J|A|C|S
COMMUNICATIONS

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 antitermination[2] 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 pM in solution at physiologic concentrations of monovalent cation, significantly stronger than previously determined by gel mobility shift (GMSA)[4] and polyacrylamide gel coelectophoresis (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}−λboxB). Kinetic data from stopped-flow measurements indicate that greater stability of the $P22_{N21}$−P22boxB complex relative to that of the λ_{N22}−λboxB complex results from a slower rate of dissociation (k_{off})($P22_{N21}$−P22boxB: $k_{off} = 0.01$ s^{-1}; λ_{N22}−λboxB: $k_{off} = 0.7$ 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}−λboxB: $k_{on} \approx 7 \times 10^8$ 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-fold[9] in the pentaloops of P22boxB and λboxB hairpins, respectively. In both $P22_{N21}$−P22boxB and λ_{N22}−λboxB 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}−λboxB 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 λboxB 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 binding affinities of these substituted peptides for λboxB 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}−λboxB 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

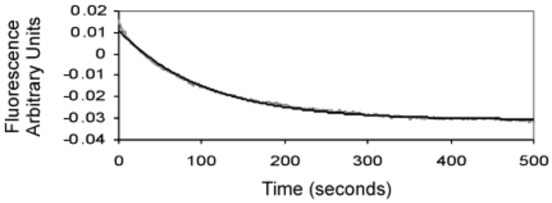
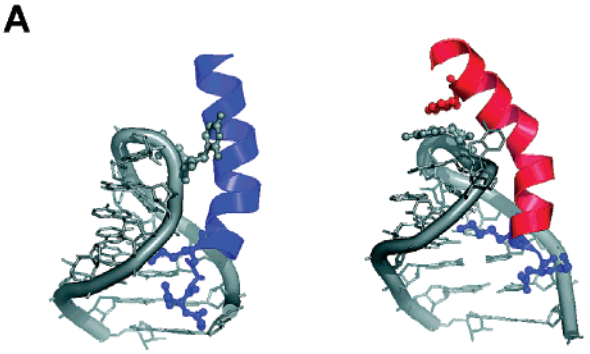


Figure 1. Stopped-flow measurement of $P22_{N21}$−P22boxB complex k_{off} infused with 100× concentration of unlabeled P22boxB at time zero (data shown in gray, fit shown in black) Experiments were performed at 20° C in 50 mM KOAc, 20 mM Tris-OAc buffer, pH 7.5.



B

			K_d(150) (nM)	K_d(70) (pM)
$P22_{N21}$	GNAKTRRHERR	RKLAIERDTIgy	0.20	5
λ_{N22}	MDAQTRRRERR	AEKQAQWKAAN	22	1,900
λ_{N22}(N2)	MNAQTRRRERR	AEKQAQWKAAN	2.6	170
λ_{N22}(N2,K4)	MNAKTRRRERR	AEKQAQWKAAN	2.2	97
λ_{N22}(N2,R4)	MNARTRRRERR	AEKQAQWKAAN	0.56	23
λ_{N22}(G1,N2,K4)	GNAKTRRRERR	AEKQAQWKAAN	0.88	34
λ_{N22}(G1,N2,R4)	GNARTRRRERR	AEKQAQWKAAN	0.30	12
11-36	MDAQTRRRERR	AMERATLPQVLgy	27	3,000
11-36(G1,N2,R4)	GNARTRRRERR	AMERATLPQVLgy	0.89	33

Figure 2. (A) Structural models of $P22_{N21}$−P22boxB (blue) and λ_{N22} (M1G, D2N, Q4K)−λboxB (red).[10,18] The extruded third base in P22boxB loop interfaces with $P22_{N21}$, and the second base in the λboxB 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 λboxB 2AP-2 (red) hairpins at 150 and 70 mM salt were determined from salt dependences as described in Supporting Information.

---

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

10.1021/ja026610b CCC: $22.00

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
λ_{N22}(N2)	−1.2	3.5 ± 0.1	44
λ_{N22}(N2,K4)	−1.3	4.1 ± 0.2	38
λ_{N22}(N2,K4)	−2.1	4.2 ± 0.2	41
λ_{N22}(G1,N2,K4)	−1.8	4.3 ± 0.2	52
λ_{N22}(G1,N2,R4)	−2.5	4.2 ± 0.2	44

[a] Data are listed for cognate complexes ($P22_{N21}$−P22boxB, λ_{N22}−λboxB, $\lambda_{N22(N2)}$−λboxB, 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 M1G, D2N, Q4R substitution set to the peptide 11−36, a previously selected high-specificity binder of λboxB,[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}−λboxB 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 λboxB loop consistent with Trp18 stacking and base stacking interactions, while a large increase in fluorescence is observed for loop position 4 of λboxB (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 λboxB 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}−λboxB native complex (Supporting Information).

Interestingly, circular dichroism measurements indicate that substituted peptides bind λboxB 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}−λboxB complexes, demonstrating that favorable binding affinity exhibited from these substitutions has an electrostatic component (Table 1). Conversely, M1G 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}−λboxB and 11−36−λboxB 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 techniques[17] 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

(1) Weiss, M. A.; Narayana, N. *Biopolymers* **1998**, *48*, 167−180.
(2) Lazinski, D.; Grzadzielska, E.; Das, A. *Cell* **1989**, *59*, 207−218.
(3) (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.
(4) Tan, R.; Frankel, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5282−5286.
(5) Cilley, C. D.; Williamson, J. R. *RNA* **1997**, *3*, 57−67.
(6) Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630−5641.
(7) Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12374−12378.
(8) Peptides and unlabeled RNA were synthesized as described in Supporting Information. $P22boxB_L$: gGCGCU GACAA AGCGCc; $\lambda boxB_R$: gGCCCU GAAAA AGGGCc. Loop residues are underlined.
(9) Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191−194.
(10) (a) Su, L.; Radek, J. T.; Hallenga, K.; Hermanto, P.; Chan, G.; Labeots, 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.
(11) (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.
(12) (a) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1992**, *31*, 8932−8946. (b) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1997**, *36*, 7272−7279.
(13) $F = F_F / F_O$; F_O = fluorescence of free RNA.; F_F = fluorescence of RNA following titration with peptide.
(14) (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.
(15) Trp18 imino peak shifts from 10.0 ppm in free substituted peptides to 9.2 ppm in peptide−λboxB complexes.
(16) (a) Draper, D. E. *J. Mol. Biol.* **1999**, *293*, 255−270. (b) GuhaThakurta, D.; Draper, D. E. *J. Mol. Biol*. **2000**, *295*, 569−580.
(17) Cheng, A. C.; Calabro, V.; Frankel, A. D. *Curr. Opin. Struct. Biol.* **2001**, *11*, 478−484.
(18) The PyMOL Molecular Graphics System; DeLano Scientific: San Francisco, CA, 2002; http://www.pymol.org.

JA026610B