TARGETING OF RNA-BINDING PROTEINS WITH MACROCYCLIC PEPTIDES

Dissertation zur Erlangung des akademischen Grades Doktor der Naturwissenschaften

(Dr. rer. Nat.)

Eingereicht an der

Fakultät Chemie und Chemie und Chemische Biologie

Technische Universität Dortmund

Vorgelegt von Stefan Schmeing, M. Sc.

Geboren am 04.06.1995 in Bocholt

Matrikelnummer 166510

Erstgutachter: Prof. Dr. Dr. hc. Herbert Waldmann

Zweitgutachterin: Prof. Dr. Susanne Brakmann

Für meine Familie

DANKSAGUNG – ACKNOWLEDGEMENTS

ZUSAMMENFASSUNG

Die Arzneimittelforschung konzentrierte sich lange Zeit auf die Entwicklung von Modulatoren für Enzyme, da viele Enzyme zielgerichtete, kleine Taschen und katalytische Stellen aufweisen, für die kleine Moleküle auf rationale Weise entworfen oder gescreent werden konnten. Auch Protein-Protein-Wechselwirkungen, für die verschiedene chemische Räume erforscht werden mussten, wurden erfolgreich ins Visier genommen. In den letzten Jahren wurde das Interesse an vielfältigeren makromolekularen Komplexen wie Nukleinsäure-Protein-Interaktionen geweckt. Bei dieser Arbeit liegt der Schwerpunkt auf der Entdeckung von Modulatoren von Protein-RNA-Interaktionen. Das Projekt konzentrierte sich auf den Spleißfaktor polypyrimidine tract binding protein 1 (PTBP1), der bei Überexpression verschiedene Krankheiten wie Krebs, Herzerkrankungen, Diabetes und den Abbau von Nervenzellen verursacht. Dieses Protein besteht aus vier RNA bindenden Domänen, die unterschiedliche Konsensussequenzen mit hoher Ähnlichkeit binden und durch die Tertiärstruktur des Proteins und der RNA eine komplexe Bindung an die Ziel-RNAs ermöglichen. Zunächst wurde versucht, einen auf *E. coli* basierenden zellulären Assay zur Quantifizierung der Interaktion von PTBP1 mit seinen Ziel-RNAs zu entwickeln. Dieser Assay sollte verwendet werden, um eine Bibliothek genetisch kodierter Peptidmakrozyklen auf Inhibitoren dieser Interaktion zu untersuchen. Später wurde die Funktion einer kürzlich berichteten transienten α-Helix zwischen den *N*-terminalen Domänen von PTBP1 untersucht und genutzt, um stabilisierte Peptide zu entwickeln, die mit der nativen, in das Protein eingebetteten Helix *in vitro* konkurrieren. Es konnte gezeigt werden, dass die Makrozyklen die Kooperativität zwischen RRM1 und RRM2 für eine RNA mit zwei Bindungstellen hemmen. Die Peptide weisen einen einstelligen mikromolaren K_I in den durchgeführten Experimenten auf. Diese Moleküle erwiesen sich als mäßig zellpermeabel, lysatstabil und modulierten das Spleißen mit geringen Auswirkungen. Darüber hinaus wurde der Bindungsmodus eines Peptids durch Ko-Kristallisation mit RRM1 validiert.

In einem weiteren Projekt wurden makrozyklische Peptid-Inhibitoren für ein nicht-konventionelles RNA-Bindungsprotein WDR5 strukturell untersucht, um ihre Wirkungsweise zu validieren.

ABSTRACT

Drug discovery traditionally focused on the development of modulators for enzymes for a long time because many enzymes provide targetable small pockets and catalytic sites for which small molecules could be rationally designed or screened. Also, protein-protein interactions, for which different chemical spaces needed to be explored, were targeted successfully. In the recent years the interest in more diverse macromolecular complexes like nucleic acid-protein interactions awaked interest. For this work, the focus is on the discovery of modulators of protein-RNA interactions. The project focused on the splicing factor polypyrimidine tract binding protein 1 (PTBP1) which drives several diseases including cancer, cardiac diseases, diabetes and neuronal degradation when overexpressed. This protein consists of four RNA recognition motifs which bind to distinct consensus sequences with high similarities, which orchestrate a complex binding mode to target RNAs through the tertiary structure of the protein and RNA. First, the development of an *E. coli* based cellular assay for the quantification of the interaction of PTBP1 with its target RNAs was attempted. This assay was planned to be used to screen a library of genetically encoded peptide macrocycles as inhibitors of this interaction. Later, the function of a recently reported transient α-helix between the *N*-terminal domains of PTBP1 was investigated and utilized to design stapled peptides to compete with the native helix *in vitro*. Here, it was shown that the macrocycles inhibit the cooperativity between RRM1 and RRM2 which compete with a single RNA molecule with two binding sites with a single digit micromolar K_I. These molecules were validated to have a sufficient cell-permeability, lysate stability, and modulated splicing with low effects. Further on, the binding mode of one peptide was validated through co-crystallization with RRM1 lacking the native helix.

In another project, macrocyclic peptide inhibitors for a non-conventional RNA binding protein WDR5 was structurally studied to validate their mode of action.

Results from this dissertation contributed to the following publications:

S. Schmeing, G. Amrahova, K. Bigler, J.Y. Chang, J. Openy, S. Pal, L. Posada, R. Gasper and P. 't Hart, "Rationally designed stapled peptides allosterically inhibit PTBP1-RNA-binding", *Chemical Science*, **2023**, 14, 8269-8278, DOI: 10.1039/D3SC00985H

J.Y. Chang, C. Neugebauer, S. Schmeing, G. Amrahova and P. 't Hart, „Macrocyclic peptides as inhibitors of WDR5-lncRNA interactions", *Chemical Communications*, **2023**, accepted manuscript, DOI: 10.1039/D3CC03221C

J. Nowacki, M. Malenica, S. Schmeing, D. Schiller, B. Buchmuller, G. Amrahova and P. 't Hart, „A translational repression reporter assay for the analysis of RNA-binding protein consensus sites", *RNA Biology*, **2023,** Jan;20(1):85-94, DOI: 10.1080/15476286.2023.2192553

Results from this dissertation can also be found in the following master thesis:

K. Bigler, "Evaluation of PTBP1 RNA Interactions", Master thesis, **2022**

TABLE OF CONTENT

1. INTRODUCTION

1.1. SPLICING

Through evolution, all organisms were constantly driven to change and diversify for the needed adaption to live in the changing environments. For this, the multicellular organisms evolved complex macromolecular structures, which are running highly regulated programs allowing their development starting from a fertilized egg.[1] Phenotypes of multicellular organisms are somehow encoded in the genome of these organisms. After sequencing the human genome, it was astonishing that the number of phenotypes of human cells (e.g., different cell types) and the complexity were not reflected by the number of genes encoding for proteins.[2,3] It is thought that alternative splicing plays a considerable role in increasing the number and diversity of proteins in cells, leading to different cell types. Human exons are 50 – 300 nt long and have an average length of 137 nt.[4] Larger introns flank them with average sizes of 3400 nt.[5] A large complex called the spliceosome, composed of around 170 proteins and five small nuclear RNAs (snRNAs), is guided to specific sites on the exons and introns. After recognition of the exon-intron boundaries, the flanking introns are removed during pre-mRNA synthesis by RNA polymerase II.[6]

Exons contain three major sequence elements:

i. The 5' splice site
ii. The 3' splice site
iii. The branchpoint

These elements are recognized and the spliceosomal complex assembles stepwise on the nascent pre-mRNA. The 5' splice site is bound by U1 snRNP. Afterwards, splicing factor 1 binds to the branch point and induces the binding of U2AF at the 3' splice site (E complex). The spliceosome transits from the E complex to the A complex by substituting SF1 with U2snRNP. By binding more factors, the A complex transitions into the active B complex, which catalyzes the transesterification reaction that remove the intron and joins two exons.[7]

1.1.1. MECHANISM OF SPLICING

In the assembled spliceosome, intron removal from the pre-mRNA happens through two transesterification reactions (**Figure 1**). The two steps of splicing are:

i. The nucleophilic attack of the 2' hydroxyl group of the branchpoint adenosine at the 5' splice site leads to a 2'-5' phosphodiester linkage between the branchpoint and the 5'-terminal nucleotide of the intron.

ii. The released 3' hydroxyl of the 5' exon attacks the 3' splice site, forming a new phosphodiester bond between the 5' exon and the 3' exon.

The final products are the connected exons and a free intron released as a lariat. After the second step, the spliceosome dissociates and can perform more turnovers.[8]
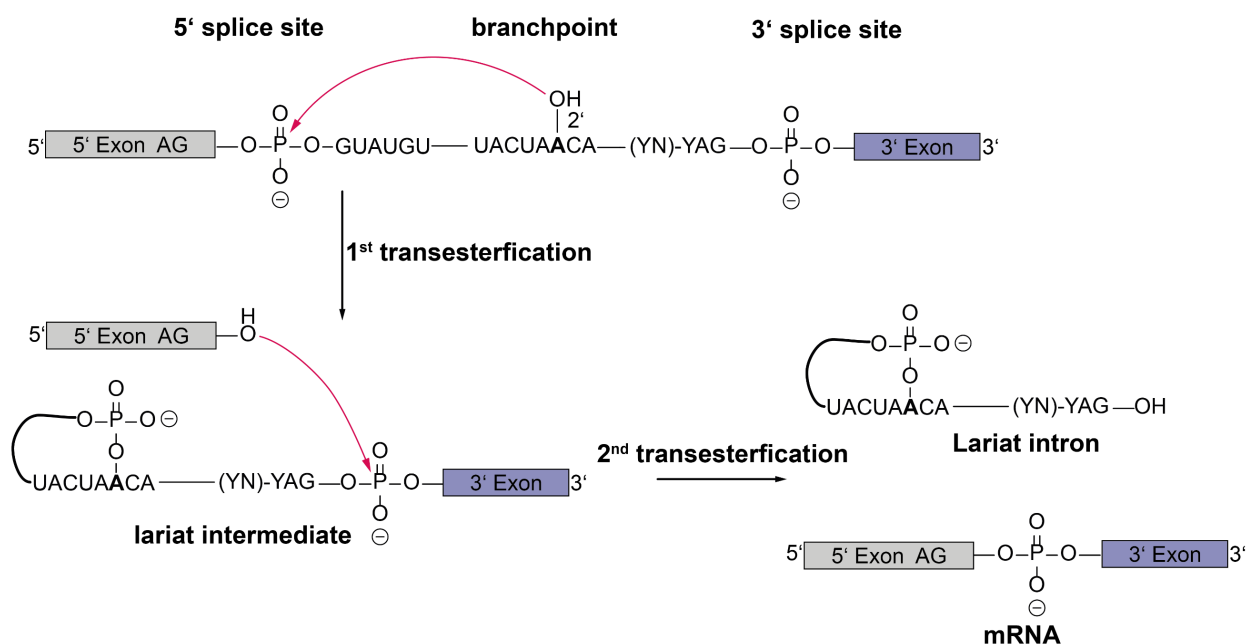
Figure 1: Mechanism of splicing. Figure adapted from [8].

So, why do those simple reactions require so many protein cofactors? The proteins in the spliceosomal complex have essential roles in recognizing splice sites and providing catalytic functions, including conformational changes needed for splicing.[9,10] The 5' splice site, branchpoint, and 3' splice site of pre-mRNA introns contain poorly conserved consensus sequences compensated by *trans*-acting factors that fold the RNA. By this, recognition, the pairing of splice sites within the myriad of consensus sequences, and the right atomic orientation of these sites are achieved. For this, a large number of proteins in the spliceosome is needed.[11,12]

Splice sites in higher eukaryotes are relatively degenerated, and alternative splicing is predominant. So, the recognition of splice sites is influenced mainly by flanking regulatory sequences. Those sequences are called intronic and exonic splicing enhancers or silencers, which have either positive or negative effects on the usage of the splice sites in close proximity.[13] *Trans*-acting regulators bind to those *cis*-acting enhancer elements and recruit snRNP subunits of the

2

spliceosome to the adjacent splice site. In opposite, negative regulators prevent the snRNP subunits from binding. Exonic splicing enhancers (ESEs) are bound mainly by serine-arginine-rich (SR) proteins, while heterogeneous nuclear ribonucleoproteins (hnRNPs) bind mainly exonic splicing silencers (ESSs). An interplay of positive and negative effectors decides whether the spliceosome recognizes or skips a splice site.[12]

1.1.2. ALTERNATIVE SPLICING

First thought of as a less abundant mechanism of gene regulation with low appearance, alternative splicing (AS) is now accepted as an important mechanism for most human RNAs. The detailed sequencing of human tissues indicates that >90 % of human genes with introns are alternatively spliced.[14] Often, AS events are regulated in a tissue-specific manner, increasing multicellular organisms' complexity. For example, the *Drosophila* Dscam gene can theoretically be spliced into 38016 different isoforms by four sets of mutually exclusive exons with 12, 48, 33, and 2 variants, respectively. The Dscam gene encodes a membrane protein essential for neuronal wiring in brain development. Shortly, Dscam confers single neurons the possibility to discriminate themselves from neighboring neurons, which is vital for brain development. Thousands of isoforms are needed for development, but not the complete set of 38016.[15,16] Such a diversity of AS events needs regulation through splicing factors. Many *cis*-acting factors affect the recognition and selection of splice sites, as shown in the following:

i.   Splice site consensus sequences:

Sites with stronger consensus sequences are observed to be more functional.[17]

ii. Spatial/temporal proximity of splice sites:

If splice sites are equal, those which are closer together are paired. Also, a "first-come, first-served" effect was observed because splicing is co-transcriptional.[18]

iii. Extreme proximity:

If the branchpoint of a downstream exon is in extreme proximity to the 5' splice site of an upstream exon, the fusing of those exons is prevented (one mechanism of mutually exclusive splicing).[19]

iv. Exon size:

Exons shorter than defined in the exon definition hypothesis (50 – 300 nt) suffer from a steric clash between splicing factors. Exons longer than 300 nt might need positive regulators for effective splicing.[20,21]

v. Auxiliary splicing regulatory elements (SREs):

>   ESEs, ESSs, ISEs, and ISSs influence exon inclusion positively or negatively by the association of splicing factors.[22]

vi. RNA secondary structure:

>   The secondary structures of the RNA can mask enhancers or splice site elements. Also, base pairing might bring distal splice sites next to each other, leading to the assembly of inhibitory or activating complexes.[23]

vii. Polymerase processivity and elongation speed:

>   As mentioned in ii., splicing happens co-transcriptionally, and in rapid translation, weak exons might not have enough time for recognition by the spliceosome.[24] By this, exons are skipped, while slower transcription leads to exon inclusion.[8]

viii. Epigenetic modifications of chromatin:

>   By slowing down RNA polymerase II, inclusions of weak exons can be enhanced. Also, some epigenetic modifications on histones can lead to the recruitment of regulatory proteins.[25,26]

The previously listed factors lead to this extend of alternative splicing events in pre-mRNAs. Five different splicing patterns were identified as typical AS events (**Figure 2A**). The most common mode in mammalian AS is exon skipping, where an exon gets spliced out of the transcript. If one of two exons is included in the transcript but not both, those are called mutually exclusive exons. Also, an alternative 5' splice site might be used, which leads to different 3' boundaries of the upstream exon. Like this, an alternative 3' splice site can be used, which changes the 5' boundaries of the downstream exons. In the rarest AS event of intron retention, a sequence is simply retained. The difference to exon skipping is the lack of flanking introns at the retained sequence.[27,28]
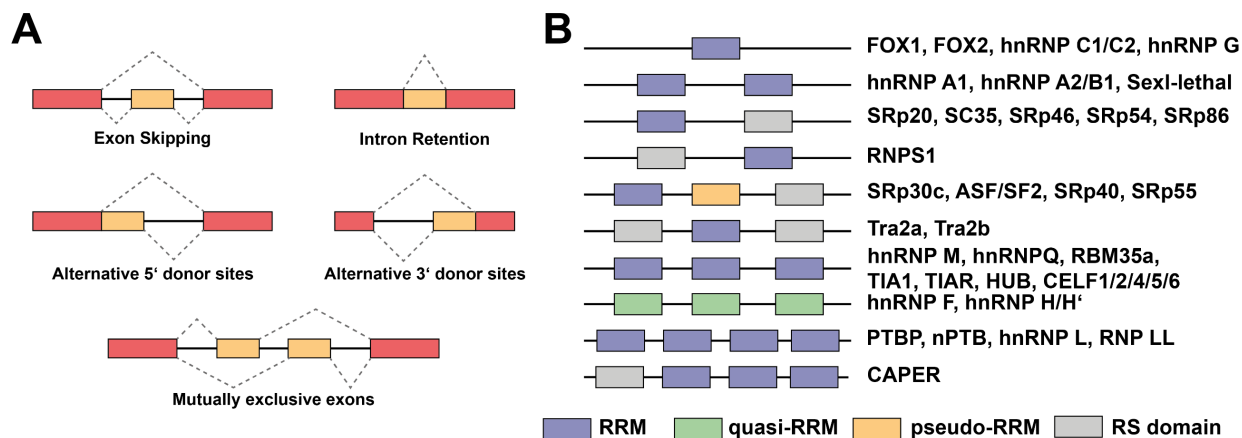
Figure 2: (A) Depiction of different AS events from pre-mRNAs. **(B)** List of example proteins containing RNA binding domains.

1.1.3. STRUCTURAL INSIGHTS INTO PROTEINS INVOLVED IN SPLICING REGULATION

As mentioned, the spliceosome and associated proteins resemble a dynamic protein-RNA complex. It spans a network of RNA-RNA, protein-RNA, and protein-protein interactions. In total, ~170 spliceosome-associated factors are known for human cells. Those proteins are classified into core spliceosomal proteins and non-core splicing factors. Those splicing factors bind *cis*-acting elements on the pre-mRNA and influence the splicing of *trans*-located splice sites positively or negatively. Splicing factors are categorized into three families:

i.   SR proteins that facilitate splice-site recognition.
ii.   hnRNP proteins often antagonize splicing.
iii. tissue-specific splicing factors, which can do both.

All three families are RNA binding proteins that contain either one or more sequence-specific RNA binding domains, namely RNA recognition motifs (RRMs), K homology domain (KH domains), or zinc fingers (**Figure 2B**). In the following chapters, structural insights on RRMs will be given as these are the focus of this thesis.

1.1.4. STRUCTURE OF THE RNA RECOGNITION MOTIF

The RRM (also known as: RBD or RNP) is the most abundant RNA binding domain in higher vertebrates and is part of 0.5 – 1% of human genes.[29] This type of domain is known to bind nucleic acids, both RNA and DNA, and also plays a crucial role in higher organized protein-RNA complexes by facilitating protein-protein interactions.[30,31] RRMs are approximately 90 amino acids

in size with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology (see **Figure 3A**). This structure forms a four-stranded β-sheet packed against two α-helices. Most protein-RNA interactions are located on the sheet surface, but interestingly more diverse, non-conserved binding modes were observed even though the fold shows high conservation (**Figure 3B**).[30,31] Structural studies of several RRMs showed a variety of different binding modes for different domains. In some cases, like PTBP1 RRM1 and RRM3, RNA binds over the whole sheet, while for example hTra2 and U1A don't bind nucleotides on the β_4 strand. Also, the orientation of the termini of the RNA is different in those cases (**Figure 3B**).[32]
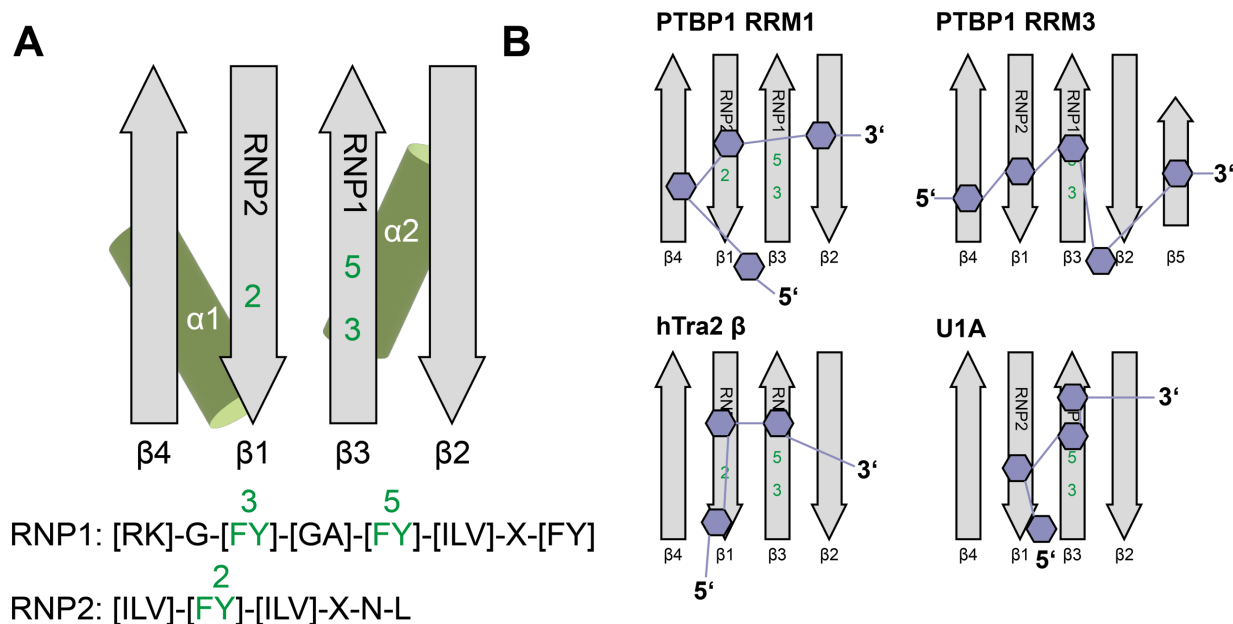
Figure 3: (**A**) Schematic representation of the RRM fold with positions of RNP-residues on the canonical RRM fold. (**B**) Similar representation of the RNA modes of four different RRMs with the RNA nucleotides as purple hexagons. Figure adapted from [8] and [32].

The typical interaction between RNA and RRM involves three aromatic side chains exposed on the β-sheet surface binding two nucleotides. The 5' and 3' nucleotide bases stack on the aromatic rings on β_1 in position 2 and β_3 in position 5 of the RNP residues (**Figure 3B**). The third aromatic side chain is located on β_3 in RNP position 3 and inserted between the sugars of the dinucleotide. Some RRMs differ from this mode of action by having additional *N*- and *C*-terminal extensions that contact the RNA or make contacts via the interdomain-linkers between two different RRMs in the same protein.[8] In the rarer case of a protein containing a single RRM, the specificity and affinity are expected to be lower than in a multi-RRM protein, which seems to be the case for SRp20. This SR protein binds to 4 nucleotides (consensus sequence CNNC) with an affinity of ~20 µM. The low affinity and specificity were explained by evolutionary pressure through the universal abundance of the consensus sequence.[33] But, whenever one finds those rules, exceptions are

found as well. For example, Fox-1 does have a single RRM and sub-nanomolar affinities to GCAUG motifs. The canonical RRM is extended by long loops with aromatic residues connecting secondary structure elements leading to the binding capacity of seven nucleotides.[34]

An astonishing fact about RRMs is the specificity of the protein-RNA interactions mediated by the canonical and well-conserved fold. In the past, a binding code was hypothesized, and the solved structures of RRM-RNA interactions indicate some preferences of the nucleotide binding sites. The N1 and N2 pockets favor C and A or G with *syn* conformation (N: number of the bound nucleotide). It was shown that 52 % of bound cytosines and 35 % of adenines locate in the N1 pocket, while 30 % of guanosines are situated in the N2 pocket with *syn* conformation.[35,36] The most frequently bound nucleotide is U which can bind to all pockets (including N1 and N2).[37–43] Albeit this slight preference of RRMs for specific nucleotides, the canonical fold cannot fully explain the sequence specificity. For example, the unusual versatility of RNA recognition is encoded in the non-conserved features like the length of loops and secondary structure elements, extensions of the *N*- and *C*-termini, or the interplay of several domains.[36]

Several structures of RRMs of proteins containing multiple RRMs have been solved in the past. These structures showed that RRMs separated by smaller loops bind to similar sequences but with different affinities and secondary structure preferences. Most splicing regulators contain several RNA binding domains, which were for a long time thought to be independently binding to their specific binding sites on the RNA, for example, *cis*-acting sites. The inter-domain linkers of these RBDs were ignored and considered to be flexible linkers for a while.[30] In several recent studies, the role of those linkers was investigated and found to be more than just a spacer. The linkers can bind and block the RNA binding site, allow the perfect spacing of two tandem RRMs to fold onto each other, extend the sheet's surface or take part in intra-RRM contacts of the same protein chain.[43–46] Extension of the binding sites usually happens by secondary structure elements that fold or unfold upon RNA binding and either extend or block the RNA binding site.[44] Further, these single motifs could mediate inter-domain contacts, which change the tertiary fold of the protein and, by this, change the specificity for the long pre-mRNA by organizing the structure of the splicing factor. An example of an extended RRM is DND1s RRM1 domain, which is *N*-terminally extended by a β-hairpin (β-1.0) which folds on an α0-helix. RRM1 canonically binds to the RNA, but RRM2 sandwiches the RNA over a non-conserved positively charged surface without using RNP residues, while the short intra-domain linker makes additional contacts to the RNA over the backbone and sidechain residues.[47] A protein from the same hnRNP family, Syncrip, does provide a similar *N*-terminal extension which binds to a non-canonical RBD and increases the affinity to the target miRNA by this. The other two RRMs of Syncrip were thought to work in tandem

but do not have a stable fold in the absence of RNA.[48,49] As discussed, some RRMs of multidomain RBPs interact, but there are also examples where the domains act independently, like Syncrip. Another example are the tandem RRMs of Sex-lethal, which do not interact with each other in the RNA-bound and -unbound state.[50]

Another interesting hnRNP protein is PTBP1, which is the main topic of this work and will be discussed in more detail in chapter 1.2.

1.2. POLYPYRIMIDINE TRACT-BINDING PROTEIN (PTBP)

1.2.1. STRUCTURE OF PTBP AND ROLES OF THE DOMAINS

Polypyrimidine tract-binding proteins (PTBP) are essential in splicing and are expressed in many cell types. Different proteins of the PTBP family share a high similarity but show different biological functions and expression patterns in different cell types. PTBP1 (also known as hnRNP I and PTB) and its paralog PTBP2 (also known as brPTB or nPTB) are trans-acting splicing repressors binding to polypyrimidine tracts on target mRNAs.[51–53] PTBP1 is highly expressed in most cell types, while PTBP2 is less expressed but shows high expression levels in cells lacking PTBP1, like neuronal cells. During neuronal differentiation, PTBP1 is downregulated while PTBP2 levels increase; it is assumed that PTBP2 is important in neuronal maturation.[54] Proteins of the PTBP family utilize four highly conserved RRMs to bind the CU-rich consensus RNA sequence (**Figure 4A/B**).[55] The structures of the RRMs were solved independently and shared the same consensus sequence but bind to different target mRNAs with different affinities.[56] Analyses by small-angle X-ray scattering (SAXS) and NMR without RNA revealed that the two *N*-terminal domains RRM1 and RRM2 tumble free in solution, while RRM3 and 4 interact with each other.[57] The NMR structure of RRM1 was solved bound to a stem-loop RNA containing the consensus UCUU motif. There, the RRM binds three pyrimidines (C11-U12-U13) on the conventional β-sheet with F98 (RNP1) and a
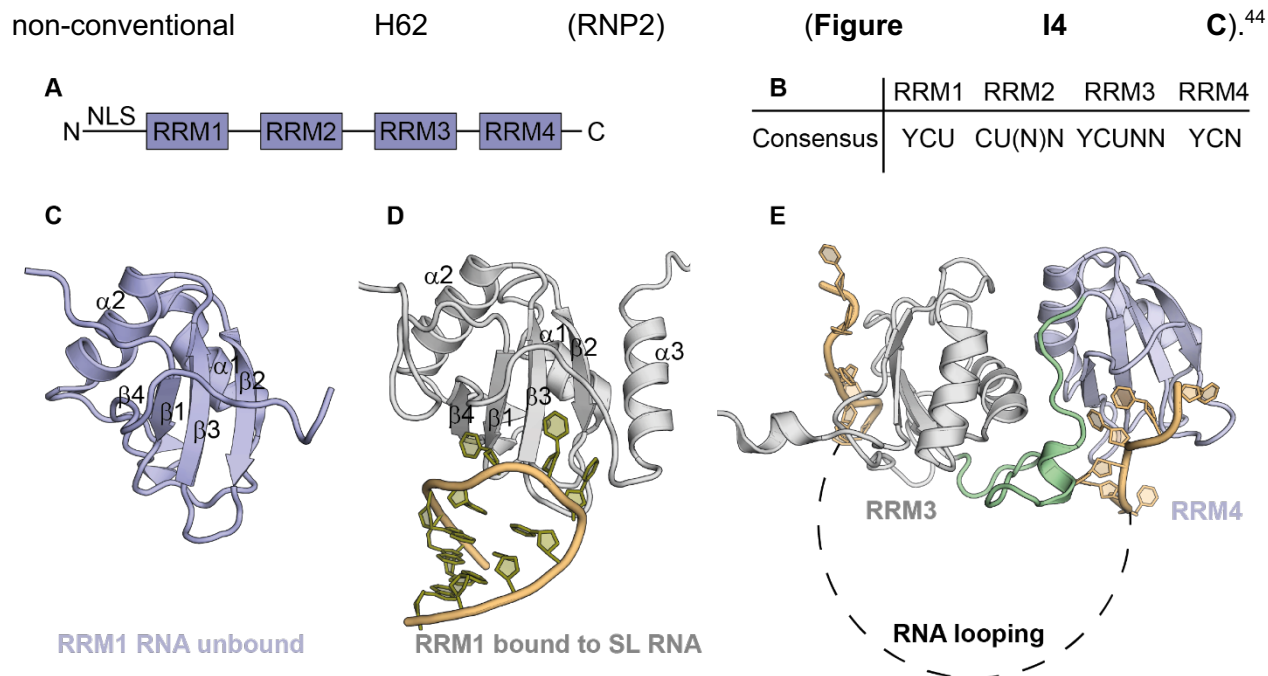
non-conventional   H62   (RNP2)   (**Figure   I4   C**).[44]

Figure 4: (**A**)Domain structure of PTBP1. The protein contains four highly conserved RRMs and an *N*-terminal nuclear localization signal (NLS). (**B**) Consensus sequences of RNAs for the individual RRMs. (**C**) NMR structure of RRM1 in the RNA-unbound state (PDB: 1SJQ) (**D**) NMR structure of PTBP1 RRM1 bound to a stem-loop RNA with an additional α3-Helix packed to β2 and α1 (PDB: 2N3O). (**E**): NMR structure of the RRM34 tandem bound to RNA (PDB: 2ADC).

First, it was thought that RRM1 folds into the canonical RRM1 fold, but later the group of Allain found that the unstructured, long interdomain-linker of RRM1 and RRM2 forms a transient α3-helix which stacks onto RRM1s β3 and α1 helix when RNA is bound (**Figure 4D**). This helix was present in low abundance in the RNA-unbound state and was hypothesized to first sense the fold, bind the RNA and then stably fold in the RNA-bound state.[44] In a recent pre-print, the same group further investigated the tertiary structure of PTBP1 bound to the EMCV IRES RNA. The authors revealed several intramolecular interactions of the interdomain-linker of RRM1 and 2, which places the two domains in proximity. This proximity is mainly caused by α3 folding back, which leads to RRM1 facing the β-sheet of RRM2. Also, the α2-β4 loop of RRM1 interacts with the β3-α2 of RRM1 while the linker contacts the α1-α2 surface of RRM1.[45]

The RNA binding of RRM2 is relatively conserved. It binds four pyrimidines with the consensus sequence CU(N)N and extends the conventional RRM-fold by a fifth β-strand allowing the binding of an additional nucleotide. The RNP1 motif does play a weak role in the binding process and is replaced by hydrophobic chains located in β2. Also, the solved structures don't show any conformational change of the protein upon RNA binding, except stabilization of loops after binding

9

like the β4-β5 loop.[55] It was demonstrated that RRM2 is the minimal component for many protein-protein interactions. RRM2 contains a dorsal binding site for several PTBP1-RRM2 interaction peptides (PRIs).[58] The first interaction partner found was Raver1, overexpression of which promotes exon 3 skipping on the target Tpm1.[59] The role of PRI interactions is unclear, as the knockdown of Raver1 did not show changes in the transcriptome. But the specificity of this interaction suggests that the interaction might play a physiological role. Matrin 3, another PRI-containing protein, plays a role in regulating alternative splicing. Transcriptome analysis showed that Matrin 3 principally represses cassette exons, acting as a splicing regulator. Comparison with PTBP1 target exons showed 18% of shared targets, so Matrin 3 acts mainly independent of PTBP1.[58]

Like RRM2, the third domain is extended by a *C*-terminal β-strand and, by this, binds five pyrimidines with the consensus motif YCUNN. The fourth domain can bind three pyrimidines with the consensus motif YCN. Both *C*-terminal domains can bind RNA independently when expressed and purified as single domains, but interestingly they act in tandem. Both domains are monomers when independently analyzed, but the pair shows a globular fold in SAXS analyses indicating inter-domain contacts even without RNA.[57] The globular fold was further proven by solving the structure of the tandem in the presence and absence of RNA.[55,60] In the RNA-unbound conformation, domains 3 and 4 share the same fold between the tandem and individually solved structures. But they bind to each other by interactions between the helices α1 and α3 of RRM3 and α2 of RRM4 (**Figure 4E**). Additional interactions by the interdomain-linker are observed, and in total, the interface involves 27 side chains. The main drivers are contacts of α2 (RRM4) with the loop α2-β4 of RRM3 and the *C*-terminal part of α2 (RRM4) with the *N*-terminal region of α1 (RRM3) and L335 of the linker. Another hotspot of the interaction are several sidechains of the linker interacting with α2 of RRM3.[60] In the RNA-bound state, the conformation of the protein does not significantly change, indicating the importance of the orientation of the structures for PTBP1 specificity. The pre-organization of the domains allows the binding of two CUCUCU hexamers with a 15 nt spacer between them. Following, RRM34 brings two distant pyrimidine tracts into a 30 Å distance. By this, the tandem can introduce RNA looping, and thus PTBP1 could sequester a short alternative exon or branch point (**Figure 4E**).[55]

For PTBP1, by sequentially deleting the RRMs of heterologously expressed PTBP1, it was shown on specific target RNAs *in vitro* that RRM3 represents the minimal domain for specific RNA recognition. At the same time, the roles of RRM1 and RRM4 are minimal.[61] It was found that flexible linkers between RRMs 1, 2, and 3 play a part in the specific interaction with target RNAs.

Furthermore, the stiff linker between RRMs 3 and 4 enabled back-to-back packing of both domains without the presence of RNA, further confirming that they act as one binding unit.[55]

1.2.2. THE ROLE OF PTBP1

The polypyrimidine tract binding protein plays a role in diverse RNA-dependent cellular processes like alternative splicing, polyadenylation, RNA transport, localization, stability, replication, and IRES-mediated translation (**Figure 5**).[62] First, it was thought that RRM3 binds to polypyrimidine-rich elements of introns antagonizing U2 auxiliary factor.[63,64] This was later found to be unlikely by showing that PTBP1 can repress exons without additional proteins.[65,66] Exemplarily, PTBP1 regulates tropomyosin β-TM exon 7 in non-muscle cells while it inhibits α-TM exon 3 in muscle cells.[67] Further, mutually exclusive splicing of α-actin upstream of a non-muscular exon and downstream of a smooth muscle pointing exon generates these muscle subtypes.[68] The alternative splicing of GABAγ2 exon N and c-src exon N1 in neuronal cells was also heavily studied and was shown to be regulated by PTBP1. PTB and KH splicing regulatory proteins, for example, interact with the c-src transcript, and those interactions are changed with the cell's differentiation state.[52,53] For neuronal cells, it is also important that the exclusion of exon 11 in PTBP1 leads to nonsense mediated decay (NMD) of the mRNA and that PTBP1 leads to the exclusion of this exon on its own.[67] Similarly, PTBP1 leads to the exclusion of exon 10 of PTBP2 in non-neuronal cells, while PTBP2 leads to the exclusion of exon 11 of PTBP1 in neuronal cells. By this, the tissue-specific expression level of the paralogs is regulated.[52] As mentioned, PTBP1 does not only play a role in alternative splicing events. By binding to U-rich elements upstream of poly-adenylation signals, the poly-adenylation of mRNA is promoted.[69] Furthermore, the *N*-terminal 55 amino acids of PTBP1 contain nuclear localization and export signals. Through this, PTBP1 is shuttled between the nucleus and the cytoplasm and takes part in the export and localization of mRNA.[70] Also, it was shown to play a role in the localization of RNA in *Xenopus* oocytes.[71] It also affects the output and stability of several mRNAs, like the activation of hepatitis B virus mRNA or the stabilization of mRNA by promoting the dissociation of RNA helicase UPF1 from 3'-UTRs.[72–74] In pancreatic β-cells, PTBP1 expression is regulated in a glucose-dependent manner. Upon glucose uptake, PTBP1 expression is increased, stabilizing the insulin mRNA by binding to its 3'-UTR, which increases insulin production.[75–77]
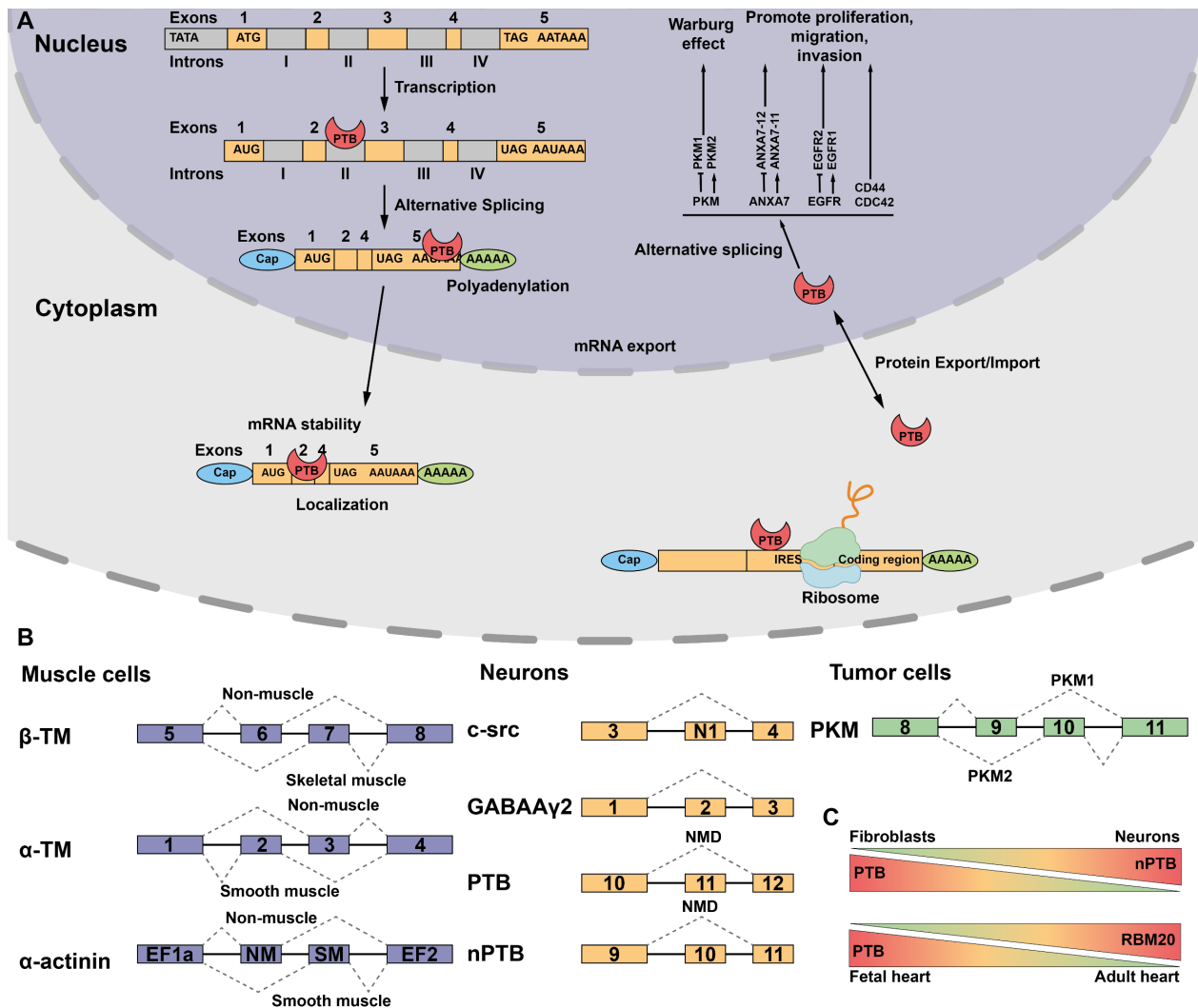
Figure 5: Roles of PTBP1: effect on localization (**A**) and specific examples for alternative splicing events (**B**).

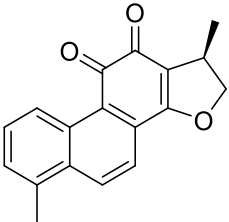## 1.3.   TARGETING OF SPLICING AND RNA RECOGNITION MOTIFS
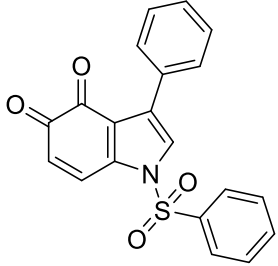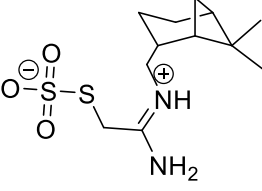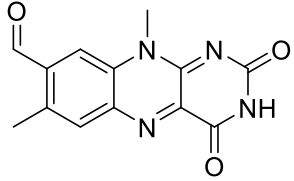
As described earlier, alternative splicing is well orchestrated, and misregulation can lead to cancer, neurodegeneration, metabolic, cardiovascular, and immune diseases.[78] An increase of 30% in AS events in tumor cells compared to healthy tissues was recently observed.[79] Also, transcriptomics revealed cancer-specific splicing patterns and NMD events caused by defects in AS in thousands of genes of cancer tissue.[80–85] In many cases, the differently spliced transcripts play a role in pathways promoting proliferation, evading apoptosis, influencing telomere length, cell cycle regulation, tumor metabolism, and angiogenesis.[86–90] The upcoming question is: Which defects cause this misregulation? There are several possibilities like mutations in intronic or exonic *cis*-acting sequences, over- and under-expression, or altered copy numbers of splicing factors.[91,92]

Lost regulation of those splicing factors from the SR and hnRNP family causes cancer because many of those proteins act as oncoproteins or tumor suppressors.[93–95]

The first discovered anti-cancer drugs targeting splicing were small molecules that interfere directly with the spliceosome activity or assembly. Their target is the SF3b subcomplex of U2 snRNPs, responsible for recognizing branchpoint sequences in pre-mRNA and is part of the catalytic core of the spliceosome. By binding to SF3b, spliceostatin A and pladienolide B inhibit splicing *in vitro* and promote pre-mRNA accumulation.[96,97] The first splicing modulator entering clinical studies, E7107, is a pladienolide B derivative and showed splicing modulation in phase I studies, but alas, the lack of clinical efficacy and the high toxicity of the drug led to discontinuation of the clinical studies.[98,99] The same was true for another small molecule H3B-8800, which showed splicing changes, but no clinical effect was observed.[100] These failures could be caused by the cancers losing their dependency on the splicing defects by additional mutations or a needed higher dose and treatment time which could not be accomplished because of the toxic effects. It is assumed that general splicing inhibitors are less specific for individual targets and are more toxic. There is a need to target the disease relevant subsection of splicing events. This could be achieved through inhibition of RNA binding domains of splicing factors instead of the spliceosome. So, the direct targeting of splicing factors inhibiting their RNA binding activity would be an ideal mode of action for the modulation of splicing.[101]

The specific targeting of individual RRMs can be approached in two ways: i) A nucleic acid-focused approach binding to the splice site or mRNA, targeted by the RRM, or ii) a protein-focused approach with ligands for the RRMs. Utilities for the first approach could be siRNAs, CRISPR/Cas9 editing or antisense oligonucleotides (ASOs). Those ASOs are short oligonucleotides that base pair with the target mRNA and, by this, alter the AS of the target. Several clinical studies are currently performed utilizing splice-switching oligonucleotides (SSOs), which sterically block the spliceosome or splicing factors and influence the balance of isoforms from a given pre-mRNA.[102] The second, protein-focused strategy can also be executed with oligonucleotides. But in this case, decoy oligonucleotides are designed to bind to specific RRMs of target proteins, inhibiting binding activity to target RNAs. The group of Allain demonstrated this approach on the splicing factors SRSF1, RBFOX1/2, and PTBP1.[103]

Table 1: List of reported example compounds inhibiting RRM domains of different proteins.

Compound	Affinities	Cellular
HuR		
DHTS[105]	IC$_{50}$ 149 nM (EMSA) K$_D$ 3.74 nM (Alpha screen)	IC$_{50}$ 0.84 µM (Viability)
6a [106]	K$_I$ 12.8 nM (AlphaScreen) K$_D$ 4.8 µM (DMR)	EC$_{50}$ 24 - 30 nM (viability in cancer cells)
U2AF		
NSC 194308[104]	K$_D$ 100 µM (FP)	EC$_{50}$ 10 – 50 µM (Splicing Assays)
Musashi		
Ro 08-2750 [107]	IC$_{50}$ 2.7 µM (FP) K$_D$ 12.3 µM (MST)	EC$_{50}$ 2.6 µM (Viability)

Several groups have attempted to accomplish RRM targeting using small molecules, which can overcome some of the weaknesses of oligonucleotide-based drugs such as poor *in vivo* stability, cell permeability, and bioavailability.
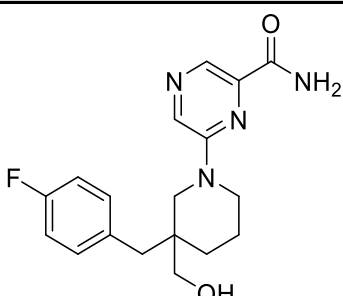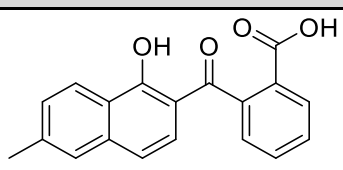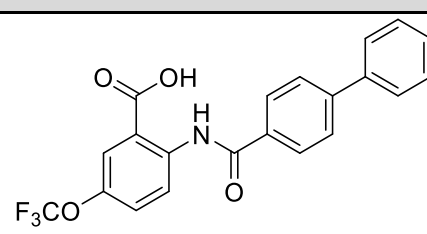
Several small molecule modulators for RRMs involved in diverse cellular pathways were found in the last years (**Table 1** and **Table 2**), reviewed in the following paragraphs. Considering how often RRM domains are encoded in genes, the following examples contribute only to a small fraction of RRMs in the human genome. A few examples of small molecules targeting U2AF, HuR, Musashi1/2, TDP-43, La, IMP2, NONO, and hnRNP A2/B1 are given here. The first RNA-competitive RRM inhibitors were discovered for the protein HuR, an essential factor for stabilizing AU-rich elements (ARE).

These ARE elements are involved in inflammatory processes, cell division, migration, and metabolism; their misregulation has been linked to several diseases. HuR contains two RRMs that bind the RNA in a non-conserved way by "sandwiching" it between both RRMs. This inter-domain surface was identified early on as a druggable pocket, and several nano- to micromolar inhibitors like DHTS (and derivatives) were identified in high throughput screenings and SAR studies. Here, the compounds bind on the β-sheet of RRM1 and not only compete with the RNA but also inhibit the dimerization and cytoplasmic localization of HuR, which is needed to perform its cellular function.[108,109]

Other proteins were also targeted using small molecules binding between two domains, like U2AF with NSC 194308 with a higher 100 μM affinity.[104] Here, the ligand is hypothesized to interact between the two RRMs and induce an RNA-incompatible conformation without interacting with the RNP residues. Controversially, the initial screening results suggested increased RNA binding in the presence of the ligand, while the follow-up evaluations showed an inhibiting effect. Further on, the biophysical and animal model data do not fully explain the mode of action of this compound since effects were achieved well below the in vitro K_D values.[104]

Table 2: Continued from Table 1. List of reported example compounds inhibiting RRM domains of different proteins.

Compound	Affinities	Cellular
TDP-43		
rTRD01 [110]	K_D 90 μM (MST) IC$_{50}$ 150 μM (AlphaScreen	n.d.
La		
C2.01 [111]	K_D > 1.3 mM (EMSA)	EC$_{50}$ > 50 μM (viability)
IMP2		
Cmpd 4 [112]	K_I 81.3 μM (FP)	Not quantified

The proteins of the Musashi family (MSI1 and MSI2) could be targeted using molecules that bind to an "invisible" pocket occupied by its native allosteric regulators of the oleic acid class. Because MSI1 and MSI2 play a role in regulating lipid synthesis pathways, first ω-9 fatty acids were identified in a high throughput screening (HTS) to inhibit RNA binding of RRM1 of MSI1 with nanomolar K_Is.[113] Later, small molecules were identified through HTS and *in silico* docking approaches to inhibit RNA binding, probably binding to the same pockets or interacting with the RNP residues with high nanomolar to low micromolar K_Is.[107,114,115]

In silico docking on the RNA binding protein TDP-43 involved in neurodegenerative diseases also identified high micromolar binders and inhibitors. The small molecule rTRD01 was identified to bind to the RNP residues on the RNA binding surface with 90 μM affinity and competes with RNA in an orthogonal assay (IC$_{50}$

150 µM). [110] Disputably, this compound does not compete with the consensus RNA sequences of the RRM but competes with the disease-related RNA probe, and cellular activity was validated with a very general *in vivo* larvae turning assay.[110] Three more small molecules for TDP-43 were found in NMR-based fragment screenings (Hit 1-3) binding to RRM2 at the nucleotide-binding side, but no further biophysical or cellular assays were performed to validate the mode of action because of the poor affinities.[116]

Further, a fluorescence polarization HTS was used to identify inhibitors of IMP2, from which the best hit Cmpd4 competed with RNA binding for different RNAs with a Ki of 81.3 µM.[112] The hit was validated with orthogonal *in vitro* methods in thermal shift assays and STD-NMR, but the binding site could not be identified with *in silico* methods. Cmpd4 also showed cellular effects in cell viability assays on cancer cells, inhibited cell growth, and reduced the mRNA levels of IMP2 target transcripts.[112]

More reliable data was collected for the compound (*R*)-SKBG-1, which inhibits the NONO protein. This compound was found in a phenotypic screening approach searching for potential covalent binders which influence the androgen receptor mRNA levels.[117] Activity-based protein profiling with mass spectrometry identified a hit B21 binding to NONO. This hit was further optimized, resulting in the lead compound (*R*)-SKBG-1, validated to bind covalently to C145 in NONO, which lies between two RRM domains. The authors proved a NONO-dependent mode of action by performing cellular assays with knockdowns and C145A mutations of NONO. They confirmed that the presence of WT NONO is needed for sensitivity to the compound and performed eCLIP assays that indicated (*R*)-SKBG-1 stabilizes NONO-RNA interactions, and a compensatory effect of ortholog proteins PSPC1 and SFPQ was discovered.[117]
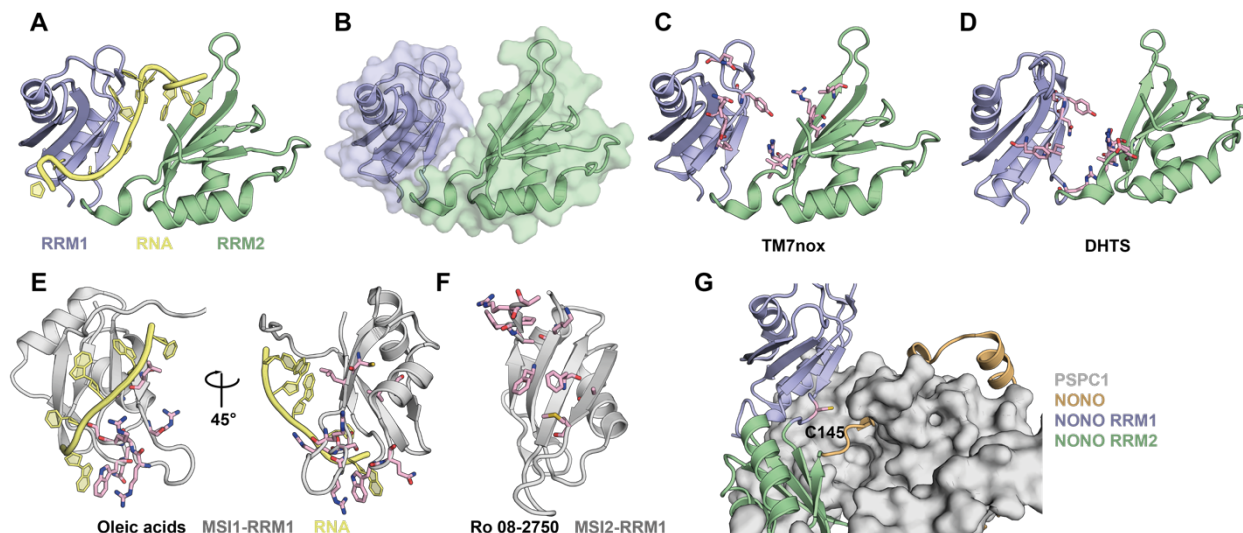
Figure 6: (**A**) Structure of the HuR RRM1 and RRM2 tandem with RNA bound and (**B**) surface representation of both domains. (**C**) Nucleotides involved in TM7nox binding (pink) and for (**D**) DHTS. (**E**) RNA bound structure of MSI1-RRM1 with amino acids involved in oleic acid binding in pink). (**F**) MSI2-RRM2 structure with residues involved in Ro 08-2750 binding highlighted in pink. (**G**) Structure of the PSPC1-NONO complex showing the position of C145 which is covalently bound to covalent NONO inhibitors.

Generally, the presented molecules either bind to the RNP residues on the β-sheet, which take part in the RNA binding, or bind to specific non-conserved pockets (**Figure 6**). Molecules inhibiting the protein-RNA interaction by binding to the RNP residues might lack specificity because of the high conservation of the fold and location of the important residues. Unfortunately, this property was not checked in the reported studies, except for the MSI2 binder Ro 08-2750, where the compound was also active on other RRM-containing proteins, albeit with significantly lower K_{DS}.[107] A less predictable mode of action was achieved for the presented compounds binding to pockets outside the traditional RNA binding site. This mode of action seems to be a practical approach for identifying specific molecules because of the lower conservation of the pockets. The reported small molecules either inhibit the oligomerization, localization of the target proteins, or lower their affinity to RNA by altering the conformation of the RRM. Also, proteins harboring several RRMs could be targeted with molecules binding between the domains. Comparison of both modes of actions in terms of affinities does not allow us to judge the success, as for both concepts, nano- to micromolar affinities could be achieved.

In summary, some RRM-containing proteins were targeted with varying compounds with different modes of action. Generally, the specificity and affinity of RRM ligands are weaknesses of the reported compounds. To improve this, several strategies could be used:

First, unique features of RRMs should be identified. Small molecules designed for those potential allosteric pockets could have a high selectivity for the target protein while also having moderate affinities (like MSI and NONO binders). Second, transient structural elements like conformational changes needed for RNA binding could be blocked or glued to the RRMs. Thus, the RNA affinity could be altered, as was shown for the HuR binders, which bind between the domains. Third, the oligomerization and co-localization of the proteins could be altered by small molecules, as demonstrated for HuR.

1.4. MACROCYCLIC PEPTIDES – INHIBITORS FOR PROTEIN-PROTEIN INTERACTIONS

During the last few decades, a rapidly increasing number of drug discovery projects focused on protein-protein interactions (PPIs) were performed.[118] These interactions were especially challenging until recent developments because < 500 Da small molecules could not interact with shallow or flat, non-hydrophobic, and large surface areas of PPIs. Also, the amino acid composition of typical small-molecule interfaces and PPI interfaces differ in hotspot areas; hence regular small-molecule libraries don't suit perfectly.[119–123] The recent development of biologics like antibodies, or other biomacromolecules resulted in the targeting of previous undruggable interactions. Typically, biologicals have molecular weights of > 5000 Da, high affinities, and selectivity for the targets but lack bioavailability, stability, and cell permeability.[124,125] The gap between the properties of biologicals and small molecules can be filled with therapeutic peptides, which often provide high affinities and selectivity and can be optimized to overcome the problems of cell penetration and bioavailability.[118] Their numbers in clinical studies portray the clinical relevance of peptides: Until March 2017, 484 advanced medicinal peptides were part of pre-clinical or clinical studies in the major pharmaceutical markets. Of those, 54% were discontinued, 32% were in active clinical trials, and 12% were approved.[126]

Helical peptides are, together with β-sheets, the main secondary structure element in PPI surfaces. Often, the helices are not stable in the absence of the full-length protein or the binding partner, and the peptides are sensitive to proteolysis, have poor cell permeability, and, by this, are poor drug candidates.[127–129] Therefore, peptides are often modified to avoid those negative characteristics. Macrocyclization is one of the most used modifications of peptides, which is inspired by many peptide-based natural products found in bacteria or funghi. Through macrocyclization, often the affinity to a biological target is increased because of the lowered

flexibility of the peptide. Furthermore, cyclic peptides are often very resistant against proteolytic degradation.[130] Peptides can be cyclized through many different strategies including head-to-side-chain, side chain-to-side chain, side chain-to-tail, or head-to-tail linkages (**Figure 7**).

Stabilization of helices through macrocyclization is achieved through introduction of non-canonical amino acids into peptide sequences, which can be exploited to covalently lock the peptides into helical conformations by linking two amino acids on the same side of the helix.[126,129,131,132] The resulting peptides are called "stapled" peptides.
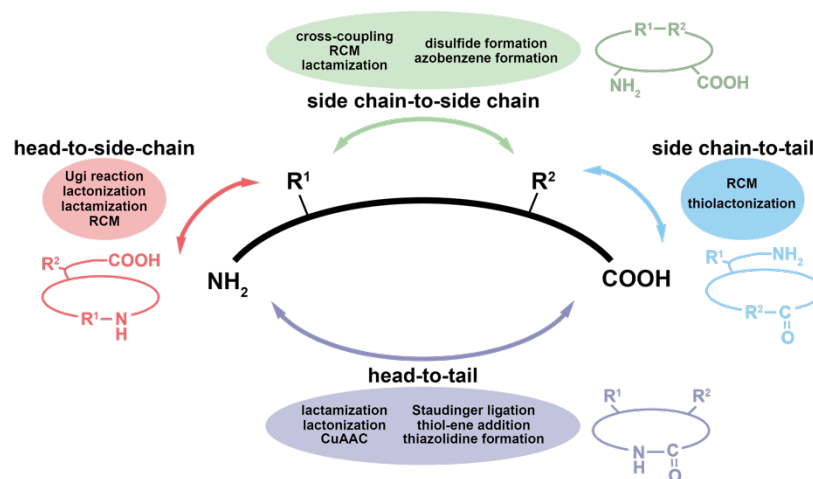

Figure 7: Macroyclization strategies of peptides. Peptides can be cyclized in four different manners: head-to-side-chain, side chain-to-side chain, side chain-to-tail, and head-to-tail. Figure adapted from [133].

First, Blackwell and Grubbs cross-linked O-allylserine residues in template peptides using Grubbs-catalysts.[134] This was later optimized by using α,α-disubstituted amino acids with olefin tethers, a strategy that is now one of the most used stapling techniques.[132] The first reported stapled peptides targeting a biological pathway were hydrocarbon stapled BH3 peptides binding to Bcl-2/3 domains which regulate apoptosis. Those peptides have higher stability, proteolysis resistance, and cellular permeability than their linear counterparts.[135] The design and synthesis of stapled peptides are relatively straightforward processes if structural or biochemical data of the binding site is known. The stapling amino acids are usually placed on the side of the helix, which does not play a role in the binding process to prevent binding disrupting interactions. Also, by designing the peptides based on a native binding partner, the stapled peptide should represent a "negative" version of the protein surface.[120] Initially, screening different stapling positions and strategies is recommended, thus scanning all sides of the helical surface. Also, using one-component (direct bond between the amino acids) and two-component stapling reactions (linking

two amino acids with a bifunctional linker) is possible.[127] The most used one-component stapling strategy involves *S*-pentenylalanine at *i*, *i+4* positions (or *R*-pentenylalanine at *i, i+4*) for one helical turn or *R*-octenylalanine with *S*-octenylalanine for *i, i+7* linkages covalently linked by ring-closing metathesis (RCM).[132] Other strategies with *i, i+3,* or *i, i+11* linkages were also reported (**Figure 8A**).[127,135–137] Further on, peptides can also be stapled at multiple positions, and a central bis-pentenylglycine can form a junction between two staples in *i,i+4+4,* or *i,i+4+7,* or *i, i+3+4+7* configurations.[138] The mentioned strategies all rely on a RCM reaction to generate hydrocarbon linkers, but covalent linkers between staples can also be achieved using lactamization,[139,140] cycloadditions,[141–143] thioethers,[144] or reversible disulfide bridges (**Figure 8B**).[145]
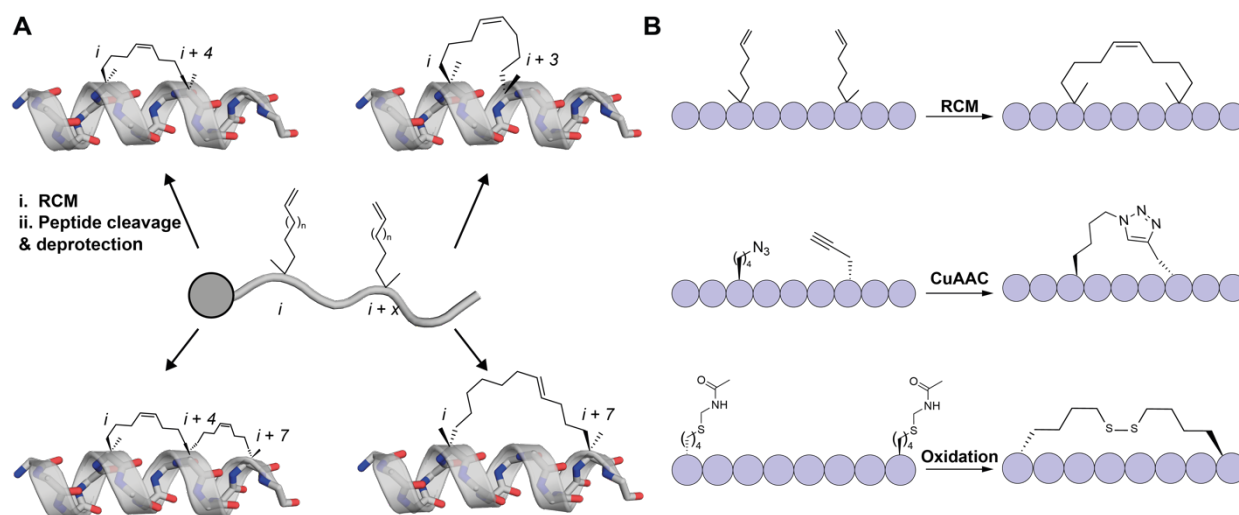

Figure 8: (**A**) Stapling strategies with different linker lengths for RCM like. (**B**) Examples for stapling of peptides through RCM, CuAAC, and reversible disulfide bridges.

MATERIALS AND METHODS

1. MATERIALS

Reagents

Reagents	Manufacturer
Acetic Acid	VWR International GmbH
Acrylamid/Bis Solution, 37.5:1 (30% w/v)	SERVA Electrophoresis GmbH
Agarose	SERVA Electrophoresis GmbH
DMEM	Sigma-Aldrich®
DTT (Dithiothreitol)	BioFroxx, neolab Migge GmbH
10 mM dNTP mix	Kapa Biosystems
EDTA (Ethylenediaminetetraacetic acid)	VWR International GmbH
Ethanol	Thermo Fisher Scientific
GeneRuler^{TM} 1 kb Plus DNA Ladder	Thermo Fisher Scientific
GeneRuler^{TM} DNA Ladder Mix	Thermo Fisher Scientific
Glycerol	VWR International GmbH
HEPES (2-[4-[2-hydroxyethyl)piperazin-1-1yl]ethanesulfonic acid)	Thermo Fisher Scientific
HCl (Hydrochloric acid)	Thermo Fisher Scientific
Imidazole	Carl Roth GmbH + Co. KG
IPTG (Isopropyl-β-D-thiogalactopyranosid)	Carl Roth GmbH + Co. KG
Isopropanol (Propan-2-ol)	Carl Roth GmbH + Co. KG
Lipofectamine^{TM} RNAiMAX Transfection Reagent	Thermo Fisher Scientific
NaCl (Natrium chloride)	Carl Roth GmbH + Co. KG
NaOH (Natrium hydroxide)	VWR International GmbH
NaF	Sigma-Aldrich®
PageRuler^{TM} prestained protein ladder	Thermo Fisher Scientific
PBS (cell culture)	Gibco, Thermo Fisher Scientific
Phenol:Chloroform:Iso-amyl alcohol (125:24:1)	Sigma-Aldrich®
PMSF	neoLab MIgge GmbH
Na$_2$HPO$_4$ (Disodium hydrogen phosphate)	Thermo Fisher Scientific
NaH$_2$PO4 (Monosodium phosphate)	VWR International GmbH
Skim Milk powder blotting grade	BioFroxx, neolab Migge GmbH
SYBR^{TM} Safe DNA Gel Stain	Thermo Fisher Scientific
TEMED (N,N,N',N'-tetramethylethane-1,2-diamine)	Fisher BioReagents
6x TriTrack DNA loading dye	Thermo Fisher Scientific
Tris (Tris-(hydroxymethyl)-aminomethan)	Carl Roth GmbH + Co. KG
Trypan Blue solution	Sigma-Aldrich®
Sodium Dodecyl Sulfate (SDS)	Sigma-Aldrich®
Tween 20	Acros Organics, Thermo Fisher Scientific

Kits

Kit	Manufacturer
GeneJET Plasmid Miniprep	Thermo Fisher Scientific
QIAquick PCR Purification	Qiagen
QIAquick Gel Extraction	Qiagen
RNeasy Mini Kit	Qiagen
Midi Kit	Qiagen
Trans-Blot Turbo RTA Mini 0.2 µM Nitrocellulose Transfer kit	BioRad
High-Capacity cDNA Reverse transcription kit	Thermo Fisher Scientific

Enzymes

Enzyme	Manufacturer
Phusion Polymerase	NEB, Thermo Fisher, in house
DpnI	Thermo Fisher
BamHI	Jena Bioscience
XhoI	Jena Bioscience
EcoRV	NEB
NdeI	Jena Bioscience
MfeI	NEB
Shrimp alkaline phosphatase	NEB
HotStart Taq DNA polymerase	NEB
T4 DNA polymerase	Thermo Fisher Scientific
Lysozyme from egg white	VWR International GmbH
DNAse I (protein purifications)	Sigma Aldrich
DNAse I (molecular biology)	Roche
T4 DNA Ligase	NEB
Taq DNA polymerase	Sigma Aldrich
TEV Protease	PCF Dortmund
3C precision protease	PCF Dortmund

Instruments

Instruments	Manufacturer
ÄKTA Explorer / ÄKTA Prime	GE Healthcare
Cell Density Meter, WPA Biowave, CO 8000	Biochrom
JASCO J-815 CD spectrophotometer	Jasco
Centrifuge 5417R	Eppendorf SE
Centrifuge 5424	Eppendorf SE
Centrifuge 5804R	Eppendorf SE
Centrifuge Avanti J-25	Beckman Coulter
Centrifuge Avanti J-26XP	Beckman Coulter
ChemiDoc MP Imaging System	Bio-Rad Laboratories
CleneCub	Herolab GmbH
Countess™ II automated cell counter	Thermo Fisher Scientific
Cytiva HisTrap™ HP prepacked column, 5 mL	Thermo Fisher Scientific
Cytiva HiTrap™ heparin HP column, 5 mL	Thermo Fisher Scientific
Dragonfly® Crystal	SPT Labtech
Drying / Heating oven ED53	Binder GmbH
Electrophoresis chamber (Mini-PROTEAN Tetra Vertical Electrophoresis Cell)	Bio-Rad Laboratories

Finnpipette^{TM} F1 Multichannel Pipette	Thermo Fisher Scientific
HiLoad^{TM} 16/600 Superdex^{TM} 75	Cytiva
HiLoad^{TM} 26/60 Superdex^{TM} 75	Cytiva
Incubator Shaker: New Brunswick Scientific^{TM} Innova^{TM} 42	Eppendorf SE
Incubator Shaker: New Brunswick Scientific^{TM} Innova^{TM} 4430	Eppendorf SE
JA-25.50 Fixed-Angle Aluminum Rotor	Beckman Coulter
JLA-8.1000 Fixed-Angle Aluminum Rotor	Beckman Coulter
Mosquito® Crystal (Pipetting robot)	SPT Labtech
NanoDrop 2000c Spectrophotometer	Thermo Fisher Scientific
Nalgene^{TM} Reusable Bottle Top Filters	Thermo Fisher Scientific
Orbital Shaker 3005	GFL
PCR Thermal Cycler – Mastercycler®	Eppendorf SE
Peristaltic pump LA-900	Landgraf Laborsysteme HLL GmbH
PowerPac^{TM} Basic Power Supply	Bio-Rad Laboratories
Rock Imager® (Crystal imager)	Formulatrix
Eppendorf Research plus Single Channel Pipettes	Eppendorf SE
Spark® multimode microplate reader	Tecan trading AG
Steri-Cycle^{TM} CO$_2$ Incubator	Thermo Fisher Scientific
SW shaking water bath	JULABO GmbH
ThermoMixer® comfort	Eppendorf SE
Trans-Blot Turbo Transfer System	Bio-Rad Laboratories
Vortex laboratory shaker RS-VA 10	Phoenix Insturment
ZEISS LSM 800 confocal microscope	Zeiss

Medium

Medium	Ingredients
LB medium (1 l)	10 g bacterial tryptone 5 g yeast 10 g NaCl
SOC medium (1 l)	20 g tryptone 5 g yeast extract 0.58 g NaCl 0.19 g KCl 2.03 g MgCl$_2$ x 6 H$_2$O 2.64 g MgSO$_2$ x 7 H$_2$O 40 mL glucose solution (50 % w/v)
SOB medium	20 g/l tryptone 5 g/l yeast extract 0.58 g/l NaCl 0.19 g/l KCl 2.03 g/l MgCl$_2$x 6 H$_2$O 2.46 g/l MgSO$_4$ x 7 H$_2$O
M9 Medium	200 ml/l 5X M9 salts 1 ml/l 1M MgSO$_4$ 8 ml/l 50% glucose 1 ml/l 0.1M CaCl$_2$
5X M9 salts	42.55 g/l Na$_2$HPO$_4$ 15.0 g/l KH$_2$PO$_4$

2.5 g/l NaCl	
5.0 g/l NH$_4$Cl	

Consumables

Consumable	Manufacturer
Adhesive foil for 96-well plates	VWR International GmbH
Amicon® Ultra-0.5 Centrifugal Filter Unit	Millipore®
Amicon® Ultra-4 Centrifugal Filter Unit	Millipore®
Amicon® Ultra-15 Centrifugal Filter Unit	Millipore®
Cell counting chips	Thermo Fisher Scientific
Cuvettes (Polystyrene, 10x4x45 mm)	Sarstedt AG & Co. KG
Dialysis membrane Spectra/Por® 3	VWR International GmbH
Falcon tubes (15 mL, 50 mL)	Sarstedt AG & Co. KG
Finntips	Thermo Fisher Scientific
Membrane Filters ME24 (mixed cellulose ester, 0.2 μm)	Whatman Cytiva
Omnifix® syringes	B. Braun SE
PCR strip of 8 (100 μL)	Sarstedt AG & Co. KG
Pipette tips (1 mL, 200 μL, 10 μL)	Sarstedt AG & Co. KG
SafeSeal reaction tubes (1.5 mL, 2 mL)	Sarstedt AG & Co. KG
Sterican® single-use hypodermic needles	B. Braun SE
Syringe filter, Filtropur S, 0.45 μM	Sarstedt AG & Co. KG
384-Well Solid Black Microplate	Corning
Serological Pipettes	Sarstedt AG & Co. KG
96-well qPCR plates	BioRad

Buffers

Buffer Name	Composition
Buffer A$_{PTBP1}$	50 mM Tris pH 8.0, 200 mM NaCl
Buffer B$_{PTBP1}$	50 mM Tris pH 8.0, 200 mM NaCl, 500 mM Imidazole
Low Salt$_{PTBP1}$	50 mM Tris pH 8.0
High Salt$_{PTBP1}$	50 mM Tris pH 8.0, 2 M NaCl
SEC buffer$_{PTBP1}$	50 mM Tris, 200 mM NaCl
CD buffer	20 mM sodium phosphate pH 8.0, 10 mM NaF
Buffer A$_{SRSF1}$	50 mM Na$_2$HPO$_4$ pH 8.0, 300 mM KCl, 50 mM L-Arg, 50 mM L-Gly, 1.5 mM MgCl$_2$
Buffer B$_{SRSF1}$	50 mM Na$_2$HPO$_4$ pH 8.0, 300 mM KCl, 50 mM L-Arg, 50 mM L-Gly, 1.5 mM MgCl$_2$, 500 mM Imidazole
Wash buffer$_{SRSF1}$	50 mM Na$_2$HPO$_4$ pH 8.0, 300 mM KCl, 50 mM L-Arg, 50 mM L-Gly, 1.5 mM MgCl$_2$, 40 mM Imidazole
Storage buffer$_{SRSF1}$	20 mM Na$_2$HPO$_4$ pH 8.0, 150 mM KCl, 50 mM L-Arg, 50 mM L-Gly, 1.5 mM MgCl$_2$, 0.2 mM EDTA, 1 mM TCEP

Buffer A$_{hnRNPA2/B1}$	50 mM HEPES pH 8.0, 300 mM NaCl, 20 mM Imidazole, 1 mM TCEP
Wash buffer$_{hnRNPA2/B1}$	50 mM HEPES pH 8.0, 300 mM NaCl, 30 mM Imidazole, 1 mM TCEP
Elution buffer$_{hnRNPA2/B1}$	50 mM HEPES pH 8.0, 300 mM NaCl, 500 mM Imidazole, 1 mM TCEP
Storage buffer$_{hnRNPA2/B1}$	50 mM HEPES pH 8.0, 100 mM NaCl, 1 mM TCEP
Buffer A$_{WDR5}$	50 mM HEPES pH 8.0, 300 mM NaCl, 30 mM imidazole, 1 mM TCEP
Buffer B$_{WDR5}$	50 mM HEPES pH 8.0, 300 mM NaCl, 500 mM imidazole, 1 mM TCEP
SEC buffer$_{WDR5}$	50 mM HEPES pH 8.0, 300 mM NaCl, 1 mM TCEP
FP buffer$_{PTBP1}$	20 mM Sodium phosphate buffer pH 8.0, 50 mM NaCl, 0.001% Tween20
FP-buffer$_{SRSF1}$	20 mM Na_2HPO_4 pH 7.0, 150 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 50 mM L-Glu, 50 mM L-Arg, 1 mM TCEP, 0.01 % Triton X-100
HKR solution	5mM HEPES pH 7.4, 137 mM NaCl, 2.68 mM KCl, 2.05 mM $MgCl_2$, 1.8 mM $CaCl_2$, 1g/l glucose
Coomassie Staining Solution	0.1% w/v Coomassie R-250, 40% Ethanol, 10% acetic acid
Destaining Solution	10% acetic acid
1X SDS Running Buffer /2 l	28.8 g glycine, 6.04 g Tris, 2 g SDS
PBS pH 7.4	137 mM NaCl, 27 mM KCl, 10 mM Sodium phosphate dibasic, 18 mM potassium phosphate monobasic
TBS	50 mM Tris-HCl pH 7.5, 150 mM NaCl
TBS-T	TBS + 0.1% Tween20
5x Laemmli Buffer	10 % w/v SDS, 50% v/v glycerol, 30 mM Tris-HCl pH 6.8, 0.05 % w/v bromophenol blue, 100 mM DTT
Stacking Gel Buffer	0.5 M Tris-HCl, pH 6.8
Resolving Gel Buffer	1.5 M Tris-HCl, pH 8.8
NP40 buffer	50 mM Tris pH 8.0, 150 mM NaCl, 1.0% NP-40
50X TAE / 1l	242g Tris, 57.1 ml acectic acid, 100 ml 0.5 M EDTA (pH 8.0)

Oligonucleotides

Primer Pair	Primer 1	Primer 2	Purpose
RRM1234	TTTCAGGGCCATATGGGATCCCCCTCTAGAGTGATCCACATCCG	CTCAAGCTTGAATTCCTCGAGTTAGACCGCCTCCTCCACGGAG	Cloning
RRM12	TTTCAGGGCCATATGGGATCCCCCTCTAGAGTGATCCACATCCG	CTCAAGCTTGAATTCCTCGAGTTAGACGTTGAGGCTGGTGAGCT	Cloning
RRM34	TTTCAGGGCCATATGGGATCCCCCCTGGCCATCCCC	CTCAAGCTTGAATTCCTCGAGTTAGACCGCCTCCTCCACGG	Clonging
RRM1*2	cccatctacatccagttcatcctccacaaggaggcgaagaccgacagctctcc	ggagagctgtcggtcttcgcctccttgtggaggatgaactggatgtagatggg	Mutagenesis
RRM12*	ggcacagtgttgaagatcagcaccgccacctcgaacaaccagttccaggcc	ggcctggaactggttgttcgaggtggcggtgctgatcttcaacactgtgcc	Mutagenesis
RRM12_L151G	ttcaccgcctgcccggccgcctgggc	gcccaggcggccgggcaggcggtgaa	Mutagenesis
RRM1_L151G	ttcaccgcctgcccggccgcctgggc	gcccaggcggccgggcaggcggtgaa	Mutagenesis
RRM1	TTTCAGGGCCATATGGGATCcggaaatgacagcaagaagttcaaagg	CTCAAGCTTGAATTCCTCGAaaTTAggccaggttccccg	Cloning
RRM2	TTTCAGGGCCATATGGGATCCatggccgggcagagc	CTCAAGCTTGAATTCCTCGAGTTAgtccccggaaggcagg	Cloning
PTBP2 e10	AGCTGCTGCTGGCCGAGTG	GATTGGTTTCCATCAGCCATCTG	RT-PCR
Rod1 e2	CTTCCTCTGCTCGCGGTTAG	AGGTCCGTTAATGATGCCAGA	RT-PCR
GPC2	GAGACTGAGGCCACCTTCCG	CTTCGCTGACCACATTTCTTCCA	RT-PCR
PTCH1	GACAGGACCGGGGAGTGC	TGTAAAACAGCAGAAAATATCCAGT	RT-PCR
TNIK	CAAAGGCGAGAGAAGGAGCTG	TCTTGCTTTAGCTGCCTCTGC	RT-PCR
ZNF711	TCCCAAAGAAAATAACATAGCACAGAAG	AGTCTGGCGTGTGCAATCCAAG	RT-PCR
FLAG-WDR5	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGATTATAAAGATGATGATGATAAATCCGCCACTCAGAGCAAGC	GCTTGCTCTGAGTGGCGGATTTATCATCATCATCTTTATAATCCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTC	Cloning
TRAP-mir29b2-d2	gaattcaaaagatcttttaagaaggagatatacatatgagATTTTTCCATCTTTGTATCgcgagtagcgaagacgttatc	gataacgtcttcgctactcgcGATACAAAGATGGAAAAATctcatatgtatatctccttcttaaaagatcttttgaattc	Cloning
TRAP-mir29b2-d5	gaattcaaaagatcttttaagaaggagatatacatatgagataATTTTTCCATCTTTGTATCgcgagtagcgaagacgttatc	gataacgtcttcgctactcgcGATACAAAGATGGAAAAATtatctcatatgtatatctccttcttaaaagatcttttgaattc	Cloning
TRAP-mir29b2-d8	gaattcaaaagatcttttaagaaggagatatacatatgagatacagATTTTTCCATCTTTGTATCgcgagtagcgaagacgttatc	gataacgtcttcgctactcgcGATACAAAGATGGAAAAATctgtatctcatatgtatatctccttcttaaaagatcttttgaattc	Cloning
TRAP-mir29b2-d14	gaattcaaaagatcttttaagaaggagatatacatatgagatacagatacgaATTTTTCCATCTTTGTATCgcgagtagcgaagacgttatc	gataacgtcttcgctactcgcGATACAAAGATGGAAAAATtcgtatctgtatctcatatgtatatctccttcttaaaagatcttttgaattc	Cloning
TRAP-RRM1	TAGCAaactggggcacaagcttaat GGC GGT GGC GGT TCG GTCCCCTCTAGAGTGATCCAC	TACTTATCGTCGTCATCCTTGTAATC CGCCTGGTTGGGAGAG	Cloning
TRAP-RRM2	TAGCAaactggggcacaagcttaat GGC GGT GGC GGT TCG GCGATGGCCGGGCA	TACTTATCGTCGTCATCCTTGTAATC GACGTTGAGGCTGGTGAG	Cloning
TRAP-RRM12	TAGCAaactggggcacaagcttaat GGC GGT GGC GGT TCG GTCCCCTCTAGAGTGATCCAC	TACTTATCGTCGTCATCCTTGTAATC GACGTTGAGGCTGGTGAG	Cloning
TRAP-RRM34	TAGCAaactggggcacaagcttaat GGC GGT GGC GGT TCGCCCCTGGCCATCCCTC	TACTTATCGTCGTCATCCTTGTAATC GACCGCCTCCTCCACGGA	Cloning
TRAP-GABA-d0	gaattcaaaagatcttttaagaaggagatatacat atg CAATTCTCTTTTCTGTCTACAAATCCAAAG gcgagtagcgaagacgttatc	gataacgtcttcgctactcgcCTTTGGATTTGTAGACAGAAAAGAGAATTGcatatgtatatctccttcttaaaagatcttttgaattc	Cloning
TRAP-GABA-d3	gaattcaaaagatcttttaagaaggagatatacat atg agtCAATTCTCTTTTCTGTCTACAAATCCAAAG gcgagtagcgaagacgttatc	gataacgtcttcgctactcgcCTTTGGATTTGTAGACAGAAAAGAGAATTGactcatatgtatatctccttcttaaaagatcttttgaattc	Cloning

26

TRAP-GABA-d6	gaattcaaaagatcttttaagaaggagatatacat atg agtggcCAATTCTCTTTTCTGTCTACAAATC CAAAG gcgagtagcgaagacgttatc	gataacgtcttcgctactcgcCTTTGGATTTGTAGACAGAA AAGAGAATTGgccactcatatgtatatctccttcttaaaagatcttttg aattc	Cloning
TRAP-GABA-d9	gaattcaaaagatcttttaagaaggagatatacat atg agtggcagtCAATTCTCTTTTCTGTCTACAAA TCCAAAG gcgagtagcgaagacgttatc	gataacgtcttcgctactcgcCTTTGGATTTGTAGACAGAA AAGAGAATTGactgccactcatatgtatatctccttcttaaaagatct tttgaattc	Cloning
TRAP-GABA-d12	gaattcaaaagatcttttaagaaggagatatacat atg agtggcagtggcCAATTCTCTTTTCTGTCTAC AAATCCAAAG gcgagtagcgaagacgttatc	gataacgtcttcgctactcgcCTTTGGATTTGTAGACAGAA AAGAGAATTGgccactgccactcatatgtatatctccttcttaaaag atcttttgaattc	Cloning

FP Probes

Name	Sequence	Vendor
RNA-1	FAM-GUCUUAA	Sigma Aldrich
RNA-2	FAM-AUUUUUCCAUCUUUGUAUC	IDT
Unlabeled-RNA2	AUUUUUCCAUCUUUGUAUC	IDT
RNA-S	Cy5-AGAAGAACAGAAGAACAGAAGAAC	Sigma Aldrich
RNA-AB	AAGGACUAGC-Cy5	Sigma Aldrich

Antibodies

Target	Vendor	Catalog #
Mouse anti-PTBP1	Santa Cruz Biotechnology, Inc.	sc-56701
Mouse anti-actin	Santa Cruz Biotechnology, Inc.	sc-8432
Goat anti-mouse HRP-conjugate	Abcam PLC	ab205719

siRNAs

Target	Vendor	Catalog #
PTBP1	Horizon Discovery Group plc	L-003528-00-0005
Control siRNA	Horizon Discovery Group plc	D-001810-01-05

Plasmid backbones

Backbone	Origin	Catalog #
pMAL-c6T	NEB	E8201
pBbA5k	Jay Keasling, Addgene	35282
pBbE8c	Jay Keasling, Addgene	35269
pBbB2a	Jay Keasling, Addgene	35343
pOPIN-His_SUMO	PCF Dortmund	-
pOPIN-NHis-3C	PCF Dortmund	-
pOPIN-His-MBP	PCF Dortmund	-
pOPIN-His-TRX	PCF Dortmund	-
pOPIN-His-GST	PCF Dortmund	-
pET19	PCF Dortmund	-

2. MOLECULAR BIOLOGY

Preparation of chemically competent cells

A pre-culture of a single colony of the *E. coli* strain of interest was grown in selective LB media at 37 °C, 160 rpm overnight. The next day, 2 l SOB media without antibiotics were inoculated 1:100 with the pre-culture and cultivated (37 °C, 160 rpm) until an OD_{600} of 0.4 was reached. The following steps were performed in chilled instruments and tubes, and buffers were chilled on ice. First, the culture was pelleted by centrifugation (5000 x g, 4 °C, 10 min) and the media was replaced with 10 % culture volume of 0.1 M $MgCl_2$. The suspension was centrifuged twice, and the pellet was resuspended in 10 % culture volume of 0.1 M $CaCl_2$. After the final centrifugation, the pellet was resuspended in 1 % of the culture volume of 0.1 M $CaCl_2$ supplemented with 15 % glycerol. The suspension was split into 50 µl aliquots, flash-frozen in liquid nitrogen, and stored at -80 °C until use.

Preparation of electrocompetent cells

Cultures were cultivated like for chemically competent cells. The cultures were harvested by centrifugation and the medium was replaced three times with 0.1 culture volume of 10% glycerol followed by centrifugation. Cells were flash-frozen in liquid nitrogen and stored at -80 °C in 50 µl until use.

Bacterial transformations by heat shock

Chemically competent *E. coli* strains (50 µl) were thawed on ice and 10 – 300 ng of plasmid DNA were added (10 min, 0 °C). Afterward, plasmid uptake was induced by heat shock (42 °C, 45 sec) followed by immediate cooling of the mixture by incubation on ice for 2 min. The bacteria were recovered by adding 300 µl SOC media and incubated at 37 °C for 45 min. The cells were pelleted by centrifugation (1000 x g for 2 min) and the medium was removed, followed by resuspension of the cell pellet in the residual 50 µl and plating on selective LB-agar plates.

Bacterial transformations with electroporation

Electrocompetent *E. coli* cells (50 µl) were thawed on ice and 10 – 300 ng of plasmid DNA were added. The cells were immediately transformed using a BioRad electroporator in 1 mm cuvettes with bacteria settings. The cells were rescued by adding 300 µl SOC media, followed by incubation at 37 °C for 45 min and plating on selective LB-agar plates.

Polymerase chain reactions

Polymerase chain reactions (PCR) were performed with Phusion polymerase (NEB) according to the manufacturer's protocol (**Table 3**).

Table 3: Setup of a standard PCR using Phusion polymerase (top) and the thermocycler (bottom) setup.

Component	Final amount or concentration
5X HF Buffer	1x
10 mM dNTP mix	0.2 mM
10 µM forward primer	0.5 µM
10 µM reverse primer	0.5 µM
Phusion polymerase	1 U
Template	5 – 50 ng
Nuclease free H_2O	Fill up to 20 or 50 µl

Step	Temperature / °C	Time / mm:ss	
Melting	95	00:30	
Melting	95	00:30	
Annealing	Primer dependent	0:20	25 x
Extension	72	20 s / kb	
Hold	4	∞	

Side-directed mutagenesis

Side-directed mutagenesis was performed like standard PCRs in two 25 µl reactions with complementary reverse and forward primers containing the mutation of interest. The cycling was performed like in a PCR, but both primers were added to their own reaction tube. After the initial three cycles, both reactions were pooled, and the reaction continued for ten more cycles. The reaction mix was afterward DpnI digested by adding 5 U DpnI to the mix. It was transformed into chemical competent *E. coli* cloning strains followed by plating on selective LB-agar plates.

Subcloning with sequence and ligation-independent cloning

Sequence and ligation independent cloning (SLIC) was performed as published by Li et al.[146] The vector was either linearized through inverse PCR or by restriction digest and PCR purified (PCR) or gel extracted (Restriction digest or PCR) depending on the quality of the product. If the vector was generated by PCR, it was further DpnI digested by adding 5 U of DpnI to the mix and incubating at 37 °C for 1 h. Insert was generated by PCR with primers containing homology

regions in the 5' overhang and was also DpnI digested. From both vector and insert, 1 μg was digested with 0.5 U T4 DNA polymerase in 1X NEBuffer 2 (Rt, 30 min). The reaction was quenched by adding 0.1 V 10 mM dCTP. Afterward, 100 ng vector was mixed with insert in 1:1, 1:3, or 1:5 ratio in 1X T4 DNA ligase buffer (37 °C, 30 min). The mixture was immediately transformed into chemical competent *E. coli* cloning strains and plated on selective agar plates. Plasmids from single clones were isolated through MiniPrep from 5 ml cultures in selective LB-medium and validated through Sanger sequencing.

Subcloning with circular polymerase extension cloning

If SLIC cloning failed in different vector:insert ratios, a subsequent circular polymerase extension cloning (CPEC) was tried.[147,148] Briefly, CPEC was performed like a PCR with fusion polymerase with 100 ng vector and a 1:1 molar ratio of vector:insert. No primers are added in this method because the insert and vector are priming each other (**Table 4**). The reactions were afterward directly transformed into *E. coli* cloning strains and plated onto selective LB-agar plates. Plasmids from single clones were isolated through MiniPrep from 5 ml cultures in selective LB-medium and validated through Sanger sequencing.

Table 4: Setup of a standard CPEC reaction using Phusion polymerase (top) and the thermocycler (bottom) setup.

Component	Final amount or concentration
5X HF Buffer	1x
10 mM dNTP mix	0.2 mM
Insert	Equimolar to vector
Phusion polymerase	1 U
Linearized vector	100 ng
Nuclease free H_2O	Fill up to 20 μl

Step	Temperature / °C	Time /mm:ss	
Melting	95	00:30	
Melting	95	00:30	
Annealing	58-65 (depends on construct)	0:20	10 x
Extension	72	20 s / kb (final product)	
Hold	4	∞	

Ethanol Precipitation

DNA samples were further purified or concentrated through ethanol precipitation. Therefore, 0.1 volumes of 3 M sodium acetate, 2 µl glycoblue (ThermoFisher Scientific) and 2 volumes of ice-cold 100 % ethanol were added. The sample was stored at -80 °C for at least 1 hour to increase precipitation. The DNA was pelleted by centrifugation (21000 x g, 4 °C, 30 min) and the pellet was washed twice with 500 µl ice-cold 70 % ethanol. The supernatant was removed, and the pellet was dried by storing the tube with an open lid at Rt. The pellet was then resuspended in nuclease-free water in a volume appropriate for the following steps.

SICLOPPS library cloning

The generation of the SICLOPPS library was adapted from Tavassoli and Benkovic for the pBbB2a vector.[149] The *Synechocystis* sp PCC6803 DnaE split intein gene was synthesized as gene blocks and subcloned into the pBbB2a vector. Therefore, the vector was linearized by inverse PCR with primers pBbB2a_lin_fw and pBbB2a_lin_rev. The two inserts were hybridized in 1xHF buffer and a fused insert generated by standard PCR using Phusion polymerase and primers SICLOPPS_fw and SICLOPPS_rev. This insert was cloned into the linearized vector with SLIC in a 1:3 ratio (vector:insert).

The pBbB2a_SICLOPPS vector was amplified in 50 ml LB+Amp, isolated by Midi-Prep (Qiagen), and the whole batch triple digested with the restriction enzymes MfeI, XhoI, and EcoRV and simultaneously dephosphorylated using shrimp alkaline phosphatase. The reaction was inactivated (80 °C, 10 min) and separated in a 1 % agarose-TAE gel. The band corresponding to the linearized vector was extracted with a Gelextraction kit.

The insert containing the randomized $(NNS)_5$ sequence was generated by PCR with Phusion polymerase PCR in 5xHF buffer using CBD-r_XhoI and C5_library_primer and the pBbB2a_SICLOPPS plasmid as template. The PCR product was used in a subsequent PCR with Zipper_primer and CBD-r_XhoI after PCR purification and subsequently double-digested with MfeI and XhoI.

Insert and vector were further purified through ethanol precipitation and subsequently ligated using 125 ng vector and 100 ng insert in 10 x 50 µl reactions and 1 U ligase per 100 ng vector (16 °C, overnight). The reaction was purified by ethanol precipitation and the pellet was resuspended in 10 µl DNAse free water. The product was transformed into electrocompetent *E. coli DH5alpha* cells (ThermoFisher Scientific) by adding 1 µl of DNA to 50 µl of cells. The recovered cells were pooled and plated onto selective agar plates in dilution series of up to $1*10^6$ and the residual

culture was used to inoculate 100 ml LB+Amp. After one night at 37 °C, 180 rpm, the plasmid library was isolated using a MidiPrep kit (Qiagen). Twenty individual clones from the plates used for the dilution series were sequenced to validate the sequence of the library containing stock by Sanger sequencing.

Table 5: Oligonucleotides used for the generation of a $(NNS)_5$ SICLOPPS library.

Primer Name	Sequence
pBbB2a_lin_fw	ggatccaaactcgagtaaggatc
pBbB2a_lin_rev	atgtatatctccttcttaaaagatcttttgaatt
SICLOPPS_fw	aattcaaaagatcttttaagaaggagatatacatATGGTTAAAGTTATCGGTCGTCGTTC
SICLOPPS_rev	gatccttactcgagtttggatccttattgaag
CBD-r_XhoI	CTCGAGTCATTGAAGCTGCCACAAGG
Zipper_primer	GGAATTCGCCAATGGGGCGATCGCC
C5_library_primer	GGAATTCGCCAATGGGGCGATCGCCCACAATTGCNNSNNSNNSNNSNNSTGCTTAAGTTTTGGC

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed to analyze the purity of purified proteins and for Western blotting. Gels were manually cast using the BioRad Mini-PROTEAN cells. First, the resolving gel was cast between two glass-slides and topped with isopropanol, and then the stacking gel was cast on top after removing the isopropanol. The recipe for the gels is shown in **Table 6**.

Table 6: Recipe for SDS-PAGEs.

	Stacking Gel 4%	Resolving Gel X%
Component	*Volume*	*Volume*
30% Acrylamide/bis	1.98 ml	0.5 x X ml
Stacking Gel buffer	3.78 ml	-
Resolving Gel buffer	-	3.75 ml
10% SDS	150 µl	150 µl
H_2O	9 ml	11.03-(0.5 x X) ml
TEMED	15 µl	7.5 µl
10% APS	75 µl	75 µl

Protein samples were diluted to appropriate concentrations for loading and Laemmli buffer was added to a final concentration of 1X. From those samples, 10 µl were injected into the pockets of the gel and the PAGE was run in 1x Running buffer at 140 V. Gels were removed from the glass slides, washed, and stained with Coomassie staining solution by microwaving shortly and incubating for 3 min on a rocking shaker. The gel was destained using either water or destaining solution after microwaving. Gels were imaged using a ChemiDoc MP imaging system.

Agarose gel electrophoresis for nucleic acids

Agarose gels were prepared by melting 1 – 2% w/v agarose in 1x TAE buffer and pre-stained with Ethidium bromide or SybrSafe according to the manufacturer's instructions. Gels were cast using the BioRad Wide Mini-Sub Cell GT System or Mini-Sub Cell GT System. Samples were prepared by mixing 10 µl of pre-diluted DNA with 6x loading dye to reach 1x loading dye concentration. The gels were run at 100 V (const.) until the lowest loading dye front reached the bottom of the gel or until a sufficient separation was reached.

TRAP assay

Chemical competent *E. coli* Top10F' were co-transformed with two assay plasmids: the reporter and the protein construct, and selected on LB-Amp with kanamycin (50 µg/ml) and chloramphenicol (34 µg/ml). Single colonies were picked and used to inoculate a pre-culture of 1 ml in LB with the same antibiotics. The cultures were sealed with a semi-permeable lid and cultivated in a deep-well 96-well block overnight at 37°C, 160 rpm. These pre-cultures were used to inoculate M9 medium 1:19 to a final volume of 200 µl, and the assay cultures were cultivated (37 °C, 160 rpm) until an OD_{600} of ~0.2 was reached. The expression of reporter and expression construct was induced by adding 1 mM IPTG and titration of L-arabinose (L-ara) with final concentrations of 0, 0.125, 0.25, 0,5, and 1%. The culture was cultivated in a TECAN spark plate reader at 30 °C while shaking and readout of the OD_{600} and fluorescence intensities of both fluorophores every 20 min. Data analysis was performed as described by Katz et al..[150] Briefly, RBP expression as measured through the TagBFP signal, was normalized to the OD_{600} (1).

$$Averaged\ TagBFP = \frac{1}{N} \sum_{t=T_0}^{T_{final}} \frac{TagBFP(t)}{OD_{600}(t)} \qquad (1)$$

The reporter rate of production was calculated with equation (2).

$$reporter\ production\ rate = \frac{reporter(T_{final}) - reporter(T_0)}{\int_{T_0}^{T_{final}} dt \cdot OD(t)} \qquad (2)$$

Finally, the reporter production rate was plotted against the averaged TagBFP.

Fluorescence-assisted cell sorting

Pre-cultures were grown like for the TRAP assay, and the assay culture was cultivated in LB-media supplemented with antibiotics. At an OD_{600} of 0.2, expression was induced using 1 mM IPTG and 0.5 % L-ara, and expression was performed for 6 h at 30 °C. The expression culture was pelleted by centrifugation (4500 x g, 4 °C, 5 min), cells were washed twice with 1x PBS, and the final pellet was resuspended in 800 µl PBS. The samples were analyzed using an SH800SFP Cell sorter (Sony Biotechnology) with a 70 µm microfluidic chip, and reporter and expression fluorescence intensities were recorded and compensated using respective single-color controls.

3. PROTEIN PURIFICATION

Expression of PTBP1 variants in *E. coli*

All proteins were purified using the following protocol: *E. coli* BL21 DE3 RIL were transformed with plasmids encoding for the protein of interest by heat shock. Afterward, a pre-culture in LB medium (100 µg/ml Amp) was inoculated with a fresh colony and grown overnight (37 °C, 180 rpm). An expression culture was inoculated 1:200 with the pre-culture and cultivated at 37°C, 180 rpm until an OD_{600} of ca. 0.6 was reached. After chilling (4 °C, 30 min), protein expression was induced with 200 µM IPTG overnight at 20 °C, 180 rpm. Cultures were harvested at 5000 x g.

Purification of PTBP1 variants

The harvested cell pellets were suspended in Buffer A_{PTBP1} with 0.1 mM PMSF, and then lysozyme and DNAse I (Sigma Aldrich) were added. The suspension was incubated at 4 °C for 30 minutes while stirring and then sonicated for lysis (70 % amplitude, 10 sec on, 10 sec off, 10 min). The solution was cleared by centrifugation (60000 x g, 1 hour, 4 °C) and filtered through a 0.22 µm filter. All FPLC-based methods were performed using an ÄKTA Explorer or ÄKTA prime FPLC system (GE Healthcare). The protein was purified by nickel affinity chromatography (His-Trap 5 ml, GE Healthcare) using a gradient of Buffer B_{PTBP1}. If necessary, tag-cleavage was performed during dialysis against Buffer A_{PTBP1} with either His-tagged TEV-protease or His-tagged 3C-protease. Non-cleaved protein and protease were removed by collecting the unbound fraction of a Ni-NTA column. If further purification with ion-exchange was needed, the solution was then diluted in 50 mM Tris pH 8.0 to reach a NaCl concentration of ≤ 50 mM and separated on a Heparin column (Heparin HP 5 ml, GE Healthcare) using a gradient of high salt$_{PTBP1}$ buffer. Finally, size

exclusion chromatography (Superdex 75 16/60 or 26/60) was performed with SEC buffer$_{PTBP1}$. Therefore, the mixture was concentrated to a final volume ≤ 3 ml and particles were removed by centrifugation (16000 x g, 4 °C, 10 min) and injected to the column using a 5 ml loop. Protein homogeneity was determined by the size exclusion chromatogram, and purity was confirmed to be ≥ 90% by SDS-PAGE analysis if not otherwise mentioned.

Expression of SRSF1

After subcloning the gene that encodes for SRSF1-RRM12 (1-195) into the pOPIN-His expression vector, expression was performed in *E. coli* BL21 (DE3). To do this, a pre-culture of transformed bacteria was created, followed by expression in LB-Medium with 100 µg/ml ampicillin and 1 mM IPTG at 18 °C and 160 rpm overnight, after growing the culture at 37 °C and 160 rpm to an OD$_{600}$ of 0.6. Finally, the cells were collected by centrifugation (5000 x g, 4 °C, 15 min).

Purification of SRSF1

Protein purifications of SRSF1 were performed by Gulshan Amrahova ('t Hart Group, CGC Dortmund).

After resuspension of cells in Buffer A$_{SRSF1}$ supplemented with 1 mM PMSF, lysis was performed through sonification. The protein was purified by Ni-affinity chromatography (HisTrap HP 5 ml column (GE-Healthcare)) with a gradient of Buffer B$_{SRSF1}$. This was followed by dialysis into wash buffer$_{SRSF1}$ and a subsequent second affinity chromatography step. The protein-containing fractions were dialyzed into wash buffer, and the affinity tag was cleaved using His-tagged 3C protease. The cleaved construct was subsequently used for a third affinity column run collecting the unbound fraction to remove uncleaved products, protease, and the low molecular weight tag. The final product was dialyzed into storage buffer$_{SRSF1}$.

Expression of hnRNP A2/B1

The MBP-tagged hnRNP A2/B1 construct (1-251) was expressed from a pOPIN-His-MBP vector in *E. coli* BL21 (DE3) RIPL cells. The expression culture was inoculated from a pre-culture in TB (0.01% lactose, 2 mM MgSO$_4$, 100 µg/ml ampicillin, 50 µg/ml chloramphenicol) with a starting OD$_{600}$ of 0.05 and cultivated at 37 °C, 180 rpm. After autoinduction at 37 °C after 4 h, expression was performed overnight at 25 °C, 180 rpm. Subsequently, the culture was harvested by centrifugation (5000 x g, 15 min, 4 °C).

Purification of hnRNP A2/B1

Protein purifications of hnRNP A2/B1 were performed by Gulshan Amrahova ('t Hart Group, CGC Dortmund).

Harvested cells were resuspended in Buffer $A_{hnRNPA2/B1}$ supplemented with 1 mM PMSF and lysed with a microfluidizer. The first purification step was done with a HisTrap FF crude 5 ml column equilibrated with Buffer $A_{hnRNPA2/B1}$. After washing with wash buffer$_{hnRNPA2/B1}$, the protein eluted with elution buffer$_{hnRNPA2/B1}$. The protein-containing fractions were further purified with size exclusion chromatography (Superdex 75 26/60 prep grade column) into storage buffer$_{hnRNPA2/B1}$.

Expression of WDR5 variants

The WDR5 constructs were expressed from a pET19 vector with a *C*-terminal 3C-protease recognition site followed by a His-tag. A preculture of a single transformed colony of *E. coli* BL21 (DE3) RIL was used 1:200 to inoculate an expression culture. This culture was cultivated until the OD_{600} reached 0.6 (37 °C, 180 rpm) and induced with 200 μM IPTG following expression overnight (20 °C, 180 rpm). The cells were pelleted by centrifugation (5000 x g, 15 min, 4 °C), the supernatant was removed, and the pellet was stored at -80 °C after flash freezing in liquid nitrogen until use.

Purification of WDR5 variants

The harvested cultures (5000 x g, 15 min, 4 °C) were resuspended in Buffer A_{WDR5} supplemented with 1 mM PMSF and lysed by sonification. The cleared lysate (60000 x g, 45 min, 4 °C) was filtered and purified by nickel affinity chromatography using a HisTrap HP 5 ml column (GE Healthcare) and a gradient of Buffer B_{WDR5}. The protein-containing fractions were pooled and diluted 1:5 to avoid precipitation and dialyzed against Buffer A_{WDR5} with 3C protease for tag-cleavage (4 °C, overnight). Afterward, the unbound fraction of a subsequent affinity chromatography run was collected and further purified by size exclusion chromatography (Superdex 75 16/60, GE Healthcare) using SEC buffer$_{WDR5}$. The eluted protein was concentrated using spin columns, flash-frozen in liquid nitrogen, and stored at -80 °C until use.

4. BIOPHYSICAL METHODS

Fluorescence polarization assay with PTBP1

After buffer exchange of stored protein with spin columns into FP buffer$_{PTBP1}$, the protein was serially diluted 1:1 in FP buffer$_{PTBP1}$ for protein-RNA binding assays. Fluorescently FAM-labeled

RNA was added to a final concentration of 5 nM in a total volume of 20 µl. After equilibrating (20 °C., 30 min), the fluorescence polarization was measured in a plate reader (TECAN Spark, monochromators: Exc. 490 ± 10 nm; Em. 520 ± 10 nm). To perform competitive assays, protein concentrations corresponding to 50-70 % binding of the probe were added to 10 nM RNA. Competitors were dissolved in FP buffer$_{PTBP1}$ and 1:1 serially diluted before adding pre-incubated protein-RNA complex (final RNA concentration 5 nM), followed by equilibration (20 °C, 30 min) and read out as mentioned before. Experiments were performed in triplicates.

Data was normalized to inhibition (%) using the following equation $Inhibition\% = \frac{FP-FP_{min}}{FP_{max}-FP_{min}} * 100$ and fitted using GraphPad Prism.

Fluorescence polarization assays with purified PTBP1 constructs and FITC-labeled peptides were performed in FP buffer$_{PTBP1}$. The protein buffers were exchanged into FP buffers using spin columns and serially diluted in the same buffer. Afterward, 20 nM fluorescently labeled **P-6F1** was added to reach a final concentration of 10 nM probe. Fluorescence polarization was read out in a TECAN Spark with the settings mentioned previously.

Fluorescence polarization assay with SRSF1

Fluorescence polarization assays with SRSF1 were performed by Gulshan Amrahova ('t Hart Group, CGC Dortmund).

Purified proteins were buffer exchanged into FP-buffer$_{SRSF1}$ using spin columns. For the competitive assays, 1 nM Cy5-labeled RNA-S and 400 nM SRSF1 (corresponding to 50 – 70% binding) were added to serially diluted competitors dissolved in FP-buffer$_{SRSF1}$. Fluorescence polarization was measured in a plate reader (TECAN Spark; filters: Exc. 610 ± 20 nm; Em. 670 ± 25 nm).

Fluorescence polarization assay with hnRNP A2/B1

Fluorescence polarization assays with hnRNP A2/B1 were performed by Gulshan Amrahova ('t Hart Group, CGC Dortmund).

Using spin columns, purified hnRNP A2/B1 was exchanged into FP-buffer$_{hnRNPA2/B1}$ (25 mM HEPES pH 8.0, 100 mM NaCl, 1 mM TCEP, 0.01 % Triton X-100). For the competitive assay 10 nM hnRNP A2/B1 and 1 nM of Cy5-labeled RNA-AB were added to serially diluted competitors in FP-buffer$_{hnRNPA2/B1.}$ Fluorescence polarization was measured in a plate reader (TECAN Spark; filters: Exc. 610 ± 20 nm; Em. 670 ± 25 nm).

Microscale thermophoresis

The protein constructs underwent buffer exchange using spin columns to transfer them to the FP buffer. Then, they were diluted serially into the same buffer. Peptide **P-6F1** was dissolved in FP buffer containing 0.04 % Tween 20. Next, **P-6F1** was added to the serially diluted protein to achieve a final concentration of 100 nM. The solutions were loaded into Monolith NT.115 premium capillaries and subsequently measured with the blue light source and high MST-Power on a Monolith NT.115 instrument. The data obtained were analyzed at time points with the best signal-to-noise ratio and evaluated using the software provided by the manufacturer. All experiments were conducted in duplicates.

Thermal shift assays

Thermal shift assays were performed using SYPRO Orange in white plate, transparent bottom qPCR plates (BioRad) with a total reaction volume of 50 µl. Peptide solutions (25 µl in SEC buffer) were added to the plates. Afterward, Protein-SYPRO Orange mix with final concentrations of 1 – 10 µM PTBP1 and 1 – 3 X SYPRO Orange were added. Melting curves were measured using a BioRad CFX96 qPCR cycler and analyzed using CFX Maestro Software (BioRad). The melting points were determined as the minima of the first derivative of the fluorescent signal (-dF/dT) plotted against the temperature.

Circular Dichroism Spectroscopy

Peptide solutions of 50 µM were measured for circular dichroism in CD buffer in triplicates at 20°C with a JASCO J-815 CD spectrometer and a 1 mm pathlength. JASCO multivariate secondary structure analysis with a reference data set was used to calculate the secondary structure composition of the peptides.

For experiments with PTBP1 constructs, the protein was buffer exchanged into CD buffer using spin columns, and measurements were performed in this buffer under similar conditions with protein concentrations of 0.2 mg/ml.

Co-Crystallisation of RRM1Δα3 with P-6

Protein construct RRM1Δα3 was purified using the aforementioned protocol and concentrated to 14 mg/ml. **P-6** was dissolved in SEC buffer and added in a 1.5 molar excess to the protein solution. After removing particles via centrifugation (20000 x g, 4 °C, 10 min), sitting-drop experiments were initiated by adding 200 nl of protein/ligand to 100 nl of reservoir solution (1.89 M $(NH_4)_2SO_4$, 0.1 M HEPES pH 6.86, 2% v/v PEG400) in MRC 3-drop plates. Sealed plates were incubated at 20

°C using a Formulatrix RockImager, and crystals were obtained after several days. Before flash-freezing, the crystals were cryo-conserved by adding 0.5 μl reservoir solution containing 20% v/v glycerol to the drop.

X-ray data collection and processing of RRM1Δα3 with P-6

X-ray diffraction data was collected at X10SA (PXII) at Swiss Light Source. Datasets were integrated using the XDS package and scaled using XSCALE.[151] The structure was solved using phaser (Phenix) and a truncated Alphafold model of RRM1 with the helical peptide present.

Structure solution and refinement of RRM1Δα3 with P-6

RRM1Δα3•**P-6** crystallized in P 2_1 2_1 2 spacegroup with dimensions 244.37 x 76.83 x 94.19 Å with 32 molecules in the asymmetric unit (16 dimers of protein bound to peptide). The structure was solved using a truncated model of human PTBP1 RRM1 from AlphaFold with Phaser (Phenix Suite).[152,153] The resulting model was refined in iterations of phenix.refine and manual model building in Coot.[153,154] The structure was refined to a final R_{free} of 34% at 2.9 Å.

Co-Crystallisation of WDR5 with WDR5-P1

Purified WDR5 was concentrated to 23.25 mg/ml, and ligand **WDR5-P1** dissolved in gel filtration buffer was added in a 1.5-fold molar excess. Crystallization was set up in MRC-3 drop plates by adding 100 nl protein/ligand complex to 100 nl reservoir solution, and the plates were stored at 20°C. Crystals were obtained after a few days with reservoir conditions of 0.2 M Li_3-citrate and 20 % w/v PEG3350 and were cryo-conserved in reservoir solution supplemented with 20 % glycerol.

X-ray data collection and processing of WDR5 with WDR5-P1

Diffraction data was collected at beamline ID30A-3 at ESRF Grenoble (acquisition date 18.11.2022). Datasets were integrated using XIA2/DIALS (ccp4).[155,156] The structure was solved using Phaser (Phenix) and an AlphaFold model of apo-WDR5.[152,153]

Structure solution and refinement of WDR5 with WDR5-P1

WDR5-P1 co-crystallized in spacegroup P 4_3 2_1 2 with dimensions 82.1737 x 82.1737 x 201.707 Å with two protein-peptide dimers in the asymmetric unit. The solved structure was refined by iterations of phenix.refine and manual model building in coot.[153,154] The structure was refined to a final Rfree of 24 % at 1.84 Å.

5. CELL CULTURE

Subculturing of cells

All cell lines were subcultured in 10 cm dishes containing 10 ml DMEM supplemented with 10% FBS with 1x Pen-Strep (Gibco) at 37 °C, 5% CO_2, and subcultured before reaching 80% confluency. Adherent cells were washed with 10 ml 1x PBS. After removal of PBS, 3 ml 0.05% Trypsin-EDTA (Gibco) was added, and the plate was incubated at 37 °C, 5% CO_2, until fully detached. The proteolysis was quenched by topping up the suspension with 7 ml medium. Next, the cells were pelleted by centrifugation (500 x *g*, 3 min, Rt), following the removal of the medium, leaving 300-500 μl of liquid. The cells were carefully resuspended with a 1250 μl pipette tip and counted using a Countess II instrument after 1:1 trypan blue staining. Finally, cells were seeded into 10 ml of growth media at the anticipated density.

RT-PCR assays

HEK293T cells were seeded with a 20000 cells/well density in 96-well plates. Subsequently, cells were treated with a final 100-300 μM peptide concentration with a final DMSO concentration of 0.5%. After 24 and 48 hours of treatment, total RNA was isolated using RNeasy Mini Kits. The RNA was precipitated by adding 0.1 volume 3M NaAc, 3 volumes ice-cold ethanol, and GlycoBlue™ co-precipitant (ThermoFisher Scientific). After chilling at -20°C for 30 min, the RNA was pelleted by centrifugation in a table-top centrifuge (30 min, 4 °C, 14000 x *g*). The pellet was washed twice with 500 μl ice-cold 70 % ethanol followed by centrifugation (10 min, 4 °C, 14000 x *g*). After removing the supernatant, the pellet was air-dried in an RNase-free environment, and the pellet was resolubilized using 10 μl nuclease-free water. Next, 0.5 μg of RNA was reverse transcribed using a High-Capacity cDNA Reverse transcription kit (Thermo Fisher). One microliter of the cDNA mix was amplified with Hot Start Taq DNA polymerase (NEB) for 45 cycles. From these reactions, 20 μl were analyzed on 2% agarose gels (1x TAE) pre-stained with ethidium bromide (Sigma Aldrich).

Cell Viability assays

The day before the experiment, HEK293T cells were seeded in 96-well plates at a density of 30000 cells/well in 75 μl DMEM supplemented with 10% FBS and 1x Pen-Strep (Gibco). Once adhered after overnight, the cells were treated with a dilution series of peptides in DMSO (final DMSO concentration 0.5%). Cell viability was determined using CellTiterGlo 2.0 (Promega) according to the manufacturer's instructions, and the results were expressed as a percentage of the signal normalized to the DMSO controls.

Cell Permeability by confocal microscopy

The day before the experiment, HEK293T cells were seeded at a concentration of 10,000 cells in 8-well µ-Slides (ibidi). The cells were washed with 1x PBS and HKR solution the following day.[157] Next, FITC-labeled peptides were added to a final concentration of 10 µM and incubated for 1 hour. After 5 washes with HKR buffer, the cells were fixed with 4% formaldehyde (ROTHI®Histofix, ROTH) at 37 °C for 10 minutes, then washed with HKR buffer and stored in 1x PBS. Cells were imaged using a ZEISS LSM 800 at 40x magnification with oil immersion.

Lysate stability assay

Lysate stability assays were performed together with Gulshan Amrahova ('t Hart Group, CGC Dortmund).

To measure peptide stability in HEK293T lysates, cell pellets were stored at -80 °C until use. They were then resuspended in 1x PBS and lysed through 3 cycles of freezing in liquid nitrogen and thawing at 37 °C. After centrifugation (16873 x g, 10 min), the supernatant was used for the stability experiments. Time course experiments were carried out by incubating 140 µl of a mixture of 1 mM peptide and 1 mg/ml lysate at 37 °C. At specific intervals, 10 µl samples were taken and mixed immediately in equal volumes with ice-cold ethylparaben in MeOH (0.05 mg/ml). The samples were then chilled on ice for 15 min and centrifuged (16873 x g, 4 °C, 10 min) to remove insoluble material. The supernatant was analyzed by LCMS with a gradient of H_2O (0.1% TFA) and ACN (0.1% TFA) from 5 to 95 % over 20 min. The area under the curve of the peaks was integrated and normalized to the ethylparaben standard.

Knockdown with siRNA

For knockdowns, 30000 cells/well HEK293T cells were seeded in 6 well plates. After adhering overnight, knockdowns with PTBP1-siRNA (Dharmacon) or control siRNA (Dharmacon) were performed with Lipofectamine RNAimax (ThermoFisher) according to the manufacturer's instructions. Cells were cultivated for 48 h, and lysates were prepared for Western blotting.

Western Blotting

After culturing of cells in 6-well plates, the cultures were washed with ice-cold PBS twice. Then, 200 µl NP40-Buffer supplemented with 1X Complete protease inhibitor cocktail (Takara Bio) were added to the cells and incubated for 20 min at 4 °C on a rocking table. The mixture was resuspended, added to microcentrifuge tubes, and further incubated on a rotary shaker at 4 °C for 30 min. Lysis was completed by rigorous resuspension with a P20 pipette, and the debris was

41

removed by centrifugation (20000 x g, 4 °C, 10 min). The protein concentration of the lysate was determined using the DC Assay (BioRad) according to the manufacturer's protocol, and 50 µg total lysate was used in SDS-Gel-Electrophoresis (15 % SDS-PAGE, Tris-Glycine). Proteins were transferred onto nitrocellulose membranes by using semi-dry blotting (BioRad TransBlot® Turbo™ Transfer System). The membrane was blocked with 1X TBST (5 % skim Milk powder) and washed 3x with TBST. First and secondary antibodies were used at recommended dilutions in TBST with milk and washed 3x between and after incubations (1 h, 20 °C, rocking shaker). The blots were read out using Pierce™ ECL Western Blotting-Substrate (ThermoFisher).

NanoClick cell permeability assay

The assay was performed with HEK239T cells like reported by Peier et al..[158] HEK293T cells were diluted to reach a concentration of 200000 cells/ml in 10 ml and transfected with 15 µl FuGENE® transfection reagent (Promega) using 0.5 µg BRET plasmid (NanoBRET™ positive control vector; Promega N1581) and 4.5 µg carrier DNA (Promega N1581) mixed with 480 µl OptiMEM (no phenol red, Gibco). From this transfected mix, 100 µl were seeded into 96-well plates (Greiner white transparent bottom, **655094**) and were cultivated for 24 h. Then, the medium was replaced with 90 µl assay buffer (OptiMem without phenol red + 1 % FBS), and DIBAC-CA was added to reach a final concentration of 3 µM. After 1 h of treatment, the cells were washed twice with HBSS (with Mg and Ca), and the medium was replaced with 80 µl assay buffer. Directly, the cells were treated with titrations of azido-peptides (**P-6-Az**, **P-6-Az**, **R8-Az,** and **ONEG-Az**) and incubated for 20 h. The next day, the NB618AZ dye (Promega) was added to a final concentration of 10 µM, and the cells were treated for 1 h. Then, 50 µl of 3X Intracellular TE Nano-Glo® Substrate/Inhibitor mix (Promega N2162) was added, and luminescence was read out using a TECAN Spark plate reader (415 – 430 nm and 595 – 635 nm). BRET ratios were calculated with background correction according to equation (1) as described in the Promega® protocol for the use of the Intracellular TE Nano-Glo® Substrate/Inhibitor.[159]

$$BRET\ ratio = \left(\frac{Acceptor_{Sample}}{Donor_{Sample}} - \frac{Acceptor_{NoTracerControl}}{Donor_{NoTracerControl}} \right) \cdot 1000 \qquad (3)$$

6. SYNTHETIC METHODS

Reagents

Unless otherwise noted, all solvents and reagents were commercially sourced and used without further purification. Solid phase peptide synthesis was performed manually using 20 ml polypropylene syringe reactors from MultiSynTech GmbH. Peptidyl resin suspensions were agitated with a shaker at room temperature, or with argon using polypropylene fritted syringe reactors mounted on a Vac-Man® Laboratory Vacuum Manifold equipped with polypropylene three-way valves.

Intermediate evaluation during solid phase peptide synthesis was performed by cleaving a small amount of resin followed by analysis using either an Agilent 1200 HPLC equipped with an analytical EC HPLC column (NUCLEOSHELL RP 18, 5 µm, 50x4.6 mm) or 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). Purity analysis of final peptides was performed using an Agilent infinity UHPLC equipped with an Agilent InfinityLab Poroshell 120 EC-C18, 2.1 x 150 mm, 2.7 µm, narrow bore LC column. HRMS data was recorded using electrospray ionization with an LTQ Orbitrap in tandem with an HPLC-System fitted with a 50 mm x 1 mm, 1.9 µm Hypersyl GOLD column.

Synthesis of Peptides

The synthesis, purification and analytics of peptides was performed by Jen-Yao Chang ('t Hart Group, CGC Dortmund), Dr. Joseph Openy ('t Hart Group, CGC Dortmund), Dr. Sunit Pal ('t Hart Group, CGC Dortmund). NanoClick reagents and control peptides were synthesized by Dr. Laura Posade ('t hart Group, CGC Dortmund).

Linear Peptide Synthesis

The synthesis of peptides was conducted on Rink-Amide-AM resin (1% DVB; 100-200 mesh, Carbolution Chemicals GmbH) as solid support, typically on a 50-100 µmole scale. The resin was initially swollen using DCM (1 min × 1) and then washed twice with DMF (2 × 30s). Fmoc-deprotection was performed using 1:4 Piperidine in DMF (1 × 5 min; 1 × 10 min) and washed with DMF (4 × 30s). The first Fmoc-Xaa-OH (5.0 eq.) was dissolved in a freshly prepared solution of PyBOP (5.0 eq.) and DIPEA (10.0 eq.). This was then added to the Fmoc-deprotected resin and shaken for 45 min at room temperature. Subsequently, the resin was washed with DMF (3x), DCM (3x), and diethyl ether (3x) and dried under high vacuum. Afterwards, the loading efficiency of the

resin was checked by measuring the UV-absorption at 300 nm of a small quantity of resin deprotected using the piperidine solution. Then, an appropriate amount of resin was swollen again as before and extended using Fmoc-protected amino acid (4 eq), PyBOP (4 eq), and DIPEA (8 eq) for 45 min to 1 h followed by washing with DMF (3x), DCM (3x), and DMF (3x). For the coupling of unnatural amino acid building blocks (R_5, S_5 and B_5), Fmoc-Xaa-OH (2.0 eq.) was dissolved in DMF in the presence of PyBOP (2.0 eq.), Oxyma (1.0 eq.) and DIPEA (4.0 eq.), added to the resin and shaken for overnight at room temperature. If necessary, amino acids following the non-natural amino acids used for stapling were coupled twice.

Acetylation

The Fmoc-deprotected peptidyl resin was suspended in DMF followed by the addition of Ac_2O (10 eq) and DIPEA (10 eq). The suspension was shaken for 30 min and washed with DMF (2 × 30s).

FITC Labelling

The syringe reactor containing a linear peptide on resin with Fmoc protection in DMF suspension was loaded with Fmoc-O2Oc-OH (2 eq), PyBOP (2 eq), and DIPEA (4 eq) and agitated overnight. The resin was washed with DMF (2 × 30s), deprotected with 1:4 Piperidine in DMF (1 × 5 min; 1 × 10 min), washed with DMF (4 × 30s) and resuspended in DMF. Subsequently, the syringe reactor was charged with FITC (1.5 eq) and DIPEA (4 eq) in DMF and shaken overnight. The peptidyl resin was then washed with DMF (4 × 30s), DCM (2 × 30s), and Et_2O (1 × 1 min) and dried under vacuum.

Ruthenium Olefin Metathesis

Either one of the following protocols was used for the ring-closing metathesis reaction:

(i) A solution of Grubbs Catalyst® 1st Generation (0.1 equivalent) was made in degassed DCM (3 ml). This solution was then drawn into a 10 ml syringe reactor containing the Fmoc-protected substrate on resin. The sealed reactor was shaken at rt for 2 hours, followed by discharging the supernatant. The substrate was treated by the same procedure with freshly prepared Grubbs Catalyst® 1st Generation solution for 3 more times. Afterwards, the resin was washed with DCM (2 x 30 s), dried under vacuum.

(ii) A suspension of Fmoc-protected peptide on resin was placed in a syringe reactor. A solution of Hoveyda-Grubbs Catalyst® 2nd Generation (20 mol%) in DCE (2 ml) was then drawn into the reactor. The reaction vessel was equipped with an open two-way valve and shaken for 2 h. Afterward, the supernatant was removed, and the reactor was recharged with a fresh catalyst

solution and agitated for an additional 2 hours. The peptidyl resin was washed with DMF (4 × 30s), DCM (2 × 30s), and Et$_2$O (1 × 1 min) and dried under vacuum.

Peptide cleavage

The peptides were separated from the resin by treating it with a solution of TFA/TIPS/H$_2$O (95:2.5:2.5) for 1 hour. The resin was then filtered, and the filtrate was added to ice-cold Et$_2$O to obtain the crude peptide by centrifugation. The pellet was resuspended in fresh ice cold Et$_2$O followed by centrifugation. This procedure was repeated once more.

Stapled Peptide Reduction

If the double bond of the hydrocarbon staple needed to be reduced, the substrate was first cleaved from the resin using the cleavage protocol previously described. The cleaved substrate was then precipitated using cold diethyl ether followed by dissolution in MeOH/AcOH (3 mL, 1:1) before being transferred to a round-bottom flask containing a stir bar and 10% Pd/C (30 mg). The sealed flask was filled with argon. Next, the argon was exchanged with hydrogen and the mixture was stirred at rt for 2 h. After completion, the reaction was filtered through a pad of celite. The residue was washed with MeOH, the filtrates combined, concentrated, and purified by preparative HPLC.

RESULTS AND DISCUSSION

PART A: USING AN *E. COLI* BASED TRANSLATION REPRESSION ASSAY FOR THE IDENTIFICATION OF MACROCYCLIC INHIBITORS OF PTBP1

Compared to protein-protein interactions, RNA–protein interactions are more challenging to study *in cellulo* because of the increasing complexity of assay systems. Many RNA-protein interactions were studied *in vitro* by EMSA, pull-downs, or CLIP. These methods have been demonstrated to be powerful tools for investigating biophysical and cellular interactions. Still, they all lack the possibility of easy upscaling and miniaturization to screen for inhibitors. Optimal assays should have a low number of components to not rely on numerous adapters in hybrid screening methods, and be performed *in cellulo* to avoid the need for purification of individual components.

The two-hybrid assays are some of the most advanced and most commonly used techniques for screening for PPIs and identifying inhibitors.[160,161] Two-hybrid assays rely on a protein-protein interaction between a bait and a prey protein for the cellular readout. The RNA counterpart, a three-hybrid assay, utilizes three interactions between a DNA-RNA adapter protein, a bait RNA, and a prey protein. These hybrid assays benefit from the accessibility of molecular genetics nowadays; thus, they are a powerful tool for dissecting molecular interactions. Both assays can be used to identify inhibitors of interactions between biomacromolecules, but three hybrid assays require high-affinity interactions because of low signal-to-noise ratios.[162]

This chapter's objective is to develop a bacterial assay for quantifying RNA-protein interactions in cellular environments based on the translational repression of a reporter protein by the protein-RNA interaction of PTBP1 and target RNAs. For this, the **T**ranslational **R**epression **A**ssay **P**rocedure (TRAP) was chosen.[150,163,164] The TRAP assay is established for several high-affinity RNA-protein interactions, in which an RNA target sequence is cloned in the spatial environment of a reporter gene's ribosome binding site (RBS). Upon expression of the RNA binding protein of interest, the binding of the ribosome to the RBS is sterically hindered, and therefore the amount of reporter is reduced (**Figure 9A**). The assay design consists of two plasmids: First, a reporter plasmid, where an RNA of interest is cloned either before the RBS or in frame of the protein after the start codon of the reporter gene. Second, protein expression is encoded on the other plasmid. In this case, the RBP of interest is also fluorescently labeled with an orthogonal fluorescent protein to measure the expression level of the RBP. (**Figure 9B**) The reporter and expression plasmids contain an origin of replication (ORI) from different families and antibiotic resistances to avoid

incompatibilities (Reporter pBbA5k: p15a and KanR, and Expression pBbE8c: colE1 and CamR). Furthermore, the RBP expression is arabinose-inducible using an araBAD promoter, while the reporter expression is IPTG inducible through the lac-operon. By this, the expression levels can be titrated against each other to optimize conditions.
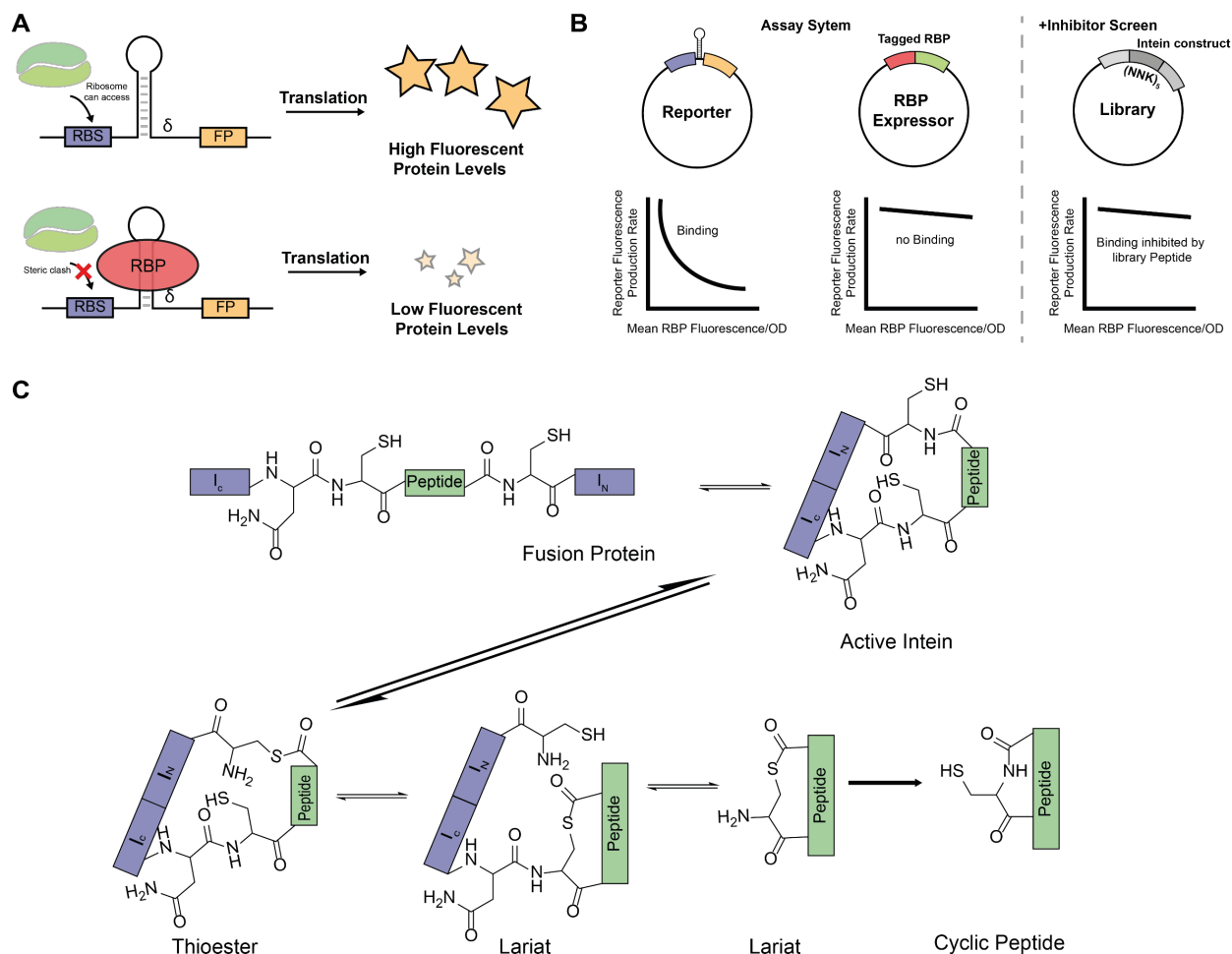
Figure 9: (**A**) Schematic explanation of Translational Repression Assay Procedure (TRAP). (**B**) *E. coli* based screening system with the addition of an inhibitor library encoded on a third plasmid. (**C**) Split intein mediated cellular generation of cyclic peptides by SICLOPPS.

After establishing the TRAP assay, it was planned to screen for macrocyclic peptide inhibitors of the RBP using genetically encoded libraries. This can be done using split-intein circular ligation of peptides and proteins (SICLOPPS).[165,166] In this method, cloning introduces a randomized peptide sequence into the Synechocystis sp PCC6803 DnaE split intein gene. Expression of the split intein yields an active *cis*-intein which releases a cyclized peptide or protein through an intramolecular splicing reaction through the formation of an intermediate thioester (**Figure 9C**). The peptide sequence has few limitations but needs an *N*-terminal cysteine or serine for nucleophilic attack during the splicing reaction.[149] In most reported cases, six-membered macrocycles are used with

five randomized positions encoded by NNS (N: A, C, G or T; S: C or G) to avoid the ochre and opal stop codons from the library and to ensure the coverage of the library which is limited by the transformation efficiency of *E. coli*.[149,167,168]

A.1: GENERATION OF THE SICLOPPS LIBRARY

The *Synechocystis* sp PCC6803 DnaE split intein gene was subcloned into the pBbB2a vector and used for the SICLOPPS library's following generation. The library encodes for a cyclic peptide containing a cysteine and five randomized codons (NNS) to generate a six-membered macrocycle. The plasmid backbone was prepared with a triple digest with EcoRV/XhoI/MfeI. The additional cut with EcoRV was needed because the two products of a double digest with XhoI/MfeI had similar sizes. The insert was generated by PCR. When the primary PCR product was used in restriction/ligation reactions, many frameshifts were observed, presumably caused by the high diversity of the randomized NNS codons that yielded impure PCR products after several cycles. To avoid this, a "zipper PCR" with primers binding at the insert's 5' and 3' ends was performed, producing clean dsDNA products.[149] The reported protocol from Tavassoli et al. described dialysis of ligation products before the final electroporation step. This yielded too low transformation efficiencies of < 1000 colonies in our hands. Because of this, ten ligation reactions containing 100 ng vector each were pooled, purified, and concentrated by ethanol precipitation. The pellet was solubilized in 10 µl nuclease-free water and transformed into 10 x 50 µl electrocompetent DH5alpha cells, which were again pooled and plated onto selective agar plate in serial dilutions. The plates with dilutions of 10^5 and 10^6 contained approximately 400 and 40 colonies, respectively, corresponding to $40*10^6$ clones which covers the $32^5 = 33.55*10^6$ possible products roughly 1.2 times. This number is on the lower side, but previous reports suggest similar coverages should be sufficient for inhibitor screenings.[149,167,169]

A.2: ESTABLISHING OF THE TRAP ASSAY

We performed the TRAP assay in a 96-well format in total volumes of 200 µl with *E. coli* Top10F', which were transformed with an expressing and a reporting plasmid. After reaching an OD_{600} of ~ 0.2, the cells were induced with IPTG and L-ara to induce reporter and RBP expression. In a further overnight expression, the fluorescence intensities of the reporter and tagged RBP and the optical density were measured in a plate reader. The fluorescence data was normalized to the optical density first. Then, the reporter production rate was calculated and plotted against the integrated intensity of the fluorescently tagged RBP. More details are to be found in the materials and methods.

First, we used the red fluorescent protein mRFP as the reporter and the blue fluorescent protein TagBFP for labeling the RBP of interest. Therefore, we cloned the sequence of the miRNA mir29b2, a known PTBP1 binder behind the start codon of mRFP on the reporter plasmid.[170] Further, we subcloned several different constructs of PTBP1 (RRM1, RRM2, RRM12, and RRM34) to be *N*-terminally TagBFP tagged. We varied the distances of mir29b2 to the start codon to evaluate the optimal space to the RBS (**Figure 10**).
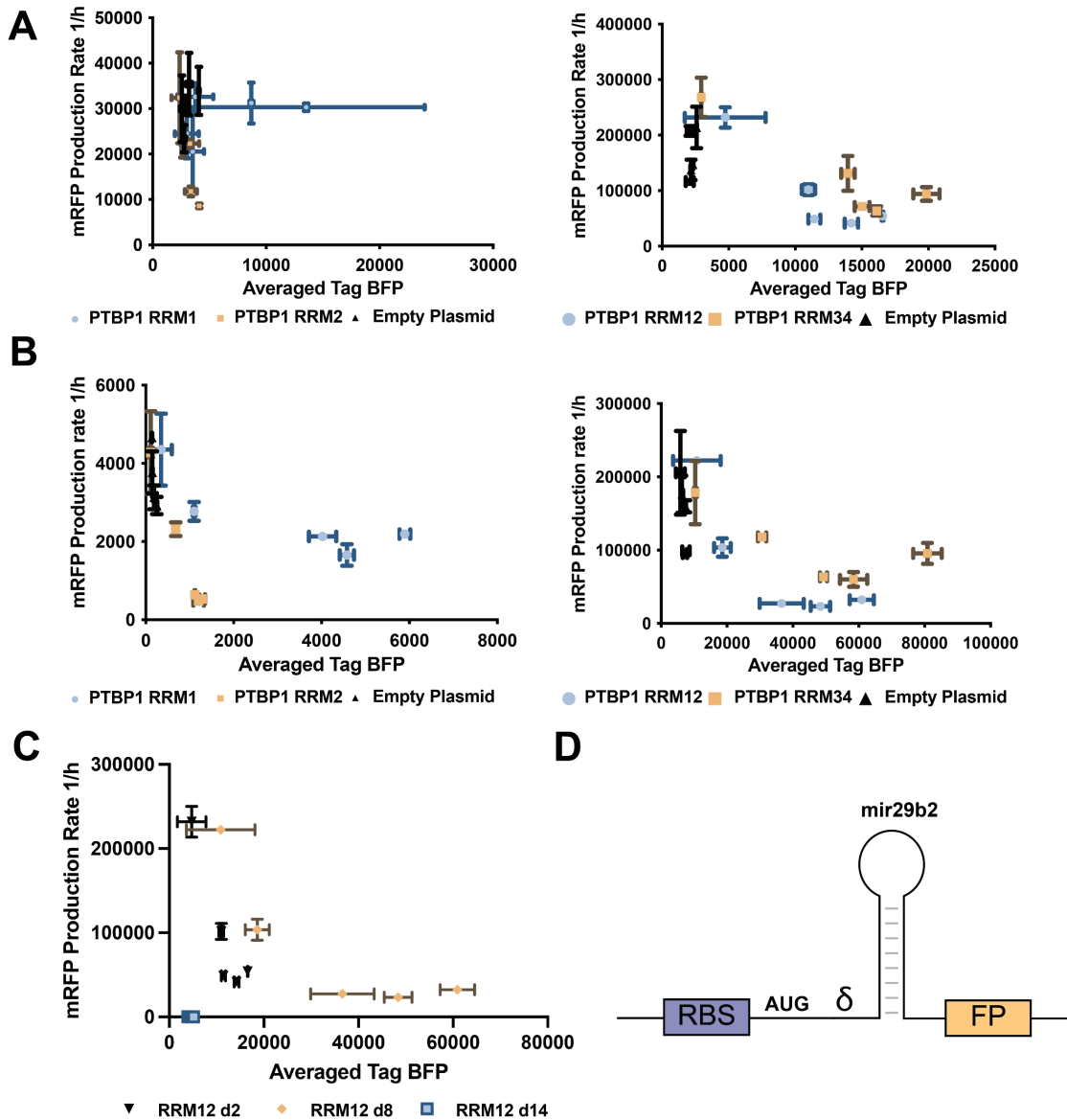
Figure 10: TRAP assay performed with mir29b2 RNA placed with a distance of 2 nt from the start codon (A) for RRM1 and RRM2 (left) and RRM12 and RRM34 tandems (right). (B) Same as A with a distance of δ = 8 nt. (C) Representative representation of the dependency of the distance δ from the RBS for the mir29b2 constructs after the start codon with RRM12. (D) Depiction of the placement of the RNA of interest in the reporter construct.

49

For the three mir29b2 constructs placed behind the start codon of mRFP, we could identify several combinations with repressed translation of the repressor at increasing levels of RBP expression. With the lowest distance (δ=2 nt; **Figure 10A**), we could only recognize modest repression for the tandem constructs RRM12 and RRM34, which was mildly over the vector control. With increased distance (δ=8 nt; **Figure 10B**), RRM1, RRM2, RRM12, and RRM34 showed a dose-dependent translation repression. When we increased the distance to δ=14 nt, we could no longer observe the reporter's translation. As the assay's dynamic range compared to the vector control was not big, we tried to optimize the assay using a different PTBP1 consensus sequence. Therefore, we subcloned the reported GABA$_A$ γ2 RNA sequence behind the start codon of the mRFP and performed the TRAP assay with RRM12 and RRM34.[56]

A

B
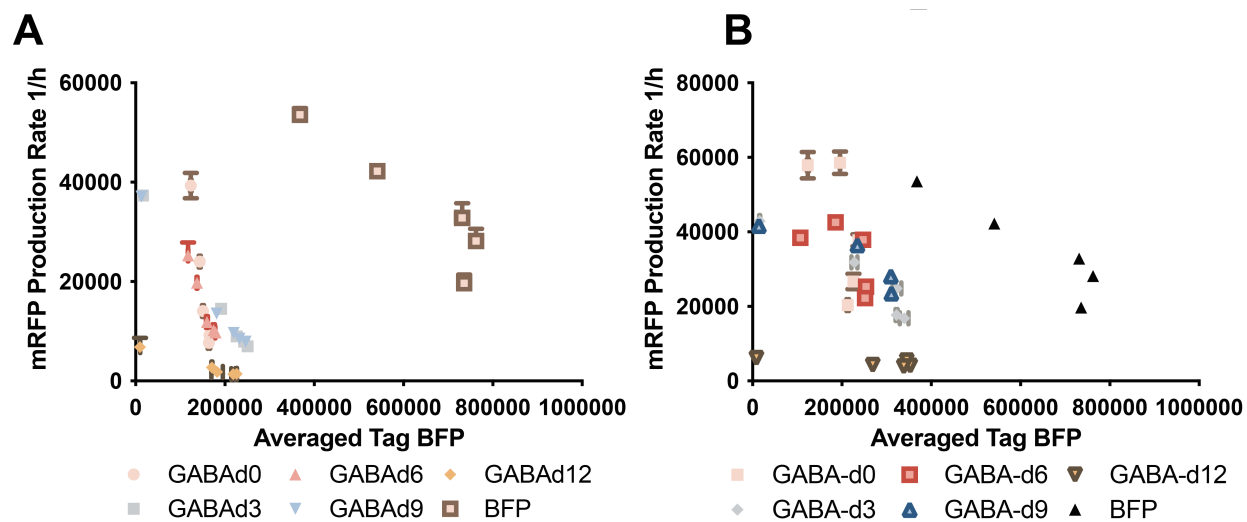
Figure 11: (**A**) TRAP assay performed with GABA RNA cloned after the start codon of mRFP with RRM12 (**A**) and RRM34 (**B**).

We could identify translation repression for the RRM12 construct for the GABA$_A$ γ2 RNA with δ=0, 3, 6, and 9 nt, while again, the longest distance decreased the initial fluorescence significantly (**Figure 11A**). For RRM12, the curve shape was comparable to the expected and reported one. When we performed the assay with the RRM34 construct, again, we could not observe any fluorescence for the longest distance (δ=12 nt), while the sorter ones (δ=0, 3, 6, 9 nt) did show repression. Still, the curve shape and the maximum repression were similar to the vector control (**Figure 11B**). To our surprise, our results indicated binding of RRM12, but not RRM34, to the GABA$_A$ γ2 RNA, while Clerte and Hall reported it contrariwise.[56]

To this end, we decided to use flow cytometry to detect the fluorescence of individual cells and sort them based on the mRFP and TagBFP levels with the mir29b2d8 reporter plasmid. Through this, we could differentiate between individual cells in a future screening using a SICLOPPS

library. We, therefore, performed the assay similarly to the plate reader-based assay and sorted the cells using a Sony SH800 cell sorter (**Figure 12**).
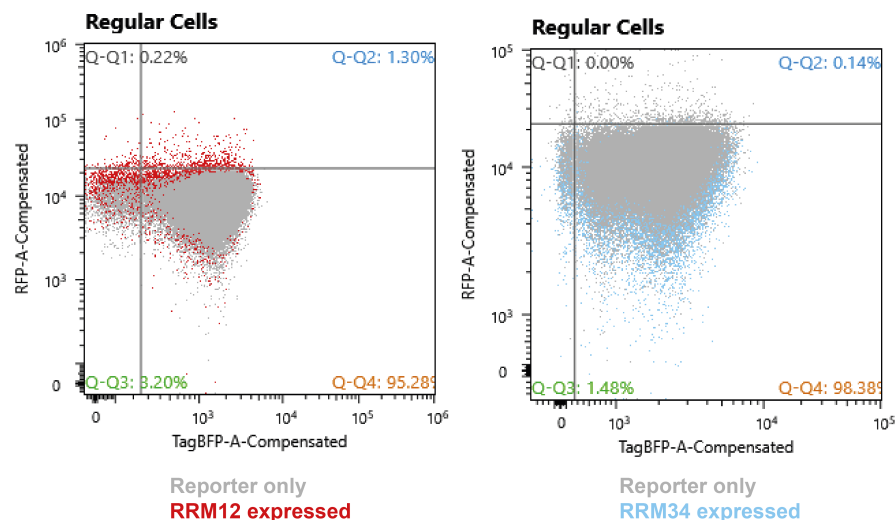
Figure 12: FACS analysis of *E. coli* Top10 F' co-transformed with reporter and expressing plasmids of the TRAP assay with no protein expression control (grey) and RRM12 (red) and RRM34 co-expression (blue).

In the FACS experiment, we could not recognize reasonable differences between RBP-expressing cells and the negative control without any protein co-expression. This could be caused by the low maturation time of the mRFP, which was just in the timeframe of the experiment. Therefore, we changed the reporter fluorescent protein to a faster maturating, compatible sfGFP.[171] We replaced the mRFP in the reporter plasmid with sfGFP and verified the assay again (**Figure 13**).

When we performed the TRAP assay with the sfGFP reporter, we recognized a ~2-fold decrease of reporter expression under co-expression of RRM34. To our surprise, we detected a significant population of bacteria simultaneously downregulated TagBFP and sfGFP expression in cell sorting, which is inconclusive with the assay hypothesis. When we sequenced the individual colonies in Q3 (**Figure 13, right**), we identified a loss of RRM34 expression, probably caused by recombination. This loss of expression could be caused by too much stress caused by the over-expression of the protein of interest or the toxicity of inducing agents.

Outlook and Perspectives

At this point, we decided to discontinue the development of an assay to screen a genetically encoded library of potential inhibitors. In our hands, the TRAP assay worked for other protein-RNA combinations, yielding much higher ratios of translational repression than reported in this project. Although the mRFP-based plate reader-based assay looked promising with PTBP1,

mRFP was incompatible with FACS in several other projects in the group.[172] We decided to change the strategy to a rational, structure-based design based on unique, non-canonical structural features of PTBP1.
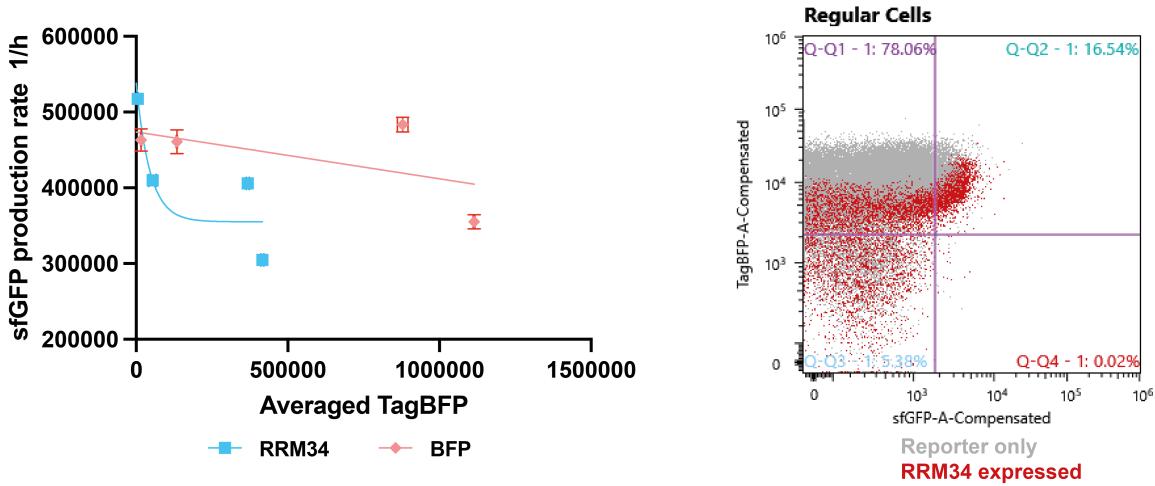
Figure 13: TRAP assay performed with the mir29b2d8 reporter gene with sfGFP and PTBP1 RRM34 (left) and analysis of a population from a similar experiment in FACS (right).

PART B: STAPLED PEPTIDES AS INHIBITORS FOR OF PTBP1

Parts of this chapter have been published in <u>S. Schmeing</u>, et al., "Rationally designed stapled peptides allosterically inhibit PTBP1-RNA-binding", *Chemical Science*, **2023**.

B.1: PROTEIN PURIFICATION

To perform biophysical evaluation of the role of the individual domains of PTBP1 during RNA binding and to validate synthesized inhibitors, different domains and mutants of PTBP1 were heterologously expressed in *E. coli* BL21 DE3 RIL cells using a pMAL expression vector. The constructs were *N*-terminally His-MBP-tagged, and the affinity tag can be cleaved off using TEV protease. The general purification strategy was to first perform an affinity enrichment step with either the *N*-terminal His$_6$ or MBP-tag using immobilized metal affinity chromatography (IMAC) or an amylose resin for the tags, respectively. The tag was then cleaved during dialysis using TEV protease, and the non-converted protein and tags were removed using a subsequent affinity enrichment of the side products and collection of the unbound fraction. The resulting samples were finally purified, and aggregates were removed using a Superdex 75 size exclusion column. The purity of the samples was validated to be >90% by SDS-PAGE and evaluation of the chromatograms if not otherwise stated.

RRM12

The RRM12 domain tandem was expressed using *E. coli* Bl21 DE3 RIL cells overnight at 20 °C. The protein could be purified using amylose resin or Ni-NTA affinity chromatography through the MBP tag. Shown in **Figure 14** is the purification using the MBP tag, which consisted of the following steps: i. Gravity amylose column, ii. Dialysis with tag cleavage using TEV-protease, iii. reverse Ni-NTA, and iv. a final gel filtration. The overexpression and first affinity column were successful, as the soluble fraction of the lysate contained a very prominent band at 70 kDa corresponding to the molecular weight of MBP-tagged RRM12. This protein was, after lysis, successfully enriched and eluted using amylose beads and 10 mM glucose as competitor (**Figure 14A**). The His-MBP-tag in the protein-containing fractions was afterward removed using TEV-protease, which resulted in a mixture of MBP-RRM12 (~ 70 kDa), MBP (~ 55 kDa), and RRM12 (~ 22 kDa) (**Figure 14B,** TEV o.n.). The side products were removed by IMAC collecting the unbound fraction, and pure RRM12 was enriched by this (**Figure 14B**, Unb. Frac.), while the side products were eluting with ≥ 25 mM imidazole. The already pure RRM12-containing fractions were further purified by size exclusion chromatography to remove protein aggregates, and >90% pure RRM12 was obtained in reasonable yields (**Figure 14C**).
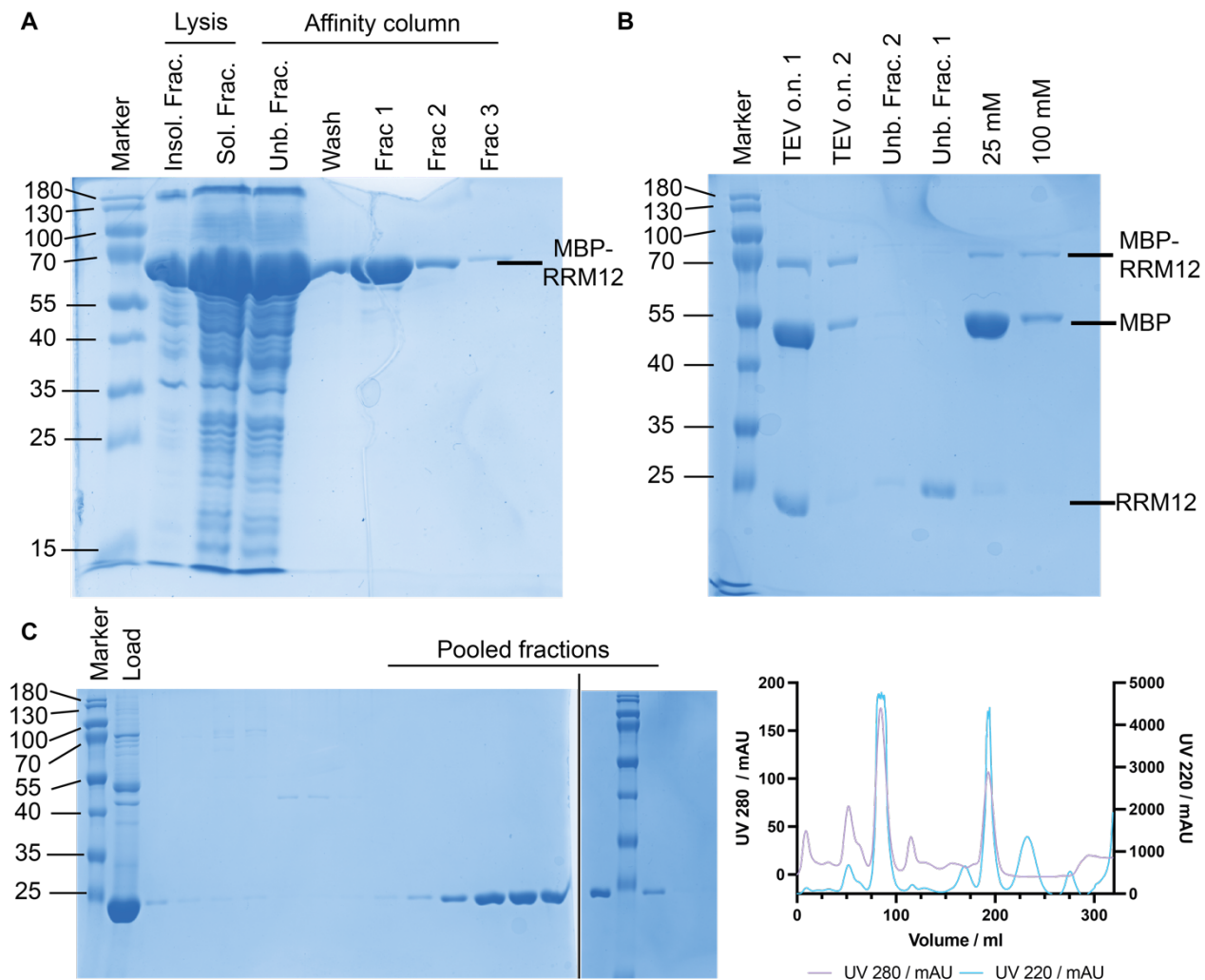
Figure 14: Purification of PTBP1-RRM12 tracked with 15 % SDS-PAGE and chromatograms (280 and 220 nm channel). (**A**) Amylose affinity column. (**B**) Reverse IMAC column after tag cleavage. (**C**) Size exclusion chromatography with a Superdex 75 26/60 column; fractions were loaded on two gels, and the cropping was indicated by a black bar.

RRM12-L151G

To verify the role of the α3-helix in the inter-domain linker between RRM1 and RRM2 of PTBP1, an L151G mutation in this helix was introduced by site-directed mutagenesis and the mutant was purified similarly to the RRM12 construct.[44] The affinity column was chosen as an IMAC for more straightforward handling purposes. Again, the L151G mutant eluted very pure from the first affinity column (**Figure 15A**). After tag cleavage in dialysis overnight, the full-length construct was fully converted into MBP, and the RRM12-L151G construct (**Figure 15B**, -TEV/+TEV lanes) with RRM12-L151G running at ~ 22 kDa and MBP at ~ 55 kDa corresponding to their molecular weight. The tag-cleaved construct was the only visible band in the unbound fraction of the reverse IMAC,

and significant amounts of this domain bound to the resin unspecifically but could be eluted with 10 % Buffer B with low quantities of MBP tag co-eluting (**Figure 15B**; Unb. Fraction/Wash).
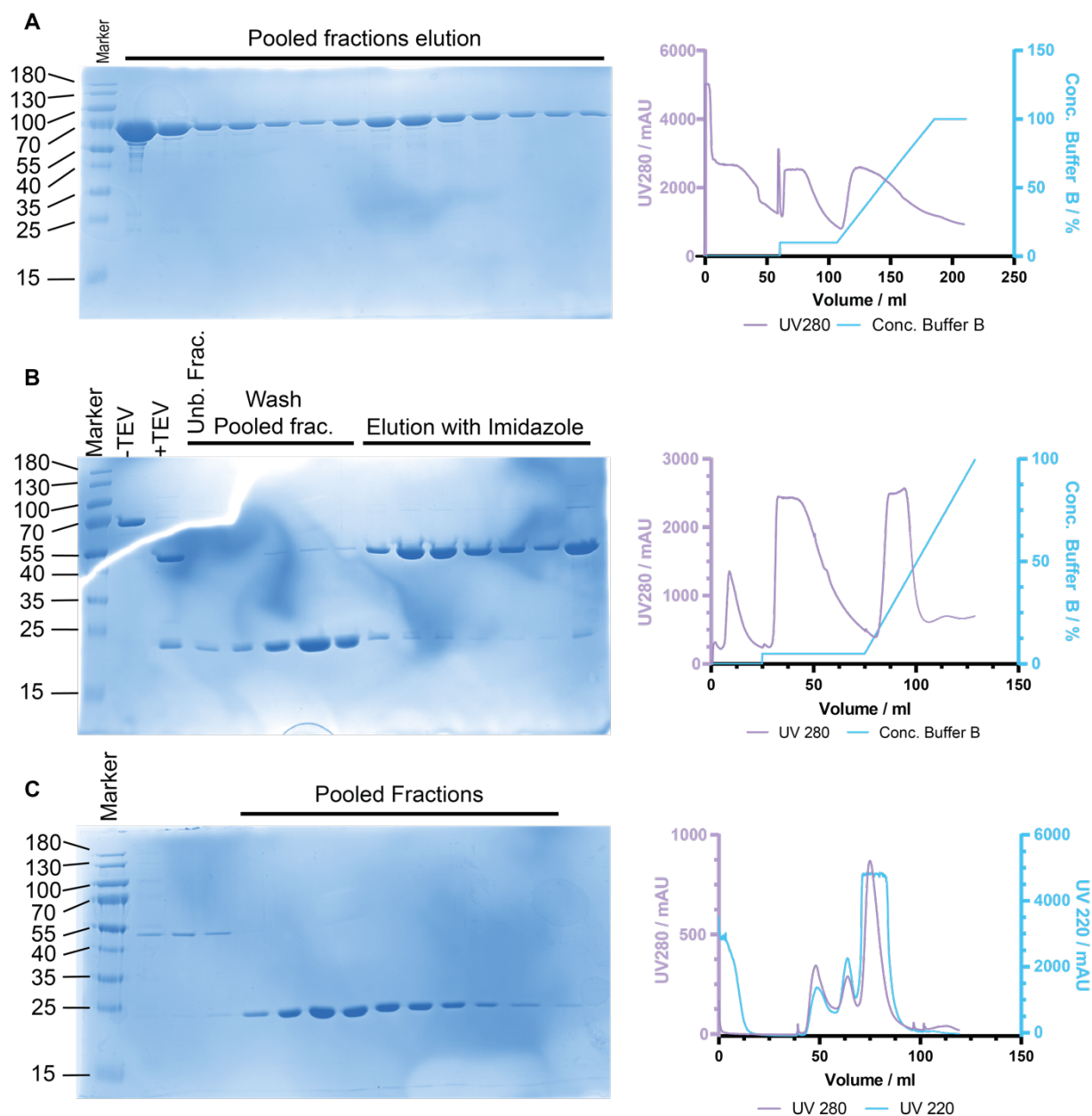
Figure 15: Purification of RRM12-L151G mutant observed by 15 % SDS-PAGE and chromatograms of the respective FPLC steps. **(A)** IMAC using HisTrap HP column after lysis. **(B)** Tag-cleavage overnight using TEV-protease and enrichment with reverse IMAC using a HisTrap HP column. **(C)** Size exclusion chromatography with a Superdex 75 16/60 column.

The pooled unbound and wash fractions were concentrated and injected onto a Superdex 75 16/60 column, from which RRM12-L151G eluted in a homogeneous peak with >90% purity in modest amounts (**Figure 15C**).

RRM1

The protein was expressed similarly to the previous RRM12 constructs and could be detected in significant amounts in the soluble fraction after lysis (**Figure 16A**; Sol. Frac). The binding of MBP-RRM1 to the IMAC column was not optimal, as very intense bands corresponding to MBP-RRM1 were detected in the unbound and wash fractions (**Figure 16A**, Unb. Frac. / Wash). The construct could be eluted in sufficient amounts by applying an imidazole gradient, and elution was observed starting from 20 mM. The protein-containing fractions were concentrated and applied onto a Superdex 75 16/60, where a homogeneous pure peak was observed, which contained pure MBP-RRM1 (**Figure 16B**).
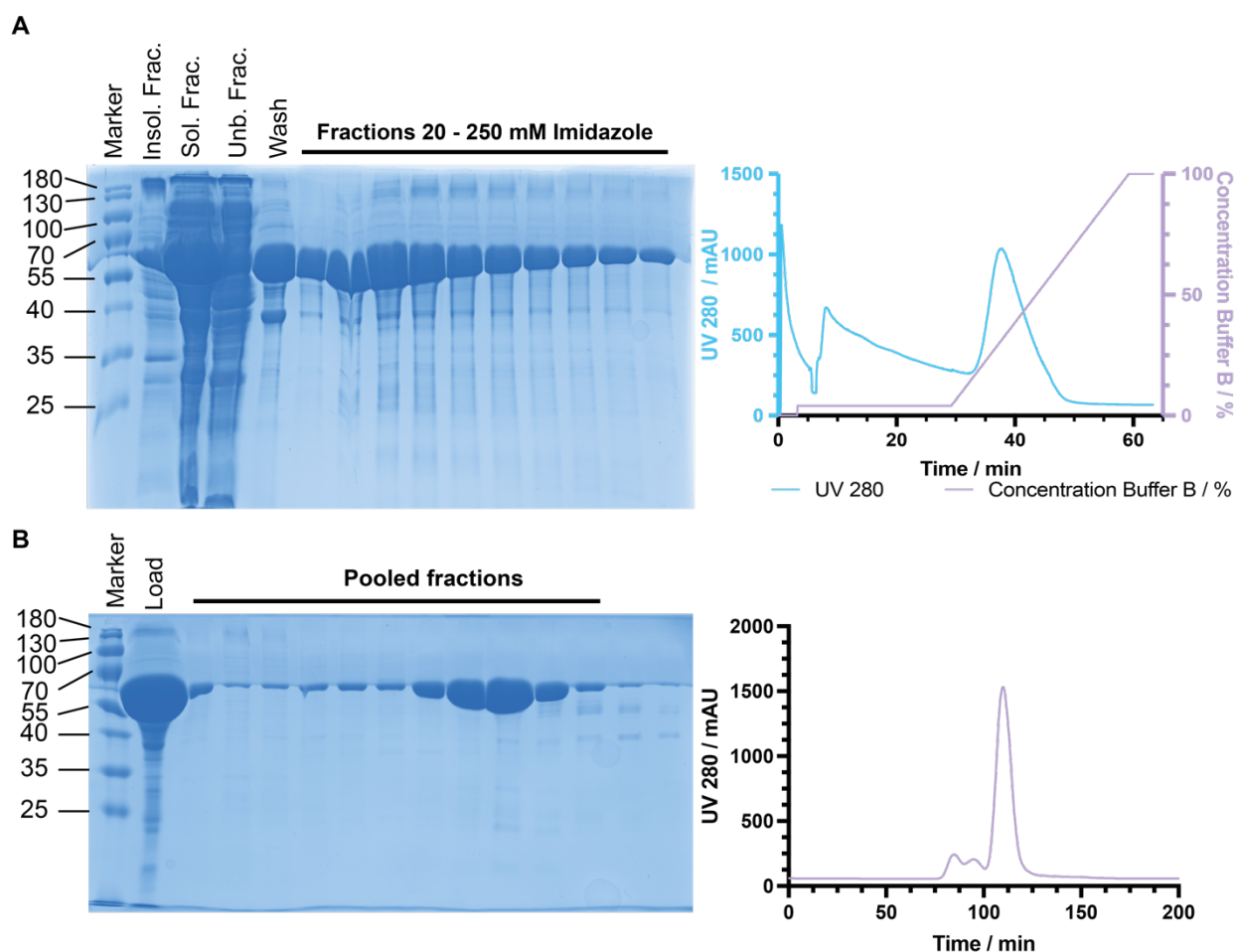
Figure 16: Purification of MBP-RRM1 tracked by 12 % SDS-PAGE and chromatograms of the corresponding FPLC runs. (**A**) IMAC using a HisTrap HP 5 ml column. (**B**) Size exclusion chromatography using a Superdex 75 16/60.

RRM1-L151G

To verify the roles of the α3-helix in RRM1 orchestrating the RNA binding of PTBP1, an RRM1 construct with L151G mutation was cloned by site-directed mutagenesis and expressed and purified like the RRM1 construct with an IMAC followed by size exclusion chromatography. Again, the protein of interest eluted from the first affinity column in decent amounts and high purity (**Figure 17A**). The protein-containing fractions were then concentrated and injected onto a Superdex 75 16/60 gel filtration column for further purification and removal of aggregates. The protein co-eluted with a second peak containing a degradation product, leading to a significant yield decrease (**Figure 17 B**). The purest fractions were concentrated and stored at -80 °C until use.
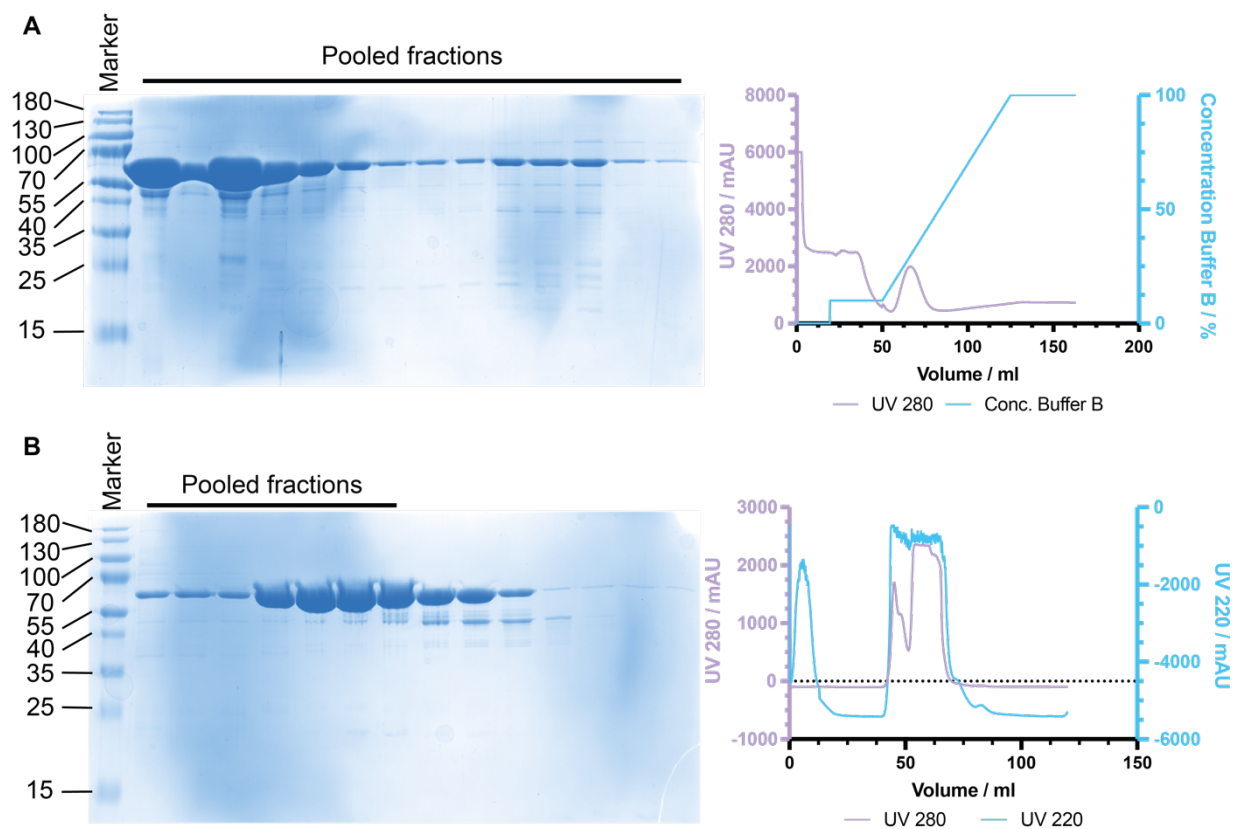
Figure 17: Purification of MBP-RRM1-L151G mutant. (**A**) 12 % SDS-PAGE of an IMAC after cell lysis (left) and chromatogram of the FPLC run (right). (**B**) 12 % SDS-PAGE of the size exclusion chromatography with a Superdex 75 16/60 column (left) and the corresponding chromatogram (right).

Further on, half of the batch was further processed to be tag-cleaved in overnight dialysis with TEV-protease (**Figure 18A**) and could be isolated in good yields after gelfiltration (**Figure 18B**).
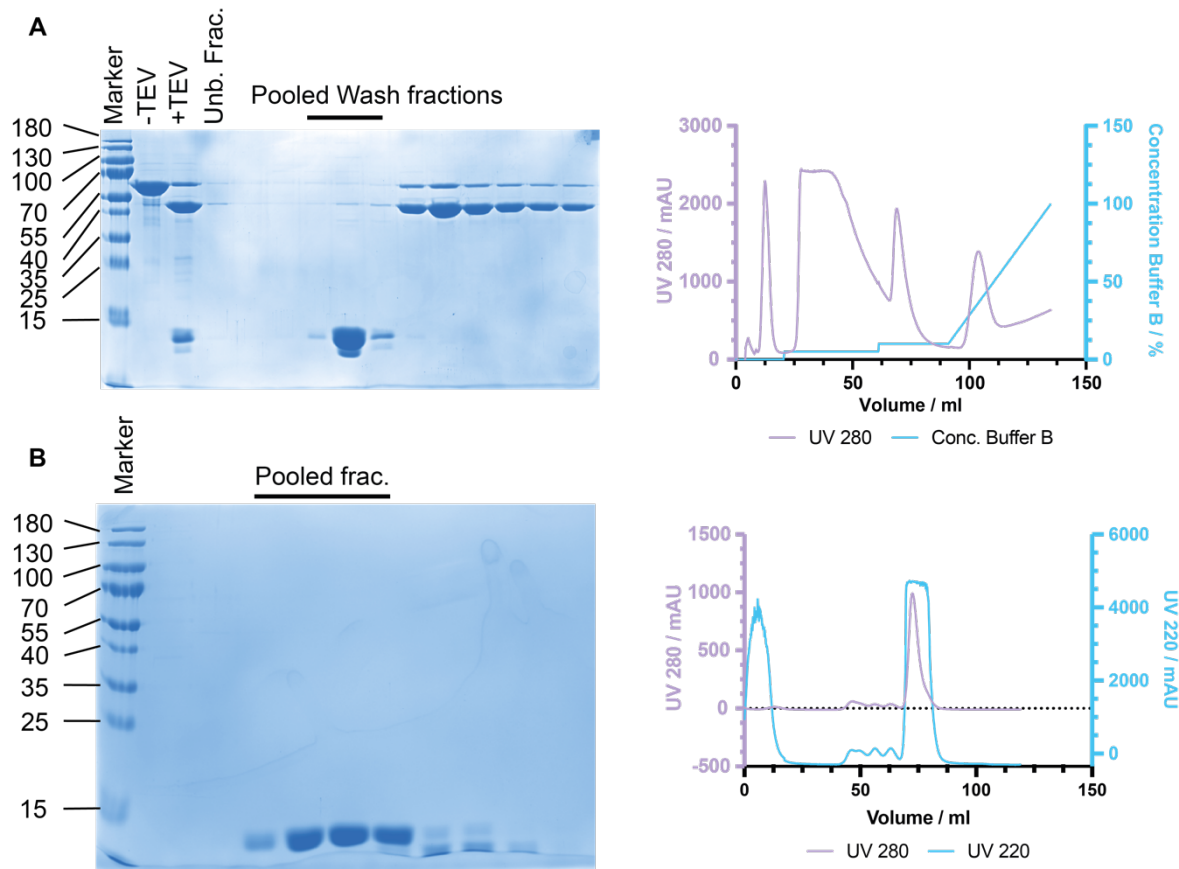
Figure 18: Purification of tag-cleaved RRM1-L151G. (**A**) Reverse IMAC after TEV treatment in dialysis overnight. (**B**) Size exclusion chromatography using a Superdex 75 16/60 column.

RRM2

The PTBP1-RRM2 construct was purified in the same way as the RRM1 construct. First, the overexpression and enrichment through IMAC and elution with an imidazole gradient were successful (**Figure 19A**). The fractions containing MBP-RRM2 were pooled, concentrated, and further purified using a Superdex 75 16/60. There, the protein of interest eluted in a homogeneous peak that contained >90 % pure MBP-RRM2. The protein was concentrated and stored at -80 °C after flash freezing until use.
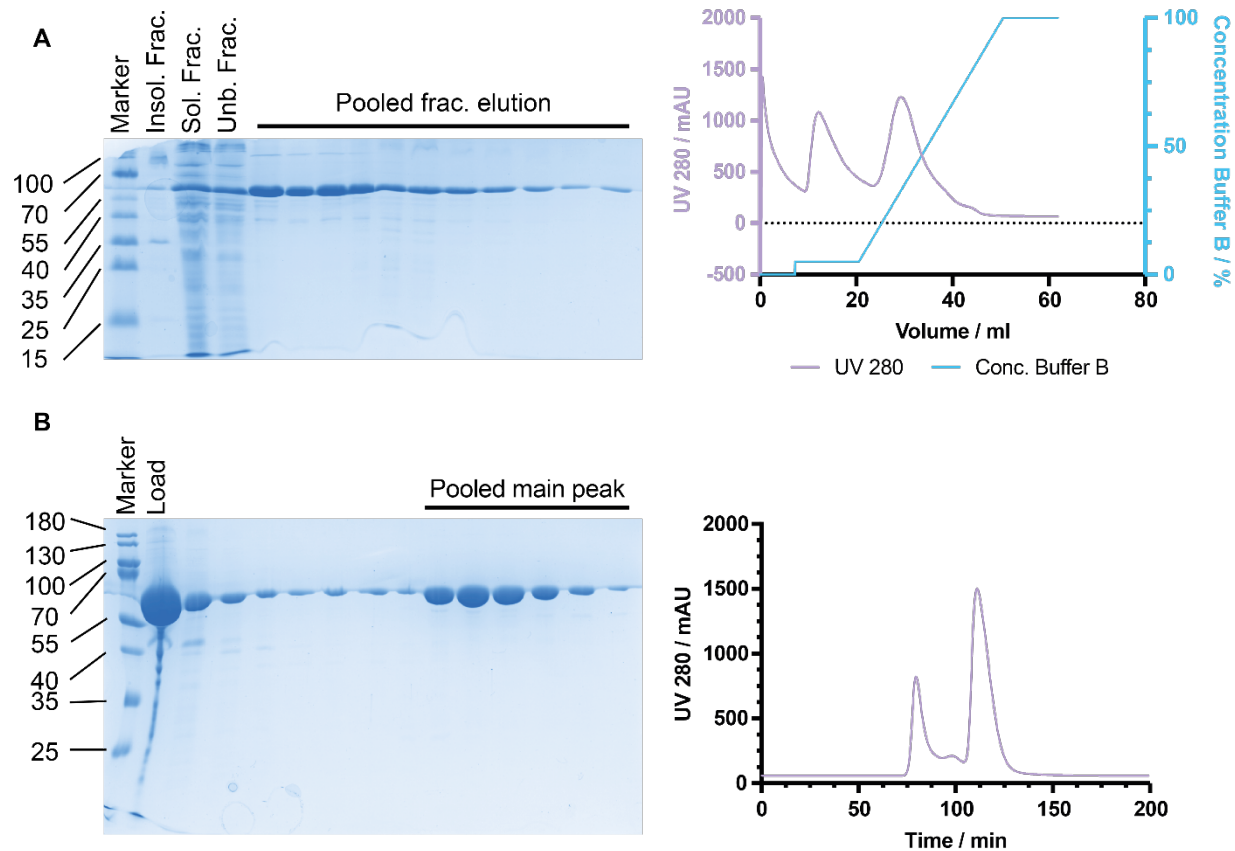
Figure 19: Purification of MBP-RRM2 using (**A**) a HisTrap HP 5 ml IMAC and elution with an imidazole gradient and (**B**) a Superdex 75 16/60 size exclusion chromatography column.

RRM1234

A protein construct containing all four RRMs of PTBP1 was subcloned, expressed, and purified like the previous constructs to be used in biophysical experiments to validate inhibitors of PTBP1. Therefore, the His-MBP-RRM1234 construct was enriched using gravity flow amylose beads which yielded a prominent band on the SDS-PAGE at ca. 100 kDa corresponding to the fusion protein (**Figure 20A**). Tag cleavage in dialysis with TEV-protease was quantitative, and the two product bands at ~ 51 kDa (PTBP1-RRM1234), and 45 kDa (His-MBP) were detected. During the reverse IMAC, added TEV-protease (ca. 70 kDa) was removed. Still, the His-MBP tag could not be successfully removed (**Figure 20B**), while the RRM1234 construct co-eluted with TEV-protease in a 20 mM imidazole wash. The unbound fractions were concentrated and further purified using a Superdex 75 26/60 gel filtration, and the POI RRM1234 co-eluted with significant contaminations of His-MBP. Therefore, the purification was just partially successful. Nevertheless, the RRM1234 containing fractions were concentrated, frozen in liquid nitrogen, and stored at -80 °C until use.
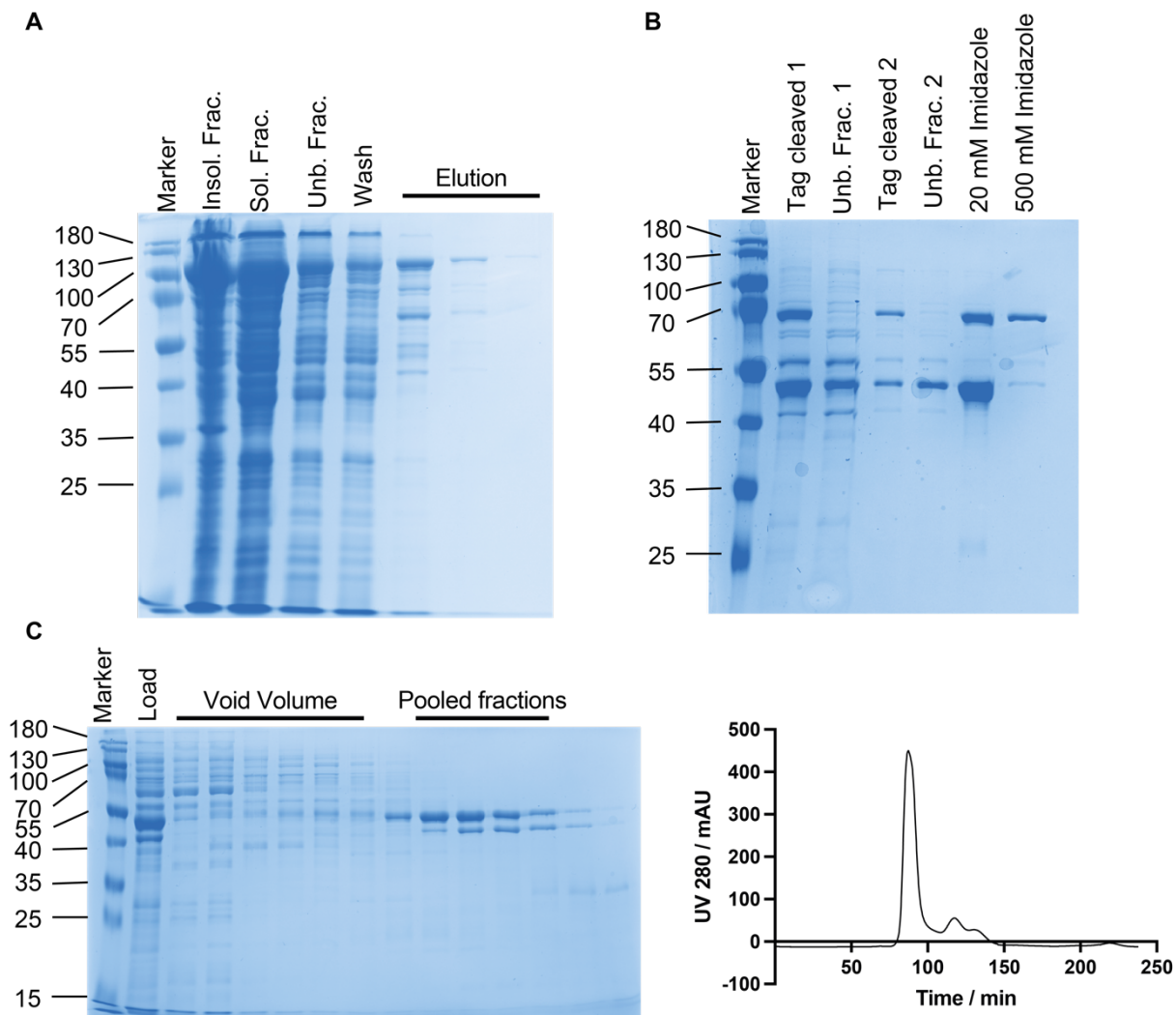
Figure 20: Purification of RRM1234 tracked by 12% SDS-PAGE and the corresponding chromatograms of the FPLC runs. (**A**) Lysis and affinity chromatography using gravity flow amylose beads. (**B**) Reverse IMAC using a HisTrap HP 5ml column and (**C**) size exclusion chromatography with a Superdex 75 26/60 column.

RRM1*2

To investigate the role of the individual RRMs in the PTBP1-RNA binding event, an RRM12 tandem mutant was subcloned by site-directed mutagenesis, which contains an inactive RRM1 domain (**RRM1*2;** S131I, N132L, L136A).[174] This protein construct was overexpressed and purified like the RRM12 tandem. During the first affinity column (**Figure 21A**), the POI bound and eluted to the column in an imidazole gradient. However, the POI co-eluted with two more significant proteins, degradation products containing the His-MBP tag. The tag-cleavage during dialysis was nearly quantitative because the His-MBP-RRM1*2 band at ~70 kDa almost disappeared. At the same time, the side product His-MBP and the POI RRM1*2 at ~22 kDa

significantly increased their prominence in the reaction check via SDS-PAGE (**Figure 21B**). The protein-containing but side-product contaminated unbound fraction of the reverse-IMAC was concentrated and further purified with a Superdex 75 16/60 size exclusion column, where the RRM1*2 protein eluted as a homogeneous main peak with minor impurities. The protein-containing fractions were concentrated, flash-frozen in liquid nitrogen, and stored at -80 °C until use.
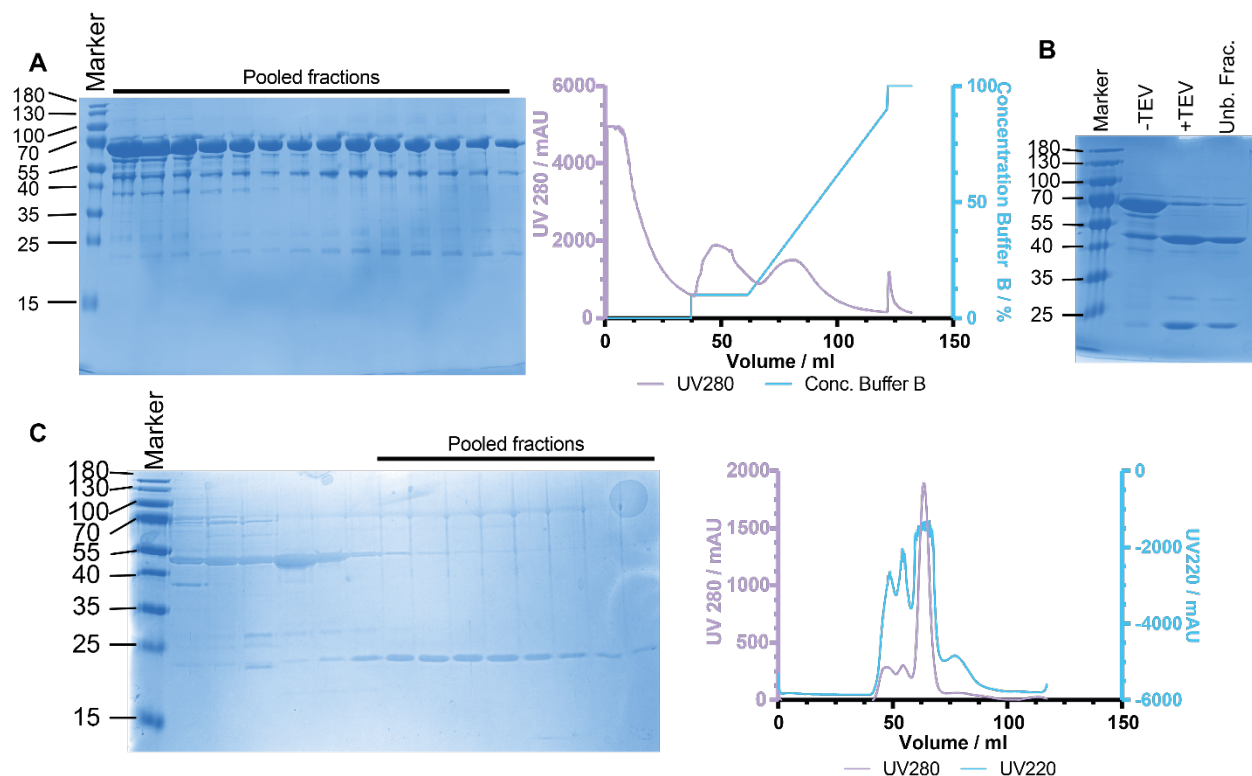
Figure 21: Purification of an RRM12 tandem with an inactive RRM1 (**RRM1*2**). (**A**) First purification step using an IMAC to enrich His-MBP-**RRM1*2**. (**B**) Tag-cleavage in dialysis with TEV-protease. (**C**) Size exclusion chromatography with a Superdex 75 16/60 column.

RRM12*

For the same reason as for RRM1*2, to evaluate the roles of the individual domains in the RNA-binding event, a PTBP1-RRM12 mutant with an inactive RRM2 was cloned by site-directed mutagenesis and expressed and purified with the same strategy (**RRM12***, I214S, F216A, K218S).[174] Similarly to the previous purifications, the protein of interest was successfully overexpressed, bound to the column in reasonable amounts, and eluted with an imidazole gradient in a pure manner (**Figure 22A**). The POI-containing fractions were dialyzed overnight in the presence of TEV-protease, and the fusion protein was processed quantitatively (**Figure 22B**, lanes -TEV/+TEV). The unbound fraction from the reverse IMAC (**Figure 22B,** lanes Unb. Frac. /

Wash 0% B) was pooled and further purified with size exclusion chromatography with a Superdex 75 26/60, where **RRM12*** eluted as a homogeneous peak with no determinable contaminations. The protein was concentrated, frozen in liquid nitrogen, and stored at -80 °C until use.
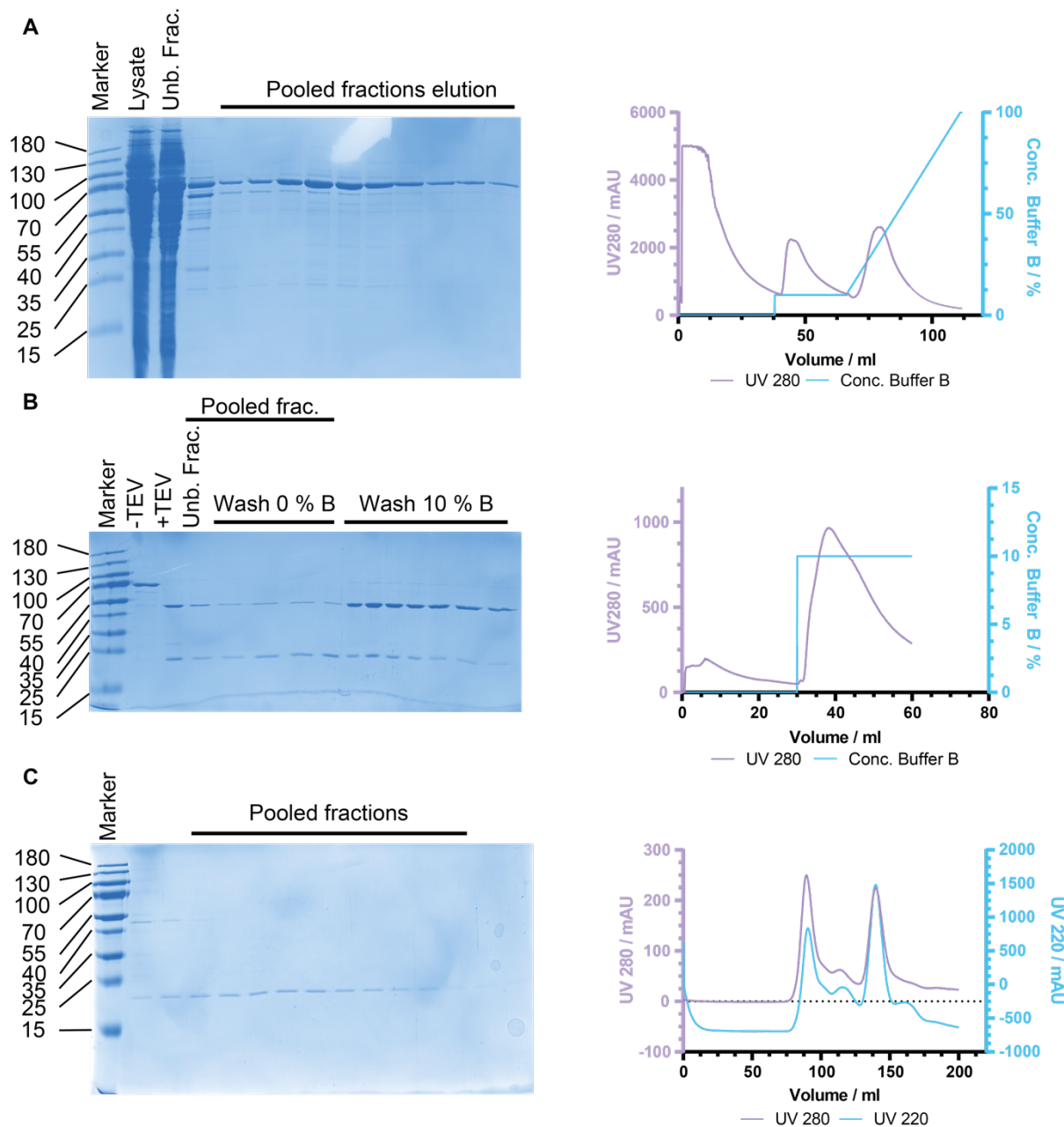
Figure 22: Purification of an RRM12 tandem with an inactive RRM2 (RRM12*) visualized by 12% SDS-PAGE and chromatograms of the FPLC runs. (**A**) Lysis and IMAC using a HisTrap HP 5 ml column. (**B**) Reverse IMAC after tag-cleavage in dialysis overnight using TEV-protease. (**C**) Size exclusion chromatography using a Superdex 75 26/60 column.

B.2: BIOPHYSICAL EVALUATION OF THE COOPERATIVITY OF RRM1 AND RRM2

As already mentioned in 1. Introduction, PTBP1 binds to RNA through the four RRMs with differing specificities of each domain for either the polypyrimidine rich sequences or secondary structures. We were particularly interested in the molecular mechanisms of RRM1 and RRM2 binding to ssRNA with several binding sites. This is of interest, because it is well reported, that the individual domains bind to longer stretches of RNA in a sort of cooperative manner to provide specificity for target pre-mRNAs.[175] Previous studies focused more on the role of the *C*-terminal domains RRM3 and RRM4, because they are forming a stable globular unit which orchestrates RNA looping and plays a critical role in the cellular functions of PTBP1. The specificity caused by the quaternary structure of the protein-RNA complex could be exploited to generate specific inhibitors for cellular events like splicing which show a higher dependency on the RRM1 and RRM2 domains than for the *C*-terminal domains. We therefore hypothesized, that the transient α3-helix in the interdomain linker of RRM1 and RRM2 might influence the activity of full length PTBP1 in a cellular context during RNA binding. To investigate the role of the different domains, we purified the wildtype RRM1, RRM2, RRM12-tandem and mutants interfering with either the activity of individual domains or a potential cooperativity of both domains (**RRM1*2, RRM12*** and **RRM1-L151G, RRM12-L151G**, respectively). Those protein constructs were then used in fluorescence polarization experiments with fluorescently labeled RNAs with either one or two binding sites. By this, we were able to measure the binding constants of the individual domains in context of the RRM12-tandem and validate if any cooperative or avidity-related effects occur.

First, we evaluated whether the purified PTBP1 constructs were correctly folded. Although, the mutants showed comparable elution profiles during size exclusion chromatography, the fold could still be influenced significantly by the introduced mutations. Therefore, we measured the circular dichroism spectra of the relevant mutants and compared them to the WT spectra. CD spectroscopy is a technique, by which the secondary structure of proteins or other macromolecules can be determined. Plane polarized light is made up of two circular polarized components with equal magnitude (counter- and clockwise). Optically active chiral chromophores, or chromophores in a chiral environment, absorb the two components in different amounts $\Delta A = A_L - A_R$, and by this the light turns into an ellipse. In CD spectroscopy this is then calculated into ellipticity $\theta = 32.98 * \Delta A$. For comparison this value gets normalized to molar ellipticity with the law of Lambert Beer. Proteins contain several chromophores, from which the peptide bond is the most important, because it absorbs light in the near-UV region which is used to predict secondary

structures, while the far-UV region where the aromatic amino acids absorb gives information about a fingerprint of a fold.[176]

The near-UV CD spectra of **RRM12**, **RRM1*2**, **RRM12***, and **RRM12-L151G** are shown **in Figure 23A**. The spectra of RRM12, RRM12* and RRM12-L151G are very similar in shape and amplitude, so all of them seem to be folded similarly and the introduced mutations do not seem to influence the fold drastically. While the spectrum of RRM1*2 shows a very similar curve, the amplitude differs significantly. Having the same shape, the protein seems to be folded in general, but because of the difference in amplitude it can't be ruled out that parts of the protein are not functional. It could be the case that the introduced inactivating residues influence the fold of RRM1. The single domains of interest RRM1, RRM1-L151G and RRM2 have matching curves, so the L151G mutation does not influence RRM1s fold. Further on, RRM2 seems to have the same fold as RRM1, which is expected because of the high conservation of RRMs.
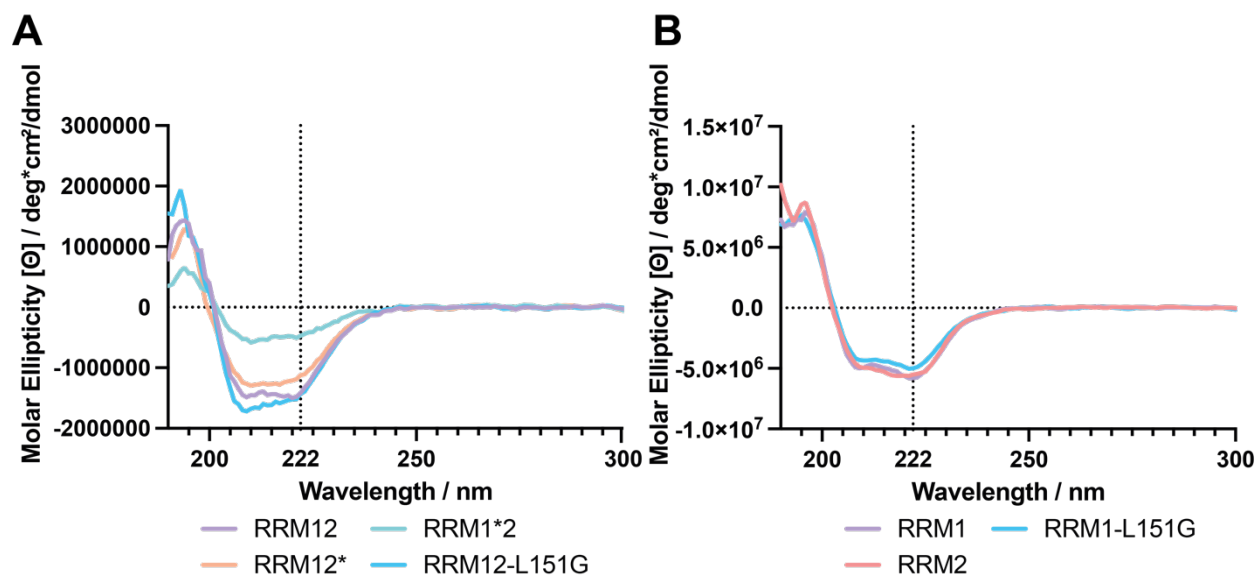
Figure 23: Near-UV CD spectra of (**A**) the tandem constructs RRM12, RRM1*2, RRM12*, and RRM12-L151G and (**B**) of RRM1, RRM2, and RRM1-L151G.

After validating the constructs, we performed fluorescence polarization assays to determine the dissociation constants of each individual domain for two model RNAs. The first, **RNA-1** contains a single GUCUUAA binding motif and is 5' labeled with FAM. The second, **RNA-2** is longer and contains two binding motifs for PTBP1 (A**UUUUU**CCA**UCUUU**GUAUC). When measuring the affinities for **RRM1**, **RRM2**, and **RRM12** with **RNA-1** we observed that the K_Ds were 2.213, 3.802 and 0.881 μM, respectively indicating that RRM1 binds stronger than RRM2 and that a positive allosteric, cooperative or avidity effect increases the binding when RRM1 and RRM2 are present in a tandem (**Figure 24B/C**). Because of this observation, we further tested those three constructs

with **RNA-2**, where the same trend was observed with 90, 218 and 17 nM binding constants for **RRM1**, **RRM2**, and **RRM12**, respectively. The general increase of affinity was to be expected, as it was reported previously for PTBP1 and other RRM-containing RBPs that repeated binding sites on the RNA lead to increased binding constants.[103] To further investigate the role of the α3-helix in RRM1, we performed the same FP experiments with the **RRM1-L151G** and **RRM12-L151G** constructs with a mutation which hinders the α3-helix from forming,[44] where the K_D for **RRM1-L151G** with **RNA-1** decreased 3.6-fold to 7.989 μM, and 2.4-fold to 218 nM for **RNA-2** in comparison to RRM1. The results for **RRM1-L151G** validate the change of affinity for this mutant reported by the Allain group.[44] The lower decrease of affinity to **RNA-2** compared to **RNA-1** for the **RRM12-L151G** mutant could be explained by RRM2 compensating the loss of affinity of RRM1. But it shows further that the inaccessibility of α3 uncouples both domains and might interfere with the correct orientation and assembly of the complex.

Construct	Domains	RNA-1 K_D (μM)	RNA-2 K_D (μM)
RRM1	RRM1 α3	2.213 ± 0.068	0.090 ± 0.023
RRM1Δα3	RRM1	> 50	6.454 ± 1.361
RRM1-L151G	RRM1 α3 L151G	7.989 ± 1.724	0.218 ± 0.011
RRM2	RRM2	3.802 ± 0.296	0.721 ± 0.308
RRM12	RRM1 α3 — RRM2	0.881 ± 0.074	0.017 ± 0.002
RRM12-L151G	RRM1 α3 L151G — RRM2	8.391 ± 2.273	0.042 ± 0.001
RRM1*2	RRM1 α3 — RRM2	n.d.	0.272 ± 0.009
RRM12*	RRM1 α3 — RRM2	n.d.	0.117± 0.024

B RNA-1   FAM-G<u>UCUU</u>AA

RNA-2   FAM-A<u>UUUUUCC</u>A<u>UCUUU</u>GUAUC

Figure 24: Fluorescence polarization binding experiments of PTBP1 constructs and two different RNAs. (**A**) Alphafold model of the RRM12 tandem with the α3 helix in blue. The model is not accurate in the orientation of the domains to each other. RNA-binding sites were visualized by overlaying the alphafold model with reported holo NMR-structures of RRM1 and RRM2 (PDB 2N3O and 2ADB, respectively). (**B**) Sequences of the used RNA-tracers of the FP experiments. (**C**) PTBP1 constructs analyzed in the FP-experiments with two different RNAs and the determined binding constants ± SD calculated from triplicates.

Next, we tried to uncouple both RRMs binding to **RNA-2** by using our previously purified inactive mutants of the RRM12 tandem (**RRM1*2**, and **RRM12***). Both bound **RNA-2** with significantly decreased binding constants of 272 and 117 nM (16- and 6.9-fold difference compared to **RRM12**). The activity of **RRM12*** nearly equals the activity of **RRM1** with a 1.3-fold difference, which shows that the first domain is more active in presence of the second domain while binding

to a bidentate ligand. The 2.6-fold difference of **RRM1*2** to **RRM2** is inconclusive in this case but could also be caused by the elevated error of this experiment (2.6 ± 1.1-fold).

Furthermore, we hypothesized a positioning of RRM2 to RRM1 over the α3-helix in the inter-domain linker. Unfortunately, typical biophysical methods for the determination of particle sizes and oligomeric states like SEC-MALS or mass photometry were not possible because of the expected low affinity, the dynamic properties of this interaction and the low molecular weight of the final complex. Therefore, we performed another fluorescence polarization assay with **RNA-2** and RRM1, where we added constant concentrations of 0, 10 and 25 µM of inactive **RRM2*** (**Figure 25**). We determined a K_D of 4.92 µM without addition of **RRM2*** and saw a decrease of affinity to 10.79 µM in presence of 25 µM **RRM2***. Simultaneously, the hill slope increased from 0.5 to 1.5 with increasing concentrations of RRM2*. Typically, hill slopes ≠ 1.0 indicate positive or negative allosteric effects, or they are different from 1.0 if the binding stoichiometry of receptor and ligand is not 1:1.[177] Without the presence of **RRM2***, a 1:2 (probe:protein) stoichiometry with a hill coefficient of 0.5 is observed, which is expectable from the design of **RNA-2**. The fact, that the hill slope increases after addition of **RRM2*** indicates, that the probe is bound by less **RRM1** molecules. This is probably because the ratio changes to 1:1 caused by either allosterically induced conformational changes, or **RRM2*** binding to **RRM1** sterically blocking more **RRM1** molecules from binding to the same probe.

RRM2* / µM	K_D / µM	Hill Slope
0	4.92 ± 1.54	0.56
10	9.28 ± 1.44	1.42
25	10.79 ± 0.63	1.52
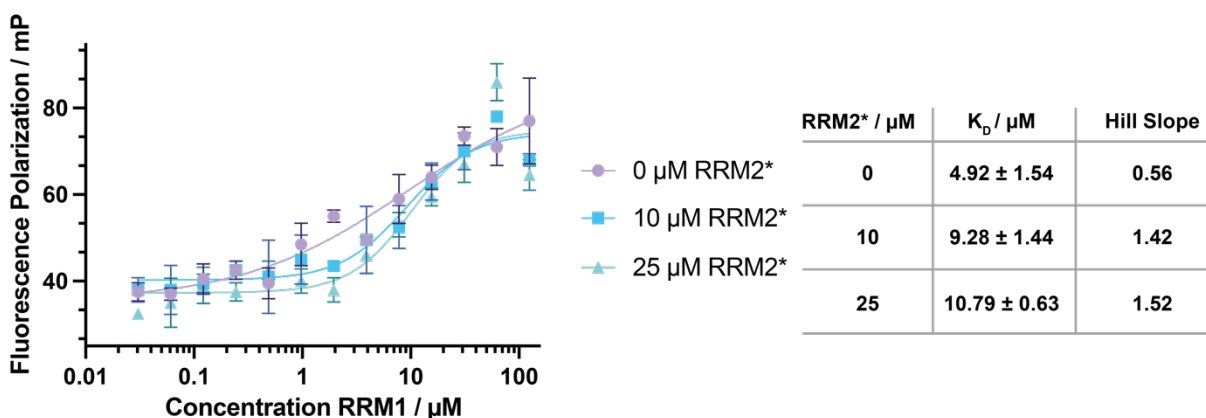
Figure 25: Fluorescence polarization binding experiment of **RRM1** with **RNA-2** in presence of 0, 10, 25 µM **RRM2*.**

All in all, we analyzed the cooperativity of the RRM12-tandem of PTBP1 by a compendium of fluorescence polarization binding experiments with two RNA probes of different length. We found that the affinity of the tandem increases with the number of binding sites on the RNA and identified a positive effect on binding if both domains are active. We validated this hypothesis by introducing several mutations in RRM1, RRM2, and the inter-domain linker to decipher the role of each

individual component. Our results validate a previous report on the role of the α3-helix from the Allain group, who used the L151G mutation first, and extends this work to the RRM12-tandem.[44] Further on, we could show that not only the α3-helix, but also the activity of the second domain is indeed influencing the binding of RRM1 to a longer RNA and vice-versa. Generally, the affinities of RRM1 for RNA are higher than for RRM2. Because of this, we hypothesize that the transient α3-helix forms upon RNA-binding, as reported by the Allain group, and shortens the distance between both domains and by this positively influences the affinity of the second domain. Further on, it could be possible that the α3-helix is able to bind to RRM2, introducing intra-domain contacts which would orchestrate the binding of both domains together (**Figure 26**). This hypothesis is strengthened by a recent pre-print from the Allain group who identified a variety of intra-domain contacts between the RRMs of PTBP1 in an integrative structural biology approach. Several of those interactions were found to be mediated by the α3-helix in the intra-domain linker of RRM1 and RRM2 which can bind to the inter-domain linker of RRM2 and RRM3. In other models the RRM1 α2-β4 loop interacts with the β3-α2-loop of RRM2 and the inter-domain linker makes contacts with the α1-α2 surface of RRM1. All of those models lead to RRM1 and RRM2 facing each other in a way that allows binding to adjacent RNA binding sites.[175] In contrast to this, Simpson et al. reported in 2004 that apo-PTBP1 is an elongated particle in solution, where all domains are working independent of each other.[57] The discrepancy of those models is explained by the absence of RNA in Simpson et. al.'s experiments.
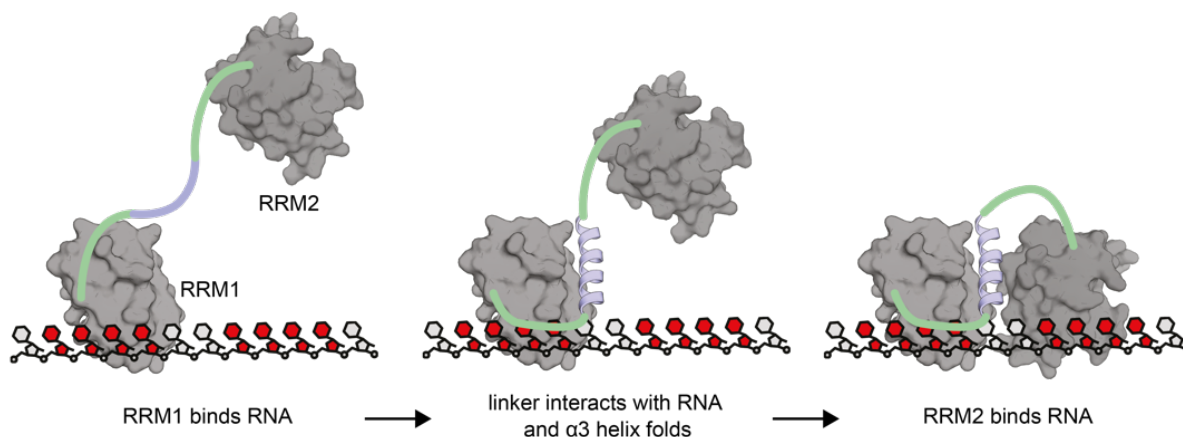
Figure 26: Model of RNA binding of the RRM12-tandem. RRM1 binds to the RNA first because of the higher affinity and upon α3-helix formation RRM2 gets positioned to bind to an adjacent RNA binding site.

B.3 STAPLED PEPTIDES INHIBIT THE INTERACTION OF A RRM12 TANDEM WITH RNA

Rational design and competitive fluorescence polarization assays of RRM12 inhibitors

After validating the role of the inter-domain linker between RRM1 and RRM2 of PTBP1, we hypothesized that α3-mimics could influence the RNA-association of an RRM12 tandem in a positive or negative manner. By binding to the potential α3 interaction surface, the active conformation could either be induced (molecular glue) or inhibited by blocking possible interaction sites.[178] We chose to prove this concept by synthesizing peptides designed on the native sequence of the α3-helix. We decided that peptides would be good candidates because of the big surface area, which is proposed to be a "typical" protein-protein interaction site, and good
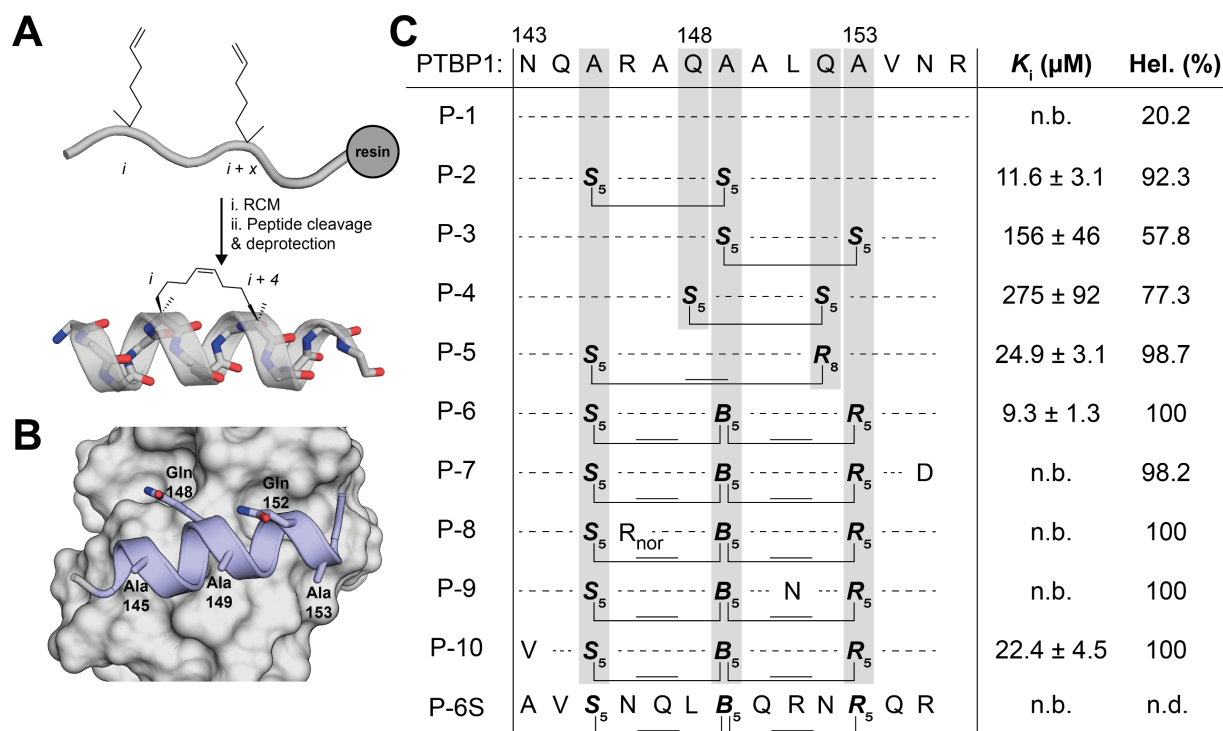
Figure 27: Design of stapled peptides for the inhibition of the RRM12-RNA interaction. (**A**) Schematic depiction of the synthesis of hydrocarbon stapled peptides with on-resin RCM followed by peptide cleavage and deprotection. (**B**) Surface representation of PTBP1-RRM1 (grey) and cartoon representation of the α3-helix. Key amino acids for stapling are represented in sticks. (**C**) Schematic representation of the discussed stapled peptides with inhibitory constants in competitive FP experiments and helical content of the peptides determined by CD spectroscopy.

possibilities for the design of peptide mimics for those. We designed hydrocarbon staples, which is an established strategy to stabilize helical conformations of peptides by covalently linking amino-acids on the same side of the helix using unnatural, disubstituted amino-acids bearing an alkene for a subsequent ring-closing metathasis reaction (RCM) for covalent macrocyclization.[132,179] The

peptides were synthesized by standard Fmoc-based SPPS and the macrocyclization performed with Grubbs-catalyst on resin (**Figure 27A**).
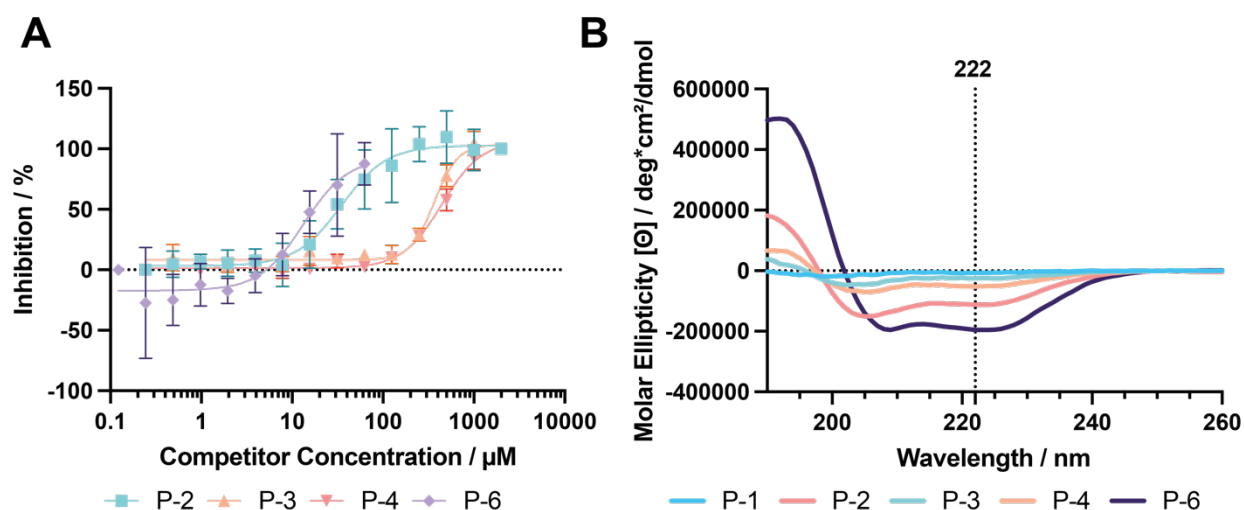
Figure 28: (A) Competitive FP assays of **RRM12** with **RNA**-2 and peptides **P-2**, **P-3**, **P-4**, and **P-6**. **(B)** Circular dichroism spectra of peptides P-1, P-2, P-3, P-4, and P-6.

Initially, we focused on i, $i + 4$ staples using two (S)-pentenyl alanines (S_5), with three amino acids in between the staples. We picked three stapling positions, designed so that the hydrocarbon linker is facing away from the binding site on the RRM1 surface, based on the orientation observed in the NMR structure. This design is represented in **P-1** (Ala145/Ala149), **P-2** (Gln148/Gln152), and **P-3** (Ala149/Ala153) (**Figure 27B/C**). These three peptides were further hydrogenated in the linker region to avoid a mixture of E/Z-diastereomers using an in-solution hydrogenation protocol.[180] We also synthesized the linear, native, unmodified native peptide (**P-1**) which had to be extended with an additional Arg on the C-terminus to ensure solubility. As the C-terminus is facing away from the binding site we did not expect any changes in binding affinity. Then, we performed a competitive FP assay between PTBP1-RRM12 and **RNA-2** using constant protein concentrations of 50 nM which represents 50 – 70% binding of the complex to the RNA in equilibrium and titrated the peptides from a high concentrated stock in FP-buffer. While the linear peptide **P-1** did not compete with RNA under the assay conditions, peptide **P-2** performed best with a K_I of 13.1 µM. The other peptides **P-3** and **P-4** showed lower activities with 156 and 275 µM respectively. Because of the favored stapling position of **P-2** at Ala145/Ala149 with i, $i+4$ staples, we believed that this orientation of the staple was most optimal, so we decided to increase the space of the linker by introducing longer staples using Ala145/Q152 (**P-5**), and Ala145/Ala149/Ala152 (**P-6**) in i, $i+7$, and i, $i+4$, $i+8$ stapling approaches respectively.[138,179] Both peptides had low micromolar K_I's of 24.9 and 9.3 µM,

respectively, further confirming that a staple spanning from Ala145 is the optimal position. Peptide **P-5** shows a weaker affinity to RRM1 by combining the stapling positions of **P-2** and the weaker
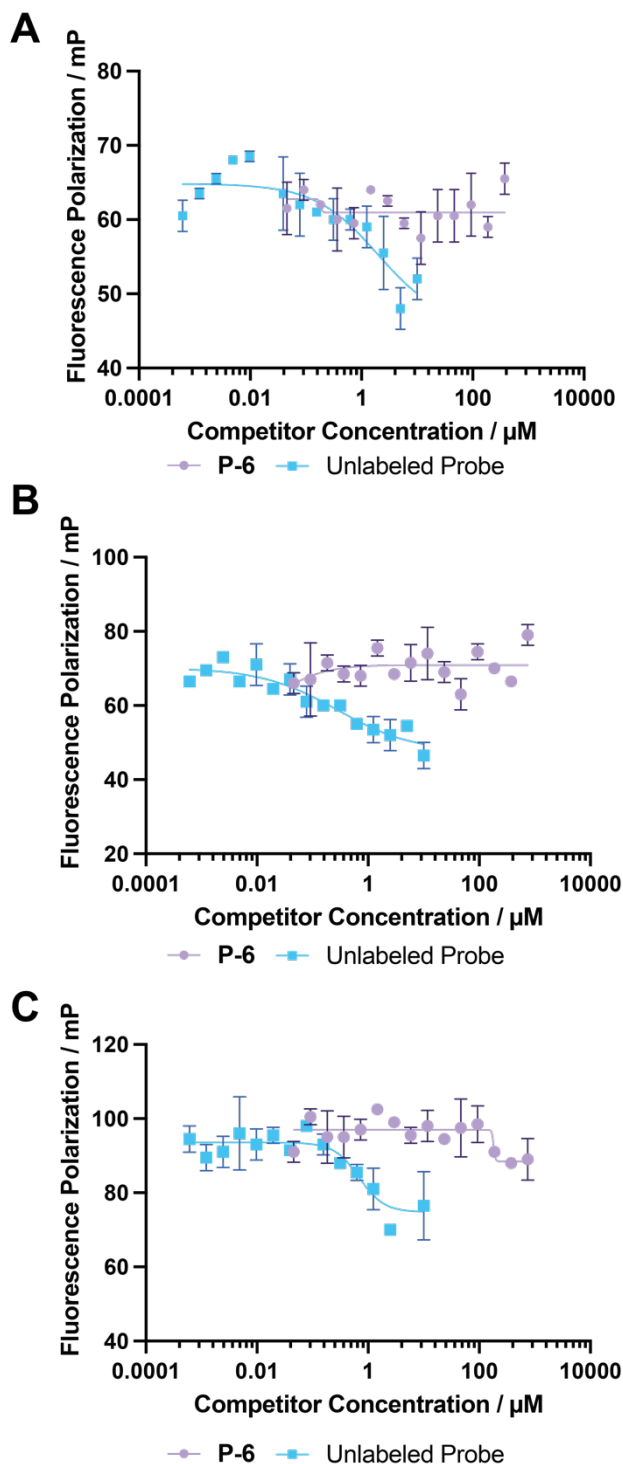
P-4, while **P-6** shows the best K_I through using a so called stitched macrocyclization strategy with a double tether being introduced at position 149. Additionally, we synthesized a scrambled peptide of the best binder **P-6** (**P-6S**) and performed the same assay. We could not observe any inhibition of the **RNA-2**-RRM12 interaction with this compound, excluding any unspecific, sequence related effects.

Macrocyclization increases the affinity of peptides because of several reasons. First, the binding enthalpy can improve by providing a better conformation for optimal interactions between the peptide and the target. Second, an entropic gain in the binding event can occur because of the pre-organization through the covalent "lock" of the oligomer in the correct binding pose. Hydrocarbon staples are reported to stabilize the helical conformation of peptides and generally an increase of helicity should be observed compared to the native, linear sequence. Such conformational stabilization can be observed by CD-spectroscopy, which we performed with the synthesized peptides to evaluate if a correlation between helicity and inhibitory effect can be observed (**Figure 27C** and **Figure 28B**).[176] We observed the lowest

helicity of 20.2% with linear peptide **P-1**, while the affinity of the *i, i+4* stapled peptides **P-2, P-3,**

Figure 29: Competitive FP assays for the RNA-2-**PTBP complexes (A)** RRM1**, (B)** RRM12***, (C)** RRM12-L151G**.**

and **P-4** correlated well with the helicity of 92.3, 57.8 and 77.3 % respectively, indicating that an *N*-terminal staple nucleates the helical fold of the peptide better than *C*-terminal staples. Peptides **P-5** and **P-6** with different macrocyclization strategies showed 98.7 and 100% helicity respectively, which is higher than the preorganization of **P-2**, but only **P-6** showed also a higher inhibition of the RRM12-RNA interaction. These small discrepancies could be caused by the staple influencing the binding sterically or by the flexibility of a longer *i, i+7* linker in **P-5**.

We performed further competitive FP experiments of **RNA-2** with RRM1, RRM12-L151G, and RRM12* to investigate if the mode of action of the peptides **P-6** and **P-2** is dependent on single domains or works only in the tandem configuration (**Figure 29**). To our surprise, the peptides were not able to inhibit the **RNA-2**-**RRM1** interaction. Our conclusion is that the α3-helix coordinates domains 1 and 2 in a way, where both domains bind onto the same RNA. Furthermore, we could not observe any inhibition on the **RRM12-L151G** construct, where both domains are uncoupled from each other through the previously discussed helix breaking mutation in α3. This hypothesis is further confirmed by a missing inhibition of RNA binding of the **RRM12*** construct through **P-6**.

Summed up, we rationally designed and synthesized stapled peptides based on the native sequence of the α3-helix in the inter-domain linker between PTBP1 RRM1 and RRM2. We hypothesized a possible inhibition of **RNA-2** binding to RRM12 based on previous reports by the Allain group and our own investigation of this interaction. We generated a stitched peptide **P-6**, and an *i, i+4* stapled peptide **P-2** with 100 and 92% helical conformation which inhibited this interaction with low micromolar K_Is. To our knowledge, this is the first reported stapled peptide with a novel mode of action of binding an allosteric, transient, intra-molecular protein-protein interaction of an RRM containing RBP. We were able to decipher the mode of action to be dependent on the α3-helix in the intra-domain linker, and also on the activity of RRM1 and RRM2 acting together, as the RNA binding of the single RRM1, and our deficient mutants **RRM12-L151G** and **RRM12*** could not be affected by **P-6**.

Direct binding experiments and evaluation using mutant PTBP1 constructs

To further verify and investigate the mode of action of our synthesized PTBP1 inhibitors, we performed binding experiments of fluorescently labeled **P-6F1**, which is extended *N*-terminally by a lysine which is used to covalently link a FITC moiety. The peptide was tested using the **RRM1**, **RRM12** and the **RRM12-L151G** constructs. We hypothesized, that despite the missing inhibition of RNA binding **on RRM1**, and **RRM12-L151G**, the peptides should still be able to bind to those constructs. Also, the peptides should have an increased affinity for constructs harboring the L151G

mutation, or on the single RRM1 domain because of decreased intra-molecular competition in the binding site.

In the fluorescence polarization experiment with fluorescently labeled **P-6F1**, we observed binding of all three used constructs (RRM1, RRM12, RRM12-L151G). As expected, the peptide bound to **RRM1** and **RRM12-L151G** with best affinities of 18.6 and 20.4 µM, where we barely reached the saturation of the equilibrium binding experiment, but could not further increase protein concentration due to solubility limits (**Figure 30A/E**). We observed a very comparable, mildly right shifted curve for **RRM12** without reaching the plateau of the binding event. Because of this, we were not able to determine a binding constant, but expect it to be moderately bigger than for the RRM1 and RRM12-L151G constructs. These observations fit to our hypothesis of the mode of action of **P-6**, as the intramolecular competition with the native α3-helix and its binding sites should be decreased in the RRM1 and RRM12-L151G constructs.
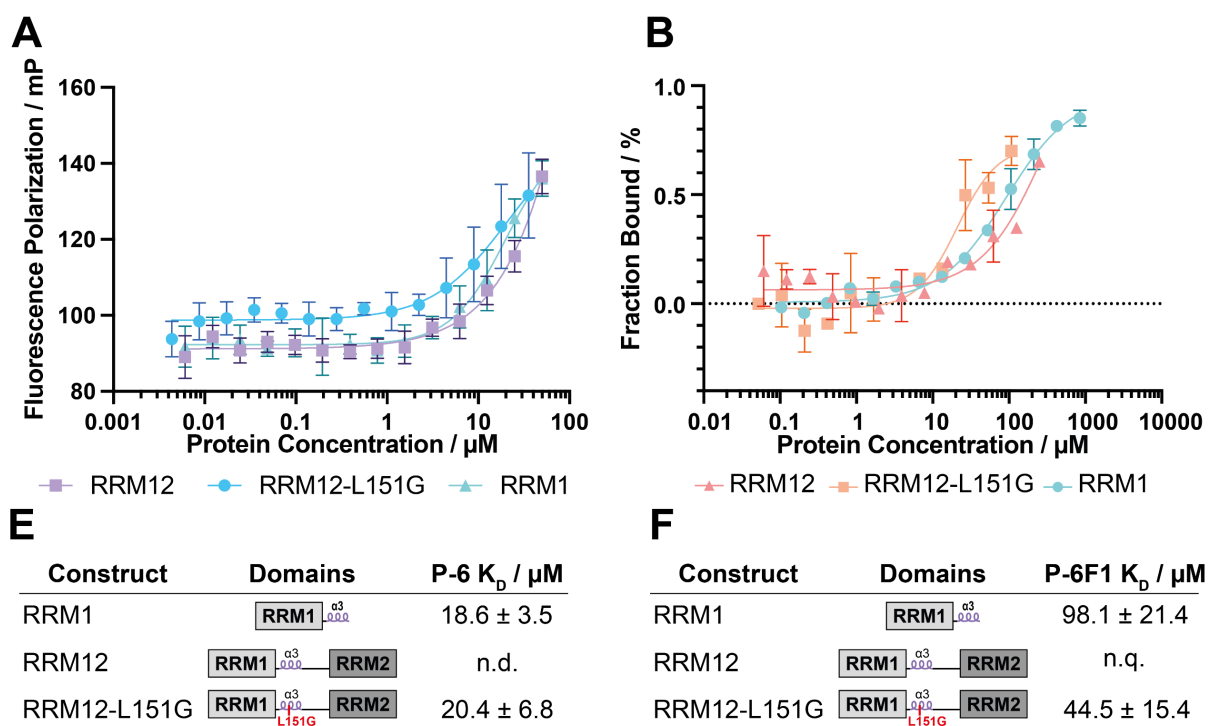
Figure 30: Binding experiments of different PTBP1 constructs with the probe **P-6F1**. (**A**) Titration of protein constructs **RRM1**, **RRM12**, and **RRM12-L151G** in a fluorescence polarization assay using fluorescent probe **P-6F1**. (**B**) Microscale thermophoresis assay titrating **RRM1** and **RRM12-L151G**, and **RRM12** in presence of probe **P-6F1**. (**E**) Binding constants determined in the FP assay. (**F**) Binding constants determined in the MST assay.

To further verify the binding of peptide **P-6F1** with PTBP1 we performed microscale thermophoresis (MST) (**Figure 30B/F**). MST is based on thermophoresis which is the directed

movement of macromolecules induced by a temperature gradient. This movement is influenced by many macromolecular properties like size, hydration shell, conformation or charge. During the experiment, a temperature gradient is applied through an infrared laser and the directed movement of molecules is detected by covalently linked or intrinsic fluorophores.[181] We decided to perform this assay with a fluorescently labeled peptide **P6-F1** and titrated the protein constructs of interest. We observed binding of the probe for both the **RRM1** and **RRM1-L151G** constructs. In compliance with the previous performed FP assay and our hypothesis, the **RRM1-L151G** mutant showed an increased binding constant compared to the WT **RRM1** with 30.0 µM and 86.7 µM respectively. The ~7-fold increase of K_D of the RRM1 construct in comparison to FP could be explained by using a different technique, which in our hands usually reports higher binding constants than most other techniques.

Further, we wanted to prove ligand binding in an orthogonal assay, which is not based on fluorescent labels on the peptide. Therefore, we performed thermal shift assays with different protein constructs and **P-6**. Because of the lack of Trp residues in the **RRM1** and **RRM12** constructs, we used SYPRO orange as a fluorescent dye for determining the protein integrity in a thermal gradient in presence of different concentrations of **P-6**. We observed a destabilization of **RRM1** and **RRM12** in presence of high concentrations of **P-6** based on the shift of the melting point to lower degrees. For **RRM1** the melting point shifted from 57 to 54 °C in lower concentrations in range of the measured K_I, while **RRM12** needed highly elevated concentrations of **P-6** for a more moderate shift from 59 to 58 °C (**Figure 31**). Typically, the thermal stability of proteins gets increased during ligand binding. But several opposite examples have been experimentally found.[182]
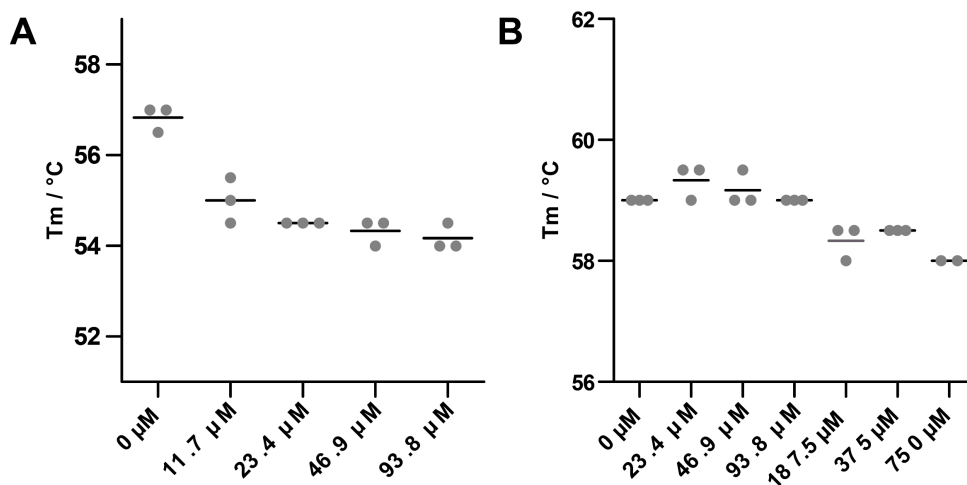
Figure 31: Thermal shift assay of **P-6** with **RRM1** (**A**) and **RRM12** (**B**).

We hypothesize that ligand binding "opens" the fold of RRM1 and RRM12 by occupying the binding sites and thus destabilizes the protein.

Summarized, we were able to observe binding and determine binding constants for our best peptide **P-6** in a less complex assay than the competitive FP with fluorescently labeled RNA. We showed that the peptide can bind the **RRM12** tandem and the **RRM1** construct. In contrast to the competitive FP assay with a fluorescently labeled RNA probe (**RNA-2**), we were also able to observe an effect with the individual **RRM1** domain, and the peptide was able to bind to the L151G mutants of RRM12 and RRM1. This strengthens our understanding of the binding event of RRM12 to an RNA with >1 binding sites for PTBP1 like previously discussed.

Selectivity of P-6 for PTBP1

The RRM domain is one of the most abundant RNA binding domains in the proteome and exhibits a high conservation between different genes. Therefore, we expected that designing specific inhibitors for those is challenging.[20,30] Although, some small molecule inhibitors for RRMs have been reported in the past (reviewed in the introduction), most of those reports, except Musashi-1 inhibitor Ro 08-2750, didn't investigate the specificity of the compounds.[104,106,107,110,111,113,114,183–186]
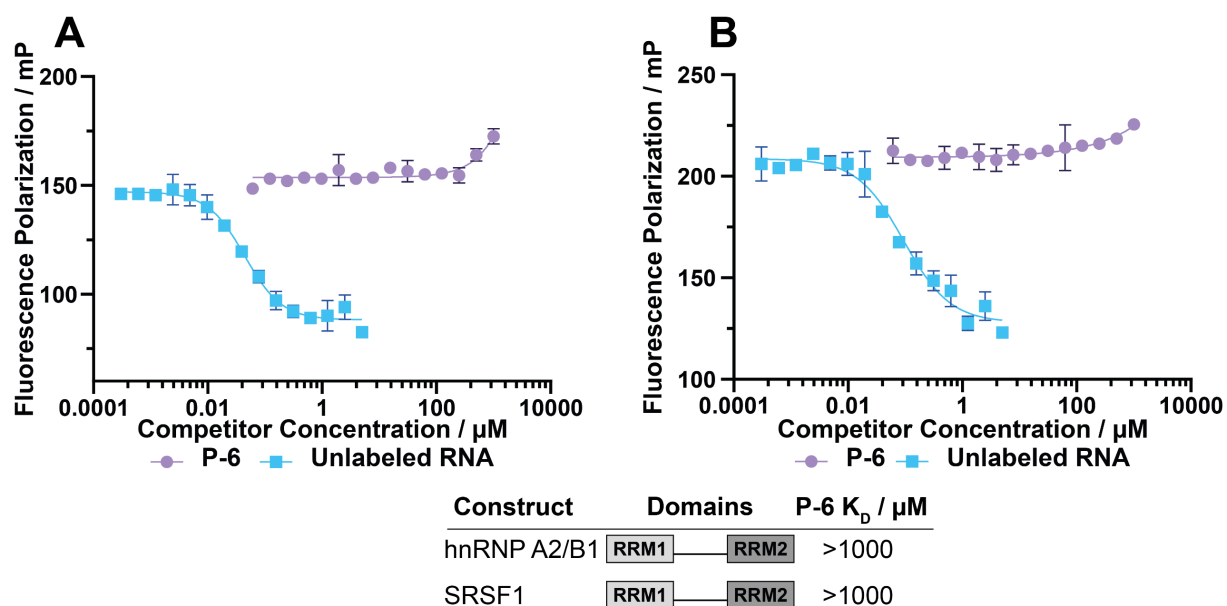
Figure 32: Competitive fluorescence polarization assays of (**A**) hnRNP A2/B1 with **P-6** and the unlabeled probe control and (**B**) SRSF1 with **P-6** and the unlabeled probe control.

Accordingly, we tested **P-6** in competitive fluorescence polarization assays against hnRNP A2/B1, another dual RRM protein of the hnRNP family, and one protein with two RRMs from the SR

protein family (SRSF1) with previously reported RNAs recognized by those proteins (**Figure 32**).[172]

For both protein constructs of the hnRNP and SR protein family, we could not observe inhibition caused by **P-6** up to 1000 μM of used compound. Although, the set of proteins is rather small, the specificity of **P-6** is good enough to not inhibit the general RRM fold. In case of the Musashi-1 inhibitor Ro 08-2750, the specificity of the small molecule was worse and the compound was able to inhibit several RRM containing proteins.[107] The higher specificity of **P-6** compared to this small molecule inhibitor could be caused by the bigger, more complex interaction site of the peptide on the target protein surface. But a full characterization of the specificity, for example through proteomic approaches, would be needed to fully validate a high specificity. The binding mode of the native α3-helix seems to be unconserved enough to supply enough specificity.[44] Taking our biophysical investigation of the mode of action of the α3-helix in the native complex into account as well as the fact, that the activity of **P-6** is dependent on the full RRM12 tandem, these results are consistent with our previously discussed findings.

Co-Crystallisation of P-6 with RRM1Δα3

Even though we were confident that our rationally designed peptides were binding on the expected binding site on RRM1, we tried to co-crystallize **P-2** and **P-6** with a RRM1Δα3 construct omitting the native helix to validate the binding mode. We chose to truncate the domain to ensure a homogeneous conformation and avoid the possibility of the native helix blocking the binding site. For peptide **P-6** we observed co-crystals in space group $P2_1\,2_1\,2$ which diffracted to a resolution of 2.9 Å (**Table 7**). The dataset was integrated using the XDS package and scaled using XSCALE.[151] We solved the crystal using a truncated alphafold model of PTBP1-RRM1 and identified 8 dimers in the asymmetric unit after phasing with phaser (Phenix).[152,153] After initial building of the model and manually placing eight more dimers in the asymmetric unit, we finally obtained an ASU without significant gaps with a total of 16 protein-peptide dimers (**Figure 33A**). Further, we observed a dense packing in the crystal by looking at the symmetry related molecules in 50 Å environment of the

Table 7: Table 1 for the PTBP1- **RRM1-P-6** crystal containing statistics about data collection and Refinement.

	RRM1Δ3-P6
Data collection	
Space group	P 21 21 2
Cell dimensions	
a, b, c (Å)	244.2, 76.63, 94.13
α, β, γ (°)	90, 90 ,90
Resolution (Å)	43.92-2.9 (3.004 – 2.9)
R_{merge}	0.0653 (1.497)
$I\,/\,\sigma$	14.58 (1.17)
Completeness (%)	99.31 (99.34)
Redundancy (Multiplicity)	4.6 (4.8)
CC1/2	0.999 (0.554)
Refinement	
Resolution (Å)	2.9
No. Reflections	39815 (3904)
R_{work} / R_{free}	0.2431 / 0.3403
No. Atoms	11682
Protein	11264
Ligand/ion	418
Water	0
B-factors	
Protein	103.52
Ligand/ion	93.53
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.39
Ramachandran favored (%)	90.092
Ramachandran allowed (%)	8.27
Ramachandran outliers (%)	0.80
Rotamer outliers (%)	5.88
Clashscore	25.86

ASU (**Figure 33B**). Because of the dense packing of the crystal, we are certain that all protein-peptide dimers were identified in the final model.

The resulting model was refined with iterations of phenix.refine and manual model building in Coot (CCP4) to a final R_{free} of 34% at 2.9 Å.[153,154] We observed a high disorder in the crystal, which is also represented in the general high B-factors, and especially high B-factors of the individual chains I/i, J/j, N/n,O/o, and P/p (**Figure 33C**).

76
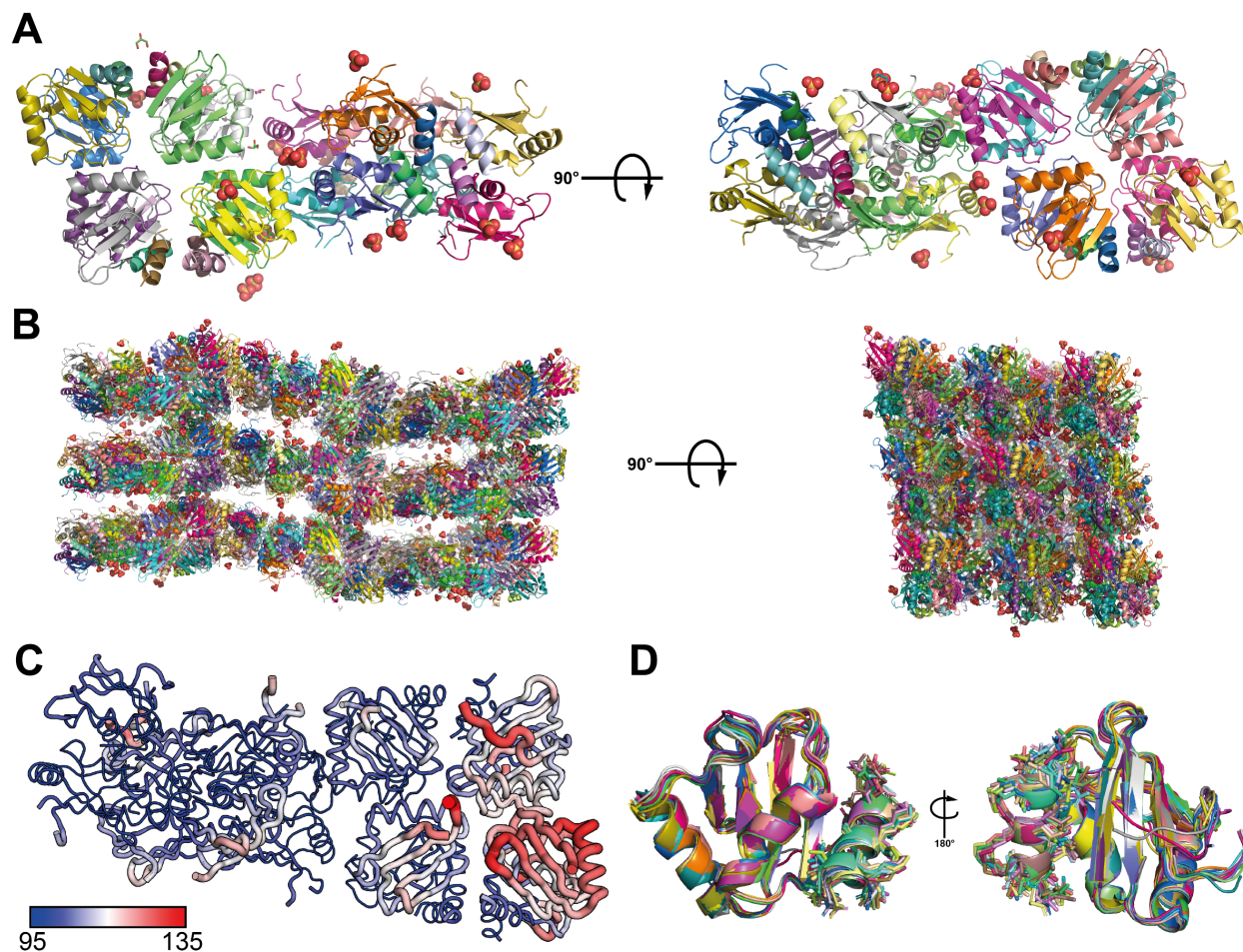
Figure 33: (**A**) Asymmetric unit of the **RRM1Δα3-P-6** crystal containing 16 protein-peptide dimers. (**B**) Symmetry related molecules in 50 Å environment of the ASU shows a dense packing of the crystal. (**C**) B-factor putty plot of the ASU reveals high disorder of the chains I/I, J/j, N/n, O/o, and P/p. (**D**) Overlay of the 16 protein-peptide dimers with similar fold and binding modes of **P-6**.

Most of the chains with high B-factors were also not identified in the first phasing. The high B-factors of the model where also represented in an overall weak density of the side chain orientations. We further identified a potential domain-swap between two chains A/B at residues 118 -127 at low occupancy. When we tried to build this domain swap, the R-factors and density quality indicated, that the non-swapped conformation is in higher occupancy, so we decided against building the swapped conformation. When we compared the 16 RRM1 monomers against each other, we observed a reasonable similarity between them with RMSDs of 0.45 – 0.89 Å after aligning all chains onto chain A (**Figure 33D, Table 8).**

Table 8: Comparison of the 16 monomers of the ASU by RMSD values with each other and to an RNA-bound structure (PDB 2n3o) and RNA-unbound structure (PDB 1sjq).

Chain	RMSD to chain A / Å	RMSD to RNA bound / Å	RMSD to RNA unbound / Å
A	-	1.329	1.595
B	0.45	1.45	1.58
C	0.687	**1.144**	**1.482**
D	0.718	1.48	1.643
E	0.728	1.306	1.549
F	0.691	1.382	1.612
G	0.712	1.431	1.534
H	0.697	1.513	1.659
I	0.723	1.357	1.587
J	0.776	1.42	1.73
K	0.596	1.418	1.66
L	0.684	1.399	1.536
M	0.894	1.423	1.563
N	0.725	1.464	1.734
O	0.886	1.564	1.554
P	0.882	1.523	1.678

Most differences between the chains were observed in the unstructured loops or orientations of the side chains. Next, we compared the fold of the RRM1 domain with published NMR structures of RRM1 bound and unbound to RNA (PDB 2n3o, 1sjq). After aligning all 16 monomers, we observed RMSD values of 1.14 – 1.56 and 1.48 – 1.73 Å respectively (**Figure 34B/C, Figure 35B, Table 8**). These RMSD values indicate, that the domain resembles the RNA-bound conformation more upon binding of **P-6**. We observed a high similarity between our and both reported structures and were not able to identify significant changes in the binding surface of the α3-helix (**Figure 34B**), and the RNA-binding pocket (**Figure 34C**). Next, we analysed if **P-6** bound into the same pocket as the native helix, which was the case (**Figure 34C**). We were able to identify the ligand through an obvious helical density including electron density of the two hydrocarbon linkers (**Figure 35A**).
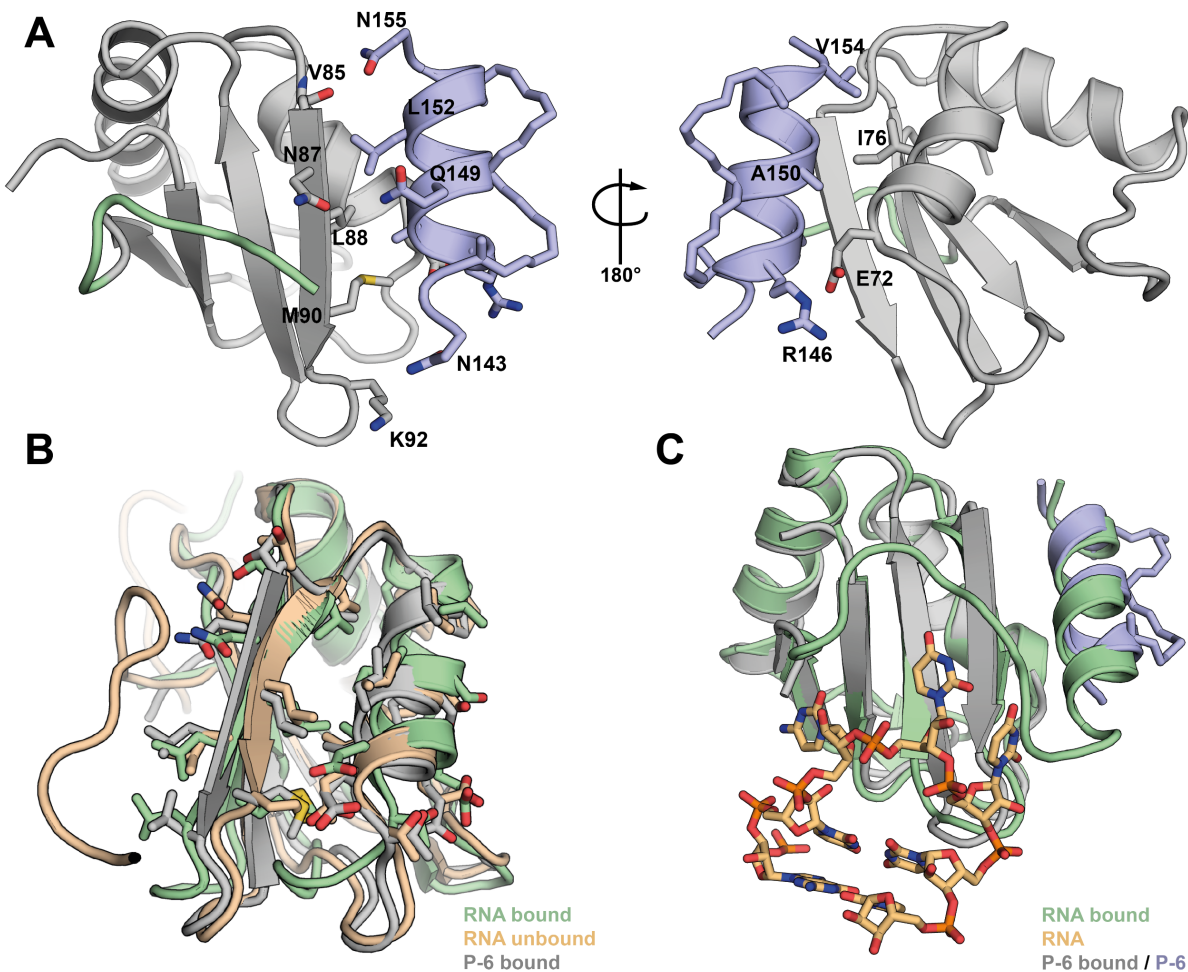
Figure 34: (**A**) Chain E/e of the solved crystal structure of RRM1Δα3 (grey) bound to **P-6** (blue). (**B**) Comparison of the α3 binding surface of different RRM1 structures RNA bound (green), RNA unbound (orange) and **P-6** bound (grey). (**C**) Visualization of the RNA binding site from the RNA-bound structure (green/orange) compared to the **P-6** bound structure shows no influence on the RNA binding site.

Our ligand bound RRM1 with the same hydrophobic binding core of I76, L80, V85, M90 and L88 on the RRM1 domain, and A150, L152 and V154 on **P-6** (**Figure 34A**). We identified additional polar interactions between E72 (RRM1) and R146 (**P-6**), the carbonyl oxygen of V85 (RRM1) and N155 (**P-6**), and side-chain-side-chain interactions of N87 (RRM1) and Q148 (**P-6**). We also checked if core amino acids in the RNP motif of RRM1 were still oriented comparable to the RNA-bound and -unbound conformation and recognized no changes for V60, F98, and L89 (**Figure 35B**).[55]
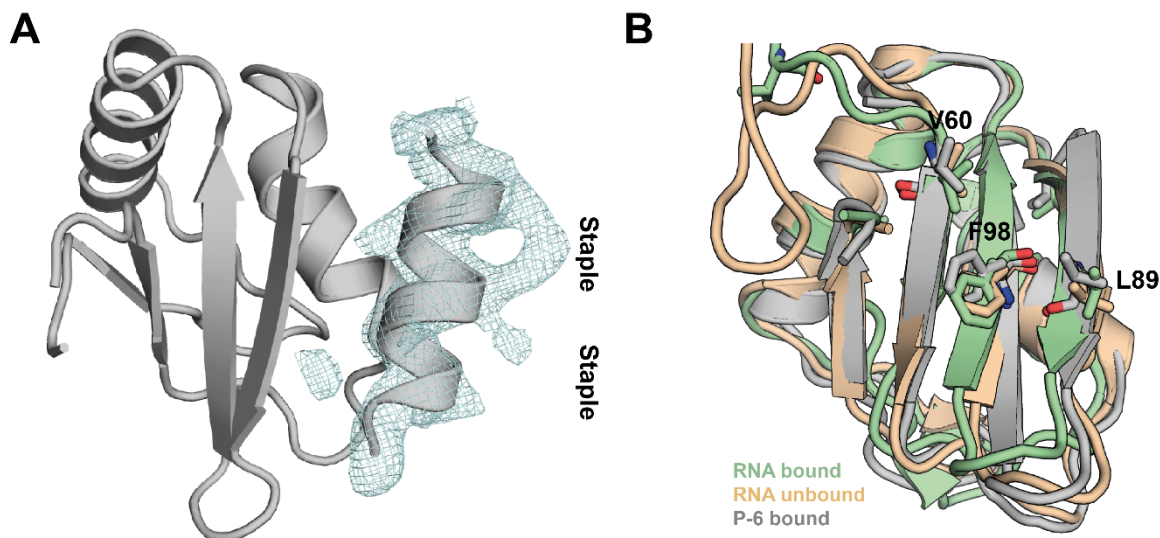
Figure 35: (**A**) $2F_O\text{-}F_C$ map shaped around the ligand to visualize the presence of **P-6**. (**B**) Comparison of key residues of the RNA-binding site of RRM1 between RNA-bound (green), RNA-unbound (orange) and **P-6** bound (grey) structures.

Summarized, we were able to validate the proposed binding mode, which we expected from our rational design of the peptide. Like expected, we observed the peptide bound to the native surface of the α3-helix. We were not able to identify any conformational changes of RRM1 upon binding of **P-6**. Because of this, we conclude that the inhibition of **P-6** observed in the *in vitro* experiments is based on an allosteric effect between the two *N*-terminal domains of PTBP1. We hypothesize, that binding of **P-6** to RRM1 blocks the intra-molecular association of RRM2. We can't conclude more from this structure, as the native helix is omitted in this construct.

B.4 Cellular investigation of P-2 and P-6

Our inhibitors of the RRM12 tandem of PTBP1 showed low micromolar activity *in vitro*. Because of this, we thought would be sufficient to investigate the cellular effects. First, we investigated the cell permeability through confocal microscopy with fluorescently labeled peptides **P-2F** and **P-6F2**, which are *N*-terminally labeled with FITC over a O2Oc linker.

We compared those peptides to a fluorescently labeled, well known, cell penetrating peptide (CPP) Tat (**Tat-F**) and a highly negatively charged, non-permeable peptide (**P11-F**) which was available inhouse. We treated HEK293T and HeLa cells with the labeled peptides for 1 h, washed the cells excessively, and fixed them for confocal microscopy following a protocol from Holm et al.(**Figure 36**).[157] For both cell lines, we observed uptake of **Tat-F**, which was well distributed in the cytoplasm, but also accumulated in granular, punctate structures in the cells. The molecular uptake mechanism of Tat and other Arg-rich CPPs is still under discussion. Although, many CPPs are frequently used for delivery of cargos into cells, the uptake mechanism of most are poorly understood.[187] For Tat many cellular uptake mechanisms were reported. The peptide can enter the cells via direct penetration, arginine induced micropinocytosis or diverse endocytosis uptake mechanisms, but also energy- and temperature independent pathways were reported for some Arg-rich CPPs.[188–192] For cargo-loaded CPPs all of the reported pathways might be used, but also individual pathways could not work because of the molecular properties of the cargo. Our negatively charged negative control **P-11F** was not able to penetrate both cell lines, while cells treated with fluorescently labeled PTBP1 inhibitors **P-2F** and **P-6F2** showed a modest fluorescence with possible cytoplasmic distribution. Here, we could observe a low signal, which indicates a low uptake during the 1 h of treatment. We could also identify some granular bodies in the cells, which could either be endosomally trapped peptide or an artifact from the low signal to noise ratio caused by the low uptake. One of the most controversially discussed properties of stapled peptides is their ability to penetrate the cell membrane.[193–196] For now, no strict rules or guidelines for optimizing the penetration of stapled peptides could be proposed. Despite this lack, it was shown that it is mainly depended on stapling type, staple position, the formal charge of the

peptide and that penetration happens probably through an ATP-dependent endocytosis pathway.[196]
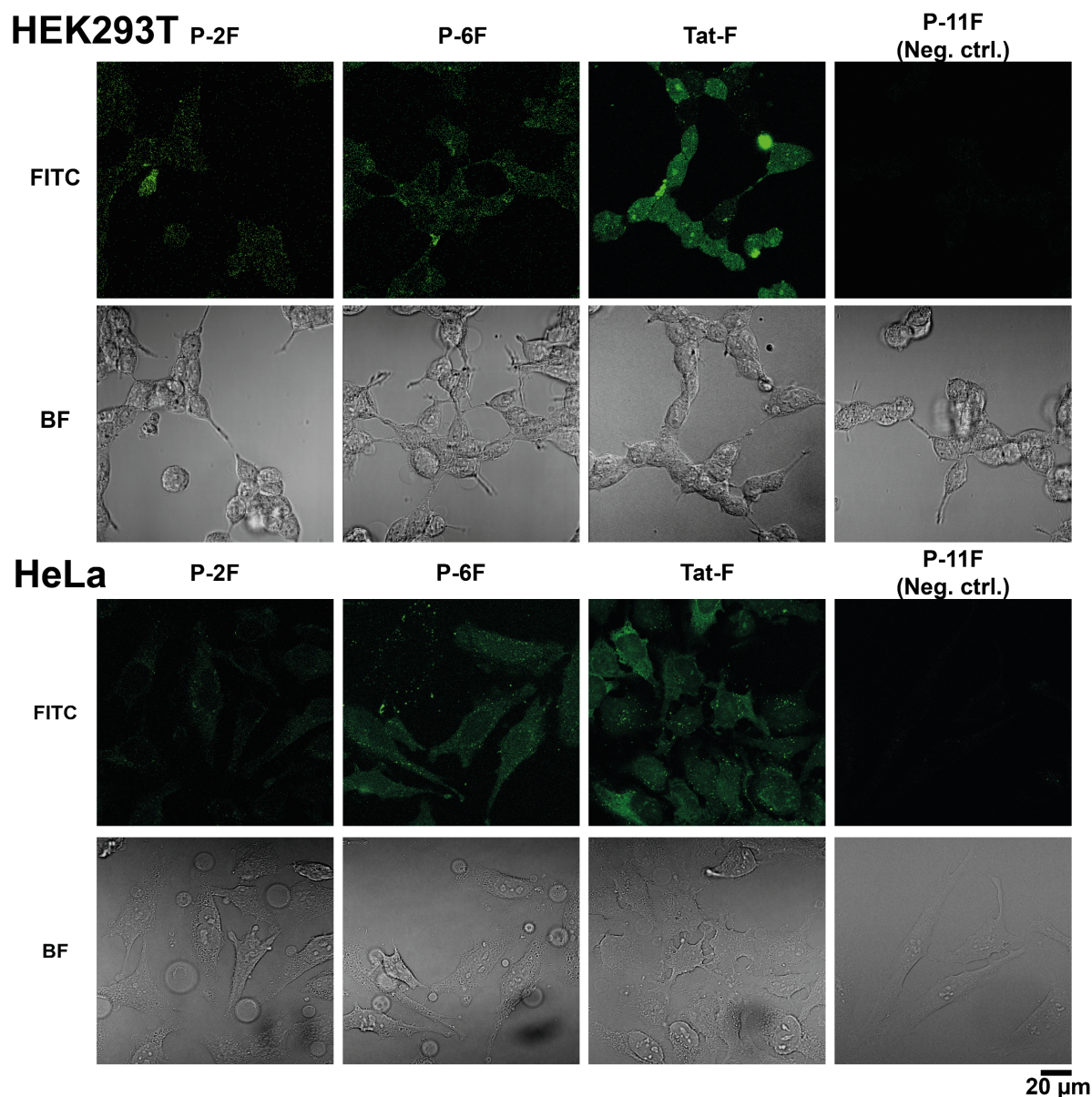
Figure 36: Confocal microscopy of HEK293T (top) and HeLa cells (bottom) at 40X magnification treated with inhibitors **P-2F**, and **P-6F2**. The used controls are a reported CPP **Tat-F** and an in-house available highly negatively charged, non-permeable peptide **P-11F**. Scale: 20 µm.

Because of the possibility that peptides **P-2F** and **P-6F2** could be endosomally trapped, we used azidolysine derivatized variants of **P-6** and the scrambled **P-6S** (**P-6-Az** and **P-6S-Az**) and performed an orthogonal nano-click assay (**Figure 37**).[158] In this assay, cells are transfected to express a NanoLuc-HaloTag fusion protein and relies on a quantitative copper-free [3 + 2] cycloaddition between the azide-modified peptides and dibenzoazacyclooctyne (DIBAC) which is

linked to the HaloTag domain. After treatment with the peptides, the cells are washed, and treated with an Azido-NanoBRET 618 dye, which generates a bioluminescence resonance energy transfer (BRET) signal with the luciferase. If the peptides of interest are available in the cytoplasm, this BRET signal is inhibited. Through this assay, the used peptides don't need to be modified with bulk moieties influencing the permeability, and all reactions are biorthogonal.[158] The inhibition of the BRET can only happen, if the peptides enter the cytosol of the cell, while endosomally trapped peptides are not measured and the assay provides a good tool to verify and quantify cell permeability. We performed the experiment with peptides **P-6-Az**, **P-6S-Az**, **azido-octa-arginine** (positive control), and **azido-ONEG** (negative control) in HEK293T cells with a 24 h treatment of peptides.

Peptide	EC_{50} / nM
Pos. ctrl.	103.02 ± 108.44
Neg. ctrl.	n.a.
P-6-Az	35.31 ± 20.05
P-6S-Az	159.92 ± 85.21
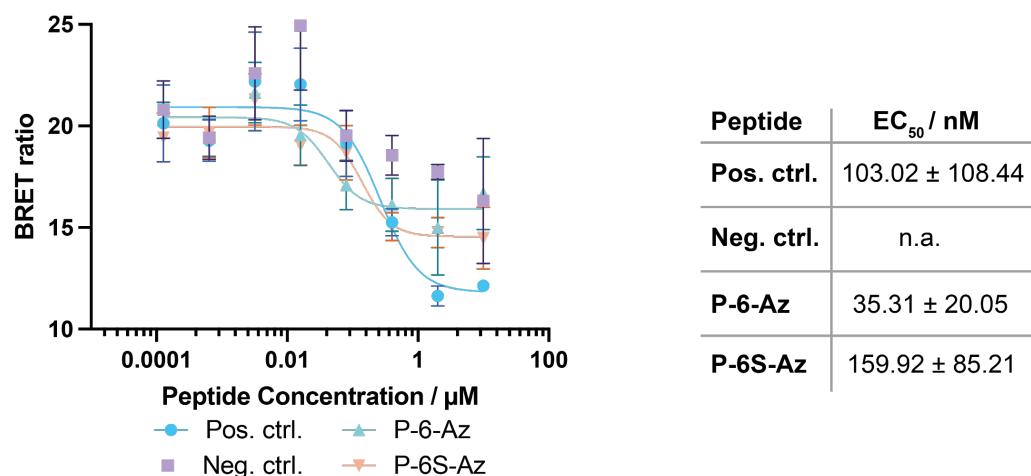
Figure 37: NanoClick cell permeability assay with **P-6Az** and the **P-6** derived scrambled peptide **P-6S-Az**. Both show a high cell permeability exceeding the positive control.

Our determined activities of the peptides in the NanoClick assay are well in conformity with the reported activities of the positive and negative control. To our surprise, peptides **P-6-Az** is significantly surpassing the permeability of the reported positive control **octa-arginine-Az** with 35.31 nM compared to the control with 103 nM. Our scrambled peptide **P-6S-Az** also seems to be slightly less permeable than the positive control, but exceeds the negative control. In contrast to the confocal microscopy results, **P-6-Az** has a high cytoplasmic abundance, which we did not expect to this extend. The differences in results could be explained by either the FITC label from the confocal microscopy probe hindering cellular uptake, or the difference in treatment time. Peptides passing the cell membrane in the NanoClick assay are consumed through the click-reaction after entering the cytoplasm and accumulate in the cell over the duration of the treatment time. By this, any active or passive transport could be driven towards cellular uptake.

Nevertheless, a treatment of 24 h for the cell permeability represents the longer incubation times of *in cellulo* experiments more, than the short 1 h treatment in the confocal microscopy experiments. In conclusion, we could observe a modest cellular uptake by fluorescence microscopy using FITC labeled probes derived from **P-6** and **P-2** and we could further prove, that **P-6** can enter the cytoplasm in treatment-relevant time periods through the NanoClick assay using azido-lysine derivatives.

Next, we were interested in the stability and toxicity of **P-6** and **P-2**. Those two parameters are critical parameters of drug candidates because any potential cellular effects could be highly limited by instable or toxic compounds. Stability is troublesome for peptide derived candidates, as cellular proteases can metabolize and thus inactivate the molecules. Therefore, we compared the stability of **P-6** and **P-2** with the linear counterpart **P-1** in HEK293T cell lysate via LC-MS (**Figure 38A**).
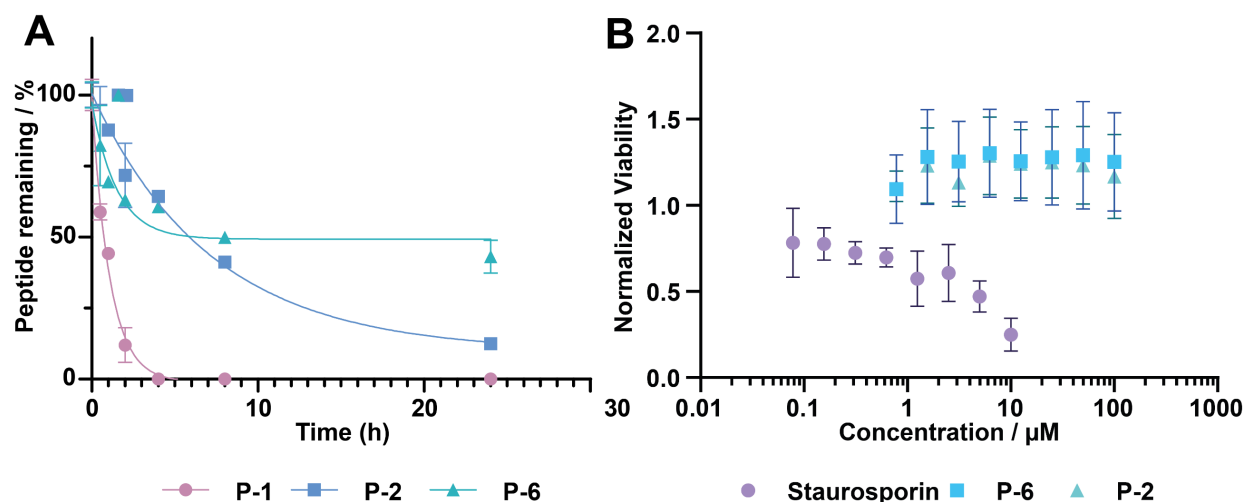
Figure 38: (**A**) Stability of **P-1**, **P-2**, and **P-6** in HEK293T lysate determined by LC-MS. Remaining peptide was calculated through normalization to an internal ethylparabene standard (**B**) Cell viability of HEK293T cells treated with staurosporin, **P-2**, and **P-6** after 48 h determined with the CellTiterGlo 2.0 assay.

Like expected, the linear peptide **P-1** was fully digested after 4 hours, while the monocyclic peptide **P-2** was still 12% intact after 24 hours. The stitched peptide **P-6** was significantly more stable than the other two with 43% being intact after 24 hours. We expected an increased protease stability induced by the macrocyclization and were confirmed in our hypothesis that the more constrained stitched peptide is stabilized through a long staple spanning nearly the whole peptide. In general, unmodified peptides are prone to degrade very fast in cellular lysates with half-lifes of a few minutes to hours.[197–199] Therefore, we deemed that the stability of **P-6** and **P-2** is sufficient for further cellular experiments, albeit significant amounts of peptides were degraded.

Further we determined if the cell viability of HEK293T cells is influenced by peptide **P-2** and **P-6**. Therefore, we treated HEK293T cells with varying concentrations of **P-2**, **P-6**, and staurosporin (positive control). We measured the cell viability using the reported CellTiterGlo 2.0 assay from Promega and could not observe any influences of the peptides on the cell viability (**Figure 38B**). Therefore, we can exclude toxicity caused by inhibiting PTBP1 or induced by non-specific interactions.

Because the peptides showed reasonable permeability, stability and no toxicity in HEK293T cells, we tested if inhibition of the RRM12 tandem of PTBP1 can lead to detectable changes in the splicing patterns of PTBP1 target RNAs. The protein PTBP2 has an AS spliced cassette exon 10 in a conserved region equivalent to the autoregulated exon 11 of PTBP1 (**Figure 39**).[200,201] The exon inclusion of e10 is negatively regulated by PTBP2 itself and the RNA undergoes NMD if it's not included. Further, the exon skipping is also regulated by PTBP1. It was shown, that PTBP1 depletion increases PTBP2 expression by shifting the equilibrium towards e10 inclusion and the opposite could be caused by a overexpression rescue experiment.[202–204] A similar effect was shown for PTBP2 on PTBP1 exon 11, but with weaker effects.[52]
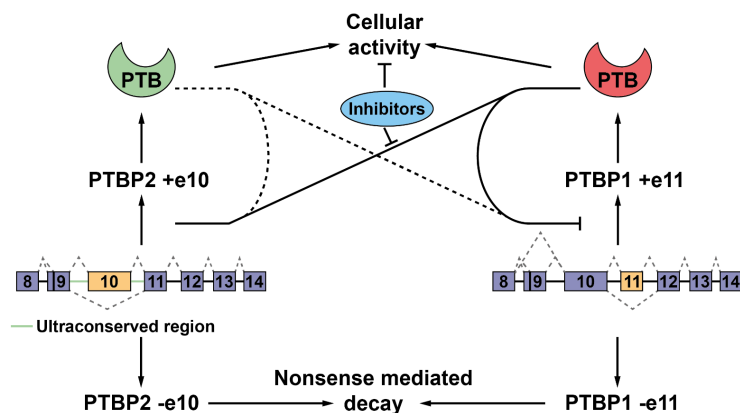
Figure 39: Post-transcriptional regulation of PTBP1 and PTBP2. The inclusion of the autoregulated exons e10 (PTBP2) and e11 (PTBP1) leads to NMD of the respective transcript. E10 and e11 are crossregulated by the paralog counterparts respectively. This effect is most evident in the e10 inclusion of PTBP2 regulated by PTBP1. PTBP1 inhibitors could be able to de-regulate this feedback loop and could lead to increased exon 10 inclusion in PTBP2. Figure adapted from [52].

We chose to perform a reverse transcription (RT) based splicing assay in which total RNA of treated cells was isolated, reverse transcribed, and the cDNA of this known PTBP1 target pre-mRNA was amplified using specific primers for the detection of exon inclusions. The primers are designed to specifically bind before and after an exon, and thus PCR amplification leads to differently sized products of an exon inclusion or exclusion event. We chose to validate the mode

of action of **P-6** on the exon 10 inclusion of the paralog PTBP2, which is one of the best studied targets of PTBP1 (**Figure 40**). First, we validated the PTBP1 dependent regulation of PTBP2 exon 10 and the possibility to detect the splicing event specifically using the chosen RT assay. Therefore, we transfected HEK293T cells with siRNA against PTBP1, isolated the total RNA, reverse transcribed it into cDNA and performed PCR reactions specific for this transcript. We analyzed the PCR products using agarose gel electrophoresis and could detect a full exon 10 inclusion induced by the PTBP1 knockdown (**Figure 40A**). Western blot analysis of PTBP1 demonstrates knockdown by detecting decreasing amounts of PTBP1 in the whole cell lysate, although 100% reduction was not reached (**Figure 40B**). Then, we treated cells with 100 and 300 µM **P-6** and **P-6S**. We chose to use the scrambled **P-6S** as our negative control to exclude any effects, which could be caused by the high peptide concentrations in our treatments. We observed a significant 5% increase of exon 10 inclusion induced by 100 µM **P-6** after 24 h. As the effect was similar in a treatment with 300 µM we expect that we already reached saturation. While we could observe a change for **P-6**, our negative control **P-6S** did not have any significant changes compared to the DMSO control, indicating that the effect is caused by specific inhibition of PTBP1.
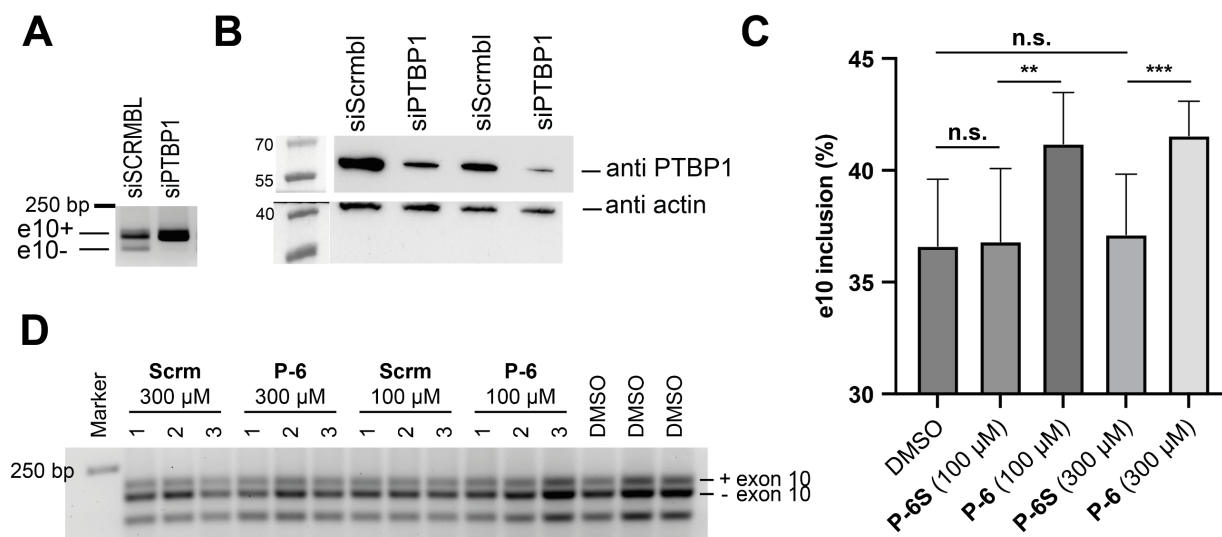
Figure 40: Reverse transcription PCR assay of PTBP2 exon 10 inclusion. (**A**) Exon 10 inclusion of PTBP2 increases after knockdown of PTBP1 with siRNA. (**B**) Western blot analysis of PTBP1 siRNA knockdown. (**C**) Quantification of exon 10 inclusion after treatment with **P-6S** (scramble control) and **P-6** shows increased exon 10 inclusion after treating HEK293T cells for 24 h. (**D**) Representative 2 % agarose gel stained with ethidium bromide of the RT assay quantified in (**C**). **: $p = 0.0056$, ***: $p = 0.0008$, n.s.: not significant, by students t-test.

The modest change of splicing could be caused by several factors. Mainly, this could be caused by the incomplete inhibition of PTBP1, as RRM3 and RRM4 are still active components and could

rescue the missing activity of the RRM12 tandem. On the other hand, the investigated splicing event could not be dependent on the RRM12 to that extend, as previous reports indicated certain specificities of AS events for the individual RRMs.

Further, we tested other PTBP1 target pre-mRNAs which were reported to be regulated by PTBP1.[205] Again, we first compared the knockdown of PTBP1 and could identify significant splicing changes in GPC2 and Rod1 (PTBP3). Further smaller changes were detectable for ANXA8, TNIK, and ZNF711 (**Figure 41A**).
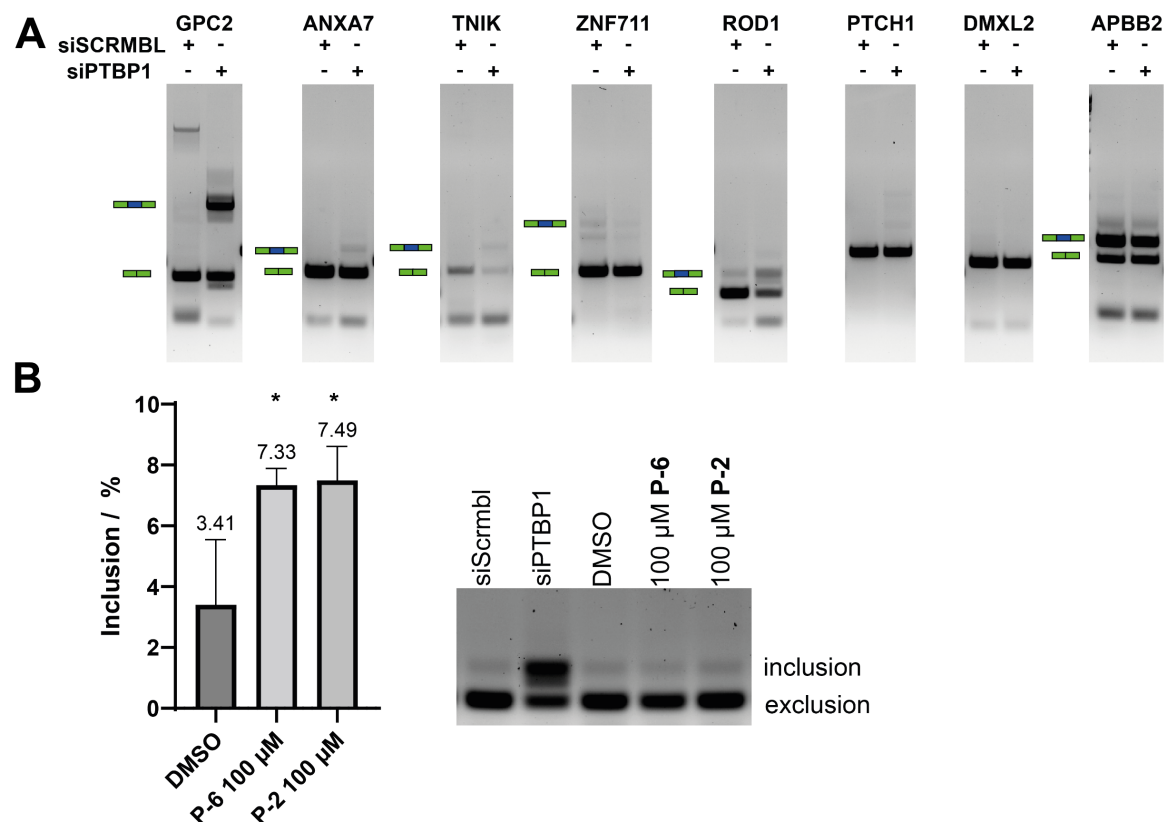
Figure 41: (**A**) Regulation of splicing of reported PTBP1 mediated exon inclusions validated by knockdown of PTBP1 using siRNA compared to a scramble RNA. (**B**) RT-PCR analysis of Rod1 exon2 inclusion (from **A**) after 48h treatment at 100 µM with **P-6** and **P-2**. * **P-6:** p = 0.00371, * **P-2**: p = 0.0428; unpaired students t-test.

After treatment with **P-2**, and **P-6** we could only identify changes in the Rod1 exon2 splicing after 48 h (**Figure 41B**). This, and the modest changes identified for PTBP2 e10 and Rod1 e2 indicate, that not all reported PTBP1 mediated splicing changes are dependent on RRM12, but rather could be dependent on RRM34. Further, inhibition of individual domains could be rescued through the other domains. Exemplarily, it was shown that deletion of parts of the linker between RRM2 and 3 can influence the splicing on several targets, while others were independent of this.[205]

B.5: CONCLUSIONS AND PERSPECTIVES

In summary, we were able to investigate the role of the two RRMs 1 and 2 from PTBP1 for binding to a bi-valent RNA containing two PTBP1 consensus sites. We found that RRM1 and RRM2 bind a longer RNA through both domains and could identify an allosteric or cooperative effect which influences the affinity positively. We further proved that a helix breaking mutation in the α3-helix lowers the affinity of RRM1 and the RRM12 tandem, indicating that the allosteric effect is driven by this transient helix. We further generated inactive mutants RRM1*2 and RRM12* to show that high affinity is reached in presence of both active domains. Based on these results, we synthesized stapled peptides, which bind to domains RRM1 and the tandem RRM12, but also to the different L151G mutants with higher affinity. Further on, we were satisfied to see that the RRM12-RNA interaction is inhibited after treatment with our stapled peptides. We validated the binding mode of our strongest inhibitor through X-ray crystallography and could not determine any conformational changes of RRM1. We hypothesize that the α3-helix forms after or during RNA binding of RRM1 and mediates intra-domain contacts with RRM2 which increases the affinity to the RNA by pre-organization of the complex.

Further on, we evaluated the cellular activity of **P-6** by determining a modest cell permeability, lysate stability and low toxicity. Driven by those results, we performed RT-PCR assays to analyze the alternative splicing of two well-known splicing targets of PTBP1: PTBP2 and Rod1(PTBP3). We were able to measure mild changes in the AS of both, which is on par with literature reports.

Next, the stapled peptides could be used to generate compounds which inactivate full length PTBP1. For example, they could be modified into proteolysis targeting chimeras (PROTACs). The low effect of **P-6** in cells might be caused by the inhibition of only 50% of the domains of PTBP1, while the rest of the protein stays active and might rescue the activity. Through degradation of PTBP1 with PROTACs, full inhibition could be achieved.

Further, the low micromolar peptides could be used to screen for small molecules, which compete them from RRM1. By this, small molecules with the same mode of action could be identified, and a wider chemical space would be available.

More active inhibitors of PTBP1 could be used to decipher the role of the RRM12 tandem in cellular context. Inhibitors with sufficient specificity could be used as probes to inactivate the RNA binding function without disturbing the protein-protein interactions of the RRM12 tandem. Those probes would be better tools for understanding the role of individual domains in splicing regulation. Recently discussed is the possibility of glia-to-neuron conversion and a one-step conversion of

astrocytes to functional neurons caused by PTBP1 depletion.[206,207] This conversion, which was claimed to reverse the Parkinson's disease phenotype in mice is highly debated because of possible artifacts from the used genetic methods and the methods for detection of conversion.[208,209] Specific inhibitors with reasonable potencies or PROTACs with high activities could be useful tools bypassing the needs of genetic methods for PTBP1 depletion in cellular or *in vivo* disease models. Our stapled peptide could contribute to this by acting as a precursor to more active compounds.

Conceptionally, this work reports a novel strategy for the inhibition of RNA-binding proteins with several RRMs by using macrocyclic peptides as mimics of transient secondary structure elements to allosterically inhibit RNA binding. Most previously reported inhibitors of RRMs either have unknown mechanisms, compete directly with RNA, or manipulate the oligomerization of the targets. Allosteric inhibitors of Musashi and NONO, which differ a lot in the mode of action through binding to a hydrophobic pocket for lipids or through covalently binding a cysteine, were identified in screening campaigns. Our strategy could be adapted for other RRM containing proteins, which hold special secondary structure features, especially those where thesesecondary structure elements are transient.

PART C: OTHER RNA BINDING PROTEINS

C.1: WDR5

Introduction

A significant population of the cellular RNA does not code for any proteins and thus is called non-coding RNA (ncRNA). Part of those are the long non-coding RNAs (lncRNAs), which are longer than 200 nucleotides and reported to be without a protein-coding function.[206] Many lncRNAs have been identified to have regulatory roles in epigenetic processes at specific genomic locations.[207–209] They often regulate histone modifications for activation or repression of gene transcription. Until now, several studies reported the therapeutic chances for disease treatment by targeting lncRNA-protein interactions. But only very few molecules were reported to do so, mainly focusing on the EZH2-lncRNA interactions.[210,211]

This project focused on the scaffold protein WDR5, part of the WRAD-complex with RbBP5, ASH2l, and DPY30. This complex is responsible for the recruitment of the SET1/MLL methyltransferase for histone 3 methylation at K4.[212–214] The overactivity of the SET1/MLL complex is linked to several cancers and was previously validated as an excellent therapeutic target.[215,216]

The structure of WDR5 has a conserved WD40 fold, which commonly scaffolds protein complexes with two PPI binding sites.[217,218] The two binding sites are in the "top" and "bottom" of a central cavity in the "donut-shaped" protein. From both, the WBM site binds to RbBP5, c-Myc, and KANSL, and more interesting to several lncRNAs. LncRNA binding was further shown to drive the activity of the MLL complex at specific genomic loci.[219–222] One well-studied, cancer-driving lncRNA binding to the WBM site is HOTTIP.[219,223]

Many inhibitors for the WIN site were reported in the past, but the number of small molecules binding to the WBM site is very limited.[224–228] We could not find reports about WBM ligands affecting the WDR5-lncRNA interaction.

This chapter will focus on the purification of the protein WDR5 and crystallographic analysis of a ligand **P-WDR5-1** synthesized by Jen-Yao Chang (Chemical Genomics Centre III, Group of Dr. Peter 't Hart) binding to the WBM site.

Purification

The cDNA of human WDR5(22-334) was subcloned into a pET19 vector. The protein was heterologously expressed in *E. coli* BL21 (DE3) RIPL and purified using a HisTrap HP 5 ml column (**Figure 42A**). The protein of interest eluted in high purity from the affinity column, according to the very prominent and single band at ~37 kDa. Subsequently, the protein was dialyzed in the presence of 3C-prescission protease to cleave the *C*-terminal His-tag, concentrated, and further purified using a Superdex 75 16/60 column. The isolated product was >90% pure with yields of up to 30 mg/liter expression. The concentrated protein was flash-frozen in liquid nitrogen and stored at -80 °C until use in binding experiments or crystallography.

Similarly, a FLAG-tagged construct was expressed and purified for RIP experiments. The POI was less pure after the affinity chromatography (**Figure 42C**) but was sufficiently pure after a reverse IMAC and following gel filtration (**Figure 42**).
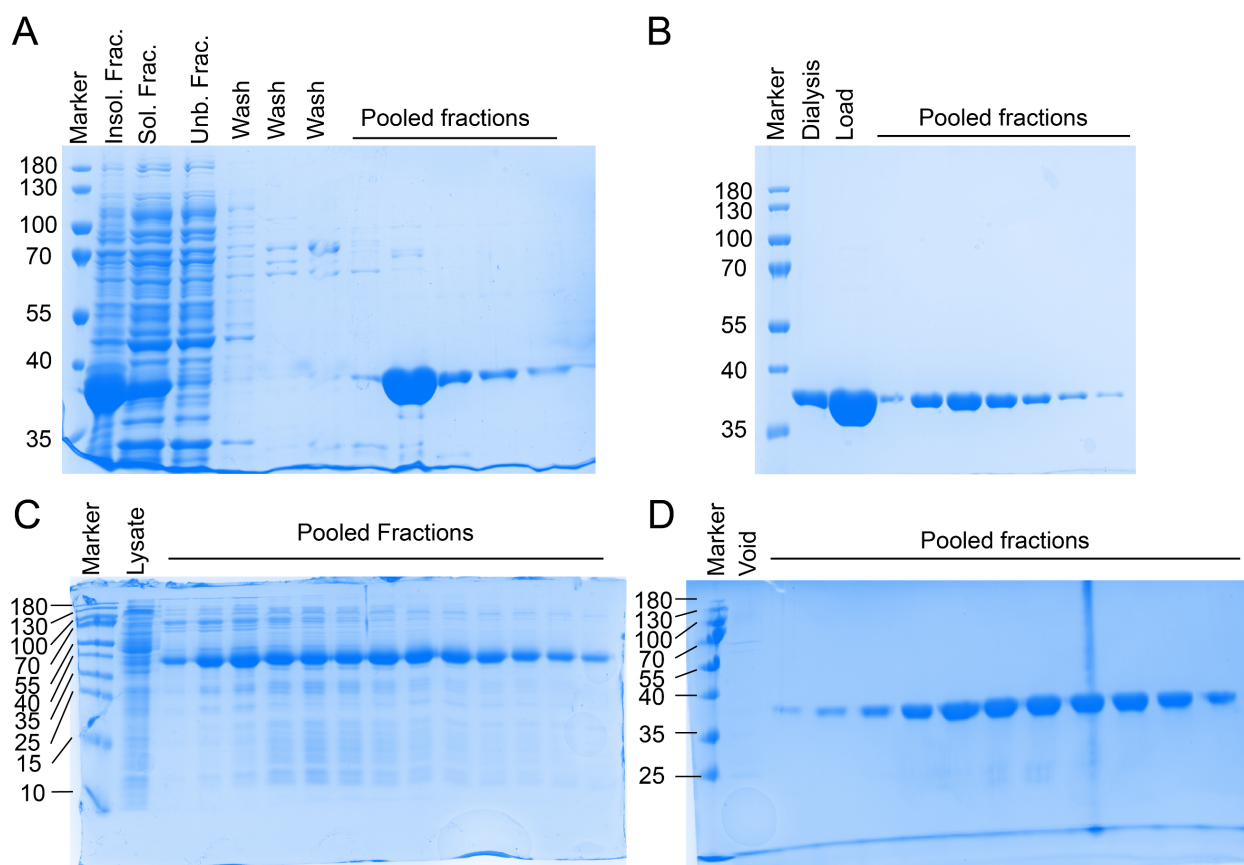
Figure 42: Purification of WDR5 constructs. (**A**) 12% SDS-PAGE of a Ni-NTA column of WDR5-His and (**B**) Size exclusion chromatography using a Superdex 75 16/60 column of tag-cleaved WDR5. (**C/D**) Ni-NTA and Superdex 75 16/60 of His-tagged WDR5, which got *N*-terminally FLAG-tagged for pull-down and RIP experiments.

Co-Crystallization

To validate the binding pose of the synthesized peptides, we expressed and purified WDR5 with the previously mentioned method. We co-crystallized a protein-peptide complex by adding a 1.5-fold molar excess of the peptide to the concentrated protein. Initially, we observed many crystals in different conditions but got big, cubic crystals in only one condition. One measured crystal with the space group P 4_3 2_1 2 diffracted to 1.8 Å. We were able to solve the crystal using an AlphaFold model and Phaser (Phenix). We identified two protein-peptide dimers in the asymmetric unit (**Table 9**; **Figure 43A**). We observed a dense packing of monomers when depicting the symmetry-related neighbors at a 20 Å distance (**Figure 44**) indicating a correct model as no significant gaps could be found. To our surprise, we recognized a significant discrepancy in the quality of the $2F_O$-F_C map between the two WDR5 chains, probably caused

Table 9: Data collection and refinement data for the reported WDR5 crystal.

	WDR5:WDR5-P-1
Data collection	
Space group	P 4_3 2_1 2
Cell dimensions	
a, b, c (Å)	82.1737 82.1737 201.707
α, β, γ (°)	90, 90 ,90
Resolution (Å)	42.98 - 1.843 (1.909 - 1.843)
R_{merge}	0.1688 (3.632)
I / sI	8.05 (0.26)
Completeness (%)	96.59 (66.37)
Redundancy (Multiplicity)	26.7 (27.4)
CC1/2	0.995 (0.361)
Refinement	
Resolution (Å)	1.84
No. Reflections	58517 (3960)
R_{work} / R_{free}	0.1975 / 0.2429
No. Atoms	5614
Protein	5179
Ligand/ion	0
Water	435
B-factors	
Protein	56.22
Water	57.80
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.25
Ramachandran favored (%)	95.19
Ramachandran allowed (%)	4.65
Ramachandran outliers (%)	0.16
Rotamer outliers (%)	2.36
Clashscore	8.32

by a higher disorder and flexibility in chain B, which corresponds to the increased B-factors of this dimer (**Figure 43B**). Both chains share the typical fold of seven beta-propeller blades with four-stranded antiparallel sheets.[229] Overall, the two WDR5 chains still have a high similarity with an RMSD of 0.285 Å (**Figure 43C**). The most significant differences between both chains were observed in the *N*- and *C*-terminal regions, where we observed a higher ordered β-sheet conformation for chain B, while the loops of chain A point away from each other, which could be caused by additional contacts to symmetry-related molecules. The regions with the high

discrepancy have high B-Factors of ~150, while the average B-factor of the structure is 56, caused by the increased flexibility of the termini, which is also represented in a bad $2F_O$-F_C density.
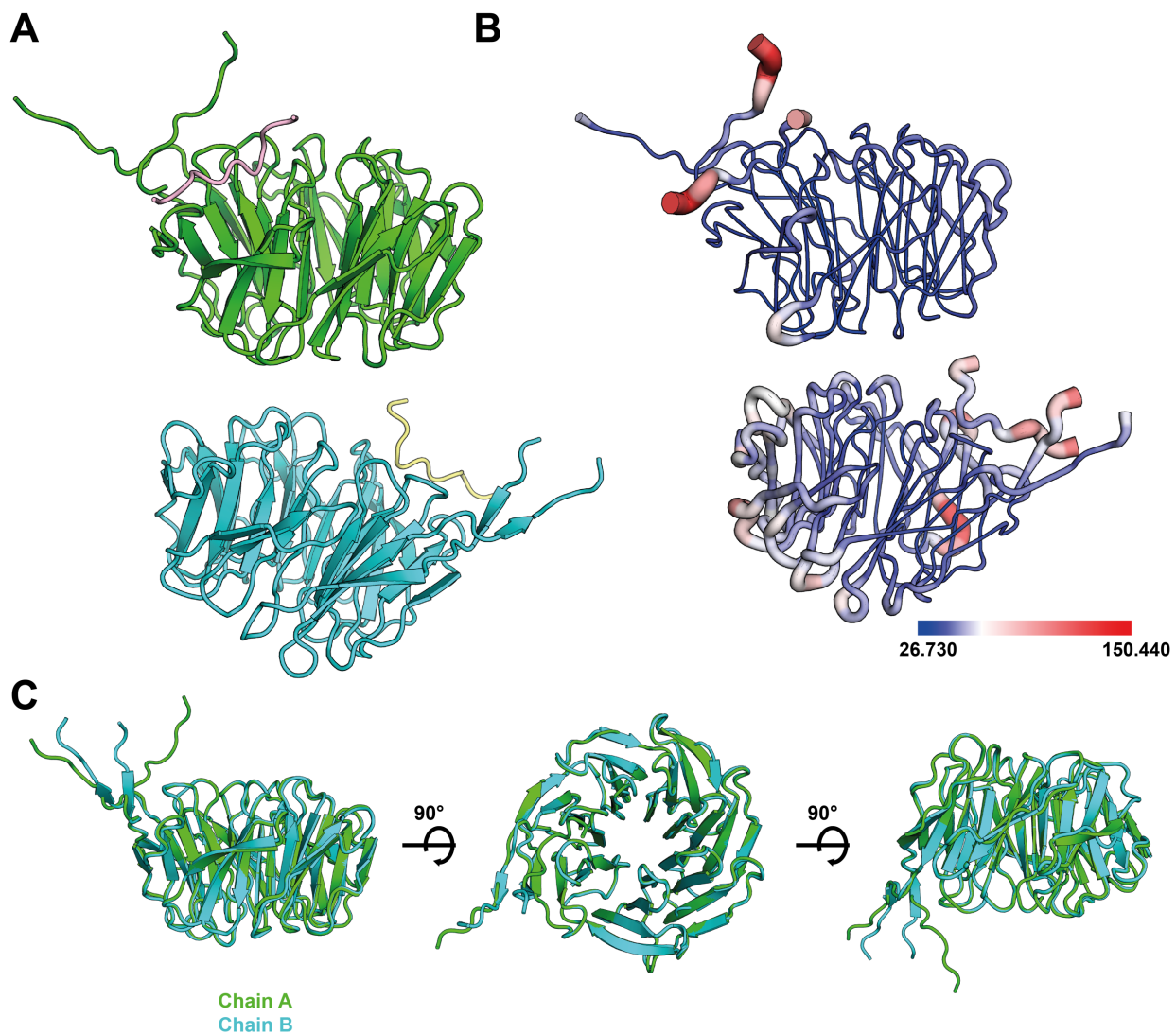
Figure 43: (**A**) Asymmetric unit (ASU) of the WDR5:**P-WDR5** co-crystal containing two protein-peptide dimers. (**B**) Depiction of the B-factors of both dimers A/a (top) and B/b (bottom) shows that the termini of the peptide ligands and chain B have higher B-factors than chain A, corresponding to the quality of electron density. (**C**) Overlay of the protein chains A and B indicate a high similarity of both chains except the termini, which point away from each other.

A well-resolved density of the central amino acids of **P-WDR5**, including the position of macrocyclization, verified the presence of ligand **P-WDR5**. But the density of the terminal Glu and Asp was less well resolved, indicating higher flexibility of the terminal amino acids, which was also observed by previously reported structures of linear peptides bound to the WBM site (**Figure 45C/D**). Further, ligands a and b bind identically to the reported native peptide of Rbbp5 (**Figure 45A/B**).

Figure 44: Symmetry mates in 20 Å distance show that the crystal packing of the structure is high, as expected.

The binding of **P-WDR5** to the WBM site does not lead to any significant changes in the conformation of this site and the protein compared to the reported apo-structure (PDB 2H14) (**Figure 45E**). Although no significant conformational changes were identified, the peptide fits into the WBM pocket and binds over several residues to K250, K227, N225, L249, and Q289. Furthermore, we observed numerous structured waters in the structure, most of which were located in the central cavity of the protein, which was also observed in previous reports (**Figure 45F**).[230] To our surprise, we recognized a potential disulfide bridge of chains A and B with a symmetry-related molecule of chains B and A, respectively. The *C*-terminally located C334 is located within a 2.0 Å distance of the symmetry-related chain, and the density indicates a possible bond.

Figure 45: (**A/B**) The WBM binding site with P-WDR5 bound overlays well with a peptide with a similar binding mode of Rbbp5 (PDB 2XL2) in the core regions of the peptide. (**C/D**) $2F_O$-F_C-maps of ligands a and b indicate the correct posing of the ligand in the WBM site. (**E**) Compared to a published apo-structure, the ligand P-WDR5 binds to the WBM interface without obstructing the apo-fold (PDB 2H14). (**F**) Like in previous structures, the channel in the WD-fold is filled with many structured waters.

95

ABBREVIATIONS

snRNA	Small nuclear RNA
snRNP	Small nuclar ribonucleoprotein particle
ESE	Exonic splicing enhancer
SR	Serine-arginine-rich
hnRNP	Heterogeneous nuclear ribonucleoprotein
ESS	Exonic splicing silencer
AS	Alternative splicing
SRE	Splicing regulatory elements
RRM	RNA recognition motif
KH	K homology domain
RBD	RNA binding domain
RBP	RNA binding protein
PTBP	Polypyrimidine tract-binding protein
SAXS	small-angle X-ray scattering
PRI	PTBP1-RRM2 interaction peptides
TM	tropomyosin
NMD	nonsense mediated decay
nPTB	Neuronal PTB / PTBP2
UTR	Untranslated region
ASO	antisense oligonucleotide
SSO	splice-switching oligonucleotide
ARE	AU-rich element
HTS	High throughput screening
PPI	Protein-protein interaction
PTM	Post-translational modifcation
PG	Protecting group
Fmoc	9-fluorenylmethyloxycarbonyl
RCM	ring-closing metathesis
PCR	Polymerase chain reaction
SLIC	Sequence and ligation independent cloning
CPEC	circular polymerase extension cloning
Amp	Ampicillin
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TRAP	Translational Repression Assay Procedure
RBS	Ribosomal binding site
ORI	origin of replication
KanR	Kanamycin resistance
CamR	Chloramphenicol resistance
AmpR	Ampicillin resistance
SICLOPPS	split-intein circular ligation of peptides and proteins
IMAC	immobilized metal affinity chromatography
FP	Fluorescence polarization
S_5	(*S*)-pentenyl alanine
CPP	cell penetrating peptide
RT	Reverse transcription

DIBAC	dibenzoazacyclooctyne
BRET	bioluminescence resonance energy transfer
ncRNA	Non-coding RNA
lncRNA	long non-coding RNA

REFERENCES

1.   Bush, S. J., Chen, L., Tovar-Corona, J. M. & Urrutia, A. O. Alternative splicing and the evolution of phenotypic novelty. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 1–7 (2017).

2.   Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457–463 (2010).

3.   Hahn, M. W. & Wray, G. A. The g-value paradox. *Evol. Dev.* **4**, 73–75 (2002).

4.   Berget, S. M. Exon Recognition in Vertebrate Splicing. *J. Biol. Chem.* **270**, 2411–2414 (1995).

5.   Long, M. & Deutsch, M. Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. *Mol. Biol. Evol.* **16**, 1528–1534 (1999).

6.   Behzadnia, N. *et al.* Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *EMBO J.* **26**, 1737–1748 (2007).

7.   Kelemen, O. *et al.* Function of alternative splicing. *Gene* **514**, 1–30 (2013).

8.   Stamm, S., Smith, C. & Lührmann, R. *Alternative pre-mRNA Splicing*. (Wiley-VCH Verlag GmbH & Co. KGaA, 2012). doi:10.1002/9783527636778.

9.   Abelson, J. Is the spliceosome a ribonucleoprotein enzyme? *Nat. Struct. Mol. Biol.* **15**, 1235–1237 (2008).

10.   Pyle, A. M. Translocation and Unwinding Mechanisms of RNA and DNA Helicases. *Annu. Rev. Biophys.* **37**, 317–336 (2008).

11.   Will, C. L. & Luhrmann, R. Spliceosome Structure and Function. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).

12.   Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136**, 701–718 (2009).

13.   Singh, R. & Valcárcel, J. Building specificity with nonspecific RNA-binding proteins. *Nat. Struct. Mol. Biol.* **12**, 645–653 (2005).

14.   Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).

15.   Schmucker, D. *et al.* Drosophila Dscam Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity. *Cell* **101**, 671–684 (2000).

16.   Hattori, D. *et al.* Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. *Nature* **461**, 644–648 (2009).

17.   Ast, G. How did alternative splicing evolve? *Nat. Rev. Genet.* **5**, 773–782 (2004).

18.   Nogués, G., Muñoz, M. J. & Kornblihtt, A. R. Influence of Polymerase II Processivity on Alternative Splicing Depends on Splice Site Strength. *J. Biol. Chem.* **278**, 52166–52171 (2003).

19.   Smith, C. W. J. & Nadal-Ginard, B. Mutually exclusive splicing of α-tropomyosin exons enforced by an unusual lariat branch point location: Implications for constitutive splicing. *Cell* **56**, 749–758 (1989).

20.   Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

21.   Dominski, Z. & Kole, R. Selection of splice sites in pre-mRNAs with short internal exons. *Mol. Cell. Biol.* **11**, 6075–6083 (1991).

22.   Goren, A. *et al.* Comparative Analysis Identifies Exonic Splicing Regulatory Sequences—The Complex Definition of Enhancers and Silencers. *Mol. Cell* **22**, 769–781 (2006).

23.   Buratti, E. & Baralle, F. E. Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. *Mol. Cell. Biol.* **24**, 10505–10514 (2004).

24.   KORNBLIHTT, A. R. Multiple links between transcription and splicing. *RNA* **10**, 1489–1498 (2004).

25.   Alló, M. *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **16**, 717–724 (2009).

26.   Luco, R. F. *et al.* Regulation of Alternative Splicing by Histone Modifications. *Science* **327**,

996–1000 (2010).

27.   Sammeth, M., Foissac, S. & Guigó, R. A general definition and nomenclature for alternative splicing events. *PLoS Comput. Biol.* **4**, (2008).

28.   Black, D. L. Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* **72**, 291–336 (2003).

29.   doi:10.1126/science.1058040.

30.   Cléry, A., Blatter, M. & Allain, F. H. T. RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298 (2008).

31.   Maris, C. & Dominguez, C. The RNA recognition motif , a plastic RNA-binding platform to regulate post-transcriptional gene expression. **272**, 2118–2131 (2005).

32.   Muto, Y. & Yokoyama, S. Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems: Structural insight into RNA recognition motifs. *Wiley Interdiscip. Rev. RNA* **3**, 229–246 (2012).

33.   Bourgeois, C. F., Lejeune, F. & Stévenin, J. Broad Specificity of SR (Serine ⁄ Arginine) Proteins in the Regulation of Alternative Splicing of Pre-Messenger RNA. in *Progress in Nucleic Acid Research and Molecular Biology* vol. 78 37–88 (Academic Press, 2004).

34.   Auweter, S. D. *et al.* Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J.* **25**, 163–173 (2006).

35.   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).

36.   Afroz, T., Cienikova, Z., Cléry, A. & Allain, F. H. T. One, Two, Three, Four! How Multiple RRMs Read the Genome Sequence. in *Methods in Enzymology* vol. 558 235–278 (Elsevier, 2015).

37.   Afroz, T. *et al.* A fly trap mechanism provides sequence-specific RNA recognition by CPEB proteins. *Genes Dev.* **28**, 1498–1514 (2014).

38.   Jenkins, H. T., Malkova, B. & Edwards, T. A. Kinked β-strands mediate high-affinity recognition of mRNA targets by the germ-cell regulator DAZL. *Proc. Natl. Acad. Sci.* **108**, 18266–18271 (2011).

39.   Cieniková, Z., Damberger, F. F., Hall, J., Allain, F. H.-T. & Maris, C. Structural and Mechanistic Insights into Poly(uridine) Tract Recognition by the hnRNP C RNA Recognition Motif. *J. Am. Chem. Soc.* **136**, 14536–14544 (2014).

40.   Sickmier, E. A. *et al.* Structural Basis for Polypyrimidine Tract Recognition by the Essential Pre-mRNA Splicing Factor U2AF65. *Mol. Cell* **23**, 49–59 (2006).

41.   Wang, X. & Tanaka Hall, T. M. Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat. Struct. Biol.* **8**, 141–145 (2001).

42.   Teplova, M., Song, J., Gaw, H. Y., Teplov, A. & Patel, D. J. Structural Insights into RNA Recognition by the Alternate-Splicing Regulator CUG-Binding Protein 1. *Structure* **18**, 1364–1377 (2010).

43.   Wang, H. *et al.* The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 373–380 (2013).

44.   Maris, C., Jayne, S., Damberger, F. F., Beusch, I. & Allain, H. A transient alpha-helix in the N-terminal RNA recognition motif of polypyrimidine tract binding protein senses RNA secondary structure. *Nucleic Acids Res* 1–17 (2020) doi:10.1093/nar/gkaa155.

45.   Dorn, G. *et al.* Integrative solution structure of a PTBP1-viral IRES complex reveals strong compaction and ordering with residual conformational flexibility. *bioRxiv* 1–53 (2022).

46.   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).

47.   Duszczyk, M. M. *et al.* The solution structure of Dead End bound to AU-rich RNA reveals an unusual mode of tandem RRM-RNA recognition required for mRNA regulation. *Nat. Commun.* **13**, 5892 (2022).

48.   Hobor, F. *et al.* A cryptic RNA-binding domain mediates Syncrip recognition and exosomal

partitioning of miRNA targets. *Nat. Commun.* **9**, 831 (2018).

49.   Chen, Y. *et al.* SYNCRIP, a new player in pri-let-7a processing. *RNA* **26**, 290–305 (2020).

50.   Crowder, S. M., Kanaar, R., Rio, D. C. & Alber, T. Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. *Proc. Natl. Acad. Sci.* **96**, 4892–4897 (1999).

51.   Kafasla, P. *et al.* Defining the roles and interactions of PTB. *Biochem. Soc. Trans.* **40**, 815–820 (2012).

52.   Keppetipola, N., Sharma, S., Li, Q. & Black, D. L. Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit. Rev. Biochem. Mol. Biol.* **47**, 360–378 (2012).

53.   Licatalosi, D. D. *et al.* Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. *Genes Dev.* **26**, 1626–1642 (2012).

54.   Li, Q. *et al.* The splicing regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. *eLife* **3**, (2014).

55.   Oberstrass, F. C. *et al.* Structure of PTB bound to RNA: Specific binding and implications for splicing regulation. *Science* **309**, 2054–2057 (2005).

56.   Clerte, C. & Hall, K. B. The domains of polypyrimidine tract binding protein have distinct RNA structural preferences. *Biochemistry* **48**, 2063–2074 (2009).

57.   Simpson, P. J. *et al.* Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure* **12**, 1631–1643 (2004).

58.   Coelho, M. B. *et al.* Functional interactions between polypyrimidine tract binding protein and PRI peptide ligand containing proteins. *Biochem. Soc. Trans.* **44**, 1058–1065 (2016).

59.   Gromak, N. *et al.* The PTB interacting protein raver1 regulates α-tropomyosin alternative splicing. *EMBO J.* **22**, 6356–6364 (2003).

60.   Vitali, F. *et al.* Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. *EMBO J.* **25**, 150–162 (2006).

61.   Pérez, I., McAfee, J. G. & Patton, J. G. Multiple RRMs contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein. *Biochemistry* **36**, 11881–11890 (1997).

62.   Dai, S. *et al.* PTB: Not just a polypyrimidine tract-binding protein. *J. Cell. Physiol.* 1–17 (2022) doi:10.1002/jcp.30716.

63.   Saulière, J., Sureau, A., Expert-Bezançon, A. & Marie, J. The Polypyrimidine Tract Binding Protein (PTB) Represses Splicing of Exon 6B from the β-Tropomyosin Pre-mRNA by Directly Interfering with the Binding of the U2AF65 Subunit. *Mol. Cell. Biol.* **26**, 8755–8769 (2006).

64.   Izquierdo, J. M. *et al.* Regulation of Fas Alternative Splicing by Antagonistic Effects of TIA-1 and PTB on Exon Definition. *Mol. Cell* **19**, 475–484 (2005).

65.   Sharma, S., Kohlstaedt, L. A., Damianov, A., Rio, D. C. & Black, D. L. Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nat. Struct. Mol. Biol.* **15**, 183–191 (2008).

66.   Sharma, S., Falick, A. M. & Black, D. L. Polypyrimidine Tract Binding Protein Blocks the 5′ Splice Site-Dependent Assembly of U2AF and the Prespliceosomal E Complex. *Mol. Cell* **19**, 485–496 (2005).

67.   Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. J. Autoregulation of Polypyrimidine Tract Binding Protein by Alternative Splicing Leading to Nonsense-Mediated Decay. *Mol. Cell* **13**, 91–100 (2004).

68.   Southby, J., Gooding, C. & Smith, C. W. J. Polypyrimidine Tract Binding Protein Functions as a Repressor To Regulate Alternative Splicing of α-Actinin Mutally Exclusive Exons. *Mol. Cell. Biol.* **19**, 2699–2711 (1999).

69.   Lou, H., Gagel, R. F. & Berget, S. M. An intron enhancer recognized by splicing factors activates polyadenylation. *Genes Dev.* **10**, 208–219 (1996).

70.   Ghetti, A., Pinol-Roma, S., Michael, W. M., Morandi, C. & Dreyfuss, G. hnRNP 1, the polyprimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.* **20**, 3671–3678 (1992).

71.   Cote, C. A. *et al.* A Xenopus Protein Related to hnRNP I Has a Role in Cytoplasmic RNA Localization. *Mol. Cell* **4**, 431–437 (1999).

72.   Li, B. & Yen, T. S. B. Characterization of the Nuclear Export Signal of Polypyrimidine Tract-binding Protein. *J. Biol. Chem.* **277**, 10306–10314 (2002).

73.   Fritz, S. E., Ranganathan, S., Wang, C. D. & Hogg, J. R. The RNA-binding protein PTBP1 promotes ATPase-dependent dissociation of the RNA helicase UPF1 to protect transcripts from nonsense-mediated mRNA decay. *J. Biol. Chem.* **295**, jbc.RA120.013824 (2020).

74.   Ge, Z., Quek, B. L., Beemon, K. L. & Hogg, J. R. Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. *eLife* **5**, e11155 (2016).

75.   Jeong, D. E. *et al.* Glucose Controls the Expression of Polypyrimidine Tract-Binding Protein 1 via the Insulin Receptor Signaling Pathway in Pancreatic β Cells. **41**, 909–916 (2018).

76.   Knoch, K.-P. *et al.* Polypyrimidine tract-binding protein promotes insulin secretory granule biogenesis. *Nat. Cell Biol.* **6**, 207–214 (2004).

77.   Tillmar, L., Carlsson, C. & Welsh, N. Control of Insulin mRNA Stability in Rat Pancreatic Islets. *J. Biol. Chem.* **277**, 1099–1106 (2002).

78.   Zhang, Y., Qian, J., Gu, C. & Yang, Y. Alternative splicing and cancer: a systematic review. *Signal Transduct. Target. Ther.* **6**, 78 (2021).

79.   Kahles, A. *et al.* Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* **34**, 211-224.e6 (2018).

80.   Graubert, T. A. *et al.* Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat. Genet.* **44**, 53–57 (2012).

81.   Ferreira, P. G. *et al.* Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res.* **24**, 212–226 (2014).

82.   Jung, H. *et al.* Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat. Genet.* **47**, 1242–1248 (2015).

83.   Brooks, A. N. *et al.* A Pan-Cancer Analysis of Transcriptome Changes Associated with Somatic Mutations in U2AF1 Reveals Commonly Altered Splicing Events. *PLoS ONE* **9**, e87361 (2014).

84.   Furney, S. J. *et al.* SF3B1 Mutations Are Associated with Alternative Splicing in Uveal Melanoma. *Cancer Discov.* **3**, 1122–1129 (2013).

85.   PCAWG Transcriptome Core Group *et al.* Genomic basis for RNA alterations in cancer. *Nature* **578**, 129–136 (2020).

86.   Bechara, E. G., Sebestyén, E., Bernardis, I., Eyras, E. & Valcárcel, J. RBM5, 6, and 10 Differentially Regulate NUMB Alternative Splicing to Control Cancer Cell Proliferation. *Mol. Cell* **52**, 720–733 (2013).

87.   Wang, L. *et al.* Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell* **30**, 750–763 (2016).

88.   Dominguez, D. *et al.* An extensive program of periodic alternative splicing linked to cell cycle progression. *eLife* **5**, e10288 (2016).

89.   Ben-Hur, V. *et al.* S6K1 Alternative Splicing Modulates Its Oncogenic Activity and Regulates mTORC1. *Cell Rep.* **3**, 103–115 (2013).

90.   Pradella, D. *et al.* A ligand-insensitive UNC5B splicing isoform regulates angiogenesis by promoting apoptosis. *Nat. Commun.* **12**, 4872 (2021).

91.   Frampton, G. M. *et al.* Activation of MET via Diverse Exon 14 Splicing Alterations Occurs in Multiple Tumor Types and Confers Clinical Sensitivity to MET Inhibitors. *Cancer Discov.* **5**, 850–859 (2015).

92.   Lee, S. C. W. & Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat. Med.* **22**, 976–986 (2016).

93.   Karni, R. *et al.* The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* **14**, 185–193 (2007).

94.   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).

95.   Jbara, A., Siegfried, Z. & Karni, R. Splice-switching as cancer therapy. *Curr. Opin. Pharmacol.* **59**, 140–148 (2021).

96.   Kaida, D. *et al.* Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* **3**, 576–583 (2007).

97.   Kotake, Y. *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat. Chem. Biol.* **3**, 570–575 (2007).

98.   Eskens, F. A. L. M. *et al.* Phase I Pharmacokinetic and Pharmacodynamic Study of the First-in-Class Spliceosome Inhibitor E7107 in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* **19**, 6296–6304 (2013).

99.   Hong, D. S. *et al.* A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) splicesome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest. New Drugs* **32**, 436–444 (2014).

100.   Steensma, D. P. *et al.* Phase I First-in-Human Dose Escalation Study of the oral SF3B1 modulator H3B-8800 in myeloid neoplasms. *Leukemia* **35**, 3542–3550 (2021).

101.   Bashari, A., Siegfried, Z. & Karni, R. Targeting Splicing Factors for Cancer Therapy. *RNA* rna.079585.123 (2023) doi:10.1261/rna.079585.123.

102.   Kole, R., Krainer, A. R. & Altman, S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* **11**, 125–140 (2012).

103.   Denichenko, P. *et al.* Specific inhibition of splicing factor activity by decoy RNA oligonucleotides. *Nat. Commun.* **10**, (2019).

104.   Chatrikhi, R. *et al. A synthetic small molecule stalls pre-mRNA splicing by promoting an early-stage U2AF2-RNA complex. Cell Chemical Biology* (2021). doi:10.1016/j.chembiol.2021.02.007.

105.   D'Agostino, V. G. *et al.* Dihydrotanshinone-I interferes with the RNA-binding activity of HuR affecting its post-transcriptional function. *Sci. Rep.* **5**, 16478 (2015).

106.   Manzoni, L. *et al.* Interfering with HuR–RNA Interaction: Design, Synthesis and Biological Characterization of Tanshinone Mimics as Novel, Effective HuR Inhibitors. *J. Med. Chem.* **61**, 1483–1498 (2018).

107.   Minuesa, G. *et al.* Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia. *Nat. Commun.* **10**, 2691 (2019).

108.   Cheng, Y.-C. *et al.* MPT0B098, a Novel Microtubule Inhibitor That Destabilizes the Hypoxia-Inducible Factor-1α mRNA through Decreasing Nuclear–Cytoplasmic Translocation of RNA-Binding Protein HuR. *Mol. Cancer Ther.* **12**, 1202–1212 (2013).

109.   Blanco, F. F. *et al.* Impact of HuR inhibition by the small molecule MS-444 on colorectal cancer cell tumorigenesis. *Oncotarget* **7**, 74043–74058 (2016).

110.   François-Moutal, L. *et al.* Small Molecule Targeting TDP-43's RNA Recognition Motifs Reduces Locomotor Defects in a Drosophila Model of Amyotrophic Lateral Sclerosis (ALS). *ACS Chem. Biol.* acschembio.9b00481 (2019) doi:10.1021/acschembio.9b00481.

111.   Sommer, G. *et al.* Applying a high-throughput fluorescence polarization assay for the discovery of chemical probes blocking La:RNA interactions in vitro and in cells. *PLOS ONE* **12**, e0173246 (2017).

112.   Dahlem, C. *et al.* First Small-Molecule Inhibitors Targeting the RNA-Binding Protein IGF2BP2/IMP2 for Cancer Therapy. *ACS Chem. Biol.* **17**, 361–375 (2022).

113.   Clingman, C. C. *et al.* Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *eLife* **2014**, 1–26 (2014).

114.   Lan, L. *et al.* Natural product (−)-gossypol inhibits colon cancer cell growth by targeting RNA-binding protein Musashi-1. *Mol. Oncol.* **9**, 1406–1420 (2015).

115.   Lan, L. *et al.* Identification and Validation of an Aspergillus nidulans Secondary Metabolite Derivative as an Inhibitor of the Musashi-RNA Interaction. *Cancers* **12**, 2221 (2020).

116.   Nshogoza, G. *et al.* NMR fragment-based screening against tandem RNA recognition motifs of TDP-43. *Int. J. Mol. Sci.* **20**, (2019).

117.   Kathman, S. G. *et al.* Remodeling oncogenic transcriptomes by small molecules targeting NONO. *Nat. Chem. Biol.* 1–12 (2023) doi:10.1038/s41589-023-01270-0.

118.   Ali, A. M., Atmaj, J., Van Oosterwijk, N., Groves, M. R. & Dömling, A. Stapled Peptides Inhibitors: A New Window for Target Drug Discovery. *Comput. Struct. Biotechnol. J.* **17**, 263–281 (2019).

119.   Yan, C., Wu, F., Jernigan, R. L., Dobbs, D. & Honavar, V. Characterization of Protein–Protein Interfaces. *Protein J.* **27**, 59–70 (2008).

120.   Verdine, G. L. & Hilinski, G. J. Stapled Peptides for Intracellular Drug Targets. in *Methods in Enzymology* vol. 503 3–33 (Elsevier, 2012).

121.   Pérot, S., Sperandio, O., Miteva, M. A., Camproux, A.-C. & Villoutreix, B. O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discov. Today* **15**, 656–667 (2010).

122.   Ivanov, A. A., Khuri, F. R. & Fu, H. Targeting protein–protein interactions as an anticancer strategy. *Trends Pharmacol. Sci.* **34**, 393–400 (2013).

123.   Smith, M. C. & Gestwicki, J. E. Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev. Mol. Med.* **14**, e16 (2012).

124.   Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The Future of Peptide-based Drugs: **Peptides in Drug Development**. *Chem. Biol. Drug Des.* **81**, 136–147 (2013).

125.   Bakail, M. & Ochsenbein, F. Targeting protein–protein interactions, a wide open field for drug design. *Comptes Rendus Chim.* **19**, 19–27 (2016).

126.   Lau, J. L. & Dunn, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg. Med. Chem.* **26**, 2700–2707 (2018).

127.   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).

128.   Jenssen, H. & Aspmo, S. I. Serum Stability of Peptides. in *Peptide-Based Drug Design* (ed. Otvos, L.) vol. 494 177–186 (Humana Press, 2008).

129.   Tsomaia, N. Peptide therapeutics: Targeting the undruggable space. *Eur. J. Med. Chem.* **94**, 459–470 (2015).

130.   Wills, R., Adebomi, V. & Raj, M. Site-Selective Peptide Macrocyclization. *ChemBioChem* **22**, 52–62 (2021).

131.   Jubb, H., Higueruelo, A. P., Winter, A. & Blundell, T. L. Structural biology and drug discovery for protein–protein interactions. *Trends Pharmacol. Sci.* **33**, 241–248 (2012).

132.   Schafmeister, C. E., Po, J. & Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* **122**, 5891–5892 (2000).

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

134.   Blackwell, H. E. & Grubbs, R. H. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angew. Chem. Int. Ed.* **37**, 3281–3284 (1998).

135.   Walensky, L. D. *et al.* Activation of Apoptosis in Vivo by a Hydrocarbon-Stapled BH3 Helix. *Science* **305**, 1466–1470 (2004).

136.   Walensky, L. D. & Bird, G. H. Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress: Miniperspective. *J. Med. Chem.* **57**, 6275–6288 (2014).

137.   Kneissl, S., Loveridge, E. J., Williams, C., Crump, M. P. & Allemann, R. K. Photocontrollable Peptide-Based Switches Target the Anti-Apoptotic Protein Bcl-x $_L$. *ChemBioChem* **9**, 3046–3054 (2008).

138.   Hilinski, G. J. *et al.* Stitched α-Helical Peptides via Bis Ring-Closing Metathesis. *J. Am. Chem. Soc.* **136**, 12314–12322 (2014).

139.   Felix, A. M. *et al.* Synthesis, biological activity and conformational analysis of cyclic GRF analogs. *Int. J. Pept. Protein Res.* **32**, 441–454 (2009).

140.   Khoo, K. K. *et al.* Lactam-Stabilized Helical Analogues of the Analgesic µ-Conotoxin KIIIA. *J. Med. Chem.* **54**, 7558–7566 (2011).

141.   Scrima, M. *et al.* Cu ^{I} -Catalyzed Azide–Alkyne Intramolecular *i* -to-(*i* +4) Side-Chain-to-Side-Chain Cyclization Promotes the Formation of Helix-Like Secondary Structures. *Eur. J. Org. Chem.* **2010**, 446–457 (2010).

142.   Kawamoto, S. A. *et al.* Design of Triazole-Stapled BCL9 α-Helical Peptides to Target the β-Catenin/B-Cell CLL/lymphoma 9 (BCL9) Protein–Protein Interaction. *J. Med. Chem.* **55**, 1137–1146 (2012).

143.   Madden, M. M. *et al.* Synthesis of cell-permeable stapled peptide dual inhibitors of the p53-Mdm2/Mdmx interactions via photoinduced cycloaddition. *Bioorg. Med. Chem. Lett.* **21**, 1472–1475 (2011).

144.   Brunel, F. M. & Dawson, P. E. Synthesis of constrained helical peptides by thioether ligation: application to analogs of gp41. *Chem. Commun.* 2552 (2005) doi:10.1039/b419015g.

145.   Jackson, D. Y., King, D. S., Chmielewski, J., Singh, S. & Schultz, P. G. General approach to the synthesis of short .alpha.-helical peptides. *J. Am. Chem. Soc.* **113**, 9391–9392 (1991).

146.   Lie, M. Z. & Stephen J., E. SLIC: A Method for Sequence- and Ligation-Independent Cloning. in *Gene Synthesis* (ed. Peccoud, J. pdf) vol. 852 51–59 (Humana Press, 2012).

147.   Quan, J. & Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE* **4**, (2009).

148.   Quan, J. & Tian, J. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat. Protoc.* **6**, 242–251 (2011).

149.   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).

150.   Katz, N., Cohen, R., Atar, O., Goldberg, S. & Amit, R. An Assay for Quantifying Protein-RNA Binding in Bacteria. *J. Vis. Exp.* 1–9 (2019) doi:10.3791/59611.

151.   Kabsch, W. XDS.å *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

152.   McCoy, A. J. *et al.* ıt Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

153.   Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).

154.   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

155.   Winter, G. *xia2* : an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).

156.   Winn, M. D. *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

157.   Holm, T. *et al.* Studying the uptake of cell-penetrating peptides. *Nat. Protoc.* **1**, 1001–1005 (2006).

158.   Peier, A. *et al.* NanoClick: A High Throughput, Target-Agnostic Peptide Cell Permeability Assay. *ACS Chem. Biol.* **16**, 293–309 (2021).

159.   https://www.promega.de/-/media/files/resources/protocols/technical-manuals/500/intracellular-te-nanoglo-substrate-inhibitor-and-nanoglo-vivazine-inhibitor-protocol.pdf?rev=128d75627a404a8eb0a1cbd7d9b92c1b&sc_lang=en. TECHNICAL MANUAL Intracellular TE Nano-Glo® Substrate/Inhibitor and Intracellular TE Nano-Glo® Vivazine™/Inhibitor.

160.   Colas, P. The eleven-year switch of peptide aptamers. *J. Biol.* **7**, 2 (2008).

161.   Groner, B., Borghouts, C. & Kunz, C. Peptide Aptamer Libraries. *Comb. Chem. High Throughput Screen.* **11**, 135–145 (2008).

162.   Stockert, O. M., Gravel, C. M. & Berry, K. E. A bacterial three-hybrid assay for forward and reverse genetic analysis of RNA–protein interactions. *Nat. Protoc.* **17**, 941–961 (2022).

163.   Katz, N. *et al.* An in Vivo Binding Assay for RNA-Binding Proteins Based on Repression of a Reporter Gene. *ACS Synth. Biol.* **7**, 2765–2774 (2018).

164.   Katz, N. *et al.* Synthetic 5′ UTRs Can Either Up- or Downregulate Expression upon RNA-Binding Protein Binding. *Cell Syst.* **9**, 93-106.e8 (2019).

165.   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).

166.   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).

167.   Tavassoli, A. & Benkovic, S. J. Genetically Selected Cyclic-Peptide Inhibitors of AICAR Transformylase Homodimerization. *Angew. Chem. Int. Ed.* **44**, 2760–2763 (2005).

168.   Horswill, A. R., Savinov, S. N. & Benkovic, S. J. A systematic method for identifying small-molecule modulators of protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15591–15596 (2004).

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

170.   Treiber, T. *et al.* A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Mol. Cell* **66**, 270-284.e13 (2017).

171.   Balleza, E., Kim, J. M. & Cluzel, P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods* **15**, 47–51 (2018).

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

173.   Schmeing, S. *et al.* Rationally designed stapled peptides allosterically inhibit PTBP1 RNA-binding. *Chem. Sci.* 10.1039.D3SC00985H (2023) doi:10.1039/D3SC00985H.

174.   Kafasla, P., Lin, H., Curry, S. & Jackson, R. J. Activation of picornaviral IRESs by PTB shows differential dependence on each PTB RNA-binding domain. *RNA* **17**, 1120–1131 (2011).

175.   Dorn, G. *et al. Integrative solution structure of a PTBP1-viral IRES complex reveals strong compaction and ordering with residual conformational flexibility.* http://biorxiv.org/lookup/doi/10.1101/2022.07.08.498958 (2022) doi:10.1101/2022.07.08.498958.

176.   Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1751**, 119–139 (2005).

177.   Prinz, H. Hill coefficients, dose–response curves and allosteric mechanisms. *J. Chem. Biol.* **3**, 37–44 (2010).

178.   Schreiber, S. L. The Rise of Molecular Glues. *Cell* **184**, 3–9 (2021).

179.   Cromm, P. M., Spiegel, J. & Grossmann, T. N. Hydrocarbon Stapled Peptides as Modulators of Biological Function. *ACS Chem. Biol.* **10**, 1362–1375 (2015).

180.   Yuen, T. Y. *et al.* Stereoisomerism of stapled peptide inhibitors of the p53-Mdm2 interaction: An assessment of synthetic strategies and activity profiles. *Chem. Sci.* **10**, 6457–6466 (2019).

181.   Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).

182.   Cimmperman, P. *et al.* A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. *Biophys. J.* **95**, 3222–3231 (2008).

183.   Lal, P. *et al.* Regulation of HuR structure and function by dihydrotanshinone-I. *Nucleic Acids Res.* **45**, 9514–9527 (2017).

184.   Wu, X. *et al.* Identification and Validation of Novel Small Molecule Disruptors of HuR-mRNA Interaction. *ACS Chem. Biol.* **10**, 1476–1484 (2015).

185.   Meisner, N. C. *et al.* Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat. Chem. Biol.* **3**, 508–515 (2007).

186.   Kaur, K. *et al.* The fungal natural product azaphilone-9 binds to HuR and inhibits HuR-RNA interaction in vitro. *PLOS ONE* **12**, e0175471 (2017).

187.   Madani, F., Lindberg, S., Langel, Ü., Futaki, S. & Gräslund, A. Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. *J. Biophys.* **2011**, 1–10 (2011).

188.   Richard, J. P. *et al.* Cell-penetrating Peptides. *J. Biol. Chem.* **278**, 585–590 (2003).

189.   Guterstam, P. *et al.* Elucidating cell-penetrating peptide mechanisms of action for

membrane interaction, cellular uptake, and translocation utilizing the hydrophobic counter-anion pyrenebutyrate. *Biochim. Biophys. Acta BBA - Biomembr.* **1788**, 2509–2517 (2009).

190. Säälik, P. *et al.* Protein Cargo Delivery Properties of Cell-Penetrating Peptides. A Comparative Study. *Bioconjug. Chem.* **15**, 1246–1253 (2004).

191. Wadia, J. S., Stan, R. V. & Dowdy, S. F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* **10**, 310–315 (2004).

192. Thorén, P. E. G. *et al.* Uptake of analogs of penetratin, Tat(48–60) and oligoarginine in live cells. *Biochem. Biophys. Res. Commun.* **307**, 100–107 (2003).

193. Li, Y.-C. *et al.* A Versatile Platform to Analyze Low-Affinity and Transient Protein-Protein Interactions in Living Cells in Real Time. *Cell Rep.* **9**, 1946–1958 (2014).

194. Sun, T.-L., Sun, Y., Lee, C.-C. & Huang, H. W. Membrane Permeability of Hydrocarbon-Cross-Linked Peptides. *Biophys. J.* **104**, 1923–1932 (2013).

195. Okamoto, T. *et al.* Stabilizing the Pro-Apoptotic BimBH3 Helix (BimSAHB) Does Not Necessarily Enhance Affinity or Biological Activity. *ACS Chem. Biol.* **8**, 297–302 (2013).

196. Chu, Q. *et al.* Towards understanding cell penetration by stapled peptides. *MedChemComm* **6**, 111–119 (2015).

197. Huhmann, S. & Koksch, B. Fine-Tuning the Proteolytic Stability of Peptides with Fluorinated Amino Acids: Fine-Tuning the Proteolytic Stability of Peptides with Fluorinated Amino Acids. *Eur. J. Org. Chem.* **2018**, 3667–3679 (2018).

198. Proctor, A., Wang, Q., Lawrence, D. S. & Allbritton, N. L. Metabolism of peptide reporters in cell lysates and single cells. *The Analyst* **137**, 3028 (2012).

199. Rozek, A., Powers, J.-P. S., Friedrich, C. L. & Hancock, R. E. W. Structure-Based Design of an Indolicidin Peptide Analogue with Increased Protease Stability ,. *Biochemistry* **42**, 14130–14138 (2003).

200. Rahman, L., Bliskovski, V., Kaye, F. J. & Zajac-Kaye, M. Evolutionary conservation of a 2-kb intronic sequence flanking a tissue-specific alternative exon in the PTBP2 gene. *Genomics* **83**, 76–84 (2004).

201. Rahman, L., Bliskovski, V., Reinhold, W. & Zajac-Kaye, M. Alternative Splicing of Brain-Specific PTB Defines a Tissue-Specific Isoform Pattern That Predicts Distinct Functional Roles. *Genomics* **80**, 245–249 (2002).

202. Boutz, P. L. *et al.* A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.* **21**, 1636–1652 (2007).

203. Spellman, R., Llorian, M. & Smith, C. W. J. Crossregulation and Functional Redundancy between the Splicing Regulator PTB and Its Paralogs nPTB and ROD1. *Mol. Cell* **27**, 420–434 (2007).

204. Zheng, S. *et al.* PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat. Neurosci.* **15**, 381–388 (2012).

205. Gueroussov, S. *et al.* An alternative splicing event amplifies evolutionary differences between vertebrates. *Science* **349**, 868–873 (2015).

206. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).

207. Lee, J. T. Epigenetic Regulation by Long Noncoding RNAs. *Science* **338**, 1435–1439 (2012).

208. Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci.* **106**, 11667–11672 (2009).

209. Mercer, T. R. & Mattick, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **20**, 300–307 (2013).

210. Li, Y. *et al.* A Compound AC1Q3QWB Selectively Disrupts HOTAIR-Mediated Recruitment of PRC2 and Enhances Cancer Therapy of DZNep. *Theranostics* **9**, 4608–4623 (2019).

211. Pedram Fatemi, R. *et al.* Screening for Small-Molecule Modulators of Long Noncoding

RNA-Protein Interactions Using AlphaScreen. *SLAS Discov.* **20**, 1132–1141 (2015).

212. Wysocka, J. *et al.* WDR5 Associates with Histone H3 Methylated at K4 and Is Essential for H3 K4 Methylation and Vertebrate Development. *Cell* **121**, 859–872 (2005).

213. Guarnaccia, A. & Tansey, W. Moonlighting with WDR5: A Cellular Multitasker. *J. Clin. Med.* **7**, 21 (2018).

214. Ernst, P. & Vakoc, C. R. WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Brief. Funct. Genomics* **11**, 217–226 (2012).

215. Rao, R. C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334–346 (2015).

216. Alicea-Velázquez, N. L. *et al.* Targeted Disruption of the Interaction between WD-40 Repeat Protein 5 (WDR5) and Mixed Lineage Leukemia (MLL)/SET1 Family Proteins Specifically Inhibits MLL1 and SETd1A Methyltransferase Complexes. *J. Biol. Chem.* **291**, 22357–22372 (2016).

217. Gori, F., Divieti, P. & Demay, M. B. Cloning and Characterization of a Novel WD-40 Repeat Protein That Dramatically Accelerates Osteoblastic Differentiation. *J. Biol. Chem.* **276**, 46515–46522 (2001).

218. Hu, X.-J. *et al.* Prokaryotic and Highly-Repetitive WD40 Proteins: A Systematic Study. *Sci. Rep.* **7**, 10585 (2017).

219. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).

220. Dias, J. *et al.* Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. *Genes Dev.* **28**, 929–942 (2014).

221. Thomas, L. R. *et al.* Interaction with WDR5 Promotes Target Gene Recognition and Tumorigenesis by MYC. *Mol. Cell* **58**, 440–452 (2015).

222. Avdic, V. *et al.* Structural and Biochemical Insights into MLL1 Core Complex Assembly. *Structure* **19**, 101–108 (2011).

223. Yang, Y. W. *et al.* Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *eLife* **3**, e02046 (2014).

224. Macdonald, J. D. *et al.* Discovery and Optimization of Salicylic Acid-Derived Sulfonamide Inhibitors of the WD Repeat-Containing Protein 5–MYC Protein–Protein Interaction. *J. Med. Chem.* **62**, 11232–11259 (2019).

225. Ding, J. *et al.* Discovery of Potent Small-Molecule Inhibitors of WDR5-MYC Interaction. *ACS Chem. Biol.* **18**, 34–40 (2023).

226. Chacón Simon, S. *et al.* Discovery of WD Repeat-Containing Protein 5 (WDR5)–MYC Inhibitors Using Fragment-Based Methods and Structure-Based Design. *J. Med. Chem.* **63**, 4315–4333 (2020).

227. Han, Q. *et al.* Discovery, evaluation and mechanism study of WDR5-targeted small molecular inhibitors for neuroblastoma. *Acta Pharmacol. Sin.* **44**, 877–887 (2023).

228. Chen, X. *et al.* Targeting WD Repeat-Containing Protein 5 (WDR5): A Medicinal Chemistry Perspective. *J. Med. Chem.* **64**, 10537–10556 (2021).

229. Couture, J.-F., Collazo, E. & Trievel, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat. Struct. Mol. Biol.* **13**, 698–703 (2006).

230. Patel, A., Dharmarajan, V. & Cosgrove, M. S. Structure of WDR5 Bound to Mixed Lineage Leukemia Protein-1 Peptide. *J. Biol. Chem.* **283**, 32158–32161 (2008).

SUPPLEMENTAL INFORMATION

CHEMISTRY

P-1

Ac——N—Q—A—R—A—Q—A—A—L—Q—A—V—N—R——NH$_2$
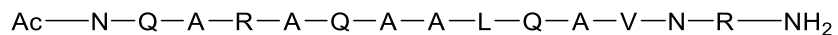
LRMS: Exact mass (calculated) [M+H]$^+$: 776.43; Observed mass: 776.60

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 16% to 40% over 40 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
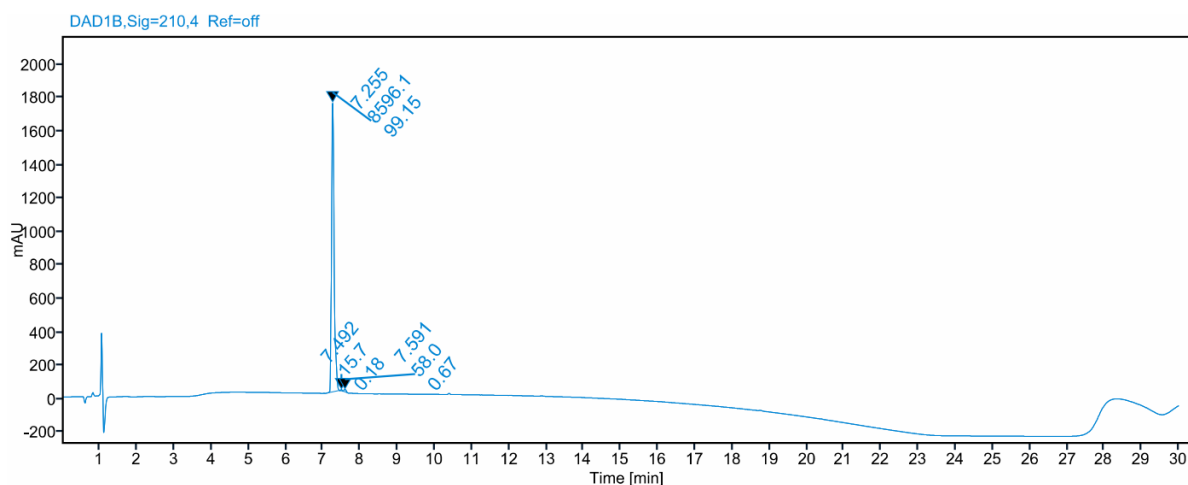
P-2
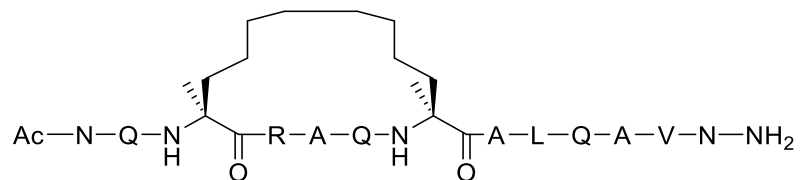
HRMS: Exact mass (calculated) [M+H]$^+$: 1505.8547; Observed mass: 1505.8480

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 16% to 40% over 40 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
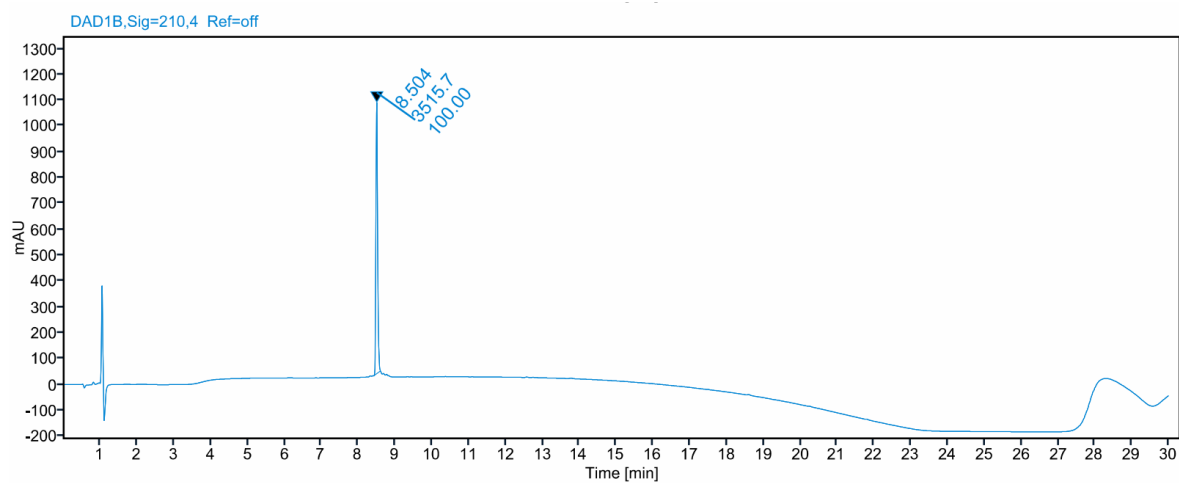
P-3
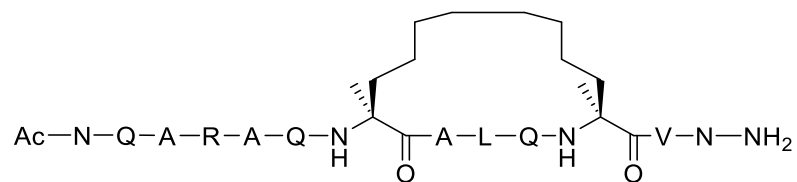
HRMS: Exact mass (calculated) [M+H]$^+$: 1505.8547; Observed mass: 1505.8581

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5μM, 125 x 21 mm) was used for purification applying a linear gradient from 16% to 40% over 40 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
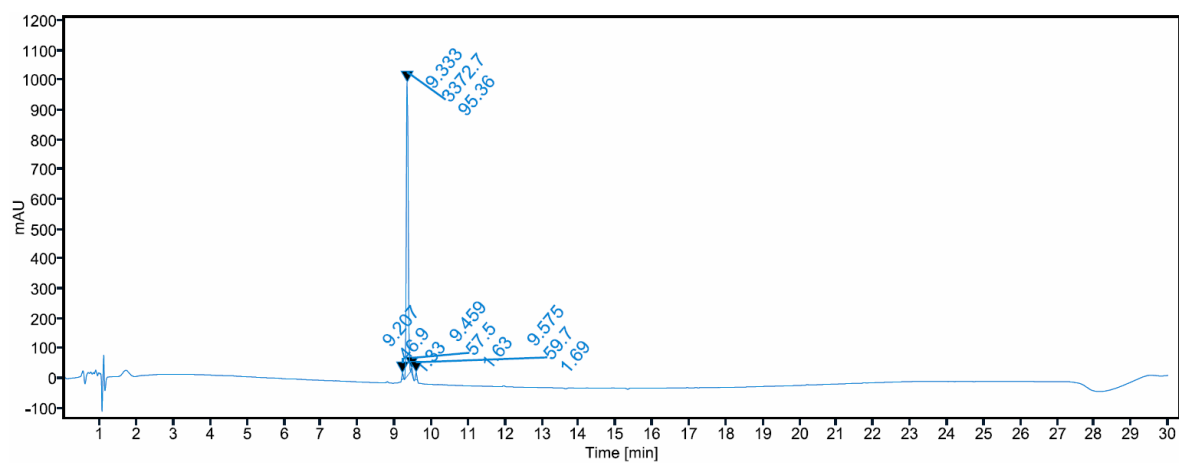
P-4
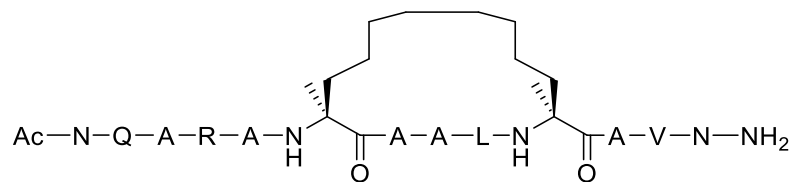
HRMS: Exact mass (calculated) [M+H]$^+$: 1391.8118; Observed mass: 1391.8148

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 16% to 40% over 40 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
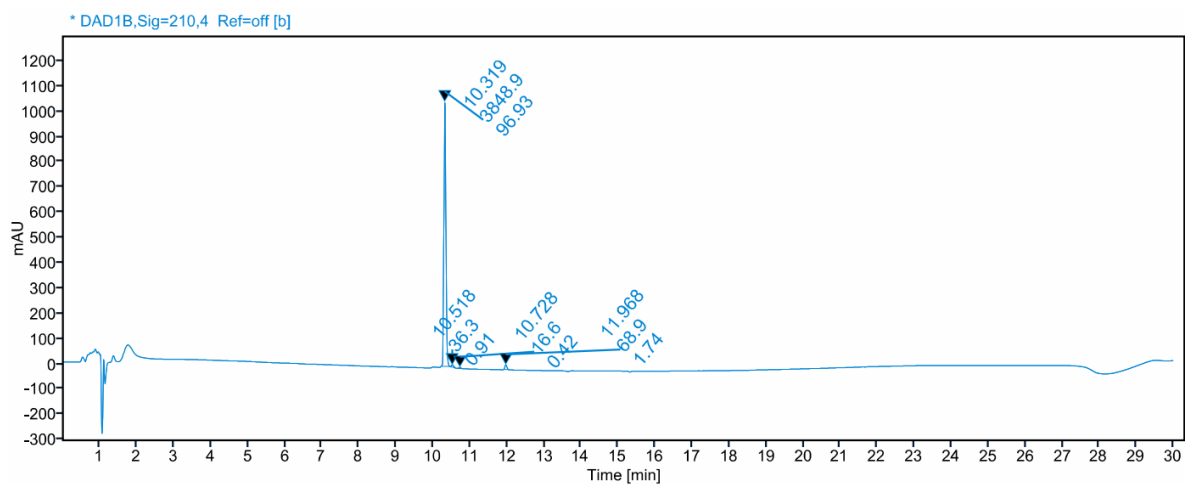
P-5
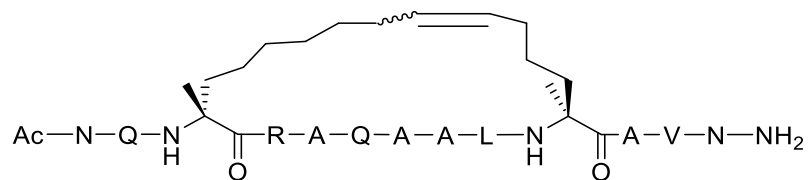
HRMS: Exact mass (calculated) [M+H]$^+$: 1488.8651; Observed mass: 1488.8670

Purification: An Agilent Infinity II LC-MS system, using a Zorbax 300SBC18 column (Agilent, Santa Clara, USA) was used for purification applying a linear gradient from 5% to 50% over 30 minutes at a flow rate of 6 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
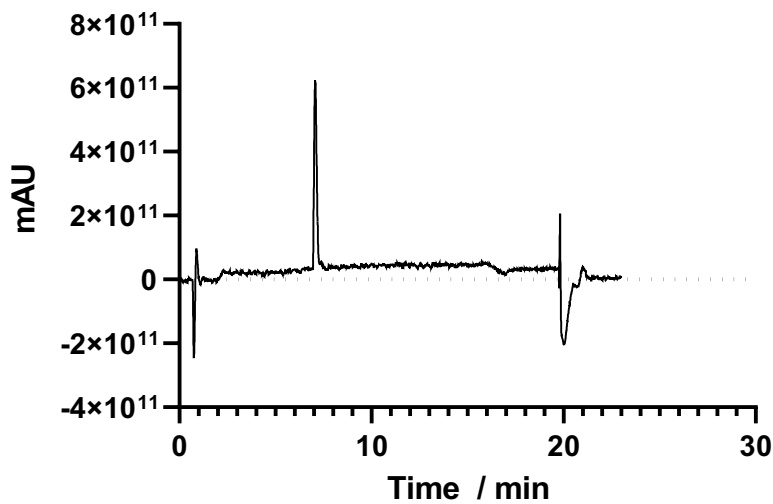
P-6
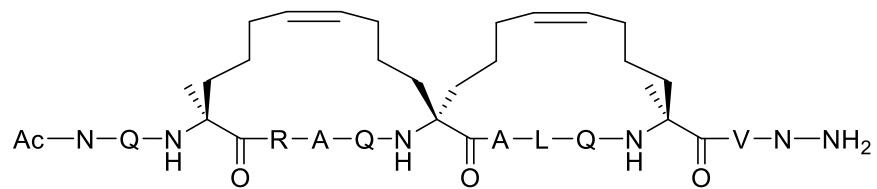
HRMS: Exact mass (calculated) $[M+H]^+$: 1597.9178; Observed mass: 1597.9205

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
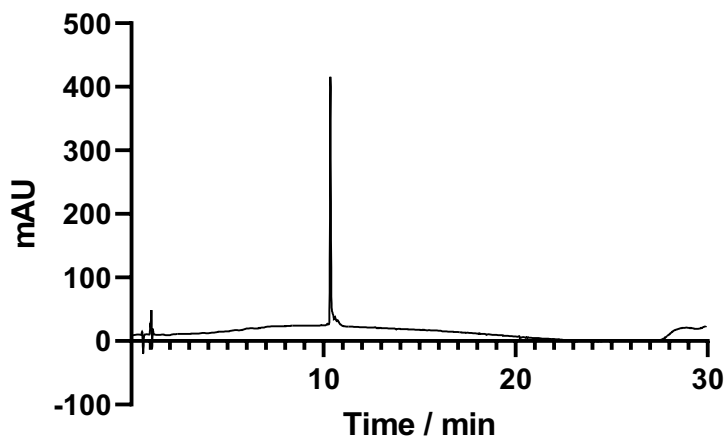
P-7
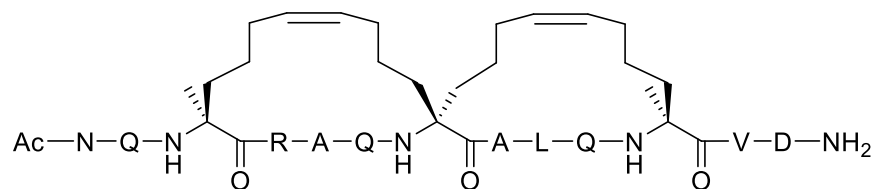
HRMS: Exact mass (calculated) [M+H]$^+$: 1598.9019; Observed mass: 1598.9013

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
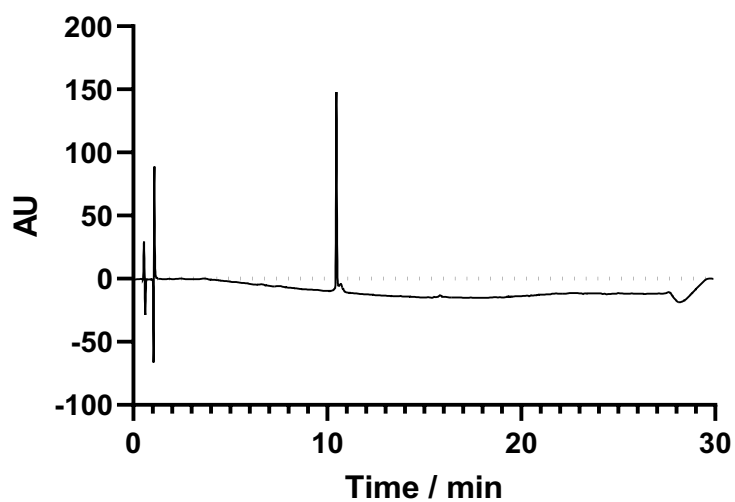
P-8
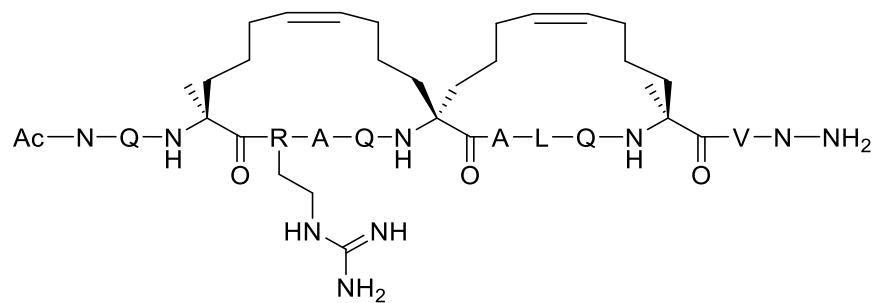
HRMS: Exact mass (calculated) [M+H]$^+$: 1583.9022; Observed mass: 1583.9056

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5μM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
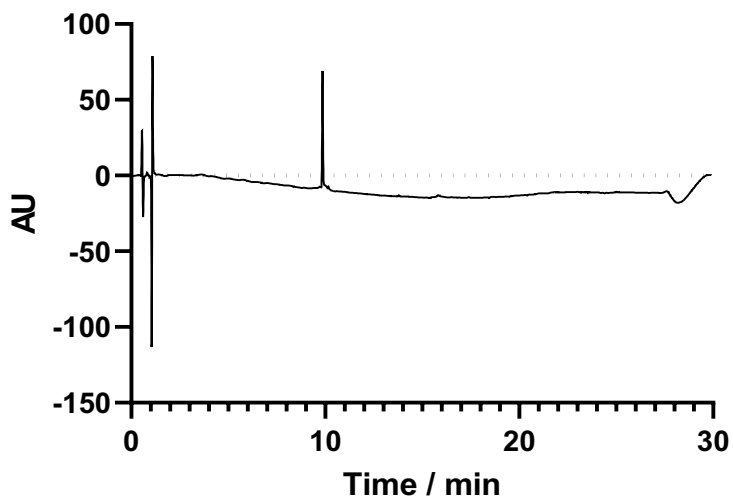
P-9
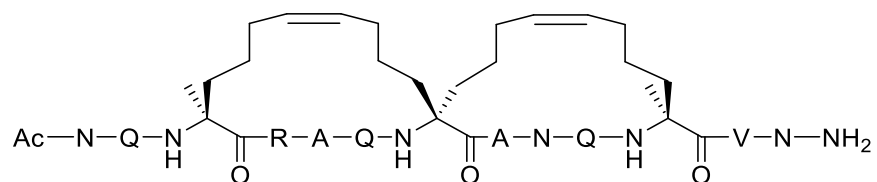
HRMS: Exact mass (calculated) $[M+H]^+$: 1598.8767; Observed mass: 1598.8805

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
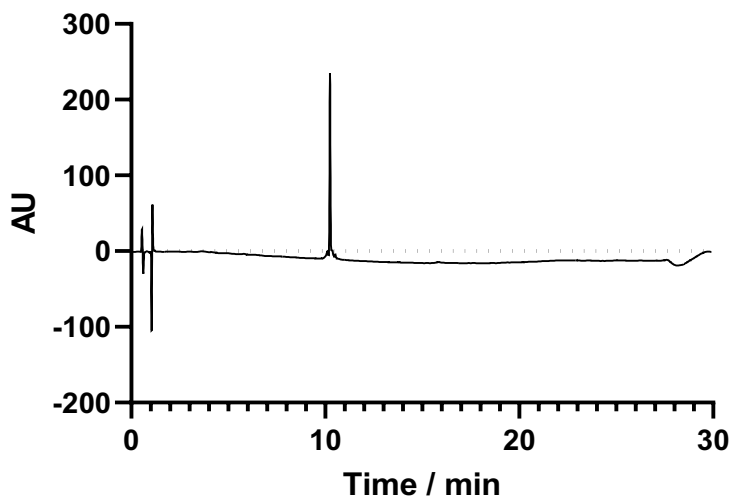
P-10
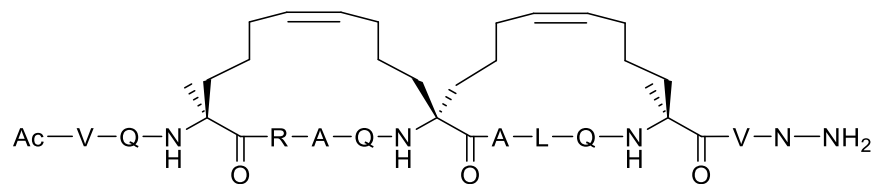
HRMS: Exact mass (calculated) [M+2H]$^{2+}$: 791.9756; Observed mass: 791.9787

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
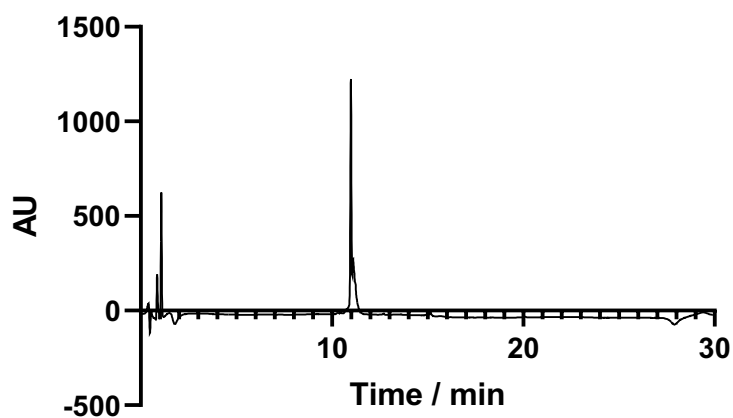
P-2F
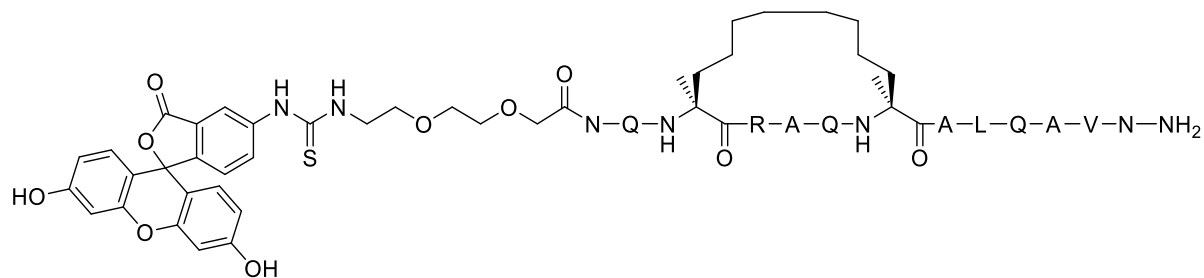
HRMS: Exact mass (calculated) $[M+2H]^{2+}$: 999.4811; Observed mass: 999.4843

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
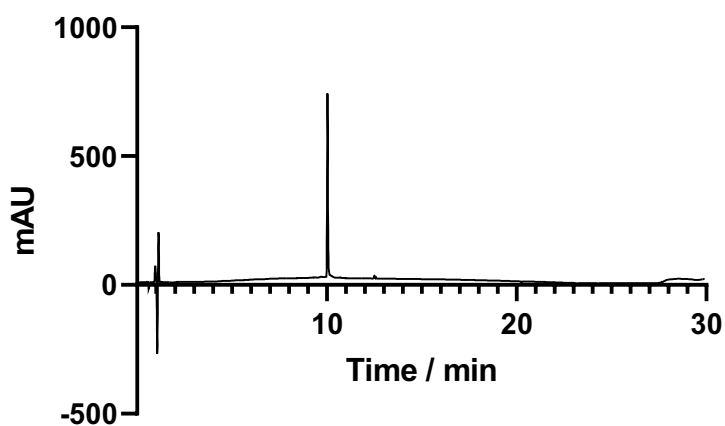
P-6F1
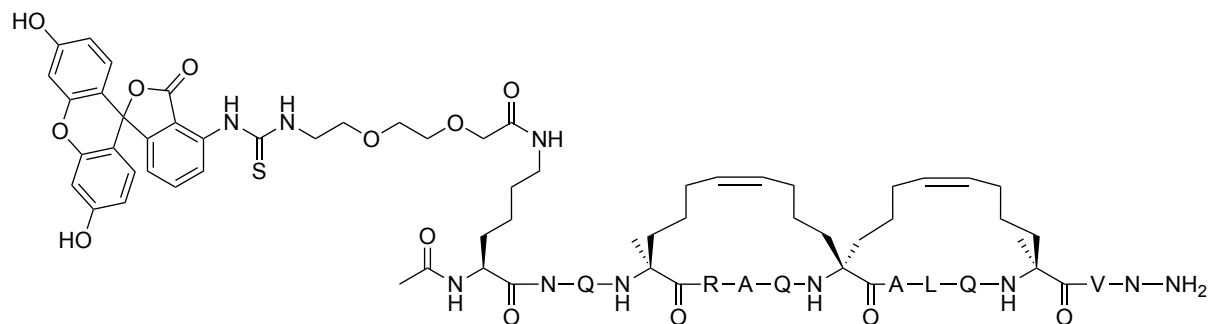
HRMS: Exact mass (calculated) $[M+2H]^{2+}$: 1130.56461; Observed mass: 1130.56743

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
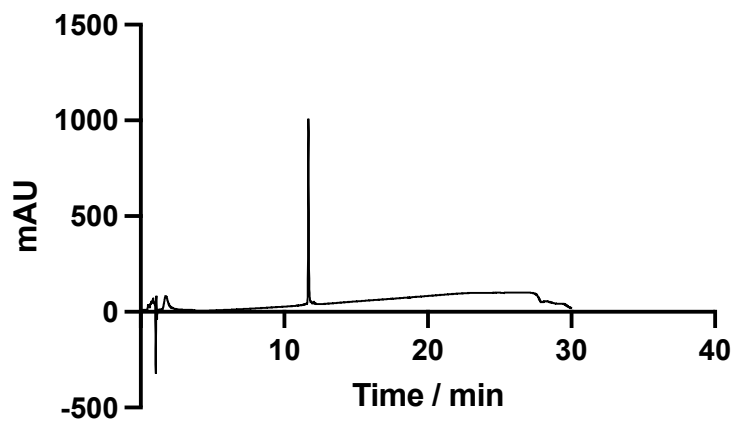
P-6F2
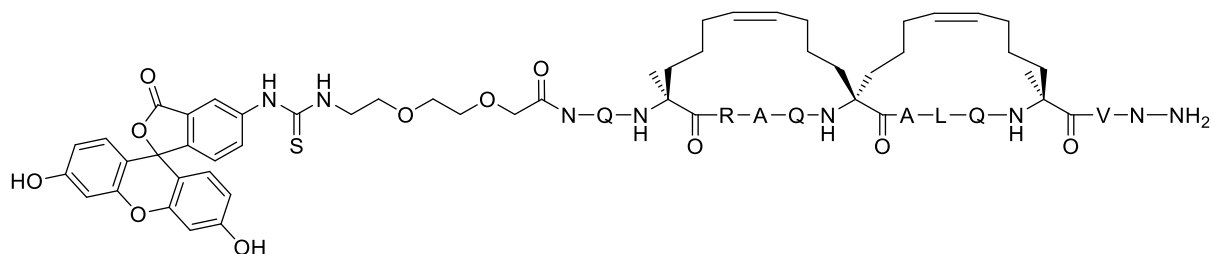
HRMS: Exact mass (calculated) $[M+2H]^{2+}$: 1045.5124; Observed mass: 1045.5160

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
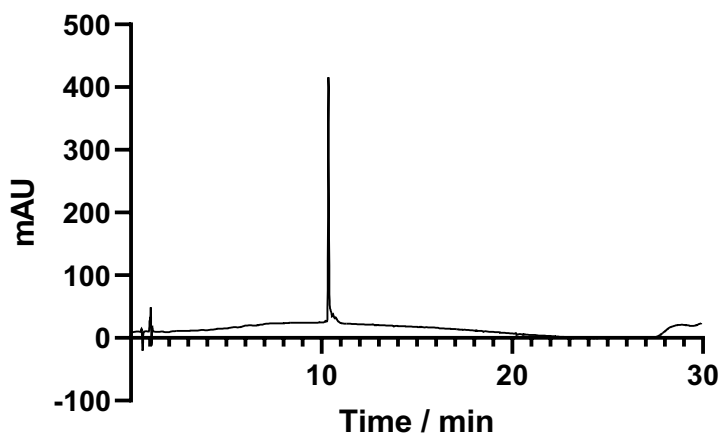
P-6S
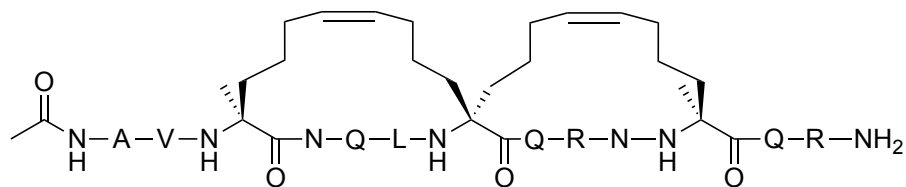
HRMS: Exact mass (calculated) $[M+2H]^{2+}$: 799.46228 ; Observed mass: 799.46410

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
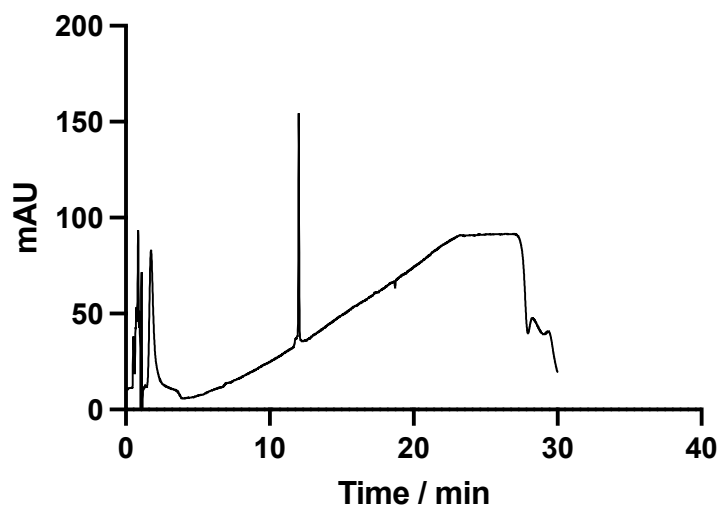
P-6S-Az
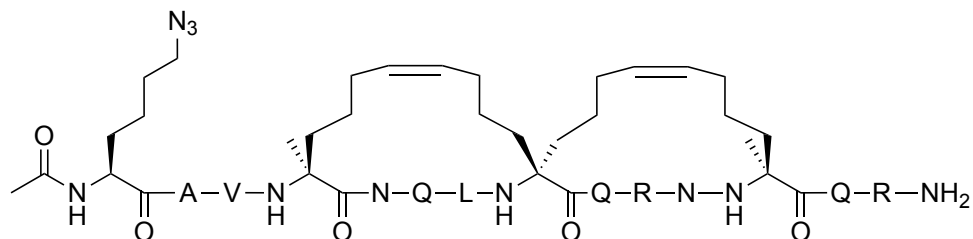
HRMS: Exact mass (calculated) [M+2H]$^{2+}$: 876.50501; Observed mass: 876.50711

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5μM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
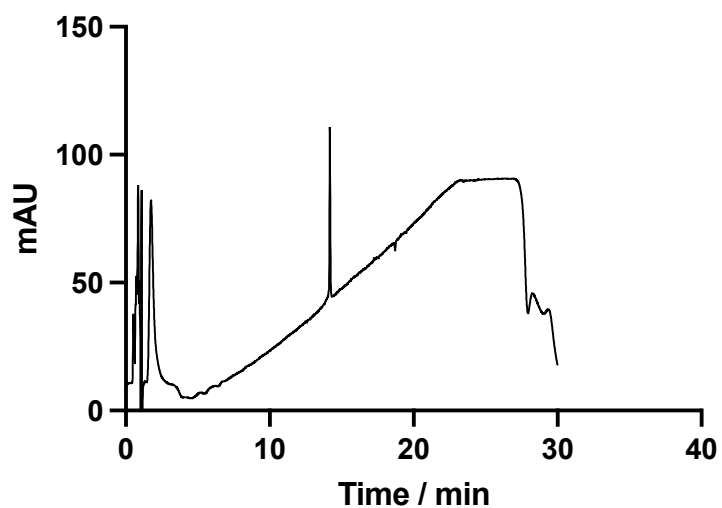
P-6-Az
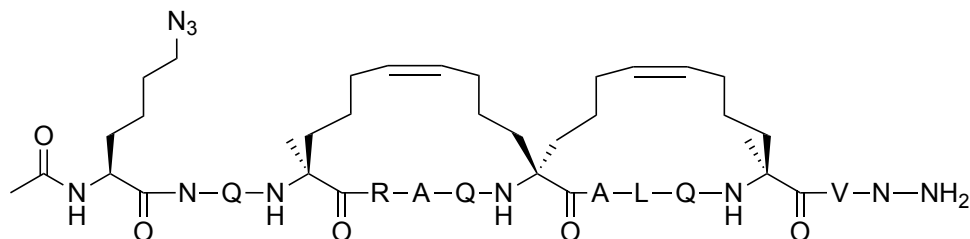
HRMS: Exact mass (calculated) $[M+2H]^{2+}$: 876.50501; Observed mass: 876.50697

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
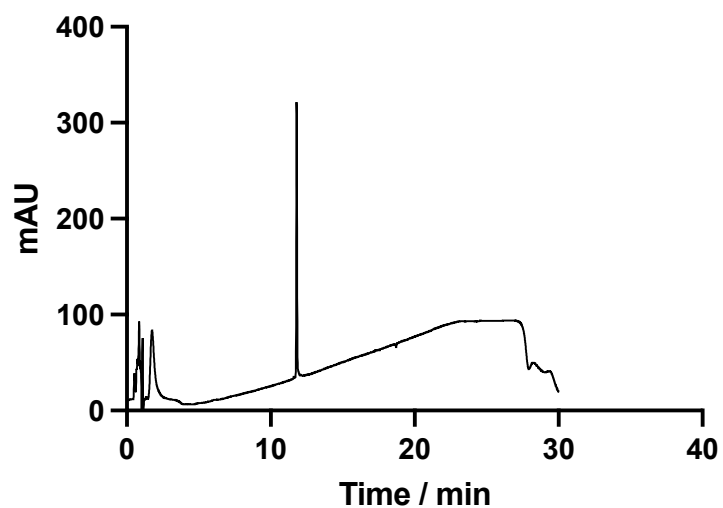
P-11F
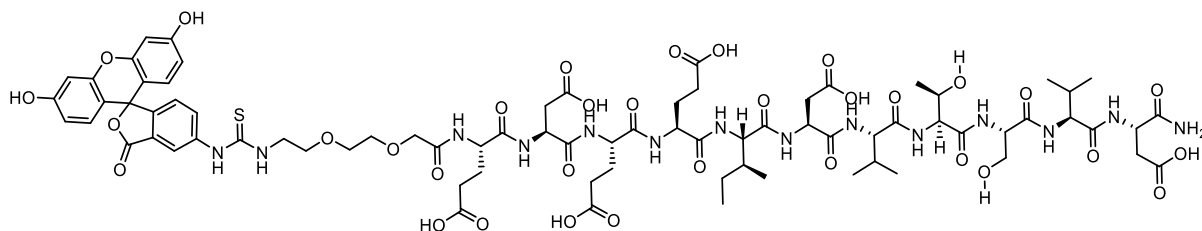
HRMS: Exact mass (calculated) [M+H]$^+$: 1783.6527; Observed mass: 1783.6517

Purification: A Büchi Pure C-850 Flash/Prep system, using a C18 column (Macherey-Nagel, 5µM, 125 x 21 mm) was used for purification applying a linear gradient from 0% to 95% over 60 minutes at a flow rate of 20 ml/min, with a mobile phase composed of eluent A (99.9% v/v H$_2$O, 0.1% v/v TFA) and eluent B (99.9% v/v ACN and 0.1% v/v TFA).
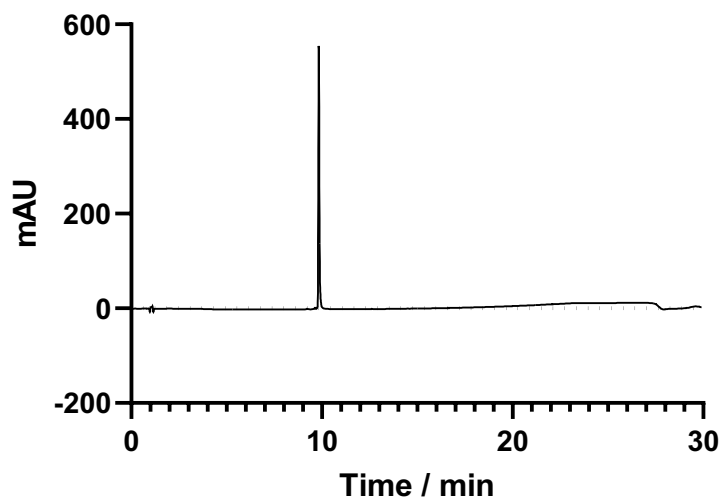
Tat-F:
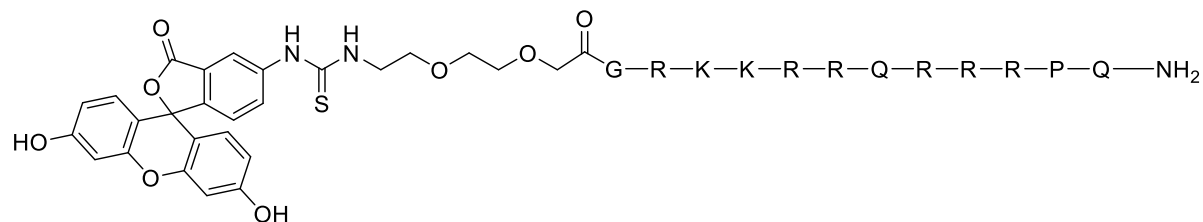
HRMS: Exact mass (calculated) $[M+3H]^{3+}$: 719.0492; Observed mass: 719.0485

Synthesis:

Synthesis was performed on ChemMatrix Rink Amide resin using 4 eq. of amino acid, 3.6 eq. COMU, 4 eq. oxyma pure, 8 eq. DIPEA in DMF containing 0.4M LiCl, where each coupling was performed 2x 30 min at r.t. Fmoc removal was done with 20% piperidine in DMF for 5 min, then repeated for 10 min at r.t. Fmoc-AEEA-OH (4 eq.) was double-coupled to the peptide with, 3.6 eq. COMU, 4 eq. oxyma pure, 8 eq. DIPEA in DMF containing 0.4M LiCl (2x 30 min, rt). Labelling was done with FITC isomer I (2 eq.) in presence of DIPEA (4 eq.) in DMF over 1h at rt, repeated overnight.

Cleavage from resin and global side chain deprotection was achieved by treatment of the resin with TFA/H_2O/DODT/TIPS (90 : 5 : 2.5 : 2.5 v/v) over 16h at rt, followed by trituration and washing in cold Et_2O, lyophylisation and purification with preparative HPLC, where MeCN + 0.1% TFA and H_2O + 0.1% TFA were used as buffers.

Purification:

Preparative scale HPLC purification of the peptides was carried out either on an Agilent Infinity II LC-MS system (Agilent Technologies, USA) equipped with a 125 mm x 21 mm, 5 µm Macherey-Nagel Nucleodur C18 Gravity column (Macherey-Nagel GmbH & Co. KG, Germany) and detection at 210 nm. Purity of the final products was determined at 210 nm using an Agilent Infinity HPLC system with 50 mm x 3 mm, 1.8 µm Macherey-Nagel Nucleodur C18 Gravity column, with a flow rate of 0.56 ml/min, using elution system: 5% → 65% MeCN (0.1% TFA) in H_2O (0.1% TFA) over 14 min. HRMS analyses were performed on a LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, USA) using electrospray ionization.
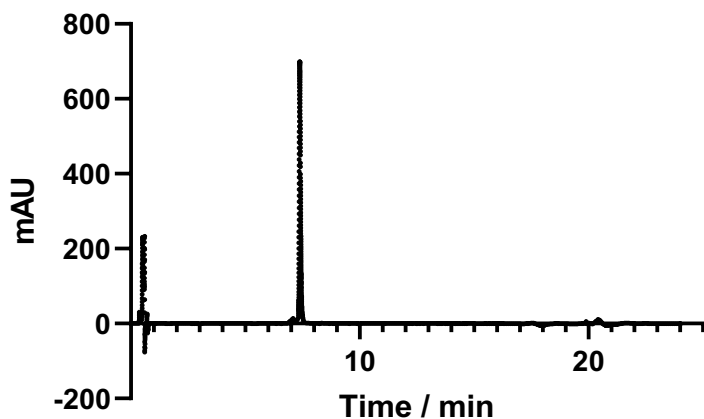
DIBAC-CA
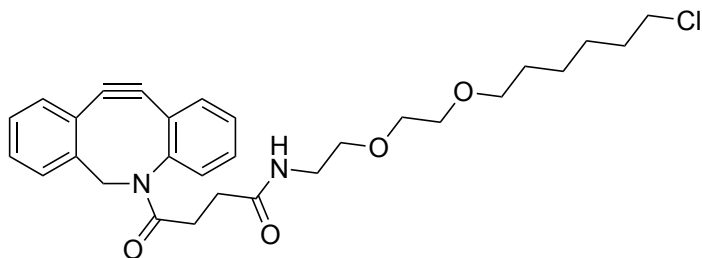
LRMS: Exact mass (calculated) [M+H]$^+$: 511.23; Observed mass [M+H]$^+$: 511.80
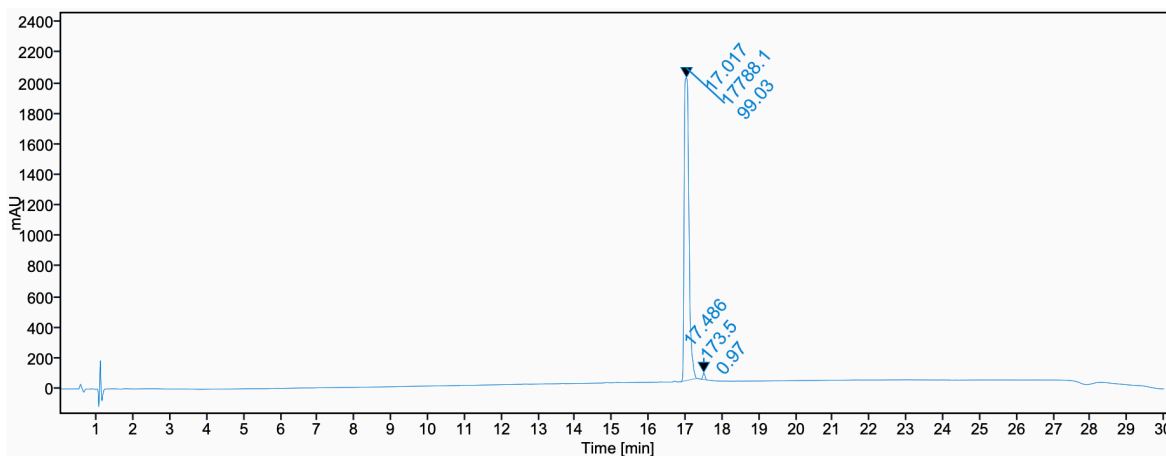
Synthesis of dibenzoazacyclooctyne Chloroalkane (DIBAC-CA)

Synthesis was based on the protocol described by Peier et al.[158] DIBAC free acid (150.0 mg, 0.45 mmol) was dissolved in DMF and HATU (256.5 mg, 0.67 mmol), 2-(2-((6-chlorohexyl)oxy)ethoxy)ethan-1-amine (117 mg, 0.45 mmol) and N,N-diisopropylethylamine (234 uL, 1.35 mmol) were added at 0 °C and the mixture was stirred for 10 minutes. Then, the reaction mixture was allowed to reach room temperature and stirred for 4hs. The solvent was removed under reduced pressure, and the resulting residue was dissolved in H_2O:MeCN (50:50, 0.01%TFA) and purified using a Büchi Pure C850 FlashPrep Chromatography system equipped with a NUCLEODUR C18 Gravity, 5 μm, 125x10 mm column, mobile phase H_2O:MeCN 0.01%TFA, from 0 to 50%MeCN in 8 min, 50 to 80% at 20 min and 100% in 30 min. The product was then lyophilized and stored protected from light. Yield: 181 mg (79.0%) as a colorless oil.

azido-ONEG
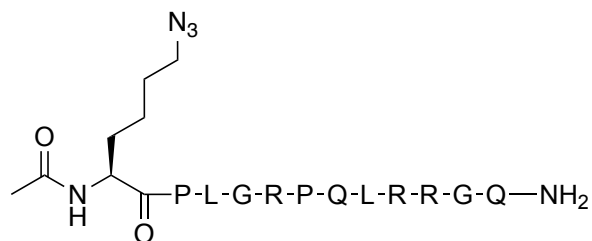
LRMS: Exact mass (calculated) $[M+2H]^{2+}$: 736.95, detected mass $[M+2H]^{2+}$: 737.00
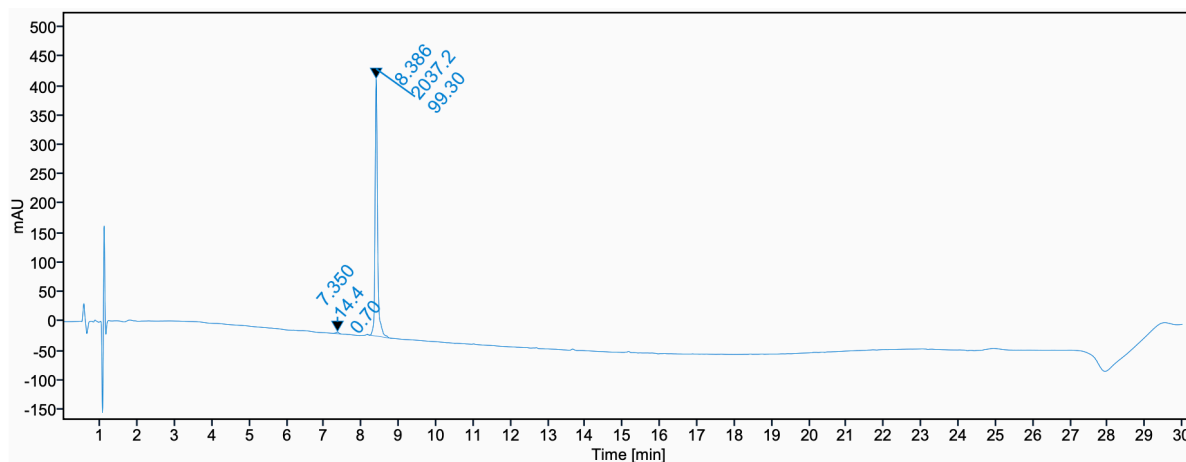
The peptide was deprotected and cleaved with a TFA:TIS:DCM (95:2.5:2.5) treatment for 1 hour followed by precipitation in cold Et_2O and centrifugation. The supernatant was removed, and the pellet resuspended in cold Et_2O followed by centrifugation. This procedure was repeated twice. The resulting pellet was dissolved in H_2O/MeCN (1:1) and lyophilized. Purification was carried out using a Büchi Pure C850 FlashPrep Chromatography system equipped with a NUCLEODUR C18 Gravity, 5 µm, 125x10 mm column, mobile phase H_2O:MeCN 0.01%TFA, from 0 to 50% MeCN in 30 min.

Azido-octa-arginine
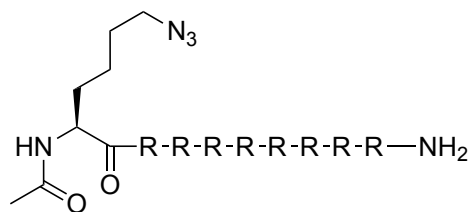
LRMS: Exact mass (calculated) $[M+3H]^{3+}$: 488.3, detected mass $[M+3H]^{3+}$: 488.6

azido-octa-arginine was treated with a TFA:TIS:DCM (95:2.5:2.5) solution for 2 hours to avoid possible remaining Pbf protecting groups. The TFA was removed under vacuum and the crude dissolved in H_2O/MeCN (1:1) and lyophilized before purification. Purification was carried out using a Büchi Pure C850 FlashPrep Chromatography system equipped with a NUCLEODUR C18 Gravity, 5 μm, 125x10 mm column, mobile phase H_2O:MeOH 0.01%TFA, from 0 to 30% MeOH in 30 min, and 30 to 100% until 40 min.

PLASMID MAPS
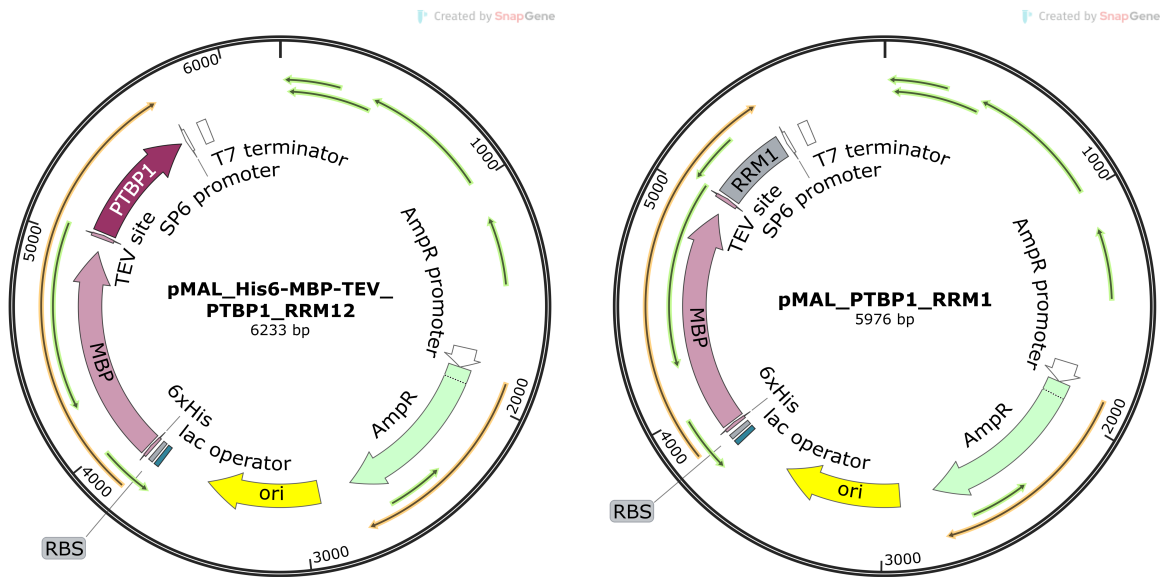
Supplemental P1: Plasmid maps of the expression vectors for PTBP1 RRM12 and RRM1 in the pMAL backbone.
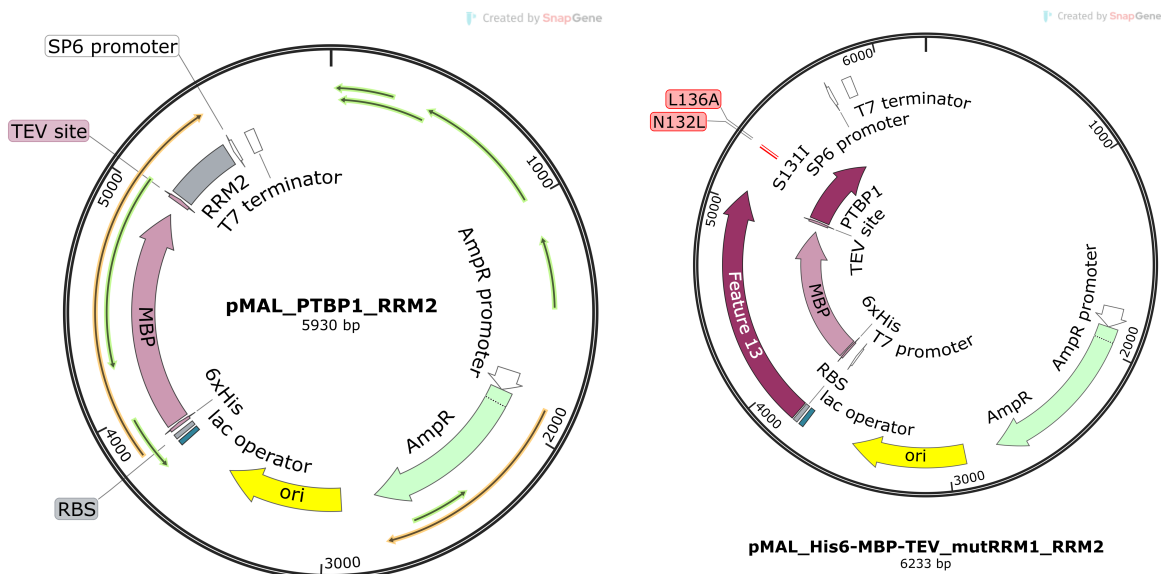
Supplemental P2: Plasmid maps of the expression vectors for PTBP1 RRM2 and RRM1*2 in the pMAL backbone.
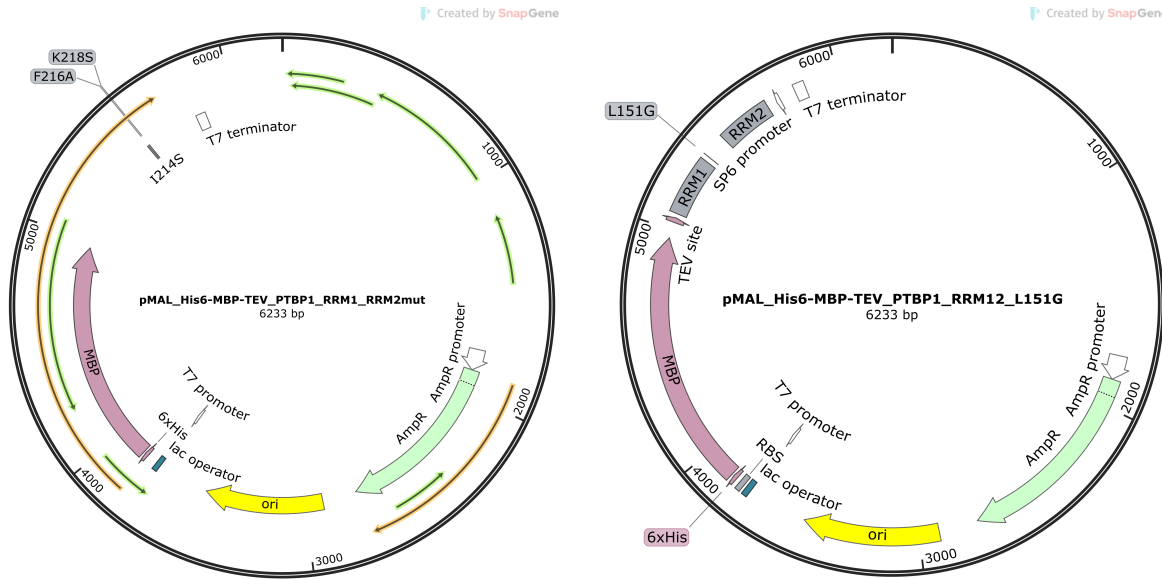
Supplemental P3: Plasmid maps of the expression vectors for PTBP1 RRM12* and RRM12-L151G in the pMAL backbone.
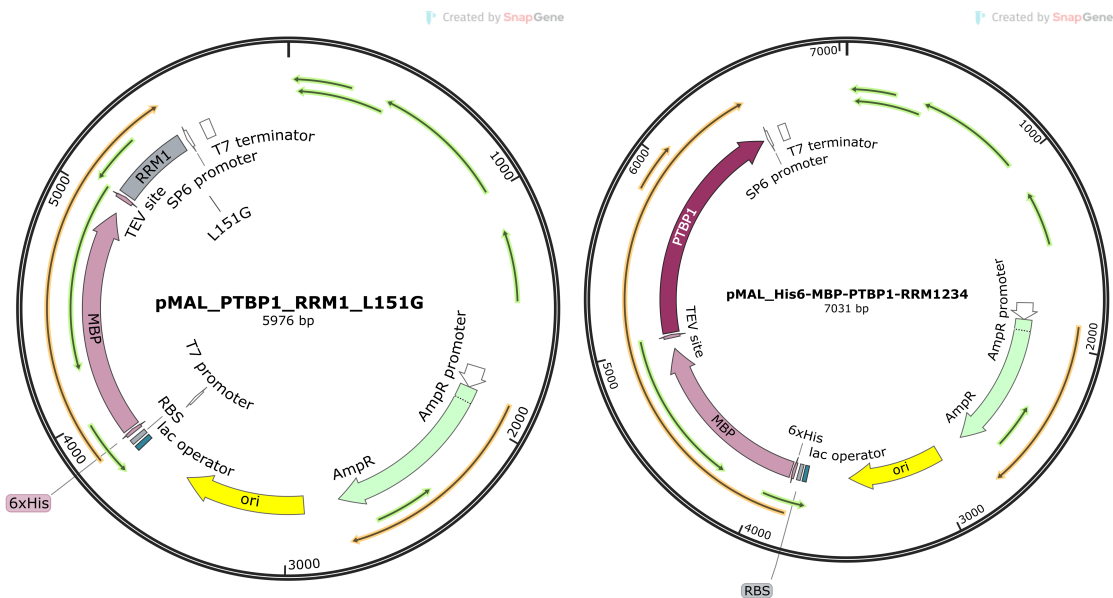
Supplemental P4: Plasmid maps of the expression vectors for PTBP1 RRM1-L151G and PTBP1-RRM1234 in the pMAL backbone.
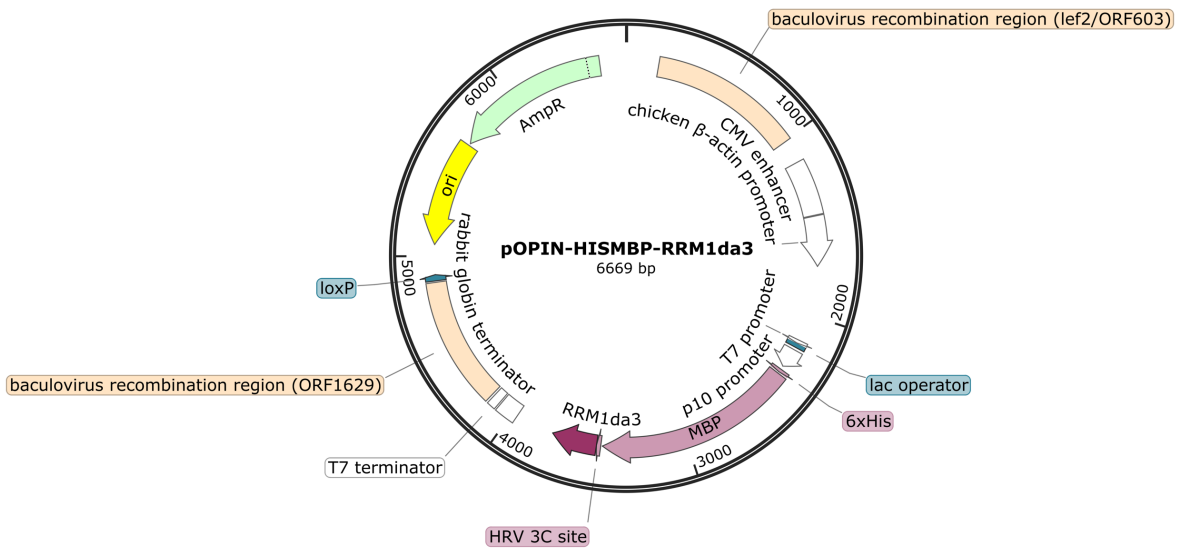
Supplemental P5: Expression vector of PTBP1-RRM1Δα3 in the pOPIN backbone.
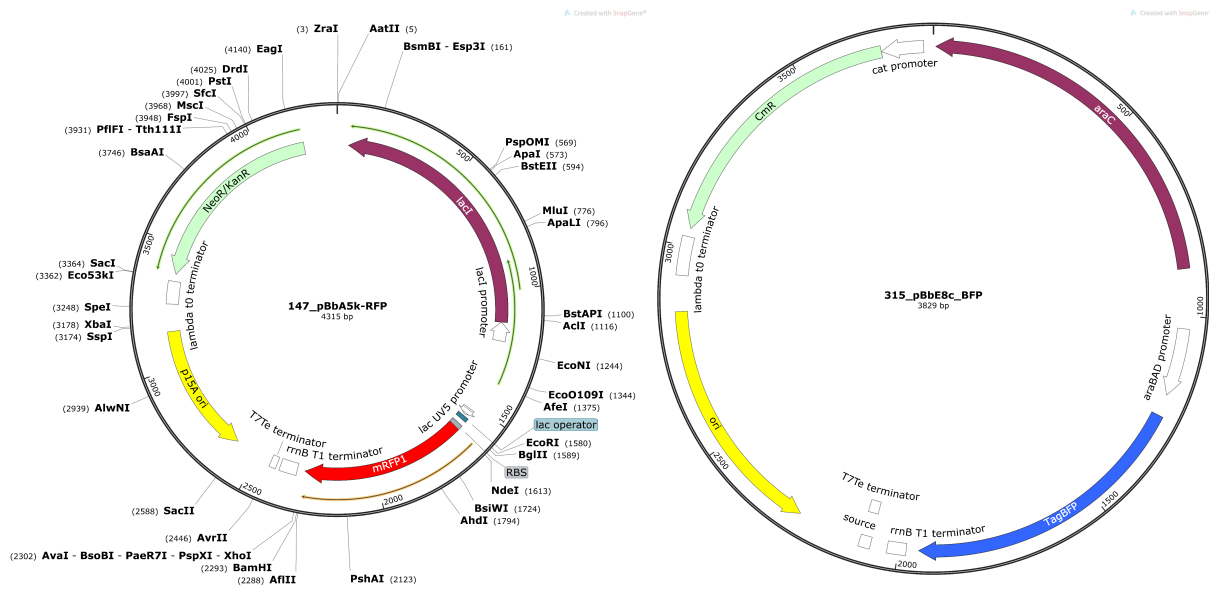
Supplemental P6: Plasmids pBbA5k-RFP and pBbE8C-BFP used in the TRAP assay.
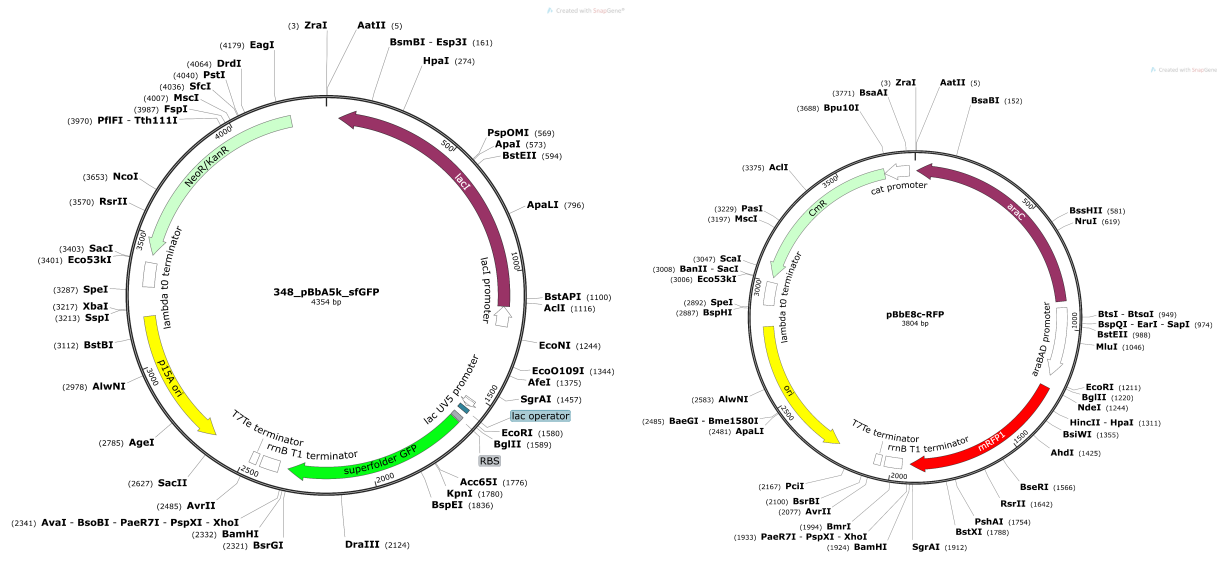
Supplemental P7: Plasmids pBbA5k-sfGFP and pBbE8c-RFP used in the TRAP assay.
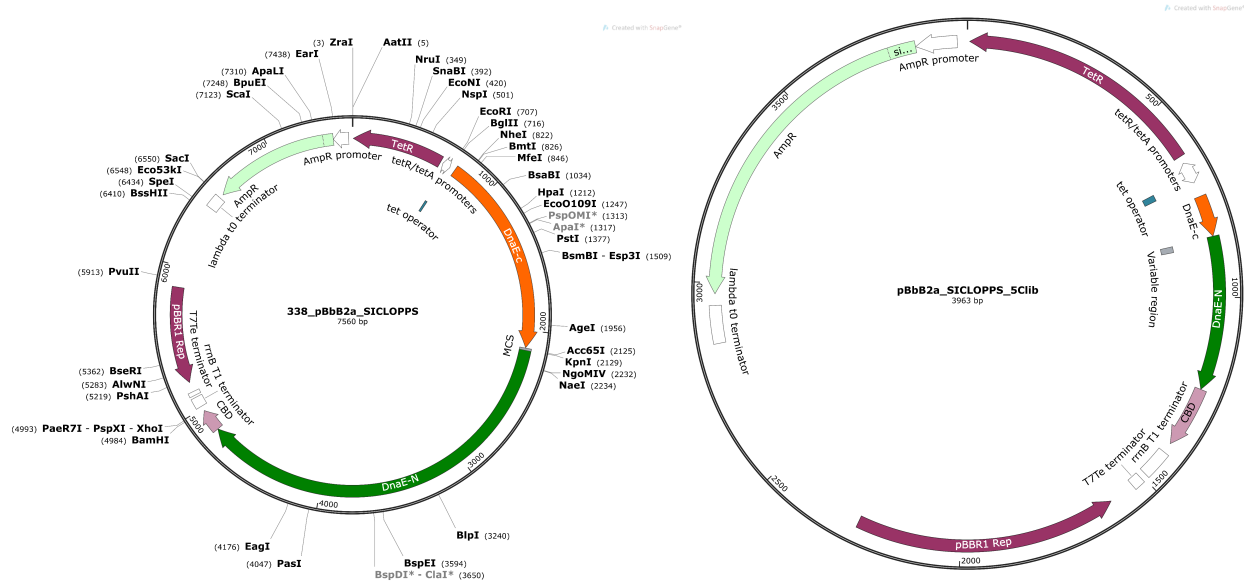
Supplemental P8: Plasmids pBb2a_SICLOPPS used as the backbone for library generation (left) and the resulting library plasmid map after cloning (right).

LIST OF FIGURES

EIDESSTATTLICHE VERSICHERUNG (AFFIDAVIT)

<u>Schmeing, Stefan</u>

Name, Vorname
(Surname, first name)

<u>166510</u>

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar. Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

<u>Dortmund,_____</u>

Ort, Datum
(Place, date)

_____

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

<u>TARGETING OF RNA-BINDING PROTEINS WITH MACROCYCLIC PEPTIDES</u>

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.

Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

<u>Dortmund,_____</u>

Ort, Datum
(Place, date)

_____

Unterschrift
(Signature)