Chapter 4: A link between structural and functional plasticity in an evolvable peptide sequence-space

Ryan J. Austin[†]*, Karin A. Crowhurst*[‡]*, Scott A. Ross*[‡]*, and Richard W. Roberts*[§]*

† Biochemistry and Molecular Biophysics Option, California Institute of Technology, Pasadena CA 91125.

‡ Howard Hughes Medical Institute and Division of Biology, California Institute of Technology, Pasadena CA 91125.

§ Department of Chemistry and Mork Family Department of Chemical Engineering, University of Southern California, Los Angeles, CA 90089.

* Corresponding Author: richrob@usc.edu

Abstract

The mechanism by which a protein evolves to serve a new biological function is widely thought to proceed through mutations conferring protein structural and functional plasticity. Here we present NMR structural and dynamical characterization of a two-step walk through peptide sequence space; from the λ N peptide sequence, which functions to activate transcription antitermination, to a peptide sequence that inhibits antitermination. A single arginine substitution in the wild-type peptide sequence fundamentally alters the α-helical fold, but maintains peptide biological activity through conformational exchange with a wild-type like structure. The work illustrates a link between protein mutability and structural and functional plasticity.

Introduction

A biological mechanism by which specific protein recognition processes evolve in parallel with new protein functions has been suggested by directed evolution and structural analysis of these molecules. These experiments demonstrate a correlation between protein structural plasticity and functional plasticity, indicating that natural proteins may evolve from one specific function into another via conformationally-flexible intermediates [1, 2]. One example of conformationally-flexible protein sequences that may exhibit a high degree of evolvability is provided by arginine-rich viral peptides, which participate in a wide variety of RNA binding interactions (Reviewed in [3-6]). The structural plasticity and mutability of these peptides have been well demonstrated [7], however evidence directly linking mutations in peptide sequences with shifts in the structural and functional plasticity of these peptides has been lacking.

One arginine-rich peptide-RNA complex that has been studied in some detail is the bacteriophage λ N peptide-λboxB$_R$ RNA hairpin. This complex nucleates viral transcription antitermination via formation of a shape specific bent α-helix/pseuo-tetraloop structure (Fig. 4.1) [8, 9]. Recent, directed evolution experiments on the λ N peptide have identified two critical mutations (Q15R and K14E), that when coupled increase the binding specificity of the peptide for λboxB$_R$ [10, 11]. These two substitutions have been shown to shift the structural populations of bound N peptide, from a stacked Trp18 conformation into an unstacked conformation and, in so doing, reverse the peptide's biological function from an activator of transcription anti-termination to an inhibitor [12-14]. The short mutagenic distance between the

functionally unique $E_{14}R_{15}$ and wild-type peptides presents an opportunity to characterize the mechanism of fold and function evolution between these two distinct molecular recognition solutions.

Here we present NMR structural and dynamical analysis of the R_{15} and $E_{14}R_{15}$ peptide mutant complexes (Fig. 4.1C); illustrating the linkage between peptide sequence, fold, and function, in a two-step evolutionary walk. In step-one, a single mutation of Gln15 to Arg15 (Q15R) fundamentally alters the fold of the bound N peptide from a 17 residue bent α-helix into a shortened, unbent 13 residue α-helix. This compact α-helical fold shows structural plasticity, retaining wild-type activity through conformational exchange with a stacked Trp18 conformation in the peptide C-terminus. In step-two, introduction of the covariant Glu14 substitution (K14E) reinforces the peptide's compact helical fold and increases structural degeneracy of the C-terminus, which disfavors stacking of the Trp18 residue, switching the peptide's biological function from activator to inhibitor.

Results and Discussion

Structural analysis of wild-type, R_{15}, and $E_{14}R_{15}$ peptide-λboxB$_R$ complexes

We have assigned the backbone and side-chain resonances for the wild-type, R_{15}, and $E_{14}R_{15}$ peptide-λboxB$_R$ complexes (Supplemental Tables S4.4-S4.6). The resonance assignments for wild-type N peptide free in solution and bound to λboxB$_R$ are similar to previously published values and indicate dramatic peptide structural rearrangement upon binding RNA [8, 9]. The R_{15} and $E_{14}R_{15}$ peptide mutants likewise exhibit large structural rearrangements upon binding λboxB$_R$ (Fig. 4.2). Backbone resonances of conserved amino-acid residues 2-7 in the wild-type, R_{15}, and $E_{14}R_{15}$ peptide-λboxB$_R$ complexes, show close alignment, indicating conservation of N-terminal peptide structure in these complexes. Distinctly, amide-backbone resonances for residues 8-17 of the R_{15}, and $E_{14}R_{15}$ peptide-λboxB$_R$ complexes shift away from the wild-type resonances on linear vectors suggesting a conformational shift in the structural fold of this peptide region.

The peptide fold induced upon binding λboxB$_R$ can be characterized by the shift in ^{13}Cα resonances between the free and bound peptide, where shifts greater > 1 ppm are indicative of α-helical structure [15]. ^{13}Cα shifts (ΔδCα) for the peptide complexes are shown in Figure 3. The bent α-helical structure of wild-type N peptide bound to λboxB$_R$ is illustrated nicely by the reduced ΔδCα value at residue 11 of this peptide. All of the peptides appear to adopt similar α-helical structures in their N-terminal regions, exhibiting negative ΔδCα and positive ΔδCβ values at residue 2, consistent with N-terminal helix capping [16]. However, further from the N-terminus, the R_{15} peptide adopts a fundamentally different structural fold, forming an unbent α-helix that spans

from residues 3-16. While some level of wild-type structure appears preserved adjacent to this compact helix from residues 16-19 in the R_{15} complex, this structure is collapsed in the $E_{14}R_{15}$ mutant, which reinforces the compact α-helical fold of R_{15}. The adopted secondary structure of wild-type and $E_{14}R_{15}$ peptides is consistent with α-helicity values previously reported from circular dichroism spectra of bound and unbound peptides, which estimated 16.4 and 13.8 residues of helicity in the wild-type and $E_{14}R_{15}$ peptides, respectively [12].

The mechanism by which the Q15R mutation alters the structural fold of λ N peptide is likely attributable to the disruption of Gln15 bend-stabilizing intra-molecular contacts and the elecrostatics of the Arg15 residue. Structural analyses of the wild-type complex suggest that the amino-side chain of Gln15 stabilizes the 120 ° peptide bend via intra-molecular hydrogen-bond contacts with the main-chain carbonyl of Arg11 (Fig. 4.1B) [8], which would be disrupted by a Q15R substitution. A separate structural analysis of the λ N peptide-boxB complex has suggested that inter-molecular hydrophobic interactions between Gln15 and nucleotides A7 and A8 stabilize the helix-bend-helix fold of the peptide [9], however it is unclear whether the Q15R substitution would disrupt such contacts. Previous mutagenic studies using the uncharged arginine isostere, citrulline, have indicated that the unique conformation of the R_{15} peptide is a result of Arg15 charge and not side-chain sterics [14]. The Q15R mutation introduces an additional guanidinium group into the vicinity of a positively charged arginine-rich cluster composed of Arg7, Arg8, and Arg11 side-chains. Alleviation of the sharp helical bend and termination of peptide-helicity at residue 16 could mitigate charge-charge repulsions and generate a favorable electrostatic interaction between Arg15 and the C-

terminal macroscopic-dipole of the compact helix. It is interesting to note that a corresponding Arg15 residue is found in the HK022 Nun peptide, which shows significant sequence and structural conservation with the wild-type λ peptide [17]. The HK022 Nun peptide, however, binds λboxB$_R$ as a bowed helix, without the helix-disrupting bend of the wild-type λ N peptide.

Dynamic analysis of peptide-λboxB$_R$ complexes

We have performed auto-relaxation and NOE saturation experiments on the peptide complexes to access the structural plasticity of the peptide fold (Fig. 4.4). Most residues of the wild-type peptide exhibit $\{^1H\}$-^{15}N NOE values > 0.65, indicating that the peptide is well ordered on the picosecond to nanosecond time scale [18]. The slightly smaller $\{^1H\}$-^{15}N NOE values at residues 2 (0.39 ± 0.14), 21 (0.53 ± 0.04), and 22 (0.43 ± 0.01) are indicative of an increasing distribution of conformations at the wild-type peptide termini. Separately, the $\{^1H\}$-^{15}N NOE values of R_{15} and $E_{14}R_{15}$ peptides indicate a fold that maintains structural order in residues 16-18 beyond the compact α-helix (residues 3-16). Both R_{15} and $E_{14}R_{15}$ structures degenerate after residue 18, showing increasingly negative $\{^1H\}$-^{15}N NOE values in the C-termini, indicative of highly disordered peptide conformations. $\{^1H\}$-^{15}N NOE values for the Trp18 indole proton were also measured and show increasing structural degeneracy with the Q15R and K14E mutations (Trp18 Hε $\{^1H\}$-^{15}N NOE: wild-type 0.67 ± 0.02; R_{15} 0.33 ± 0.01; $E_{14}R_{15}$ 0.16 ± 0.01) consistent with Trp18 stacking populations determined by transient fluorescence measurements (population Trp18 stacked: wild-type 0.66; R_{15} 0.22; $E_{14}R_{15}$ 0.11 [14]).

Transverse relaxation values (R_2), which include contributions from chemical shift anisotropy (CSA) and conformational exchange processes, show larger variability in the α-helical regions of these peptides than the NOE values. The α-helical region of the wild-type peptide is divided into two R_2 populations: a fast relaxation population consisting of residues 2-9 (R_2 = 11.4 ± 1.6 s^{-1}) and a slower relaxation population consisting of residues 10-19 (R_2 = 7.5 ± 0.6 s^{-1}). Measurements of the exchange free transverse cross-relaxation rate (η_{xy}) indicate that fast R_2 rates in residues 2-9 of the wild-type peptide are a consequence of anisotropic relaxation mechanisms (Supplemental Table S4.1). The R_{15} and $E_{14}R_{15}$ peptides exhibit a more gradual decline in CSA dependent R_2 rates from residue 2 to 18, consistent with bowed, rather than bent peptide geometries in these complexes. Plots of the S^2 order parameter, determined from model-free fitting of the relaxation data, show values approaching 1.0 in wild-type peptide residues 2-19 (S^2 = 0.90 ± 0.07), consistent with a well ordered structural fold. R_{15} and $E_{14}R_{15}$ peptides were fit to lower S^2 values at residues 2-18 (S^2: R_{15} = 0.75 ± 0.05; $E_{14}R_{15}$ = 0.74 ± 0.04) indicating less ordered structure in these peptides. Unlike the wild-type and R_{15} peptides, relaxation arising from conformational exchange is not predicted to contribute to the relaxation of C-terminal residues 19-22 in the $E_{14}R_{15}$ peptide (Supplemental Fig. 4.1 and Table S4.9).

Previous studies have demonstrated a correlation between Trp18 stacking and N peptide antitermination activity. Up-conversion and transient fluorescent experiments indicate that the stacked Trp18 population is in exchange with an unstacked peptide conformation or conformations [13, 14]. Proton-NMR and TCSPC fluorescence measurements have demonstrated that this exchange rate is faster than milliseconds, as a

single Trp18 imino-resonance is observed to be in fast exchange across a 0-600 Hz chemical shift ($\Delta\omega$) in the peptide complexes [12], but slower than nanoseconds, as multiple stable populations are observed in TCSPC experiments [13]. These experiments therefore suggest a Trp18 exchange rate on the order of microseconds. Both protein folding events and conformational exchange processes have been shown to operate on microsecond time scales [19]. Our present findings demonstrate a correlation between unstacking of the Trp18 residue and unfolding of the carboxy-terminus in R_{15} and $E_{14}R_{15}$ peptide mutants. This suggests that Trp18 stacking is a function of peptide backbone folding, rather than a simple two-state conformational exchange in these complexes. While the carboxy-terminus of the $E_{14}R_{15}$ peptide degenerates into isotropic tumbling, the unfolded carboxy-terminus of R_{15} retains populations in exchange ($R_{ex} > 0$). Such structural plasticity in R_{15} may account for the retained wild-type function of the peptide, suggesting a remarkably subtle linkage between structural and functional plasticity in the peptides studied.

In vitro selection experiments have demonstrated the importance of the Q15R and K14E substitutions in evolving new specific peptide-$\lambda boxB_R$ binding solutions. These two substitutions are present in 86% (Arg15) and 39% (Glu14, Arg15) of sequences evolved from a pool of λ N peptides randomized at positions 13-22 [11]. A diverse family of peptides containing the Glu14-Arg15 sequence all exhibit similar α-helical and loop-binding conformations [20], evidencing that these residues are critical for evolution of the compact α-helical fold. A similar set of coupled amino-acid substitutions Q4K and R8H have been shown to dramatically alter the binding specificity of the λ N peptide towards the boxB hairpins of bacteriophage P22 and ϕ21 [21, 22]. The evolvability of λ

N peptide reflects a general theme in viral peptide-RNA interactions [7]. These arginine-rich peptide-RNA complexes should provide a simple platform for further characterization of the relationships between protein mutability and structural and functional plasticity.

Experimental Procedures

RNA preparation

The λboxB$_R$ sequence is 5'-GCCCU GAAGA AGGGC-3'. λboxB$_R$ was transcribed *in vit*ro using T7 polymerase [23]. The DNA sequence of the template used was 5'-GCC CTT TTT CAG GGC TAT AGT GAG TCG TAT TAC GAA TT-3'; the promoter sequence was 5'-AAT TCG TAA TAC GAC TCA CTA TA-3'. λboxB$_R$ was purified by 20% denaturing Urea-PAGE, concentrated by ELUTRAP (*Schleicher and Schuell*) and desalted by NAP25 sephadex column (*Amersham Biosciences*).

Peptide preparation

N peptides were expressed as Ubiquitin (Ub) fusions [24]. λ wild-type N peptide DNA sequence (5'-ATG GAT GCA CAA ACA CGC CGC CGC GAA CGT CGC GCA GAG AAA CAG GCT CAA TGG AAA GCA GCA AAT CCC CTG TTG GTT GGG GTA AGC GCA-3') was amplified (primers: *5' SacII*, 5'-AAG GCT CCG CGG TGG TAT GGA TGC ACA ACC CGC CGC-3'; *3' BamH1* 5'-GCA GTC GAA TCC TTA ATT TGC TGC TTT CCA TTG CGC CTG-3') and cloned into a ubiquitin expression vector H6MUBQ (5.6 kB). H6MUBQ is a pET11c (*New England Biolabs*) based vector, containing a synthetic amino-terminal His-6 tag-ubiquitin construct with a carboxy-terminal *SacII* cloning site. H6MUBQ was adapted from the MUBQ, an ubiquitin expression plasmid obtained from Steven Mayo. R_{15} and $E_{14}R_{15}$ N peptide mutant expression sequences were produced using the Quik Change Site-Directed mutagenesis kit (*Stratagene*) (primers: *5'KR* 5'-CGA ACG TCG CGC AGA GAA ACG TGC TCA

ATG GAA AGC A-3'; *3'KR* 5'-TGC TTT CCA TTG AGC ACG TTT CTC TGC GCG ACG TTC G-3'; *5'ER* 5'-GCG AAC GTC GCG CAG AGG AAC GTG CTC AAT GGA AAG CAG-3'; *3'ER* 5'-CTG CTT TCC ATT GAG CAC GTT CCT CTG CGC GAC GTT CGC-3'). Expression plasmids were transformed into the BL21(DE3) *E. coli* strain (*Novagen*). Cells were grown on minimal media supplemented with either $^{15}NH_4Cl$ (*Isotech*) or $^{15}NH_4Cl$ and $^{13}C_6$-*d*-glucose (*Isotech*)(1L minimal media w/ ^{15}N: 17 g Na_2HPO_4-$7H_2O$, 4 g KH_2PO_4, 3 g NaCl, 2 g $^{15}NH_4Cl$, 2 g Na_2SO_4, 8 g glucose; 2 ml 1 M $MgSO_4$, 100 μl $CaCl_2$, 1x *Gibco* Eagle Basal Media Vitamin Mix, 1x antibiotic. 1 L minimal media w/ ^{15}N and ^{13}C: 17 g Na_2HPO_4-$7H_2O$, 4 g KH_2PO_4, 3 g NaCl, 2 g $^{15}NH_4Cl$, 2 g Na_2SO_4, 2 g ^{13}C-glucose; 2 ml 1M $MgSO_4$, 100 μl $CaCl_2$, 1x *Gibco* Eagle Basal Media Vitamin Mix, 1x antibiotic)[25]. Cells were grown at 37 °C to OD_{600} = 0.6 and induced with 1.0 mM IPTG for 3 hours. Cells were collected by centrifugation and lysed by French Press. Ub-N peptide fusions were purifed on Ni-NTA (*QIAGEN*). Ub-N peptide Ni-NTA elution fractions were treated directly with 5 nM ubiquitin hydrolase (*Boston Biochem*); fusion cleavage was complete after 30 minutes. The mixture was concentrated and N peptide was purified by HPLC (*Beckman Instruments*) using a C18 reverse-phase semi-prep column. (*Vydac*). N peptide fractions were assessed to be >95% pure by analytical HPLC and by MALDI-TOF mass spectrometry.

NMR sample preparation

Samples were prepared by aliquoting concentrated peptide into RNA, resuspending in 280 μl total volume and transferring to a D_2O matched microtube (*Shigemi*). NMR buffer was 20 mM KH_2PO_4 (pH 6.0), 50 mM NaCl, 0.5 mM EDTA, 1.0

mM DSS (H_20/D_20 v/v 95/5). Peptide and RNA concentrations were measured by UV spectrophotometry at 280 nm and 260 nm respectively.

NMR experiments

NMR experiments were performed at 25 °C on a *Varian Inova* 600 MHz instrument. ^{15}N-HSQC, ^{13}C-HSQC, HNCACB, CBCACONH, CCONH-TOCSY, HCCONH-TOCSY, and HCCH-COSY experiments were performed on each sample. ^{15}N longitudinal auto-relaxation (R_1) rates were measured with *biopack's* gNhsqc inversion recovery pulse sequence (12 experiments: relaxation delays (ms) = 10, 50, 2x(100), 150, 200, 250, 400, 2x(600), 800, 1000). Transverse auto-relaxation (R_2) rates were determined with *biopack* gNhsqc CPMG spin-echo experiments (12 experiments: relaxation delays (ms) = 10, 30, 2x(50), 70, 90, 110, 130 150, 2x(210), 250); ^{15}N nuclear Overhauser effects (NOE) were measured in triplicate using an ^{1}H saturation pulse of 3 seconds. Longitudinal (η_z) and transverse (η_{xy}) ^{1}H-^{15}N dipolar /^{15}N CSA cross-correlation relaxation rates were measured from a ^{15}N labeled sample using modified pulse sequences provided by Mikeal Akke's lab [26-30]. 64 transients were averaged per t_1 increment in the auto-correlation experiment, while 16 transients were averaged per t_1 increment in the cross-correlation experiment. Relaxation delays used for the η_z measurements were 0.15, 0.225, 0.30, 0.375, and 0.450 seconds. The range of delay τ was determined using autorelaxation data [31]. Relaxation delays used for the η_{xy} measurements were 0.0320, 0.0534, 0.0748, 0.0961, and 1.068 seconds assuming a coupling constant of 94 Hz [31].

NMR data analysis

Spectra were processed using nmrPipe [32] and resonance assignments were obtained using NMRView 5.0 (*One Moon Scientific*) [33]. The secondary chemical shift of Cα was calculated as $\Delta\delta C\alpha = \delta C\alpha_{bound} - \delta C\alpha_{free}$, where $\delta C\alpha_{bound}$ is the shift in the peptide-RNA complex and $\delta C\alpha_{free}$ is the shift in the free peptide alone. *R1* and *R2* relaxation rates were determined from peak intensities using the Levenberg-Marquardt non-linear least squares fitting program Curvefit (www.palmer.hs.columbia.edu). The heteronuclear $\{^1H\text{-}^{15}N\}$ NOE was calculated from the average of three measurements as $NOE = I_{sat}/I_{unsat}$. The cross-correlation relaxation rate constants η_z and η_{xy} were obtained by Levenberg-Marquard non-linear optimization of the function:

$$I_{cross}(\tau)/I_{auto}(\tau) = \tanh(\eta_i\tau) \qquad \textit{(Eq. 4.1)}$$

where I_{cross} is the intensity from the cross-correlation experiment and I_{auto} is the intensity from the auto-correlation experiment [26, 27]. The exchange-free transverse relaxation rate, R_2^0, is calculated from the equation:

$$R_2^0 = (R_1 - 1.249\sigma_{NH})(\eta_{xy}/\eta_z) + 1.079\sigma_{NH} \qquad \textit{(Eq. 4.2)}$$

Where σ_{NH} is the heteronuclear cross-relaxation rate constant, which is related to the steady state $[^1H\}\text{-}^{15}N$ *NOE* by

$$\sigma_{NH} = R_1(\gamma_N/\gamma_H)(NOE - 1) \qquad \textit{(Eq. 4.3)}$$

In which γ_H and γ_N are the gyromagnetic ratios of ^1H and ^{15}N respectively. [18, 31]

A 3 parameter Model free analysis of R_1, R_2, and *NOE* relaxation rates was performed in *Mathematica*, using notebooks developed by the Spyracopoulos lab [34]. Order parameters were fit to one of five models, indicated in supplemental tables S4.7-S4.9.

Acknowledgements

We would like to thank P. Huang and T. Takahashi for generously providing the peptide-ubiquitin expression vector MUBQ. We would also like to thank M. Akke for the cross-relaxation scripts used to measure η_{xy} and η_z values, and S. Mayo for use of the *Varian Inova* 600 MHz instrument. This work has been supported by the National Institutes of Health (GM60416) and the Beckman Foundation.

References

1. Joyce, G.F. (1997). Evolutionary chemistry: getting there from here. Science *276*, 1658-1659.
2. James, L.C., and Tawfik, D.S. (2003). Conformational diversity and protein evolution-a 60-year-old hypothesis revisited. Trends Biochem Sci *28*, 361-368.
3. Weiss, M.A., and Narayana, N. (1998). RNA recognition by arginine-rich peptide motifs. Biopolymers *48*, 167-180.
4. Draper, D.E. (1999). Themes in RNA-protein recognition. J Mol Biol *293*, 255-270.
5. Patel, D.J. (1999). Adaptive recognition in RNA complexes with peptides and protein modules. Curr Opin Struct Biol *9*, 74-87.
6. Frankel, A.D. (2000). Fitting peptides into the RNA world. Curr Opin Struct Biol *10*, 332-340.
7. Das, C., and Frankel, A.D. (2003). Sequence and structure space of RNA-binding peptides. Biopolymers *70*, 80-85.
8. Legault, P., Li, J., Mogridge, J., Kay, L.E., and Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell *93*, 289-299.
9. Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S., and Rosch, P. (2000). Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. Eur J Biochem *267*, 2397-2408.
10. Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T., and Roberts, R.W. (2001). Large libraries reveal diverse solutions to an RNA recognition problem. Proc Natl Acad Sci U S A *98*, 12374-12378.
11. Barrick, J.E., and Roberts, R.W. (2002). Sequence analysis of an artificial family of RNA-binding peptides. Protein Sci *11*, 2688-2696.
12. Xia, T., Frankel, A., Takahashi, T.T., Ren, J., and Roberts, R.W. (2003). Context and conformation dictate function of a transcription antitermination switch. Nat Struct Biol *10*, 812-819.
13. Xia, T., Becker, H.-C., Wan, C., Frankel, A., Roberts, R.W., and Zewail, A.H. (2003). The RNA-protein complex: Direct probing of the interfacial recognition dynamics and its correlation with biological functions. Proc Natl Acad Sci U S A *100*, 8119-8125.
14. Xia, T., Wan, C., Roberts, R.W., and Zewail, A.H. (2005). RNA-protein recognition: single-residue ultrafast dynamical control of structural specificity and function. Proc Natl Acad Sci U S A *102*, 13013-13018.
15. Wishart, D.S., and Sykes, B.D. (1994). The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR *4*, 171-180.
16. Gronenborn, A.M., and Clore, G.M. (1994). Identification of N-terminal helix capping boxes by means of 13C chemical shifts. J Biomol NMR *4*, 455-458.
17. Faber, C., Scharpf, M., Becker, T., Sticht, H., and Rosch, P. (2001). The structure of the coliphage HK022 Nun protein-lambda-phage boxB RNA complex.

Implications for the mechanism of transcription termination. J Biol Chem *276*, 32064-32070.
18. Tjandra, N., A, S., and Bax, A. (1996). Protein backbone dynamics and 15N chemical shift anisotropy from quantitative measurement of relaxation interference effects. J Am Chem Soc *118*, 6986-6991.
19. Kay, L.E. (2005). NMR studies of protein structure and dynamics. J Magn Reson *173*, 193-207.
20. Barrick, J.E., and Roberts, R.W. (2003). Achieving specificity in selected and wild-type N peptide-RNA complexes: the importance of discrimination against noncognate RNA targets. Biochemistry *42*, 12998-13007.
21. Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. (2002). Designed Arginine-Rich RNA-Binding Peptides with Picomolar Affinity. J. Am. Chem. Soc. *124*, 10966-10967.
22. Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. (2003). Differential modes of recognition in N peptide-boxB complexes. Biochemistry *42*, 14957-14967.
23. Milligan, J.F., and Uhlenbeck, O.C. (1989). Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol *180*, 51-62.
24. Pilon, A.L., Yost, P., Chase, T.E., Lohnas, G.L., and Bentley, W.E. (1996). High-level expression and efficient recovery of ubiquitin fusion proteins from Escherichia coli. Biotechnol Prog *12*, 331-337.
25. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory Press).
26. Evenas, J., Forsen, S., Malmendal, A., and Akke, M. (1999). Backbone dynamics and energetics of a calmodulin domain mutant exchanging between closed and open conformations. J Mol Biol *289*, 603-617.
27. Malmendal, A., Evenas, J., Forsen, S., and Akke, M. (1999). Structural dynamics in the C-terminal domain of calmodulin at low calcium levels. J Mol Biol *293*, 883-899.
28. Evenas, J., Malmendal, A., and Akke, M. (2001). Dynamics of the transition between open and closed conformations in a calmodulin C-terminal domain mutant. Structure (Camb) *9*, 185-195.
29. Akke, M. (2002). NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis. Curr Opin Struct Biol *12*, 642-647.
30. Palmer, A.G., 3rd, Kroenke, C.D., and Loria, J.P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol *339*, 204-238.
31. Kroenke, C.D., Loria, J.P., Lee, L.K., Rance, M., and Palmer, A.G., 3rd (1998). Longitudinal and transverse 1H-15N dipolar/15N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. J Am Chem Soc *120*, 7905-7915.
32. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.

33. Johnson, B.A., and Blevins, R.A. (1994). Journal of Biomolecular NMR *4*, 603-614.
34. Spyracopoulos, L. (2006). A suite of Mathematica notebooks for the analysis of protein main chain 15N NMR relaxation data. J Biomol NMR *36*, 215-224.

Figures

Figure 4.1 Bacteriophage λ N peptide-λboxB$_R$ complex. (A) Model of the λ N peptide-λboxB$_R$ complex. The λ N peptide, represented in dark gray, adopts a bent α-helical conformation, which is accommodated along the 5' edge of the λboxB$_R$ hairpin major groove (white). Gln15 and Trp18 residues of λ N peptide, along with the hairpin adenine 7 base (A7) are highlighted in ball-and-stick representation. π-stacking of the Trp18 residue upon A7 extends the pseudo-GNRA tetraloop base stack of the λboxB$_R$ loop, and is critical for the antitermination function of the complex (Model was generated from the Protein Data Bank file 1QFQ [9]). **(B)** An intra-molecular hydrogen bond between the amide side-chain of Gln15 and the main-chain carbonyl of Arg11 stabilizes a 120° bend in the α-helix. **(C)** Sequences are shown for the λboxB$_R$ hairpin construct (left) and N peptide wild-type, R_{15} and $E_{14}R_{15}$ mutants (right).

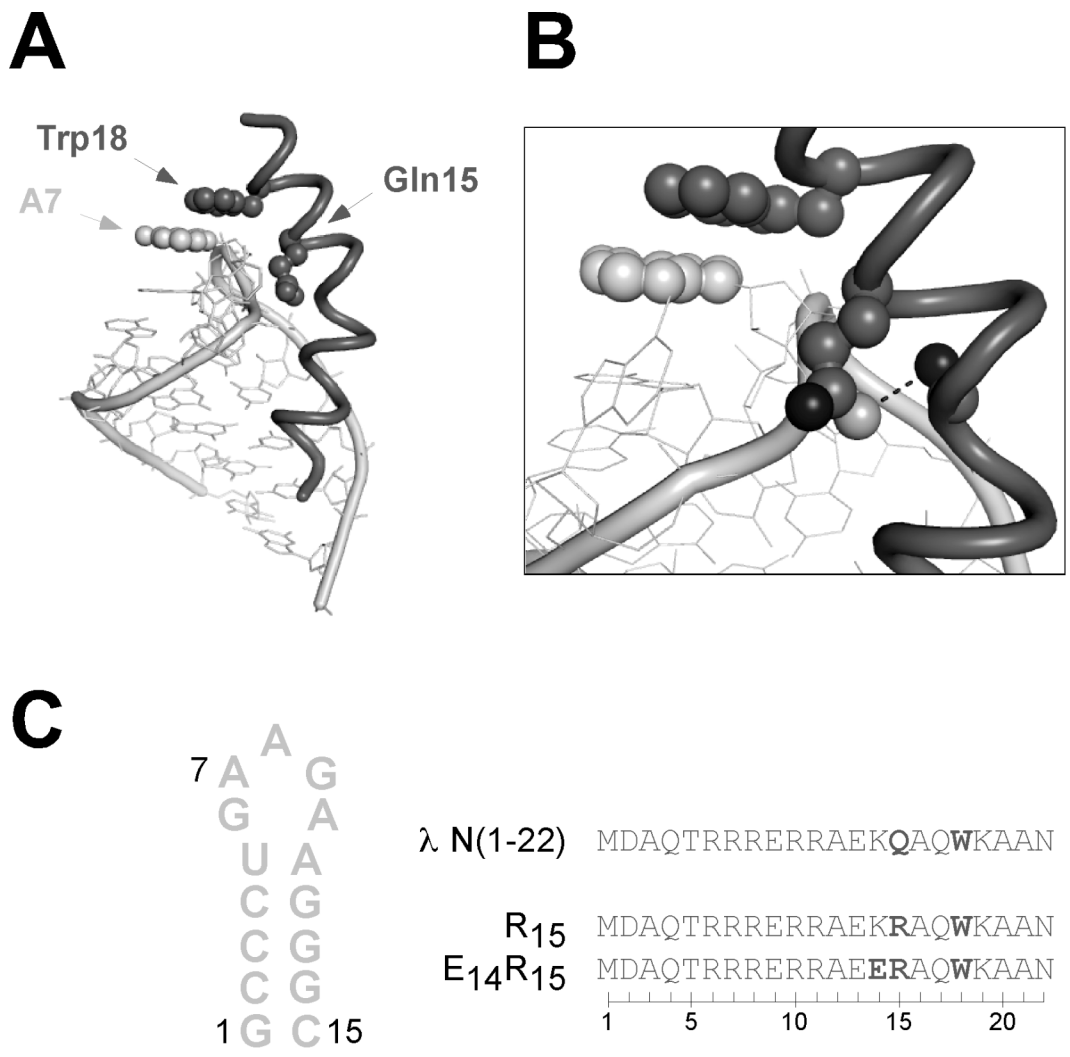

A
Trp18
A7
Gln15
B
C
7
A
A G
G A
U A
C G
C G
C G
1 G C 15
λ N(1-22) MDAQTRRRERRAEKQAQWKAAN
R15 MDAQTRRRERRAEKRAQWKAAN
E14R15 MDAQTRRRERRAEERAQWKAAN
1 5 10 15 20

Figure 4.2 ^{1}H-^{15}N HSQC NMR spectrum of the $E_{14}R_{15}$ λ N peptide mutant. Peptide free in solution (black) and complexed with λboxB_R (gray). Resonance assignments for the backbone amide groups are numerated. The Trp18 indole resonance for both free and bound peptides are overlapping and lie off diagram (^{1}H = 10.11 ppm; ^{15}N = 129.15). Amide resonances of the λ N wild-type, R_{15}, and $E_{14}R_{15}$ peptide-boxB_R complexes overlap at backbone positions 2-7 (Amide proton, amide nitrogen, and side chain Cα and Cβ resonance assignments are tabulated in supplemental tables S4.4-S4.6).

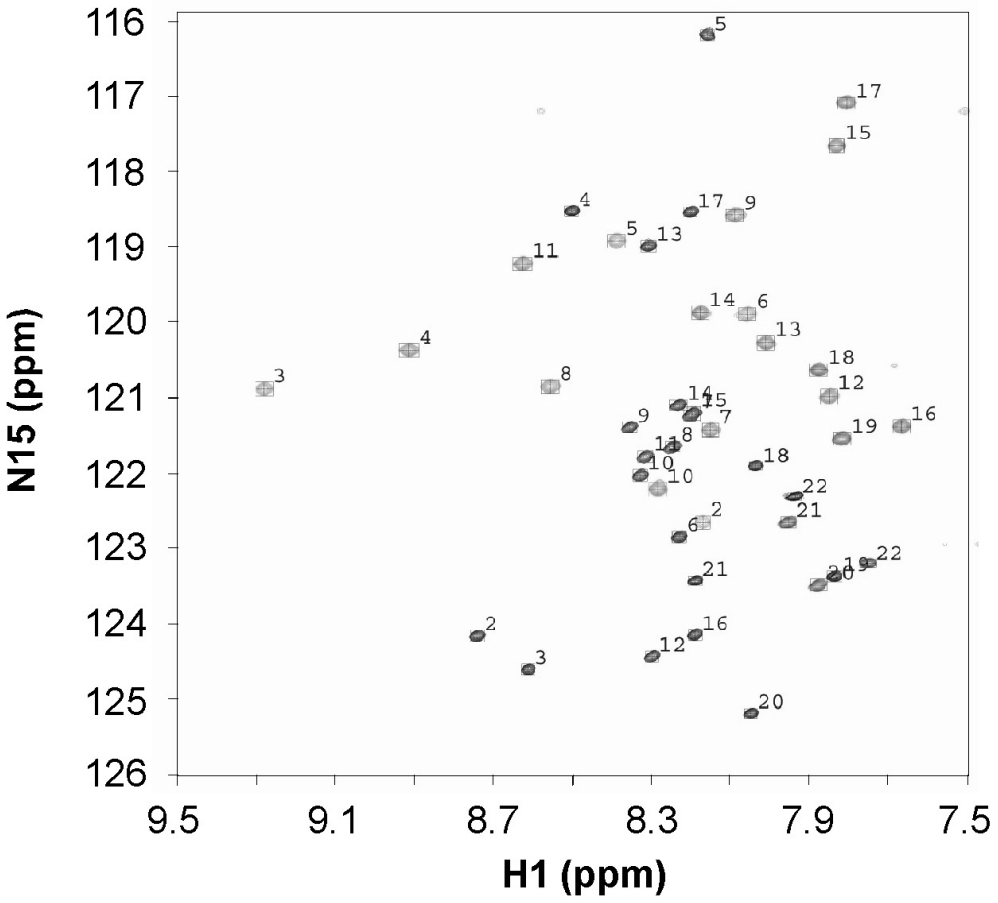

N15 (ppm)
116
117
118
119
120
121
122
123
124
125
126
9.5
9.1
8.7
8.3
7.9
7.5
H1 (ppm)

Figure 4.3 Backbone $^{13}C\alpha$ secondary chemical shifts ($\Delta\delta C\alpha$) for peptide $\lambda boxB_R$. $\Delta\delta C\alpha$ > 1 ppm are indicative of α-helical peptide conformations. Values approaching zero indicate an absence of peptide secondary structure. Smoothed line plots are shown to guide the eye.

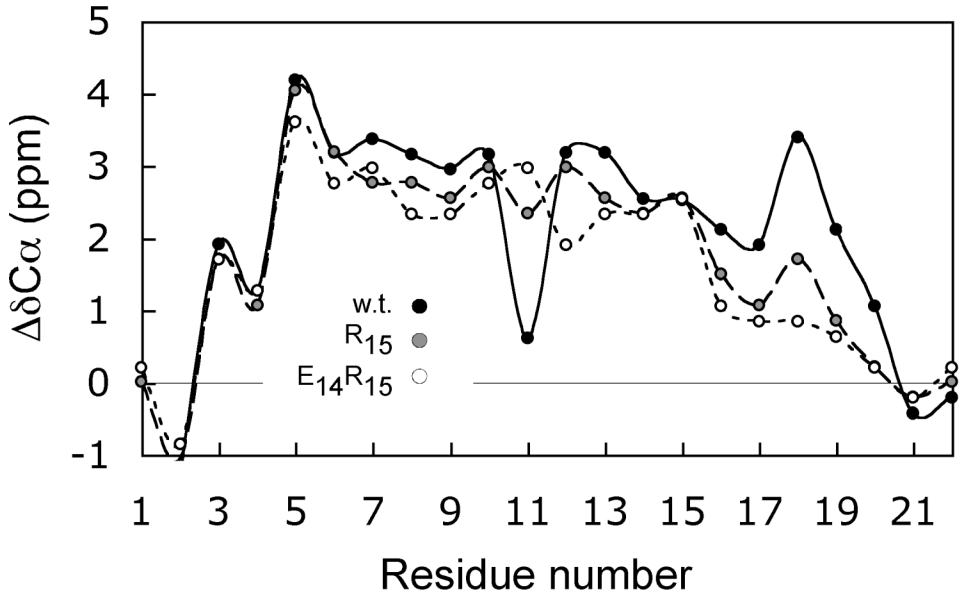
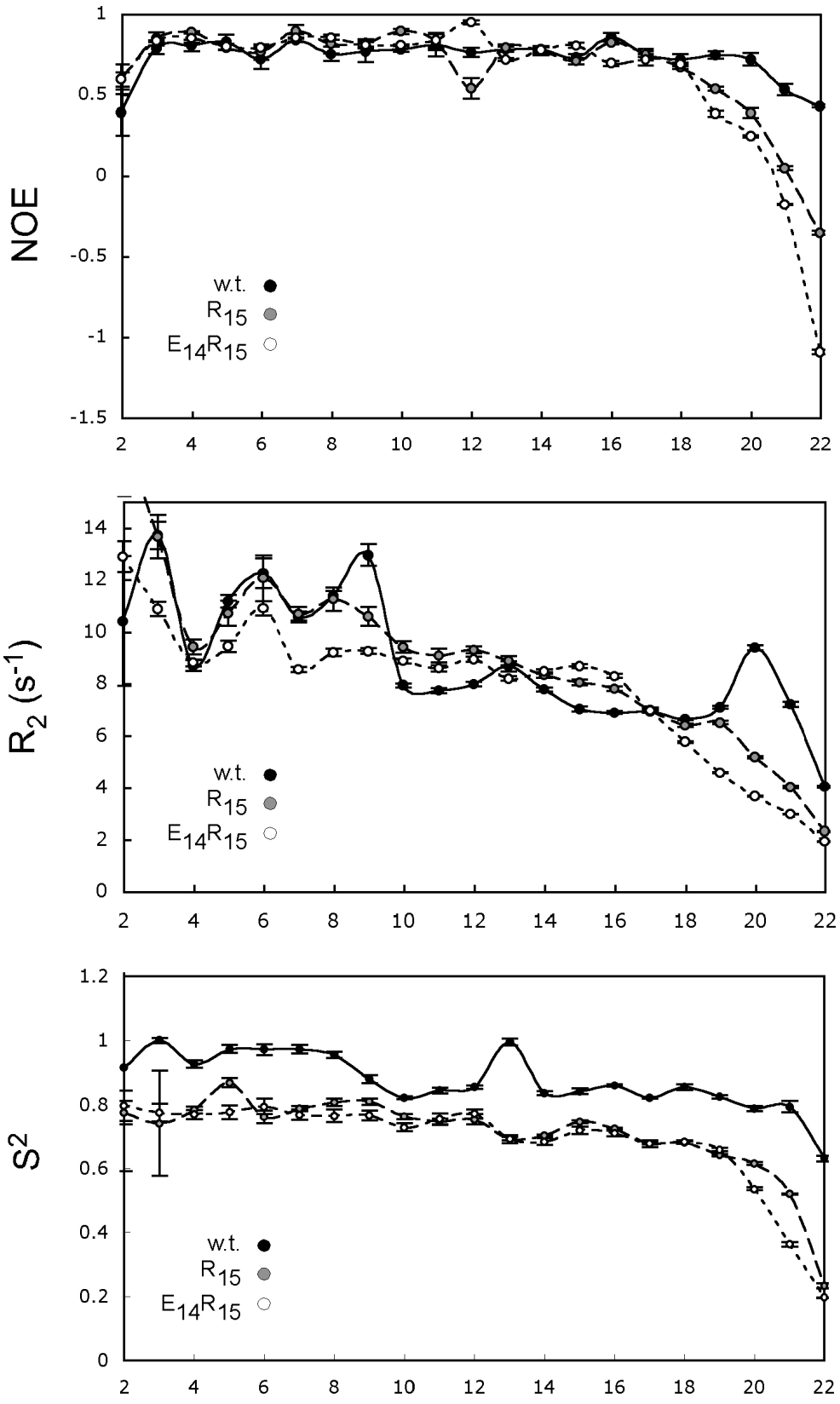

5
4
3
2
1
0
-1
ΔδCα (ppm)
w.t.
R15
E14R15
1
3
5
7
9
11
13
15
17
19
21
Residue number

Figure 4.4 Peptide backbone relaxation measurements. Measurements of backbone $\{^1H\}$-^{15}N NOE (top), ^{15}N R_2 (middle), and the Model free order parameter S^2 (bottom). NOE values > 0.65 indicate an ordered peptide conformation; NOE values < 0 and R_2 values approaching 0 are characteristic of disordered peptide backbone conformations. S^2 Model free order parameters were determined from ^{15}N spin relaxation measurements as described in experimental procedures. Smoothed line plots are shown to guide the eye.


NOE
w.t.
R15
E14R15
R2 (s-1)
S2
Residue Number

Supplemental Figures and Tables

Figure S4.1 R_{ex} values for the wild-type, R_{15}, and $E_{14}R_{15}$ peptide-λboxB$_R$ complexes. Positive R_{ex} values are observed in residues 19-22 of the wild-type and R_{15} peptide complexes, but not in the $E_{14}R_{15}$ complex. Cross-relaxation experiments indicate that large R_{ex} values in the peptide N-terminus are attributable to anisotropic relaxation mechanisms. Model free order parameters for the peptides are tabulated in supplemental tables S4.7-S4.9.

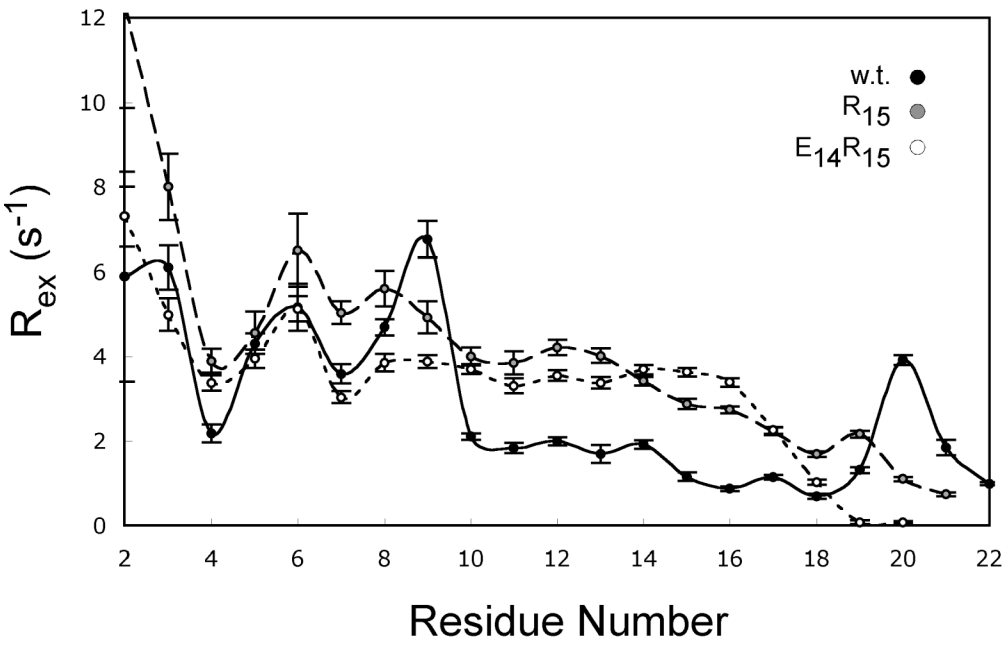

w.t.
R15
E14R15
Rex (s-1)
Residue Number
0
2
4
6
8
10
12
14
16
18
20
22

Table S4.1: NOE and Relaxation Rates for the wild-type λN peptide-boxB_R complex

Residue	NOE	R_1 (s^{-1})	R_2 (s^{-1})	η_{xy} (s^{-1})	η_z (s^{-1})
D2	0.388±0.142	2.385±0.201	10.397±2.491	3.885±2.638	1.338±1.589
A3	0.786±0.038	2.535±0.049	13.687±0.536	2.682±1.753	1.349±0.288
Q4	0.801±0.037	2.434±0.032	8.689±0.204	9.535±1.723	0.952±0.097
T5	0.824±0.045	2.556±0.033	11.171±0.248	8.052±2.074	0.878±0.166
R6	0.719±0.062	2.510±0.050	12.231±0.571	15.879±5.949	1.297±0.393
R7	0.835±0.009	2.533±0.033	10.564±0.212	8.307±0.644	1.197±0.071
R8	0.749±0.044	2.527±0.030	11.371±0.188	8.548±0.323	1.149±0.093
E9	0.765±0.068	2.326±0.037	12.928±0.417	9.676±1.284	1.473±0.433
R10	0.780±0.017	2.151±0.015	7.923±0.067	9.841±0.409	0.850±0.080
R11	0.804±0.067	2.232±0.022	7.725±0.106	6.502±0.079	1.037±0.013
A12	0.759±0.029	2.260±0.014	7.966±0.079	7.410±0.268	1.191±0.014
E13	0.779±0.018	2.628±0.032	8.656±0.189	7.249±0.077	1.341±0.192
K14	0.777±0.013	2.210±0.017	7.757±0.083	7.013±0.054	1.253±0.052
Q15	0.730±0.018	2.225±0.022	7.003±0.101	5.904±0.194	1.369±0.175
A16	0.855±0.026	2.273±0.010	6.877±0.042	7.028±0.136	1.279±0.060
Q17	0.744±0.036	2.167±0.009	6.900±0.039	7.261±0.166	1.149±0.028
W18	0.714±0.035	2.260±0.009	6.617±0.031	6.824±0.108	1.302±0.059
K19	0.743±0.024	2.179±0.013	7.070±0.059	5.753±0.037	1.227±0.040
A20	0.715±0.038	2.083±0.012	9.382±0.103	5.531±0.102	1.155±0.047
A21	0.528±0.035	2.095±0.014	7.201±0.105	5.590±0.036	0.954±0.016
N22	0.429±0.010	1.623±0.006	4.007±0.025	3.372±0.051	0.844±0.018
W18-Hε	0.674±0.023	1.876±0.012	5.314±0.048	4.556±0.051	0.736±0.015

Table S4.2: NOE and Relaxation Rates for the R_{15} peptide-boxB_R

Residue	NOE	R_1 (s^{-1})	R_2 (s^{-1})	η_{xy} (s^{-1})	η_z (s^{-1})
D2	0.593±0.093	2.016±0.091	17.714±2.503	0.355±0.372	
A3	0.843±0.040	1.866±0.045	13.638±0.832	1.619±1.360	
Q4	0.884±0.007	2.060±0.024	9.406±0.297	11.449±9.878	
T5	0.787±0.016	2.281±0.038	10.695±0.486	1.887±3.031	
R6	0.760±0.049	1.966±0.057	12.048±0.870	-1.728±3.018	
R7	0.890±0.039	2.051±0.022	10.674±0.264	4.509±0.860	
R8	0.811±0.058	2.131±0.034	11.242±0.436	21.334±20.139	
E9	0.819±0.023	2.138±0.026	10.576±0.364	23.185±15.229	
R10	0.890±0.015	1.998±0.020	9.401±0.209	5.520±2.722	
R11	0.827±0.052	1.977±0.026	9.066±0.271	11.939±3.032	
A12	0.536±0.064	1.993±0.016	9.288±0.145	8.577±1.147	
E13	0.789±0.018	1.833±0.014	8.874±0.174	6.604±1.262	
K14	0.771±0.028	1.859±0.009	8.325±0.104	9.973±0.843	
R15	0.705±0.020	1.972±0.012	8.038±0.120	8.655±1.587	
A16	0.820±0.012	1.915±0.008	7.795±0.085	8.226±0.644	
Q17	0.749±0.016	1.791±0.009	6.979±0.091	8.082±0.621	
W18	0.670±0.013	1.800±0.007	6.375±0.062	4.143±0.243	
K19	0.533±0.013	1.695±0.008	6.482±0.072	6.372±0.097	
A20	0.385±0.030	1.625±0.006	5.155±0.051	2.904±0.157	
A21	0.044±0.011	1.373±0.006	3.992±0.040	4.629±0.121	
N22	-0.358±0.014	1.072±0.004	2.302±0.022	2.789±0.118	
W18-Hε	0.329±0.013	1.353±0.009	3.966±0.055	2.655±0.047	

Table S4.3: NOE and Relaxation Rates for the $E_{14}R_{15}$ peptide-boxB_R

Residue	NOE	R_1 (s^{-1})	R_2 (s^{-1})	η_{xy} (s^{-1})	η_z (s^{-1})
D2	0.592±0.045	2.074±0.115	12.865±0.604	2.691±1.159	
A3	0.830±0.006	1.951±0.068	10.871±0.273	-5.883±6.692	
Q4	0.846±0.030	2.025±0.043	8.795±0.139	7.421±0.899	
T5	0.795±0.017	2.041±0.057	9.434±0.213	8.495±0.748	
R6	0.788±0.007	2.048±0.063	10.894±0.267	9.097±0.256	
R7	0.851±0.024	2.002±0.042	8.544±0.106	7.180±0.526	
R8	0.850±0.014	2.022±0.053	9.196±0.148	7.772±0.723	
E9	0.803±0.019	2.026±0.040	9.240±0.119	7.328±0.656	
R10	0.806±0.002	1.910±0.034	8.867±0.101	9.253±0.522	
R11	0.835±0.045	1.996±0.046	8.581±0.122	8.342±1.004	
A12	0.946±0.012	2.037±0.035	8.932±0.103	9.003±0.272	
E13	0.714±0.009	1.829±0.033	8.182±0.089	11.977±7.268	
E14	0.774±0.019	1.809±0.029	8.465±0.083	8.638±0.356	
R15	0.802±0.015	1.903±0.031	8.659±0.089	7.182±0.234	
A16	0.691±0.015	1.877±0.026	8.284±0.073	7.421±0.147	
Q17	0.715±0.035	1.788±0.023	6.960±0.063	6.326±0.153	
W18	0.684±0.013	1.805±0.016	5.739±0.037	5.023±0.063	
K19	0.381±0.019	1.794±0.015	4.543±0.026	3.626±0.080	
A20	0.243±0.007	1.467±0.009	3.657±0.021	2.721±0.021	
A21	-0.182±0.002	1.346±0.008	2.945±0.016	1.859±0.034	
N22	-1.095±0.014	0.906±0.007	1.906±0.012	1.016±0.025	
W18-Hε	0.156±0.008	1.417±0.015	3.000±0.024	1.760±0.030	

Table S4.4: Chemical Shifts (^{1}H, ^{13}C, ^{15}N ppm) of wild-type λN peptide-boxB_R complex

Residue	NH1	N(H1)	Cα	Cβ	Others
M1					
D2	8.24	122.29	53.13	42.73	
A3	9.23	120.96	55.47	18.75	
Q4	8.89	120.49	58.22	26.81	
T5	8.43	118.48	67.56	67.99	
R6	7.96	120.18	60.13	31.27	
R7	8.07	122.02	59.92	29.57	
R8	8.35	119.40	59.71	30.85	
E9	7.61	117.02	59.50	29.78	
R10	8.00	120.34	59.71	31.06	
R11	7.65	119.12	57.16	31.06	
A12	7.30	123.45	56.10	19.17	
E13	8.94	118.24	59.71	29.36	
K14	7.59	119.22	59.29	33.39	
Q15	7.67	119.59	58.65	27.45	
A16	8.17	122.22	55.04	18.32	
Q17	7.99	117.53	58.22	28.72	
					NH 9.32
W18	7.71	121.17	61.20	29.36	N(H) 128.97
K19	7.98	119.69	58.22	32.97	
A20	7.55	120.13	53.56	18.75	
A21	7.32	120.41	52.07	20.02	
N22	7.22	123.65	54.62	41.03	

Table S4.5: Chemical Shifts (^{1}H, ^{13}C, ^{15}N ppm) of the R_{15} peptide-boxB_R

Residue	NH1	N(H1)	Cα	Cβ	Others
M1	8.75	124.33	55.25	33.18	
D2	8.75	124.41	54.19	41.45	
A3	8.60	124.78	53.76	19.38	
Q4	8.50	118.66	57.16	29.14	
T5	8.14	115.99	63.53	69.47	
R6	8.24	123.03	57.16	30.84	
R7	8.22	121.56	57.16	30.84	
R8	8.28	121.99	56.95	30.84	
E9	8.38	121.82	56.95	30.21	
R10	8.37	122.43	56.74	30.84	
R11	8.34	122.39	56.74	31.05	
A12	8.33	125.22	52.92	19.38	
E13	8.31	120.44	56.74	30.42	
K14	8.34	122.66	56.74	33.18	
R15	8.30	122.34	56.52	31.05	
A16	8.33	125.22	52.92	19.38	
Q17	8.28	119.18	56.31	29.57	
					NH
W18	8.10	122.37	57.80	29.78	N(H)
K19	7.84	123.74	56.10	33.81	
A20	8.07	125.52	52.70	19.59	
A21	8.22	123.62	52.49	19.59	
N22	7.93	122.73	54.83	40.61	

Table S4.6: Chemical Shifts (^{1}H, ^{13}C, ^{15}N ppm) of the $E_{14}R_{15}$ peptide-boxB_R

Residue	NH1	N(H1)	Cα	Cβ	Others
M1			55.46	33.39	
D2	8.17	122.75	53.13	42.73	
A3	9.28	120.99	55.68	18.75	
Q4	8.91	120.48	58.43	26.60	
T5	8.39	119.04	67.56	68.62	
R6	8.06	120.00	60.13	31.06	
R7	8.15	121.53	59.71	29.78	
R8	8.56	120.96	59.71	30.84	
E9	8.09	118.69	59.50	29.36	
R10	8.29	122.32	59.71	29.99	
R11	8.63	119.34	59.92	30.63	
A12	7.85	121.10	55.46	18.11	
E13	8.01	120.38	59.28	29.36	
E14	8.18	119.99	59.28	29.57	
R15	7.84	117.78	59.28	30.84	
A16	7.67	121.49	54.19	18.53	
Q17	7.81	117.21	57.16	28.93	
					NH 10.11
W18	7.88	120.74	58.65	29.36	N(H) 129.15
K19	7.82	121.65	56.74	33.39	
A20	7.88	123.58	52.92	19.38	
A21	7.96	122.75	52.49	19.59	
N22	7.76	123.30	54.83	40.82	

Table S4.7: Model Free order parameters of the wild-type λN peptide-boxB_R complex

Residue	S^2	Error	τ_{ex} (ps)	Error	R_{ex} (s^{-1})	Error	Model #
M1							
D2	0.9161	0.3236			5.887	2.476	4
A3	1	0.0079			6.097	0.5288	3
Q4	0.9254	0.0117			2.176	0.2121	3
T5	0.9727	0.0123			4.296	0.2417	3
R6	0.9718	0.0173			5.158	0.5646	3
R7	0.9722	0.0129			3.59	0.225	3
R8	0.9558	0.0103			4.687	0.1896	3
E9	0.8799	0.0126			6.771	0.4352	3
R10	0.8205	0.0049			2.101	0.0713	3
R11	0.8441	0.0093			1.825	0.1213	3
A12	0.8546	0.0054			1.992	0.0973	3
E13	0.9949	0.0107			1.689	0.2115	3
K14	0.8357	0.0057			1.917	0.0952	3
Q15	0.8415	0.0087	22.90	9.39	1.154	0.1001	4
A16	0.8592	0.0038			0.8706	0.0553	3
Q17	0.8209	0.0038			1.145	0.0487	3
W18	0.8547	0.0082	34.33	21.47	0.693	0.057	4
K19	0.8239	0.0051			1.311	0.0664	3
A20	0.7881	0.00865	21.67	15.08	3.912	0.1163	4
A21	0.7932	0.0171	91.55	23.67	1.847	0.1875	4
N22	0.4490	0.00939			0.9887	0.0358	5
(Model 5) parameters			τ_s	Error			

Table S4.8: Model Free order parameters for the R_{15} peptide-boxB_R complex

Residue	S^2	Error	τ_{ex} (ps)	Error	R_{ex} (s^{-1})	Error	Model #
M1							
D2	0.7745	0.0367	61.73	34.3	12.3	2.411	4
A3	0.7419	0.1645			8.01	0.7862	3
Q4	0.7826	0.0095			3.893	0.2821	3
T5	0.8681	0.0143			4.56	0.5007	3
R6	0.7612	0.0207			6.508	0.8614	3
R7	0.7871	0.0077			5.029	0.2599	3
R8	0.8063	0.0124			5.604	0.4249	3
E9	0.8089	0.0102			4.918	0.3731	3
R10	0.7622	0.0078			3.994	0.2089	3
R11	0.7475	0.0091			3.842	0.2769	3
A12	0.7535	0.0157	71.35	21.3	4.205	0.1825	4
E13	0.6937	0.0051			4.017	0.1769	3
K14	0.7029	0.0031			3.414	0.1011	3
R15	0.7459	0.005	20.31	5.94	2.872	0.1238	4
A16	0.7242	0.0032			2.733	0.0863	3
Q17	0.6784	0.0032			2.223	0.0877	3
W18	0.6807	0.0035	22.53	3.24	1.688	0.064	4
K19	0.6411	0.0037	43.13	2.29	2.157	0.0796	4
A20	0.6149	0.0053	61.54	4.51	1.098	0.0569	4
A21	0.5197	0.0023	77.67	1.27	0.745	0.0433	4
N22	0.2325	0.00843	0.4742	0.0168			5
(Model 5) parameters			τ_s	Error			

Table S4.9: Model Free order parameters for the $E_{14}R_{15}$ peptide-boxB_R complex

Residue	S^2	Error	τ_{ex} (ps)	Error	R_{ex} (s^{-1})	Error	Model #
M1							
D2	0.7967	0.0478	70.18	33.69	7.302	0.7041	4
A3	0.7753	0.0271			4.986	0.3878	3
Q4	0.7692	0.014			3.376	0.1774	3
T5	0.7765	0.0211			3.946	0.2203	3
R6	0.7929	0.0241			5.122	0.298	3
R7	0.7683	0.016			3.033	0.1486	3
R8	0.7647	0.0204			3.848	0.208	3
E9	0.7664	0.0144			3.878	0.1506	3
R10	0.7284	0.0125			3.699	0.1194	3
R11	0.7547	0.018			3.306	0.1679	3
A12	0.7704	0.0134			3.547	0.125	3
E13	0.6923	0.0126	13.63	1.91	3.374	0.131	4
E14	0.6839	0.0113			3.685	0.1128	3
R15	0.7198	0.0113			3.626	0.1034	3
A16	0.7096	0.009	20.56	3.37	3.383	0.1005	4
Q17	0.6774	0.0108	12.67	7.16	2.251	0.0854	4
W18	0.6824	0.0062	19.6	2.82	1.03	0.0599	4
K19	0.6587	0.0061	81.14	4.2	0.075	0.0499	4
A20	0.5358	0.0037	65.53	1.16	0.0776	0.0347	4
A21	0.3625	0.0066	0.3	0.025			5
N22	0.1954	0.0011	97.15	1.26			2
(Model 5) parameters			τ_s	Error			