Published in final edited form as:

Chembiochem. 2019 April 01; 20(7): 931–939. doi:10.1002/cbic.201800645.

A small cyclic β -hairpin peptide mimics the Rbfox2 RRM and binds to the precursor miRNA 20b

Yi-Ting Sun^[a], Matthew D. Shortridge^[a], and Gabriele Varani^[a]

^[a]Department of Chemistry, University of Washington, 4000 15th Ave NE, Bagley Hall, Seattle, WA 98195-1700 (USA)

Abstract

The RRM (RNA-recognition motif), the most abundant RNA-binding motif in eukaryotes, is a well-structured domain of about 90 amino acids, yet the β 2- β 3 hairpin corresponding to strands 2 and 3 of the β -sheet, and the intervening loop makes essential interactions with RNA in many RRM complexes. We designed a series of small cyclic peptides mimics of the β 2- β 3 hairpin of Rbfox2 protein, that recognize the terminal loop of precursor miR-20b, to investigate whether we could mimic the full RNA-binding protein with a minimal structurally pre-organized peptide. Within a small library of 7 cyclic peptides, we identified a peptide with low μ -molar affinity for the miR-20b precursor. NMR titration data suggest that this peptide specifically targets the apical loop of pre-miR-20b. This work shows that it is possible to mimic RNA-binding proteins with designed stable peptides, which provide a starting point for designing or evolving small peptide mimetics of RRM proteins.

Graphical Abstract

Rbfox2 RRM recognizes the terminal loop of pre-miR-20b, with its β 2 β 3 hairpin inserts into the opened-up terminal loop. A short cyclic peptide CPfox4 was designed to mimic the core of recognition elements of RRM and target pre-miR-20b.

Keywords

peptide design; peptide mimics; Rbfox2; RNA recognition motif; pre-miR-20b

Introduction

Several RNA binding proteins regulate processing of miRNAs, for example the highly conserved RRM of Rbfox proteins (Rbfox2 RRM), which inhibits the maturation of miR-20b and miR-107 by binding to the terminal loop of the precursor miRNAs.^[1] The RNA recognition motif (RRM) is the most abundant RNA-binding motif in eukaryotes.^[2] RRM-containing proteins are involved in all aspects of post-transcriptional gene regulation, including pre-mRNA processing, mRNA export and stability.^[3] RRMs are typically 80–100 amino acids long and fold to form a four-stranded antiparallel β -sheet, packed against two α -helices, with a split $\alpha\beta$ topology.^[3–4] Despite the similarity in structure, RRMs are able to bind to a wide variety of RNAs that have different sequence and length, in some cases with high affinity and specificity, although this most often is found in proteins containing multiple domains.^[5] The versatile sequence specificity provides RRM-containing protein with diverse biological functions by endowing them with the ability to target many cellular RNAs.^[3–4] Typically, RRM proteins recognize RNA through the β -sheet surface, at the center of which are strands $\beta 1$ and $\beta 3$ that contain the conserved signature sequences that define the domain.^[4, 6] However, non-canonical interactions through the α -helices^[7], protein loop^[8], and N- or C-terminal extensions^[9] are also found in RRM-RNA complexes, to boost affinity or specificity of individual domains and even to redirect RRMs to bind to proteins.

Among the RNA whose activity is regulated by RRM proteins are miRNAs^[10], small non-coding RNAs of 21–23 nucleotides that play key regulatory roles in protein expression by binding target mRNAs and inhibiting translation or inducing mRNA degradation.^[11] The primary miRNA (pri-miRNA), the long precursor of miRNA transcribed by RNA polymerases, is cleaved by the Drosha/DGCR8 microprocessor complex in the nucleus to release the precursor miRNA (pre-miRNA), a 60–70-nucleotides-long stem-loop embedding the mature miRNA. The pre-miRNA is then cleaved into a ~22 nt miRNA:miRNA* duplexes by the RNase III enzyme Dicer in the cytoplasm, from which the mature miRNA is loaded onto the RNA-induced silencing complex (RISC) to guide the repression of target mRNAs.^[10] Several RRM proteins regulate miRNA processing by binding at or near the apical region of the stem-loop precursors species, which is common to both pre- and pri-miRNAs,^[12] including the highly conserved RRM of Rbfox proteins (Rbfox2 RRM), which inhibits the maturation of miR-20b and miR-107 by binding to the terminal loop of the precursor miRNAs.^[1]

Rbfox proteins are tissue-specific alternative splicing factors,^[8a, 13] and include three family members in mammals with nearly identical N-terminal RRMs: Rbfox1(A2BP1), Rbfox2(RBM9) and Rbfox3(NeuN).^[13b] The RRM of Rbfox proteins targets RNA specifically through the conserved (U)GCAUG element, a well-known cis-element for the regulation of alternative splicing^[8a, 13a] which is also found in the apical loop of the two miRNA precursors targeted by Rbfox2.^[1] Through this interaction, the Rbfox2 RRM inhibits the processing of the two pri-miRNAs *in vitro* and in mammalian cells.^[1] The down-regulation of miR-20b by Rbfox2 increases the expression of downstream targets,^[14] such as the tumor suppressor PTEN (Phosphatase and Tensin Homolog).^[1] Conversely, up-

regulation of miR-107 by knocking-down Rbfox2 decreases the expression of Dicer protein, [1] a downstream target of miR-107.^[15]

The structure of the complex of Rbfox2 RRM and pre-miR-20b was determined by nuclear magnetic resonance (NMR) spectroscopy (Figure 1).^[11] In the complex, three base pairs in the apical stem immediately below the loop open up, drastically enlarging the size of the terminal loop. The $\beta 2\beta 3$ hairpin inserts into the opened-up terminal loop of pre-miR-20b, anchoring the RNA loop-helix junction against the protein surface. Inspired by this structure, we designed a small set of short cyclic peptide mimics of the $\beta 2\beta 3$ and investigated whether it was possible to mimic the core recognition elements of RRM proteins with a minimal peptide and target pre-miR-20b.

In addition to the intrinsic interest in protein engineering and evolution of RNA-binding proteins, which might have emerged from much smaller β -hairpins by subsequent duplication and fusion, peptides provide much more advantageous starting points to discover miRNA inhibitors^[16], an attractive prospect since mis-expression of miRNAs is observed in many human diseases^[10b, 17], often as a result of post-transcriptional mis-regulation of their maturation. Peptides potentially combine the specificity and binding activity of proteins, but are smaller in size and cheaper to manufacture and deliver^[16], and indeed peptidic inhibitors of miRNA processing have been reported.^[18] Compared to linear peptides, cyclic peptides have better binding affinity, specificity and stability due to conformational rigidity and are less susceptible to proteolysis within cells and organisms.^[19] Cyclic peptides mimic of Arginine-rich domain proteins that potently disrupt the HIV TAR RNA-Tat protein interactions have been discovered in our group^[20], and, while this work was being written-up, a peptide mimic of an RRM stabilized by a perfluoroaryl side chain linkage was discovered through evolution and selection.^[21] We wished to use a similar design strategy with the Rbfox RRM2, but unlike the Tat protein, which forms a short β -hairpin that can immediately be mimicked by cyclic peptides, RRM proteins such as Rbfox2 are much larger and have complex structures. By mimicking the $\beta 2\beta 3$ hairpin of Rbfox2 based on the structure of Fig.1, we report the discovery of a cyclic peptide mimetic of this protein which selectively binds to pre-miR-20b with a dissociation constant (Kd) $\sim 50 \mu\text{M}$, and opens the way for the optimization of this new peptide by rational design and genetic selection.

Results

Peptide Design

Cyclic β -hairpin peptides were designed to mimic the $\beta 2\beta 3$ hairpin in the Rbfox2 RRM structure (Figure 1B). This part of the protein forms a β -hairpin in all RRMs, and amino acids from it, including both the strands and intervening loop, are often involved in sequence-specific RNA recognition. For example, the specificity of U1A and U2B" proteins can be switched by swapping $\beta 2$ among the two proteins.^[22] Given its role in sequence-specific RNA recognition, we sought to mimic these hairpins with synthetic peptides, cyclized to form a stable β -hairpin with a sequence and structure matching those of the $\beta 2\beta 3$ hairpin of Rbfox2. We reasoned that the peptide would retain some of the activity of the complete protein, and that full activity could later be restored by further cycles of design and/or selection.

To provide proof of principle that this goal can be achieved, seven cyclic peptides were designed (Table 1) varying in length (14, 16 and 18 residues), as was originally done when we sought to mimic the BIV Tat protein.^[20a] It was later found that although the 14-amino-acid-long cyclic peptide are more stable, the 16-amino-acid-long cyclic peptides mimic the protein structure best, as described herein. The ^DPro-^LPro type II' β -turn was introduced into all sequences to promote β -hairpin formation.^[20a] The β 2 β 3 loop in Rbfox2 RRM is composed of Asn151, Glu152, Arg153 and Gly154. We transplanted this loop and part of the β 2 and β 3 strands onto the type II' β -turn to mimic the RNA interacting segment observed in Rbfox2-RNA structure (Figure 2). Direct transplantation of residue 148–159 and 146–161 from Rbfox2 RRM generate cyclic peptides CPfox1 and CPfox3, respectively. The basic side chain of residue 156 provides a salt bridge interaction to the RNA backbone^[3]; therefore, Lys or Arg were used in position 11 in CPfox peptides. Aromatic residues on strand β 3 (Phe158 and Phe160) which often interact with the RNA sugar in RRM/RNA complexes^[3–4] were left unchanged in all mimetic peptides (position 13 and 15). Gly residues on strand 2 were substituted with Arg in CPfox2 and CPfox4–7 (position 12 and/or 14) to increase β -sheet propensity and solubility and to rigidify the structure. Position 2 in CPfox5–7 was changed into Lys to increase the basicity of peptide, and for the same reason, position 5 in CPfox6 was changed to Arg. Finally, in CPfox7, the Glu in the center of the RNA interacting loop (position 7) was mutated to Gln to remove the acidic side chain while retaining side chain size and hydrophilicity.

Structure of the Peptides

The ability of peptides to bind the target RNA is affected by their structure. Folded peptides which pre-organize the interacting groups often give better binding affinity and especially specificity than unfolded peptides. Thus, we investigated the structure of each peptide mimetic in aqueous solution by NMR spectroscopy to confirm their rigidity. 1D and 2D NMR spectra, including 1H NOESY, TOCSY and natural abundance HSQC spectra, were obtained for all but peptide CPfox3, for which 2D spectra could not be recorded because of its low solubility in water.

The 1D and 2D ¹H NMR spectra of CPfox1 show that the peptide is unfolded (Figure 3 and Supplementary Figure 1A), as revealed by the narrow diversion of HN peaks and multiple conformations are present in exchange, as revealed by the broad peak shape. In contrast, the spectra of CPfox2 demonstrate formation of a highly stable β -hairpin structure (Figure 3 and 4A). The HN peaks are widely dispersed, and most of the ³J_{HN-H α} constants for residues within the putative β -strands are >9.0 Hz, as expected for a β -sheet structure. In addition, many intra-strand and cross-strand NOE signals expected for β -sheet structures are observed in the NOESY spectrum of CPfox2. For CPfox3, the 1D ¹H NMR spectra show a moderate dispersion of HN peaks, suggesting a less stable β -hairpin structure compared to CPfox2 (Figure 3). Unfortunately, no 2D NMR information could be collected for CPfox3 because of its low solubility. The 1D and 2D NMR spectra show that 16-mers CPfox4–7 have β -sheet structure (Figure 3, 4 and Supplementary Figure 1B–D). The dispersed HN peaks and β -sheet NOE signals found in 2D NOESY spectra of these peptides indicate the formation of β -hairpins. However, the observation of fewer cross-strand NOEs and moderate dispersion of HN peaks, compared to CPfox2, suggests that the β -hairpins are moderately structured or

perhaps experiencing conformational exchange. This is not inconsistent with their length, less favorable than the 14-residue sequences of CPfox2 for β -hairpin formation.

Altogether, the NMR studies demonstrate that CPfox1 and CPfox2 are mostly unfolded and folded peptides, respectively, while the rest of the peptide mimetics fold into moderately stable β -hairpins.

Structure calculations were performed for CPfox2 and CPfox4–7 using NOE-derived restraints (see Method section). The results are shown in Figure 5 and summarized in Table 2. The 10 lowest structures converge well for each peptide, with no NOE restraint violation greater than 0.2 Å and high structural definition. The backbone RMSD values of the 5 calculated peptide structures are all between 1.0 and 1.4 Å (Table 2).

The calculations confirm that the 14-mer CPfox2 is a highly structured antiparallel β -hairpin with type-II and type-II' β -turns at each end of the hairpin (Figure 5A). However, the hydrogen-bonding residues observed in the structure of this peptide differ from those observed in the Rbfox2 RRM protein, suggesting that CPfox2 is a very poor structural mimic of $\beta 2\beta 3$ in Rbfox2. Namely, residue Ile1, Phe3, Glu5, Ser8, Arg10 and Arg12 are hydrogen-bonded across the β -strand, while the corresponding residues in Rbfox2 RRM are not. As a consequence, the side chains in the peptide point in the opposite direction compared to the corresponding residue in the Rbfox2 RRM (Figure 6A). We encountered a similar problem with BIV Tat and were unable to simply transplant the BIV ARM sequence onto a $^D\text{Pro-LPro}$ template, but had instead to change its sequence to mimic the peptide structure observed in the complex with RNA.^[20a]

The structure calculation of CPfox4 revealed a less structured β -hairpin compared to CPfox2 (Figure 5B), with a longer and more flexible loop. However, this 16-mer peptide presents the same hydrogen-bonding residues and side chain geometry as Rbfox2 RRM (Figure 6B). The structures calculated for the other 16-mer peptides also showed the side chain geometry observed in the Rbfox2-RNA structure, suggesting that 16-mer peptides can mimic the $\beta 2\beta 3$ hairpin of the Rbfox2 RRM, while 14-mer peptides cannot do so because of intrinsic structural preferences induced by the $^D\text{Pro-LPro}$ turn. Among all 16-mers, both CPfox4 and CPfox7 possess a flat round loop (Figure 5B and E), while the loops of CPfox5 and CPfox6 are more rigidly structured (Figure 5C and D).

Peptide Binding to Pre-miR-20b

Electrophoretic mobility shift assays (EMSA) were used to test the interaction between the peptide mimetics and pre-miR-20b RNA. Among the designs, none of the 14- or 18-mer peptides (CPfox1–3) showed any detectable binding (Figure 7 and Supplementary Figure 2) at concentrations of 200–500 μM . This is not in retrospect surprising, once the structures of the peptides were analyzed, since these are poor structural mimetic of the $\beta 2\beta 3$ hairpin of Rbfox2, with distinct cross-strand hydrogen bonding pattern and therefore different displays of amino acid side chains to the RNA. CPfox5 and CPfox6 bind to RNA, but the interactions appears to be non-specific, since it can be significantly reduced when binding assays are carried out in the presence of 250-fold excess of competitor tRNA (Supplementary Figure 3). On the other hand, peptidomimetics CPfox4 and CPfox7 show binding to pre-miR-20b in

the presence of competitor tRNA (Figure 7), albeit with a relatively weak apparent K_d of $\sim 50 \mu\text{M}$ for CPfox4 and weaker for CPfox7. However, EMSA measurements are not accurate at this K_d because of the fast exchange between the bound and free forms of the RNA on the gels; as a result, the band corresponding to the CPfox7/pre-miR-20b complex was broadly distributed in the lane (Supplementary Figure 2). Precipitation was also observed in both CPfox4/pre-miR-20b and CPfox7/pre-miR-20b mixture at peptide concentrations above $100 \mu\text{M}$, leading to reduced band intensity on the gels. It is therefore very likely that the K_d 's are underestimated and the actual binding is stronger than reported by EMSA.

NMR studies of peptide-RNA interaction

The interaction between pre-miR-20b and CPfox4 and CPfox7 were further investigated by NMR titration experiments. 1D ^1H NMR spectra of pre-miR-20b RNA with increasing amounts of CPfox4 or CPfox7 were recorded (Figure 8 and Supplementary Figure 4). Some peaks shifted and broadened when the concentration of peptide increased, confirming binding as observed by EMSAs, while some of the peaks remained unchanged during the titration, indicating that the interaction occurs at a specific site on the RNA.

2D TOCSY spectra of free RNA and the CPfox4/RNA complex (at 2.5:1 ratio) were also recorded (Figure 9). The peak corresponding to the C30 H5–H6 resonances in pre-miR-20b, a nucleotide in the RNA apical loop, shifts significantly after the addition of CPfox4 peptide. The H5–H6 resonance peaks of U25, U26 and U36, which are immediately below the apical loop (Figure 1), shift as well. The result suggests a direct interaction between the peptide and this RNA apical loop.

On the other hand, only weak changes in chemical shifts were found in the TOCSY spectrum of CPfox7/RNA (2.5:1 ratio) mixture (Supplementary Figure 5). Direct interactions between pre-miR-20b apical loop and CPfox7 were not observed by NMR, though this might be due to the more limited solubility and weak affinity of this peptide.

Discussion

RRM protein recognize RNAs through a variety of binding modes^[4, 5b], but most use its large relatively featureless β -sheet surface; thus, it is difficult to mimic RRM/RNA binding with small molecules for the purpose of therapeutic RNA targeting. However, peptide analogues that mimic key protein structural elements can provide specific and potent therapeutic leads; for example, α -helical regions of DNA-binding proteins can be mimicked with stapled peptides^[23], whereby the stapling stabilizes the α -helical structure that is necessary for recognition. Since the $\beta 2\beta 3$ hairpin and its intervening loop often provides critical contacts in RRM-RNA complexes, we reasoned that it might be possible to use cyclic β -hairpin peptides to mimic the binding activity of the much larger complete domain. We used the Rbfox2 RRM and its interaction with the pre-miR-20b RNA as a starting point for design, in part because this protein binds to a structured RNA, which would make it more likely to observe direct binding compared to single stranded RNA targets recognized by other RRMs, including Rbfox2 binding to splicing signals.

Based on the structure of Rbfox2 RRM/pre-miR-20b complex, we designed peptide mimetics of $\beta 2\beta 3$ hairpin, tested their folding with NMR, measured binding affinities by EMSA gel shift assays. The first synthetic peptide mimetic of the $\beta 2\beta 3$ hairpin in Rbfox2 RRM was a straightforward transplant of residues Ile148-Gly159 of Rbfox2 onto a hairpin-inducing $^D\text{Pro-LPro}$ template (Table 1). By pre-structuring the mimetic peptide into a conformation close to that observed the bound complex with the templating amino acids, we expected to favor binding by reducing entropic losses associated with peptide folding from an unstructured state.^[20a] However, the resulting design, CPfox1, does not adopt a β -hairpin conformation, most likely because of the high Gly-content of the sequence, which introduces excessive conformational flexibility.

In order to improve the stabilities of the design, two different strategies were adopted. The first was replacing Gly residues with Arg, which were introduced to improve the β -sheet propensity and its solubility and potentially provide improved binding to RNA. The resulting mimetic CPfox2 forms a stable β -hairpin structures in aqueous solution (Figure 5A), but failed to bind to RNA. Structural analysis revealed that the side chains in this peptide point in a direction opposite to that observed in the Rbfox2 RRM (Figure 6A). This occurs because the ($^D\text{Pro-LPro}$) template forms a type-II' β -turn, which dictates the conformation of the rest of the peptide, making residue 1, 3, 5, 8, 10 and 12 form cross-strand backbone hydrogen bonds. The corresponding residues in Rbfox2 RRM are non-hydrogen-bonding residues instead. When we elongated the β -strand by four amino acids, to provide more cross-strand hydrogen bonds, the resulting CPfox3 peptide was better structured than CPfox1, but has low solubility in water and could not be fully analyzed by NMR.

This situation was also encountered when we first transplanted the sequence of BIV Tat onto the template to mimic the BIV Tat-TAR complex.^[20a] It is necessary to mimic the structure of the peptide, not just its sequence, to observe any binding. Based on the structure of CPfox2, we synthesized 16-mer peptide mimetics rather than 14-mers to restore the hydrogen bonding pattern observed in the complex. It is well-known that the length of cyclic peptide influences the stability of β -hairpin structures.^[24] Cyclic peptides with $4n+2$ residues form more stable β -hairpins compared to those with $4n$ residues. Thus, it is not surprising that CPfox4–7, 16-mer cyclic peptides, are all less structured than the 14-mer CPfox2 from which they originate. However, by increasing the peptide length by 2 residues and placing the turn in phase with the correct residues, the side chains of the 16-mer peptides adopt conformations consistent with the $\beta 2\beta 3$ hairpin within Rbfox2 RRM (Figure 6B). Mid-low μM binding to the pre-miR-20b RNA was observed for two of these peptides (CPfox4 and CPfox7), although their K_d could not be accurately measured by EMSA due to the fast exchange of the complex and limited solubility under high peptide concentration ($>100\mu\text{M}$).

CPfox4, with a flat and flexible loop, corresponding to the $\beta 2\beta 3$ loop in the Rbfox2 RRM, binds to pre-miR-20b best among all peptidomimetics. Flexibility can help improve the RNA binding affinity by allowing direct interactions through conformational adaptation.^[25] The flexibility in the peptide loop can thus help improve its binding to RNA, particularly in the initial phases of design when binding is expected to be weak and interactions sub-optimal. Consistent with this hypothesis, CPfox5 and CPfox6, which possess a more

structured loop than CPfox4, only bind to RNA non-specifically (binding activity is abrogated in the presence of excess tRNA). Thus, the structure and binding results support the hypothesis that flexibility in the binding loop facilitates binding. We further used a linear peptide analogue LPfox4 (EIIFNERGSKRFRFP), to prove that a pre-structured β -hairpin is required for the interaction to occur, presenting each amino acid in the correct orientation with respect to the RNA stem-loop. As expected, the linear version of CPfox4 failed to bind to RNA (Supplementary Figure 6).

When the binding between pre-miR-20b and CPfox4 was examined by NMR, gradual peak shifts were observed in 1D ^1H NMR spectra upon addition of the peptides, suggesting that the bound form of pre-miR-20b is in fast exchange with the free form, consistent with the mid- μM affinity observed by EMSA. In addition, the TOCSY spectra showed that the peaks of four nucleotides located in the apical loop shifts significantly after the addition of CPfox4 (Figure 9). The peak shift of C30 indicates a direct interaction between the RNA apical loop and the peptide also observed in the structure of Rbfox2 RRM bound to pre-miR-20b (Figure 10).

Although CPfox4 is basic, the interaction with RNA is not simply the result of non-specific electrostatic interactions. First of all, the binding affinity of CPfox7 is weaker than CPfox4, although CPfox7 is more positively charged; second, CPfox4 cannot be competed away by tRNA. To further confirm the specificity of CPfox4 binding to pre-miR-20b, an EMSA experiment with CPfox4 and HIV TAR was conducted and no binding affinity was detected (Supplementary Figure 7), despite the RNA's similar size, secondary structure and loop size. Binding between CPfox4 and mutant pre-miR-20b (C30A), and between mutant CPfox4 (cyclo-EIIFNEAGSKRFRFP) and pre-miR-20b were also examined. By mutating the Cytosine in RNA or Arginine in peptide involved in the interacting interface (Figure 10), the CPfox4 and pre-miR-20b binding affinity was erased (Supplementary Figure 8).

The results presented here demonstrate the successful mimicking of a full RNA-binding protein domain with a small cyclic peptide, and adds a second approach, beyond direct evolution, to generate small peptide mimics of relatively large RRM proteins.^[21] We created a specific, although weak binder for the pre-miR-20b apical loop by grafting the sequence of the $\beta 2\beta 3$ loop of Rbfox2 RRM onto a β -hairpin promoting template, and ensuring through the design that the correct amino acid side chains are exposed to the RNA. In the Rbfox2 RRM and in many such proteins, residues within $\beta 1$, the $\beta 1\alpha 1$ loop and the C-terminal loop also interact with RNA, creating a large interface (Figure 1). Thus, it is not surprising that affinity is decreased 3 orders of magnitude, since we have discarded half of the interacting functional groups in the interface and disrupted a stable protein fold. Clearly, this peptide represents a starting not a final point for the successful design of more potent and specific peptides, as we did successfully with the BIV and HIV TAR RNAs, by using structure-based peptide design^[20a-c] or evolution/selection methods.^[21] This work also showed the possibility of structure-based design strategy to discover RNA-binding ligands. By understanding the crucial interactions between a protein and RNA, we can mimic the binding interface with designed peptides. Since an increasing number of protein-RNA structures are becoming available in the Protein Data Bank (PDB),^[26] more RNA-binding peptides could be discovered through a better understanding of protein-RNA interactions.

The discovery of CPfox peptides might provide leads for the development of miRNA maturation inhibitors, and proof of principle that even complex interfaces, as observed between RRM and RNA, can be mimicked by much smaller fundamental peptide structures.

Experimental Section

Peptide Synthesis

Peptides were synthesized by solid phase peptide synthesis using Fmoc chemistry.^[20a, 27] Resin and protected amino acids were purchased from AappTec and P3 Biosystems. Arg(Pbf)-preloaded 2-chlorotrityl resin was used for the synthesis. Linear peptide synthesis was performed on a Liberty Blue™ automated microwave peptide synthesizer with DMF as the solvent. Piperidine (20% in DMF) was used for deprotection, and DIC/Oxyma for coupling. The linear precursor was cleaved from the resin with 5% acetic acid in DCM for 2 hours. Cyclization was performed with 6 equivalents of PyAOP/HOAt and 2,4,6-trimethylpyridine in DMF solution, and progress was monitored by HPLC. The peptide was then treated with TFA:triisopropylsilane:H₂O (95:2.5:2.5) for 2 hours for side-chain deprotection.

Cyclic peptides were purified by reverse phase HPLC with a semi-preparative C18 column. The gradients were from 7.5% to 100% isopropanol in H₂O with 0.1% TFA and 5% to 100% Acetonitrile in H₂O with 0.1% TFA. Each peptide showed a single peak (>95% purity) in HPLC. The identity of all peptides was confirmed with Bruker Esquire Ion Trap mass spectrometry (data shown in Supplementary Material).

RNA Preparation and Binding Assays

Pre-miR-20b RNA (G19 to C41) was prepared by *in vitro* transcription with in-house purified T7 RNA polymerase with a commercially synthetic oligonucleotide template (IDT).^[28] The RNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis (PAGE), electroeluted and concentrated by ethanol precipitation.

In order to prepare labeled RNA for binding assays, purified RNA was 5'-dephosphorylated with Calf Intestinal protein (CIP) then re-phosphorylated with [γ -32P]-ATP and T4 polynucleotide kinase. RNA was annealed and snap cooled before use. Peptide and RNA were incubated in buffer containing 12.5 mM Tris-HCl (pH 7.4), 12.5 mM KCl, 250-fold excess tRNA, and 0.25% Triton X-100 for 30 minutes. The samples were loaded onto 12% native polyacrylamide gel in 0.5X TBE buffer and electrophoresed at 300 V and 4 W for 2 hours. Dried gels were exposed to a phosphor imaging plate and scanned by GE Typhoon FLA 9000 gel imager.

NMR Binding Studies

NMR titration experiments were performed on Bruker DRX 500 MHz or Avance III 600 MHz spectrometer, equipped with cryocooler probes. Free RNA samples were prepared at a concentration of 200 μ M in 10 mM phosphate buffer (pH 6.5, H₂O:D₂O= 9:1). The 1D 1H NMR spectra at different ratios of peptide to RNA (0:1, 0.25:1, 0.5:1, 1:1, 1.5:1, 2:1, 2.5:1)

were collected. The 2D TOCSY spectra were collected for free RNA and peptide/RNA (2.5:1) sample. All spectra were collected at 5 °C unless specified.

NMR Spectroscopy and Structure Determination

Peptide samples at a concentration of 2–5 mg/mL were prepared in H₂O:D₂O (9:1) 10 mM phosphate buffer at pH 6.5. 1D and 2D NMR spectra were collected on Bruker DRX 500 MHz or Avance III 600 MHz spectrometer equipped with cryo-cooled probes. 2D NMR data (TOCSY, NOESY, HSQC) were processed with TOPSPIN (Bruker) and analyzed in NMRFAM-SPARKY.^[29] Two mixing time (100 ms and 300 ms) were used for 2D 1H NOESY spectra to assist quantitative evaluation of internuclear distances.

Peptide structures were calculated with Xplor-NIH using torsional angle dynamics and simulated annealing from a starting peptide template generated with randomized torsion angles ^[30]. Interproton distance restraints were derived from NOE peaks in 2D 1H NOESY spectra and sorted into strong (2.5 ± 0.7 Å), medium (3.5 ± 1.2 Å) and weak (4.5 ± 1.6 Å) bins based on relative peak intensities. For each peptide, 200 structures were calculated and 10 structures with the lowest energy were selected for analysis from the population of converged structures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work is supported by NIH through grants R35GM126942 and RO1 GM103834.

References

- [1]. Chen Y, Zubovic L, Yang F, Godin K, Pavelitz T, Castellanos J, Macchi P, Varani G, Nucleic Acids Res. 2016, 44, 4381–4395. [PubMed: 27001519]
- [2]. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al., Science 2001, 291, 1304–1351. [PubMed: 11181995]
- [3]. Maris C, Dominguez C, Allain FH, FEBS J. 2005, 272, 2118–2131. [PubMed: 15853797]
- [4]. Clery A, Blatter M, Allain FH, Curr. Opin. Struct. Biol 2008, 18, 290–298. [PubMed: 18515081]
- [5]. a)Burd CG, Dreyfuss G, Science 1994, 265, 615–621; [PubMed: 8036511] b)Afroz T, Cienikova Z, Clery A, Allain FH, Methods Enzymol. 2015, 558, 235–278. [PubMed: 26068744]
- [6]. Maris C, Dominguez C, Allain FH, FEBS. J 2005, 272, 2118–2131. [PubMed: 15853797]
- [7]. a)Tintaru AM, Hautbergue GM, Hounsflow AM, Hung ML, Lian LY, Craven CJ, Wilson SA, EMBO Rep. 2007, 8, 756–762; [PubMed: 17668007] b)Clery A, Sinha R, Anczukow O, Corriero A, Moursy A, Daubner GM, Valcarcel J, Krainer AR, Allain FH, Proc. Natl. Acad. Sci. U. S. A 2013, 110, E2802–2811. [PubMed: 23836656]
- [8]. a)Auweter SD, Fasan R, Reymond L, Underwood JG, Black DL, Pitsch S, Allain FH, EMBO J. 2006, 25, 163–173; [PubMed: 16362037] b)Skrisovska L, Bourgeois CF, Stefl R, Grellscheid SN, Kister L, Wenter P, Elliott DJ, Stevenin J, Allain FH, EMBO Rep. 2007, 8, 372–379; [PubMed: 17318228] c)Dominguez C, Fiset JF, Chabot B, Allain FH, Nat. Struct. Mol. Biol 2010, 17, 853–861. [PubMed: 20526337]
- [9]. a)Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL, et al., Science 2005, 309, 2054–2057; [PubMed: 16179478] b)Tsuda K, Kuwasako K, Takahashi M, Someya T, Inoue M, Terada T, Kobayashi N, Shirouzu

- M, Kigawa T, Tanaka A, et al., *Nucleic Acids Res.* 2009, 37, 5151–5166; [PubMed: 19553194]
 c)Clery A, Jayne S, Benderska N, Dominguez C, Stamm S, Allain FH, *Nat. Struct. Mol. Biol* 2011, 18, 443–450. [PubMed: 21399644]
- [10]. a)Krol J, Loedige I, Filipowicz W, *Nat. Rev. Genet* 2010, 11, 597–610; [PubMed: 20661255]
 b)Ha M, Kim VN, *Nat. Rev. Mol. Cell Biol* 2014, 15, 509–524. [PubMed: 25027649]
- [11]. a)Bartel DP, *Cell* 2009, 136, 215–233; [PubMed: 19167326] b)Fabian MR, Sonenberg N, Filipowicz W, *Annu. Rev. Biochem* 2010, 79, 351–379. [PubMed: 20533884]
- [12]. a)Michlewski G, Guil S, Semple CA, Caceres JF, *Mol. Cell* 2008, 32, 383–393; [PubMed: 18995836] b)Newman MA, Thomson JM, Hammond SM, *RNA* 2008, 14, 1539–1549. [PubMed: 18566191]
- [13]. a)Jin Y, Suzuki H, Maegawa S, Endo H, Sugano S, Hashimoto K, Yasuda K, Inoue K, *EMBO J.* 2003, 22, 905–912; [PubMed: 12574126] b)Kuroyanagi H, *Cell. Mol. Life Sci* 2009, 66, 3895–3907. [PubMed: 19688295]
- [14]. a)Zhou W, Shi G, Zhang Q, Wu Q, Li B, Zhang Z, *Cell Biosci.* 2014, 4, 62; [PubMed: 25364498]
 b)Zhu J, Chen L, Zou L, Yang P, Wu R, Mao Y, Zhou H, Li R, Wang K, Wang W, et al., *Hum. Immunol* 2014, 75, 348–353. [PubMed: 24468585]
- [15]. a)Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, Enzo E, Guzzardo V, Rondina M, Spruce T, et al., *Cell* 2010, 141, 1195–1207; [PubMed: 20603000] b)Li Y, Mao L, Gao Y, Baral S, Zhou Y, Hu B, *Sci. Rep* 2015, 5, 13316; [PubMed: 26294080] c)Ristori E, Lopez-Ramirez MA, Narayanan A, Hill-Teran G, Moro A, Calvo CF, Thomas JL, Nicoli S, *Dev. Cell* 2015, 32, 546–560. [PubMed: 25662174]
- [16]. Craik DJ, Fairlie DP, Liras S, Price D, *Chem. Biol. Drug Des* 2013, 81, 136–147. [PubMed: 23253135]
- [17]. a)Croce CM, *Nat. Rev. Genet* 2009, 10, 704–714; [PubMed: 19763153] b)Im HI, Kenny PJ, *Trends Neurosci.* 2012, 35, 325–334. [PubMed: 22436491]
- [18]. Sakamoto K, Otake K, Umemoto T, *Bioorg. Med. Chem. Lett* 2017, 27, 826–828. [PubMed: 28109790]
- [19]. Gongora-Benitez M, Tulla-Puche J, Albericio F, *Chem. Rev* 2014, 114, 901–926. [PubMed: 24446748]
- [20]. a)Athanassiou Z, Dias RL, Moehle K, Dobson N, Varani G, Robinson JA, *J. Am. Chem. Soc* 2004, 126, 6906–6913; [PubMed: 15174860] b)Athanassiou Z, Patora K, Dias RL, Moehle K, Robinson JA, Varani G, *Biochemistry* 2007, 46, 741–751; [PubMed: 17223695] c)Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G, *Proc. Natl. Acad. Sci. U. S. A* 2009, 106, 11931–11936; [PubMed: 19584251] d)Lalonde MS, Lobritz MA, Ratcliff A, Chamanian M, Athanassiou Z, Tyagi M, Wong J, Robinson JA, Karn J, Varani G, et al., *PLoS Pathog.* 2011, 7, e1002038. [PubMed: 21625572]
- [21]. Belashov IA, Crawford DW, Cavender CE, Dai P, Beardslee PC, Mathews DH, Pentelute BL, McNaughton BR, Wedekind JE, *Nucleic Acids Res.* 2018.
- [22]. Price SR, Evans PR, Nagai K, *Nature* 1998, 394, 645–650. [PubMed: 9716128]
- [23]. Frank AO, Vangamudi B, Feldkamp MD, Souza-Fagundes EM, Luzwick JW, Cortez D, Olejniczak ET, Waterson AG, Rossanese OW, Chazin WJ, et al., *J. Med. Chem* 2014, 57, 2455–2461. [PubMed: 24491171]
- [24]. Gibbs AC, Kondejewski LH, Gronwald W, Nip AM, Hodges RS, Sykes BD, Wishart DS, *Nat. Struct. Biol* 1998, 5, 284–288. [PubMed: 9546219]
- [25]. Leulliot N, Varani G, *Biochemistry* 2001, 40, 7947–7956. [PubMed: 11434763]
- [26]. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE, *Nucleic Acids Res.* 2000, 28, 235–242. [PubMed: 10592235]
- [27]. Fields GB, Noble RL, *Int. J. Pept. Protein Res* 1990, 35, 161–214. [PubMed: 2191922]
- [28]. a)Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC, *Nucleic Acids Res.* 1987, 15, 8783–8798; [PubMed: 3684574] b)Beckert B, Masquida B, *Methods Mol. Biol* 2011, 703, 29–41. [PubMed: 21125481]
- [29]. Lee W, Tonelli M, Markley JL, *Bioinformatics* 2015, 31, 1325–1327. [PubMed: 25505092]
- [30]. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM, *J. Magn. Reson* 2003, 160, 65–73. [PubMed: 12565051]

Figure 1. (A) The NMR structure of the complex of Rbfox2 RRM protein and pre-miR-20b RNA. (B) A closer look into the Rbfox2 RRM and pre-miR-20b interface.^[1] (PDB: 2N82). The $\beta 2\beta 3$ loop mimicked in our designed cyclic peptides is colored in magenta in both images.

Figure 2.

Illustration of CPfox1 cyclic peptide design. Residues 148–159 of the Rbfox2 RRM, including the $\beta 2\beta 3$ hairpin, are transplanted onto the $^D\text{Pro-LPro}$ template to generate a cyclic peptide of equivalent structure that presents the same amino acids in the apical loop. Subsequent designs are based on this simple transplant.

Figure 3.
Overlay of 1D NMR spectra of seven cyclic peptide mimics.

Figure 4. HN-Ha and HN-HN regions of the 2D NOESY spectra for CPfox2 (A) and CPfox4 peptides (B). Main chain Ha-HN and HN-HN NOE assignments are shown.

Figure 5. Superposition of the 10 NMR structures of lowest energy for (A) CPfox2, (B) CPfox4, (C) CPfox5, (D) CPfox6 and (E) CPfox7. In each image, the $\beta 2\beta 3$ loop mimic is on the left, and $DPro-LPro$ template on the right.

Figure 6. Superposition of (A) the NMR structure of CPfox2 (lowest energy, green) and Rbfox2 RRM $\beta 2\beta 3$ (grey) and (B) the structure of CPfox4 (lowest energy, cyan) and Rbfox2 RRM $\beta 2\beta 3$ (grey). Amino acid side chains are represented as spheres.

Figure 7.

EMSA of Rbfox2 peptide mimetics binding to pre-miR-20b. Various concentration of peptide and 30 nM RNA were incubated in Tris buffer (pH7.4) containing 250-fold excess tRNA, and 0.25% Triton X-100 for 30 minutes. The samples were loaded onto 12% native polyacrylamide gel in 0.5X TBE buffer and electrophoresed at 300 V and 4 W for 2 hours. Because peptide-RNA solubility is limited above 100 μM , the band intensity decreases above this concentration.

Figure 8.

1D ^1H NMR titration pre-miR-20b RNA with CPfox4. Peptide and RNA ratios are 0:1 (black), 0.25:1 (red), 0.5:1 (dark green), 1:1 (purple), 1.5:1 (brown), 2:1 (orange), 2.5:1 (green). The spectrum of free CPfox4 is showed in blue.

Figure 9. 2D TOCSY spectra of free pre-miR-20b (blue) and CPfox4/pre-miR-20b (2.5:1) mixture (red). Peaks of U25, U26, C30 and U36 shift after the addition of CPfox4.

Figure 10.

The interface between Rbfox2 RRM $\beta 2\beta 3$ loop and pre-miR-20b has C30 in the center, in direct contact with the amino acid side chains of the protein (PDB:2N82).

Table 1.Amino acid sequence of Rbfox2 RRM (residue 146–161) and all 7 CPfox mimetic peptides ^[a],^[b],^[c]

Peptide/Protein	Position																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Rbfox2(146–161)	V	E	I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	K	G	F	G	F	V		
CPfox1			I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	K	G	F	G	p	P		
CPfox2			I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	K	R	F	R	p	P		
CPfox3	V	E	I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	K	G	F	G	F	V	p	P
CPfox4		E	I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	K	R	F	R	F	p	P	
CPfox5		K	I	I	F	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	R	G	F	R	F	p	P	
CPfox6		K	I	I	R	<u>N</u>	<u>E</u>	<u>R</u>	<u>G</u>	S	R	G	F	R	F	p	P	
CPfox7		K	I	I	F	<u>N</u>	<u>Q</u>	<u>R</u>	<u>G</u>	S	R	R	F	R	F	p	P	

^a. All peptides are cyclized following synthesis, as described in the methods section.^b. The lower-case p represents DPro.^c. Residues corresponding to the $\beta 2\beta 3$ loop in the Rbfox2 RRM are underlined.

Table 2.

Experimental restraints and structural statistics for the 10 lowest energy structures of CPfox mimetic peptides.

	CPfox2	CPfox4	CPfox5	CPfox6	CPfox7
No. of NOE-derived restraints	167	131	220	183	158
Intraresidue	85	70	119	113	94
Sequential	51	43	63	43	36
medium range (2 $\leq$ i-j $\leq$ 5)	24	7	27	18	19
long range (i-j $\geq$ 6)	7	11	11	9	9
No. of NOE violation					
number (> 0.2 Å)	0	0	0	0	0
maximum (Å)	0	0	0	0	0
mean RMSD value (Å)					
all heavy atoms	1.89 $\pm$ 0.46	2.30 $\pm$ 0.40	2.08 $\pm$ 0.60	2.26 $\pm$ 0.85	2.16 $\pm$ 0.18
backbone atoms	1.00 $\pm$ 0.34	1.39 $\pm$ 0.21	1.06 $\pm$ 0.38	1.29 $\pm$ 0.49	1.29 $\pm$ 0.22