technische universität dortmund

fakultät für chemie und chemische biologie

MAX PLANCK INSTITUTE OF MOLECULAR PHYSIOLOGY

CHEMICAL GENOMICS CENTRE OF THE MAX PLANCK SOCIETY

Targeting protein-RNA interactions with peptide inhibitors

Dissertation

for attainment of the academic degree of

Doctor in Natural Sciences

Dr. rer. nat.

Submitted to

Department of Chemistry and Chemical Biology

TU Dortmund University

Submitted by

Gulshan Amrahova

Dortmund, 14th of August, 2024

The work in this thesis was carried out starting from October 2019 to October 2023 and supervised by Prof. Dr. Dr. h.c. Herbert Waldmann and Dr. Peter ‘t Hart at the TU Dortmund University, Faculty of Chemistry and Chemical Biology and the Max Planck Institute of Molecular Physiology.

This research was funded by German Academic Exchange Service (DAAD), the Max-Planck Institute of Molecular Physiology, Pfizer, AstraZeneca, and Merck KGaA.

Dean: Prof. Dr. Stefan M. Kast

First Referee: Prof. Dr. Dr. h.c. Herbert Waldmann

Second Referee: Prof. Dr. Susanne Brakmann

Table of Contents

I. Acknowledgement v
II. Abstract vii
III. Zusammenfassung vii
IV. List of Abbreviations viii
1. Introduction 1
1.1. The complex lives of mRNAs, RNA splicing and alternative splicing 1
1.2. RNA-binding proteins, roles of RBPs in RNA life cycle 3
1.2.1. How RNA-binding proteins interact with RNA 4
1.3. Roles of RBPs in diseases 4
1.3.1. RBPs in diabetes and cardiovascular disease 4
1.3.2. Roles of RBPs in cancer 5
1.3.3. RBPs and their role in neurodegenerative disease 5
1.4. Peptides as RNA-protein interaction modulators 6
1.5. Peptide cyclization strategies 8
1.5.1. Amide bond formation 8
1.5.2. Disulfide cyclization 8
1.5.3. C-C double bond formation: Alkene metathesis 9
1.6. Cyclic peptide libraries 9
1.6.1. Genetically encoded libraries 10
1.7. Peptide nucleic acids 12
1.8. Allosteric regulation, regulation with small molecules and peptides 13
2. Modulating the interaction between RBM20 and RNA 15
2.1. Introduction 15
2.2. Aim 17
2.3. Results and discussions 18
2.3.1. Targeting RBM20-RNA interaction with stapled peptides 18
2.3.2. Design of peptides 19
2.3.3. Synthesis of the linear and stapled peptides 19
2.3.4. Circular dichroism spectroscopy of the stapled peptides 21
2.3.5. Direct fluorescence polarization with RNAs 23
2.3.6. Direct fluorescence polarization assay with stapled peptides 25
2.3.7. Competition fluorescence polarization with RNA 28
2.3.8. Evaluation of the peptides in competition FP assay 29
2.3.9. Point mutations for further optimization of peptides 32

2.3.10. Fluorescence polarization and total intensity 33
3. Targeting the interaction of hnRNP A2/B1 with RNA 36
3.1. Introduction 36
3.2. Aim 39
3.3. Results 40
3.3.1. Targeting hnRNP A2/B1-RNA interaction with peptide nucleic acids 40
3.3.2. Peptide nucleic acid synthesis 40
3.3.3. Direct fluorescence polarization assay with RNAs 41
3.3.4. Competition FP experiment with RNAs 43
3.3.5. Direct fluorescence polarization assay with PNAs 44
3.3.6. PNA monomer synthesis 45
3.3.7. Synthesis of 10mer and 8mer PNAs 45
3.3.8. Competition experiment with mononucleotides and PNA variants 47
3.3.9. Identification of cyclic peptide inhibitors of the hnRNP A2/B1 49
3.3.10. Synthesis of the macrocyclic peptides of the hnRNP A2/B1 49
3.3.11. Direct fluorescence polarization assay with RNA 51
3.3.12. Evaluation of the hnRNP A2/B1 peptides in competition FP assay 53
3.3.13. Alanine scanning for the hit peptide selected from the SICLOPPS screening 56
3.3.14. Evaluation of the hit peptides in competition FP assay 57
3.3.15. Microscale thermophoresis with selected hits from SICLOPPS and alanine scanning 61
3.3.16. Investigation of the interaction of hnRNP A2/B1 peptide with RRM1 and RRM2 domains 63
3.3.17. Investigation of hnRNP A2/B1 and RNA interaction with mass photometry ... 64
3.4. Summary and outlook 67
4. Materials, reagents, chemicals, devices 68
4.1. Materials 68
4.2. Devices 68
4.3. Software 69
4.4. Reagents 69
4.5. Medium and buffers 71
4.6. Protein buffers 72
4.7. Assay buffers 74
4.8. Antibiotics and induction agents 75
4.9. RNAs 76
4.10. Amino acids 75
4.11. PNA monomers 76

4.12. SPPS reagents/solvents 76
5. Methods 77
5.1. Preparation of bacterial cultures 77
5.2. Preparation of glycerol stock 77
5.3. Bacterial Transformation 77
5.4. RBM20 protein expression and purification 77
5.4.1. Expression of RBM20 delta_a3 77
5.4.2. Purification of RBM20 delta_a3 77
5.4.3. Expression of RBM20_a3 78
5.4.4. Purification of RBM20_a3 78
5.5. hnRNP A2/B1 protein expression and purification 78
5.5.1. MBP-tagged hnRNP A2/B1 expression 78
5.5.2. MBP-tagged hnRNP A2/B1 purification 78
5.6. Protein concentration determination by Bradford assay 79
5.7. Direct fluorescence polarization assay 79
5.8. Competition Fluorescence polarization 79
5.9. Solid Phase Peptide Synthesis 79
5.9.1. Loading amino acid on Rink amide resin 79
5.9.2. Loading amino acid to ChemMatrix resin 80
5.9.3. Loading amino acid to 2-chlorotrityl chloride (2-CTC) resin 80
5.9.4. 2-CTC resin reactivation 80
5.9.5. Resin loading determination 80
5.10. Synthesis of the RBM20 peptides on solid phase 81
5.10.1. C-terminal FITC labeling 81
5.10.2. Stapled Peptide synthesis 81
5.10.3. RBM20 peptide cleavage from solid support 82
5.11. Peptide nucleic acid synthesis 82
5.11.1. Cleavage 82
5.11.2. Synthesis of PNA variants 82
5.11.3. PNA monomer synthesis 83
5.12. Synthesis of the hnRNP A2/B1 and SRSF1 peptides 83
5.12.1. Linear peptide synthesis 83
5.12.2. Cleavage of the peptides from solid support 83
5.12.3. Cyclization in solution 84
5.12.4. Cleavage of protecting groups 84
5.13. Peptide purification 84

5.13.1. Preparative RP-HPLC ... 84
5.13.2. LC-MS purity analysis ... 84
5.13.3. High-resolution mass spectrometry ... 84
5.14. Circular Dichroism Spectroscopy ... 84
5.15. Microscale thermophoresis ... 84
5.16. Mass photometry ... 85
6. Literature ... 86
7. Appendix ... 99
7.1. Table of figures ... 99
7.2. Table of tables ... 104
7.3. Table of supplementary ... 105
7.4. Supplementary figures ... 106
7.4.1. RBM20 peptide sequences, masses (calculated and determined) and purities 106
7.4.2. PNA structures, masses (calculated and determined) and purity data ... 115
7.4.3. hnRNP A2/B1 peptide purity data ... 118
7.4.4. Mass photometry figures ... 125
7.4.5. Used RNAs ... 125
7.4.6. CD data ... 126

I. Acknowledgement

I want to start by thanking Prof. Dr. Dr. h.c. Herbert Waldmann for offering me the chance to do my Ph.D. in Department IV and CGC and for accepting to be the first reviewer of my doctoral dissertation.

I thank Prof. Dr. Susanne Brakmann for accepting to be the second examiner for my thesis.

Then, I would like to express my deepest thanks to my supervisor Dr. Peter ‘t Hart, who gave me the chance to be a part of his research group and offered me a great project to work on. I highly appreciate my time spent within the group, where I had a great chance to be involved in scientific discussions and learned new techniques, which all helped me to bring my experience to a higher level. I would also like to thank Peter for always being reachable and open to discussions on any topic. Additionally, I want to thank Peter for investing time to read this thesis.

Next, I highly appreciate being a part of AG Hart group and want to express my thanks for being an openminded group, giving a lot of support to each other, successful team work and beneficial discussions. I especially want to thank Sunit Pal, Joseph Openy, Jen-Yao Chang, and Jessica Nowacki for being good collaboration partners.

In addition, I am very thankful for the financial support from DAAD, which enabled me to perform my doctoral studies successfully.

I also want to thank my lab mates Adrian Krzyzanowski, Romain Tessier, Sasikala Thawam, Cristina Alcala, and Hikaru Aoyama, with whom we had a lot of discussions, from which I took a lot.

I want to thank to Jens Warmers for always being available and solving any problems related to the instruments at a very good time.

I want to thank the protein facility, especially Raphael Gasper-Schönenbrücher and Petra Geue, for all the help and input while using the instrumental devices from the department.

Arina Khamzina is the next person I want to thank for her contributions to the project.

I highly appreciate having a chance to work both in CGC and Department IV and be acquainted with many people with whom I had a great time.

I want to specifically mention that, it was an excellent opportunity to access all the instruments available at Max-Planck Institute. This saved me a lot of time and allowed me to meet new people from different departments.

I also want to thank Christa Hornemann for all her support with my concerns on several non-scientific topics. Also, thank Frau Rose and Birgit Apprecht for helping with the documentation

and paperwork. Thank Debora Bruzzese for her perfect organization of various events within the CGC.

In addition, I thank the TGIF team for organizing get-together events and gathering us all together, which I enjoyed my time a lot.

Lastly, I want to thank my family for their full support, understanding, and patience. Without their support, it would be challenging to reach this stage.

II. Abstract

Protein-RNA interactions (PRIs) play an essential role in many cellular functions. Improved methods are needed to systematically develop modulators of PRIs. Therefore, having a widely applicable strategy to develop methods that target such interactions is highly valuable.

RNA binding motif 20 (RBM20) is a protein that plays a vital role in heart function and mutations in RBM20 associated with dilated cardiomyopathy (DCM). Previous studies have identified the structure of the RBM20 domain that is bound to the specific RNA sequence. RNA binding is regulated by an RNA recognition motif (RRM), with a well-folded C-terminal helix. Targeting RNA-RBM20 interaction can be a good starting approach in treating cardiac disease, yet direct targeting of the PRI can be challenging. However, allosteric inhibition of the target protein would provide a way of selective inhibition. Stapled peptides of RBM20 were synthesized and evaluated in the biophysical experiments. As a result, synthesized peptides showed binding affinity to the target protein. Further evaluation needs to be carried out for the effect of stapled peptides on titin splicing and on how alternative splicing relates to the diastolic heart failure.

Heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNP A2/B1), which is also an RNA binding protein, is upregulated in various types of cancer where it affects the splicing of tumor suppressors and oncogenes. The target makes an attractive starting point as it shows a promising mode of action in cancer therapy, and several crystal structures of its interaction with RNA are available. Peptide nucleic acids (PNAs) are proposed to be used as a scaffold to carry fragments that target a PRI binding site. By exchanging the PNA nucleobases for small molecule fragments, novel chemical matter with the potential to inhibit the interaction can be identified. In addition, another strategy was explored, called translational repression assay procedure (TRAP), to track the recognition of the RNA by hnRNP A2/B1. This assay was then employed in combination with the split-intein circuit ligation of peptides and proteins (SICLOPPS) technique for the screening of the macrocyclic peptides as inhibitors for hnRNP A2/B1. Selected peptides were first synthesized and then evaluated by applying biophysical experiments that defined two peptides that are binding to hnRNP A2/B1. Optimization through alanine scanning revealed a derivate with a ~3-fold increase in binding affinity. Further investigation and analysis need to be carried out to obtain a potential peptide inhibitor of hnRNP A2/B1.

III. Zusammenfassung

Protein-RNA-Interaktionen (PRIs) spielen eine wesentliche Rolle bei vielen zellulären Funktionen. Es werden daher Methoden benötigt, um Modulatoren zu entwickeln, die systematisch PRIs beeinflussen können. Die Entwicklung einer allgemein anwendbaren Strategie zur Modulation solcher PRIs wäre sehr wertvoll.

Das RNA-Bindungsmotiv 20 (RBM20) ist ein Protein, das eine wichtige Rolle bei der Herzfunktion spielt, und Mutationen in RBM20 werden mit dilatativer Kardiomyopathie (DCM) in Verbindung gebracht. Frühere Studien haben die Struktur der RBM20-Domäne identifiziert, die an die spezifische RNA-Sequenz gebunden ist. Die RNA-Bindung wird durch ein RNA-Erkennungsmotiv (RRM), die eine C-terminalen Helix aufweist, reguliert. Die gezielte Beeinflussung der RNA-RBM20-Interaktion kann ein guter Ansatz für die Behandlung von Herzkrankheiten sein, jedoch kann die direkte Beeinflussung der PRI eine Herausforderung darstellen. Die allosterische Hemmung des Zielproteins würde eine Möglichkeit der selektiven Hemmung bieten. In dieser Arbeit wurden stabilisierte Peptide der C-terminalen Helix von RBM20 synthetisiert und die Bindungsaffinität mittels biophysikalischeer Experimente untersucht. Die Wirkung der stabilisierten Peptide auf alternative Spleißenprozesse im Zusammenhang mit diastolischer Herzinsuffizienz müssen noch weiter untersucht werden.

Das heterogene nukleare Ribonukleoprotein A2/B1 (hnRNP A2/B1), welches ebenfalls ein RNA-bindendes Protein ist, wird bei verschiedenen Krebsarten hochreguliert, wo es das Spleißen von Tumorsuppressorgenen und Onkogenen beeinflusst. hnRNP A2/B1 ist ein attraktivess Zielprotein, da es eine vielversprechende Wirkungsweise in der Krebstherapie aufweist und bereits mehrere Kristallstrukturen seiner Interaktion mit RNA verfügbar sind. In zweiten Teil dieser Arbeit werden Peptidnukleinsäuren (PNA) als Gerüst für Fragmente verwendet, die auf eine PRI-Bindungsstelle abzielen. Durch den Austausch der PNA-Nukleobasen gegen kleine Molekülfragmente können neue chemische Moleküle gebildet werden, die das Potenzial haben, die Interaktion zu hemmen. Darüber hinaus wurde eine weitere Strategie erforscht, die die “translational repression assay procedure“ (TRAP)mit der biotechnologischen Methode „Split-intein Circuit Ligation of Peptides and Proteins“ (SICLOPPS) kombinierte, um nach makrozyklischen Peptiden als Inhibitoren für hnRNP A2/B1 zu screenen. Ausgewählte Peptide wurden zunächst chemsich synthetisiert und dann mit Hilfe von biophysikalischen Experimenten bewertet. Es wurden zwei vielversprechende Peptidkandidaten gefunden, die eine moderate Affinität an hnRNP A2/B1 aufwiesen. Die Optimierung durch Alanin-Scanning ergab ein Derivat mit einer ~3-fachverbesserter Bindungsaffinität. Es müssen jedoch weitere Untersuchungen durchgeführt werden, um den Peptidkandidaten als Inhibitor von hnRNP A2/B1 zu verifizieren.

IV. List of Abbreviations

3'-UTR	Three prime untranslated region
3' SS	3' splice site
5' SS	5' splice site
ALS	Amyotrophic lateral sclerosis
AS	Alternative splicing
BP	Branch point
CCR5	C-C chemokine receptor type 5
cDNA	complementary DNA
CD	Circular dichroism
CLIP	Cross Linking and Immunoprecipitation
CPP	Cell penetrating peptides
CVD	Cardiovascular disease
CtBP	C-terminal binding protein
DCM	Dilated cardiomyopathy
DECLs	DNA-encoded chemical libraries
DM	Diabetes mellitus
dsRBD	Double stranded RNA-binding domain
DTS	DNA-templated synthesis
ECs	Endothelial cells
eIF4E	Eukaryotic initiation factor 4E
ESEs	Exonic splicing enhancers
ESRPs	Epithelial splicing regulatory proteins
ESSs	Exonic splicing silencers
FACS	Fluorescence activated cell sorting

FDA	US Food and Drug Administration
FH	Fractional helicity
GAGs	Glycosaminoglycans
GPR55	G protein-coupled receptor 55
HDAg	Hepatitis delta antigen
HER2	Human epidermal growth factor receptor 2
HFpEF	Heart failure with preserved ejection fraction
HGF	Hepatocyte growth factor
HIF	Hypoxia inducible factor
hnRNPs	Heterogeneous nuclear ribonucleoproteins
HOTAIR	HOX Transcript Antisense Intergenic RNA
HuR	Human antigen R
IAjxn	I-band-A-band junction
IGF2BP1	Insulin like growth factor 2 mRNA binding protein 1
EMT	Epithelial-mesenchymal transition
Intein	Internal protein segment
IP	Immunoprecipitation
ISEs	Intronic splicing enhancers
ISSs	Intronic splicing silencers
KH domain	K-homology domain
LPI	Lysophosphatidylinositol
MP	Mass photometry
miRNA	MicroRNA
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein particles
NAM	Negative allosteric modulation

NGS	Next generation sequencing
OBOC	One-bead one-compound
PAM	Positive allosteric modulation
PPIs	Protein-protein interactions
pre-mRNA	Precursor mRNA
PRI	Protein-RNA interaction
RBD	RNA binding domain
RBM20	RNA recognition motif 20
RBP	RNA binding protein
RCM	Arg-Gly-Gly rich box
RGG	Ring closing metathesis
RIP	RNA immunoprecipitation
RNA	Ribonucleic acid
RRM	RNA recognition motif
RTHS	Reverse two-hybrid system
RT-PCR	Reverse transcription polymerase chain reaction
SICLOPPS	Split-intein circuit ligation of peptides and proteins
SMA	Spinal muscular atrophy
SMN2	Survival of Motor Neuron 2
snRNPs	Small nuclear ribonucleoproteins
SR	Serine arginine rich proteins
SRSF1	Serine/arginine-rich splicing factor 1
STAR	Signal transduction and activation of RNA
TIA1	T cell-restricted intracellular antigen 1
TGF-β	Transforming growth factor beta
UV	Ultraviolet

VEGF	Vascular endothelial growth factor
VdW	Van der Waals
ZnF	Zinc finger

1. Introduction

1.1. The complex lives of mRNAs, RNA splicing and alternative splicing

In eukaryotes, transcription and translation are performed separately in the nucleus and the cytoplasm, respectively, which allows eukaryotes to apply post-transcriptional processing of pre-mRNA, producing a variety of mRNAs[1]. mRNAs are first generated in the nucleus as pre-mRNAs, which are then subjected to several processing steps, including 5'-end capping, splicing, 3'-end cleavage, and polyadenylation. Then, mature mRNAs are exported from the nucleus and transported into the cytoplasm, where translation by ribosome takes place, resulting in protein synthesis[2]. Often, protein-coding sequences are interrupted by introns (non-coding segments), which must be removed from pre-mRNA, and as a result, exons (protein-coding segments) must be joined to form mature mRNA[3]. This process is known as pre-mRNA splicing. Introns are characterized by three major sites, which are the 3' splice site (3' SS), 5' splice site (5' SS), and branch point (BP). The splicing process takes place between the transcription and translation steps and happens through two transesterification reactions catalyzed by the spliceosome[4]. It starts with the attack of the 2' hydroxyl group of an adenosine residue to the 5' SS resulting in a cleaved 5' exon and a lariat intermediate. In the second step, the 3' hydroxyl group attacks the phosphodiester at the 3' splice site, forming the lariat intron and ligated exon (Figure 1)[5,6].

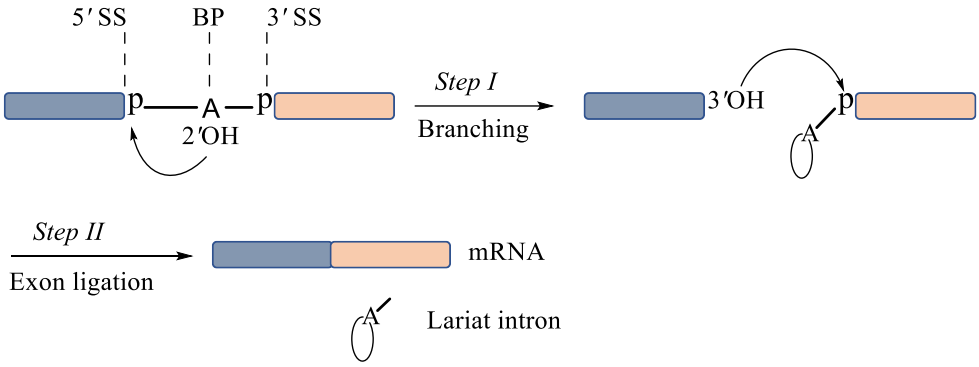


Figure 1: RNA splicing process. Pre-mRNA is transformed into spliced mRNA, in which introns are removed through two transesterification reactions: branching and exon ligation. Adapted from Wilkinson et al[3].

Alternative splicing (AS) which is a crucial step in mRNA processing, leads to the production of various mRNA variants from a single gene[7,8]. Alternative splicing is an important process leading to gene expression complexity and results in cellular differentiation. Alternative splicing has six main types: exon skipping, mutually exclusive exons, alternative 3' splice site, alternative 5' splice site, back splicing, and intron retention (Figure 2).

In the most frequent alternative splicing events, particular exons are skipped, forming new mRNA variants. Nevertheless, based on the latest investigations, intron retentions also appear in 75% of the genes[9]. The relative abundance of each type of AS event differs remarkably

between eukaryotes[10,11]. For instance, in metazoans, exon skipping is widespread, while in plants and fungi, intron retention is most abundant[12,13].

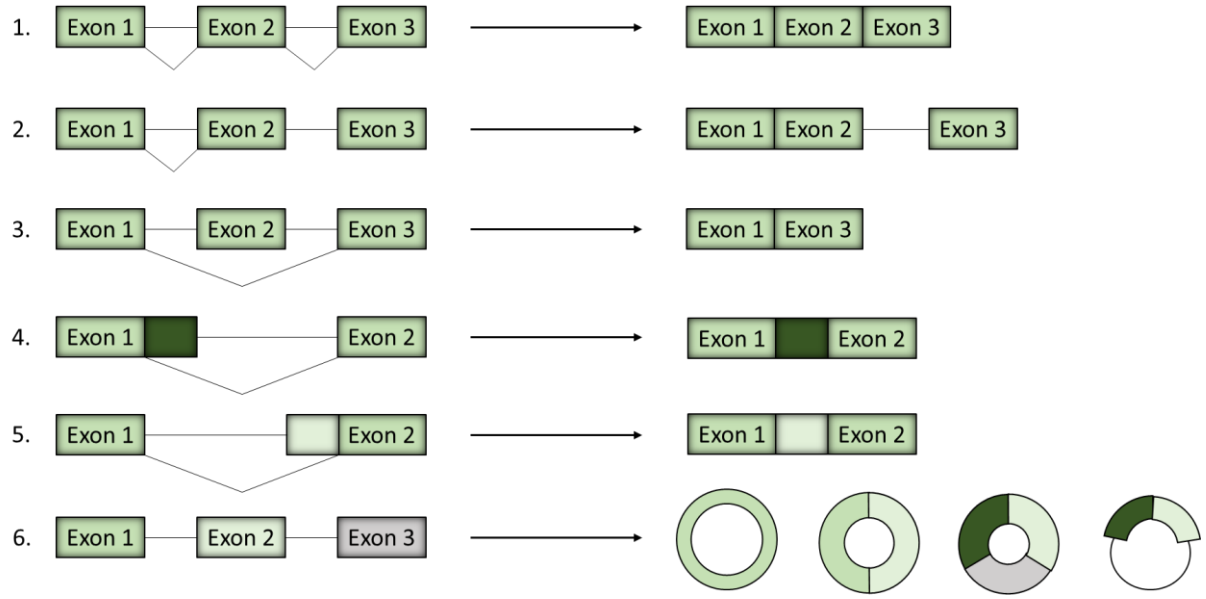


Figure 2: Constitutive splicing and five different modes of alternative splicing: 1. Constitutively splicing 2. Intron retention 3. Exon skipping 4. Alternative 5' splice site 5. Alternative 3' splice site 6. Back splicing. Adapted from Ren et al[14].

The splicing process has two main steps: Spliceosome assembly and actual splicing of the pre-mRNA[15]. The spliceosome itself recognizes consensus sequences at 5'SSs, 3'SSs, and BP sites. Alternative splicing can be regulated by cis-regulatory and trans-regulatory factors. Cis-regulatory factors comprise exonic (ESEs) and intronic splicing enhancers (ISEs) which are the sequences within the exons and introns, respectively[16,17]. They interact with positive-acting regulatory elements, for example, serine/arginine-rich family proteins (SR), while exonic splicing silencers (ESSs) and intronic splicing silencers (ISS), interact with negative-acting regulatory elements, for instance, heterogeneous nuclear ribonucleoproteins (hnRNPs) (Figure 3)[18]. Cooperation between these elements causes enhancement or inhibition of spliceosome assembly at weak splice sites[19,20].

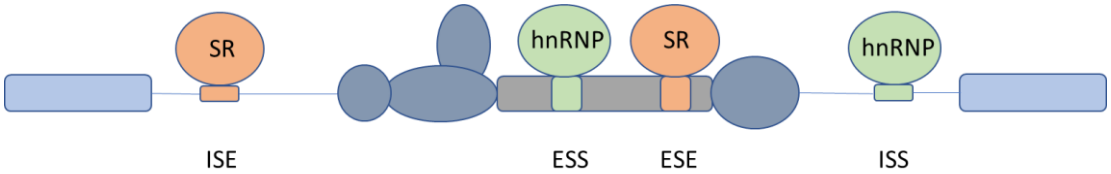


Figure 3: Alternative splicing regulatory factors and role of two splicing factor families (hnRNP and SR) in alternative splicing. The enhancer elements are recognized by activator proteins, and the silencer elements are bound by repressor proteins. These two protein families promote or inhibit spliceosome assembly at weak splice sites. Adapted from Wang et al[15].

SR proteins and hnRNPs are two protein families of splicing factors that regulate alternative splicing of many target genes[18]. Recent studies have uncovered that the regulation properties

of SR or hnRNP proteins that bind to pre-mRNA are often specific to cellular conditions[18]. However, SR and hnRNP proteins are not the only splicing regulatory proteins, as there are other classes of proteins that bind to RNA and regulate splicing. Some examples of these are the signal transduction and activation of RNA (STAR) proteins[21], NOVA proteins[22], epithelial splicing regulatory proteins (ESRPs), T cell-restricted intracellular antigen 1 (TIA1)[23], and many others. Splicing factor proteins aroused great attention and are now considered as main regulators of cellular processes. As splicing can be a major contributor to human disease, obtaining a better understanding of the AS regulatory proteins is gaining more interest[24].

1.2. RNA-binding proteins, roles of RBPs in RNA life cycle

The pre-mRNA processing reactions are regulated by RBPs and trans-acting RNAs[1]. In yeast, 570 different proteins have the capacity to bind RNA, and this number is predicted to be even higher in humans, as even the RRM, which is one type of RNA binding domain (RBD), is found in approximately 500 different human genes. There are some other well-known RBDs, including the K-homology (KH) domain, RGG (Arg-Gly-Gly) box, zinc finger (ZnF), and double-stranded RNA-binding domain (dsRBD)[25,26].

To better understand the RBP function one of the important aspects is subcellular localization, which can be both nuclear or cytoplasmic[1]. Several RBPs are able to shuttle between the nucleus and the cytoplasm and play a crucial role in the RNA life cycle which involves processes such as RNA splicing, stability, translation, degradation, and localization (Figure 4).

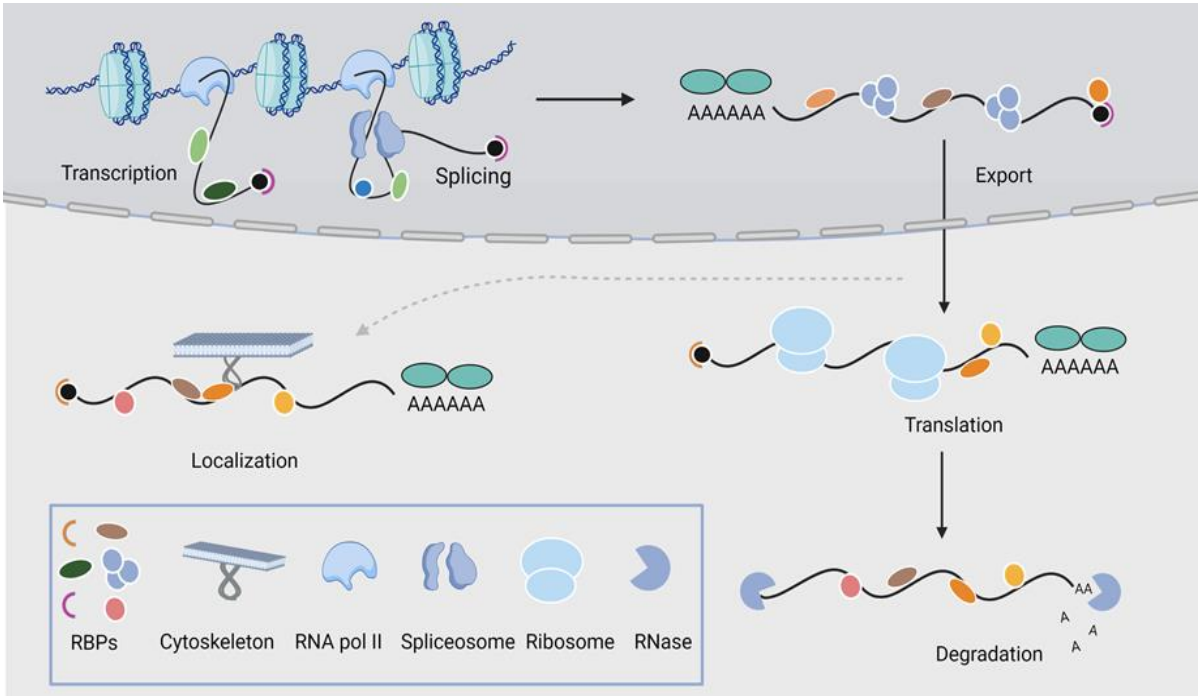


Figure 4: RBPs are involved in controlling RNA life. They are involved in splicing, polyadenylation, and transcription, besides transport, translation, localization, and degradation of mRNA. Adapted from Gebauer et al[27].

One example of an RBP involved in splicing, stabilization, and translation of many mRNA species is the human antigen R (HuR) protein[28]. The function of HuR starts in the nucleus, where HuR associates with introns and influences splicing. Then, HuR's role keeps going in the cytoplasm to stabilize the mRNAs, and it modulates the mRNA translation machinery.

1.2.1. How RNA-binding proteins interact with RNA

Most of the PRIs are distinctive, and identification of how they interact with each other is necessary. To understand a PRI, a closer look at the chemical fragments of proteins and RNAs is needed[29]. Binding between proteins and RNAs may take place throughout different interactions such as hydrogen bonding, Van der Waals interactions, hydrophobic interactions, electrostatic interactions, and π-interactions.

Hydrogen bonding is a type of bonding that can ensure major stability to the protein-RNA complex. Hydrogen bonds form at 2.4–3.0 Å distances between an electronegative atom bound to a hydrogen atom, whose electrons are not shared equally[30]. Van der Waals (VdW) interactions which are weak electrostatic interactions occur above ~3.0 Å can also be considered as weak hydrogen bonds. The 2'-OH and the phosphodiester backbone of RNAs can interact with protein through hydrogen bonds and VdWs[31].

Hydrophobic interactions, which form at distances of 3.8–5.0 Å, can be important stabilizing factors for the PRIs as they may result in the formation of a "hydrophobic core". Hydrophobic interactions that have been studied at protein-RNA interfaces represent around 50% of all interactions[32].

π-interactions are rather strong and can occur at the nitrogenous base ring and an aromatic ring containing amino acid[33]. These interactions are forming most frequently with a distance of 2.7–4.3 Å and often result in π-π stacking interactions. As a result of the structural analyses, multiple π-interactions occurring in protein-RNA crystal structures have been found, and these kinds of interactions can participate in the stabilization of protein-RNA interaction.

1.3. Roles of RBPs in diseases

1.3.1. RBPs in diabetes and cardiovascular disease

Several RBPs have been linked to human pathologies, from vascular disorders to cancer and neurogenerative diseases[27]. Diabetes mellitus (DM), a widespread global health condition[34] is hallmarked by chronic hyperglycemia, which results in systemic damage to the vascular system that triggers cardiovascular disease (CVD). In vascular endothelial cells (ECs), hyperglycemia was shown to cause marked alterations in gene expression. This is regulated by transcriptional and post-transcriptional control mechanisms. It is vital for the healthy performance of vascular ECs to regulate posttranscriptional modifications in the RNA network. A wide

variety of RNA networks regulated by RBPs are participating in vascular endothelial disorders in diabetes[35]. For instance, RNA Binding Fox-1 Homolog 2 (RBFOX2), upregulated in diabetic hearts, regulates alternative splicing[36]. It binds to its target RNA motif and controls the splicing of genes implicated in diabetic cardiomyopathy. In addition, HuR is a highly expressed RBP that is upregulated in the presence of high-grade diabetes.

Apart from that, in the adult heart, changes in splicing patterns have also been shown to disturb cardiac function. The expression levels of several splicing factors are found to be repressed in dysfunctional cardiomyocytes. Studies have identified that the expression levels of RNA-binding motif protein 20 (RBM20) which is the most abundantly expressed protein in the human heart, can lead to dilated cardiomyopathy in humans[37,38].

1.3.2. Roles of RBPs in cancer

RBPs are strongly correlated with tumor formation[39], and many investigations have indicated that RBP dysregulation has a major contribution to cancer progression[40].

Recent studies have defined that splicing factors play an important role in cancer progression[18]. For example, splicing factor proteins can bind to cis-regulatory elements to mediate alternative splicing reactions[39]. Compared to normal cells, in cancer cells aberrant alternative splicing has been observed which can drive cancer development[18].

Localization of an mRNA is a significant process in the regulation of mRNA stability and translation. RBPs that are linked to cancer mainly interact with RNAs and regulate their localization and translation[41]. For instance, IGF2BP1 (insulin-like growth factor 2 mRNA binding protein 1) regulates mRNA localization, which is linked to cell motility and influences cancer invasion and metastasis[42,43].

RBPs are implicated in different stages of the translation process, including termination, elongation, and initiation. Eukaryotic translation initiation factor E4 (eIF4E) is an oncogene and a tumor-related RBP and is over-expressed in cancer cells[44]. HuR is another RBP that binds mRNAs having AU-rich elements (ARE) in their 3′-UTR and as a result, increases the ProTαmRNA translation. An increased level of ProTαmRNA has been reported to have a relation with anti-apoptotic effects in cancer cells[45].

RBPs can be a promising target in cancer therapy as they act as a critical regulator of cancers. Therefore, expanding the understanding of the roles of RBPs in cancers will help to provide better insight into therapeutic perspectives.

1.3.3. RBPs and their role in neurodegenerative disease

Recent investigations suggest that RBPs are playing a key role in the maintenance and integrity of neurons. Changes in RBP function and RNA metabolism can trigger various

neurodegenerative diseases. For instance, studies have revealed that in amyotrophic lateral sclerosis (ALS), there is a clear genetic correlation between mutations of RBP-encoding genes and the development of the disease[46]. In addition, Fragile X Messenger Ribonucleoprotein 1 (FRM1) mutation leads to an intellectual disorder where the common mutation occurs in the 5' untranslated region of a gene on Chromosome X that encodes the Fragile X mental retardation protein (FMRP)[47,48]. FMRP has multiple RBDs[49] binding to the range of mRNAs and suppresses translation[50]. In fragile X syndrome, over-translation of RNA to protein presumably results in a neurological disorder.

1.4. Peptides as RNA-protein interaction modulators

One way to target PRIs is by using small molecules to modulate RNA-binding proteins. As an example, Park *et al.* investigated small molecule inhibitors for Lin28–let-7 interaction[51]. A newly designed compound increased the let-7 miRNA level and decreased the let-7 gene expression. However, it might not always be straightforward to find small molecule inhibitors due to the high surface area of the target protein and the lack of well-defined binding pockets[52]. Due to the limitations of their molecular size, small molecules occupy only 300-1000 $Å^2$ of the protein surface[53].

Apart from small molecules, peptide-based inhibitors are promising candidates to target protein-protein interactions (PPIs) and PRIs effectively[54–56]. Peptides offer higher selectivity and efficacy, along with low levels of toxic properties. However, therapeutic peptides also have two intrinsic drawbacks. Those are poor permeability and stability *in vivo*. Several factors may affect membrane permeability, including the length of the peptide and the amino acid combination. The failure of passing through the membrane limits the application of the peptides in drug development. A second drawback is the poor stability of the peptide candidates *in vivo*. Amide bonds of the peptides can be hydrolyzed by proteases and this leads to the short half-life and fast elimination[57]. However, these drawbacks have led to more interest in optimizing them, which has resulted in a series of peptide drugs.

1.4.1. Improving the physicochemical properties of peptides

Different strategies such as modifications on the backbone, on the side chain, peptide termini modification, peptide stapling, conjugation with cell-penetrating peptides, and lipid conjugation are some of the examples that helped to improve the peptide properties (Figure 5)[58].

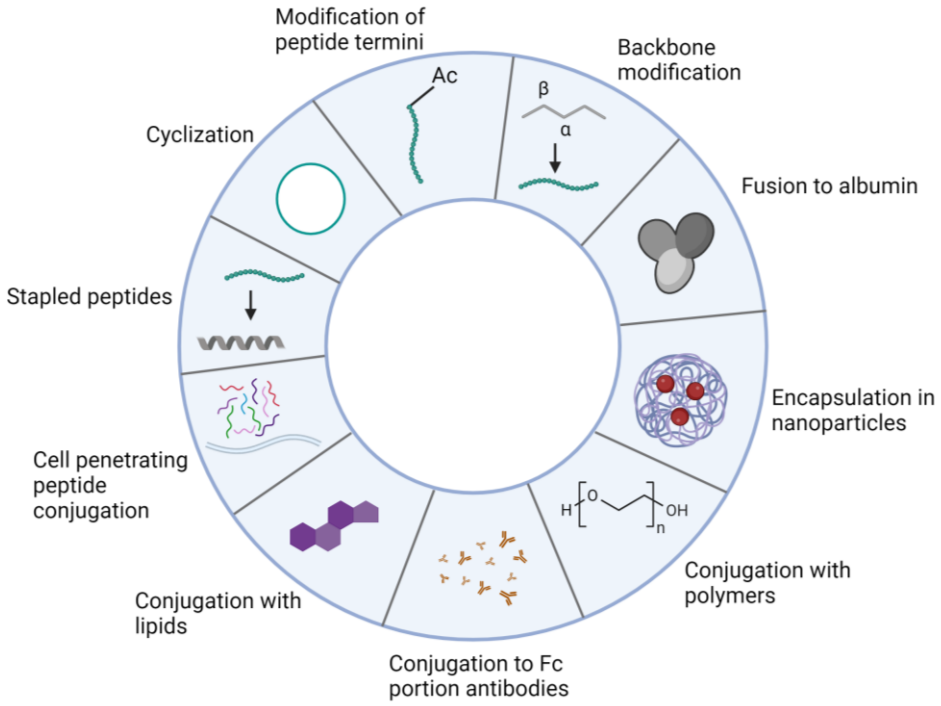


Figure 5: Developed methods to enhance the physicochemical characteristics of peptides. Adapted from Lee et al[58].

Backbone and side chain modifications

With backbone modification, the proteolytic stability of peptides can be improved while modifications on side the chain are applied to improve the binding capacity and selectivity of the peptides. There are several ways of introducing backbone modifications: L- to D-amino acid substitution, inserting N-alkylated amino acids, α-substituted or β-substituted amino acid insertion. In side chain modifications, natural amino acids are exchanged with their analogs[59]. For example, homoarginine and β-phenylalanine are both suitable for modifying side chain of peptides. To increase the peptide stability, mainly ones which are sensitive to enzymes, peptide termini modifications are also a good choice.

Cell penetrating peptides and lipid conjugations

Cell-penetrating peptides (CPPs) can travel through biological membranes. Basic residues such as arginine and lysine are abundant in CPPs, and they enhance the recognition by the glycosaminoglycans (GAGs), which enable CPPs to travel through the membrane. The conjugation of CPPs with other peptides can also enhance the membrane permeability; for example, the conjugation of human insulin to the TAT peptide improved the permeability of the human insulin through the membrane[60,61].

Another used technique, conjugation with lipids results in reduced toxicity and improved oral bioavailability. For example, fatty acids that contain carboxylic acid and hydrocarbon chains can be conjugated with drugs. Stearic acid, squalenoic acid, and palmitic acid are examples of fatty acids used for conjugation[62].

Cyclization

Cyclization is a widely used technique to improve the properties of peptides[63,64]. Cyclization can increase affinity and proteolytic stability. Via cyclization, stabilization of α-helices and β-sheets can be achieved[65,66]. Cyclization improves binding affinity as cyclic molecules have a lower loss of entropy compared to their linear analogues[67]. Furthermore, macrocycles are can keep some of the properties of potent small molecules, such as metabolic stability and a lack of immunogenicity. All these properties make the macrocyclic peptides advantageous tools in the developing therapeutics. The peptide cyclization can be a challenging synthesis step since a specific structure has to be obtained. The numerous strategies that have been reported for successful peptide cyclization include amide bond formation, disulfide bond formation, alkene metathesis, ether/thioether formation, CH activation, and photocatalyzed reactions[68].

1.5. Peptide cyclization strategies

1.5.1. Amide bond formation

Cyclization via amide bond formation is classified into four general groups which are head-to-tail, side chain-to-side chain, side chain-to-tail, and head-to-side chain cyclization (Figure 6)[69]. For cyclization based on the formation of the amide bond, the coupling protocol used for linear peptides can be applied[70]. The three main types of reagents (phosphonium, carbodiimides, and aminium-/uranium-iminium) have been optimized to reduce racemization during amide bond formation[71]. However, head-to-tail cyclization is still challenging due to C-terminal epimerization and dimerization, which is common for peptides shorter than seven residues. If the coupling reagent and additives are carefully chosen these side products can be reduced[72]. For instance, PyBOP was employed for the cyclomarin C synthesis[73], while for the teixobactin synthesis, HATU/HOAt/Oxyma Pure/DIEA was preferred[74].

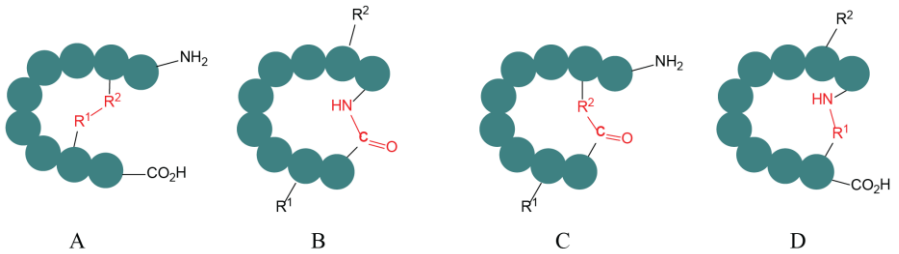


Figure 6: Illustration of the amide cyclization. A) Side chain-to-side chain; B) Head-to-tail; C) Tail-to-side chain; D) Head-to-side chain cyclization. Adapted from Hayes et al[69].

1.5.2. Disulfide cyclization

Disulfide bond formation is another widely used strategy to design cyclic peptides. Cysteines, which can be easily included into the peptide chain are used for disulfide cyclization. Disulfides are formed under oxidative conditions, but naturally, disulfide bonds are unstable in reducing

environments such as the cytosol. Stability can be increased by changing disulfides with other analogs such as thioether, triazole, lactam, etc[68].

1.5.3. C-C double bond formation: Alkene metathesis

Stapled peptides are peptides that can mimic the structures at the interface of PPIs with an α-helical shape[75]. Due to the high hydrophobicity from the helical structure, the peptide can pass through biological membranes[76]. With peptide stapling, an extended plasma half-life and proteolytic resistance can be achieved.

Alkene metathesis has been commonly applied to synthesize stapled peptides as it is compatible with various functional groups and mostly results in good yields. The reaction happens in the presence of two alkenes and is also referred to as ring-closing metathesis (RCM) (Figure 7). RCM was initially applied by Grubbs and Blackwell for stapling peptides and reported for solution-phase metathesis. However, RCM can be carried out both in solid and solution phases. Typically, a Grubbs-type catalyst is used, which are classified as Hoveyda-Grubbs catalyst, 1st, 2nd [77], and 3rd generation Grubbs catalyst[78]. Different catalyst concentrations are required depending on the reaction conditions, e.g. in the solution phase or in the solid phase. In the solution phase, lower concentrations are applied to avoid dimer formation. In the solid phase, microwave irradiation is usually applied in order to obtain the cyclized product. For successful stapling, alkene modified specific side-chain amino acids are used. Hydrocarbon stapling is now optimized for i, i + 3/4/7 spacings and has been applied to peptides to stabilize α-helical structures, which are able to mimic binding motifs from proteins.

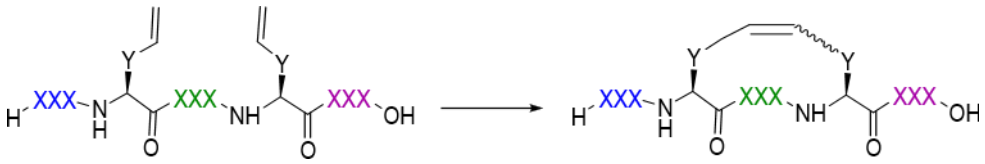


Figure 7: Ring closing metathesis scheme. X = amino acid, Y = CH_2, $(CH_2)_3$, CH_2SCH_2. Adapted from Bechtler et al[68].

Song *et al.* reported a peptide as an example of successful RCM, where a six-fold enhanced peptide affinity for the initiation factor eIF4E was observed[79]. In addition, an insulin analog defined by van Lierop *et al*. in which substitution of the disulfide with alkene retained an affinity for insulin receptors and exhibited increased efficacy in mice[80].

1.6. Cyclic peptide libraries

Macrocyclic peptides as regulators of protein-protein and protein-RNA interactions is a progressing field, and since they are advantageous compounds as potential drugs suitable methodologies for cyclic peptide library generation are required[68]. The available technologies can be classified based on the methods used to prepare the cyclic peptides, which are either

chemically or genetically prepared[81]. As an example of methods that are based on chemical synthesis include DNA-encoded libraries while split intein circular ligation of peptides and proteins (SICLOPPS), mRNA display, and phage display are genetically encoded methods.

1.6.1. Genetically encoded libraries

Phage display libraries

Phage display allows the screening of 10^9 peptide variants[82]. In phage display, the phenotype is linked to genotype, and the peptide ligand which is displayed is encoded by the phage genome (as a fusion to a coat protein) (Figure 8)[83]. For the first time, Smith *et al*. introduced this technology in which he utilized synthetic peptide chains on the bacteriophage surface protein[84]. Briefly, DNA sequences, derived from fragments of specific genes or as synthetic DNA libraries, are inserted into the phage genome that then codes for the peptide fused to the N- or C-terminus of one of its coat proteins. It is necessary to mention that phage display has some limitations since it can only use proteinogenic amino acids as building blocks. In this case, the obtained peptides are sensitive to proteolytic degradation. Conformational flexibility, to a high degree, also affects the binding affinity or macrocyclization specificity. Moreover, in phage display the diversity of the population can be limited due to consecutive library amplifications[85]. The first generated macrocycle libraries were peptides cyclized via a disulfide bond between two cysteine residues. As a result of the screening of such cyclic peptide libraries, a biological probe to illuminate the binding interface between Protein A (a component of the *Staphylococcus aureus* cell wall) and the antibody Fc fragment has been discovered[86].

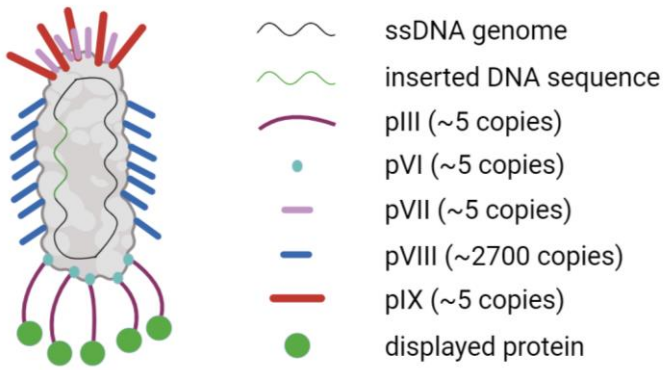


Figure 8: Principle of the phage display technology using M13 bacteriophage. M13 bacteriophage has a single-stranded DNA genome, and five different coat proteins. The DNA sequence that encodes for the peptide is inserted into the genome sequence, as a fusion to the gene which encodes for the phage coat protein (pIII). In the bacterial host gene expression starts and the peptide fragment is displayed on the phage surface as a combined product of genes which are involved in coat protein encoding. Adopted from Jaroszewicz et al[87].

mRNA display libraries

As in phage display, in mRNA display, the peptides in the library are covalently linked to their encoding mRNA strand, which allows identification of the peptide hits via sequencing[88,89]. The

RNA-peptide conjugates are exposed to affinity-based screening, reverse transcription, amplification, and transcription, and several cycles are required to find a tight binder. This technique avoids E. coli transformation by the application of in vitro transcription and translation. As a result, mRNA display can generate much larger libraries compared to phage display[90]. With mRNA display, generation of artificial amino acid-containing peptides libraries is possible which is one of the major advantages when compared with other methods.

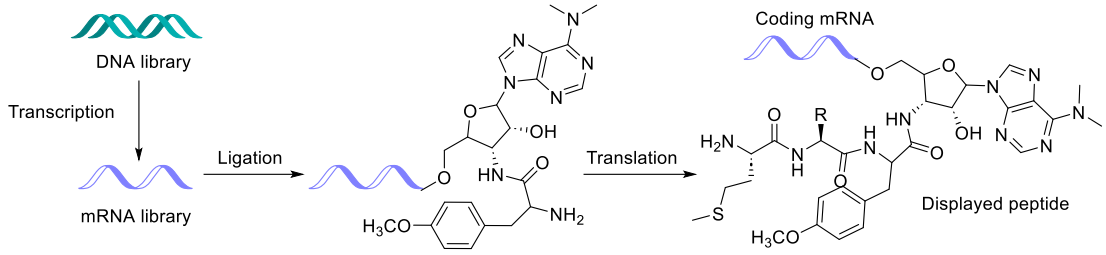


Figure 9: Synthesis scheme of an mRNA display peptide library. Adapted from Dougherty et al[91].

Intein-based libraries

Benkovic and co-workers developed a technology called SICLOPPS[92]. This method allows the intracellular cyclization of peptides and a library of up to 10^8 peptides can be produced[91]**.** In SICLOPPS intein self-excision followed by C- and N-terminal segments ligation results in a cyclic peptide[93]. Reconstitution of the C- and N-terminal parts of an intein forms an active intein, and as a result a thioester is formed. That undergoes a transesterification process at the C-terminal intein junction to form a lariat intermediate. Then, the further rearrangement releases the lactone as a cyclic product, and an X-to-N acyl shift results in the formation of a cyclic peptide (Figure 10). The only requirement for SICLOPPS to function efficiently is that a cysteine or serine must be the first residue[94]. For the selection of active hits, SICLOPPS can be combined with a cell assay such as a bacterial reverse two-hybrid system (RTHS)[95]. Tavassoli and colleagues reported a peptide binding HIF-1α PAS-B domain and as a result, inhibits the hypoxia-inducible factor (HIF)[96].

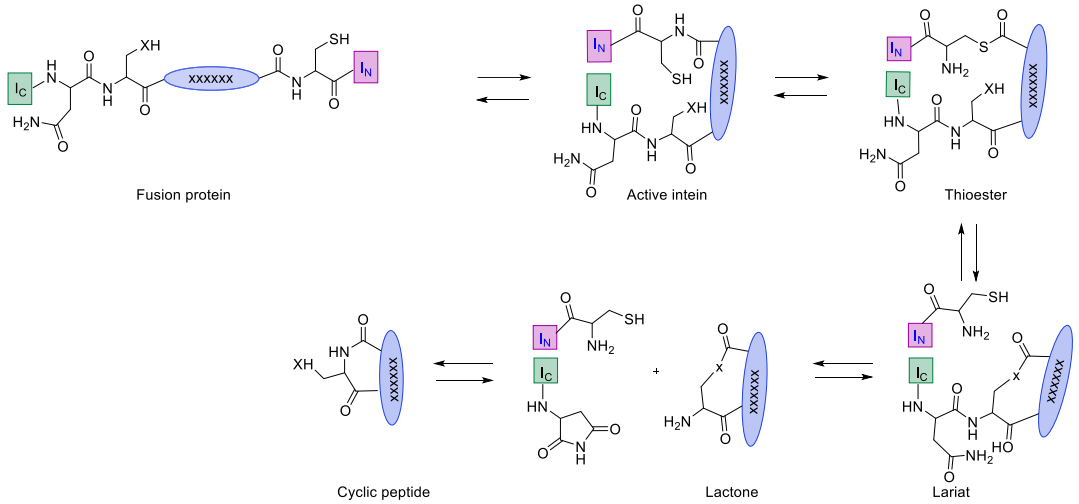


Figure 10: Representation of the cyclic peptide generation mechanism by SICLOPPS. X = O or S. Adopted from Tavassoli et al[97].

1.7. Peptide nucleic acids

PNAs have potential as therapeutics, molecular tools, anticarcinogenic agents, biosensors, and in nano-technology. PNA technology was first invented by Peter and his colleagues in 1991 as a DNA analog[98]. The aim of this investigation was to design a ligand that can recognize double-stranded DNA through Hoogsteen base pairing.

PNAs consist of N-(2-aminoethyl) glycine building blocks, which form the peptide backbone and replace the phosphoribose backbone of DNA. The purine and pyrimidine nucleobases are bound to this backbone which is achieved by methylene carbonyl linkages[99].

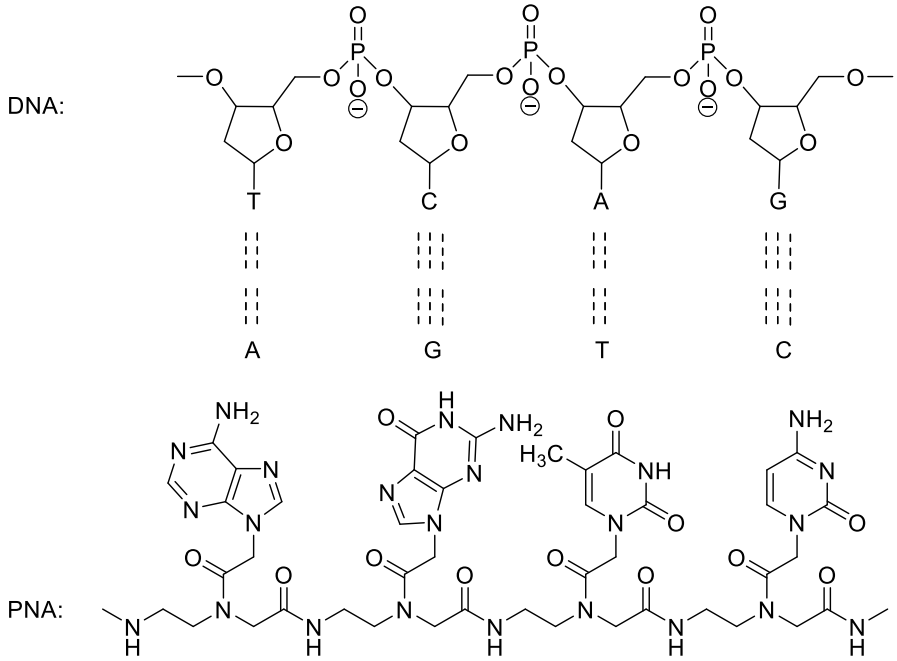


Figure 11: Structure of DNA and PNA. A: adenine, G: guanine, C: cytosine and T: thymine. Adapted from Wu et al[100].

PNA synthesis is similar to peptide synthesis[169], which can be synthesized by applying solid-phase synthesis protocols. The backbone of PNA gives it a unique hybridization characteristic, and unlike DNA and RNA, the backbone of the PNA is not charged. Hereby, there isn't electrostatic repulsion while PNA hybridization to its target nucleic acid which provides high stability to the PNA-RNA or PNA-DNA duplexes. PNAs are resistant to enzymatic degradation, which gives them a longer lifetime both in vitro and in vivo. As the peptide backbone of PNAs is not easily recognized by nucleases or proteases, it can be biostable in human serum, bacterial cell extracts, and tumor cell extracts.

Their specific optimized properties, such as higher affinity, strong specificity, and low toxicity, make PNAs strong candidates in therapeutic drug development. It has been reported that gene expression in P. falciparum can be controlled by using antisense PNA molecules. For the first time, Kolevzo and coworkers showed that using PNA technology to treat malaria infections is

possible[101]. They conjugated antisense PNA to an octa-D-lysine CPP and, as a result, achieved downregulation of P. falciparum gene expression. Targeting pre-mRNA splicing is also a promising strategy for cancer drug discovery. Some PNA conjugates have been identified which can target the 5'- or 3'- splice sites and inhibit pre-mRNA splicing[102]. Nielsen *et al.* reported PNAs that regulate the pre-mRNA splicing of the Mouse double minute 2 homolog (MDM2) gene[102]. One of the PNAs showed splicing inhibition of intron 3 while another showed splicing inhibition of intron 2 and intron 3. These results demonstrate PNAs can be a good modulator of splicing events.

1.8. Allosteric regulation, regulation with small molecules and peptides

Allosteric regulation is one of the ways the interactions between biomolecules can be regulated[103]. Allosteric regulation can be classified as negative (NAMs) and positive allosteric modulation (PAMs) (Figure 12). For NAMs, the allosteric inhibitor binds to the allosteric binding site, which leads to structural change with the protein that results in the inhibition of the binding of the ligand at the active site. For PAMs, the allosteric activator binds to the allosteric binding site and, through structural reorganization of the protein improves the binding affinity of the ligand to the active site.

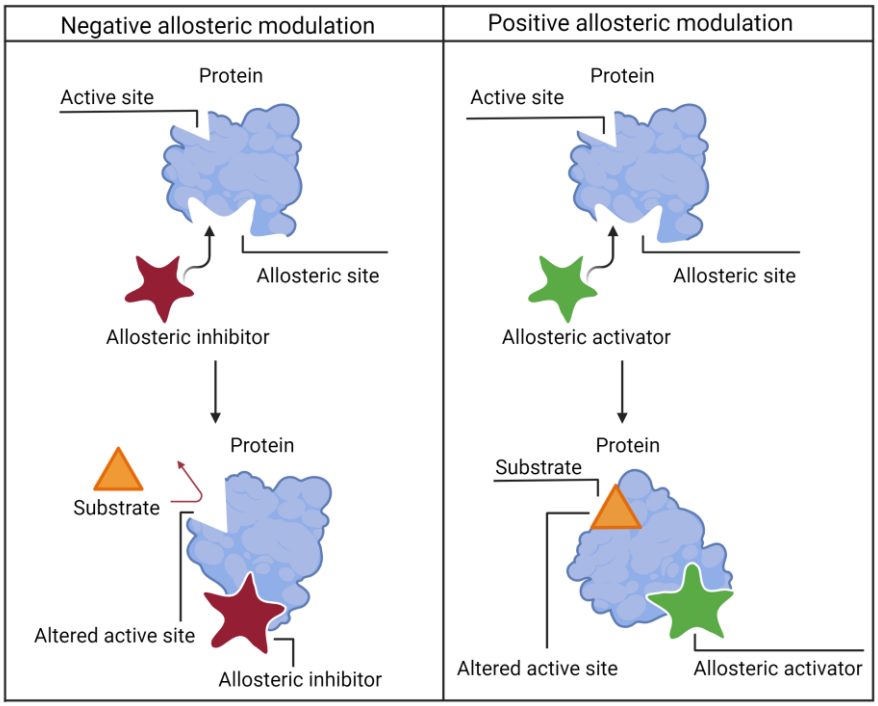
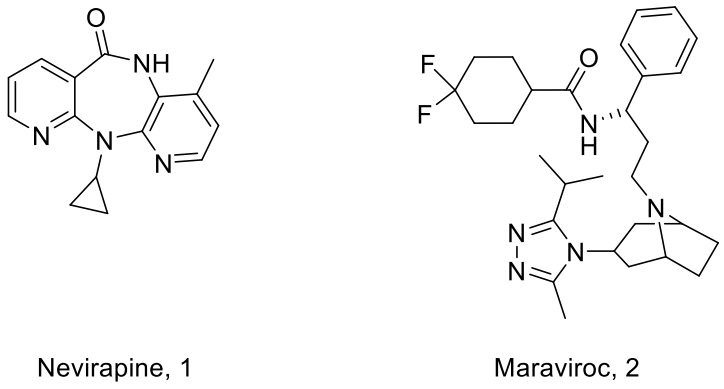


Figure 12: Allosteric regulation. Negative allosteric modulation (left) and positive allosteric modulation (right) is explained. In NAM the active site of the protein (blue) is altered by an allosteric inhibitor (red). Therefore, binding of the substrate (orange) is prohibited. In PAM (right panel), the affinity for the substrate (orange) enhances upon the allosteric activator (green) binding to the active site of the protein. Adapted from Ni et al[103].

Many proteins undergo allosteric regulation, and one of the examples for this is hemoglobin, where a negative allosteric modulation process take place[104]. Hemoglobin is an oxygen

transport protein that is found in blood cells. Oxygen binds to the hemoglobin on one subunit, resulting in a conformational change that induces the binding of oxygen to the rest of the active hemoglobin subunit.

There are some drugs that have been already approved as allosteric regulators by the US Food and Drug Administration (FDA). One of these drugs is a Nevirapine (1) (Figure 13, 1) which hinders virus infections caused by immunodeficiency. This immune deficiency (HIV/AIDS) is acquired by reverse transcriptase enzyme inhibition[105]. Unlike other inhibitors, nevirapine has an allosteric binding away from the active site. Another well-known allosteric drug is Maraviroc (2) (Figure 13, 2), which works based on the negative modulation mechanism and is used in the HIV infection treatment. Chemokine receptor type 5 (CCR5) is required for the viral entry into the cell and maraviroc can cause a conformational change of CCR5 allosterically that prevents viral entry[106].



Nevirapine, 1 Maraviroc, 2

Figure 13: Allosteric drugs approved by FDA. Adapted from Ni et al[103].

Apart from small molecules, peptides can also act as allosteric regulators. Peptide sequences that mimic endogenous allosteric modulators and can influence the activity of the target can be selected through structure-based drug design or computational studies. Several allosterically acting peptides have been discovered[107]. For example, a macrocyclic peptide was reported that inhibits hepatocyte growth factor (HGF)-induced MET receptor activation allosterically which happens through binding to the active form of HGF selectively[108]. As another example, a cyclic peptide that allosterically modulates lysophosphatidylinositol (LPI)-induced endocytosis and inhibits G protein-coupled receptor 55 (GPR55)-mediated proliferation[81]. In a recently published paper, a stapled peptide that is targeting Polypyrimidine Tract Binding Protein 1 (PTBP1) was reported to bind to the RBD domain of the PTBP1 and, therefore, allosterically disrupt the protein-RNA interaction[109]. All these discoveries indicate that allosteric regulation is a convincing approach to target protein regulation selectively to control cellular processes.

2. Modulating the interaction between RBM20 and RNA

2.1. Introduction

RNA binding motif protein 20 (RBM20) is a protein that binds RNA and is found to be involved in DCM. DCM is a heart muscle disease that is related to the enlargement of the heart ventricles (Figure 14) and is one of the major causes of heart transplantation[110]. Survival chances are low since DCM progresses over time and as a result may lead to death.

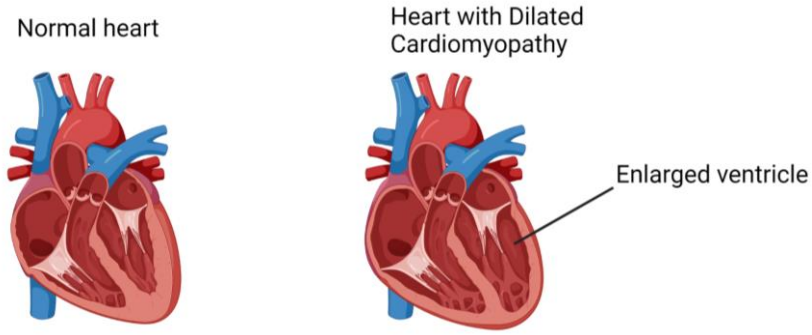


Figure 14: Image showing how DCM effects the normal heart. In normal heart (left) chambers relax and fill, then contract and pump. In the heart with DCM (right) muscle fibers have stretch, heart chambers enlarge. Adapted from "Medical gallery of Blausen Medical 2014".

The first RBM20-related human cardiomyopathy was reported in 2009 and by then the main cause of cardiomyopathy was considered to be mutations in the RBM20 protein[38]. Mutations that lead to DCM have been found in a number of genes and RBM20 is one of them that shows a mutation in 2-3% of DCM cases. RBM20 was found to be selectively expressed in heart tissue and acts as a splicing factor protein affecting the genes that are necessary for cardiac function. 25% of DCM mutations are related to the TTN gene which is a human gene with 364 exons, the highest exon number in mammals[111,112]. The TTN gene encodes the titin protein which is the largest sarcomeric protein and is specific for muscles[113]. It has been shown that titin elasticity can be pathologically modified in heart failure and cardiomyopathy. Unstable titin elasticity is linked to titin isoform switching which is a result of alternative splicing[114,115]. It was known that the TTN gene can undergo alternative splicing, however, regulation of TTN splicing was not investigated until it was reported that RBM20 is involved[116]. It was stated that the rescue of HFpEF-like symptoms happens when the titin splicing factor RBM20 is inhibited in mice[111]. HFpEF is a type of heart failure that results in complications such as exercise intolerance, diastolic dysfunction, and concentric hypertrophic remodeling[117,118]. Inhibition of RBM20 resulted in restored diastolic function and improved exercise tolerance when using Ttn^{IAjxn} (I-band - A-band junction) mouse model.

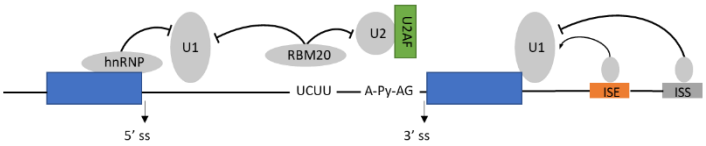


Figure 15: The regulatory mechanism of the RBM20 binding to cis-regulatory elements in TTN. Adapted from Guo et al[119].

The RBM20 complying gene carries 14 exons and is located on chromosome 10[120,121]. The protein consists of 1227 amino acid and the annotated domains include a leucine-rich region, two zinc finger domains, an RRM domain, a serine- and arginine-rich region (SR region), and a glutamate-rich region (Figure 16).

Figure 16: RBM20 protein structure, including functional domains. Adopted from Watanabe et al[120].

The RNA recognition motif of RBM20 recognizes UCUU RNA sequences specifically. When the UCUU is mutated to CGUU or CGCG the binding is completely lost, while with the UCUG modification a reduction of the binding affinity was observed[122,123]. Based on structural information the RRM domain is extended upon interaction with RNA. Upon closer inspection of the RNA-bound structure it is visible that the C-terminus of the protein folds into a helix (helix 3) upon binding of RBM20 to RNA with a UCUU sequence, see PDB structures 6so9 and 6soe (Figure 17). On the other hand, when the helix is removed the binding affinity of the protein for RNA is reduced[122]. The structural studies confirm that even though helix 3 does not directly interact with RNA it plays an unusual role in uridine nucleotide binding to the protein complex.

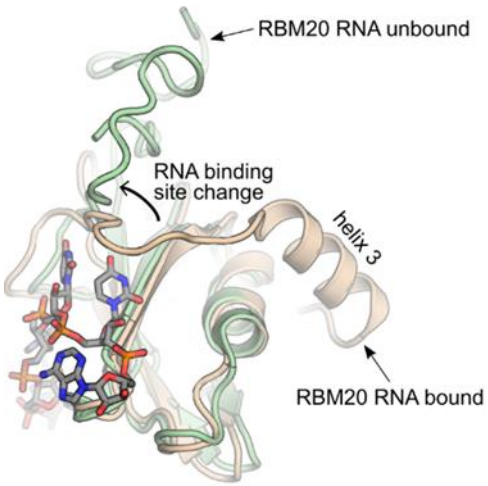


Figure 17: Overlay of the structure of RNA (AUCUUA) bound and unbound to RBM20. Shown in beige is the RNA bound RRM domain of RBM20 (PDB: 6so9), in green the unbound RBM20 (PDB: 6soe). Helix 3 is indicated and makes a dramatic conformational shift upon RNA binding. Structures are generated with the mouse protein. Adapted from Upadhyay et al[122].

ITC experiment confirms the interaction of RNA with RBM20

The interaction of RRM domain of murine RBM20 with RNA was confirmed with an isothermal titration calorimetry (ITC) experiment in which a low micromolar binding affinity to RBM20 was observed[122]. The experiment was carried out between RBM20 (513–621) binding to AUCUUA RNA, which showed a 1:1 interaction with a K_D of 5.7 µM. To study binding specific mutations were made and the new RNA derivates were tested by ITC. When either the first or last base was mutated, there were no significant change in the affinity to RBM20. Mutating the first uracil to cytosine led to a 2-fold increase in K_D, while mutation in the remaining CUU led to 16–36 times higher K_D values than the original motif (Table 1), which indicates the importance of the UCUU for interaction.

Table 1: K_D values obtained from ITC measurements of RBM20 with RNAs with various point mutations (underlined nucleotide). Adapted from Upadhyay et al[122]. For mRBM20 Δα3 the α3 helical region has been removed.

Protein	RNA	K_D (µM)
mRBM20 α3 (513–621)	AUCUUA	5.7 ± 0.2
mRBM20 α3 (513–621)	GUCUUA	4.0 ± 0.1
mRBM20 α3 (513–621)	ACCUUA	13 ± 2
mRBM20 α3 (513–621)	AUUUUA	140 ± 30
mRBM20 α3 (513–621)	AUCCUA	210 ± 3
mRBM20 α3 (513–621)	AUCUCA	95 ± 1
mRBM20 α3 (513–621)	AUCUUG	4.8 ± 0.1
mRBM20 Δα3 (513–609)	AUCUUA	43 ± 1

2.2. Aim

In this chapter, the regulation of the RBM20-RNA interaction was targeted. For this purpose, we planned to allosterically inhibit RBM20 with stapled peptides. The RBM20 helix 3 is unique and plays a significant role in the protein-RNA interaction. Therefore, allosteric inhibition at this site would provide a way of selective inhibition of RBM20 (Figure 18).

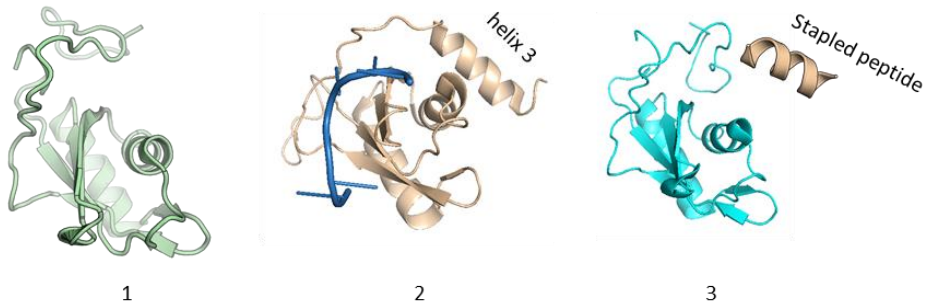


Figure 18: 1. Structure of RBM20 RRM domain in RNA unbound (green); 2. RNA bound states (tint); 3. together with stapled peptide (brown).

2.3. Results and discussions

2.3.1. Targeting RBM20-RNA interaction with stapled peptides

While peptide-based drugs are gaining more attention, there are still difficulties in using these compounds as therapeutics[76]. Over the past years, several techniques have been investigated to overcome problems with peptide-based pharmaceuticals. Using hydrocarbon stapled peptides is one way that could serve as a solution since stable α-helical peptides can be obtained via stapling, which is one of the most common secondary structures found in proteins[124]. A stapled peptide can have increased cellular uptake compared to its linear counterpart. Inserting a staple enhances the hydrophobicity, which might hypothetically play a role in cellular uptake. However, the actual mechanism of increased cell permeability is still being investigated[76]. Moreover, the α-helical structure of stapled peptides can increase protease resistance, as enzymes cannot reach target sites on the peptide.

Figure 19: Example of peptide stapling corresponding to one staple in the peptide chain. Adopted from Moiola et al[76].

The side chains of two amino acids can be linked together to force the peptide to form an α-helical structure. However, more than one staple can be installed within a long peptide using three or four non-natural amino acids depending on targeting to obtain two separate or connected staples (Figure 20, B)[125]. Hydrocarbon stapling uses non-native amino acids carrying side chains with a terminal alkene that can be coupled via RCM (Figure 20, A)[126]. One of the fundamental principles is that the amino acids selected for stapling are at the proper distance from each other. Usually, amino acids in position i and i + 4 or i + 7 are ideal for stapling. For i, i + 4 stapling, at least 7 or 8 carbon chains are used, while 10-12 carbon chains are needed for i, i + 7 stapling to form a brace between the amino acids[127,128].

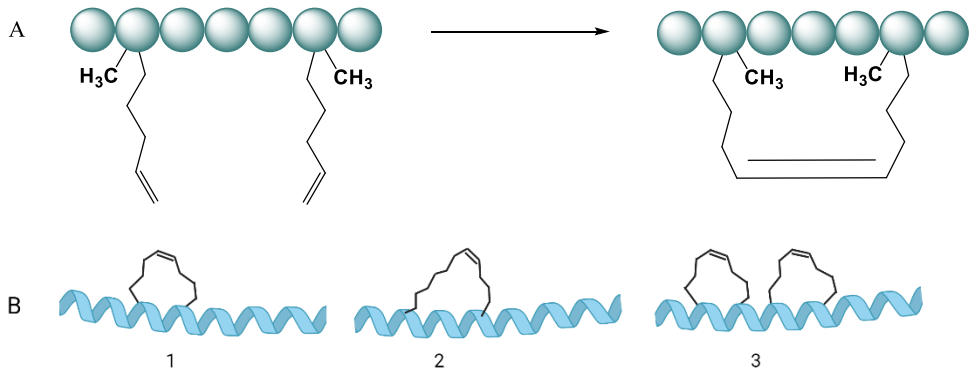


Figure 20: A) Example of RCM with the non-native amino acids bearing α-olefins; B) Examples of (1) one i, i + 4 stapling; (2) one i, i + 7 stapling; (3) double i, i + 4 stapling. Adopted from Walensky et al[128].

2.3.2. Design of peptides

After a close look at the structure of RBM20 protein residues, it was identified that an extra α-helix (helix 3) is formed upon RNA binding (Figure 21, A). We hypothesized that the helix could be mimicked by a stapled peptide (Figure 21, B), which could bind its site and allosterically inhibit RNA binding.

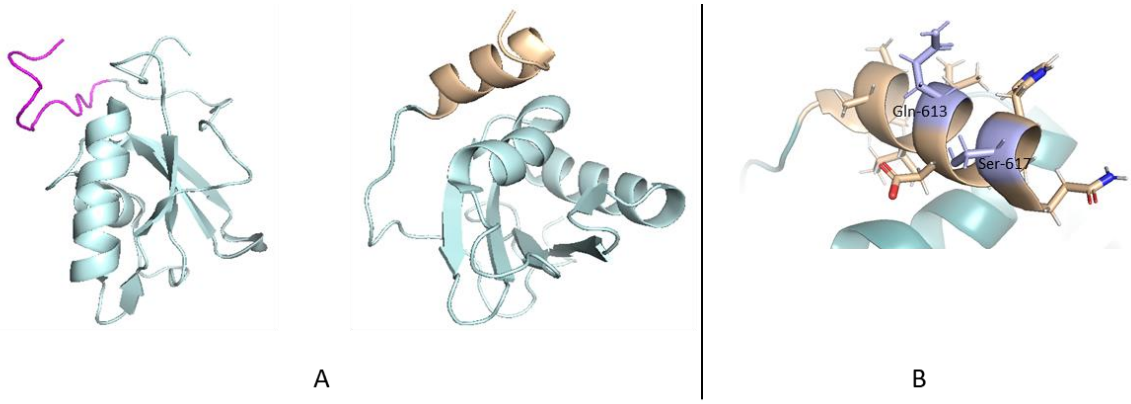


Figure 21: A) RBM20 loop in an unbound state shown in magenta and RBM20 helix 3, which is formed upon RNA binding shown in beige; B) Helix of RBM20 that were planned to be mimicked by stapled peptides. Stapling positions were shown in light blue.

Based on the structure of the extra helix 3 three peptides were designed for synthesis. An extra arginine (R) was added to improve the solubility properties of the peptides. The first was the original linear peptide sequence from RBM20 (Figure 21, magenta: Figure 22, RB1 peptide), and two were stapled peptides that were designed for mimicking the α-helix (Figure 21, beige; Figure 22, RB2–RB3 peptides). These two peptides were designed with the non-natural amino acids in a position that was not involved in binding according to the structure.

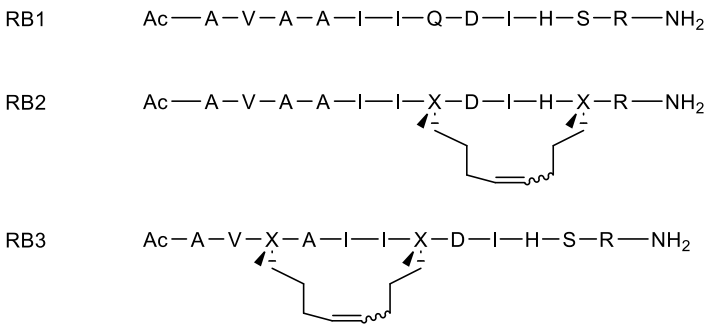


Figure 22: Designed mimics of RBM20 helix 3. X represents the non-natural amino acid that was placed to perform hydrocarbon stapling.

2.3.3. Synthesis of the linear and stapled peptides

Synthesis of the linear peptide is based on SPPS which was first reported by Merrifield[129]. SPPS starts with the attachment of the first amino acid to solid support by a covalent bond, followed by the assembly of the rest of the amino acids, forming a chain attached to the solid polymer until the desired peptide sequence is obtained. One of the important requirements of

SPPS is the selection of a suitable polymer that has to be insoluble in solvents used for the synthesis. As the peptide sequence will stay attached to the insoluble solid support during synthesis, this allows rapid filtration and washing of the excess reagents and solvents after each step of the synthesis[130].

Polystyrene-based Rink Amide AM resin was utilized to synthesize linear peptides. First, the Fmoc group on the resin was removed with 20% piperidine in DMF (Figure 23). Coupling was carried out in the presence of Fmoc-amino acid, PyBOP, and DIPEA in DMF. Next, the Fmoc deprotection and coupling steps were repeated until the desired peptide sequence was obtained. To remove the excess reactants the resin was washed with DMF after every step. If needed, acetylation of the N-terminus with acetic anhydride and DIPEA in DMF was carried out. To be able to use the peptide as a probe in biophysical experiments, FITC labeled versions were synthesized as well. For FITC labeling, a linker was attached to the peptide N-terminus by applying the above-mentioned standard coupling conditions and followed by Fmoc group deprotection. The peptide was then treated with FITC isomer I in the presence of DIPEA and DMF followed by cleavage from the resin with TFA/TIPS/H_2O. Liquid chromatography-mass spectrometry (LC-MS) analysis confirmed the success of the synthesis, and the purification was accomplished using reversed-phase high-performance liquid chromatography (RP-HPLC).

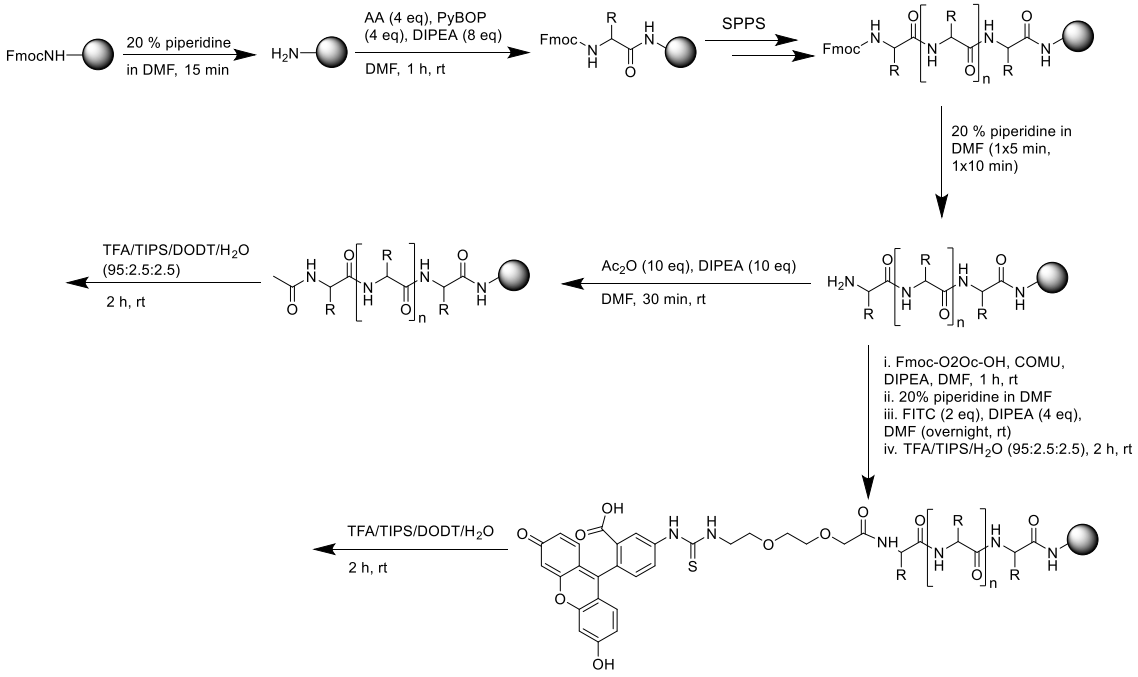


Figure 23: Synthesis scheme of the RBM20 linear peptides on solid phase. Rink amide resin was used as a solid support.

For the synthesis of stapled peptides (Figure 24), a Rink amide AM LL (low-loading) resin was chosen as a solid support. First, the linear peptide sequence was completed following the protocol mentioned above. However, in this case, non-native amino acids with alkene side chains (Fmoc-S_5-OH) for covalent connection were placed at two specific positions (see section 2.3.2.). Next, RCM[126] was applied to obtain stapled peptides from linear precursors. Initially, Hoveyda-Grubbs 2nd generation catalyst was used to efficiently link two alkyl groups located at the side chain of the amino acids[131,132]. To this end, the Hoveyda-Grubbs 2nd generation catalyst was added to the resin in argon-flushed DCE, and followed by microwave irradiation (MWI) for 10 min at 120 °C (see section 5.10.2.). Subsequently, the resin-bound stapled peptides were split into two portions, where one portion was acetylated while the other was labeled with FITC prior to cleavage and global deprotection[133]. Next, the peptides were purified using RP-HPLC.

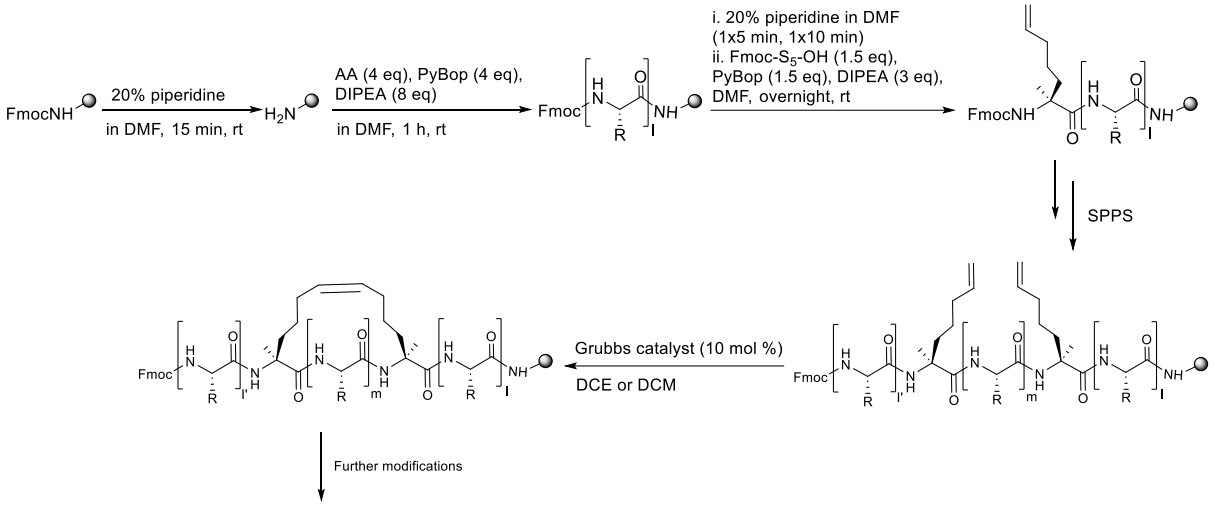


Figure 24: A synthesis scheme of the stapled peptides on the solid phase. Rink amide low loading resin was used as a solid support. Fmoc-(S)-2-(4-pentenyl) Ala-OH (Fmoc-S_5-OH) was chosen as a non-native amino acid for stapling.

To improve yield after cyclization, a slightly different approach was applied where stapling was carried out using Grubbs 1st generation catalyst[134,135], and the reaction was repeated overall four times for 2 h at room temperature (see section 5.10.2.). Once the cyclization was complete, the success of the stapling was monitored by ultra-high-performance liquid chromatography (UHPLC).

2.3.4. Circular dichroism spectroscopy of the stapled peptides

Circular dichroism (CD) spectroscopy is a method that is widely used in structural biology to determine the secondary structure of peptides and proteins[136]. Biomolecules are made up of chiral monomers that can produce signals upon the absorbance via amide and carbonyl groups of the polypeptide backbones when exposed to circularly polarized light[137]. CD spectroscopy can use the information on the absorption of polarized light by such chromophores, and this

can be employed to obtain structural information about the peptide and protein conformations. Secondary structure determination of proteins and peptides by CD can be performed in the UV range from 190 to 300 nm. From the CD measurement, it can be defined whether peptide or protein has a random coil, α-helical or β-sheet conformations as depicted in Figure 25[138].

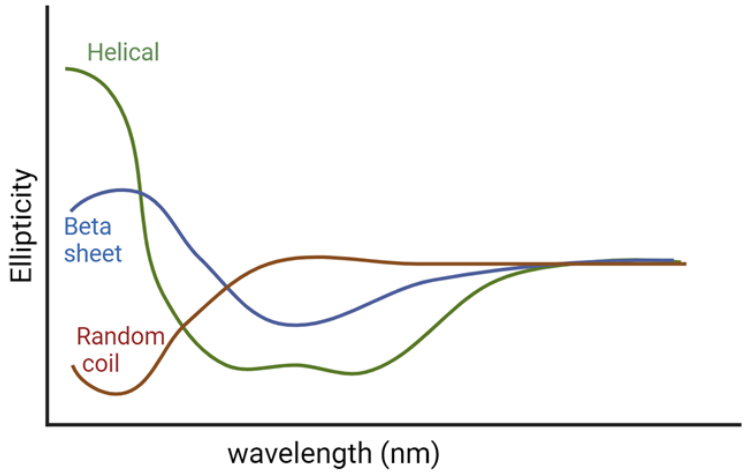


Figure 25: CD spectra of proteins with α-helix (red), β-sheet (blue) and random (pink) structures. Adapted from Sahoo et al [138].

To define the α-helicity of synthesized peptides, CD spectroscopy measurements were carried out. It is essential to follow a well-designed protocol to obtain high-quality and reproducible CD data. A sodium phosphate buffer, which is the buffer used for other experiments using RBM20, was used for this measurement. However, the buffer should not contain any components that can absorb in the used wavelength range. Therefore, buffer components were modified accordingly (see section 5.14.).

The CD molar ellipticity data at 222 nm can be used for quantifying the helical content of protein and peptides, where helical structure exhibits a characteristic minimum[139,140]. The fractional helicity (FH) of a protein in solution can be estimated from the CD response at 222 nm by equation (1) where θ_λ^{exp} is the experimentally observed mean residue ellipticity for a given wavelength (λ), and θ_λ^u and θ_λ^h correspond to the ellipticity for a peptide with 0% and 100% helical content at wavelength (λ) which are typically experimentally or theoretically estimated.

$$(1) \qquad FH = (\theta_\lambda^{exp} - \theta_\lambda^u) / (\theta_\lambda^h - \theta_\lambda^u)$$

Synthesized peptides were measured by CD spectroscopy and percentage helicity of the peptides were calculated based on equation (1).

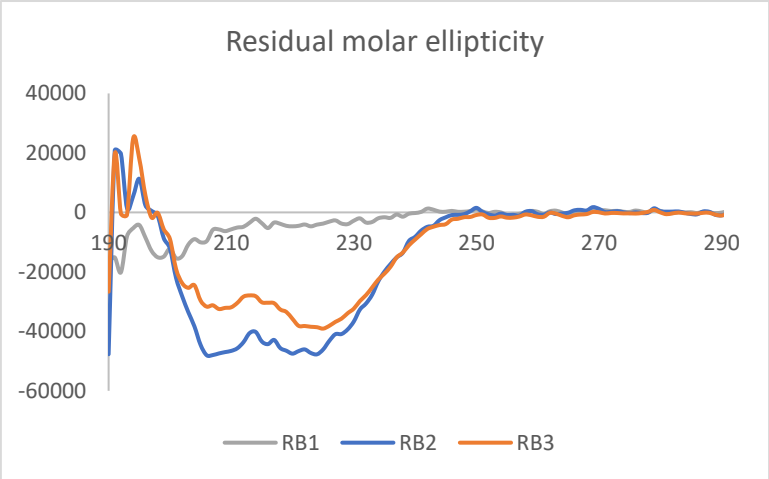


Figure 26: Representation of the CD spectra for acetylated RB1 (linear), RB2 (stapled) and RB3 (stapled) peptides. Buffer was used for a baseline correction.

Table 2: Percentage helicity of the peptides calculated based on the ellipticity values obtained from CD spectroscopy measurements.

Peptide	% Helicity
RB1	4.0
RB2	14.8
RB3	12.8

Linear peptide spectra seemed more similar to the random coil conformation, while stapled peptides had a higher α-helicity (Figure 26). The percentage of helicity calculated from the ellipticity values obtained from CD spectroscopy measurement confirmed this assumption. Calculations (Table 2) showed that the linear RB1 has the lowest percentage of helicity (4.0). In contrast, for RB3, a >3-fold increase in the percentage of helicity (12.8) was obtained, which was even higher for peptide RB2 (14.8).

2.3.5. Direct fluorescence polarization with RNAs

Two human RBM20 constructs, RBM20 α3 (residues 511–619) with helix 3 and RBM20 Δα3 (residues 511–601) without helix 3 (see section 5.4.) were chosen to be used in the fluorescence polarization (FP) experiment. To establish the binding affinity of RNA with each RBM20 construct a particular nucleotide sequence keeping the central UCUU motif has to be selected. The RNA nucleotides 3'FAM-GUCUUAA (R1) and 3'FAM GUCUUAGUCUUAGUCUUA (R2) were tested in the FP experiment. A three-times repeated UCUU variant was utilized to improve the binding affinity which would require lower protein concentrations in competitive experiments. Direct FP experiments were carried out with both protein constructs RBM20 α3 and RBM20 Δα3 (Figure 27-28). The labeled RNAs (1 nM) were titrated with a 2-fold dilutions series of protein in

384 well plate. FP was measured on a plate reader after 60 minutes of incubation at room temperature.

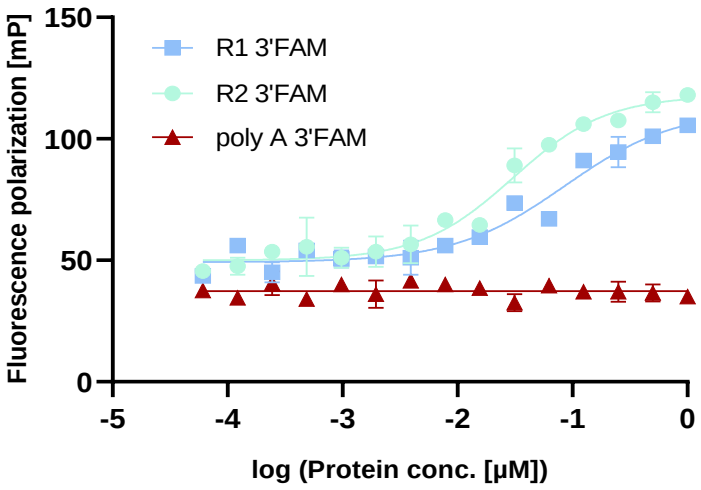


Figure 27: Fluorescence polarization curves of R1 and R2 RNAs titrated with RBM20 α3 (511–619). Polarization was measured after a 60-minute incubation period and presented as an average of technical replicates (N=2). Poly A (AAAAAAA) RNA was chosen as a negative control.

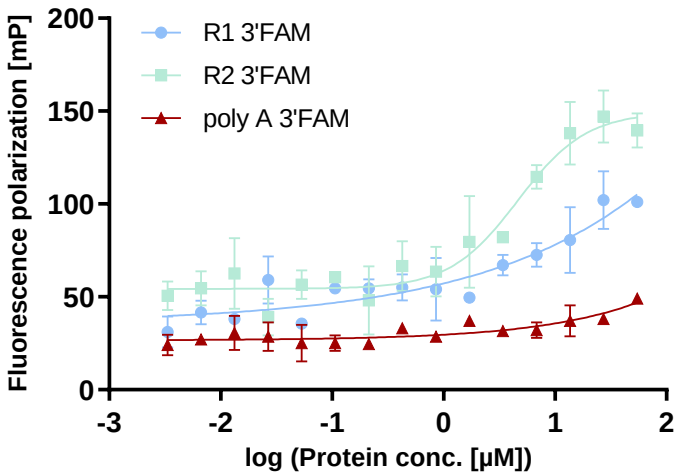


Figure 28: Fluorescence polarization curves of R1 and R2 RNAs titrated with RBM20 Δα3 (511–601). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2). Poly A (AAAAAAA) RNA was employed as a negative control.

Table 3: K_D values of the R1 and R2 RNAs for their binding affinity to the RBM20 protein with and without helix.

Protein	RNA	K_D (µM)
RBM20 α3	R1	0.087 ± 0.01
RBM20 α3	R2	0.030 ± 0.004
RBM20 α3	Poly A	-
RBM20 Δα3	R1	>50

RBM20 Δα3	R2	4.63 ± 0.32
RBM20 α3	Poly A	-

For both protein constructs, binding between RNAs and RBM20 is observable. However, the RBM20 Δa3 construct shows weaker binding affinity compared to the RBM20 α3. This data correlates with literature stating that helix 3 is required for tighter RNA binding, and absence of the helix results in up to 10-fold decrease in the binding affinity[122]. The K_D value for R1 with RBM20 Δα3 could not be determined as no saturation was reached with an increasing protein concentration. For R2 RNA, the K_D was in the low micromolar range for RBM20 Δα3, while with RBM20 α3, a low nanomolar K_D was observed. The control RNA (Poly A) does not show any significant change in FP, indicating no binding to the protein.

2.3.6. Direct fluorescence polarization assay with stapled peptides

FITC-labeled peptides (Figure 29, see section 2.3.3. and 5.10. for the synthesis) were evaluated in a direct FP assay for the determination of their binding affinity to RBM20 Δa3 (see Figure 30).

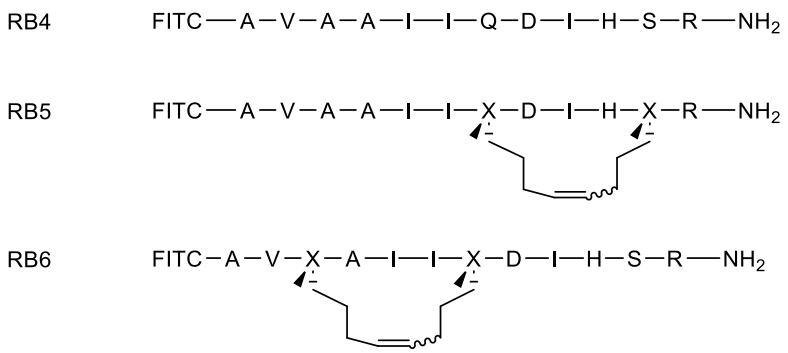


Figure 29: N-terminally FITC labeled peptides of derived from RBM20 helix 3. (X) represents the non-natural amino acid which was used to perform hydrocarbon stapling.

FITC labeled peptides were titrated with a 2-fold dilution series of the RBM20 Δa3 protein in 384 well plates. FP was measured after 60 minutes of incubation at room temperature and using a plate reader.

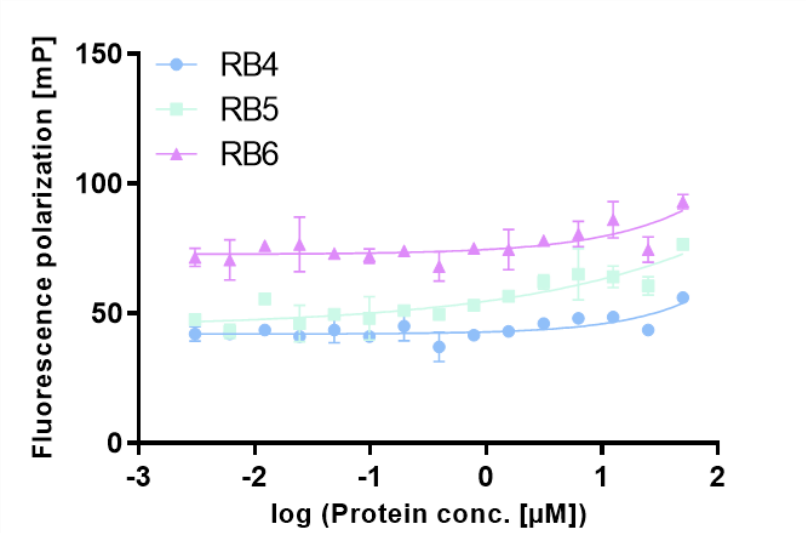


Figure 30: Fluorescence polarization curves of N-terminally FITC labeled peptides RB4, RB5, RB6 measured with RBM20 Δa3 (RBM20 without helix 3). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2).

Measurements showed a slight upward trend in FP with peptide; however, top-value saturation was not achieved. Hence, the data did not allow the calculation of accurate K_D values. The experiment was also carried out with C-terminally labeled peptides to exclude the possibility that the label might influence the binding. C-terminally labeled peptides (Figure 31) were synthesized applying the standard SPPS protocol[141] (see section 2.3.3.), but an additional lysine with an Mtt-protected side chain was added. After the ring-closing metathesis reaction, the Mtt group was removed from the resin-bound peptides by treatment with HFIP in DCM[142], as stated in section 5.10.1. Next, the peptide was labeled with FITC using the standard protocol.

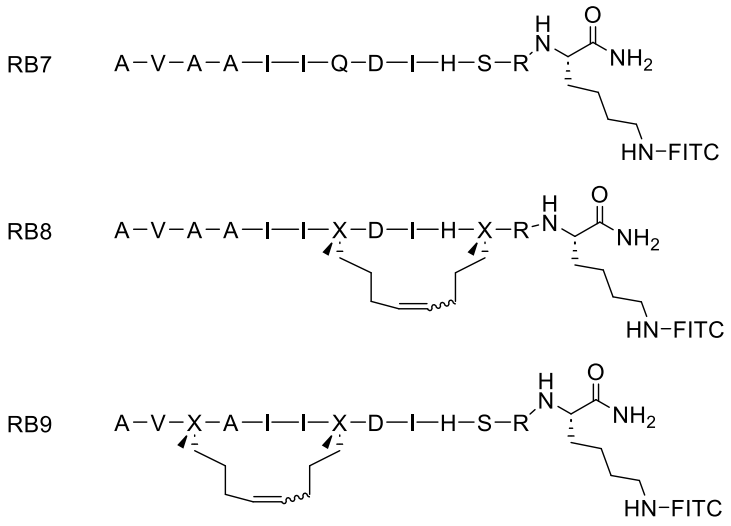


Figure 31: C-terminally FITC labeled peptides of RBM20 protein. Fmoc-Lys(Mtt)-OH was loaded to the solid support for C-terminal labeling.

C-terminally labeled peptides were evaluated in the FP experiment (Figure 32) under the same conditions applied before (see section 2.3.5.).

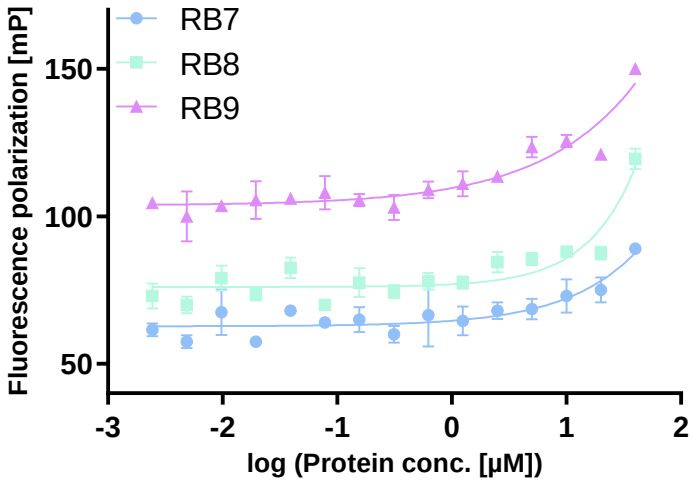


Figure 32: Fluorescence polarization curves of C-terminally FITC labeled peptides RB7, RB8 and RB9 measured with RBM20 Δa3 (RBM20 without helix 3). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2).

Even though an increased FP with C-terminally labeled peptides was achieved, top-value saturation could not be reached. Therefore, K_D values of the peptides could not be calculated. To be able to obtain a fitted curve, the experiment was repeated with an increased protein starting concentration (110 µM). N-terminally FITC-labeled peptides (RB4–RB6) were evaluated in the experiment.

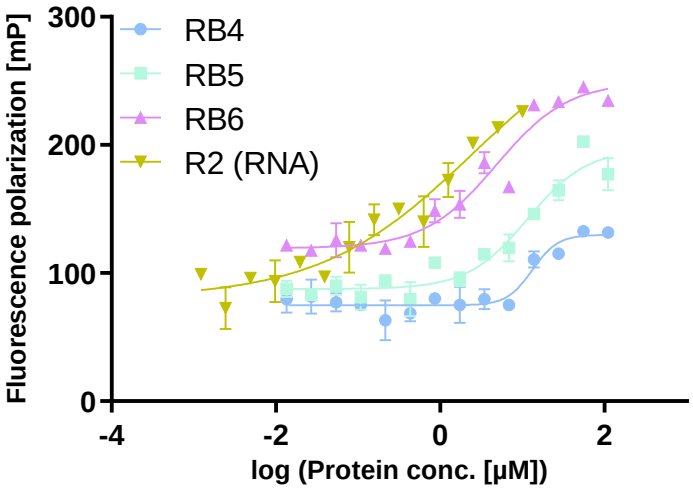


Figure 33: Fluorescence polarization curves of peptides RB4, RB5, RB6 and R2 RNA measured with RBM20 Δa3 (RBM20 without helix 3). As a positive control, R2 RNA was used in this experiment. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2).

Table 4: K_D values of the FITC-labeled peptides RB4, RB5, RB6 and the R2 RNA from the FP assay of the RBM20 Δa3 protein.

protein	Hit	K_D (µM)
RBM20 Δα3	R2	1.55 ± 1.07
RBM20 Δα3	RB4	14.32 ± 2.04
RBM20 Δα3	RB5	11.495 ± 2.03
RBM20 Δα3	RB6	4.78 ± 0.04

Binding between RBM20 Δa3 and peptides was observed with an increased protein concentration. K_D values were calculated, which are given in Table 4. Improved binding affinity with both stapled peptides was observed in comparison to the linear peptide, which could be explained by the helical structure of the stapled peptide that could better mimic the protein and, therefore, improve binding affinity to the protein.

2.3.7. Competition fluorescence polarization with RNA

Next, we wanted to evaluate the inhibitory properties of the peptides in the competition FP experiment against the RBM20/RNA interaction. First, a competition experiment with unlabeled RNAs was performed (Figure 34). Labeled R1 and R2 RNAs were used as tracers in the experiment in the presence of a fixed concentration of RBM20 α3. A 2-fold dilution series of unlabeled equivalent RNAs was added to the plate. The appropriate protein concentration for each RNA was determined based on the K_D values obtained from direct FP experiments[143].

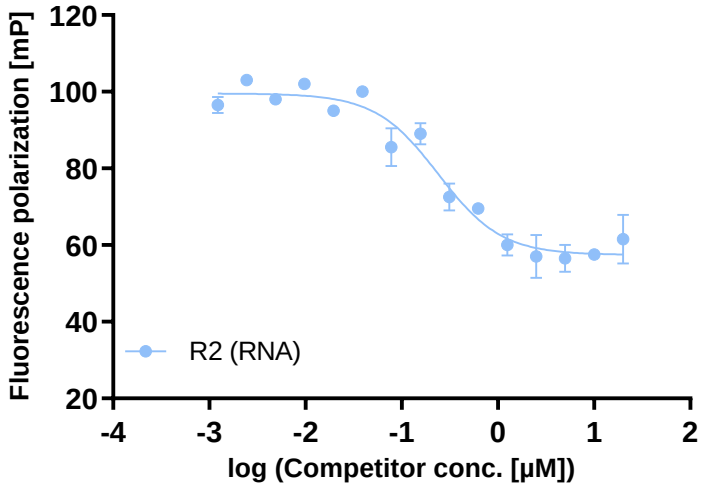


Figure 34: Competition FP curve of the 1nM of FAM labeled R2 RNA with 60 nM RBM20 α3. Unlabeled RNA was used as a competitor in this experiment. Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2).

Table 5: IC_{50} values calculated from the competition assay of RBM20 α3 with RNAs. RNA sequences are shown in section 2.4.5.

Protein	RNA	IC_{50} (µM)	K_I (µM)
RBM20 α3	R1	1.63 ± 0.59	0.536 ± 0.2
RBM20 α3	R2	0.251 ± 0.03	0.074 ± 0.01

Competition between labeled and unlabeled RNAs was observed. However, the R1 RNA shows a higher IC_{50} value than the R2 RNA, and a >6-fold difference between the two RNA motifs was determined. R2 RNA, which has a low nanomolar IC_{50}, can be used as a control in the competition experiments to evaluate the inhibitory properties of the peptides.

2.3.8. Evaluation of the peptides in competition FP assay

Three peptides (RB1-RB3, Figure 22) that were synthesized and N-terminally acetylated were tested in the competition FP experiment for their inhibitory properties against the interaction between 3'FAM labeled R2 RNA and RBM20 (Figure 35). Peptides with a 2-fold dilution series were incubated with premixed protein/labeled RNA with the appropriate protein concentration[143].

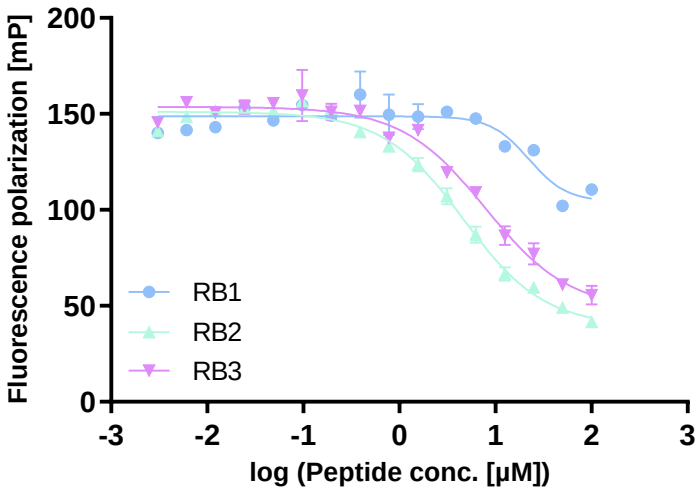


Figure 35: Competition FP of 100 µM of RB1, RB2 and RB3 with 60 nM RBM20 α3. Polarization was measured after a 30-minute incubation and presented as technical replicates (N=2).

Table 6: IC_{50} values of the RBM20 peptides calculated from the competition FP assay of RBM20 α3 and FAM labeled R2 RNA.

Protein	Peptide	IC_{50} (µM)	K_I (µM)
RBM20 α3	RB1	~100	~100
RBM20 α3	RB2	4.71 ± 1.14	1.567 ± 0.38
RBM20 α3	RB3	8.03 ± 1.69	2.681 ± 0.57

All the peptides showed a decrease in FP with an increasing peptide concentration. RB2 and RB3 peptides showed low micromolar IC_{50} values. However, since RB1 was not properly soluble, the experiment needed to be repeated for RB1; at the same time its solubility had to be improved for accurate determination of its IC_{50}. To improve the solubility of linear peptide (RB1), one of the amino acids, which is expected not to contribute to binding, was exchanged with arginine, which is predicted to improve solubility (Figure 36). Indeed, solubility was improved for the modified RB10 peptide. For the purity data, see Supplementary Figure 4.

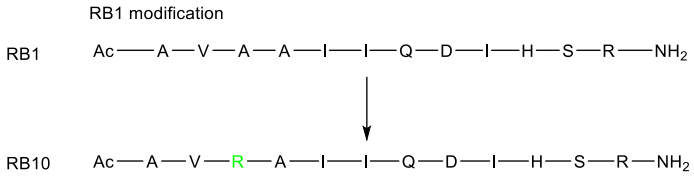


Figure 36: RB10 peptide, which is a modified analog of the RB1 to improve solubility. Modified position was shown in green.

Next, RB10 was tested in the competition FP experiment (Figure 37) to evaluate its inhibitory properties for the RBM20-RNA interaction. The same conditions mentioned above were applied to set up the competition experiment.

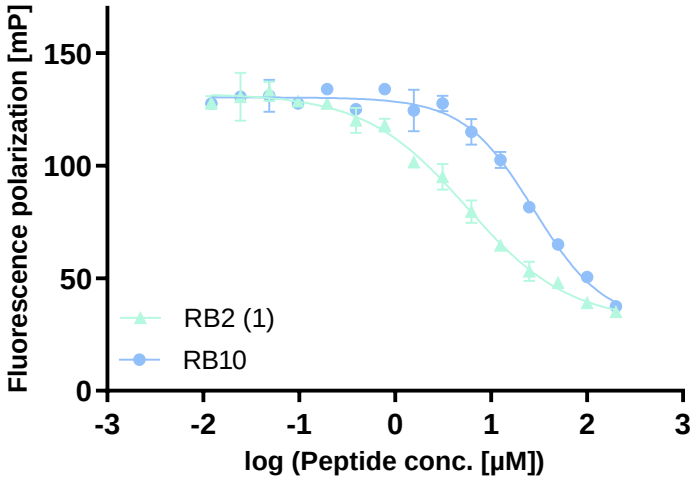


Figure 37: Competition fluorescence polarization of 100 µM of RB10 and RB2 with 60 nM RBM20 α3. As a positive control, RB2 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2).

In addition, as the stapled peptides showed a low micromolar inhibition for the RBM20-RNA interaction (Figure 35), they could be further optimized to improve the inhibitory properties of the peptides. For this purpose, creating a peptide library would be an option. However, preparing a library with a stapled peptide is synthetically challenging. Helicity can also be achieved by introducing α-aminoisobutyric acid (AIB) at specific positions in the peptide sequence[144]. Since peptides containing AIB residues are linear, they could be used to do an alanine scan and to

synthesize a library. AIB-containing peptides can be designed in a way that the AIB is placed in stapling positions (Figure 38).

Aib containing peptide

RB11 Ac—A—V—X—A—I—I—X—D—I—H—X—R—NH_2

Figure 38: AIB containing RB11 peptide. X represents α-aminoisobutyric acid.

RB11 was synthesized and involved in the competition FP experiment (Figure 39). The same conditions were applied to set up the experiment.

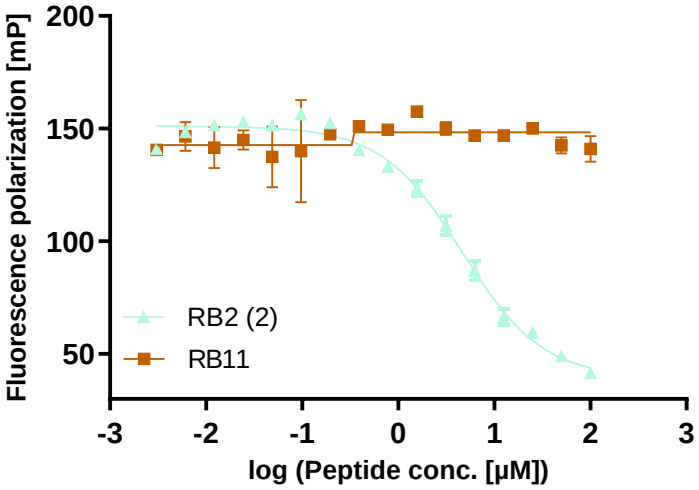


Figure 39: Competition fluorescence polarization of 100 µM RB2 and RB11 with 60 nM RBM20 α3 protein. As a positive control RB2 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2).

Table 7: IC_{50} values of the RBM20 peptides calculated from the competition FP assay of RBM20 α3 and FAM labeled R2 RNA.

Protein	Peptide	IC_{50} (µM)	K_I (µM)
RBM20 α3	RB2 (1)	5.66 ± 0.64	1.89 ± 0.22
RBM20 α3	RB10	27. 95 ± 0.82	9.36 ± 0.27
RBM20 α3	RB2 (2)	4.71 ± 1.14	1.567 ± 0.38
RBM20 α3	RB11	-	-

RB10 was able to compete with the RNA (Figure 37), while RB11 did not show any inhibition for the RNA-RBM20 interaction (Figure 39). Peptides (RB10 and RB11) were also spectroscopically analysed, and the obtained CD spectra (Figure 40) allowed the calculation of the percentage of helicity (Table 8). The calculated percentage helicity showed that RB10 and RB11 are similar to RB2 and RB3.

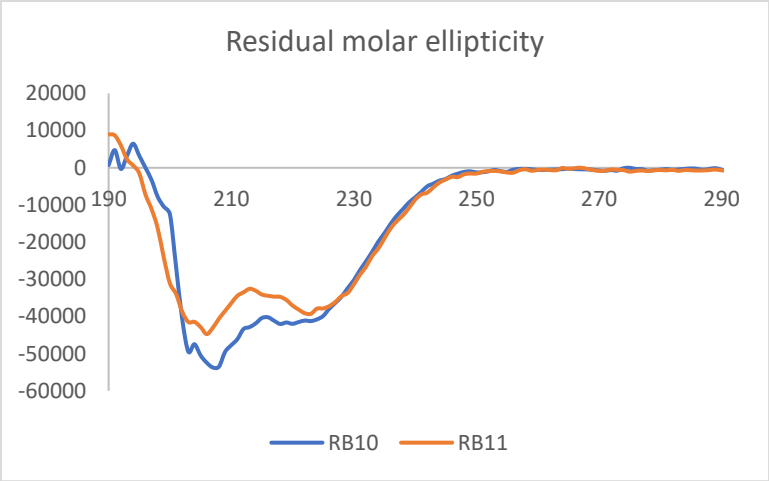


Figure 40: Representation of the CD spectra for RB10 and RB11. Buffer was used for a baseline correction.

Table 8: The percentage helicity of the peptides calculated based on the ellipticity values obtained from CD spectroscopy measurement.

Peptide	% Helicity
RB10	13.55
RB11	13.04

As working with AIB peptide for library design was not an option and instead RB10 had good solubility and maintained IC_{50} values in a low micromolar range it was chosen for further optimization.

2.3.9. Point mutations for further optimization of peptides

RB10 peptide was selected for optimization in which point mutations at specific positions were made based on NMR structures to achieve better interactions with RBM20. NMR analysis can be applied to obtain information about peptides and their intermolecular interactions[145]. Synthesis of the linear peptide is advantageous as it is time and cost-effective. Once improved binding affinity with a linear peptide was obtained, the mutation can be transferred to the stapled peptide to enhance binding affinity further. Accordingly, α-helicity, cell permeability, and stability can also be improved[76]. New peptides with point mutations (Figure 41) were used and purified for further investigations.

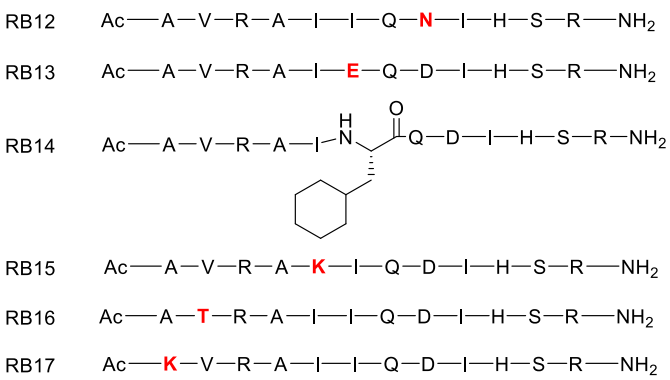


Figure 41: Variants of the RB10 peptides. Point mutations are highlighted in red.

A competition FP experiment was conducted with the mutated peptides to define their inhibitory properties for RBM20-RNA interaction (Figure 42). The same conditions (section 2.3.8.) were used to set up the competition experiment.

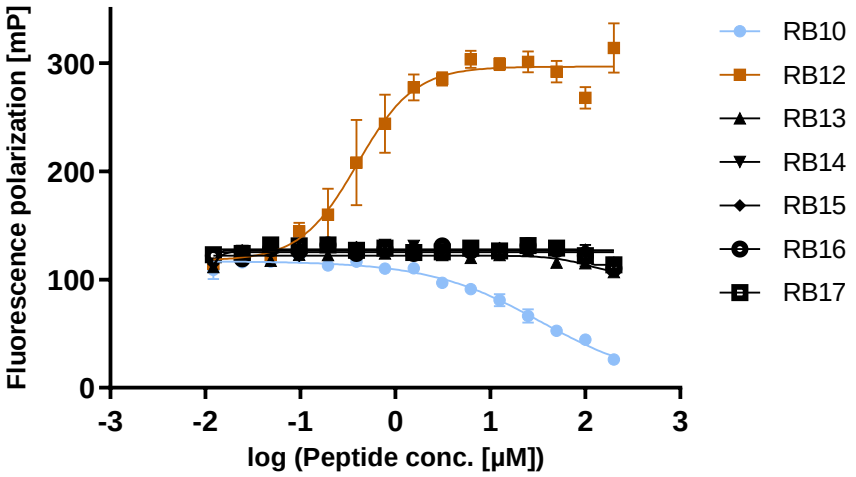


Figure 42: Competition FP of 100 µM of mutated linear peptides RB12-RB17 with 60 nM RBM20 α3. As a positive control, RB10 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2).

While competition for RB10 was observed, similar competition could not be reached for point mutations. In addition, for the RB12 peptide, a binding-like curve was obtained where the fluorescence polarization increased with an increasing peptide concentration, which might be related to the aggregation.

2.3.10. Fluorescence polarization and total intensity

When a fluorescent molecule is excited by plane-polarized light, it results in the emission of largely depolarized light due to the rotational diffusion during the lifetime of the fluorescence[146]. The emitted light passes through the excitation polarizing filter (the parallel polarizing filter) and a filter oriented at 90° to the excitation filter (the perpendicular polarizing filter) (Figure 43)[143]. To measure the FP, emission intensities can be collected from the channels that are parallel and perpendicular to the electric vector. Equation (2), which is defined as the difference between the parallel ($I_{\parallel}$) and perpendicular ($I_{\perp}$) emission intensities normalized by the total fluorescence intensity, can be used to determine FP.

$$(2) \quad FP = III - I_{\perp} / III + I_{\perp}$$

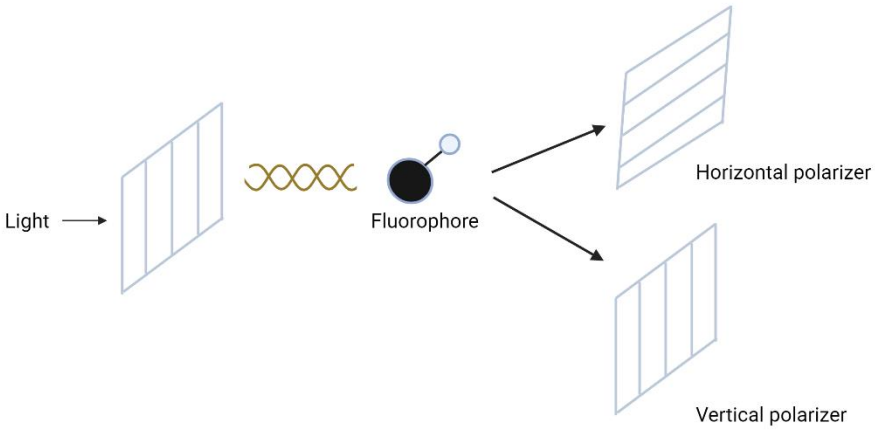
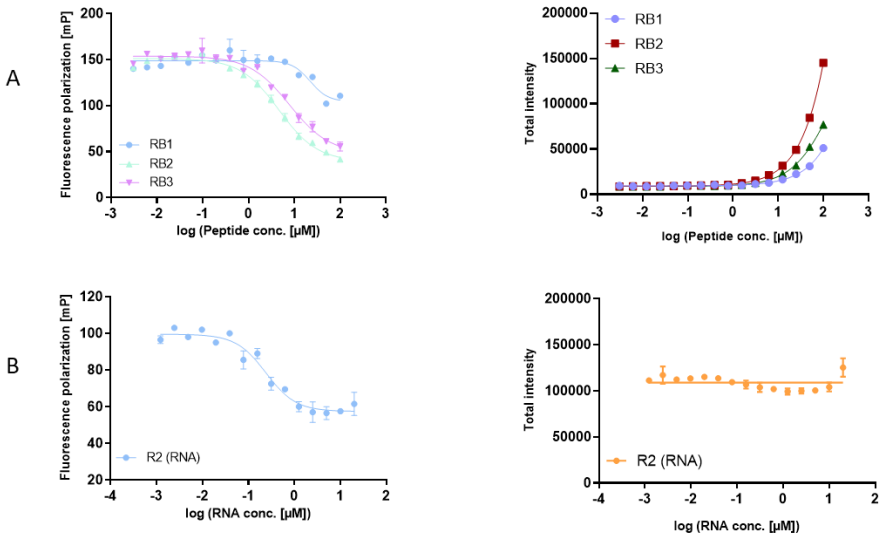


Figure 43: Principle of FP. A fluorophore is excited with light by passing through an excitation polarizing filter; the polarized fluorescence is measured either parallel or perpendicular to the exciting light's plane of polarization. Two intensity measurements are obtained ($I_{\perp}$ and $I_{\parallel}$) and used for the calculation of fluorescence polarization. Adopted from Lea and Simeonov et al[147].

It can be seen from Equation 2 that the FP value is independent of fluorophore concentration as it is not dependent on the absolute intensities of the emission light collected at either orientation[147]. Deviation from such independence of FP on the concentration of the fluorophore reagent can indicate fluorescence probe aggregation (anomalous FP increase and premature fluorescence intensity saturation upon increase of fluorophore concentration). On the other hand, a change in the fluorescence intensity of a fluorophore may result in significantly different contributions to the total fluorescence intensity of the sample, which can complicate the interpretation of FP measurements[148,149].

When we analyzed the raw data of the competition FP measurement, we observed that there was a significant change in the total fluorescence intensity with an increasing peptide concentration (Figure 44, A) which was not seen with the RNA in which total intensity stayed steady (Figure 44, B). Evaluation also demonstrates that with the peptides, the total intensity increased while the fluorescence polarization decreased and reached the highest value for RB2 (Figure 44, A).



C

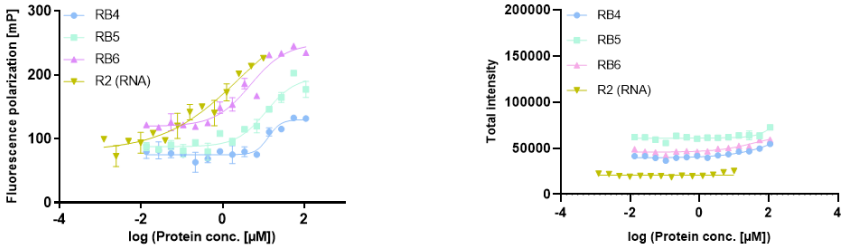


Figure 44: A) Competition FP and total intensity of RB1, RB2 and RB3 with RBM20 α3 (see Figure 35); B) Competition FP of R2 RNA (see Figure 34) with RBM20 α3 and total intensity change with increasing RNA concentration; C) Direct FP of RB4, RB5 and RB6 with RBM20 Δα3 (see Figure 33) and total intensity change with an increasing peptide concentration

Even though the direct FP experiment successfully confirmed the binding between designed peptides and RBM20 with a low micromolar K_D (see Figure 33) and no significant change was observed for total intensity (Figure 44, C) the same cannot be said for the competition experiment. Although peptides (RB1-RB3) in competition experiments initially showed inhibition, this data was not reliable due to the unstable total intensity values. Therefore, this experiment was not suitable to interpret the inhibitory properties of the peptides. In conclusion, stapled peptides showed binding to the RBM20 protein, but their inhibitory properties for the RNA-RBM20 interaction could not be proven.

It has previously been shown that RBM20 regulates titin splicing.[150] Titin truncation variants are the most frequent cause of dilated cardiomyopathy, one of the main causes of heart failure and heart transplant. Therefore, evaluating the stapled peptides on TTN splicing would be an interesting target. However, first biophysical experiments such as MST, ITC, or pull-down can be carried out to further study binding properties of peptide to RBM20.

3. Targeting the interaction of hnRNP A2/B1 with RNA

3.1. Introduction

hnRNPs, the heterogeneous nuclear ribonucleoproteins, are RNA-binding proteins involved in various RNA metabolic pathways such as pre-mRNA processing, mRNA stabilization, and splicing which makes them crucial in the regulation of gene expression[151]. Although hnRNP family proteins have different combinations of various domains, they show similar properties to each other. Structural investigation of hnRNPs clarifies that proteins are mainly composed of the RRM, qRRM (quasi-RNA recognition motif), glycine-rich domains, KH domains, and an RGG box domains consisting of Arg-Gly-Gly repeats. hnRNPs are named in alphabetic order, starting with hnRNP A1 and ending with hnRNP U, having molecular weights in a range between 34 and 120 kDa (Figure 45)[152].

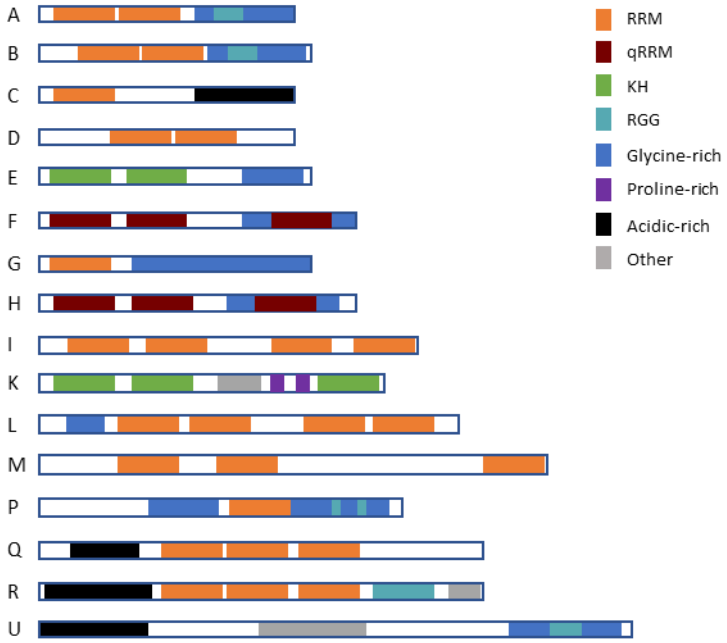


Figure 45: Domain construction of the different hnRNP proteins. Adapted from Geuens et al[152].

Some hnRNP proteins alter alternative splicing patterns, resulting in different isoforms caused by exon skipping or other alternative splicing mechanisms[152]. This modulation can be direct for some proteins such as hnRNP H and A1 or indirect, for example, in hnRNP M. hnRNP H is one of the splicing factors resulting in the formation of human epidermal growth factor receptor 2 (HER2) splice variants, which is important for the HER2 signaling pathway (Figure 46)[153]. hnRNP M is found in breast tumors and involves in the activation of an AS switch, thus regulating CD44 splice isoforms[154]. In addition, hnRNP A1 involves in alternative splicing of various genes such as c-Src, caspase-2, SMN2, and even itself[155]. hnRNP A1 collaborates with other splicing factors by interacting with splicing regulatory elements, protein domains, and splicing sites[156].

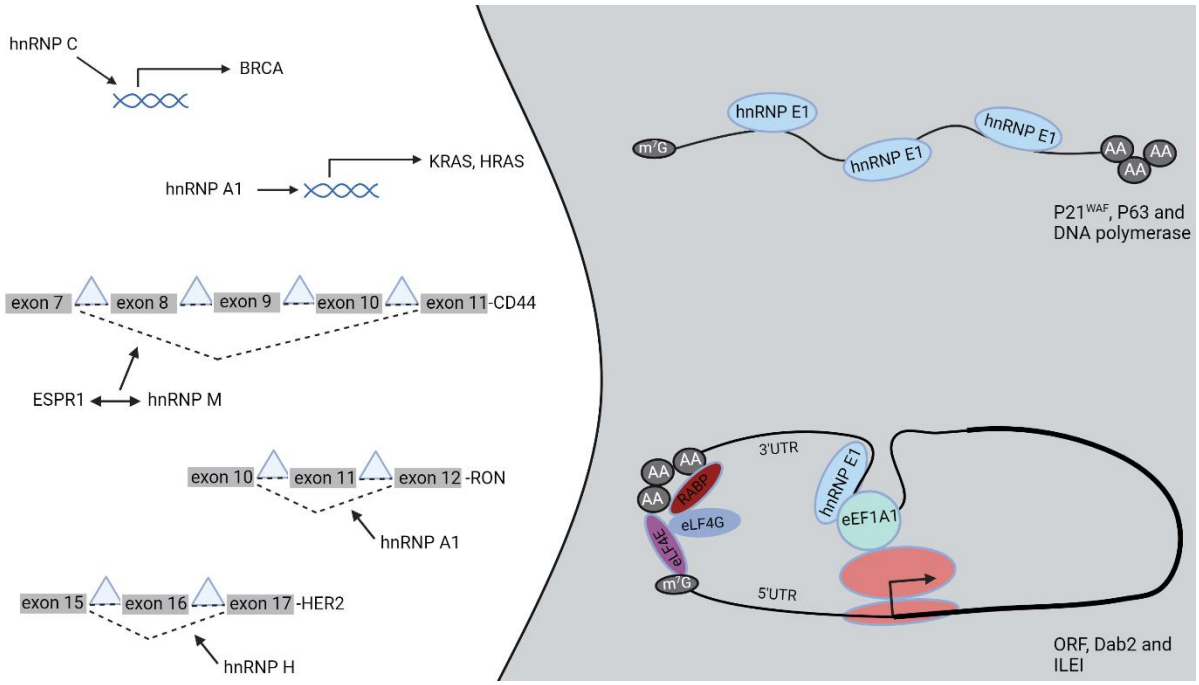


Figure 46: hnRNPs that are involved in cancer and metastasis. hnRNP C and A1 are acting on the promotor of BRCA and KRAS, HRAS, respectively[157,158]*. hnRNP E1 is known to be involved in the regulation of mRNA half-life and mRNA stability has a relation with the growth of cancer*[159–161]*. The recent discoveries on hnRNP E1 which plays a key role in Transforming growth factor beta (TGF-β)-mediated epithelial-mesenchymal transition (EMT) has clarified that hnRNP E1 is regulating oncogenes such as Dab-2 and ILEI*[162]*. Adapted from Geuens et al*[152]*.*

One of the subfamilies of the hnRNP proteins is the hnRNP A/B family, which has four members: hnRNP A1, A2/B1, A3, and A0 of which A2/B1 are the most studied proteins, and they are found to be involved in mRNA translation[163] and splicing[164]. Furthermore, hnRNP A2/B1 plays a significant part in neuronal mRNA trafficking[165]. hnRNP A/B is an important member of the hnRNPs, and increasing evidence shows that they play a role in tumor development[166]. In previous studies, overexpression of hnRNP A2/B1 in lung and breast cancers was observed[167,168] and in breast cancer cells, knockdown of hnRNP A2/B1 induced apoptosis[169]. Moreover, recent studies found that the splicing factors hnRNP A1 and A2/B1 can modulate the alternative splicing of PKM2[169]. This activates a metabolic switch to aerobic glycolysis, which is a necessary process for cancer cells.

hnRNP A2/B1 exists as two isoforms, where an additional 12 amino acids at the N-terminus differentiate the B1 from the A1 isoform. Both have an RBD consisting of RRM1 and RRM2, as well as a Gly-rich domain including a prion-like domain (PrLD), PY-motif and an RGG box (Figure 47)[170,171].

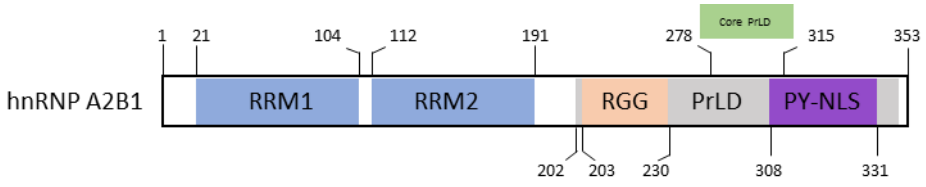


Figure 47: Domain structure represented based on the hnRNP B1 isoform. Adapted from Wu et al[172]*.*

hnRNP A2/B1 can bind specific RNA sequences and as a result, affects mRNA processing. In addition to the mRNAs, it also takes part in the regulation of other RNAs. For instance, it binds to HIV-1 RNA, leading to vRNA nuclear retention (vault RNA) and microRNAs[173,174]. hnRNP A2/B1 can also boost the long non-coding RNA HOTAIR (HOX Transcript Antisense Intergenic RNA) association to intervene in HOTAIR-dependent heterochromatin initiation[175].

To uncover the RNA-binding specificity and characteristics structural studies of hnRNP A2/B1 in combination with various RNAs have been carried out[172]. The crystal structures show a well-defined RNA binding groove that recognizes UAG and AGG by the RRM2 and RRM1 domains, respectively.

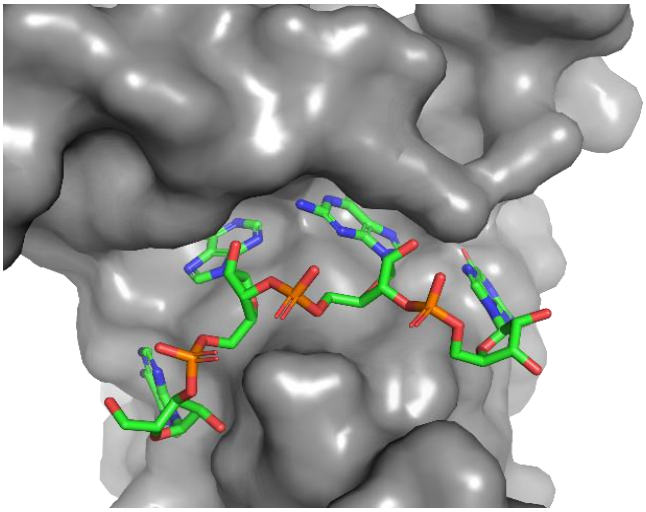


Figure 48: Representation of RRM1 domain of the protein bound to AAGG RNA. Adapted from Wu et al[172].

It was demonstrated that, mutation of the UAG motif which is recognized by RRM2, resulted in just a slight reduction in binding affinities, while mutating the AGG interacting with RRM1, such as 2nd G to C or 3rd G to C had a more apparent influence on the binding affinities. Crystallography studies propose that RRM1 mainly recognizes purine-rich motifs, whereas RRM2 seems to recognize RNA sequences with less restrictions, containing either UAG, GAG, or pyrimidine containing RNAs.

It is also reported that N6-methyladenosine (m^6A), one of the most abundant internal modifications of messenger RNA, may provide hnRNP A2/B1 access to certain binding sites[176]. For example, the presence of m^6A may ease hnRNP A2/B1 ability to improve primary microRNA processing[177,178]. Molecular studies have demonstrated that m^6A can impact alternative splicing[179]. This finding can connect the correlation between hnRNP A2/B1 and m^6A and implicate hnRNP A2/B1 as a nuclear reader and effector of the m^6A mark. However, further investigations are needed to better understand the mechanism of action[178].

3.2. Aim

The aim of this work is to develop inhibitors for hnRNP A2/B1 protein which is an RNA-binding protein and is upregulated in various types of cancer. Initially, PNAs which would mimic the RNA to target the hnRNP A2/B1-RNA binding site were synthesized and evaluated.

In the second part of the chapter, cyclic peptide hits of hnRNP A2/B1, identified from the screening-based approach, were synthesized and evaluated for their inhibitory properties. Peptides were first evaluated in fluorescence polarization assay, and selected hits were further evaluated for their activity in the microscale thermophoresis assay.

3.3. Results

3.3.1. Targeting hnRNP A2/B1-RNA interaction with peptide nucleic acids

The importance of the hnRNP A2/B1 protein in alternative splicing and its upregulation in various cancer types[169,180,181] as well as RNA binding features of RRM domains of hnRNP A2/B1 have been already discussed in the previous sections[172]. To target the hnRNP A2/B1-RNA binding site a strategy using PNA scaffolds modified with fragments was chosen. RNA can be mimicked by PNA where the deoxyribose phosphate backbone is exchanged by a peptide backbone (Figure 49, see section 1.7)[99].

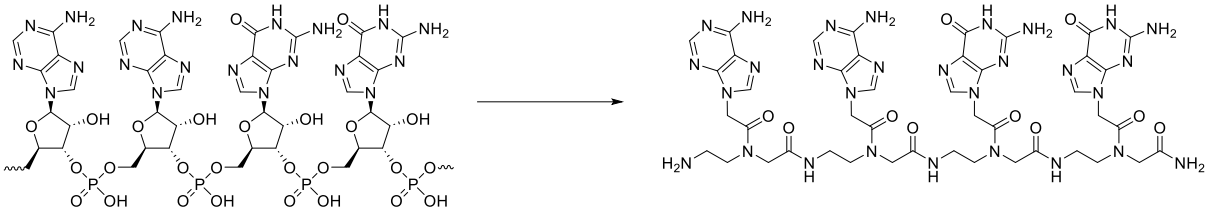


Figure 49: Mimicking RNA oligonucleotide by peptide nucleic acid where the phosphodiester backbone of RNA was replaced by a peptide backbone. Adopted from Wu et al[100].

3.3.2. Peptide nucleic acid synthesis

Initially, the synthesis of a 4mer PNA with AAGG nucleobases linked to the backbone was planned. The 4mer PNA would mimic the AAGG motif of the RNA (10mer: AAGGACUAGC) shown bind to the RRM1 domain of the hnRNP A2/B1 protein (see section 3.1.)[172].The synthetic strategy for PNAs relies on standard solid phase based synthesis (Figure 50, see section 5.11.)[129,182]. The Fmoc group is used as N-terminal protecting group while the benzhydryloxycarbonyl (Bhoc) group is used for the amino groups of nucleobases.

H-Rink Amide ChemMatrix resin was chosen for the synthesis. Once the amino acid loading, which was carried out in the presence of Fmoc-amino acid, COMU, and DIPEA (see section 5.9.2.), was complete, capping of the free amino group was performed with Ac2O and DIPEA in DMF to prevent any side products. After resin loading determination (see section 5.9.5.) coupling of the next monomer was carried out using the corresponding Fmoc-protected PNA monomer, COMU, and DIPEA. Coupling was followed by Fmoc group removal with 20% pyrrolidine in DMF, and washing between every step was performed in DMF. The steps mentioned above were repeated until the desired sequence was obtained. For fluorescent labeling, a linker was attached to the N-terminus of the PNA applying the above-mentioned coupling conditions, and this was followed by attaching the fluorophore (Fluorescein I) using FITC isomer, and DIPEA in DMF. The success of the synthesis was monitored by LC-MS, which was followed by purification in RP-HPLC.

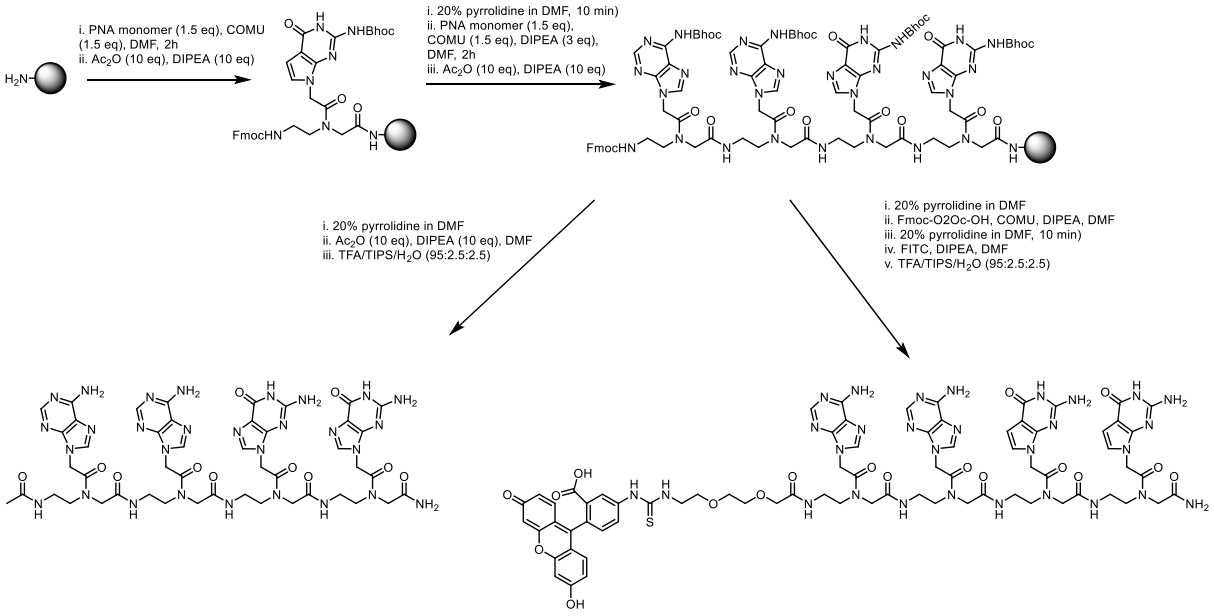


Figure 50: Synthesis scheme of the 4mer PNA (AAGG) referring to the solid phase-based synthesis protocol[129].

3.3.3. Direct fluorescence polarization assay with RNAs

Isothermal titration calorimetry (ITC) data have been reported in the literature for the interaction of hnRNP A2/B1 with an appropriate RNA nucleotide sequence[172]. To confirm the ITC results from the previous studies an FP experiment was designed. For the direct FP experiment 3'FAM labelled RNAs with the same RNA nucleotide sequences as reported in the literature were selected which were AGGACUGC (8mer) and AAGGACUAGC (10mer) (Figure 51). hnRNP A2/B1 (residues 1-251) with an N-terminal MBP tag was used where the MBP tag was added to improve purification success and yields. Labeled RNAs (1 nM) were used and titrated with a 2-fold dilutions series of protein. Polarization was measured on a plate reader after 30 minutes of incubation at room temperature.

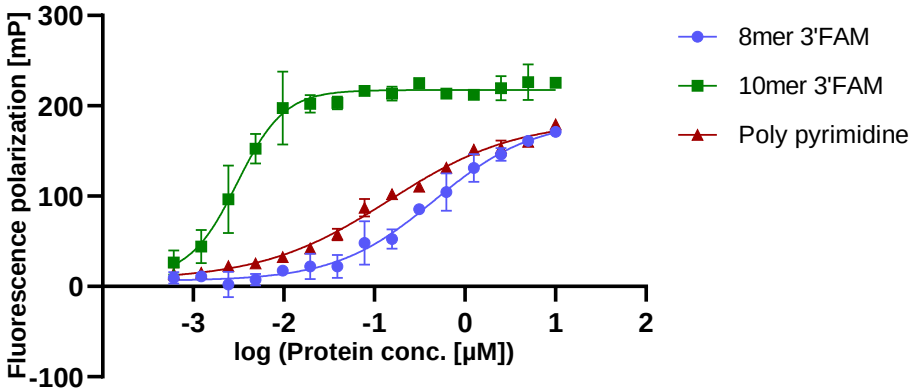


Figure 51: Direct FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 3'FAM labeled AGGACUGC (8-mer) in blue and AAGGACUAGC (10-mer) in green. As a control, FAM labeled polypyrimidine was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2).

For both RNAs a binding curve can be fitted which confirms the binding event between hnRNP A2/B1 and the RNA. The K_D values calculated from the obtained data are shown in Table 9. In

addition, a polypyrimidine sequence (AUUUUUCCAUCUUUGUAUC) that was intended as a negative control gives a binding curve meaning that the protein interacts with the pyrimidine sequence.

Table 9: K_D values for the 8-mer and 10mer RNAs tested in FP assay for binding affinity with hnRNP A2/B1.

RNA	K_D (µM)
8mer 3'FAM	0.457 ± 0.202
10mer 3'FAM	0.00314 ± 0.0001
Poly pyrimidine 3'FAM	0.147 ± 0.005

Next, to investigate the contribution of the label to the binding affinity 5'FAM labeled RNA was also tested under the same conditions (Figure 52). Since the polypyrimidine bound to the target protein and therefore cannot be used as a negative control it was excluded from the experiment. Instead the CCCCCCCC RNA sequence (Poly C) was used for the following experiment as a negative control.

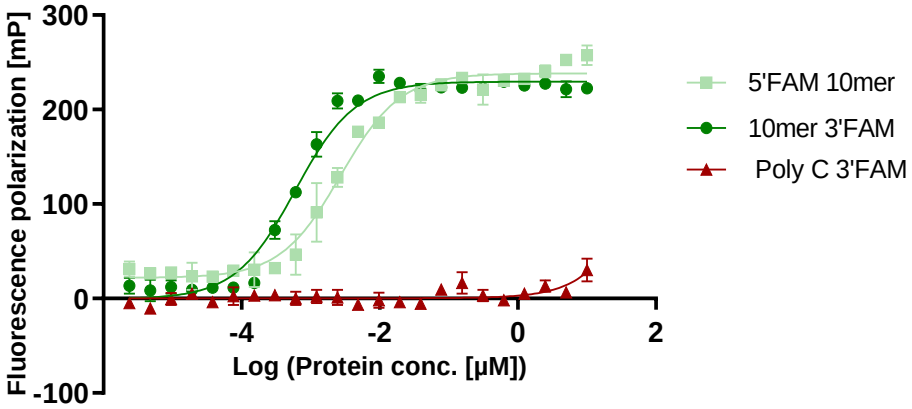


Figure 52: Direct FP curves of N-terminal MBP tagged hnRNP A2/B1(residues 1-251) with 3'FAM (dark green) and 5'FAM (light green) labeled 10mer RNA (AAGGACUAGC). As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation period and presented as a technical replicate.

Table 10: K_D values for 3'FAM and 5'FAM labeled 10mer RNA tested in FP assay for binding affinity with hnRNP A2/B1.

RNA	K_D (µM)
10mer 3'FAM	0.0064 ± 0.0001
5'FAM 10mer	0.0026 ± 0.0005
Poly C 3'FAM	-

From the binding curves of this FP measurement the K_D values were calculated and indicated that the linkage of the FAM label on either side does not affect the binding affinity of RNA to the target protein to a significant degree (Figure 52 and Table 10). Therefore, 3'FAM RNA was used for the further experiments. On the other hand, Poly C RNA did not show a significant increase in fluorescence polarization, which makes it a suitable control for the direct FP experiment.

3.3.4. Competition FP experiment with RNAs

Next, a competition FP experiment (Figure 53-54) for the hnRNP A2/B1 and RNA was performed where the unlabeled RNAs with the same nucleotide sequence were used as competitors. This would allow to determine how RNAs compete with each other and this can be utilized as a control later in the competition FP with other inhibitor molecules. For this purpose, two unlabeled RNAs with the same sequence (10mer: AAGGACUAGC and 8mer: AGGACUGC) and MBP-tagged hnRNP A2/B1 (residues 1-251) were used. The unlabeled RNAs were added to the plate as a 2-fold dilution series and were titrated with a premixed protein/labeled RNA with fixed protein and RNA concentrations[143]. According to the data from the direct FP measurements (shown in Table 10) a 5 nM and 625 nM of a protein concentration should be used for the 10mer and 8mer RNAs, respectively.

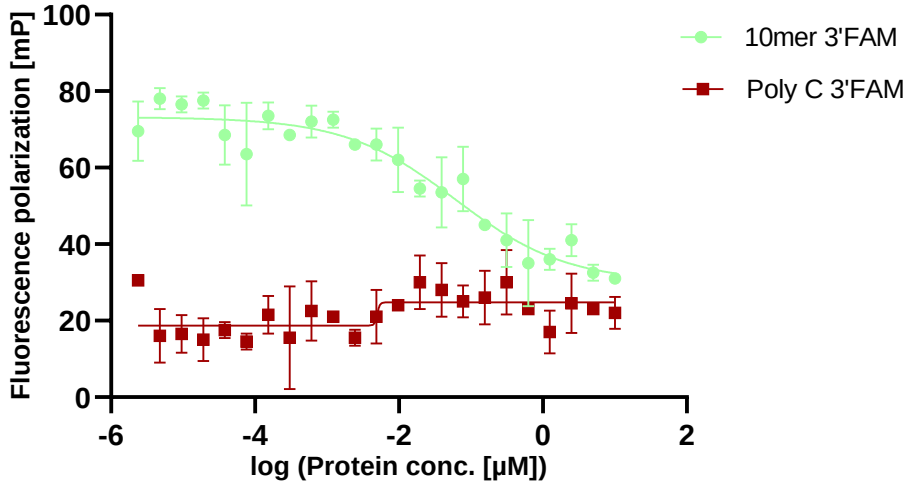


Figure 53: Competition FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 1 nM 10mer 3'FAM RNA (AGGACUGC). 10 µM of unlabeled 10mer RNA was used as a competitor. As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2).

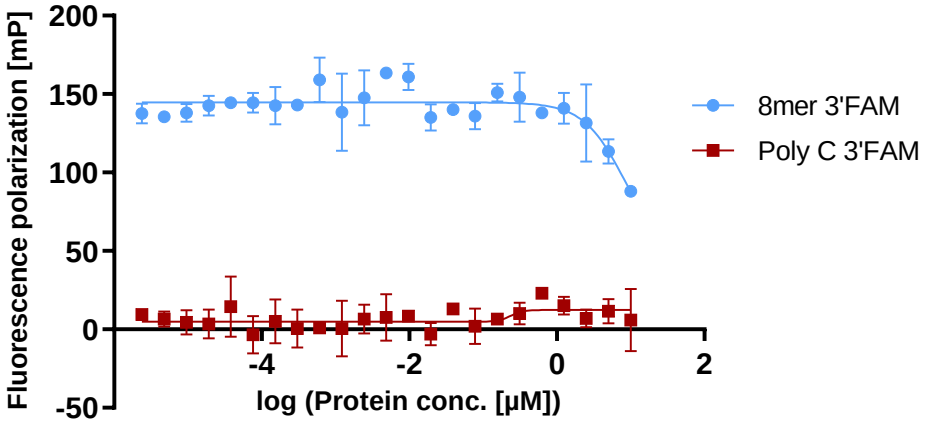


Figure 54: Competition FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 1 nM 8-mer 3'FAM RNA (AAGGACUAGC). 10 µM of unlabeled 8mer RNA was used as a competitor. As a negative control, poly C was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2).

For both RNAs a decrease in fluorescence polarization was observed with an increasing unlabeled RNA concentration. However, for the 8mer RNA only a slight decrease in fluorescence polarization was observed and a competition curve could not be fitted. For the 10mer RNA the competition curve can be fitted which indicates the competition between labeled and unlabeled RNAs. The IC_{50} value for 10mer RNA shown in Table 11 was obtained. Since the 10mer RNA results in better IC_{50}, it can be used as a positive control in competition experiments to evaluate the activity of the molecules.

Table 11: Competition FP derived IC_{50} values for the 8mer and 10mer RNAs.

Labeled RNA	IC_{50} (µM)
8mer RNA 3'FAM	>10
10mer RNA 3'FAM	0.067 ± 0.063
Poly C 3'FAM	-

3.3.5. Direct fluorescence polarization assay with PNAs

After the successful synthesis of the fluorescently labelled PNA it was tested in a direct FP assay to investigate the affinity of the PNA for hnRNP A2/B1 (Figure 55). For this purpose, different hnRNP A2/B1 constructs, spanning residues 1-104, 12-195, or 1-251 covering RRM1 only, RRM1 and RRM2 or RRM1 and 2 and part of the RGG box, respectively, were used. Proteins were applied as a 2-fold dilution series in the plate and titrated with FITC labeled peptides (1 nM).

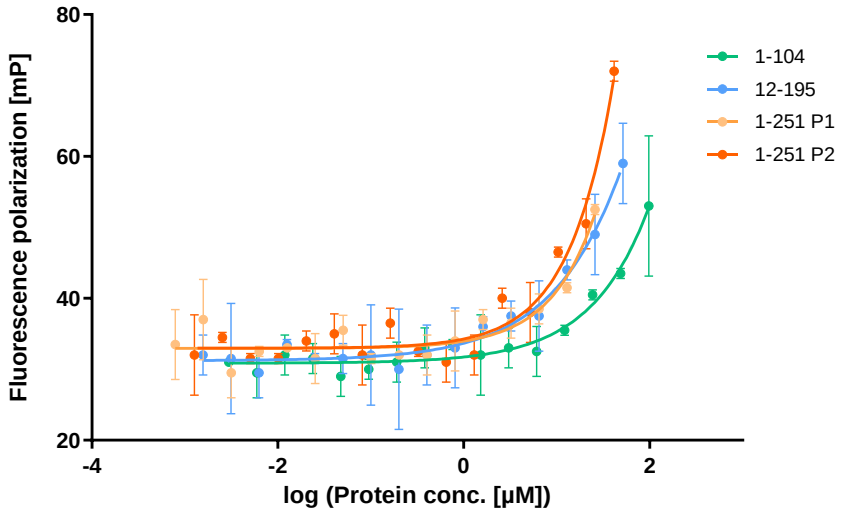


Figure 55: Direct FP curves of hnRNP A2/B1 constructs, 1-104 (RRM1), 12-195 (RRM1 and RRM2), 1-251 (construct containing RGG box: P1 and P2) with FITC labeled 4mer PNA (AAGG). During the protein purification hnRNP A2/B1 (residues 1-251) was obtained as two separate fractions (possibly monomer and dimer) and both were tested. Measurements were performed as technical replicates (N=2).

There was an increase in fluorescence polarization with an increasing PNA concentration. However, accurate K_D values could not be calculated as full curves could not be fitted. This could be explained by the loss of binding affinity with a short PNA having only four nucleobases in the

sequence. Therefore, it would be interesting to test the binding affinity of longer PNAs that mimic the 10mer and 8mer RNAs which both have a nanomolar binding affinity to the target protein[172].

3.3.6. PNA monomer synthesis

For the synthesis of 10mer and 8mer PNAs a new PNA monomer (Compound 4, Fmoc-PNA-U-OH) was needed which is not commercially available. Fmoc-PNA-U-OH was synthesized based on previously reported strategies (Figure 56, see section 5.11.3.)[183,184]. Initially, to a solution of uracil acetic acid (1) HBTU and DIPEA were added. After stirring for 2 min, Fmoc protected aminoethyl glycine (2) was added to the reaction mixture. To remove the *tert*-butyl protecting group from compound (3) TFA was used. Once the reaction was complete the solvents were evaporated and the product (4) purified by column chromatography resulting in a white powder.

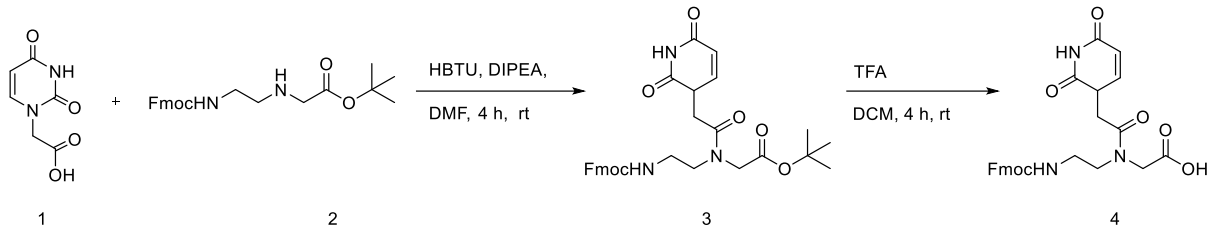


Figure 56: Synthesis scheme of the PNA monomer (Fmoc-PNA-U-OH) needed for 10mer and 8mer PNA synthesis in solid phase.

3.3.7. Synthesis of 10mer and 8mer PNAs

For the synthesis of the longer PNAs (Figure 57) the same solid phase synthesis strategy as used previously was applied (see section 3.3.2. and 5.11.)[129,182]. However, the synthesis of the 10mer and 8mer PNA sequences was unsuccessful. Although synthesis of short PNA sequences was successful it becomes challenging to complete the full sequence after five or six monomers have been connected.

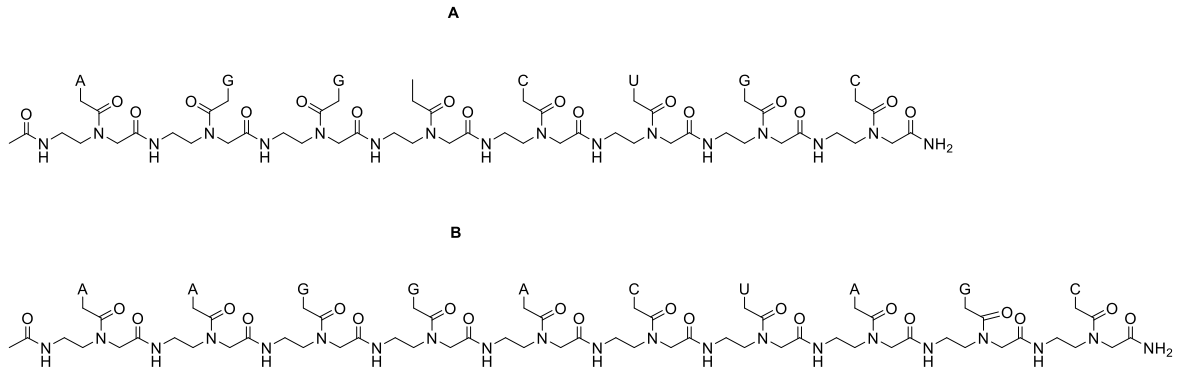


Figure 57: Representation of 8mer and 10mer PNA sequences: A) 8-mer PNA sequence; B) 10-mer PNA sequence.

It was necessary to develop a strategy which allows easier synthesis and should have better binding properties than PNA. Based on the RNA sequences, adenine and guanine seemed to be the most important nucleotides for recognition by hnRNP A2/B1 and would probably serve as the

main source of affinity. As an alternative, we hypothesized that the starting point of an inhibitor could be found in these single nucleotides. Once inhibitory effects were observed for such a molecule it could be further optimized by addition of amino acids. The other amino acids would serve to improve potency and create a selective molecule to avoid binding to similar proteins. The first step would be to see how mononucleotides can inhibit the protein-RNA interaction. These would include adenosine (5), adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7), guanosine (8), guanosine 3'monophosphate (9), guanosine 5'-monophosphate (10) (Figure 58).

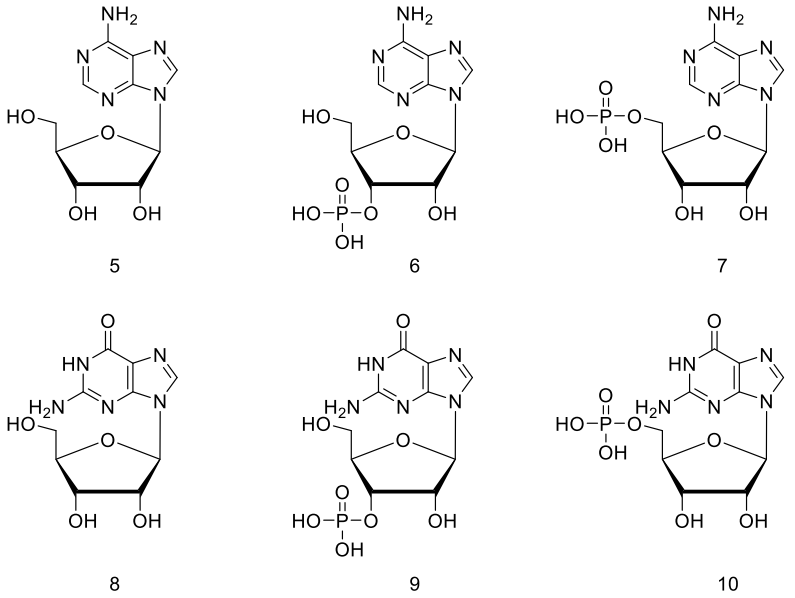


Figure 58: Structures of the selected mononucleotides represented from 5 to 10. Guanosine 3'-monophosphate (9) was not commercially available and therefore, was not involved in further experiments.

Next, it would be interesting to see whether the amide variants 11 and 12 (Figure 59) can inhibit the interaction between RNA and hnRNP A2/B1.

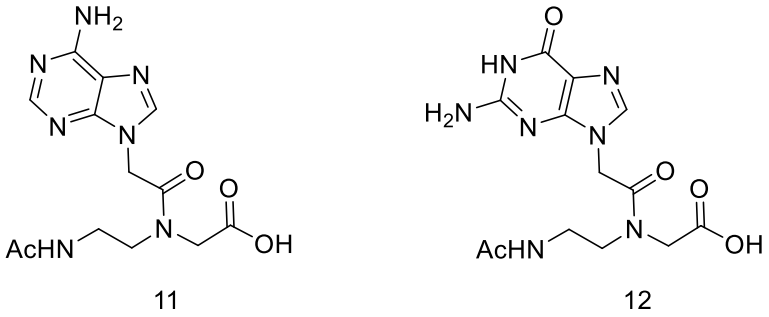


Figure 59: Structure of the synthesized PNA variants having adenine (11) and guanine (12) on the side chain. The synthesis was carried out in solid phase.

PNA variants were synthesized on solid phase following the SPPS protocol (see section 5.11.2.)[129]. 2-Chlorotrityl chloride (2-CTC) resin was chosen as a solid support for the synthesis. After loading the PNA monomer the N-terminus Fmoc group was deprotected and the free amino group was acetylated with acetic anhydride. PNA variants were cleaved from resin using TFA in DCM. Solvents were removed and purification was carried out using preparative HPLC.

3.3.8. Competition experiment with mononucleotides and PNA variants

Once the compounds were obtained either commercially or synthetically they were evaluated in the competition FP assay against the 10mer RNA (Figure 60). MBP-tagged hnRNP A2/B1 (residues 1-251) with an appropriate concentration[143] and FAM labeled 10mer RNA (1 nM) were involved in the experiment. Unlabeled RNA or mononucleotides were added to the plate as a 2-fold dilution series and FP was measured after 30 minutes incubation.

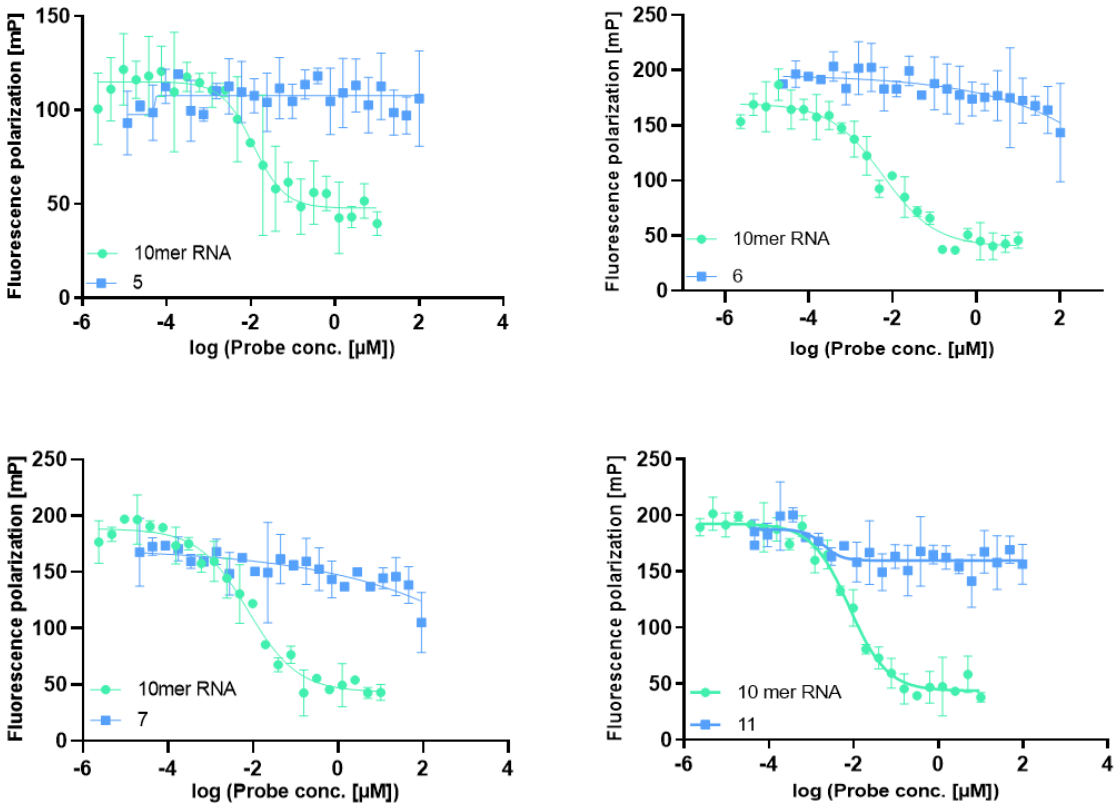


Figure 60: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: adenosine (5) adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7) and PNA variant (11) against the 5 nM hnRNP A2/B1 interaction with FAM labelled 10mer RNA. As a positive control unlabeled 10mer RNA was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates.

Since there was a slight decrease observed in fluorescence polarization for adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7), the experiment was repeated with an increased compound concentration (Figure 61).

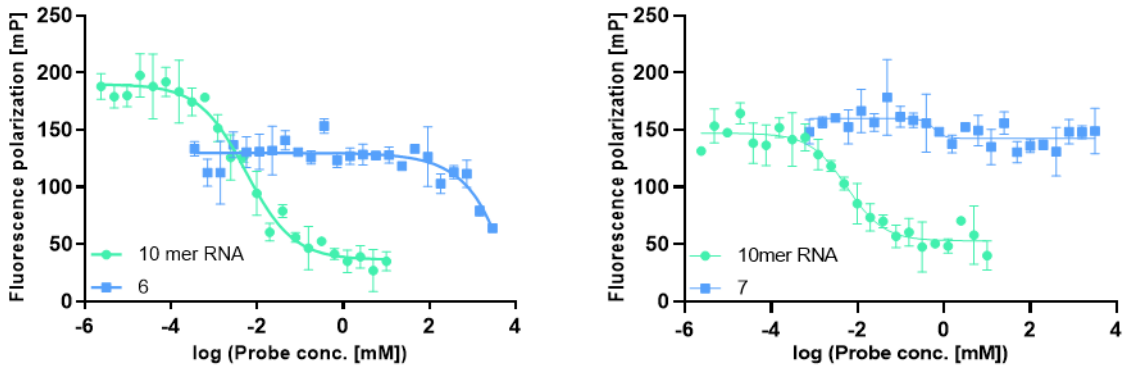


Figure 61: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7) against the 5 nM hnRNP A2/B1 interaction with FAM labeled 10mer RNA. Mononucleotides were used in mM concentration range. As a positive control, unlabeled 10mer RNA was used and shown as technical replicates.

Since hnRNP A2/B1 can bind guanosine as well, guanosine variants were also tested in the competition FP experiment (Figure 62). This time a higher compound concentration (mM range) was applied directly.

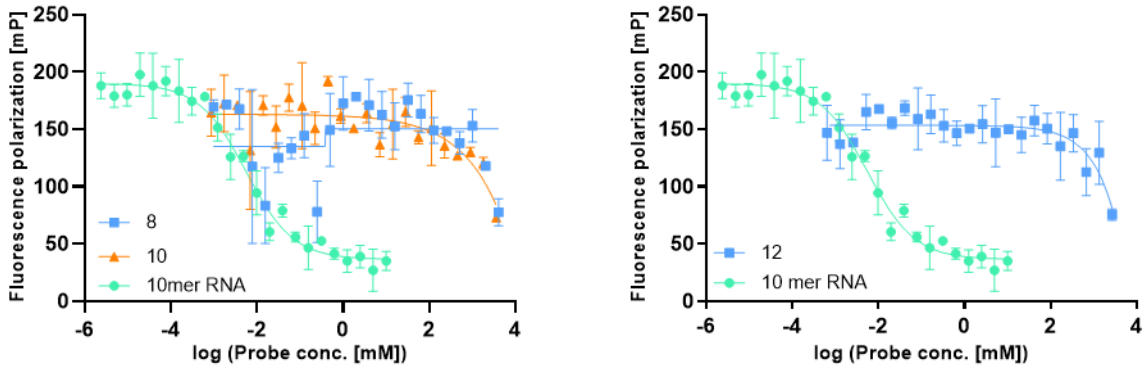


Figure 62: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: guanosine (8), guanosine 5'-monophosphate (10) and PNA variant (12) against the 5 nM hnRNP A2/B1 interaction with FAM labeled 10mer RNA. Unlabeled compounds were used in mM concentration range. Unlabeled 10mer RNA was used as a positive control and shown as technical replicates.

In the presence of the compounds there is small indication of inhibition, however, the concentration of the probes used for the experiment was in the mM range. These results show that the given compounds either do not show binding or only weak binding to the target protein, therefore, could not compete with RNA.

In conclusion, there is a high affinity of RNA to the hnRNP A2/B1 and 4mer PNA didn't have a desired binding affinity as well as the synthesis of the longer PNAs was not straightforward. Another option would be trying to design new molecules which has phosphate groups on it to improve the binding affinity. For that, a good control for both PNA synthesis and nucleotide synthesis is needed.

As this approach was not successful, a screening-based approach was performed next.

3.3.9. Identification of cyclic peptide inhibitors of the hnRNP A2/B1

The assay discussed in this chapter has been developed and performed by Dr. Jessica Nowacki-Hansen. The SICLOPPS library has been prepared by Dr. Stefan Schmeing according to the reports by Benkovic and Tavassoli[185].

Cyclic peptide inhibitors (Table 12) of the hnRNP A2/B1-RNA interaction have been identified by a screening-based approach (detailed information can be found in "dissertation of Jessica Nowacki-Hansen"). For this purpose, a SICLOPPS[97] plasmid library encoding for of hexameric cyclic peptides was combined with the TRAP assay variant developed in our group[186]. The assay was encoded on two plasmids: one expresses a Green fluorescent protein (GFP) tagged hnRNP A2/B1 (AA1-251). The other plasmid expresses a BFP RNA construct that contained the 10mer RNA twice distanced by a GGG linker (AAGGACUAGCGGGAAGGACUAGC) in its 5'-UTR in front of the ribosomal binding site. Previously, an interaction between the protein and the RNA construct could be detected leading to repression of the reporter gene (low BFP levels). An inhibitor of the interaction, therefore, should lead to the restoration of the reporter gene (high BFP levels). Escherichia coli cells were used for the screening and cells with an inhibited phenotype (restored BFP production) were selected for using fluorescence activated cell sorting (FACS)[187]. SICLOPPS plasmids were isolated from the hit population and analyzed by next generation sequencing[188,189]. The screening procedure is schematically depicted in Figure 63.

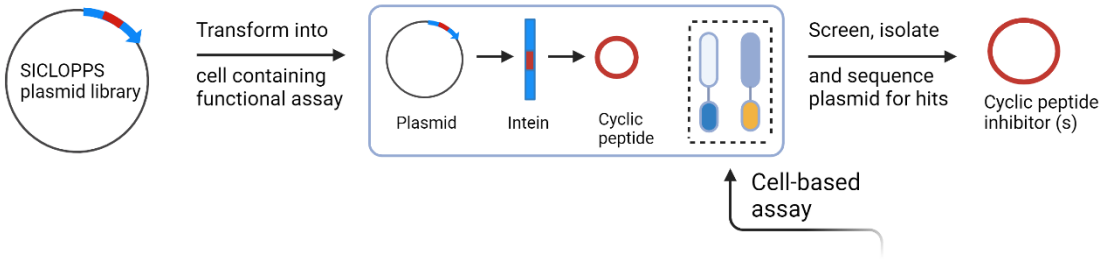


Figure 63: Representation of the SICLOPPS peptide library generation and screening. The SICLOPPS plasmids and TRAP assays plasmids can be transformed into E. coli cells including screening assay and screened by FACS. SICLOPPS plasmids with a desired phenotype can be isolated from cells and sequenced to identify the macrocyclic peptide. Adopted from Tavassoli, 2017[139].

3.3.10. Synthesis of the macrocyclic peptides of the hnRNP A2/B1

Synthesis protocol described in this chapter was optimized by Dr. Peter 't Hart.

Peptide synthesis was carried out using an SPPS protocol[129]. 2-CTC resin was used as a solid support for the synthesis. To start the synthesis the first amino acid of the appropriate peptide sequence was attached to the resin using Fmoc-amino acid and DIPEA in DCM (see section 5.9.3.). After successful loading of the first amino acid, removal of the Fmoc-protecting group was performed in 20% piperidine/DMF, which was then followed by coupling with a second amino acid resulting in a first peptide bond formation. Coupling was carried out with appropriate Fmoc-protected amino acid, PyBOP, and DIPEA in DMF. The Fmoc-deprotection and coupling

reactions were repeated until the desired peptide sequence had been achieved and the peptide chain was selectively cleaved from the solid support without removing the side chain protecting groups (see section 5.12.2.). The peptide went through a head-to-tail cyclization in solution in the presence of PyBOP, Oxyma pure, and DIPEA (see section 5.12.3.). After cyclization the success of the reaction was monitored by LC-MS and a final cleavage was performed in TFA/TIPS/H_2O/DODT to remove all the protecting groups (see section 5.12.4.). Once peptide was obtained it was purified using preparative HPLC. The synthesis route for the cyclic peptides is shown in Figure 64.

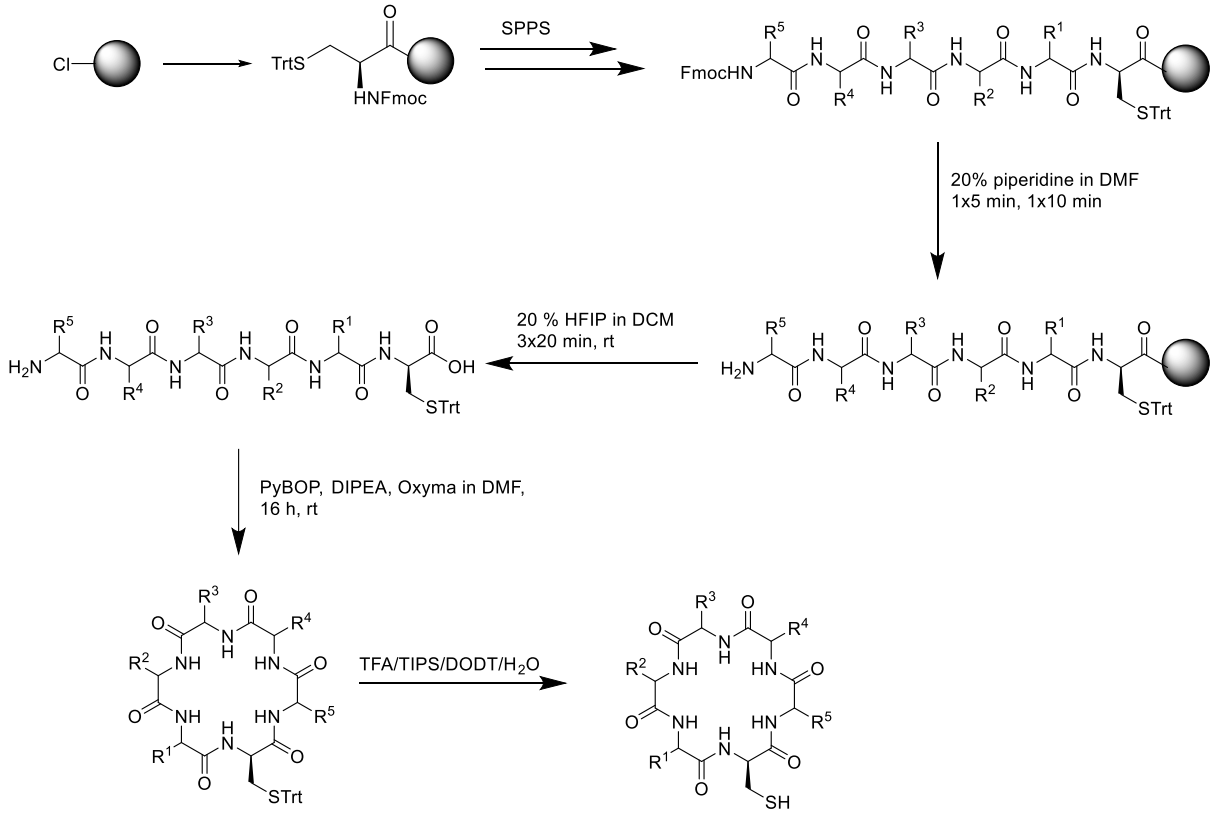


Figure 64: Synthesis scheme of the SICLOPPS peptides.

Certain peptide candidates were obtained as isomers as they were giving two peaks during LC-MS analysis both having the corresponding peptide mass which could be explained as a result of racemization during cyclization.

Table 12: Cyclic peptides derived from the SICLOPPS screening of hnRNP A2/B1.

Peptide	Peptide sequence	Name
1	X X X X X X	13
2	X X X X X X	14
3	X X X X X X	15
4	X X X X X X	16

5	X X X X X X	17
6	X X X X X X	18

The successfully synthesized peptides were then tested in the competition FP experiment to evaluate the inhibitory properties of the peptides against the hnRNP A2/B1-RNA interaction.

3.3.11. Direct fluorescence polarization assay with RNA

To develop a competition experiment for the peptides we first performed a direct FP experiments with RNA to define the optimal protein concentration for the competition FP assay[143]. Previous studies showed that hnRNP A2/B1 binds to RNA with an AAGG motif with low nanomolar affinity as measured by ITC[172]. To establish an assay that measures hnRNP A2/B1 binding to a particular nucleotide sequence, 3'Cy5 labeled RNA1 (AAGGACUAGC) and 5'FAM labeled RNA2 (AAGGACUAGCGGGAAGGACUAGC) oligonucleotides were employed. First, the direct FP assay was performed using 3'Cy5 labeled RNA1 and MBP-tagged hnRNP A2/B1 (residues 1-251). Labeled RNA (1 nM) was titrated with 2-fold dilutions series of protein in 384 well plate.

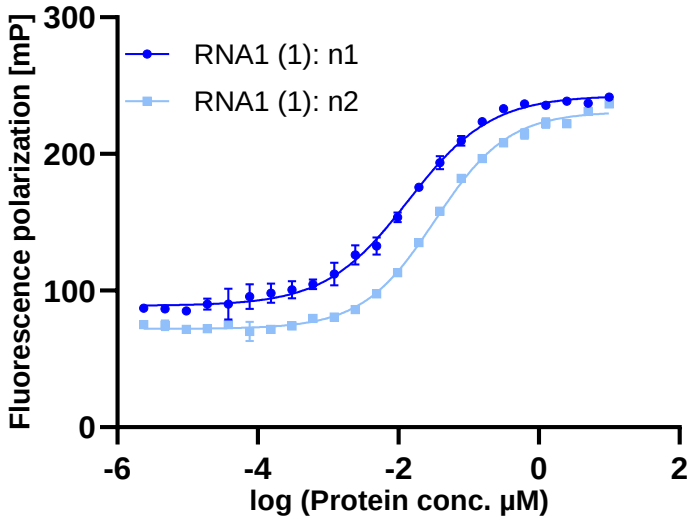


Figure 65: Fluorescence polarization curves of 3'Cy5 labeled RNA1 with hnRNP A2/B1 (1-251) measured after 30 minutes of incubation. n1 and n2 represents two biological replicates, respectively.

Binding between RNA and the protein was observed for both biological replicates (Figure 65). Direct FP experiment was carried out once again involving the 3'Cy5 labeled Poly C (CCCCCCCC) RNA as a negative control (Figure 66).

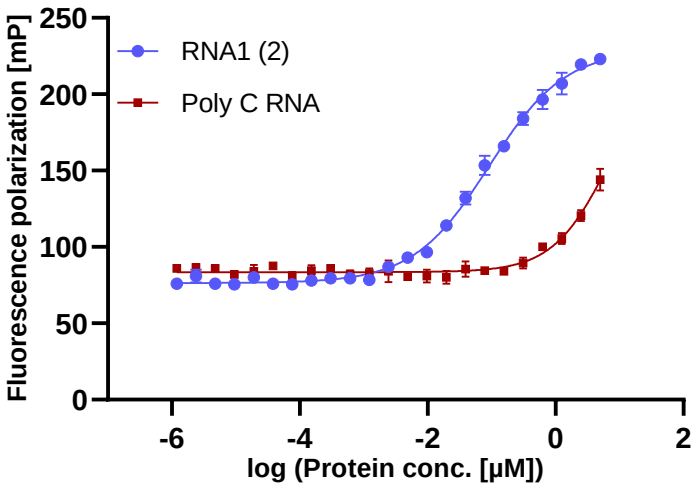


Figure 66: Fluorescence polarization curve of 3'Cy5 labeled RNA1 and Poly C RNA with hnRNP A2/B1 (residues 1-251). As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation and shown as technical replicates.

Table 13: Averaged K_D values for the 3'Cy5 labeled RNA1 tested in FP assay for binding affinity with hnRNP A2/B1.

Replicate	Labeled RNA	K_D (µM)
(1)	RNA1 3'Cy5	0.0236 ± 0.0114
(2)	RNA1 3'Cy5	0.093 ± 0.018
	Poly C 3'Cy5	>10

For the RNA1, an increase in fluorescence polarization with an increasing protein concentration was observed which indicated binding between protein and RNA. The Poly C RNA also showed a noticeable increase in fluorescence polarization at higher protein concentrations which could be explained with the weak interaction between cytosine residues and the protein. The direct FP experiment was carried out under the same conditions in the presence of 5'FAM labeled RNA2 (AAGGACUAGCGGGAAGGACUAGC) oligonucleotide. MBP-tagged hnRNP A2/B1 (residues 1-251) was used in the experiment.

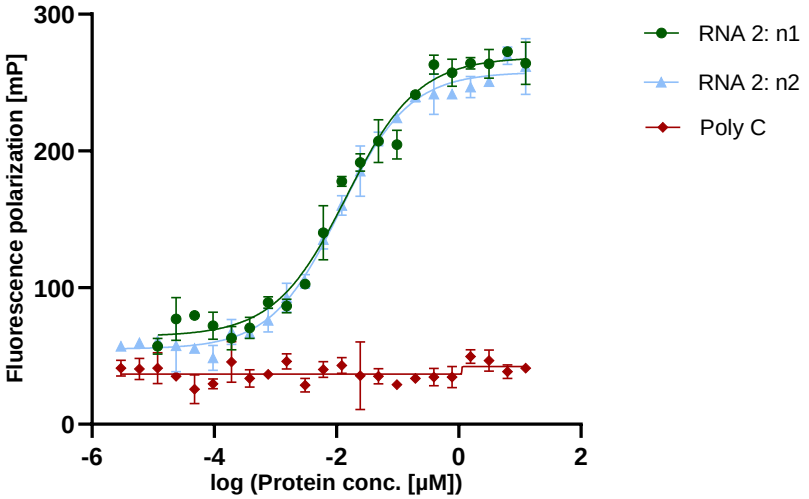


Figure 67: Fluorescence polarization curves of 5'FAM labeled RNA2 and Poly C RNAs with hnRNP A2/B1 (residues 1-251). Polarization was measured after a 30-minute incubation period. n1 and n2 represents two biological replicates, respectively. As a control, 5'FAM Poly C RNA was utilized.

For both biological replicates binding between RNA and the protein is observable (Figure 67). The control RNA (Poly C) does not show any significant change in fluorescence polarization which indicates that there is no binding to the protein.

Table 14: Averaged K_D values for the 5'FAM labeled RNA2 tested in FP assay for binding affinity with hnRNP A2/B1.

Labeled RNA	K_D (µM)
5'FAM RNA2	0.0131± 0.003
5'FAM Poly C	-

3.3.12. Evaluation of the hnRNP A2/B1 peptides in competition FP assay

The synthesized peptides for the hnRNP A2/B1 SICLOPPS screening (Table 15) were evaluated in the competition FP assay.

Table 15: Peptides from the hnRNP A2/B1 SICLOPPS screening.

Peptide	Sequence	Name	Note
1	X X X X X X	13	
2	X X X X X X	14	
3	X X X X X X	15	
4	X X X X X X	16	
5	X X X X X X	17-1	1st isomer
5_2	X X X X X X	17-2	2nd isomer
6	X X X X X X	18	

Competition FP experiment was carried out in the presence of the FAM labeled RNA2 and MBP-tagged hnRNP A2/B1 (residues 1-251) (Figure 68-70). Labeled RNA2 at a final concentration of 1 nM was used in the experiment. Unlabeled molecules (RNA1 and peptide) was applied as a 2-fold dilution series and titrated with protein/labeled RNA with the fixed protein concentration[143].

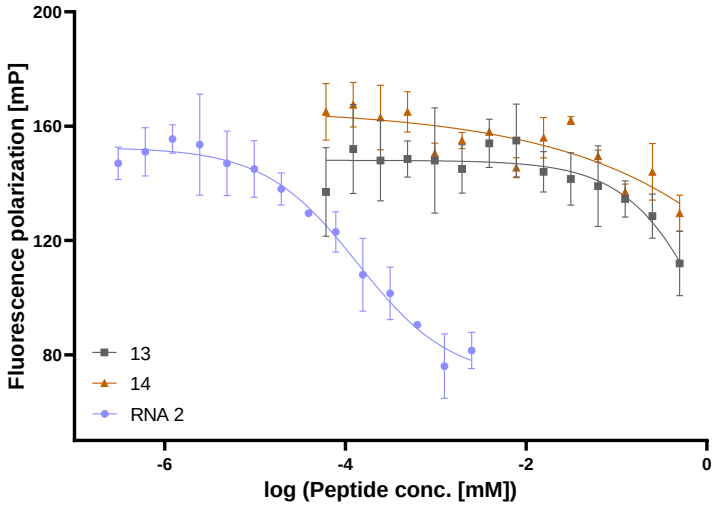


Figure 68: Curves of 13, 14 peptides and unlabeled RNA1 from the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was was used in this experiment. Polarization was measured after a 60-minute incubation and shown as a technical replicate (N=2). The measurements were performed at the same time on the same plate however the graphs are shown separately for easier interpretation.

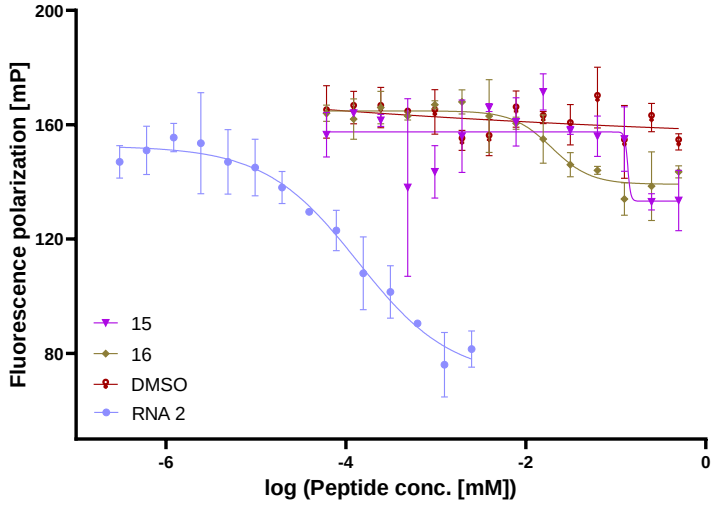


Figure 69: Curves of 15, 16 peptides and unlabeled RNA1 the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was used and DMSO was used as a negative control in this experiment. A4 was dissolved in DMSO and diluted with buffer to a final DMSO concentration of 1%. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2).

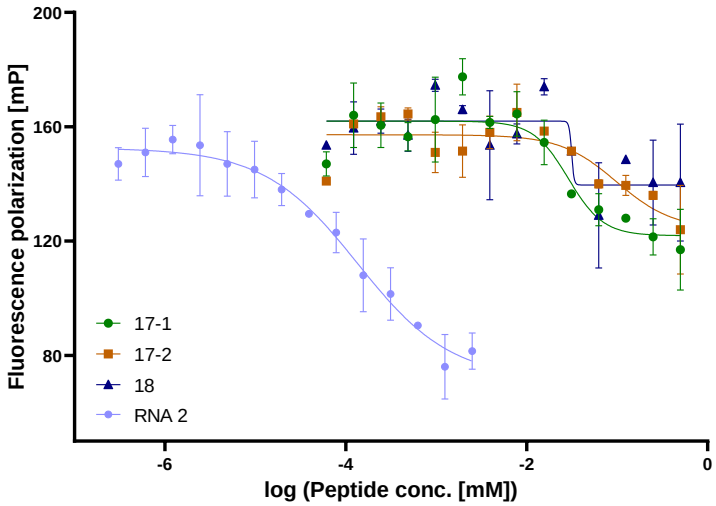


Figure 70: Curves of 17-1, 17-2, 18 peptides and unlabeled RNA1 from the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was used. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2).

Excitingly, we observed competition of the peptides with the FAM labeled RNA2. Peptides 13, 14, 15 and 18 showed only slightly decreased polarization and curves could not accurately be fitted. 16 and 17 seem to compete with the RNA with low micromolar IC_{50} values (Table 16). However, in comparison to 17-2, 17-1 gives ~5-fold lower IC_{50} indicating 17-1 is likely the correct isomer. Therefore, 16 and 17-1 were selected as the best candidates from the SICLOPPS screening. Based on the IC_{50} values K_I inhibition constants were calculated[190].

Table 16: IC_{50} values and calculated K_I values derived from competition FP of the peptides tested for their inhibitory properties for the interaction of RNA with hnRNP A2/B1.

Compound	IC_{50} (µM)	K_I (µM)
RNA2	0.15 ± 0.10	0.05 ± 0.02
13	>1000	-
14	>1000	-
15	-	-
16	22.91 ± 15.46	7.83 ± 5.29
17-1	46.92 ± 29.91	16.04 ± 10.23
17-2	261.09 ± 312.56	89.26 ± 106.86
18	>1000	-
DMSO control	-	-

In conclusion, two hit peptides have low µM IC_{50} and K_I values. Out of these two peptides, 16 had very poor solubility which is why it was not involved in further experiments. Therefore, 17-1 was

selected for further optimization to improve the inhibitory properties of the hit for hnRNP A2/B1 and RNA interaction.

3.3.13. Alanine scanning for the hit peptide selected from the SICLOPPS screening

Data represented in this chapter has been obtained together with Arina Khamzina.

Alanine scanning is a technique that can be used to identify the contribution of the side chain of each amino acid on the bioactivity of a peptide[191]. Alanine is selected for this purpose as it lacks a bulky side chain with a functional group while maintaining the main chain structure that many other amino acids possess. Glycine would also be an option for the exchange since it does not have a sidechain, however, the lack of chirality at the α-carbon could introduce flexibility and conformational changes making the interpretation more complicated[191]. Therefore, alanine scanning was selected to determine the function of sidechains with certain properties such as hydrophobicity, polarity, or charge that might play a function in the interaction with protein. As peptide candidate 17-1 was the most promising in terms of IC_{50} and solubility it was used for alanine scanning.

For the purpose of alanine scanning, each amino acid of 17-1 was exchanged with alanine one at a time, which gave six new derivatives with a different sequence represented in Table 17. In this way, amino acid binding properties to the target protein can be determined which would allow to perform appropriate modifications for further optimization. Each alanine-substituted peptide was separately assessed in the biophysical assays.

Table 17: Peptides derived from the alanine scanning of A5-1. Point mutations are marked in red.

Peptide	Sequence	Name
1	X X X X X X	17-1
2	A X X X X X	19
4	X A X X X X	20
5	X X A X X X	21
6	X X X A X X	22
7	X X X X A X	23
8	X X X X X A	24

The peptides were synthesized applying the same synthesis protocol (see section 5.12.) based on SPPS[141].

Peptide 19 and 20 both eluted as two signals during HPLC purification with the same corresponding mass indicating that racemization occurred during cyclization. To be able to test both isomers in the FP experiment they were isolated separately. While both isomers were obtained peptide 19, only a single one of the isomers could be isolated (20-1) due to a very low yield of the other and therefore it was not included in the FP experiment. However, since it is unknown which isomer is missing for 20 the data is needs to be interpreted with care.

3.3.14. Evaluation of the hit peptides in competition FP assay

To characterize the competitive properties of each mutated peptide competition FP assays were performed using FAM labeled RNA2 and MBP-tagged hnRNP A2/B1 (residues 1-251). Unlabeled molecules (RNA1 or peptide) were applied as a 2-fold dilution series and were titrated with protein[143]/labeled RNA (Figure 71-73).

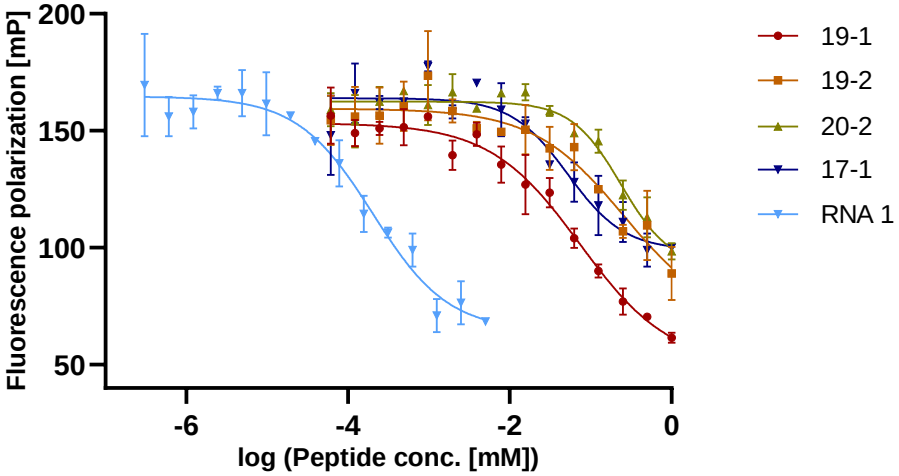


Figure 71: Competition FP curves of peptides 17-1, 19-1, 19-2, 20-2 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 60-minute incubation period, polarization was measured and shown as a technical replicate (N=2). The experiment was carried out at once in the same plate but are shown separately for better interpretation.

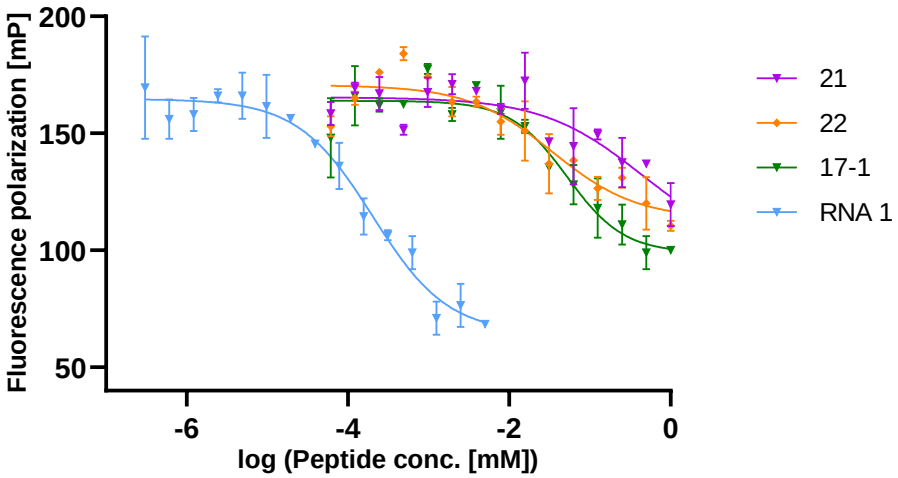


Figure 72: Competition FP curves of peptides 17-1, 21, 22 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 60-minute incubation period, polarization was measured and shown as a technical replicate (N=2).

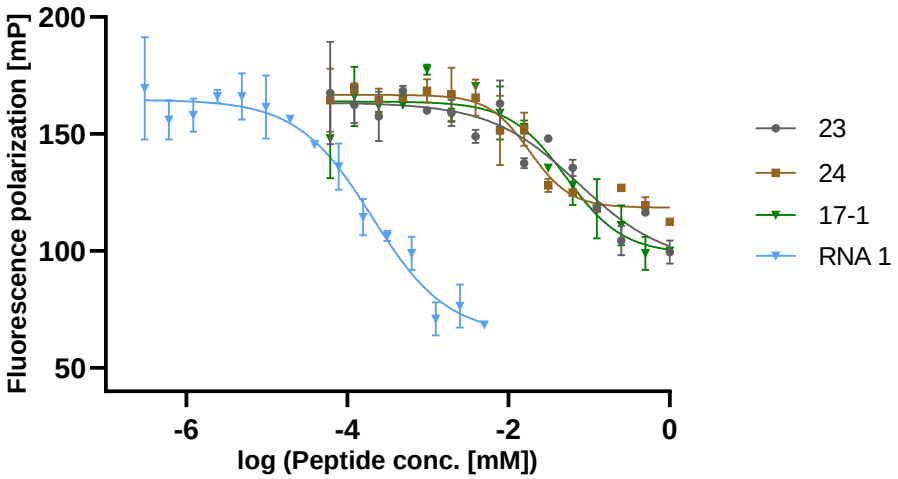


Figure 73: Competition FP curves of peptides 17-1, 23, 24 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 30-minute incubation period, polarization was measured and shown as a technical replicate (N=2).

Competition between FAM labeled RNA2 and all the tracers was observable. Competition curves could be fitted and an IC_{50} values were determined for 19-1, 20-2, 22, 23 and 24 peptides. For 19-2 and 21 bottom saturation could not be reached and therefore IC_{50} values could not be determined. Mutations at several positions lead to the loss of inhibition properties. Surprisingly, in comparison to 17-1, 24 shows a >2-fold improved IC_{50} value. Based on the IC_{50} values derived from the nonlinear regression fit the inhibition constants (K_I) were calculated (Table 18).

Table 18: IC_{50} and K_I values calculated from alanine scanning of peptides evaluated against hnRNP A2/B1-FAM labeled 10merggg10mer RNA interaction.

Compound	IC_{50} (µM)	K_I (µM)
RNA2	0.20 ± 0.04	0.065 ± 0.01
17-1	52.9 ± 11.2	18.1 ± 3.8
19-1	74 ± 2.1	25.2 ± 7.3
19-2	>1000	>500
20-2	324.5 ± 197.1	110.8 ± 67.3
21	>1000	>1000
22	48.2 ± 51.6	16.4 ± 17.6
23	156.9 ± 130.9	53.6 ± 44.7
24	19 ± 3.3	6.5 ± 1.2

The competition FP experiment was repeated using Cy5 labeled RNA1 as a tracer to verify the results as previous experiment in our group demonstrated that using the FAM label can lead to false positive results (Figure 74-76).

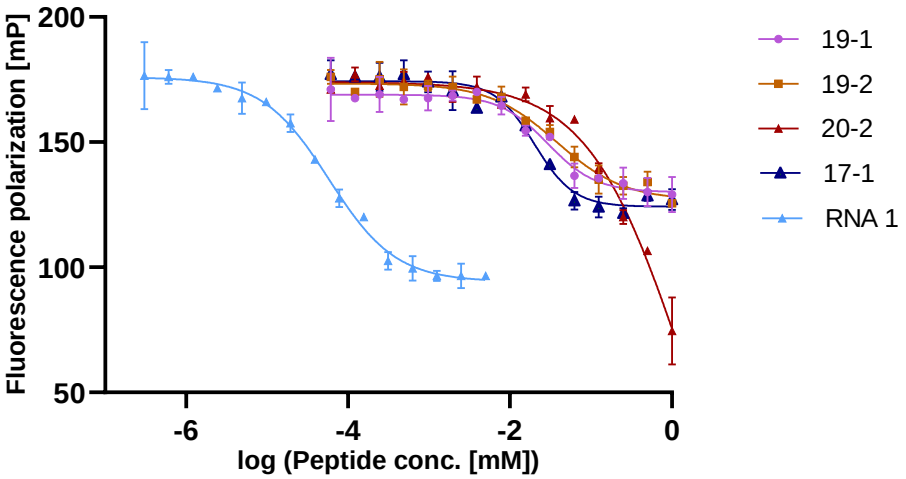


Figure 74*: Competition FP curves of peptides 17-1, 19-1, 19-2, 20-2 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2). The measurements were performed at the same time on the same plate however the graphs are shown separately for easier interpretation.*

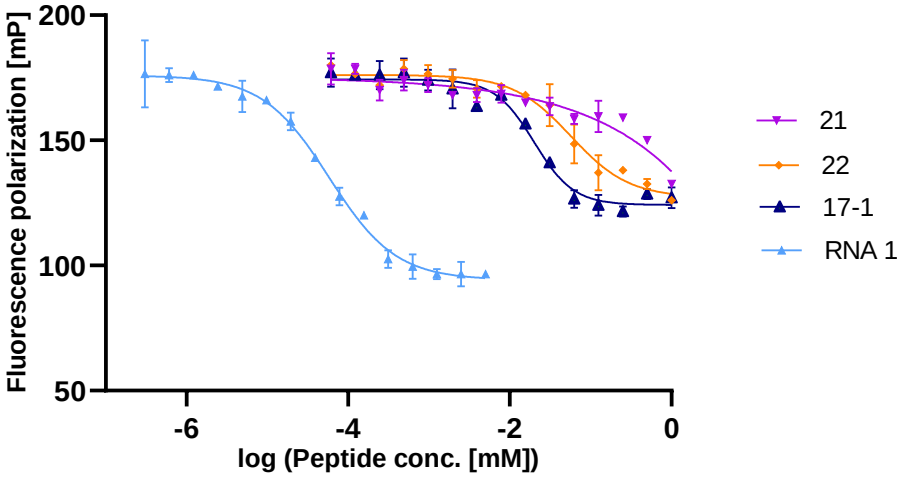


Figure 75: Competition FP curves of peptides 17-1, 21, 22 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2).

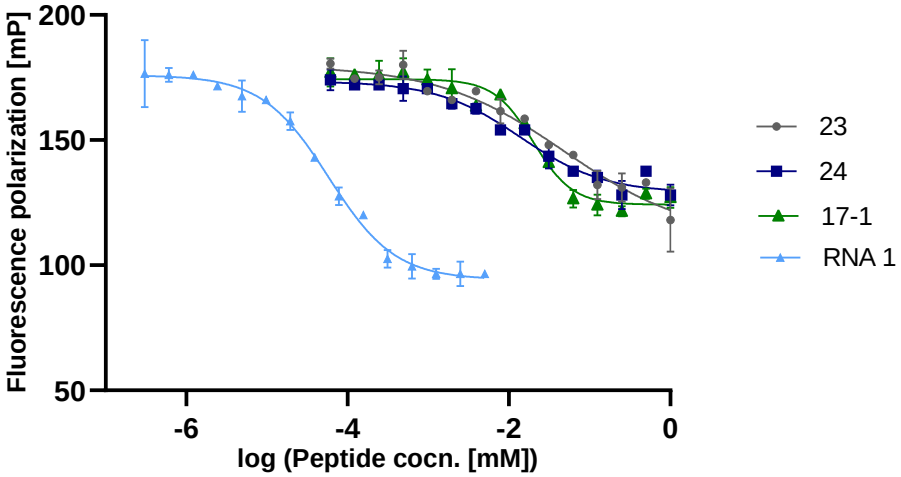


Figure 76: Competition FP curves of peptides 17-1, 23, 24 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2).

Table 19: IC_{50} and K_I values calculated for the peptides from alanine scanning evaluated against hnRNP A2/B1-Cy5 labeled 10mer RNA interaction.

Compound	IC_{50} (µM)	K_I (µM)
RNA1	0.058 ± 0.002	0.02 ± 0.001
17-1	20. 07 ± 1.01	8.14 ± 0.41
19-1	29.6 ± 7.42	12.01 ± 3.01
19-2	43.2 ± 16.79	17.53 ± 6.82
20-2	>1000	>1000
21	>1000	>1000
22	60.73 ± 31.01	24.65 ± 12.59
23	24.09 ± 5.76	9.77 ± 2.34
24	13. 49 ± 1.75	5.47 ± 0.71

Similar as for the experiment using the FAM labeled RNA, competition between the Cy5 labelled RNA and the peptides was observed. For the peptides 19-1, 19-2, 22, 23 and 24 a competition curve can be fitted and an IC_{50} values fall within a lower µM range. For peptides 20-2 and 21 lack of the lower plateau does not allow the determination of the accurate IC_{50} values. Based on the IC_{50} values derived from the nonlinear regression fit the inhibition constants (K_I) can be calculated. For 24 ~1.5-fold improved IC_{50} and K_I values were observed compared to 17-1, which is similar to the results obtained usinf FAM-labelled RNA2. In summary, except for 20-2 and 21 there was no significant change in IC_{50} and K_I values which indicates that the amino acids in these two positions are very important for binding to the target protein. It can be noticed that the peptide curves from FP are not giving complete inhibition, which is most likely related to RNA that interacts with both domains of hnRNP A2/B1[172]. It is probable that the peptide candidate is binding to

one domain of the hnRNP A2/B1 and therefore, the other domain of the protein is still bound to the RNA.

3.3.15. Microscale thermophoresis with selected hits from SICLOPPS and alanine scanning

Some of the MST measurements were performed by Dr. Jessica Nowacki-Hansen

Microscale thermophoresis (MST) is a biophysical assay technique used to investigate the interaction between molecules[192]. This assay can be used to quantify protein-protein and protein-small molecule interactions. The method allows the detection of the movement of fluorescently labelled molecules over a temperature gradient. For an MST measurement, capillaries are loaded with a sample and the temperature gradient is produced by an infrared laser. The MST instrument than monitors emitted fluorescence over time. For MST, the protein is usually covalently labeled with a fluorescent dye, but the ligand could be labelled as well. The fluorescent molecule is mixed with the ligand at different concentrations, and after a certain incubation time the mixtures of molecules are loaded into separate capillaries. When the temperature gradient is applied to the capillaries with samples the movement of the fluorescent molecule is observed and recorded. The different concentrations of the ligand effect the movement of the fluorescent molecule at different rates and this feature can be used to evaluate the interaction between molecules. When the fluorescent molecule interacts with the ligand, there will be a change in some properties of the molecules, such as size, charge or hydration shell and this will influence the movement of the molecules.

Peptides 17-1 selected from SICLOPPS screening and 24 selected from alanine scanning were used in the MST experiment to investigate the binding properties to hnRNP A2/B1 (Figure 77-78). RED fluorescent dye NT-647-NHS was chosen to label the protein following the protocol from the manufacturer. The dye contains an NHS-ester, which reacts with primary amines that are found on lysine residues of proteins to form dye-protein-conjugates.

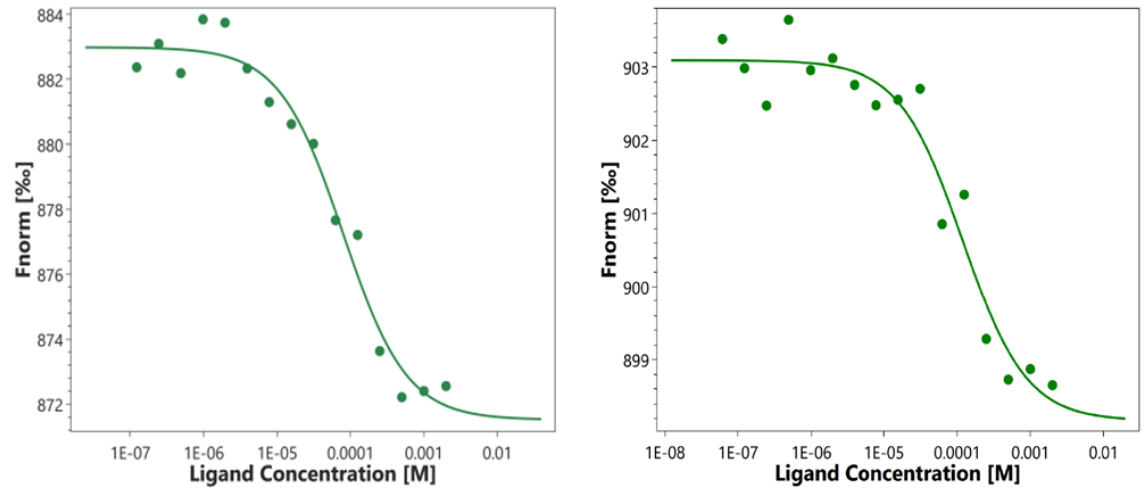


Figure 77: MST curve indicating the interaction between 17-1 and hnRNP A2/B1 (residues 1-251) protein. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates.

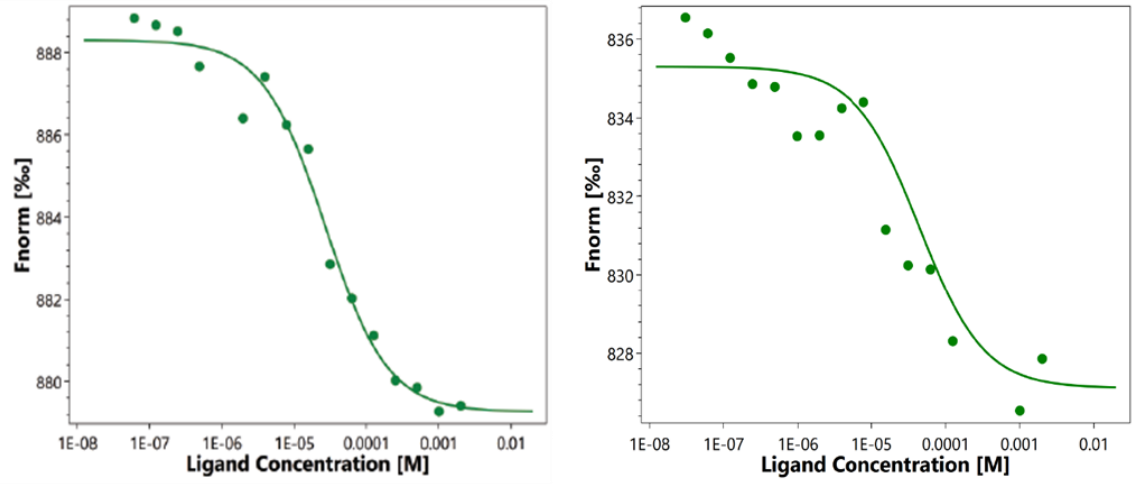


Figure 78: MST curve indicating the interaction between 24 and hnRNP A2/B1 (residues 1-251) protein. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates.

Table 20: Dissociation constant (K_D) of the peptides with hnRNP A2B1 protein from the MST experiment.

Peptide	**K_D [µM]**
17-1	99 ± 28.28
24	35.4 ± 12.45

For both peptides the K_D values were calculated by NanoTemper MST Monolith Nt.115 software. Based on this analysis 24 had an almost >2.5-fold improved K_D when compared with the A17-1 peptide for hnRNP A2/B1.

3.3.16. Investigation of the interaction of hnRNP A2/B1 peptide with RRM1 and RRM2 domains

MST measurements were performed by Dr. Jessica Nowacki-Hansen

Next, to investigate whether the hit peptide (24) is binding to the RRM1 or the RRM2 domain of hnRNP A2/B1 the binding affinity for the single domains was evaluated by MST (Figure 79-80).

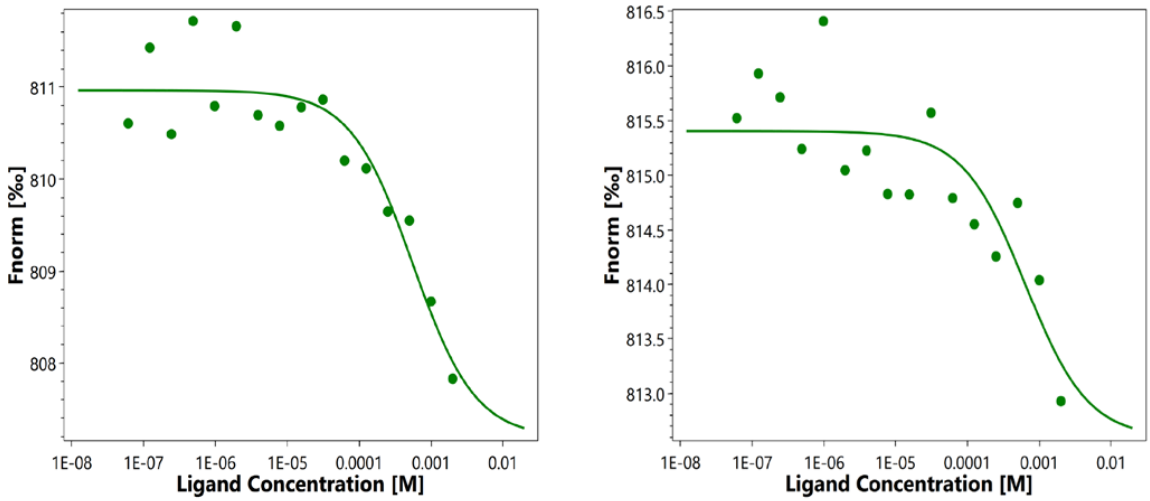


Figure 79: MST curve of the peptide 24 and hnRNP A2/B1 RRM1(residues 1-104) domain. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates.

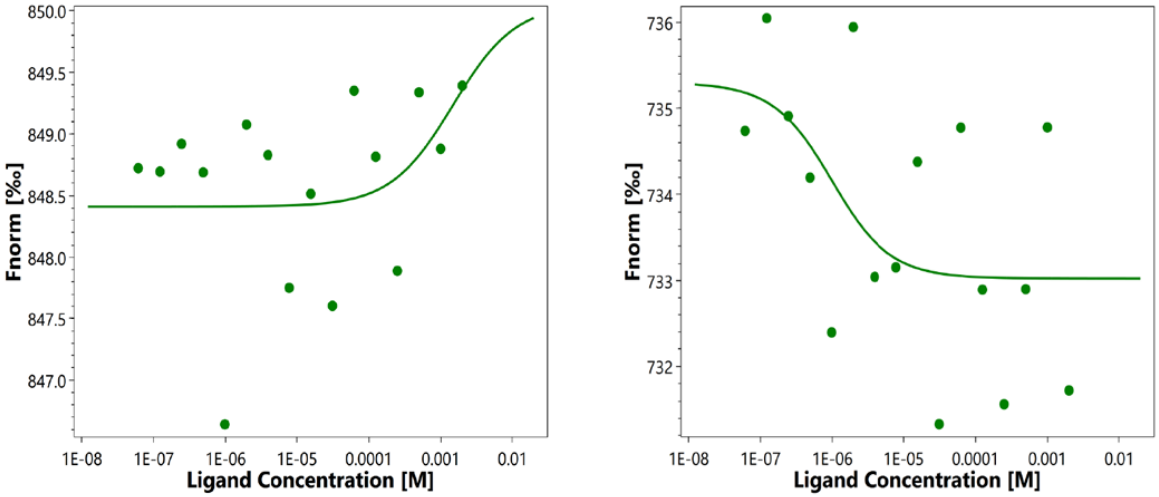


Figure 80: MST curve of peptide 24 and hnRNP A2/B1 RRM2 (residues 112-195) domain. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates.

Table 21: Dissociation constant (K_D) of the peptide 24 with individual hnRNP A2B1 RRM domains as measured by the MST.

Protein	Peptide	K_D [µM]
hnRNP A2/B1 RRM1	24	598.5 ± 57.28
hnRNP A2/B1 RRM2	24	-

Binding affinity between 24 and hnRNP A2/B1 RRM1 domain was monitored from the MST curve and the K_D value calculated using the NanoTemper MST Monolith Nt.115 software. Binding to RRM1 was observed with a significantly higher K_D value than for the construct containing both RRMs. Furthermore, for RRM2 no binding was observed which suggests that 24 binds to RRM1 but RRM2 is required for high affinity binding[193,194].

3.3.17. Investigation of hnRNP A2/B1 and RNA interaction with mass photometry

The crystal structure of hnRNP A2/B1 shows that both RRM domains of separate hnRNP A2/B1 recognized by RNA (Figure 81)[172]. To investigate this, we used mass photometry (MP) which is able to measure the accurate mass of molecules in solution in a label-free manner[195]. These features enable the application of mass photometry to the quantification of protein-protein interactions by providing the mass of bound and unbound molecules in solution.

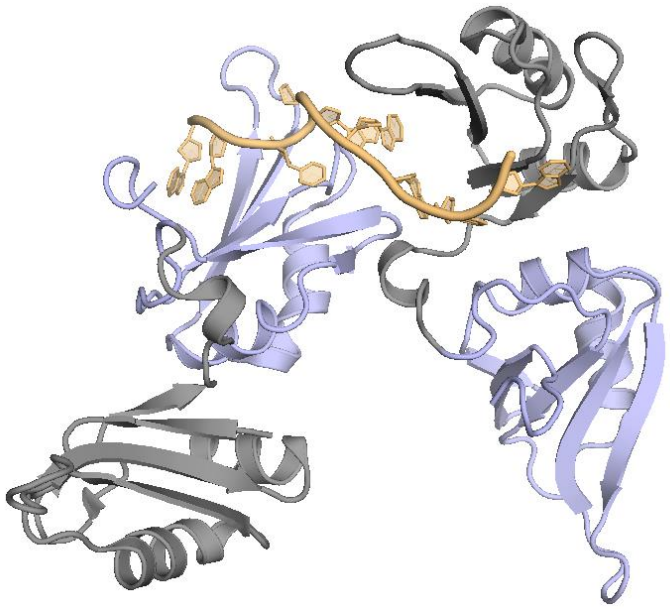


Figure 81: Representation of the RRM domains of the hnRNP A2/B1 in complex with RNA1 (AAGGACUAGC). RNA1 accommodated into a (+) charged groove of the RRM domains of two hnRNP A2/B1. Adopted from Wu et al[172].

MP detects single biomolecules or their complexes by their light scattering as they bind nonspecifically to a surface, a glass coverslip (Figure 82, A)[196]. Each binding event leads to a change at the glass/water interface, which effectively alters the local reflectivity and can be detected by taking advantage of optimized interference between scattered and reflected light (Figure 82, B). Detection also allows to measure the concentration of a sample by counting individual molecules. To explore this capability monomers that form dimer and produce mass distributions with the expected major bands can be chosen (Figure 83, C).

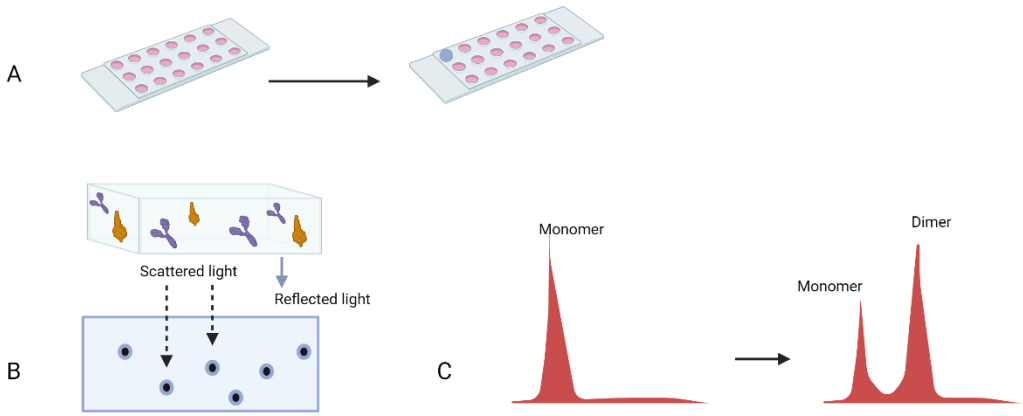


Figure 82: A) glass cover slip with (right, blue) and without sample (left), B) Label-free detection by imaging the scattered and reflected light over time. b) Scatter plot of single-molecule contrast and mass distribution for a mixture of monomer/dimer. Adopted from Soltermann et al[196]*.*

Mass photometry was applied to investigate the interaction between RNA and hnRNP A2/B1. If in case the RNA binds to two hnRNP A2/B1 molecules there would be ~2-fold mass difference between samples with and without RNA [197,198]. Different RNA concentrations were applied and (Figure 84, Supplementary figure 11-12) MBP tagged hnRNP A2/B1 construct (residues 1-251, MW: 70.306 kDa) was used.

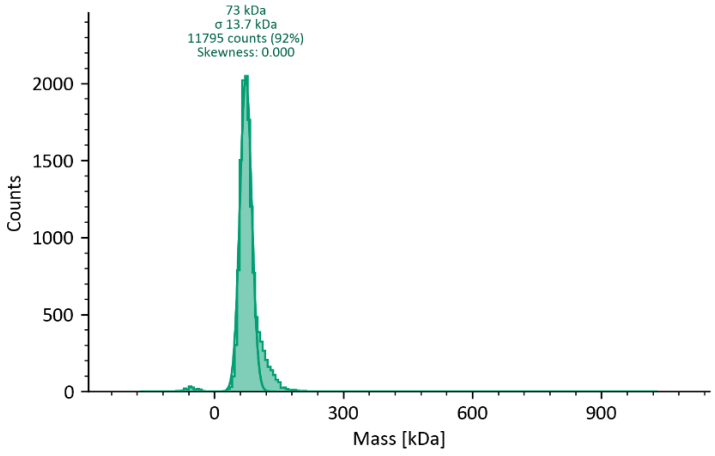


Figure 83: Mass photometry data observed from the measurement for hnRNP A2/B1 protein alone.

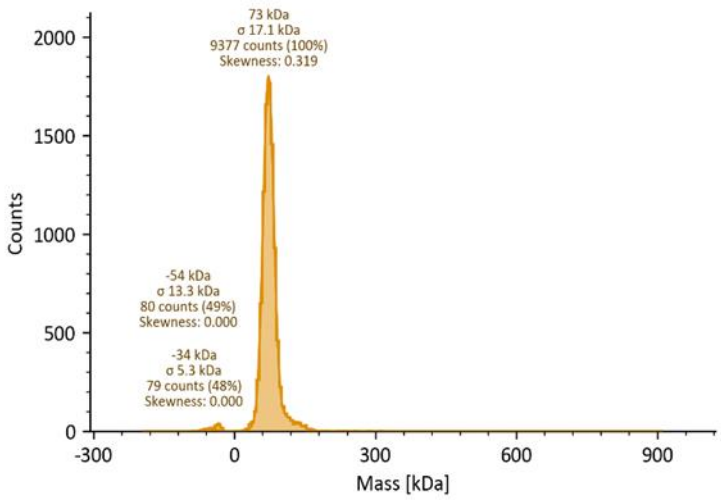


Figure 84: Mass photometry data observed from the measurement for hnRNP A2/B1 protein with Cy5 labeled 10mer RNA. The protein concentration was 50 nM protein and the RNA concentration 25 nM. The measurement was performed after 30 min incubation at rt.

Under all conditions (Figure 84, Supplementary figure 11-12) only the mass which corresponding to the corresponding protein construct (1-251, MW: 70.306 kDa) was obtained. Since RNA-binding was confirmed by various other biophysical techniques this indicated that mass photometry is not suitable to clarify if the RNA is interacting with two hnRNP A2/B1 through different RBDs. Other experiments are therefore still needed to clarify how the interaction between hnRNP A2/B1 and RNA take place in solution which would allow further investigation of the identified inhibitors.

3.4. Summary and outlook

Peptide inhibitors for hnRNP A2/B1 were selected from a screening-based approach followed by synthesis using SPPS and tested in a fluorescence polarization assay. As a result, two low micromolar hit peptides were found that inhibit the interaction of hnRNP A2/B1 with the corresponding RNA. Since one of the peptides (16) had a poor solubility and therefore could not be involved in the further experiments, peptide A17-1 was selected for further optimization studies. To verify the interaction between peptide and the target protein, another biophysical assay, microscale thermophoresis experiment was performed. MST experiment confirmed the protein-peptide binding with a mid-micromolar K_D value with the selected peptide. Next, an alanine scan was applied to identify the amino acids important for the binding. Interestingly, alanine scanning revealed a derivate (24) with a ~1.5-fold increase in affinity, indicating further structure-activity relationship studies to be promising. Next, the MST experiment with the single domains (RRM1 and RRM2) of hnRNP A2/B1 revealed that 24 mainly bind to the RRM1 domain but it needs both domains or even the whole hnRNP A2/B1 protein (1-251) for proper binding. Solving the crystal structure of the protein together with the hit candidate would give a better insight into the protein-peptide interaction and would allow further optimization of the hit peptide. Apart from that, optimization could be done by substitution of amino acids that doesn't show binding with other amino acids. In addition to that, optimization could be done on the important amino acids by substituting their side chains with close analogs.

Once a peptide candidate with desired activity has been identified it can be involved in cell-based experiments to investigate its splicing inhibitory properties. For that purpose, the permeability properties of the cyclic peptides needed to be investigated as well.

4. Materials, reagents, chemicals, devices

4.1. Materials

Material/Chemical	Supplier
Amicon® Ultra	Millipore®
Coomassie Protein-Assay-Kit	Thermo Fisher Scientific
Coomassie stain	Serva
Eppendorf tube	Sarsted AG & Co. KG
Falcons	Sarsted AG & Co. KG
HPLC vials	VWR International GmbH
Microplate, 96-well	Greiner Bio-One
Microplate, 384 well, 4514	Corning
Multichannel Pipettes	Thermo Fisher Scientific
Needles	Braun
Pipettes	Eppendorf
Protein Labeling Kit RED - NHS	NanoTemper Technologies
Sereological pipettes	Sarsted AG & Co. KG
Syringes	Braun
Syringe filters	Fischer Scientific
Syringe reactors	Multisyntech GmbH
SYPRO Orange protein stain (5000x)	Thermo Fischer Scientific

4.2. Devices

Device	Supplier
Äkta Explorer	GE Healthcare
Centrifuges	Eppendorf
Cell disruptor sonicator	Fischerbrand

ChemiDoc MP Imaging System	Bio-Rad Laboratories
DNA/RNA UV cleaner box	LTF Labortechnik
FinnpipetteTM F1 Multichannel Pipette	Thermo Fisher Scientific
HiLoadTM 16/600 SuperdexTM 75	Cytiva
HiLoadTM 26/60 SuperdexTM 75	Cytiva
Incubator	Binder Drying and heating chamber ED 23
Lyophilizer	Christ
LC-MS	Agilent Technologies
Monolith Nt.115	NanoTemper Technologies
NanoDrop 2000c Spectrophotometer	Thermo Fisher Scientific
Plate reader Tecan Sparks	Tecan trading AG
Peptide synthesizer	Multisyntech GmbH, Syro I peptide synthesizer
	Gyros Protein Technologies, PurePep® Chorus
Peristaltic pump	Landgraf Laborsysteme HLL GmbH
Rotor evaporator	Büchi
RP-HPLC	Agilent Technologies
Incubator shaker	Eppendorf SE
Thermomixer	Eppendorf Thermomixer
Thermal cycler	Bio-Rad
Shaker	IKA Labortechnik
Sonicator	Bandelin
Vacuum Manifold	Promega

4.3. Software

Software	Supplier
ChemDraw	PerkinElmer
GraphPad Prism 9	GraphPad Software, Inc.
SnapGene	GSL Biotech LLC
Xcalibur	Thermo Fisher Scientific

4.4. Reagents

Reagent	Supplier
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (HEPES)	Thermo Fisher Scientific
Disodium hydrogen phosphate (Na_2HPO_4)	Thermo Fisher Scientific
Glycerol	Thermo Fisher Scientific
Hydrochloric acid (HCl)	Thermo Fisher Scientific
Imidazole	Carl Roth GmbH
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	Carl Roth GmbH
Isopropanol	Carl Roth GmbH
PageRulerTM prestained protein ladder	Thermo Fisher Scientific
Phenylmethylsulfonyl fluoride (PMSF)	Applichem
Sodium chloride (NaCl)	Carl Roth GmbH
Sodium hydroxide (NaOH)	VWR International GmbH
Sodium fluoride (NaF)	Sigma-Aldrich
Sodium Dodecyl Sulfate (SDS)	Sigma-Aldrich
Sodium dihydrogen phosphate (NaH_2PO_4)	VWR International GmbH
tris(2-carboxyethyl) phosphine (TCEP)	TCI
Triton X-100	Serva
Tween-20	Fischer BioReagents

4.5. Medium and buffers

Medium	Components
LB medium	0.5% Yeast extract 1% Tryptone 171 mM NaCl pH 7.4
Super Optimal Broth with catabolite repression (SOC)	0.5% Yeast extract 2% Trytone 2.5 mM KCl 10 mM NaCl 10 mM MgSO4 10 mM MgCl2 20 mM Glucose 200 mM Tris (pH 6.8)
5x SDS sample buffer	40% Glycerol 8% SDS 10 mM DTE 0.007 mM Bromophenol blue
SDS running buffer	200 mM Glycine 25 mM Tris 1% SDS
Coomassie staining solution	

Destaining solution	0.1% Coomassie R250 10% Acetic acid 40% MeOH 10% acetic acid in H2O

4.6. Protein buffers

RBM20 buffers

Buffers	Components
Lysis buffer	50mM Tris 500mM NaCl 5% glycerol 1 mM PMSF pH 7.5
Wash buffer	50mM Tris 500mM NaCl 5mM Imidazole 5% glycerol pH 7.5
Elution buffer	50mM Tris 500mM NaCl 500mM Imidazole 5% glycerol pH 7.5
Storage buffer	50mM Tris 500mM NaCl 5% glycerol

	pH 7.5

hnRNP A2/B1 buffer

Buffers	Components
Lysis buffer	50 mM HEPES 300 mM NaCl 1 mM PMSF pH 8
Washing buffer	300 mM NaCl 50 mM HEPES 1 mM TCEP 20 mM Imidazol pH 8
Eluting buffer	50 mM HEPES 300 mM NaCl 1 mM TCEP 500 mM Imidazol pH 8
Storage buffer	100 mM NaCl 25 mM HEPES 1 mM TCEP

	pH 8

4.7. Assay buffers

Buffers	Components
hnRNP A2/B1 buffer	100 mM NaCl 25 mM HEPES 1 mM TCEP 0.01 Triton pH 8
SRSF1 buffer	20 mM Na_2HPO_4 150 mM KCl 50 mM L-Arg 50 mM L-Glu 1.5 mM $MgCl_2$ 0.2 mM EDTA 1 mM TCEP 0.01% Triton pH 8
RBM20 buffer	20 mM Sodium Phosphate (NaH_2PO_4/Na_2HPO_4) 50 mM NaCl 2.5 mM TCEP 0.01% Tween-20 pH 7.5

4.8. Antibiotics and induction agents

Antibiotic/Induction agent	Supplier
Ampicillin	Roth
IPTG	Biochemica

4.9. Amino acids

Amino acid	Supplier
Fmoc-Ala-OH	Merck
Fmoc-Ile-OH	Carbolution
Fmoc-Leu-OH	Chem-Impex
Fmoc-Asp (OtBu)-OH	Novabiochem
Fmoc-His (Trt)-OH	Roth
Fmoc-Met-OH	Carbolution
Fmoc-Ser (tBu)-OH	Carbolution
Fmoc-Cys(Trt)-OH	Carbosynth
Fmoc-Val-OH	Carbolution
Fmoc-Arg (Pbf)-OH	Carbosynth
Fmoc-Pro-OH	Novabiochem
Fmoc-Glu(OtBu)-OH	Novabiochem
Fmoc-Trp(Boc)-OH	Novabiochem
Fmoc-Phe-OH	Carbosynth
Fmoc-Gln(Trt)-OH	Carbolution
Fmoc-Thr(tBu)-OH	Novabiochem
Fmoc-Tyr (tBu)-OH	Novabiochem
Fmoc-Lys (Mtt)-OH	Novabiochem
Fmoc-(S)-2(4-pentenyl) Ala-OH	Fluorochem

4.10. PNA monomers

PNA monomer	Supplier
Fmoc-PNA-A(Bhoc)-OH	ASM Research Chemicals
Fmoc-PNA-G(Bhoc)-OH	ASM Research Chemicals
Fmoc-PNA-C(Bhoc)-OH	ASM Research Chemicals

4.11. SPPS reagents/solvents

Reagents/solvents	Supplier
1-Cyano-2-ethoxy-2-oxoethylidenaminooxy (COMU)	Biosynth
2,2′- (Ethylendioxy)diethanthiol (DODT)	Sigma-Aldrich
Acetonitrile	Roth
Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBop)	Carbosynth
Dimethylformamide (DMF)	VWR Chemicals
Dichloromethane (DCM)	Thermo Fisher Scientific
Dichloroethane (DCE)	Sigma Aldrich
Grubbs 1st generation	TCI
Grubbs 2nd generation	TCI
Hexafluorisopropanol	Thermo Fisher Scientific
N,N-Diisopropylethylamine (DIPEA)	Roth
Oxyma pure	Carbosynth
Piperidine	Acros
Pyrrolidine	TCI
Trifluoroacetic acid (TFA)	Roth
Triisopropyl silane (TIPS)	Sigma Aldrich

4.12. RNAs

All RNAs were obtained from Sigma-Aldrich.

5. Methods

5.1. Preparation of bacterial cultures

Medium with a desired volume was prepared in an Erlenmeyer-flask (~4x volume of medium). The antibiotic was added in a working concentration, followed by cells that were picked from the agar plate and were utilized to inoculate the medium. The culture was incubated at 37 °C for overnight. The next day, the saturated culture was used for protein expression.

5.2. Preparation of glycerol stock

A culture of a single colony was grown in LB medium with appropriate antibiotics at 37 °C (160 rpm) overnight. From that culture, 500 µL was mixed with 50% glycerol solution in a 1:1 ratio. The glycerol stock was frozen in liquid N_2 and stored at -80 °C until further use.

5.3. Bacterial Transformation

Chemically competent cells were placed on ice for approximately 10 minutes. Agar plates from storage at 4 °C were warmed up to room temperature. 1–2 µL of plasmid was mixed with cells and incubated on ice for 15 minutes. The cells were exposed to the heat shock in a heat block for 45 seconds at 42 °C. Meanwhile, the Super Optimal Broth with catabolite repression (SOC medium, see section 4.5.) was warmed up to 37 °C in the thermomixer. The cell suspension was placed back into ice for 2 min, and 300 µL of SOC medium was added. The culture was incubated in shaking incubator (300 rpm) for 1 h at 37 °C. Then, cell suspension was added to a plate containing the corresponding antibiotic under sterile conditions and grown overnight at 37 °C.

5.4. RBM20 protein expression and purification

Protein expression and purification was carried out either by Dr. Stefan Schmeing or work shared referring to the report from Upadhyay *et al*[122].

5.4.1. Expression of RBM20 delta_a3

pMAL-RBM20-RRM-delta_a3 was transformed into *E. coli* BL21DE3 (RIL) cells and plated on an agar plate with Ampicillin (Amp). The next day, a single colony was utilized to inoculate the preculture which was grown overnight at 37 °C. For the expression, 20 mL of the preculture was added to the 2 L LB medium. The culture was grown at 37 °C until an OD_{600} of ~0.7 was reached. The culture was cooled at 4 °C for 30 min and was then induced with 300 µM IPTG. Expression was carried out overnight at 20 °C and 180 rpm. The next day, cells were pelleted by centrifugation for 10 min at 5000xg at 4 °C.

5.4.2. Purification of RBM20 delta_a3

Cell pellets were resuspended in lysis buffer (for buffer components, see section 4.6.), and 1 mM PMSF was added. This was followed by lysis of the cells for 10 min using a cell disruptor sonicator with a 70% amplitude, 10 sec on, 10 sec off, and a pulse temperature of 0 °C on ice. The

supernatant was collected and filtered before it was loaded on a 5 mL Ni-NTA HisTrap (GE) column, which was equilibrated with wash buffer (see section 4.6.) beforehand. The protein was eluted with elution buffer (see section 4.6.), and fractions were collected. Fractions containing protein were combined and further purified by size exclusion chromatography (SEC) using HiLoad 26/60 superdex75 column eluting with protein storage buffer (see section 4.6.).

5.4.3. Expression of RBM20_a3

pMAL-RBM20-RRM-a3 was transformed into *E. coli* BL21 DE3 (RIL) cells and plated on an agar plate with Amp. The same protocol (see section 5.4.1.) was followed to grow the expression culture. Once an OD_{600} of ~0.55 was reached, the culture was cooled at 4 °C for 45 min and was induced with 300 µM IPTG. Expression was carried out overnight at 20°C, 160 rpm. The next day, the culture was centrifuged for 10 min at 5000xg, 4 °C, and the pellet was harvested.

5.4.4. Purification of RBM20_a3

The purification was done following the same protocol from section 5.4.2.

5.5. hnRNP A2/B1 protein expression and purification

MBP-tagged hnRNP A2/B1 (residues 1–251) was expressed and purified applying the protocol from the protein chemistry facility (PCF) in the MPI Dortmund as described in the following sections.

5.5.1. MBP-tagged hnRNP A2/B1 expression

E. coli BL21 (DE3) cells were used to express MBP-tagged hnRNP A2/B1 (residues 1–251). A previously prepared *E. coli* glycerol stock with the corresponding plasmid was used for protein expression. To prepare a preculture, a small amount of the glycerol stock was added to the LB medium with Amp and incubated at 37 °C overnight. The next day, the preculture was used to inoculate a 5 L expression culture (5 L) and incubated until an OD600 of 0.6-0.8 was obtained. 1 mM IPTG was added to induce the protein expression and the culture was incubated at 18 °C and 180 rpm overnight.

5.5.2. MBP-tagged hnRNP A2/B1 purification

The next day, bacteria were harvested by centrifugation (5000 x g, 15 min, 4 °C) and resuspended in lysis buffer (see section 4.6.). Lysis was performed on ice using a cell disrupter sonicator with 60% amplitude, 10 sec on, 10 sec off. The collected supernatant was loaded onto a 5 mL Ni-based HisTrapTM HP column. The column was connected to the ÄKTA System, which was already equilibrated with wash buffer (see section 4.6.). The protein was eluted with elution buffer (see section 4.6.) and protein containing fractions were combined and dialyzed overnight in the storage buffer (see section 4.6.). The next day, the concentrated protein was purified on SEC with a HiLoad 26/60 Superdex 75 column eluting using storage buffer. Fractions were monitored by SDS gel electrophoresis, and pure fractions were combined and concentrated. The

protein concentration was monitored by Nanodrop (2000c Spectrophotometer), and protein aliquots were stored at -80 °C.

5.6. Protein concentration determination by Bradford assay

The concentration of protein was determined by using the Coomassie® Protein Assay Reagent Kit following the microplate procedure. BSA was used as a protein standard, and measurements were performed at 595 nm in a plate reader (TECAN Spark).

5.7. Direct fluorescence polarization assay

100 µM RNA stock solutions were prepared and stored at -20 °C for later use. 384 well-plates (black, 20 µL) with a round bottom were used for the experiment. To determine the protein concentration the Bradford assay (see section 5.6.) or a NanoDrop 2000c spectrophotometer was used. All assays were performed in duplicates. See section 4.7. for the buffer components used for the experiment.

Fluorescently labeled (FAM or Cy5) RNAs were used for the experiment, and protein with an appropriate concentration was applied in the 384 well-plates as a 2-fold dilution series in the assay buffer. Labeled RNA at a 1 nM final concentration was added (1:1 v/v), and the plate was protected from light. After incubation for 30 or 60 min (if not mentioned otherwise), the fluorescence was measured, with filters set at 490 nm/520 nm (Ex/Em) for FAM labeled RNA and 610 nm/670 nm (Ex/Em) for Cy5 labeled RNA on a plate reader (TECAN Spark).

5.8. Competition Fluorescence polarization

Assay buffer was used for the experiment (see section 4.7.). The protein concentration was determined either spectrophotometrically (NanoDrop 2000c Spectrophotometer) or with the Bradford assay (see section 5.6.). Fluorescently labeled RNAs (1 nM final concentration) were mixed with a protein at the desired concentration corresponding to a 50-80% of binding as determined in the direct FP assay for the given RNA[143]. 2-fold serial dilutions of the unlabeled competitor were added into the 384 well-plates (black, 20µL), which was then mixed 1:1 with pre-incubated labeled RNA/protein mixture. After incubation for 30 minutes or 60 minutes, if not mentioned otherwise, the fluorescence was measured setting either 490 nm/520 nm for FAM or 610 nm/670 nm (Ex/Em) for Cy5 label on a Tecan plate reader (TECAN Spark).

5.9. Solid Phase Peptide Synthesis

5.9.1. Loading amino acid on Rink amide resin

The resin was added into the syringe reactor and swelled with DCM for 20 min. The solvent was exchanged for DMF, and the resin was suspended in 20% piperidine in DMF and agitated for 15 min. In a separate flask, 4 eq of the Fmoc-amino acid was dissolved in a small amount of DMF. The same eq of PyBOP (or COMU, depending on the purpose) and 8 eq of DIPEA was prepared

in DMF and added to the syringe reactor. The mixture was agitated until the coupling reagent dissolved. Next, the amino acid solution was added, and the mixture was shaken for 1 h at rt. Reagents were drained, and the resin was washed with DMF, followed by drying under vacuum. The resin loading was determined using a NanoDrop 2000c Spectrophotometer (see section 5.9.5.).

5.9.2. Loading amino acid to ChemMatrix resin

The same protocol as described in section 5.9.1. was applied.

5.9.3. Loading amino acid to 2-chlorotrityl chloride (2-CTC) resin

The resin was swollen for 30 minutes in DCM (3 mL). DCM was removed, and Fmoc-amino acid (4 eq) and DIPEA (8 eq) in DCM were added into the syringe reactor and shaken for 1 h. The resin was washed with DCM (4 x 1 min) and dried under vacuum.

5.9.4. 2-CTC resin reactivation

For 2-CTC resin that was stored for a longer amount of time a reactivation step was performed before use. In a round bottom flask placed on ice, resin was swollen in DCM (10 mL), and 2.4 eq of pyridine were added, followed by addition of 1.2 eq thionyl chloride, dropwise[199,200]. The temperature was elevated to rt, followed by refluxing the reaction mixture overnight. The next day, the mixture was transferred to a syringe reactor and washed with DCM (4 x 1 min).

5.9.5. Resin loading determination

5 mg of dry resin was weighed in an Eppendorf tube. 1 mL of 20% piperidine in DMF was added to the resin and agitated for 10 min. In a 1 mL quartz cuvette, 1 mL of 20% piperidine in DMF solution was placed, and the UV/Vis was used to measure the blank at 301 nm. The agitated mixture was diluted with 20% piperidine/ DMF (10 or 20-fold dilution was applied depending on the loading efficiency) solution, and the absorption was measured using a NanoDrop 2000c Spectrophotometer. The recorded absorption value, the resin weight, and the volume were entered into an equation (3) derived from Lambert–Beer's law[201], together with the absorption coefficient (7800 $M^{-1}cm^{-1}$) at 301 nm, in order to calculate the resin loading.

$$(3)\ S = E_{301}\, V\, D / \varepsilon\, m\, l$$

S = Substitution [mmol g^{-1}]

ε_{301} = Molar absorption coefficient at 301 nm (M^{-1} cm^{-1})

E_{301} = Absorption of the sample solution at 301 nm

m = Sample weight of the resin [mg]

V = Sample volume

l = Optical path length of the cell [cm]

D = Dilution factor

5.10. Synthesis of the RBM20 peptides on solid phase

Reagents and solvents were commercially available. Peptides were manually synthesized and resins were agitated using either a Vacuum Manifold or shaker. Linear peptide synthesis was performed applying SPPS protocol[129].

The synthesis was carried out on a Rink Amide AM resin (substitution: 0.7 mmol/g, mesh 100-200). After the first amino acid loading (see section 5.9.1.), capping was performed to prevent any side products. For the capping, 10 eq Ac_2O and 12 eq DIPEA in DMF were used, and the reaction was agitated at room temperature for 30 min. Fmoc removal was performed with 20% piperidine in DMF for 1 x 5 min and 1 x 10 min. For coupling, Fmoc-amino acid (4 eq), PyBOP (4 eq), and DIPEA (8 eq) were mixed in DMF. The mixture was added to the syringe reactor, and agitated for 1 h. After draining the solvents, the Fmoc-cleavage solution was added, and the resin was agitated as before. DMF wash (4 x 30 sec) was applied to wash the leftover reagents between every step. If needed, acetylation was carried out using 10 eq Ac_2O and 12 eq DIPEA in DMF (30 min at rt). For the FITC labeling of the peptides, the Fmoc-AEEA-OH linker (4 eq) was attached to the peptide, applying standard coupling protocol. This was followed by FITC labeling overnight at rt using 2 eq FITC isomer and 4 eq DIPEA. Synthesis was monitored by UHPLC. In addition to manual synthesis, synthesis can be automatically performed using the peptide synthesizer (Syro I).

5.10.1. C-terminal FITC labeling

For the C-terminal labeling of RBM20 peptides, Fmoc-L-Lys(Mtt)-OH was loaded (see section 5.9.1.) on Rink Amide AM resin. This was followed by coupling the rest of the amino acids by applying the protocol from section 5.10. Once the linear sequence was completed, Fmoc group removal was carried out, which was followed by acetylation of the N-terminus. The N-(4-methoxytriphenylmethyl) (Mtt) group was deprotected using 40% HFIP in DCM. Additionally, 2.5% TIPS was used as a scavenger. Mtt deprotection was done for 4 x 20 min at room temperature[142]. Then, the Fmoc-AEEA-OH linker was attached to the free amino group of lysine which was followed by FITC labeling (see section 5.10.).

5.10.2. Stapled Peptide synthesis

Either one of two methods were used for peptide stapling:

(i) Rink Amide AM LL resin (substitution: 0.20 - 0.45 mmol/g, 100-200 mesh) was used for the synthesis of the stapled peptides. For the stapling, the dry resin with the N-terminus of the linear peptide still Fmoc-protected (synthesized referring to section 5.10.) was swelled in DCM for 20 min while flushing with argon. To the mixture flushing with argon 1st generation Hoveyda-Grubbs catalyst in 10 mol% was added, and the reaction was carried out for 2 h at rt. This step was repeated three more times. Solvents were removed between steps, and

fresh ones were added under the argon flow. After the cyclization, the resin was washed with DCM twice for 1 min. Completion of the reaction was monitored by UHPLC or FLEET-HPLC.

(ii) DCE was used to swell the dry resin for 20 min under argon. To the mixture under argon Grubbs 2nd generation catalyst (10 mol %) was added. The reaction was heated under MWI at 120 °C for 10 min. This step was repeated once more with fresh catalyst (10 mol %) and after completion, DCM (4 x 1 min) and DMF (4 x 1 min) were utilized to wash the resin.

5.10.3. RBM20 peptide cleavage from solid support

A fresh solution of TFA/TIPS/H_2O (95:2.5:2.5, v/v) was prepared and added to the resin attached peptide. The reaction was agitated for 2 h, and the solution was dropwise added to the ice-cold diethyl ether. For each mL of peptide solution, a 10-fold ether volume was used. To recover remaining peptide, the syringe reactor was shaken with fresh TFA, and the solution was added dropwise to the ice-cold ether. The mixture was centrifuged for 5 min at 4000 rpm at 4 °C. The supernatant was removed, and fresh ether was added and the peptide resuspended. The centrifugation step was repeated three times in total. The dried pellet was dissolved in 50% ACN/50% H_2O/0.1% TFA, followed by lyophilization. Crude peptides were stored at -20 °C until further use.

5.11. Peptide nucleic acid synthesis

As a solid support for PNA synthesis, ChemMatrix resin (substitution: 0.40 - 0.60 mmol/g, mesh 35 - 100) was chosen. The PNA monomer loading was performed using 1.5 eq Fmoc protected PNA monomer, 1.5 eq COMU and 3 eq DIPEA for 2 h at rt. Then, capping was performed with 10 eq Ac_2O and 12 eq DIPEA in DMF for 30 min at rt. Fmoc removal was carried out with 20% pyrrolidine in DMF for 10 min. The resin was washed with DMF (4 x 1 min) between the steps, and the cycle was repeated until the desired PNA sequence was complete. To perform acetylation and FITC labeling the same protocol from section 5.10. was followed. However, in this case, COMU was used as the activating reagent. The synthesis was monitored by UHPLC.

5.11.1. Cleavage

The same cleavage protocol was applied as in section 5.10.3.

5.11.2. Synthesis of PNA variants

PNA variants (compound 11-12) were synthesized on solid support using 2-chlorotrityl chloride resin (substitution: 1.00 - 1.80 mmol/g, 100 - 200 mesh). PNA monomer was loaded to the resin (see section 5.9.3.), and this was followed by capping with DCM: MeOH: DIPEA (80:15:5, v/v) for 15 min. Afterward, the resin was washed with DMF, and the Fmoc group was deprotected. After the resin was washed, the standard capping protocol was applied for the acetylation of N-terminus. Resin was washed with 2 x DMF and 2 x DCM and dried under vacuum. Amide variants

were cleaved from resin with 95% TFA in DCM for 1 h at rt. TFA was evaporated, and the PNA monomer went through purification using RP-HPLC.

5.11.3. PNA monomer synthesis

Fmoc-U-OH was synthesized based on previously reported strategies[183,184]. Initially, to a solution of uracil acetic acid (0.4 mmol) in DMF (5.5 mL), HBTU (0.54 mmol) and DIPEA (0.94 mmol) were added at 0 °C. After stirring for 2 min, Fmoc-protected aminoethyl glycine (0.53 mmol) was added to the reaction mixture at the same temperature. The resulting mixture was allowed to warm to room temperature and stirred for 4 h. The reaction mixture was extracted with ethyl acetate three times. The combined organic layers were washed with 10% aqueous $NaHSO_4$, saturated aqueous $NaHCO_3$, brine, dried over anhydrous sodium sulfate, and concentrated *in vacuo* after filtration.

To remove the tert-butyl group, TFA was slowly added at 0 °C to a solution of the compound in DCM (1:1, v/v). The reaction mixture was then warmed to room temperature and stirred for 4 h. Then, solvents were evaporated, and residual TFA was co-evaporated with chloroform 3-4 times. The obtained product (Fmoc-U-OH) was purified by column chromatography (DCM/MeOH, 5:1), resulting a white powder.

5.12. Synthesis of the hnRNP A2/B1 and SRSF1 peptides

SPPS was performed either manually or automated on a PurePep® Chorus peptide synthesizer.

Evaluation of the synthesis was monitored by performing a test cleavage where a small amount of resin was cleaved and followed by analysis using an Agilent Infinity UHPLC equipped with an Agilent ZORBAX Eclipse Plus column (2.1 mm x 50 mm, 1.8 µm Zorbax Eclipse C18 Rapid Resolution).

5.12.1. Linear peptide synthesis

The peptides were synthesized on 2-CTC resin (substitution: 1.00 - 1.80 mmol/g, 100-200 mesh). After the first amino acid loading (see section 5.9.3.), capping was performed to prevent any side products (see section 5.11.2.). DCM (5 x 1 min) was used to wash the resin, and the solvent was exchanged to DMF to perform the Fmoc group removal (20% piperidine in DMF, 1 x 5 min and 1 x 10 min). The resin was washed with DMF (4 x 1 min), and coupling was carried out as in section 5.10. The cycle was repeated until the desired sequence was obtained. The success of the synthesis was monitored in the UHPLC.

5.12.2. Cleavage of the peptides from solid support

First, the Fmoc group was removed from the N-terminus of the peptide with 20% piperidine in DMF. The resin was agitated with DMF (4 x 1 min) and DCM (2 x 1 min). 20% HFIP/DCM (5 mL) was drawn into the syringe reactor and agitated for 20 min. The procedure was repeated 2 more

times. The solution was drained into a round bottom flask, and the solvents were removed in a vacuum. Remaining traces of solvents were removed under high vacuum overnight.

5.12.3. Cyclization in solution

Dry peptides were suspended either in DCM or DMF depending on the solubility, and 2 eq Py-BOP, 2 eq Oxyma pure, and 4 eq DIPEA were added. Peptide concentration in the solution should was kept below 1 mM. The mixture was agitated overnight, and the solvents were evaporated using a rotary evaporator. A rotary evaporator connected to a high vacuum pump was used to evaporate DMF.

5.12.4. Cleavage of protecting groups

A fresh solution of TFA/TIPS/H_2O/DODT (95:2.5:2.5:5, v/v) was prepared and added to the protected cyclic peptide. The protocol from section 5.10.3. was followed to obtain the peptide as a powder.

5.13. Peptide purification

5.13.1. Preparative RP-HPLC

An Agilent Infinity II LC-MS system equipped with 125 mm x 21 mm, 5 µm or 125 mm x 10 mm, 5 µm Macherey-Nagel Nucleodur C18 Gravity column was used for the purification of peptides. A 20 mL/min flow rate was applied and detection at UV detection was performed at 210 nm. As an elution solvent, H_2O and CH_3CN supplemented with 0.1% TFA were used.

5.13.2. LC-MS purity analysis

The fractions from preparative HPLC were collected and analyzed using an Agilent Infinity HPLC system equipped with 50 mm x 3 mm, 1.8 µm Macherey-Nagel Nucleodur C18 Gravity column. Analysis was carried out applying 5-95% ACN (0.1% TFA) for 30 min and with a flow rate of 0.56 mL/min.

5.13.3. High-resolution mass spectrometry

LTQ-XL Orbitrap mass spectrometer with Hypersyl GOLD (50 mm x 1 mm, 1.9 µm column) was used for high-resolution mass spectrometry (HRMS).

5.14. Circular Dichroism Spectroscopy

For circular dichroism analysis, 100 µM unlabeled peptides were prepared in CD buffer (20 mM sodium phosphate buffer, 10 mM NaF, pH 8), and the absorbance was measured in the UV range from 190 to 300 nm using JASCO CD spectrometer at 20 °C and with 1 mm pathlength.

5.15. Microscale thermophoresis

MST was carried out with MBP tagged hnRNP A2/B1 (residues 1-251) in 100 mM NaCl, 25 mM HEPES, 0.05% Tween-20, pH 8. The protein was labelled with RED fluorescent dye NT-647-

NHS and labeling was performed according to the manufacturer's instructions (Monolith NT™ Protein Labeling Kit RED-NHS). Labeling efficiency was determined using a Nanodrop 2000c Spectrophotometer [202]. Then, 0.5 mL reaction tubes or PCR tubes were used to prepare 10 µL of 2-fold serial dilutions of the peptide. 10 µL of labeled protein was added into each tube (1:1, v/v) to reach a final protein concentration of 250 nM. The mixture was incubated for 30 minutes. Afterward, samples were loaded into the premium capillaries (Monolith Premium Capillaries), and Monolith Nt.115 device was used for the measurements.

5.16. Mass photometry

Mass photometry experiments were performed using MBP-tagged hnRNP A2/B1 (residues 1-251) in storage buffer (25 mM HEPES, 100 mM NaCl, 1 mM TCEP, pH 8). 50 nM protein and 12.5 nM, 25 nM, and 50 nM Cy5 labelled RNA were used. Protein and RNA were mixed and incubated for 30 minutes at room temperature. Calibration was done using 5 µM bovine serum albumin (BSA) and 5 µM thyroglobulin (TG) protein. Measurement was carried out using a Refeyn TwoMP instrument.

6. Literature

1. Glisovic, T., Bachorik, J. L., Yong, J. & Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* **582**, 1977–1986 (2008).

2. Moore, M. J. From birth to death: The complex lives of eukaryotic mRNAs. *Science (80-.).* **309**, 1514–1518 (2005).

3. Wilkinson, M. E., Charenton, C. & Nagai, K. RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **89**, 359–388 (2020).

4. Cook-Andersen, H. & Wilkinson, M. F. Splicing does the two-step. *Nature* **521**, 300–301 (2015).

5. Chen, M. & Manley, J. L. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **10**, 741–754 (2009).

6. Crotti, L. B. & Horowitz, D. S. Exon sequences at the splice junctions affect splicing fidelity and alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18954–18959 (2009).

7. Tao, Y., Zhang, Q., Wang, H., Yang, X. & Mu, H. Alternative splicing and related RNA binding proteins in human health and disease. *Signal Transduct. Target. Ther.* **9**, (2024).

8. Liu, Q., Fang, L. & Wu, C. Alternative Splicing and Isoforms: From Mechanisms to Diseases. *Genes (Basel).* **13**, (2022).

9. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **17**, 19–32 (2016).

10. Zhang, C., Yang, H. & Yang, H. Evolutionary Character of Alternative Splicing in Plants. *Bioinform. Biol. Insights* **9s1**, 47–52 (2015).

11. Schwartz, S. *et al.* Large-scale comparative analysis of splicing signals and their corresponding splicing factors in eukaryotes. *Genome Res.* 88–103 (2008) doi:10.1101/gr.6818908.cruitment.

12. Gonzalez-Hilarion, S. *et al.* Intron retention-dependent gene regulation in Cryptococcus neoformans. *Sci. Rep.* **6**, 1–13 (2016).

13. Zheng, S. Alternative splicing and nonsense-mediated mRNA decay enforce neural specific gene expression. *Int. J. Dev. Neurosci.* **55**, 102–108 (2016).

14. Ren, X., Zhang, K., Deng, R. & Li, J. RNA Splicing Analysis: From In Vitro Testing to Single-Cell Imaging. *Chem* **5**, 2571–2592 (2019).

15. WANG, Y. *et al.* Mechanism of alternative splicing and its regulation. *Biomed. Reports* **3**, 152–158 (2015).

16. Blencowe, B. J. Exonic splicing enhancers: Mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.* **25**, 106–110 (2000).

17. Rogalska, M. E., Vivori, C. & Valcárcel, J. Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects. *Nat. Rev. Genet.* **24**, 251–269 (2023).

18. Dvinge, H., Kim, E., Abdel-Wahab, O. & Bradley, R. K. RNA splicing factors as oncoproteins and tumour suppressors. *Nat. Rev. Cancer* **16**, 413–430 (2016).

19. Wang, Z., Xiao, X., Van Nostrand, E. & Burge, C. B. General and Specific Functions of Exonic Splicing Silencers in Splicing Control. *Mol. Cell* **23**, 61–70 (2006).

20. Wang, Z. & Burge, C. B. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *Rna* **14**, 802–813 (2008).

21. Feracci, M. *et al.* Structural basis of RNA recognition and dimerization by the STAR proteins T-STAR and Sam68. *Nat. Commun.* **7**, (2016).

22. Huelga, S. C. *et al.* of alternative splicing by hnRNP proteins. **1**, 167–178 (2012).

23. Waris, S. *et al.* TIA-1 RRM23 binding and recognition of target oligonucleotides. *Nucleic Acids Res.* **45**, 4944–4957 (2017).

24. Fredericks, A. M., Cygan, K. J., Brown, B. A. & Fairbrother, W. G. RNA-binding proteins: Splicing factors and disease. *Biomolecules* **5**, 893–909 (2015).

25. Chen, Y. & Varani, G. Protein families and RNA recognition. *FEBS J.* **272**, 2088–2097 (2005).

26. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: Modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* **8**, 479–490 (2007).

27. Gebauer, F., Schwarzl, T., Valcárcel, J. & Hentze, M. W. RNA-binding proteins in human genetic disease. *Nat. Rev. Genet.* **22**, 185–198 (2021).

28. Srikantan, S. & Gorospe, M. UneCLIPsing HuR Nuclear Function. *Mol. Cell* **43**, 319–321 (2011).

29. Corley, M., Burns, M. C. & Yeo, G. W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell* **78**, 9–29 (2020).

30. Auweter, S. D., Oberstrass, F. C. & Allain, F. H. T. Sequence-specific binding of single-stranded RNA: Is there a code for recognition? *Nucleic Acids Res.* **34**, 4943–4959 (2006).

31. Teplova, M. *et al.* Structure-function studies of STAR family quaking proteins bound to their in vivo RNA target sites. *Genes Dev.* **27**, 928–940 (2013).

32. Hu, W., Qin, L., Li, M., Pu, X. & Guo, Y. A structural dissection of protein-RNA interactions based on different RNA base areas of interfaces. *RSC Adv.* **8**, 10582–10592 (2018).

33. Wilson, K. A., Holland, D. J. & Wetmore, S. D. Topology of RNA-protein nucleobase-amino acid π-π interactions and comparison to analogous DNA-protein π-π contacts. *Rna* **22**, 696–708 (2016).

34. Saeedi, P. *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* **157**, 107843 (2019).

35. Saggar, a *et al.* Multiple Susceptibility Variants. *Science (80-.).* **316**, 1341–1345 (2007).

36. Nutter, C. A. *et al.* Dysregulation of RBFOX2 Is an Early Event in Cardiac Pathogenesis of Diabetes. *Cell Rep.* **15**, 2200–2213 (2016).

37. Li, D. *et al.* Identification of novel mutations in RBM20 in patients with dilated cardiomyopathy. *Clin. Transl. Sci.* **3**, 90–97 (2010).

38. Brauch, K. M. *et al.* Mutations in RNA Binding Protein Gene Cause Familial Dilated Cardiomyopathy. *J Am Coll Cardiol* **54**, 930–941 (2010).

39. Kang, D., Lee, Y. & Lee, J.-S. RNA-Binding Proteins in Cancer : Functional and. *Cancers (Basel).* **12**, 2699 (2020).

40. Wang, J., Liu, Q. & Shyr, Y. Dysregulated transcription across diverse cancer types reveals the importance of RNA-binding protein in carcinogenesis. *BMC Genomics* **16**, 1–10 (2015).

41. Degrauwe, N., Suvà, M. L., Janiszewska, M., Riggi, N. & Stamenkovic, I. Imps: An RNA-binding protein family that provides a link between stem cell maintenance in normal development and cancer. *Genes Dev.* **30**, 2459–2474 (2016).

42. Farina, K. L., Hüttelmaier, S., Musunuru, K., Darnell, R. & Singer, R. H. Two ZBP1 KH domains facilitate β-actin mRNA localization, granule formation, and cytoskeletal attachment. *J. Cell Biol.* **160**, 77–87 (2003).

43. Gu, W. *et al.* Regulation of local expression of cell adhesion and motility-related mRNAs in breast cancer cells by IMP1/ZBP1. *J. Cell Sci.* **125**, 81–91 (2012).

44. Hsieh, A. C. *et al.* HHS Public Access. **16**, 4914–4920 (2020).

45. Lal, A. *et al.* Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO J.* **23**, 3092–3102 (2004).

46. Zhao, M., Kim, J. R., Bruggen, R. Van & Park, J. Molecules and Cells RNA-Binding Proteins in Amyotrophic Lateral Sclerosis. *Mol. Cells* **41**, 818–829 (2018).

47. Nelson, D. L., Orr, H. T. & Warren, S. T. The unstable repeats-Three evolving faces of neurological disease. *Neuron* **77**, 825–843 (2013).

48. Cookson, M. R. RNA-binding proteins implicated in neurodegenerative diseases. *Wiley Interdiscip. Rev. RNA* **8**, 1–18 (2017).

49. Vasilyev, N. *et al.* Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5391–E5400 (2015).

50. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).

51. Byun, W. G., Lim, D. & Park, S. B. Discovery of Small-Molecule Modulators of Protein–RNA Interactions by Fluorescence Intensity-Based Binding Assay. *ChemBioChem* **21**, 818–824 (2020).

52. White, A. W., Westwell, A. D. & Brahemi, G. Protein - Protein interactions as targets for small-molecule therapeutics in cancer. *Expert Rev. Mol. Med.* **10**, 1–14 (2008).

53. Wang, L. *et al.* Therapeutic peptides: current applications and future directions. *Signal Transduct. Target. Ther.* **7**, (2022).

54. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: Progressing toward the reality. *Chem. Biol.* **21**, 1102–1114 (2014).

55. Chang, Y. S. *et al.* Stapled α-helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).

56. Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of α-helix-mediated protein-protein interactions using designed molecules. *Nat. Chem.* **5**, 161–173 (2013).

57. Diao, L. & Meibohm, B. Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides. *Clin. Pharmacokinet.* **52**, 855–868 (2013).

58. Lee, M. F. & Poh, C. L. Strategies to improve the physicochemical properties of peptide-based drugs. *Pharm. Res.* **40**, 617–632 (2023).

59. Masri, E., Ahsanullah, Accorsi, M. & Rademann, J. Side-Chain Modification of Peptides Using a Phosphoranylidene Amino Acid. *Org. Lett.* **22**, 2976–2980 (2020).

60. Wagstaff, K. M. & Jans, D. A. Protein Transduction: Cell Penetrating Peptides and Their Therapeutic Applications. *Front. Med. Chem. (Volume 5)* 98–126 (2012) doi:10.2174/978160805208011005010098.

61. Low, S. C., Nunes, S. L., Bitonti, A. J. & Dumont, J. A. Oral and pulmonary delivery of FSH-Fc fusion proteins via neonatal Fc receptor-mediated transcytosis. *Hum. Reprod.* **20**, 1805–1813 (2005).

62. Irby, D., Du, C. & Li, F. Lipid-Drug Conjugate for Enhancing Drug Delivery. *Mol. Pharm.* **14**, 1325–1338 (2017).

63. White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509–524 (2011).

64. Pelay-Gimeno, M., Tulla-Puche, J. & Albericio, F. 'Head-to-side-chain' cyclodepsipeptides of marine origin. *Mar. Drugs* **11**, 1693–1717 (2013).

65. Van Der Knaap, M. *et al.* Design, synthesis and structural analysis of mixed α/β-peptides that adopt stable cyclic hairpin-like conformations. *Tetrahedron* **68**, 2391–2400 (2012).

66. Ricardo, M. G., Vasco, A. V., Rivera, D. G. & Wessjohann, L. A. Stabilization of Cyclic β-Hairpins by Ugi-Reaction-Derived N-Alkylated Peptides: The Quest for Functionalized β-Turns. *Org. Lett.* **21**, 7307–7310 (2019).

67. Choi, J. S. & Joo, S. H. Recent trends in cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **28**, 18–24 (2020).

68. Bechtler, C. & Lamers, C. Macrocyclization strategies for cyclic peptides and peptidomimetics. *RSC Med. Chem.* **12**, 1325–1351 (2021).

69. Hayes, H. C., Luk, L. Y. P. & Tsai, Y. H. Approaches for peptide and protein cyclisation. *Org. Biomol. Chem.* **19**, 3983–4001 (2021).

70. Bhat, A., Roberts, L. R. & Dwyer, J. J. Lead discovery and optimization strategies for peptide macrocycles. *Eur. J. Med. Chem.* **94**, 471–479 (2015).

71. Taylor, J. W. The synthesis and study of side-chain lactam-bridged peptides. *Biopolym. - Pept. Sci. Sect.* **66**, 49–75 (2002).

72. El-Faham, A. & Albericio, F. Peptide coupling reagents, more than a letter soup. *Chem. Rev.* **111**, 6557–6602 (2011).

73. Wen, S. J. & Yao, Z. J. Total synthesis of cyclomarin C. *Org. Lett.* **6**, 2721–2724 (2004).

74. Jin, K. *et al.* Total synthesis of teixobactin. *Nat. Commun.* **7**, (2016).

75. LaBelle, J. L. *et al.* A stapled BIM peptide overcomes apoptotic resistance in hematologic cancers. *J. Clin. Invest.* **122**, 2018–2031 (2012).

76. Moiola, M., Memeo, M. G. & Quadrelli, P. Stapled peptides-a useful improvement for peptide-based drugs. *Molecules* **24**, (2019).

77. Haney, C. M. & Horne, W. S. Oxime side-chain cross-links in an α-helical coiled-coil protein: Structure, thermodynamics, and folding-templated synthesis of bicyclic species. *Chem. - A Eur. J.* **19**, 11342–11351 (2013).

78. Haney, C. M. & Horne, W. S. Dynamic covalent side-chain cross-links via intermolecular oxime or hydrazone formation from bifunctional peptides and simple organic linkers. *J. Pept. Sci.* **20**, 108–114 (2014).

79. Song, J. M., Gallagher, E. E., Menon, A., Mishra, L. D. & Garner, A. L. The role of olefin geometry in the activity of hydrocarbon stapled peptides targeting eukaryotic translation initiation factor 4E (eIF4E). *Org. Biomol. Chem.* **17**, 6414–6419 (2019).

80. Van Lierop, B. *et al.* Insulin in motion: The A6-A11 disulfide bond allosterically modulates structural transitions required for insulin activity. *Sci. Rep.* **7**, 1–14 (2017).

81. Sohrabi, C., Foster, A. & Tavassoli, A. Methods for generating and screening libraries of genetically encoded cyclic peptides in drug discovery. *Nat. Rev. Chem.* **4**, 90–101 (2020).

82. Parmley, S. F. & Smith, G. P. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**, 305–318 (1988).

83. Scott, J. K. & Smith, G. P. Searching for peptide ligands with an epitope library. *Science (80-.).* **249**, 386–390 (1990).

84. Smith, G. P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science (80-.).* **228**, 1315–1317 (1985).

85. Castel, G., Chtéoui, M., Heyd, B. & Tordo, N. Phage display of combinatorial peptide libraries: Application to antiviral research. *Molecules* **16**, 3499–3518 (2011).

86. DeLano, W. L., Ultsch, M. H., De Vos, A. M. & Wells, J. A. Convergent solutions to binding at a protein-protein interface. *Science (80-.).* **287**, 1279–1283 (2000).

87. Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L. & Węgrzyn, G. Phage display and other peptide display technologies. *FEMS Microbiol. Rev.* **46**, 1–25 (2022).

88. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408 (1997).

89. Roberts, R. W. & Szostakt, J. W. RNA-Peptide Fusions for the in vitro Selection of Peptides and Proteins Author (s): Richard W . Roberts and Jack W . Szostak Source : Proceedings of the National Academy of Sciences of the United States of Published by : National Academy of Sciences Sta. **94**, 12297–12302 (1997).

90. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J. Am. Chem. Soc.* **141**, 4167–4181 (2019).

91. Dougherty, P. G., Qian, Z. & Pei, D. Macrocycles as protein-protein interaction inhibitors. *Biochem. J.* **474**, 1109–1125 (2017).

92. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C. & Benkovic, S. J. Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 13638–13643 (1999).

93. Paulus, H. The chemical basis of protein splicing. *Chem. Soc. Rev.* **27**, 375–386 (1998).

94. Scott, C. P., Abel-Santos, E., Jones, A. D. & Benkovic, S. J. Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chem. Biol.* **8**, 801–815 (2001).

95. Lennard, K. R. & Tavassoli, A. Peptides come round: Using SICLOPPS libraries for early stage drug discovery. *Chem. - A Eur. J.* **20**, 10608–10614 (2014).

96. Miranda, E. *et al.* A cyclic peptide inhibitor of HIF-1 heterodimerization that inhibits hypoxia signaling in cancer cells. *J. Am. Chem. Soc.* **135**, 10418–10425 (2013).

97. Sohrabi, C., Foster, A. & Tavassoli, A. Methods for generating and screening libraries of genetically encoded cyclic peptides in drug discovery. *Nat. Rev. Chem.* **4**, 90–101 (2020).

98. Nielsen, P. E. & Egholm, M. An introduction to peptide nucleic acid. *Curr. Issues Mol. Biol.* **1**, 89–104 (1999).

99. Sharma, C. & Awasthi, S. K. Versatility of peptide nucleic acids (PNAs): role in chemical biology, drug discovery, and origins of life. *Chem. Biol. Drug Des.* **89**, 16–37 (2017).

100. Wu, J. *et al.* Recent advances in peptide nucleic acid for cancer bionanotechnology. *Nat. Publ. Gr.* **38**, 798–805 (2017).

101. Kolevzon, N., Nasereddin, A., Naik, S., Yavin, E. & Dzikowski, R. Use of peptide nucleic acids to manipulate gene expression in the malaria parasite Plasmodium falciparum. *PLoS One* **9**, (2014).

102. Shiraishi, T., Eysturskard, J. & Nielsen, P. E. Modulation of mdm2 pre-mRNA splicing by 9-aminoacridine-PNA (peptide nucleic acid) conjugates targeting intron-exon junctions. *BMC Cancer* **10**, (2010).

103. Ni, D., Liu, N. & Sheng, C. *Allosteric modulators of protein–protein interactions (PPIs)*. *Advances in Experimental Medicine and Biology* vol. 1163 (2019).

104. Monod, J., Wyman, J. & Changeux, J. P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).

105. Schauer, G. D., Huber, K. D., Leuba, S. H. & Sluis-Cremer, N. Mechanism of allosteric inhibition of HIV-1 reverse transcriptase revealed by single-molecule and ensemble fluorescence. *Nucleic Acids Res.* **42**, 11687–11696 (2014).

106. Levy, J. A. HIV pathogenesis: 25 years of progress and persistent challenges. *Aids* **23**, 147–160 (2009).

107. Mannes, M., Martin, C., Menet, C. & Ballet, S. Wandering beyond small molecules: peptides as allosteric protein modulators. *Trends Pharmacol. Sci.* **43**, 406–423 (2022).

108. Leroux, A. E. & Biondi, R. M. Renaissance of Allostery to Disrupt Protein Kinase Interactions. *Trends Biochem. Sci.* **45**, 27–41 (2020).

109. Schmeing, S. *et al.* Rationally designed stapled peptides allosterically inhibit PTBP1-RNA-binding. *Chem. Sci.* **14**, 8269–8278 (2023).

110. Lennermann, D., Backs, J. & van den Hoogenhof, M. M. G. New Insights in RBM20 Cardiomyopathy. *Curr. Heart Fail. Rep.* **17**, 234–246 (2020).

111. Bull, M. *et al.* Alternative splicing of titin restores diastolic function in an HFpEF-like genetic murine model (Ttn ΔIAjxn). *Circ. Res.* **119**, 764–772 (2016).

112. Morales, A. & Hershberger, R. E. Genetic evaluation of dilated cardiomyopathy. *Curr. Cardiol.*

Rep. **15**, (2013).

113. Gigli, M. *et al.* A Review of the Giant Protein Titin in Clinical Molecular Diagnostics of Cardiomyopathies. *Front. Cardiovasc. Med.* **3**, 1–9 (2016).

114. Linke, W. A. Sense and stretchability: The role of titin and titin-associated proteins in myocardial stress-sensing and mechanical dysfunction. *Cardiovasc. Res.* **77**, 637–648 (2008).

115. Granzier, H. L. & Labeit, S. The Giant Protein Titin: A Major Player in Myocardial Mechanics, Signaling, and Disease. *Circ. Res.* **94**, 284–295 (2004).

116. Guo, W. *et al.* Europe PMC Funders Group RBM20 , a gene for hereditary cardiomyopathy , regulates titin splicing. *Nat. Med.* **18**, 766–773 (2013).

117. Butler, J. *et al.* Developing therapies for heart failure withpreservedejection fraction. Current state and future directions. *JACC Hear. Fail.* **2**, 97–112 (2014).

118. Haq, M. A. ul. Therapeutic interventions for heart failure with preserved ejection fraction: A summary of current evidence. *World J Cardiol* **6**, 67 (2014).

119. Guo, W. & Sun, M. RBM20, a potential target for treatment of cardiomyopathy via titin isoform switching. *Biophys. Rev.* **10**, 15–25 (2018).

120. Watanabe, T., Kimura, A. & Kuroyanagi, H. Alternative splicing regulator RBM20 and cardiomyopathy. *Front. Mol. Biosci.* **5**, 1–11 (2018).

121. Murayama, R. *et al.* Phosphorylation of the RSRSP stretch is critical for splicing regulation by RNA-Binding Motif Protein 20 (RBM20) through nuclear localization. *Sci. Rep.* **8**, 1–14 (2018).

122. Upadhyay, S. K. & MacKereth, C. D. Structural basis of UCUU RNA motif recognition by splicing factor RBM20. *Nucleic Acids Res.* **48**, 4538–4550 (2021).

123. Maatz, H. *et al.* RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *J. Clin. Invest.* **124**, 3419–3430 (2014).

124. Guarracino, D. A. *et al.* Macrocyclic Control in Helix Mimetics. *Chem. Rev.* **119**, 9915–9949 (2019).

125. Lau, Y. H. *et al.* Functionalised staple linkages for modulating the cellular activity of stapled peptides. *Chem. Sci.* **5**, 1804–1809 (2014).

126. Lau, Y. H., De Andrade, P., Wu, Y. & Spring, D. R. Peptide stapling techniques based on different macrocyclisation chemistries. *Chem. Soc. Rev.* **44**, 91–102 (2015).

127. Schafmeister, C. E., Po, J. & Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides [8]. *J. Am. Chem. Soc.* **122**, 5891–5892 (2000).

128. Walensky, L. D. & Bird, G. H. Hydrocarbon-Stapled Peptides: Principles, Practice, and

Progress. (2014).

129. Merrifield. Solid-Phase Peptide Synthesis. 1905–1909 (1963).

130. Coin, I., Beyermann, M. & Bienert, M. Solid-phase peptide synthesis: From standard procedures to the synthesis of difficult sequences. *Nat. Protoc.* **2**, 3247–3256 (2007).

131. Shim, S. Y., Kim, Y. W. & Verdine, G. L. A new i, i + 3 peptide stapling system for α-helix stabilization. *Chem. Biol. Drug Des.* **82**, 635–642 (2013).

132. Li, X., Chen, S., Zhang, W. D. & Hu, H. G. Stapled Helical Peptides Bearing Different Anchoring Residues. *Chem. Rev.* **120**, 10079–10144 (2020).

133. Li, X., Chen, S., Zhang, W. D. & Hu, H. G. Stapled Helical Peptides Bearing Different Anchoring Residues. *Chem. Rev.* (2020) doi:10.1021/acs.chemrev.0c00532.

134. Kim, Y. W., Grossmann, T. N. & Verdine, G. L. Synthesis of all-hydrocarbon stapled ±-helical peptides by ring-closing olefin metathesis. *Nat. Protoc.* **6**, 761–771 (2011).

135. Blackwell, H. E. & Grubbs, R. H. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Angew. Chemie - Int. Ed.* **37**, 3281–3284 (1998).

136. Ladokhin, A. S., Fernández-Vidal, M. & White, S. H. CD spectroscopy of peptides and proteins bound to large unilamellar vesicles. *J. Membr. Biol.* **236**, 247–253 (2010).

137. Miles, A. J., Janes, R. W. & Wallace, B. A. Tools and methods for circular dichroism spectroscopy of proteins: A tutorial review. *Chem. Soc. Rev.* **50**, 8400–8413 (2021).

138. Sahoo, H. & Molecules, B. *Optical Spectroscopic and Microscopic Techniques*. *Optical Spectroscopic and Microscopic Techniques* (2022). doi:10.1007/978-981-16-4550-1.

139. Tavassoli, A. SICLOPPS cyclic peptide libraries in drug discovery. *Curr. Opin. Chem. Biol.* **38**, 30–35 (2017).

140. Wei, Y., Thyparambil, A. A. & Latour, R. A. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230 nm. *Biochim. Biophys. Acta - Proteins Proteomics* **1844**, 2331–2337 (2014).

141. Merrifield, R. B. Solid phase peptide synthesis. *Excerpta Med., I.C.S.* **No.374**, 29–39 (1976).

142. Wester, A., Hansen, A. M., Hansen, P. R. & Franzyk, H. Perfluoro-tert-butanol for selective on-resin detritylation: a mild alternative to traditionally used methods. *Amino Acids* **53**, 1455–1466 (2021).

143. Moerke, N. J. Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Curr. Protoc. Chem. Biol.* **1**, 1–15 (2009).

144. Isabella L. Karle, B. P. Structural Characteristics of «-Helical Peptide Molecules Containing Aib Residues. 6747–6756 (1990).

145. Otvos, L. *Peptide-Based Drug Design: Preface*. *Methods in Molecular Biology* vol. 494 (2008).

146. Hall, M. D. *et al.* Fluorescence polarization assays in high-throughput screening and drug discovery. **4**, 1–41 (2017).

147. Article, R., Floden, A, Combs, C. & Article, R. Fluorescence Polarization Assays in Small Molecule Screening. *Bone* **23**, 1–7 (2008).

148. Dandliker, W. B., Hsu, M. L., Levin, J. & Rao, B. R. Equilibrium and Kinetic Inhibition Assays Based upon Fluorescence Polarization. *Methods Enzymol.* **74**, 3–28 (1981).

149. Lundblad, J. R., Laurance, M. & Goodman, R. H. Fluorescence of Protein-DNA Interactions Polarization Analysis and Protein-Protein. *Library (Lond).* 607–612 (2008).

150. Dauksaite, V. & Gotthardt, M. Molecular basis of titin exon exclusion by RBM20 and the novel titin splice regulator PTB4. *Nucleic Acids Res.* **46**, 5227–5238 (2018).

151. Serafin Piiiol·Roma & Gideon Dreyfuss. Shuttling of pre.mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**, 730–732 (1992).

152. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Hum. Genet.* **135**, 851–867 (2016).

153. Gautrey, H. *et al.* SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells. *RNA Biol.* **12**, 1139–1151 (2015).

154. Xu, Y. *et al.* Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. *Genes Dev.* **28**, 1191–1203 (2014).

155. Suzuki, H. & Matsuoka, M. hnRNPA1 autoregulates its own mRNA expression to remain non-cytotoxic. *Mol. Cell. Biochem.* **427**, 123–131 (2017).

156. Feng, J., Zhou, J., Lin, Y., Huang, W. & Brook, M. hnRNP A1 in RNA metabolism regulation and as a potential therapeutic target. 1–17 (2022) doi:10.3389/fphar.2022.986409.

157. Jean-Philippe, J., Paz, S. & Caputi, M. hnRNP A1: The Swiss Army Knife of gene expression. *Int. J. Mol. Sci.* **14**, 18999–19024 (2013).

158. Anantha, R. W. *et al.* Requirement of Heterogeneous Nuclear Ribonucleoprotein C for BRCA Gene Expression and Homologous Recombination. *PLoS One* **8**, (2013).

159. Waggoner, S. A., Johannes, G. J. & Liebhaber, S. A. Depletion of the poly(C)-binding proteins αCP1 and αCP2 from K562 cells leads to p53-independent induction of cyclin-dependent kinase inhibitor (CDKN1A) and G1 arrest. *J. Biol. Chem.* **284**, 9039–9049 (2009).

160. Ren, C., Cho, S. J., Jung, Y. S. & Chen, X. DNA polymerase η is regulated by poly(rC)-binding protein 1 via mRNA stability. *Biochem. J.* **464**, 377–386 (2014).

161. Cho, S. J., Jung, Y. S. & Chen, X. Poly (C)-Binding Protein 1 Regulates p63 Expression

through mRNA Stability. *PLoS One* **8**, 1–7 (2013).

162. Hussey, G. S. *et al.* Identification of an mRNP Complex Regulating Tumorigenesis at the Translational Elongation Step. *Mol. Cell* **41**, 419–431 (2011).

163. Park, S. J. *et al.* Heterogeneous nuclear ribonucleoprotein A1 post-transcriptionally regulates Drp1 expression in neuroblastoma cells. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 1423–1431 (2015).

164. Mayeda, A. & Krainer, A. R. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**, 365–375 (1992).

165. Shan, J., Munro, T. P., Barbarese, E., Carson, J. H. & Smith, R. A molecular mechanism for mRNA trafficking in neuronal dendrites. *J. Neurosci.* **23**, 8859–8866 (2003).

166. Carpenter, B. *et al.* The roles of heterogeneous nuclear ribonucleoproteins in tumour development and progression. *Biochim. Biophys. Acta - Rev. Cancer* **1765**, 85–100 (2006).

167. Kim, M. K. *et al.* Heterogeneous nuclear ribonucleoprotein A2/B1 regulates the ERK and p53/HDM2 signaling pathways to promote the survival, proliferation and migration of non-small cell lung cancer cells. *Oncol. Rep.* **46**, 1–10 (2021).

168. Zhou, J. *et al.* Differential expression of the early lung cancer detection marker, heterogeneous nuclear ribonucleoprotein-A2/B1 (hnRNP-A2/B1) in normal breast and neoplastic breast cancer. *Breast Cancer Res. Treat.* **66**, 217–224 (2001).

169. Golan-Gerstl, R. *et al.* Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. *Cancer Res.* **71**, 4464–4472 (2011).

170. Harrison, A. F. & Shorter, J. RNA-binding proteins with prion-like domains in health and disease. *Biochem. J.* **474**, 1417–1438 (2017).

171. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).

172. Wu, B. *et al.* Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1. *Nat. Commun.* **9**, (2018).

173. Bériault, V. *et al.* A late role for the association of hnRNP A2 with the HIV-1 hnRNP A2 response elements in genomic RNA, Gag, and Vpr localization. *J. Biol. Chem.* **279**, 44141–44153 (2004).

174. Villarroya-Beltri, C. *et al.* Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nat. Commun.* **4**, 1–10 (2013).

175. Emily, K. M., Maggie, M. B., Karla, S., Krystal, H. & Aaron, M. J. An RNA matchmaker protein regulates the activity of the long noncoding RNA HOTAIR. *Rna* **22**, 995–1010 (2016).

176. Wu, B. *et al.* Molecular basis for the specific and multivariant recognitions of RNA substrates by

human hnRNP A2/B1. *Nat. Commun.* **9**, (2018).

177. Alarcón, C. R., Lee, H., Goodarzi, H., Halberg, N. & Tavazoie, S. F. N6-methyladenosine marks primary microRNAs for processing. *Nature* **519**, 482–485 (2015).

178. Goodarzi, H. *et al.* Processing Events. **162**, 1299–1308 (2015).

179. Zhu, Z., Pei, D. & Zhu, Z. Crosstalk between m6A modification and alternative splicing during cancer progression Crosstalk between m6A modification and alternative splicing during cancer progression. (2023) doi:10.1002/ctm2.1460.

180. Dowling, P. *et al.* Abnormal levels of heterogeneous nuclear ribonucleoprotein A2B1 (hnRNPA2B1) in tumour tissue and blood samples from patients diagnosed with lung cancer. *Mol. Biosyst.* **11**, 743–752 (2015).

181. Hu, Y. *et al.* Splicing factor hnRNPA2B1 contributes to tumorigenic potential of breast cancer cells through STAT3 and ERK1/2 signaling pathway. *Tumor Biol.* **39**, 1–11 (2017).

182. Hyrup, B. & Nielsen, P. E. Peptide nucleic acids (PNA): Synthesis, properties and potential applications. *Bioorganic Med. Chem.* **4**, 5–23 (1996).

183. Manicardi, A. *et al.* PNA bearing 5-azidomethyluracil: A novel approach for solid and solution phase modification. *Artif. DNA PNA XNA* **3**, 1–10 (2012).

184. Oquare, B. Y. & Taylor, J. S. Synthesis of peptide nucleic acid FRET probes via an orthogonally protected building block for post-synthetic labeling of peptide nucleic acids at the 5-position of uracil. *Bioconjug. Chem.* **19**, 2196–2204 (2008).

185. Tavassoli, A. & Benkovic, S. J. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E. coli. *Nat. Protoc.* **2**, 1126–1133 (2007).

186. Nowacki, J. *et al.* A translational repression reporter assay for the analysis of RNA-binding protein consensus sites. *RNA Biol.* **20**, 85–94 (2023).

187. Cheng, L. *et al.* Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease. *Protein Sci.* **16**, 1535–1542 (2007).

188. Behjati, S. & Tarpey, P. S. What is next generation sequencing? *Arch. Dis. Child. Educ. Pract. Ed.* **98**, 236–238 (2013).

189. Devarakonda, S., Morgensztern, D. & Govindan, R. Role of Next-Generation Sequencing Technologies. *Target. Ther. Oncol.* **122**, 1–15 (2013).

190. Brandt, R. B., Laux, J. E. & Yates, S. W. Calculation of inhibitor Ki and inhibitor type from the concentration of inhibitor for 50% inhibition for Michaelis-Menten enzymes. *Biochem. Med. Metab. Biol.* **37**, 344–349 (1987).

191. Morrison, K. L. & Weiss, G. A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **5**, 302–307 (2001).

192. Hicks, G. R. & Zhang, C. *Plant Chemical Genomics Methods and Protocols Second Edition Methods in Molecular Biology 2213*. (2021).

193. Cattoni, D. I., Chara, O., Kaufman, S. B. & Flecha, F. L. G. Cooperativity in binding processes: New insights from phenomenological modeling. *PLoS One* **10**, 1–14 (2015).

194. Stefan, M. I. & Le Novère, N. Cooperative Binding. *PLoS Comput. Biol.* **9**, 2–7 (2013).

195. Foley, E. D. B., Kushwah, M. S., Young, G. & Kukura, P. Mass photometry enables label-free tracking and mass measurement of single proteins on lipid bilayers. *Nat. Methods* **18**, 1247–1252 (2021).

196. Soltermann, F. *et al.* Quantifying Protein–Protein Interactions by Molecular Counting with Mass Photometry. *Angew. Chemie - Int. Ed.* **59**, 10774–10779 (2020).

197. Cole, D., Young, G., Weigel, A., Sebesta, A. & Kukura, P. Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics* **4**, 211–216 (2017).

198. Piliarik, M. & Sandoghdar, V. Direct optical sensing of single unlabelled proteins and super-resolution imaging of their binding sites. *Nat. Commun.* **5**, 1–8 (2014).

199. Redwan, I. N. & Grøtli, M. Method for activation and recycling of trityl resins. *J. Org. Chem.* **77**, 7071–7075 (2012).

200. Román, T. *et al.* Improving 2-Chlorotrityl Chloride (2-CTC) Resin Activation. *Methods Protoc.* **6**, 1–10 (2023).

201. Eissler, S. *et al.* Substitution determination of Fmoc-substituted resins at different wavelengths. *J. Pept. Sci.* **23**, 757–762 (2017).

202. Jiang, H., Cole, P. A. & Sciences, M. Protein Interactions Using Labeled Protein. **1**, 1–44 (2022).

7. Appendix

7.1. Table of figures

Figure 1: RNA splicing process. Pre-mRNA is transformed into spliced mRNA, in which introns are removed through two transesterification reactions: branching and exon ligation. Adapted from Wilkinson et al[3]. 1
Figure 2: Constitutive splicing and five different modes of alternative splicing: 1. Constitutively splicing 2. Intron retention 3. Exon skipping 4. Alternative 5' splice site 5. Alternative 3' splice site 6. Back splicing. Adapted from Ren et al[14]. 2
Figure 3: Alternative splicing regulatory factors and role of two splicing factor families (hnRNP and SR) in alternative splicing. The enhancer elements are recognized by activator proteins, and the silencer elements are bound by repressor proteins. These two protein families promote or inhibit spliceosome assembly at weak splice sites. Adapted from Wang et al[15]..... 2
Figure 4: RBPs are involved in controlling RNA life. They are involved in splicing, polyadenylation, and transcription, besides transport, translation, localization, and degradation of mRNA. Adapted from Gebauer et al[27]. 3
Figure 5: Developed methods to enhance the physicochemical characteristics of peptides. Adapted from Lee et al[58]. 7
Figure 6: Illustration of the amide cyclization. A) Side chain-to-side chain; B) Head-to-tail; C) Tail-to-side chain; D) Head-to-side chain cyclization. Adapted from Hayes et al[69]. 8
Figure 7: Ring closing metathesis scheme. X = amino acid, Y = CH_2, $(CH_2)_3$, CH_2SCH_2. Adapted from Bechtler et al[68]. 9
Figure 8: Principle of the phage display technology using M13 bacteriophage. M13 bacteriophage has a single-stranded DNA genome, and five different coat proteins. The DNA sequence that encodes for the peptide is inserted into the genome sequence, as a fusion to the gene which encodes for the phage coat protein (pIII). In the bacterial host gene expression starts and the peptide fragment is displayed on the phage surface as a combined product of genes which are involved in coat protein encoding. Adopted from Jaroszewicz et al[87]. 10
Figure 9: Synthesis scheme of an mRNA display peptide library. Adapted from Dougherty et al[91]. 11
Figure 10: Representation of the cyclic peptide generation mechanism by SICLOPPS. X = O or S. Adopted from Tavassoli et al[97]. 11
Figure 11: Structure of DNA and PNA. A: adenine, G: guanine, C: cytosine and T: thymine. Adapted from Wu et al[100]. 12
Figure 12: Allosteric regulation. Negative allosteric modulation (left) and positive allosteric modulation (right) is explained. In NAM the active site of the protein (blue) is altered by an allosteric inhibitor (red). Therefore, binding of the substrate (orange) is prohibited. In PAM (right panel), the affinity for the substrate (orange) enhances upon the allosteric activator (green) binding to the active site of the protein. Adapted from Duan Ni et al[103]. 13
Figure 13: Allosteric drugs approved by FDA. Adapted from Duan Ni et al[103]. 14
Figure 14: Image showing how DCM effects the normal heart. In normal heart (left) chambers relax and fill, then contract and pump. In the heart with DCM (right) muscle fibers have stretch, heart chambers enlarge. Adapted from "Medical gallery of Blausen Medical 2014". 15
Figure 15: The regulatory mechanism of the RBM20 binding to cis-regulatory elements in TTN. Adapted from Guo et al[119]..... 16
Figure 16: RBM20 protein structure, including functional domains. Adopted from Watanabe et al[120]. 16
Figure 17: Overlay of the structure of RNA (AUCUUA) bound and unbound to RBM20. Shown in beige is the RNA bound RRM domain of RBM20 (PDB: 6so9), in green the unbound RBM20 (PDB: 6soe). Helix 3 is indicated and makes a dramatic conformational shift upon RNA binding. Structures are generated with the mouse protein. Adapted from Upadhyay et al[122]..... 16

Figure 18: 1. Structure of RBM20 RRM domain in RNA unbound (green); 2. RNA bound states (tint); 3. together with stapled peptide (brown). 17
Figure 19: Example of peptide stapling corresponding to one staple in the peptide chain. Adopted from Moiola et al[76]. 18
Figure 20: A) Example of RCM with the non-native amino acids bearing α-olefins; B) Examples of (1) one i, i + 4 stapling; (2) one i, i + 7 stapling; (3) double i, i + 4 stapling. Adopted from Walensky et al[128]. 18
Figure 21: A) RBM20 loop in an unbound state shown in magenta and RBM20 helix 3, which is formed upon RNA binding shown in beige; B) Helix of RBM20 that were planned to be mimicked by stapled peptides. Stapling positions were shown in light blue. 19
Figure 22: Designed mimics of RBM20 helix 3. X represents the non-natural amino acid that was placed to perform hydrocarbon stapling. 19
Figure 23: Synthesis scheme of the RBM20 linear peptides on solid phase. Rink amide resin was used as a solid support. 20
Figure 24: A synthesis scheme of the stapled peptides on the solid phase. Rink amide low loading resin was used as a solid support. Fmoc-(S)-2-(4-pentenyl) Ala-OH (Fmoc-S_5-OH) was chosen as a non-native amino acid for stapling. 21
Figure 25: CD spectra of proteins with α-helix (red), β-sheet (blue) and random (pink) structures. Adapted from Sahoo et al [138]. 22
Figure 26: Representation of the CD spectra for acetylated RB1 (linear), RB2 (stapled) and RB3 (stapled) peptides. Buffer was used for a baseline correction. 23
Figure 27: Fluorescence polarization curves of R1 and R2 RNAs titrated with RBM20 α3 (511–619). Polarization was measured after a 60-minute incubation period and presented as an average of technical replicates (N=2). Poly A (AAAAAAA) RNA was chosen as a negative control. 24
Figure 28: Fluorescence polarization curves of R1 and R2 RNAs titrated with RBM20 Δα3 (511–601). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2). Poly A (AAAAAAA) RNA was employed as a negative control. 24
Figure 29: N-terminally FITC labeled peptides of derived from RBM20 helix 3. (X) represents the non-natural amino acid which was used to perform hydrocarbon stapling. 25
Figure 30: Fluorescence polarization curves of N-terminally FITC labeled peptides RB4, RB5, RB6 measured with RBM20 Δa3 (RBM20 without helix 3). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2). 26
Figure 31: C-terminally FITC labeled peptides of RBM20 protein. Fmoc-Lys(Mtt)-OH was loaded to the solid support for C-terminal labeling. 26
Figure 32: Fluorescence polarization curves of C-terminally FITC labeled peptides RB7, RB8 and RB9 measured with RBM20 Δa3 (RBM20 without helix 3). Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2). 27
Figure 33: Fluorescence polarization curves of peptides RB4, RB5, RB6 and R2 RNA measured with RBM20 Δa3 (RBM20 without helix 3). As a positive control, R2 RNA was used in this experiment. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2). 27
Figure 34: Competition FP curves of the 1nM of FAM labeled R1 and R2 RNAs with 60 nM RBM20 α3. Unlabeled RNAs were used as a competitor in this experiment. Polarization was measured after a 60-minute incubation period and presented as technical replicates (N=2). 28
Figure 35: Competition FP of 100 μM of RB1, RB2 and RB3 with 60 nM RBM20 α3. Polarization was measured after a 30-minute incubation and presented as technical replicates (N=2). 29
Figure 36: RB10 peptide, which is a modified analog of the RB1 to improve solubility. Modified position was shown in green. 30

Figure 37: Competition fluorescence polarization of 100 µM of RB10 and RB2 with 60 nM RBM20 α3. As a positive control, RB2 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2). 30
Figure 38: AIB containing RB11 peptide. X represents α-aminoisobutyric acid. 31
Figure 39: Competition fluorescence polarization of 100 µM RB2 and RB11 with 60 nM RBM20 α3 protein. As a positive control RB2 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2). 31
Figure 40: Representation of the CD spectra for RB10 and RB11. Buffer was used for a baseline correction. 32
Figure 41: Variants of the RB10 peptides. Point mutations are highlighted in red. 32
Figure 42: Competition FP of 100 µM of mutated linear peptides RB12-RB17 with 60 nM RBM20 α3. As a positive control, RB10 was used. After a 30-minute incubation period, polarization was measured and presented as technical replicates (N=2). 33
Figure 43: Principle of FP. A fluorophore is excited with light by passing through an excitation polarizing filter; the polarized fluorescence is measured either parallel or perpendicular to the exciting light's plane of polarization. Two intensity measurements are obtained ($I_{\perp}$ and $I_{\parallel}$) and used for the calculation of fluorescence polarization. Adopted from Lea and Simeonov et al[147]. 34
Figure 44: A) Competition FP and total intensity of RB1, RB2 and RB3 with RBM20 α3 (see Figure 35); B) Competition FP of R2 RNA (see Figure 34) with RBM20 α3 and total intensity change with increasing RNA concentration; C) Direct FP of RB4, RB5 and RB6 with RBM20 Δα3 (see Figure 33) and total intensity change with an increasing peptide concentration. 35
Figure 45: Domain construction of the different hnRNP proteins. Adapted from Geuens et al[152]. 36
Figure 46: hnRNPs that are involved in cancer and metastasis. hnRNP C and A1 are acting on the promotor of BRCA and KRAS, HRAS, respectively[157,158]. hnRNP E1 is known to be involved in the regulation of mRNA half-life and mRNA stability has a relation with the growth of cancer[159–161]. The recent discoveries on hnRNP E1 which plays a key role in Transforming growth factor beta (TGF-β)-mediated epithelial-mesenchymal transition (EMT) has clarified that hnRNP E1 is regulating oncogenes such as Dab-2 and ILEI[162]. Adapted from Geuens et al[152]. 37
Figure 47: Domain structure represented based on the hnRNP B1 isoform. Adapted from Wu et al[172]. 37
Figure 48: Representation of RRM1 domain of the protein bound to AAGG RNA. Adapted from Wu et al[172]. 38
Figure 49: Mimicking RNA oligonucleotide by peptide nucleic acid where the phosphodiester backbone of RNA was replaced by a peptide backbone. Adopted from Wu et al[100]. 40
Figure 50: Synthesis scheme of the 4mer PNA (AAGG) referring to the solid phase-based synthesis protocol[129]. 41
Figure 51: Direct FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 3'FAM labeled AGGACUGC (8-mer) in blue and AAGGACUAGC (10-mer) in green. As a control, FAM labeled polypyrimidine was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2). 41
Figure 52: Direct FP curves of N-terminal MBP tagged hnRNP A2/B1(residues 1-251) with 3'FAM (dark green) and 5'FAM (light green) labeled 10mer RNA (AAGGACUAGC). As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation period and presented as a technical replicate. 42
Figure 53: Competition FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 1 nM 10mer 3'FAM RNA (AGGACUGC). 10 µM of unlabeled 10mer RNA was used as a

competitor. As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2). 43
Figure 54: Competition FP curves of N-terminal MBP tagged hnRNP A2/B1 (residues 1-251) with 1 nM 8-mer 3'FAM RNA (AAGGACUAGC). 10 µM of unlabeled 8mer RNA was used as a competitor. As a negative control, poly C was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates (N=2). 43
Figure 55: Direct FP curves of hnRNP A2/B1 constructs, 1-104 (RRM1), 12-195 (RRM1 and RRM2), 1-251 (construct containing RGG box: P1 and P2) with FITC labeled 4mer PNA (AAGG). During the protein purification hnRNP A2/B1 (residues 1-251) was obtained as two separate fractions (possibly monomer and dimer) and both were tested. Measurements were performed as technical replicates (N=2). 44
Figure 56: Synthesis scheme of the PNA monomer (Fmoc-PNA-U-OH) needed for 10mer and 8mer PNA synthesis in solid phase. 45
Figure 57: Representation of 8mer and 10mer PNA sequences: A) 8-mer PNA sequence; B) 10-mer PNA sequence. 45
Figure 58: Structures of the selected mononucleotides represented from 5 to 10. Guanosine 3'-monophosphate (9) was not commercially available and therefore, was not involved in further experiments. 46
Figure 59: Structure of the synthesized PNA variants having adenine (11) and guanine (12) on the side chain. The synthesis was carried out in solid phase. 46
Figure 60: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: adenosine (5) adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7) and PNA variant (11) against the 5 nM hnRNP A2/B1 interaction with FAM labelled 10mer RNA. As a positive control unlabeled 10mer RNA was used. Polarization was measured after a 30-minute incubation period and presented as technical replicates. 47
Figure 61: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: adenosine 3'-monophosphate (6), adenosine 5'-monophosphate (7) against the 5 nM hnRNP A2/B1 interaction with FAM labeled 10mer RNA. Mononucleotides were used in mM concentration range. As a positive control, unlabeled 10mer RNA was used and shown as technical replicates. 48
Figure 62: Competition FP curves of the unlabelled 10mer RNA and mononucleotides: guanosine (8), guanosine 5'-monophosphate (10) and PNA variant (12) against the 5 nM hnRNP A2/B1 interaction with FAM labeled 10mer RNA. Unlabeled compounds were used in mM concentration range. Unlabeled 10mer RNA was used as a positive control and shown as technical replicates. 48
Figure 63: Representation of the SICLOPPS peptide library generation and screening. The SICLOPPS plasmids and TRAP assays plasmids can be transformed into E. coli cells including screening assay and screened by FACS. SICLOPPS plasmids with a desired phenotype can be isolated from cells and sequenced to identify the macrocyclic peptide. Adopted from Tavassoli, 2017[139]. 49
Figure 64: Synthesis scheme of the SICLOPPS peptides. 50
Figure 65: Fluorescence polarization curves of 3'Cy5 labeled RNA1 with hnRNP A2/B1 (1-251) measured after 30 minutes of incubation. n1 and n2 represents two biological replicates, respectively. 51
Figure 66: Fluorescence polarization curve of 3'Cy5 labeled RNA1 and Poly C RNA with hnRNP A2/B1 (residues 1-251). As a negative control, Poly C was used. Polarization was measured after a 30-minute incubation and shown as technical replicates. 52
Figure 67: Fluorescence polarization curves of 5'FAM labeled RNA2 and Poly C RNAs with hnRNP A2/B1 (residues 1-251). Polarization was measured after a 30-minute incubation

period. n1 and n2 represents two biological replicates, respectively. As a control, 5'FAM Poly C RNA was utilized. 53
Figure 68: Curves of 13, 14 peptides and unlabeled RNA1 from the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was was used in this experiment. Polarization was measured after a 60-minute incubation and shown as a technical replicate (N=2). The measurements were performed at the same time on the same plate however the graphs are shown separately for easier interpretation. 54
Figure 69: Curves of 15, 16 peptides and unlabeled RNA1 the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was used and DMSO was used as a negative control in this experiment. A4 was dissolved in DMSO and diluted with buffer to a final DMSO concentration of 1%. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2). 54
Figure 70: Curves of 17-1, 17-2, 18 peptides and unlabeled RNA1 from the competition FP of 25 nM of hnRNP A2/B1. As a positive control, unlabeled RNA was used. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2)..... 55
Figure 71: Competition FP curves of peptides 17-1, 19-1, 19-2, 20-2 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 60-minute incubation period, polarization was measured and shown as a technical replicate (N=2). The experiment was carried out at once in the same plate but are shown separately for better interpretation. 57
Figure 72: Competition FP curves of peptides 17-1, 21, 22 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 60-minute incubation period, polarization was measured and shown as a technical replicate (N=2). 57
Figure 73: Competition FP curves of peptides 17-1, 23, 24 and unlabeled RNA1 against the 25 nM of hnRNP A2/B1 interaction with 1 nM of FAM labeled RNA2. As a positive control, 17-1 and unlabeled RNA were used. After a 30-minute incubation period, polarization was measured and shown as a technical replicate (N=2). 58
Figure 74: Competition FP curves of peptides 17-1, 19-1, 19-2, 20-2 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2). The measurements were performed at the same time on the same plate however the graphs are shown separately for easier interpretation. 59
Figure 75: Competition FP curves of peptides 17-1, 21, 22 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2)..... 59
Figure 76: Competition FP curves of peptides 17-1, 23, 24 and unlabeled RNA1 against the 35 nM of hnRNP A2/B1 interaction with 1 nM of Cy5 labeled RNA1. 17-1 and unlabeled RNA were used as positive control in this experiment. Polarization was measured after a 60-minute incubation period and shown as a technical replicate (N=2)..... 60
Figure 77: MST curve indicating the interaction between 17-1 and hnRNP A2/B1 (residues 1-251) protein. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates. . 62
Figure 78: MST curve indicating the interaction between 24 and hnRNP A2/B1 (residues 1-251) protein. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates..... 62
Figure 79: MST curve of the peptide 24 and hnRNP A2/B1 RRM1(residues 1-104) domain. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates..... 63

Figure 80: MST curve of peptide 24 and hnRNP A2/B1 RRM2 (residues 112-195) domain. RED fluorescent dye NT-647-NHS was used to label the protein. Measurement took place after 30 minutes of incubation and the graphs represent two biological replicates. 63
Figure 81: Representation of the RRM domains of the hnRNP A2/B1 in complex with RNA1 (AAGGACUAGC). RNA1 accommodated into a (+) charged groove of the RRM domains of two hnRNP A2/B1. Adopted from Wu et al[172]. 64
Figure 82: A) glass cover slip with (right, blue) and without sample (left), B) Label-free detection by imaging the scattered and reflected light over time. b) Scatter plot of single-molecule contrast and mass distribution for a mixture of monomer/dimer. Adopted from Soltermann et al[196]. 65
Figure 83: Mass photometry data observed from the measurement for hnRNP A2/B1 protein alone. 65
Figure 84: Mass photometry data observed from the measurement for hnRNP A2/B1 protein with Cy5 labeled 10mer RNA. The protein concentration was 50 nM protein and the RNA concentration 25 nM. The measurement was performed after 30 min incubation at rt. 66

7.2. Table of tables

Table 1: K_D values obtained from ITC measurements of RBM20 with RNAs with various point mutations (underlined nucleotide). Adapted from Upadhyay et al[122]. For mRBM20 Δα3 the α3 helical region has been removed. 17
Table 2: Percentage helicity of the peptides calculated based on the ellipticity values obtained from CD spectroscopy measurements. 23
Table 3: K_D values of the R1 and R2 RNAs for their binding affinity to the RBM20 protein with and without helix. 24
Table 4: K_D values of the FITC-labeled peptides RB4, RB5, RB6 and the R2 RNA from the FP assay of the RBM20 Δa3 protein. 28
Table 5: IC_{50} values calculated from the competition assay of RBM20 α3 with RNAs. RNA sequences are shown in section 2.4.5. 29
Table 6: IC_{50} values of the RBM20 peptides calculated from the competition FP assay of RBM20 α3 and FAM labeled R2 RNA. 29
Table 7: IC_{50} values of the RBM20 peptides calculated from the competition FP assay of RBM20 α3 and FAM labeled R2 RNA. 31
Table 8: The percentage helicity of the peptides calculated based on the ellipticity values obtained from CD spectroscopy measurement. 32
Table 9: K_D values for the 8-mer and 10mer RNAs tested in FP assay for binding affinity with hnRNP A2/B1. 42
Table 10: K_D values for 3'FAM and 5'FAM labeled 10mer RNA tested in FP assay for binding affinity with hnRNP A2/B1. 42
Table 11: Competition FP derived IC_{50} values for the 8mer and 10mer RNAs. 44
Table 12: Cyclic peptides derived from the SICLOPPS screening of hnRNP A2/B1. 50
Table 13: Averaged K_D values for the 3'Cy5 labeled RNA1 tested in FP assay for binding affinity with hnRNP A2/B1. 52
Table 14: Averaged K_D values for the 5'FAM labeled RNA2 tested in FP assay for binding affinity with hnRNP A2/B1. 53
Table 15: Peptides from the hnRNP A2/B1 SICLOPPS screening. 53
Table 16: IC_{50} values and calculated K_I values derived from competition FP of the peptides tested for their inhibitory properties for the interaction of RNA with hnRNP A2/B1. 55
Table 17: Peptides derived from the alanine scanning of A5-1. Point mutations are marked in red. 56

Table 18: IC_{50} and K_I values calculated from alanine scanning of peptides evaluated against hnRNP A2/B1-FAM labeled 10merggg10mer RNA interaction. 58
Table 19: IC_{50} and K_I values calculated for the peptides from alanine scanning evaluated against hnRNP A2/B1-Cy5 labeled 10mer RNA interaction. 60
Table 20: Dissociation constant (K_D) of the peptides with hnRNP A2B1 protein from the MST experiment. 62
Table 21: Dissociation constant (K_D) of the peptide 24 with individual hnRNP A2B1 RRM domains as measured by the MST 63

7.3. Table of supplementary

Table of supplementary figures

Supplementary figure 1: Unlabeled RB1, RB2 and RB3 peptide sequences, purity and calculated/determined masses. 106
Supplementary figure 2: N-terminally FITC labeled RB4, RB5 and RB6 peptide sequences, purity and calculated/determined masses. 107
Supplementary figure 3: C-terminally FITC labeled RB7, RB8 and RB9 peptide sequences, purity and calculated/determined masses. 109
Supplementary figure 4: A to R modified RB10 and AIB containing RB11 peptide sequences, purity and calculated/determined masses. 110
Supplementary figure 5: Sequences, purity and calculated/determined masses of point mutated linear peptides. 111
Supplementary figure 6: Structures, purity and calculated/determined masses of unlabeled and FITC labeled 4mer PNAs. 115
Supplementary figure 7: Structure, purity and calculated/determined masses of PNA monomer. 116
Supplementary figure 8: Structures, purity and calculated/determined masses of PNA amide variants. 116
Supplementary figure 9: Purity data of macrocyclic peptides selected from SICLOPPS screening. 118
Supplementary figure 10: Purity data of the macrocyclic peptides from alanine scanning of A5-1 121
Supplementary figure 11: Mass photometry data observed from the measurement of MBP-tagged hnRNP A2/B1 (residues 1-251, MW: 70.306 kDa) protein with RNA1. 50 nM protein was mixed with 50 nM RNA. The measurement was taken after 30 min incubation at rt. 125
Supplementary figure 12: Mass photometry data observed from the measurement of MBP-tagged hnRNP A2/B1 protein (residues 1-251, MW: 70.306 kDa) with RNA1. 50 nM protein was mixed with 12.5 nM RNA. The measurement was taken after 30 min incubation at rt. 125
Supplementary figure 13: CD chromatogram of peptide mutations of stapled peptides. 126

Table of supplementary tables

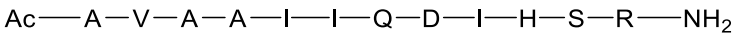
Supplementary table 1: RBM20 RNAs 125
Supplementary table 2: hnRNP A2/B1 RNAs 125
Supplementary table 3: Percentage of helicity of mutated stapled peptides. 126

7.4. Supplementary figures

7.4.1. RBM20 peptide sequences, masses (calculated and determined) and purities

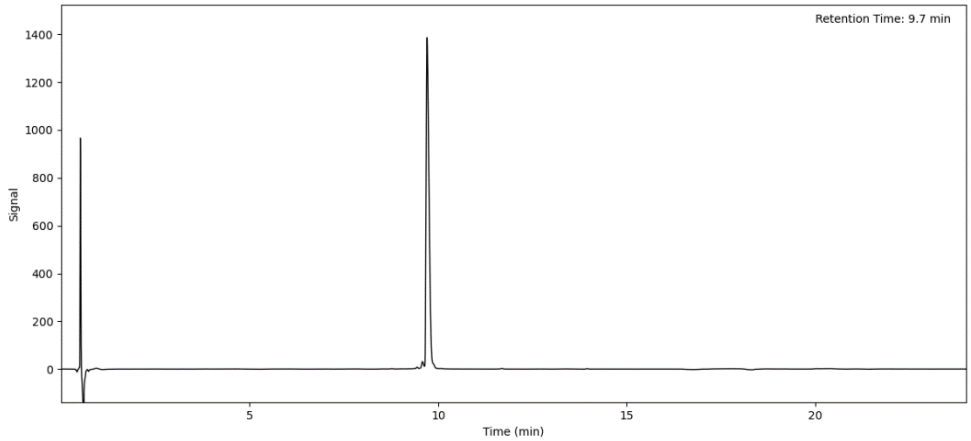
Supplementary figure 1: Unlabeled RB1, RB2 and RB3 peptide sequences, purity and calculated/determined masses.

RB1

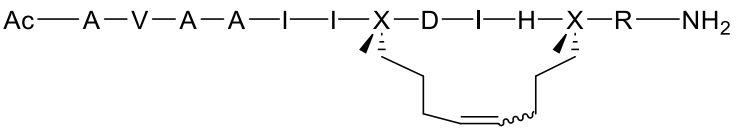


HRMS ESI: m/z [C58H99N19O17]

Calculated: 1333.7466 $[M+H]^+$, measured: 668.88317 $[M+2H]^{2+}$.



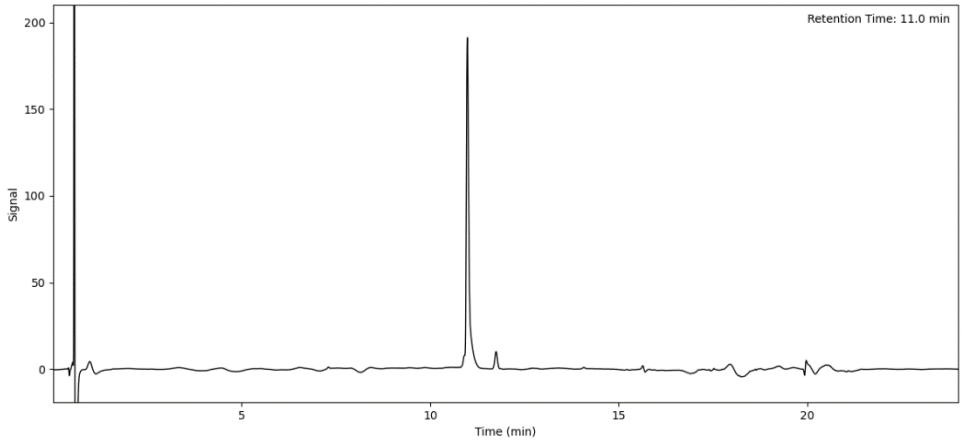
RB2

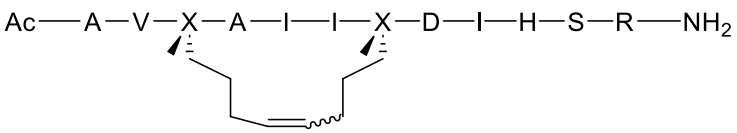


HRMS ESI: m/z [C64H108N18O15]

Calculated: 1368.8242 $[M+H]^+$, measured: 685.41956 $[M+2H]^{2+}$.

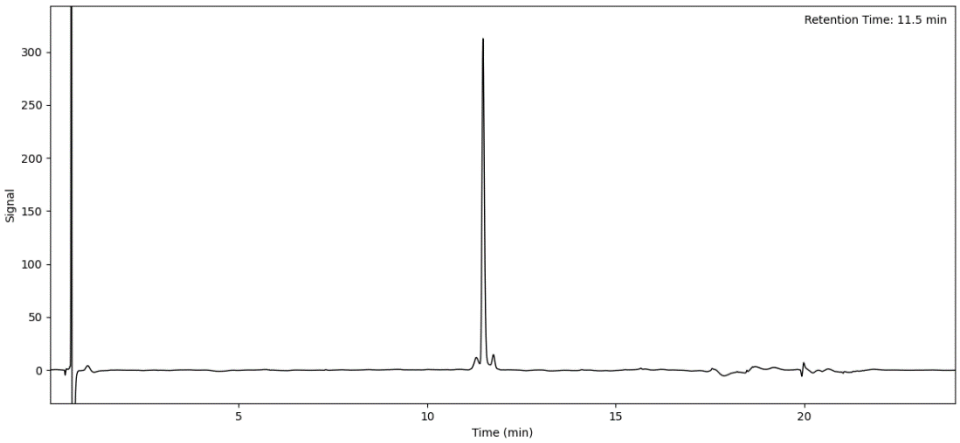
RB3





HRMS ESI: m/z [C64H108N18O16]

Calculated: 1384.8191 $[M+H]^+$, measured: 693.41705 $[M+2H]^{2+}$.



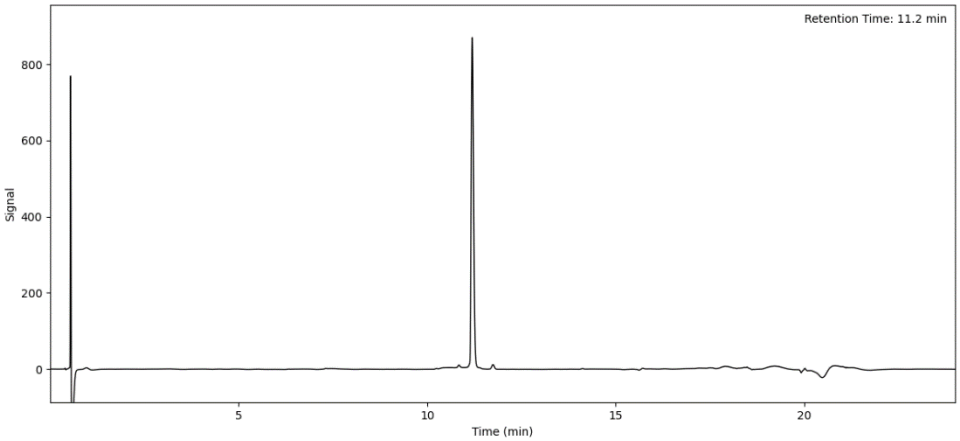
Supplementary figure 2: N-terminally FITC labeled RB4, RB5 and RB6 peptide sequences, purity and calculated/determined masses.

RB4

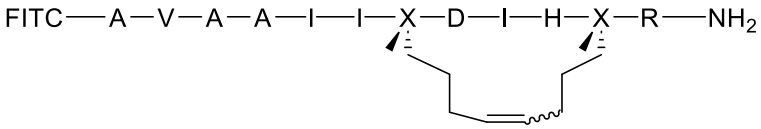
FITC—A—V—A—A—I—I—Q—D—I—H—S—R—NH_2

HRMS ESI: m/z [C83H121N21O24S]

Calculated: 1827.8614 $[M+H]^+$, measured: 914.93326 $[M+2H]^{2+}$.

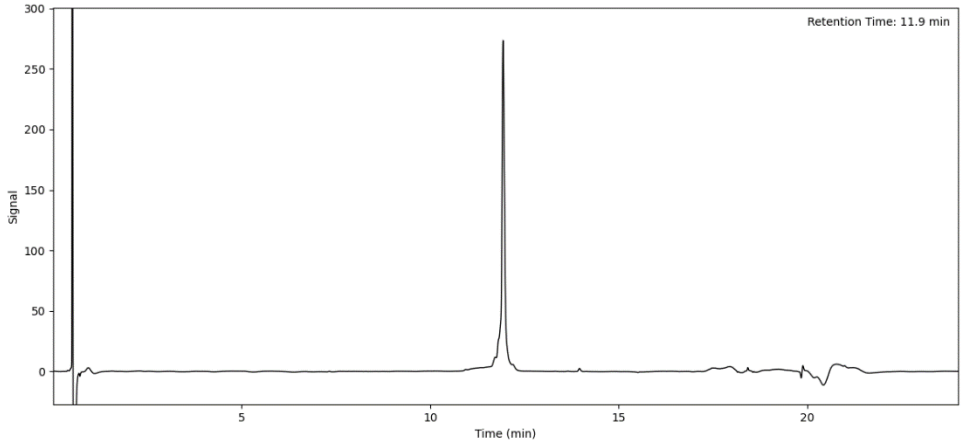


RB5

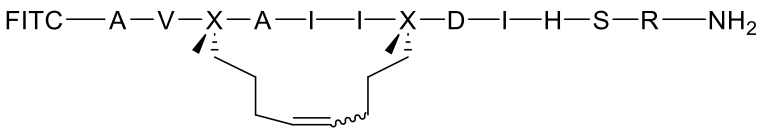


HRMS ESI: m/z [C89H128N20O22S]

Calculated: 1860.9233 $[M+H]^+$, measured: 931.46958 $[M+2H]^{2+}$.

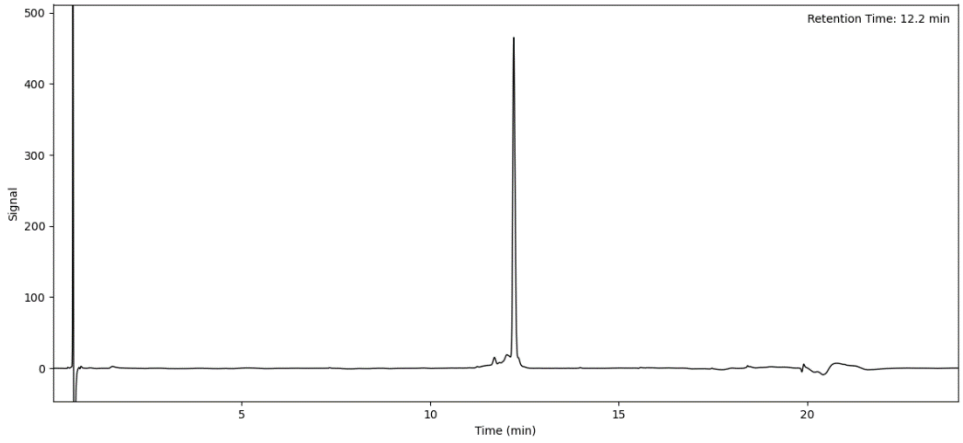


RB6



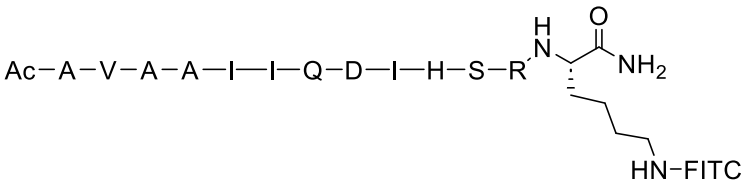
HRMS ESI: m/z [C89H128N20O23S]

Calculated: 1876.9182 $[M+H]^+$, measured: 939.46700 $[M+2H]^{2+}$.



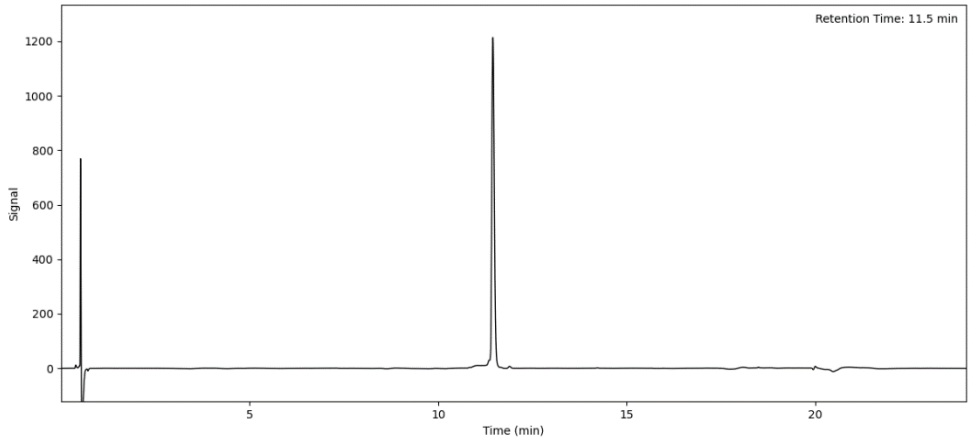
Supplementary figure 3: C-terminally FITC labeled RB7, RB8 and RB9 peptide sequences, purity and calculated/determined masses.

RB7

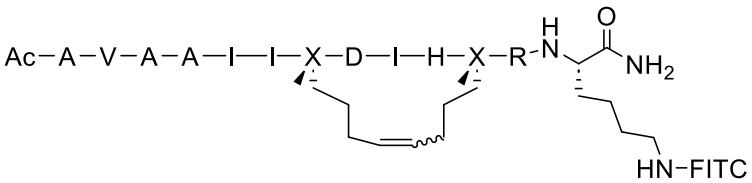


HRMS ESI: m/z [C85H122N22O23S]

Calculated: 1850.8774 $[M+H]^+$, measured: 617.96586 $[M+3H]^{3+}$.

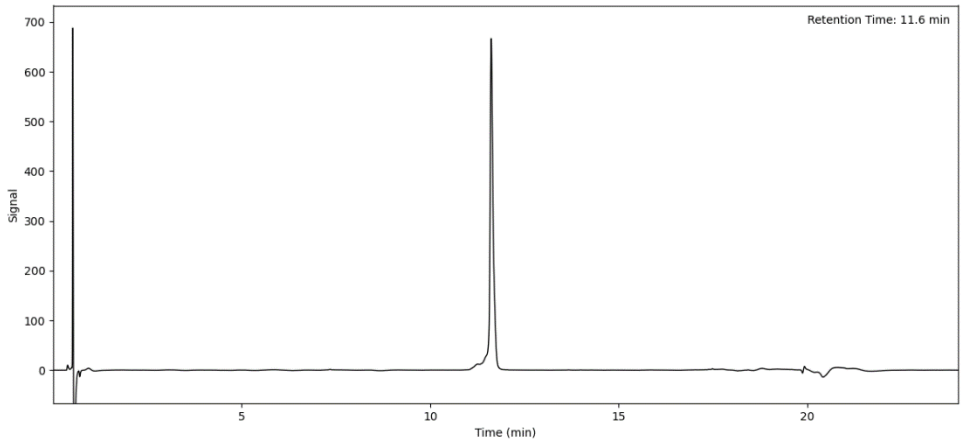


RB8

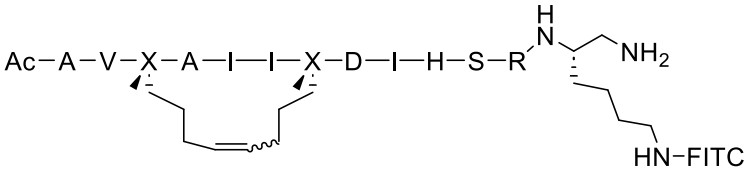


HRMS ESI: m/z [C91H131N21O21S]

Calculated: 1885.9549 $[M+H]^+$, measured: 943.98479 $[M+3H]^{2+}$.

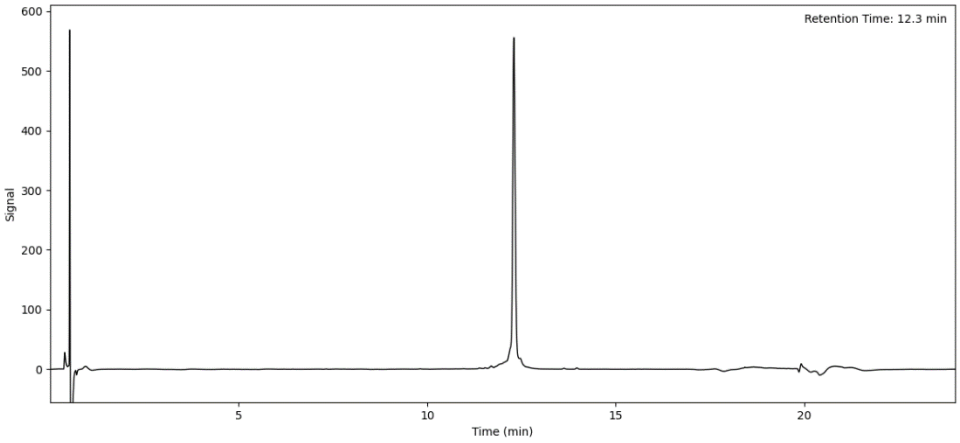


RB9



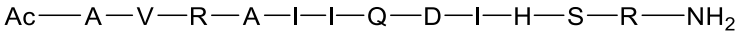
HRMS ESI: m/z [C91H131N21O22S]

Calculated: 1901.9498 $[M+H]^+$, measured: 634.99046 $[M+3H]^{3+}$.



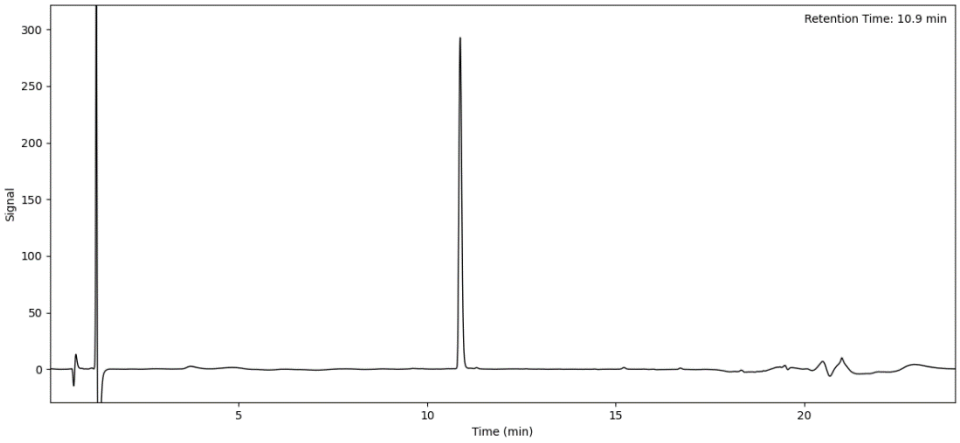
Supplementary figure 4: A to R modified RB10 and AIB containing RB11 peptide sequences, purity and calculated/determined masses.

RB10



HRMS ESI: m/z [C61H108N22O17]

Calculated: 1420.8263 $[M+H]^+$, measured: 710.41466 $[M+3H]^{2+}$, 473.94524 $[M+3H]^{3+}$.



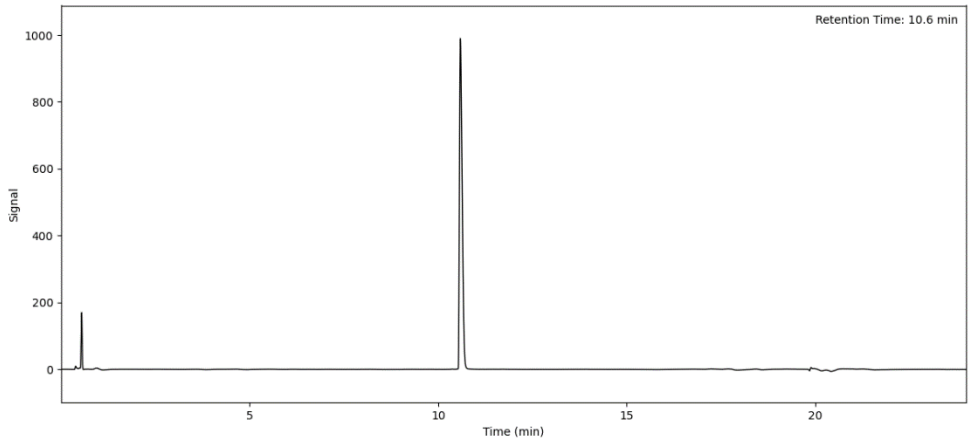
RB11

Ac—A—V—X—A—I—I—X—D—I—H—X—R—NH_2

X= Aib

HRMS ESI: m/z [C59H102N18O15]

Calculated: 1302.7772 $[M+H]^+$, measured: 652.39673 [M+3H]2+.



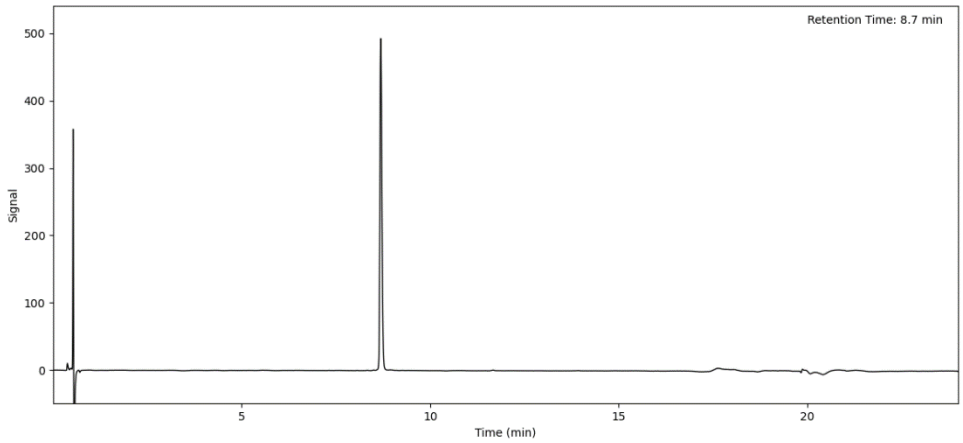
Supplementary figure 5: Sequences, purity and calculated/determined masses of point mutated linear peptides.

RB12

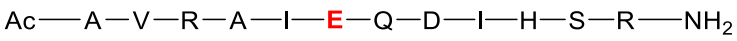
Ac—A—V—R—A—I—I—Q—**N**—I—H—S—R—NH_2

HRMS ESI: m/z [C61H107N23O16]

Calculated: 1417.8266 $[M+H]^+$, measured: 709.92258 $[M+3H]^{2+}$, 473.61718 $[M+3H]^{3+}$.

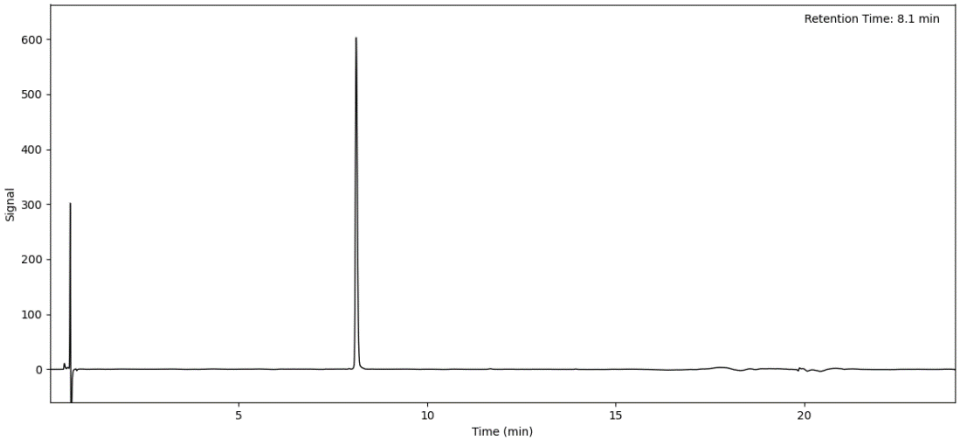


RB13

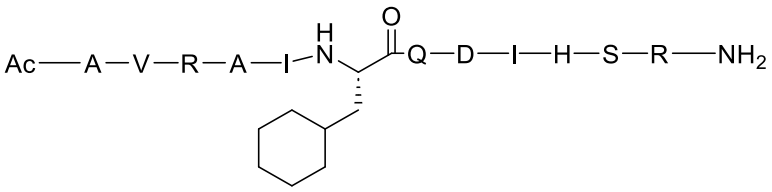


HRMS ESI: m/z [C60H102N22O19]

Calculated: 1434.7692 $[M+H]^+$, measured: 718.39328 $[M+3H]^{2+}$, 479.26442 $[M+3H]^{3+}$.

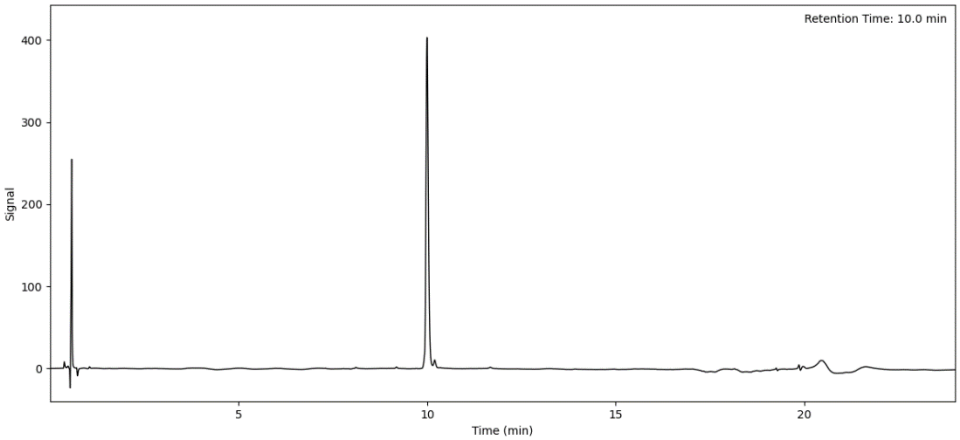


RB14



HRMS ESI: m/z [C64H110N22O17]

Calculated: 1458.8419 $[M+H]^+$, measured: 730.43027 $[M+3H]^{2+}$, 487.28893 $[M+3H]^{3+}$.

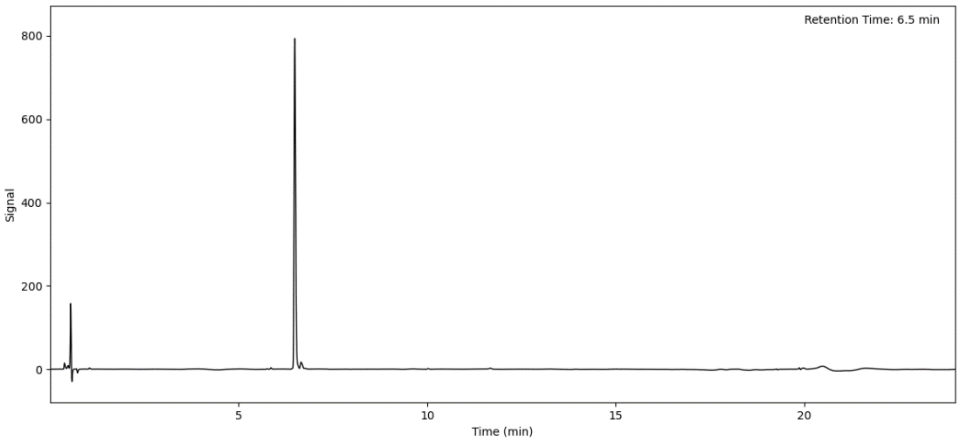


RB15

Ac—A—V—R—A—**K**—I—Q—D—I—H—S—R—NH_2

HRMS ESI: m/z [C61H107N23O17]

Calculated: 1433.8215 $[M+H]^+$, measured: 717.92006 $[M+3H]^{2+}$, 478.94867 $[M+3H]^{3+}$.

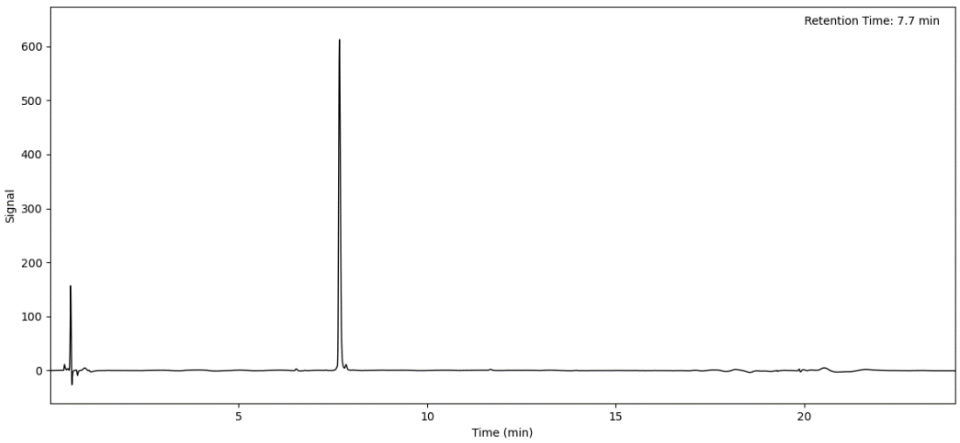


RB16

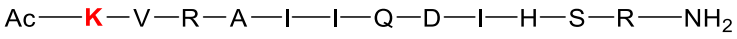
Ac—A—**T**—R—A—I—I—Q—D—I—H—S—R—NH_2

HRMS ESI: m/z [C60H104N22O18]

Calculated: 1420.7899 $[M+H]^+$, measured: 711.40424 $[M+3H]^{2+}$, 474.60496 $[M+3H]^{3+}$.

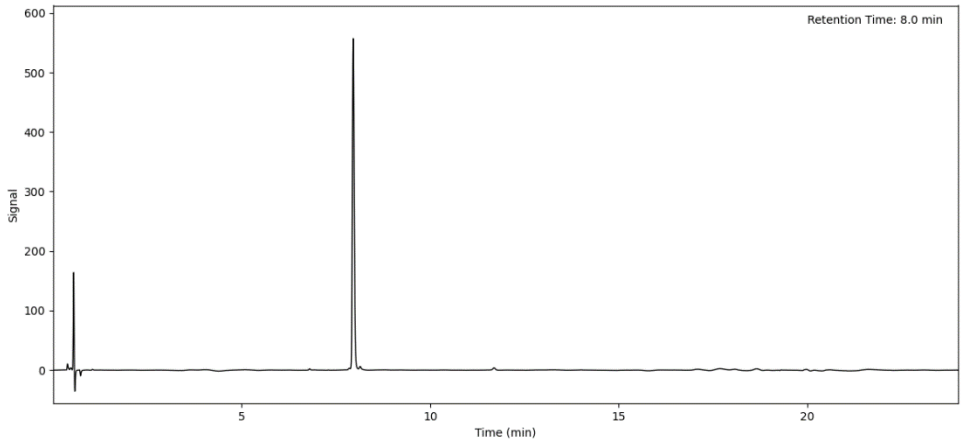


RB17



HRMS ESI: m/z [C64H113N23O17]

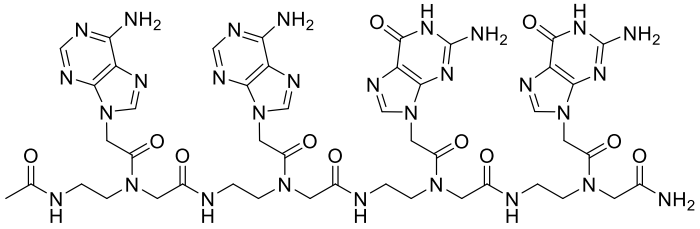
Calculated: 1475.8685 $[M+H]^+$, measured: 738.94402 $[M+3H]^{2+}$, 492.96474 $[M+3H]^{3+}$.



7.4.2. PNA structures, masses (calculated and determined) and purity data

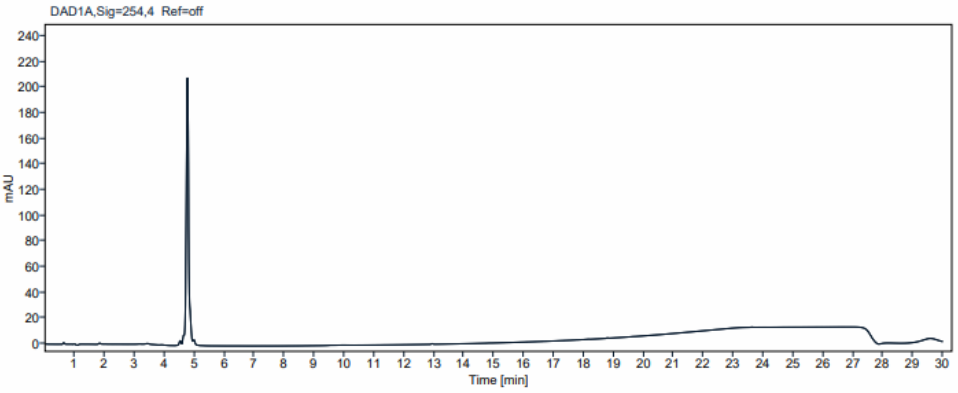
Supplementary figure 6: Structures, purity and calculated/determined masses of unlabeled and FITC labeled 4mer PNAs.

Unlabeled PNA

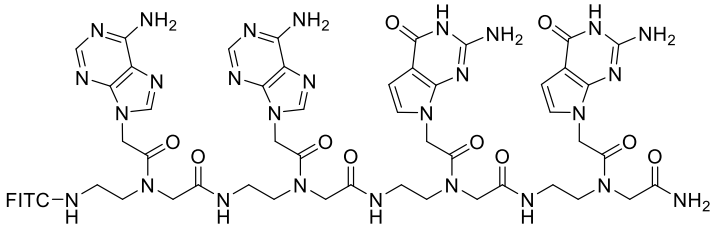


HRMS ESI: m/z [C46H57N29O11]

Calculated: 1191.4792 $[M+H]^+$, measured: 1192.48628.

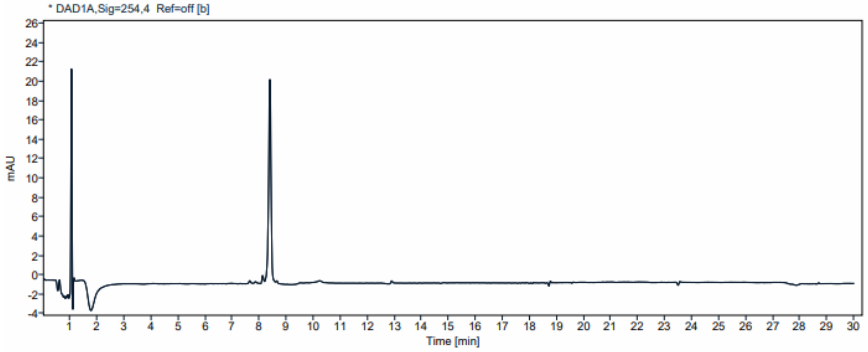


FITC-labeled PNA



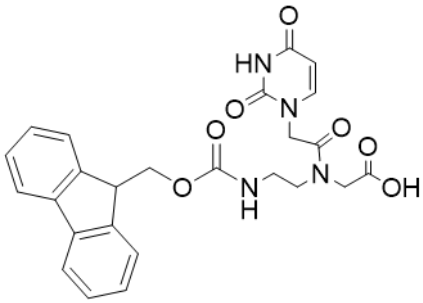
HRMS ESI: m/z [C71H77N31O18S]

Calculated: 1683.5784 $[M+H]^+$, measured: 842.79627 $[M+2H]^{2+}$.



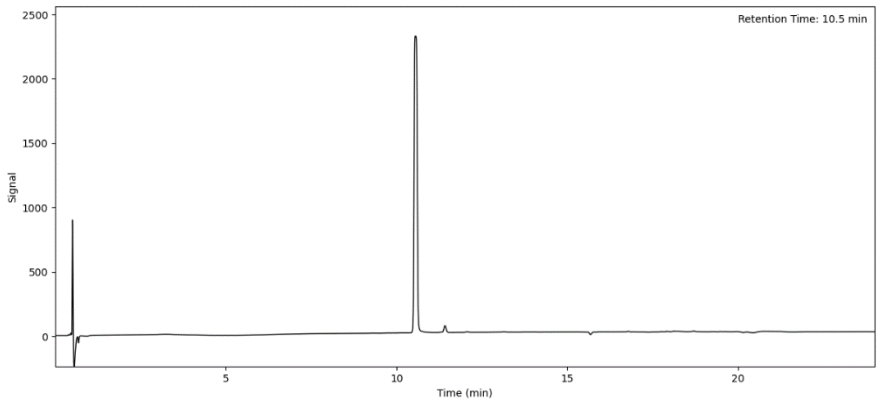
Supplementary figure 7: Structure, purity and calculated/determined masses of PNA monomer.

Compound 4 (Fmoc-PNA-U-OH)



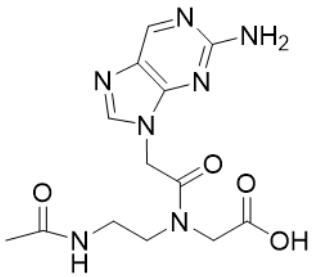
HRMS ESI: m/z [C25H24N4O7]

Calculated: 492.1645 $[M+H]^+$, measured: 493.17215.



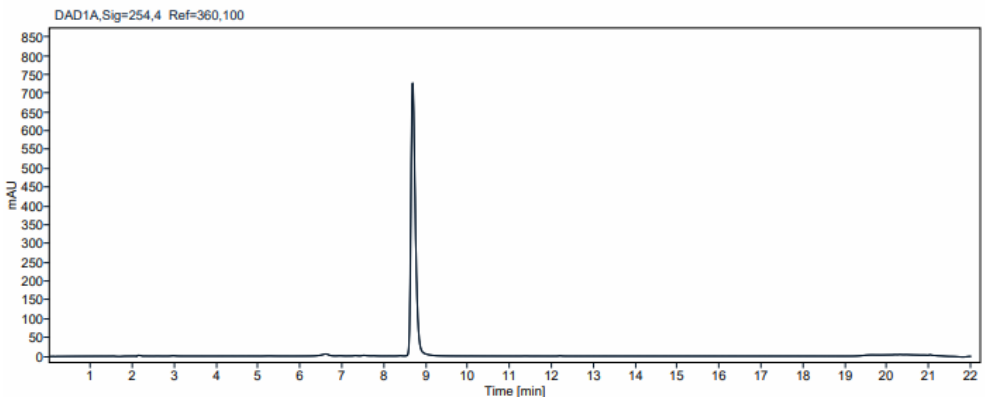
Supplementary figure 8: Structures, purity and calculated/determined masses of PNA amide variants.

Compound 11

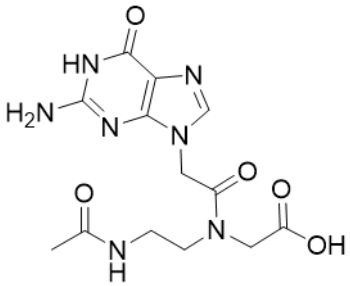


HRMS ESI: m/z [C13H17N7O4]

Calculated: 335.1342 $[M+H]^+$, measured: 336.14163.

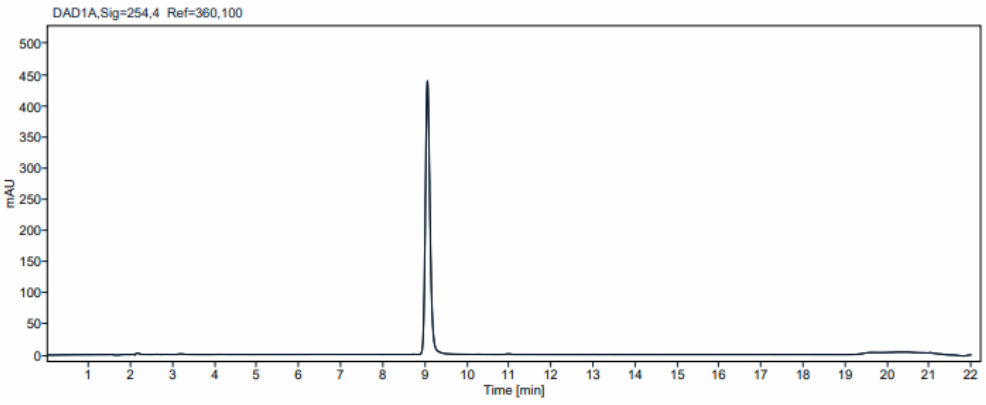


Compound 12



HRMS ESI: m/z [C13H17N7O5]

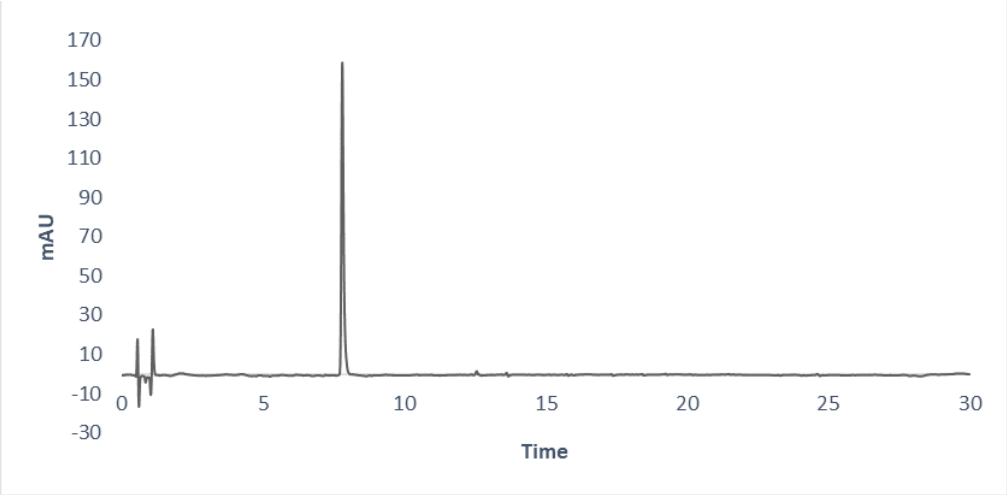
Calculated: 351.1291 $[M+H]^+$, measured: 352.13655.



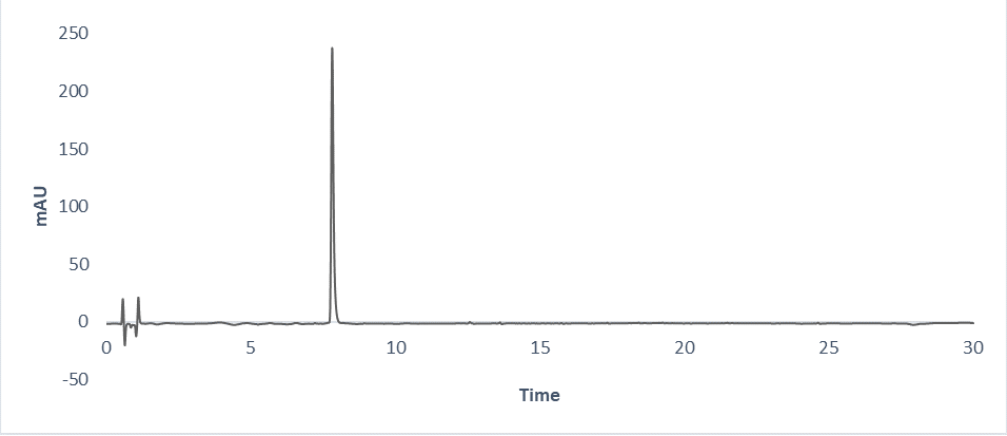
7.4.3. hnRNP A2/B1 peptide purity data

Supplementary figure 9: Purity data of macrocyclic peptides selected from SICLOPPS screening.

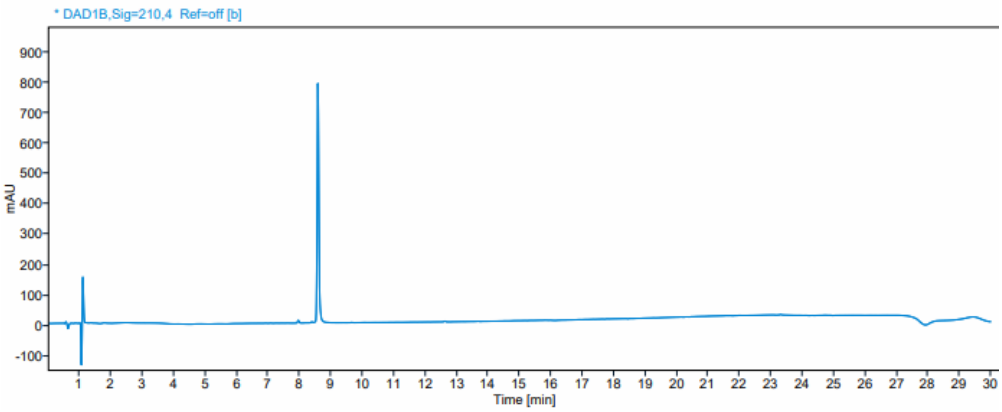
13



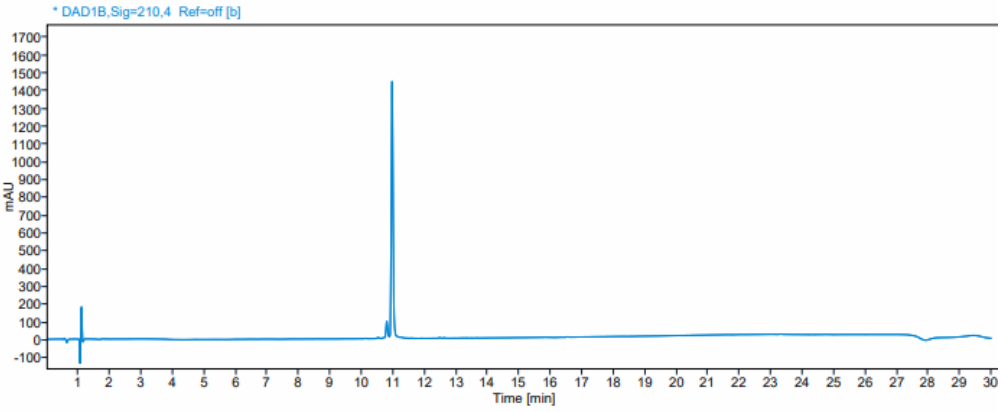
14



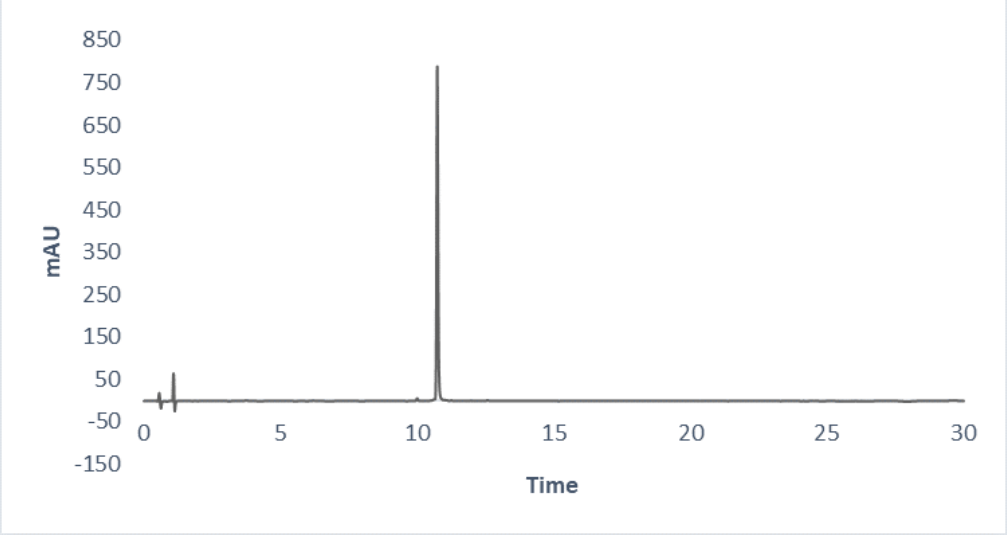
15


* DAD1B,Sig=210,4 Ref=off [b]
mAU
Time [min]

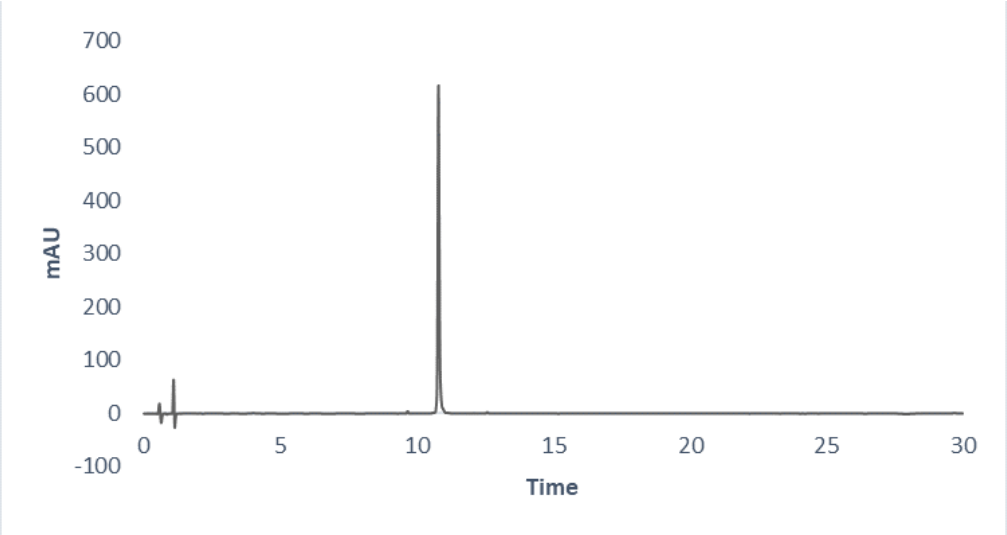
16


* DAD1B,Sig=210,4 Ref=off [b]
mAU
Time [min]

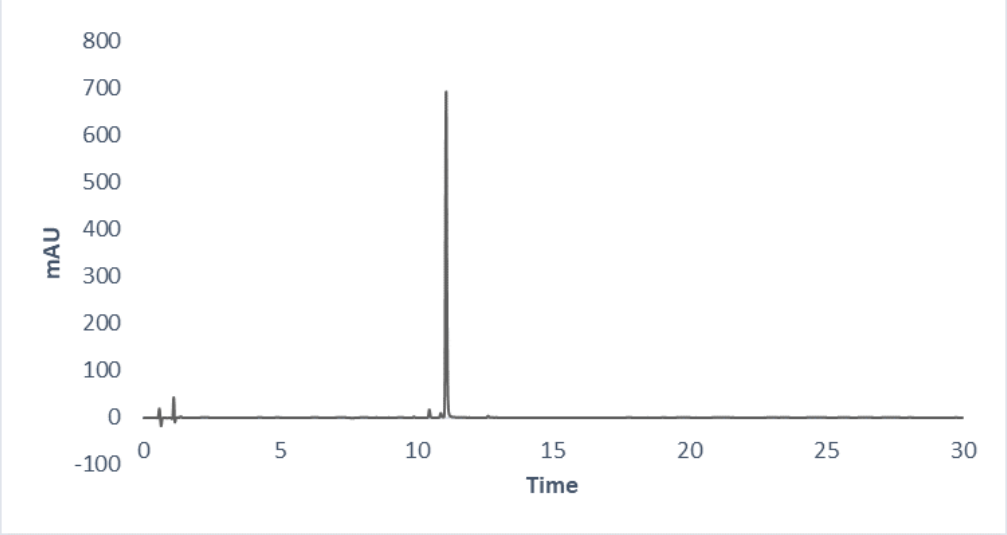
17-1


850
750
650
550
450
350
250
150
50
-50
-150
mAU
0
5
10
15
20
25
30
Time

17-2

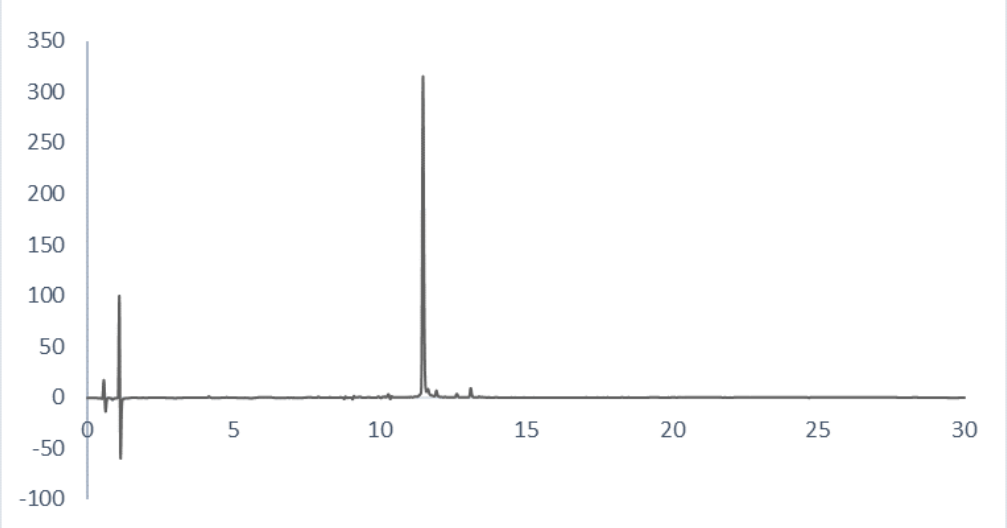

700
600
500
400
300
200
100
0
-100
mAU
0
5
10
15
20
25
30
Time

18

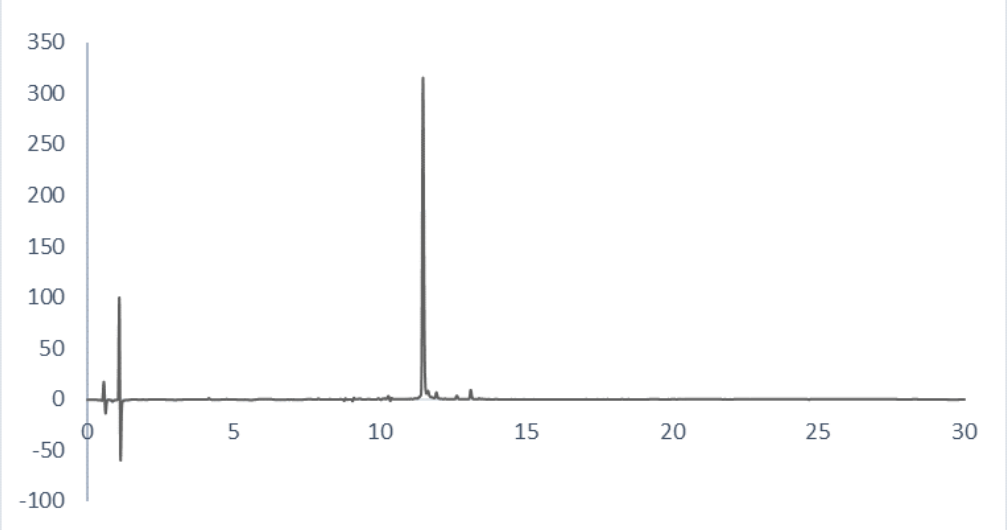


Supplementary figure 10: Purity data of the macrocyclic peptides from alanine scanning of A5-1.

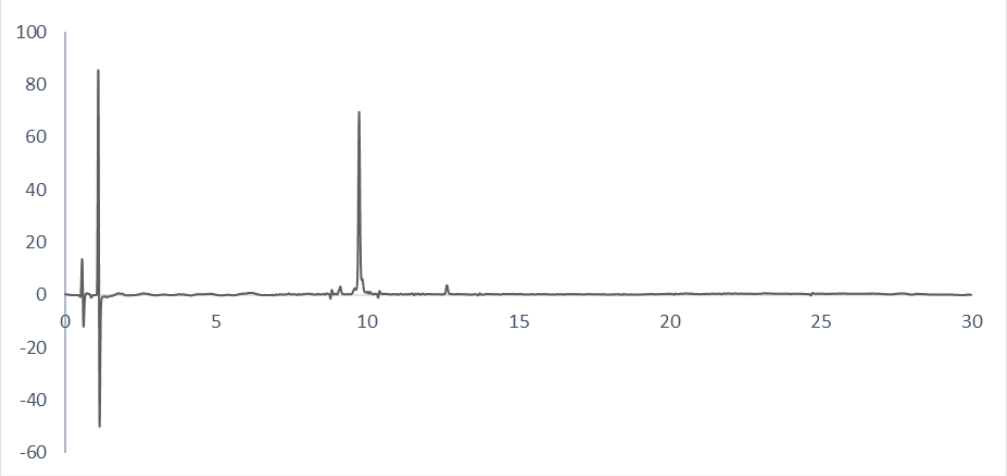
19-1



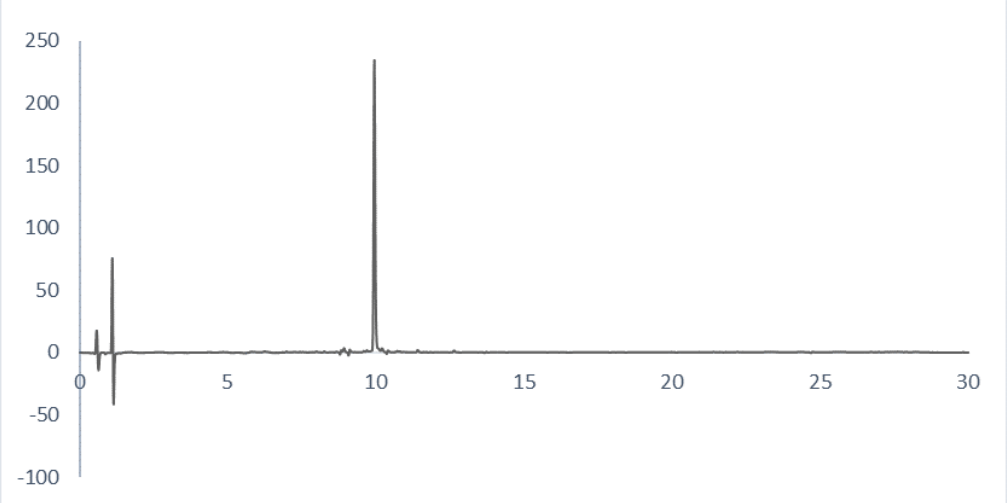
19-2


350
300
250
200
150
100
50
0
-50
-100
0
5
10
15
20
25
30

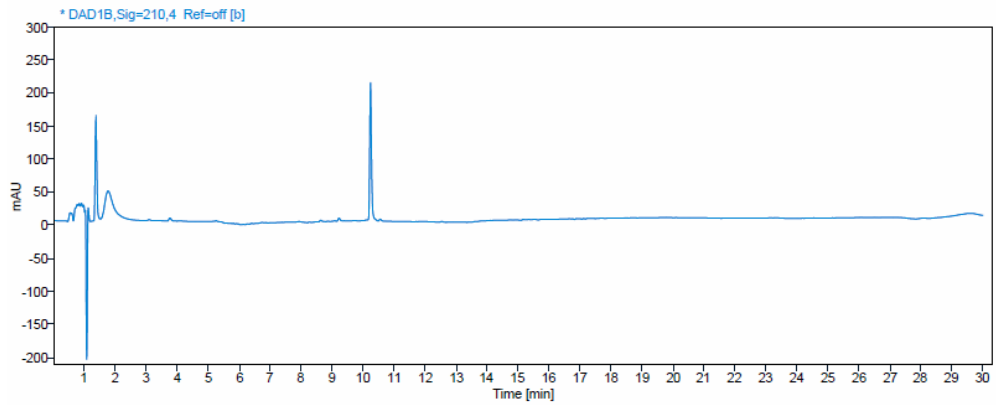
20-1


100
80
60
40
20
0
-20
-40
-60
0
5
10
15
20
25
30

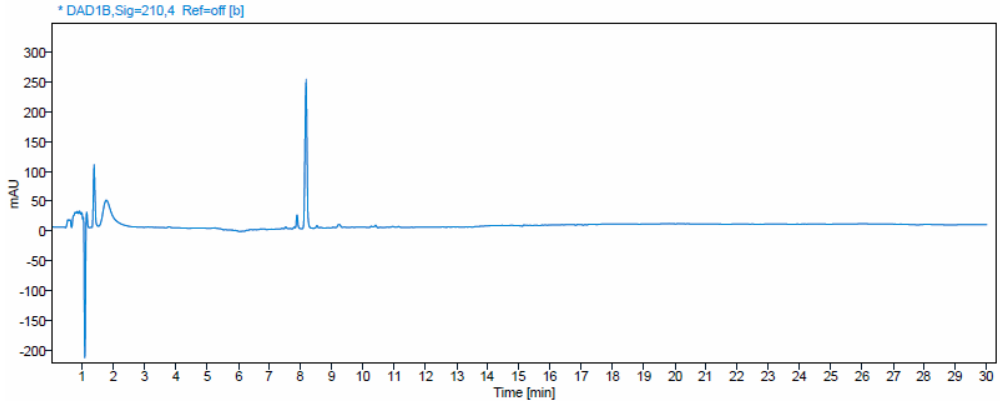
20-2


250
200
150
100
50
0
-50
-100
0
5
10
15
20
25
30

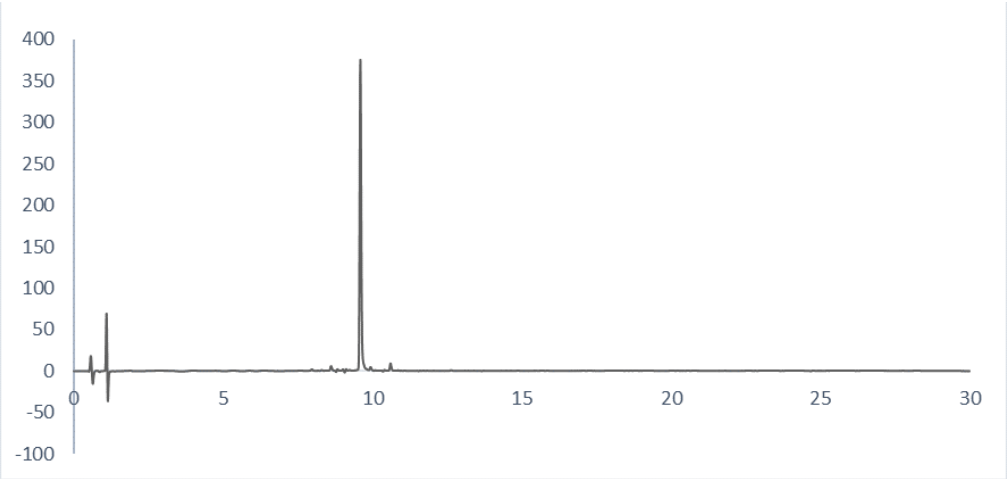
21


* DAD1B,Sig=210,4 Ref=off [b]
mAU
Time [min]

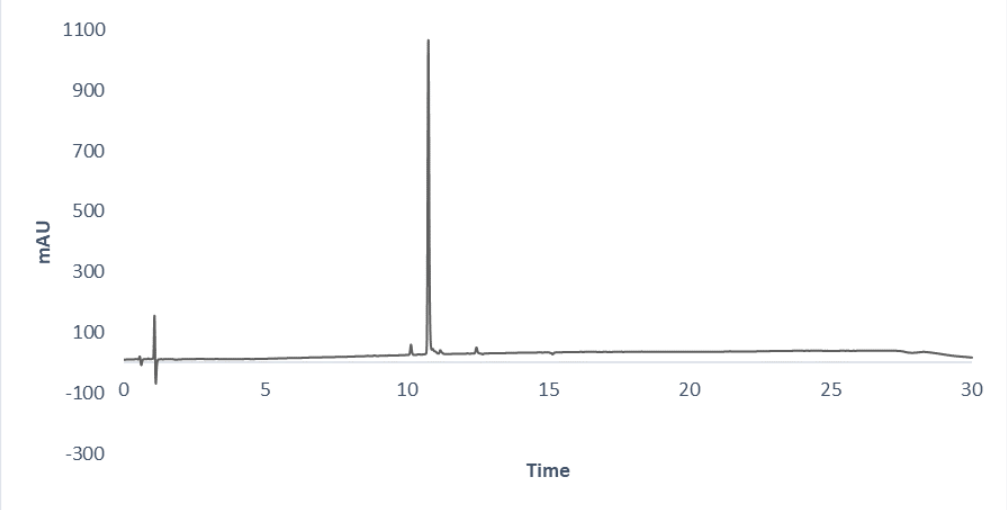
22


* DAD1B,Sig=210,4 Ref=off [b]
mAU
Time [min]

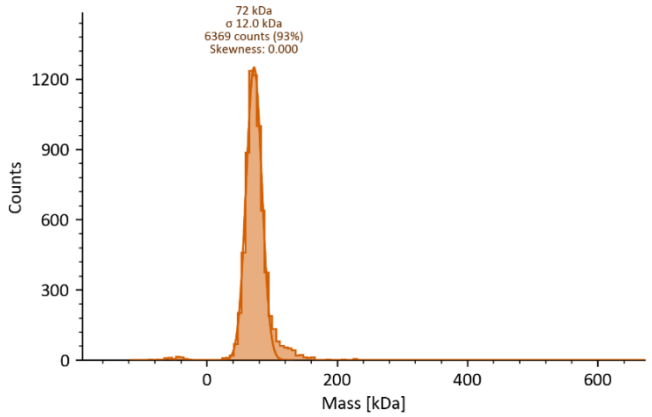
23


400
350
300
250
200
150
100
50
0
-50
-100
0
5
10
15
20
25
30

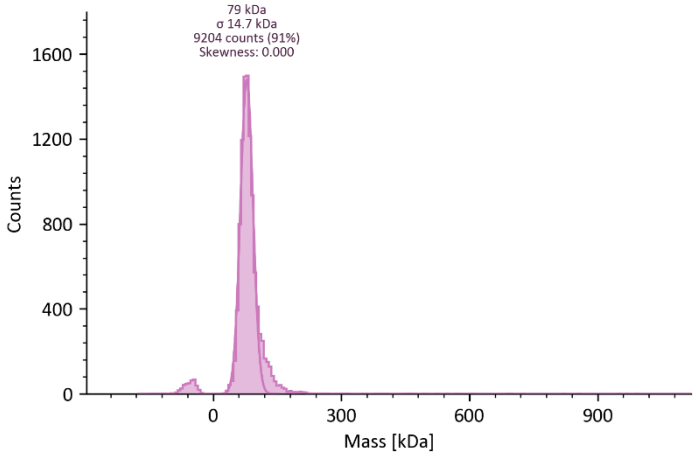
24


1100
900
700
500
300
100
-100
-300
mAU
0
5
10
15
20
25
30
Time

7.4.4. Mass photometry figures



Supplementary figure 11: Mass photometry data observed from the measurement of MBP-tagged hnRNP A2/B1 (residues 1-251, MW: 70.306 kDa) protein with RNA1. 50 nM protein was mixed with 50 nM RNA. The measurement was taken after 30 min incubation at rt.



Supplementary figure 12: Mass photometry data observed from the measurement of MBP-tagged hnRNP A2/B1 protein (residues 1-251, MW: 70.306 kDa) with RNA1. 50 nM protein was mixed with 12.5 nM RNA. The measurement was taken after 30 min incubation at rt.

7.4.5. Used RNAs

Supplementary table 1: RBM20 RNAs

Name	Sequence	Modification
R1	GUCUUAA	3'FAM
R2	GUCUUAGUCUUAGUCUUA	3'FAM
Poly A	AAAAAAA	3'FAM

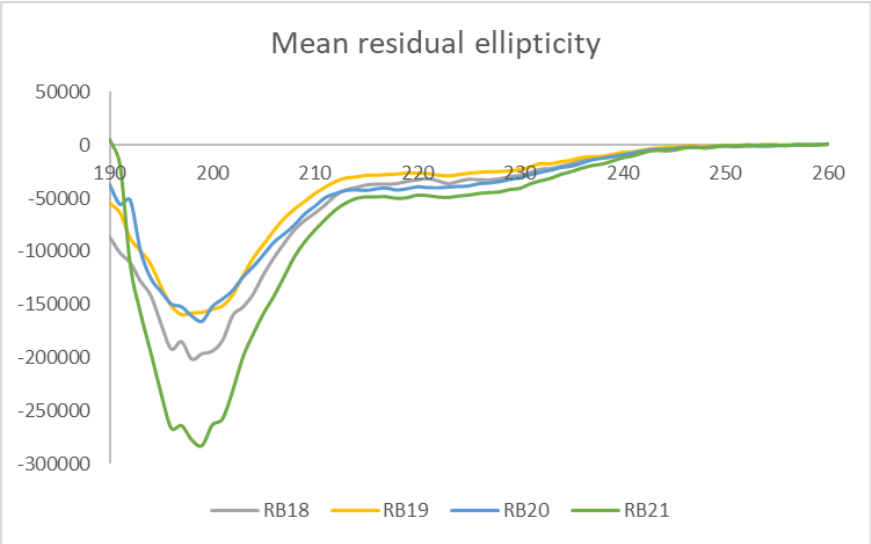
Supplementary table 2: hnRNP A2/B1 RNAs

Name	Sequence	Modification
8mer	AGGACUGC	3'FAM
10mer	AAGGACUAGC	3'FAM

10mer	AAGGACUAGC	5'FAM
RNA1	AAGGACUAGC	3'Cy5
Poly pyrimidine	AUUUUUCCAUCUUUGUAUC	3'FAM
Poly C	CCCCCCCC	3'FAM
Poly C	CCCCCCCC	5'FAM
Poly C	CCCCCCCC	3'Cy5
RNA2	AAGGACUAGCGGGAAGGACUAGC	5'FAM

7.4.6. CD data

Supplementary figure 13: CD chromatogram of peptide mutations of stapled peptides.



Supplementary table 3: Percentage of helicity of mutated stapled peptides.

Peptide	% Helicity
RB18	6.66
RB19	12.34
RB20	14.56
RB21	13.18