

Designed Arginine-Rich RNA-Binding Peptides with Picomolar Affinity

Ryan J. Austin, Tianbing Xia, Jinsong Ren, Terry T. Takahashi, and Richard W. Roberts*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received April 19, 2002

Arginine-rich peptide motifs (ARMs) capable of binding unique RNA structures play critical roles in transcription, translation, RNA trafficking, and RNA packaging.¹ Bacteriophage ARMs necessary for transcription antitermination² bind to distinct boxB RNA hairpin sequences with a characteristic induced α -helical structure.³ Characterization of ARMs from lambdoid phages reveals that the dissociation constant (K_d) of the P22 bacteriophage model-antitermination complex (P22_{N21}-P22boxB) is 200 ± 56 μ M in solution at physiologic concentrations of monovalent cation, significantly stronger than previously determined by gel mobility shift (GMSA)⁴ and polyacrylamide gel coelectrophoresis (PACE),⁵ and 2 orders of magnitude stronger than the tightest known native ARM-RNA interaction at physiological salt.⁶ Here, we use a reciprocal design approach to enhance the binding affinity of two separate α -helical ARM-RNA interactions, one derived from the native λ phage antitermination complex and a second isolated using mRNA display selection experiments targeting boxB RNA.⁷

Steady-state K_d 's over a range of salt concentrations were determined by monitoring the change in fluorescence of a 2-aminopurine (2AP) base analogue substituted at different adenine positions within the loop of the P22boxB RNA target.⁸ At physiologic concentrations of monovalent salt (150 mM [M^+]) the P22 cognate complex is 100 times (2.7 kcal mol⁻¹) more stable than its counterpart antitermination complex in λ phage (λ_{N22} - λ boxB). Kinetic data from stopped-flow measurements indicate that greater stability of the P22_{N21}-P22boxB complex relative to that of the λ_{N22} - λ boxB complex results from a slower rate of dissociation (k_{off}) (P22_{N21}-P22boxB: $k_{off} = 0.01$ s⁻¹; λ_{N22} - λ boxB: $k_{off} = 0.7$ s⁻¹) (Figure 1), while the rate of association (k_{on}) is consistent with a diffusion-controlled process in both complexes (P22_{N21}-P22boxB, λ_{N22} - λ boxB: $k_{on} \approx 7 \times 10^8$ s⁻¹ M⁻¹).

The NMR solution structures of P22 and λ antitermination complexes reveal adaptation of similar but distinct RNA hairpin folds, extruding either the third base pyrimidine or fourth base purine to generate a highly stable GNRA-fold⁹ in the pentaloops of P22boxB and λ boxB hairpins, respectively. In both P22_{N21}-P22boxB and λ_{N22} - λ boxB complexes the ARM peptides adopt bent α -helical structures that bind within the major groove of their cognate hairpins.¹⁰ The amino-terminal residues of P22_{N21} and λ_{N22} peptides are highly conserved and make similar contacts with the hairpin stems of their cognate boxB RNA targets.¹¹

The principle of our design was to use peptide sequence information from the P22_{N21}-P22boxB complex to increase the stability of the λ_{N22} - λ boxB complex without disrupting its unique loop-binding structure. To this end, we focused on nonconserved residues within the amino-terminal helical segment of λ_{N22} that interact primarily with the λ boxB hairpin stem in the cognate complex. Nonconserved residues within the amino-terminal helix of λ_{N22} were substituted with reciprocal residues from P22_{N21}. The

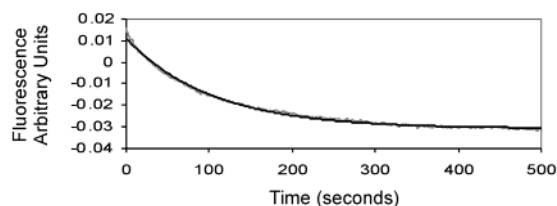


Figure 1. Stopped-flow measurement of P22_{N21}-P22boxB complex k_{off} infused with 100 \times concentration of unlabeled P22boxB at time zero (data shown in gray, fit shown in black) Experiments were performed at 20 $^\circ$ C in 50 mM KOAc, 20 mM Tris-OAc buffer, pH 7.5.

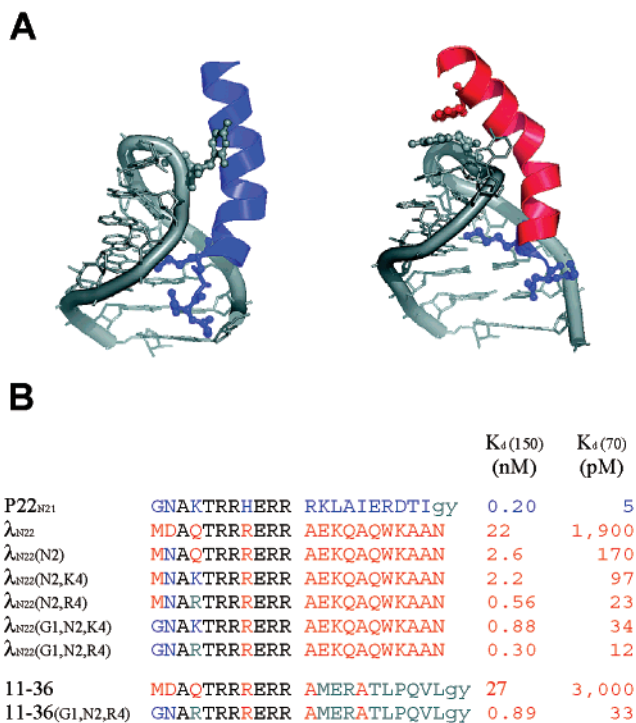


Figure 2. (A) Structural models of P22_{N21}-P22boxB (blue) and λ_{N22} (M1G, D2N, Q4K)- λ boxB (red).^{10,18} The extruded third base in P22boxB loop interfaces with P22_{N21}, and the second base in the λ boxB loop stacks with Trp18 of λ_{N22} in the cognate complexes. (B) Sequence alignment of P22_{N21} (blue), λ_{N22} (red), and substituted λ_{N22} peptides; conserved residues are shown in black, non-native residues are shown in gray. Binding constants for peptides against P22boxB 2AP-2 (blue) and λ boxB 2AP-2 (red) hairpins at 150 and 70 mM salt were determined from salt dependences as described in Supporting Information.

binding affinities of these substituted peptides for λ boxB at 150 and 70 mM salt are shown in Figure 2. Three reciprocal substitutions (M1G, D2N, Q4K) in the λ_{N22} peptide increase the stability of the λ_{N22} - λ boxB substituted complex by 1.8 kcal mol⁻¹ (Table 1). Substitution of Lys4 with an arginine residue further increases binding stability by 0.7 kcal mol⁻¹, in line with previous in vitro evolution experiments showing a glutamine-to-arginine change at

* Address correspondence to this author E-mail roberts@its.caltech.edu.

Table 1. Thermodynamic Parameters and Helicity of Peptides^a

peptide	$\Delta\Delta G$ (kcal mol ⁻¹) ^b	δ/δ^c	α -helicity (%) ^d
P22 _{N21}	-2.7	5.2 ± 0.2	65
λ_{N22}	0	3.2 ± 0.2	57
$\lambda_{N22}(N2)$	-1.2	3.5 ± 0.1	44
$\lambda_{N22}(N2,K4)$	-1.3	4.1 ± 0.2	38
$\lambda_{N22}(N2,K4)$	-2.1	4.2 ± 0.2	41
$\lambda_{N22}(G1,N2,K4)$	-1.8	4.3 ± 0.2	52
$\lambda_{N22}(G1,N2,R4)$	-2.5	4.2 ± 0.2	44

^a Data are listed for cognate complexes (P22_{N21}-P22boxB, λ_{N22} - λ boxB, $\lambda_{N22}(N2)$ - λ boxB, etc.). ^b Free energy values ($\Delta\Delta G$) were calculated from binding constants at 150 mM [M⁺]; $\Delta G = -10.3$ kcal mol⁻¹. ^c Salt dependence values $\delta/\delta = (\delta \log K_{obs})/(\delta \log [M^+])^{-1}$ were determined from five or more measurements over a range of 50-500 mM KOAc. ^d Helicities were calculated from circular dichroism measurements in 10 mM phosphate buffer.

this position (J.E.B and R.W.R, personal communication) and with work indicating that arginine oligomers bind RNA more tightly than those made of lysine.¹² Similarly, application of the MIG, D2N, Q4R substitution set to the peptide 11-36, a previously selected high-specificity binder of λ boxB,⁷ increases the binding stability of this peptide by 2.0 kcal mol⁻¹.

In addition to providing a binding signal, the relative fluorescence (F)¹³ of 2AP in the peptide-boxB complex confers information about the structure of the bound RNA. A decrease in fluorescence ($F < 1$) results from base stacking of the 2AP probe, while an increase in fluorescence ($F > 1$) results from exposure of 2AP to aqueous solvent.¹⁴ Titration of the native λ_{N22} - λ boxB complex gives a fluorescence signature consistent with the known solution structure of the complex. A large decrease in fluorescence is observed for 2AP label at positions 2 (2AP-2: $F = 0.31$) and 3 (2AP-3: $F = 0.22$) of the λ boxB loop consistent with Trp18 stacking and base stacking interactions, while a large increase in fluorescence is observed for loop position 4 of λ boxB (2AP-4: $F = 2.35$), indicative of base extrusion into solvent. This native structure is maintained in the substituted peptide complexes (2AP-2: $F = 0.32$ -0.36; 2AP-3: $F = 0.12$ -0.22; 2AP-4: $F = 2.36$ -2.60). Stacking of Trp18 on the λ boxB loop is also observed in NMR imino spectra of the substituted peptide complexes.¹⁵ Thermodynamic and kinetic measurements together with fluorescence and NMR data demonstrate that our reciprocal substitutions stabilize the adopted loop structure of the unique λ_{N22} - λ boxB native complex (Supporting Information).

Interestingly, circular dichroism measurements indicate that substituted peptides bind λ boxB with less α -helical character than native λ_{N22} , in large part attributable to the D2N substitution, which decreases helicity of the bound peptide by 13%. We speculate that this decrease in helicity results from loss of macrodipole stabilization by the negatively charged Asp2 residue.

Salt dependence data ($\delta/\delta = (-\delta \log K_{obs})/(\delta \log [M^+])^{-1}$) are useful in investigating the electrostatic contributions of different amino acid residues in binding events.¹⁶ D2N and Q4K reciprocal substitutions increase δ/δ values in the λ_{N22} - λ boxB complexes, demonstrating that favorable binding affinity exhibited from these substitutions has an electrostatic component (Table 1). Conversely, MIG and K4R replacements do not significantly affect δ/δ values, indicating that enhanced binding affinity from these substitutions is not purely electrostatic in nature.

We have determined that the native P22_{N21}-P22boxB complex and engineered λ_{N22} - λ boxB and 11-36- λ boxB complexes bind with picomolar affinity at physiological salt. These pairs are the most stable peptide-RNA complexes reported to our knowledge. Such examples provide evidence that short peptides can be found that bind unique RNA hairpins with affinity that rivals that of macromolecular complexes, providing a potential means to control RNA-dependent cellular processes through chemical methods. The reciprocal engineering approach presents a complementary tool to combinatorial or computational design techniques¹⁷ and should be of particular utility where all the space of possible sequences greatly exceeds the total number of sequences that may be generated and examined.

Acknowledgment. We thank Dr. Adam Frankel for his comments on the manuscript. This work was supported by a NSF Grant 9876246 (R.W.R), NIH Grant GM60416 (R.W.R), and training grant T32 GM08501.

Supporting Information Available: Experimental procedures accompanied by a tabulation of dissociation constants and relative fluorescence as well as CD and NMR figures (PDF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Weiss, M. A.; Narayana, N. *Biopolymers* **1998**, *48*, 167-180.
- Lazinski, D.; Grzadzilska, E.; Das, A. *Cell* **1989**, *59*, 207-218.
- (a) Van Gilst, M. R.; Rees, W. A.; Das, A.; von Hippel, P. H. *Biochemistry* **1997**, *36*, 1514-1524. (b) Frankel, A. D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 332-340.
- Tan, R.; Frankel, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5282-5286.
- Cilley, C. D.; Williamson, J. R. *RNA* **1997**, *3*, 57-67.
- Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630-5641.
- Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12374-12378.
- Peptides and unlabeled RNA were synthesized as described in Supporting Information. P22boxB_R: gGCGCU GACAA AGCGCc; λ boxB_R: gGC-CCU GAAAA AGGGCc. Loop residues are underlined.
- Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191-194.
- (a) Su, L.; Radek, J. T.; Hallenga, K.; Hermanto, P.; Chan, G.; Labeets, L. A.; Weiss, M. A. *Biochemistry* **1997**, *36*, 12722-12732. (b) Mogridge, J.; Legault, P.; Li, J.; Van Oene, M. D.; Kay, L. E.; Greenblatt, J. *Mol. Cell* **1998**, *1*, 265-275. (c) Cai, Z.; Gorin, A.; Frederick, R.; Ye, X.; Hu, W.; Majumdar, A.; Kettani, A.; Patel, D. J. *Nat. Struct. Biol.* **1998**, *5*, 203-212. (d) Legault, P.; Li, J.; Mogridge, J.; Kay, L. E.; Greenblatt, J. *Cell* **1998**, *93*, 289-299. (e) Scharpf, M.; Sticht, H.; Schweimer, K.; Boehm, M.; Hoffmann, S.; Rosch, P. *Eur. J. Biochem.* **2000**, *267*, 2397-2408.
- (a) Weiss, M. A. *Nat. Struct. Biol.* **1998**, *5*, 329-333. (b) Patel, D. J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 74-87. (c) Weisberg, R. A.; Gottesman, M. E. *J. Bacteriol.* **1999**, *181*, 359-367.
- (a) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1992**, *31*, 8932-8946. (b) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1997**, *36*, 7272-7279.
- $F = F_F/F_O$; F_O = fluorescence of free RNA; F_F = fluorescence of RNA following titration with peptide.
- (a) Millar, D. P. *Curr. Opin. Struct. Biol.* **1996**, *6*, 322-326. (b) Rachofsky, E. L.; Osman, R.; Ross, J. B. *Biochemistry* **2001**, *40*, 946-956. (c) Menger, M.; Eckstein, F.; Porschke, D. *Biochemistry* **2000**, *39*, 4500-4507. (d) Rist, M. J.; Marino, J. P. *Nucleic Acids Res.* **2001**, *29*, 2401-2408.
- Trp18 imino peak shifts from 10.0 ppm in free substituted peptides to 9.2 ppm in peptide- λ boxB complexes.
- (a) Draper, D. E. *J. Mol. Biol.* **1999**, *293*, 255-270. (b) GuhaThakurta, D.; Draper, D. E. *J. Mol. Biol.* **2000**, *295*, 569-580.
- Cheng, A. C.; Calabro, V.; Frankel, A. D. *Curr. Opin. Struct. Biol.* **2001**, *11*, 478-484.
- The PyMOL Molecular Graphics System; DeLano Scientific: San Francisco, CA, 2002; <http://www.pymol.org>.

JA026610B

Designed Arginine Rich RNA-Binding Peptides with Picomolar Affinity

Ryan J. Austin, Tianbing Xia, Jinsong Ren, Terry T. Takahashi, and Richard W. Roberts*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena CA 91125

Supplemental Information

RNA and Peptide Synthesis

Peptides were generated on an Applied Biosystems 432A peptide synthesizer using solid phase, F-Moc chemistry. Crude peptides were deprotected by TFA/ethanedithiol/thioanisole treatment and purified on a C-18 reverse phase HPLC column to a final purity greater than 95% (MALDI-TOF, Analytical C-18 HPLC). Peptides without a naturally occurring tryptophan or tyrosine residue were synthesized with a carboxy-terminal Gly-Tyr tag for quantification purposes.

Unlabeled RNA hairpins (λ box_{R15} and P22box_{BL15}) were synthesized by *in vitro* transcription using T7 RNA polymerase.¹ The RNA was purified by 20% urea-PAGE, desalted on a NAP column (Amersham Pharmacia), and freeze-dried. RNA was quantified by UV absorption at a wavelength of 260 nm.

Labeled RNA hairpins containing 2-amino purine (2AP) at loop position 2 (2AP-2), 3 (2AP-3), or 4 (2AP-4) were constructed by automated RNA synthesis using either 2-aminopurine-TOM-CE phosphoramidite or 2'-O-methyl 2-aminopurine phosphoramidite (Glen Research, Sterling, VA).

Steady-State Fluorescence Measurements

Measurements were conducted following the procedures of Barrick et. al.² Titrations were performed on a Shimadzu Spectrofluorophotometer at 20° C with Excitation/Emission wavelengths at 310/370 nm . Peptides were titrated iteratively into a constantly stirred solution of 2AP labeled RNA hairpin (20-200 nM RNA). Binding buffer contained 20 mM Tris-OAc, with a variable concentration of KOAc (15 mM-500 mM) at pH 7.5. Binding Constants were calculated for a one step binding mechanism by nonlinear least squares regression using the computer program DynaFit.³ All isotherms were fit with < 10% uncertainty. Relative 2AP fluorescence change (F) is reported as a fraction; $F = (F_F/F_O)$ where F_O and F_F are the initial and final fluorescence respectively. Binding constants required a change in fluorescence of greater than 10% for reliable fitting. Salt dependence values $(-\delta\log K_{obs})(\delta\log[M^+])^{-1}$ were determined from five or more binding constants within a range of salt concentrations (50-500mM KOAc).

Stopped-Flow Fluorescence Measurements

Experiments were conducted following the procedures of Lacourciere et al.⁴ Measurements were performed at 20° C under standard buffer conditions (20 mM Tris-OAc, 50 mM KOAc, pH 7.5) using a stop-flow device from Applied Photophysics (Surrey, U.K.) in two-syringe mode. Fluorescence excitation was performed at 310 nm and emission was measured with a filter cutoff > 360 nm. The

association rate (k_{on}) was determined for the λ_{N22} - λ boxB complex to be $7 \times 10^8 \text{ s}^{-1}\text{M}^{-1}$. A similar rate was observed for the P22_{N21}-P22boxB complex. Dissociation rate values (k_{off}) were determined for each interaction by infusing labeled complex with both 100x and 500x concentration of unlabeled competitor boxB at time zero.

Circular Dichroism Spectroscopy

Spectra were recorded on an Aviv 62 DS CD spectrofluorimeter in 10 mM potassium phosphate buffer at 20° C. Spectra for the peptide-RNA complexes were determined by subtracting spectra with the equivalent concentrations of RNA and of peptide from the spectra of the complex. Helix content was calculated from ellipticity at 222 nm (θ_{222}), using $-40,000(1-2.5/n)$ and $0 \text{ deg cm}^2 \text{ dmol}^{-1}$ as values for $\theta_{\text{helix}222}$ and $\theta_{\text{coil}222}$ respectively (n is the number of amino acids in the peptide).⁵ Helicity = $(\theta_{\text{observed}})/(\theta_{\text{helix}} + \theta_{\text{aromatic}})$. The contribution of Trp was assumed to be -2300 ± 600 (θ_{aromatic}) while the GY tag of the P22_{N21} was assumed to contribute insignificant helicity and aromaticity to the CD spectrum.⁶

NMR Spectroscopy

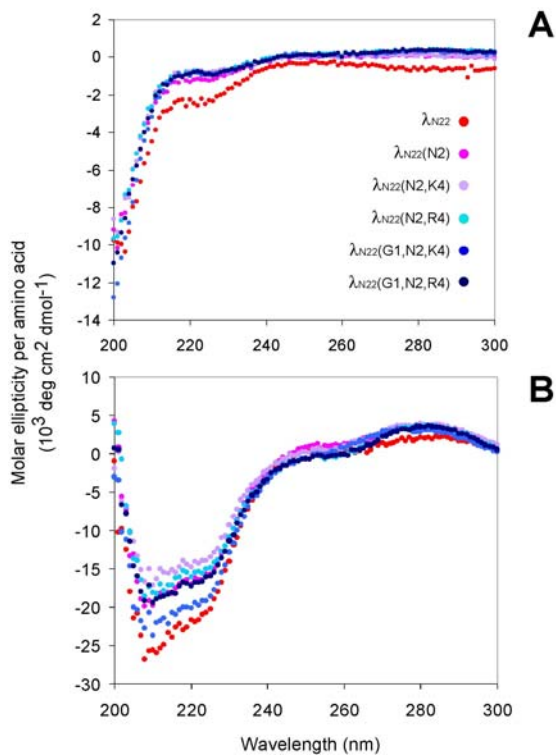
Samples were prepared in NMR buffer: 50 mM NaCl, 10 mM phosphate, pH 6 in H₂O:D₂O (90:10, vol:vol) as previously described.² NMR spectra were collected at 25°C on a Varian INOVA 600-MHz spectrometer. A modified double gradient echo Watergate solvent-suppression pulse sequence was used to suppress the solvent peak.⁷ Assignments were based on reported work.⁸

Supplemental Table 1. ^a

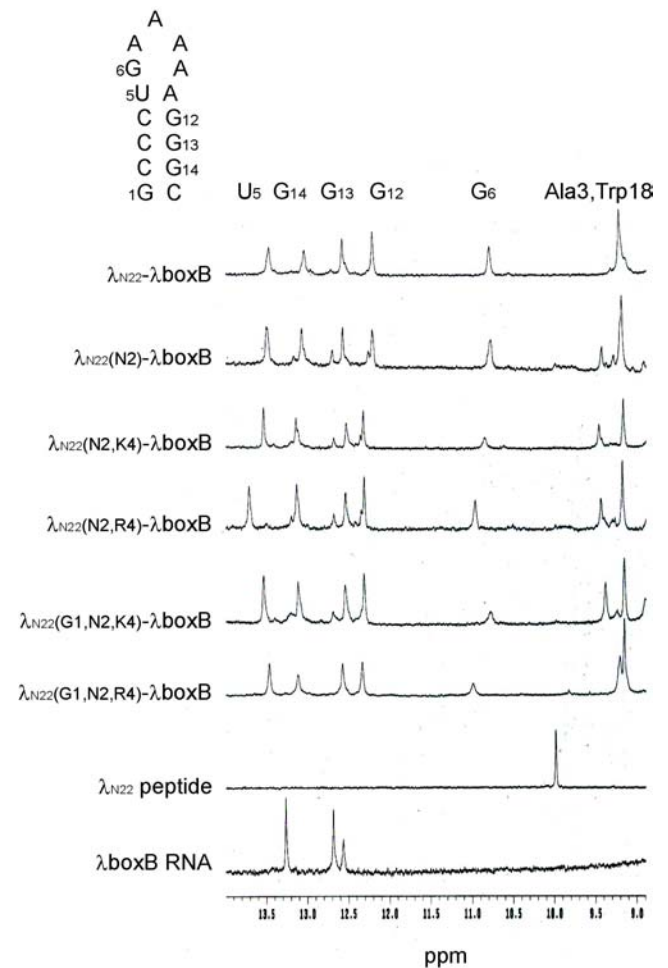
peptides	P22boxB RNA		λ boxB RNA		Fluorescence (F)		
	$K_d(70)$ (pM)	k_{off}/k_{on} (pM)	$K_d(70)$ (pM)	k_{off}/k_{on} (pM)	2AP-2	2AP-3	2AP-4
P22 _{N21}	5.0 ± 1.4	12	850 ± 290		1.60	0.53	1.00
λ_{N22}			1,900 ± 690	1,000	0.32	0.22	2.35
$\lambda_{N22}(N2)$			170 ± 18	170	0.36	0.15	2.60
$\lambda_{N22}(N2,K4)$			97 ± 22	69	0.33	0.21	2.45
$\lambda_{N22}(N2,R4)$			23 ± 4.9	17	0.35	0.14	2.50
$\lambda_{N22}(G1,N2,K4)$			34 ± 7.9	49	0.35	0.12	2.42
$\lambda_{N22}(G1,N2,R4)$			12 ± 3.2		0.32	0.13	2.39

^a Dissociation constants (K_d 's) were determined from salt dependence measurements for 2AP-2 labeled complexes at standard salt conditions (70 mM [M⁺]). Salt dependence values $\delta/\delta = (\delta \log K_{\text{obs}})/(\delta \log [M^+])^{-1}$ were determined from five or more measurements over a range of 50-500 mM KOAc. Individual measurements were fit with < 10% error. Kinetic approximations of dissociation constants, $K_d = k_{off}/k_{on}$, were calculated using a rate of association value, $k_{on} = 7 \times 10^8 \text{ s}^{-1}\text{M}^{-1}$; rate of dissociation values were fit with < 10% error.

Supplemental Figure 1. Circular dichroism measurements for λ_{N22} and reciprocal mutants in free solution (A) and in complex with λ_{boxB} (B). Measurements were taken in 10mM phosphate buffer at 20°C. Concentration of peptide in free solution was 10 μ M. Peptide-RNA complexes were prepared at 6:5 μ M stoichiometry.



Supplemental Figure 2. NMR imino-spectra of λ_{N22} - λ_{boxB} native and reciprocal complexes. The Trp18 imino-peak shifts from 10.0 ppm in the free peptides to 9.2 ppm in the λ_{boxB} complex. Samples were prepared in 50 mM NaCl, 10 mM phosphate buffer, pH 6 and titrated to 1:1 stoichiometry at 25°C. Assignments are based on reported work.⁸



References

* address correspondence to this author, E-mail rroberts@its.caltech.edu

- (1) (a) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol* **1989**, *180*, 51-62; (b) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res* **1987**, *15*, 8783-8798.
- (2) Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. *Proc Natl Acad Sci U S A* **2001**, *98*, 12374-12378.
- (3) Kuzmic, P. *Anal Biochem* **1996**, *237*, 260-273.
- (4) Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630-5641.
- (5) Chakrabarty, A.; Schellman, J. A.; Baldwin, R. L. *Nature* **1991**, *351*, 586-588.
- (6) (a) Chakrabarty, A.; Kortemme, T.; Padmanabhan, S.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 5560-5565; (b) Chakrabarty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci* **1994**, *3*, 843-852.
- (7) Liu, W.; Okajima, K.; Murakami, K.; Harada, N.; Isobe, H.; Irie, T. *J Lab Clin Med* **1998**, *132*, 432-439.
- (8) (a) Su, L.; Radek, J. T.; Hallenga, K.; Hermanto, P.; Chan, G.; Labeots, L. A.; Weiss, M. A. *Biochemistry* **1997**, *36*, 12722-12732; (b) Legault, P.; Li, J.; Mogridge, J.; Kay, L. E.; Greenblatt, J. *Cell* **1998**, *93*, 289-299.