationen in Gesamt-RNA dar und es wurden bereits zahlreiche Ψ-Loci in tRNA, rRNA, snRNA und snoRNA ermittelt. Als Modifikation in mRNA wurde es erst durch aktuelle Sequenzierungsstudien identifiziert.^{75,78,93} Hierbei wurde ausgenutzt, dass *N*-Cyclohexyl-*N'*-(4-methylmorpholinium)ethylcarbodiimid (CMC) zwar mit U, G und Ψ reagiert, aber nur das Ψ-Addukt stabil gegenüber

basischer Nachbehandlung ist. Es blockiert die reverse Transkription und kann damit mittels Sequenzierung detektiert werden (**Abbildung 9**).⁹⁴⁻⁹⁶

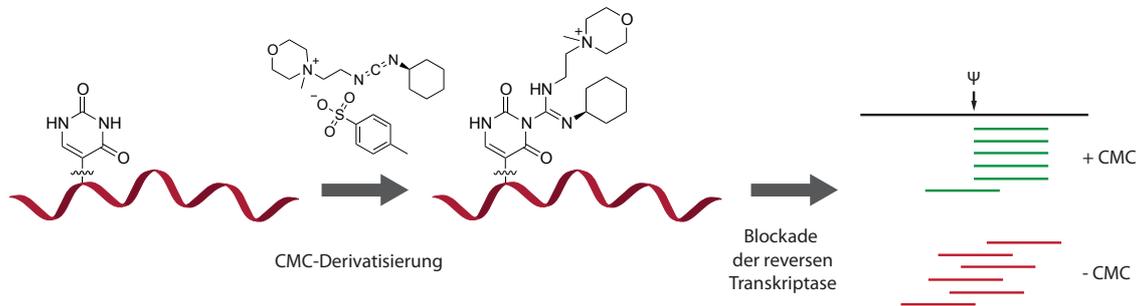


Abbildung 9: Schematischer Ablauf der Sequenzierung von Pseudouridin mit Hilfe chemischer Derivatisierung. Ψ reagiert mit CMC zu einem Addukt, das eine reverse Transkriptase an dieser Position blockiert. Bei der Erstellung der komplementären DNA-Bibliothek (cDNA) entstehen hierdurch Abbrüche an der modifizierten Position, die durch ihre Häufung eine genaue Bestimmung des derivatisierten Locus ermöglichen.

Die drei Publikationen identifizieren stark voneinander abweichende Level Ψ in mRNA aus Hefe (41 bis 466 Loci), humanen Kulturzellen (96 bis 1889 Loci) und murinen Geweben (1543 bis 1741 Loci). Die großen Unterschiede sind zum Teil auf eine variable Tiefe der Sequenzierung zurückzuführen, vor allem aber unterliegt Ψ in mRNA in allen Studien einer ausgeprägten dynamischen Regulation. Umwelteinflüsse, wie Hitzeschock⁷⁵ und Kulturdichte⁷⁸ führten zu massiven Veränderungen des Pseudouridylierungsmusters in Hefe. Der Überlapp zwischen Hirn- und Lebergewebe aus Mäusen ist mit 54 aus 1741/1532 Loci sehr gering, was auf eine ausgeprägt gewebsspezifische Regulation schließen lässt.⁹³ Die Verteilung der gefundenen Positionen deutet auf eine geringe Anzahl Modifikationen pro Transkript hin. Zudem wurde, im Gegensatz zu m⁶A, keine bevorzugte Position (5'-UTR/CDS/3'-UTR) identifiziert.

Ein einziges, globales Konsensmotiv ist nicht bestimmbar, da kein einzelner Proteinkomplex für die Modifizierung der RNA verantwortlich ist. Pseudouridin wird von Pseudouridinsynthasen (PUS) durch Isomerisierung aus U erzeugt.⁹⁷ Es existieren zwei übergeordnete Klassen, die entweder als eigenständige Enzyme mit definiertem RNA-Bindemotiv wirken oder als Proteine ohne klare Sequenzaffinität, die erst im Komplex mit *guide RNAs* (H/ACA snoRNAs) und weiteren Faktoren Spezifität erlangen.⁹⁸ Pseudouridylierung durch eigenständige (RNA-unabhängige) Synthasen konnte bislang lediglich in Hefe, nicht aber in höheren Eukaryoten nachgewiesen werden.^{79,99} Die homologen Proteine existieren aber. Die Heterogenität dieser Enzymgruppe, bzw. der steuernden snoRNAs, spiegelt sich in einer großen Vielfalt modifizierter Motive wider.

Welche Auswirkung die Pseudouridylierung auf die mRNA hat ist bislang unbekannt. Es kann durch verstärktes *stacking* zur Stabilisierung von RNA Duplexen beitragen, und geht mit allen kanonischen Basen Paarungen ein.^{63,100} *Karijolic* und *Yu* spekulierten in einer aktuellen Arbeit, dass eine $U \rightarrow \Psi$ Isomerisierung einer Rekodierung von Stopcodons dienen kann.¹⁰¹ Sie zeigten, dass UAA, UAG und UGA nach der Transformation nicht mehr als Stopcodon erkannt werden, sondern vielmehr je zwei Aminosäuren codieren. Ob die veränderte Translation vom modifizierten Codon selbst, oder von Effekten auf benachbarte Codons ausgeht ist aber umstritten.¹⁰²

Als gesichert gelten kann lediglich, dass Ψ eine konservierte und ubiquitär vorkommende mRNA Modifikation ist, die einer gesteuerten Modulation unterliegt.

1.3.2 2'-O-Methylmodifikationen

Humane mRNAs werden co-transkriptionell mit einer 5'-Capstruktur versehen, die in der Regel direkt an ein bis zwei ribosemethylierte Nukleoside anschließt (**Abbildung 10**). Sie besteht aus einem über eine Triphosphatbrücke 5'-zu-5'-verknüpften Guanosin, das an Position 7 zusätzlich methyliert ist (m^7GpppN). Das 5'-terminale Nukleosid des Transkriptes und auch das benachbarte tragen darüber häufig eine 2'-O-Methylgruppe ($m^7GpppNmpN$ und $m^7GpppNmpNm$).^{103,104} Als interne Modifikation der mRNA konnten Mitglieder dieser Gruppe bislang aber keinen definierten Positionen zugeordnet werden.

Eine Funktion der Methylierungen scheint die Unterscheidung zwischen zelleigener und zellfremder RNA zu sein.¹⁰⁵⁻¹⁰⁷ Viele Virusarten nutzen die Abschwächung der Interferonantwort durch 2'-O-Methylierung, indem sie eigene Methyltransferasen codieren. Auch in der Optimierung artifizieller siRNA zu klinischen Zwecken werden diese Modifikationen genutzt um eine Immunantwort zu minimieren.¹⁰⁸ Interessanterweise müssen die Methylierungen dabei nicht terminal sein, sondern wirken auch als interne Modifikationen. Eine Modulation der Immunantwort wäre also auch durch innerhalb des Transkriptes lokalisierte 2'-O-Methylierungen möglich. Es liegt die Vermutung nahe, dass auch zelleigene, humane mRNAs durch sie von einer Immunantwort ausgenommen werden.

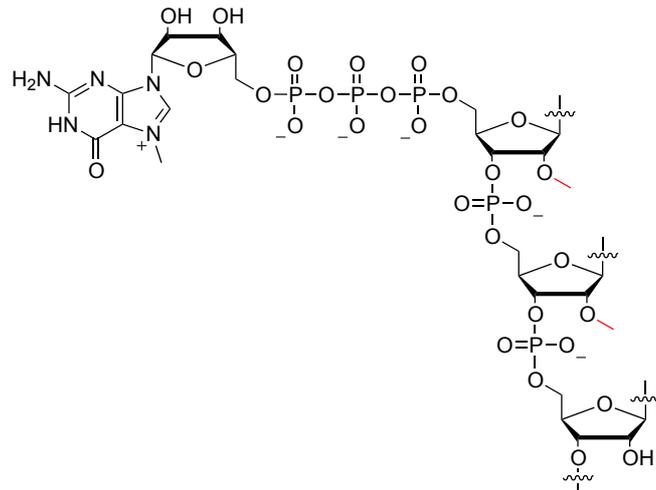


Abbildung 10: Eukaryotische *messenger* RNA trägt eine 5'-Cap-Struktur, die aus einem über ein Triphosphat 5'-zu-5' verknüpften m⁷G besteht und meist ein bis zwei 2'-O-methylierte Nukleoside als terminale Basen einschließt.

Veronika Reiter konnte in Quantifizierungsstudien zeigen, dass aus HeLa Zellen isolierte mRNA sehr hohe Level Am, Gm und Cm enthält.¹⁰⁹ Die ermittelten Werte sind durch ausschließlich ein- oder zweifache 5'-terminale Modifikation jedes Transkriptes nicht zu erklären. Sie lassen auf eine hohe, bislang unterschätzte Abundanz dieser Modifikationen in mRNA schließen. Die Interpretation der Ergebnisse wird aber dadurch erschwert, dass mRNA- bzw. poly(A)-RNA-Isolationen potentiell in erheblichem Umfang mit ribosomaler RNA verunreinigt sind. Nachdem humane rRNA eine hohe Dichte an ribosemethylierten Nukleosiden enthält könnte ein Teil der gemessenen Werte auf eine solche Kontamination zurückzuführen sein.

Die Einführung der Methylierung der Riboseeinheit wird an bekannten Loci (z.B. rRNA) durch komplementäre C/D snoRNAs gesteuert.¹¹⁰ Da zahlreichen snoRNAs bislang kein klares Ziel innerhalb der bekannten nicht-codierenden RNAs zugeordnet werden konnte, ist es möglich, dass sie die Methylierung von mRNAs kontrollieren.¹¹⁰⁻¹¹²

Bemerkenswert erscheint, dass Am erheblich langsamer zu Inosin deaminiert wird als A¹¹³ und eine humane C/D snoRNA einer mRNA-Zielregion zugeordnet werden konnte, die bereits als Ziel für *A-to-I editing* definiert wurde.¹¹² *Cavaillé et al.* leiten daraus ab, dass eine 2'-O-Modifikation an dieser Position der Unterdrückung von *A-to-I editing* an einem definierten Locus dienen könnte. Interessant ist in diesem Zusammenhang außerdem, dass die durch *Brandmayr et al.* detektierten Mengen der Modifikation Am in tRNA aus murinen und porcinen Geweben eine gegenläufige Regulation im Vergleich zu nahezu allen anderen betrachteten Modifikationen zeigten.¹¹⁴ Hohe Level von Am waren mit niedrigen Mengen anderer Modifikationen verbunden und umgekehrt. Eine wechselseitige Beeinflussung

verschiedener RNA-Modifikationen ist vorstellbar und stellt eine mögliche Erklärung für die publizierten Beobachtungen dar.

Hochdurchsatzfähige Sequenzierungsmethoden zur Detektion von 2'-*O*-Methylmodifikationen sind zwar verfügbar,^{66,115,116} bislang wurden aber keine Daten zu codierenden Transkripten publiziert. Daher bleibt die Existenz dieser Sonderbasen über die auf die Capstruktur unmittelbar folgenden Basen hinaus spekulativ.

1.3.3 5-Methylcytidin

Während m⁵dC in DNA die am besten erforschte epigenetische Modifikation darstellt, ist im Vergleich wenig über diese Basenmethylierung in mRNA (m⁵rC, im Folgenden m⁵C) bekannt. Obwohl bereits 1975 m⁵C in mRNA aus Hamsterzellen detektiert wurde,¹¹⁷ war die darauf folgende Forschung lange Zeit auf die Funktion in rRNA fokussiert.¹¹⁸ Das steigende Interesse an chemischen Modifikationen der mRNA und die große Bedeutung, die m⁵dC in der DNA-Epigenetik erlangt hat, haben wieder zu verstärkter Forschung zu diesem Derivat geführt. Eine aktuelle Sequenzierungsstudie konnte 8495 m⁵C-Loci in humaner (HeLa) mRNA identifizieren.¹¹⁹ Die Modifikation ist vor allem in untranslatierten Regionen (5'-UTR und 3'-UTR) der Transkripte zu finden und signifikant an bekannten Bindestellen der Argonaute I-IV Proteine angereichert, was eine Funktion in der posttranskriptionellen Genregulation suggeriert.

Es existieren acht bekannte humane Methyltransferasen die rC zu m⁵rC umsetzen (Nsun2-7, Trdmt1/Dnmt2 und Nop2/Nsun1). Vier Studien, in denen humane und murine mRNA-Interaktome untersucht wurden¹²⁰⁻¹²³ finden drei dieser Methyltransferasen assoziiert mit mRNA: Nsun2, Nsun5 und Nop2. Letztere könnte hierbei ein Artefakt darstellen, da sie rRNA methyliert und damit über Ribosomen mit mRNA assoziiert sein könnte. Die Vermutung liegt nahe, dass Nsun2, Nsun5 und möglicherweise Nop2 codierende Transkripte methylieren können. Für Nsun2 wurde bereits gezeigt, dass es einen Locus im 3'-UTR der p16 mRNA methyliert, wodurch deren Lebenszeit erhöht wird.¹²⁴ Zhang *et al.* kommen allerdings zu dem Schluss, dass die Nsun2-vermittelte Methylierung an einem Adenosin stattfindet und nicht an einem Cytosin.

Auch die oxidierten Derivate hm⁵rC und f⁵rC konnten in Mäusen nachgewiesen werden.¹²⁵ Die Oxidation läuft zum Teil über einen Tet-unabhängigen Mechanismus ab.¹²⁵ Nach Haag *et al.* ist mit ALKBH1 (auch ABH1) ein Mitglied der AlkB-Familie für die Oxidation verantwortlich.¹²⁶ Näheres über die Abundanz und biologische Funktion von m⁵rC und seiner

Derivate in mRNA ist nicht bekannt. *Nakano et al.* konnten zeigen, dass Nsun3 als spezifische Methyltransferase für die Bildung von f⁵C34 in humaner, mitochondrialer tRNA^{Met} essentiell ist. Das erzeugte m⁵C wird aber so schnell oxidiert, dass m⁵C34 und hm⁵C34 nicht nachweisbar sind.¹²⁷ Ein Fehlen des Formylcytosins führt zu reduzierter mitochondrialer Translation und dysfunktionalen Mitochondrien. Damit ist eine dezidierte Funktion von f⁵rC in tRNA, sowie die Existenz oxidierender Enzyme belegt und eine Funktion in mRNA zumindest vorstellbar. Neben humaner tRNA wurde f⁵rC auch in boviner, Hühner-, Ratten- und *Xenopus laevis*-tRNA nachgewiesen.^{128,129}

1.3.4 N⁶-Methyladenosin

Unter den bekannten Modifikationen der mRNA nimmt N⁶-Methyladenosin (m⁶A) eine Sonderstellung ein, als überdurchschnittlich gut in seiner Abundanz und Auswirkung erforschtes Derivat. Es wurde bereits 1974 in humaner mRNA entdeckt^{130,131} und bald darauf erste Konsensmotive ermittelt (Gm⁶AC und Am⁶AC).^{132,133} Die substöchiometrische Modifikation bekannter Loci legte früh eine spezifische Modulation der Methylierung nahe¹³⁴, aber erst die Entdeckung der ersten m⁶A-Demethylase *Fat mass and obesity-associated protein (FTO)* 2011¹³⁵ zeigte, dass sowohl Einführung, als auch Entfernung der Modifikation aktiv katalysiert werden können. In schneller Abfolge wurden darauf folgend mit ALKBH5 eine zweite Demethylase entdeckt¹³⁶ und in mehreren Arbeiten Sequenzinformationen publiziert.^{71,72,91,137} Sie zeigen, dass m⁶A eine hoch abundante, ubiquitär vorkommende Modifikation in humaner und muriner mRNA ist. Etwa 6000 bis 8000 humane mRNAs tragen durchschnittlich drei bis vier Methylierungen pro Transkript. Sowohl globale Methylierungsmuster als auch einzelne Loci sind von der Maus zum Menschen konserviert.^{71,72}

Als Konsensmotiv wurde in beiden Spezies R¹R²ACH verifiziert (R¹ = G/A; R² = G>A; H = U/C/A). Die Modifikation ist verstärkt innerhalb der *Coding sequence* (CDS), nahe des Stop Codons und innerhalb des 3'-UTR zu finden (**Abbildung 11**).

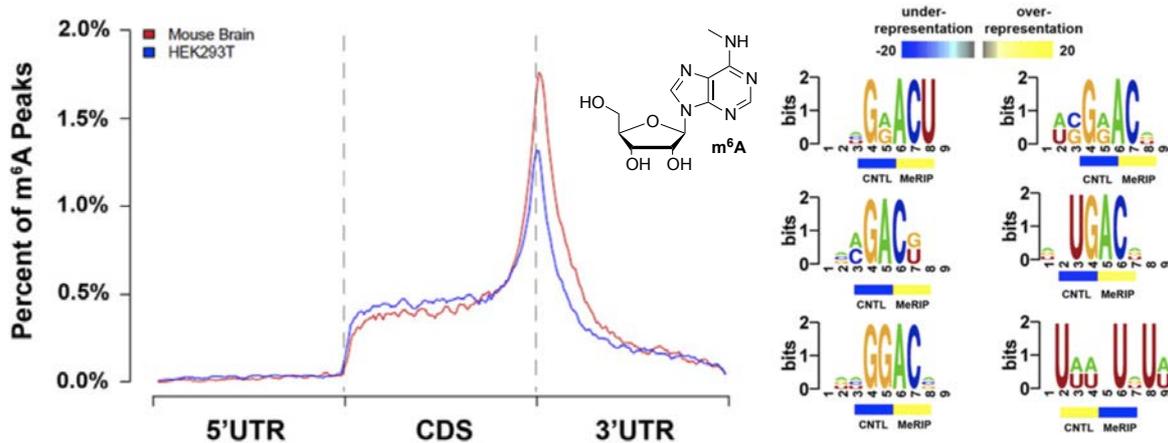


Abbildung 11: N^6 -Methyladenosin (m^6A) ist in humanen und murinen codierenden Transkripten innerhalb der CDS, nahe dem Stop Codon und innerhalb des 3'-UTR angereichert. Das dominante konservierte Konsensmotiv ist RRm^6ACH (Abbildung nach Meyer *et. al.*⁷²).

Als Faktoren, die den Grad der Methylierung steuern sind zum einen die Demethylasen FTO und ALKBH5 von Bedeutung. Sie entfernen die Methylgruppe in einem Fe^{II} und α -Ketoglutarat abhängigen Mechanismus, der zumindest für FTO schrittweise über hm^6A und f^6A als Intermediate abläuft.^{135,136,138,139} Im Gegensatz zu den oxidierten dC Derivaten hm^5dC , f^5dC und ca^5dC sind hm^6A und f^6A unter physiologischen Bedingungen nicht stabil und zerfallen hydrolytisch mit Halbwertszeiten von etwa 190 min.¹³⁸ Die Oxidation ist hier also ausreichend für die Entfernung der Methylgruppe. Die Methylierung wird durch einen Komplex aus zwei katalytisch aktiven Methyltransferasen METTL3 und METTL14 sowie einem regulativen Faktor WTAP und einem weiteren Faktor unbekannter Funktion (KIAA1429) eingeführt (Abbildung 12).¹⁴⁰⁻¹⁴³ Als Quelle der Methylgruppe dient *S*-Adenosylmethionin (SAM).

Eine biologische Funktion kann, neben einer Modulation der RNA Sekundärstruktur, nur über Interaktion mit spezifisch bindenden Proteinen (*Reader*-Proteine) stattfinden. Bislang sind sechs direkt mit m^6A interagierende Proteine bekannt: YTHDF1/2/3,¹⁴⁴ YTHDC1/2¹⁴⁵ und HNRNPA2B1¹⁴⁶. Die zur YTH Familie gehörenden Proteine verfügen alle über eine YTH-Domäne, die für die Bindung und Sequenzspezifität verantwortlich ist.^{145,147} Neben seiner Funktion als *Writer*-Protein könnte METTL3 auch eine Rolle als *Reader* einnehmen, womit es der siebte bekannte Interaktor wäre.¹⁴⁸

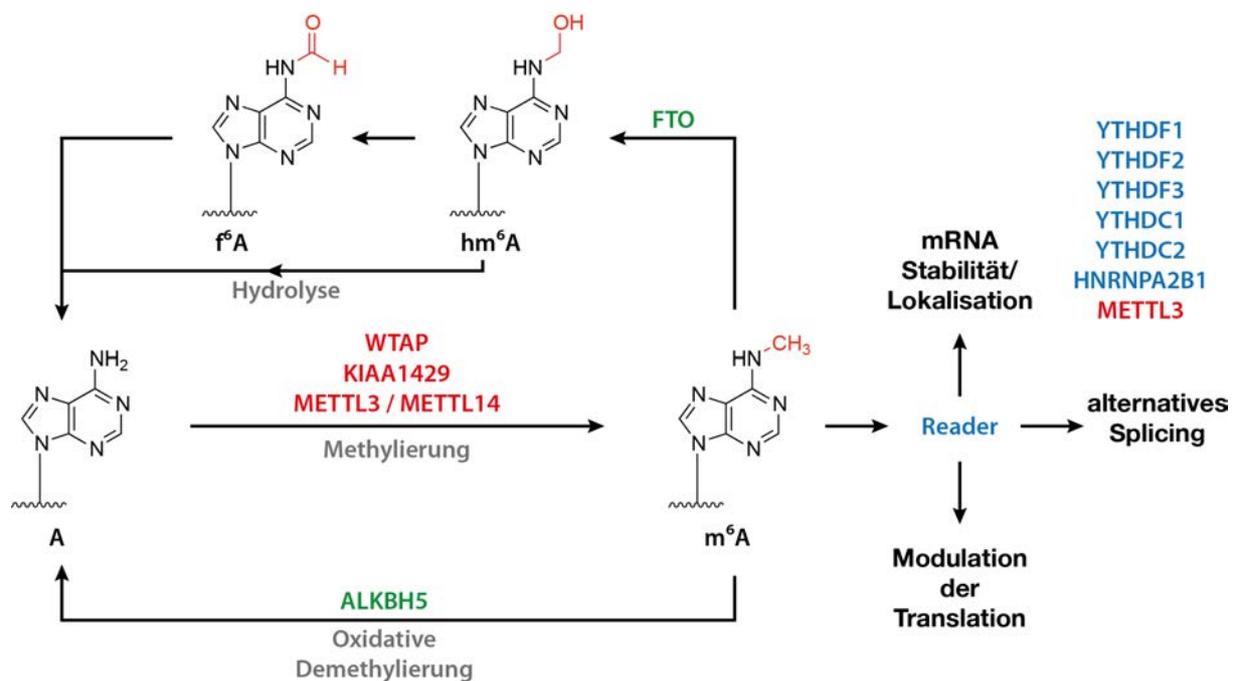


Abbildung 12: Übersicht über den bislang bekannten Metabolismus der Modifikation m⁶A: Ein Komplex aus mindestens vier Komponenten methyliert Adenosin zu m⁶A. Methyltransferaseaktivität besitzen hierbei METTL3 und METTL14, WTAP und KIAA1429 nehmen bislang nicht aufgeklärte Funktionen wahr. Die Methylgruppe kann durch FTO und ALKBH5 oxidativ entfernt werden, wobei die Intermediate hm⁶A und f⁶A bislang nur für FTO beobachtet werden konnten. Seine regulative Funktion entfaltet m⁶A über spezifische Interaktionen mit *Reader*-Proteinen, von denen bislang sieben bekannt sind: YTHDF1-3, YTHDC1-2, HNRNPA2B1 und vermutlich stellt auch METTL3 ein solches dar.

Die durch m⁶A gesteuerten regulativen Prozesse sind noch unvollständig verstanden, einige Grundmechanismen deuten sich aber an. Durch m⁶A markierte RNA wird durch YTHDF2 gebunden, dem aktiv translatierten mRNA Pool entzogen und zum Abbau in *P-bodies* relokalisiert.^{143,144,149} Eine wichtige Funktion von m⁶A ist also die gezielte Herabsetzung der Lebensdauer einzelner mRNAs, was unter anderem den Prozess der Stammzellendifferenzierung beeinflusst.¹⁴⁹⁻¹⁵² Neben der Differenzierung wird auch der circadiane Rhythmus durch m⁶A-Methylierung moduliert und m⁶A Level unterliegen selbst diesem Rhythmus.^{153,154} Dies scheint nicht nur auf den selektiven Abbau zurückzuführen zu sein, sondern auch von einer Beeinflussung des Kernexportes der ausschlaggebenden mRNAs abzuhängen.

Neben der Destabilisierung durch YTHDF2 kann eine verstärkte, YTHDF1-vermittelte Translation stattfinden.¹⁵⁵ Dabei lokalisiert YTHDF1 das Transkript zur Translationsmaschinerie und führt eine beschleunigte Translationsinitiation herbei. Die beiden konträr agierenden Proteine YTHDF1 und YTHDF2 haben sowohl voneinander verschiedene, als auch überlappende Zieltranskripte, was im Falle der von beiden gebundenen RNAs zu einer kurzen, aber intensiven Translation mit darauf folgendem Abbau führen würde.¹⁵⁵

In einem YTHDF2-abhängigen Mechanismus führt ein Hitzeschock in murinen Kulturzellen zu selektiver Methylierung im 5'-UTR von Hsp-codierenden mRNAs und einer darauf folgenden verstärkten Translationsinitiation in einem Cap-unabhängigen Mechanismus.¹⁵⁶ Meyer *et al.* führen den beobachteten Effekt auf eine direkte Interaktion zwischen eIF3 und dem methylierten Locus zurück¹⁵⁷, während Lin *et al.* eine Rekrutierung von eIF3 durch METTL3 an m⁶A Loci in der Nähe von Stop Codons beobachten konnten.¹⁴⁸ Es deutet sich also neben einer verstärkten Degradierung der mRNA auch eine Rolle von YTHDF2 in der Translationsinitiation an.

Neben nicht-codierenden Abschnitten der mRNA ist m⁶A auch innerhalb der CDS zu finden. Eine mögliche Auswirkung der Methylierung in diesem Bereich ist die Modulation der Dynamik der Translationselongation, was eine alternative cotranslationale Prozessierung der synthetisierten Proteine zur Folge haben könnte.¹⁵⁸ Diese Modulation beruht auf der Tatsache, dass m⁶A, analog zu A, mit U paart, die Bindung aber leicht geschwächt ist.^{159,160} Dadurch wird in einem weiteren Mechanismus die selektive Freigabe zunächst in Sekundärstrukturen eingebundener Bindemotive für RNA-interagierende Proteine möglich, also ein durch Methylierung aktivierter struktureller „Schalter“.^{161,162}

Dem m⁶A bindenden Protein YTHDC1 wird eine Funktion in der Regulation alternativen *Splicings* zugeschrieben.¹⁶³ Es wurde früh vermutet, dass m⁶A Splicing reguliert, nachdem gezeigt wurde, dass es bevorzugt in differenziell gespliceten Transkripten zu finden ist und ein *Knockdown* einzelner Komponenten des Methyltransferasekomplexes zu Variation im Verhältnis von Splicing-Isoformen führt.^{71,141} Zudem konnten Xiao *et al.* zeigen, dass ein Komplex aus YTHDC1 und dem serinreichen Splicingfaktor SRSF3 eine verstärkte Inklusion methylierter Exons bewirkt, indem er mit SRSF10 um die RNA-Bindung konkurriert.¹⁶³ Über die Funktion von YTHDF3 und YTHDC2 ist bislang nichts Genaues bekannt.

Zusammenfassend lässt sich also sagen, dass eine durch m⁶A modifizierte mRNA, in Abhängigkeit von der Position der Methylierung und vermutlich dem Sequenzkontext, sowohl relokalisiert und abgebaut wird, verstärkt und selektiv translatiert werden kann und sie darüber hinaus alternativ gespleißt sowie in ihrer Sekundärstruktur moduliert werden kann (**Abbildung 13**).

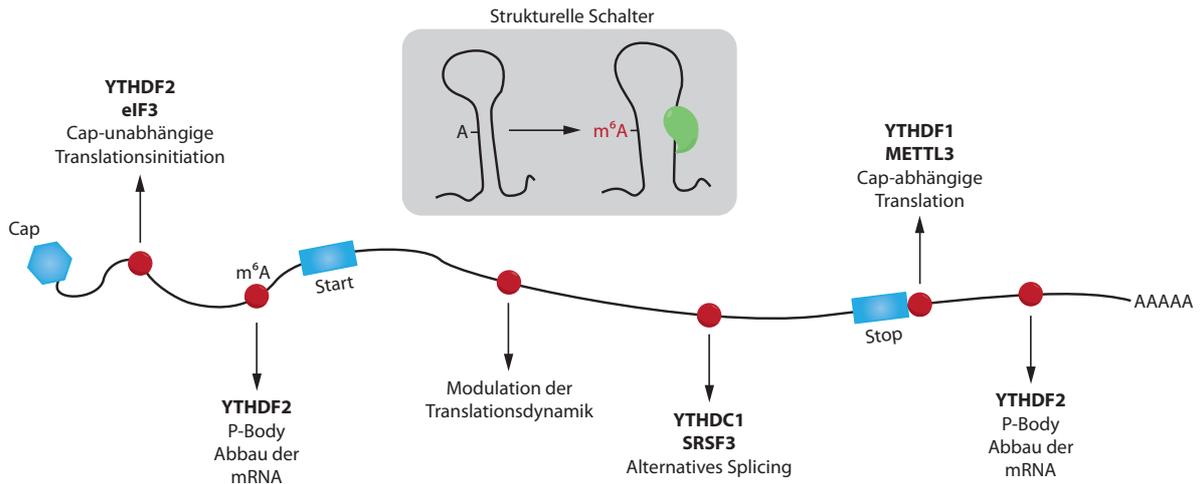
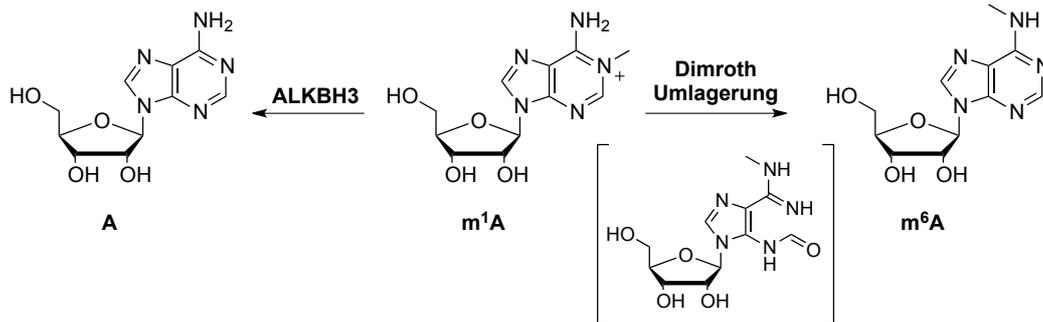


Abbildung 13: Überblick über bislang bekannte spezifische Binder und Funktionen von m^6A an unterschiedlichen Positionen eines Transkriptes in der Regulation des mRNA Metabolismus. Modifizierte Loci werden durch rote Punkte symbolisiert.

1.3.5 N^1 -Methyladenosin

N^1 -Methyladenosin (m^1A) wurde frühzeitig als Bestandteil von humaner tRNA und rRNA¹⁶⁴ identifiziert. Es besitzt bei physiologischem pH eine positive Ladung¹⁶⁵ und kann damit zu fundamentalen Änderungen der Struktur und Funktion der modifizierten Oligonukleotide führen.⁵¹ Seine chemische Stabilität ist geringer als die vieler anderer Basen, es reagiert in einer *Dimroth* Umlagerung (**Schema 1**) unter basischen Bedingungen zu N^6 -Methyladenosin (m^6A).¹⁶⁶

Als Modifikation in humaner mRNA ist m^1A erst 2016 identifiziert und lokalisiert worden.^{167,168} In Sequenzierungsstudien wurden in mRNA aus HEK293T und HeLa Zellen sowie murinem Gewebe relativ hohe, aber voneinander abweichende Mengen m^1A detektiert (887 bis 7154 Loci). Die Publikationen stimmen darin überein, dass m^1A in hoher Abundanz, aber geringer Dichte (meist ein m^1A pro Transkript) und stark gehäuft innerhalb des 5'-UTR nahe der Translationsinitiationsstelle und in der CDS bei der ersten *splice site* auftritt. Es scheint bevorzugt in GC-reichen und hoch strukturierten Segmenten vorzukommen, ein einzelnes, übereinstimmendes Konsensmotiv ist aber nicht identifizierbar. Das Vorhandensein eines m^1A korreliert mit einem erhöhten Proteinlevel, also einer verstärkten Translation.¹⁶⁷ Viele m^1A Loci sind von Mäusen zum Menschen konserviert¹⁶⁷ und die Modifikation unterliegt einer dynamischen Regulation.¹⁶⁸ Die Demethylase ALKBH3 setzt m^1A *in vitro* und *in vivo* zu A um, m^1A kann also aktiv demethyliert werden.^{168,169}



Schema 1: Die Modifikation m¹A kann unter basischen Bedingungen in einer Dimroth-Umlagerung zu m⁶A isomerisieren. Enzymatisch wird m¹A durch die α -Ketoglutarat/Fe^{II}-abhängige Oxidase ALKBH3 zu Adenosin demethyliert.

Damit deutet sich also auch für dieses methylierte Adenosinderivat eine regulative Funktion in codieren Transkripten an. Seine positive Ladung und hohe Abundanz in strukturierten RNA-Motiven können als Hinweis auf eine ausgeprägt strukturgebende Rolle gedeutet werden.

1.3.6 Inosin

Inosin wird durch *Adenosin deaminases acting on RNA* (ADARs) aus Adenosin durch hydrolytische Desaminierung in doppelsträngigen RNA Abschnitten erzeugt (*A-to-I editing*).¹⁷⁰ Inosin ist auch in DNA vorhanden, wird hier aber als stressbedingter Schaden gewertet und entsteht durch spontane hydrolytische Desaminierung.^{171,172} Da I als G gelesen wird¹⁷² ist es ein mutagener Schaden. Dies ermöglicht wiederum auf der RNA Ebene eine Funktion im *Re-coding*, also der posttranskriptionellen Umprogrammierung einzelner Codons. Eine ausgeprägte Korrelation zwischen ADAR-Expression und messbarem I-Gehalt in mRNA, zusammen mit großen gewebsspezifischen Unterschieden lassen eine überwiegend aktiv gesteuerte Desaminierung in mRNA naheliegend erscheinen.¹⁷³ Der Umfang in dem *A-to-I* Editierung im Transkriptom stattfindet ist höchst umstritten. Zahlreiche Studien zur Sequenzierung kommen zu stark voneinander abweichenden Ergebnissen.^{174,175}

Eine robuste Methode zur Lokalisierung von Inosin Loci im Transkriptom stellt die chemische Derivatisierung mit Acrylnitril dar. Diese Reaktion resultiert in einem Abbruch der reversen Transkription und ermöglicht damit eine präzise Sequenzierung.^{176,177} Auf diese Weise wurden in poly(A)-RNA aus humanem Hirngewebe nahezu 30000 editierte Loci identifiziert, die zu einem weit überwiegendem Teil in untranslatierten (3'-UTR, 5'-UTR, Introns) oder nicht annotierten Regionen zu finden sind. In Übereinstimmung mit anderen Publikationen¹⁷⁰ befinden sich die Loci häufig in *Alu* Sequenzen, also etwa 300 bp langen,

hoch abundanten, transposablen Elementen, deren Funktion oder Auswirkungen noch schlecht erforscht sind.¹⁷⁸ Es wird spekuliert, dass *A-to-I editing* in *Alus* den mRNA Stoffwechsel verändert indem es den Austausch einzelner Aminosäuren (*Re-coding*), veränderte Splicingmuster, Konkurrenz mit RNAi oder eine Retention im Nukleus verursacht.¹⁷⁹ Verschiedene Anhaltspunkte sprechen für eine dezidierte Funktion des *A-to-I editings* in Hirngewebe von Menschen und Primaten. So sind ADARs in Hirngewebe hoch exprimiert, in Korrelation mit hohen Inosinmengen in mRNA.¹⁷³ Die Editierung betrifft verstärkt Transkripte, die wichtige Proteine im Hirnstoffwechsel codieren.^{177,180}

Trotz der umfangreichen Forschungsarbeit zu Inosin in mRNA ist aber nach wie vor wenig über die molekularen Mechanismen und Dynamik des *A-to-I editings* bekannt. Inosin scheint aber die global am häufigsten vorkommende Modifikation in mRNA sein und verschiedene regulative Funktionen zu besitzen.

1.3.7 Die Isopentenylmodifikationen *N*⁶-Isopentenyladenosin und *N*⁶-Isopentenyl-2-thiomethyladenosin

1.3.7.1 Vorkommen

Neben den beschriebenen Modifikationen, die sich durch strukturell geringfügige Abweichungen von den kanonischen Basen, zumeist Methylierungen, auszeichnen gibt es eine Reihe erheblich komplexerer Derivate (siehe **Abbildung 8**), sogenannte Hypermodifikationen. Zu ihnen kann man neben Basen mit modifizierten Grundgerüsten, wie Queuosin (Q) und Wybutosin (W), die zweifach substituierte Isopentenylmodifikation *N*⁶-Isopentenyl-2-thiomethyladenosin (*ms*²*i*⁶A) zählen. Gemeinsam mit ihrem biosynthetischen Vorläufer *N*⁶-Isopentenyladenosin (*i*⁶A) und den biosynthetisch verwandten hydroxylierten Varianten *io*⁶A und *ms*²*io*⁶A bilden sie die Gruppe der Isopentenylmodifikationen (**Abbildung 14**).

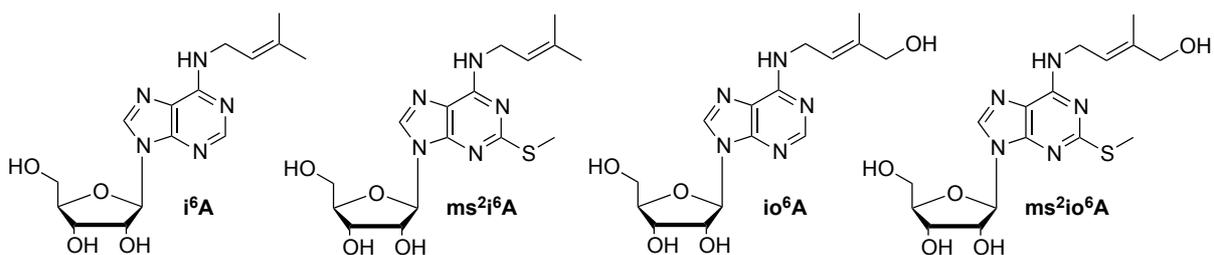


Abbildung 14: Strukturen der Bekannten Isopentenylmodifikationen *i*⁶A, *ms*²*i*⁶A und ihrer hydroxylierten Derivate *io*⁶A und *ms*²*io*⁶A.

Die Modifikationen i^6A und ms^2i^6A wurden ab 1966 als iPA und ms-iPA aus tRNA von *E. coli*¹⁸¹⁻¹⁸³ und *S. cerevisiae*^{184,185} isoliert. Ging man zunächst davon aus, dass das exklusive Vorkommen einer der beiden Modifikationen kennzeichnend für die Domänen ist, mit i^6A als eukaryotischem und ms^2i^6A als prokaryotischem Derivat, so wurde ms^2i^6A später auch in Mitochondrien gefunden.^{36,186,187} Interessanterweise ist dieses Verhältnis auch in pflanzlichen Zellen zu finden, wo ebenfalls i^6A in zytosolischen tRNAs und ms^2i^6A in tRNAs der Chloroplasten zu finden sind.^{186,188,189} Dies deutet auf einen evolutionär alten, bakteriellen Ursprung und anschließenden Transfer mittels Endosymbiose hin. Die hydroxylierten Derivate io^6A und ms^2io^6A wurden in Enterobakterien, pflanzenassoziierten Bakterien und Pflanzen selbst isoliert.¹⁹⁰⁻¹⁹⁴ Im Weiteren soll vor allem auf die beiden in dieser Arbeit untersuchten Modifikationen i^6A und ms^2i^6A eingegangen werden.

Beide kommen in tRNA ausschließlich an Position 37 vor, welche sich unmittelbar 3' des Anticodons befindet und als semikonservierte Position ausschließlich mit Purinen belegt ist.²² ms^2i^6A ist dabei in der Mehrheit aller tRNAs zu finden, die mit U beginnende Codons auslesen, wozu die tRNAs für Phenylalanin, Tryptophan, Tyrosin, Cystein, Leucin und Serin gehören.¹⁸⁷ Die tRNA^{Ser(1,IV)} stellt hierbei eine Ausnahme dar und ist nicht modifiziert.¹⁹⁵

Veronika Reiter konnte im Zuge ihrer Dissertation zeigen, dass ms^2i^6A in humanen Kulturzellen (HeLa, HMEC und MCF-7) eine Modifikation der poly(A)-RNA ist, welche über für diese RNA-Spezies geeignete Aufreinigungsmethoden aus Gesamt-RNA angereichert werden kann.^{109,196} Die absolute Quantifizierung von Modifikationen in mRNA wird erschwert durch den geringen Anteil von codierenden Transkripten an zellulärer Gesamt-RNA. Modifikationen, die in hoher Abundanz in anderen RNA-Spezies, wie rRNA vorhanden sind, können schwer abtrennbare Kontaminationen darstellen. Eine Möglichkeit dennoch zuverlässig beurteilen zu können, ob eine Modifikation in mRNA vorhanden ist, stellt die Betrachtung relativer Quantitäten dar.

Die durch *Veronika Reiter* in HeLa detektierten Mengen sind absolut und relativ in **Abbildung 15** dargestellt. Es wird deutlich, dass sowohl i^6A , als auch ms^2i^6A in geringen Mengen vorhanden sind, das Thiomethylderivat aber neben m^6A die einzige Modifikation darstellt, deren Menge durch poly(A)-Anreicherung ansteigt.

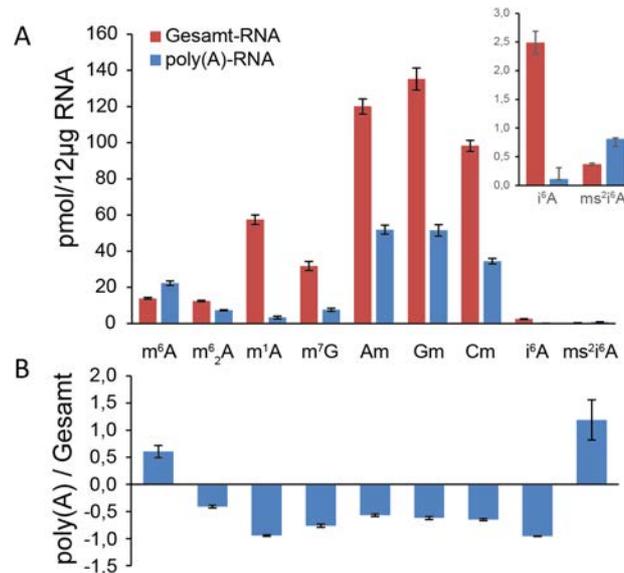


Abbildung 15: Massenspektrometrisch quantifizierte Modifikationen in HeLa Gesamt-RNA und oligo(dT)-angereicherter poly(A)-RNA. A) Absolutwerte pro 12 µg isolierter RNA. B) Relative Veränderung der Werte durch poly(A)-Anreicherung. (Daten: *Veronika Reiter*; Abbildung: *Markus Müller*).

Damit verhält es sich gegenläufig zu seiner Vorstufe i⁶A, die stark abgereichert wird. Im Einklang hiermit finden *Brückl et al.* i⁶A, nicht aber ms²i⁶A, in tRNA humaner Kulturzellen (HeLa, HCT-116).⁶⁰ In kleinen RNA-Spezies (70-100 nt) detektieren *Kellner et al.* ein analoges Verhältnis in humanem und porcinem Hodengewebe.¹⁹⁷ Darüber hinaus konnten die Verhältnisse durch *Sarah Schiffers* und *Katharina Iwan* bestätigt werden (**Abbildung 16**).

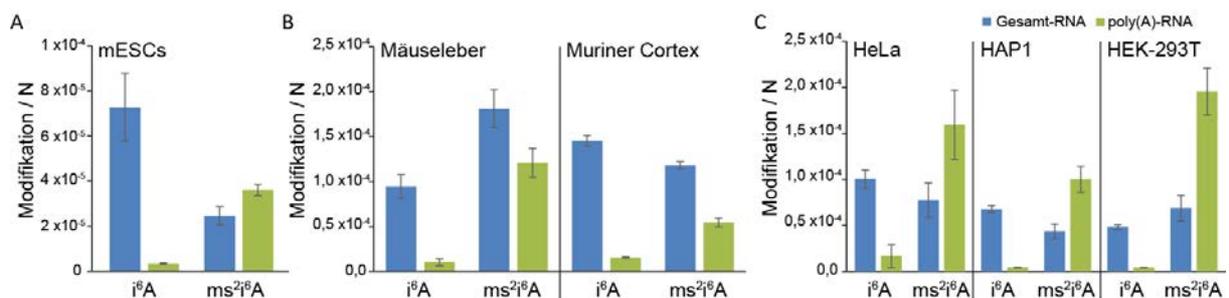


Abbildung 16: Massenspektrometrische Quantifizierung von i⁶A und ms²i⁶A in RNA aus humanen und murinen Quellen. A) Aus murinen embryonalen Stammzellen (mESCs). B) Aus Mausegeweben. C) Aus humanen Kulturzellen. Daten: *Katharina Iwan* und *Sarah Schiffers*. Abbildung: *Markus Müller*.

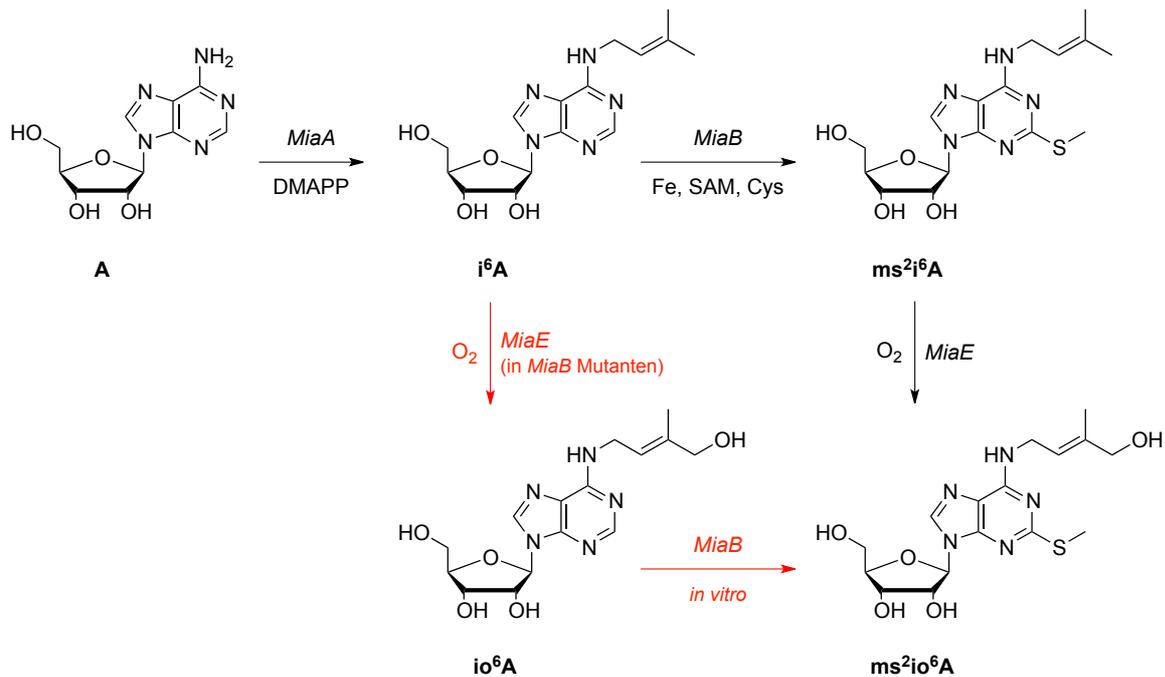
In mRNA aus humanen Kulturzellen (HeLa, HAP und HEK293T) ist deutlich mehr ms²i⁶A in poly(A)-RNA zu finden, als in Gesamt-RNA. Murine Stammzellen besitzen eine weniger ausgeprägte Differenz, die qualitativ aber noch besteht. In Leber- und Hirngewebe aus Mäusen dagegen scheinen andere RNA-Spezies bedeutende Quellen der Modifikation zu sein. Tatsächlich ist ms²i⁶A in muriner mitochondrialer tRNA beschrieben.¹⁹⁸ Aus **Abbildung 16 C** geht auch hervor, dass eine poly(A)-Anreicherung aus murinen Geweben eine erheblich

größere Reduktion der i^6A -Mengen verursacht, als das für sein Thiomethylderivat der Fall ist, was durch ein Vorhandensein in mRNA erklärbar ist.

Aus den geschilderten Daten lässt sich schlussfolgern, dass ms^2i^6A eine humane und vermutlich murine mRNA-Modifikation darstellt, während i^6A eine reine tRNA-Modifikation ist. Darüber hinaus scheint tRNA aus Mausgewebe signifikante Mengen ms^2i^6A zu enthalten, während dies für murine Kulturzellen weniger ausgeprägt und im Fall humaner Kulturzellen gar nicht der Fall ist. Damit stellt es eine Modifikation der *messenger RNA* dar, die in ihrer chemischen Komplexität bislang einzigartig ist.

1.3.7.2 Biosynthese

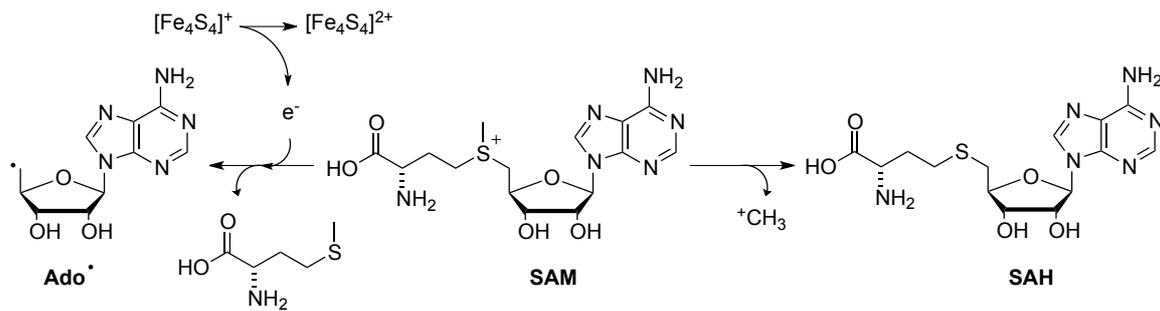
Die Biosynthese der Isopentenylmodifikationen ist gut erforscht, die einzelnen Schritte und jeweils verantwortlichen Enzyme sind in **Schema 2** dargestellt.¹⁹⁹ Die Biosynthese geht vom Adenosin aus, welches von dem bakteriellen Enzym MiaA (MiaA: *Methylthio-isopentenyl-adenosine synthesis enzyme A*; human: Trit1 (*tRNA isopentenyl transferase 1*); Hefe: Mod5; allgemein auch DMATase) zu i^6A umgesetzt wird. Es existieren Kristallstrukturen mehrerer bakterieller^{200,201} und eines Enzyms aus *S. cerevisiae*.²⁰² Zunächst wird der zu modifizierende RNA-Hairpin gebunden und so verformt, dass die Nukleotide 33-35 und 37 herausgedreht werden und insbesondere A_{37} für die Reaktion zugänglich wird. Anschließend wird Dimethylallylpyrophosphat (DMAPP) gebunden, welches als Prenyldonor fungiert, und die Isopentenylgruppe wird unter Abspaltung von Pyrophosphat auf die gebundene RNA übertragen.^{203,204} Das Pyrophosphat koordiniert dabei ein Mg^{2+} -Ion.



Schema 2: Biosynthese der Modifikationen i^6A , io^6A , $ms^{2:i^6}A$ und $ms^{2:io^6}A$ in Bakterien. Schwarze Pfeile kennzeichnen den natürlichen Syntheseweg, rote einen künstlich herbeigeführten.¹⁹⁹

In einer Folgereaktion wird nun die Thiomethylgruppe an Position 2 des Puringerüsts eingeführt. Hierfür ist das bakterielle Enzym MiaB, bzw. das humane Cdk5Rap1 (*Cdk5 regulatory subunit-associated protein 1*), verantwortlich.^{205,206} Sie gehören zur Familie der Radikal-SAM-Enzyme (SAM = *S*-Adenosylmethionin) und übertragen, mit Hilfe zweier [4Fe-4S] Cluster, die Thiomethylgruppe auf das Substrat.²⁰⁷⁻²⁰⁹ Einer dieser Cluster wird hierbei von einem CysX₃CysX₂Cys-Motiv koordiniert, während der zweite von drei N-terminalen Cysteinen (Cys10, Cys46, Cys79) koordiniert wird. Pierrel *et al.* schlussfolgern aus ihrer Beobachtung, dass in Abwesenheit von MiaA weder i^6A , noch $ms^{2:i^6}A$ detektierbar sind, dass lediglich i^6A als Substrat für MiaB dienen kann.²¹⁰ Da das Produkt einer Umsetzung von A durch MiaB aber ms^2A wäre, welches in ihrer Studie nicht betrachtet wurde, ist diese Aussage kritisch zu hinterfragen.

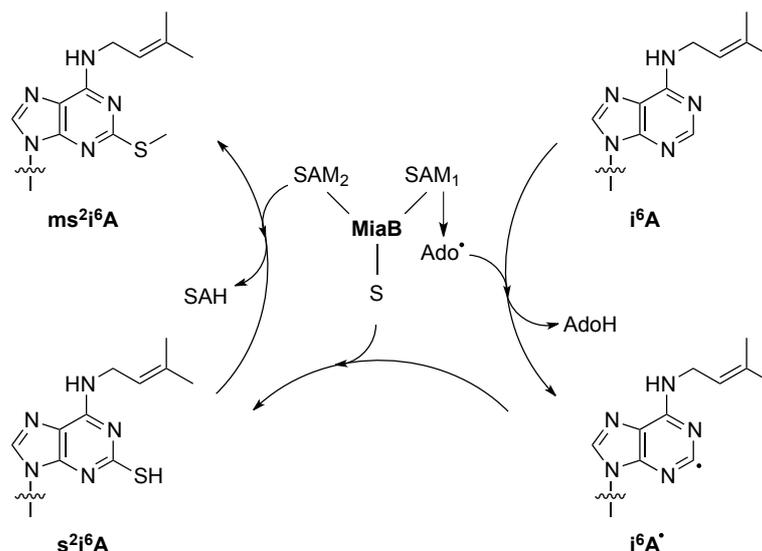
Der Mechanismus der Thiomethylierungsreaktion wurde bereits vielfach untersucht, konnte bislang aber nicht restlos aufgeklärt werden. Im Lauf der Reaktion werden mindestens zwei Äquivalente SAM benötigt. SAM kann in dieser enzymatischen Reaktion als „Radikalstarter“ (Ado[•]-Donor) bzw. als Quelle einer elektrophilen Methylgruppe (⁺CH₃-Donor) dienen, wobei zwei unterschiedliche Abbauprodukte erzeugt werden (**Schema 3**).²¹¹



Schema 3: Mögliche Reaktionswege, über die *S*-Adenosylmethionin (SAM) als Co-Substrat an der enzymatischen Reaktion von MiaB oder Cdk5Rap1 teilnimmt. Es kann als elektrophiler Methylondonor zu *S*-Adenosylhomocystein (SAH) reagieren, oder reduktiv in Methionin und ein 5'-Deoxyadenosylradikal gespalten werden.

Die Umsetzung von i^6A zu ms^2i^6A erzeugt beide Spezies, wobei etwa ein Äquivalent AdoH und deutlich mehr als ein Äquivalent SAH (2.0 bis 3.4 eq.) entstehen.^{209,212} Daraus lässt sich zunächst schließen, dass die Reaktion sowohl einen radikalischen Schritt, als auch die Übertragung einer Methylgruppe mittels Substitution einschließt. Die Entstehung von mehr als einem Äquivalent SAH erscheint überraschend, *Landgraf et al.* konnten aber zeigen, dass SAH auch in Abwesenheit eines RNA-Substrates entsteht und die Methylgruppe auf das Protein MiaB übertragen wird.²¹² In Anwesenheit eines geeigneten nukleophilen Akzeptors wird das Enzym demethyliert. Experimente mit d_3 -SAM konnten zeigen, dass das Enzym mit einer Methylgruppe „beladen“ werden kann, und diese auf sein RNA-Substrat überträgt, bevor in der Reaktion anwesendes SAM als Methylondonor herangezogen wird. Hieraus kann die Vermutung abgeleitet werden, dass ein intermediärer Akzeptor in MiaB existiert, der die Übertragung der Methylgruppe auf das Substrat vermittelt. Unter Bedingungen, die eine Nebenreaktion, z. B. eine Hydrolyse des methylierten Intermediates, erlauben, kann also eine von der eigentlichen Reaktion unabhängige Entstehung von SAH stattfinden, was die unerwartet große Menge entstehenden SAHs erklärt.

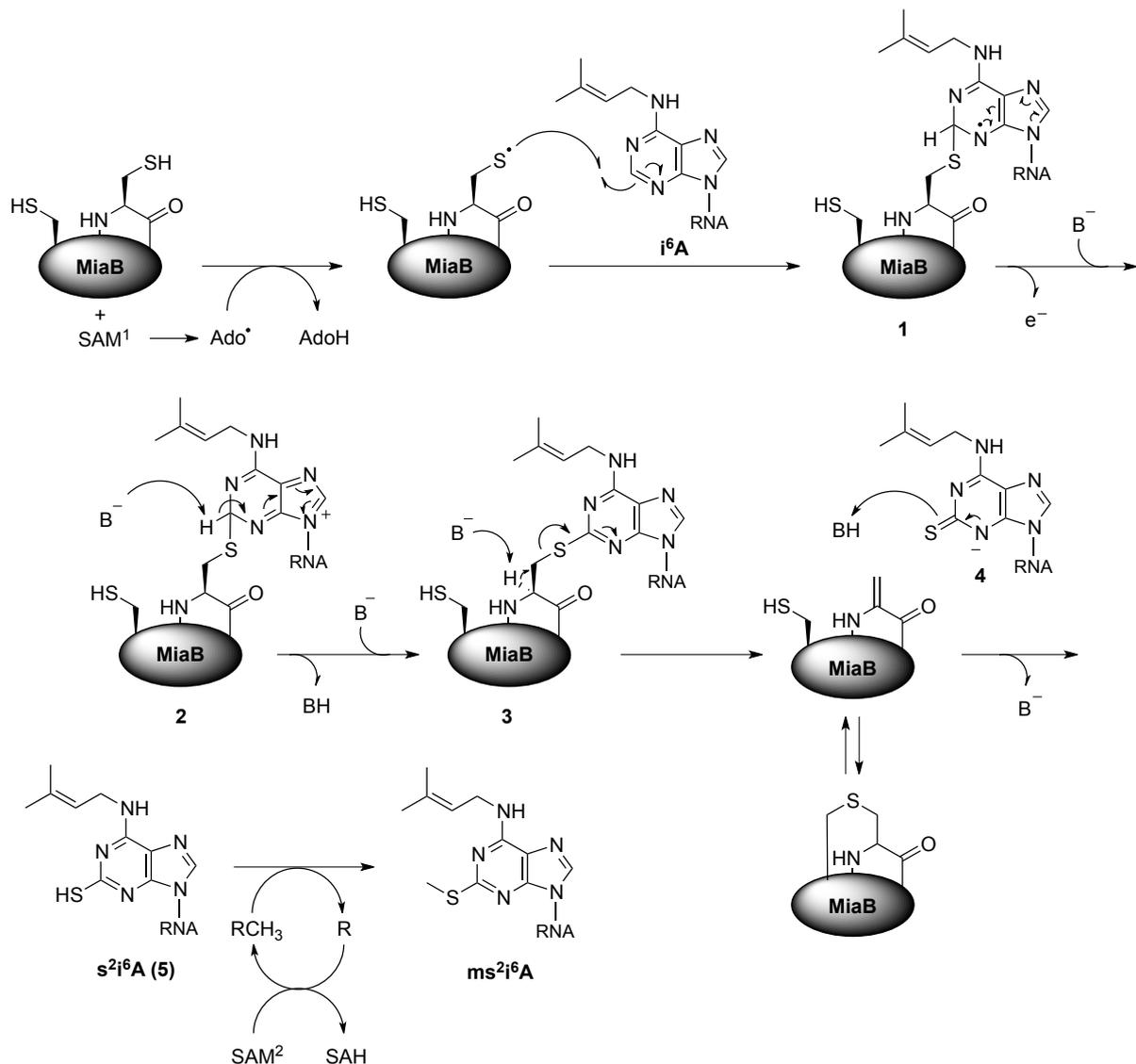
Der radikalische Schritt der Reaktion muss als erster erfolgen und die Übertragung des Schwefels beinhalten. Ein ursprünglich von *Pierrel et al.* vorgeschlagener Mechanismus²¹⁰ ist in **Schema 4** dargestellt und beinhaltet die Abstraktion eines H-Radikals an Position 2 des i^6A -Substrates.



Schema 4: Durch *Pierrel et al.* vorgeschlagener Mechanismus für die durch MiaB katalysierte Thiomethylierung von i^6A zu $ms^{2:6}i^6A$.²¹⁰

Das in diesem Mechanismus entstehende Adenosin-C2-Radikal ist aber schlecht stabilisiert und daher ein unwahrscheinliches Intermediat. Eine theoretische Studie durch *Hioe* und *Zipse*²¹³ weist darauf hin, dass eine solche Abstraktion mit $\Delta H = +30.9 \text{ kJmol}^{-1}$ deutlich endotherm und damit unwahrscheinlich wäre. Die Erzeugung eines Cysteinradikals wäre energetisch erheblich günstiger und der Verlauf der Reaktion über ein N3-Radikal, analog zur direkten Methylierung am C2 und in Anlehnung an den hierzu von *Booker et al.*²¹⁴ vorgeschlagenen Mechanismus, erscheint wahrscheinlicher.²¹³ Aus den geschilderten Überlegungen lässt sich ein hypothetischer Mechanismus ableiten, der in **Schema 5** dargestellt ist und zuerst von *Dorothea Matschkal* vorgeschlagen wurde.²¹⁵

In diesem wird zunächst unter Verbrauch eines Äquivalentes SAM ein Cysteinradikal erzeugt, welches an Position 2 des Nukleosids i^6A angreift und ein mesomeriestabilisiertes N3-Radikal (**1**) erzeugt. Durch eine Oxidationsreaktion, in welcher der zweite (auxiliäre) FeS-Cluster als Elektronenakzeptor dienen könnte und anschließende Deprotonierung entsteht das kovalent mit dem Protein verknüpfte Thioadenosinderivat **2**. Dieses wird durch basenkatalysierte Eliminierung vom Protein gelöst (**3**), um nach einer Reprotonierung $s^{2:6}i^6A$ (**4** zu **5**) zu erhalten. Dieses wird im abschließenden Schritt durch ein zweites SAM-Molekül elektrophil methyliert, wodurch das Endprodukt $ms^{2:6}i^6A$ entsteht. Das als Schwefelquelle dienende Cystein wird in der Reaktion in ein Dehydroalanin, bzw. einen Thioether mit einem benachbarten Cystein überführt.



Schema 5: Von Dorothea Matschkal vorgeschlagener Mechanismus für die durch MiaB katalysierte Thioethylierung von i^6A zu ms^2i^6A .²¹⁵ B = Base; R = unbekannter, intermediärer Methylakzeptor.

Im Einklang mit diesem Mechanismus ist bekannt, dass als Schwefelquelle in der Reaktion nicht SAM, sondern das Protein selbst dient.²¹⁰ In Abwesenheit eines exogenen Schwefeldonors (wie Na₂S oder NaSCH₃), katalysiert das rekonstituierte Enzym lediglich eine Umsetzung²¹², in seiner Anwesenheit *in vitro* aber deutlich mehr.²⁰⁹ Das hypothetische Intermediat s^2i^6A wurde bislang noch nicht nachgewiesen, auch wenn *Agris et al.* Hinweise auf seine Existenz fanden.²¹⁶

In Analogie zum Mechanismus für das paraloge Enzym RimO, welches die Konversion von Aspartat zu 2-Thioaspartat katalysiert, ist alternativ auch denkbar, dass durch den auxiliären FeS-Cluster koordinierter exogener Schwefel die Quelle des 2-Thiosubstituenten darstellt.²⁰⁹ Dies würde erklären, warum bei Zusatz von Selenidquellen (MeSe⁻, PhSe⁻)

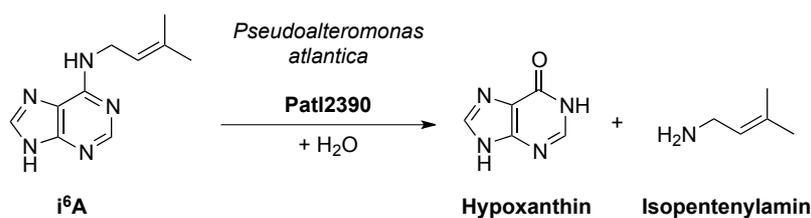
massenspektrometrisch kein Selenocystein im Protein MiaB detektierbar war (Experimente durchgeführt von *Martin Rossa*). Ein solches Selenocystein wäre zu erwarten, wenn eine Reaktion des Enzyms über eine Dehydroalanin-Zwischenstufe verlief, welche als *Michael*-System gegenüber starken Nukleophilen, wie PhSe^- reaktiv sein müsste.

Abschließend lässt sich also sagen, dass die von MiaB, bzw. Cdk5Rap1 katalysierte Reaktion nach einem bislang nicht restlos geklärten Mechanismus abläuft, in welchem in zwei Schritten zunächst Schwefel in die C2-H-Bindung des Substrates insertiert wird und anschließend eine Methylgruppe übertragen wird. Ein C2-Radikal (**Schema 4**) erscheint als Intermediat unwahrscheinlich.

Ausgehend von $\text{ms}^2\text{i}^6\text{A}$ kann in Bakterien durch Hydroxylierung $\text{ms}^2\text{io}^6\text{A}$ gebildet werden. Die Reaktion wird unter aeroben Bedingungen durch das MiaE katalysiert, welches in einigen Bakterienstämmen, wie *Salmonella thyphimurium* und *Pseudomonas aeruginosa*, nicht aber in *E. coli* existiert.^{199,217} Im Einklang hiermit kann *E. coli* kein $\text{ms}^2\text{io}^6\text{A}$ nachgewiesen werden.¹⁹⁴ MiaE ist eine Nicht-Häm-Zwei-Eisen-Oxygenase, die unter anaeroben Bedingungen inaktiv ist.^{218,219} Als Substrate können sowohl i^6A , als auch $\text{ms}^2\text{i}^6\text{A}$ dienen, $\text{ms}^2\text{i}^6\text{A}$ wird aber stark bevorzugt und gilt als natürliches Substrat.^{217,220,221} Der Mechanismus der Reaktion konnte noch nicht aufgeklärt werden, auch wenn erste Studien zu kinetischen Parametern publiziert wurden.^{222,223}

Eine phylogenetische Verteilung der für die Biosynthese verantwortlichen Enzyme ist in **Abbildung 17** dargestellt.¹⁹⁹ Alle analysierten Taxa verfügen über MiaB-ähnliche Gene, nicht aber über MiaA-Homologe, was durch die ubiquitäre Verbreitung von ms^2 -Modifikationen erklärt werden kann, die nicht zwingend Isopentenylsubstituenten tragen. *Anantharaman et al.* schließen aus der Verteilung, dass MiaB-Paralogue durch evolutionär frühe Duplikationen entstanden und sich zu Thiomethyltransferasen für unterschiedliche ms^2 -Modifikationen, wie $\text{ms}^2\text{t}^6\text{A}$, $\text{ms}^2\text{i}^6\text{A}$ oder $\text{ms}^2\text{m}^6\text{A}$ diversifiziert haben.²²⁴ Die geringe Verbreitung von $\text{ms}^2\text{io}^6\text{A}$ spiegelt sich in der ebenfalls geringen Verbreitung von MiaE-Homologen wider.

Tatsächlich konnte *MacLennan* 1975 zeigen, dass in bovinem Zellextrakt eine enzymatische Aktivität zu existieren scheint, welche den Isopentenylsubstituenten einer i^6A -modifizierte tRNA entfernt.²³⁴ Die Identität der Produkte konnte nicht bestimmt werden, aber durch radioaktive Markierungen konnte gezeigt werden, dass die Isopentenylgruppe von der Base getrennt wurde. Weiterhin wurde in einer aktuelleren Publikation eine bakterielle Cytokinindeaminase (Pat12390) identifiziert, die i^6A zu Hypoxanthin hydrolysiert (**Schema 6**).²³⁵ Als Substrat getestet wurde lediglich die freie Base, nicht das Nukleosid. Neben i^6A können auch abweichend N6-substituierte Adenine verwertet werden, unsubstituiertes Adenin, oder Adenosin sind aber schlechte Substrate. Unter den untersuchten Verbindungen stellt i^6A das beste Substrat dar, was als Hinweis darauf gedeutet werden kann, dass es das natürliche Substrat darstellt. Die Autoren deuten die beobachtete Aktivität als Möglichkeit Cytokinine durch Desaminierung zu Hypoxanthin dem *Purinnukleotid-Salvage-Pathway* zuzuführen, also nicht in erster Linie als Möglichkeit zur funktionellen Inaktivierung der Substrate.

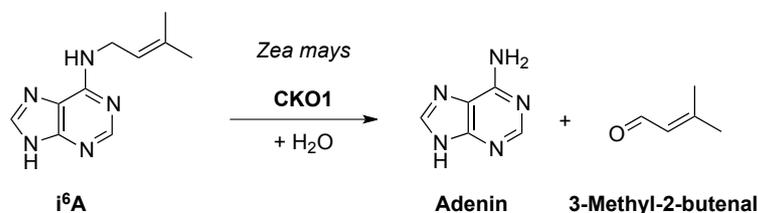


Schema 6: Von der Deaminase Pat12390, isoliert aus *Pseudoalteromonas atlantica*, katalysierte Hydrolyse von freiem i^6A zu Hypoxanthin.²³⁵

Hall et al. gelang es, aus Hühner-Knochenmark ein Enzym teilweise aufzureinigen, welches i^6A zu Inosin desaminieren kann und eine Präferenz für das Nukleosid gegenüber der freien Base aufweist, darüber hinaus aber auch Adenosin effektiver desaminiert als i^6A , was das Isopentenylderivat als unwahrscheinliches Substrat erscheinen lässt.²³⁶

In einer metabolischen Studie konnten *Chheda* und *Mittelman* zeigen, dass intravenös oder oral verabreichtes, radioaktiv markiertes i^6A Nukleosid im Menschen innerhalb weniger Stunden zu einer Reihe unterschiedlicher Metaboliten abgebaut wird.²³⁷ Darunter befand sich auch Adenin, welches über einen alternativen Mechanismus entstehen muss. Diese Beobachtung muss nicht auf einen dezidierten Metabolismus für i^6A zurückzuführen sein, sondern kann das Resultat unspezifischer enzymatischer Degradierung des in höheren Mengen verabreichten i^6A sein.

Neben den beschriebenen Hinweisen auf enzymatische Desaminierung von i^6A zu Inosin, bzw. Hypoxanthin sind in Pflanzen Cytokinin-Oxidasen bekannt, welche freies i^6A zu Adenin und 3-Methyl-2-butenal hydrolysieren²³⁸ (**Schema 7**) und sowohl Adenosin, als auch verschiedene Adenosinderivate sind bekannte Metabolite des Abbaus von i^6A in Tabakpflanzen.^{239,240}



Schema 7: Von der Cytokinin-Oxidase CKO1 (*Zea mays*) katalysierte Reaktion von i^6A zu Adenin und 3-Methyl-2-butenal.

Über eine analoge Umsetzung von ms^2i^6A ist nichts bekannt. Das Produkt einer solchen oxidativen Deaminierung wäre ms^2A , welches als tRNA-Modifikation bereits früh mittels Dünnschichtchromatografie in *Bacillus subtilis* und *Bacillus stearothermophilus* nachgewiesen wurde.^{241,242} In einer aktuelleren Studie gelang es auch *Kellner et al.* ms^2A in kleinen RNA-Spezies (70-100 nt) aus *E. coli*, Schwein und dem Menschen zu detektieren.¹⁹⁷ Es ist anzunehmen, dass diese Isolation zum überwiegenden Teil aus tRNA besteht. Nachdem keine Thiomethyltransferase bekannt ist, die A in ms^2A umwandelt, kann aus den geschilderten Indizien die Hypothese abgeleitet werden, dass die geringen Mengen detektierbaren ms^2A s das Resultat einer Metabolisierung einer ms^2x^6A Modifikation in RNA darstellt. Alle bekannten RNA Adenosinmodifikationen, welche einen Thiomethylsubstituenten an Position 2 tragen und somit als Substrat für eine solche Umsetzung in Frage kämen sind in **Abbildung 18** dargestellt.^{36,37} Im Menschen wurden bislang nur ms^2i^6A und ms^2t^6A nachgewiesen und die sie erzeugenden Thiomethyltransferasen identifiziert.^{196,198,243}

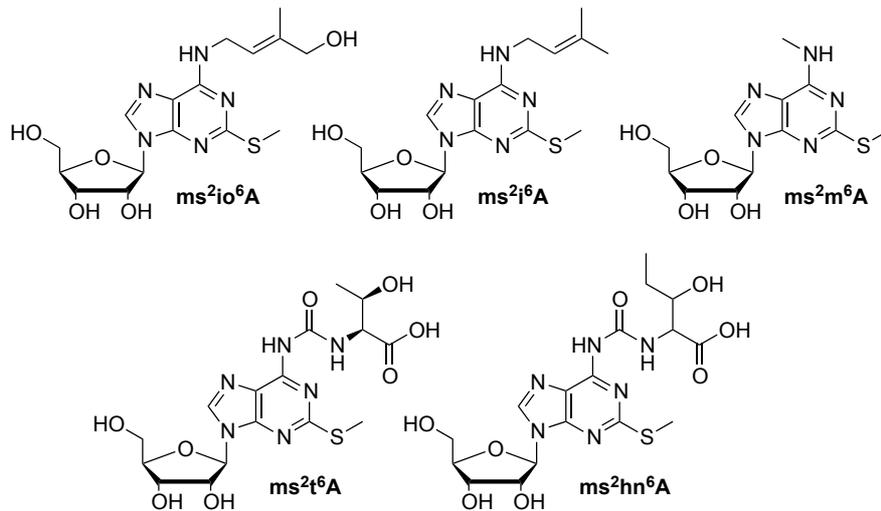


Abbildung 18: Bekannte ms^2A -Modifikationen aus allen drei Domänen des Lebens.

Ein gezielter Abbau der Modifikationen i^6A und ms^2i^6A ist in Eukaryoten bislang also nicht belegbar, die hierfür nötigen enzymatischen Reaktivitäten sind in anderen Domänen aber bekannt und ihre Existenz im Menschen ist denkbar. Die Produkte eines solchen Abbaus von i^6A , Adenosin und Inosin, sind aufgrund ihrer ubiquitären Verbreitung analytisch schwerer nachzuerfolgen, als die aus ms^2i^6A hypothetisch hervorgehenden Metaboliten ms^2A oder ms^2I . Experimentelle Evidenz existiert für ms^2A , bislang aber nicht für ms^2I . Die Möglichkeit einer regulierten Inaktivierung von ms^2i^6A oder i^6A durch eine Entfernung des Isopentenylsubstituenten stellt eine interessante, bislang wenig beachtete Facette der biologischen Wirkung dieser Modifikationen dar.

1.3.7.4 Funktion in der Translation

Wie bereits erläutert reichen die Funktionen der tRNA Modifikationen von der Festlegung ihrer Identität, über die korrekte Aminoacylierung und Faltung, beziehungsweise Stabilisierung der Sekundärstruktur bis hin zu einer Beeinflussung der Codon-Anticodon-Interaktion. Die unmittelbare Nähe der Position 37 zum Anticodon der tRNA legt die Vermutung nahe, dass hier zu findende Modifikationen einen Einfluss auf die Basenpaarung während der Translation haben. Für die Isopentenylmodifikationen i^6A und ms^2i^6A konnte in der Tat bereits früh gezeigt werden, dass sie schwache A-U Wechselwirkungen stabilisieren können.²⁴⁴ In einer NMR Studie konnten *Cabello-Villegas et al.* zeigen, dass ein i^6A modifizierter tRNA^{Phe} *Hairpin* seine Konformation zwar nur geringfügig ändert, der Isopentenylsubstituent aber einen disruptiven Einfluss auf die im unmodifizierten *Hairpin* vorliegende $U_{33} \cdot A_{37}$ Basenpaarung hat²⁴⁵, ein Effekt, der analog auch für t^6A_{37} gezeigt werden

konnte.²⁴⁶ Hierdurch wird eine Flexibilisierung der Struktur bedingt, die wohl eine bessere Codon-Anticodon Paarung zur ermöglicht. Laut *Wilson* und *Roe*, beeinflusst $ms^{2:6}A$ die Aminoacylierung der *E. coli* tRNA^{Phe} nicht, spielt für die Festlegung der Identität also keine Rolle, stabilisiert aber Wechselwirkungen zwischen Codon und Anticodon.²⁴⁷ Analog hierzu vermuteten *Diaz* und *Ehrenberg*, dass es einen stabilisierenden Einfluss auf korrekte Codon-Anticodon Paarungen hat, während es inkorrekte Paarungen destabilisiert und damit *Misreading* in der Translation reduziert.²⁴⁸ Neuere Ergebnisse zeigen, dass die Thiomethylgruppe des $ms^{2:6}A$ durch *Stacking* mit der ersten Base des auszulesenden Codons zu einer erheblichen Stabilisierung der Wechselwirkungen zwischen tRNA und mRNA in allen drei Bindestellen des Ribosoms führt und damit zu einer Stabilisierung des Leserahmens führt (**Abbildung 19**).⁴²

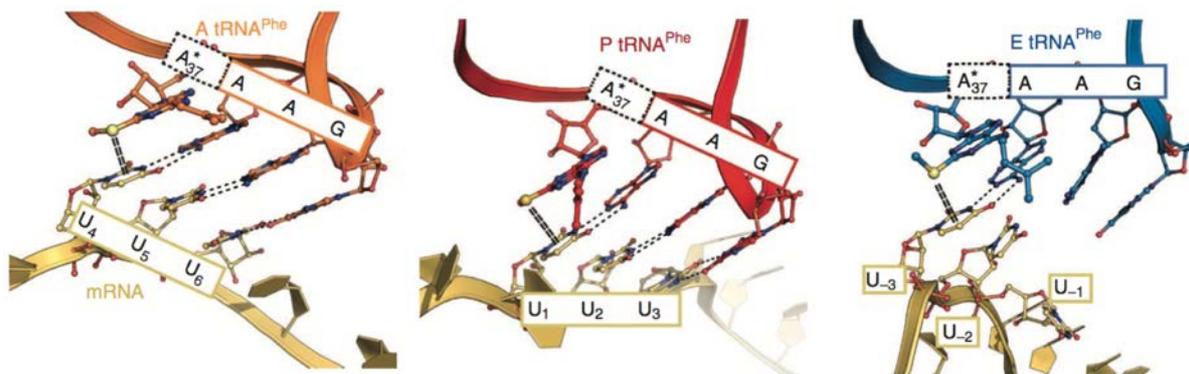


Abbildung 19: Die Modifikation $ms^{2:6}A$ stabilisiert die Paarung zwischen Codon und tRNA während der Translation in allen drei Bindestellen des Ribosoms (A, P und E). Ausschlaggebend hierfür ist ein *Stacking* zwischen dem Thiomethylsubstituenten und der ersten Base des Codons.⁴²

Nimmt man die geschilderten Ergebnisse zusammen, so lässt sich sagen, dass beide Substituenten über die Kombination zweier unterschiedlicher Auswirkungen zu dem beobachteten Resultat führen. Zunächst ist der Isopentenylsubstituent nötig um das den Anticodon umgebende $U_{33} \cdot A_{37}$ Basenpaar zu brechen und dem Anticodon *Hairpin* die nötige Flexibilität zu verleihen damit in einem zweiten Schritt die ms^2 -Gruppe eine stark stabilisierende Wechselwirkung mit der ersten Base des Codon-Triplets eingehen kann. Auch ein eigenständiger Effekt eines lediglich thiomethylsubstituierten ms^2A wäre also denkbar.

1.3.7.5 Bedeutung im Zusammenhang mit Krankheiten

Es gibt bislang keine bekannte Erkrankung des Menschen, die in einem direkten Zusammenhang mit einer Hypo- oder Hypermodifikation von ms^2i^6A bzw. i^6A Loci in RNA stehen. Allerdings sind beide als Modifikationen eukaryotischer mitochondrialer tRNA bekannt und mit MELAS (*Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes*) und MERRF (*Myoclonic epilepsy with ragged red fibers*) existieren zwei Erkrankungen, die auf Hypomodifikation spezifischer mitochondrialer tRNAs in der Wobbleposition zurückzuführen sind, welche hier in gesunden Individuen ein τm^5U bzw. τm^5s^2U tragen.²⁴⁹⁻²⁵¹ Mitochondrien verfügen über ein eigenständiges Genom mit 37 Genen, welches unter anderem mitochondriale tRNAs codiert (22 Gene). Die entsprechenden Gene sind anfällig für Punktmutationen, welche in einer Störung der mitochondrialen Translation resultieren können und schon vielfach mit Krankheiten in Verbindung gebracht wurden.²⁵²⁻²⁵⁵

Eine aktuelle Studie an Cdk5Rap1 *Knockout*-Mäusen durch *Wei et al.* konnte zeigen, dass eine ms^2i^6A -Hypomodifikation zu verstärktem *Frameshifting*, einer reduzierten mitochondrialen Translation und schließlich substanziell reduzierter Synthese einzelner Komponenten (Komplex I, II und IV) der Atmungskette führte.¹⁹⁸ Hieraus resultierte eine leicht erhöhte Belastung mit reaktiven Sauerstoffspezies (ROS), die durch aus der dysfunktionalen Atmungskette leckenden Elektronen entstehen. Darüber hinaus waren die betroffenen Mäuse deutlich anfälliger für kardiale Dysfunktion und waren physisch weniger leistungsfähig, was durch den verstärkten Effekt mitochondrialer Defekte auf Muskelgewebe erklärt werden kann.

Eine bedeutende Ursache für humane MELAS-Erkrankungen sind A3243G Mutationen, welche mt-tRNA^{Leu(UUR)} betreffen und ebenfalls zu erhöhten ROS-Werten führen.²⁵⁶ *Wei et al.* konnten zeigen, dass aus an MELAS erkrankten Patienten isolierte periphere Blutzellen, bei denen eine A3243G Mutation nachweisbar war, über ein reduziertes Maß an ms^2 -Modifikationen der tRNA^{Trp}, tRNA^{Phe}, tRNA^{Tyr} und tRNA^{Ser(UCN)} verfügten. Da diese Beobachtung nicht mit einer Reduktion der Cdk5Rap1 Expression korrelierte, erklären die Autoren sie mit einem sekundären Effekt, der durch primär erhöhte ROS-Level das Fe²⁺-abhängige Enzym Cdk5Rap1 in seiner Aktivität hemmt und damit den initialen Phänotyp verstärkt. Die Repression der messbaren ms^2 -Level konnte durch eine Behandlung mit H₂O₂, also einer Erhöhung des oxidativen Stresses, reproduziert werden.

In humanen Kulturzellen (MCF-7) konnten *Wang et al.* zeigen, dass ein *Knockdown* von Cdk5Rap1 zu einem Anstieg der ROS-Belastung und darauf folgender p-JNK-vermittelter

Apoptose führt.²⁵⁷ Auch dieser Effekt rührt von einer Störung des mitochondrialen Stoffwechsels.

Diese Forschungsergebnisse zeigen, dass ms^2i^6A für eine voll funktionale mitochondriale Translation und damit einhergehend mitochondriale Funktion wichtig sind, einen direkten Zusammenhang mit einer konkreten Erkrankung gibt es bislang aber nicht. Ein Hinweis auf einen solchen könnte eine Studie von *Shin et al.* sein, welche eine empirische Assoziation zwischen dem Vorkommen eines SNP (*Single nucleotide polymorphism*) innerhalb des *Cdk5Rap1*-Gens und der Entstehung von Vitiligo, einer Pigmentierungsstörung der Haut, feststellt. Die untersuchte Population ist mit 269 Patienten allerdings klein und eine kausale Verbindung bislang nicht ersichtlich.²⁵⁸

Weiterhin konnte für Lungenkrebspatienten gezeigt werden, dass in befallenen Gewebe deutlich niedrigere Expressionslevel der Isopentenyltransferase *Trit1* vorlagen und eine Überexpression des Proteins das Tumorwachstum hemmt.²⁵⁹ Hieraus schließen *Spinola et al.*, dass *Trit1* eine negativ-regulative Funktion in diesem Tumorgewebe zukommt, was als Hinweis auf eine Tumor-suppressive Rolle von i^6A , oder seines biosynthetischen Folgeproduktes ms^2i^6A , gedeutet werden kann. In der Tat konnte eine zytotoxische Wirkung N^6 -substituierter Adenosine, darunter auch i^6A , bereits in deutlich älteren Studien nachgewiesen werden.²⁶⁰ Eine ungesättigte Seitenkette ist dabei von Bedeutung und *trans*-Isomere zeigten sich aktiver als *cis*-Isomere. Eine Hydroxylierung an Position 4 der Isopentenylgruppe (io^6A , ms^2io^6A) erhöhte ebenfalls die Aktivität, während sie an anderen Positionen zu einer Abschwächung führte.²⁶¹

Neben der Thiomethyltransferase *Cdk5Rap1*, die ms^2i^6A aus i^6A erzeugt, existiert im Menschen ein verwandtes Enzym (*Cdkal1*, siehe auch **Abbildung 17**), welches in tRNA 6-Threonylcarbamoyladenosen (t^6A) zu ms^2t^6A (**Abbildung 18**) umsetzt.²⁴³ SNPs im *CDKAL1*-Gen korrelieren mit einer erhöhten Wahrscheinlichkeit an Typ 2 Diabetes zu erkranken²⁶²⁻²⁶⁵, was auf einen gestörten mitochondrialen Stoffwechsel zurückgeführt wird.²⁶⁶ *Ohara-Imaizumi et al.* konnten zeigen, dass ein *Cdkal1 Knockout* zu einer reduzierten Insulinexozytose und mitochondrialen ATP-Synthese führt. Vor diesem Hintergrund erscheint es möglich, dass auch *Cdk5Rap1*-Defekte zu kranken Phänotypen führen, auch eine wechselseitige Beeinflussung beider Modifikationen, bzw. ihrer biosynthetischen Enzyme ist vorstellbar.

Es existieren also Hinweise auf schädliche Veränderungen des Zellstoffwechsels durch ein Fehlen von m^2i^6A oder i^6A , eine direkte, ursächliche Relation zu einer bekannten Erkrankung wurde bislang aber noch nicht gefunden.

2. Aufgabenstellung

In der Einleitung wurde bereits ausgeführt, dass Derivate der kanonischen RNA Basen über alle Domänen des Lebens einen sehr großen chemischen Raum abdecken und seit geraumer Zeit Gegenstand der Forschung sind. Während ihr Vorkommen in nicht-kodierenden Transkripten, wie tRNA und rRNA bereits lange bekannt ist und intensiv erforscht wurde, stellt das Feld der chemischen Modifikationen in proteincodierender RNA (mRNA) einen verhältnismäßig jungen Forschungsschwerpunkt dar. Für eine spezifische Regulation der Expression einzelner Proteine, oder die Modulation alternativen Splicings, stellt mRNA aber die deutlich besser geeignete Spezies dar, da Modifikationen der tRNA oder rRNA notwendigerweise globale Effekte verursachen.

Die geringe Beachtung, die Modifikationen der mRNA bislang gefunden haben lässt sich mit der sehr niedrigen, in Zellen vorhandenen Menge von mRNA im allgemeinen, besonders aber einzelner Transkripte erklären. Die Abundanz oder genaue Lokalisierung einzelner Modifikationen ist erst durch die Entwicklung moderner Methoden zur massenspektrometrischen Quantifizierung oder Hochdurchsatz-Sequenzierung untersuchbar. Diese Verfügbarkeit verbesserter analytischer Methoden hat zu einer rapiden Entwicklung der Forschung an Modifikationen der mRNA geführt und die Annahme, mRNA würde eine nahezu unmodifizierte Spezies darstellen, musste in den letzten Jahren revidiert werden.^{32,267,268}

Die Modifikation m⁶A ist die in ihrer biologischen Wirkung bislang am besten erforschte mRNA Modifikation²⁶⁹ und ein Beispiel das zeigt, welche Bedeutung die Identifikation von spezifischen Bindern (*Reader* Proteinen) hat. Die Kenntnis dieser *Reader* stellt eine notwendige Basis zur Aufklärung der biologischen Mechanismen dar, über die eine Modifikation wirkt. Eine etablierte und für diesen Einsatzzweck gut geeignete Methode stellen *Pulldown*-Experimente aus mit stabilen Isotopen markierten Kulturzellen dar (SILAC; *Stable isotope labelling with amino acids in cell culture*).^{270,271}

Ziel dieser Arbeit war es, die Interaktome mehrerer Modifikationen, die in mRNA vorhanden sind, oder vermutet werden, zu erforschen, um weitergehende Kenntnis über ihre biologischen Funktionen zu erlangen. In der vorgelegten Arbeit sollten N⁶-Methyladenosin (m⁶A), N⁶,N⁶-Dimethyladenosin (m⁶₂A), N⁶-Isopentenyladenosin (i⁶A), N⁶-Isopentenyl-2-thiomethyladenosin (ms²i⁶A), Pseudouridin (Ψ) und die 2'-O-Methylmodifikationen aller kanonischen

Basen (Am, Gm, Cm, Um) untersucht werden. Die Identifikation von Am, Gm, Cm, ms²i⁶A, i⁶A und m⁶₂A in mRNA durch *Veronika Reiter*¹⁰⁹ warf unmittelbar die Frage nach deren biologischen Funktion auf.

Besonderes Augenmerk sollte auf m⁶A und ms²i⁶A gelegt werden. m⁶A ist die momentan am stärksten im Fokus stehende mRNA-Modifikation. Eine umfangreiche Studie ihres Interaktoms ist bislang aber nicht publiziert, was eine bedeutende Wissenslücke darstellt. Die Hypermodifikation ms²i⁶A stellt die erste Modifikation ihrer Art in mRNA dar, Derivate von vergleichbarer struktureller Komplexität konnten bislang nicht nachgewiesen werden. Für Pseudouridin wiederum sollten nach der Publikation von Sequenzierungsstudien^{75,78,81} im Verlauf dieser Dissertation erste Erkenntnisse zu spezifischen Bindern gewonnen werden.

Unerlässlich für die Durchführung von *Pulldown*-basierten Interaktomstudien sind Oligoribonukleotide mit definierter Sequenz, die als „Köder“ (*Baits*) spezifisch bindende Proteine anreichern können und hierfür die zu untersuchende Modifikation tragen müssen. Hierfür war die chemische Synthese der zu Beginn dieser Arbeit nicht kommerziell erhältlichen Modifikationen m⁶A, m⁶₂A, i⁶A und ms²i⁶A und ihrer Phosphoramidite nötig. Nach ihrer erfolgreichen Herstellung sollten sie in Oligoribonukleotide inkorporiert werden und in Kooperation mit der Gruppe von *Michiel Vermeulen* ihre *Reader* Proteine identifiziert werden.

Neben der Identifikation der entscheidenden Interaktoren war ein weiteres Ziel dieser Arbeit eine Vertiefung des Verständnisses von der biologischen Relevanz der Modifikation ms²i⁶A durch eine phänotypische Charakterisierung einer *Cdk5Rap1 Knockout*-Zelllinie.

3. Synthese der modifizierten Ribonukleoside und ihrer Phosphoramidite

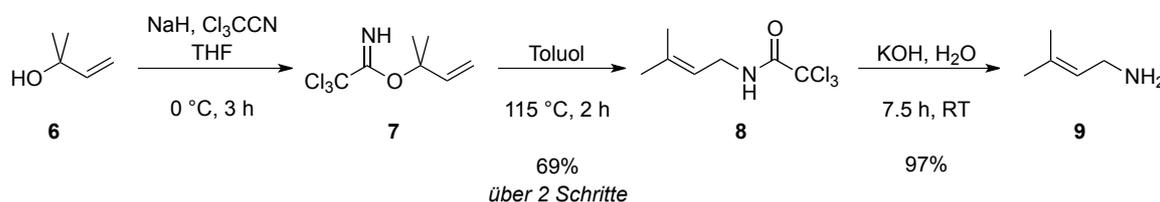
Um Proteine zu identifizieren, die spezifisch an chemisch modifizierte Nukleoside in zellulären Transkripten binden war es nötig, Oligoribonukleotide zur Verfügung zu stellen, welche die zu untersuchenden Modifikation tragen. Hierzu wurden die Modifikationen i^6A , ms^2i^6A , m^6A , m^6_2A und ihre Phosphoramidite dargestellt. Weitere Modifikationen, deren Interaktome ebenfalls betrachtet werden sollten waren kommerziell erhältlich (Am, Gm, Cm, Um, Ψ). Darüber hinaus sollte die Anwendbarkeit eines *Crosslinkers* untersucht werden, der mittels Click-Chemie kovalent an Oligonukleotide gebunden werden kann.

3.1 Synthese der Isopentenylmodifikationen i^6A und ms^2i^6A

Die Synthese der Isopentenyl substituierten Adenosinderivate i^6A und ms^2i^6A wurde in dieser Arbeitsgruppe bereits von *Tobias Brückl* und *Veronika Reiter* beschrieben und wurde hier in Anlehnung an deren Dissertationen durchgeführt.^{59,109} Beide Nukleoside wurden bereits in der Literatur beschrieben und synthetisch hergestellt.^{184,260,272-274}

3.1.1 Synthese des N^6 -Isopentenyladenosin Phosphoramidits

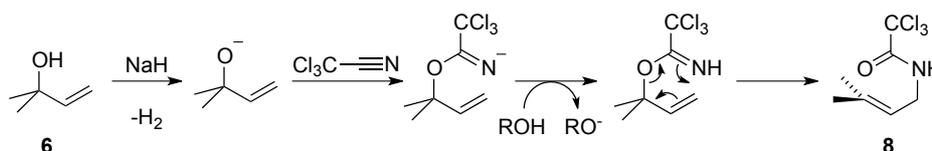
Zunächst musste für die Synthese von sowohl i^6A , als auch ms^2i^6A Isopentenyloamin dargestellt werden. Die literaturbekannte Synthese²⁷⁵ erfolgte über drei Schritte in einer Gesamtausbeute von 67% (**Schema 8**).



Schema 8: Syntheseroute zur Darstellung von Isopentenyloamin **9** ausgehen von Allylalkohol **6**.

Hierzu wurde zunächst Allylalkohol **6** nach Deprotonierung durch Natriumhydrid mit Trichloracetonitril zu Imidat **7** umgesetzt. Katalytische Mengen Base sind hier ausreichend, um die Transformation vollständig ablaufen zu lassen, da das entstehende Intermediat in der Lage ist den verbleibenden Alkohol zu deprotonieren. Das Imidat kann in einer *Overman*

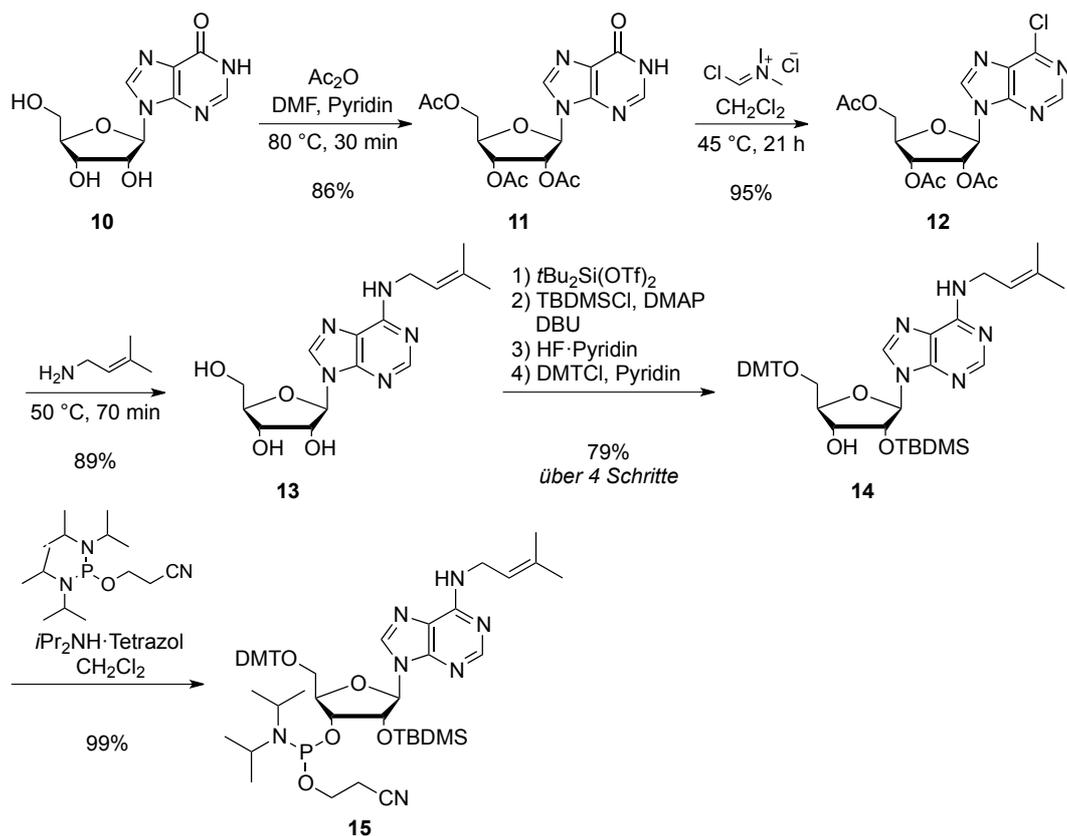
Umlagerung in das Trichloracetamid (**8**) überführt werden, welches bereits das Trichloracetyl-geschützte Isopentenylamin darstellt. Die $[3,3]$ -sigmatrope Umlagerung läuft in siedendem Toluol innerhalb von zwei Stunden nach dem in **Schema 9** abgebildeten Mechanismus ab.²⁷⁶ Das Produkt kann durch Destillation gewonnen werden, liegt bei Raumtemperatur aber als Feststoff vor. Basische Spaltung des Trichloracetamides in wässriger KOH liefert das gewünschte Amin **9** in hervorragender Ausbeute.



Schema 9: Mechanismus der Umsetzung von Allylalkohol **6** zu Trichloracetamid **8** über eine irreversible $[3,3]$ -sigmatrope *Overman* Umlagerung. Die Reaktion läuft konzertiert über einen sechsgliedrigen Übergangszustand ab.

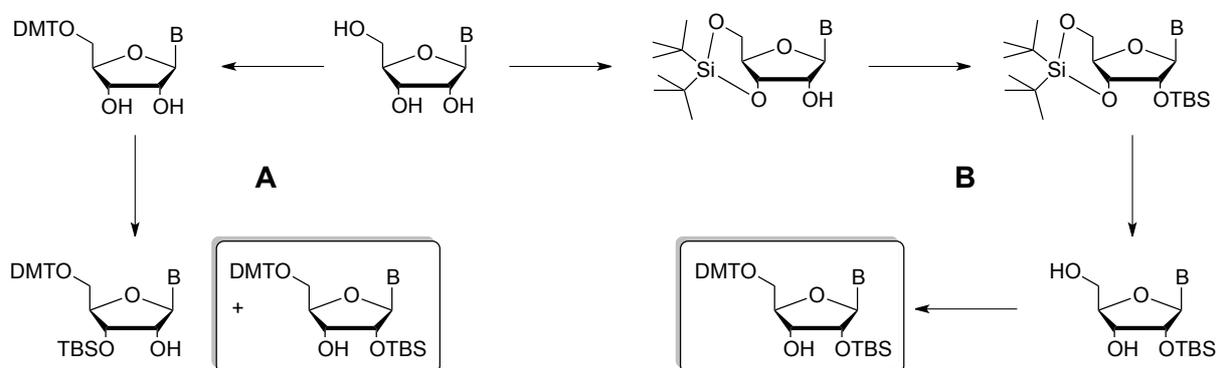
Die Synthese von i^6A , wie auch der anderen ausschließlich N^6 -modifizierten Adenosinderivate (m^6A , m^6_2A), wurde ausgehend von Inosin durchgeführt. Dieses wurde zunächst mit Acetanhydrid unter Standardbedingungen geschützt²⁷⁷ (**10** zu **11**) und anschließend unter Verwendung des *Vilsmeier*-Reagenz ((Chlormethylen)dimethyliminiumchlorid) zu seinem elektrophileren Derivat **12** umgesetzt.²⁷⁸ Diese Transformation erwies sich als deutlich stabiler in ihren Ausbeuten als die sonst übliche Chlorierung mittels $POCl_3$, mit Dimethylanilin und Tetramethylammoniumchlorid.²⁷⁷

Das in sehr guter Ausbeute gewonnene Chlorinosin wurde nun in reinem Isopentenylamin gelöst und unter erhöhtem Druck mit diesem umgesetzt, wodurch sowohl eine Substitution des Chlorides, als auch eine Entschützung der Ribose erreicht wurde. Eine alternative Route stellt die Umsetzung von **12** mit Isopentenylamin in ethanolischer Lösung^{59,109,260,279} dar. Sie nimmt etwas mehr Zeit in Anspruch und liefert in einigen Fällen ein nicht vollständig deacetyliertes Produktgemisch.



Schema 10: Synthese des N^6 -Isopentenyladenosin Phosphoramidits ausgehend von Inosin in 8 Schritten mit einer Gesamtausbeute von 57%.

Aus dem so gewonnenen Nucleosid i^6A (**13**) wurde in einer vierstufigen Sequenz der Vorläufer für die Phosphitylierung (**14**) dargestellt. Um Ribonucleoside mit den für die Festphasensynthese nötigen Schutzgruppen zu versehen gibt es zwei zumeist angewandte Routen (**Schema 11**).²⁸⁰⁻²⁸³



Schema 11: Die zwei üblichen Herangehensweisen zur Schützung von Nucleosiden, um sie in der Festphasensynthese einzusetzen. **A:** Der reaktivere primäre Alkohol wird zunächst trityliert um anschließend die 2'-Hydroxygruppe zu Schützen. Hierbei entsteht ein Gemisch aus 2'- und 3'-O-TBDMS Produkten. **B:** 5'- und 3'-Alkohole der Ribose werden durch eine Silylklammer blockiert, bevor selektiv die 2'-OH Gruppe TBDMS-geschützt wird. Spaltung des Silyldiethers und darauf folgende Reaktion mit DMT-Chlorid liefert das gewünschte Nucleosid als ein Regioisomer.

In der kürzeren Route (**A**) wird das Nukleosid zunächst DMT geschützt, wobei in der Regel selektiv die 5'-Hydroxygruppe reagiert. Im zweiten Schritt erfolgt die TBDMS-Schätzung, die ein Gemisch aus 2'- bzw. 3'-*O*-silyliertem Produkt liefert. Diese Route ermöglicht die Synthese des gewünschten Produktes in nur zwei Schritten, ist aber mit dem Nachteil behaftet eine Mischung zweier Regioisomere zu erzeugen, die in vielen Fällen nur schwer voneinander trennbar sind.^{281,282} Verwertbar für die RNA Festphasensynthese ist nur das 2'-*O*-TBDMS geschützte Produkt.

Die zweite Methode²⁸⁰ (**B**) inaktiviert in einer ersten Reaktion die reaktive 5'- und 3'-Position durch eine an *Markiewicz*²⁸⁴ angelehnte Methodik unter Verwendung eines Silylbistriflates. Dies ermöglicht eine selektive Silylierung der 2'-Hydroxygruppe im Folgeschritt. Eine anschließende Abspaltung der Silylklammer und Tritylierung des nun freien primären Alkohols liefert das gewünschte Produkt zwar nach insgesamt vier Reaktionsschritten, aber ohne signifikante Verunreinigung durch das störende Regioisomer. Bei der Synthese aller in dieser Arbeit dargestellten Modifikationen wurde auf letztere Route zurückgegriffen, da hierdurch die Produkte in aller Regel in guten Ausbeuten und sehr guter Reinheit zugänglich waren.

Bei der Handhabung und Lagerung des Produktes ist zu bedenken, dass beide Regioisomere ineinander überführbar sind und sich, abhängig von Lösemittel und Temperatur, schnell Gleichgewichte einstellen.²⁸⁵ Insbesondere methanolische Lösungen sind nicht dauerhaft stabil.

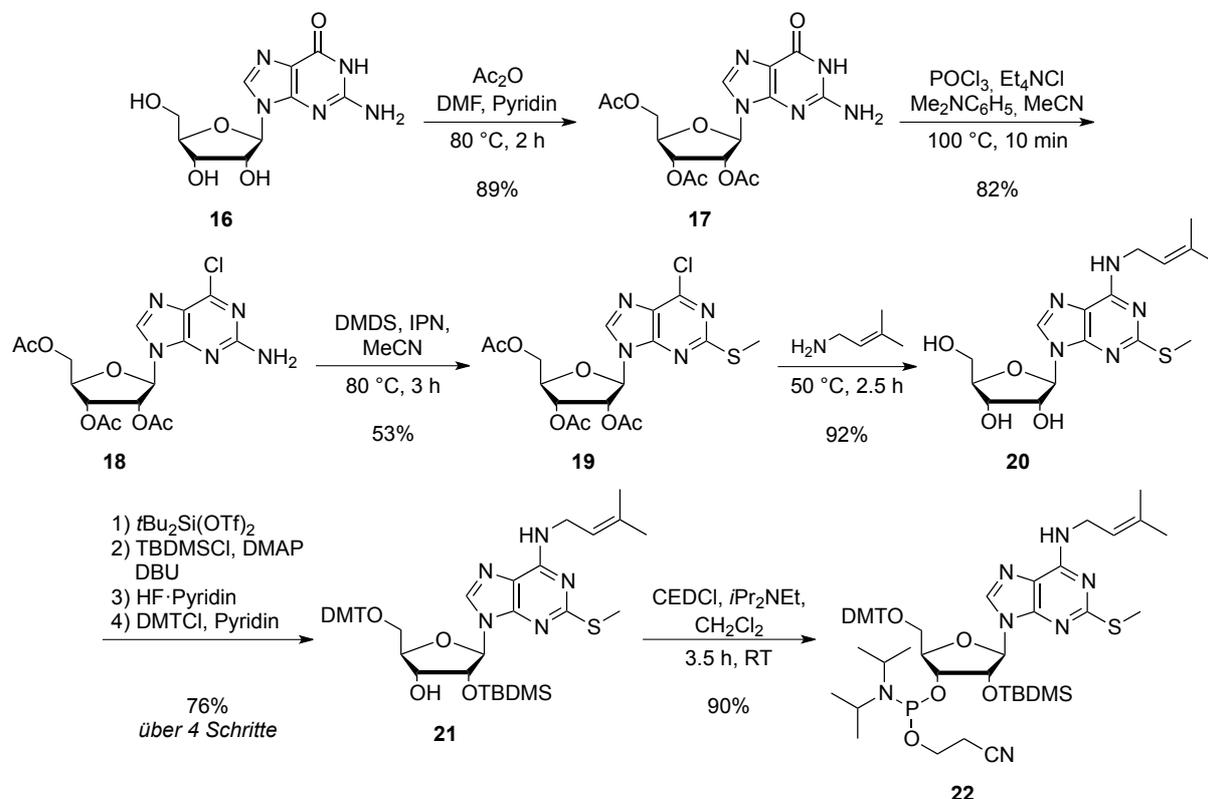
Nach dem oben beschriebenen Verfahren wurde in vier Schritten in exzellenter Ausbeute das geschützte Nukleosid **14** synthetisiert, welches wiederum unter Verwendung von 2-Cyanoethoxy-bis(*N,N*-diisopropylamino)phosphin und *N,N*-Diisopropylammoniumtetrazolid^{286,287} unter milden Bedingungen in das i⁶A-Phosphoramidit **15** überführt werden konnte.

3.1.2 Synthese des 2-Thiomethyl-*N*⁶-isopentenyladenosin Phosphoramidits

Die Darstellung des ms²i⁶A Phosphoramidites erfolgte in Teilen analog zur vorangehend beschriebenen Synthese und ist in **Schema 12** zusammengefasst. Da eine Funktionalisierung der 2-Position erfolgen sollte, ging die Synthese statt Inosin von Guanosin aus, welches unter Standardbedingungen *O*-Acetyliert wurde (**16** zu **17**).²⁷⁷ Um den Isopentenylsubstituenten einführen zu können, wurde auch dieses Intermediat chloriert. Hierbei konnte nicht auf die milde Substitution mittels *Vilsmeier* Reagenz zurückgegriffen werden, da dieses eine hohe Reaktivität am N2 zeigte und als Nebenprodukt das *N*-Dimethylformamidin-Derivat erzeugte.

Eine nach einem der *Vilsmeier-Haack*-Reaktion ähnlichen Mechanismus ablaufende Umsetzung mit POCl_3 und Dimethylanilin lieferte das Chlorpurin **18** in guter Ausbeute.²⁷⁷ Hierbei war es zentral, die Reaktionszeit kurz zu halten, da sich das Produkt unter den Bedingungen der Umsetzung rasch zu zahlreichen Nebenprodukten zersetzt. Im Anschluss erfolgte die Einführung der Thiomethylgruppe indem zunächst ein hochreaktives 2-Diazopurin erzeugt wurde, welches *in situ* eine Substitution mit Dimethyldisulfid zu **19** durchlief.²⁷³ Die N^6 -Modifikation wurde wiederum in dem substituierenden Amin als Lösemittel unter erhöhtem Druck erreicht, wobei parallel die Ribose vollständig entschützt werden konnte, um $\text{ms}^{2;6}\text{A}$ (**20**) zu erhalten.

In vier Schritten wurde darauf folgend das geschützte Nukleosid **21** dargestellt. In dieser oben bereits beschriebenen Abfolge von Schützungen und Entschützungen wird üblicherweise im zweiten Schritt die 2'-Silylgruppe durch TBDMSCl in Kombination mit Imidazol als Aktivator unter erhöhter Temperatur eingeführt.²⁸⁰ Eine mildere Variante dieser Vorgehensweise unter Verwendung von 4-Dimethylaminopyridin als Katalysator bei Raumtemperatur erwies sich als effektiver und wurde daher an dieser Stelle, wie auch für alle anderen Nukleoside, eingesetzt. Die Umsetzung zum Phosphoramidit **22** gelang mit CEDCl in sehr guter Ausbeute.



Schema 12: Synthese des $\text{ms}^{2;6}\text{A}$ Phosphoramidits **22** ausgehend von Guanosin in 9 Schritten mit einer Gesamtausbeute von 24%.

3.2 Synthese der N^6 -methylierten Adenosine m^6A und m^6_2A

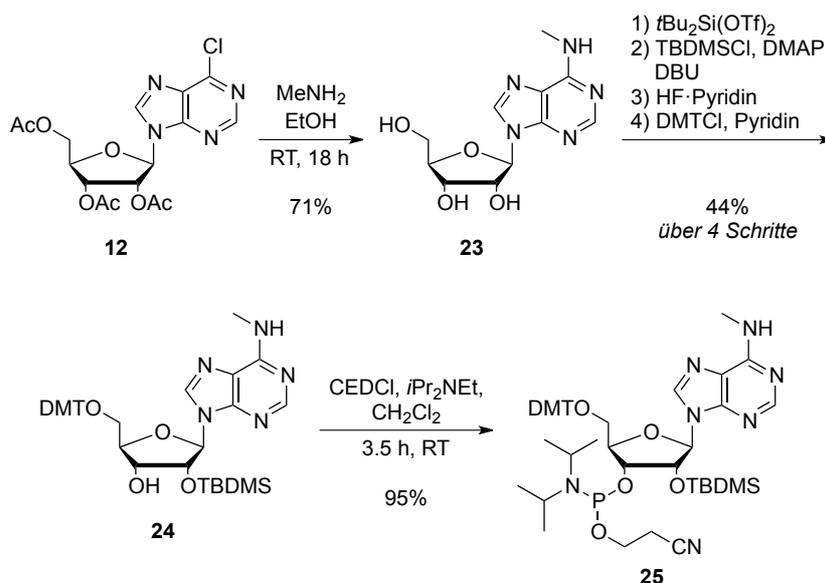
Die Synthese der Modifikationen m^6A und m^6_2A wurde in dieser Arbeitsgruppe bereits von *Daniel Globisch* durchgeführt und ist in der Literatur beschrieben.^{278,288} Die ersten Vorschriften zu ihrer Darstellung wurden 1955 (m^6_2A) und 1958 (m^6A) veröffentlicht.^{289,290}

3.2.1 Synthese des N^6 -Methyladenosin Phosphoramidits

Es sind verschiedene Herangehensweisen literaturbekannt, um N^6 -alkylierte Adenosinderivate zu erzeugen. Dazu gehören eine N^1 -Alkylierung mit darauf folgender *Dimroth* Umlagerung^{291,292}, Aminierungen von Inosin,²⁹³ N^6 -Alkylierungen unter Phasentransferbedingungen²⁹⁴ und nukleophile Substitutionen an 6-Chlorpurinen^{260,279}.

Da sich 6-Chlorinosin **12** als gemeinsamer Vorläufer für sowohl m^6A , als auch m^6_2A und i^6A eignete und in zwei Schritten in großer Quantität zugänglich ist, wurde diese Methode verwendet.

Unter Verwendung einer ethanolischen Methylaminlösung wurde eine Substitution am C6 durchgeführt und m^6A (**23**) in guter Ausbeute erhalten. Dieses wurde zunächst in der bereits beschriebenen Abfolge von Schritten geschützt (**24**) und in das am N6 ungeschützte Phosphoramidit **25** überführt (**Schema 13**).

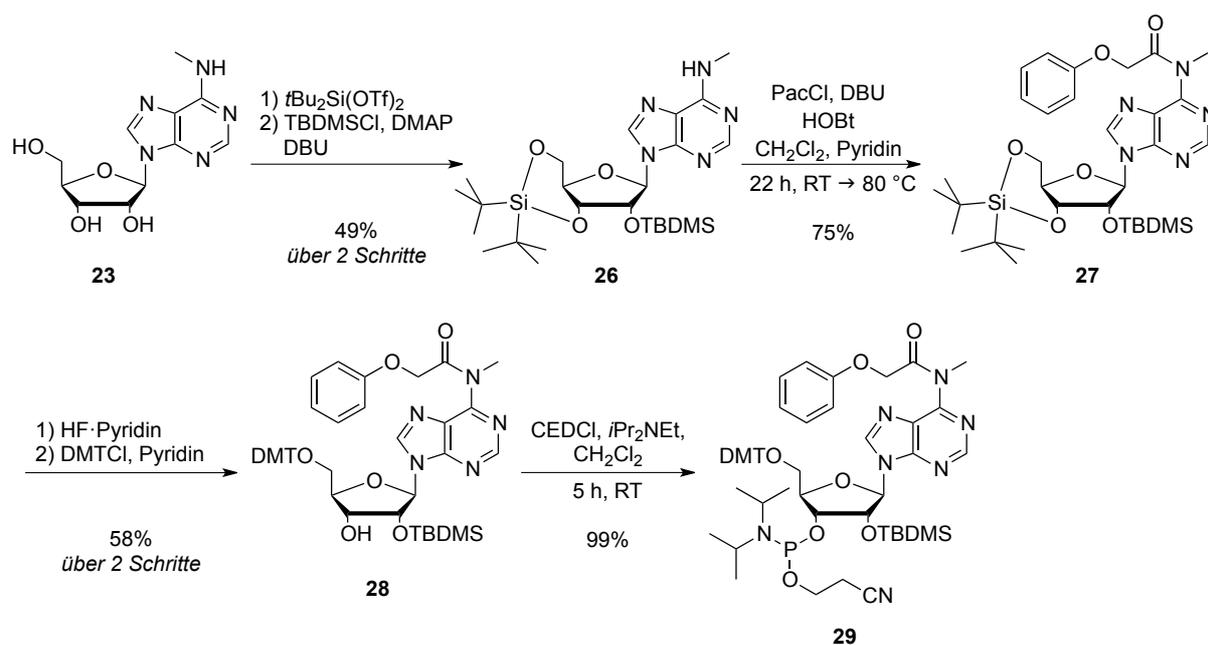


Schema 13: Synthese des ersten m^6A Phosphoramidits aus 6-Chlorinosin in 6 Stufen mit einer Gesamtausbeute von 30%.

Mit diesem konnten zwar zum Teil gute Ausbeuten in der Festphasensynthese erzielt werden und eine Verzweigung durch unerwünschte Kopplung am freien N6 konnte, entgegen der Erwartung,²⁹⁵ auch unter Verwendung von 4,5-Dicyanoimidazol (DCI) als Aktivator nicht beobachtet werden. Sowohl die Effizienz der Phosphitylierungsreaktion, als auch die Kopplungsrate in der Oligonukleotidsynthese erwiesen sich aber als sehr schlecht reproduzierbar, womit dieses Phosphoramidit kein geeignetes Monomer für die verlässliche Bereitstellung einer Vielzahl mehrfach modifizierter Sequenzen darstellte.

Neben **25** ist auch ein *N*⁶-Phenoxyacetyl (Pac) geschütztes Amidit kommerziell erhältlich (*Glen Research*), zu dem aber weder ein Patent, noch eine publizierte Synthesevorschrift verfügbar sind. Dieses würde Nebenreaktionen des methylierten Amins sowohl mit CEDCI als auch in der Festphasensynthese minimieren. Eine nachträgliche Schützung von **24** erwies sich aufgrund der geringen Reaktivität des N6 und dem ungeschützt vorliegenden 3'-OH, durch am Alkohol bevorzugt stattfindende Substitutionsreaktionen als nicht möglich.

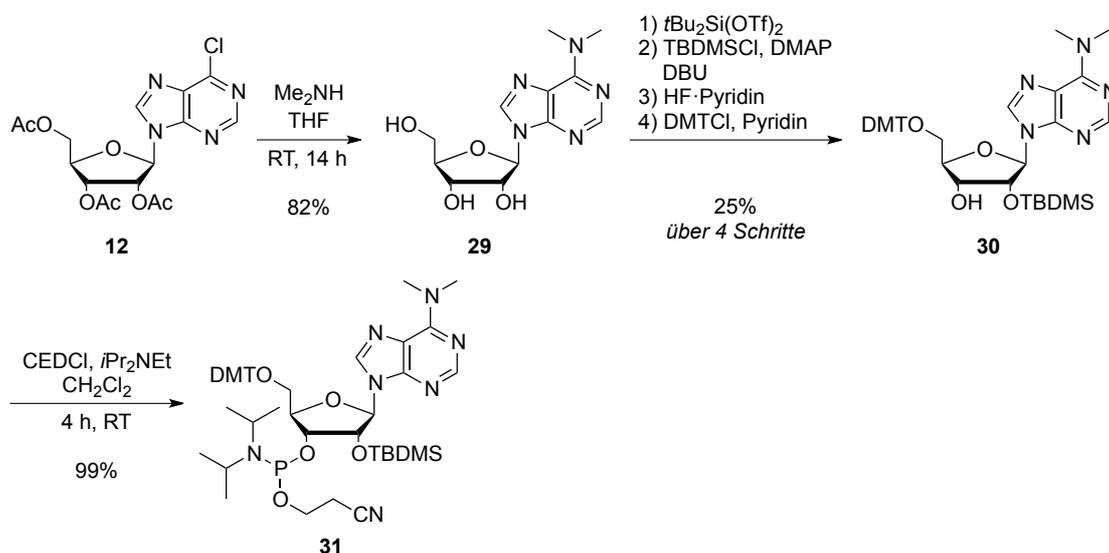
Ausgehend von m⁶A konnte das Zielmolekül dargestellt werden, indem zunächst mittels Silylbistriflat 5'- und 3'-Hydroxygruppe inaktiviert und darauf folgend der einzig freie Alkohol TBDMS-geschützt wurden (**26**). In einem weiteren Schritt wurde der Phenoxyacetylsubstituent an der N6-Position eingeführt (**27**), die Silylklammer mittels HF entfernt und der primäre Alkohol trityliert (**28**). In einer abschließenden Reaktion wurde das Phosphoramidit eingeführt, was mit CEDCI in exzellenter Ausbeute gelang, und das vollgeschützte m⁶A Monomer **29** erzeugt.



Schema 14: Synthese des vollgeschützten m⁶A Phosphoramidits aus m⁶A in 6 Stufen mit einer Gesamtausbeute von 21%.

3.2.2 Synthese des N^6,N^6 -Dimethyladenosin Phosphoramidits

m^6_2A und sein Phosphoramidit konnten analog zu m^6A dargestellt werden, wobei hier keine Schützung des N6 nötig war. 6-Chlorinosin **12** wurde durch Umsetzung mit einer Dimethylaminlösung in THF in Dimethyladenosin **29** überführt.²⁷⁸ Dieses konnte in der bereits beschriebenen Sequenz von Silylschützung und Entschützung, darauf folgender Tritylierung und Phophitylierung zu **30** und schließlich **31** umgesetzt werden (**Schema 15**).

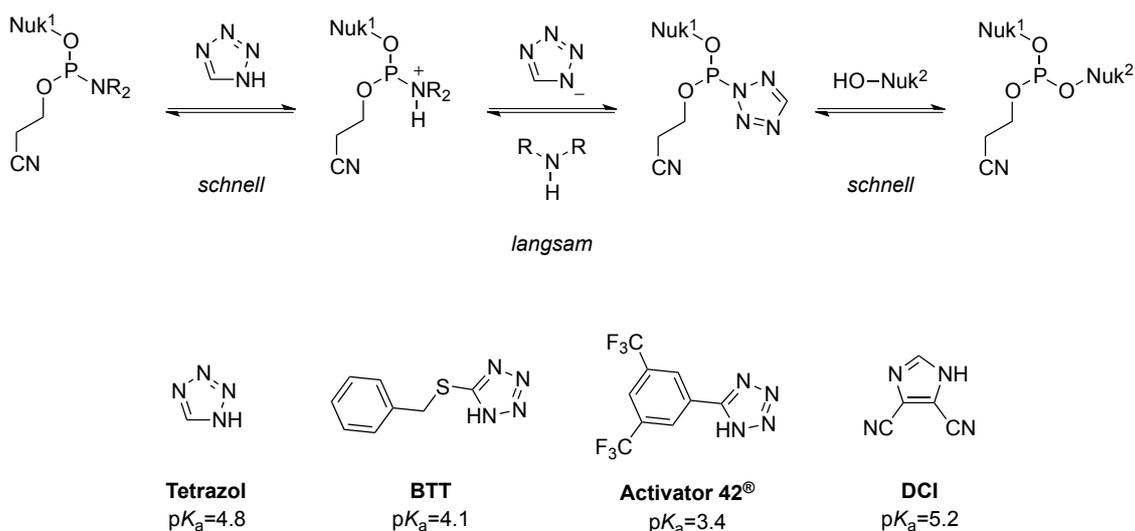


Schema 15: Synthese des m^6_2A Phosphoramidits aus 2',3',5'-*O*-Triacetyl-6-chlorinosin **12** in 6 Schritten mit einer Gesamtausbeute von 20%.

4. Synthese der Oligoribonukleotide

Die Synthese der Oligoribonukleotide (ORN) wurden auf einem *Applied Biosystems 394 DNA/RNA Synthesizer* durchgeführt, wobei Benzylthiotetrazol (BTT) als Aktivator verwendet wurde. Es wurden 2'-*O*-TBDMS Amidite mit den üblichen Schutzgruppen für exozyklische Amine (Ac-C, Bz-A, dmf-G) genutzt und Kopplungszeiten von 10 min für Standardbasen, sowie 20 min für Sonderamidite verwendet. Es wurden leicht abgewandelte Standardzyklen eingesetzt (siehe Anhang). Kurze Sequenzen, bis 25 Nukleotide, wurden mit 1.5 M Amiditlösungen und einfachen Kopplungen synthetisiert. Um die Ausbeute bei längeren ORN (>25 Nukleotide) zu maximieren, wurden hier 1.0 M Amiditlösungen und doppelte Kopplungen für jede Position der Sequenz angewandt.

Neben dem überwiegend verwendeten Aktivator BTT wurden auch der proprietäre *Activator 42*[®] (5-[3,5-Bis(trifluoromethyl)phenyl]-2*H*-Tetrazol, *Sigma Aldrich*) und DCI (4,5-Dicyanoimidazol) getestet. Die Aktivatoren in der Kopplungsreaktion der Festphasensynthese üben eine duale Rolle als Säure und Nukleophil aus, indem sie zunächst den Aminosubstituenten des Phosphoramidites protonieren und ihn anschließend substituieren (**Schema 16**).²⁹⁶⁻²⁹⁹ Ausgehend vom klassischen Aktivator Tetrazol wurden verschiedene Derivate entwickelt, welche die Nukleophilie von Tetrazol weitgehend erhalten sollen, und gleichzeitig niedrigere pK_a -Werte besitzen.²⁹⁶



Schema 16: Oben: Mechanismus der Aktivierung eines Phosphoramidits im Kopplungsschritt der Oligonukleotid-Festphasensynthese. Unten: Tetrazolderivate, die Anwendung als Aktivatoren finden (Nuk = Nukleosid; R = *i*Pr).²⁹⁶

Unter den Derivaten stellt *Activator 42*[®] einen hocheffektiven und in der DNA Synthese beliebten Aktivator dar. Während in der Festphasensynthese von DNA Oligomeren Kopplungszeiten von 30 s üblich sind, werden in der RNA Synthese deutlich längere Zeiten, von 5 bis über 20 Minuten benötigt. Innerhalb dieser Reaktionszeit ist eine Entschützung der stark säurelabilen 5'-*O*-DMT Gruppe durch saure Aktivatoren möglich, wodurch Doppelkopplungen stattfinden können.³⁰⁰ Die resultierenden Sequenzen mit alternativer Basenfolge, aber unter Umständen identischer Länge, sind in der Aufreinigung sehr schwer abtrennbar und ihre Entstehung sollte daher unbedingt vermieden werden. Da in dieser Arbeit eine Vielzahl verschiedener Sonderbasen inkorporiert werden sollten, welche zum Teil lange Kopplungszeiten benötigen, wurde *Activator 42*[®] trotz guter Kopplungsausbeuten nicht als Aktivator eingesetzt, um die beschriebenen Probleme zu umgehen. Um Nebenreaktion zu unterbinden, kann der wenig azide Aktivator DCI verwendet werden, er lieferte aber reduzierte Kopplungsausbeuten, die insbesondere bei langen Sequenzen zu einer signifikanten Reduktion der erzielbaren Gesamtausbeute führten. BTT erwies sich als Alternative, die gute Kopplungsausbeuten lieferte, ohne zu unerwünschten Reaktionen zu führen.

Die Seitengruppen der Adenosinderivate m^2i^6A und i^6A mussten in der Synthese nicht geschützt werden. Während m^6A als Baustein mit ungeschütztem N6 nutzbar ist, wurden Verzweigungsreaktionen am exozyklischen Amin berichtet²⁹⁵ und der ungeschützte Baustein erwies sich als unzuverlässiges Substrat in der Phosphorylierungsreaktion. Die Einführung einer Phenoxyacetyl-Schutzgruppe (Pac) lieferte einen gut phosphorylierbaren und zuverlässig koppelnden Baustein.

Nach der Synthese wurden die Oligonukleotide mit NH_3 /Methylamin vom Träger abgespalten und basisch entschützt, TBDMS-Gruppen wurden mit $HF \cdot NEt_3$ entfernt. Die Aufreinigung erfolgte mittels *Reversed phase* HPLC (*rp*-HPLC) und alle Produkte wurde nach der Aufreinigung entsalzt, um vorhandene Pufferreste zu entfernen. Die Produkte wurden abschließend durch MALDI-ToF und analytische *rp*-HPLC charakterisiert. In der Regel wurden Synthesen im Maßstab 1 μ mol durchgeführt, wobei die isolierten Ausbeuten stark von Sequenz und Modifikationsgrad des ORN abhingen. Nach allen Prozessierungsschritten konnten zwischen 40 nmol und 290 nmol Produkt erhalten werden, üblich waren Gesamtausbeuten von etwa 100 nmol, bzw. 10%.

Aufgrund der großen Anzahl synthetisierter Sequenzen werden hier nicht alle ORN dargestellt. Eine repräsentative Synthese zweier Stränge ist in **Abbildung 20** zusammen-

gefasst. Es handelt sich um zwei Oligonukleotide, die den dominanten Sequenzkontext, in welchem m⁶A zu finden ist, enthalten (GGACU).

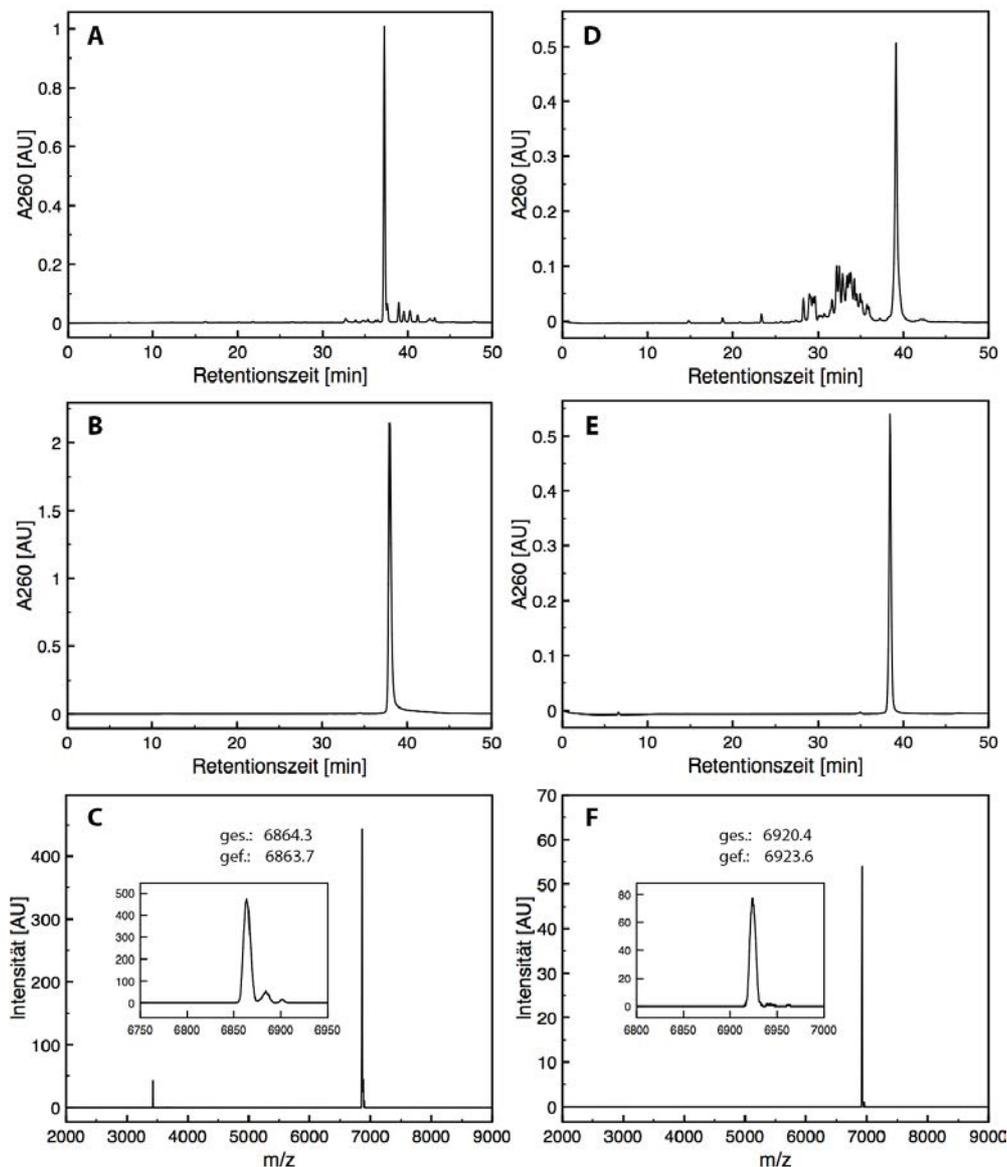


Abbildung 20: A-C: Charakterisierung des ORN 5'-Biotin GGACU GGACU GGACU GGACU-3'. A) HPLC-Chromatogramm des Rohproduktes, nach Abspaltung und Entschützung. B) HPLC-Chromatogramm des aufgereinigten ORN. C) MALDI-ToF Analyse des aufgereinigten ORN. D-F: Charakterisierung des modifizierten ORN 5'-Biotin GGm⁶ACU GGm⁶ACU GGm⁶ACU GGm⁶ACU-3'. D) HPLC-Chromatogramm des Rohproduktes, nach Abspaltung und Entschützung. E) HPLC-Chromatogramm des aufgereinigten ORN. F) MALDI-ToF Analyse des aufgereinigten ORN.

Als besonders herausfordernd erwiesen sich Oligonukleotide, die zur Ausbildung von Sekundärstrukturen neigen. Am Beispiel eines ms²i⁶A-modifizierten *Hairpins* ist das in **Abbildung 21** dargestellt. Sequenzen, die stabile Duplexe ausbilden können zum einen während der Synthese Sekundärstrukturen ausbilden, welche durch erhöhten sterischen Anspruch die Kopplungsausbeuten reduzieren und zum anderen während der Aufarbeitung gemischte Dimere und alternative Konformationen ausbilden, welche aufgrund ihres

veränderten Retentionsverhaltens nicht mit dem Volllängenprodukt eluieren und damit in der Aufreinigung verloren gehen. Mittels MALDI-ToF ist das gewünschte Produkt in gezeigten Fall bei verschiedenen Retentionszeiten nachweisbar, die nicht mit der Hauptfraktion übereinstimmen. **Abbildung 21 D** verdeutlicht auch anschaulich die durch $ms^{2:6}A$ bedingte erhebliche Veränderung der Retentionszeit des ORN. Keine der anderen in dieser Arbeit behandelten Modifikationen hatte eine vergleichbare Wirkung.

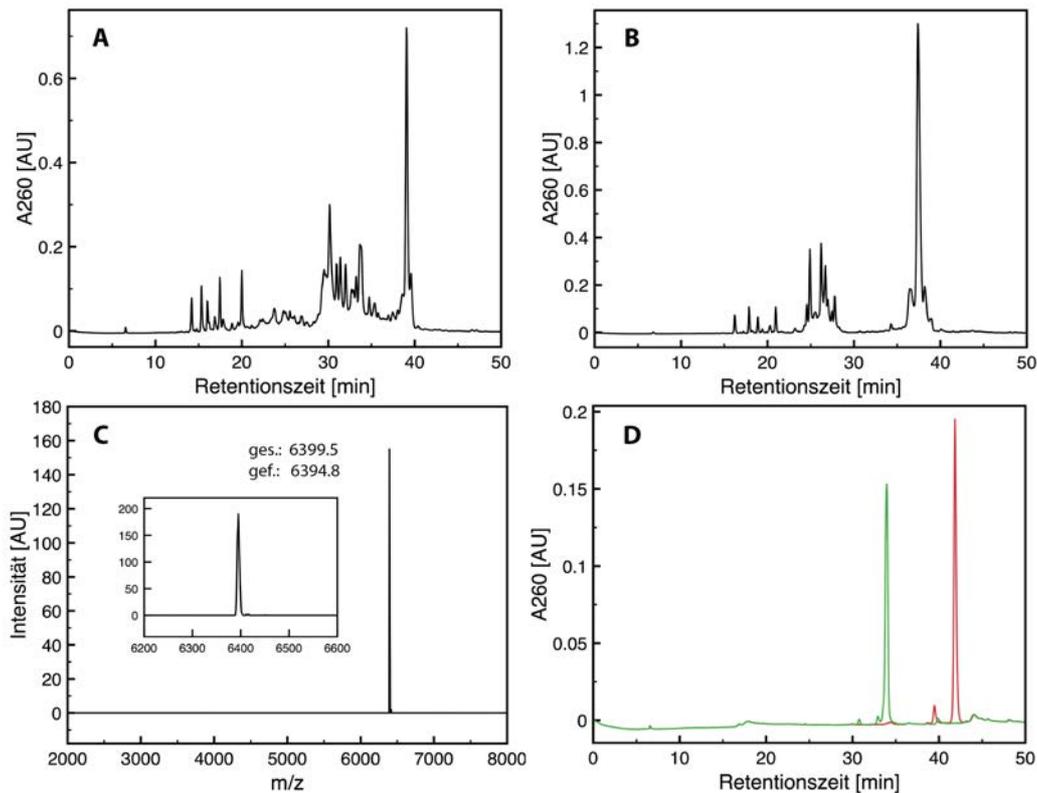


Abbildung 21: Charakterisierung der ORN-1/-1' 5'-bU-GGG GAU UGA Ams^{2:6}AA UCC CC-3' und 5'-bU-GGG GAU UGA AAA UCC CC-3' (bU = 33, Abb. 22). A) HPLC Elugramm des Rohproduktes ORN-1 nach der Festphasensynthese. B) HPLC Elugramm des Rohproduktes ORN-1' nach der Festphasensynthese. C) MALDI-ToF Analyse des aufgereinigten Produktes ORN-1 nach der Click-Reaktion D) Elugramme der aufgereinigten ORN-1 (rot) und -1' (grün).

Die Mehrzahl der in dieser Arbeit synthetisierten ORN sollte als *Baits* in Pulldown-Experimenten genutzt werden. Um das gesamte Interaktom nur teilweise definierter Sequenzmotive (z. B. ARACH, wobei R = A oder G; H = U, C oder A) analysieren können wurden degenerierte ORN benötigt. Diese wurden hergestellt, indem äquimolare Mischungen der Phosphoramidite an den zu randomisierenden Positionen verwendet wurden. In dem genannten Beispiel wurden Mischungen der Amidite G und A (R) bzw. A, C und U (H) eingesetzt, um eine Mischung verschiedener Sequenzen zu erzeugen, die alle möglichen Motive und deren Kombinationen enthalten, welche in dem aus fünf Nukleotiden bestehenden, degenerierten Motiv (A G/A A C A/C/U) enthalten sind. Um eine hohe Affinität

von *Reader* Proteinen sicherzustellen wurde das Sequenzmotiv vierfach wiederholt (5'-Biotin-(A G/A A C A/C/U)₄A-3'). Das Produkt dieser Synthese zeigt notwendigerweise kein klar definiertes Laufverhalten in der HPL Chromatographie. Das allen Komponenten des Gemisches gemeinsame Biotin am 5'-Terminus sorgt aber für eine Verschiebung der Elutionszeit, die ausreichend ist, um das Produktgemisch von den aus der Synthese stammenden Abbruchfragmenten zu trennen. Die abschließende Charakterisierung zeigt eine Massenverteilung um den erwarteten Mittelpunkt von $m/z = 7144.8$ (**Abbildung 22 C**). Die Extremwerte vom Produkt mit dem geringsten Molekulargewicht 5'-Biotin-(AAACC)₄A-3' bis zu dem mit dem höchsten 5'-Biotin-(AGACA)₄A-3' sind nicht sichtbar. Da jedes individuelle Oligonukleotid in diesem Fall acht teilrandomisierte Positionen enthält ergeben sich 1296 mögliche Sequenzen, eine geringe Abundanz von einzelnen Motivkombinationen ist also zu erwarten. Abbruchfragmente (bis $m/z = 6918.6$) sind nach der Aufreinigung nicht vorhanden und ein Vergleich der methylierten Sequenzen mit den unmethylierten Kontrollsträngen zeigt die zu erwartende Verschiebung von + 55.5 Da (**Abbildung 22 D**).

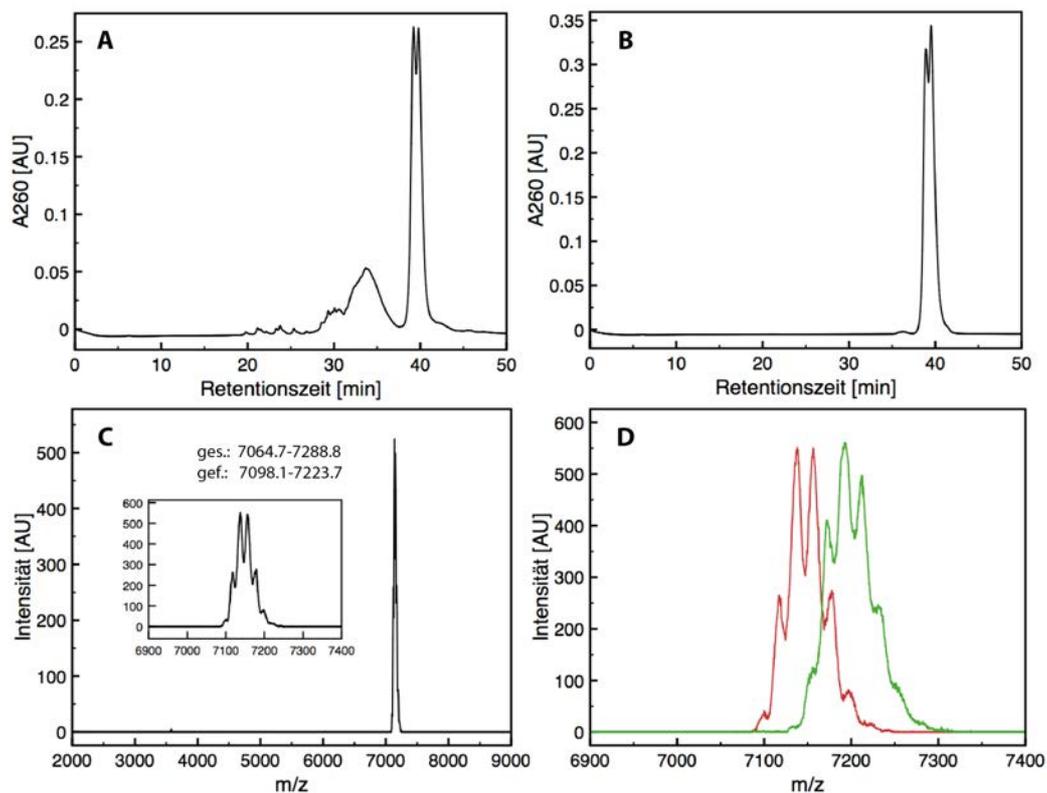


Abbildung 22: Charakterisierung des teilrandomisierten ORN 5'-Biotin-(ARACH)₄A-3'. A) HPL Chromatogramm des Rohproduktes nach Abspaltung und Entschützung. B) HPL Chromatogramm des aufgereinigten ORN. C) MALDI-ToF Analyse des aufgereinigten ORN. D) Überlagerung der MALDI-ToF Spektren des Kontrollstrang-Gemisches 5'-Biotin-(ARACH)₄A-3' (rot) mit dem korrespondierenden modifizierten ORN 5'-Biotin-(ARm⁶ACH)₄A-3' (grün).

Es konnten also zwei Populationen von Oligonukleotiden hergestellt werden, die alle möglichen Variationen des Konsensmotives enthielten und damit geeignet waren, die gesamte Bandbreite an möglichen Sequenzen in einem Experiment zu analysieren. Neben den gezeigten ORN mit dem Motiv (ARACH)₄ wurden ebenfalls solche mit dem an drei Positionen randomisierten Motiv (RRACH)₄ erfolgreich synthetisiert.

Die für Affinitätsstudien benötigten ORN wurde mit einem Biotin-Tag versehen. Drei grundlegende Ansätze wurden hierfür in Erwägung gezogen. Zum einen ist es möglich, durch einen Biotin-funktionalisierten Träger ein ORN mit 3'-Biotin zu erzeugen. Ein von *Glen Research* bezogener 3'-Biotin CPG Träger bot in der Festphasensynthese Ausbeuten, die ihn als ungeeignet erscheinen ließen (> 1% Gesamtausbeute).

Alternativ kann Biotin an das 5'-Ende angefügt werden, indem ein Biotin-Phosphoramidit (*Link Tech*, **32**, **Abbildung 23**) verwendet wird, oder ein Biotin-Azid (*Sigma Aldrich*) an eine zuvor inkorporierte Alkinbase mittels Click-Chemie gebunden wird (**33**). Beide Ansätze wurden getestet und die direkte Biotinylierung erwies sich als deutlich überlegen in den erzielbaren Ausbeuten. Ein direkter Einbau lieferte bis zu 14% isolierte Gesamtausbeute, während ein nachträglicher Click nicht über 2-3% ermöglichte. Daher wurden nahezu alle in dieser Arbeit verwendeten Sequenzen durch direkte Inkorporation biotinyliert. Ein Beispiel für ein mittels Click-Chemie biotinyliertes Paar von Strängen ist mit ORN-1 und ORN-1' in **Abbildung 21** dargestellt.

Auch ein durch Cycloaddition verknüpfter kovalenter *Crosslinker* wurde getestet. Er wurde von *Barbara Steigenberger* entwickelt und enthält einen reaktiven NHS-Ester (NHS = *N*-Hydroxysuccinimid), der bevorzugt mit ϵ -Aminogruppen von Lysinen reagiert. Der *Crosslinker* sollte in Oligonukleotiden dafür sorgen, dass bindende Proteine kovalent an den verwendeten Strang gebunden werden, was zum einen eine bessere Identifikation von transienten Bindern ermöglichen sollte. Zum anderen würde er eine Unterscheidung zwischen direkten Bindern des *Baits* und indirekten, die lediglich als Komponenten eines Komplexes angereichert werden, ermöglichen. Im besten Fall würde hierdurch eine zusätzliche Information zugänglich, die mit üblichen Methoden nicht zu ermitteln ist. Da der NHS-Ester stark hydrolyseempfindlich ist kann er nicht in der Festphasensynthese in Form eines Phosphoramidits inkorporiert werden, sondern wurde als Azid direkt vor den Interaktionsstudien mit einer Alkinbase mittels Click-Chemie verknüpft (**34**).

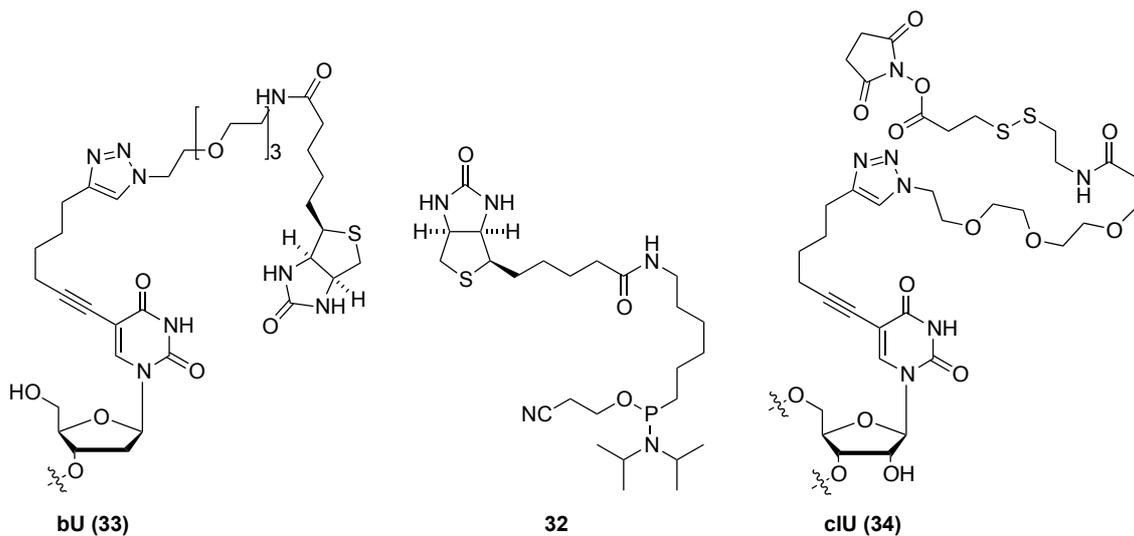


Abbildung 23: Strukturen der verwendeten Biotineinheiten und des eingesetzten kovalenten Crosslinkers.

Er sollte durch eine Reaktion mit *Reader* Proteinen eine höhere Anreicherung und gleichzeitig eine spezifische Markierung von direkt, im Gegensatz zu indirekt bindenden Proteinen leisten.

Alle im Lauf dieser Dissertation hergestellten ORN wurden nach den beschriebenen Methoden hergestellt und gereinigt. Für eine Vielzahl unterschiedlicher Anwendungen wurden Oligonukleotide unterschiedlicher Länge und Sequenz benötigt. Eine Liste ist in der folgenden Tabelle zu finden.

Tabelle 1: Sequenzen aller in dieser Dissertation hergestellten Oligoribonukleotide. (bU = Click-Biotin **33**; B = 5'-Biotin (Link Tech); cIU = Crosslinker **34**; Cy3 = Cyanine-3 (Link Tech); R = G/A; H = U/C/A).

ORN #	Sequenz
ORN-1	5'-bU GGG GAU UGA Ams ^{2;6} i ⁶ AA UCC CC-3'
ORN-1'	5'-bU GGG GAU UGA AAA UCC CC-3'
ORN-2	5'-B GGG GAU UGA Ams ^{2;6} i ⁶ AA UCC CC-3'
ORN-2'	5'-B GGG GAU UGA AAA UCC CC-3'
ORN-2''	5'-B GGG GAU UGA Ai ⁶ AA UCC CC-3'
ORN-3	5'-B GCG AUU cIU Ams ^{2;6} i ⁶ A AAU CAC C-5'
ORN-3'	5'-B GCG AUU cIU AA AAU CAC C-5'
ORN-4	5'-B GAC UUG AAms ^{2;6} i ⁶ A AUU GUA Ams ^{2;6} i ⁶ AA UCA U-3
ORN-4'	5'-B GAC UUG AAi ⁶ A AUU GUA Ai ⁶ AA UCA U-3
ORN-4''	5'-B GAC UUG AAA AUU GUA AAA UCA U-3

<i>ORN-5</i>	5'-B GG ^{ms²i⁶A} CUG G ^{ms²i⁶AC} UGG ^{ms²i⁶ACU} GG ^{ms²i⁶A} CU-3'
<i>ORN-5'</i>	5'-B GGA CUG GAC UGG ACU GGA CU-3'
<i>ORN-6</i>	5'-Cy3 GGG GAU UGA A ^{ms²i⁶AA} UCC CC-3'
<i>ORN-6'</i>	5'-Cy3 GGG GAU UGA AAA UCC CC-3'
<i>ORN-7</i>	5'-Cy3 GUG CAU UGA A ^{ms²i⁶AA} UCU AC-3'
<i>ORN-8</i>	5'-CAC UGU CAU CCC UUU GUC CUA ACC A-3'
<i>ORN-9</i>	5'-Cy3 CAC UGU CAU CCC UUU GUC CUA ACC A-3'
<i>ORN-9'</i>	5'-Cy3 CAC UGU CAU CCC UUU GUC CU ^{ms²i⁶A} ACC A-3'
<i>ORN-9''</i>	5'-Cy3 CAC UGU C ^{ms²i⁶AU} CCC UUU GUC CUA ACC A-3'
<i>ORN-10</i>	5'-bU GG ^{m⁶A} CUG G ^{m⁶AC} UGG ^{m⁶ACU} GG ^{m⁶A} CU-3'
<i>ORN-10'</i>	5'-bU GGA CUG GAC UGG ACU GGA CU-3'
<i>ORN-11</i>	5'-B GG ^{m⁶A} CUG G ^{m⁶AC} UGG ^{m⁶ACU} GG ^{m⁶A} CU-3'
<i>ORN-11'</i>	5'-B GGA CUG GAC UGG ACU GGA CU-3'
<i>ORN-12</i>	5'-B GG ^{m⁶A} CUU AAG G ^{m⁶AC} UUG AGG ^{m⁶ACU} UAG GG ^{m⁶A} CU -3'
<i>ORN-12'</i>	5'-B GGA CUU AAG GAC UUG AGG ACU UAG GGA CU -3'
<i>ORN-13</i>	5'-B GA ^{m⁶A} CUG A ^{m⁶AC} UGA ^{m⁶ACU} GA ^{m⁶A} CU-3'
<i>ORN-13'</i>	5'-B GAA CUG AAC UGA ACU GAA CU-3'
<i>ORN-14</i>	5'-B CGA AUG G ^{m⁶AC} UclUG ^{Gm⁶AC} UGA-3'
<i>ORN-14'</i>	5'-B CGA AUG GAC UclUG GAC UGA-3'
<i>ORN-15</i>	5'-B RR ^{m⁶A} CHR R ^{m⁶AC} HRR ^{m⁶ACH} RR ^{m⁶A} CHA -3'
<i>ORN-15'</i>	5'-B RRA CHR RAC HRR ACH RRA CHA -3'
<i>ORN-16</i>	5'-B AR ^{m⁶A} CHA R ^{m⁶AC} HAR ^{m⁶ACH} AR ^{m⁶A} CHA -3'
<i>ORN-16'</i>	5'-B ARA CHA RAC HAR ACH ARA CHA -3'
<i>ORN-17</i>	5'-B GG ^{m₂⁶A} CUG G ^{m₂⁶AC} UGG ^{m₂⁶ACU} GG ^{m₂⁶A} CU-3'
<i>ORN-17'</i>	5'-B GGA CUG GAC UGG ACU GGA CU-3'
<i>ORN-18</i>	5'-B GG ^{m₂⁶A} CUU AAG G ^{m₂⁶AC} UUG AGG ^{m₂⁶ACU} UAG GG ^{m₂⁶A} CU-3'
<i>ORN-18'</i>	5'-B GGA CUU AAG GAC UUG AGG ACU UAG GGA CU-3'
<i>ORN-19</i>	5'-GG ^{m₂⁶A} CU-3'
<i>ORN-20</i>	5'-GUG ^{m₂⁶Am₂⁶AC} CU-3'
<i>ORN-21</i>	5'-GG ^{m⁶A} CU-3'
<i>ORN-22</i>	5'-GUG ^{m⁶Am⁶AC} CU-3'
<i>ORN-23</i>	5'-B UGC GUmU CUC AAUm ACC UGUm CUC GA-3'
<i>ORN-23'</i>	5'-B UGC GUU CUC AAU ACC UGU CUC GA-3'
<i>ORN-24</i>	5'-B UGC GAmU CUC AAAm ACC UGAm CUC GA-3'

<i>ORN-24'</i>	5'-B UGC GAU CUC AAA ACC UGA CUC GA-3'
<i>ORN-25</i>	5'-B UGC GCmU CUC AACm ACC UGCm CUC GA-3'
<i>ORN-25'</i>	5'-B UGC GCU CUC AAC ACC UGC CUC GA-3'
<i>ORN-26</i>	5'-B UGC GmU CUC AA Gm ACC UGm CUC GA-3'
<i>ORN-26'</i>	5'-B UGC GGU CUC AAG ACC UGG CUC GA-3'
<i>ORN-27</i>	5'-B CAA UGΨ AGC AAU GΨA GCA AUG ΨAG -3'
<i>ORN-27'</i>	5'-B CAA UGU AGC AAU GUA GCA AUG UAG -3'
<i>ORN-28</i>	5'-B AGA GUΨ CAG AGU ΨCA GAG ΨUC -3'
<i>ORN-28'</i>	5'-B AGA GUU CAG AGU UCA GAG UUC -3'

5. Interaktomstudien zu den Modifikationen m⁶A, m⁶₂A, i⁶A, ms²i⁶A, Ψ, Am, Gm, Cm und Um

5.1 Allgemeine Methode des SILAC Experimentes

Das Ziel der Experimente war es, aus Zelllysaten Proteine anzureichern, die spezifisch an ausgewählte RNA-Modifikationen binden, um sie zu identifizieren und quantitativ zu analysieren. Sollen Proben, die komplexe Mischungen von Proteinen enthalten, wie Zelllysate, oder aus diesen angereicherte Proteinmischungen, miteinander verglichen werden, ist eine verlässliche Quantifizierung nicht trivial. Zunächst ist es nicht möglich, die an einen RNA-„Köder“ (*Bait*) spezifisch gebundenen Proteine durch eine Einzelmessung zuverlässig zu identifizieren. Unspezifisch gebundene Proteine, allgemein RNA-bindende Proteine oder solche, die in den verwendeten Lysaten hoch abundant sind werden in hohen Quantitäten an den verwendeten Köder binden, ohne notwendigerweise spezifische Interaktoren darzustellen. Dies macht eine Vergleichsmessung zweier Proben nötig. Die isolierten Proteine des modifizierten *Baits* müssen also zuverlässig quantitativ mit denen eines unmodifizierten Kontrollstranges verglichen werden können.

SILAC (*Stable isotope labeling with amino acids in cells culture*) stellt eine etablierte und vielfach angewandte Methode zur Mengenbestimmung mittels Isotopenmarkierung in vergleichenden *Proteomics*-Experimenten dar.³⁰¹ Zwei Kulturen identischer Zellen werden in Nährmedien kultiviert, die in einem Fall mit Schwerisotopen-markierten Aminosäuren versetzt sind, im zweiten Fall hingegen natürliche („leichte“) Aminosäuren enthalten. Geeignete und in der Regel verwendete Aminosäuren sind ¹³C₆-Lysin, ¹³C₆-¹⁵N₂-Lysin ¹³C₆-Arginin, ¹³C₆-¹⁵N₄-Arginin, ¹³C₉-Tyrosin und ¹³C-²H₃-Methionin bzw. Kombinationen aus ihnen.²⁷¹ Man erhält auf diese Weise Zelllinien, die phänotypisch identisch sind, deren Proteine aber interne Markierungen tragen, die sie massenspektrometrisch unterscheidbar machen, auch wenn sie gemeinsam in einer Probe analysiert werden. Analog zu der Verwendung eines Barcodes ist so jedes Peptid der komplexen Mischung einer Quelle (Kultur/Experiment) zuzuordnen. Hierdurch wird das Vereinigen mehrerer Experimente zu einer Probe an einem frühen Zeitpunkt im Ablauf eines Experimentes möglich, was Präzision und Reproduzierbarkeit der Analyse im Vergleich zu einer nachträglichen Markierung (z.B. iTRAQ) erhöht.³⁰²

In den hier beschriebenen Experimenten wurden modifizierte Stränge bzw. Kontrollstränge mit SILAC-Lysaten inkubiert, die *Reader*-Proteine isoliert, anschließend beide Proben vereinigt, tryptisch verdaut und gemeinsam massenspektrometrisch analysiert. Aus dem Verhältnis zwischen „schwerem“ Peptid und „leichtem“ Peptid lässt sich die relative Menge jedes identifizierten Proteins bestimmen. Um eine hohe Qualität der Ergebnisse sicherzustellen wurden Experimente jeweils „vorwärts“ und „rückwärts“ durchgeführt („*label swap* Experiment“) und die Ergebnisse gemeinsam in einem *Scatterplot* dargestellt (siehe **Abbildung 24**). Im Ergebnis wird schließlich sichtbar, welche Proteine bevorzugt an die modifizierte Sequenz binden (*Reader*, spezifische Binder) und welche bevorzugt das Kontroll-ORN binden (*repelled*, abgestoßene Proteine). Die Experimente wurden von *Raghu Edupuganti* in der Gruppe von *Michiel Vermeulen* durchgeführt.

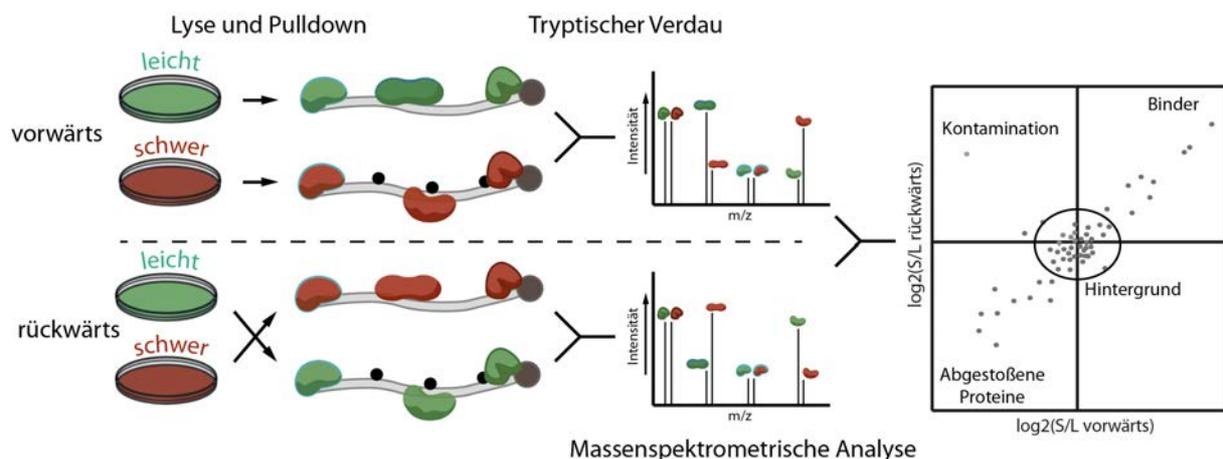


Abbildung 24: Arbeitsablauf eines SILAC-Experimentes. Zwei Zellkulturen werden in Medium kultiviert, welches entweder natürliche ("leichte", grün) oder schwerisotopenmarkierte ("schwere", rot) Aminosäuren enthält. Aus beiden werden Lysate gewonnen und diese dann genutzt um die interagierenden Proteine eines modifizierten Stranges, bzw. des korrespondierenden Kontrollstranges anzureichern. Nach einem tryptischen Verdau werden angereicherte Proteine massenspektrometrisch identifiziert und quantifiziert.

Die verwendeten Zellextrakte wurden als Mischung zahlreicher Kulturgefäße einer Zelllinie hergestellt und dann aus dieser Mischung Experimente im Triplikate durchgeführt. Die Daten resultieren also aus technischen Triplikaten, die Vorgehensweise ist bereits beschrieben worden.^{23,303}

Um unspezifische Interaktionen in *Pulldown*-Experimenten zu minimieren ist es üblich Kompetitoren einzusetzen. In allen im Weiteren beschriebenen Experimenten wurde tRNA aus Hefe hierfür eingesetzt, sofern es nicht anders beschrieben ist. Außerdem kamen, neben obligatorischen Proteaseinhibitoren, Phosphataseinhibitoren zum Einsatz, um einen nativen Zustand regulatoriver Protein-Phosphorylierung zu erhalten.

5.2 Interaktomstudien zu m⁶A

5.2.1 SILAC-Experimente

Neben der ersten Hochdurchsatzstudie zum Sequenzkontext von m⁶A in mRNA publizierten *Dominissini et al.* 2012 eine massenspektrometrische Analyse seines Interaktoms.⁷¹ Sie identifizierten YTHDF2, YTHDF3 und ELAVL1 (auch HUR) als spezifische *Reader* Proteine. Abgesehen von dieser initialen Studie gibt es bislang keine umfangreiche Untersuchung der Gesamtheit interagierender Proteine, insbesondere nicht unter Berücksichtigung unterschiedliche Sequenzkontexte. Um eine solche Analyse zu leisten, wurden verschiedene ORN mit homogenen und gemischten Sequenzmotiven für die durchgeführten SILAC Studien verwendet (**Tabelle 2**).

Tabelle 2: Für die Analyse des Interaktoms der Modifikation m⁶A hergestellte Oligoribonukleotide (bU = Click-Biotin **33**, cIU = Crosslinker **34**; B = 5'-Biotineinheit (*Link Tech*); R = G/A; H = U/C/A).

ORN	Sequenz
ORN-10	5'-bU GGm ⁶ ACU GGm ⁶ ACU GGm ⁶ ACU GGm ⁶ ACU-3'
ORN-11	5'-B GGm ⁶ ACU GGm ⁶ ACU GGm ⁶ ACU GGm ⁶ ACU-3'
ORN-12	5'-B GGm ⁶ ACU <u>UAA</u> GGm ⁶ ACU <u>UGA</u> GGm ⁶ ACU <u>UAG</u> GGm ⁶ ACU -3'
ORN-13	5'-B GAm ⁶ ACU GAm ⁶ ACU GAm ⁶ ACU GAm ⁶ ACU-3'
ORN-14	5'-B CGA AUG Gm ⁶ ACU cIUG Gm ⁶ ACU GA-3'
ORN-15	5'-B RRm ⁶ ACH RRm ⁶ ACH RRm ⁶ ACH RRm ⁶ ACH A -3'
ORN-16	5'-B ARm ⁶ ACH ARm ⁶ ACH ARm ⁶ ACH ARm ⁶ ACH A -3'

Mit ORN-10 und 11, welche sich lediglich durch die Art ihres Biotin-*Tags* unterscheiden, wurden *Pulldowns* aus den Lysaten von drei Zelltypen durchgeführt (HeLa, *mouse embryonic stem cells* (mESC) und *Neural progenitor cells* (NPC)). Sie enthalten das dominante Konsensmotiv GGACU. Zwischen den beiden verschiedenen *Tags* waren keine signifikanten Unterschiede feststellbar, daher wird im Folgenden zwischen den Ergebnissen nicht unterschieden.

Es wurden jeweils zytosolische Extrakte bzw. Zellkernextrakte verwendet, um eine grobe Lokalisation der Interaktion zu ermöglichen. Bei der Interpretation der Ergebnisse sollte aber berücksichtigt werden, dass selbst eine sorgfältig durchgeführte getrennte Lyse Fraktionen erzeugt, welche signifikante Verunreinigungen durch Proteine der abgetrennten Zellkompartimente enthalten. Die Daten zeigen zum einen, welche Proteine bevorzugt an

modifizierte Sequenzen binden, zum anderen aber auch welche von m⁶A abgestoßen werden, also mit höherer Affinität an den Kontrollstrang binden (**Abbildung 25 bis 27**).

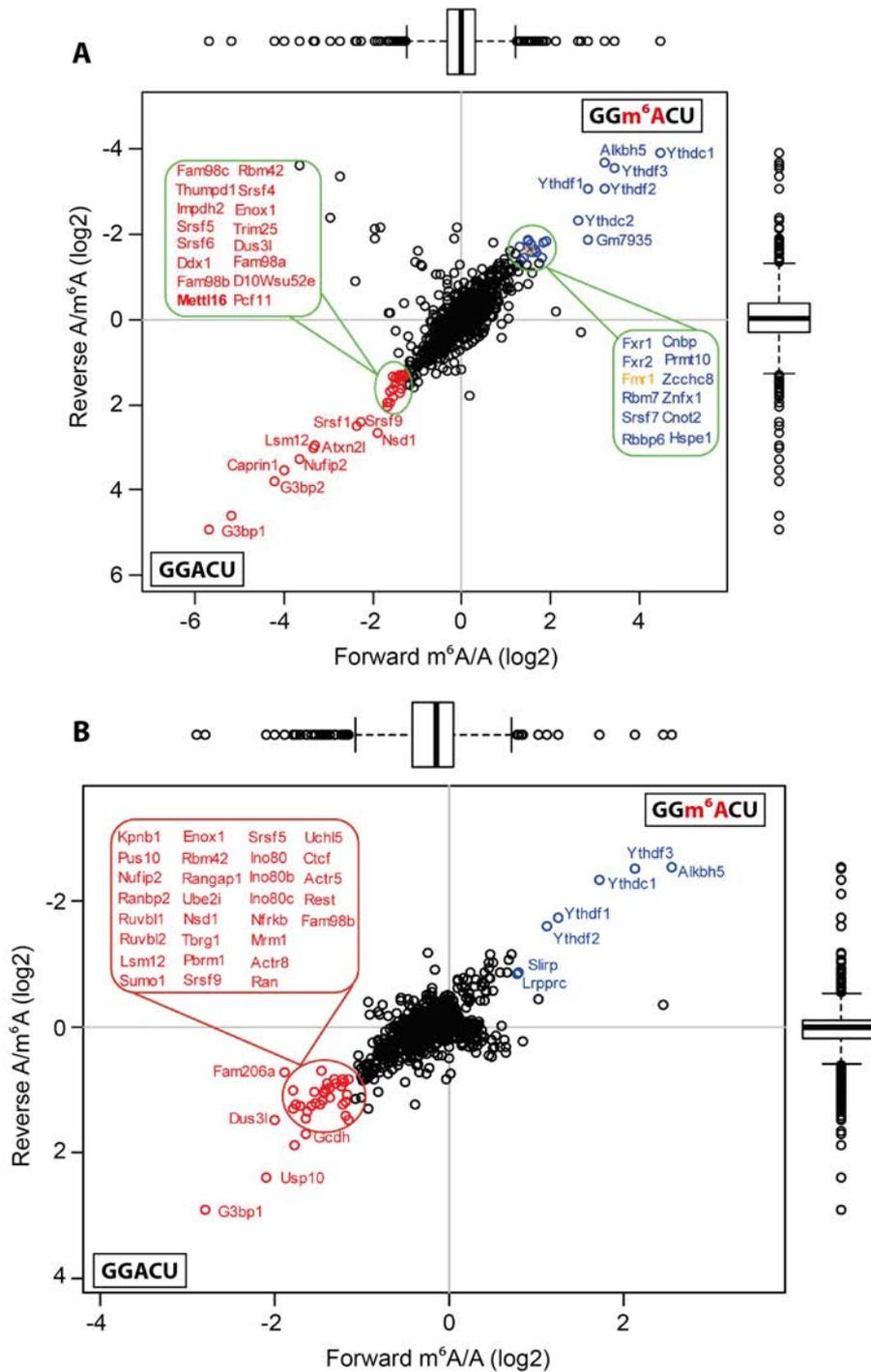


Abbildung 25: Scatterplots SILAC-basierter *Pulldowns* aus mESC Extrakten. Als *Bait* wurde ORN-11 verwendet: 5'-Biotin-(GGm⁶ACU)₄-3'. A) *Pulldown* aus zytosolischem mESC Lysat. B) *Pulldown* aus mESC Kernextrakt. Die Experimente wurden von Raghu Edupuganti durchgeführt.

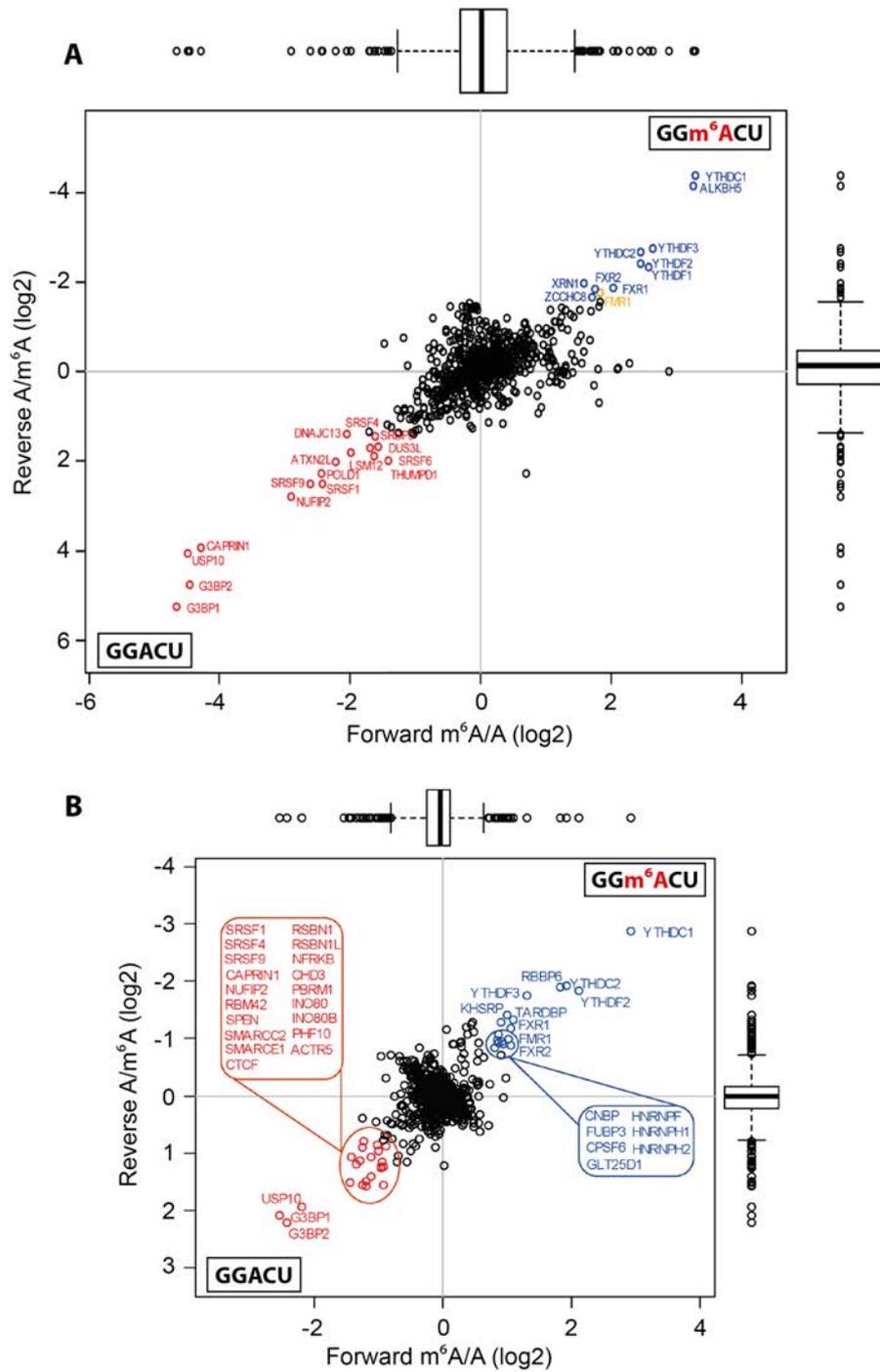


Abbildung 26: Scatterplots SILAC-basierter *Pulldowns* aus HeLa Extrakten. Als *Bait* wurde ORN-10 verwendet: 5'-Biotin-(GGm⁶ACU)₄-3'. A) *Pulldown* aus zytosolischem HeLa Lysat. B) *Pulldown* aus HeLa Kernextrakt. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

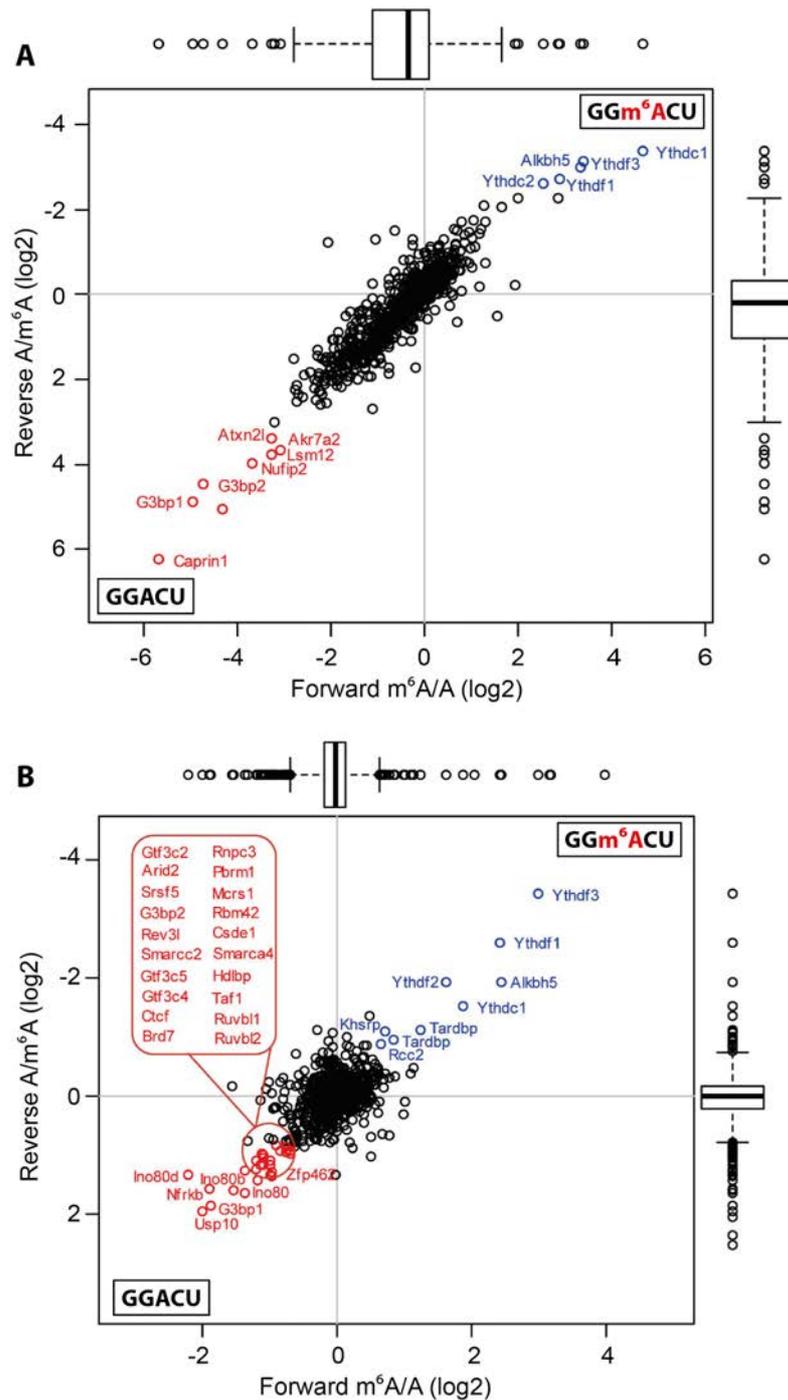


Abbildung 27: Scatterplots SILAC-basierter *Pulldowns* aus NPC Extrakten. Als *Bait* wurde ORN-11 verwendet: 5'-Biotin-(GGm⁶ACU)₄-3'. A) *Pulldown* aus zytosolischem NPC Lysat. B) *Pulldown* aus NPC Kernextrakt. Die Experimente wurden von Raghu Edupuganti durchgeführt.

Die Resultate der SILAC *Pulldowns* zeigen ein robustes Interaktom für m⁶A, welches von der Maus zum Menschen weitgehend konserviert ist. Zu den Interaktoren gehören neben den zu Beginn dieser Arbeit bereits bekannten YTHDF2 und 3 auch die anderen Mitglieder der YTH-Familie YTHDF1, YTHDC1 und YTHDC2, die inzwischen alle als spezifische Inter-

aktoren publiziert sind.^{144,147,155,163} Wie erwartet, ist die Demethylase ALKBH5 als Interaktor sichtbar, FTO hingegen nicht. Da die eingesetzten ORN im Verlauf der hier vorgestellten Experimente vor der massenspektrometrischen Analyse rigoros gewaschen werden, um unspezifische Verunreinigungen zu minimieren, können transiente Binder verloren gehen. Ein Vergleich der katalytischen Aktivität von FTO und ALKBH5 zeigt, dass nur FTO detektierbare Intermediate (hm⁶A, f⁶A) erzeugt, ALKBH5 dagegen nicht. Dies stützt die Vermutung, dass ALKBH5 m⁶A und seine oxidierten Derivate fester bindet als FTO, ohne Zwischenstufen in das Medium zu entlassen.¹³⁸

In mESC und HeLa, nicht aber NPCs, sind FMR1, FXR1 und FXR2 bislang nicht bekannte *Reader* (FMR1 = *Fragile X mental retardation protein 1*, auch FMRP; FXR1/2 = *Fragile X mental retardation syndrome-related protein 1*). Das Fragile-X-Syndrom (auch *Martin-Bell-Syndrom*) stellt eine der meistverbreiteten Ursachen für genetisch bedingte geistige Retardierung und Autismus dar.³⁰⁴ Es wird weit überwiegend durch eine Expansion von *CGG repeats* im 5'-UTR des FMR1 Gens ausgelöst, die in einer Hypermethylierung und transkriptionellen Inaktivierung des Gens resultieren. Mit dem Grad der Expansion variiert der phänotypische Schweregrad, die Lokalisation des Gens auf dem X Chromosom ist für stärker ausgeprägte Phänotypen in männlichen Individuen verantwortlich. Das Protein FMR1 bindet spezifisch mRNAs, beeinflusst ihre Lokalisation und unterdrückt ihre Translation.³⁰⁵⁻³⁰⁸ Die homologen Proteine FXR1 und FXR2³⁰⁹ sind deutlich weniger gut erforscht. Eine spezifische Attraktion von FMR1, FXR1 und FXR2 durch m⁶A stellt einen bislang unbekanntem Zusammenhang zwischen dieser Modifikation und den mit ihr assoziierten Proteinen her.

In NPCs wird die Expression von FMR1 durch die miRNA-130b unterdrückt³¹⁰, was eine Erklärung für das Fehlen von FMR1 im NPC *Pulldown* (**Abbildung 27**) sein könnte. Insgesamt gibt es eine Gruppe vom Zelltyp unabhängiger *Reader* und abgestoßener Proteine, während eine erhebliche Anzahl identifizierter Proteine spezifisch für den jeweiligen Zelltyp zu sein scheint (**Abbildung 28**). Insgesamt ist der Anteil der mindestens in zwei aus drei Experimenten identifizierten Proteinen unter den *Readern* (45%) etwas höher als unter den abgestoßenen Proteinen (32%).

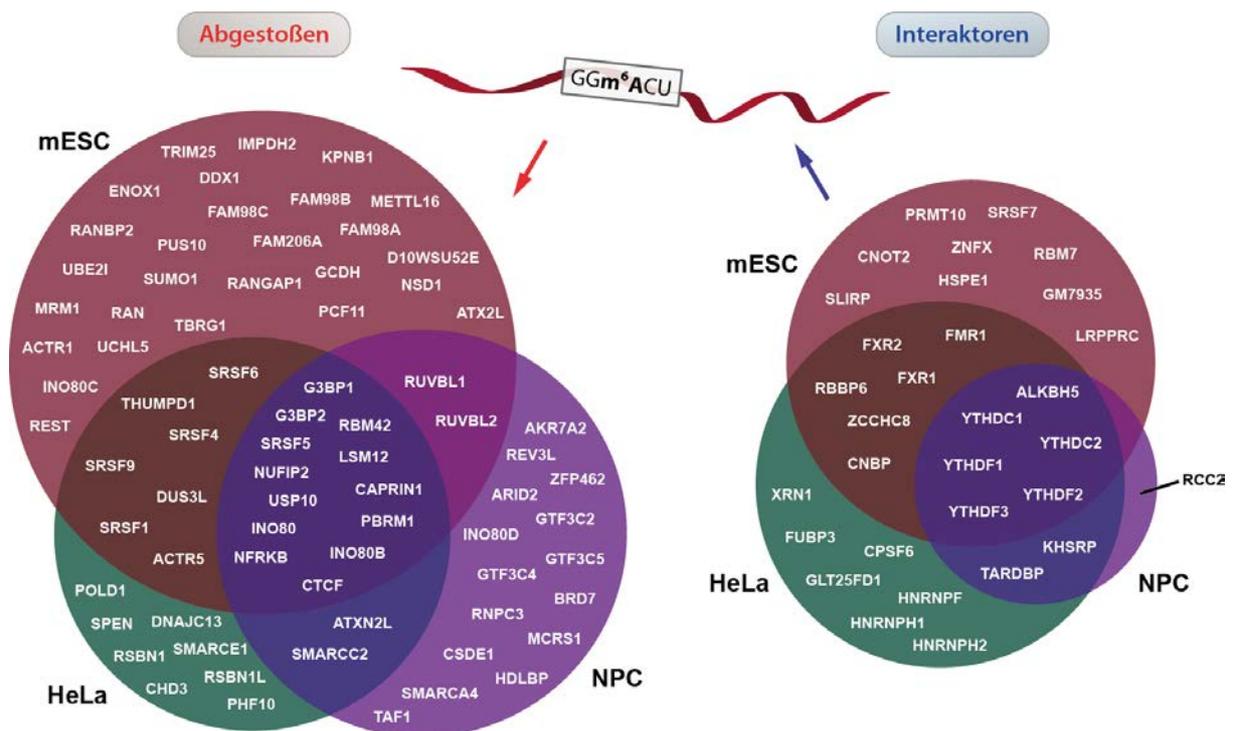


Abbildung 28: Visualisierung des Überlappungsgrades zwischen den Interaktomen für das Sequenzmotiv GGACU in drei unterschiedlichen Zelltypen (mESC, HeLa und NPC). Für die Analyse wurden identifizierte Proteine aus zytosolischen Lysaten und Kernextrakten des jeweiligen Zelltyps zusammengefasst.

Abgesehen von den spezifischen *Reader*-Proteinen können auch die von m⁶A abgestoßenen Proteine Informationen zu seiner biologischen Funktion liefern. Diese Proteine sind nur im Zusammenhang mit Sequenzinformationen sinnvoll interpretierbar, eine herabgesetzte Affinität muss nicht notwendigerweise aus der untersuchten Modifikation resultieren, sondern ist auch für anders substituierte Derivate denkbar (hier: allgemein N6-substituierte Adenosine). Für bekannte, modifizierte Sequenzmotive wird aber durchaus ein individuell moduliertes Interaktom für spezifische Motive sichtbar, welches eine funktionelle Relevanz besitzen kann, wie im Weiteren auch gezeigt werden konnte.

Hier stehen besonders G3BP1, G3BP2, CAPRIN1, USP10 und NUFIP2 ins Auge (G3BP1/2 = *Ras GTPase-activating protein-binding protein 2*; CAPRIN1 = *Cytoplasmic activation- and proliferation-associated protein 1*; USP10 = *Ubiquitin carboxyl-terminal hydrolase 10*; NUFIP2 = *Nuclear fragile X mental retardation-interacting protein 2*). Außerdem sind die Splicingfaktoren SRSF1/4/5/6/9 (*Serine/arginine-rich splicing factor 1/4/5/6/9*) in mindestens zwei untersuchten Zelltypen angereichert.

Eine molekulare Interaktion, welche einen funktionellen Zusammenhang vermitteln könnte, ist für CAPRIN1, FMR1, G3BP1/2 und USP10 bekannt³¹¹⁻³¹³, auch NUFIP2 bindet FMR1 *in vivo*.³¹⁴ Sie alle sind Komponenten von Ribonukleoprotein-Komplexen (RNPs) in Stressgranula (Stress-induzierte Aggregationen aus RNA und Proteinen innerhalb des

Zytosols).^{315,316} G3BP1 bindet USP10 und inhibiert es.³¹⁷ Es liegt also die Vermutung nahe, dass die identifizierten Proteine Komponenten eines gemeinsamen, dynamischen mRNPs sind.

Neben der Abhängigkeit des Interaktoms vom Zelltyp sollte auch untersucht werden, inwieweit es durch den Sequenzkontext beeinflusst wird. *Pulldowns* mit ORN-13 (5'-Biotin-(GAACU)₄-3') als *Bait* sollten zeigen, ob das detektierbare Interaktom von m⁶A durch den Austausch einer einzelnen Base moduliert wird. Neben GGACU stellt GAACU das wichtigste Motiv dar, in dem m⁶A zu finden ist.

Die Experimente zeigen eine partielle Übereinstimmung für beide untersuchten Kontexte. Auch im GAACU-Kontext ist die gesamte YTH-Familie im Interaktom vorhanden, ihre relativen Affinitäten verändern sich jedoch. Während YTHDC1 in zytosolischen Lysaten der stärkste Binder für GGm⁶ACU ist (**Abbildung 25-27**), scheint seine Affinität für GAm⁶ACU herabgesetzt zu sein (**Abbildung 29**). Übereinstimmend hiermit konnten *Xu et al.* zeigen, dass YTHDC1, im Gegensatz zu YTHDF1 und YTHDF2, eine signifikant geringere Affinität zu GAm⁶ACU als zu GGm⁶ACU besitzt.¹⁴⁷ Weiterhin stellten sie fest, dass von den vier untersuchten YTH-Proteinen (YTHDF1/2, YTHDC1/2) YTHDC2 am schwächsten bindet, was ebenfalls in guter Übereinstimmung mit den hier gezeigten Daten ist. Dies belegt anschaulich, dass die angewandte Methode nicht nur geeignet ist, bindende Proteine zu identifizieren, sondern auch ihre relative Affinität abbilden kann. Auffällig ist auch, dass ALKBH5 weniger ausgeprägt angereichert ist. Dies deutet an, dass auch die Demethylase eine Präferenz für GGm⁶ACU über GAm⁶ACU besitzt, das erste Motiv also schneller Demethyliert werden könnte. Im Kontrast hierzu können *Zou et al.* keinen klaren Unterschied in der katalytischen Aktivität von FTO oder ALKBH5 zwischen Gm⁶AC und Am⁶AC Motiven feststellen.³¹⁸ Relative Bindungsaffinität und katalytische Aktivität müssen aber nicht notwendigerweise korrelieren.

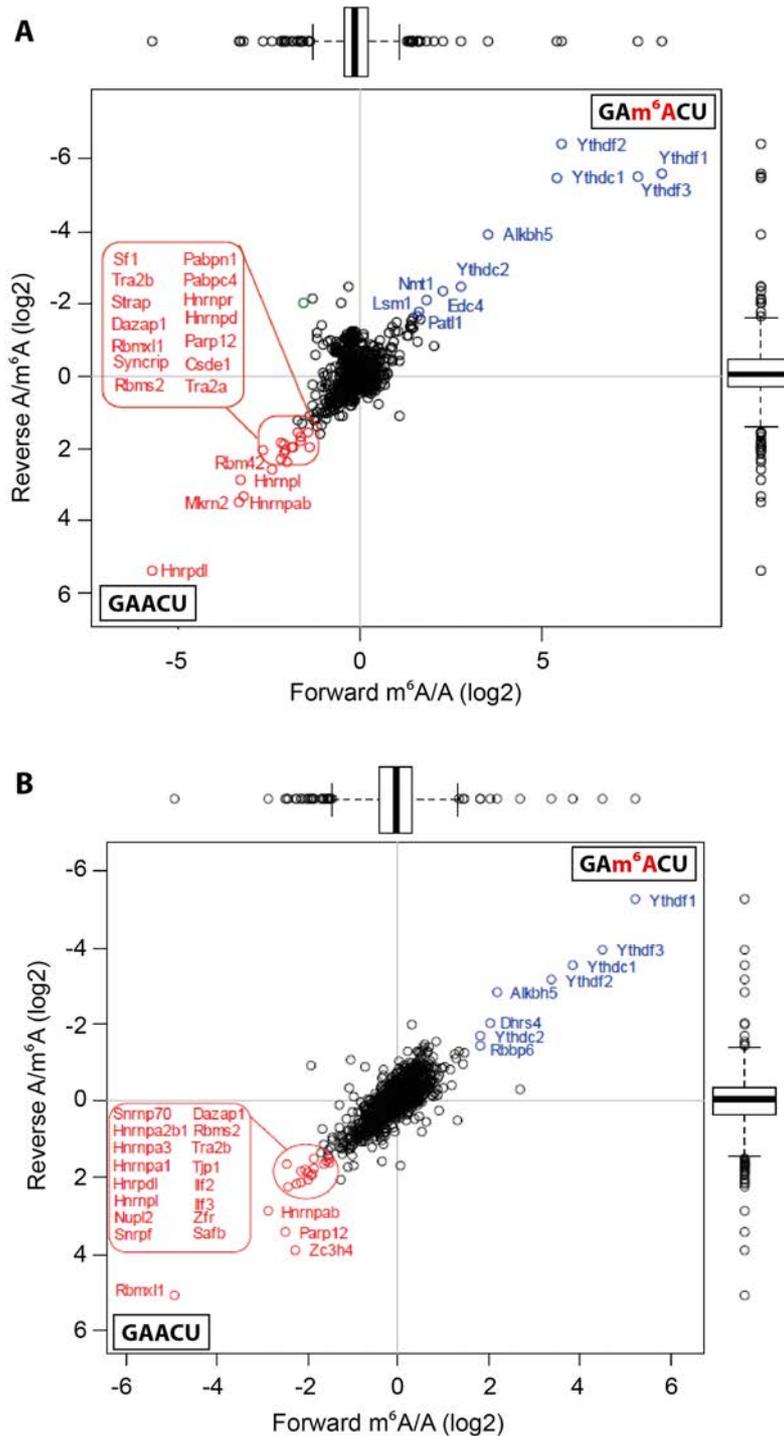


Abbildung 29: Scatterplots SILAC-basierter *Pulldowns* für ORN-13: 5'-Biotin-(GAACU)₄-3'. A) *Pulldown* aus zytosolischem mESC Lysat. B) *Pulldown* aus 3T3 (murine Fibroblasten) Kernextrakt. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Die Gruppen der abgestoßenen Proteine zeigen wenig Übereinstimmung. Für GAACU geht aus beiden Experimenten eine übereinstimmende Gruppe aus HNRNPDL, HNRNPAB, PARP12, RBMXL1, DAZAP1 und TRA2B (HNRNPDL = *Heterogeneous nuclear ribonucleoprotein D-like*; HNRNPAB = *Heterogeneous nuclear ribonucleoprotein A/B*; PARP12 = *Poly [ADP-ribose] polymerase 12*; RBMXL1 = *RNA binding motif protein, X-*

linked-like-1; DAZAP1 = *DAZ-associated protein 1*; TRA2B = *Transformer-2 protein homolog beta*). Alle sind bekannte poly(A)-RNA bindende Proteine. Als Gruppe können sie keinem bekannten Prozess zugeordnet werden, mit DAZAP1, TRA2B und RBMXL1 sind drei von ihnen in mRNA *splicing* involviert.

In Kernextrakt aus 3T3-Zellen (murinen Fibroblasten) ist HNRNPA2B1 signifikant angereichert (**Abbildung 29 B**), was überraschend ist, da *Alarcón et al.* eine direkte positive Assoziation zwischen diesem kernlokalisierten Protein und m⁶A-modifizierter RNA beobachten konnten.¹⁴⁶ Sie beziehen sich in ihrer Arbeit auf modifizierte R⁶Gm⁶AC Motive. Die hier gezeigten Ergebnisse sprechen für ein abweichendes Verhalten im RAm⁶AC Kontext.

Da m⁶A in der unmittelbaren Umgebung des Stop Codons stark angereichert ist^{71,72,137} (siehe auch **Abbildung 11**) wurde mit ORN-12 die Auswirkung dieser Assoziation untersucht. In diesem Fall enthielt der *Bait* das Sequenzmotiv GGm⁶ACU, direkt gefolgt von einem Stop Codon (unterstrichen): 5'-B GGm⁶ACU UAA GGm⁶ACU UGA GGm⁶ACU UAG GGm⁶ACU -3'.

Für diese Sequenz zeigt sich übereinstimmend in humanem und murinem Lysat ALKBH5 sowie die YTH-Familie als Interaktoren, mit YTHDC1 als stärkstem Binder (**Abbildung 30**). Die mit dem Fragilen-X-Syndrom in Verbindung stehenden Proteine fehlen. Unter den Proteinen, die stärker an den Kontrollstrang binden finden sich neben den auch ohne Stop Codon vorhandenen G3BP1, G3BP2 und USP10 noch ATIC (*Bifunctional purine biosynthesis Protein PURH*) und HDLBP (*High density lipoprotein-binding protein*, auch Vigilin). Die letztgenannten sind in Lipid- bzw. Sterolmetabolismus involviert und besitzen keine bekannte Funktion im Zusammenhang mit mRNA. HDLBP/Vigilin verfügt über die ungewöhnlich hohe Anzahl von 14 KH-Domänen, die typisch für RNA-bindende Proteine sind³¹⁹ und Sequenzspezifität vermitteln können. Beachtenswert erscheint hier, dass neben den Fragiles-X assoziierten *Readern* FMR1, FXR1 und FXR2 auch das abgestoßene CAPRIN1 fehlt und NUFIP2 lediglich in einem Experiment angereichert erscheint. Beide sind bekannte FMR1-Binder und werden mit den Fragilen-X-Syndrom in Zusammenhang gebracht.^{314,315}

Die dargestellten Interaktomstudien zeigen also auch für GGACU-Motive in unmittelbarer Nähe zu einem Stop Codon eine spezifische Gruppe interagierender und abgestoßener Proteine.

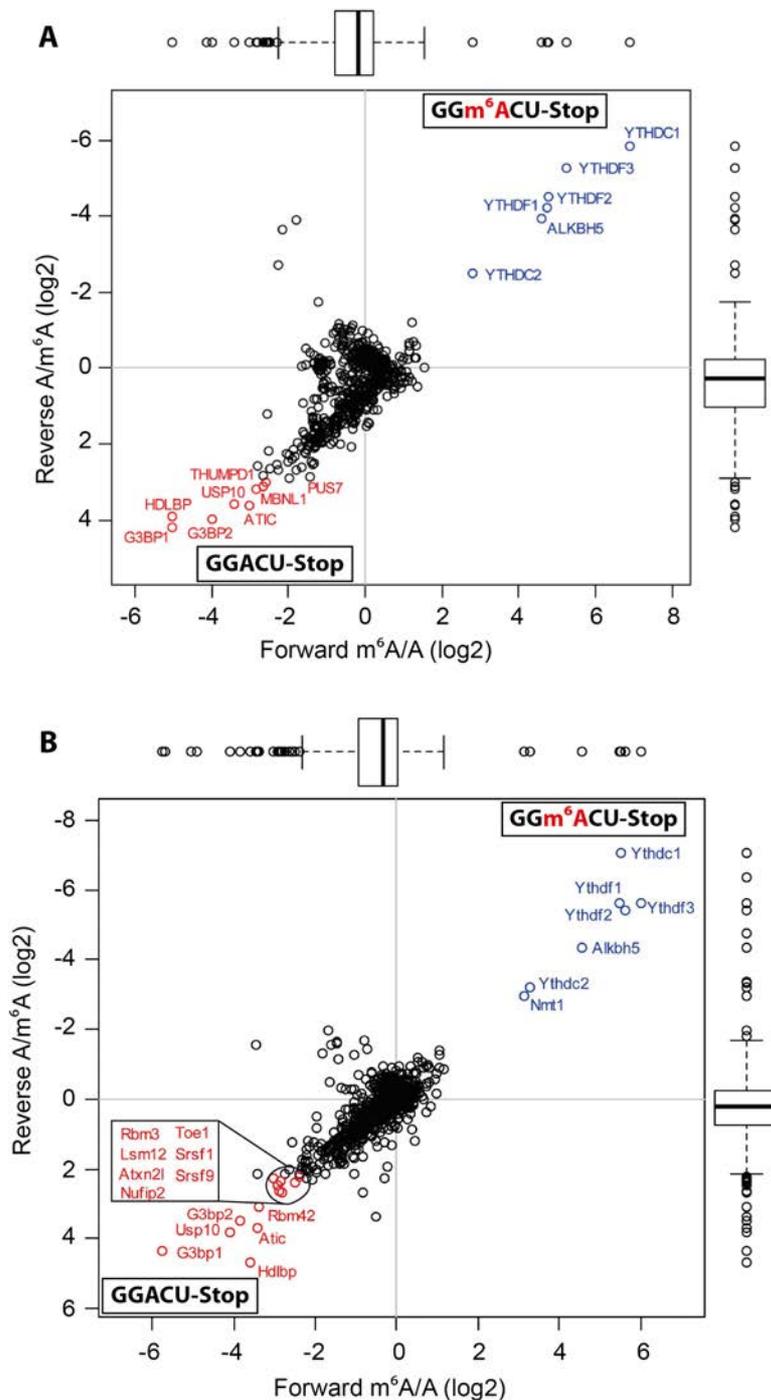


Abbildung 30: Scatterplots SILAC-basierter *Pulldowns* für ORN-12: 5'-Biotin-GGm⁶ACU UAA GGm⁶ACU UGA GGm⁶ACU UAG GGm⁶ACU-3'. A) *Pulldown* aus zytosolischem HeLa Lysat. B) *Pulldown* aus zytosolischem mESC Lysat. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Die bislang beschriebenen Experimente bieten bereits einen bislang nicht verfügbaren Überblick über das Interaktom von m⁶A im Kontext unterschiedlicher Sequenzmotive, decken aber nur einen Teil der im Konsensmotiv RRACH enthaltenen zwölf möglichen Sequenzen ab. Um weitere Binder untergeordneter Sequenzmotive (wie z. B. AGm⁶ACA) zu finden, bzw. ein „Konsens-Interaktom“ zu identifizieren sollten mit ORN-15 und ORN-16 zwei teilrandomisierte *Baits* eingesetzt werden. ORN-15 repräsentiert mit seiner Sequenz

5'-Biotin-(G/A G/A m⁶A C U/C/A)₄ A-3' die Gesamtheit aller vorhandenen Motive, einschließlich des dominanten GGACU. ORN-16 enthält eine Untergruppe möglicher Sequenzen unter Ausschluss der bereits individuell untersuchten GGACU und GAACU (**Abbildung 31**).

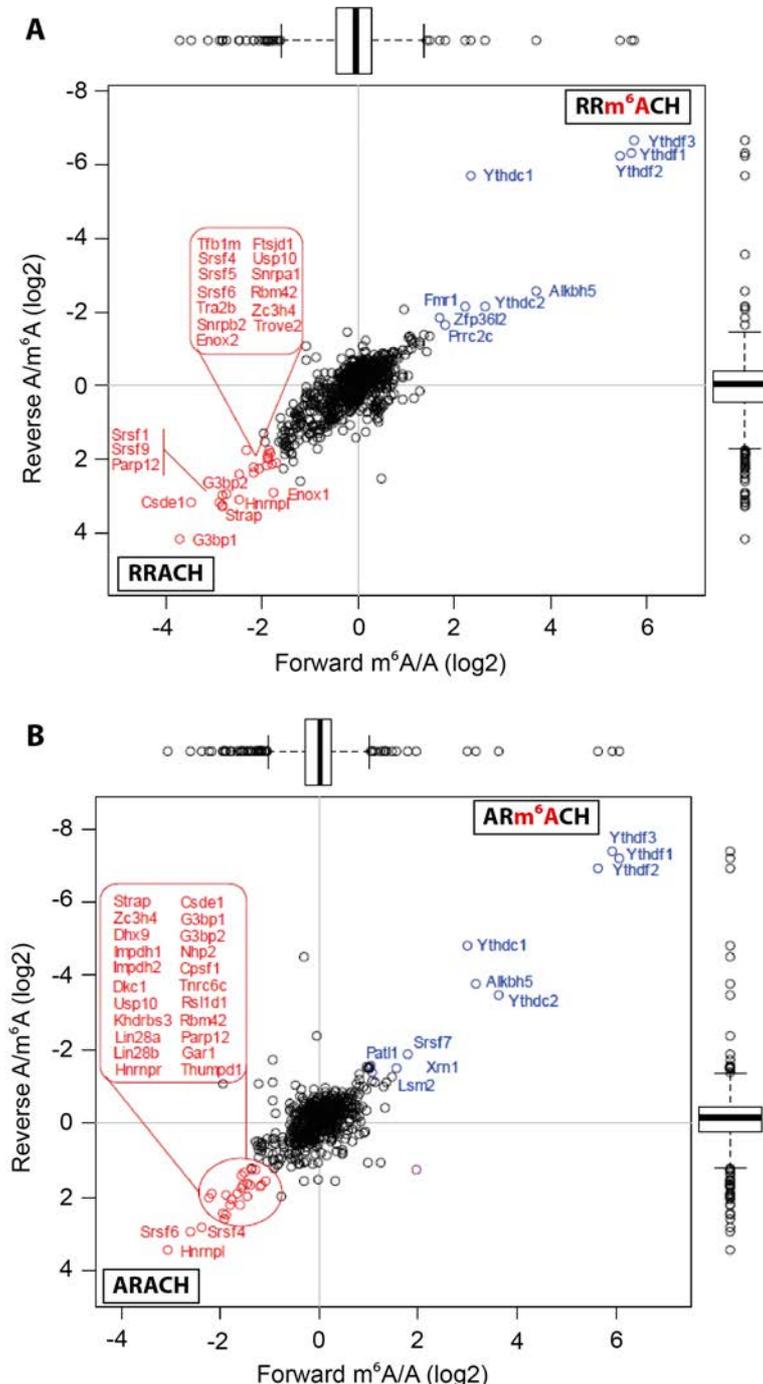


Abbildung 31: Scatterplots SILAC-basierter *Pulldowns* aus mESC Gesamtlisat für teilrandomisierte ORN-15 und -16. A) Als *Bait* wurde ORN-15 verwendet: 5'-Biotin-(RRACH)₄A-3' B) Als *Bait* wurde ORN-16 verwendet: 5'-Biotin-(ARACH)₄A-3'. R = A und G; H = A, C und U. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Die Resultate bestätigen zunächst die ausgeprägte Modulation der Affinität von YTHDC1, im Vergleich zu den anderen Mitgliedern der Familie. Auch ALKBH5 zeigt eine geringere

Bindung an die randomisierten Substrate, wie bereits für GAm^6ACU . FMR1 ist auf GGm^6ACU als Kontext angewiesen, da es ausschließlich für ORN-10 und ORN-11 (beide GGACU) als *Reader* detektierbar ist, im Falle von ORN-15 (RRACH) schwach vorhanden ist und für ORN-16 (ARACH) fehlt.

Unter den Proteinen, welche die unmodifizierte Sequenz bevorzugen scheinen G3BP1, G3BP2 und USP10 nicht zwingend auf GGACU angewiesen zu sein, ihre Abreicherung ist für ARACH aber schwächer ausgeprägt. Während CAPRIN1 und NUFIP2 für beide randomisierten ORN fehlen, zeigen sich mit STRAP, CSDE1, HNRNPL, RBM42 und den Splicingfaktoren SRSF4 und SRSF6 spezifisch angereicherte Proteine (STRAP = *Serine-threonine kinase receptor-associated protein*; CSDE1 = *Cold shock domain-containing protein E1*; HNRNPL = *Heterogeneous nuclear ribonucleoprotein L*; RBM42 = *RNA-binding protein 42*), die alle bereits in vorherigen Experimenten identifiziert wurden (siehe auch **Abbildung 35**).

Wie unter 4. bereits beschrieben wurden auch ORN eingesetzt, die einen kovalenten *Crosslinker* enthielten, um transiente Binder besser sichtbar zu machen und direkte von indirekten Bindern unterscheiden zu können. Die den *Crosslinker* enthaltenden Stränge (ORN-14/14') lieferten keine reproduzierbar angereicherten Proteine, was vermutlich auf eine zu ausgeprägte Reaktivität des kovalenten *Crosslinkers* zurückzuführen ist.

5.2.2 Validierungsexperimente zu m^6A

Die SILAC Studien gaben Hinweise auf bislang unbekanntes physische Interaktionen zwischen N^6 -Methyladenosin als mRNA-Modifikation und den mit dem Fragilen-X-Syndrom in Zusammenhang stehenden Proteinen FMR1, FXR1 und FXR2. Darüber hinaus werden mit G3BP1, G3BP2, NUFIP2, CAPRIN1 und USP10 mehrere Proteine durch m^6A abgestoßen, die in die Bildung bzw. Funktion von Stressgranula involviert sind.

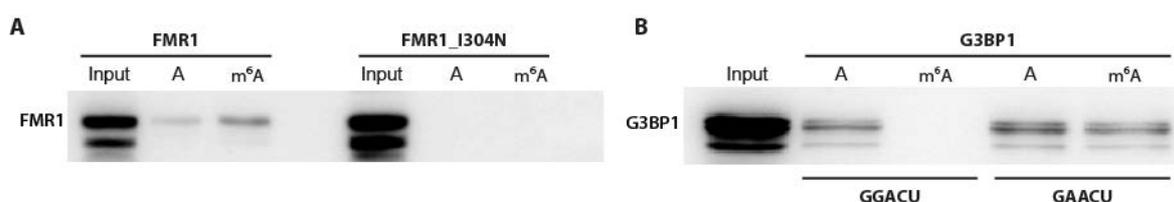


Abbildung 32: Validierungsexperimente für Interaktionen zwischen m^6A und FMR1, bzw. G3BP1. Extrakte wurden aus *E. coli* Kulturen gewonnen, die das Zielprotein rekombinant exprimierten. *Western Blots* erfolgten gegen den verwendeten GST-Tag A) *Pulldowns* mit ORN-11/11' als *Baits*. FMR1_I304N ist an Position 304 von Ile zu Asn mutiert. B) *Pulldowns* mit ORN-11/11' und ORN-13/13' als *Baits*. Die Experimente wurden von Raghu Edupuganti durchgeführt.

Eine Validierung mittels *Western Blot* (**Abbildung 32**) konnte bestätigen, dass FMR1 bevorzugt an GGm⁶ACU bindet. Darüber hinaus konnte gezeigt werden, dass FMR1_I304N diese Selektivität verliert und nur noch schwach an RNA bindet. Diese Mutation befindet sich innerhalb einer KH-Domäne des Proteins und stellt eine seltene Ursache des Fragilen-X-Syndroms dar, die mit stark ausgeprägten klinischen Symptomen verbunden ist.^{320,321} Mit G266E konnten *Myrick et al.* in einer aktuellen Studie eine weitere Punktmutation identifizieren, die sich innerhalb einer KH-Domäne befindet und zu einer Erkrankung des Patienten führte.³²² Hierdurch wird ein direkter Zusammenhang hergestellt zwischen der Fähigkeit des Proteins selektiv RNA zu binden und einer phänotypischen Erkrankung am Fragilen-X-Syndrom. FMR1 verfügt über drei KH- und eine RGG-Domäne³²³, es ist anzunehmen, dass diese Domänen gemeinsam für die RNA-Bindung verantwortlich sind.

Unter der Annahme, dass eine Regulation der Lokalisation und Translation definierter Transkripte mittels der FMR-Familie durch m⁶A moduliert wird, ist eine zumindest partielle Übereinstimmung zwischen den Bindemotiven der Proteine und bekannten m⁶A Loci zu erwarten. *Ascano et al.* zeigten in einer PAR-CLIP Studie, dass FMR1 mit ACUK und WGGA (K = G oder U; W = A oder U) zwei Konsensmotive bindet.³²⁴ Zusammengenommen bilden sie mit WGGACUK ein Motiv, welches das dominante Konsensmotiv GGACU für die Modifikation m⁶A enthält. In der Tat finden sie beide Motivfragmente häufig in Kombination in einem Locus. Zudem bestätigt eine weitere Analyse durch *Ray et al.* das Sequenzmotiv GAC bzw. GGAC als FMR1-Bindestelle. Auch FXR1 und FXR2 präferieren GAC.³²⁵ Diese Studie sagt darüber hinaus eine stark destabilisierende Wirkung auf die Transkripte durch FMR1 und FXR1 voraus. Ein Abgleich der Daten aus m⁶A-Sequenzierung und FMR1-PAR-CLIP zeigt einen Überlapp von etwa 35% für HEK293 Kulturzellen und eine enge räumliche Nähe der Loci (**Abbildung 33**). FMR1 bindet also *in vivo* m⁶A-modifizierte Loci in mRNA.

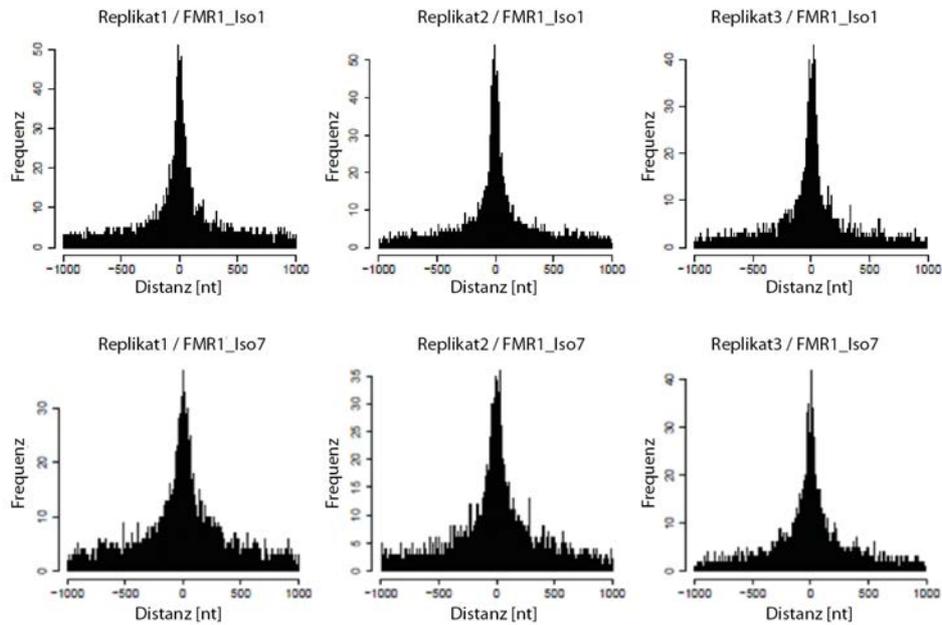


Abbildung 33: Vergleich der identifizierten RNA Loci aus m^6A Sequenzierung und FMR1-PAR-CLIP für HEK293 Zellen. Dargestellt sind Überlagerungen dreier Replikate aus m^6A -Sequenzierung mit Daten für FMR1-Isoformen 1 und 7. Die Analyse zeigt, dass FMR1-Bindestellen und m^6A -modifizierte Motive ein hohes Maß an Übereinstimmung in ihrer Lokalisation zeigen. Sie wurde durchgeführt von Raghu Edupuganti.

Weiterhin stellen *Ascano et al.* fest, dass Bindestellen der Homologen FXR1 und FXR2 zu 95% mit FMR1-Loci übereinstimmen, die Familie also nahezu identische Affinitäten besitzt, was eine Anreicherung aller drei Proteine im SILAC Experiment erklärt.³²⁴ Zusammengenommen liefern die Experimente überzeugende Hinweise auf eine Verbindung zwischen der posttranslationalen Modifikation von Transkripten durch m^6A und der Bindung durch FMR1.

Die Validierung der SILAC-Interaktome für G3BP1 zeigt dass es GGACU-Motive ausschließlich bindet, wenn sie nicht durch m^6A modifiziert sind. Dies befindet sich in Übereinstimmung mit den vorangegangenen massenspektrometrischen Analysen. Eine vergleichbare Selektivität ist für GAACU nicht zu beobachten (**Abbildung 32**). G3BP1 ist, gemeinsam mit G3BP2, CAPRIN1, USP10 und FMR1 in die Assemblierung und Funktion von Stressgranula involviert.^{311-313,316,326,327} Hierdurch erscheint die Annahme naheliegend, dass auch hier m^6A im GGACU-Kontext einer Modulierung der Interaktion zwischen mRNA und den involvierten Proteinen dient. *Ray et al.* finden für G3BP2 das Bindemotiv GGAU³²⁵, welches eine partielle Übereinstimmung mit dem m^6A -Konsensmotiv aufweist.

Stressgranula und die in ihnen enthaltenen mRNPs dienen unter anderem der Sequestrierung von mRNA und der Regulation ihrer Stabilität und Translation.^{328,329} Um eine mögliche Korrelation zwischen der Stabilität der mRNA und einer Bindung durch G3BP1/2 zu

verifizieren und ihr RNA-Bindemotiv einzuengen wurde durch die Gruppen von *Michiel Vermeulen* und *Chuan He* ein PAR-CLIP-Experiment mit beiden Proteinen durchgeführt. Die Analyse zeigte, dass beide Proteine miteinander verwandte Motive binden, die aber keine deutliche Übereinstimmung mit den SILAC-Experimenten zeigen. Sie zeigen eine Präferenz für CAACUC. In dem Fragment ist zwar ein Teil des m⁶A-Kontextes wiederzuerkennen (AACU), die für m⁶A vorhandene Bedeutung der Guanosine (GGACU / GAACU) ist aber nicht reproduziert. *Ray et al.* finden im Gegensatz hierzu GGAU als dominantes Konsensmotiv für humanes G3BP2.³²⁵ Ebenso wie die FMR-Familie binden die G3BPs bevorzugt im 3'-UTR der Transkripte. Die G3BPs scheinen zwar nur zum Teil mRNAs zu binden, die auch m⁶A-modifiziert sein können. Innerhalb dieser Untergruppe liegt die Übereinstimmung der identifizierten Loci aber bei etwa 80%. Eine Korrelation der identifizierten G3BP PAR-CLIP *Peaks* mit bekannten Daten zur Halbwertszeit der betroffenen Transkripte zeigt einen deutlichen stabilisierenden Einfluss (**Abbildung 34**). Die Modifikation m⁶A übt einen gegensätzlichen Einfluss aus. Die m⁶A-Binder und die G3BP Proteine befinden sich also in einem funktionellen Antagonismus und sie scheinen um verwandte Sequenzmotive zu konkurrieren, in Abhängigkeit von deren Modifikationsgrad.

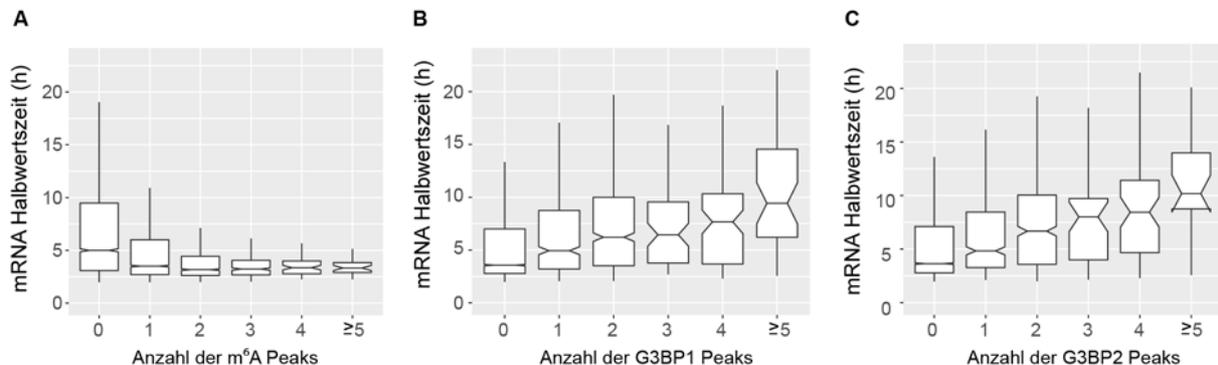


Abbildung 34: Während m⁶A die Halbwertszeit eines modifizierten Transkripts negativ beeinflusst ist für die Bindung durch G3BP1 und G3BP2 ein gegenteiliger Effekt zu beobachten. A) Korrelation zwischen bekannten m⁶A Loci mit der Halbwertszeit der modifizierten Transkripte. B) und C) Korrelation zwischen PAR-CLIP *Peaks* (identifizierten Bindestellen) der Proteine G3BP1 und G3BP2 mit der Halbwertszeit gebundener Transkripte. Die Analyse wurde durchgeführt von *Raghu Edupuganti*.

5.2.3 Zusammenfassung

Mit Hilfe der in dieser Arbeit synthetisierten m⁶A-modifizierten ORN war es möglich, in Kooperation mit der Gruppe von *Michiel Vermeulen*, das m⁶A-Interaktom in einer bislang nicht vorhandenen Tiefe aufzuklären. Die Stärke der verwendeten SILAC-Experimente zeigt sich am deutlichsten in der Tatsache, dass sich ein signifikanter Teil der neu gewonnenen

Erkenntnisse auf Proteine bezieht, die von m⁶A abgestoßen werden, d.h. bevorzugt Sequenzmotive in ihrer unmodifizierten Form binden.

In einer Zusammenschau zeigen die Analysen, dass das dominante Motiv GGACU im Zytosol von allen Mitgliedern der YTH-Familie gebunden wird, wobei YTHDC1 der stärkste Binder zu sein scheint. Im Nukleus ist das Bild nicht vollständig klar, da in murinen Extrakten (mESC, NPC) YTHDF3 die stärkste Affinität aufweist, was in HeLa hingegen nicht der Fall ist. Dies kann als Hinweis auf eine im Kern lokalisierte Funktion des Binders YTHDF3 gedeutet werden, dessen präzise Rolle bislang unklar ist. Andere Motive (GAACU, RRACH, ARACH) zeigen eine geringer ausgeprägte Affinität für YTHDC1 und weisen auf die YTHDF-Proteine als Binder mit breiterer Sequenzspezifität hin. Für alle Motive scheint YTHDC2 ein *Reader* mit niedriger relativer Bindungsstärke zu sein. Diese Beobachtungen sind in Übereinstimmung mit bereits publizierten Daten und sprechen damit für die Belastbarkeit der Methode.¹⁴⁷

Mit FMR1, FXR1 und FXR2 konnte eine bislang unbekannte Gruppe spezifischer Interaktoren identifiziert werden, die m⁶A in einen direkten Zusammenhang mit dem genetisch bedingten Fragilen-X-Syndrom bringen. Die Krankheit wird durch ein dysfunktionales FMR1 ausgelöst, also die Fehlfunktion eines RNA-bindenden Proteins. FMR1 ist verantwortlich für die gewebsspezifische Lokalisation und gesteuerte Translation der von ihm gebundenen mRNAs.^{305,306,330,331} Der Verlust dieser Kontrolle resultiert in einer Manifestierung der Erkrankung. Das Protein bindet bevorzugt Sequenzabschnitte, die eine Ähnlichkeit zum m⁶A-Konsensmotiv aufweisen^{324,325} Für den Kontext GGACU konnte die Selektivität validiert werden, für andere Derivate der RRACH-Gruppe ist sie nicht feststellbar.

FMR1 ist als Komponente von mRNPs in Stressgranula bekannt, wo es gemeinsam mit G3BP1, G2BP2, CAPRIN1 und USP10 vorkommt, die zu den von m⁶A abgestoßenen Proteinen gehören.^{311-313,316,326,327} Die Annahme liegt nahe, dass ein durch m⁶A dynamisch regulierter Antagonismus zwischen den vier letztgenannten Proteinen und den m⁶A-Reader-Proteinen für eine Steuerung der Aktivität der betroffenen mRNA verantwortlich ist. Während G3BP1 und G2BP2 einen stabilisierenden Effekt auf mRNA haben wird für FMR1 ein destabilisierender Einfluss vorhergesagt.³²⁵ Insgesamt geben bisherige Untersuchungen aber kein klares Bild des Einflusses von FMR1 auf die mRNA-Halbwertszeit.³³² Beide Proteine wirken repressiv auf die Translation der gebundenen RNAs.^{308,313}

Eine Erklärung für die beschriebenen Beobachtungen wäre eine durch m⁶A beeinflusste Reorganisation von mRNPs die als Komponenten FMR1, FXR1/2, G3BP1/2, CAPRIN1 und USP10 enthalten. Zahlreiche Komponenten von Stressgranula gehören zu den reproduzierbar von m⁶A abgestoßenen Proteinen. Hieraus lässt sich die Vermutung ableiten, dass die Modifikation und die damit einhergehende Bindung der mRNA durch die Familie der YTH-Proteine zu einer Herauslösung des Transkripts aus Stressgranula führt. Durch die bindenden Proteine könnte sie anschließend einer Relokalisierung, einer verstärkten Translation und einem schnellen Abbau zugeführt werden, was sich im Einklang mit publizierten Erkenntnissen befände.

Neben dem binären Schalter A/m⁶A ist eine Veränderung dieser Prozesse auch durch den Sequenzkontext der Modifikation möglich. So sind für GAM⁶ACU nach wie vor alle dominanten Interaktoren detektierbar, eine Selektivität von FMR1 oder G3BP1 ist aber nicht mehr vorhanden. Das randomisierte Motiv RRACH verfügt wiederum über eine eigenständige Gruppe von abgestoßenen Proteinen. Auch SRSF1 und 9, die im Verlauf dieser Dissertation von *Xiao et al.* als YTHDC1-Interaktoren beschrieben wurden, konnten in dieser Analyse identifiziert werden.

Ein Überblick über die individuellen Interaktome und ihren Überlappungsgrad bietet **Abbildung 35**. Sie zeigt, dass es eine hoch konservierte Gruppe von Interaktoren gibt (ALKBH5, YTHDC1/2, YTHDF1/2/3), der Grad der Konservierung unter den abgestoßenen Proteinen hingegen deutlich geringer ist, wie bereits geschildert. Als einziges Protein ist RBM42 (*RNA-binding protein 42*) für alle untersuchten Zelltypen und Sequenzmotive signifikant abgereichert (siehe auch **Abbildung 28**). Über seine Funktion ist sehr wenig bekannt, es gibt aber Hinweise, dass es als Komponente von Stressgranula in die Stabilisierung von mRNAs involviert ist.³³³ Darüber hinaus wurde für ein orthologes Protein aus dem Protozoon *Toxoplasma gondii* gezeigt, dass es ein essentieller Regulator des Splicings ist.³³⁴ Die nähere Untersuchung seiner Funktion im Zusammenhang mit m⁶A erscheint in hohem Maße lohnenswert, da es basierend auf den hier dargestellten Untersuchungen das alleinige ubiquitär abgestoßene Protein ist.

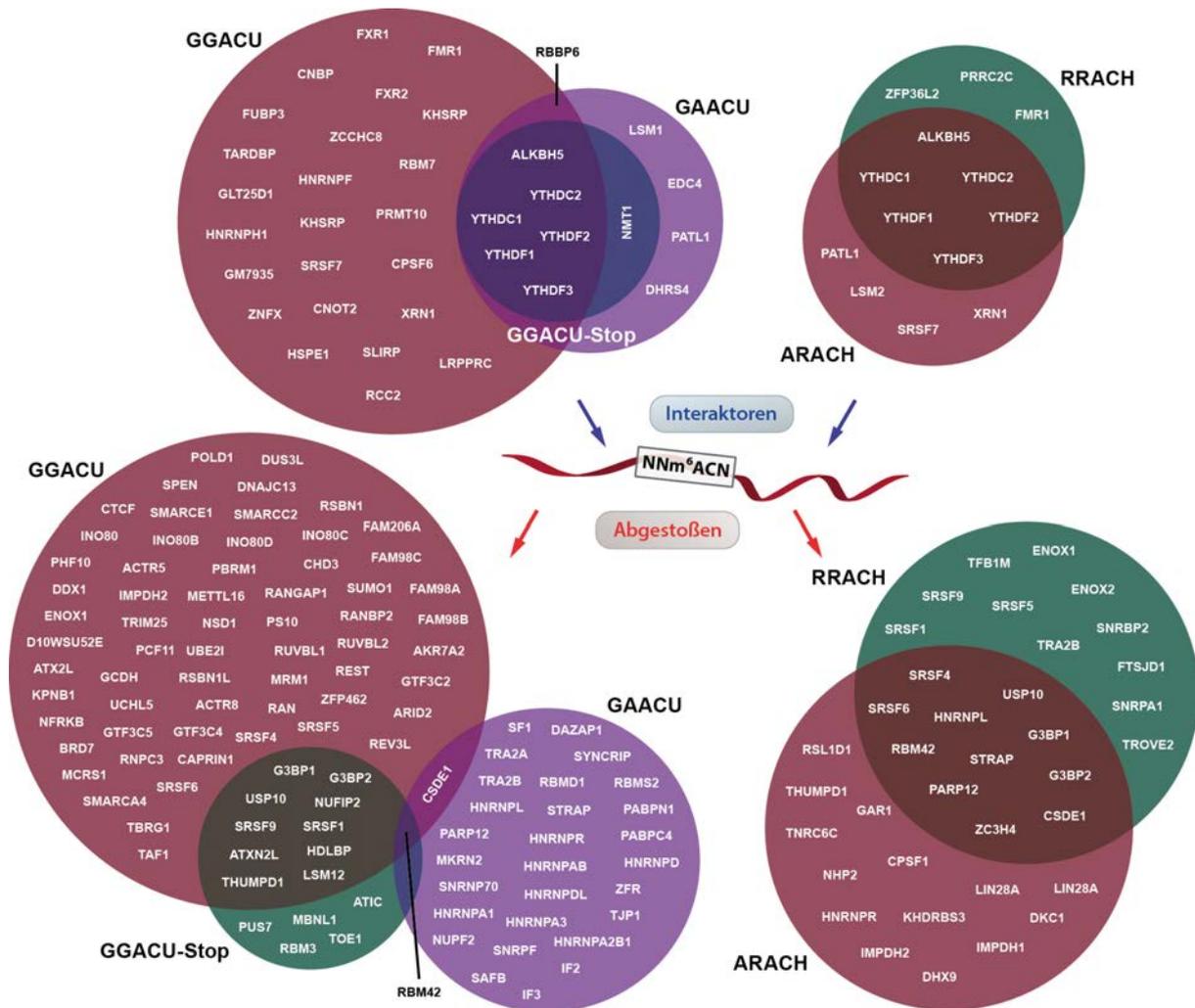


Abbildung 35: Visualisierung des Überlappungsgrades zwischen den Interaktomen für das alle untersuchten m⁶A-enthaltenden Sequenzmotive. Für die Analyse wurden identifizierte Proteine aus zytosolischen Lysaten und Kernextrakten des jeweiligen Sequenzmotives zusammengefasst. Für GGm⁶ACU wurden die Interaktome aller drei untersuchten Zelltypen (HeLa, mESC und NPC) zusammengefasst, daher ist die Gesamtzahl der Proteine für dieses Motiv erheblich größer als für die anderen Sequenzkontexte. RRACH und ARACH stellen teildegenerierte Motive dar: R = G/A und H = A/C/U.

Die Übereinstimmung mit bereits bekannten Daten ist insgesamt hoch und die generierten Interaktomstudien bilden eine vielversprechende Grundlage für weitergehende Analysen.

Abbildung 36 zeigt eine Zusammenfassung, reduziert auf die bedeutendsten Proteine, grob sortiert nach ihrer Affinität.

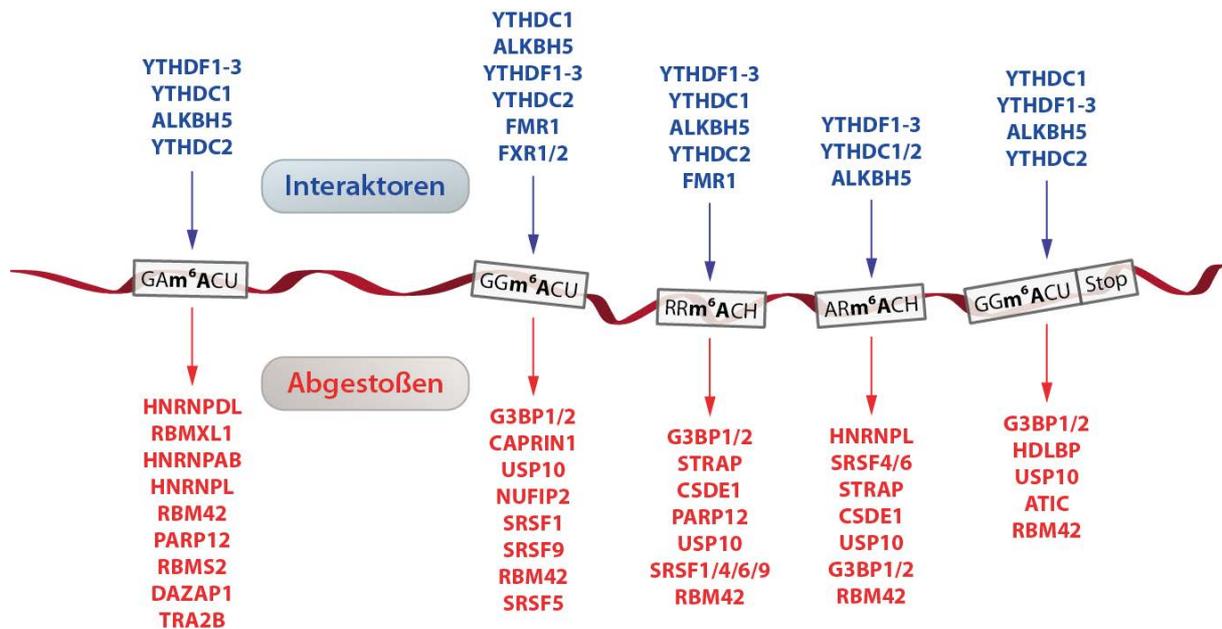


Abbildung 36: Auswahl einiger wichtiger, in SILAC-Studien identifizierter Interaktoren (selektiven Binder) und abgestoßener Proteine, d.h. der Proteine, die kanonisches Adenosin mit höherer Affinität binden als m⁶A. Die Proteine sind absteigend nach ihrer Affinität geordnet, soweit möglich.

Die Zusammenfassung macht deutlich, dass die Veränderungen des Interaktoms überwiegend innerhalb der Gruppe der abgestoßenen Proteine stattfinden, also über die selektive Verdrängung bindender Proteine durch die Einführung der Methylgruppe. Die *Reader*-Gruppen sind deutlich statischer, bzw. weniger abhängig vom Sequenzkontext. Auch hier gibt es aber kontextspezifische Veränderungen im Interaktom. So konnte für FMR1 gezeigt werden, dass es GGm⁶ACU spezifisch bindet, GAm⁶ACU dagegen nicht. Darüber hinaus beeinflusst der Sequenzkontext die relativen Affinitäten der bindenden Proteine, was den jeweils verursachten Effekt maßgeblich beeinflussen kann. Damit eröffnen die hier vorgestellten Analysen einen tiefen Einblick in das komplexe Interaktom dieser Modifikation, der über alternative Methoden nicht ohne Weiteres zu erreichen wäre.

5.3 Analyse des Interaktoms von N^6,N^6 -Dimethyladenosin

Durch *Veronika Reiter* durchgeführte Quantifizierungsstudien deuteten darauf hin, dass humane polyadenylierte RNA signifikante Mengen N^6,N^6 -Dimethyladenosin enthält.¹⁰⁹ Ausgehend von diesen Daten sollten in dieser Arbeit *Reader* der Modifikation identifiziert werden um eine möglicherweise vorhandene Funktion in mRNA aufzuklären. Im Verlauf der Experimente wurde der m^6_2A -Modifikationsgehalt von Gesamt- sowie poly(A)-RNA erneut von *Sarah Schiffers* analysiert. Dabei wurde die RNA einem rigorosen Aufreinigungsprotokoll unterzogen und die anfänglich detektierten Mengen konnten nicht bestätigt werden. Es ist davon auszugehen, dass m^6_2A keine in mRNA vorkommende RNA-Modifikation darstellt, was sich im Einklang mit der gängigen Annahme befindet, dass sein Vorkommen auf rRNA beschränkt ist.⁹³ Die Ergebnisse werden hier dennoch dargestellt, da sie Hinweise auf die Spezifität einzelner Proteine geben, die bereits für m^6A identifiziert wurden. Sie offenbaren, dass einige m^6A -Binder auch die sterisch anspruchsvollere Modifikation m^6_2A binden können und gegenüber A bevorzugen.

Zwar ist für humane rRNA ein zweifach m^6_2A -modifiziertes Motiv bekannt (GUG m^6_2A m m^6_2A ACCU). Da aber keine Sequenzinformationen zu m^6_2A in mRNA vorliegen wurde auf das bekannte Konsensmotiv für m^6A zurückgegriffen. GGACU wurde einzeln und in Kombination mit Stop Codons untersucht. Die Sequenzen (**Tabelle 3**) sind, abgesehen vom Austausch von m^6A gegen m^6_2A , identisch mit den für m^6A verwendeten ORN-11 und -12.

Tabelle 3: Für die Analyse des Interaktoms von m^6_2A hergestellte Oligoribonukleotide. B = 5'-Biotineinheit (*Link Tech*).

ORN	Sequenz
ORN-17	5'-B GG m^6_2A CU GG m^6_2A CU GG m^6_2A CU GG m^6_2A CU-3'
ORN-18	5'-B GG m^6_2A CU <u>UAA</u> GG m^6_2A CU <u>UGA</u> GG m^6_2A CU <u>UAG</u> GG m^6_2A CU-3'

Die SILAC-Experimente zeigen ein Interaktom für m^6_2A , das eine partielle Übereinstimmung mit dem für m^6A beobachteten besitzt (**Abbildung 37**). Zwar ist die gesamte Familie der YTH-Proteine nicht detektierbar, die Demethylase ALKBH5 dagegen ist unter den *Readern*. Auffällig ist zudem, dass FMR1, FXR1 und FXR2 angereichert wurden. Im Gegensatz zu den YTH-Proteinen scheinen diese *Reader* also über eine weniger strikte Spezifität zu verfügen, so dass sie trotz des höheren sterischen Anspruches, der durch die doppelte Methylierung vorhanden ist, binden können. Hieraus ließe sich ableiten, dass die identifizierten Proteine keine spezifischen *Reader* darstellen, sondern allgemein N^6 -alkylierte Adenosine binden.

ALKBH5 ist aber ein etablierter m⁶A-Interaktor, der den Metabolismus der Modifikation maßgeblich prägt. Eine Anreicherung *in vitro* durch sowohl m⁶A als auch m⁶₂A widerspricht einer auf m⁶A beschränkten biologischen Funktion also offensichtlich nicht.^{136,335}

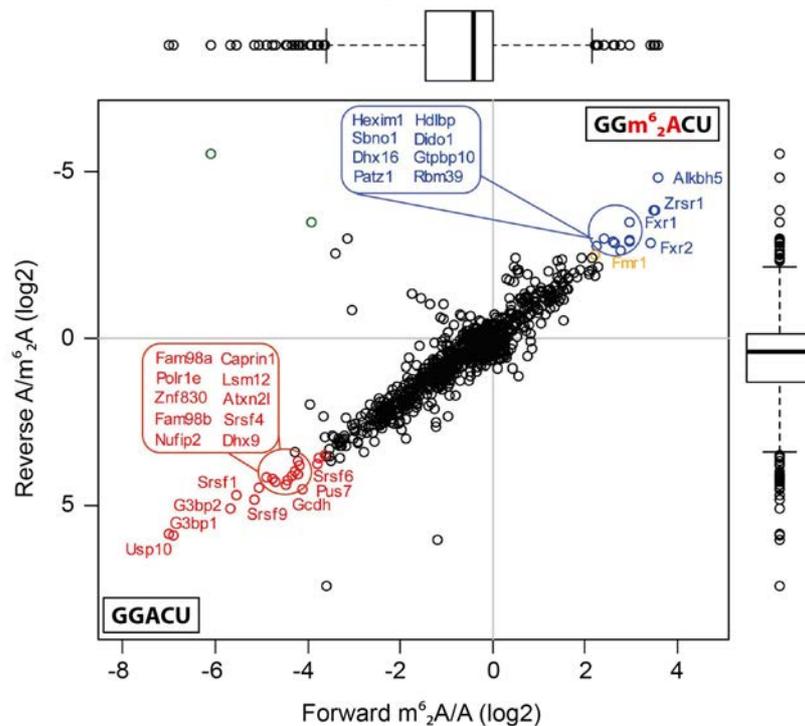


Abbildung 37: Scatterplot eines SILAC-basierten *Pulldowns* aus zytoplasmatischem mESC-Lysat für ORN-17: 5'-Biotin-(GGm⁶₂ACU)₄-3'. Das Experiment wurde von *Raghu Edupuganti* durchgeführt.

Unter den abgestoßenen Proteinen findet sich mit USP10, G3BP1/2, SRSF1/9, CAPRIN1 und NUFIP2 ein noch höheres Maß an Übereinstimmung zum m⁶A-Interaktom. Hierdurch wird bestätigt, dass es sich bei dieser Gruppe um Proteine handelt, die eine ausgeprägte Affinität zu dem Sequenzmotiv GGACU besitzen, welche durch eine Alkylierung an der N⁶-Position des zentralen Adenosins verloren geht.

Mit ORN-18 wurde m⁶₂A auch im Stop Codon-Kontext analysiert (**Abbildung 38**). In diesem Fall zeigen die Ergebnisse aber keinen ausgeprägten Überlapp mit dem analogen m⁶A-modifizierten Strang (**Abbildung 30**), lediglich ALKBH5 ist ebenfalls angereichert.

Zusammengenommen zeigen die Resultate, dass ALKBH5 offensichtlich ein breiteres Spektrum an Substraten binden kann, als andere m⁶A-Interaktoren. Es ist Teil einer Familie von DNA- bzw. RNA Reparaturenzymen, die zum Teil breite Substratspezifitäten besitzen³³⁶, was eine Erklärung für diese Beobachtung sein könnte. Da m⁶₂A nach dem derzeitigen Kenntnisstand nicht in mRNA zu finden ist sind die Ergebnisse aber von geringer

biologischer Relevanz. Eine sich aus den hier vorgestellten Resultaten ergebende Frage wäre, ob m^6_2A in seinem bislang bekannten rRNA Kontext ein mögliches Substrat für ALKBH5 darstellt.

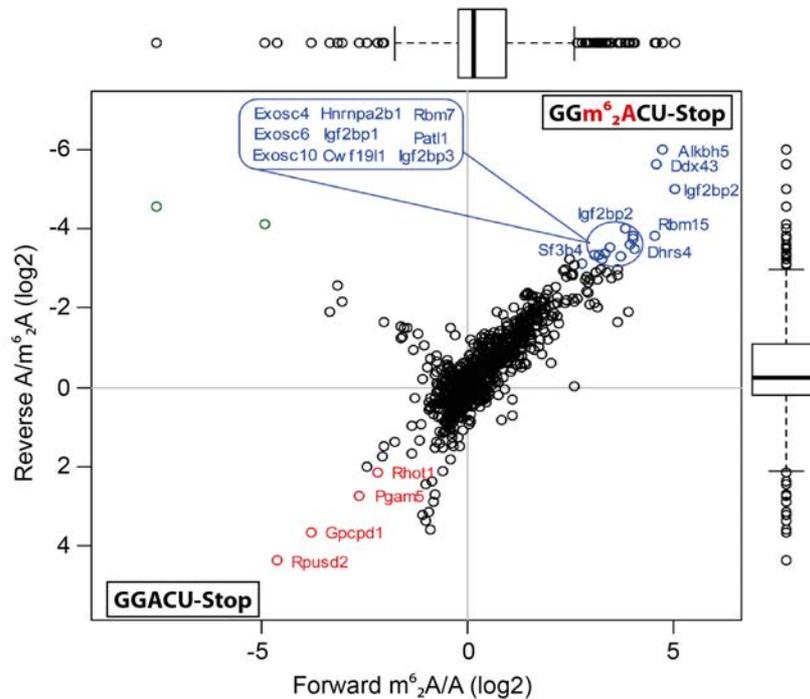


Abbildung 38: Scatterplot eines SILAC-basierten *Pulldowns* aus zytoplasmatischem mESC-Lysat für ORN-18: 5'-Biotin GGm⁶₂ACU UAA GGm⁶₂ACU UGA GGm⁶₂ACU UAG GGm⁶₂ACU-3'. Das Experiment wurde von *Raghu Edupuganti* durchgeführt.

5.4 Analyse des Interaktoms von Pseudouridin

Pseudouridin ist als mRNA-Modifikation etabliert und es sind erste Sequenzinformationen publiziert.^{75,78,93} Sie zeichnen sich durch ein hohes Maß an Heterogenität aus, was auf die Uneinheitlichkeit innerhalb der Gruppe der modifizierenden Enzyme (Pseudouridinsynthasen) zurückzuführen sein könnte.^{98,337} Es ist kein klares, globales Konsensmotiv identifizierbar. Zwei der untersuchten Pseudouridinsynthasen haben aber definierte Konsensmotive: UGΨAR für Pus7 und GUΨC für Pus4. Diese Motive wurden in ORN-27 und ORN-28 (**Tabelle 4**) in einen Sequenzkontext eingebettet, der keine Sekundärstruktur ausbilden sollte. Zur Simulation von Sekundärstrukturen wurde der *RNAfold WebServer* der Universität Wien genutzt.³³⁸ Der Einfluss von posttranslationalen Modifikationen, wie Ψ, kann hier bislang nicht berücksichtigt werden.

Tabelle 4: Für die Analyse des Pseudouridin-Interaktoms hergestellte Oligoribonukleotide. B = 5'-Biotineinheit (*Link Tech*).

ORN	Sequenz
ORN-27	5'-Biotin CAAUGΨAG CAAU GΨAG CAAUGΨAG -3'
ORN-28	5'-Biotin AGAGUΨC AGAGUΨC AGAGΨUC -3'

Mit ORN-27 durchgeführte *Pulldowns* aus murinem Kernlysat (mESC) zeigen keine reproduzierbar angereicherten spezifischen Interaktoren, aber eine große Gruppe abgestoßener Proteine (**Abbildung 39**). Darunter befinden sich mit NUDT21 (auch CPSF5) und CPSF6/7 drei bekannte Komponenten des *Cleavage Factor I_m* Komplexes (NUDT21 = *Cleavage and polyadenylation specificity factor subunit 5*; CPSF6/7 = *Cleavage and polyadenylation specificity factor subunit 6/7*). Das Heteromultimer ist involviert in die Erkennung von Polyadenylierungssequenzen (*poly(A) sites*) in prä-mRNA und fördert die Assemblierung des 3'-prozessierenden Komplexes während der Transkription.³³⁹ *Brown* und *Gilmartin* konnten zeigen, dass der CFI_m-Komplex bevorzugt 5'-UGUAA-UGUAB-3'-Motive bindet (wobei B = C/U/G).³⁴⁰ Interessant erscheint, dass das bekannte Konsensmotiv der Pseudouridinsynthase PUS7 (UGΨAR) in diesem Motiv enthalten ist. *Schwartz et al.* konnten eine PUS7-abhängige Pseudouridylierung eines solchen Motives in mRNA in Hefe zeigen⁷⁵ und die für die Polyadenylierung ausschlaggebenden Sequenzmotive gelten als hoch konserviert.³³⁹

Weiterhin sind mit CPSF1/2/4 und CSTF1/2/2t/3 zahlreichere weitere Faktoren in beiden Experimenten abgereichert, die in die Polyadenylierung von prä-mRNA involviert sind (CSTF 1/2/2t/3 = *Cleavage stimulation factor 1/2/2 tau/3*). Viele humane mRNAs besitzen mehrere Polyadenylierungssignale³⁴¹ und es ist noch unklar, wie eine Selektion des korrekten Signals abläuft. Die hier dargestellten Daten weisen darauf hin, dass eine Modifikation der CFI_m-Bindemotive durch Ψ die Rekrutierung der polyadenylierenden Faktoren beeinflussen könnte.

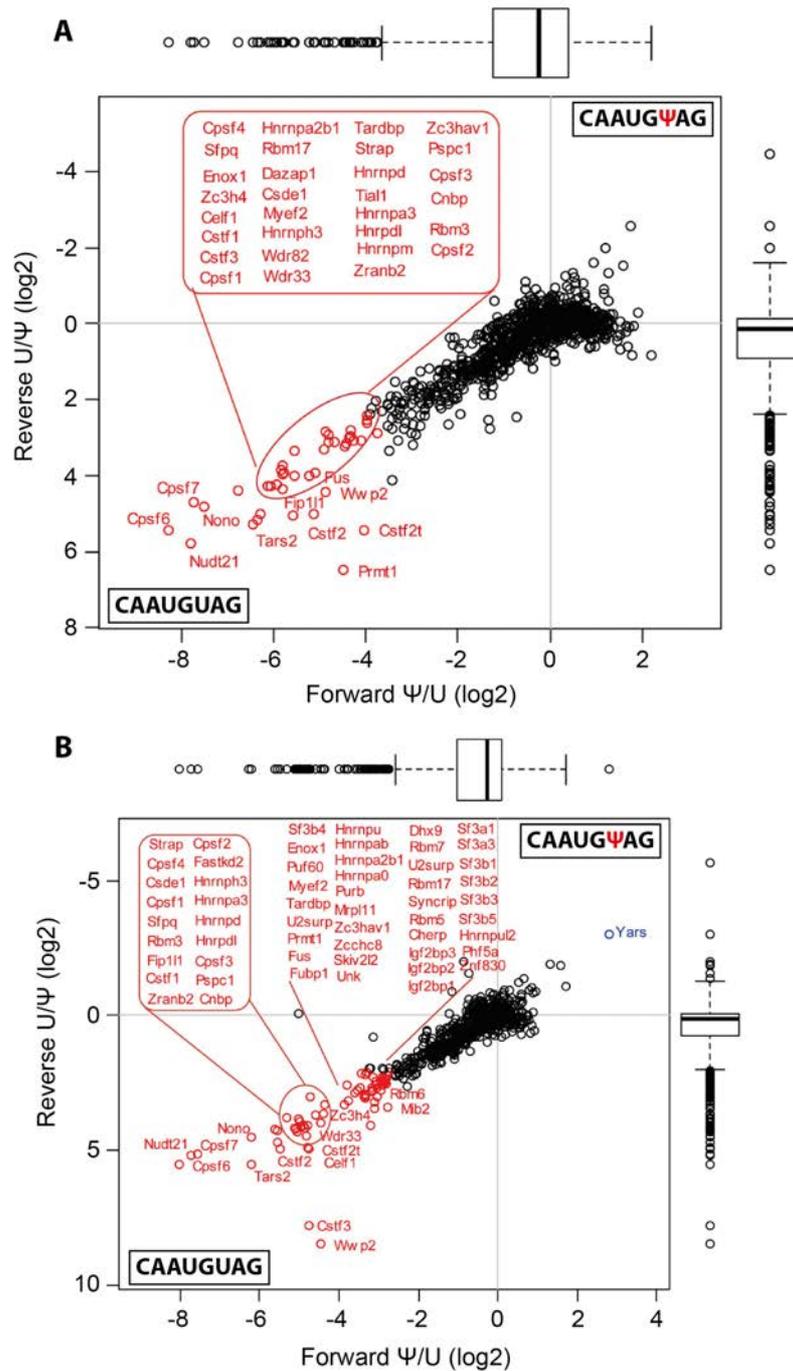


Abbildung 39: Scatterplots SILAC-basierter *Pull-downs* aus mESC-Kernextrakt (A und B) für ORN-27/27': 5'-Biotin-CAAUGΨAG CAAUGΨAG CAAUGΨAG-3'. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Zytosolischer Extrakt erbrachte keine verwertbaren Ergebnisse (**Abbildung 40**), ebenso wie alle mit ORN-28 durchgeführten SILAC-*Pull-downs*. Ein Grund hierfür können ausgeprägt unterschiedliche Sekundärstrukturen der beiden Stränge sein, welche aufgrund der stark strukturmodulierenden Wirkung von Ψ schwer vermeidbar und nicht vorherzusagen sind.^{63,100} Da Ψ die Ausbildung von RNA-Duplexen fördert und Paarungen mit allen kanonischen Basen eingehen kann, ist die Ausbildung von inter- und intramolekularen Multiplexen

vorstellbar, wodurch die erfolgreiche Durchführung der SILAC-Experimente erheblich erschwert wäre.

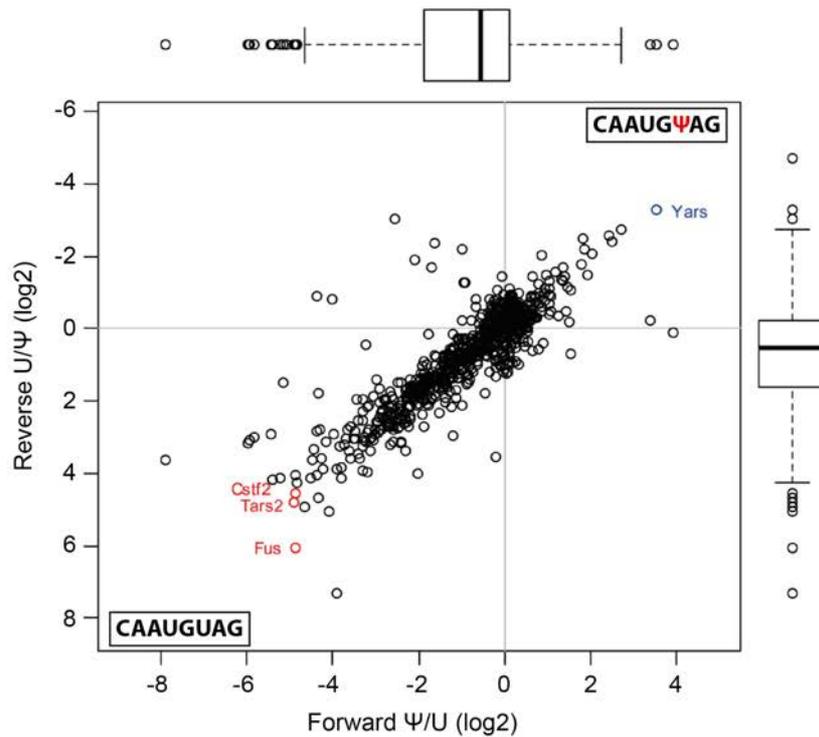


Abbildung 40: Scatterplot eines SILAC-basierten *Pulldowns* aus zytosolischem mESC-Lysat für ORN-27/27': 5'-Biotin-CAAUGΨAG CAAUGΨAG CAAUGΨAG-3'. Das Experiment wurde von *Raghu Edupuganti* durchgeführt.

Die aus Kernlysate generierten Daten zeigen einen Überlapp von 85%. Die Überwiegende Mehrheit der in Pulldown A identifizierten Proteine findet sich auch in Pulldown B wieder (**Abbildung 41**).

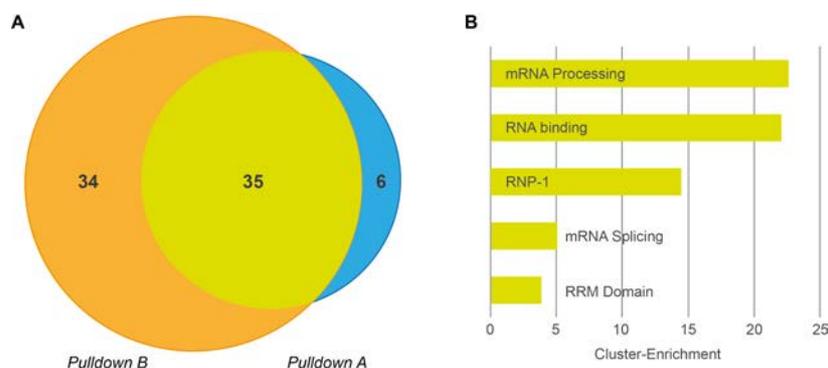


Abbildung 41: Analyse der abgestoßenen Proteingruppen für das CAAUGUAG-Motiv in mESC Kernlysate (Abbildung 26). A) Venn-Diagramm der abgestoßenen Proteine. B) DAVID *Gene Ontology* Analyse der 35 Proteine, die in beiden *Pulldowns* zu den abgestoßenen Proteinen gehören (gelb in A). RNP-1 ist eine 8 AS lange Konsenssequenz, die Teil der eukaryotischen RRM-Domäne (*RNA recognition motif*) ist.

Eine Suche nach angereicherten biologischen Funktionen oder Proteindomänen mittels DAVID (Version 6.7, *Medium Stringency*)^{342,343} zeigt, dass RNA-bindende, in mRNA-

Prozessierung und Splicing involvierte Proteine, sowie solche, die über eine RRM/RNP-1-Domäne verfügen durch Ψ abgestoßen werden.

Um zuverlässige Aussagen über Ψ -interagierende Proteine treffen zu können erscheint eine Wiederholung der Experimente mit alternativen *Baits* sinnvoll. Ein geringerer Modifikationsgrad der Oligonukleotide würde die Auswirkung des Ψ auf ihre Sekundärstruktur abmildern. Die fehlende Einheitlichkeit der modifizierten Sequenzmotive macht eine globale Analyse des Ψ -Interaktoms aber erheblich komplexer als für andere Modifikationen. Die in **Abbildung 39** dargestellten Daten bilden aber einen ersten Hinweis auf eine Funktion von Ψ in der Steuerung von prä-mRNA-Polyadenylierung, die eine weitere Untersuchung des Interaktoms dieser Modifikation lohnenswert erscheinen lässt.

5.5 Interaktomanalyse der 2'-*O*-Methylmodifikationen

Die 2'-*O*-Methylierungen stellen eine Gruppe posttranslationaler Modifikationen dar, die sowohl global, als auch in rRNA hoch abundant sind.³⁴⁴ Mit den durch *Veronika Reiter* durchgeführten Quantifizierungsstudien gibt es Hinweise darauf, dass diese ebenfalls in mRNA in signifikanten Mengen vorhanden sind,¹⁰⁹ auch wenn es bislang keine Sequenzinformationen gibt, die dies bestätigen. Es ist jedoch anzunehmen, dass humane mRNA durch 2'-*O*-Methylierungen von einer Interferonantwort ausgenommen, also als zelleigen markiert wird. Viren nutzen zum Teil eigene Methyltransferasen um virale RNA vor einer Immunreaktion zu schützen.^{105,106,345}

Da, abgesehen von der 5'-Capstruktur, keine Motive für Nm-Loci in mRNA bekannt sind wurden mit ORN 23-26 (**Tabelle 5**) Sequenzen entworfen, die keine ausgeprägte Neigung zum Ausbilden von Sekundärstrukturen aufweisen (Simulation durchgeführt mit *RNAfold WebServer*³³⁸). Um eine hohe Anreicherung von Readern zu erreichen, enthalten sie die Modifikationen dreifach, eingebettet in verschiedene Umgebungen.

Tabelle 5: Sequenzen der Oligoribonukleotide, die für Interaktomstudien der 2'-*O*-Methylmodifikationen genutzt wurden. B = 5'-Biotineinheit (*Link Tech*).

ORN	Sequenz
ORN-23	5'-B UGC GUmU CUC AAUm ACC UGUm CUC GA-3'
ORN-24	5'-B UGC GAmU CUC AAAm ACC UGAm CUC GA-3'
ORN-25	5'-B UGC GCmU CUC AACm ACC UGCm CUC GA-3'
ORN-26	5'-B UGC GGmU CUC AAGm ACC UGGm CUC GA-3'

Trotz der weitgehend identischen Sequenzumgebung zeigen die *Pulldowns* stark heterogene Ergebnisse für die untersuchten Modifikationen. Für Cm ist eine kleine Gruppe von neun spezifisch bindenden Proteinen identifizierbar (**Abbildung 42**). Die beiden höchstangereicherten, NONO und SFPQ sind in *Nuclear paraspeckles* als Komponenten von Heteromultimeren in verschiedene zelluläre Prozesse involviert, wie die DNA-Reparatur³⁴⁶⁻³⁴⁸, transkriptionelle Regulation³⁴⁹ aber auch in die Retention von mRNAs im Zellkern^{350,351} (NONO = *Non-POU domain-containing octamer-binding protein*; SFPQ = *Splicing factor, proline- and glutamine-rich*).

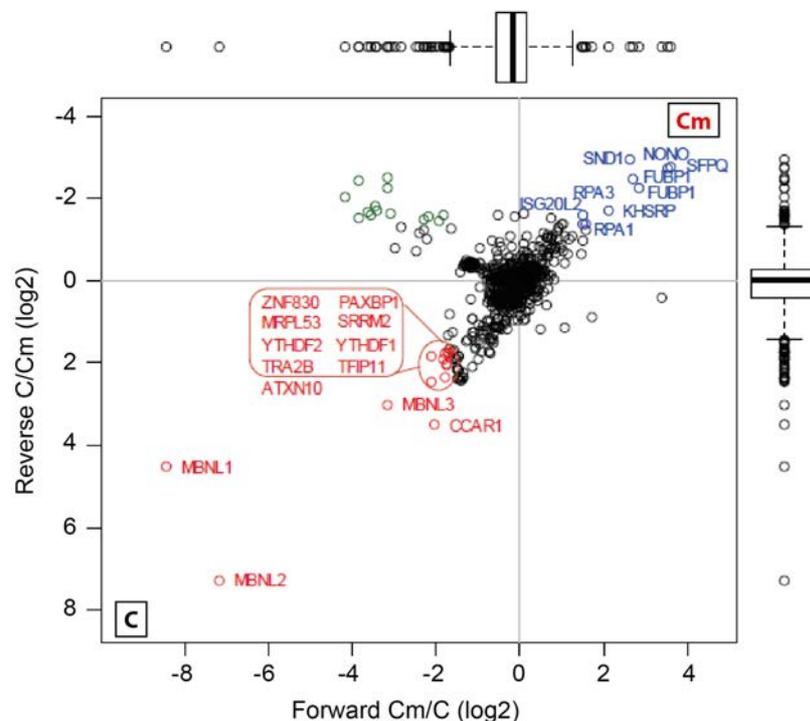


Abbildung 42: Scatterplot eines SILAC-basierten *Pulldowns* aus HeLa-Gesamtlysat für ORN-25/25': 5'-Biotin-UGC GCmU CUC AACm ACC UGCm CUC GA-3'. Das Experiment wurde von Raghu Edupuganti durchgeführt.

Auffällig ist, dass mit RPA1 und RPA3 zwei weitere Faktoren identifiziert wurden, die in DNA-Reparatur und Replikation involviert sind³⁵² (RPA1 = *Replication protein A 70 kDa*

DNA-binding subunit; RPA3 = Replication protein A 14 kDa subunit). KHSRP und eine Gruppe aus drei abgestoßenen, homologen Proteinen MBNL1-3 üben regulative Funktionen im alternativen Splicing aus³⁵³⁻³⁵⁶ (KHSRP = KH type-splicing regulatory protein; MBNL1-3 = Muscle-blind like protein 1-3) Interessant ist zudem, dass mit YTHDF1/2 zwei m⁶A-bindende Proteine schwach angereichert sind.

Dies ist ebenso für Gm der Fall (**Abbildung 43**), wo mit YTHDF1-3 die gesamte Gruppe stark durch die Modifikation abgestoßen wird. Der Effekt ist ausgeprägter als für Cm. Zudem sind auch die drei MBNL-Proteine zu finden und mit G3BP1/2 zwei Faktoren, die auch durch m⁶A abgestoßen werden. Die Bindung der RNA durch G3BPs dient vermutlich ihrer Rekrutierung zu Stressgranula und damit einhergehend ihrer Stabilisierung (siehe auch **Abbildung 34** und Kapitel 5.2).

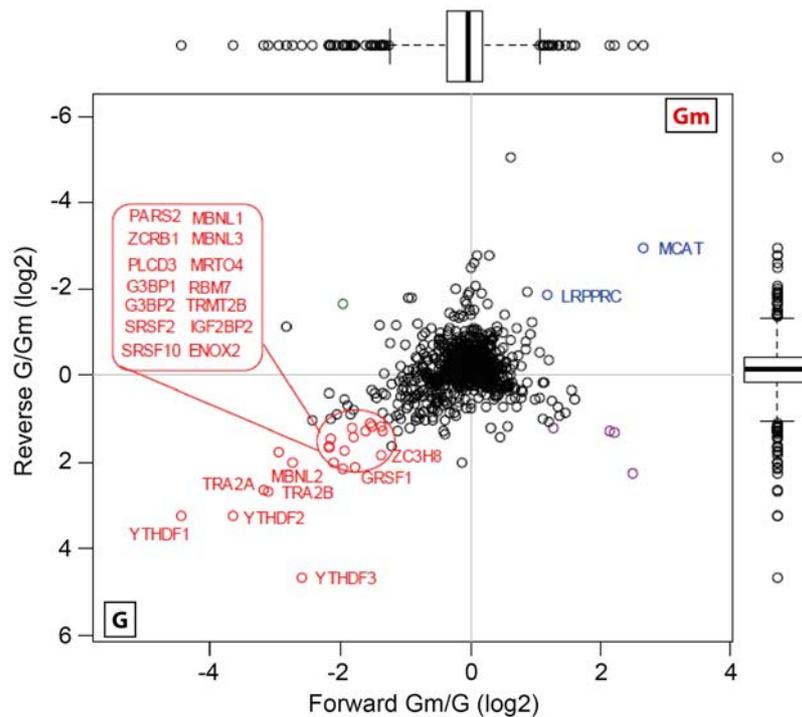


Abbildung 43: Scatterplot eines SILAC-basierten Pulldowns aus HeLa-Gesamtlysat für ORN-26/26': 5'-Biotin-UGC GGmU CUC AAGm ACC UGGm CUC GA-3'. Das Experiment wurde von Raghu Edupuganti durchgeführt.

Mit SRSF2/10 sowie TRA2A und B (*Transformer-2 protein homolog alpha* und *beta*) und den MBNL-Proteinen sind insgesamt sieben mit (alternativem) Splicing assoziierte Proteine angereichert.³⁵³⁻³⁵⁸ Für SRSF10 wurde gezeigt, dass es in einer direkten Interaktion mit YTHDC1 m⁶A-vermittelt *Exon skipping* verursacht.¹⁶³ Es besteht also ein direkter Zusammenhang zwischen einer RNA-Methylierung und der Regulation alternativen

Zudem wurden RAVER1 und 2 (*Ribonucleoprotein PTB-binding 1 und 2*) identifiziert. RAVER1 ist als Coaktivator von MDA5³⁶³ in die zelluläre Virenabwehr involviert. *Züst et al.* konnten bereits zeigen, dass MDA5, in Abhängigkeit davon, ob 2'-*O*-Methylierungen vorhanden sind, virale RNA detektiert und eine Immunantwort auslöst.¹⁰⁷ Die SILAC-Daten geben einen Hinweis darauf, dass diese Effekte durch eine RAVER1-RNA-Interaktion zustande kommen könnten, welche durch Um unterbunden wird.

Für Am, die letzte der vier Ribosemethylierungen, wurden *Pulldowns* mit nach Kern und Zytosol getrennten Lysaten durchgeführt, die Resultate überschneiden sich zum Teil (**Abbildung 45**). Vor allem die Gruppe der Binder zeigt ein hohes Maß an Übereinstimmung. Stark angereichert ist, ebenso wie für Um, HDLBP und analog zu Cm RPA-Proteine, in diesem Fall RPA1-3. Neben diesen sind auch NONO, SFPQ, KHSRP und ISG20L2 wieder als *Reader* identifiziert. Auffällig ist, dass mit FMR1 und FXR1/2 die dem Fragilen-X-Syndrom zugeordneten Proteine angereichert sind, wie auch für m⁶A.

Die bevorzugt an Adenosin bindenden Proteine geben ein heterogenes Bild ab, mit weniger klaren funktionellen Gruppen als im Fall der anderen Methylmodifikationen. Pop1/4/5/7 (*Ribonucleases P/MRP protein subunit POP1, 4, 5 und 7*) gehören zu den Untereinheiten des Ribonuklease P-Komplexes (RNase P), der in die Prozessierung von Prä-tRNA involviert ist.^{364,365} Auch RPP21/25/30/38/40 sind Komponenten der Ribonuklease P.³⁶⁵ Damit sind neun Untereinheiten dieses endoribonukleolytisch aktiven Komplexes abgereichert. Es ist möglich, dass die Erkennung des prä-tRNA Substrates durch den Ribonuklease P-Komplex von Am beeinflusst wird. Es ist bekannt, dass Modifikationen der tRNA in Eukaryoten unter anderem der Qualitätskontrolle dienen und Abbau und Lokalisierung beeinflussen.³⁶⁶⁻³⁶⁹ Für Am wurde dies bislang nicht gezeigt, es gibt aber Hinweise auf einen Zusammenhang zwischen dem Am-Level in tRNA und ihrem Modifikationsgrad durch andere Basenderivate.¹¹⁴ Die Degradierung hypomodifizierter tRNA anderen Nukleasen zugeschrieben (RAT1/XRN1). Eine Rolle der Modifikation in der Reifung von prä-tRNA ist aber denkbar. Neben tRNAprozessiert RNase P auch einige andere, hauptsächlich nicht-codierende RNA-Spezies.³⁷⁰ Auch diese Prozesse könnten durch Am moduliert werden.

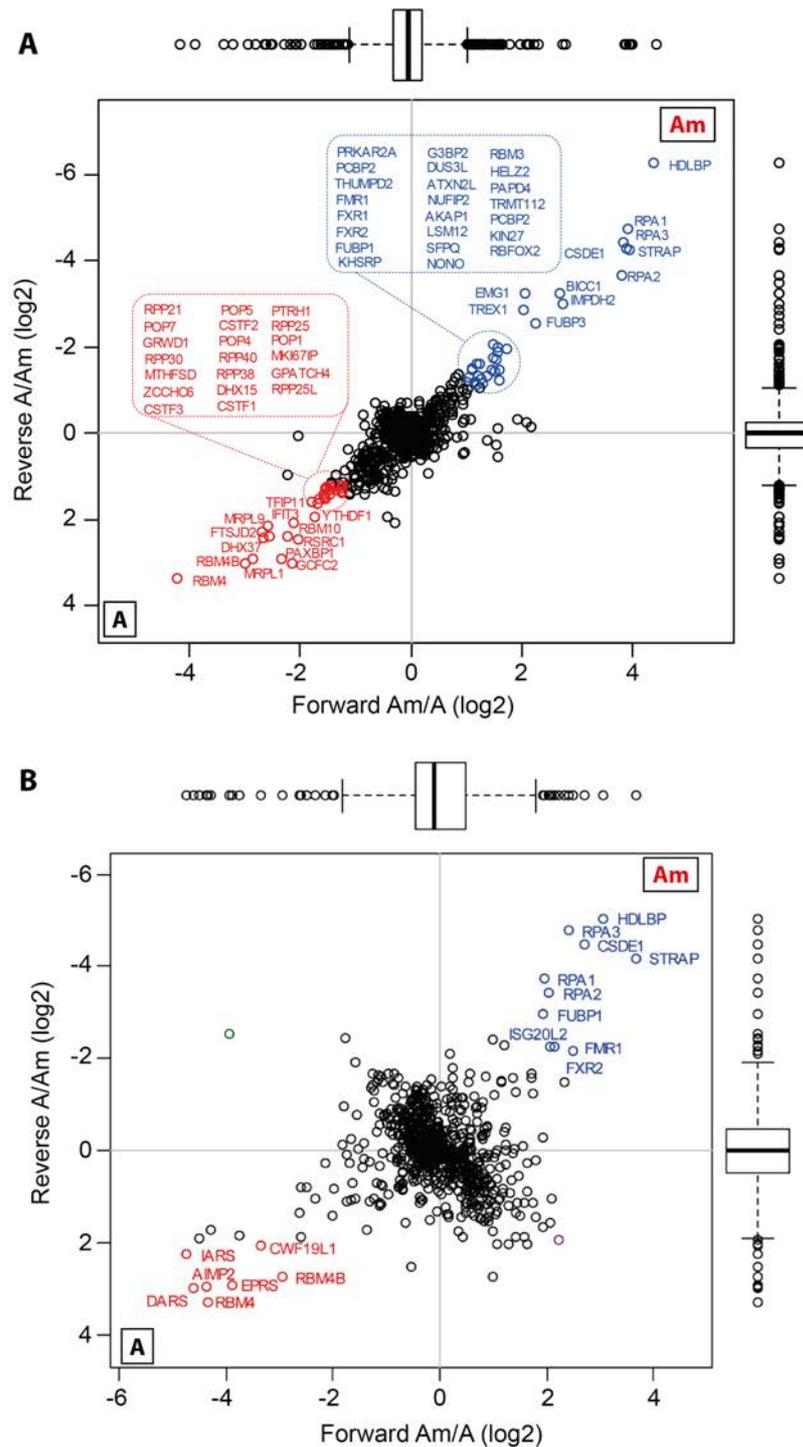


Abbildung 45: Scatterplots SILAC-basierter *Pulldowns* aus HeLa Extrakten für ORN-24/24': 5'-Biotin-UGC GAmU CUC AAAm ACC UGAm CUC GA-3'. A) *Pulldown* aus zytosolischem Lysat. B) *Pulldown* aus Kernextrakt. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Eine bioinformatische Analyse der Ergebnisse kann Trends in den Daten zeigen, also eine Anreicherung oder Abreicherung funktioneller Cluster, die sich einer manuellen Suche nicht notwendigerweise erschließen. Für die Analyse der Interaktome aller 2'-*O*-Methylierungen

wurde DAVID (Version 6.8)^{342,343} verwendet. Die *Functional Annotation Cluster* sind in **Abbildung 46** dargestellt.

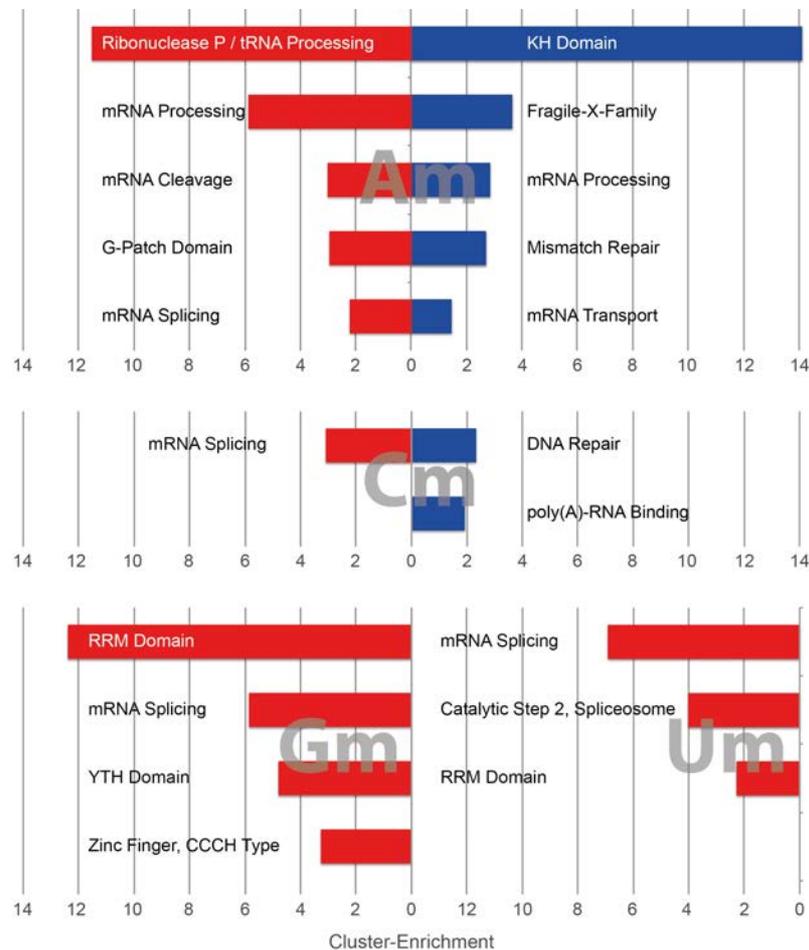


Abbildung 46: *Gene Ontology* Analyse der angereicherten (blau) bzw. abgereicherten (rot) Proteine der SILAC-Experimente für Am, Cm, Gm und Um. Die Analyse wurde mit DAVID (Version 6.8) durchgeführt. Angereicherte Proteine für Gm und Um wurden aufgrund ihrer geringen Anzahl nicht mit einbezogen. Die ersten fünf an-/abgereicherten Cluster sind dargestellt, sofern vorhanden. Bei der Aggregation der identifizierten Attribute (*Functional Clustering*) wurde die Einstellung *Medium Stringency* gewählt. Für alle Cluster gilt $p < 0.05$ (Benjamini-Hochberg).

Die mit den individuellen Ribosomethylierungen assoziierten Proteine sind in uneinheitliche funktionelle Cluster gruppiert. Eine einheitliche Funktion aller vier Modifikationen, beschränkt auf die Markierung zelleigener mRNA, erscheint damit unwahrscheinlich. Die hohe Abundanz von mit mRNA verbundenen biologischen Prozessen stützt die These, dass es sich um Basen handelt, deren Vorkommen sich nicht auf rRNA und tRNA beschränkt. Für alle analysierten Modifikationen finden sich Splicingfaktoren unter den abgestoßenen Proteinen. Als Gruppe könnten Ribosomethylierungen also die Aufgaben besitzen, die Affinität von Splicingfaktoren zur RNA zu mindern.

Da die analysierten Gruppen zum Teil nur wenige Proteine umfassen, ist nicht in jedem Fall eine starke Anreicherung funktioneller Gruppen möglich. Darüber hinaus sind individuelle

Interaktionen (wie die zwischen SFPQ und NONO) in dieser Analyse nicht sichtbar, obwohl sie von hoher Relevanz sein können.

Eine grafische Zusammenfassung einiger wichtiger Proteine aus den SILAC-Analysen ist in **Abbildung 47** dargestellt. Während *Reader* erkennbar sind, die mehrere der Nm-Derivate binden, wie HDLBP und ISG20L2, so wird doch sichtbar, dass insgesamt nur ein geringer Grad an Übereinstimmung vorhanden ist.

Hervorzuheben ist die Abstoßung von YTHDF-Proteinen durch Cm und, deutlich ausgeprägter, Gm. Da das m⁶A enthaltende Konsensmotiv in mRNA (GGACU) zwei Guanosine enthält^{71,72,91} ist eine Veränderung seines Interaktoms durch Gm vorstellbar. Die selektive Reduktion der Affinität der YTHDF-Proteine würde eine Verringerung des Interaktoms auf YTHDC1/2 bewirken und damit eine starke Veränderung der Auswirkung eines m⁶As auf die mRNA. Gemeinsam ist Gm und Cm auch die negative Interaktion mit den MBNL-Proteinen, die alternatives Splicing beeinflussen.³⁵³⁻³⁵⁶

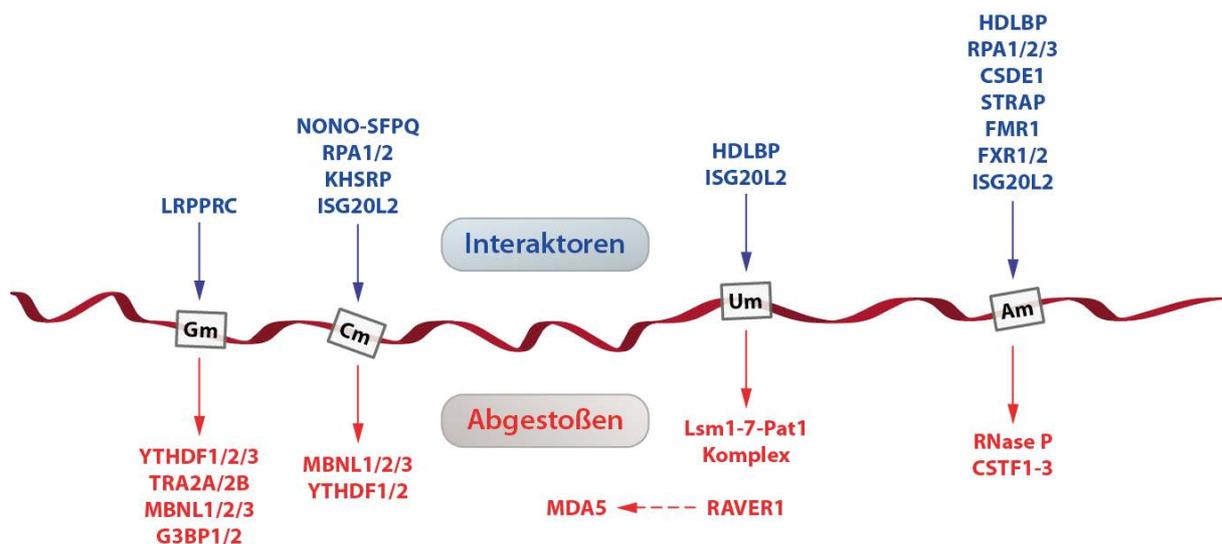


Abbildung 47: Auswahl der wichtigsten in SILAC-Studien identifizierten Interaktoren (selektiven Binder) und abgestoßenen Proteine für Gm, Cm, Um und Am. Die Proteine sind grob absteigend nach ihrer Affinität geordnet.

Die einzige bislang publizierte Hypothese zur Funktion von 2'-O-Methylmodifikationen in mRNA ist die Markierung von zelleigenen Transkripten zur Vermeidung einer Immunantwort.¹⁰⁵⁻¹⁰⁷ Daher galt ein besonderes Augenmerk in der Analyse der Ergebnisse der Identifikation von Proteinen, die in die antivirale Reaktion der Zelle auf eine Infektion involviert sind. In diesem Zusammenhang ist lediglich RAVR1 auffällig, da es mit MDA5 gemeinsam virale RNA bindet und eine Immunantwort auslöst.³⁶³ Die Selektion von viraler

RNA durch RAVER1, vermittelt durch Um, ist denkbar und würde eine interessante Präzisierung der aktuellen Theorie darstellen.

Außerdem ist hervorzuheben, dass für Um und Am mit dem Lsm1-7-Pat1 Komplex, RNase P und der CSTF1-3 (*Cleavage stimulation factor subunit 1-3*) mehrere nukleolytisch aktive Komplexe abgereichert sind, von denen Lsm/Pat und das CSTF-Trimer eine bekannte Aktivität gegenüber mRNA besitzen. Beide Modifikationen könnten hierdurch den mRNA-Metabolismus beeinflussen. Die parallele Abreicherung zahlreicher Komponenten der bekannten Komplexe in den individuellen Analysen spricht für eine Relevanz der Ergebnisse.

Die Analysen zeigen, dass alle vier Modifikationen individuelle Interaktome besitzen, die auf Funktionen in der Kontrolle der mRNA-Lokalisierung, des *mRNA turnover*, dem alternativen Splicing und der Unterscheidung von zelleigener und zellfremder RNA hinweisen und einen hervorragenden Ausgangspunkt für weitergehende Untersuchungen bilden.

5.6 Studien zur Funktion von ms²i⁶A

5.6.1 Interaktomanalyse

Die Untersuchung des Interaktoms von ms²i⁶A und seiner biologischen Relevanz stellte einen Schwerpunkt dieser Arbeit dar. Die ersten Hinweise auf ms²i⁶A in mRNA fand *Veronika Reiter*,^{109,196} womit es die erste Hypermodifikation in mRNA darstellt. Im Vergleich zu den bislang betrachteten Derivaten der kanonischen Basen ist es ein strukturell erheblich komplexeres Ribonukleosid. In seiner zweistufigen Biosynthese entsteht zunächst i⁶A aus Adenosin, durch Übertragung einer Isopentenylgruppe aus Dimethylallylpyrophosphat. Anschließende Thiomethylierung an Position 2 erzeugt schließlich ms²i⁶A. Dabei ist das Intermediat i⁶A auch als eigenständige Modifikation bekannt (siehe auch Kapitel 1.3.7.2).

Um Näheres zu seiner biologischen Relevanz in Erfahrung zu bringen, wurden SILAC-Studien mit i⁶A und ms²i⁶A in unterschiedlichen Sequenzkontexten durchgeführt (**Tabelle 6**) durchgeführt. Als Ausgangspunkt diente ORN-1, das bereits von *Veronika Reiter*, ohne Biotin-Tag, hergestellt und untersucht wurde. Es stellt einen Ausschnitt der *E. coli* tRNA^{Phe} dar, die ein ms²i⁶A trägt. Die Thiomethylierung wird durch Cdk5Rap1 katalysiert und es ist nicht unwahrscheinlich dass der *Hairpin*, den ORN-1/2/2'' ausbilden eine essentielle Voraussetzung für die Erkennung als Substrat darstellt.

Tabelle 6: Oligoribonukleotide, die für SILAC-Interaktomstudien zu den Modifikationen i⁶A und ms²i⁶A verwendet wurden (B = 5'-Biotineinheit (*Link Tech*); bU = Click-Biotin **33**).

ORN	Sequenz
ORN-1	5'-bU GGG GAU UGA Ams ² i ⁶ AA UCC CC-3'
ORN-2	5'-B GGG GAU UGA Ams ² i ⁶ AA UCC CC-3'
ORN-2''	5'-B GGG GAU UGA Ai ⁶ AA UCC CC-3'
ORN-4	5'-B GAC UUG Ams ² i ⁶ A AUU GUA Ams ² i ⁶ AA UCA U-3
ORN-4'	5'-B GAC UUG AAi ⁶ A AUU GUA Ai ⁶ AA UCA U-3
ORN-5	5'-B GGms ² i ⁶ ACU GGms ² i ⁶ ACU GGms ² i ⁶ ACU GGms ² i ⁶ ACU-3'

Interaktomstudien in HeLa Zellextrakten zeigen komplexe Gruppen interagierender und abgestoßener Proteine für ORN-1 (**Abbildung 48**), wobei auffällt, dass aus Kernextrakten deutlich weniger Proteine identifiziert wurden. Dies legt die Vermutung nahe, dass relevante Interaktionen zwischen RNA-bindenden Faktoren und der Modifikation überwiegend im Zytosol stattfinden.

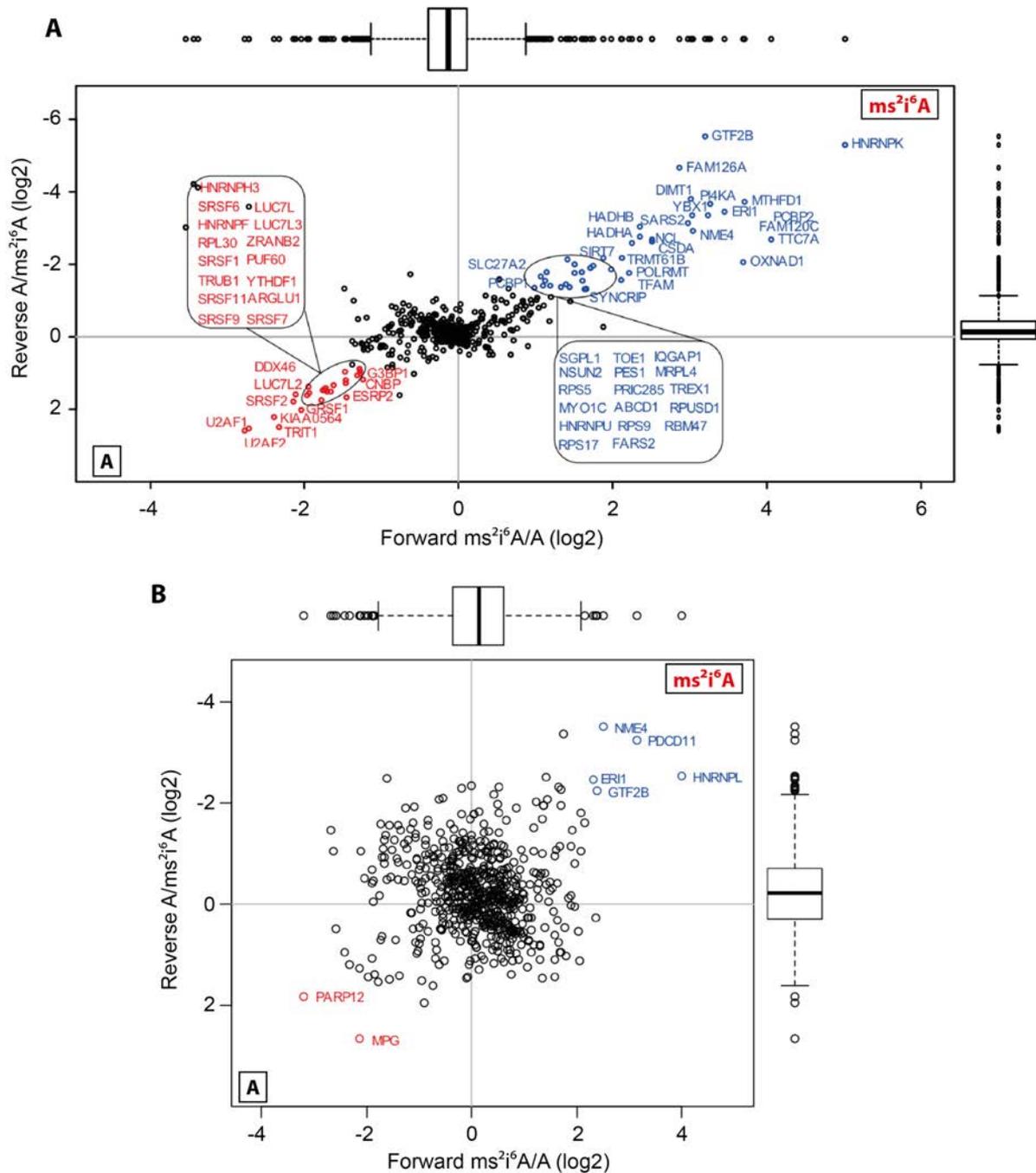


Abbildung 48: Scatterplots SILAC-basierter *Pulldowns* aus HeLa Extrakten für ORN-1/1': 5'-bUGGGGA UUGAAms²i⁶AAUCCCC-3'. A) *Pulldown* aus zytosolischem Zelllysate. B) *Pulldown* aus Kernextrakt. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Die höchstangereicherten *Reader*-Proteine HNRNPK (*Heterogeneous nuclear ribonucleoprotein K*) und GTF2B (*General transcription factor 2B*, auch TF-2B oder TF-IIB) stehen in keinem offensichtlichen Zusammenhang. GTF2B ist ein essentieller Faktor in der Transkriptionsinitiation.³⁷¹ HNRNPK ist ein Bestandteil von Ribonukleoproteinkomplexen, also Partikeln aus unterschiedlichen Proteinen, mRNA und snRNA, die mit gerade

entstehenden Transkripten assoziiert sind.³⁷² TRIT1, die humane Dimethylallyltransferase wird von m^2i^6A abgestoßen, was nicht überrascht, da der Kontrollstrang sein natürliches (unmodifiziertes) Substrat darstellt.

Pulldowns aus murinem Stammzelllysat bestätigen den Trend der vornehmlich zytosolischen Interaktoren (**Abbildung 49**) und reproduzieren die hoch angereicherten Binder HNRNPK, GTF2B und FAM120C (*Constitutive coactivator of PPAR-gamma-like protein 2*).

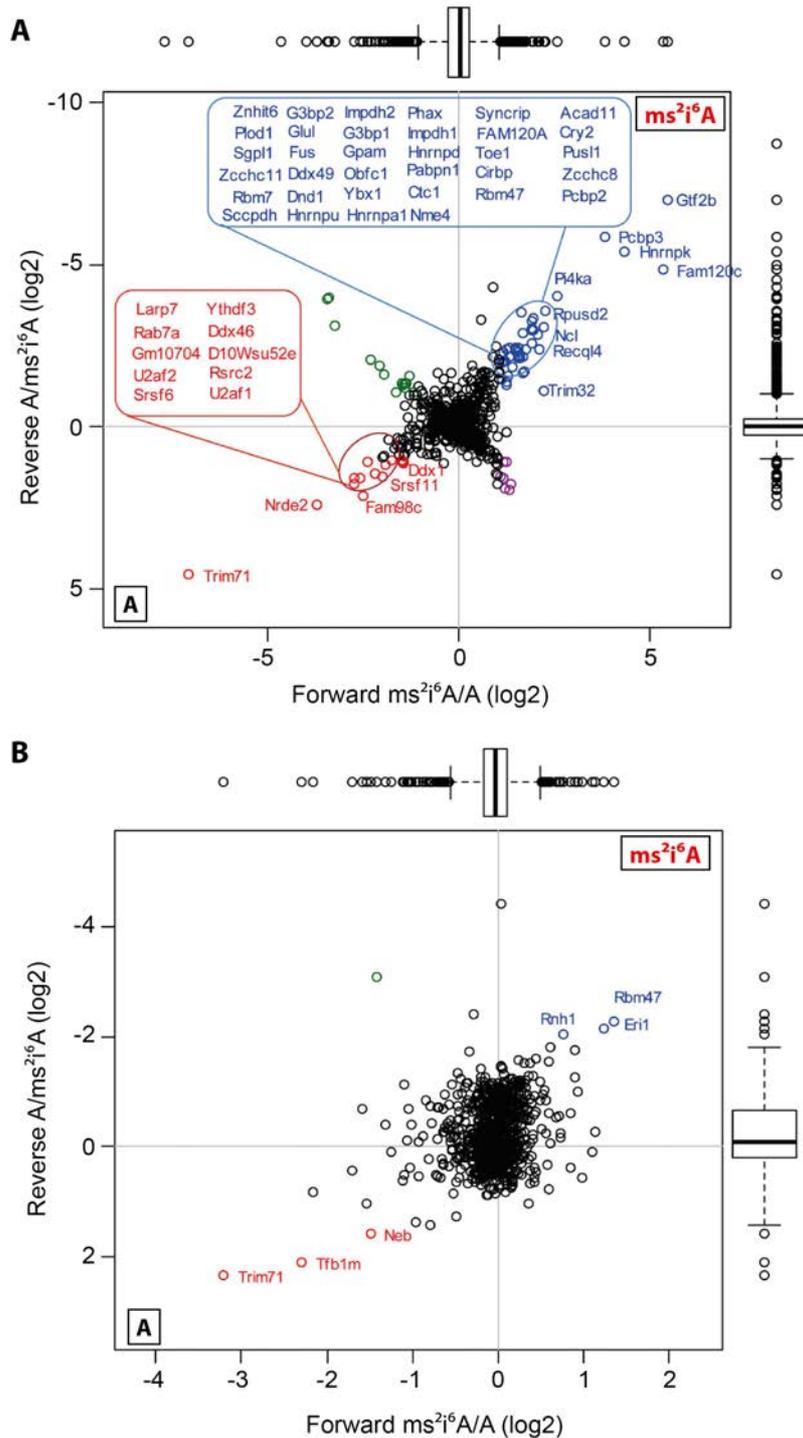


Abbildung 49: Scatterplots SILAC-basierter *Pulldowns* aus mESC Extrakten für ORN-1/1': 5'-bUGGGGA UUGAAms²i⁶AAUCCCC-3'. A) *Pulldown* aus zytosolischem Zelllysate. B) *Pulldown* aus Kernextrakt. Die Experimente wurden von *Raghu Edupuganti* durchgeführt.

Insgesamt ist der Überlapp zwischen mESC und HeLa Interaktom in zytosolischem Lysat mit 19 aus 88 (22%) moderat. Eine Analyse der an- und abgereicherten Proteine mittels DAVID (Version 6.7)^{342,343} zeigt, dass mitochondriale und mit Splicing assoziierte Proteine unter den Readern überrepräsentiert sind (**Abbildung 50**). Splicing ist als funktionelle Kategorie auch

unter den abgestoßenen Proteinen vertreten, zusammen mit der teilweise redundanten Kategorie der mRNA-Prozessierung.

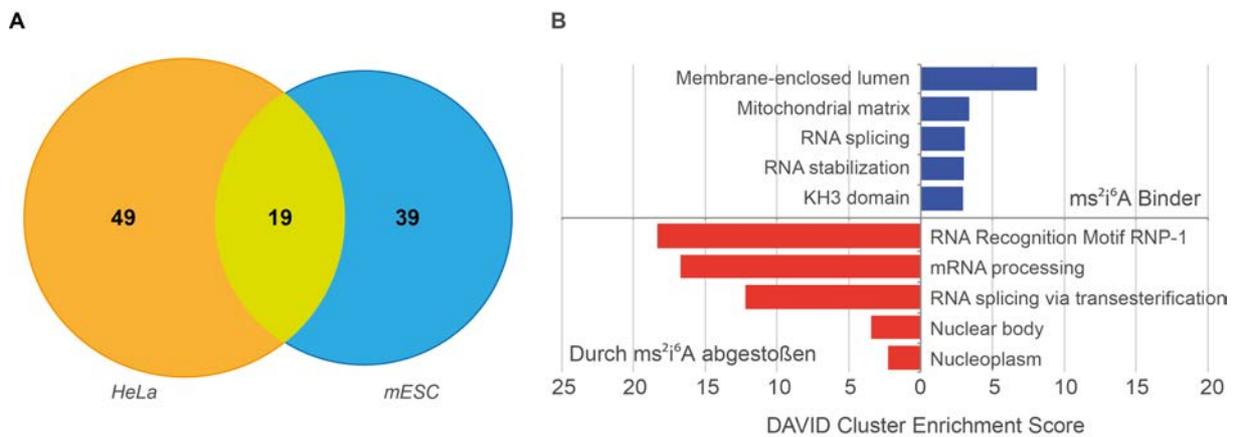


Abbildung 50: Analyse der an- und abgestoßenen Proteingruppen für ORN-1/1' in zytosolischem HeLa- und mESC-Lysat. A) Venn-Diagramm der identifizierten Proteine. B) DAVID *Gene Ontology* Analyse der 49 Proteine, die in HeLa-Lysat identifiziert wurden.

Gemeinsam mit HNRNPK, das in HeLa und mESC zu den dominanten *Readern* gehört, sind weitere Proteine angereichert, die *in vivo* mit ihm interagieren, darunter YBX1³⁷³, NCL³⁷⁴, PCBP1^{375,376}, PCBP2^{377,378} und SYNCRIP³⁷⁹ (YBX1 = *Y-box binding protein 1*, auch *Nuclease-sensitive element binding protein 1*; NCL = *Nucleolin*; PCBP1 = *Poly(rC)-binding protein 1*, auch HNRNPE1; PCBP2 = *Poly(rC)-binding protein 2*, auch HNRNPE2; SYNCRIP = *Heterogeneous nuclear ribonucleoprotein Q*, auch HNRNPQ). Viele der gemeinsam angezogenen Faktoren sind in die Regulation der Translation, alternatives Splicing und die Lokalisation gebundener Transkripte involviert.

HNRNPK verfügt über drei KH-Domänen, die kooperativ seine RNA-Spezifität bestimmen.³⁸⁰ Weiterhin besitzt es zahlreiche posttranslationale Modifikationen, die seiner Regulation und Lokalisation dienen. Es ist an mehreren Tyrosinen und Serinen phosphoryliert^{376,381-383}, durch die Kinasen c-Src, ERK und PKC δ . Darüber hinaus trägt es fünf methylierte Arginine.³⁸⁴

HNRNPK bindet, ebenso wie PCBP1-3, C-reiche RNA-Sequenzen, darunter auch mitochondriale Transkripte.^{385,386} Gemeinsam mit PCBP1 und 2 beeinflusst es den Metabolismus von Transkripten, indem es sie zu translationell reprimierenden mRNPs lokalisiert, ihre Translation durch Bindung an IRES-Elemente (*Internal ribosomal entry sites*) stimulieren und alternatives Splicing beeinflussen.³⁷⁷ Es konkurriert direkt mit U2AF2 (*Splicing factor U2AF 65 kDa subunit*, auch U2AF65) um eine Bindestelle zweier mRNAs (Snap25 und Runx1), um an diesen Loci Splicing zu unterdrücken,³⁸⁷ was besonders

interessant erscheint, da zum einen HNRNPK in den SILAC-Experimenten hoch angereichert ist, zum anderen U2AF2 und seine weitere Untereinheit U2AF1 (auch U2AF35) stark abgereichert sind. Die durch *Cao et al.* beschriebene exklusive Bindung derselben Sequenzen ist eine mögliche Erklärung für diese Beobachtung. Eine vergleichbare regulative Relation wurde auch für U2AF2 und HNRNPC gezeigt.³⁸⁸

Neben U2AF1/2 sind mit SRSF1/2/6/7/9/11 und PUF60 weitere Splicingfaktoren abgereichert. Eine mögliche funktionelle Verbindung des Transkriptionsfaktors GTF2B zu den restlichen Interaktoren stellt seine Assoziation mit dem Lsm2-8-Komplex dar,³⁸⁹ der in die Lokalisation der U6 snRNA in Hefe involviert ist.³⁹⁰

Neben alternativem Splicing ist eine Beeinflussung des mRNA-Metabolismus möglich, unter den spezifischen Interaktoren befinden sich mit YBX1, HNRNPU und SYNCRIP drei Komponenten des *Coding region instability determinant*-bindenden Komplexes, der mRNAs bindet und stabilisiert.³⁹¹

Um zu beurteilen, ob die identifizierten *Reader* spezifisch für ms^2i^6A sind, oder lediglich auf die Isopentenylseitenkette angewiesen sind, wurden auch mit i^6A Interaktomstudien durchgeführt (**Abbildung 51**). Im Gegensatz zu allen anderen hier diskutierten *Pulldowns* wurde in diesem Fall keine tRNA aus Hefe als Kompetitor eingesetzt, da sie i^6A enthält und damit Interaktionen unterdrücken könnte. Es wurde stattdessen DNA aus Lachssperma verwendet. Die zur Quantifizierung notwendige Isotopenmarkierung wurde mittels *Dimethyl labeling* eingeführt.³⁹²

Es konnten zwei Binder identifiziert werden. PABPC1 und 4 (*Polyadenylate-binding protein 1 und 4*) sind schwach angereichert und wurden für ms^2i^6A nicht gefunden. Die Gruppe der abgestoßenen Proteine ist mit 44 deutlich größer, der Überlapp mit den für ms^2i^6A in HeLa abgereicherten ist mit vier Proteinen (HNRNPH3, ESRP2, HNRNPF, GRSF1) gering. Dagegen werden sieben (FARS2, DIMT1, NME4, SARS2, POLRMT, HADHA und HADHB) durch i^6A abgereichert und von ms^2i^6A angezogen, verhalten sich also gegensätzlich.

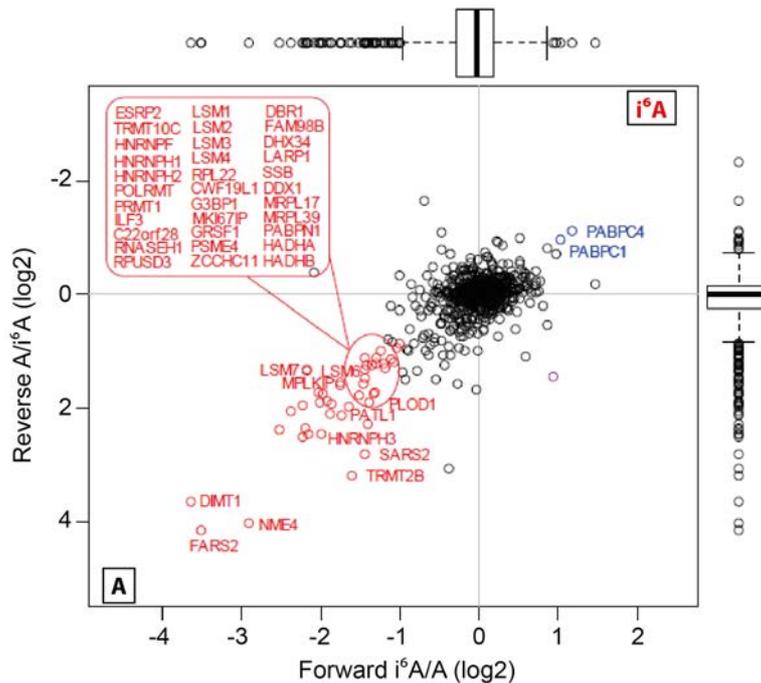


Abbildung 51: Scatterplot des *Pulldowns* aus zytosolischem HeLa-Extrakt für ORN-2'2'': 5'-Biotin GGGGAUUG AAi⁶AAUCCCC-3'. Das Experiment wurde von *Raghu Edupuganti* durchgeführt. Abweichend von der üblichen Versuchsdurchführung wurde hier mit *Salmon Sperm DNA* als Kompetitor gearbeitet, anstelle von tRNA aus Hefe. Die Isotopenmarkierung zur quantitativen Analyse erfolgte mittels *Dimethyl labeling*.³⁹²

Auffällig ist die Abstoßung von LSM1/2/3/4/6/7 und PATL1, Komponenten des Lsm1-7-Pat1-Komplexes,³⁶¹ der auch durch Um angereichert wurde (**Abbildung 44**).

Das am stärksten abgestoßene FARS2 (*Phenylalanine-tRNA ligase, mitochondrial*) belädt mitochondriale tRNA^{Phe} mit ihrer Aminosäure. Die Ergebnisse legen nahe, dass die Einführung der Thiomethylgruppe durch Cdk5Rap1 eine wichtige Markierung zur Qualitätskontrolle in der Biogenese der mt-tRNA^{Phe} darstellt, da i⁶A die Ligase abstößt, während ms²i⁶A zu einer stärkeren Bindung führt.

Der Vergleich der Interaktome zeigt, dass die große Mehrzahl der positiv, wie negativ mit ms²i⁶A interagierenden Proteine auf die Anwesenheit beider Substituenten angewiesen sind und nicht an i⁶A binden.

Die bislang beschriebenen Experimente wurde mit ORN durchgeführt, die *Hairpins* mit hohen Schmelzpunkten ausbilden.¹⁰⁹ Es ist anzunehmen, dass die Struktur des Oligonukleotids einen maßgeblichen Einfluss auf etwaige Interaktionen hat. Um zu analysieren, ob spezifische Binder die Modifikation ms²i⁶A auch außerhalb dieses Kontextes binden wurden weitere Interaktionsstudien mit ORN-4/4'/4'' durchgeführt, die das zentrale Motiv der ursprünglichen ORN (5'-UUGAAms²i⁶AAUC) enthalten, denen aber die

Möglichkeit zur Ausbildung der Sekundärstruktur fehlt. In diesem Fall wurde die Quantifizierung ohne Isotopenmarkierung durchgeführt (*label-free quantification*, LFQ). Der Überlappungsgrad ist mit 18 reproduzierten Readern (FAM120C, HNRNPU, HADHA, HADHB, FAM126A, DIMT1, PI4KA, MTHFD1, TTC7A, OXNAD1, NCL, TRMT61B, TFAM, SGPL1, NSUN2, ABCD1 und FARS2) moderat. Mit TFAM ist ein mitochondrialer Transkriptionsfaktor unter den gemeinsamen Interaktoren, der auch mt-tRNAs bindet.³⁹³ Die dominanten Reader HNRNPK und GTF2B fehlen, FAM120C dagegen bleibt erhalten. Für die überwiegende Mehrzahl der beobachteten Interaktionen scheint die Sekundärstruktur des *Hairpins* also essentiell zu sein.

Zusammenfassend lässt sich sagen, dass $ms^{2:6}A$ über ein komplexes Interaktom zu verfügen scheint, dass auf eine Funktion in der Steuerung des mRNA-Metabolismus und des alternativen Splicings hinweist. Die RNA-Sekundärstruktur, in die es eingebettet ist spielt in der Steuerung der Interaktionen eine zentrale Rolle.

5.6.2 Validierungsexperimente

Die beschriebenen Interaktomanalysen geben einen guten Einblick in die mit $ms^{2:6}A$ positiv, wie negativ interagierenden Proteingruppen. Um aus diesen Informationen belastbare funktionelle Hypothesen ableiten zu können, wurde versucht, wichtige Interaktoren in ihrer selektiven Bindung zu validieren. Als *Reader* mit der, neben GTF2B höchsten Affinität wurde zunächst HNRNPK untersucht.

5.6.2.1 Electromobility Shift Assay (EMSA)

Eine etablierte Methode, um Interaktionen zwischen Oligonukleotiden und Proteinen zu untersuchen, stellt der EMSA (*Electrophoretic mobility shift assay*) dar.³⁹⁴ In der Regel wird hier mit fluoreszent, oder radioaktiv markierten Oligonukleotiden gearbeitet und diese mit ansteigenden Mengen eines aufgereinigten Proteins versetzt. Findet eine Bindung zwischen beiden Partnern statt, äußert sich dies in einer veränderten elektrophoretischen Mobilität des markierten Substrates und es ist eine Verschiebung der zugehörigen Bande in einem geeigneten Gel sichtbar.

Um die Analyse für HNRNPK durchführen zu können wurde sein Gen in einen geeigneten Vektor (pDEST007) zur bakteriellen Expression kloniert, in *E. coli* exprimiert und aus diesem System mittels seines Strep-Tags aufgereinigt. Als Substrate wurden verschiedene ORN synthetisiert, die in **Tabelle 7** zusammengefasst sind. Die Aufreinigung des Proteins und alle folgenden Versuche wurden in Kooperation mit *Martin Rossa, Matthias Kurz* und *Jessica Steinbacher* durchgeführt.

Tabelle 7: Für die Validierungsexperimente zur Modifikation ms²ⁱ⁶A hergestellte Oligoribonukleotide.

ORN	Sequenz
ORN-6	5'-Cy3 GGG GAU UGA Ams ²ⁱ⁶ AA UCC CC-3'
ORN-6'	5'-Cy3 GGG GAU UGA AAA UCC CC-3'
ORN-7	5'-Cy3 GUG CAU UGA Ams ²ⁱ⁶ AA UCU AC-3'
ORN-8	5'-CAC UGU CAU CCC UUU GUC CUA ACC A-3'
ORN-9	5'-Cy3 CAC UGU CAU CCC UUU GUC CUA ACC A-3'
ORN-9'	5'-Cy3 CAC UGU CAU CCC UUU GUC CUms ²ⁱ⁶ A ACC A-3'
ORN-9''	5'-Cy3 CAC UGU Cms ²ⁱ⁶ AU CCC UUU GUC CUA ACC A-3'

Mit ORN-6/6' wurden Sequenzen hergestellt, die den *Baits* aus den SILAC-Experimenten entsprechen, ORN-7 enthält das zentrale Motiv der *Hairpins* (5'-AUUGAAms²ⁱ⁶AAUC-3'), kann aber keine Sekundärstruktur mehr ausbilden und ORN-9/9'/9'' entsprechen den Sequenzabschnitten der Snap25 mRNA, die von HNRNPK und U2AF2 in Konkurrenz gebunden wird.³⁸⁷ Hier wurden Adenosine durch ms²ⁱ⁶A substituiert, die dem für die Bindung ausschlaggebenden C-reichen Abschnitt am nächsten liegen, um zu analysieren ob die regulative Konkurrenz zwischen den beiden in SILAC-Studien für ms²ⁱ⁶A hoch an- bzw. abgereicherten Proteinen durch die Modifikation an dieser Position beeinflusst wird.

Zunächst wurde die Interaktion zwischen HNRNPK und ORN-6/6' untersucht. Es wurden zahlreiche Bedingungen, wie unterschiedliche Puffer, Salzkonzentration und Inkubationszeit bzw. -temperatur getestet, aber in keinem Fall konnte eine Bindung des Proteins an die eingesetzten Substrate detektiert werden (**Abbildung 52**).

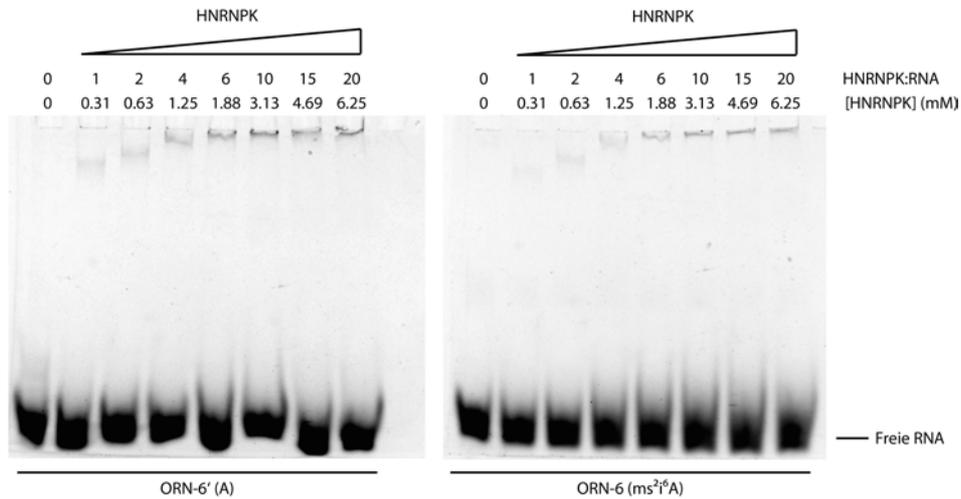


Abbildung 52: EMSA mit aus *E. coli* aufgereinigtem HNRNPK_StrepII. Als Substrate wurden ORN-6 und ORN-6' eingesetzt. Sie lagen in einer konstanten Konzentration (0.31 mM) vor. Die Reaktion wurde für 60 min bei 37 °C in VM-Puffer durchgeführt. Die Gelelektrophorese erfolgte bei 4 °C.

Parallel wurden ebenfalls Versuche mit ORN-9/9'/9'' durchgeführt, für das eine starke Bindung des Proteins beobachtet werden konnte (**Abbildung 53**).

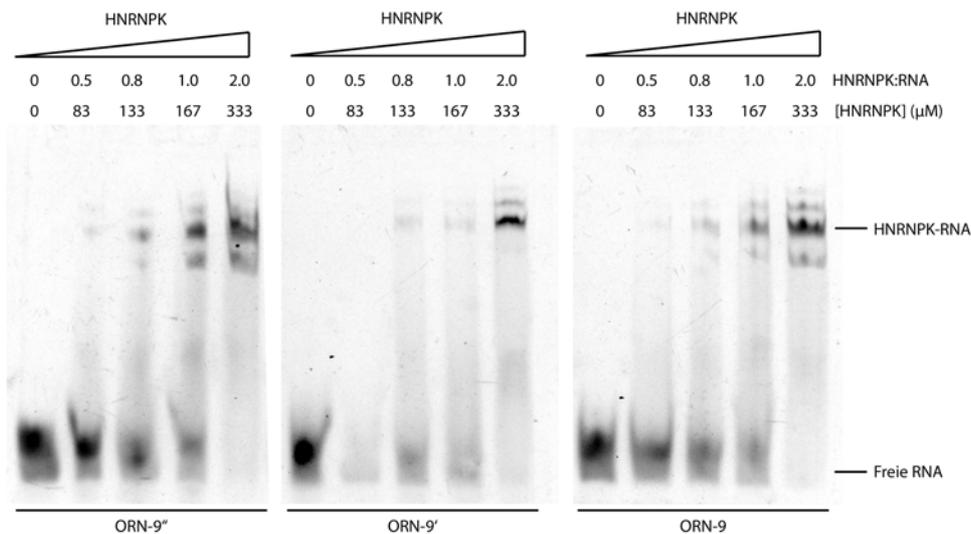


Abbildung 53: EMSA mit aus *E. coli* aufgereinigtem HNRNPK_StrepII. Als Substrate wurden ORN-9, ORN-9' und ORN-9'' eingesetzt. Sie lagen in einer konstanten Konzentration (167 µM) vor. Die Reaktion wurde für 30 min bei 25 °C in VM-Puffer durchgeführt. Die Gelelektrophorese erfolgte bei 4 °C.

Es wird deutlich, dass das Protein in einer Form aufgereinigt wurde, die fähig ist an RNA zu binden und unter den verwendeten Bedingungen aktiv ist. Es bindet alle betrachteten Varianten der von der Snap25-mRNA abgeleiteten Sequenz mit vergleichbarer Affinität. Bereits ein geringer Überschuss (2 eq.) ist ausreichend, um die eingesetzte RNA quantitativ zu verschieben.

Da HNRNPK eine hohe Affinität zu der linearen Sequenz der ORN-9/9'/9'' besitzt, die *Hairpins* ORN-6 und 6' aber nicht bindet, wurde ihr zentrales Motiv in ORN-7 in eine Umgebung eingebettet, die keine Sekundärstruktur bildet, um auszuschließen, dass diese einen negativen Einfluss auf die Bindung ausübt. Der direkte Vergleich zeigt, dass auch dieses Kernmotiv nicht, oder nur unwesentlich gebunden wird (**Abbildung 54**). Neben der Sekundärstruktur fehlt in ORN-7 auch ein C-reicher Abschnitt, über den alle sonstigen ORN verfügen und der ein Grund für Bindung durch HNRNPK sein kann,^{325,380,386,395} aber offensichtlich nicht ausreichend ist, beziehungsweise offen vorliegen muss.

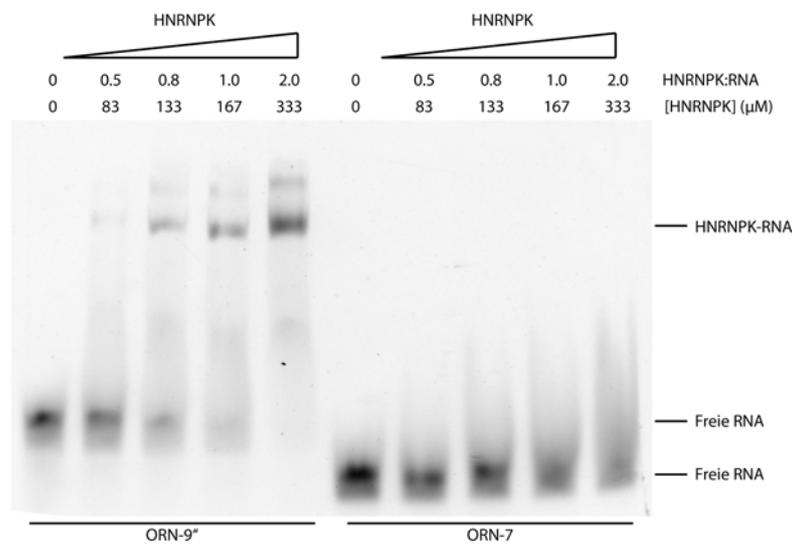


Abbildung 54: EMSA mit aus *E. coli* aufgereinigtem HNRNPK_StrepII. Als Substrate wurden ORN-9'' und ORN-7 eingesetzt. Sie lagen in einer konstanten Konzentration (167 μM) vor. Die Reaktion wurde für 30 min bei 25 °C in EMSA-A Puffer durchgeführt. Die Gelelektrophorese erfolgte bei 4 °C.

Eine Alternative zur Verwendung ansteigender Proteinkonzentrationen, wie in den bislang beschriebenen EMSAs stellt die Verdrängung des fluoreszenzmarkierten Substrates aus der Bindung durch das Zielprotein mit Hilfe eines nicht markierten und damit nicht sichtbaren Kompetitors dar. Da hierfür eine deutliche Bindung Voraussetzung ist wurde dieses Experiment nur mit ORN-9/9'/9'' durchgeführt, um eine möglicherweise vorhandene Präferenz für ein ms²i⁶A-modifiziertes Substrat sichtbar zu machen. Als Kompetitor wurde das unmodifizierte ORN-8 verwendet.

Auch in dieser Untersuchung war kein deutlicher Unterschied in der Bindung von HNRNPK an seine Substrate erkennbar (**Abbildung 55**). Alle fluoreszenzmarkierten Oligonukleotide wurden in vergleichbarem Umfang aus der Bindung durch HNRNPK verdrängt, indem ORN-8 in ansteigender Konzentration zugegeben wurde.

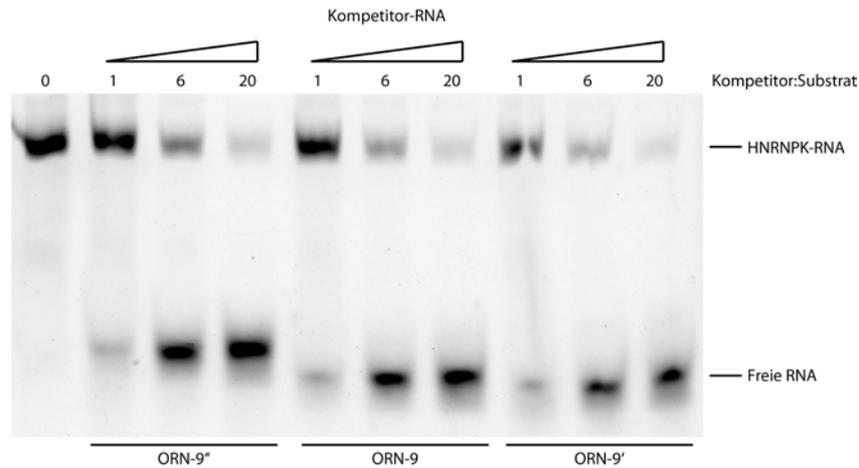


Abbildung 55: EMSA mit aus *E. coli* aufgereinigtem HNRNPK_StrepII. Als Substrate wurden ORN-9/9'/9'' eingesetzt. Sie lagen in einer konstanten Konzentration (167 μ M) vor. HNRNPK lag in einer ebenfalls konstanten Konzentration von 333 μ M (2 eq.) vor. Die Reaktion wurde für 30 min bei 30 °C in VM-Puffer durchgeführt, wobei von Beginn an sowohl das fluoreszenzmarkierte ORN, als auch der Kompetitor zugegeben wurden. Die Gelelektrophorese erfolgte bei 4 °C.

Aus den EMSA-Analysen lässt sich zusammenfassend ableiten, dass das isolierte Protein HNRNPK spezifisch RNA bindet, wobei es eine ausgeprägte Affinität gegenüber den von Snap25 abgeleiteten Oligonukleotiden besaß. Für diese war keine Präferenz zwischen den verwendeten Sequenzen sichtbar. Die Affinität des Proteins wird durch m^2i^6A also weder gesteigert, noch verringert. Alle anderen Stränge wurden von HNRNPK nicht in detektierbarem Umfang gebunden, was für ORN-6/6' überrascht, da sie den in den SILAC-Experimenten verwendeten Sequenzen entsprechen, mit denen HNRNPK stark angereichert wurde.

5.6.2.2 Fluoreszenzpolarisation

Eine mögliche Problemquelle der Bindungsstudie mittels EMSA ist die Analyse per Gelelektrophorese. Während der Optimierung der Bedingungen für die Durchführung der Trennung wurde deutlich, dass nur eine sorgfältig in ihrer Dauer und Temperatur abgestimmte Elektrophorese reproduzierbare Bindungsmuster lieferte. Eine Erhöhung der Temperatur während der Trennung führe zu einem teilweisen, bis vollständigen Verlust der Bindung und einem unklaren Ergebnis. Das Bindungsverhalten im Gel spiegelt nicht notwendigerweise die Affinitäten in Lösung wieder und eine präzise Bestimmung von kinetischen Parametern ist nicht ohne weiteres möglich.

Eine sensitive Methode, um eine vergleichbare Analyse in Lösung durchzuführen stellt die Fluoreszenzpolarisation (auch *Fluorescence anisotropy*) dar. Kurz zusammengefasst basiert sie auf der Tatsache, dass mit polarisiertem Licht angeregte Fluorophore ebenfalls

polarisiertes Licht ausstrahlen. In Abhängigkeit von ihrer Bewegungsfreiheit nimmt der Polarisationsgrad des emittierten Lichtes dabei ab. Wird ein fluoreszenzmarkiertes Substrat durch einen Reaktionsteilnehmer gebunden sinkt dessen Beweglichkeit in Lösung und der Polarisationsgrad des emittierten Lichtes steigt an.³⁹⁶⁻³⁹⁸

Die Messung wurde mit ORN-9/9'/9'' durchgeführt, dabei wurde eine konstante Konzentration des Oligonukleotids mit ansteigenden Konzentrationen des aufgereinigten Proteins inkubiert. Aus den Rohdaten ist unmittelbar ersichtlich, dass auch unter diesen Versuchsbedingungen (37 °C, Messung in Puffer) keine unterschiedlichen Affinitäten feststellbar sind (**Abbildung 56**), da die Kurven nahezu deckungsgleich sind. Ihr Wendepunkt definiert den messbaren K_d -Wert und wäre im Falle unterschiedlicher Affinitäten verschoben.

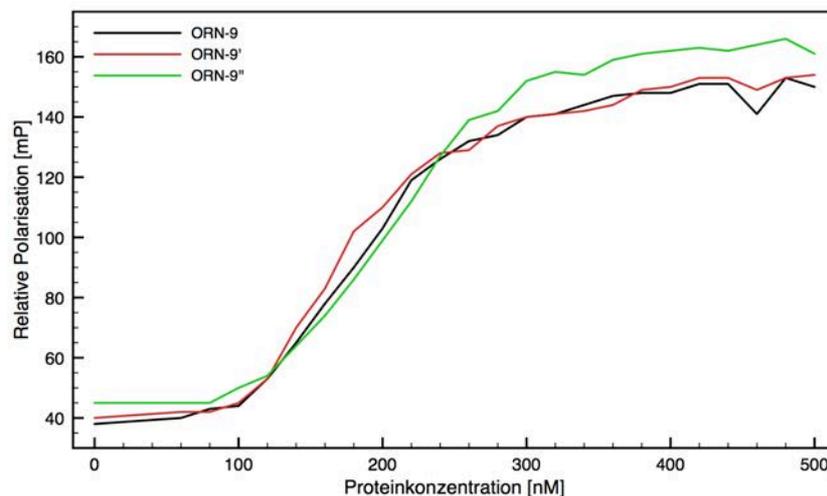


Abbildung 56: Fluoreszenzpolarisationsmessung für die Bindung von ORN-9/9'/9'' durch aus *E. coli* aufgereinigtes HNRNPK. Die Reaktion wurde in VM-Puffer durchgeführt. Die Konzentration des ORN betrug 80 nM, das Protein wurde in Konzentrationen zwischen 0 nM und 500 nM eingesetzt und die Reaktion für 30 min bei 25 °C inkubiert, bevor die Messung erfolgte.

5.6.2.3 GFP-Pulldown Assay

Eukaryotische Proteine, die aus bakteriellen Expressionssystemen rekombinant exprimiert werden, büßen unter Umständen ihre Funktionalität, oder Teile davon ein, wenn sie hierfür auf posttranslationelle Modifikationen (PTMs) angewiesen sind. HNRNPK trägt solche Modifikationen in Form von Arginin-Methylierung^{384,399} und Phosphorylierungen an Serinen und Tyrosinen.^{376,381-383} Es wurde gezeigt, dass diese Modifikationen seine Lokalisation und seine Fähigkeit RNA zu binden regulieren.^{376,399}

Um auszuschließen, dass die nicht vorhandene Selektivität im Falle des rekombinanten HNRNPKs aus fehlenden PTMs resultiert, wurde eine Bindungsstudie unter Verwendung von humanen Kulturzellen als Expressionssystem durchgeführt. Da es ungleich fordernder ist, große Quantitäten human exprimierter Proteine durch klassische Aufreinigung zu erhalten, wurde als alternativer Ansatz ein direkter *Pulldown* des Proteins aus Zellextrakt gewählt. HNRNPK wurde in einen geeigneten Vektor (pDEST47) kloniert und mit GFP-Tag in HEK293-Zellen exprimiert. Die Zellen wurden lysiert, der Zellextrakt mit fluoreszenzmarkiertem Oligonukleotid versetzt und nach einer Inkubation das HNRNPK mittels Anti-GFP-Beads (*ChromoTek*) aus dem Lysat angereichert. Bei einer Bindung der RNA durch das Protein wurde das ORN ebenfalls angereichert und konnte anschließend auf einem denaturierenden PAGE-Gel visualisiert werden (**Abbildung 57**). Neben dem Erhalt des natürlichen PTM-Musters bietet diese Herangehensweise den Vorteil, dass auch Komplexe aus mehreren Proteinen im Extrakt ausgebildet werden können, falls diese für eine Bindung an das Substrat notwendig sind.

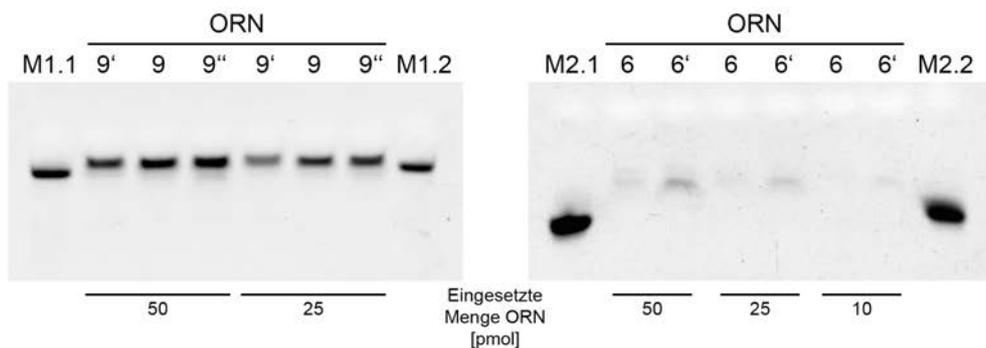


Abbildung 57: Harnstoff-PAGE Gel der *Pulldowns* von HNRNPK-GFP in Kombination mit fluoreszenzmarkierten ORN-6/6' und ORN-9/9'/9''. HNRNPK-GFP-exprimierende HEK293-Zellen wurden lysiert und die Extrakte mit den Oligonukleotiden versetzt. Nach einer Inkubation bei 0 °C für 30 min wurde das rekombinante Protein mittels *GFP-Trap* (*ChromoTek*) isoliert, mehrfach gewaschen und mit Gelladepuffer gelöst. Die anschließende Analyse erfolgte auf einem Harnstoff-PAGE-Gel (15%). Als Marker wurde ORN-9'' (M1.1 = 2.50 pmol; M1.2 = 1.25 pmol) oder ORN-6' (M2.1 = 2.50 pmol; M2.2 = 1.25 pmol) verwendet.

Wie aus **Abbildung 57** ersichtlich, zeigte auch dieses Experiment keine eindeutigen Unterschiede in der Bindung der Snap25-Derivate ORN-9/9'/9'' und keine Bindung der *Hairpin*-Sequenzen ORN-6/6', die über dem Niveau einer Negativkontrolle (nicht gezeigt) lag.

5.6.2.4 Gesamtlysate-EMSA

Ein *Tag*, wie der in dem vorhergehenden Experiment verwendete GFP-*Tag* bietet neben der Möglichkeit einer Aufreinigung oder Anreicherung weitere Vorteile, wie die fluoreszenzmikroskopische Lokalisierbarkeit und die Möglichkeit einer einfachen Quantifizierung des rekombinanten Proteins, auch in heterogenen Proben. Die hier eingesetzte Markierung (GFP) ist aber mit dem Nachteil behaftet, im Verhältnis zum untersuchten Protein groß zu sein. Dies beeinträchtigt möglicherweise die Funktion des Proteins, oder seine Möglichkeit Komplexe mit weiteren Faktoren zu bilden. Um eine vollkommen native Interaktion zwischen HNRNPK und den zu untersuchenden Oligonukleotiden zu ermöglichen, die solche Fehlerquellen ausschließt, wurde ein EMSA-Experiment mit Zellextrakt, anstelle von aufgereinigtem Protein durchgeführt. In diesem Fall ist es zwar möglich Änderungen der elektrophoretischen Mobilität in einem nativen Proteingel sichtbar zu machen, die Identität der bindenden Faktoren ist aber nicht ohne Weiteres zugänglich.

Die Untersuchung wurde mit Lysat aus unbehandelten HEK293-Zellen durchgeführt und die *Hairpin*-Sequenzen ORN-6 und ORN-6' als Substrate eingesetzt. Oligonukleotide wurden in konstanter Konzentration zugegeben und mit ansteigenden Konzentrationen von unfraktioniertem Zellextrakt versetzt. Nach einer kurzen Inkubationszeit wurden die Proben auf einem nativen Proteingel analysiert (**Abbildung 58**). Die Experimente wurden in Zusammenarbeit mit *Martin Rossa* durchgeführt.

Das sehr ähnliche Bandenmuster für beiden Sequenzen legt nahe, dass sie durch die gleichen Proteine oder Proteinkomplexe gebunden werden. Es bestand aber ein sichtbarer Unterschied in der Lysatmenge, die nötig war um für einen quantitativen *Shift* zu sorgen. Die modifizierte Sequenz ORN-6 wurde reproduzierbar mit höherer Affinität gebunden und änderte bei geringeren Proteinkonzentrationen ihr Migrationsverhalten.

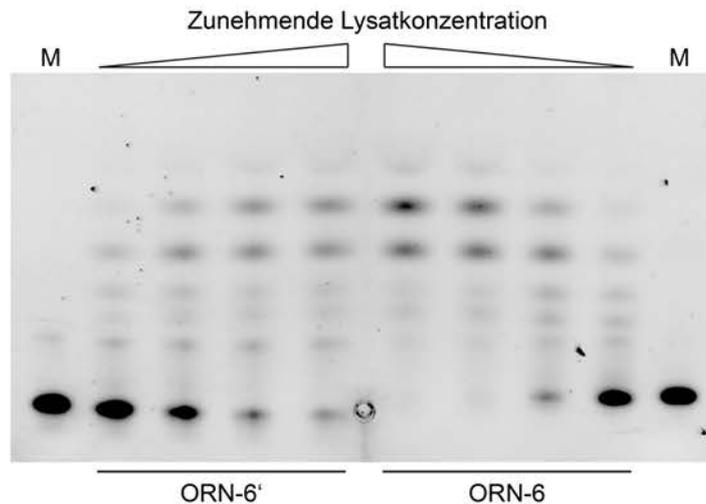


Abbildung 58: EMSA-Experiment mit gesamtem Zellextrakt für *Hairpin*-Sequenzen ORN-6/6'. HEK293-Zellysate wurde in VM-Puffer mit Oligonukleotid versetzt. Die Proben wurden bei 25 °C und 4 °C inkubiert und auf einem nativen PAGE-Gel (4-20%) bei 4 °C analysiert. Als Marker diente jeweils das ungebundene ORN.

Um die Proteine zu identifizieren, durch welche die eingesetzten Oligonukleotide gebunden wurden kamen zwei unterschiedliche Methoden zum Einsatz. Zum einen ist dies möglich durch einen *Supershift-Assay*,⁴⁰⁰ in dem identisch behandelte Proben zusätzlich mit Antikörpern gegen Kandidatenproteine versetzt werden, um sie zusätzlich zu verschieben. Eine neue Bande, die aus einem spezifisch an das Substrat bindenden Protein besteht, würde hierdurch in eine zweite, höher laufende verschoben. Es wurden Antikörper gegen HNRNPK, FAM120C und U2AF2, also zwei ms^2i^6 A-Reader und ein abgestoßenes Protein eingesetzt. In keinem dieser Fälle resultierte ein sichtbarer Unterschied im Bandenmuster durch die Zugabe des Antikörpers (nicht gezeigt).

Eine weitere Möglichkeit zur Identifikation von Bindern aus dem initialen Experiment stellt die Exzision von einzelnen Banden und die darauf folgende massenspektrometrische Analyse dar. Alle in **Abbildung 58** sichtbaren Banden wurden ausgeschnitten und in mehreren Replikaten einer solchen Analyse unterzogen. Durch die Komplexität der aufgetragenen Proben aus dem gesamten Lysat, in Kombination mit der geringen Trennleistung des nativen PAGE-Gels war es nicht möglich mit ausreichender Zuverlässigkeit den beobachteten Banden einzelne Proteine zuzuordnen.

5.6.2.5 Western Blot

Da alle bislang beschriebenen Versuche, die selektive Bindung von HNRNPK an ms^2i^6 A-modifizierte RNA zu bestätigen keinen Erfolg zeigten, wurde der Versuch unternommen den

Versuchsablauf der SILAC-Experimente möglichst unverändert zu replizieren und das Ergebnis mittels *Western Blot*, anstelle einer massenspektrometrischen Analyse, zu visualisieren. Hierfür wurden Oligonukleotide verwendet die mit einem 5'-Biotin-Tag ausgestattet waren und den für die SILAC-*Pulldowns* eingesetzten Sequenzen entsprachen (ORN-2/2'). Sie wurden, analog zu der Vorgehensweise von *Raghu Edupuganti*, zunächst auf Streptavidin-*Beads* geladen, dann mit HeLa-Zellextrakt bei 25 °C und 4 °C inkubiert und nach mehrfachem Waschen alle gebundenen Proteine mit SDS-Probenpuffer gelöst und auf einem SDS-PAGE Gel aufgetrennt. Zielproteine wurden abschließend mittels *Western Blot* mit geeigneten Antikörpern identifiziert. Es wurden Antikörper gegen HNRNPK, FAM120C und U2AF2 verwendet, um sowohl spezifische Binder, als auch ein von ms²i⁶A abgestoßenes Protein analysieren zu können. Die Experimente wurden in Zusammenarbeit mit *Jessica Steinbacher* und *Matthias Kurz* durchgeführt.

Zunächst wurden die Sensitivität und Spezifität der kommerziellen Antikörper überprüft (**Abbildung 59** und **60**). Für HNRNPK zeigte sich eine deutlich reduzierte Abundanz in zytosolischem Lysat im Vergleich zu Kernlysate und eine Detektionsgrenze von unter 100 ng Protein. Der Antikörper zeigte eine einzelne Bande, die ein Molekulargewicht leicht über dem für HNRNPK erwarteten von 51 kDa zeigte.

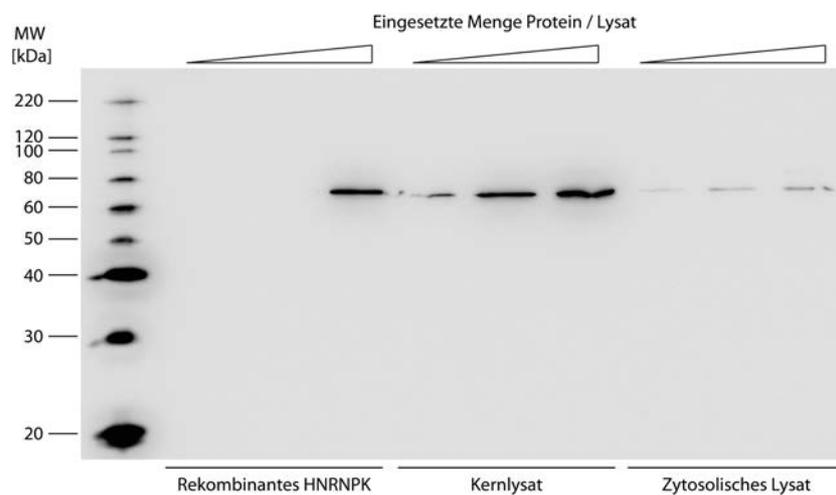


Abbildung 59: *Western Blot* (*Wet Blot*) zur Bestimmung der Sensitivität und Spezifität des eingesetzten HNRNPK-Antikörpers (*Thermo Fisher*, PA5-27522). Marker: MagicMark XP WB Standard (*Life Tech*). Es wurden ansteigende Mengen aus *E. coli* aufgereinigtes HNRNPK_StrepII (1 ng, 20 ng, 100 ng) und fraktioniertes HeLa-Zellysate geladen (je 10 µg, 15 µg und 20 µg).

Für U2AF2 und FAM120C waren die Signale deutlich schwächer in den analysierten Lysaten; aufgereinigte Proteine standen nicht zur Verfügung (**Abbildung 60**). FAM120C zeigte zwei kaum detektierbare Banden bei etwa 120 kDa und 47 kDa. Es existieren zwei

Isoformen mit erwarteten Molekulargewichten von 121 kDa und 25 kDa. U2AF2 zeigte eine schwache Bande bei etwa 62 kDa, das erwartete Molekulargewicht liegt bei 53 kDa.

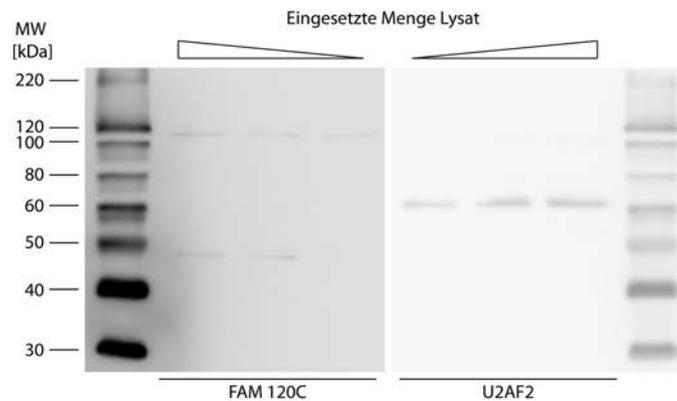


Abbildung 60: *Western Blot (Wet Blot)* zur Bestimmung der Sensitivität und Spezifität der eingesetzten Antikörper gegen FAM120C (*Sigma-Aldrich*, SAB2107574) und U2AF2 (*Sigma-Aldrich*, SAB2701232). Marker: MagicMark XP WB Standard (*Life Tech*). Es wurden ansteigende Mengen HeLa-Gesamtlysat geladen (je 10 µg, 15 µg und 20 µg).

Zunächst wurden *Pulldown*-Experimente mit zytosolischem Lysat und tRNA aus Hefe als Kompetitor (15 eq.) durchgeführt. In diesem Fall konnte für kein Protein ein Signal in den isolierten Proben detektiert werden, weshalb die Menge des eingesetzten Kompetitors reduziert wurde (3 eq.). Darüber hinaus wurde im Weiteren mit unfraktioniertem Zellextrakt gearbeitet, da die erreichbaren Proteinkonzentrationen für diesen deutlich höher waren, was sich als förderlich für die Detektion von Signalen im *Western Blot* erwies.

Hier war ein schwaches Signal für U2AF2 detektierbar (**Abbildung 61**), aber keines für HNRNPK, oder FAM120C. Auch ein abschließendes Experiment ohne Kompetitor lieferte keine sichtbare Bindung für HNRNPK und keinen deutlichen Unterschied für U2AF2 (**Abbildung 61**).

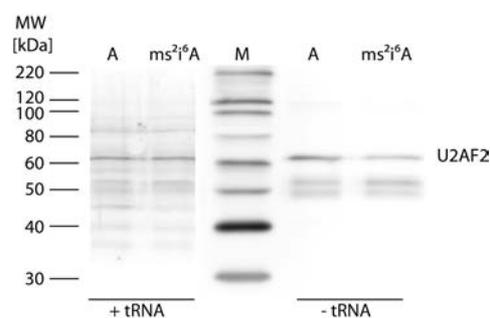


Abbildung 61: *Western Blot (Wet Blot)* zur Analyse zweier *Pulldowns* mit ORN-2 (ms^2i^6A) und ORN-2' (A). Die Experimente wurden mit HeLa-Gesamtlysat in Anwesenheit von 3 eq. tRNA aus Hefe als Kompetitor (links), oder ohne Kompetitor (rechts) durchgeführt. Marker: MagicMark XP WB Standard (*Life Tech*). Antikörper: SAB2701232 (*Sigma-Aldrich*)

Parallel wurden in der Gruppe von *Michiel Vermeulen* analoge Validierungsexperimente durchgeführt, die zu gleichen Ergebnissen kamen.

Es muss also abschließend festgestellt werden, dass eine Validierung der im SILAC-Experiment festgestellten selektiven Bindung für die Proteine HNRNPK, FAM120C und U2AF2 nicht gelang, trotz einer Vielzahl unterschiedlicher Ansätze, um potentielle Fehlerquellen zu eliminieren.

5.6.3 Lokalisierungsstudien zu ms^2i^6A Bindern

Für ms^2i^6A wurde ein komplexes Interaktom in zytosolischem Zelllysat ermittelt (**Abbildung 48**). Einige der identifizierten Proteine sind als kernlokalisiert bekannt, andere vornehmlich im Zellplasma zu finden. Um zu überprüfen, ob sich ihre Lokalisierung im Einklang mit den massenspektrometrischen Experimenten befindet und Vorarbeit für mögliche Co-Lokalisationsstudien zu leisten wurden YBX1, HNRNPU, HNRNPK, SYNCRIP, TOE1, PCBP1 und GTF2B in einen geeigneten Expressionsvektor mit Fluoreszenz-Tag (Cycle 3-GFP) für humane Kulturzellen (pDEST47) kloniert. Die Vektoren wurden in HeLa-Zellen transfiziert und diese nach 24 h und 48 h konfokalmikroskopisch untersucht (**Abbildung 62**).

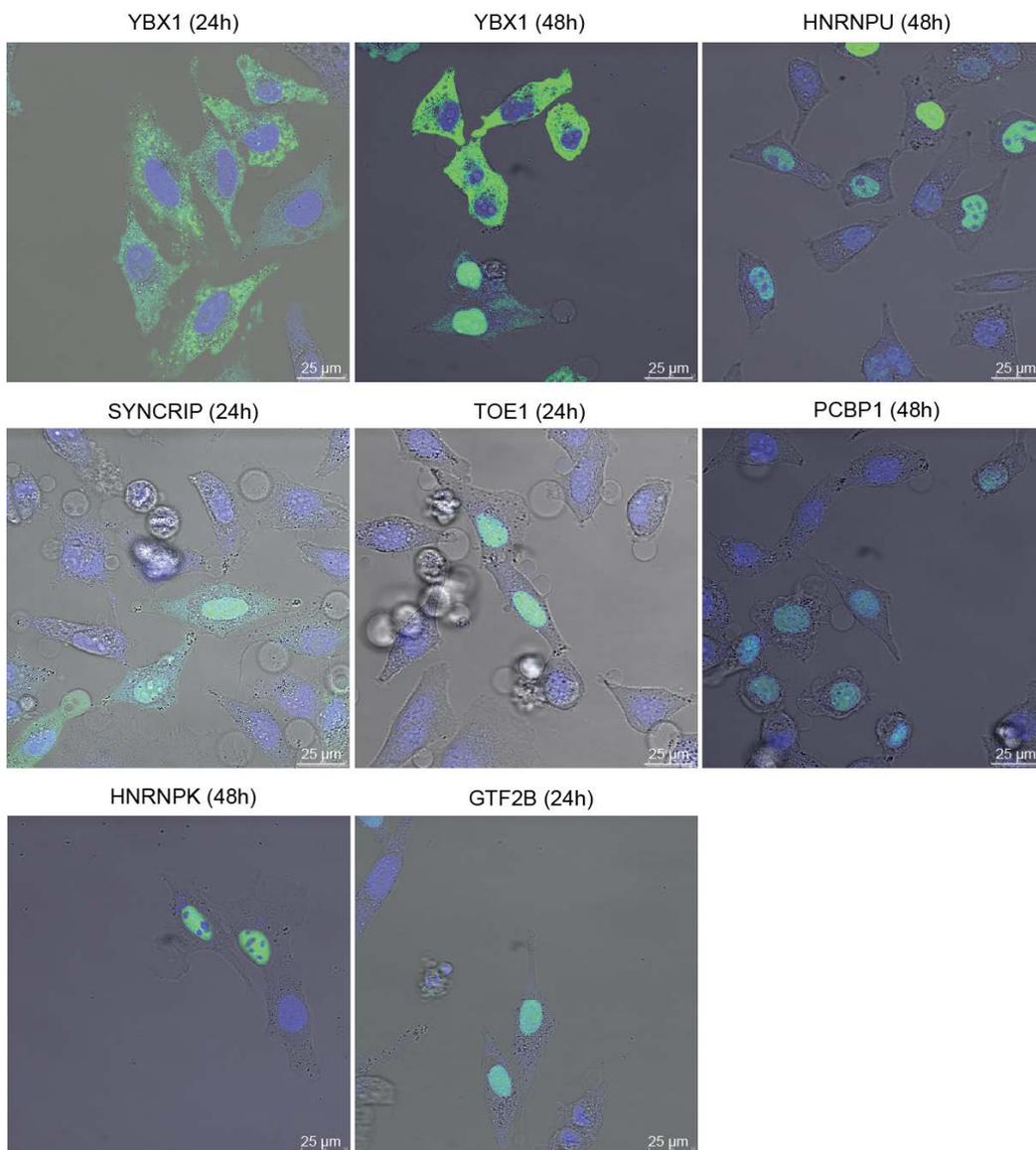


Abbildung 62: Konfokalmikroskopische Lokalisierungsstudien zu ausgewählten ms^2i^6A -Bindern. Die Proteine wurden in einen humanen Expressionsvektor mit GFP-Tag (pDEST47) kloniert und in HeLa-Kulturzellen transfiziert. Die mikroskopische Analyse erfolgte nach 24 h und 48 h.

YBX1 und HNRNPU und HNRNPK sind in ihrer GFP-markierten Form erheblich stärker exprimiert als SYNCRIP, TOE1 oder GTF2B. In ihrer individuellen Verteilung innerhalb der betrachteten HeLa-Zellen zeigt sich kein einheitliches Bild. Da die untersuchten Proteine auf dem verwendeten Vektor unter der Kontrolle eines artifiziellen Promoters (CMV) stehen, ist hieraus aber keine Aussage zu ihrer natürlichen Abundanz abzuleiten. Mit YBX1 ist ein einziges Protein deutlich zytosolisch lokalisiert. Es findet sich in kleinen, definierten Partikeln außerhalb des Zellkerns, mit einer kleinen Subpopulation an Zellen, die eine deutliche Kernlokalisierung aufweist (nach 48 h). Es ist bekannt, dass es Funktionen in beiden Kompartimenten ausübt,^{401,402} also in der Lage sein muss sich zwischen Kern und Zellplasma hin- und herzubewegen. Auch ein passiver Transport ist möglich. SYNCRIP (auch HNRNPQ) ist in beiden Teilen der Zelle zu beobachten, wobei die Lokalisation im Kern zu überwiegen scheint.

Die schwach exprimierten Proteine TOE1, PCBP1 und GTF2B sind eindeutig im Kern vorzufinden, ohne eine definierte Exklusion eines Unterbereiches aufzuweisen. Die verbleibenden HNRNPK und HNRNPU sind ebenfalls deutlich kernlokalisiert, sind dabei aber nicht gleichmäßig verteilt, sondern besitzen eine deutlich geringere Intensität in Nucleoli.

Da die untersuchten Proteine alle zu den aus zytosolischem Lysat angereicherten Proteinen gehören, war in diesem Experiment keine Kernlokalisierung zu erwarten. Dennoch ist sie für HNRNPK, HNRNPU, TOE1, GTF2B und PCBP1 zu beobachten. Proteine sind nicht zwingend ausschließlich in einem Zellkompartiment zu finden und insbesondere prä-mRNA-bindende Faktoren können häufig zwischen Kern und Zytosol *shutteln*. Insbesondere für HNRNP-Proteine^{359,378}, darunter HNRNPK³⁹⁹ und darüber hinaus das U2AF-Heterodimer⁴⁰³ wurde dies auch bereits gezeigt. Ihre Lokalisation kann vom Zustand der betrachteten Zelle und externen Stimuli abhängen. Dennoch scheint für einen Teil der untersuchten Proteine die Kernlokalisierung dominant zu sein.

5.6.4 Charakterisierung der Cdk5Rap1-Knockout Zelllinie

Um mehr über die zellübergreifende biologische Rolle von m^2i^6A in Erfahrung zu bringen, wurde eine humane Knockout-Zelllinie charakterisiert. Die Zellen wurden von *Horizon Discovery* bezogen und der Knockout durch *Angie Kirchner* validiert. Sie sind abgeleitet von der annähernd haploiden Linie HAP1, die wiederum aus KBM-7 Zellen (chronische

myeloische Leukämie) generiert wurde.⁴⁰⁴ Sie sind adhärent und besitzen eine Fibroblasten-ähnliche Morphologie.

Eine Analyse der Gesamt-RNA der KO-Zelllinie bestätigte ihre Unfähigkeit ms^2i^6A biosynthetisch herzustellen. Ihr Vorläufer i^6A dagegen ist in deutlich größeren Mengen vorhanden, akkumuliert also in gewissem Maß (**Abbildung 63 A**, durchgeführt durch *Katharina Iwan* und *Angie Kirchner*). Die Zellen sind signifikant kleiner (34%) als der Wildtyp und eine Färbung der Mitochondrien zeigt keine eindeutige Änderung der ihrer Morphologie (**Abbildung 63 B und C**).

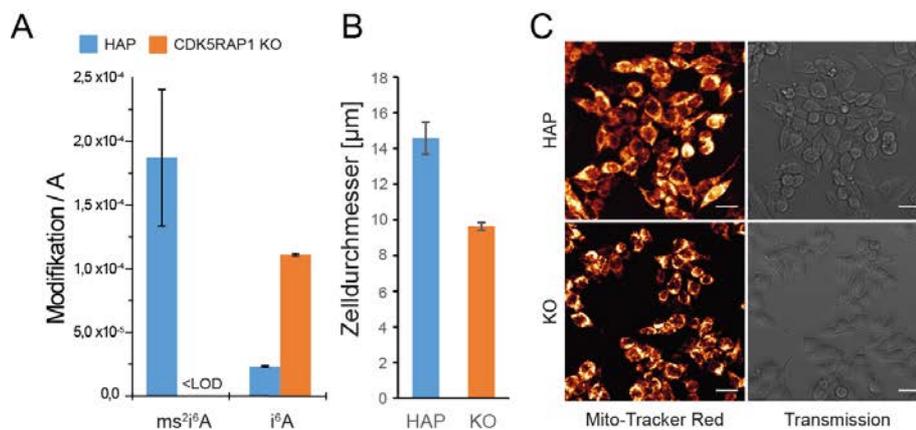


Abbildung 63: Charakterisierung der Cdk5Rap1-Knockout-Zelllinie. A) Quantifizierung des Gehaltes an ms^2i^6A und i^6A in beiden Zelllinien (LOD = *Limit of detection*; Detektionslimit). B) Messung des Durchmessers beider Zelllinien. C) Mikroskopische Aufnahmen der Zellen: Transmission (rechts) und Färbung der Mitochondrien mittels *MitoTracker Red* (links). Maßstab: 20 μ m. (Daten/Abbildung: *Angie Kirchner* und *Katharina Iwan*).

Zur weiteren Charakterisierung wurden das Transkriptom und das Gesamtproteom beider Zelllinien miteinander verglichen. Zunächst wurde durch RNA-Sequenzierung im Transkriptom nach signifikanten Unterschieden gesucht. Die Experimente und Analysen wurden durch *Olga Kotljarova* und *Markus Müller* durchgeführt. Dabei zeigte sich eine signifikant verringerte Anzahl mitochondrialer *Reads*, was als Hinweis auf einen Defekt in der mitochondrialen Transkription gedeutet werden kann (**Abbildung 64 A**).

Eine Analyse nach an- bzw. abgereicherten funktionellen Gruppen (DAVID, Version 6.7)^{342,343} unter den identifizierten Transkripten zeigt wenig ausgeprägte Überexpressionsmuster in KO-Zellen, dagegen aber mehrere stark in Wildtypzellen angereicherte funktionelle Gruppen. Kernlokalisierte Proteine, in mRNA-Splicing, Chromatinorganisation, Prozessierung von rRNA und Transkription involvierte Faktoren sind in Wildtypzellen angereichert, bzw. werden durch die Abwesenheit von ms^2i^6A in ihrer Abundanz reduziert

(Abbildung 64 B). Die Kategorie *non-membrane organelle* (nicht von einer Membran umgebene Organellen) schließt Ribosomen, das Zytoskelett und Chromosomen ein. Nukleare und mit Splicing assoziierte Proteine sind auch unter den ms^2i^6A -interagierenden Proteinen zu finden (Abbildung 48).

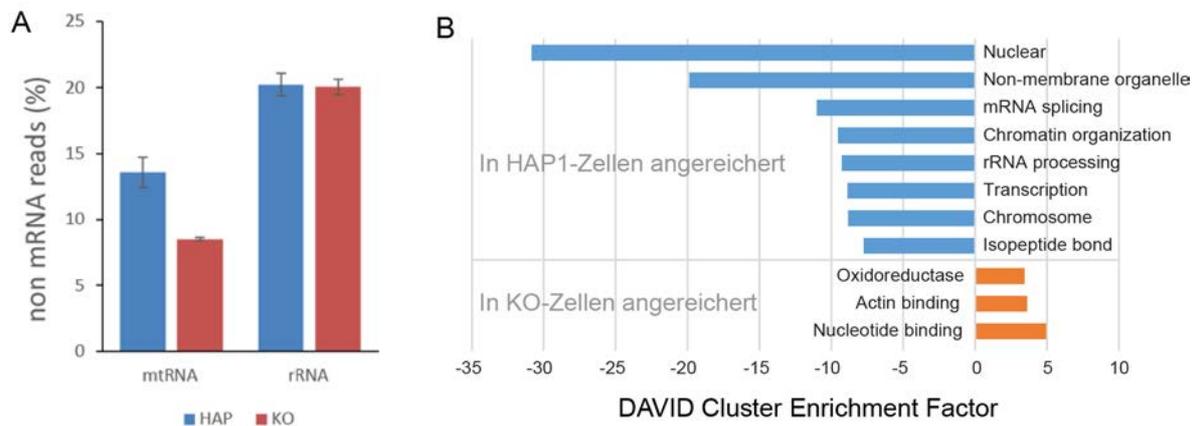


Abbildung 64: Transkriptomanalyse der HAP1-Wildtyp und Cdk5Rap1-Knockout Zelllinien. A) Quantifizierung der nicht mRNA zuzuordnenden Reads für mitochondriale und ribosomale RNA. B) Gene Ontology Analyse (DAVID, Version 6.7) der signifikant an- und abgereicherten Transkripte beider Zelllinien (A: Analyse und Abbildung: Markus Müller).

Die Analyse des Transkriptoms wurde durch eine Proteomanalyse ergänzt, da zum einen der beobachtete Phänotyp bestätigt werden sollte und zum anderen eine Modulation auf RNA-Ebene nicht notwendigerweise eine identische Veränderung auf der Proteomebene zur Folge hat, eine weitergehende Analyse den Phänotyp also präzisieren kann. Um eine möglichst hohe analytische Abdeckung des Proteoms zu erreichen, wurden Zellen einer fraktionierten Lyse unterzogen (Zytosol, Nukleus und Mitochondrien). Die so gewonnenen Fraktionen wurden mit Hilfe von *StageTips*⁴⁰⁵ aufgereinigt und wiederum fraktioniert eluiert. Jede der gewonnenen Proben wurde separat massenspektrometrisch analysiert und die gewonnenen Daten zu einem Datensatz zusammengefügt. Jeder dieser Datensätze resultierte aus der Kombination von sechs einzelnen Messungen. Die Analyse wurde in biologischen Triplikaten durchgeführt, in Zusammenarbeit mit *Jessica Steinbacher* und *Michael Stadlmeier*. Die Zellen wurden von *Angie Kirchner* zur Verfügung gestellt.

Die mehrfache Fraktionierung ermöglichte eine gute Abdeckung von 4940 Proteinen, nach der Anwendung aller Qualitätsfilter. Die Daten zeigen eine gute Übereinstimmung mit den Sequenzdaten, wie **Abbildung 65** zeigt. Die große Mehrzahl der in beiden Analysen identifizierten Proteine, bzw. Transkripte zeigt ein gleichbleibendes Verhalten (blau/orange).

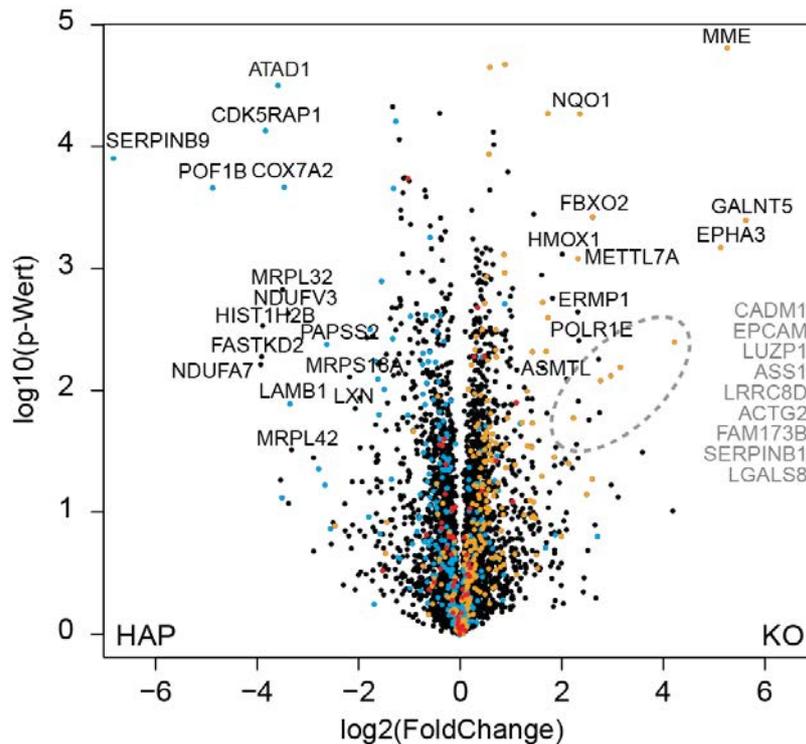


Abbildung 65: Proteomanalyse der HAP1-Wildtyp und Cdk5Rap1-Knockoutzellen. Signifikante Proteine sind mit Namen bezeichnet (Signifikanzgrenzen: FoldChange > ±2; p < 0.05). Proteine sind farblich hervorgehoben, wenn sie in der Transkriptomanalyse im Wildtyp überrepräsentiert (blau), bzw. unterrepräsentiert (orange) sind. Proteine, die im SILAC Interaktom eines ms²-A-modifizierten *Baits* signifikant an- oder abgereichert wurden, sind rot markiert (Abbildung: Markus Müller).

Eine *Gene Ontology* Analyse der Proteomdaten zeigt, dass mitochondriale Prozesse auf dieser Ebene stark betroffen sind. Proteine der mitochondrialen Translation und Faktoren der Atmungskette sind relativ im Wildtyp überrepräsentiert (**Abbildung 66 A**). Dieser Phänotyp deckt sich mit den Beobachtungen von *Wei et al.*, die in murinen Kulturzellen zeigen konnten, dass ein Cdk5Rap1-Knockout zu einer gestörten mitochondrialen Translation führt, von dem insbesondere Faktoren der Atmungskette betroffen sind.¹⁹⁸ Sie führen diese Beobachtung auf hypomodifizierte mt-tRNAs zurück und schließen eine zytosolische Funktion der Thiomethyltransferase aus.

Neben dem ausgeprägten mitochondrialen Phänotyp sind auch ribosomale Proteine massiv im *Knockout* abgereichert, was als Hinweis auf einen globalen Transkriptionsdefekt gedeutet werden kann. Dieser scheint sich stärker auf mitochondriale Transkripte als rRNA zu auszuwirken (**Abbildung 64 A**), betrifft aber insgesamt stärker andere funktionelle Gruppen (**Abbildung 64 B**). Wie bereits in der Untersuchung des Transkriptoms sind im *Knockout* deutlich weniger definierte Gruppen von Proteinen überrepräsentiert als im Wildtyp.

Die beschriebenen funktionellen Klassifizierungen mit Hilfe von DAVID erlauben eine schnelle Gruppierung der identifizierten Proteine nach zellulären Funktionen, Lokalisation in

Zellkompartimenten und anderen Kriterien, beziehen dabei aber ihre relative An- oder Abreicherung (**Abbildung 48: Fold Change**) nicht mit ein. Um eine entsprechende Analyse unter Berücksichtigung dieser Information durchzuführen, wurden alle zuverlässig quantifizierten Proteine ($p < 0.05$) aus der Proteomanalyse nach funktioneller Zuordnung gruppiert (*Perseus*, Version 1.5.0.9) und mit Hilfe der relativen Verschiebung Schwerpunkte der Subpopulationen bestimmt (**Abbildung 66 B**).

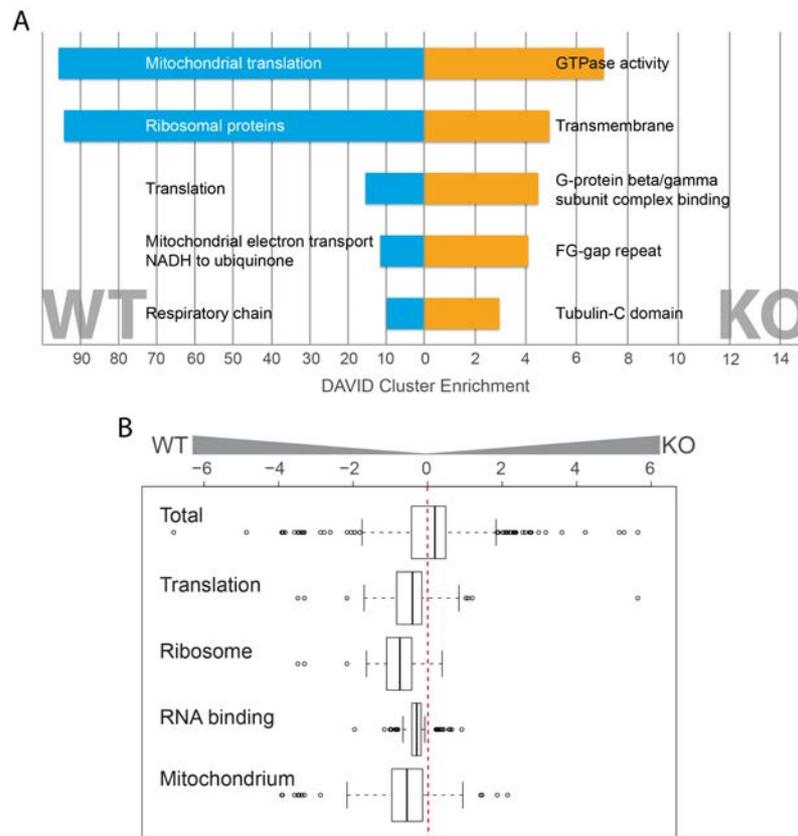


Abbildung 66: Analyse der proteomweiten massenspektrometrischen Quantifizierung. A) DAVID *Gene Ontology* Analyse (Version 6.8) der signifikant in Wildtyp (blau), bzw. *Knockout* überrepräsentierten Proteine (Signifikanzgrenzen: $FoldChange > \pm 0.25$; $p < 0.05$). Das *Functional Clustering* wurde mit der Einstellung *High Stringency* durchgeführt. B) Vergleich der Schwerpunkte funktioneller Subpopulationen. Zuordnung (von oben nach unten): Attribut GOBP enthält „translation“; Attribut Corum enthält „ribosom“; Attribut GOMF enthält „RNA-Binding“; Attribut GOCC enthält „mitoch“. Die Gruppen wurden einem Welch-Test unterzogen und signifikant abweichende Gruppen mit Hilfe von R (Rstudio; Version 0.99.489) grafisch dargestellt. (Statistische Analyse für B): *Andrea Künzel*

Wie erwartet sind mitochondriale Proteine und der Translation zugeordnete im *Knockout* abgereichert, darüber hinaus auch RNA-bindende Proteine. Mit ribosomaler Funktion verbundene Proteine sind aber unter Einbeziehung quantitativer Information die am deutlichsten von einem *Cdk5Rap1-Knockout* betroffene Subpopulation (grafisch dargestellt in **Abbildung 67**). Die gesamte Population aller zellulären Proteine ist dabei nicht signifikant

verschoben, es handelt sich also um Änderungen innerhalb der beschriebenen Untergruppen, welche sich im Gegensatz zum allgemeinen Trend befinden.

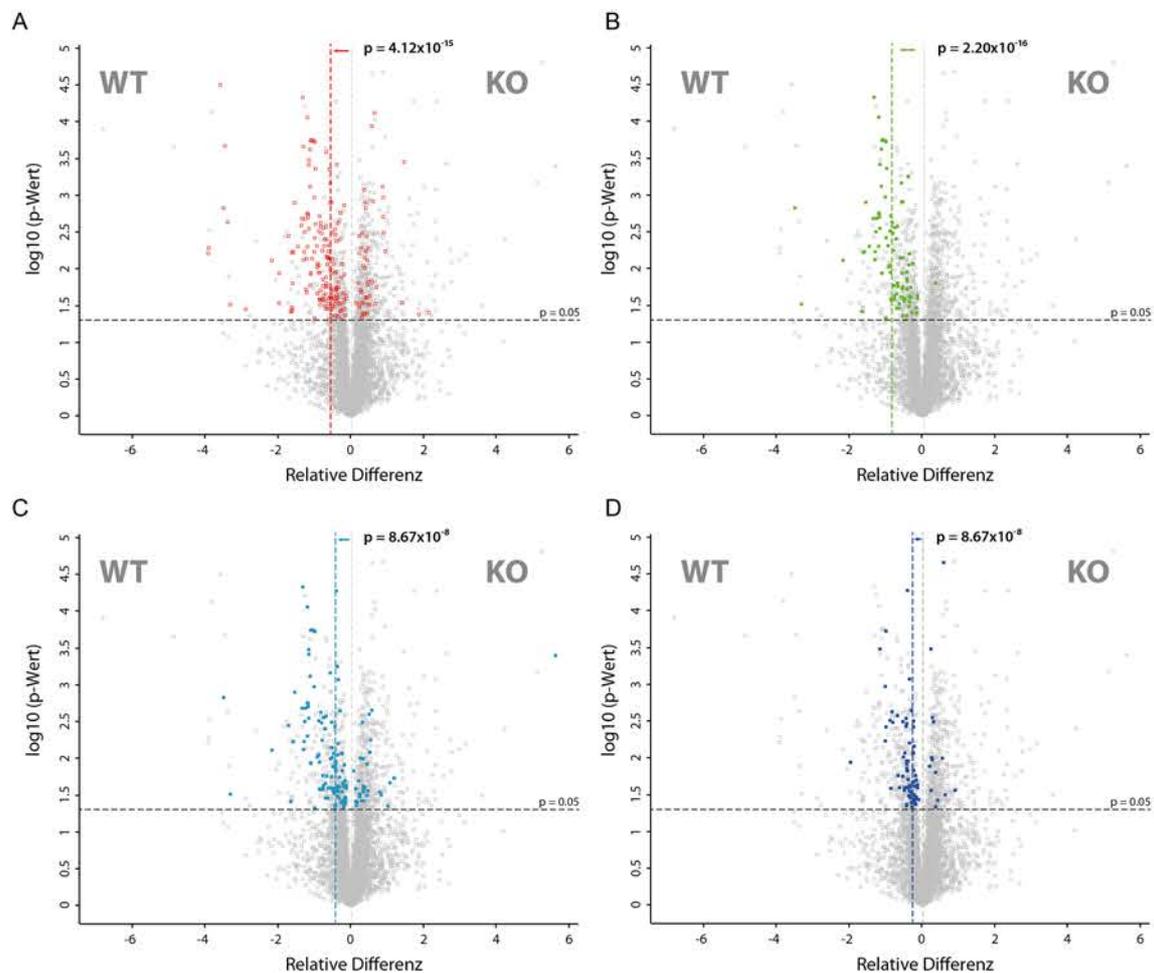


Abbildung 67: Grafische Darstellung der funktionellen Subpopulationen innerhalb der Gesamtproteom-Daten. Untergruppen wurden wie oben beschrieben (**Abbildung 59**) identifiziert und farblich hervorgehoben. A) Rot: Mit Mitochondrien assoziierte Proteine. B) Grün: Mit Ribosomen assoziierte Proteine. C) Türkis: Mit Translation assoziierte Proteine. D) Blau: RNA-bindende Proteine.

Da viele der mit m^2i^6A interagierenden Proteine in (alternatives) Splicing involviert sind, wurde innerhalb der Transkriptomdaten nach einem direkten Nachweis für veränderte Splicingmuster gesucht. Die Analyse wurde von *Markus Müller* durchgeführt. Es wurden 92 differenziell gesplicete Gene identifiziert, von denen 52 bekannte Splicing-Isoformen besitzen. Zwei von ihnen wurden mittels qPCR validiert. LARP1 (*La-related protein1*) ist ein RNA-bindendes Protein, das regulativ in die Translation eingreift indem es den 3'-Terminus von poly(A)-RNA bindet.⁴⁰⁶⁻⁴⁰⁸ Es beeinflusst vor allem Transkripte, die ribosomale und andere in die Translation involvierte Proteine codieren, also Gruppen, die durch einen *Cdk5Rap1-Knockout* erheblich in ihrer Abundanz reduziert werden.^{409,410} Seine mRNA ist in einer Variante mit einem verkürzten Exon im *Knockout* abgereichert (**Abbildung 68 A**).

Zusätzlich wurde TERF1 (*Telomeric repeat-binding factor 1*) bestätigt, das zur Stabilisierung humaner Telomere beiträgt.⁴¹¹ Seine Abundanz wird durch U2AF2 reguliert, das zu den von ms^2i^6A abgestoßenen Proteinen gehört.⁴¹² Hier wird im KO ein zusätzliches Exon verstärkt mit eingeschlossen (**Abbildung 68 D**).

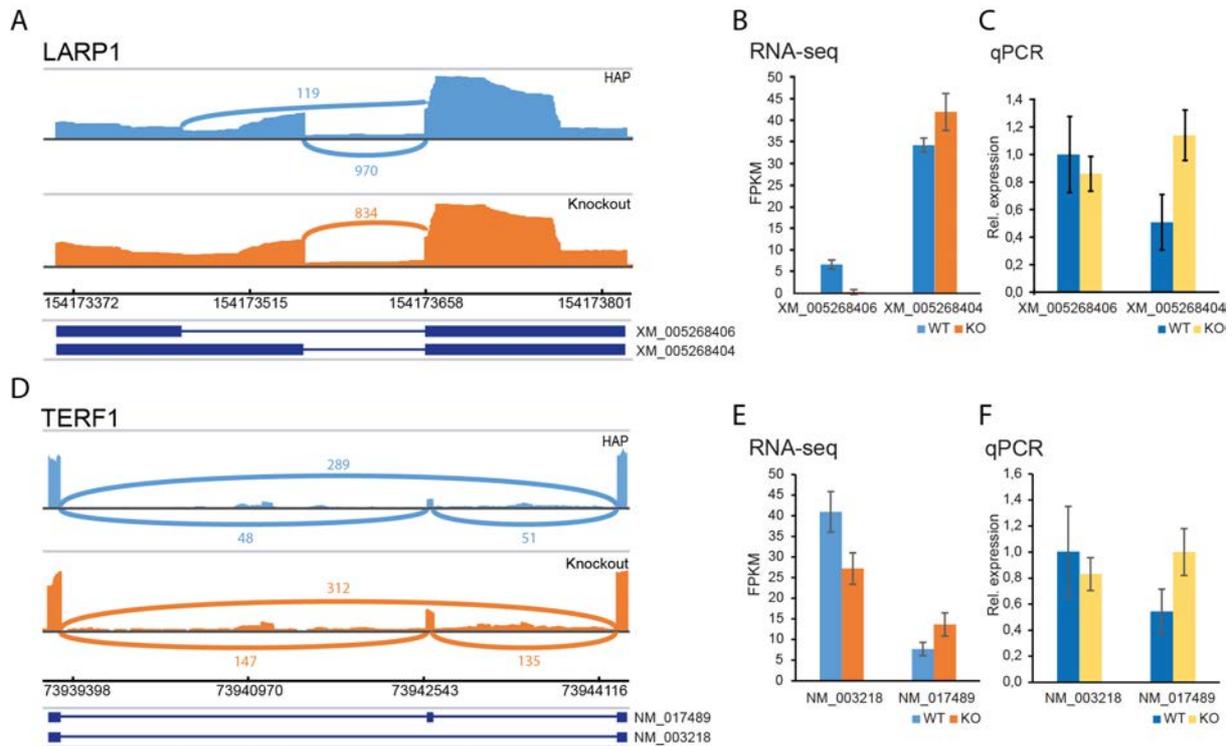


Abbildung 68: Validierung der variablen Splicingmuster in HAP1-Wildtyp- und Cdk5Rap1-Knockout-Zelllinien. A) Vergleich der Splicingmuster für LARP1. B) Verhältnis der Isoformen von LARP1 in beiden Zelllinien, anhand der Sequenzierungsdaten. C) Validierung der LARP1-Isoformen mittels qPCR. A) Vergleich der Splicingmuster für TERF1. B) Verhältnis der Isoformen von TERF1 in beiden Zelllinien, anhand der Sequenzierungsdaten. C) Validierung der TERF1-Isoformen mittels qPCR. (Analyse und Abbildung: Markus Müller; Validierung: Angie Kirchner)

Bei einer Regulation alternativen Splicings durch ms^2i^6A ist eine Retention der Modifikation in codierenden Transkripten in Abhängigkeit von einer Prozessierung durch das Spliceosom möglich. Um dieser Hypothese nachzugehen wurden durch Katharina Iwan und Angie Kirchner Quantifizierungsstudien mit dem Splicing-Inhibitor Isoginkgetin durchgeführt. Dieser führt zur Akkumulation von präspliceosomalen A-Komplexen, indem er die Rekrutierung des U4/U5/U6 snRNP hemmt (**Abbildung 69 A**).⁴¹³ HAP1-Zellen wurden mit dem Inhibitor behandelt, ihre poly(A)-RNA isoliert und mit der aus unbehandelten Kontrollzellen verglichen. Die Analyse zeigte, dass die behandelten Zellen prä-mRNA akkumulieren und parallel der Gehalt an i^6A und ms^2i^6A in isolierter poly(A)-RNA zunimmt (**Abbildung 69 B und C**). Die Zunahme ist gering, aber signifikant (i^6A : $p = 0.0095$; ms^2i^6A : $p = 0.022$).

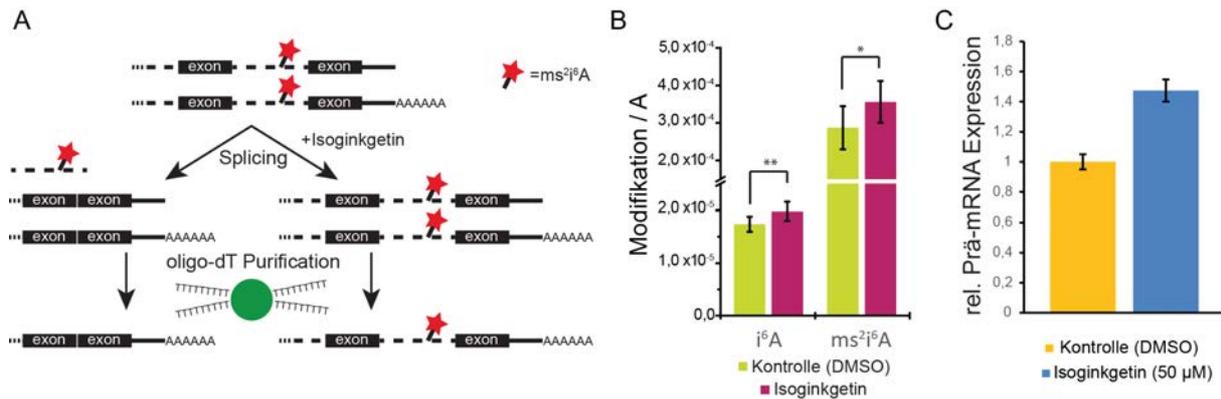


Abbildung 69: Beeinflussung des Splicings durch Isoginkgetin und damit zusammenhängende Veränderung des Gehaltes an i⁶A und ms²i⁶A in poly(A)-RNA. A) Prinzip des Experimentes. Der Inhibitor Isoginkgetin unterdrückt Splicing und führt damit zu einer Akkumulation ungespliceter prä-mRNA. Modifikationen, die sich innerhalb von Introns befinden und üblicherweise im Zuge des Splicing ausgeschnitten werden, werden ebenfalls angereichert. B) Massenspektrometrische Quantifizierung von i⁶A und ms²i⁶A in poly(A)-RNA aus Zellen, die mit Isoginkgetin behandelt wurden und Kontrollzellen, die mit DMSO behandelt wurden. C) Quantifizierung des Verhältnisses zwischen prä-mRNA und prozessierter mRNA durch qPCR. (Experimente: Katharina Iwan und Angie Kirchner; Abbildung: Markus Müller)

Die Analyse legt also nahe, dass mit der Akkumulation ungespliceter Introns auch ms²i⁶A und sein biosynthetischer Vorläufer i⁶A in poly(A)-RNA angereichert werden. Die Zunahme ist geringer als der globale Anstieg unprozessierter prä-mRNA. Da nicht davon auszugehen ist, dass jedes ungesplicete Intron oder ausgelassene Exon ms²i⁶A enthält, ist dieses Verhältnis nachvollziehbar.

5.6.5 Addendum zum ms^2i^6A Interaktom

Nach dem Verfassen dieser Arbeit wurden durch die Gruppe von *Michiel Vermeulen* die initialen *Pulldown*-Experimente wiederholt, da sich entscheidende Interaktionen nicht validieren ließen. Hierzu wurden erneut Oligonukleotide mit *Hairpin*-Sequenz (ORN-1/1') und mittels Click verknüpftem Biotin hergestellt und als *Baits* eingesetzt. Die Isotopenmarkierung erfolgte mittels *Dimethyl-labeling*³⁹² (**Abbildung 70**)

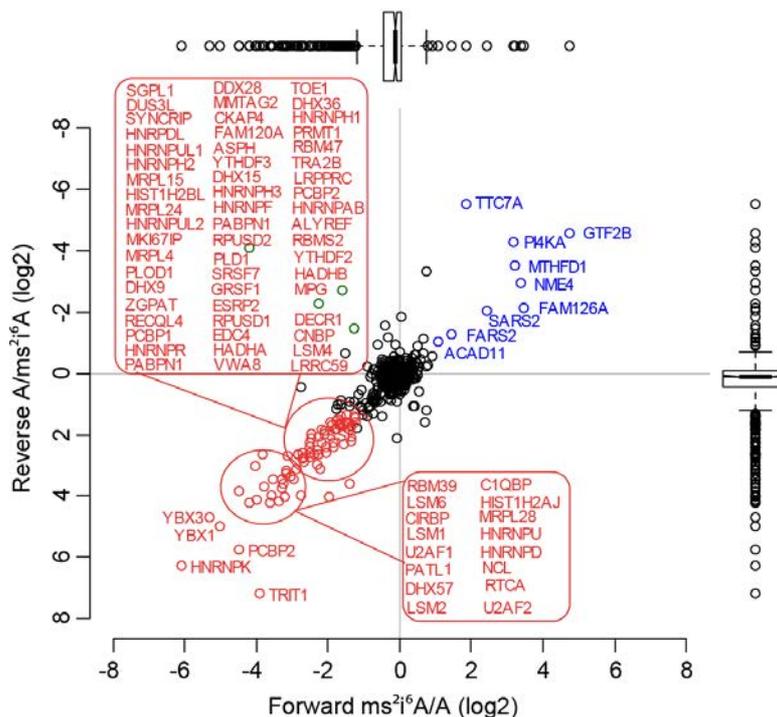


Abbildung 70: Scatterplot des *Pulldowns* aus zytoplasmischem HeLa-Extrakt für ORN-1/1': 5'-XGGGGAUUG AA ms^2i^6A AUCCCC-3'. Das Experiment wurde von *Raghu Edupuganti* durchgeführt. Die Isotopenmarkierung zur quantitativen Analyse erfolgte mittels *Dimethyl labeling*.³⁹²

Wie aus der Abbildung ersichtlich ist, geht aus dieser neuerlichen Analyse ein erheblich anderes Interaktom hervor als aus den vorherigen Analysen. Während die wenigen hier identifizierten Interaktoren alle in mindestens einer der vorherigen Studien bereits identifiziert wurden, ist HNRNPK in dieser Analyse das am stärksten abgereicherte Protein, was eine Erklärung für die in zahlreichen Validierungsexperimenten nicht feststellbare selektive Bindung an ms^2i^6A darstellt. Darüber hinaus sind mit YBX1, PCBP2, SYNCRIP und TOE1 weitere Proteine unter den abgestoßenen zu finden, die zuvor als Binder gefunden wurden. Dies erscheint überraschend, da andere *Baits* (siehe auch Kapitel 5.2) im Allgemeinen gut reproduzierbare und für FMR1 und G3BP1 auch validierte Ergebnisse lieferten. Es erscheint

an dieser Stelle sinnvoll, lediglich die reproduzierbaren *Reader* zu betrachten (**Abbildung 71**).

Aus dieser Gruppe wurden FAM126A, PI4KA, TTC7A, FARS2 und MTHFD1 ebenfalls in LFQ *Pulldowns* unter Verwendung von ORN-4/4'' identifiziert, es ist also davon auszugehen, dass es sich bei ihnen um Interaktoren handelt, die nicht abhängig von der *Hairpin*-Sekundärstruktur binden. GTF2B, welches reproduzierbar die *Hairpins* bindet, konnte nicht als *Reader* der linearen Sequenzen detektiert werden. Eine GO-Analyse der verbleibenden Interaktoren aus **Abbildung 70** zeigt keine angereicherte funktionelle Gruppe, mit FARS2 und SARS2 sind lediglich zwei tRNA Ligasen identifiziert. Ihre Substrate sind mit mt-tRNA^{Phe} und mt-tRNA^{Ser} mitochondriale tRNAs, die ms²i⁶A als Modifikation tragen, das Adenosinderivat könnte also der Erkennung dienen.

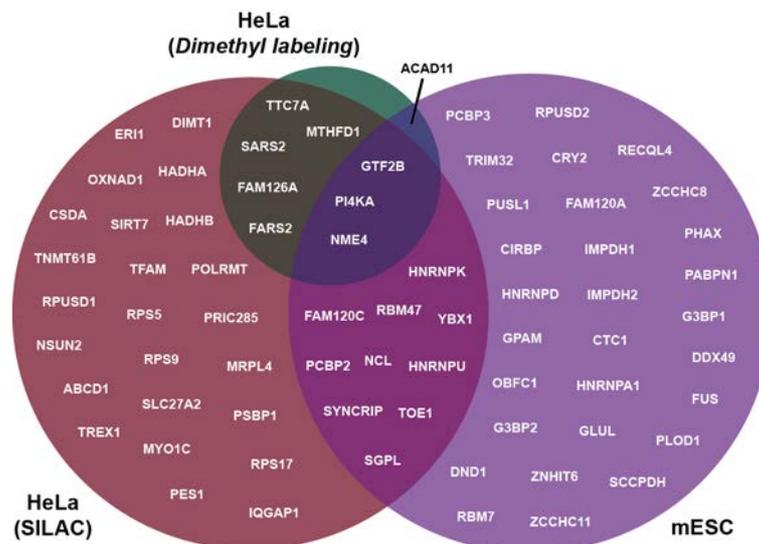


Abbildung 71: Visualisierung des Überlappungsgrades zwischen den *Reader*-Proteinen für das ORN-Paar 1/1' in drei unterschiedlichen Experimenten (siehe auch **Abbildung 48** und **70**).

Mit U2AF1/2 sind auch zwei Splicingfaktoren wiederholt abgestoßen, da sich die selektive Bindung an den unmodifizierten Strang bislang nicht validieren ließ, erscheint eine weitere Interpretation an dieser Stelle nicht sinnvoll.

6. Experimenteller Teil

6.1 Biochemische und mikrobiologische Methoden

6.1.1 Material und allgemeine Methoden

Alle Puffer, Medien und Lösungen wurden, wenn nicht anders angegeben in bidest. Wasser hergestellt. Chemikalien wurden von den Anbietern *Sigma-Aldrich*, *Merck*, *AppliChem* in den üblichen Qualitätsstufen *puriss*, *p.a.*, *purum* und *molecular biology grade* bezogen. Zur Herstellung von bidest. Wasser wurde eine *arium Pro DI* Anlage der Firma *Sartorius Stedim* verwendet.

6.1.1.1 Medien, Lösungen und Puffer

Tabelle 8: In dieser Dissertation verwendete Puffer und Lösungen.

Bezeichnung	Zusammensetzung
DMEM Medium	500 mL DMEM - <i>high glucose</i> (4500 mg/L Glukose) Fertigmedium (<i>Sigma Aldrich</i>), 56 mL (= 10%) FBS (<i>Gibco</i>), 10.25 mL (= 4 mM) L-Alanyl-L-Glutamin (200 mM, <i>Sigma- Aldrich</i>), 5.6 mL (= 1% v/v) Antibiotic-Antimycotic (<i>Sigma- Aldrich</i>) oder 5.6 mL (= 1% v/v) Penicillin-Streptomycin (<i>Sigma Aldrich</i>)
RPMI Medium	500 mL RPMI-1640 Fertigmedium (<i>Sigma-Aldrich</i>), 56 mL (= 10%) FBS (<i>Gibco</i>), 10.25 mL (= 4 mM) L-Alanyl-L-Glutamin (200 mM, <i>Sigma-Aldrich</i>), 5.6 mL (= 1% v/v) Antibiotic- Antimycotic (<i>Sigma-Aldrich</i>) oder 5.6 mL (= 1% v/v) Penicillin- Streptomycin (<i>Sigma Aldrich</i>)
IMDM	500 mL IMDM-Fertigmedium (<i>Sigma-Aldrich</i>), , 56 mL (= 10%) FBS (<i>Gibco</i>), 10.25 mL (= 4 mM) L-Alanyl-L-Glutamin (200 mM, <i>Sigma-Aldrich</i>), 5.6 mL (= 1% v/v) Penicillin-Streptomycin (<i>Sigma Aldrich</i>)
DPBS	Dulbecco's PBS (<i>Sigma Aldrich</i>)
Trypsin-EDTA	Trypsin-EDTA Fertiglösung (<i>Sigma-Aldrich</i>), 0.25% (w/v) Trypsin, 0.2 g/L Na ₄ ·EDTA
LB Medium	1.0% (w/v) Pepton, 0.5% (w/v) Hefeextrakt, 1.0 % (w/v) NaCl; pH = 7.5; nach dem Autoklavieren 0.1% Antibiotika-Stocklösung

LB Agar	1.5% (w/v) Agar-Agar in LB-Medium
SOC Medium	0.5% (w/v) Pepton, 0.5% (w/v) Hefeextrakt, 10 mM NaCl, 2.5 mM KCl; nach dem Autoklavieren 5.0 mM MgCl ₂ , 5.0 mM MgSO ₄
EMSA-Lysispuffer	10 mM Tris·HCl (pH 7.5), 150 mM KCl, 0.5 mM EDTA, 0.5% (v/v) NP-40, 1 × Roche Complete Protease Inhibitor (ohne EDTA), 1% (v/v) Phosphatase Inhibitor Cocktail 2 und 3 (<i>Sigma Aldrich</i>)
EMSA-A Puffer	10 mM HEPES (pH 8.0), 50 mM NaCl, 1 mM EDTA, 5% (v/v) Glycerol
VM-Puffer	10 mM Tris·HCl (pH 7.5), 150 mM KCl, 1.5 mM MgCl ₂ , 0.1% (v/v) NP-40, 0.5 mM DTT, 1 × Roche Complete Protease Inhibitor (ohne EDTA), 1% (v/v) Phosphatase Inhibitor Cocktail 2 und 3 (<i>Sigma Aldrich</i>), 0.4 u/μLRNasin (nur in Inkubationsschritten)
Pulldown-Lysispuffer	10 mM Tris·HCl (pH 7.5), 150 mM NaCl, 0.5 mM EDTA, 0.5% (v/v) NP-40, 1 × Roche Complete Protease Inhibitor (ohne EDTA), 1% (v/v) Phosphatase Inhibitor Cocktail 2 und 3 (<i>Sigma Aldrich</i>)
Dilution Buffer	10 mM Tris·HCl (pH 7.5), 150 mM NaCl, 0.5 mM EDTA
Proteinase K Puffer	10 mM Tris·HCl (pH 8.0), 1.0 mM EDTA, 0.5% (w/v) SDS
Proteinase K Stock	50% (v/v) Glycerin, 20 mM Tris·HCl (pH 8.0), 1 mM CaCl ₂ , 2 mg/mL Proteinase K (<i>GeneON</i>)
10 × TBE	890 mM Tris·HCl (pH 8.0), 890 mM NaH ₂ BO ₃ , 20 mM EDTA
2 × RNA Ladepuffer	2 × TBE (178 mM Tris·HCl (pH 8.0), 178 mM NaH ₂ BO ₃ , 4 mM EDTA), 12% (w/v) Ficoll 400, 7 M Harnstoff

6.1.1.2 Antibiotika

Antibiotika wurden als 1000 × Stocklösungen vorbereitet und bei -20 °C gelagert.

Carbenicillin:

100 mg/mL in H₂O

Chloramphenicol:

34 mg/mL in EtOH_{abs}

Kanamycin:

50 mg/mL in H₂O

Spectinomycin:

100 mg/ml in H₂O

6.1.1.3 Antikörper

Tabelle 9: In dieser Dissertation verwendete Antikörper und ihre Hersteller.

Antikörper	Hersteller	Bestellnummer
Anti-Rabbit IgG Peroxidase	Sigma-Aldrich	A6667
hnRNP K, polyklonal	Thermo Fisher	PA5-27522
U2AF2	Sigma-Aldrich	SAB2701232
FAM120C	Sigma-Aldrich	SAB2107574

6.1.1.4 Plasmide

Plasmide mit Inserts zur Expression von HNRNPK, YBX1, SYNCRIP, GTF2B, HNRNPU, PCBP1 und TOE1 wurden von *DNASU* (Tempe, AZ) bezogen und sind in der nachfolgenden Tabelle aufgelistet.

Tabelle 10: In dieser Dissertation verwendete Plasmide.

Insert	Produktnummer	Vektor	Resistenz	Quelle
HNRNPK	HsCD00622816	pENTR223	Spectinomycin	DNASU, Tempe, AZ
YBX1	HsCD00040576	pDONR221	Kanamycin	DNASU, Tempe, AZ
SYNCRIP	HsCD00513121	pENTR223	Spectinomycin	DNASU, Tempe, AZ
GTF2B	HsCD00509797	pENTR223	Spectinomycin	DNASU, Tempe, AZ
PCBP1	HsCD00288587	pENTR223	Spectinomycin	DNASU, Tempe, AZ
TOE1	HsCD00399441	pDONR223	Spectinomycin	DNASU, Tempe, AZ
HNRNPU	HsCD00603898	pANT7_cGST	Carbenicillin	DNASU, Tempe, AZ
-	-	pDONR201	Kanamycin	Life Tech
-	12281-010	pDEST47	Carbenicillin	Life Tech
HNRNPK	-	pDEST007	Carbenicillin	Eigen

6.1.1.5 Geräte

Tabelle 11: Liste der in dieser Dissertation verwendeten Laborgeräte.

Gerät	Hersteller
Agarose Gelelektrophoresekammer	<i>peqlab</i> , Erlangen
Analysenwaage	<i>Sartorius</i> , Göttingen
ÄKTA Purifier Chromatographiesystem	<i>GE Healthcare</i> , München
Autoklav Vakulab S3000 V65	<i>Systec</i> , Gießen
Biofuge pico	<i>Heraeus</i> , Hanau
LAS-3000 Image Analyser	Fujifilm
Gelscanner IDA	<i>Raytest</i> , Straubenhardt
Inkubator 1S	<i>Noctua</i> , Wiesloch
Inkubator 44R	<i>New Brunswick</i> , Nürtingen
Mini Protean 3 Cell	<i>Bio-Rad</i> , München
Multicaster	<i>Bio-Rad</i> , München
NanoDrop ND-1000 Spektrophotometer	<i>peqlab</i> , Erlangen
pH-Meter MP220	<i>Mettler Toledo</i> , Gießen
PowerPac Basic, Netzteil f. Elektrophorese	<i>Bio-Rad</i> , München
SDS Gelkammern	<i>Bio-Rad</i> , München
Sorvall Zentrifuge, Evolution RC	<i>Thermo Life Science</i> , Dreieich
Speedvac	<i>Thermo Life Science</i> , Dreieich
Storm 860	<i>GE Healthcare</i> , München
Themomixer Comfort	<i>Eppendorf</i> , Hamburg
Tiefkühlschrank (-86 °C)	<i>SANYO</i> , Bad Nenndorf
Tischzentrifuge 5415R	<i>Eppendorf</i> , Hamburg
Vortexer, Vortex-Genie 2	<i>Scientific Industries</i> , Bohemia
Zentrifuge 5810R	<i>Eppendorf</i> , Hamburg
Zentrifuge 5424R	<i>Eppendorf</i> , Hamburg
Reinstwasseranlage arium Pro DI	<i>Sartorius Stedim</i> , Göttingen
Genios Pro Microplate Reader	<i>TECAN</i>
TC-20 Zellzähler	<i>Bio-Rad</i> , München

6.1.2 Transformation von chemisch kompetenten *E. coli* Zellen

Chemisch kompetente Zellen (50 µL) wurden aus der Lagerung (-80 °C) entnommen und auf Eis ca. 5 min aufgetaut. Zu den Zellen wurden 50 bis 300 ng Plasmid-DNA bzw. 1 µL Clonase-Reaktion gegeben und die Mischung 30 min auf Eis inkubiert. Nach einem Hitzeschock (42 °C, 30 s) wurden die Zellen wieder auf Eis gekühlt und mit SOC Medium (450 µL) versetzt, danach bei 37 °C für 1 h inkubiert und 50 bis 150 µL der Reaktion auf LB-Agarplatten ausgestrichen. Die Platten wurden über Nacht (ca. 16 h) bei 37 °C inkubiert.

6.1.3 Bakterien-Übernachtskultur

Für Bakterien-Übernachtskulturen wurden 4 bis 6 mL LB-Medium in ein 14 mL-Kulturröhrchen (*Greiner bio-one*) oder 50 mL LB-Medium in einen 250 mL Erlenmeyerkolben gegeben und mit dem zur Selektion nötigen Antibiotikum versetzt. Mit einer sterilen Pipettenspitze wurde eine einzelne Kolonie selektiert und die Flüssigkultur damit angeimpft. Die Kultur wurde bei 37 °C und 200 rpm für 14 bis 16 h inkubiert.

6.1.4 Plasmidpräparation

Um kleine Mengen (5 bis 18 µg pro Isolation) Plasmid DNA aus Bakterienzellen zu isolieren wurde ein *QIAprep Spin Miniprep Kit* (*Qiagen*) verwendet. Aus einer Übernachtskultur *E. coli* in LB Medium wurden 2 mL entnommen und in einem Eppendorf Reaktionsgefäß zentrifugiert (5000 rpm, 5 min, RT). Der Überstand wurde verworfen und das Pellet in Puffer P1 (250 µL) resuspendiert. Nach Zugabe von Puffer P2 (250 µL) wurde die Probe zügig durchmischt, indem 4 bis 6 mal invertiert wurde. Die Reaktionen wurden 5 min bei RT inkubiert, je 350 µL Puffer N3 zugegeben und wiederum durch mehrmaliges Invertieren durchmischt. Die präzipitierte genomische DNA wurde durch Zentrifugation pelletiert (21130 g, 10 min, RT) und der Überstand in eine *QIAprep 2.0 spin column* überführt. Es wurde erneut zentrifugiert (21130 g, 60 s, RT) und der Durchfluss verworfen. Die an die Säulen gebundene Plasmid DNA wurde mit Puffer PB gewaschen (2 × 500 µL / 21130 g, 60 s, RT). Nachdem der Durchfluss des zweiten Waschschrittes verworfen wurde, erfolgte eine weitere Zentrifugation um Reste des Waschpuffers zu entfernen und es wurde schließlich mit H₂O (2 × 25 µL) in ein frisches *Eppendorf* Reaktionsgefäß eluiert.

Wurden größere Mengen Plasmid DNA (100 bis 150 µg) benötigt, wurde zu einem *QIAGEN Plasmid Midi Kit* gegriffen. Aus einer Übernachtskultur *E. coli* wurden 50 mL in ein Falcon Reaktionsgefäß überführt und die Bakterienzellen durch Zentrifugation pelletiert (6000 g, 15 min, RT). Das Pellet wurde in Puffer P1 (4 mL) resuspendiert, Puffer P2 (4 mL) zugegeben und zügig durch mehrmaliges Invertieren durchmischt. Nach einer Inkubationszeit von 5 min bei RT wurde eisgekühlter Puffer P3 (4 mL) zugegeben und wieder durch Invertieren gemischt. Die Probe wurde zentrifugiert (20 000 g, 30 min, 4 °C) und das Pellet verworfen. Ein *QIAGEN-tip 100* wurde mit 4 mL Puffer QBT äquilibriert (keine Zentrifugation) und daraufhin mit dem Überstand der Zentrifugation beladen. Die gebundene DNA wurde mit Puffer QC (2 × 10 mL) gewaschen und schließlich mit Puffer QF (5 mL) eluiert. Das Eluat wurde mit 3.5 mL Isopropanol versetzt und die präzipitierte DNA durch Zentrifugation pelletiert (15000 g, 30 min, 4 °C). Das Pellet wurde mit 70% Ethanol (2 × 2 mL) gewaschen und in ddH₂O (250 µL) resuspendiert.

6.1.5 Konzentrationsbestimmung von Nukleinsäuren

Zur Konzentrationsbestimmung von isolierter Plasmid DNA und synthetisierter Oligoribonukleotide wurde ein NanoDrop Spektrophotometer ND-1000 (*peqlab*). Im Allgemeinen besitzen kanonische Nukleotide ein Absorptionsmaximum bei etwa 260 nm, so dass die Konzentration einer Nukleinsäurelösung vereinfacht über folgende Relation bestimmt werden kann:

$$1 A_{260}\text{-Einheit} = 50 \mu\text{g/mL dsDNA} = 40 \mu\text{g/mL RNA in H}_2\text{O}$$

Hierbei wird ein durchschnittlicher Extinktionskoeffizient über alle Nukleotide angenommen, was hinreichend genau ist, wenn die Probe über mehrere tausend Basen(paare) verfügt, wie im Falle eines Plasmides. Für kurze Sequenzen aus der Festphasensynthese (20-30 nt) ergeben sich Abweichungen von dieser Näherung, die eine sequenzspezifische Berechnung nötig machen. Für individuelle Sequenzen wurde die Gesamtmenge anhand der folgenden Gleichung ermittelt:

$$n [\mu\text{mol}] = \frac{E_{260}/V [\text{mL}]}{\sum(\epsilon_{260} \times \text{Anzahl der Basen}) \times 1000}$$

Dabei wurden publizierte⁴¹⁴ und im Arbeitskreis Carell (*Katharina Iwan*) ermittelte Extinktionskoeffizienten verwendet. Für randomisierte Positionen wurden Durchschnittswerte herangezogen (**Tabelle 12**). Für Basen, deren Extinktionskoeffizient nicht bekannt war wurde der Extinktionskoeffizient des strukturell ähnlichsten Derivates verwendet (z. B. U für Ψ).

Tabelle 12: Für die präzise Berechnung der Konzentration von synthetisch hergestellten Oligoribonukleotiden verwendete Extinktionskoeffizienten.

Base	ϵ_{260} [L mol ⁻¹ cm ⁻¹]
rA	1.5×10^{-2}
rG	1.2×10^{-2}
rC	7.1×10^{-3}
U	9.7×10^{-3}
m ⁶ A	1.6×10^{-2}
m ⁶ ₂ A	1.3×10^{-2}
i ⁶ A	2.0×10^{-2}
R	1.4×10^{-2}
H	1.1×10^{-2}

Da aromatische Aminosäuren eines Proteins ein Absorptionsmaximum bei 280 nm besitzen kann über das Verhältnis 260 nm/280 nm der Probe eine Verunreinigung der DNA/RNA mit Protein detektiert werden. Bei einem Wert von 1.8 bis 2.0 wurde eine ausreichende Reinheit der Nukleinsäuren angenommen.

Zur Konzentrationsbestimmung wurde der Probenhalter des Gerätes zunächst mit ddH₂O gereinigt, dann mit 1.0 μ L ddH₂O kalibriert und mit dem Lösemittel (1.0 μ L) in dem sich die zu messende Probe befand eine Leerwertmessung durchgeführt. Anschließend konnte mit 1.0 μ L Probenlösung die Absorption ermittelt werden.

6.1.6 Gateway-Klonierung

Das *Gateway Cloning* System basiert auf homologer Rekombination zwischen zwei Vektoren. Es ist abgeleitet von den Rekombinationsreaktionen, die bei der Insertion der Phagen DNA des Bakteriophagen λ in das Genom des befallenen Wirtes stattfinden.^{415,416} Es findet hierbei ein Austausch an definierten Sequenzabschnitten des bakteriellen Chromosoms

(attB) und der Phagen DNA (attP) statt, die über einen sieben Nukleotide langen homologen Abschnitt verfügen. Der natürliche Prozess erzeugt zwei hybride Sequenzen „links“ (attL) und „rechts“ (attR) des insertierten Abschnittes.

Unter Verwendung von Plasmiden, die über die nötigen orthogonalen *att sites* verfügen, lassen sich in LR-, bzw. BP-Reaktionen schnell und zuverlässig Expressionsvektoren erzeugen (**Abbildung 72**). Die verwendeten *Destination Vectors* ermöglichen durch ihre *ccdB* (*control of cell death*) Kasette eine Selektion der wie gewünscht klonierten Plasmide, da nur im Falle einer erfolgreichen LR Reaktion ein Plasmid entsteht, das über das Resistenzgen des *Destination* Plasmids verfügt, ohne gleichzeitig das *ccdB* Toxin zu codieren.

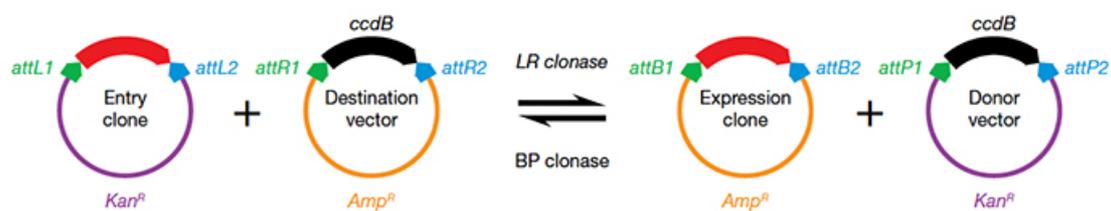


Abbildung 72: Schematische Darstellung des Gateway Cloning Systems.⁴¹⁵

Ziel der Gateway Klonierung war es, alle Inserts (*HNRNPK*, *YBX1*, *SYNCRIP*, *GTF2B*, *HNRNPU*, *PCBP1* und *TOE1*) mit einem C-terminalen GFP Tag zu versehen und in einen für Säugerzellen geeigneten Expressionsvektor zu klonieren. Aufgrund dieser Vorgaben wurde der Vektor pDEST47 (*Thermo Scientific*) als Expressionsplasmid ausgewählt.

HNRNPU wurde zunächst in einer BP-Reaktion aus dem erhältlichen Vektor (pANT7_cGST), der lediglich über *attB sites* verfügt, in einen geeigneten intermediären Vektor (pDONR201) kloniert. Darauf folgend konnten alle Gene in LR-Reaktionen in den Zielvektor pDEST47 kloniert werden (**Abbildung 73**).

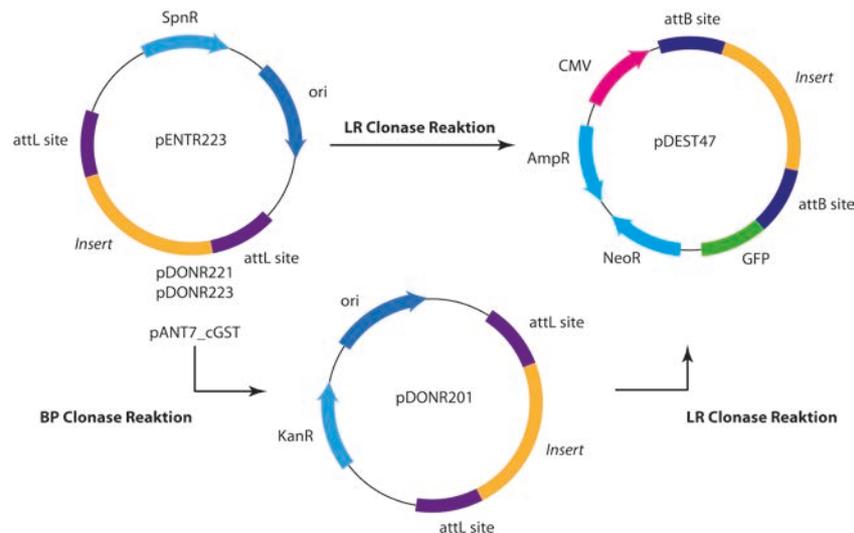


Abbildung 73: Schematische Darstellung der Vorgehensweise zur Erzeugung der gewünschten Expressionsklone für die Zielgene *HNRNPK*, *YBX1*, *SYNCRIP*, *GTF2B*, *HNRNPU*, *PCBP1* und *TOE1*.

Die Komponenten der BP-Reaktion (**Tabelle 13**) wurden zusammengegeben und die Reaktion 1 h bei 25 °C inkubiert. Um die Reaktion zu unterbrechen wurde Proteinase K (2 µL, 2 µg/µL) zugegeben und 10 min bei 37 °C inkubiert. Aus diesem Reaktionsgemisch wurde 1 µL entnommen und zu einem 50 µL Stock geeigneter Zellen (NEB 5-alpha, *New England BioLabs*) gegeben und chemisch transformiert. Der Rest der Reaktionsmischung wurde bei -20 °C aufbewahrt um die Transformation wiederholen zu können.

Tabelle 13: Zusammensetzung einer BP Clonase-Reaktion.

Komponente	Volumen [µL]
HNRNPU-pANT7_cGST (327 ng/µL)	1.5 µL
pDONR201 (150 ng/µL)	2.0 µL
5 × Clonase Reaction Buffer	4.0 µL
TE Puffer (pH 8.0)	8.5 µL
BP Clonase enzyme mix	4.0 µL

Die LR-Reaktionen wurden durchgeführt, indem ebenfalls alle Komponenten zusammengegeben wurden und die Reaktionen für 4 h bei 25 °C inkubiert wurden. Um die Reaktionen zu unterbrechen wurde Proteinase K (1 µL, 2 µg/µL) zugegeben und 10 min bei 37 °C inkubiert. Aus dieser Mischung wurde 1 µL entnommen und auf einen 50 µL Stock geeigneter Zellen (NEB 5-alpha, *New England BioLabs*) gegeben und chemisch transformiert. Der Rest der Reaktionsmischung wurde bei -20 °C aufbewahrt um die Transformation wiederholen zu können.

Tabelle 14: Zusammensetzung einer LR Clonase-Reaktion.

Komponente	Volumen [μL]
Entry Clone (100 ng/Reaktion)	0.27 μL – 0.84 μL
pDEST47 (150 ng/ μL)	1.0 μL
TE Puffer (pH 8.0), <i>ad</i> 8.0 μL	6.16 μL – 6.73 μL
LR Clonase II enzyme mix	2.0 μL

6.1.7 DNA Restriktion

Um den Erfolg der Klonierungsreaktionen zu überprüfen wurden die erhaltenen Plasmide einem Restriktionsverdau unterzogen. Die DNA (400 ng/Probe) wurde in einem Gesamtvolumen von 10 μL in dem vom Hersteller empfohlenen Puffer für 1.5 h bis 3 h bei 37 °C verdaut und anschließend mittels Agarosegelelektrophorese analysiert

6.1.8 DNA-Agarosegelelektrophorese

Von Restriktionsenzymen verdaute DNA wurde je nach erwarteter Fragmentgröße in 1.0 bis 1.5%igen Agarosegelen der Größe nach aufgetrennt. Zur Herstellung der Gele wurde Agarose in 50 mL TAE-Gelpuffer in einer Mikrowelle erhitzt. Nach kurzem Abkühlen wurde die Lösung mit Ethidiumbromid (8 μL , 1 mg/mL) und in eine Gelapparatur (*peqlab*) gegossen. DNA-Proben wurden vor dem Laden mit DNA-Auftragspuffer versetzt. Die Elektrophorese erfolgte bei einer konstanten Spannung von 90 V in TAE-Laufpuffer. Banden wurden unter UV-Licht ($\lambda = 254 \text{ nm}$) visualisiert.

6.1.9 Sequenzierung

Zusätzlich zu einem Restriktionsverdau wurden die erzeugten Plasmide zum Teil durch die Firma *GATC* (Konstanz) sequenziert. Hierzu wurde gemäß den Vorgaben des Anbieters 1.6 μg Plasmid (20 μL , 80 ng/ μL) je Probe versandt. Die erhaltenen Sequenzen wurden mit *SerialCloner* (Version 2.6.1) auf Übereinstimmung mit der theoretischen Sequenz überprüft. Die verwendeten Primer sind in der unten stehenden Tabelle aufgeführt.

Tabelle 15: Für die Sequenzierung (*GATC*) von Plasmiden verwendete Primer.

Plasmid	Primer	Sequenz
HNRNPU-pANT7_cGST	<i>Custom</i>	5'-TGAAGGATGCCCAGAAG-3'
<i>Insert</i> -pENTR223	UP-40	5'-GTTTTCCCAGTCACGAC-3'
<i>Insert</i> -pDONR221	UP-40	5'-GTTTTCCCAGTCACGAC-3'
<i>Insert</i> -pDONR223	UP-40	5'-GTTTTCCCAGTCACGAC-3'
<i>Insert</i> -pDEST47	pc3.1GFP-Topo-RP	5'- CCCATTAACATCACC-3'
HNRNPK-pENTR223	M13-RP	5'-CAGGAAACAGCTATGACC-3'
HNRNPU-pANT7_cGST	pESP-RP	5'-TCCAAAAGAAGTCGAGTGG-3'
HNRNPU-pDONR201	pDONOR-RP	5'- GCAATGTAACATCAGAGAT-3'

6.1.10 Konzentrationsbestimmung von Proteinen mittels Bradford Assay

Der Bradford Assay ist eine etablierte photometrische Methode zur Konzentrationsbestimmung wässriger Proteinlösungen. Er basiert auf der Verschiebung des Absorptionsmaximums des Farbstoffes Coomassie-Brilliant-Blau G-250 bei Komplexbildung mit Proteinen. Eine Koordinierung an Proteine in Lösung stabilisiert das Anion des Chromophors und sorgt damit für eine Verschiebung des Absorptionsmaximums von 465 auf 595 nm.⁴¹⁷

Zur Durchführung des Tests wurde zunächst eine 20%ige (v/v) Verdünnung des Farbstoffkonzentrates (Protein Assay Dye Reagent Concentrate, *Bio-Rad*) hergestellt. In einer Einmalküvette (*Brand*) wurden 1-3 µL der Proteinlösung in 47-49 µL ddH₂O (Gesamtvolumen 50 µL) gegeben. Zu dieser Lösung wurden 950 µL der Farbstoffverdünnung gegeben und gründlich durchmischt. Es wurde unter Lichtausschluss bei RT für 5 min inkubiert und die Absorption bei 595 nm in einem Spektrophotometer (BioPhotometer, *Eppendorf*) bestimmt. Um eine Beeinflussung der Messung durch Pufferbestandteile zu minimieren wurde eine zur Probe identische Puffermenge (1-3 µL) als *Blank* verwendet.

6.1.11 Zellkultur

6.1.11.1 HEK293 und HeLa Zellkultur

Adhärente HeLa und HEK293 Kulturen wurden in Dulbecco's Modified Eagle's Medium (DMEM, HeLa) oder RPMI-1640 (HEK293) bei 37 °C und 5% CO₂ kultiviert. Alle drei bis vier Tage wurden Zellpopulationen im Verhältnis 1:5 bzw. 1:10 passagiert.

Hierzu wurde das Medium entfernt, die Kultur mit DPBS gewaschen und Trypsin-EDTA Lösung zugegeben. Nach Inkubation für 3 min bei 37 °C und 5% CO₂ wurden die Zellen in Medium aufgenommen, bei 300 g für 3 min pelletiert, in Medium resuspendiert, vereinzelt und in vorgewärmtes Medium ausgesät.

Zur Langzeitlagerung wurden die Zellen in FBS mit 10% DMSO aufgenommen, für 15 min bei RT äquilibriert, und je 1×10^6 Zellen in *CryoTubes* (VWR) gefroren.

6.1.11.2 HAP1 Zellkultur

Die Kultivierung und Ernte aller HAP1 Kulturen wurde von *Angie Kirchner* durchgeführt. Humane HAP1-Zellen (*Horizon*) und CDK5RAP1-Knockout Zellen (*Horizon*) wurden in Iscove's Modified Dulbecco's Medium (IMDM, *Sigma-Aldrich*) bei 37 °C und 5% CO₂ kultiviert. Die Zellen wurden alle zwei bis drei Tage im Verhältnis 1:6 bzw. 1:12 passagiert. Die adhärennten Zellen wurden in IMDM-Wachstumsmedium bis zu 80% Konfluenz kultiviert, das Medium abgenommen und mit DPBS (*Sigma-Aldrich*) gewaschen. Anschließend wurden die Zellen mit Trypsin-EDTA Lösung (*Sigma-Aldrich*) versetzt und 5 min bei 37 °C und 5% CO₂ inkubiert. Das Trypsin wurde mit IMDM-Medium inaktiviert, die Zellen vereinzelt und 3 min bei 300 g pelletiert.

Zur Langzeitlagerung in flüssigem Stickstoff wurden die Zellen in IMDM-Wachstumsmedium aufgenommen und die gleiche Menge (v/v) Einfriermedium, bestehend aus 20% DMSO, 40% FBS und 40% IMDM, langsam zugegeben. Die Zellen wurden 15 min bei RT äquilibriert und zur anschließenden Lagerung in *CryoTubes* (VWR) aliquotiert.

6.1.12 Interaktionsassays

6.1.12.1 Electrophoretic Mobility Shift Assay (EMSA) unter Verwendung von gereinigtem hnRNP K

Um die Interaktion des potentiellen ms²i⁶A Reader Proteins HNRNP K mit unterschiedlichen RNA Sequenzen zu untersuchen sollten u. a. *Electrophoretic Mobility Shift Assays* (EMSAs) durchgeführt werden. Der EMSA stellt eine etablierte Methode zur Analyse von Protein-DNA bzw. Protein-RNA Interaktionen dar und beruht auf einer Änderung der elektrophoretischen Mobilität eines Protein-Oligonukleotid-Komplexes gegenüber einem ungebundenen Oligonukleotid.^{394,400} Eine Variante stellt der Kompetitionsassay dar, in dem ein (fluoreszenz-) markiertes Oligonukleotid unter Verwendung einer nicht-markierten („kalten“)

Sequenz aus einem bestehenden Komplex verdrängt werden soll, was wiederum die Retention herabsetzt. Hierfür wurde zunächst HNRNPK_StrepII in Zusammenarbeit mit *Martin Rossa* und *Jessica Steinbacher* in *E. coli* heterolog exprimiert und aus transformierten Bakterienzellen aufgereinigt. Es wurden EMSAs mit insgesamt 6 unterschiedlichen ORN und darüber hinaus Konkurrenzexperimente mit unmarkierter RNA durchgeführt. Standardexperimente enthielten je eine Probe ohne Protein und mehrere Proben mit ansteigender Proteinkonzentration. Konkurrenzexperimente enthielten Proben mit ausreichend Protein für eine quantitative Verschiebung (2.0 Äquivalente) und ansteigende Konzentrationen einer nicht fluoreszenzmarkierten RNA der gleichen, unmodifizierten Sequenz.

Das Zielprotein HNRNPK lag nach der heterologen Expression in einer Konzentration von 3.5 µg/µL (*Bradford Assay*) vor. Sowohl bei -80 °C, als auch bei -20 °C aufbewahrtes Protein zeigte in Kontrollexperimenten eine vergleichbare Aktivität wie frisch aufgereinigtes. Nach dem Auftauen wurde aliquotiertes Protein nicht mehr gefroren, sondern auf Eis bei 4 °C gelagert. Vor jedem Experiment wurde die Proteinlösung mit dem Reaktionspuffer (EMSA-A oder VM-Puffer) auf eine Konzentration von 35 ng/µL verdünnt. Alle verwendeten Cy3-ORN wurden in einer Konzentration von 2.5 µM in H₂O eingesetzt.

Für die einzelnen Reaktionen wurden zunächst Eppendorf Reaktionsgefäße auf Eis gekühlt, Reaktionspuffer (EMSA-A oder VM-Puffer) vorgelegt und anschließend Protein zugegeben. Die Proben wurden für 10 min auf Eis äquilibriert und dann RNA zugegeben (2.5 pmol). Das Gesamtvolumen der Reaktionen betrug 15 µL. Jede Reaktion wurde 30 min bei 25 °C inkubiert und wieder auf Eis gekühlt. Nach 5 min wurde 6 × *Native PAGE* Ladepuffer zugegeben (3.0 µL), weitere 5 min auf Eis inkubiert und die Proben anschließend auf ein natives PAGE Gel (6%, 0.75 mm) geladen. Das Gel wurde vorbereitet, indem Gelapparat und Puffer (0.5 × TBE) auf Eis gekühlt wurden und das gekühlte Gel im unbeladenen Zustand 30 min bei 130 V konstanter Spannung einlief.

Nach dem Laden der Proben wurde die Elektrophorese zunächst bei 300 V/0 °C für 7 min und anschließend bei 130 V/0 °C für 40 min durchgeführt. Das Gel wurde rasch aus der Apparatur entfernt und die fluoreszenten Banden in einem *LAS-3000 Image Analyser (Fujifilm)* visualisiert. Für Cy3-markierte ORN wurden dabei folgende Filtereinstellungen verwendet: $\lambda_{\text{ex}} = 520 \text{ nm}$, $\lambda_{\text{abs}} = 605 \text{ nm}$. Die Zusammensetzung eines typischen Ansatzes ist in **Tabelle 16** dargestellt.

Tabelle 16: Zusammensetzung eines typischen EMSA Experimentes mit aufgereinigtem HNRNPK

	1	2	3	4	5
<i>Cy3-ORN</i> [pmol]	2.5	2.5	2.5	2.5	2.5
<i>HNRNPK</i> [pmol]	0.00	1.25	2.00	2.50	5.00
<i>HNRNPK</i> [ng]	0	69	110	137	274
<i>HNRNPK : Cy3-ORN</i>	0.0	0.5	0.8	1.0	2.0
<i>HNRNPK</i> [μ L]	0.00	1.96	3.13	3.92	7.83
<i>Cy3-ORN</i> [μ L]	1.0	1.0	1.0	1.0	1.0
<i>Puffer</i> [μ L]	14.04	12.08	10.91	10.12	6.21
<i>Gesamtvolumen</i> [μ L]	15.0	15.0	15.0	15.0	15.0

Für Konkurrenzexperimente wurde ebenfalls Puffer auf Eis vorgelegt, Protein zugegeben und alle Proben für 10 min auf Eis äquilibriert. Nach Zugabe der markierten RNA wurde 30 min bei 25 °C inkubiert und anschließend unmarkierte Kompetitor-RNA zugegeben. Diese wurde in einer Verdünnungsreihe vorbereitet, um in jede Probe das gleiche Volumen (1 μ L) geben zu können. Das Gesamtvolumen der Proben betrug 16 μ L. Es wurde weitere 5 min bei 25 °C inkubiert, die Proben wieder auf Eis gestellt und nach 5 min 6 \times Native PAGE Ladebuffer (3.2 μ L) zugegeben. Im Weiteren wurde wie oben beschrieben verfahren. Die Zusammensetzung eines typischen Ansatzes ist in **Tabelle 17** dargestellt.

Tabelle 17: Zusammensetzung eines typischen EMSA Experimentes mit aufgereinigtem HNRNPK und Verwendung einer Kompetitor-RNA.

	1	2	3	4	5
<i>Cy3-ORN</i> [pmol]	2.5	2.5	2.5	2.5	2.5
<i>HNRNPK</i> [pmol]	5.00	5.00	5.00	5.00	5.00
<i>HNRNPK</i> [ng]	274	274	274	274	274
<i>HNRNPK : Cy3-ORN</i>	2.0	2.0	2.0	2.0	2.0
<i>Komp-ORN</i> [pmol]	0.00	1.25	2.50	5.00	7.50
<i>Cy3-ORN : Komp-ORN</i>	0.0	0.5	1.0	2.0	3.0
<i>HNRNPK</i> [μ L]	7.83	7.83	7.83	7.83	7.83
<i>Cy3-ORN</i> [μ L]	1.0	1.0	1.0	1.0	1.0
<i>Puffer</i> [μ L]	6.21	6.21	6.21	6.21	6.21
<i>Komp-ORN</i> [μ L]	1.00	1.00	1.00	1.00	1.00
<i>Gesamtvolumen</i> [μ L]	16.0	16.0	16.0	16.0	16.0

6.1.12.2 Electrophoretic Mobility Shift Assay (EMSA) unter Verwendung von Zellysat und Supershift-Assay

Heterolog in *E. coli* exprimierte Proteine verfügen nicht über posttranslationale Modifikationen, wie Methylierungen oder Phosphorylierungen, was ihre intermolekularen Wechselwirkungen mit RNA oder anderen Proteinen beeinflussen kann. Die Arbeit mit gereinigtem Protein ermöglicht es darüber hinaus nicht, Wechselwirkungen zwischen Oligonukleotiden und Proteinkomplexen zu untersuchen, deren Zusammensetzung nicht, oder nur teilweise bekannt ist. Um solche Einschränkungen zu umgehen ist es möglich EMSAs mit (fraktioniertem) Zellysat durchzuführen.⁴⁰⁰ Wenn hier ein *Shift* beobachtbar ist, muss das interagierende Protein noch durch eine anschließende Analyse identifiziert werden, was durch Exzision der gewünschten Bande aus dem Gel und darauf folgende massenspektrometrische Identifikation erfolgen kann. Eine weitere Möglichkeit zur Identifikation stellt der *Supershift* dar, bei welchem zum Zellextrakt ein gegen ein bekanntes Zielprotein gerichteter Antikörper gegeben wird und der im Experiment sichtbarer Komplex eine weitere Verschiebung (*Supershift*) durch den Antikörper erfährt.

Die EMSA Experimente mit Zellysat wurden stets mit frisch hergestellten Extrakten durchgeführt. Aufgrund der erheblich höheren Ausbeuten bezüglich der Gesamtproteinmenge wurde mit HEK293 Zellen gearbeitet. Für ein typisches Experiment wurden 5×10^7 bis 8×10^7 Zellen verwendet. Zellen wurden durch Trypsinieren unter Standardbedingungen geerntet, das Zellpellet mit DPBS (3×25 mL) gewaschen und anschließend in eisgekühltem EMSA-Lysispuffer (1.5-faches Volumen des Zellpellets) aufgenommen. Die Probe wurde 60 min auf Eis inkubiert und dabei alle 5 min auf und ab pipettiert. Das so erhaltene rohe Lysat wurde durch Zentrifugation (21130 g, 15 min, 4 °C) geklärt und die Proteinkonzentration mittels *Bradford Assay* bestimmt. Interaktionsassays konnten nur erfolgreich mit hochkonzentrierten Lysaten ($> 10 \mu\text{g}/\mu\text{L}$) durchgeführt werden.

Vor der Reaktion wurde das konzentrierte Lysat mit VM-Puffer auf eine Arbeitskonzentration von etwa $4 \mu\text{g}/\mu\text{L}$ verdünnt. Eppendorf-Reaktionsgefäße wurden zunächst auf Eis vorgekühlt, VM-Puffer (*ad* $10 \mu\text{L}$) vorgelegt und anschließend Lysat zugegeben. Zuletzt wurde das Cy3-ORN zugegeben ($2.8 \text{ pmol/Reaktion}$) und die Reaktion inkubiert (30 min, 25 °C, 300 rpm und anschließend 90 min, 4 °C, 300 rpm). Die Proben wurden mit vorgekühltem $6 \times$ Native PAGE Ladepuffer ($2.2 \mu\text{L}$) versetzt und auf ein vorbereitetes Gradientengel geladen (Mini-PROTEAN TGX 4-20%, *Bio-Rad*). Das Gel wurde vorbereitet indem es im Laufpuffer auf Eis gekühlt wurde und im unbeladenen Zustand 30 min bei 130 V konstanter Spannung

einlief. Nach dem Laden der Proben wurde die Elektrophorese zunächst bei 300 V/0 °C für 7 min und dann bei 130 V/0 °C für 60 min durchgeführt. Gelbanden wurden wie im vorherigen Abschnitt beschrieben visualisiert. Ein Beispiel ist in **Tabelle 18** dargestellt.

Tabelle 18: Zusammensetzung eines typischen EMSA Experimentes unter Verwendung von gesamtem Zelllysats.

	1	2	3	4	5
<i>Cy3-ORN</i> [pmol]	2.8	2.8	2.8	2.8	2.8
<i>Zelllysats</i> [µg]	0.0	2.5	5.0	7.5	10.0
<i>VM-Puffer</i> [µL]	9.00	8.32	7.65	6.97	6.30
<i>Cy3-ORN</i> [µL]	1.0	1.0	1.0	1.0	1.0
<i>Zelllysats (3.7 µg/µL)</i> [µL]	0.00	0.68	1.35	2.03	2.70
<i>Gesamtvolumen</i> [µL]	10.0	10.0	10.0	10.0	10.0

Die Durchführung der *Supershift* Experimente verlief analog. Als Kontrollexperiment diente je ein Ansatz, der lediglich Cy3-ORN und Antikörper enthielt. Hier sollte keine Interaktion stattfinden. Zunächst wurde auf Eis VM-Puffer vorgelegt und Lysat zugegeben. Die Reaktion wurde mit Antikörper versetzt (1 µg/µL Verdünnung in VM-Puffer) und 30 min bei 4 °C und 300 rpm inkubiert. Anschließend wurde fluoreszenzmarkierte RNA zugegeben und erneut inkubiert (30 min, 25 °C, 300 rpm und anschließend 90 min, 4 °C, 300 rpm). Es wurde vorgekühlter 6 × Native PAGE Ladepuffer (2.2 µL) hinzugefügt und eine Gelelektrophorese durchgeführt wie oben beschrieben.

Tabelle 19: Zusammensetzung eines typischen EMSA Experimentes unter Verwendung von gesamtem Zelllysats und Zusatz von spezifischen Antikörpern (*Supershift Assay*).

	1	2	3
<i>Cy3-ORN</i> [pmol]	2.8	2.8	2.8
<i>Zelllysats</i> [µg]	0.0	10.0	10.0
<i>Antikörper</i> [µg]	1.5	0.3	1.5
<i>VM-Puffer</i> [µL]	7.50	6.00	4.80
<i>Cy3-ORN</i> [µL]	1.0	1.0	1.0
<i>Zelllysats (3.7 µg/µL)</i> [µL]	0.00	2.70	2.70
<i>Antikörper (1 µg/µL)</i> [µL]			
<i>Gesamtvolumen</i> [µL]	10.0	10.0	10.0

6.1.12.3 RNA Pulldown mit GFP-markiertem HNRNPK

Statt modifizierte Oligonukleotide als Zielmolekül (*Bait*) für Protein-*Pulldown* Experimente zu verwenden ist auch eine umgekehrte Vorgehensweise denkbar, also die spezifische Anreicherung einer visualisierbaren RNA durch das immobilisierte Zielprotein. Wenn ein markiertes Protein (His₆, GST, GFP o. ä.) in einem geeigneten Zelltyp exprimiert wird, kann anschließend mit Hilfe eines fluoreszenzmarkierten ORN ohne vorherige Aufreinigung eine Bindungsanalyse durchgeführt werden. Da der *Pulldown* direkt aus Zellextrakten durchgeführt werden kann, können auch Interaktionen, die von posttranslationalen Modifikationen abhängig sind, oder Proteinkomplexe benötigen, sichtbar gemacht werden.

Um GFP-markiertes HNRNPK in nativem Zustand zu exprimieren wurden 2×10^5 HEK293 Zellen in *6-wells* (9.6 cm²) ausgesät und 24 h später transfiziert. Nach einer Inkubationszeit von 5 bis 6 h wurde das Transfektionsmedium entfernt und gegen RPMI Medium ausgetauscht das Natriumbutyrat (4 mM) enthält. Dieses Medium verblieb etwa 18 h auf den Zellen und wurde schließlich durch Standardmedium ersetzt. Nach weiteren 20 bis 24 h wurden die transfizierten Zellen geerntet und aufgeschlossen.

Hierzu wurde das Medium entfernt und die Zellen mit DPBS gewaschen (2×4 mL). Die Kulturplatten wurden auf Eis gekühlt und direkt mit vorgekühltem EMSA-Lysispuffer (150 µL) versetzt. Mit einem Zellschaber wurden noch haftende Zellen restlos abgelöst und die Suspension in ein Eppendorf Reaktionsgefäß überführt. Die Probe wurde eine Stunde auf Eis inkubiert, wobei alle 5 min auf und ab pipettiert wurde. Das trübe Lysat wurde durch Zentrifugation (21130 g, 15 min, 4 °C) von unlöslichen Bestandteilen befreit und die Proteinkonzentration mittels *Bradford Assay* bestimmt. Eine erfolgreiche Transfektion ist an diesem Punkt bereits an einer leichten Grünfärbung erkennbar. Darüber hinaus sollte der Extrakt unter einer laborüblichen UV-Lampe ($\lambda_{\text{ex}} = 366$ nm) deutlich sichtbar fluoreszieren.

Pro Reaktion wurde 15 µL *Bead Slurry* (GFP-Trap_M, *ChromoTek*) vorbereitet, indem die Beads mit Dilution Buffer (2×250 µL) und VM-Puffer (250 µL) gewaschen wurden und schließlich in VM-Puffer (500 µL) resuspendiert wurden.

Für jeden Ansatz wurden 100 µg Zellextrakt mit fluoreszenzmarkierter RNA (10 bis 50 pmol) versetzt und für 30 min auf Eis inkubiert. Im Anschluss wurde die vorbereitete Beadsuspension in VM-Puffer zugegeben und die Proben zunächst für 30 min bei RT und dann für 90 min bei 4 °C (Rotator) inkubiert. Der Überstand wurde abgenommen und verworfen, die Beads wurden mit VM-Puffer (3×500 µL) gewaschen und in Proteinase K Puffer (25 µL) aufgenommen. Nach Zugabe des Verdauenzym (Proteinase K, 1 µL, 1 µg/µL) wurde inkubiert (45 min, 37 °C, 300 rpm) und die Reaktion durch Zugabe von

2 × RNA Ladepuffer (7 µL) unterbrochen. Zur Elution wurden die Reaktionen kurz erhitzt (2 min, 50 °C), der Überstand abgenommen und mit ddH₂O (50 µL) gewaschen. Die vereinigten Überstände wurden evaporiert und der Rückstand in ddH₂O gelöst. Direkt vor dem Beladen des Gels wurde nochmals kurz erhitzt (2 min, 85 °C) und die Proben auf ein denaturierendes Harnstoff-PAGE Gel (15%) geladen. Die Elektrophorese wurde bei konstanten 130 V und RT durchgeführt, bis der Marker (Xylen-Cyanol FF) etwa zwei Drittel des Gels zurückgelegt hatte. Die Analyse der Gelbanden erfolgte an einem *LAS-3000 Image Analyser (Fujifilm)* ($\lambda_{\text{ex}} = 520 \text{ nm}$, $\lambda_{\text{abs}} = 605 \text{ nm}$).

6.1.12.4 Fluoreszenz-Polarisationsmessungen

Fluorophore, die durch polarisiertes Licht angeregt werden emittieren ebenfalls polarisiertes Licht. Nachdem zwischen Anregung und Emission eine kurze zeitliche Pause (Lebensdauer des angeregten Zustandes) liegt, geht ein Teil der Polarisation des eingestrahlt Lichtes durch Brownsche Rotation verloren. Je immobilere der Fluorophor ist, also je langsamer seine Rotation stattfindet, desto höher ist der Polarisationsgrad des emittierten Lichtes. Findet eine Bindung eines Oligonukleotides durch ein interagierendes Protein statt wird dadurch seine Mobilität, bzw. Rotationsfreiheit reduziert. Dies führt folglich dazu, dass ein Anstieg des Polarisationsgrades messbar ist, der parallel zur Proteinkonzentration ansteigt.³⁹⁶⁻³⁹⁸ Ein großer Vorteil dieser Art von Assay stellt die direkte Analyse der Interaktion in Lösung dar, die durch EMSA oder *Pulldown* Experimente nicht realisierbar ist.

Die Experimente wurden mit heterolog in *E. coli* exprimiertem und gereinigtem HNRNPK_StrepII und 5'-Cy3-markierten ORN durchgeführt (siehe **Tabelle 7**, Kapitel 5.6.2.1). Um einen sinnvollen Konzentrationsbereich zu ermitteln wurden zunächst Vorexperimente durchgeführt in denen ein Konzentrationsbereich ohne signifikante Bindung und die Grenze für eine Sättigung der Bindung identifiziert wurden. Dann wurde eine Reihe von zwölf Messpunkten ausgewählt, die den gesamten Bereich abdeckten. Die einzelnen Reaktionen wurden in einem Gesamtvolumen von 100 µL in einer 96-well Platte (96 F Nunclon Delta, *Thermo Scientific*) durchgeführt. Die Platten wurden auf Eis vorgekühlt und VM-Puffer (*ad* 50 µL) vorgelegt. In aufsteigender Konzentration (0 bis 500 nM) wurde HNRNPK zugegeben und die Proben für 5 min auf Eis äquilibriert. Anschließend erfolgte die Zugabe des Cy3-ORN (80 nM). Die Reaktionen wurden für 30 min bei 25 °C inkubiert und die Fluoreszenzpolarisation an einem GeniusPro (*Tecan*) Fluoreszenzphotometer gemessen

($\lambda_{\text{ex}} = 535 \text{ nm}$, $\lambda_{\text{abs}} = 586 \text{ nm}$, G-Faktor = 0.629). Jede Messung wurde im technischen Triplikate vorgenommen.

6.1.13 Proteomweite *Label-Free Quantification* (LFQ)

6.1.13.1 Fraktionierte Lyse für die proteomweite LFQ

Die humane HAP1-Zellen (WT und CDK5RAP1-Knockout, *Horizon*) wurden trypsinisiert, pelletiert (300 g, 3 min), mit 10 mL DPBS gewaschen und wieder pelletiert (300 g, 3 min). Etwa 5×10^7 Zellen wurden in 2.5 mL Extraktionspuffer (10 mM HEPES-KOH pH 7.9, 1.5 mM MgCl_2 , 10 mM KCl) resuspendiert und 10 min auf Eis inkubiert. Anschließend wurde in einen Dounce Homogenisator überführt und mit ca 40 Stößen homogenisiert, bis weniger als 50% lebende Zellen zu detektieren waren (*Bio-Rad* TC20). Anschließend wurde bei 600 g 10 min bei 4 °C pelletiert. Das Pellet besteht an diesem Punkt hauptsächlich aus Zellkernen. Cytosol (mit Mitochondrien) befindet sich im Überstand. Zellkerne wurden durch Zugabe von 200 μL CellLytic M (*Sigma-Aldrich*, mit $1 \times$ Roche Complete Protease Inhibitor) für 30 min auf Eis lysiert und dabei regelmäßig auf und ab pipettiert. Aus dem Überstand wurden Mitochondrien durch Zentrifugation abgetrennt (11 000 g, 10 min, 4 °C). Das mitochondriale Pellet wurde nun durch Zugabe von 50 μL CellLytic M und Inkubation für 30 min auf Eis lysiert. Der Überstand besteht überwiegend aus Cytosol.

Die auf diese Weise erhaltenen drei Fraktionen des Zellextraktes (Zellkerne, Mitochondrien, Cytosol) wurden anschließend durch Zentrifugation geklärt (21 130 g, 15 min, 4 °C) und die Proteinkonzentration mittels *Bradford Assay* bestimmt.

6.1.13.2 Probenvorbereitung für die proteomweite LFQ

Für die proteomweite Analyse der HAP Extraktfraktionen wurden je 20 μg Zelllysate verwendet. Die Reduktion der Proben erfolgte durch Zugabe von TCEP (100 mM) und einer Inkubation bei 60 °C und 650 rpm für 1 h. Anschließend wurden die Proben durch Zugabe von Iodacetamid (200 mM) und Inkubation bei Raumtemperatur für 30 min alkyliert. Der pH-Wert der Proben wurde mit pH-Indikatorpapier überprüft und durch Zugabe von TEAB ($\text{Et}_3\text{N} \cdot \text{H}_2\text{CO}_3$, 1 M in H_2O , *Sigma-Aldrich*) auf etwa pH = 8 eingestellt. Der Verdau der Proben erfolgte durch Zugabe 0.5 μg Trypsin (*Promega*) bei 37 °C für 16 h. Um die Reaktion zu beenden wurde anschließend Phenylmethylsulfonylfluorid (PMSF, 1 mM) zugegeben. Die Aufreinigung und fraktionierte Elution der Proben erfolgte über *StageTips*.⁴⁰⁵

6.1.13.3 LC-MS/MS Analyse

Die Proben wurden mit einem UltiMate 3000 RSLCNano Flüssigkeitschromatographie-system (Dionex, Thermo Fisher Scientific) und einem Q Exactive HF Massenspektrometer (Thermo Fisher Scientific) analysiert. Mit Hilfe einer μ -Vorfiltersäule (PepMap100, C₁₈, 5 μ M, 100 Å, size 300 μ m i.d. x 5 mm (Dionex, Thermo Fisher Scientific)) wurden die Proben aufkonzentriert und entsalzt und anschließend über eine selbstgepackte analytische Chromatographiesäule (ReproSil-Pur 120 C18-AQ, C₁₈, 1.9 μ M, 120 Å (Dr. Maisch GmbH) in einem 75 μ m ID x 150 mm Silica Picotip Emitter mit einem 8 μ m Tip (New Objective) injiziert. Die Proben wurden mit Hilfe eines 240 min dauernden Gradienten bei einer Flussrate von 300 nL/min aufgetrennt (siehe **Tabelle 20**).

Tabelle 20: Trenngradient für die Label-Free Quantification auf Proteomebene. Puffer A = H₂O + 0.1% Ameisensäure; Puffer B = Acetonitril + 0.1% Ameisensäure; Flussrate = 300 nL/min.

Zeit [min]	Puffer A	Puffer B [%]
0	99%	1%
3	99%	1%
5	92%	8%
200	68%	32%
215	40%	60%
220	15%	85%
230	15%	85%
235	99%	1%
240	99%	1%

Es wurden ausschließlich *LC-MS grade* Lösungsmittel verwendet. Die massenspektrometrische Analyse erfolgte im *Full Mass Scan* zwischen den Grenzen $m/z = 300$ und $m/z = 1750$ bei einer Auflösung von 120000 und einer maximalen Ioneneinjektionszeit von 120 ms im *profile* Modus. Diesem Suchscan folgend wurden die 20 intensivsten Ionen selektiert, fragmentiert und im *centroid* Modus mit den folgenden Parametern gemessen: einer Auflösung von 30000, ein AGC Target von $2e5$ Ladungen, einer maximalen Ioneneinjektionszeit von 70 ms, einem Isolationsfenster von 2 m/z mit einem Abstand von $+0.4 m/z$ und einer normalisierten HCD Energie von 28%.

Signale mit unbekanntem Ladungszustand oder einem Ladungszustand von 1, 7, 8 oder höher wurden nicht für die Fragmentierung ausgewählt. Um *supersampling* zu vermeiden, wurden die Signale für 40 s ausgeschlossen, nachdem sie für eine Isolation und Fragmentierung selektiert wurden. Die *Peptide-match* Einstellung wurde auf "bevorzugt" gestellt und die *exclude isotope* Einstellung wurde auf "an" gestellt.

6.1.13.4 Datenauswertung für die LFQ

Für die Aufbereitung der LFQ-Proben der Gesamtproteome wurde die Software MaxQuant (Version 1.5.0.25) verwendet. Für die Quantifizierung wurden je drei biologische Replikate herangezogen.

Die Andromeda Suchmaschine wurde in Kombination mit der Uniprot Datenbank (*H. Sapiens*). Es wurden maximal zwei *missed cleavage sites* erlaubt. Die *peptide tolerance* während der Hauptsuche wurde auf 4.5 ppm gestellt. Carbamidomethyl (C) wurde als statische Modifikation eingestellt. Variable Modifikationen waren Acetylgruppen (Protein N-term) und Oxidationen (M). Der LFQ Algorithmus wurde mit *default* Einstellungen angewendet. Die Möglichkeit "*match between runs*" wurde ebenfalls verwendet.

Die LFQ Daten wurden anschließend mit der Perseus Software (Version 1.5.0.9) analysiert. Die LFQ Intensitäten wurden logarithmiert und es wurden nur Proteine einbezogen die in allen 3 Replikaten eines Datensatzes zu finden waren.

6.1.14 Konfokalmikroskopische Lokalisationsstudien

Für die subzelluläre Lokalisation der klonierten Binderproteine *HNRNPK*, *YBX1*, *SYNCRIP*, *GTF2B*, *HNRNPU*, *PCBP1* und *TOE1* wurden durch *Gateway*-Klonierung erzeugte GFP-Fusion Konstrukte (C-terminal) verwendet. Als einheitlicher Vektor für alle Konstrukte wurde pDEST47 (*Thermo Fisher*) verwendet.

In einer Platte der Firma *Ibidi* (μ -Slide 8-well) mit einer Oberfläche von 1 cm² pro well wurden 3×10^4 HeLa Zellen 24 h in 200 μ L Standardmedium unter Normalbedingungen kultiviert. Mit Hilfe des Transfektionsreagenzes *JetPrime* (0.4 μ L) der Firma *PeqLab* wurden die Zellen jeweils mit 200 ng Plasmid DNA in 10 μ L Reaktionspuffer transfiziert. Nach 4 h Inkubationszeit wurde das Medium ausgetauscht und die Zellen für weitere 20 h inkubiert. Nach 24 h sowie nach 48 h wurden die Zellen unter einem *Leica TCS SPE* Konfokalmikroskop betrachtet. Zur Visualisierung der Zellkerne wurde der Farbstoff *Draq5* (5 μ M)

von *Cell Signaling Technology* der Herstellerangabe folgend verwendet. Als Anregungswellenlängen wurden für das GFP-Signal 488 nm und für den Farbstoff *Draq5* 635 nm verwendet.

6.2 Oligonukleotidsynthese

Die in dieser Arbeit verwendeten Oligoribonukleotide wurden mittels RNA Festphasensynthese synthetisiert und per HPLC und MALDI-ToF charakterisiert, bzw. auf ihre Reinheit geprüft. Standardamidite wurden von der Firma *Link Tech* bezogen, Sonderamidite von *Link Tech* (5'-Biotin, Am, Gm, Cm, Um, Cy3) und *Glen Research* (Ψ). Nicht kommerziell erhältliche Sonderamidite wurden selbst synthetisiert (m^6A , m^6_2A , i^6A , ms^2i^6A).

6.2.1 Festphasensynthese

Die RNA Festphasensynthese wurde auf einem *Applied Biosystems 394 DNA/RNA Synthesizer* im 1 μ mol Maßstab durchgeführt. Die TBDMS-geschützten RNA Standardamidite (Bz-A, Ac-C, dmf-G, U) und die CPG-Träger (*SynBase CPG 1000/110*, 1000 Å Porengröße) wurden von der Firma *LinkTech* bezogen. Phosphoramidite wurden in Acetonitril der Firma *Riedel de Haen* mit einem Wassergehalt < 10 ppm gelöst. Die Konzentration betrug, je nach verwendetem Zyklus 0.15 M, bzw. 0.10 M. Für alle weiteren Lösungen wurde Acetonitril (*HiPerSolv CHROMANORM*, *VWR*) in HPLC Qualität verwendet. Als Aktivator wurde BTT (0.3 M in Acetonitril, *Link Tech*) verwendet, wobei zum Teil auch mit DCI (0.25 M in Acetonitril, *Sigma Aldrich*) oder Aktivator 42 (0.25 M in Acetonitril, *Sigma Aldrich*, 5-[3,5-Bis(trifluoromethyl)phenyl]-1H-tetrazol) befriedigende Ergebnisse erzielt wurden. Dichlormethan zur Detritylierung war von technischer Qualität und wurde vor Gebrauch destilliert. Die Synthese erfolgte mit einem angepassten Syntheseprotokoll (Synthesesyklen siehe Anhang). Die Kopplungszeit betrug 10 min pro Base bzw. Sonderbase. Für die RNA Festphasensynthese wurden folgende Reagenzien verwendet:

Phosphoramidite:	0.1 M oder 0.15 M in MeCN
Aktivator:	0.30 M BTT in MeCN
Detritylierung:	3% Dichloressigsäure (v/v) in CH ₂ Cl ₂
Oxidation:	25 mM Iod in 65:30:6 (v/v) in MeCN:H ₂ O:Lutidin
Capping A:	20% Acetanhydrid in 50:50 (v/v) MeCN:Lutidin
Capping B:	20% Methylimidazol in MeCN

Nach Abschluss der Synthese wurden die Kartuschen kurz im Hochvakuum getrocknet und bei -20 °C bis zur weiteren Verwendung gelagert.

6.2.2 Entschützung, Aufreinigung und Charakterisierung

Das Trägermaterial wurde aus den Kartuschen entfernt und in *Falcon* Reaktionsgefäße (15 mL) überführt. Nach Zugabe von NH₄OH (28%, aq., 500 µL) und Methylamin (40%, aq., 500 µL) wurde für 5 min bei RT unter konstanter Agitation inkubiert und anschließend 5 min bei 65 °C und 750 rpm in einem *Eppendorf Thermocycler* geschüttelt. Die Proben wurden auf Eis gekühlt und zentrifugiert (1150 g, 5 min). Der Überstand wurde abgenommen und das Trägermaterial zwei Mal mit ddH₂O (500 µL) gewaschen. Die vereinigten Überstände wurden zunächst mittels *SpeedVac* eingeengt und schließlich *in vacuo* getrocknet. Der vollständig getrocknete Rückstand wurde in DMSO (100 µL, BioChemica, *AppliChem*) aufgenommen, mit Et₃N·3HF (125 µL, *Sigma Aldrich*) versetzt und für 90 min bei 65 °C und 800 rpm inkubiert. Es wurde NaOAc (3.0 M, aq., 25 µL) zugegeben, gründlich durchmischt, *n*BuOH (1 mL) zugegeben und für 12 h auf -80 °C gekühlt. Anschließend wurden die Proben bei 4 °C, 21130 g, 60 min zentrifugiert, der Überstand verworfen und das Pellet in *HPLC Puffer A* (1 mL) angelöst.

Die synthetisierten Oligonucleotide wurden mittels semipräparativer HPLC gereinigt. Als Puffersystem dienten *HPLC Puffer A* (0.1 M Triethylammoniumacetat in ddH₂O) und *HPLC Puffer B* (0.1 M Triethylammoniumacetat in 80/20 MeCN zu ddH₂O). Die Komposition wurde je nach Sequenz und Modifikationsgrad variiert. Typischerweise wurden Gradienten zwischen 0% → 22% *HPLC Puffer B* und 0% → 40% *HPLC Puffer B* in 45 min verwendet.

Nach erfolgter Trennung wurden die einzelnen Fraktionen mittels analytischer HPLC auf ihre Reinheit geprüft und geeignete Fraktionen vereinigt und lyophilisiert. Um Reste des HPLC Puffers zu entfernen wurden Oligonucleotide vor weiterer Verwendung mittels Sep-Pak Kartusche (C18 Classic Cartridge, *Waters*) entsalzt.

Eine Sep-Pak Kartusche wurde zunächst mit MeCN (10 mL) äquilibriert und anschließend mit ddH₂O (20 mL) gewaschen. Auf die vorbereitete Kartusche wurde das in 1 mL ddH₂O gelöste ORN langsam (< 1 mL/min) aufgetragen. Vorhandene Salze wurden durch Spülen mit ddH₂O (10 mL) entfernt und das ORN anschließend mit ddH₂O/MeCN (20:80, 6 mL) eluiert.

Die Identität und Reinheit des Endproduktes wurde durch MALDI-ToF-MS und analytische HPLC überprüft, sowie die Konzentration an einem *NanoDrop Spectrophotometer ND-1000 UV/VIS (peqlab)* bestimmt.

6.3 Chemische Synthesen

6.3.1 Material und allgemeine Methoden

Reaktionen wurden, wenn nicht anders erwähnt, ausschließlich unter Schutzgasatmosphäre (N_2 oder Ar) in mehrfach ausgeheizten Glasapparaturen durchgeführt. Bei stark oxidationsempfindlichen Substanzen wurden alle Lösemittel vor Verwendung bzw. Reaktionslösungen mehrfach entgast, wobei die „Freeze-Pump-Thaw“-Methode angewendet wurde. Lösemittel für Reaktionen wurden ohne weitere Aufreinigung verwendet und von *AppliChem*, *Sigma-Aldrich*, *Acros*, *ABCR* und *Fluka* in den Qualitätsstufen puriss, purum oder p.a. bezogen. Lösemittel für Extraktionen und säulenchromatographische Aufreinigung waren von technischer Qualität und wurden vor Gebrauch destilliert. Zur Evaporierung von Lösemitteln wurden *Laborota 4000* Rotationsverdampfer (*Heidolph*) in Kombination mit *PC2001 Vario* Vakuumpumpen (*Vacuubrand*) verwendet. Reaktionen wurden magnetisch gerührt und der Fortschritt mittels Dünnschichtchromatographie überprüft. Hierzu wurden *Merck DC Kieselgel 60 F₂₅₄* Aluminiumplatten genutzt. Zur Visualisierung diente UV Licht (254 nm und 366 nm) und/oder wässrige Lösungen von Ammoniumcermolybdat (CAM, sauer, in H_2O) Kaliumpermanganat (basisch, in H_2O), Vanillin (sauer, in Ethanol) oder Ninhydrin (sauer, in Butanol). Flash Chromatographie wurde, wenn nicht anders beschrieben, mit Kieselgel (*Geduran Si 60*, *Merck*, Partikelgröße 40-63 μm) als stationäre Phase unter einem erhöhten Druck von ca. 1,3-1,5 bar durchgeführt. Isolierte Produkte wurden im Hochvakuum (10 mbar) getrocknet. Ausbeuten beziehen sich auf chromatographisch und spektroskopisch reines Material.

1H NMR Spektren wurden auf *Varian Oxford 200* (200 MHz), *Bruker AC 300* (300 MHz), *Varian XL 400* (400 MHz), *Bruker Avance III HD* (400 MHz), *Bruker AMX 600* (600 MHz) oder *Bruker Avance III HD* (800 Mhz) Spektrometern gemessen. Chemische Verschiebungen (δ) werden in ppm, relativ zu Tetramethylsilan, mit Lösemittel als internem Standard angegeben ($CDCl_3$ bei 7.26 ppm, Aceton- d_6 bei 2.05 ppm, Methanol- d_4 bei 3.31 ppm und DMSO- d_6 bei 2.50 ppm). ^{13}C NMR Spektren wurden auf *Bruker AC 300* (75 MHz), *Varian XL 400* (100 MHz), *Bruker Avance III HD* (100 MHz), *Bruker AMX 600* (150 MHz) oder *Bruker Avance III HD* (200 Mhz) Spektrometern gemessen. Lösemittelsignale dienten als interner Standard ($CDCl_3$ bei 77.0 ppm, Aceton- d_6 bei 29.8 ppm, Methanol- d_4 bei 49.0 ppm und DMSO- d_6 bei 39.5 ppm). ^{31}P NMR Spektren wurden auf *Varian Oxford 200* (81 MHz) oder *Bruker Avance III HD* (162 MHz) Spektrometern aufgenommen, wobei H_3PO_4 (85%) als externer Standard verwendet wurde.

Signalmultiplizitäten sind als Singulett (s), Dublett (d), Triplett (t), Quartett (q), breit (br) und deren Kombinationen angegeben.

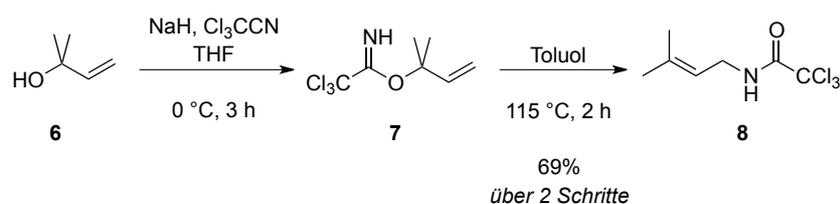
Hochauflösende Massenspektren wurden auf *Thermo Finnigan LTQ FT* (ESI) und *Finnigan MAT 95* (EI) Geräten gemessen.

IR Spektren wurden auf einem *Spectrum BX FT-IR* (*Perkin Elmer*) mit einem *DuraSamplIR II* Messkopf von *Smiths Detection* ohne weitere Präparation aufgenommen. Signale werden als schwach (w), medium (m) und stark (s) bezeichnet.

6.3.2 Durchführung der Synthesen

6.3.2.1 Synthese des Isopentenylamins

N-(3-Methyl-2-butenyl)-2',2',2'-trichloracetamid (**8**)²⁷⁵



Eine Lösung von 2-Methyl-3-buten-2-ol (48.5 mL, 464 mmol, 1.00 eq.) in THF (40 mL) wurde langsam zu einer auf 0 °C gekühlten Suspension von Natriumhydrid (2.90 g, 121 mmol, 0.26 eq.) in THF (10 mL) getropft. Die Reaktion wurde für 1 h bei 0 °C gerührt und anschließend eine Lösung von Trichloracetonitril (46.5 mL, 464 mmol, 1.00 eq.) in Et₂O (250 mL) zugegeben. Es wurde für weitere 2 h bei 0 °C gerührt und die Lösemittel *in vacuo* entfernt (330 mbar, 40 °C). Der Rückstand wurde mit Toluol (300 mL) versetzt und die braune Lösung für 2 h auf 115 °C erhitzt. Nach Evaporierung des Lösemittels wurde das Produkt durch fraktionierte Destillation (Sdp. 84 °C, 0.45 mbar) als weißer, kristalliner Feststoff erhalten (74.4 g, 318 mmol, 69%).

$R_f = 0.58$ (Hexan/Aceton 4:1)

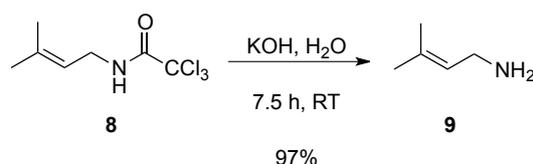
¹H NMR (300 MHz, CDCl₃) $\delta = 6.62$ (br s, 1H, NH), 5.25-5.20 (m, 1H, CH), 3.93 (t, $J = 6.0$ Hz, 2H, CH₂), 1.75 (s, 3H, CH₃), 1.71 (s, 3H, CH₃).

¹³C NMR (75 MHz, CDCl₃) $\delta = 161.7$ (CO), 138.7 (C(CH₃)₂), 118.3 (CH), 92.8 (CCl₃), 39.6 (CH₂), 25.8 (CH₃), 18.1 (CH₃).

HRMS-ESI (m/z): ber. für C₇H₁₁Cl₃NO⁺ [M+H]⁺: 229.9901. Gefunden: 229.9902.

IR (ATR): $\nu = 3292$ (m), 3050 (w), 3016 (w), 2970 (w), 2937 (w), 2914 (w), 2855 (w), 2730 (w), 2180 (w), 2024 (w), 1689 (s), 1674 (s), 1653 (w), 1527 (s), 1440 (m), 1376 (m), 1383 (w), 1376 (m), 1365 (w), 1326 (w), 1268 (m), 1254 (m), 1196 (m), 1112 (w), 1051 (m), 1025 (w), 990 (w), 902 (w), 847 (m), 825 (s), 782 (w), 726 (m).

Isopentenylamin (**9**)²⁷⁵



Verbindung **8** wurde bei 60 °C geschmolzen und bei RT tropfenweise zu einer wässrigen Lösung KOH (3 M, 700 mL) gegeben. Die resultierende Emulsion wurde für 7.5 h gerührt und mit Et₂O (4 × 300 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und das Lösemittel *in vacuo* entfernt. Das Produkt wurde als farbloses, intensiv riechendes Öl (28.3 g, 332 mmol, 97%) durch fraktionierte Destillation bei Normaldruck (Sdp. 89 °C) gewonnen.

¹H NMR (400 MHz, *d*₆-DMSO) $\delta = 5.17$ -5.14 (m, 1H, CH), 3.08-3.06 (m, 2H, CH₂), 1.62 (s, 3H, CH₃) 1.55 (s, 3H, CH₃).

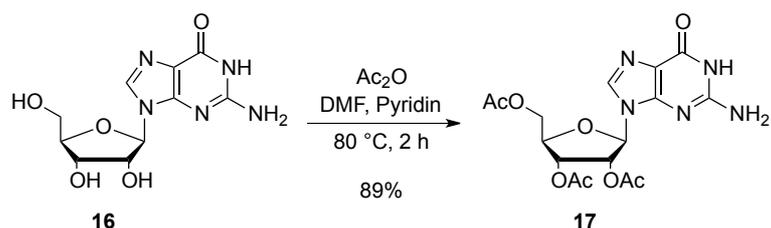
¹³C NMR (101 MHz, *d*₆-DMSO) $\delta = 130.9$ (Cq), 127.7 (CH), 79.7 (CH₂), 25.9 (CH₃), 17.9 (CH₃).

HRMS-ESI (*m/z*): ber. für C₅H₁₂N⁺ [M+H]⁺: 86.0964, gefunden 84.0965.

IR (ATR): $\nu = 3364$ (s), 3301 (s), 2971 (s), 2916 (s), 2875 (m), 2144 (w), 1644 (m), 1595 (m), 1449 (s), 1376 (m), 1340 (w), 1319 (w), 1226 (w), 1104 (w), 1005 (m), 893 (m).

6.3.2.2 Synthese des *ms*²*i*⁶A-CE Phosphoramidits

2',3',5'-Tri-*O*-acetylguanosin (**17**)²⁷⁷



Guanosin (10.0 g, 35.3 mmol, 1.00 eq.) wurde in Pyridin (20 mL) und DMF (25 mL) gelöst. Nach Zugabe von Ac₂O (43.3 g, 423 mmol, 12.0 eq.) wurde die Reaktion für 2 h auf 80 °C erwärmt. Nach Abkühlen auf Raumtemperatur und Zugabe von EtOH (23 mL) wurde die gelbe Lösung evaporiert. Der feste Rückstand wurde aus *i*PrOH (250 mL) umkristallisiert, filtriert und mit Aceton gewaschen. Das Produkt wurde als weißes Pulver (12.9 g, 31.5 mmol, 89%) erhalten.

$R_f = 0.26$ (CH₂Cl₂/MeOH 5:1)

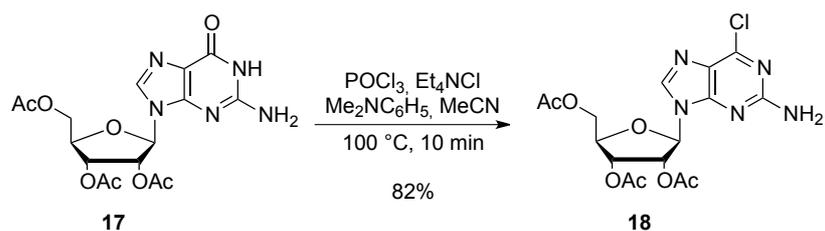
¹H NMR (400 MHz, *d*₆-DMSO) $\delta = 10.74$ (s, 1H, NH), 7.93 (s, 1H, H₈), 6.53 (br s, 2H, NH₂), 5.99 (d, $J = 6.1$ Hz, 1H, H_{1'}), 5.77 (t, $J = 6.0$ Hz, 1H, H_{2'}), 5.47 (dd, $J = 4.1$ Hz, 5.8 Hz, 1H, H_{3'}), 4.38 (dd, $J = 11.2$ Hz, 3.7 Hz, 1H, H_{5'}), 4.34-4.29 (m, 1H, H_{4'}), 4.26 (dd, $J = 11.2$ Hz, 5.6 Hz, 1H, H_{5'}), 2.11 (s, 3H, COCH₃), 2.04 (s, 3H, COCH₃), 2.04 (s, 3H, COCH₃).

¹³C NMR (100 MHz, *d*₆-DMSO) $\delta = 170.1$ (CH₃CO), 169.4 (CH₃CO), 169.3 (CH₃CO), 156.6 (C₄), 153.9 (C₂), 151.1 (C₆), 135.6 (C₈), 116.83 (C₅), 84.4 (C_{1'}), 79.5 (C_{4'}), 72.1 (C_{2'}), 70.3 (C_{3'}), 63.1 (C_{5'}), 20.5 (CH₃CO), 20.4 (CH₃CO), 20.2 (CH₃CO).

HRMS-ESI (*m/z*): ber. für C₁₆H₂₀O₈N₅⁺ [M+H]⁺: 410.1306, Gefunden: 410.1310.

IR (ATR): $\nu = 3431$ (w), 3178 (w), 2713 (w), 1742 (s), 1707 (s), 1634 (m), 1596 (m), 1538 (w), 1482 (w), 1384 (m), 1319 (w), 1190 (s), 1171 (m), 1089 (m), 1045 (m), 923 (w), 869 (w), 809 (w), 782 (m), 736 (w), 679 (m).

6-Chlor-2',3',5'-tri-*O*-acetylguanosin (**18**)²⁷⁷



Tetraethylammoniumchlorid (3.75 g, 22.7 mmol, 2.00 eq.) und Acetylguanosin **17** (4.63 g, 11.30 mmol, 1.00 eq.) wurden in MeCN (23 mL) gelöst. *N,N*-Dimethylanilin (1.44 g, 11.9 mmol) und POCl₃ (10.4 g, 67.9 mmol, 6.00 eq.) wurden zugegeben und die gelbe Lösung für 10 min. auf 100 °C erhitzt. Das Lösemittel wurde *in vacuo* entfernt und der braune Rückstand in H₂O (100 mL) suspendiert. Es wurde mit CHCl₃ (5 × 100 mL) extrahiert, die vereinigten organischen Phasen mit einer Lösung von NaHCO₃ (aq., sat., 100 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert. Der erhaltene braune Schaum wurde einer säulen-

chromatographischen Aufreinigung (CH₂Cl₂/MeOH 50:1) unterzogen, um das Produkt als weißen bis blaugelben Schaum zu erhalten (3.96 g, 9.27 mmol, 82%).

$R_f = 0.48$ (CH₂Cl₂/MeOH 20:1)

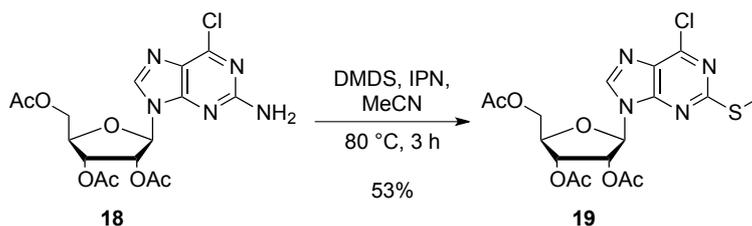
¹H NMR (600 MHz, CDCl₃) $\delta = 7.88$ (s, 1H, H₈), 6.00 (d, $J = 5.9$ Hz, 1H, H_{1'}), 5.94 (t, $J = 6.0$ Hz, 1H, H_{2'}), 5.73 (t, $J = 6.0$ Hz, 1H, H_{3'}), 5.33 (br s, 2H), 4.45-4.41 (m, 2H, H_{4'}, H_{5'}) 4.38-4.34 (m, 1H, H_{5'}), 2.13 (s, 3H, COCH₃), 2.09 (s, 3H, COCH₃), 2.07 (s, 3H, COCH₃).

¹³C NMR (150 MHz, CDCl₃) $\delta = 170.6$ (CH₃CO), 169.7 (CH₃CO), 169.5 (CH₃CO), 159.2 (C₄), 153.2 (C₂), 151.9 (C₁), 140.8 (C₈), 125.9 (C₅), 86.7 (C_{1'}), 80.1 (C_{4'}), 72.8 (C_{2'}), 70.6 (C_{3'}), 63.0 (C_{5'}), 20.8 (CH₃CO), 20.7 (CH₃CO), 20.5 (CH₃CO)

HRMS-ESI (m/z): ber. für C₁₆H₁₈O₇N₅ClNa⁺ [M+Na⁺]: 450.0787. Gefunden: 450.0790.

IR (ATR, neat): $\nu = 3468$ (w), 3320 (w), 3209 (w), 2936 (w), 2180 (w), 2024 (w), 1741 (s), 1610 (s), 1561 (s), 1512 (m), 1467 (m), 1407 (m), 1371 (m), 1190 (s), 1147 (m), 1092 (m), 1044 (s), 999 (m), 907 (s), 801 (w), 785 (m), 733 (m), 698 (w).

2-Methylthio-6-chlor-2',3',5'-tri-*O*-acetylguanosin (**19**)²⁷³



Zu einer Lösung von chloriertem Nucleosid **18** (5.17 g, 12.1 mmol, 1.00 eq.) in MeCN (98 mL) wurden Dimethyldisulfid (11.4 g, 121 mmol, 10.0 eq.) und Isopentylnitrit (7.07 g, 60.4 mmol, 5.00 eq.) gegeben. Die Reaktion wurde für 3 h auf 80 °C erwärmt, das Lösemittel *in vacuo* entfernt und der braune, ölige Rückstand in NaHCO₃ (aq., sat., 100 mL) aufgenommen. Die Mischung wurde mit CH₂Cl₂ (4 × 150 mL) extrahiert und die vereinigten organischen Phasen mit einer Lösung von NaCl (aq., sat., 80 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert. *Flash* Chromatographie (CH₂Cl₂/Aceton 100:1) lieferte das Produkt als blaugelben Schaum (2.92 g, 6.37 mmol, 53%).

$R_f = 0.70$ (CH₂Cl₂/MeOH 20:1)

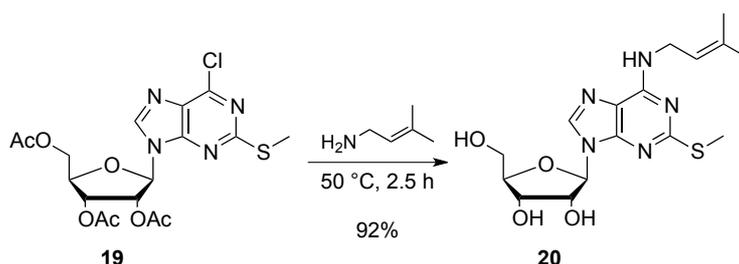
¹H NMR (600 MHz, CDCl₃) $\delta = 8.09$ (s, 1H, C₈), 6.11 (d, $J = 6.0$ Hz, 1H, H_{1'}), 5.98 (t, $J = 6.0$ Hz, 1H, H_{2'}), 5.63 (t, $J = 6.0$ Hz, 1H, H_{3'}), 4.44-4.41 (m, 2H, H_{4'}, H_{5'}) 4.32-4.29 (m, 1H, C_{5'}), 2.63 (s, 3H, SCH₃), 2.14 (s, 3H, COCH₃), 2.09 (s, 3H, COCH₃), 2.08 (s, 3H, COCH₃).

^{13}C NMR (150 MHz, CDCl_3) δ = 170.4 (CH_3CO), 169.6 (CH_3CO), 169.4 (CH_3CO), 167.5 (C_2), 152.0 (C_6), 151.3 (C_4), 142.3 (C_8), 129.2 (C_5), 87.2 ($\text{C}_{1'}$), 80.1 ($\text{C}_{4'}$), 73.0 ($\text{C}_{2'}$), 70.2 ($\text{C}_{3'}$), 62.8 ($\text{C}_{5'}$), 20.8 (CH_3CO), 20.6 (CH_3CO), 20.5 (CH_3CO), 15.0 (SCH_3).

HRMS-ESI (m/z): ber. für $\text{C}_{17}\text{H}_{20}\text{O}_7\text{N}_4\text{ClS}^+$ [$\text{M}+\text{H}^+$]: 459.0736. Gefunden: 459.0742.

IR (ATR, neat): ν = 3473 (w), 3114 (w), 2933 (w), 2349 (w), 2024 (w), 1741 (s), 1594 (m), 1547 (s), 1495 (w), 1428 (w), 1361 (s), 1325 (m), 1190 (s), 1173 (m), 1136 (m), 1090 (m), 1044 (s), 971 (m), 952 (m), 901 (m), 871 (m), 821 (w), 784 (w), 767 (w), 741 (w), 683 (w).

2-Methylthio- N^6 -isopentenyladenosin



19 (2.76 g, 6.02 mmol, 1.00 eq.) wurde in Isopentenylamin (10 mL, 93.9 mmol, 15.6 eq.) gelöst und für 2.5 h in einem Druckrohr auf 50 °C erhitzt. Die Lösung wurde *in vacuo* getrocknet und der Rückstand mit CH_2Cl_2 (2×100 mL) und Benzol (2×50 mL) coevaporiert um ein gelbes Pulver zu erhalten. Umkristallisation aus zunächst MeCN (15 mL) und schließlich H_2O (15 mL) gab das Produkt als grauweißes Pulver (2.12 g, 5.56 mmol, 92%).

R_f = 0.13 ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 10:1)

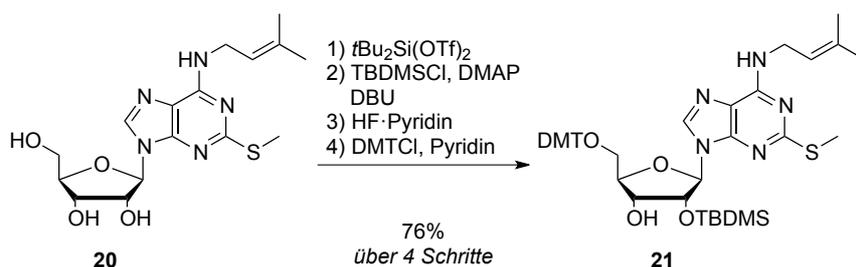
^1H NMR (400 MHz, d_6 -DMSO) δ = 8.21 (s, 1H, H_8), 7.97 (br s, 1H, NH), 5.83 (d, J = 4.0 Hz, 1H, $\text{H}_{1'}$), 5.40 (d, J = 4.0 Hz, 1H, $\text{C}_2'\text{-OH}$), 5.28 (td, J = 6.7 Hz, 3.4 Hz, 1H, CH_2CHC), 5.16 (d, J = 4.0 Hz, 1H, $\text{C}_3'\text{-OH}$), 5.03 (t, J = 4.0 Hz, 1H, $\text{C}_5'\text{-OH}$), 4.59 (q, J = 4.59 Hz, 1H, H_2'), 4.14 (dd, J = 4.0 Hz, 8.0 Hz, 1H, H_3'), 4.04 (br s, 2H, NHCH_2CH), 3.92 (q, J = 4.0 Hz, 1H, H_4'), 3.69-3.59 (m, 1H, H_5'), 3.53 (ddd, J = 11.9 Hz, 6.1 Hz, 4.2 Hz, 1H, H_5'), 2.48 (s, 3H, SCH_3), 1.70 (s, 3H, $\text{C}(\text{CH}_3)_2$), 1.67 (s, 3H, $\text{C}(\text{CH}_3)_2$).

^{13}C NMR (101 MHz, d_6 -DMSO) δ = 164.1 (C_2), 153.5 (C_4), 149.4 (C_6), 138.5 (C_8), 133.3 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 121.9 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 117.2 (C_5), 87.3 ($\text{C}_{1'}$), 85.5 ($\text{C}_{4'}$), 73.3 ($\text{C}_{2'}$), 70.5 ($\text{C}_{3'}$), 61.6 ($\text{C}_{5'}$), 37.6 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 25.4 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 17.9 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 13.8 (SCH_3).

HRMS-ESI (m/z): ber. für $\text{C}_{16}\text{H}_{23}\text{N}_5\text{O}_4\text{SNa}$ [$\text{M}+\text{Na}^+$]: 404.1363. Gefunden: 404.1368.

IR (ATR, neat): $\nu = 3409$ (s), 3254 (s), 2953 (m), 2926 (m), 2219 (vw), 2106 (vw), 2024 (w), 1690 (s), 1578 (s), 1532 (w), 1478 (m), 1448 (m), 1422 (m), 1378 (w), 1361 (w), 1345 (m), 1313 (m), 1297 (s), 1282 (m), 1253 (s), 1228 (s), 1204 (s), 1160 (w), 1132 (m), 1119 (w), 1079 (s), 1043 (m), 987 (w), 977 (w), 957 (w), 928 (w), 900 (s), 863 (m), 855 (m), 848 (m), 836 (m), 798 (w), 785 (s), 777 (m), 758 (s), 707 (m), 692 (s), 679 (s), 673 (s).

5'-O-(4,4'-Dimethoxytrityl)-2'-O-(tert-Butyldimethylsilyl)-2-methylthio-*N*⁶-isopentenyladenosin²⁸⁰



$\text{ms}^{2:6}\text{A}$ (**20**, 448 mg, 1.17 mmol, 1.00 eq.) wurde unter leichtem Erwärmen in DMF (4 mL) gelöst und die Lösung anschließend auf 0 °C gekühlt. Di-*tert*-butylsilylbis(trifluormethansulfonat) (0.42 mL, 1.29 mmol, 1.10 eq.) wurde tropfenweise zugegeben und die Lösung für 3 h im Eisbad gerührt. Das Reaktionsgemisch wurde mit EtOAc (200 mL) und einer NaHCO_3 -Lösung (aq., sat., 5 mL) verdünnt, die Phasen getrennt, die organische Phase mit NaHCO_3 (aq., semi-sat., 4 × 50 mL) und NaCl (aq., sat. 50 mL) gewaschen, anschließend über Na_2SO_4 getrocknet und evaporiert.

Der resultierende weiße Schaum wurde in MeCN (6 mL) gelöst und mit DBU (0.35 mL, 2.35 mmol, 2.00 eq.), TBDMSCl (265 mg, 1.76 mmol, 1.50 eq.) und DMAP (14 mg, 117 μmol , 0.10 eq.) versetzt. Die Reaktion wurde für 15 h bei Raumtemperatur gerührt und erneut DBU (0.35 mL, 2.35 mmol, 2.00 eq.) und TBDMSCl (221 mg, 1.47 mmol, 1.25 eq.) zugegeben. Nach weiteren 4 h wurde MeOH (1 mL) zugegeben und das Lösemittel *in vacuo* entfernt. Der Rückstand wurde in EtOAc (150 mL) aufgenommen, mit NaHCO_3 -Lösung (aq., semi-sat. 2 × 40 mL) und NaCl (aq., sat., 40 mL) gewaschen, über Na_2SO_4 getrocknet und evaporiert.

Der Rückstand wurde in CH_2Cl_2 (6 mL) und Pyridin (0.7 mL) gelöst, die Lösung auf 0 °C gekühlt und tropfenweise mit HF·Pyridin (135 μL , 5.20 mmol, 4.44 eq.) versetzt. Nach 50 min. wurde zusätzliches HF·Pyridin (300 μL , 11.55 mmol, 9.88 eq.) zugegeben und die Reaktion für weitere 15 min. gerührt. Nach Zugabe von Methoxytrimethylsilan (1.62 mL, 20.6 mmol, 17.6 eq.) wurde die Lösung auf Raumtemperatur erwärmt und die Lösemittel *in*

vacuo entfernt. Das erhaltene Rohprodukt wurde in CH₂Cl₂ (100 mL) gelöst und mit NaHCO₃ (aq., sat., 2 × 40 mL) gewaschen. Die vereinigten wässrigen Phasen wurden mit CH₂Cl₂ (50 mL) extrahiert und die vereinigten organischen Phasen mit einer Lösung von NaCl (aq., sat., 40 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert.

Der zurückbleibende gelbe Schaum wurde in Pyridin (4.5 mL) gelöst, die Lösung auf 0 °C gekühlt und DMTCl (438 mg, 1,29 mmol, 1.10 eq.) zugegeben. Nach einer Reaktionszeit von 5.5 h wurde nochmals DMTCl (40 mg, 118 μmol 0.10 eq.) zugegeben und weitere 75 min. gerührt. Methanol (1 mL) wurde zugegeben, die Reaktion auf Raumtemperatur gewärmt und das Lösemittel *in vacuo* entfernt. Der Rückstand wurde in CH₂Cl₂ (150 mL) gelöst mit NaHCO₃ (aq., sat., 50 mL) und NaCl (aq., sat., 50 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert.

Das Rohprodukt wurde einer säulenchromatographischen Aufreinigung (Säule mit CH₂Cl₂ + 0.5% Pyridin gepackt und beladen, Elution mit Hexan/Aceton 10:1 + 0.5% Pyridin bis 8:1 + 0.5% Pyridin) unterzogen um das gewünschte Nucleosid als weißen Schaum zu erhalten (716 mg, 897 μmol, 76%).

$R_f = 0.46$ (Hexan/Aceton 2:1)

¹H NMR (600 MHz, CDCl₃) δ = 7.81 (s, 1H, H₈), 7.44-7.42 (m, 2H, H_{ortho}Ph), 7.33-7.30 (m, 4H, H_{ortho}PhOMe), 7.28-7.25 (m, 2H, H_{meta}Ph), 7.22-7.19 (m, 1H, H_{para}Ph), 6.82-6.79 (m, 4H, H_{meta}PhOMe), 5.97 (d, $J = 4.8$ Hz, 1H, H_{1'}), 5.54 (br s, 1H, NH), 5.37-5.33 (m, 1H, CH₂CHC(CH₃)₂), 4.91 (t, $J = 5.1$ Hz, 1H, H_{2'}), 4.34 (q, $J = 4.7$ Hz, 1H, H_{3'}), 4.21 (q, $J = 4.0$ Hz, 1H, H_{4'}), 4.19 (br s, 2H, CH₂CHC(CH₃)₂), 3.78 (s, 6H, PhOCH₃), 3.46 (dd, $J = 10.5$ Hz, 3.4 Hz, 1H, H_{5'}), 3.41 (dd, $J = 10.5$ Hz, 4.4 Hz, 1H, H_{5'}), 2.71 (d, $J = 4.8$ Hz, 1H, C₃'OH) 2.45 (s, 3H, SCH₃), 1.77-1.73 (m, 6H, CH₂CHC(CH₃)₂), 0.86 (s, 9H, Si(CH₃)₂C(CH₃)₃), 0.00 (s, 3H, Si(CH₃)₂C(CH₃)₃), -0.10 (s, 3H, Si(CH₃)₂C(CH₃)₃).

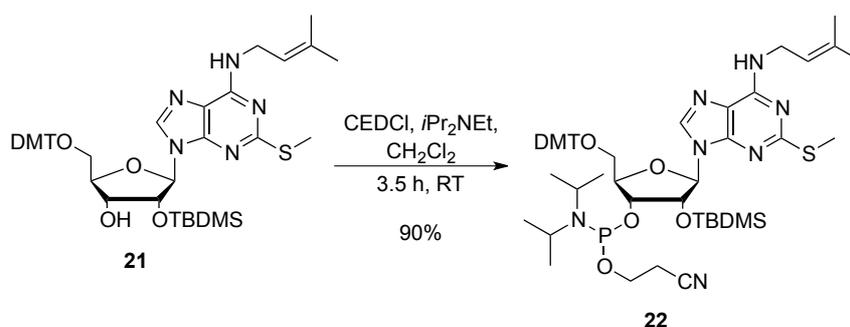
¹³C NMR (150 MHz, CDCl₃) δ = 166.2 (C₂), 158.7 (C₄), 153.9 (C₆), 144.7 (C_{ipso}Ph), 137.4 (C₈), 136.9 (CH₂CHC(CH₃)₂), 135.8 (C_{ipso}PhOMe), 135.8 (C_{ipso}PhOMe), 130.2 (C_{ortho}PhOMe), 128.3 (C_{ortho}Ph), 128.0 (C_{meta}Ph), 127.0 (C_{para}Ph), 120.4 (CH₂CHC(CH₃)₂), 117.8 (C₅), 113.3 (C_{meta}PhOMe), 88.3 (C_{1'}), 86.7 (CPh(PhOMe)₂), 83.9 (C_{4'}), 75.7 (C_{2'}), 71.7 (C_{3'}), 63.8 (C_{5'}), 55.3 (OCH₃), 38.6 (CH₂CHC(CH₃)₂) 25.8 (CH₂CHC(CH₃)₂), 25.8 (Si(CH₃)₂C(CH₃)₃), 18.2 (CH₂CHC(CH₃)₂), 18.1 (Si(CH₃)₂C(CH₃)₃), 14.6 (SCH₃), -4.7 (Si(CH₃)₂C(CH₃)₃), -5.1 (Si(CH₃)₂C(CH₃)₃).

HRMS-ESI (m/z): ber. für: C₄₃H₅₅N₅O₆SSiNa [M+Na⁺]: 820.3535. Gefunden: 820.3543.

IR (ATR, neat): ν = 3328 (w), 2928 (m), 2856 (w), 2050 (w), 1608 (s), 1582 (s), 1509 (s), 1462 (m), 1362 (w), 1328 (m), 1300 (s), 1250 (s), 1222 (s), 1176 (s), 1118 (m), 1083 (m),

1065 (m), 1035 (s), 1005 (m), 966 (w), 908 (w) 870 (w), 850 (s), 781 (s), 754 (m), 726 (w), 702 (m), 682 (w).

5'-O-(4,4'-Dimethoxytrityl)-2'-O-(*tert*-Butyldimethylsilyl)-2-methylthio-*N*⁶-isopentenyladenosin 3'-(*N,N*-diisopropyl-2-cyanoethyl)phosphoramidit (22**)**



Eine Lösung von **21** (100 mg, 128 μmol , 1.00 eq.) in CH_2Cl_2 (0.53 mL) wurde zweifach entgast, $i\text{Pr}_2\text{NEt}$ (89 μL , 511 μmol , 4.00 eq.) zugegeben, nochmals entgast, CEDCl (114 μL , 511 μmol , 4.00 eq.) zugegeben und zwei weitere Male entgast. Nach 3.5 h Reaktionszeit bei Raumtemperatur wurde die Reaktionsmischung direkt auf eine Silicasäule geladen und per *Flash* Chromatographie aufgereinigt (Hexan/EtOAc 1:2 + 0.5% Pyridin). Das oxidationsempfindliche Produkt wurde als weißer Schaum erhalten (183 mg, 183 μmol , 90%, 1:0.9 Mischung aus zwei Diastereomeren, bestimmt via NMR) und ist unter Luftausschluss einige Wochen lagerfähig.

$R_f = 0.46$ (Hexan/Aceton 2:1)

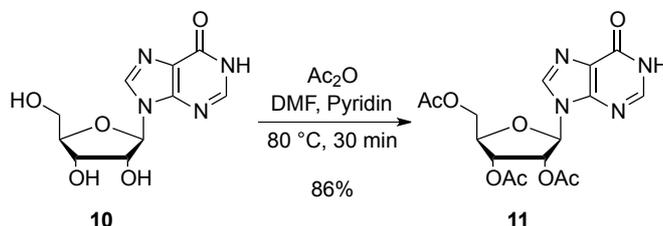
¹H NMR (400 MHz, CDCl_3) $\delta = [7.88$ (s) und 7.85 (s), 1H], 7.48-7.45 (m, 2H), 7.37-7.18 (m, 7H), 6.85-6.81 (m, 4H), [6.04 (d, $J = 6.2$ Hz) und 6.00 (d, $J = 5.9$ Hz), 1H], 5.66 (br s, 1H), 5.37-5.34 (m, 1H), 4.99-4.94 (m, 1H), 4.43-4.36 (m, 2H), 4.20 (br s, 2H), 4.02-3.87 (m, 2H) 3.80-3.78 (m, 6H), 3.72-4.47 (m, 3H), 3.38-3.34 (m, 1H), 2.73-2.56 (m, 2H), [2.46 (s) und 2.44 (s), 3H], [1.75 (s) und 1.74 (s), 6H], 1.24-1.18 (m, 12H), [0.82 (s), und 0.81 (s), 9H], [0.00 (s) und -0.02 (s), 3H], -0.15 (s, 3H).

³¹P NMR (162 MHz, CDCl_3) $\delta = 150.7, 149.0$ ppm.

HRMS-ESI (m/z): ber. für: $\text{C}_{52}\text{H}_{73}\text{N}_7\text{O}_7\text{PSSi}^+$ [$\text{M}+\text{H}$]⁺: 998.4794. Gefunden: 998.4802.

6.3.2.3 Synthese des N⁶-Isopentenyladenosin-CE Phosphosphoramidits

2',3',5'-Tri-*O*-acetylinosin (11)²⁷⁷



Inosin (10.0 g, 37.3 mmol, 1.00 eq.) wurde in einer Lösung aus Pyridin (10 mL) und DMF (14 mL) suspendiert, mit Ac₂O (19.4 mL, 0.21 mol, 5.50 eq.) versetzt und die Mischung für 30 min. auf 80 °C erhitzt. Die Lösung wurde evaporiert, der Rückstand aus *i*PrOH (80 mL) umkristallisiert, mit Aceton gewaschen und im Hochvakuum getrocknet. Das Produkt wurde als weißes Pulver erhalten (12.7 g, 86 %).

$R_f = 0.73$ (CH₂Cl₂/MeOH 5:1)

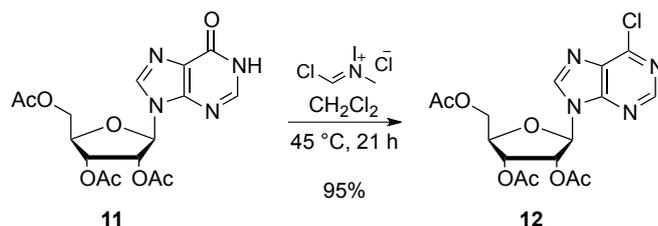
¹H-NMR (400 MHz, *d*₆-DMSO) $\delta = 12.48$ (br s, 1H, NH), 8.31 (s, 1H, H₈), 8.10 (s, 1H, H₂), 6.19 (d, $3J = 4.0$ Hz, 1H, H_{1'}), 5.91 (t, $3J = 4.0$ Hz, 1H, H_{2'}), 5.55 (dd, $3J = 4.0$ Hz, 1H, H_{3'}), 4.41 - 4.35 (m, 2H, H_{4'}, H_{5'}), 4.27 - 4.22 (m, 1H, H_{5'}), 2.11 (s, 3H, COCH₃), 2.04 (s, 3H, COCH₃), 2.02 (s, 3H, COCH₃).

¹³C-NMR (100 MHz, *d*₆-DMSO) $\delta = 170.0$ (CH₃CO), 169.4 (CH₃CO), 169.2 (CH₃CO), 156.4 (C₆), 147.9 (C₄), 146.3 (C₂), 139.3 (C₈), 124.8 (C₅), 85.6 (C_{1'}), 79.5 (C_{4'}), 72.2 (C_{2'}), 69.9 (C_{3'}), 62.8 (C_{5'}), 20.5 (CH₃CO), 20.3 (CH₃CO), 20.2 (CH₃CO).

HRMS-ESI (*m/z*): ber. für: C₁₆H₁₉N₄O₈⁺ [M+H]⁺: 395.1198. Gefunden: 395.1200.

IR (ATR, neat): $\nu = 3125$ (w), 3047 (w), 2993 (w), 2943 (w), 2830 (w), 1743 (s), 1702 (s), 1588 (m), 1553 (m), 1511 (w), 1475 (w), 1418 (m), 1382 (m), 1373 (m), 1342 (m), 1323 (w), 1299 (w), 1209 (s), 1224 (s), 1157 (m), 1113 (m), 1096 (m), 1066 (m), 1055 (m), 1016 (m), 963 (w), 908 (m), 859 (m), 801 (m), 788 (m), 762 (w), 723 (m), 706 (m), 693 (w), 666 (w).

6-Chlor-(2',3'5'-tri-*O*-acetyl)inosin (**12**)²⁷⁸



(Chlormethylen)dimethyliminiumchlorid (3.26 g, 25.43 mmol, 2.00 eq) wurde zu einer Lösung von Acetylinosin **11** (5.00 g, 12.71 mmol, 1.00 eq) in CH₂Cl₂ (30 mL) gegeben und die Reaktion für 21 h auf 45 °C erhitzt. Nach Zugabe einer Lösung von NaHCO₃ (aq, sat, 200 mL) wurde die organische Phase abgetrennt und die wässrige mit EtOAc (1 × 200 mL, 2 × 100 mL) extrahiert. Die vereinigten organischen Phasen wurden mit NaCl gewaschen (aq., sat., 50 mL), über MgSO₄ getrocknet und evaporiert. *Flash* Chromatographie (Kieselgel, CH₂Cl₂/MeOH 40:1) lieferte das Produkt als blassgelben Schaum (5.00 g, 12.11 mmol, 95%) $R_f = 0.75$ (CH₂Cl₂/MeOH 50:1)

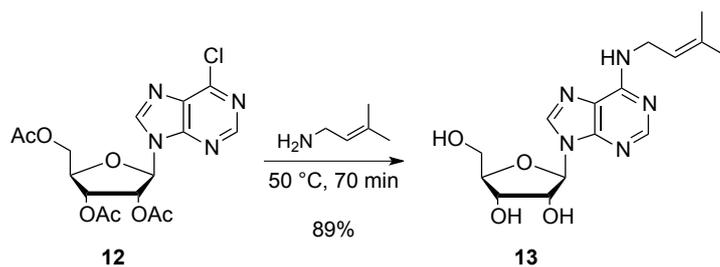
¹H-NMR (300 MHz, CDCl₃) $\delta = 8.76$ (s, 1H, H₂), 8.29 (s, 1H, H₈), 6.22 (d, $J = 5.1$ Hz, 1H, H_{1'}), 5.94 (t, $J = 5.4$ Hz, 1H, H_{2'}), 5.63 (t, $J = 5.4$ Hz, 1H, H_{3'}), 4.49 – 4.34 (m, 3H, H_{4'}, H_{5'}), 2.14 (s, 3H, CH₃COO_{3'}), 2.14 (s, 3H, CH₃COO_{5'}), 2.07 (s, 3H, CH₃COO_{2'}).

¹³C-NMR (75 MHz, CDCl₃) $\delta = 170.3$ (CH₃COO_{5'}), 169.6 (CH₃COO_{3'}), 169.4 (CH₃COO_{2'}), 152.4 (C₂), 151.8 (C₅), 151.4 (C₄), 143.7 (C₈), 132.5 (C₆), 87.0 (C_{1'}), 80.7 (C_{4'}), 73.2 (C_{2'}), 70.6 (C_{3'}), 63.0 (C_{5'}), 20.9 (CH₃CO), 20.6 (CH₃CO), 20.5 (CH₃CO).

HRMS-ESI (m/z): ber. für: C₁₆H₁₇ClN₄O₇Na⁺ [M+Na]⁺: 435.0678. Gefunden 435.0678.

IR (ATR): $\nu = 2948$ (w), 1740 (s), 1591 (m), 1561 (m), 1492 (w), 1437 (w), 1425 (m), 1404 (w), 1370 (m), 1339 (m), 1210 (s), 1198 (s), 1163 (m), 1146 (m), 1092 (m), 1044 (s), 958 (m), 938 (m), 901 (m), 860 (m), 816 (m), 792 (m), 734 (m), 701 (w).

*N*⁶-Isopentenyladenosin (**13**)



Eine Lösung von Inosinderivat **12** (5.00 g, 11.9 mmol, 1.00 eq.) in Isopentenylamin **9** (10.3 mL, 95.5 mmol, 8.00 eq.) wurde in einem Druckrohr für 70 min. auf 50 °C erhitzt. Das Reaktionsgemisch wurde mit CH₂Cl₂ coevaporiert, bis ein bräunlicher Feststoff zurückblieb. Dieser wurde in ammoniakalischem MeOH (7 N NH₃, 75 mL) aufgenommen und für 8 h bei Raumtemperatur gerührt. Das Lösemittel wurde *in vacuo* entfernt und der Rückstand zunächst aus CH₂Cl₂ (100 mL) und dann aus MeOH (100 mL) umkristallisiert um das Produkt als weißes Pulver zu erhalten (3.56 g, 10.6 mmol, 89%).

$R_f = 0.56$ (CH₂Cl₂/MeOH 10:1)

¹H NMR (400 MHz, *d*₄-Methanol) $\delta = 8.24$ (s, 1H, H₈), 8.22 (br s, 1H, H₂), 5.95 (d, $J = 6.4$ Hz, 1H, H_{1'}), 5.42-5.35 (m, 1H, CH₂CHC(CH₃)₂), 4.74 (dd, $J = 6.4$ Hz, 5.2 Hz, 1H, H_{2'}), 4.32 (dd, $J = 5.1$ Hz, 2.5 Hz, 1H, H_{3'}), 4.18 (br s, 2H, CH₂CHC(CH₃)₂) 4.17 (q, $J = 2.5$ Hz, 1H, H_{4'}), 3.89 (dd, $J = 12.6$ Hz, 2.5 Hz, 1H, H_{5'}), 3.74 (dd, $J = 12.5$ Hz, 2.6 Hz, 1H, H_{5'}), 1.77 (s, 6H, CH₂CHC(CH₃)₂).

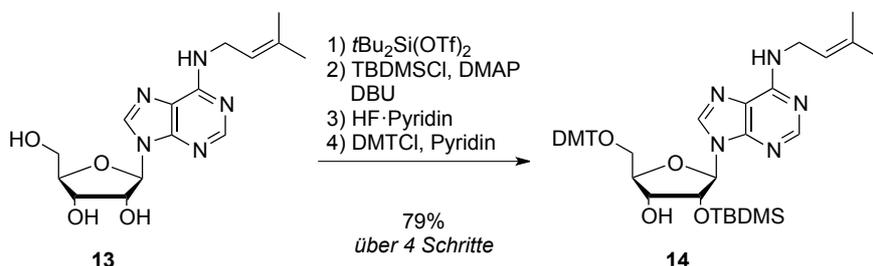
¹³C NMR (101 MHz, *d*₄-Methanol) $\delta = 156.0$ (C₆), 153.5 (C₂), 149.0 (C₄), 141.4 (C₈), 137.2 (CH₂CHC(CH₃)₂), 121.5 (2×C, CH₂CHC(CH₃)₂ und C₅), 101.4, 91.3 (C_{1'}), 88.2 (C_{4'}), 75.5 (C_{2'}), 72.7 (C_{3'}), 67.9, 63.5 (C_{5'}), 39.5 (CH₂CHC(CH₃)₂), 25.8 (CH₂CHC(CH₃)₂), 18.0 (CH₂CHC(CH₃)₂).

HRMS-ESI (*m/z*): ber. für: C₁₅H₂₂N₅O₄⁺ [M+H]⁺: 336.1667. Gefunden: 336.1664.

IR (ATR, neat): $\nu = 3145$ (w), 2914 (w), 1644 (m), 1623 (s), 1586 (m), 1535 (w), 1476 (w), 1424 (w), 1377 (m), 1336 (m), 1308 (m), 1292 (m), 1218 (m), 1183 (w), 1140 (m), 1123 (m), 1096 (m), 1078 (m), 1057 (s), 1029 (m), 985 (m), 946 (w), 989 (w), 866 (m), 823 (m), 792 (m), 764 (m), 732 (m).

5'-O-(4,4'-Dimethoxytrityl)-N⁶-isopentenyl-2'-O-(tert-butylidimethylsilyl)
(14)²⁸⁰

adenosin



N⁶-Isopentenyladenosin (500 mg, 1.49 mmol, 1.00 eq.) wurde in DMF (5 mL) gelöst, die Lösung auf 0 °C gekühlt und tropfenweise Di-*tert*-butyl-silylbis(trifluormethansulfonat) (0.54 mL, 1.64 mmol, 1.10 eq.) zugegeben. Die Reaktion wurde für 90 min. bei 0 °C gerührt und schließlich durch Zugabe von NaHCO₃ (aq., sat., 9 mL) beendet, wobei sich ein Niederschlag bildete. Diese Mischung wurde in EtOAc (200 mL) aufgenommen, die Phasen getrennt und die organische Phase mit NaHCO₃ (aq., semi-sat., 4 × 50 mL) und NaCl (50 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert.

Der Rückstand wurde in MeCN (6 mL) gelöst und mit DBU (0.89 mL, 5.96 mmol, 4.00 eq.), TBDMSCl (337 mg, 2.24 mmol, 1.50 eq.) und DMAP (27 mg, 221 μmol, 0.15 eq.) versetzt. Die Reaktion wurde für 15 h bei Raumtemperatur gerührt, durch Zugabe von MeOH (5 mL) beendet und die Lösemittel *in vacuo* entfernt. Der resultierende Feststoff wurde in EtOAc (200 mL) aufgenommen und mit NaHCO₃ (aq., sat., 40 mL) und NaCl (aq., sat., 40 mL) gewaschen. Die vereinigten wässrigen Phasen wurde mit EtOAc (2 × 50 mL) extrahiert, die vereinigten organischen Phasen über Na₂SO₄ getrocknet und das Lösemittel am Rotationsverdampfer entfernt.

Das Rohprodukt wurde in ein Polypropylengefäß überführt, in CH₂Cl₂ (7 mL) und Pyridin (1 mL) gelöst und auf 0 °C gekühlt. Es wurde tropfenweise HF·Pyridin (0.35 mL, 13.5 mmol, 9.04 eq.) zugegeben und die Lösung für 1 h bei 0 °C gerührt. Nach Zugabe von Methoxytrimethylsilan (1.5 mL, 10.9 mmol, 7.30 eq.) wurde auf Raumtemperatur erwärmt und EtOAc (200 mL) zugegeben. Die Lösung wurde mit NaHCO₃ (aq., sat. 50 mL) und NaCl (aq., sat., 50 mL) gewaschen, über Na₂SO₄ getrocknet und evaporiert.

Der Rückstand wurde in Pyridin (6 mL) gelöst, die Lösung auf 0 °C gekühlt und DMTCl (609 mg, 1.79 mmol, 1.20 eq.) zugegeben. Nach 5.5 h wurde MeOH (1 mL) zugegeben und die Mischung auf Raumtemperatur erwärmt. Das Lösemittel wurde *in vacuo* entfernt und das Produkt durch säulenchromatographische Aufreinigung (Säule mit CH₂Cl₂ + 0.5% Pyridin

gepackt und beladen, Elution mit Hexan/Aceton 10:1 + 0.5% Pyridin bis 7:1 + 0.5% Pyridin) als weißer Schaum erhalten (885, 1.18 mmol, 79%).

$R_f = 0.56$ (hexane/acetone 2:1)

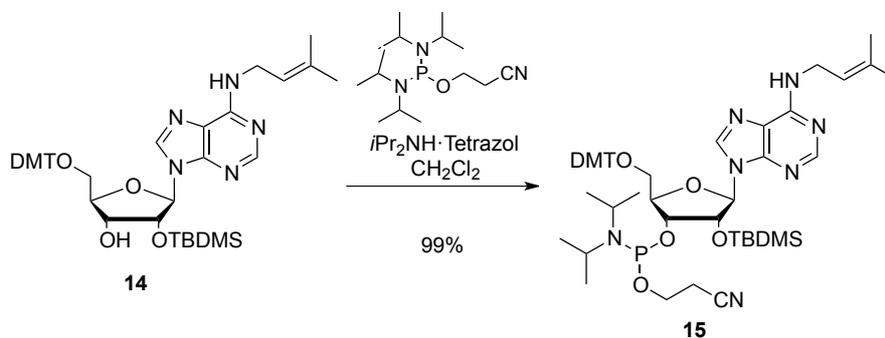
$^1\text{H NMR}$ (400 MHz, CDCl_3) $\delta = 8.31$ (s, 1H, H_2), 7.94 (s, 1H, H_8), 7.46-7.44 (m, 2H, $\text{H}_{\text{orthoPh}}$), 7.36-7.32 (m, 4H, $\text{H}_{\text{orthoPhOMe}}$) 7.29-7.25 (m, 2H, H_{metaPh}), 7.23-7.19 (m, 1H, H_{paraPh}), 6.83-6.79 (m, 4H, $\text{H}_{\text{metaPhOMe}}$), 6.01 (d, $J = 2.0$ Hz, 1H, $\text{H}_{1'}$), 5.56 (s, 1H, NH), 5.40-5.36 (m, 1H, $\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 5.00 (t, $J = 4.0$ Hz, 1H, H_2), 4.33 (q, $J = 2.0$ Hz, 1H, H_3), 4.25 (q, 1H, $J = 3.6$ Hz, H_4), 4.22 (br s, 2H, $\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 3.78 (s, 3H, PhOCH_3), 3.78 (s, 3H, PhOCH_3), 3.52 (dd, $J = 4.0$ Hz, 1H, H_5), 3.37 (dd, $J = 4.0$ Hz, 1H, H_5), 2.72 (d, $J = 4.0$ Hz, 1H, C_3OH), 1.76 (s, 3H, $\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 1.74 (s, 3H, $\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 0.84 (s, 9H, $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), -0.01 (s, 3H, $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), -0.15 (s, 3H, $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$).

$^{13}\text{C NMR}$ (100 MHz, CDCl_3) $\delta = 158.7$ ($\text{C}_{\text{paraPhOMe}}$), 154.7 (C_6), 153.4 (C_2), 149.0 (C_4), 144.8 (C_{ipsoPh}), 138.6 (C_8), 137.1 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 135.8 ($\text{C}_{\text{ipsoPhOMe}}$), 135.8 ($\text{C}_{\text{ipsoPhOMe}}$), 130.2 ($\text{C}_{\text{orthoPhOMe}}$), 128.3 ($\text{C}_{\text{orthoPh}}$), 128.0 (C_{metaPh}), 127.0 (C_{paraPh}), 120.3 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 120.2 (C_5), 113.3 ($\text{C}_{\text{metaPhOMe}}$), 88.3 ($\text{C}_{1'}$), 86.7 ($\text{CPh}(\text{PhOMe})_3$), 84.2 (C_4'), 75.7 (C_2'), 71.7 (C_3'), 63.6 (C_5'), 55.3 (PhOCH_3), 35.6 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 32.8 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), 25.8 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), 25.7 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), 18.1 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), 18.0 ($\text{CH}_2\text{CHC}(\text{CH}_3)_2$), -4.8 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), -5.1 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$).

HRMS-ESI (m/z): ber. für: $\text{C}_{42}\text{H}_{54}\text{N}_5\text{O}_6\text{Si}^+$ $[\text{M}+\text{H}]^+$: 752.3838. Gefunden: 752.3832.

IR (ATR, neat): $\nu = 3321$ (w), 2929 (w), 2857 (w), 1672 (w), 1613 (s), 1580 (m), 1508 (m), 1464 (m), 1445 (m), 1410 (w), 1376 (w), 1360 (w), 1328 (m), 1297 (m), 1249 (s), 1223 (m), 1175 (m), 1117 (m), 1064 (m), 1034 (s), 912 (m), 865 (m), 865 (m), 832 (s), 780 (s), 754 (m), 726 (m), 701 (m).

5'-O-(4,4'-Dimethoxytrityl)-N⁶-isopentenyl-2'-O-(tert-butyl dimethylsilyl) adenosin 3'-(N,N-diisopropyl-2-cyanoethyl)phosphoramidit (15)



Eine Lösung von **14** (96 mg, 128 μmol , 1.00 eq.) in CH_2Cl_2 (1 mL) wurde zweifach entgast, Diisopropylammonium Tetrazolid (13 mg, 76 mmol, 0.60 eq.) zugegeben, nochmals entgast, mit 2-Cyanoethoxy-bis-(N,N-diisopropylamino)phosphin (47 μL , 166 mmol, 1.15 eq.) und ein letztes Mal entgast. Die Reaktion wurde für 20 h bei Raumtemperatur gerührt und direkt auf eine Silicasäule (Hexan/EtOAc 1:2 + 0.5% Pyridin) geladen, um das gewünschte Phosphoramidit (122 mg, 127 μmol , 99%) als weißen Schaum und Mischung aus zwei Diastereomeren zu erhalten.

$R_f = 0.68$ (Hexan/Aceton 2:1)

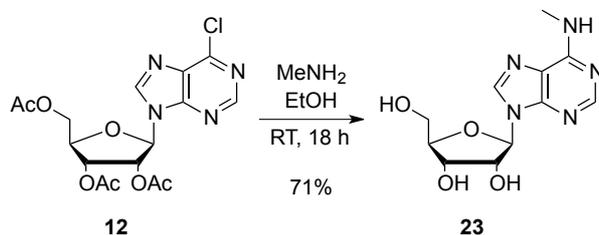
¹H NMR (400 MHz, CDCl_3) $\delta = [8.33$ (s) und 8.30 (s), 1H], $[8.00$ (s) und 7.97 (s), 1H], 7.52-7.48 (m, 2H), 7.41-7.37 (m, 4H), 7.30-7.23 (m, 3H), 6.85-6.82 (m, 4H), $[6.03$ (d, $J = 6.5$ Hz) und 5.99 (d, $J = 6.0$ Hz), 1H], 5.71-5.66 (m, 1H), 5.44-5.38 (m, 1H), 5.15-5.11 (m, 1H), 4.45-4.20 (m, 4H), 4.17-3.87 (m, 1H), 3.80 (s, 6H), 3.70-3.57 (m, 4H), 3.43-3.31 (m, 1H), $[2.70$ -2.65 (m) und 2.35 -2.32 (m), 2H] $[1.78$ (s) und 1.76 (s), 6H], 1.23-1.08 (m, 12 H), 0.81 (s, 9H), $[-0.02$ (s) und -0.02 (s), 3H], $[-0.15$ (s) und -0.17 (s), 3H].

³¹P NMR (162 MHz, CDCl_3) $\delta = 150.9, 148.9$ ppm.

HRMS-ESI (m/z): ber. für: $\text{C}_{51}\text{H}_{71}\text{N}_7\text{O}_7\text{PSi}^+$ $[\text{M}+\text{H}]^+$: 952.4917. Gefunden: 952.4905.

6.3.2.4 Synthese des m⁶A-CE Phosphoramidits

N⁶-Methyladenosin (23)



Chlorinosin **12** (8.00 g, 19.38 mmol, 1.00 eq.) wurde in EtOH (100 mL) gelöst und eine ethanolische Methylaminlösung (50 mL, 33% (w/w), 401.6 mmol, 20.7 eq.) zugegeben. Die Reaktion wurde für 18 h bei RT gerührt und das Lösemittel *in vacuo* entfernt. Zweimalige Umkristallisation aus EtOH lieferte das Produkt als weißes Pulver (3.87 g, 13.76 mmol, 71%). $R_f = 0.15$ (CH₂Cl₂/MeOH 10:1)

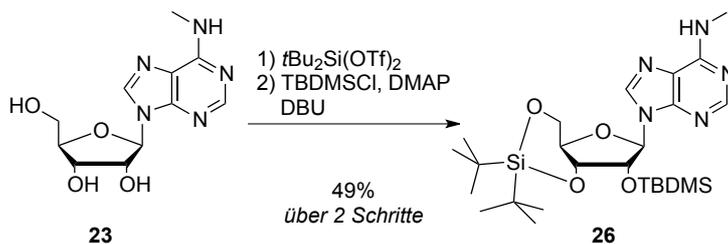
¹H NMR (400 MHz, *d*₆-DMSO) δ = 8.33 (s, 1H, H₈), 8.22 (br s, 1H, H₂), 7.80 (br s, 1H, NH), 5.88 (d, $J = 6.2$ Hz, 1H, H_{1'}), 5.43-5.39 (m, 2H, C_{2'}-OH, C_{5'}-OH) 5.16 (d, $J = 4.6$ Hz, 1H, C_{4'}-OH), 4.61 (td, $J = 6.2, 4.9$ Hz, 1H, H_{2'}), 4.15 (td, $J = 4.8, 3.0$ Hz, 1H, H_{3'}), 3.96 (q, $J = 3.5$ Hz, 1H, H_{4'}), 3.67 (dt, $J = 12.1, 4.1$ Hz, 1H, H_{5'}), 3.55 (ddd, $J = 12.1, 7.3, 3.7$ Hz, 1H, H_{5'}), 2.96 (br s, 3H, NCH₃).

¹³C NMR (101 MHz, *d*₆-DMSO) δ = 155.1 (C₆), 152.4 (C₂), 148.1 (C₄), 139.6 (C₈), 119.9 (C₅), 87.9 (C_{1'}), 85.9 (C_{4'}), 73.5 (C_{2'}), 70.7 (C_{3'}), 61.7 (C_{5'}), 27.0 (NCH₃).

HRMS-ESI (*m/z*): ber. für: C₁₁H₁₆N₅O₄⁺ [M+H]⁺: 282,1197. Gefunden 287.1198.

IR (ATR): ν = 3183 (m), 2926 (w), 1626 (s), 1543 (w), 1447 (w), 1376 (m), 1331 (m), 1309 (m), 1227 (m), 1191 (m), 1098 (s), 1082 (s), 1058 (s), 1032 (m), 986 (m), 945 (w), 864 (m), 819 (m), 792 (m), 769 (m), 739 (m), 706 (m), 667 (s).

5',3'-O-(Di-*tert*-butylsilylandiyl)-*N*⁶-methyl-2'-O-(*tert*-butyldimethylsilyl)adenosin (26)



Eine Lösung von *N*⁶-Methyladenosin (1.50 g, 5.33 mmol, 1.00 eq.) in DMF (4 mL) wurde auf 0 °C gekühlt und tropfenweise Di-*tert*-butyl-silylbis(trifluormethansulfonat) (1.91 mL, 5.87 mmol, 1.10 eq.) zugegeben. Nach 75 min. Rühren bei 0 °C wurde nochmals Di-*tert*-butyl-silylbis(trifluormethansulfonat) (0.3 mL, 921 μmol 0.17 eq.) zugegeben und die Reaktion nach weiteren 3 h durch Zugabe von NaHCO₃ (aq. sat., 15 mL) unterbrochen. Die Suspension wurde mit EtOAc (250 mL) versetzt, die Phasen getrennt und die wässrige Phase mit EtOAc (2 × 20 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und das Lösemittel *in vacuo* entfernt.

Der hieraus resultierende blassgelbe Schaum wurde in Acetonitril (15 mL) aufgenommen und bei Raumtemperatur mit DBU (2.79 mL, 18.67 mmol 3.50 eq.), TBDMSCl (2.01 g, 13.33 mmol, 2.50 eq.) und DMAP (65 mg, 530 μmol, 0.10 eq.) versetzt. Nach 15-stündigem Rühren bei Raumtemperatur wurde Methanol (20 mL) zugegeben und weitere 30 min. gerührt. Die Lösemittel wurden im Vakuum entfernt, der Rückstand in EtOAc (250 mL) aufgenommen, die Lösung mit NaHCO₃ (aq., 4 × 50 mL) und NaCl (aq. sat., 50 mL) gewaschen, über Na₂SO₄ getrocknet und das Lösemittel evaporiert. Das gewünschte Produkt **26** konnte nach Säulenchromatographie (Säule mit CH₂Cl₂ + 0.5% Pyridin gepackt und beladen, Elution mit Hexan/Aceton 5:1 + 0.5% Pyridin bis 4:1 + 0.5% Pyridin) als weißer Schaum erhalten werden (1.41 g, 2.63 mmol, 49%).

$R_f = 0.54$ (Hexan/Aceton 4:1)

¹H NMR (800 MHz, CDCl₃) $\delta = 8.35$ (br s, 1H, H₂), 7.75 (s, 1H, H₈), 6.28 (bs, 1H, NH), 5.89 (s, 1H, H_{1'}), 4.60 (d, $J = 4.8$ Hz, 1H, C_{2'}), 4.53 (dd, $J = 9.6, 4.8$ Hz, 1H, H_{3'}), 4.46 (dd, $J = 9.3, 5.2$ Hz, 1H, H_{5'}), 4.19 (ddd, $J = 10.4, 5.6, 4.8$ Hz, 1H, H_{4'}), 4.01 (dd, $J = 10.4, 9.4$ Hz, 1H, H_{5'}), 3.15 (br s, 3H, NCH₃), 1.06 (s, 9H, Si(C(CH₃)₃)₂), 1.02 (s, 9H, Si(C(CH₃)₃)₂), 0.91 (s, 9H, Si(CH₃)₂C(CH₃)₃), 0.14 (s, 3H, Si(CH₃)₂C(CH₃)₃), 0.12 (s, 3H, Si(CH₃)₂C(CH₃)₃).

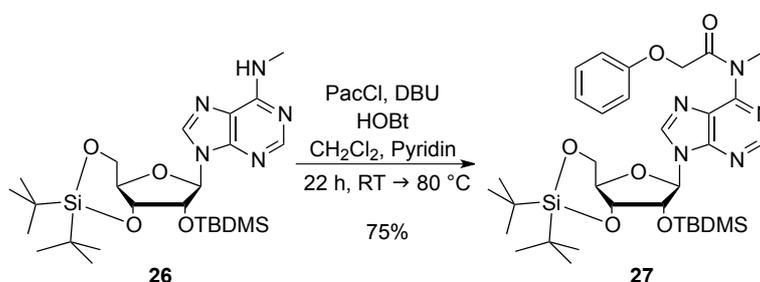
¹³C NMR (201 MHz, CDCl₃) $\delta = 155.6$ (C₆), 153.5 (C₂), 148.2 (C₄), 138.0 (C₈), 120.7 (C₅), 92.5 (C_{1'}), 75.9 (C_{3'}), 75.6 (C_{2'}), 74.8 (C_{4'}), 67.9 (C_{5'}), 27.6 (Si(C(CH₃)₃)₂), 27.5 (NCH₃) 27.1

(Si(C(CH₃)₃)₂), 26.0 (Si(CH₃)₂C(CH₃)₃), 22.8 (Si(C(CH₃)₃)₂), 20.5 (Si(C(CH₃)₃)₂), 18.4 (Si(CH₃)₂C(CH₃)₃), -4.2 (Si(CH₃)₂C(CH₃)₃), -4.9 (Si(CH₃)₂C(CH₃)₃).

HRMS-ESI (*m/z*): ber. für C₂₅H₄₆N₅O₄²⁸Si₂⁺ [M+H]⁺: 536.3083. Gefunden: 536.3083.

IR (ATR, neat): ν = 3284 (w), 3234 (w), 2931 (w), 2895 (w), 2859 (w), 2360 (w), 2341 (w), 1637 (m), 1583 (w), 1541 (w), 1472 (m), 1413 (w), 1388 (w), 1363 (w), 1346 (w), 1330 (w), 1304 (w), 1252 (m), 1226 (w), 1203 (w), 1163 (m), 1131 (s), 1105 (m), 1076 (m), 1050 (s), 1022 (w), 1001 (s), 938 (m), 907 (w), 889 (m), 833 (s), 826 (s), 793 (m), 783 (s), 751 (m), 740 (w), 720 (w), 683 (w), 668 (w).

5',3'-O-(Di-*tert*-butylsilylandiyl)-N⁶-methyl-N⁶-phenoxyacetyl-2'-O-(*tert*-butyldimethylsilyl) adenosin (**27**)



Das geschützte Adenosinderivat **26** (736 mg, 1.37 mmol, 1.00 eq.) wurde in Acetonitril (11 mL) und Pyridin (4.5 mL) aufgenommen und die Lösung auf 0 °C gekühlt. Nach Zugabe von DBU (1.03 mL, 6.87 mmol, 5.00 eq.), HOBt (928 mg, 6.87 mmol, 5.00 eq.) und PacCl (1.90 mL, 13.74 mmol, 10.00 eq.) wurde die Mischung 40 h bei Raumtemperatur gerührt. Es wurde kurz auf 85 °C erwärmt und die Reaktion anschließend durch Zugabe einer Lösung von NaHCO₃ (aq. sat., 50 mL) unterbrochen. Die Suspension wurde mit EtOAc extrahiert (250 mL) und die organische Phase mit NaOH (aq., 5% w/v, 50 mL), NH₄Cl (aq. sat., 50 mL) und NaCl (aq. sat., 50 mL) gewaschen, über Na₂SO₄ getrocknet und das Lösemittel in *vacuo* entfernt. Durch Säulenchromatographie (Hexan/Aceton 20:1 bis 8:1) konnte das Produkt als weißer Schaum (694 mg, 1.04 mmol, 75%) erhalten werden.

R_f = 0.63 (Hexan/Aceton 4:1)

¹H NMR (800 MHz, CDCl₃) δ = 8.67 (s, 1H, H₂), 8.03 (s, 1H, H₈), 7.20-7.16 (m, 2H, H_{meta}Ph), 6.90 (t, J = 7.3 Hz, 1H, H_{para}Ph), 6.67-6.65 (m, 2H, H_{ortho}Ph), 5.99 (s, 1H, H_{1'}), 5.15 (q_{AB}, $\Delta\delta_{AB}$ = 0.026, J_{AB} = 15.2 Hz, 2H, COCH₂O), 4.57 (d, J = 4.6 Hz, 1H, H_{2'}), 4.52 (dd, J = 9.3, 5.2 Hz, 1H, H_{5'}), 4.42 (dd, J = 9.7, 4.6 Hz, 1H, H_{3'}), 4.26 (td, J = 10.1, 5.2 Hz, 1H, H_{4'}), 4.04 (dd, J = 10.4, 9.5 Hz, 1H, H_{5'}), 3.75 (s, 3H, NCH₃), 1.08 (s, 9H, (Si(C(CH₃)₃)₂)), 1.05 (s,

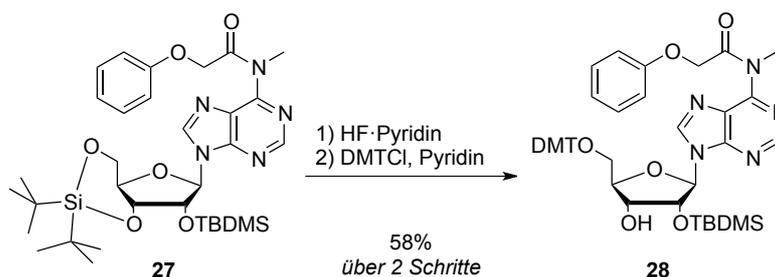
9H, (Si(C(CH₃)₃)₂)), 0.96 (s, 9H, (Si(CH₃)₂C(CH₃)₃)), 0.19 (s, 3H, (Si(CH₃)₂C(CH₃)₃)), 0.17 (s, 3H, (Si(CH₃)₂C(CH₃)₃)).

¹³C NMR (151 MHz, CDCl₃) δ = 170.7 (COCH₂O), 157.9 (C_{ipso}Ph), 153.1 (C₆), 152.2 (C₄), 151.7 (C₂), 141.5 (C₈), 129.5 (C_{meta}Ph), 126.0 (C₅), 121.5 (C_{para}Ph), 114.5 (C_{ortho}Ph), 92.6 (C_{1'}), 76.0 (C_{3'}), 75.6 (C_{2'}), 74.9 (C_{4'}), 69.0 (COCH₂O), 67.9 (C_{5'}), 35.1 (NCH₃), 27.6 (Si(C(CH₃)₃)₂), 27.2 (Si(C(CH₃)₃)₂), 26.0 (Si(CH₃)₂C(CH₃)₃), 22.9 (Si(C(CH₃)₃)₂), 20.5 (Si(C(CH₃)₃)₂), 18.5 (Si(CH₃)₂C(CH₃)₃), -4.1 (Si(CH₃)₂C(CH₃)₃), -4.8 (Si(CH₃)₂C(CH₃)₃).

HRMS-ESI (*m/z*): ber. für C₃₃H₅₂N₅O₆²⁸Si₂⁺ [M+H]⁺: 670.3451. Gefunden: 670.3453.

IR (ATR, neat): ν = 2931 (w), 2889 (w), 2858 (w), 2360 (w), 2341 (w), 1697 (m), 1599 (w), 1570 (m), 1495 (w), 1471 (m), 1464 (m), 1419 (w), 1386 (w), 1361 (w), 1351 (w), 1328 (w), 1294 (w), 1243 (m), 1220 (m), 1201 (m), 1184 (w), 1167 (m), 1136 (s), 1040 (s), 1021 (m), 1001 (m), 940 (w), 894 (m), 826 (s), 785 (m), 751 (s), 717 (w), 688 (m).

5'-O-(4,4'-Dimethoxytrityl)-N⁶-methyl-N⁶-phenoxyacetyl-2'-O-(*tert*-butyldimethylsilyl)adenosin (28)



Eine Lösung von Nucleosid **27** (645 mg, 963 μmol, 1.00 eq.) in CH₂Cl₂ (10 mL) und Pyridin (2 mL) wurde in einem Polypropylengefäß zunächst auf 0 °C gekühlt und anschließend mit HF·Pyridin (88 μL, 3.39 mmol, 3.52 eq.) versetzt. Nach 60 min. wurde erneut HF·Pyridin (88 μL, 3.39 mmol, 3.52 eq.) zugegeben und nach weiteren 45 min. ein drittes Mal (40 μL, 1.54 mmol, 1.60 eq.). Die Lösung wurde nochmals 30 min. gerührt und anschließend Methoxytrimethylsilan (1.19 mL, 8,64 mmol, 8.97 eq.) zugegeben. Die Lösung wurde auf Raumtemperatur erwärmt und das Lösemittel evaporiert.

Der so erhaltene klebrige Schaum wurde in Pyridin gelöst und auf 0 °C gekühlt. Nach Zugabe von DMTCl (536 mg, 1.58 mmol, 1.10 eq.) wurde zunächst für 5.5 h gerührt, eine weitere Portion DMTCl (49 mg, 158 μmol, 0.10 eq.) zugegeben, erneut für 60 min. gerührt und die Reaktionsmischung schließlich mit MeOH versetzt (1 mL). Entfernen des Lösemittels *in vacuo* und anschließende säulenchromatographische Aufreinigung (Hexan/Aceton

10:1 + 0.5% Pyridin bis 5:1 + 0.5% Pyridin) lieferte das Produkt als weißen Schaum (463 mg, 556 μmol , 58%). Um die Lagerfähigkeit des Produktes sicherzustellen ist es von Bedeutung etwaig verbleibendes Pyridin restlos zu entfernen.

$R_f = 0.28$ (Hexan/Aceton 4:1)

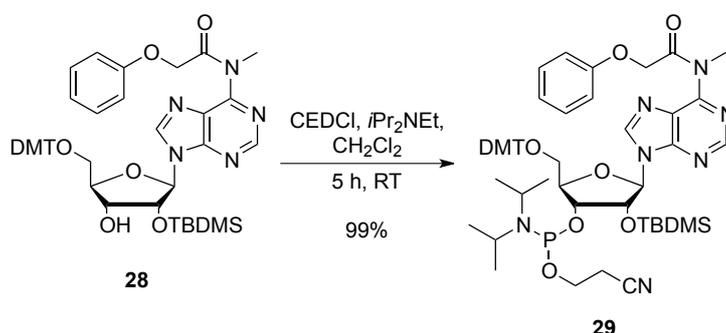
$^1\text{H NMR}$ (599 MHz, CDCl_3) $\delta = 8.62$ (s, 1H, H_2), 8.28 (s, 1H, H_8), 7.46 – 7.44 (m, 2H, $\text{H}_{\text{orthoPh-DMT}}$), 7.37 – 7.32 (m, 4H, $\text{H}_{\text{orthoPhOMe}}$), 7.31 – 7.16 (m, 5H, $\text{H}_{\text{metaPac}}$, $\text{H}_{\text{metaPh-DMT}}$, $\text{H}_{\text{paraPh-DMT}}$), 6.94 – 6.89 (m, 1H, $\text{H}_{\text{paraPac}}$), 6.85 – 6.80 (m, 4H, $\text{H}_{\text{metaPhOMe}}$), 6.73 – 6.68 (m, 2H, $\text{H}_{\text{orthoPac}}$), 6.12 (d, $J = 4.8$ Hz, 1H, H_1), 5.16 (q_{AB} , $\Delta\delta_{\text{AB}} = 0.020$, $J_{\text{AB}} = 15.6$ Hz, 2H, COCH_2O), 4.91 (dd, $J = 4.9, 4.9$ Hz, 1H, H_2), 4.35 (dd, $J = 4.5, 4.5$ Hz, 1H, H_3), 4.28 (dd, $J = 3.6, 7.2$ Hz, 1H, H_4), 3.78 (s, 3H, OCH_3), 3.78 (s, 3H, OCH_3), 3.76 (s, 3H, NCH_3), 3.56 (dd, $J = 10.7, 3.0$ Hz, 1H, H_5), 3.42 (dd, $J = 10.8, 3.8$ Hz, 1H, H_5), 0.86 (s, 9H, $(\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3)$), 0.04 (s, 3H, $(\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3)$), -0.08 (s, 3H, $(\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3)$).

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) $\delta = 170.7$ (COCH_2O), 158.8 ($\text{C}_{\text{paraPhOMe}}$), 158.0 ($\text{C}_{\text{ipsoPac}}$), 153.0 (C_6), 152.9 (C_4), 151.7 (C_2), 144.6 ($\text{C}_{\text{ipsoPh-DMT}}$), 142.2 (C_8), 135.7 ($\text{C}_{\text{ipsoPhOMe}}$), 130.2 ($\text{C}_{\text{orthoPhOMe}}$), 129.5 ($\text{C}_{\text{metaPac}}$), 129.3, 128.2 ($\text{C}_{\text{orthoPh-DMT}}$), 128.1 ($\text{C}_{\text{metaPh-DMT}}$), 127.1 ($\text{C}_{\text{paraPh-DMT}}$), 125.8 (C_5), 121.5 ($\text{C}_{\text{paraPac}}$), 114.6 ($\text{C}_{\text{orthoPac}}$), 113.4 ($\text{C}_{\text{metaPhOMe}}$), 88.6 (C_1), 86.9 ($\text{CPh}(\text{PhOMe})_2$), 84.3 (C_4), 76.0 (C_2), 71.4 (C_3), 69.1 (COCH_2O), 63.3 (C_5), 55.3 (OCH_3), 35.1 (NCH_3), 25.7 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), 18.0 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), -4.8 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$), -5.00 ($\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$).

HRMS-ESI (m/z): ber. für $\text{C}_{46}\text{H}_{54}\text{N}_5\text{O}_8^{28}\text{Si}^+ [\text{M}+\text{H}]^+$: 832.3736. Gefunden: 832.3738.

IR (ATR, neat): $\nu = 2928$ (w), 2854 (w), 1684 (w), 1599 (m), 1570 (s), 1508 (m), 1495 (m), 1463 (m), 1445 (w), 1420 (w), 1374 (w), 1361 (w), 1328 (w), 1300 (m), 1249 (s), 1219 (s), 1175 (s), 1130 (m), 1078 (w), 1064 (w), 1035 (s), 1007 (m), 911 (m), 867 (m), 832 (s), 782 (s), 752 (s), 726 (m), 701 (m), 690 (m).

5'-O-(4,4'-Dimethoxytrityl)-N⁶-methyl-N⁶-phenoxyacetyl-2'-O-(tert-butyl dimethylsilyl)adenosin 3'-(N,N-diisopropyl-2-cyanoethyl)phosphoramidit (XX)



Eine Lösung von **28** (300 mg, 361 μmol , 1.00 eq.) in CH_2Cl_2 (3 mL) wurde zweifach entgast, $i\text{Pr}_2\text{NEt}$ (250 μL , 1.44 mmol, 4.00 eq.) zugegeben, nochmals entgast, CEDCl (320 μL , 1.44 mmol, 4.00 eq.) zugegeben und zwei weitere Male entgast. Nach 5 h Reaktionszeit bei Raumtemperatur wurde die Reaktionsmischung direkt auf eine Silicasäule geladen und per *Flash* Chromatographie aufgereinigt (Hexan/EtOAc 1:2 + 0.5% Pyridin). Das oxidationsempfindliche Produkt wurde als weißer Schaum erhalten (368 mg, 357 μmol , 99%, 1:2.5 Mischung aus zwei Diastereomeren, bestimmt via NMR) und ist unter Luftausschluss einige Wochen lagerfähig.

$R_f = 0.54$ (Benzol/Et₂O 4:1.5)

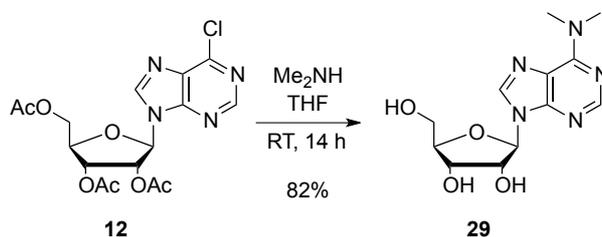
¹H NMR (400 MHz, CDCl_3) $\delta = [8.77$ (s) und 8.76 (s), 1H], $[8.48$ (s) und 8.45 (s), 1H], 7.66-7.61 (m, 2H), 7.55-7.50 (m, 4H), 7.45-7.34 (m, 5H), 7.09-7.05 (m, 1H), 7.00-6.96 (m, 4H), 6.91-6.89 (m, 2H), $[6.30$ (d, $J = 5.9$ Hz) und 6.24 (d, $J = 5.7$ Hz), 1H], 5.34-5.25 (m, 2H), 5.20-5.18 (m, 1H), 4.64-4.56 (m, 2H), 4.15-3.97 (m, 2H), 3.94-3.90 (m, 9H), 3.80-3.71 (m, 3H), 3.54-3.49 (m, 1H), $[2.84$ -2.75 (m) und 2.48 -2.44 (m), 2H] 1.37-1.21 (m, 12H) 0.93 (s, 9H), $[0.18$ (s) und 0.15 (s), 3H], $[0.02$ (s) und 0.01 (s), 3H].

³¹P NMR (162 MHz, CDCl_3) $\delta = 150.7, 149.3$.

HRMS-ESI (m/z): ber. für $\text{C}_{55}\text{H}_{71}\text{N}_7\text{O}_9^{31}\text{P}^{28}\text{Si}^+$ $[\text{M}+\text{H}]^+$: 1032.4815. Gefunden: 1032.4822.

6.3.2.5 Synthese des m⁶A-CE Phosphoramidits

N⁶,N⁶-Dimethyladenosin (29)



Inosinderivat **12** (5.01 g, 12.13 mmol, 1.00 eq.) wurde in einer Lösung von Me₂NH in THF (2 M, 100 mL, 16.5 eq.) aufgenommen und für 14 h bei RT gerührt. Das Lösemittel wurde *in vacuo* entfernt und der Rückstand zweifach aus CH₂Cl₂ (90 mL) umkristallisiert um das Produkt als weißen Feststoff zu erhalten (2.94 g, 9.95 mmol, 82%).

$R_f = 0.26$ (CH₂Cl₂/MeOH 4:1.5)

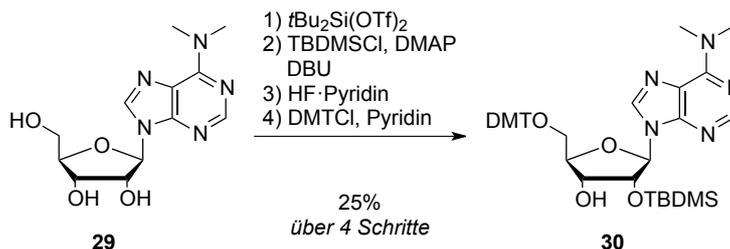
¹H NMR (400 MHz, *d*₆-DMSO) $\delta = 8.37$ (s, 1H, H₈), 8.21 (s, 1H, H₂), 5.91 (d, $J = 6.0$ Hz, 1H, H_{1'}), 5.44 (d, $J = 5.8$ Hz, 1H, C_{2'}-OH), 5.35 (dd, $J = 6.9$ Hz, 4.6 Hz, 1H, C_{5'}-OH), 5.17 (d, $J = 4.6$ Hz, 1H, C_{3'}-OH), 4.57 (q, $J = 5.5$ Hz, 1H, H_{2'}), 4.14 (q, $J = 4.2$ Hz, 1H, H_{3'}), 3.96 (q, $J = 3.5$ Hz, 1H, H_{4'}), 3.67 (dt, $J = 12.1, 4.1$ Hz, 1H, H_{5'}), 3.58-3.52 (m, 1H, H_{5'}), 3.36 (br s, 6H, N(CH₃)₂).

¹³C NMR (101 MHz, *d*₆-DMSO) $\delta = 154.3$ (C₅), 151.7 (C₂), 149.9 (C₄), 138.6 (C₈), 119.8 (C₆), 87.8 (C_{1'}), 85.7 (C_{4'}), 73.5 (C_{2'}), 70.5 (C_{3'}), 61.5 (C_{5'}), 37.9 (NCH₃).

HRMS-ESI (m/z): ber. für C₁₂H₁₈N₅O₄ [M+H]⁺: 296,1353. Gefunden: 296.1355.

IR (ATR, neat): $\nu = 3246$ (w), 3112 (w), 2986 (w), 2925 (w), 2881 (w), 2360 (w), 2339 (w), 1599 (s), 1570 (m), 1539 (w), 1486 (w), 1457 (w), 1423 (m), 1408 (w), 1394 (w), 1382 (w), 1342 (m), 1307 (m), 1271 (w), 1226 (m), 1213 (w), 1162 (w), 1130 (m), 1096 (m), 1067 (s), 1033 (s), 986 (w), 963 (w), 947 (w), 931 (w), 895 (w), 882 (w), 863 (w), 814 (m), 789 (m), 766 (w), 707 (w), 693 (w), 682 (m), 668 (w).

5'-O-(4,4'-Dimethoxytrityl)-N⁶,N⁶-dimethyl-2'-O-(tert-butyldimethylsilyl) adenosin (30)



$m^6_2\text{A}$ (900 mg, 3.05 mmol, 1.00 eq.) wurde in DMF (9 mL) gelöst, die Lösung auf 0 °C gekühlt und tropfenweise Di-*tert*-butyl-silylbis(trifluormethansulfonat) (1.10 mL, 3.38 mmol, 1.11 eq.) zugegeben. Die Reaktion wurde für 2 h bei 0 °C gerührt und durch Zugabe von NaHCO_3 (aq., sat., 20 mL) unterbrochen. Die wässrige Phase wurde mit EtOAc (150 mL) extrahiert, die organische Phase mit NaHCO_3 (aq., semi-sat., 4 × 30 mL) und NaCl (30 mL) gewaschen, über Na_2SO_4 getrocknet und evaporiert.

Der Rückstand wurde in MeCN (10 mL) aufgenommen und mit TBDMSCl (689 mg, 4.57 mmol, 1.50 eq.), DBU (0,91 mL, 6.10 mmol, 2.00 eq) und DMAP (37 mg, 305 μmol , 0.10 eq.) versetzt und für 16 h bei RT gerührt. Es wurden erneut TBDMSCl (345 mg, 2.29 mmol, 0.75 eq.) und DBU (456 μL , 3.05 mmol, 1.00 eq.) zugegeben und für weitere 6 h gerührt. Die Reaktion wurde durch Zugabe von MeOH (5 mL) unterbrochen und das Lösemittel *in vacuo* entfernt. Der erhaltene Feststoff wurde in EtOAc (150 mL) gelöst und mit NaHCO_3 (aq., sat., 15 mL) und NaCl (aq., sat., 15 mL), über Na_2SO_4 getrocknet und das Lösemittel evaporiert.

Das Rohprodukt wurde in ein Polypropylengefäß überführt und in CH_2Cl_2 (7 mL) und Pyridin (1 mL) gelöst. Die Mischung wurde auf 0 °C gekühlt und tropfenweise HF·Pyridin (190 μL , 7.32 mmol, 2.40 eq.) zugegeben. Nach einer Reaktionszeit von 2 h bei 0 °C wurde Methoxytrimethylsilan (1.30 mL, 9.46 mmol, 3.10 eq.) zugegeben, 30 min. bei RT gerührt und das Lösemittel evaporiert.

Der Rückstand wurde in Pyridin (8 mL) gelöst, auf 0 °C gekühlt und DMTCl (1.14 g, 3.36 mmol, 1.10 eq.) zugegeben. Die Reaktion wurde für 6 h gerührt, mit MeOH (5 mL) versetzt und evaporiert. Das Produkt wurde durch säulenchromatographische Aufreinigung (Hexan/Aceton 10:1 + 0.5% Pyridin bis 4:1 + 0.5% Pyridin) als weißer Schaum erhalten (539 mg, 757 μmol , 25%).

$R_f = 0.44$ (Hexan/Aceton 4:1)

$^1\text{H NMR}$ (599 MHz, CDCl_3) $\delta = 8.26$ (s, 1H, H₂), 7.93 (s, 1H, H₈), 7.47-7.45 (m, 2H, H_{ortho}Ph-DMT), 7.37-7.32 (m, 4H, H_{ortho}PhOMe), 7.29-7.24 (m, 2H, H_{meta}Ph-DMT), 7.23-7.16

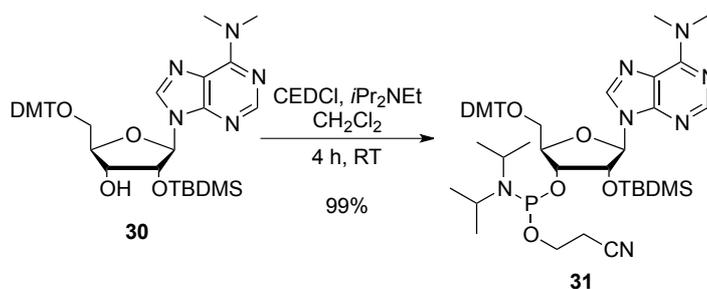
(m, 1H, H_{para}Ph-DMT), 6.83-6.79 (m, 4H, H_{meta}PhOMe), 6.03 (d, *J* = 5.5 Hz, 1H, H_{1'}), 4.97 (t, *J* = 5.3 Hz, 1H, H_{2'}), 4.30 (q, *J* = 4.0 Hz, 1H, H_{3'}), 4.23 (q, *J* = 3.7 Hz, 1H, H_{4'}), 3.78 (s, 6H, PhOCH₃), 3.54 (br s, 6H, N(CH₃)₂), 3.52 (dd, *J* = 10.6 Hz, 3.3 Hz, 1H, H_{5'}), 3.35 (dd, *J* = 10.6, 4.1 Hz, 1H, H_{5'}), 2.71 (d, *J* = 4.0 Hz, 1H, C_{3'}OH), 0.84 (s, 9H, (Si(CH₃)₂C(CH₃)₃)), -0.01 (s, 3H, (Si(CH₃)₂C(CH₃)₃)), -0.12 (s, 3H, (Si(CH₃)₂C(CH₃)₃)).

¹³C NMR (151 MHz, CDCl₃) δ = 158.6 (C_{para}PhOMe), 155.0 (C₆), 152.6 (C₂), 150.7 (C₄), 144.8 (C_{ipso}Ph), 137.0 (C₈), 135.9 (C_{ipso}PhOMe), 135.9 (C_{ipso}PhOMe), 130.2 (C_{ortho}PhOMe), 128.3 (C_{ortho}Ph-DMT), 128.0 (C_{meta}Ph-DMT), 127.0 (C_{para}Ph-DMT), 120.7 (C₅), 113.3 (C_{meta}PhOMe), 88.1 (C_{1'}), 86.6 (CPh(PhOCH₃)₂), 84.1 (C_{4'}), 75.5 (C_{2'}), 71.6 (C_{3'}), 63.6 (C_{5'}), 55.3 (PhOCH₃), 38.7 (N(CH₃)₂), 25.8 (Si(CH₃)₂C(CH₃)₃), 18.1 (Si(CH₃)₂C(CH₃)₃), -4.8 (Si(CH₃)₂C(CH₃)₃), -5.1 (Si(CH₃)₂C(CH₃)₃).

HRMS-ESI (*m/z*): ber. für C₃₉H₅₀N₅O₆Si⁺ [M+H]⁺: 712.3525. Gefunden: 712.3526.

IR (ATR, neat): ν = 3535 (w), 2950 (w), 2929 (w), 2857 (w), 2836 (w), 1592 (s), 1568 (m), 1508 (m), 1487 (w), 1463 (w), 1445 (w), 1424 (w), 1406 (w), 1377 (w), 1362 (w), 1334 (w), 1297 (m), 1249 (s), 1249 (s), 1219 (m), 1175 (m), 1133 (m), 1088 (m), 1065 (m), 1031 (s), 927 (w), 908 (w), 861 (w), 833 (s), 781 (s), 754 (w), 726 (w), 700 (m), 676 (w).

5'-*O*-(4,4'-Dimethoxytrityl)-*N*⁶,*N*⁶-dimethyl-2'-*O*-(*tert*-butyldimethylsilyl) adenosin 3'-*(N,N*-diisopropyl-2-cyanoethyl)phosphoramidit (31**)**



Eine Lösung von Dimethyladenosinderivat **30** (247 mg, 347 μmol, 1.00 eq.) wurde entgast, *i*Pr₂NEt (0.24 mL, 1.39 mmol, 4.00 eq.) zugegeben, zweifach entgast, die Lösung mit CEDCl (0.31 mL, 1.39 mmol, 4.00 eq.) versetzt, ein weiteres Mal entgast und schließlich für 4 h bei Raumtemperatur gerührt. Zur Aufreinigung wurde die Reaktion direkt auf eine Silicasäule (Hexan/EtOAc 1:2 + 0.5% Pyridin) geladen und das Produkt als weißer Schaum (318 mg, 346 μmol, 99%) erhalten.

*R*_f = 0.45 (Hexan/Aceton 2:1)

¹H NMR (400 MHz, CDCl₃) δ = [8.24 (s) und 8.21 (s), 1H], [7.95 (s) und 7.91 (s), 1H], 7.48-7.44 (m, 2H), 7.37-7.33 (m, 4H), 7.31-7.24 (m, 2H), 7.23-7.16 (m, 1H), 6.82-6.79 (m, 4H), [6.02 (d, *J* = 6.5 Hz) und 5.98 (d, *J* = 6.1 Hz), 1H], 5.09-5.04 (m, 1H), 4.40-4.32 (m, 2H), 4.00-3.82 (m, 2H), [3.79 (s) und 3.78 (s), 6H], 3.66-3.45 (m, 9H), 3.31-3.24 (m, 1H), [2.68-2.62 (m) und 2.30-2.26 (m), 2H], 1.21-1.15 (m, 12H), [0.78 (s) und 0.77 (s), 9H], [-0.02 (s) und -0.06 (s), 3H], [-0.18 (s) und -0.20 (s), 3H]

³¹P NMR (162 MHz, CDCl₃) δ = 150.7, 148.8.

HRMS-ESI (*m/z*): ber. für C₄₈H₆₇N₇O₇PSi⁺ [M+H]⁺: 912.4603. Gefunden: 912.4616.

7. Anhang

7.1 Abkürzungen

3T3	Murine Fibroblastenzelllinie
μ	mikro (10 ⁻⁶)
A	Adenosin, DNA: 2'-Desoxyadenosin
Å	Angström
Abb.	Abbildung
abs	absolut
ADAR	<i>adenosine deaminase acting on RNA</i> ; RNA-spezifische Adenosindeaminase
AdoH	5'-Desoxyadenosin
aq.	<i>aqueous</i> , wässrig
ATR	<i>attenuated total reflection</i> , abgeschwächte Totalreflexion
AU	<i>arbitrary units</i> , willkürliche Einheiten
ber.	berechnet
bp	Basenpaare
BTT	5-Benzylthio-1- <i>H</i> -tetrazol
C	Cytidin, DNA: 2'-Desoxycytidine
CAM	<i>cerium ammonium molybdate</i> , Ammoniumcermolybdat
CDS	<i>coding sequence</i> , proteincodierender Teil einer mRNA
CE	<i>N,N</i> -Diisopropylcyanoethylphosphoramidit
CED	Cyanoethoxy- <i>N,N</i> -diisopropylchlorophosphoramidit
CMC	<i>N</i> -Cyclohexyl- <i>N'</i> -(4-methylmorpholinium)ethylcarbodiimid
CMV	Cytomegalovirus
CPG	<i>controlled pore glass</i> , Glasgranulat mit definierter Porengröße
Cy3	Cyanin 3
δ	chemische Verschiebung (NMR)
d	Dublett
Da	Dalton
DBU	Diazabicyclo[5.4.0]undec-7-en
DCI	4,5-Dicyanoimidazol

ddH ₂ O	bidestilliertes Wasser
DMA	<i>N,N</i> -Dimethylannilin
DMAP	<i>N,N</i> -Dimethylaminopyridin
DMDS	Dimethyldisulfid
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxid
DMF	<i>N,N</i> -Dimethylformamid
DMT	4,4'-Dimethoxytrityl
DNA	<i>desoxyribonucleic acid</i> , Desoxyribonucleinsäure
DTT	Dithiothreitol
EDTA	Ethylendiammintetraessigsäure
eq.	Äquivalente
EMSA	<i>electrophoretic mobility shift assay</i> , Affinitätselektrophorese
ESI	<i>Electro Spray Ionisation</i> , Elektrosprayionisation
<i>et. al.</i>	<i>et alii, et aliae</i> , und andere
G	Guanosin, DNA: 2'-Desoxyadenosin
g	Gramm; Erdbeschleunigung
GFP	<i>green fluorescent protein</i> , grün fluoreszierendes Protein
GST	Glutathion-S-Transferase
h	Stunde
HEK	<i>human embryonic kidney</i>
HeLa	<i>Henrietta Lacks</i>
HPLC	<i>High Performance Liquid Chromatography</i> , Hochdruckchromatographie
HR	<i>high resolution</i> , hochauflösend
Hz	Hertz
IPN	Isopentylnitrit, Isoamylnitrit
IRES	<i>internal ribosomal entry site</i> , Interne Ribosomale Eintrittsstelle
KO	Knockout
L	Liter
LB	<i>lysogeny broth</i>
LFQ	<i>label-free quantification</i>
Lsg.	Lösung
μ	mikro (10 ⁻⁶)

m	milli (10^{-3}) bzw. Meter
M	Mega (10^6)
mA	Milliampere
MALDI	<i>Matrix Assisted Laser Desorption Ionisation</i>
mESC	<i>mouse embryonic stem cell(s)</i> , murine embryonale Stammzelle(n)
min.	Minute(n)
MS	Massenspektrometrie
n	nano (10^{-9})
NMR	<i>nuclear magnetic resonance</i> , Kernmagnetresonanz
NPC	<i>neural progenitor cell(s)</i> , neuronale Vorläuferzellen
OD	optische Dichte
ORN	Oligoribonukleotid
p	pico (10^{-12})
Pac	Phenoxyacetyl
PAGE	Polyacrylamid-Gelelektrophorese
ppm	<i>parts per million</i> (10^{-6})
PTC	<i>peptidyl transferase center</i> , Peptidyl-Transferasezentrum (des Ribosoms)
PTM	Posttranslationale Modifikation
PUS	Pseudouridinsynthese
q	Quartett
RP	<i>reversed phase</i> , Umkehrphase(n)
RPMI	<i>Roswell Park Memorial Institute</i>
RT	Raumtemperatur
rpm	<i>rounds per minute</i> , Umdrehungen pro Minute
s	Sekunde(n) bzw. Singulett
SAH	<i>S-Adenosylhomocystein</i>
SAM	<i>S-Adenosylmethionin</i>
sat.	<i>saturated</i> , gesättigt
Sdp	Siedepunkt
SDS	<i>Sodium dodecyl sulfate</i> , Natriumdodecylsulfat
SILAC	<i>stable isotope labelling with amino acids in cell culture</i>
t	Triplett
T	Thymidin

Tab.	Tabelle
TBDMS	<i>tert</i> -Butyldimethylsilyl
TCEP	Tris(2-carboxyethyl)phosphin
TEAB	Triethylammoniumhydrogencarbonat
TEMED	<i>N,N,N',N'</i> -Tetramethyldiamin
THF	Tetrahydrofuran
ToF	<i>time of flight</i>
Tris	Tri(Hydroxymethyl)-aminomethan
U	Unit bzw. Uridine
UTR	<i>untranslated region</i> , nicht translatierter Teil einer mRNA
UV	ultraviolett
V	Volt
v/v	<i>volume per volume</i> , Volumenanteil
WT	Wildtyp
w/v	<i>weight per volume</i> , Massenteil pro Volumen

7.2 Syntheseyklen für die Oligonucleotidsynthese

Syntheseyklus für Einfachkopplung / 1.5 M Amiditlösungen:

Step	Function	Number	Time	Active
1	Begin	106		
2	18 to Waste	64	3,0	
3	14 to column	40	10,0	AGCT5678
4	Reverse Flush	2	10,0	AGCT5678
5	Block Flush	1	4,0	AGCT5678
6	Phos Prep	101	3,0	
7	Column 1 On	140		
8	Block Vent	111	2,0	AGCT5678
9	Tet to Waste	58	1,7	
10	B + Tet to column	33	2,5	AGCT5678
11	Tet to Column	34	1,0	AGCT5678
12	B + Tet to Column	33	2,5	AGCT5678
13	Push to Column	43		
14	Column 1 Off	141		
15	Column 2 On	142		
16	18 to Waste	64	4,0	
17	Block Flush	1	3,0	AGCT5678
18	Block Vent	111	2,0	AGCT5678
19	Tet to Waste	58	1,7	
20	B + Tet to Column	33	2,5	AGCT5678
21	Tet to Column	34	1,0	AGCT5678
22	B + Tet to column	33	2,5	AGCT5678
23	Push to Column	43		
24	Column 2 Off	143		
25	Column 3 On	144		
26	18 to Waste	64	4,0	
27	Block Flush	1	3,0	AGCT5678
28	Block Vent	111	2,0	AGCT5678
29	Tet to Waste	58	1,7	
30	B + Tet to column	33	2,5	AGCT5678
31	Tet to Column	34	1,0	AGCT5678
32	B + Tet to Column	33	2,5	AGCT5678
33	Push to Column	43		
34	Column 3 Off	145		
35	Column 4 On	146		
36	18 to Waste	64	4,0	
37	Block Flush	1	3,0	AGCT5678
38	Block Vent	111	2,0	AGCT5678
39	Tet to Waste	58	1,7	
40	B + Tet to Column	33	2,5	AGCT5678
41	Tet to Column	34	1,0	AGCT5678
42	B + Tet to column	33	2,5	AGCT5678

43	Push to Column	43		
44	Column 4 Off	147		
45	Wait	103	600,0	AGCT5678
46	Wait	103	600,0	5678
47	Cap Prep	102	3,0	
48	18 to Waste	64	4,0	
49	Reverse Flush	2	7,0	AGCT5678
50	Block Flush	1	3,0	AGCT5678
51	Cap to Column	39	10,0	AGCT5678
52	Wait	103	5,0	AGCT5678
53	18 to Waste	64	4,0	
54	Reverse Flush	2	7,0	AGCT5678
55	Block Flush	1	3,0	AGCT5678
56	15 to Column	41	8,0	AGCT5678
57	18 to Waste	64	4,0	
58	Block Flush	1	3,0	AGCT5678
59	Wait	103	15,0	AGCT5678
60	18 to Column	42	10,0	AGCT5678
61	Flush to Waste	4	6,0	AGCT5678
62	18 to Column	42	10,0	AGCT5678
63	Reverse Flush	2	7,0	AGCT5678
64	Block Flush	1	3,0	AGCT5678
65	Start Detrityl	105		
66	18 to Waste	64	4,0	
67	18 to Column	42	10,0	AGCT5678
68	Reverse Flush	2	7,0	AGCT5678
69	Block Flush	1	3,0	AGCT5678
70	If monitoring	167		
71	19 to Column	44	35,0	AGCT5678
72	14 to Column	40	3,0	AGCT5678
73	Monitor Trytils	135		
74	14 to Column	40	85,0	AGCT5678
75	Wait	103	3,0	AGCT5678
76	Monitor Noise	136		
77	14 to Column	40	10,0	AGCT5678
78	Stop Monitor	137		
79	18 to Column	42	10,0	AGCT5678
80	Reverse Flush	2	8,0	AGCT5678
81	If not Montring	168		
82	14 to Column	40	6,0	CT5678
83	Trityl Flush	3	5,0	CT5678
84	14 to Column	40	6,0	AGCT5678
85	Wait	103	5,0	AGCT5678
86	Trityl Flush	3	5,0	AGCT5678
87	14 to Column	40	6,0	AGCT5678
88	Wait	103	5,0	AGCT5678
89	Trityl Flush	3	5,0	AGCT5678

90	14 to Column	40	6,0	AGCT5678
91	Wait	103	5,0	AGCT5678
92	Trityl Flush	3	5,0	AGCT5678
93	18 to Column	42	10,0	AGCT5678
94	Trityl Flush	3	8,0	AGCT5678
95	End Monitoring	169		
96	18 to Column	42	8,0	AGCT5678
97	Reverse Flush	2	7,0	AGCT5678
98	Block Flush	1	4,0	AGCT5678
99	End	107		

Synthesezyklus für Doppelkopplung / 1.0 M Amiditlösungen:

Step	Function	Number	Time	Active
1	Begin	106		
2	18 to Waste	64	3,0	
3	14 to column	40	10,0	AGCT5678
4	Reverse Flush	2	10,0	AGCT5678
5	Block Flush	1	4,0	AGCT5678
6	Phos Prep	101	3,0	
7	Column 1 On	140		
8	Block Vent	111	2,0	AGCT5678
9	Tet to Waste	58	1,7	
10	B + Tet to column	33	2,5	AGCT5678
11	Tet to Column	34	1,0	AGCT5678
12	B + Tet to Column	33	2,5	AGCT5678
13	Push to Column	43		
14	Column 1 Off	141		
15	Column 2 On	142		
16	18 to Waste	64	4,0	
17	Block Flush	1	3,0	AGCT5678
18	Block Vent	111	2,0	AGCT5678
19	Tet to Waste	58	1,7	
20	B + Tet to Column	33	2,5	AGCT5678
21	Tet to Column	34	1,0	AGCT5678
22	B + Tet to column	33	2,5	AGCT5678
23	Push to Column	43		
24	Column 2 Off	143		
25	Column 3 On	144		
26	18 to Waste	64	4,0	
27	Block Flush	1	3,0	AGCT5678
28	Block Vent	111	2,0	AGCT5678
29	Tet to Waste	58	1,7	
30	B + Tet to column	33	2,5	AGCT5678
31	Tet to Column	34	1,0	AGCT5678
32	B + Tet to Column	33	2,5	AGCT5678
33	Push to Column	43		

34	Column 3 Off	145		
35	Column 4 On	146		
36	18 to Waste	64	4,0	
37	Block Flush	1	3,0	AGCT5678
38	Block Vent	111	2,0	AGCT5678
39	Tet to Waste	58	1,7	
40	B + Tet to Column	33	2,5	AGCT5678
41	Tet to Column	34	1,0	AGCT5678
42	B + Tet to column	33	2,5	AGCT5678
43	Push to Column	43		
44	Column 4 Off	147		
45	Wait	103	300,0	AGCT5678
46	Wait	103	300,0	5678
47	Column 1 On	140		
48	Block Vent	111	2,0	ACGT
49	Tet to Waste	58	1,7	
50	B + Tet to column	33	2,5	ACGT
51	Tet to Column	34	1,0	ACGT
52	B + Tet to Column	33	2,5	ACGT
53	Push to Column	43		
54	Column 1 Off	141		
55	Column 2 On	142		
56	18 to Waste	64	4,0	
57	Block Flush	1	3,0	ACGT
58	Block Vent	111	2,0	ACGT
59	Tet to Waste	58	1,7	
60	B + Tet to Column	33	2,5	ACGT
61	Tet to Column	34	1,0	ACGT
62	B + Tet to column	33	2,5	ACGT
63	Push to Column	43		
64	Column 2 Off	143		
65	Column 3 On	144		
66	18 to Waste	64	4,0	
67	Block Flush	1	3,0	ACGT
68	Block Vent	111	2,0	ACGT
69	Tet to Waste	58	1,7	
70	B + Tet to column	33	2,5	ACGT
71	Tet to Column	34	1,0	ACGT
72	B + Tet to Column	33	2,5	ACGT
73	Push to Column	43		
74	Column 3 Off	145		
75	Column 4 On	146		
76	18 to Waste	64	4,0	
77	Block Flush	1	3,0	ACGT
78	Block Vent	111	2,0	ACGT
79	Tet to Waste	58	1,7	
80	B + Tet to Column	33	2,5	ACGT

81	Tet to Column	34	1,0	ACGT
82	B + Tet to column	33	2,5	ACGT
83	Push to Column	43		
84	Column 4 Off	147		
85	Wait	103	300,0	ACGT
86	Cap Prep	102	3,0	
87	18 to Waste	64	4,0	
88	Reverse Flush	2	7,0	AGCT5678
89	Block Flush	1	3,0	AGCT5678
90	Cap to Column	39	10,0	AGCT5678
91	Wait	103	5,0	AGCT5678
92	18 to Waste	64	4,0	
93	Reverse Flush	2	7,0	AGCT5678
94	Block Flush	1	3,0	AGCT5678
95	15 to Column	41	8,0	AGCT5678
96	18 to Waste	64	4,0	
97	Block Flush	1	3,0	AGCT5678
98	Wait	103	15,0	AGCT5678
99	18 to Column	42	10,0	AGCT5678
100	Flush to Waste	4	6,0	AGCT5678
101	18 to Column	42	10,0	AGCT5678
102	Reverse Flush	2	7,0	AGCT5678
103	Block Flush	1	3,0	AGCT5678
104	Start Detrityl	105		
105	18 to Waste	64	4,0	
106	18 to Column	42	10,0	AGCT5678
107	Reverse Flush	2	7,0	AGCT5678
108	Block Flush	1	3,0	AGCT5678
109	If monitoring	167		
110	19 to Column	44	35,0	AGCT5678
111	14 to Column	40	3,0	AGCT5678
112	Monitor Trytils	135		
113	14 to Column	40	85,0	AGCT5678
114	Wait	103	3,0	AGCT5678
115	Monitor Noise	136		
116	14 to Column	40	10,0	AGCT5678
117	Stop Monitor	137		
118	18 to Column	42	10,0	AGCT5678
119	Reverse Flush	2	8,0	AGCT5678
120	If not Montring	168		
121	14 to Column	40	6,0	CT5678
122	Trityl Flush	3	5,0	CT5678
123	14 to Column	40	6,0	AGCT5678
124	Wait	103	5,0	AGCT5678
125	Trityl Flush	3	5,0	AGCT5678
126	14 to Column	40	6,0	AGCT5678
127	Wait	103	5,0	AGCT5678

128	Trityl Flush	3	5,0	AGCT5678
129	14 to Column	40	6,0	AGCT5678
130	Wait	103	5,0	AGCT5678
131	Trityl Flush	3	5,0	AGCT5678
132	18 to Column	42	10,0	AGCT5678
133	Trityl Flush	3	8,0	AGCT5678
134	End Monitoring	169		
135	18 to Column	42	8,0	AGCT5678
136	Reverse Flush	2	7,0	AGCT5678
137	Block Flush	1	4,0	AGCT5678
138	End	107		

8. Literaturverzeichnis

1. Levene, P. A. The Structure of Yeast Nucleic Acid. *J. Biol. Chem.* **40**, 415–424 (1919).
2. Avery, O. T., MacLeod, C. M. & McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. *J. Exp. Med.* **79**, 137–158 (1944).
3. Hershey, A. D. & Chase, M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36**, 39–56 (1952).
4. Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209 (1950).
5. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids. *Nature* **171**, 737–738 (1953).
6. Wilkins, M. H. F., Stokes, A. R. & Wilson, H. R. Molecular structure of deoxypentose nucleic acids. *Nature* **171**, 738–740 (1953).
7. Watson, J. D. & Crick, F. H. C. Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature* **171**, 964–967 (1953).
8. Meselson, M. & Stahl, F. W. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **44**, 671–682 (1958).
9. Crick, F. H. C. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).
10. Crick, F. H. C. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).
11. Holliday, R. Epigenetics: An Overview. *Dev. Gen.* **15**, 453–457 (1994).
12. Holliday, R. Epigenetics: A Historical Overview. *Epigenetics* **1**, 76–80 (2014).
13. Wyatt, G. R. & Cohen, S. S. A new pyrimidine base from bacteriophage nucleic acids. *Nature* **170**, 1072–1073 (1952).
14. Wyatt, G. R. & Cohen, S. S. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem. J.* **55**, 774–782 (1953).
15. Pfaffeneder, T. *et al.* The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA. *Angew. Chem. Int. Ed.* **50**, 7008–7012 (2011).
16. Ito, S., Shen, L., Dai, Q., Wu, S. C. & Collins, L. B. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
17. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
18. Pfaffeneder, T. *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **10**, 574–581 (2014).
19. Baxter, E., Windloch, K., Gannon, F. & Lee, J. S. Epigenetic regulation in cancer progression. *Cell Biosci.* **4**, 1–11 (2014).
20. Klose, R. J. & Bird, A. P. Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* **31**, 89–97 (2006).
21. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33**, 245–254 (2003).
22. Carell, T. *et al.* Structure and Function of Noncanonical Nucleobases. *Angew. Chem. Int. Ed.* **51**, 7110–7131 (2012).
23. Spruijt, C. G. *et al.* Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell* **152**, 1146–1159 (2013).
24. Du, J., Johnson, L. M., Jacobsen, S. E. & Patel, D. J. DNA methylation pathways

- and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* **16**, 519–532 (2015).
25. McGinty, R. K. & Tan, S. Nucleosome Structure and Function. *Chem. Rev.* **115**, 2255–2273 (2015).
 26. Lee, J.-S., Smith, E. & Shilatifard, A. The Language of Histone Crosstalk. *Cell* **142**, 682–685 (2010).
 27. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science* **293**, 1–8 (2001).
 28. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
 29. Gardner, K. E., Allis, C. D. & Strahl, B. D. OPERating ON Chromatin, a Colorful Language where Context Matters. *J. Mol. Biol.* **409**, 36–46 (2011).
 30. Huang, H., Lin, S., Garcia, B. A. & Zhao, Y. Quantitative Proteomic Analysis of Histone Modifications. *Chem. Rev.* **115**, 2376–2418 (2015).
 31. He, C. RNA epigenetics? *Nat. Chem. Biol.* **6**, 863–865 (2010).
 32. Roundtree, I. A. & He, C. RNA epigenetics - chemical messages for posttranscriptional gene regulation. *Curr. Op. Chem. Biol.* **30**, 46–51 (2016).
 33. Sibbritt, T., Patel, H. R. & Preiss, T. Mapping and significance of the mRNA methylome. *WIREs RNA* **4**, 397–422 (2013).
 34. Wang, X. & He, C. Dynamic RNA Modifications in Posttranscriptional Regulation. *Mol. Cell* **56**, 5–12 (2014).
 35. Limbach, P. A., Crain, P. F. & McCloskey, J. A. Summary: the modified nucleosides of RNA. *Nucl. Acids Res.* **22**, 2183–2196 (1994).
 36. Cantara, W. A. *et al.* The RNA modification database, RNAMDB: 2011 update. *Nucl. Acids Res.* **39**, D195–D201 (2010).
 37. Machnicka, M. A. *et al.* MODOMICS: a database of RNA modification pathways - 2013 update. *Nucl. Acids Res.* **41**, D262–D267 (2012).
 38. Sprinzl, M. & Vassilenko, K. S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucl. Acids Res.* **33**, D139–D140 (2004).
 39. Phizicky, E. M. & Hopper, A. K. tRNA biology charges to the front. *Genes Dev.* **24**, 1832–1860 (2010).
 40. Laten, H., Gorman, J. & Bock, R. M. Isopentenyladenosine deficient tRNA from an antisuppressor mutant of *Saccharomyces cerevisiae*. *Nucl. Acids Res.* **5**, 4329–4342 (1978).
 41. Dihanich, M. E. *et al.* Isolation and characterization of MOD5, a gene required for isopentenylation of cytoplasmic and mitochondrial tRNAs of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **7**, 177–184 (1987).
 42. Jenner, L. B., Demeshkina, N., Yusupova, G. & Yusupov, M. Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat. Struct. Mol. Biol.* **17**, 555–560 (2010).
 43. Björk, G. R. *et al.* A primordial tRNA modification required for the evolution of life? *EMBO J.* **20**, 231–239 (2001).
 44. Yacoubi, El, B. *et al.* The universal YrdC/Sua5 family is required for the formation of threonylcarbamoyladenosine in tRNA. *Nucl. Acids Res.* **37**, 2894–2909 (2009).
 45. Lecointe, F. *et al.* Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop. *J. Biol. Chem.* **273**, 1316–1323 (1998).
 46. Gerber, A. P. & Keller, W. An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **286**, 1146–1149 (1999).
 47. Björk, G. R., Huang, B., Persson, O. P. & Bystrom, A. S. A conserved modified

- wobble nucleoside (mcm⁵s²U) in lysyl-tRNA is required for viability in yeast. *RNA* **13**, 1245–1255 (2007).
48. Davanloo, P., Sprinzl, M. & Watanabe, K. Role of ribothymidine in the thermal stability of transfer RNA as monitored by proton magnetic resonance. *Nucl. Acids Res.* **6**, 1571–1581 (1979).
 49. Nobles, K. N., Yarian, C. S., Liu, G., Guenther, R. H. & Agris, P. F. Highly conserved modified nucleosides influence Mg²⁺-dependent tRNA folding. *Nucl. Acids Res.* **30**, 4751–4760 (2002).
 50. Sengupta, R. *et al.* Modified constructs of the tRNA T Ψ C domain to probe substrate conformational requirements of m¹A₅₈ and m⁵U₅₄ tRNA methyltransferases. *Nucl. Acids Res.* **28**, 1374–1380 (2000).
 51. Helm, M., Giegé, R. & Florentz, C. A Watson–Crick Base-Pair-Disrupting Methyl Group (m¹A₉) Is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNA^{Lys}. *Biochemistry* **38**, 13338–13346 (1999).
 52. Pütz, J., Florentz, C., Benseler, F. & Giegé, R. A single methyl group prevents the mischarging of a tRNA. *Nat. Struct. Mol. Biol.* **1**, 580–582 (1994).
 53. Aström, S. U. & Bystrom, A. S. Rit1, a tRNA backbone-modifying enzyme that mediates initiator and elongator tRNA discrimination. *Cell* **79**, 535–546 (1994).
 54. Nameki, N., Asahara, H., Shimizu, M., Okada, N. & Himeno, H. Identity elements of *Saccharomyces cerevisiae* tRNA^{His}. *Nucl. Acids Res.* **23**, 389–394 (1995).
 55. Rudinger, J., Florentz, C. & Giegé, R. Histidylolation by yeast HisRS of tRNA or tRNA-like structure relies on residues – 1 and 73 but is dependent on the RNA context. *Nucl. Acids Res.* **22**, 5031–5037 (1994).
 56. Chan, C. T. Y. *et al.* A Quantitative Systems Approach Reveals Dynamic Control of tRNA Modifications during Cellular Stress. *PLoS Genet* **6**, e1001247–9 (2010).
 57. Chan, C. T. Y. *et al.* Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nat. Comms.* **3**, 937–9 (2012).
 58. Preston, M. A., D'Silva, S., Kon, Y. & Phizicky, E. M. tRNA^{His} 5-methylcytidine levels increase in response to several growth arrest conditions in *Saccharomyces cerevisiae*. *RNA* **19**, 243–256 (2013).
 59. Brückl, T. Synthesis of natural and isotope-labeled tRNA nucleosides and their mass spectrometric quantification. 1–247 (2010).
 60. Brückl, T., Globisch, D., Wagner, M., Müller, M. & Carell, T. Parallel Isotope-Based Quantification of Modified tRNA Nucleosides. *Angew. Chem. Int. Ed.* **48**, 7932–7934 (2009).
 61. Piekna-Przybylska, D., Decatur, W. A. & Fournier, M. J. The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucl. Acids Res.* **36**, D178–D183 (2007).
 62. Decatur, W. A. & Fournier, M. J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**, 344–351 (2002).
 63. Davis, D. R. Stabilization of RNA stacking by pseudouridine. *Nucl. Acids Res.* **23**, 5020–5026 (1995).
 64. Kawai, G. *et al.* Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* **31**, 1040–1046 (1992).
 65. Xue, S. & Barna, M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat. Rev. Mol. Cell Biol.* **13**, 355–369 (2012).
 66. Krogh, N. *et al.* Profiling of 2'-O-Me in human rRNA reveals a subset of

- fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucl. Acids Res.* **44**, 7884–7895 (2016).
67. Long, K. S., Poehlsgaard, J., Kehrenberg, C., Schwarz, S. & Vester, B. The Cfr rRNA Methyltransferase Confers Resistance to Phenicol, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A Antibiotics. *Antimicrob. Agents Chemother.* **50**, 2500–2505 (2006).
 68. Skinner, R., Cundliffe, E. & Schmidt, F. J. Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J. Biol. Chem.* **258**, 12702–12706 (1983).
 69. Thompson, J., Schmidt, F. & Cundliffe, E. Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J. Biol. Chem.* **257**, 7915–7917 (1982).
 70. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193**, 651–669 (2013).
 71. Dominissini, D. *et al.* Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* **485**, 201–206 (2012).
 72. Meyer, K. D. *et al.* Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons. *Cell* **149**, 1635–1646 (2012).
 73. Liu, N. *et al.* Probing N⁶-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* **19**, 1848–1856 (2013).
 74. Yu, Y.-T., Shu, M.-D. & Steitz, J. A. Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.* **17**, 5783–5795 (1998).
 75. Schwartz, S. *et al.* Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162 (2014).
 76. Patton, J. R. Pseudouridine formation in small nuclear RNAs. *Biochimie* **76**, 1129–1132 (1994).
 77. Patton, J. R., Jacobson, M. R. & Pederson, T. Pseudouridine formation in U2 small nuclear RNA. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3324–3328 (1994).
 78. Carlile, T. M. *et al.* Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143–146 (2014).
 79. Massenet, S., Motorin, Y. & Lafontaine, D. Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase Pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol. Cell. Biol.* **19**, 2142–2154 (1999).
 80. Gu, J., Patton, J. R., Shimba, S. & Reddy, R. Localization of modified nucleotides in *Schizosaccharomyces pombe* spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *RNA* **2**, 909–918 (1996).
 81. Karijolich, J., Yi, C. & Yu, Y.-T. Transcriptome-wide dynamics of RNA pseudouridylation. *Nat. Rev. Mol. Cell Biol.* **16**, 581–585 (2015).
 82. Yu, B. *et al.* Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**, 932–935 (2005).
 83. Li, J., Yang, Z., Yu, B., Liu, J. & Chen, X. Methylation Protects miRNAs and siRNAs from a 3'-End Uridylation Activity in Arabidopsis. *Current Biology* **15**, 1501–1507 (2005).
 84. Chen, X. Small RNAs in development – insights from plants. *Curr. Op. Genet. Dev.* **22**, 361–367 (2012).
 85. Ren, G., Chen, X. & Yu, B. Small RNAs meet their targets: When methylation defends miRNAs from uridylation. *RNA Biol.* **11**, 1099–1104 (2014).
 86. Ji, L. & Chen, X. Regulation of small RNA stability: methylation and beyond.

- Cell Res.* **22**, 624–636 (2012).
87. Kirino, Y. & Mourelatos, Z. Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. *Nat. Struct. Mol. Biol.* **14**, 347–348 (2007).
 88. Houwing, S. *et al.* A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish. *Cell* **129**, 69–82 (2007).
 89. Yang, W. *et al.* Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat. Struct. Mol. Biol.* **13**, 13–21 (2005).
 90. Nishikura, K. Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat. Rev. Mol. Cell Biol.* **7**, 919–931 (2006).
 91. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Meth.* **12**, 767–772 (2015).
 92. Davis, F. F. & Allen, F. W. Ribonucleic acids from yeast which contain a fifth nucleotide. *J. Biol. Chem.* **227**, 907–915 (1957).
 93. Li, X. *et al.* Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* **11**, 592–597 (2015).
 94. Bakin, A. & Ofengand, J. Mapping of the 13 pseudouridine residues in *Saccharomyces cerevisiae* small subunit ribosomal RNA to nucleotide resolution. *Nucl. Acids Res.* **23**, 3290–3294 (1995).
 95. Bakin, A. & Ofengand, J. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762 (1993).
 96. Ofengand, J. & Bakin, A. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaeobacteria, mitochondria and chloroplasts. *J. Mol. Biol.* **266**, 246–268 (1997).
 97. Spedaliere, C. J., Ginter, J. M., Johnston, M. V. & Mueller, E. G. The Pseudouridine Synthases: Revisiting a Mechanism That Seemed Settled. *J. Am. Chem. Soc.* **126**, 12758–12759 (2004).
 98. Hamma, T. & Ferré-D'Amaré, A. R. Pseudouridine Synthases. *Chemistry & Biology* **13**, 1125–1135 (2006).
 99. Ma, X., Zhao, X. & Yu, Y.-T. Pseudouridylation of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *EMBO J.* **22**, 1889–1897 (2003).
 100. Kierzek, E. *et al.* The contribution of pseudouridine to stabilities and structure of RNAs. *Nucl. Acids Res.* **42**, 3492–3501 (2014).
 101. Karijolich, J. & Yu, Y.-T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395–398 (2012).
 102. Svidritskiy, E., Madireddy, R. & Korostelev, A. A. Structural Basis for Translation Termination on a Pseudouridylated Stop Codon. *J. Mol. Biol.* **428**, 2228–2236 (2016).
 103. Wei, C. M., Gershowitz, A. & Moss, B. Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell* **4**, 379–386 (1975).
 104. Langberg, S. R. & Moss, B. Post-transcriptional modifications of mRNA. *J. Biol. Chem.* **256**, 10054–10060 (1981).
 105. Daffis, S. *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* **468**, 452–456 (2010).
 106. Szretter, K. J. *et al.* 2'-O Methylation of the Viral mRNA Cap by West Nile Virus Evades Ifit1-Dependent and -Independent Mechanisms of Host Restriction In Vivo. *PLoS Pathog.* **8**, e1002698–13 (2012).
 107. Züst, R. *et al.* Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat.*

- Immun.* **12**, 137–143 (2011).
108. Judge, A. D., Bola, G., Lee, A. C. H. & MacLachlan, I. Design of Noninflammatory Synthetic siRNA Mediating Potent Gene Silencing in Vivo. *Molecular Therapy* **13**, 494–505 (2006).
109. Reiter, V. Quantifizierung eukaryotischer RNA-Modifikationen und Untersuchungen zum humanen Homolog der bakteriellen Methylthiotransferase *MiaB*. 1–175 (2013).
110. Bachellerie, J. P., Cavaillé, J. & Hüttenhofer, A. The expanding snoRNA world. *Biochimie* **84**, 775–790 (2002).
111. Jády, B. E. & Kiss, T. Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-*O*-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs. *Nucl. Acids Res.* **28**, 1348–1354 (2000).
112. Cavaillé, J. *et al.* Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14311–14316 (2000).
113. Yi-Brunozzi, H. Y., Easterwood, L. M., Kamilar, G. M. & Beal, P. A. Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2. *Nucl. Acids Res.* **27**, 2912–2917 (1999).
114. Brandmayr, C. *et al.* Isotope-Based Analysis of Modified tRNA Nucleosides Correlates Modification Density with Translational Efficiency. *Angew. Chem. Int. Ed.* **51**, 11162–11165 (2012).
115. Birkedal, U. *et al.* Profiling of Ribose Methylations in RNA by High-Throughput Sequencing. *Angew. Chem. Int. Ed.* **54**, 451–455 (2015).
116. Marchand, V., Blanloeil-Oillo, F., Helm, M. & Motorin, Y. Illumina-based RiboMethSeq approach for mapping of 2'-*O*-Me residues in RNA. *Nucl. Acids Res.* **44**, e135–e135 (2016).
117. Dubin, D. T. & Taylor, R. H. The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucl. Acids Res.* **2**, 1653–1668 (1975).
118. Motorin, Y., Lyko, F. & Helm, M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucl. Acids Res.* **38**, 1415–1430 (2010).
119. Squires, J. E. *et al.* Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucl. Acids Res.* **40**, 5023–5033 (2012).
120. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**, 1393–1406 (2012).
121. Kwon, S. C. *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 (2013).
122. Baltz, A. G. *et al.* The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Mol. Cell* **46**, 674–690 (2012).
123. Conrad, T. *et al.* Serial interactome capture of the human cell nucleus. *Nat. Comms.* **7**, 1–11 (2016).
124. Zhang, X. *et al.* The tRNA methyltransferase NSun2 stabilizes p16^{INK4} mRNA by methylating the 3'-untranslated region of p16. *Nat. Comms.* **3**, 1–11 (2012).
125. Huber, S. M. *et al.* Formation and Abundance of 5-Hydroxymethylcytosine in RNA. *ChemBioChem* **16**, 752–755 (2015).
126. Haag, S. *et al.* NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. *EMBO J.* **35**, 2104–2119 (2016).
127. Nakano, S. *et al.* NSUN3 methylase initiates 5-formylcytidine biogenesis in human mitochondrial tRNAMet. *Nat. Chem. Biol.* **12**, 546–551 (2016).

128. Takemoto, C., Ueda, T. & Miura, K. Nucleotide sequences of animal mitochondrial tRNAsMet possibly recognizing both AUG and AUA codons. *Nucl. Ac. Symp. Series* **42**, 77–78 (1999).
129. Moriya, J., Yokogawa, T., Wakita, K. & Ueda, T. A novel modified nucleoside found at the first position of the anticodon of methionine tRNA from bovine liver mitochondria. *Biochemistry* **33**, 2234–2239 (1994).
130. Desrosiers, R. C., Friderici, K. H. & Rottman, F. M. Characterization of Novikoff hepatoma mRNA methylation and heterogeneity in the methylated 5' terminus. *Biochemistry* **14**, 4367–4374 (1975).
131. Desrosiers, R. & Friderici, K. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. in **71**, 3971–3975 (1974).
132. Wei, C. M., Gershowitz, A. & Moss, B. 5'-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. *Biochemistry* **15**, 397–401 (1976).
133. Wei, C. M. & Moss, B. Nucleotide sequences at the N^6 -methyladenosine sites of HeLa cell messenger ribonucleic acid. *Biochemistry* **16**, 1672–1676 (1977).
134. Horowitz, S., Horowitz, A., Nilsen, T. W., Munns, T. W. & Rottman, F. M. Mapping of N^6 -methyladenosine residues in bovine prolactin mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5667–5671 (1984).
135. Jia, G. *et al.* N^6 -Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 (2011).
136. Zheng, G. *et al.* ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Mol. Cell* **49**, 18–29 (2013).
137. Chen, K. *et al.* High-resolution N^6 -methyladenosine (m^6A) map using photo-crosslinking-assisted m^6A sequencing. *Angew. Chem. Int. Ed. Engl.* **54**, 1587–1590 (2015).
138. Fu, Y. *et al.* FTO-mediated formation of N^6 -hydroxymethyladenosine and N^6 -formyladenosine in mammalian RNA. *Nat. Comms.* **4**, 1798 (2013).
139. Gerken, T. *et al.* The Obesity-Associated FTO Gene Encodes a 2-Oxoglutarate-Dependent Nucleic Acid Demethylase. *Science* **318**, 1469–1472 (2007).
140. Bokar, J. A., Rath-Shambaugh, M. E., Ludwiczak, R., Narayan, P. & Rottman, F. Characterization and partial purification of mRNA N^6 -adenosine methyltransferase from HeLa cell nuclei. *J. Biol. Chem.* **269**, 17697–17704 (1994).
141. Ping, X.-L. *et al.* Mammalian WTAP is a regulatory subunit of the RNA N^6 -methyladenosine methyltransferase. *Cell Res.* **24**, 177–189 (2014).
142. Schwartz, S. *et al.* Perturbation of m^6A Writers Reveals Two Distinct Classes of mRNA Methylation at Internal and 5' Sites. *Cell Rep.* **8**, 284–296 (2014).
143. Liu, J. *et al.* A METTL3–METTL14 complex mediates mammalian nuclear RNA N^6 -adenosine methylation. *Nat. Chem. Biol.* **10**, 93–95 (2013).
144. Wang, X. *et al.* N^6 -methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2013).
145. Wang, X. *et al.* Structural basis for selective binding of m^6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* **10**, 1–4 (2014).
146. Alarcón, C. R. *et al.* HNRNPA2B1 Is a Mediator of m^6A -Dependent Nuclear RNA Processing Events. *Cell* **162**, 1299–1308 (2015).
147. Xu, C. *et al.* Structural Basis for the Discriminative Recognition of N^6 -Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J. Biol. Chem.* **290**, 24902–24913 (2015).
148. Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. The m^6A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol. Cell* **62**, 335–345 (2016).

149. Wang, Y. *et al.* N^6 -methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **16**, 1–10 (2014).
150. Batista, P. J. *et al.* A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. *Cell Stem Cell* **15**, 1–13 (2014).
151. Geula, S. *et al.* m^6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **347**, 1002–1006 (2015).
152. Zhao, B. S. & He, C. Fate by RNA methylation: m^6A steers stem cell pluripotency. *Genome Biol.* **16:43**, (2015).
153. Fustin, J.-M. *et al.* RNA-Methylation-Dependent RNA Processing Controls the Speed of the Circadian Clock. *Cell* **155**, 793–806 (2013).
154. Wang, C.-Y., Yeh, J.-K., Shie, S.-S., Hsieh, I.-C. & Wen, M.-S. Circadian rhythm of RNA N^6 -methyladenosine and the role of cryptochrome. *Biochem. Biophys. Res. Commun.* **465**, 88–94 (2015).
155. Wang, X. *et al.* N^6 -methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**, 1388–1399 (2015).
156. Zhou, J. *et al.* Dynamic m^6A mRNA methylation directs translational control of heat shock response. *Nature* **526**, 591–594 (2015).
157. Meyer, K. D. *et al.* 5'-UTR m^6A Promotes Cap-Independent Translation. *Cell* **163**, 999–1010 (2015).
158. Choi, J. *et al.* N^6 -methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat. Struct. Mol. Biol.* **2**, 110–115 (2016).
159. Roost, C. *et al.* Structure and Thermodynamics of N^6 -Methyladenosine in RNA: A Spring-Loaded Base Modification. *J. Am. Chem. Soc.* **137**, 2107–2115 (2015).
160. Kierzek, E. The thermodynamic stability of RNA duplexes and hairpins containing N^6 -alkyladenosines and 2-methylthio- N^6 -alkyladenosines. *Nucl. Acids Res.* **31**, 4472–4480 (2003).
161. Liu, N. *et al.* N^6 -methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature* **518**, 560–564 (2015).
162. Zhou, K. I. *et al.* N^6 -Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding. *J. Mol. Biol.* **428**, 822–833 (2016).
163. Xiao, W. *et al.* Nuclear m^6A Reader YTHDC1 Regulates mRNA Splicing. *Mol. Cell* **61**, 507–519 (2016).
164. Iwanami, Y. & Brown, G. M. Methylated bases of ribosomal ribonucleic acid from HeLa cells. *Arch. Biochem. Biophys.* **126**, 8–15 (1968).
165. Sierzputowska-Gracz, H., Gopal, H. D. & Agris, P. F. Comparative structural analysis of 1-methyladenosine, 7-methylguanosine, ethenoadenosine and their protonated salts IV: 1H , ^{13}C , and ^{15}N NMR studies at natural isotope abundance. *Nucl. Acids Res.* **14**, 7783–7801 (1986).
166. Macon, J. B. & Wolfenden, R. 1-Methyladenosine. Dimroth rearrangement and reversible reduction. *Biochemistry* **7**, 3453–3458 (1968).
167. Dominissini, D. *et al.* The dynamic N^1 -methyladenosine methylome in eukaryotic messenger RNA. *Nature* **530**, 441–446 (2016).
168. Li, X. *et al.* Transcriptome-wide mapping reveals reversible and dynamic N^1 -methyladenosine methylome. *Nat. Chem. Biol.* **12**, 311–316 (2016).
169. Aas, P. A. *et al.* Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421**, 859–863 (2003).
170. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2015).
171. Dedon, P. C. & Tannenbaum, S. R. Reactive nitrogen species in the chemical biology of inflammation. *Arch. Biochem. Biophys.* **423**, 12–22 (2004).

172. Yasui, M. *et al.* Miscoding Properties of 2'-Deoxyinosine, a Nitric Oxide-Derived DNA Adduct, during Translesion Synthesis Catalyzed by Human DNA Polymerases. *J. Mol. Biol.* **377**, 1015–1023 (2008).
173. Paul, M. S. & Bass, B. L. Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J.* **17**, 1120–1127 (1998).
174. Nigita, G., Veneziano, D. & Ferro, A. A-to-I RNA editing: current knowledge sources and computational approaches with special emphasis on non-coding RNA molecules. *Front. Bioeng. Biotechnol.* **3**, 1–7 (2015).
175. Wulff, B.-E., Sakurai, M. & Nishikura, K. Elucidating the inosinome: global approaches to adenosine- to-inosine RNA editing. *Nat. Rev. Genet.* **12**, 81–85 (2010).
176. Sakurai, M., Yano, T., Kawabata, H., Ueda, H. & Suzuki, T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat. Chem. Biol.* **6**, 733–740 (2010).
177. Sakurai, M. *et al.* A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. *Genome Res.* **24**, 522–534 (2014).
178. Deininger, P. Alu elements: know the SINEs. *Genome Biol.* **12**:236, (2011).
179. Chen, L.-L. & Carmichael, G. G. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle* **7**, 3294–3301 (2014).
180. Jepson, J. E. C. & Reenan, R. A. RNA editing in regulating gene expression in the brain. *Biochim. Biophys. Acta - Gene Regulatory Mechanisms* **1779**, 459–470 (2008).
181. Burrows, W. J. *et al.* Cytokinin from soluble RNA of *Escherichia coli*: 6-(3-methyl-2-butenylamino)-2-methylthio-9-β-D-ribofuranosylpurine. *Science* **161**, 691–693 (1968).
182. Koehler, K. M. & Schmidt, P. G. NMR study of the modified base resonances of tRNA^{tyr}_{coli}. *Biochem. Biophys. Res. Commun.* **50**, 370–376 (1973).
183. Harada, F. *et al.* 2-Methylthio N⁶-(Δ²-isopentenyl) adenosine: a component of *E. coli* tyrosine transfer RNA. *Biochem. Biophys. Res. Commun.* **33**, 299–306 (1968).
184. Hall, R. H., Robins, M. J., Stasiuk, L. & Thedford, R. Isolation of N⁶-(γ, γ-Dimethylallyl) adenosine from Soluble Ribonucleic Acid. *J. Am. Chem. Soc.* **88**, 2614–2615 (1966).
185. Robins, M. J., Hall, R. H. & Thedford, R. N⁶-(Δ²-Isopentenyl) adenosine. A Component of the Transfer Ribonucleic Acid of Yeast and of Mammalian Tissue, Methods of Isolation, and Characterization. *Biochemistry* **6**, 1837–1848 (1967).
186. Grosjean, H., Sprinzl, M. & Steinberg, S. Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* **77**, 139–141 (1995).
187. Juhling, F. *et al.* tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucl. Acids Res.* **37**, D159–D162 (2009).
188. Guillemaut, P., Martin, R. & Weil, J. H. Purification and base composition of a chloroplastic tRNA^{Phe} from *Phaseolus vulgaris*. *FEBS Lett.* **63**, 273–277 (1976).
189. Swaminathan, S. & Bock, R. M. Subcellular localization of cytokinins in transfer ribonucleic acid. *Plant Physiol.* **59**, 558–563 (1977).
190. Hecht, S. M., Leonard, N. J. & Burrows, W. J. Cytokinin of Wheat Germ Transfer RNA: 6-(4-Hydroxy-3-methyl-2-butenylamino)-2-methylthio-9-β-D-ribofuranosylpurine. *Science* **166**, 1272–1274 (1969).
191. Playtis, A. J. & Leonard, N. J. The synthesis of ribosyl-*cis*-zeatin and thin layer chromatographic separation of the *cis* and *trans* isomers of ribosylzeatin.

- Biochem. Biophys. Res. Commun.* **45**, 1–5 (1971).
192. Chapman, R. W., Morris, R. O. & Zaerr, J. B. Occurrence of *trans*-ribosylzeatin in *Agrobacterium tumefaciens* tRNA. *Nature* **262**, 153–154 (1976).
193. Thimmappaya, B. & Cherayil, J. D. Unique presence of 2-methylthio-ribosylzeatin in the transfer ribonucleic acid of the bacterium *Pseudomonas aeruginosa*. *Biochem. Biophys. Res. Commun.* **60**, 665–672 (1974).
194. Buck, M., McCloskey, J. A., Basile, B. & Ames, B. N. *cis* 2-Methylthio-ribosylzeatin (ms²io⁶A) is present in the transfer RNA of *Salmonella typhimurium*, but not *Escherichia coli*. *Nucl. Acids Res.* **10**, 5649–5662 (1982).
195. Grosjean, H., Nicoghosian, K. & Haumont, E. Nucleotide sequences of two serine tRNAs with a GGA anticodon: the structure-function relationships in the serine family of *E. coli* tRNAs. *Nucl. Acids Res.* **13**, 5697–5706 (1985).
196. Reiter, V. *et al.* The CDK5 repressor CDK5RAP1 is a methylthiotransferase acting on nuclear and mitochondrial RNA. *Nucl. Acids Res.* **40**, 6235–6240 (2012).
197. Kellner, S. *et al.* Profiling of RNA modifications by multiplexed stable isotope labelling. *Chem. Commun.* **50**, 3516–3518 (2014).
198. Wei, F.-Y. *et al.* Cdk5rap1-Mediated 2-Methylthio Modification of Mitochondrial tRNAs Governs Protein Translation and Contributes to Myopathy in Mice and Humans. *Cell Metabolism* **21**, 428–442 (2015).
199. Kaminska, K. H. *et al.* Structural bioinformatics analysis of enzymes involved in the biosynthesis pathway of the hypermodified nucleoside ms²io⁶A37 in tRNA. *Proteins* **70**, 1–18 (2007).
200. Xie, W., Zhou, C. & Huang, R. H. Structure of tRNA Dimethylallyltransferase: RNA Modification through a Channel. *J. Mol. Biol.* **367**, 872–881 (2007).
201. Chimnaronk, S. *et al.* Snapshots of Dynamics in Synthesizing N⁶-Isopentenyladenosine at the tRNA Anticodon. *Biochemistry* **48**, 5057–5065 (2009).
202. Zhou, C. & Huang, R. H. Crystallographic snapshots of eukaryotic dimethylallyltransferase acting on tRNA: Insight into tRNA recognition and reaction mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16142–16147 (2008).
203. Caillet, J. & Droogmans, L. Molecular cloning of the *Escherichia coli* *miaA* gene involved in the formation of Δ²-isopentenyl adenosine in tRNA. *J. Bacteriol.* **170**, 4147–4152 (1988).
204. Moore, J. A. & Poulter, C. D. *Escherichia coli* dimethylallyl diphosphate: tRNA dimethylallyltransferase: a binding mechanism for recombinant enzyme. *Biochemistry* **36**, 604–614 (1997).
205. Esberg, B., Leung, H., Tsui, H. & Björk, G. R. Identification of the *miaB* Gene, Involved in Methylthiolation of Isopentenylated A37 Derivatives in the tRNA of *Salmonella typhimurium* and *Escherichia coli*. *J. Bacteriol.* **181**, 7256–7265 (1999).
206. Hernández, H. L. *et al.* MiaB, a Bifunctional Radical-S-Adenosylmethionine Enzyme Involved in the Thiolation and Methylation of tRNA, Contains Two Essential [4Fe-4S] Clusters. *Biochemistry* **46**, 5140–5147 (2007).
207. Pierrel, F., Hernandez, H. L., Johnson, M. K., Fontecave, M. & Atta, M. MiaB Protein from *Thermotoga maritima*. *J. Biol. Chem.* **278**, 29515–29524 (2003).
208. Pierrel, F., Björk, G. R., Fontecave, M. & Atta, M. Enzymatic Modification of tRNAs. *J. Biol. Chem.* **277**, 13367–13370 (2002).
209. Forouhar, F. *et al.* Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases. *Nat. Chem. Biol.* **9**, 333–338 (2013).
210. Pierrel, F., Douki, T., Fontecave, M. & Atta, M. MiaB Protein Is a Bifunctional

- Radical-S-Adenosylmethionine Enzyme Involved in Thiolation and Methylation of tRNA. *J. Biol. Chem.* **279**, 47555–47563 (2004).
211. Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
212. Landgraf, B. J., Arcinas, A. J., Lee, K.-H. & Booker, S. J. Identification of an Intermediate Methyl Carrier in the Radical S-Adenosylmethionine Methylthiotransferases RimO and MiaB. *J. Am. Chem. Soc.* **41**, 15404–15416 (2013).
213. Hioe, J. & Zipse, H. Hydrogen Transfer in SAM-Mediated Enzymatic Radical Reactions. *Chem. Eur. J.* **18**, 16463–16472 (2012).
214. Grove, T. L. *et al.* A Radically Different Mechanism for S-Adenosylmethionine-Dependent Methyltransferases. *Science* **332**, 604–607 (2011).
215. Matschkal, D. Biochemie und chemischer Mechanismus der Thiomethylierung von RNA. 1–120 (2014).
216. Agris, P. F., Armstrong, D. J., Schäfer, K. P. & Söll, D. Maturation of a hypermodified nucleoside in transfer RNA. *Nucl. Acids Res.* **2**, 691–699 (1975).
217. Persson, B. C. & Björk, G. R. Isolation of the gene (*miaE*) encoding the hydroxylase involved in the synthesis of 2-methylthio-*cis*-ribozeatin in tRNA of *Salmonella typhimurium* and characterization of mutants. *J. Bacteriol.* **175**, (1993).
218. Buck, M. & Ames, B. N. A modified nucleotide in tRNA as a possible regulator of aerobiosis: synthesis of *cis*-2-methyl-thioribosylzeatin in the tRNA of *Salmonella*. *Cell* **36**, 523–531 (1984).
219. Mathevon, C. *et al.* tRNA-modifying MiaE protein from *Salmonella typhimurium* is a nonheme diiron monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13295–13300 (2007).
220. Esberg, B. & Björk, G. R. The methylthio group (ms^2) of N^6 -(4-hydroxyisopentenyl)-2-methylthioadenosine (ms^2io^6A) present next to the anticodon contributes to the decoding efficiency of the tRNA. *J. Bacteriol.* **177**, 1967–1975 (1995).
221. Persson, B. C., Ólafsson, Ó., Lundgren, H. K., Hederstedt, L. & Björk, G. R. The $ms^2io^6A_{37}$ modification of tRNA in *Salmonella typhimurium* regulates growth on citric acid cycle intermediates. *J. Bacteriol.* **180**, 3144–3151 (1998).
222. Corder, A. L. *et al.* Peroxide-Shunt Substrate-Specificity for the *Salmonella typhimurium* O₂-Dependent tRNA Modifying Monooxygenase (MiaE). *Biochemistry* **52**, 6182–6196 (2013).
223. Subedi, B. P., Corder, A. L., Zhang, S., Foss, F. W., Jr. & Pierce, B. S. Steady-State Kinetics and Spectroscopic Characterization of Enzyme–tRNA Interactions for the Non-Heme Diiron tRNA-Monooxygenase, MiaE. *Biochemistry* **54**, 363–376 (2015).
224. Anantharaman, V., Koonin, E. V. & Aravind, L. Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucl. Acids Res.* **30**, 1427–1464 (2002).
225. Sharova, L. V. *et al.* Database for mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Res.* **16**, 45–58 (2009).
226. Yang, E. *et al.* Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Res.* **13**, 1863–1872 (2003).
227. Endres, L., Dedon, P. C. & Begley, T. J. Codon-biased translation can be regulated by wobble-base tRNA modification systems during cellular stress responses. *RNA Biol.* **12**, 603–614 (2015).

228. Thompson, K. M. & Gottesman, S. The MiaA tRNA Modification Enzyme Is Necessary for Robust RpoS Expression in *Escherichia coli*. *J. Bacteriol.* **196**, 754–761 (2014).
229. Persson, B. C. Modification of tRNA as a regulatory device. *Mol. Microbiol.* **8**, 1011–1016 (1993).
230. Mittelman, A., Evans, J. T. & Chheda, G. B. Cytokinins as chemotherapeutic agents. *Ann. N.Y. Acad. Sci.* **255**, 225–234 (1975).
231. Dassano, A. *et al.* N^6 -isopentenyladenosine and analogs activate the NRF2-mediated antioxidant response. *Redox Biol.* **2**, 580–589 (2014).
232. Voller, J. *et al.* Anticancer activity of natural cytokinins: a structure-activity relationship study. *Phytochemistry* **71**, 1350–1359 (2010).
233. Laezza, C. *et al.* N^6 -isopentenyladenosine arrests tumor cell proliferation by inhibiting farnesyl diphosphate synthase and protein prenylation. *FASEB J.* **20**, 412–418 (2006).
234. MacLennan, B. D. Enzymatic demodification of transfer RNA species containing N^6 -(Δ^2 -isopentenyl)adenosine. *Biochem. Biophys. Res. Commun.* **65**, 345–351 (1975).
235. Goble, A. M., Fan, H., Sali, A. & Raushel, F. M. Discovery of a Cytokinin Deaminase. *ACS Chem. Biol.* **6**, 1036–1040 (2011).
236. Hall, R. H., Alam, S. N., McLennan, B. D., Terrine, C. & Guern, J. N^6 -(Δ^2 -isopentenyl) adenosine: Its conversion to inosine catalyzed by enzymes from Chicken bone marrow and Calf intestinal mucosa. *Canad. J. Biochem.* **49**, 623–630 (1971).
237. Chheda, G. B. & Mittelman, A. N^6 -(Δ^2 -isopentenyl) adenosine metabolism in man. *Biochem. Pharmacol.* **21**, 27–37 (1972).
238. Kopečný, D. *et al.* Kinetic and structural investigation of the cytokinin oxidase/dehydrogenase active site. *FEBS Journal* **283**, 361–377 (2015).
239. Nakajima, H., Yokota, T. & Takahashi, N. Metabolism and biosynthesis of cytokinins in tobacco crown gall cells. *Plant Cell Physiol.* **27**, 755–763 (1986).
240. Pačes, V., Werstiuk, E. & Hall, R. H. Conversion of N^6 -(Δ^2 -isopentenyl) adenosine to adenosine by enzyme activity in tobacco tissue. *Plant Physiol.* **48**, 775–778 (1971).
241. Vold, B. S., Longmire, M. E. & Keith, D. E. Thiolation and 2-methylthio-modification of *Bacillus subtilis* transfer ribonucleic acids. *J. Bacteriol.* **148**, 869–876 (1981).
242. Agris, P. F., Koh, H. & Söll, D. The effect of growth temperatures on the *in vivo* ribose methylation of *Bacillus stearothermophilus* transfer RNA. *Arch. Biochem. Biophys.* **154**, 277–282 (1973).
243. Arragain, S. *et al.* Identification of eukaryotic and prokaryotic methylthiotransferase for biosynthesis of 2-methylthio- N^6 -threonylcarbamoyl-adenosine in tRNA. *J. Biol. Chem.* **285**, 28425–28433 (2010).
244. Gefter, M. L. & Russell, R. L. Role modifications in tyrosine transfer RNA: a modified base affecting ribosome binding. *J. Mol. Biol.* **39**, 145–157 (1969).
245. Cabello-Villegas, J., Winkler, M. E. & Nikonowicz, E. P. Solution Conformations of Unmodified and $A_{37}N^6$ -dimethylallyl Modified Anticodon Stem-loops of *Escherichia coli* tRNA^{Phe}. *J. Mol. Biol.* **319**, 1015–1034 (2002).
246. Murphy, F. V., Ramakrishnan, V., Malkiewicz, A. & Agris, P. F. The role of modifications in codon discrimination by tRNA^{Lys}_{UUU}. *Nat. Struct. Mol. Biol.* **11**, 1186–1191 (2004).
247. Wilson, R. K. & Roe, B. A. Presence of the hypermodified nucleotide N^6 -(Δ^2 -isopentenyl)-2-methylthioadenosine prevents codon misreading by *Escherichia*

- coli* phenylalanyl-transfer RNA. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 409–413 (1989).
248. Díaz, I. & Ehrenberg, M. ms²i⁶A deficiency enhances proofreading in translation. *J. Mol. Biol.* **222**, 1161–1171 (1991).
249. Kirino, Y. & Suzuki, T. Human mitochondrial diseases associated with tRNA wobble modification deficiency. *RNA Biol.* **2**, 41–44 (2005).
250. Suzuki, T., Suzuki, T., Wada, T., Saigo, K. & Watanabe, K. Taurine as a constituent of mitochondrial tRNAs: new insights into the functions of taurine and human mitochondrial diseases. *EMBO J.* **21**, 6581–6589 (2002).
251. Kirino, Y., Goto, Y. & Campos, Y. Specific correlation between the wobble modification deficiency in mutant tRNAs and the clinical features of a human mitochondrial disease. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7127–7132 (2005).
252. Noer, A. S. *et al.* A tRNA^{Lys} mutation in the mtDNA is the causal genetic lesion underlying myoclonic epilepsy and ragged-red fiber (MERRF) syndrome. *Am. J. Hum. Genet.* **49**, 715–722 (1991).
253. Shoffner, J. M. *et al.* Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA^{Lys} mutation. *Cell* **61**, 931–937 (1990).
254. Levinger, L., Morl, M. & Florentz, C. Mitochondrial tRNA 3' end metabolism and human disease. *Nucl. Acids Res.* **32**, 5430–5441 (2004).
255. Tuppen, H. A. L., Blakely, E. L., Turnbull, D. M. & Taylor, R. W. Mitochondrial DNA mutations and human disease. *Biochim. Biophys. Acta* **1797**, 113–128 (2010).
256. Ishikawa, K. *et al.* Increased reactive oxygen species and anti-oxidative response in mitochondrial cardiomyopathy. *Circ. J.* **69**, 617–620 (2005).
257. Wang, H., Wei, L., Li, C., Zhou, J. & Li, Z. CDK5RAP1 deficiency induces cell cycle arrest and apoptosis in human breast cancer cell line by the ROS/JNK signaling pathway. *Oncol. Rep.* **33**, 1089–1096 (2015).
258. Shin, M.-K. *et al.* Association between CDK5RAP1 polymorphisms and susceptibility to vitiligo in the Korean population. *Eur. J. Dermatol.* **22**, 495–499 (2012).
259. Spinola, M. *et al.* Identification and functional characterization of the candidate tumor suppressor gene TRIT1 in human lung cancer. *Oncogene* **24**, 5502–5509 (2005).
260. Fleysher, M. H. N⁶-Substituted adenosines: Synthesis, Biological Activity, and Some Structure-Activity Relations. *J. Med. Chem.* **15**, 187–191 (1972).
261. Leonard, N. J., Hecht, S. M. & Skoog, F. Cytokinins: synthesis, mass spectra, and biological activity of compounds related to zeatin. *Biochemistry* **63**, 175–182 (1969).
262. Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
263. Steinthorsdottir, V. *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770–775 (2007).
264. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
265. Scott, L. J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
266. Ohara-Imaizumi, M. *et al.* Deletion of CDKAL1 Affects Mitochondrial ATP Generation and First-Phase Insulin Exocytosis. *PLoS ONE* **5**, e15553–11 (2010).
267. Jaffrey, S. R. An expanding universe of mRNA modifications. *Nat. Struct. Mol. Biol.* **21**, 945–946 (2014).

268. Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. RNA modifications: what have we learned and where are we headed? *Nat. Rev. Genet.* **17**, 365–372 (2016).
269. Liu, N. & Pan, T. *N*⁶-methyladenosine–encoded epitranscriptomics. *Nat. Struct. Mol. Biol.* **23**, 98–102 (2016).
270. Trinkle-Mulcahy, L. *et al.* Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* **183**, 223–239 (2008).
271. Ong, S.-E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **1**, 2650–2660 (2007).
272. Fleysler, M. H. Synthesis of [8-¹⁴C] *N*⁶-(Δ^2 -isopentenyl) adenosine. *J. Labelled Comp.* **8**, 455–460 (1972).
273. Kierzek, E. & Kierzek, R. The synthesis of oligoribonucleotides containing *N*⁶-alkyladenosines and 2-methylthio-*N*⁶-alkyladenosines via post-synthetic modification of precursor oligomers. *Nucl. Acids Res.* **31**, 4461–4471 (2003).
274. Hashizume, T. *et al.* Determination of Cytokinins by Mass Spectrometry Based on Stable Isotope Dilution. *Anal. Biochem.* **92**, 111–122 (1979).
275. Jordis, U., Grohmann, F. & Kuenburg, B. Optimized synthesis of some γ,γ -disubstituted allylamines. *Org. Prep. Proc. Int.* **29**, 549–560 (1997).
276. Overman, L. E. A general method for the synthesis of amines by the rearrangement of allylic trichloroacetimidates. 1,3-Transposition of alcohol and amine functions. *J. Am. Chem. Soc.* **98**, 2901–2910 (1976).
277. Robins, M. J. & Uznański, B. Nucleic acid related compounds. 33. Conversions of adenosine and guanosine to 2,6-dichloro, 2-amino-6-chloro, and derived purine nucleosides. *Canad. J. Chem.* **59**, 2601–2607 (1981).
278. Kreutz, C. *et al.* The Synthesis of 2'-O-[(Triisopropylsilyl)oxy] methyl (*TOM*) Phosphoramidites of Methylated Ribonucleosides (*m*¹*G*, *m*²*G*, *m*²₂*G*, *m*¹*I*, *m*³*U*, *m*⁴*C*, *m*⁶*A*, *m*⁶₂*A*) for Use in Automated *RNA* Solid-Phase Synthesis. *Monatsh. Chem.* **134**, 851–873 (2003).
279. Ottria, R., Casati, S., Baldoli, E., Maier, J. A. M. & Ciuffreda, P. *N*⁶-Alkyladenosines: Synthesis and evaluation of in vitro anticancer activity. *Bioorg. Med. Chem.* **18**, 8396–8402 (2010).
280. Serebryany, V. & Beigelman, L. An efficient preparation of protected ribonucleosides for phosphoramidite RNA synthesis. *Tet. Lett.* **43**, 1983–1985 (2002).
281. Ogilvie, K. K., Beaucage, S. L., Schiffman, A. L., Theriault, N. Y. & Sadana, K. L. The synthesis of oligoribonucleotides. II. The use of silyl protecting groups in nucleoside and nucleotide chemistry. VII. *Canad. J. Chem.* **56**, 2768–2780 (1978).
282. Ogilvie, K. K., Schiffman, A. L. & Penney, C. L. The synthesis of oligoribonucleotides. III. The use of silyl protecting groups in nucleoside and nucleotide chemistry. VIII. *Canad. J. Chem.* **57**, 2230–2238 (1979).
283. Beaucage, S. L. & Iyer, R. P. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* **48**, 2223–2311 (1992).
284. Markiewicz, W. T. & Wiewiórowski, M. A new type of silyl protecting groups in nucleoside chemistry. *Nucl. Acids Res.* **1**, s185–s190 (1978).
285. Jones, S. S. & Reese, C. B. Migration of *t*-butyldimethylsilyl protecting groups. *J. Chem. Soc., Perkin Trans. 1* **0**, 2762–3 (1979).
286. McBride, L. J., Kierzek, R. & Beaucage, S. L. Nucleotide chemistry. 16. Amidine protecting groups for oligonucleotide synthesis. *J. Am. Chem. Soc.* **108**, 2040–2048 (1986).

287. Barone, A. D., Tang, J. Y. & Caruthers, M. H. In situ activation of bis-dialkylaminophosphines—a new method for synthesizing deoxyoligonucleotides on polymer supports. *Nucl. Acids Res.* **12**, 4051-4061 (1984).
288. Dai, Q. *et al.* Identification of recognition residues for ligation-based detection and quantitation of pseudouridine and N6-methyladenosine. *Nucl. Acids Res.* **35**, 6322–6329 (2007).
289. Johnson, J. A., Jr, Thomas, H. J. & Schaeffer, H. J. Synthesis of potential anticancer agents. XIII. Ribosides of 6-substituted purines. *J. Am. Chem. Soc.* **80**, 699–702 (1958).
290. Kissman, H. M., Pidacks, C. & Baker, B. R. Puromycin. Synthetic Studies. XI. D-Ribofuranosyl Derivatives of 6-Dimethylaminopurine. *J. Am. Chem. Soc.* **77**, 18–24 (1955).
291. Guy, A., Molko, D., Wagrez, L. & Téoule, R. Chemical Synthesis of Oligonucleotides Containing N⁶-Methyladenine Residues in the GATC Site. *Helvetica chimica acta* **69**, 1034–1040 (1986).
292. Jones, J. W. & Robins, R. K. Purine nucleosides. III. Methylation studies of certain naturally occurring purine nucleosides. *J. Am. Chem. Soc.* **85**, 193–201 (1963).
293. Wan, Z.-K., Wacharasindhu, S., Binnun, E. & Mansour, T. An Efficient Direct Amination of Cyclic Amides and Cyclic Ureas. **8**, 2425–2428 (2006).
294. Aritomo, K., Wada, T. & Sekine, M. Alkylation of 6-N-acylated adenosine derivatives by the use of phase transfer catalysis. *J. Chem. Soc., Perkin Trans. 1* **14**, 1837–1844 (1995).
295. Dai, Q. & He, C. Reversible m⁶A RNA Modification. *The Glen Report* **26**, 1–3 (2014).
296. Wei, X. Coupling activators for the oligonucleotide synthesis via phosphoramidite approach. *Tetrahedron* **69**, 3615–3637 (2013).
297. Nurminen, E., Mattinen, J. K. & Lönnberg, H. The Effectivity of 1H-Triazoles and-Tetrazoles as Activators in Acid-Catalyzed Phosphoramidite Alcoholysis. *Helv. Chim. Acta* **86**, 2005–2008 (2003).
298. Berner, S., Mühlegger, K. & Seliger, H. Studies on the role of tetrazole in the activation of phosphoramidites. *Nucl. Acids Res.* **17**, 853–864 (1989).
299. Dahl, B. H., Nielsen, J. & Dahl, O. Mechanistic studies on the phosphoramidite coupling reaction in oligonucleotide synthesis. I. Evidence for nucleophilic catalysis by tetrazole and rate variations with the phosphorus substituents. *Nucl. Acids Res.* **15**, 1729–1743 (1987).
300. Vargeese, C. *et al.* Efficient activation of nucleoside phosphoramidites with 4,5-dicyanoimidazole during oligonucleotide synthesis. *Nucl. Acids Res.* **26**, 1046–1050 (1998).
301. Mann, M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
302. Lau, H.-T., Suh, H. W., Golkowski, M. & Ong, S.-E. Comparing SILAC- and Stable Isotope Dimethyl-Labeling Approaches for Quantitative Proteomics. *J. Proteome Res.* **13**, 4164–4174 (2014).
303. Vermeulen, M. *et al.* Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers. *Cell* **142**, 967–980 (2010).
304. Garber, K. B., Visootsak, J. & Warren, S. T. Fragile X syndrome. *Eur. J. Hum. Genet.* **16**, 666–672 (2008).
305. Mazroui, R. *et al.* Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Hum. Mol.*

- Genet.* **11**, 3007–3017 (2002).
306. Darnell, J. C. *et al.* FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. *Cell* **146**, 247–261 (2011).
307. Bassell, G. J. & Warren, S. T. Fragile X Syndrome: Loss of Local mRNA Regulation Alters Synaptic Development and Function. *Neuron* **60**, 201–214 (2008).
308. Chen, E. & Joseph, S. Fragile X mental retardation protein: A paradigm for translational control by RNA-binding proteins. *Biochimie* **114**, 147–154 (2015).
309. Kirkpatrick, L. L., McIlwain, K. A. & Nelson, D. L. Comparative Genomic Sequence Analysis of the FXR Gene Family: FMR1, FXR1, and FXR2. *Genomics* **78**, 169–177 (2001).
310. Gong, X. *et al.* MicroRNA-130b targets Fmr1 and regulates embryonic neural progenitor cell proliferation and differentiation. *Biochem. Biophys. Res. Commun.* **439**, 493–500 (2013).
311. Wu, Y., Zhu, J., Huang, X. & Du, Z. Crystal structure of a dimerization domain of human Caprin-1: insights into the assembly of an evolutionarily conserved ribonucleoprotein complex consisting of Caprin-1, FMRP and G3BP1. *Acta Cryst.* **D72**, 718–727 (2016).
312. Solomon, S. *et al.* Distinct Structural Features of Caprin-1 Mediate Its Interaction with G3BP-1 and Its Induction of Phosphorylation of Eukaryotic Translation Initiation Factor 2 α , Entry to Cytoplasmic Stress Granules, and Selective Interaction with a Subset of mRNAs. *Mol. Cell. Biol.* **27**, 2324–2342 (2007).
313. Kedersha, N. *et al.* G3BP–Caprin1–USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J. Cell Biol.* **212**, 845–860 (2016).
314. Bardoni, B. *et al.* 82-FIP, a novel FMRP (fragile X mental retardation protein) interacting protein, shows a cell cycle-dependent intracellular localization. *Hum. Mol. Genet.* **12**, 1689–1698 (2003).
315. Fatimy, El, R. *et al.* Fragile Mental Retardation Protein Interacts with the RNA-Binding Protein Caprin1 in Neuronal RiboNucleoProtein Complexes. *PLoS ONE* **7**, e39338–15 (2012).
316. Moutaoufik, M. T. *et al.* UVC-Induced Stress Granules in Mammalian Cells. *PLoS ONE* **9**, e112742–9 (2014).
317. Soncini, C., Berdo, I. & Draetta, G. Ras-GAP SH3 domain binding protein (G3BP) is a modulator of USP10, a novel human ubiquitin specific protease. *Oncogene* **20**, 3869–3879 (2001).
318. Zou, S. *et al.* N⁶-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. *Sci. Rep.* **6**, 1–12 (2016).
319. Kanamori, H., Dodson, R. E. & Shapiro, D. J. In vitro genetic analysis of the RNA binding site of vigilin, a multi-KH-domain protein. *Mol. Cell. Biol.* **18**, 3991–4003 (1998).
320. De Boulle, K. *et al.* A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat. Genet.* **3**, 31–35 (1993).
321. Zang, J. B. *et al.* A Mouse Model of the Human Fragile X Syndrome I304N Mutation. *PLoS Genet* **5**, e1000758–16 (2009).
322. Myrick, L. K. *et al.* Fragile X syndrome due to a missense mutation. *Eur. J. Hum. Genet.* **22**, 1185–1189 (2014).
323. Myrick, L. K., Hashimoto, H., Cheng, X. & Warren, S. T. Human FMRP contains an integral tandem Agenet (Tudor) and KH motif in the amino terminal domain. *Hum. Mol. Genet.* **24**, 1733–1740 (2015).
324. Ascano, M. *et al.* FMRP targets distinct mRNA sequence elements to regulate

- protein expression. *Nature* **492**, 382–386 (2012).
325. Ray, D. *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177 (2013).
326. Reineke, L. C., Kedersha, N., Langereis, M. A., van Kuppeveld, F. J. M. & Lloyd, R. E. Stress Granules Regulate Double-Stranded RNA-Dependent Protein Kinase Activation through a Complex Containing G3BP1 and Caprin1. *mBio* **6**, e02486–14 (2015).
327. Matsuki, H. *et al.* Both G3BP1 and G3BP2 contribute to stress granule formation. *Genes Cells* **18**, 135–146 (2012).
328. Anderson, N. K. A. P. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem. Soc. Trans.* **30**, 963–969 (2002).
329. Buchan, J. R. & Parker, R. Eukaryotic Stress Granules: The Ins and Outs of Translation. *Mol. Cell* **36**, 932–941 (2009).
330. Chen, E., Sharma, M. R., Shi, X., Agrawal, R. K. & Joseph, S. Fragile X Mental Retardation Protein Regulates Translation by Binding Directly to the Ribosome. *Mol. Cell* **54**, 407–417 (2014).
331. Wang, H. *et al.* Dynamic association of the fragile X mental retardation protein as a messenger ribonucleoprotein between microtubules and polyribosomes. *Mol. Biol. Cell* **19**, 105–114 (2008).
332. De Rubeis, S. & Bagni, C. Fragile X mental retardation protein control of neuronal mRNA metabolism: Insights into mRNA stability. *Mol. Cell. Neurosci.* **43**, 43–50 (2010).
333. Fukuda, T., Naiki, T., Saito, M. & Irie, K. hnRNP K interacts with RNA binding motif protein 42 and functions in the maintenance of cellular ATP level during stress conditions. *Genes Cells* **14**, 113–128 (2009).
334. Suvorova, E. S. *et al.* Discovery of a Splicing Regulator Required for Cell Cycle Progression. *PLoS Genet* **9**, e1003305–18 (2013).
335. Feng, C. *et al.* Crystal Structures of the Human RNA Demethylase Alkbh5 Reveal Basis for Substrate Recognition. *J. Biol. Chem.* **289**, 11571–11583 (2014).
336. Fedeles, B. I., Singh, V., Delaney, J. C., Li, D. & Essigmann, J. M. The AlkB Family of Fe(II)/ α -Ketoglutarate-dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. *J. Biol. Chem.* **290**, 20734–20742 (2015).
337. Koonin, E. V. Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucl. Acids Res.* **24**, 2411–2415 (1996).
338. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA Websuite. *Nucl. Acids Res.* **36**, W70–W74 (2008).
339. Zhao, J., Hyman, L. & Moore, C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445 (1999).
340. Brown, K. M. & Gilmartin, G. M. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor I_m. *Mol. Cell* **12**, 1467–1476 (2003).
341. Beadoing, E., Freier, S., Wyatt, J. R., Claverie, J. M. & Gautheret, D. Patterns of variant polyadenylation signal usage in human genes. *Genome Res.* **10**, 1001–1010 (2000).
342. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2008).
343. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.

- Nucl. Acids Res.* **37**, 1–13 (2009).
344. Yan, M. *et al.* A High-Throughput Quantitative Approach Reveals More Small RNA Modifications in Mouse Liver and Their Correlation with Diabetes. *Anal. Chem.* **85**, 12173–12181 (2013).
345. García-Sastre, A. 2 methylate or not 2 methylate: viral evasion of the type I interferon response. *Nat. Immun.* **12**, 114–115 (2011).
346. Salton, M., Lerenthal, Y., Wang, S. Y., Chen, D. J. & Shiloh, Y. Involvement of MatrIn 3 and SFPQ/NONO in the DNA damage response. *Cell Cycle* **9**, 1568–1576 (2010).
347. Udayakumar, D. & Dynan, W. S. Characterization of DNA binding and pairing activities associated with the native SFPQ·NONO DNA repair protein complex. *Biochem. Biophys. Res. Commun.* **463**, 473–478 (2015).
348. Li, S. *et al.* Involvement of p54(nrb), a PSF partner protein, in DNA double-strand break repair and radioresistance. *Nucl. Acids Res.* **37**, 6746–6753 (2009).
349. Guillaumond, F. *et al.* Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF. *FASEB J.* **25**, 2740–2756 (2011).
350. Zhang, Z. & Carmichael, G. G. The fate of dsRNA in the nucleus: a p54^{nrb}-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106**, 465–475 (2001).
351. Chen, L.-L. & Carmichael, G. G. Altered Nuclear Retention of mRNAs Containing Inverted Repeats in Human Embryonic Stem Cells: Functional Role of a Nuclear Noncoding RNA. *Mol. Cell* **35**, 467–478 (2009).
352. Zou, Y., Liu, Y., Wu, X. & Shell, S. M. Functions of human replication protein A (RPA): From DNA replication to DNA damage and stress responses. *J. Cell. Physiol.* **208**, 267–273 (2006).
353. Min, H., Turck, C. W., Nikolic, J. M. & Black, D. L. A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes Dev.* **11**, 1023–1036 (1997).
354. Holt, I. *et al.* Muscleblind-Like Proteins. *Am. J. Pathol.* **174**, 216–227 (2009).
355. Ho, T. H. *et al.* Muscleblind proteins regulate alternative splicing. *EMBO J.* **23**, 3103–3112 (2004).
356. Squillace, R. M., Chenault, D. M. & Wang, E. H. Inhibition of Muscle Differentiation by the Novel Muscleblind-Related Protein CHCR. *Dev. Biol.* **250**, 218–230 (2002).
357. Tacke, R., Tohyama, M., Ogawa, S. & Manley, J. L. Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. *Cell* **93**, 139–148 (1998).
358. Cowper, A. E., Cáceres, J. F., Mayeda, A. & Sreaton, G. R. Serine-Arginine (SR) Protein-like Factors That Antagonize Authentic SR Proteins and Regulate Alternative Splicing. *J. Biol. Chem.* **276**, 48908–48914 (2001).
359. Mili, S., Shu, H. J., Zhao, Y. & Piñol-Roma, S. Distinct RNP Complexes of Shuttling hnRNP Proteins with Pre-mRNA and mRNA: Candidate Intermediates in Formation and Export of mRNA. *Mol. Cell. Biol.* **21**, 7307–7319 (2001).
360. Couté, Y. *et al.* ISG20L2, a novel vertebrate nucleolar exoribonuclease involved in ribosome biogenesis. *Mol. Cell. Proteomics* **7**, 546–559 (2008).
361. Sharif, H. & Conti, E. Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. *Cell Rep.* **5**, 283–291 (2013).
362. Galao, R. P. *et al.* LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. *RNA* **16**, 817–827 (2010).
363. Chen, H. *et al.* RAVER1 is a coactivator of MDA5-mediated cellular antiviral response. *J. Mol. Cell Biol.* **5**, 111–119 (2013).

364. Jarrous, N. & Reiner, R. Human RNase P: a tRNA-processing enzyme and transcription factor. *Nucl. Acids Res.* **35**, 3519–3524 (2007).
365. Jarrous, N. Human ribonuclease P: subunits, function, and intranuclear localization. *RNA* **8**, 1–7 (2002).
366. Alexandrov, A. *et al.* Rapid tRNA Decay Can Result from Lack of Nonessential Modifications. *Mol. Cell* **21**, 87–96 (2006).
367. Anderson, J. *et al.* The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.* **12**, 3650–3662 (1998).
368. Chernyakov, I., Whipple, J. M., Kotelawala, L., Grayhack, E. J. & Phizicky, E. M. 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 (2008).
369. Dewe, J. M., Whipple, J. M., Chernyakov, I., Jaramillo, L. N. & Phizicky, E. M. The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. *RNA* **18**, 1886–1896 (2012).
370. Hernandez-Cid, A., Aguirre-Sampieri, S., Diaz-Vilchis, A. & Torres-Larios, A. Ribonucleases P/MRP and the Expanding Ribonucleoprotein World. *IUBMB Life* **64**, 521–528 (2012).
371. Tran, K. & Gralla, J. D. The TFIIB Tip Domain Couples Transcription Initiation to Events Involved in RNA Processing. *J. Biol. Chem.* **285**, 39580–39587 (2010).
372. Dreyfuss, G., Matunis, M. J. & Piñol-Roma, S. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* **62**, 289–321 (1993).
373. Shnyreva, M., Schullery, D. S., Suzuki, H., Higaki, Y. & Bomsztyk, K. Interaction of two multifunctional proteins. Heterogeneous nuclear ribonucleoprotein K and Y-box-binding protein. *J. Biol. Chem.* **275**, 15498–15503 (2000).
374. Lee, P.-T., Liao, P.-C., Chang, W.-C. & Tseng, J. T. Epidermal growth factor increases the interaction between nucleolin and heterogeneous nuclear ribonucleoprotein K/poly(C) binding protein 1 complex to regulate the gastrin mRNA turnover. *Mol. Biol. Cell* **18**, 5004–5013 (2007).
375. Ostareck, D. H. *et al.* mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3′ end. *Cell* **89**, 597–606 (1997).
376. Ostareck-Lederer, A. *et al.* c-Src-Mediated Phosphorylation of hnRNP K Drives Translational Activation of Specifically Silenced mRNAs. *Mol. Cell. Biol.* **22**, 4535–4543 (2002).
377. Ostareck-Lederer, A. & Ostareck, D. H. Precision mechanics with multifunctional tools: how hnRNP K and hnRNPs E1/E2 contribute to post-transcriptional control of gene expression in hematopoiesis. *Curr. Protein Pept. Sci.* **13**, 391–400 (2012).
378. Kim, J. H., Hahm, B., Kim, Y. K., Choi, M. & Jang, S. K. Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *J. Mol. Biol.* **298**, 395–405 (2000).
379. Hegele, A. *et al.* Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Mol. Cell* **45**, 567–580 (2012).
380. Paziewska, A., Wyrwicz, L. S., Bujnicki, J. M., Bomsztyk, K. & Ostrowski, J. Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Lett.* **577**, 134–140 (2004).
381. Habelhah, H. *et al.* ERK phosphorylation drives cytoplasmic accumulation of

- hnRNP-K and inhibition of mRNA translation. *Nat. Cell Biol.* **3**, 325–330 (2001).
382. Schullery, D. S. *et al.* Regulated interaction of protein kinase C δ with the heterogeneous nuclear ribonucleoprotein K protein. *J. Biol. Chem.* **274**, 15101–15109 (1999).
383. Chang, J. W., Koike, T. & Iwashima, M. hnRNP-K is a nuclear target of TCR-activated ERK and required for T-cell late activation. *Internat. Immunol.* **21**, 1351–1361 (2009).
384. Ostareck-Lederer, A. *et al.* Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *J. Biol. Chem.* **281**, 11115–11125 (2006).
385. Ostrowski, J., Wyrwicz, L., Rychlewski, L. & Bomsztyk, K. Heterogeneous Nuclear Ribonucleoprotein K Protein Associates with Multiple Mitochondrial Transcripts within the Organelle. *J. Biol. Chem.* **277**, 6303–6310 (2002).
386. Klimek-Tomczak, K., Wyrwicz, L. S., Jain, S., Bomsztyk, K. & Ostrowski, J. Characterization of hnRNP K Protein–RNA Interactions. *J. Mol. Biol.* **342**, 1131–1141 (2004).
387. Cao, W., Razanau, A., Feng, D., Lobo, V. G. & Xie, J. Control of alternative splicing by forskolin through hnRNP K during neuronal differentiation. *Nucl. Acids Res.* **40**, 8059–8071 (2012).
388. Zarnack, K. *et al.* Direct Competition between hnRNP C and U2AF65 Protects the Transcriptome from the Exonization of Alu Elements. *Cell* **152**, 453–466 (2013).
389. Havugimana, P. C. *et al.* A Census of Human Soluble Protein Complexes. *Cell* **150**, 1068–1081 (2012).
390. Spiller, M. P., Boon, K. L., Reijns, M. A. M. & Beggs, J. D. The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucl. Acids Res.* **35**, 923–929 (2007).
391. Weidensdorfer, D. *et al.* Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA* **15**, 104–115 (2008).
392. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **4**, 484–494 (2009).
393. Brown, T. A., Tkachuk, A. N. & Clayton, D. A. Mitochondrial Transcription Factor A (TFAM) Binds to RNA Containing 4-Way Junctions and Mitochondrial tRNA. *PLoS ONE* **10**, e0142436–22 (2015).
394. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc.* **2**, 1849–1861 (2007).
395. Mikula, M., Bomsztyk, K., Goryca, K., Chojnowski, K. & Ostrowski, J. Heterogeneous Nuclear Ribonucleoprotein (HnRNP) K Genome-wide Binding Survey Reveals Its Role in Regulating 3'-End RNA Processing and Transcription Termination at the Early Growth Response 1 (EGR1) Gene through XRN2 Exonuclease. *J. Biol. Chem.* **288**, 24788–24798 (2013).
396. Jameson, D. M. & Ross, J. A. Fluorescence Polarization/Anisotropy in Diagnostics and Imaging. *Chem. Rev.* **110**, 2685–2708 (2010).
397. Pagano, J. M., Clingman, C. C. & Ryder, S. P. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* **17**, 14–20 (2010).
398. Lundblad, J. R., Laurance, M. & Goodman, R. H. Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Mol. Endocrin.* **10**, 607–612 (1996).
399. Chang, Y.-I. *et al.* Identification of the methylation preference region in

- heterogeneous nuclear ribonucleoprotein K by protein arginine methyltransferase 1 and its implication in regulating nuclear/cytoplasmic distribution. *Biochem. Biophys. Res. Commun.* **404**, 865–869 (2011).
400. Holden, N. S. & Tacon, C. E. Principles and problems of the electrophoretic mobility shift assay. *J. Pharmacol. Toxicol. Methods* **63**, 7–14 (2011).
401. Eliseeva, I. A., Kim, E. R., Guryanov, S. G., Ovchinnikov, L. P. & Lyabin, D. N. Y-box-binding protein 1 (YB-1) and its functions. *Biochemistry (Moscow)* **76**, 1402–1433 (2012).
402. Lyabin, D. N., Eliseeva, I. A. & Ovchinnikov, L. P. YB-1 protein: functions and regulation. *WIREs RNA* **5**, 95–110 (2013).
403. Gama-Carvalho, M., Carvalho, M. P., Kehlenbach, A., Valcarcel, J. & Carmo-Fonseca, M. Nucleocytoplasmic Shuttling of Heterodimeric Splicing Factor U2AF. *J. Biol. Chem.* **276**, 13104–13112 (2000).
404. Carette, J. E. *et al.* Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**, 340–343 (2011).
405. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
406. Aoki, K. *et al.* LARP1 specifically recognizes the 3' terminus of poly(A) mRNA. *FEBS Lett.* **587**, 2173–2178 (2013).
407. Tcherkezian, J. *et al.* Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. *Genes Dev.* **28**, 357–371 (2014).
408. Fonseca, B. D. *et al.* La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). *J. Biol. Chem.* **290**, 15996–16020 (2015).
409. Iadevaia, V., Caldarola, S., Tino, E., Amaldi, F. & Loreni, F. All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. *RNA* **14**, 1730–1736 (2008).
410. Hamilton, T. L., Stoneley, M., Spriggs, K. A. & Bushell, M. TOPs and their regulation. *Biochem. Soc. Trans.* **34**, 12–16 (2006).
411. de Lange, T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* **19**, 2100–2110 (2005).
412. Kim, J. & Chung, I. K. The splicing factor U2AF65 stabilizes TRF1 protein by inhibiting its ubiquitin-dependent proteolysis. *Biochem. Biophys. Res. Commun.* **443**, 1124–1130 (2014).
413. O'Brien, K., Matlin, A. J., Lowell, A. M. & Moore, M. J. The Biflavonoid Isoginkgetin Is a General Inhibitor of Pre-mRNA Splicing. *J. Biol. Chem.* **283**, 33147–33154 (2008).
414. Cavaluzzi, M. J. & Borer, P. N. Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucl. Acids Res.* **32**, 13e–9 (2004).
415. Katzen, F. Gateway® recombinational cloning: a biological operating system. *Expert Opin. Drug Discov.* **2**, 571–589 (2007).
416. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA cloning using in vitro site-specific recombination. *Genome Res.* **10**, 1788–1795 (2000).
417. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).