

Chapter 4

Large Libraries Reveal Diverse Solutions to an RNA Recognition Problem

This work has been adapted from the following publication:

Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T. and Roberts, R.W. Large libraries reveal diverse solutions to an RNA recognition problem. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12374-12378.

Abstract

RNA loops that adopt the characteristic GNRA "tetraloop" fold are common in natural RNAs. We selected peptides that bind an example of this RNA loop motif using mRNA display. Starting with the RNA recognition domain from the λ N protein, we constructed libraries containing 150 and 1,600 different peptide sequences as mRNA-peptide fusions and isolated those capable of high-affinity RNA binding. These selections have resulted in almost 20 different peptides that bind the same RNA loop with nanomolar affinity. Analysis of one peptide complex by fluorescence spectroscopy and NMR spectroscopy suggest a different binding mode from the wild-type peptide. Our work demonstrates that multiple, chemically and conformationally distinct solutions exist for a particular RNA recognition problem.

Introduction

The ability to construct high-affinity, high-specificity peptide ligands provides a means to target RNA molecules of interest. Genetic approaches have been developed to isolate novel arginine-rich RNA-binding peptides *in vivo* (1-3). These systems allow selection in the context of living systems, but limit library sizes to a maximum of $\sim 10^5$ to 10^6 sequences, allowing only four residues to be searched exhaustively ($20^4 = 1.6 \times 10^5$). *In vitro* selection experiments to isolate RNA-binding peptides have not been demonstrated. A totally *in vitro* approach affords precise control over the selection conditions and the ability to explore much larger libraries, enabling the isolation of rare sequences.

We sought to use *in vitro* selection to isolate peptides that bind RNA tetraloops. The tetraloop fold is a common element in many functional RNAs, enhancing duplex stability and participating in tertiary folding interactions (4-6). We developed mRNA display (also known as the mRNA-peptide fusion system) to perform *in vitro* selection of peptides and proteins (7). In mRNA display, cycles are carried out entirely *in vitro* and libraries larger than 10^{13} independent sequences can be constructed (Figure 4.1, (7-10)).

The arginine-rich peptide corresponding to the RNA-binding domain of the λ N protein served as a starting point for our experiments and libraries. This short peptide (22 amino acids) recognizes the *boxB* RNA hairpin with high affinity and specificity as a bent α -helix (11-14). The hairpin contains a five-base RNA loop that adopts a tetraloop fold with one base extruded (15, 16). Here, we have used mRNA display to isolate peptides that bind the *boxBR* RNA motif.

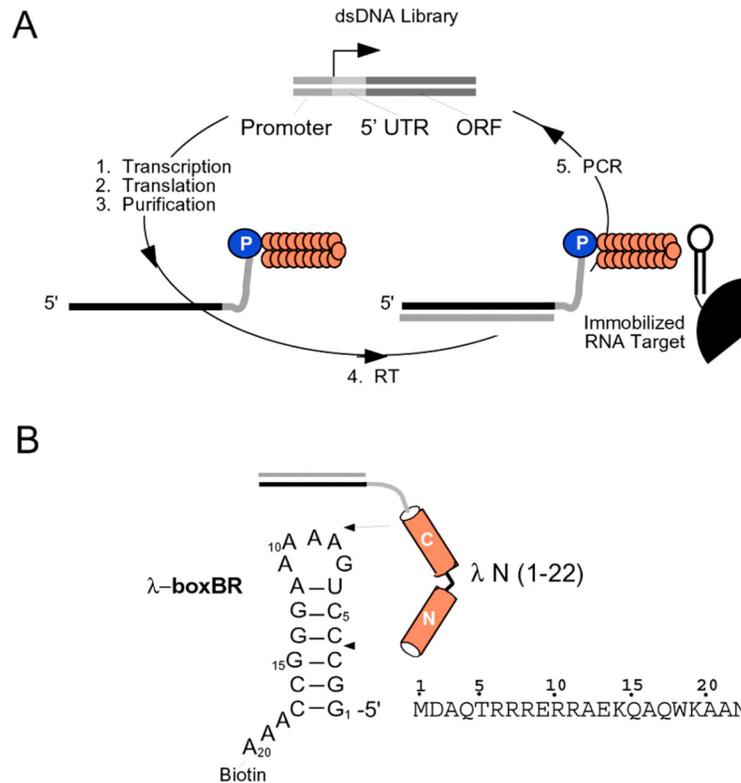


Figure 4.1. (a) Selection cycle. In the selection cycle, a double-stranded DNA library containing randomized codons is transcribed, generating a pool of mRNA templates. These templates are then ligated to a flexible DNA oligonucleotide containing puromycin at its 3' end. Translation of these ligated templates *in vitro* produces peptides covalently attached via their C-terminus to the 3' end of their own message by way of a stable amide linkage a mRNA-peptide fusion. These fusions are then converted into cDNA/mRNA hybrid fusions by using reverse transcriptase and subjected to selection on an affinity matrix. (b) Constructs used. λ N (1-22, amino acid sequence shown) binds to *boxB* as a bent α -helix.

Results and Discussion

Peptide Fusion Binding

It was unclear *a priori* whether an RNA-binding peptide would be functional as either an mRNA-peptide or cDNA/mRNA-peptide fusion. We synthesized mRNA-peptide fusions containing the λ N RNA binding domain (17) (Figure 4.1B) and tested the ability of these molecules to bind an immobilized RNA target. (Figure 4.2) shows

that both the mRNA-peptide fusion and the cDNA/mRNA-peptide fusion specifically bind the *boxBR* RNA target. These assays demonstrate that a significant fraction of the desired complex can be isolated (30-60%) whereas little (1-5%) is retained to noncognate biotinylated RNA targets (Rev Response Element, RRE; U1 RNA hairpin loop II, U1).

In Vitro Selection of λ N Peptides

We designed two selections of increasing complexity to test our ability to isolate novel tetraloop-binding peptides. In selection 1, we randomized positions 6 and 7 of the

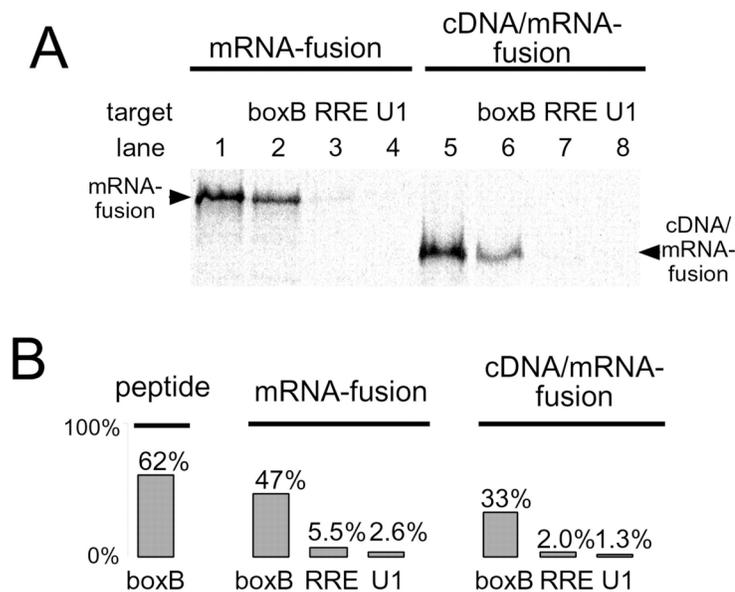


Figure 4.2. Binding and specificity of N peptide N mRNA-peptide fusion, and N cDNA/mRNA-peptide fusion constructs. **(a)** Gel analysis of binding. Binding of ^{35}S -Met-labeled N mRNA-peptide fusions and cDNA/mRNA-peptide fusions to immobilized (i) *boxBR* (lanes 2 and 6), (ii) HIV-RRE (lanes 3 and 7) and (iii) U1 hairpin (lanes 4 and 8). Lanes 1 and 5 show the total amount of fusion before binding. Complexes were eluted and resolved on SDS tricine PAGE (27). **(b)** Scintillation analysis of binding and specificity. ^{35}S -Met-labeled peptide, mRNA, and mRNA/cDNA fusions bound to immobilized *boxBR*, HIV-RRE, and U1 hairpins. The cDNA/mRNA and mRNA fusions bind the cognate *boxBR* RNA efficiently, whereas only 1-5% bind the RRE and U1 targets.

wild-type construct (R6X, R7X, henceforth R67X). Previous mutagenesis and structural work indicated a single sequence, containing arginine at both positions (R6 and R7, henceforth R6R7), should be optimal for binding (1, 13). The cassette contained 150 different amino acid combinations, only one of which was R6R7, and was designed such that the R6R7 sequence could be recognized and cut by the restriction enzyme NgoM IV (Figure 4.3). We performed selection 1 (one round), three times in parallel, using increasing concentrations of competitor yeast tRNA to provide high stringency selection. Selection under stringent conditions is expected to increase the enrichment of functional sequences (18-20).

Restriction analysis by NgoM IV indicated that a majority (>85%) of the selected clones contained R6R7 when 5 mg/mL tRNA competitor was used (Figure 4.3). Sequencing of clones selected under these conditions showed that 8/11 (72%) contained R6R7 (Figure 4.4) and 50% of other sequences contained R7, expected to be the more important of the two positions (1). Thus, increased competitor allowed good enrichment (50- to 150-fold per round) without forcing us to make the target the limiting reagent (21).

In selection 2 (Figure 4.5A, library 2) we randomized positions 7, 14, and 15, for a total of 1,600 possible combinations. In the wild-type N/*boxBR* complexes, residues 14 and 15 lie at the RNA-peptide interface (15, 16) and are important for binding (1, 14). After four rounds of selection, pool binding was similar to the wild-type sequence (Figure 4.5B). Analysis of 39 round 4 clones yielded 19 different sequences (Figure 4.5C). Each

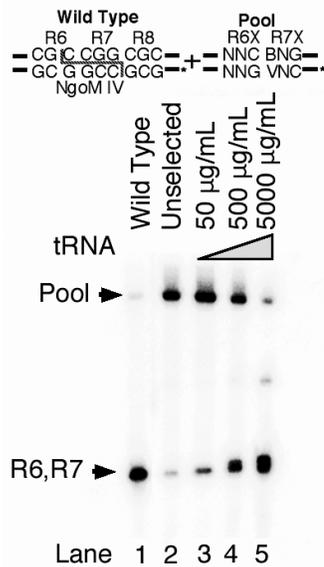


Figure 4.3. Restriction analysis of the R67X Pool. The R67X Pool (Lane 2) was subjected to one round of *in vitro* selection in the presence of increasing concentrations of competitor yeast tRNA (Lanes 3-5). DNA was PCR amplified with a ^{32}P -labeled primer and cut with NgoM IV. Only the wild-type sequence (R6R7) is cut by NgoM IV (Lane 1). Under the highest concentration of competitor tRNA (5 mg/mL) almost all (>85%) of the recovered sequences are cut by NgoM IV, and therefore correspond to the wild-type sequence.

Clone	Peptide Sequence			
	1	10	20	30
wt	•	•	•	•
Pool		XX		
1		RR		
2		FR	R	
3		RR		
5		RR		
6		ASL		
7		RR		
8		RRAGPVEGRQRLQGR	.Q	
9		RR		
11		GG	H	
G1		RR		
B2		RR	G	E

Figure 4.4. Sequencing of the R67X pool after one round of selection with 5 mg/mL tRNA competitor. Sequences corresponding to the wild-type sequence R6R7 are shown in blue. For clarity, only amino acids differing from the wild-type sequence are shown.

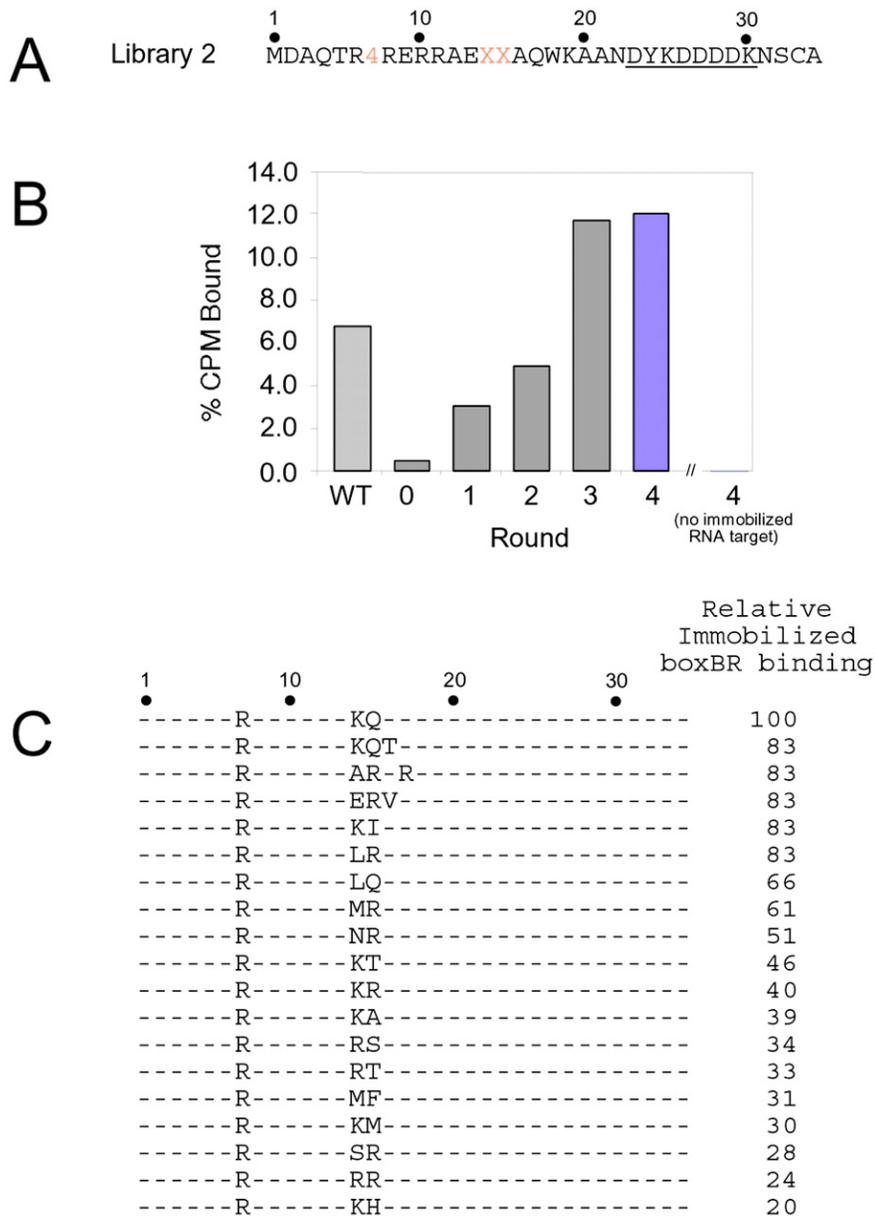


Figure 4.5. Selection of *boxBR* binding peptides from library 2. **(a)** Sequence of library 2. The target is the *boxBR* RNA (Figure 4.1B). Peptide position 7 contains a cassette encoding 4 amino acids, whereas positions 14 and 15 encode all 20 possible amino acids and one stop codon (NNG/C codons). **(b)** Binding to immobilized *boxBR* RNA for rounds 0-4 measured by using ^{35}S -Met-labeled fusions. Controls using the wild-type complex (WT) and agarose with no immobilized RNA (no *boxBR*) are shown. Material from the fourth round (colored bar) was cloned and sequenced. **(c)** Peptide sequences and peptide-binding affinities of selected clones. Binding affinities are given relative to the wild-type sequence in the same construct.

peptide was constructed by *in vitro* translation, and assayed for binding to immobilized *boxBR* RNA (Figure 4.5C) by means of a binding depletion assay (Figure 4.S3). All of the corresponding peptides bind *boxBR* and contain arginine at position 7. Aside from the wild-type (K14Q15), only three of the highest-affinity clones retain one of the two original residues. Indeed, a different consensus sequence containing R15 emerges as a dominant motif; roughly half of the selected sequences contain R15. These sequences include E14R15, which replaces lysine-14 with glutamic acid, the consensus sequence in a selection randomizing residues 13-22 of λ N (22).

Characterization of the E14R15-*boxB* Complex

Peptide affinities were determined by fluorescence titration against 2-aminopurine (2AP) labeled *boxBR*. 2-aminopurine is a fluorescent nucleoside analog of adenine that has been previously substituted in RNA (23). Peptides corresponding to wild-type λ N(1-22) and E14R15(1-22) were constructed synthetically and titrated against *boxBR* labeled at position 7 (2AP-7). Both peptides bind with high affinity to the target site (1.0 nM vs. 8.5 nM for wild-type and E14R15, respectively, Table 4.1). The difference is small considering that a single change of K14 to A14 results in loss of more than 2 kcal/mol (>30-fold) (13). The charge reversal of glutamic acid for lysine at position 14 is puzzling and may result from favorable interaction of glutamic acid with the helix dipole of residues 12-22 (24, 25).

Addition of wild-type N(1-22) produces strong quenching of 2AP fluorescence, whereas addition of E14R15 enhances fluorescence (Figure 4.6). In previous work, 2AP

quenching has been attributed to increased stacking, while fluorescence enhancement is attributed to increased exposure to aqueous solvent (23). Quenching of 2AP in the N (1-22)/*boxBR* complex is expected, as W18 stacks on position 7 (15, 16) and tryptophan fluorescence is quenched in the complex (13), and see below).

Table 4.1. Binding constants and specificity of selected and control peptides. K_d values are in nM.

Peptide	<i>boxBR</i> (2AP-7)
λ N(1-22)	1.0 ± 0.2
E14R15(1-22)	8.5 ± 2.0
Controls	
λ N(1-11)	$1,290 \pm 20$
E14R15(1-15)	303 ± 8

Binding constants were determined by fluorescence titration at 20 °C, 50 mM KOAc, 20 mM TrisOAc, pH 7.5. All peptides contain a free amino and carboxy terminus. λ N(1-11) and E14R15(1-15) contain a C-terminal GY sequence to facilitate quantitation. 2AP-7 denotes a 2'-methoxy 2-aminopurine residue inserted at the second loop position. Error estimates indicate the precision of individual fits. Data for ERV for selection 2 indicates the binding is similar to E14R15(1-22) (unpublished observation).

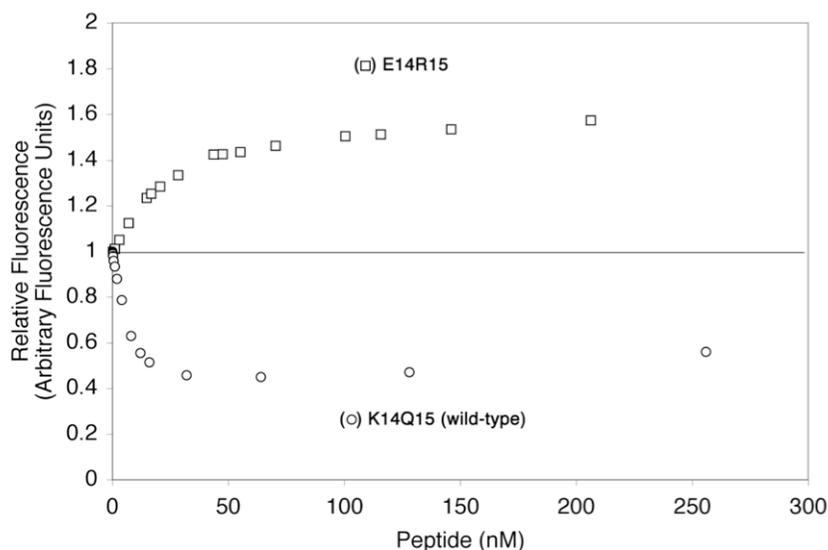


Figure 4.6. Binding isotherms for K14Q15 (wild-type) N peptide (circles) and E14R15 peptide (squares). The E14R15 peptide was added to 2AP-7 labeled *boxB* RNA resulting in an increase in fluorescence. Addition of wild-type peptide to 2AP-7 labeled RNA results in an initial decrease, followed by a gradual increase in fluorescence because of tryptophan fluorescence.

We also investigated tryptophan fluorescence in the E14R15 complex. Figure 4.7 shows that addition of *boxB* RNA to wild-type peptide results in a dramatic decrease in W18 fluorescence, as reported previously (13). However, addition of *boxB* RNA to the E14R15 peptide results in an increase of W18 fluorescence by $\sim 10\%$ relative to the unbound peptide. Taken together, these data imply that W18 is less stacked on adenine 7 in the E14R15 complex.

One- and two-dimensional NMR of the N/*boxBR* complexes provide information about the structure of both the bound peptide and RNA. Binding of the wild-type peptide results in a ~ 1 p.p.m. upfield shift of the Trp indole proton relative to the unbound peptide because of ring current effects in the stacked state (15, 16). However, in the E14R15 complex, W18 is unshifted, indicating that it is not stacked in the complex (Figure 4.8); the chemical shift of the E14R15 W18 indole proton is similar to that of the

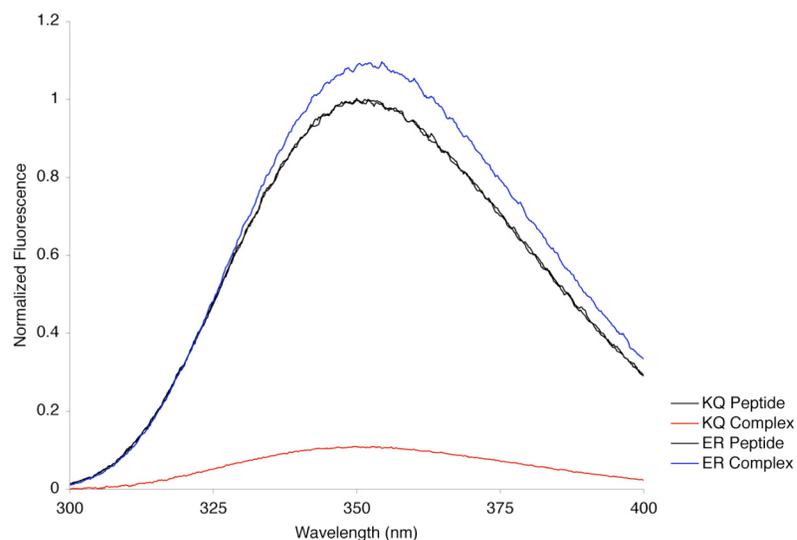


Figure 4.7. Tryptophan fluorescence of N peptide and E14R15 peptide free and in complex with *boxB* RNA. Addition of *boxB* RNA to wild-type N peptide results in a dramatic quenching of Trp fluorescence (red line). Addition of *boxB* RNA to E14R15 peptide results in a slight increase in Trp fluorescence (blue line).

free unbound peptide. Both the wild-type and E14R15 complexes show the same number of imino protons, including the G6-A10 sheared pair imino at ~ 10.8 p.p.m. (15, 16) implying that the RNA structure is similar in both complexes. The ^{15}N -HSQC of the free E14R15(A16V) peptide is similar to the free wild-type peptide (Figure 4.9). Both spectra show poor chemical dispersion in the proton dimension, characteristic of unfolded peptides. Addition of unlabeled *boxBR* RNA to E14R15 results in peak shifts and increased chemical shift dispersion in the proton dimension, suggesting a folded structure (26). The Ala3 amino proton is shifted downfield to ~ 9.3 p.p.m. in the E14R15 complex, suggesting that the N-terminal domain of the peptide binds in a manner similar to the wild-type.

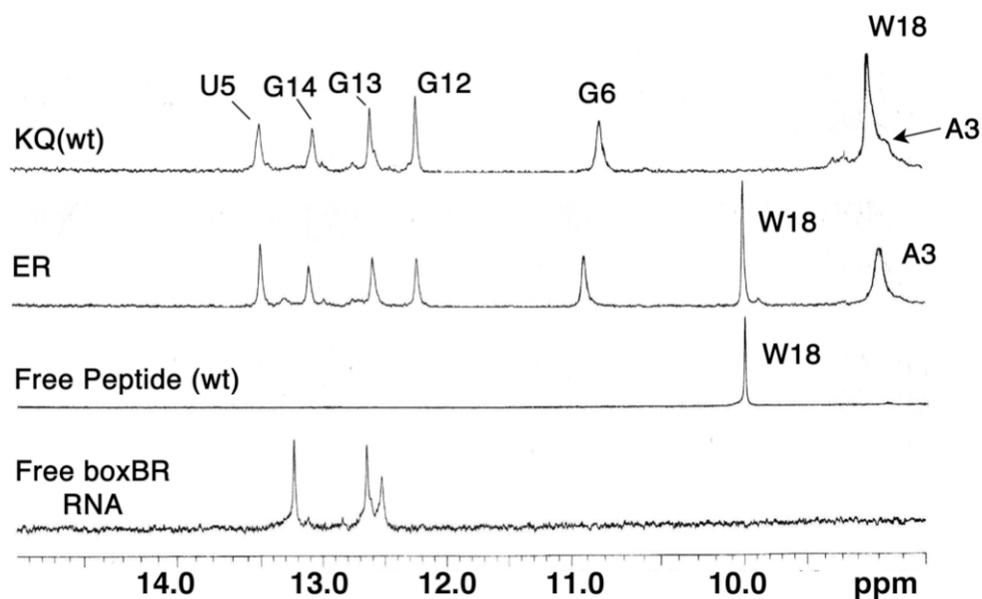


Figure 4.8. ^1H -NMR spectra of wild-type and ER peptides in complex with *boxB* RNA. Both complexes show similar patterns of imino spectra. Free wild-type peptide and *boxB* are shown as controls.

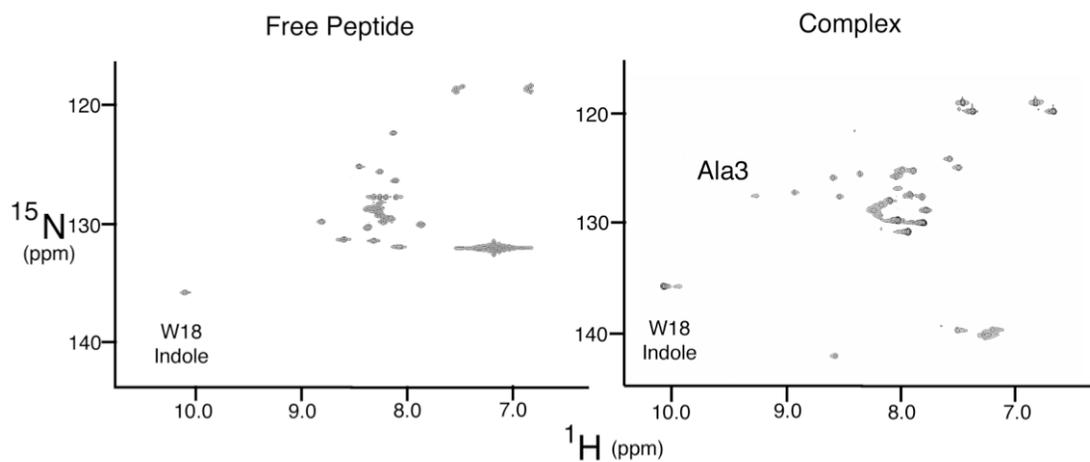


Figure 4.9. ^{15}N -HSQC of the E14R15(V16) peptide alone (left) and in complex with *boxB* RNA (right). Upon addition of the mRNA increased chemical shift dispersion in the ^1H dimension is observed, consistent with the peptide folding. Highlighted are the W18 indole and Ala3 amino peaks.

Neither peptide shows any appreciable structure in the absence of RNA as judged by circular dichroism (CD) (Figure 4.10). The difference spectra of the complexes indicate that both peptides fold into α -helices when *boxBR* RNA is added (Figure 4.10). Although globally similar, the two complexes display differences in regions indicative of peptide folding (200-225 nm) and RNA folding (260-300 nm). E14R15 is somewhat less helical than the wild-type peptide.

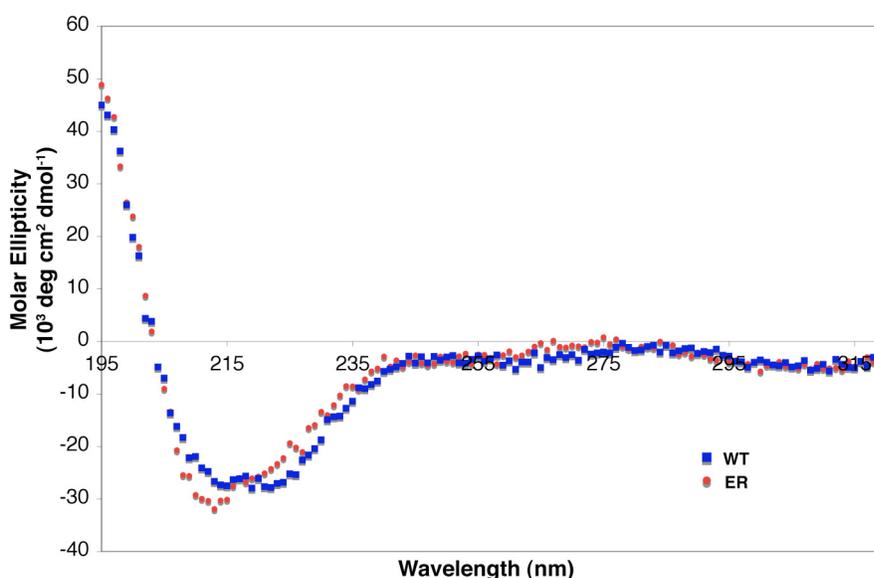


Figure 4.10. CD spectra of wild-type N and E14R15 peptides in complex with *boxB* RNA.

The biophysical data described above suggest that the E14R15 complex with *boxB* is similar in many respects to the wild-type complex. First, the E14R15 peptide binds with near wild-type affinity and folds into an α -helix upon RNA binding. In the bound structure, the RNA and the N-terminal region of E14R15 are structurally similar to the wild-type. However, the tryptophan in the E14R15 complex does not stack on the RNA, and therefore the C-terminus of E14R15 is most likely in a different conformation

than in the wild-type structure. This change in structure could disrupt native contacts between residues 16-22 of the E14R15 peptide and the RNA; the loss of the W18-RNA stacking interaction results in a loss of more than 2 kcal/mol of binding free energy (13).

We synthesized two model peptides, N (1-11) and E14R15(1-15), to address the role of the residues between positions 13 and 22. The E14R15(1-15) peptide reveals that addition of R15 cannot be the sole feature that confers low nanomolar binding to the selected peptides. Adding the sequence containing E14R15 to the N (1-11) peptide confers only a 3- to 5-fold enhancement in K_d , a G° of 0.7 to 1.0 kcal/mol (Table 4.1 E14R15(1-15)). This result indicates that the amino acids between positions 16 and 22 contribute ~ 3 kcal/mol of binding free energy to the interaction.

Conclusions

Our results demonstrate that high-affinity RNA ligands may be isolated using mRNA display. Starting with the RNA binding domain of λ N, we isolated nearly 20 peptides that retain the ability to bind the *boxB* RNA. A non-wild-type consensus, containing arginine at position 15, emerges in the selected peptides leading to the question of why nature did not select for R15 in the λ N peptide. Perhaps the E14R15 peptide suggests a potential answer: while the E14R15 peptide possesses similar affinity to the wild-type, W18 is not stacked in the complex suggesting a different structure in the C-terminus of the peptide. It is interesting that a change of only two amino acids results in such a dramatic change in structure, and this structural change has little effect on binding affinity. Future experiments will address several of these issues. These data are consistent with a model where RNA-peptide interactions are highly context dependent.

In that vein, the combinatorial approach we have used provides a powerful design strategy for interactions where rules may be idiosyncratic at best.

Materials and Methods

Construction of Fusion Template

Templates and constructs used to assay fusion binding (Figure 4.2) correspond to the N-myc construct described previously (17).

Binding Analysis

Chemically synthesized 3' biotinylated RNA oligonucleotides of the *boxBR* hairpin (5'-GGCCCUGAAAAGGGCCAAA-biotin 3', Figure 4.1B), HIV Rev RRE binding site (5'-GGUCUGGGCGCAGCGCAAGCUGACGGUACAGGCCAAA-biotin 3'), and U1 stem-loop (5'-AAUCCAUUGCACUCCGGAUUAAA-biotin 3') were immobilized on streptavidin agarose beads. Samples were added to 400 μ L of N binding buffer (10 mM HEPES pH 7.5, 0.5 mM EDTA, 100 mM KCl, 1 mM MgCl₂, 1 mM DTT, 0.01% (v/v) Nonidet P-40, 10% glycerol (v/v)) containing 50 μ g/mL yeast tRNA (Roche) and 200 pmol of immobilized RNA. This mixture was incubated at 4°C for one hour and then washed in a filter centrifuge tube with binding buffer without tRNA. Samples were eluted with either 2X tricine sample buffer (0.1 M Tris-HCl, pH 6.8, 24% glycerol (v/v), 8% SDS (w/v), 0.2 M DTT, 0.02% Coomassie blue G-250 (w/v)) or 100 mM MgCl₂.

Library Construction

Library 1 was constructed beginning with the N-FLAG-myc synthetic single-stranded DNA (5'-GGGACAATTACTATTTACAATTACAATGGACGCCAGACCNNCBNGCGCGAGCGCAGGGCCGAGAAGCAGGCCAGTGGAAGGCCGCAACGACTAC

AAGGACGACGATGACAAG-3', (17)) that was amplified using the Fmyc primer (5'-AGCGCAAGAGTTCTTGTCATCGTCGTCCTTGTAGTC-3'), and 42.108 primer (5'-TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA-3'). Library 1 contains the sequence 5' NNCBNG 3' (N = A, T, G, or C; B = C, G, or T) at codons 6 and 7 that encodes the peptide MDAQTX₆X₇RERRAEKQAQWKA ANDYKDDDDKNSCA. The random position X₆ encodes the amino acids ACDFGHILNPRSTVY while the random position X₇ encodes AEGLPQRSVW. The single nucleotide sequence that encodes R6R7 encodes an overlapping NgoM IV restriction site. Library 2 was constructed as library 1 using a ssDNA template (5'-GGGACAATTACTATTTACAATTACAATGGACGCCAGACCCGCCNGCGCGAGCGCAGGGCCGAGNNSNNSGCCAGTGGAAGGCCGCCAACGACTACAAGGACGACGATGACAAG-3'). Library 2 contains 5' CNG 3' at codon 7 and the sequence NNS (S = G, C) at codons 14 and 15. This combination encoded L, P, Q, or R at position 7 and all 20 aa at positions 14 and 15. Arginine at position 7 encodes an NgoM IV restriction site. Round 0 pools for library 1 and 2 were sequenced (see below) and are shown in **Figure 4.S1** and Figure, respectively.

Selection Experiments

The library 1 and library 2 selection rounds were carried out essentially as described (17), except that the samples were desalted by ultrafiltration (YM-30, Millipore) and not gel purified after transcription.

Restriction Digest and Quantitation

λ N wild-type template (5'-GGGACAATTACTATTTACAATTACAATGGACGCC CAGACCCGCCGGCGCGAGCGCAGGGCCGAGAAGCAGGCCCCAGTGGAAAGGCC GCCAACGACTACAAGGACGACGATGACAAG-3') and aliquots of the unselected and selected samples were amplified with 7 cycles of PCR using the 42.108 primer (5' TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA 3') and ³²P-labeled 24.108 (5' CTTGTCATCGTCGTCCTTGTAGTC 3'). The products were doubly purified using the Qiaquick PCR purification kit (Qiagen), digested with NgoMIV (NEB) for one hour, and then run on a 10% denaturing PAGE followed by quantitation by PhosphorImager (Molecular Dynamics).

Quantitation by Affinity Precipitation

Peptides for each clone were produced by ³⁵S-Met-labeled *in vitro* translation of gel purified mRNA under conditions identical to fusion synthesis (17). Ten microliters of crude lysate were mixed with 1 mL 200 nM immobilized *boxBR* in N binding buffer, washed five times with N-binding buffer and eluted with RNase A (Roche Molecular Biochemicals). The eluted peptides were run adjacent to a sample of the translation reaction on a tricine-SDS polyacrylamide gel (27), and the percentage of peptide bound was determined by PhosphorImager quantitation (Molecular Dynamics) of the respective gel bands. A representative gel is shown in Figure.

Fluorescence-Binding Measurements

Peptides were constructed by means of automated synthesis and fluorenylmethoxycarbonyl (Fmoc) or t-butyloxycarbonyl (tBoc) monomers. Crude peptides were purified as single peaks by means of reversed-phase HPLC and the identity was confirmed by MALDI-TOF mass spectrometry. RNA hairpins containing 2-aminopurine at the second loop position (denoted 2AP7, Table 4.1; substitution for A7) were constructed by automated RNA synthesis by using 2'-O-methyl 2-aminopurine phosphoramidite (Glen Research, Sterling, VA).

Fluorescence measurements were performed essentially as in Lacourciere et al. (23), with excitation and emission wavelengths of 310 and 370 nm, respectively. Concentrated peptide was added stepwise to a stirred solution of 20 nM to 800 nM 2AP-7 RNA hairpin and the temperature was maintained at 20°C.

CD Spectroscopy

Spectra were taken on an Aviv 62 DS CD spectrometer at 25°C. The samples contained 10 μ M RNA and 10 μ M peptide in 10 mM potassium phosphate buffer (pH 7.9). The spectra of the bound peptides were determined by subtracting the free RNA and excess free peptide spectra from the spectra of the complex.

NMR Sample Preparation

BoxBR RNA 5'-GCCUGAAAAAGGGC-3' (15-mer) was synthesized by *in vitro* transcription by using T7 RNA polymerase. The RNA was purified by 20% urea-PAGE, desalted on a NAP column (Amersham Pharmacia), freeze-dried, and resuspended in

NMR buffer (10 mM phosphate, pH 6, 50 mM NaCl) in H₂O/D₂O (90:10 (v/v)). Complexes between the wild-type N (1-22) or E14R15 (1-22) and *boxBR* RNA were generated by addition of concentrated (10 mM) peptide to *boxBR* RNA (280 μM) with the stoichiometry monitored by inspecting the imino-proton spectra. The final sample concentrations were 250 μM for the free RNA and 280 μM for both RNA and peptide in the complexes.

NMR Spectroscopy

NMR spectra were collected at 25°C on a Varian INOVA 600-MHz spectrometer. A modified double gradient echo Watergate solvent-suppression pulse sequence was used to suppress the solvent peak (28). Assignments were based on reported work (13, 15, 16).

Sequencing

N pool template, unselected, and material selected in 100X tRNA binding buffer were amplified with 7 cycles of PCR using the 42.108 and 24.108 primers and purified on a 4% agarose gel and extracted using the Qiaquick gel extraction kit (Qiagen). NgoMIV digestion was performed on a portion of the selected material and the uncut material was extracted and purified similarly. TA cloning was performed on the purified PCR products using the TOPO cloning kit for sequencing (Invitrogen). Single colonies were grown overnight in 5 mL of LB broth containing 100 μg/mL Ampicillin, purified using the Qiaprep Spin Miniprep kit (Qiagen) and sequenced.

Binding to BoxB RNA

N mRNA was translated in rabbit reticulocyte lysate in the presence of ^{35}S -methionine (NEN) for one hour. Five μL of translation extract was taken and incubated for one hour at 4 °C with 200 pmol of BoxB RNA, 20 μL of 50% streptavidin-agarose in 500 μL of N Binding Buffer. The beads were washed three times with 1 mL Binding buffer and the remaining beads counted by liquid scintillation counting.

Phosphorylation of Oligonucleotides

Reactions typically consisted of ~60 pmol of ($\gamma^{32}\text{-P}$)-ATP (NEN), 60 pmol of primer, and 1 μL (10 U) of polynucleotide kinase (NEB) in 25 μL of 1X kinase buffer (70 mM Tris-HCl, pH 7.6, 10 mM MgCl_2 , 5 mM DTT). Samples were incubated at 37 °C for 30 minutes followed by 10 minutes at 95 °C to kill the enzyme.

Supplemental Information

Clone	Peptide Sequence			
	1	10	20	30
wt	•	•	•	•
Pool	-----XX-----			
8	-GA-----			
10	-----GR-----			
11	-----RV-----			
12	-----PA-----			
13	-----D.-----S-----			
15	-----PG-----			
B	-----IA-SAGPSSRPSGRPPTTTRTMT			
D	-----AG-----			
G3	-----GL-----GPVEGRQLQGRR.Q			
B1	V----VG-----			
B3	-----FG-----N-			
B6	-----RL-----			

Figure 4.S1. Round 0 sequences from Selection 1 (R67X). Pool DNA was PCR amplified, T/A cloned, and sequenced (See Materials and Methods). The FLAG epitope tag is underlined. Stop codons are denoted by a red period.

Clone	Peptide Sequence			
	1	10	20	30
wt	•	•	•	•
Pool	-----4-----XX-----			
1	-----R-----LY-----			
2	-----Q-----QM-----			
4	-----R-----YS-----			
5	-----P-----F.-----			
7	-----RS-----AS-----			
13	-----P-----LT-----			
14	-----P-----AC-----			
16	-----R-----QS-----			
17	-----P-----.P-----			

Figure 4.S2. Round 0 sequences from Selection 2. Pool DNA was PCR amplified, T/A cloned, and sequenced (See Materials and Methods). The FLAG epitope tag is underlined. Stop codons are denoted by a red period. Position 7 encodes one of four amino acids (L, P, Q or R).

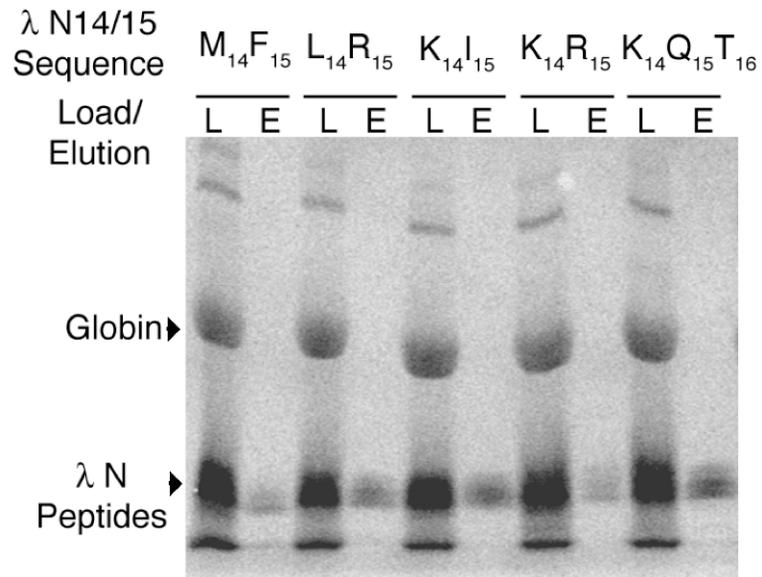


Figure 4.S3. Binding depletion of N peptides. N peptide variants were *in vitro* translated and bound to immobilized *boxB*. The beads were washed and the bound material eluted and run on a tricine gel. Crude translation extract (Load) is run on an adjacent lane to estimate the amount of peptide added to the reaction.

References

1. Franklin, N.C. Clustered arginine residues of bacteriophage lambda N protein are essential to antitermination of transcription, but their locale cannot compensate for *boxB* loop defects. (1993) *J Mol Biol* **231**, 343-360.
2. Harada, K., Martin, S.S. and Frankel, A.D. Selection of RNA-binding peptides *in vivo*. (1996) *Nature* **380**, 175-179.
3. Jain, C. and Belasco, J.G. A structural model for the HIV-1 Rev-RRE complex deduced from altered-specificity rev variants isolated by a rapid genetic strategy. (1996) *Cell* **87**, 115-125.
4. Heus, H.A. and Pardi, A. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. (1991) *Science* **253**, 191-194.
5. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Szewczak, A.A., Kundrot, C.E., Cech, T.R. and Doudna, J.A. RNA tertiary structure mediation by adenosine platforms. (1996) *Science* **273**, 1696-1699.
6. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Doudna, J.A. Crystal structure of a group I ribozyme domain: principles of RNA packing. (1996) *Science* **273**, 1678-1685.
7. Roberts, R.W. and Szostak, J.W. RNA-peptide fusions for the *in vitro* selection of peptides and proteins. (1997) *Proc Natl Acad Sci U S A* **94**, 12297-12302.
8. Roberts, R.W. Totally *in vitro* protein selection using mRNA-protein fusions and ribosome display. (1999) *Curr Opin Chem Biol* **3**, 268-273.

9. Liu, R., Barrick, J.E., Szostak, J.W. and Roberts, R.W. Optimized synthesis of RNA-protein fusions for *in vitro* protein selection. (2000) *Methods Enzymol* **318**, 268-293.
10. Cho, G., Keefe, A.D., Liu, R., Wilson, D.S. and Szostak, J.W. Constructing high complexity synthetic libraries of long ORFs using *in vitro* selection. (2000) *J Mol Biol* **297**, 309-319.
11. Lazinski, D., Grzadzielska, E. and Das, A. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. (1989) *Cell* **59**, 207-218.
12. Tan, R. and Frankel, A.D. Structural variety of arginine-rich RNA-binding peptides. (1995) *Proc Natl Acad Sci U S A* **92**, 5282-5286.
13. Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeets, L.A. and Weiss, M.A. RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda. (1997) *Biochemistry* **36**, 12722-12732.
14. Cilley, C.D. and Williamson, J.R. Analysis of bacteriophage N protein and peptide binding to *boxB* RNA using polyacrylamide gel coelectrophoresis (PACE). (1997) *Rna* **3**, 57-67.
15. Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/*boxB* RNA complex: recognition of a GNRA fold by an arginine-rich motif. (1998) *Cell* **93**, 289-299.
16. Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S. and Rosch, P. Antitermination in bacteriophage lambda. The structure of the N36 peptide-*boxB* RNA complex. (2000) *Eur J Biochem* **267**, 2397-2408.

17. Barrick, J.E., Takahashi, T.T., Balakin, A. and Roberts, R.W. Selection of RNA-binding peptides using mRNA-peptide fusions. (2001) *Methods* **23**, 287-293.
18. Rebar, E.J., Greisman, H.A. and Pabo, C.O. Phage display methods for selecting zinc finger proteins with novel DNA-binding specificities. (1996) *Methods Enzymol* **267**, 129-149.
19. Boder, E.T. and Wittrup, K.D. Optimal screening of surface-displayed polypeptide libraries. (1998) *Biotechnol Prog* **14**, 55-62.
20. Greisman, H.A. and Pabo, C.O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. (1997) *Science* **275**, 657-661.
21. Roberts, R.W. and Crothers, D.M. Specificity and stringency in DNA triplex formation. (1991) *Proc Natl Acad Sci U S A* **88**, 9397-9401.
22. Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T. and Roberts, R.W. Large libraries reveal diverse solutions to an RNA recognition problem. (2001) *Proc Natl Acad Sci U S A* **98**, 12374-12378.
23. Lacourciere, K.A., Stivers, J.T. and Marino, J.P. Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy. (2000) *Biochemistry* **39**, 5630-5641.
24. Richardson, J.S. and Richardson, D.C. Amino acid preferences for specific locations at the ends of alpha helices. (1988) *Science* **240**, 1648-1652.
25. Dasgupta, S. and Bell, J.A. Design of helix ends. Amino acid preferences, hydrogen bonding and electrostatic interactions. (1993) *Int J Pept Protein Res* **41**, 499-511.

26. Dahiyat, B.I. and Mayo, S.L. Probing the role of packing specificity in protein design. (1997) *Proc Natl Acad Sci U S A* **94**, 10172-10177.
27. Schagger, H. and Von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. (1987) *Anal Biochem* **166**, 368-379.
28. Liu, M., Mao, X., Ye, C., Huang, H., Nicholson, J.K. and Lindon, J. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. (1998) *J. Magn. Res.* **132**, 125-129.