RNA architecture dictates the conformations of a bound peptide

Xiaomei Ye[1], Andrey Gorin[1], Ronnie Frederick[1], Weidong Hu[1], Ananya Majumdar[1], Weijun Xu[1], George McLendon[2], Andrew Ellington[3] and Dinshaw J Patel[1]

Background: The biological function of several viral and bacteriophage proteins, and their arginine-rich subdomains, involves RNA-mediated interactions. It has been shown recently that bound peptides adopt either β-hairpin or α-helical conformations in viral and phage peptide–RNA complexes. We have compared the structures of the arginine-rich peptide domain of HIV-1 Rev bound to two RNA aptamers to determine whether RNA architecture can dictate the conformations of a bound peptide.

Results: The core-binding segment of the HIV-1 Rev peptide class II RNA aptamer complex spans the two-base bulge and hairpin loop of the bound RNA and the carboxy-terminal segment of the bound peptide. The bound peptide is anchored in place by backbone and sidechain intermolecular hydrogen bonding and van der Waals stacking interactions. One of the bulge bases participates in U•(A•U) base triple formation, whereas the other is looped out and flaps over the bound peptide in the complex. The seven-residue hairpin loop is closed by a sheared G•A mismatch pair with several pyrimidines looped out of the hairpin fold.

Conclusions: Our structural studies establish that RNA architecture dictates whether the same HIV-1 Rev peptide folds into an extended or α-helical conformation on complex formation. Arginine-rich peptides can therefore adapt distinct secondary folds to complement the tertiary folds of their RNA targets. This contrasts with protein–RNA complexes in which elements of RNA secondary structure adapt to fit within the tertiary folds of their protein targets.

Addresses: [1]Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA. [2]Chemistry Department, Princeton University, Princeton, NJ 08544, USA. [3]Institute of Cell and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA.

Correspondence: Dinshaw J Patel
E-mail: pateld@mskcc.org

Key words: adaptive-binding, bound peptide secondary structures, HIV-1 Rev peptide, peptide-binding RNA tertiary architectures, RNA aptamers

Received: **24 May 1999**
Revisions requested: **18 June 1999**
Revisions received: **24 June 1999**
Accepted: **8 July 1999**

Published: **13 August 1999**

Chemistry & Biology September 1999, **6**:657–669
http://biomednet.com/elecref/1074552100600657

1074-5521/99/$ – see front matter

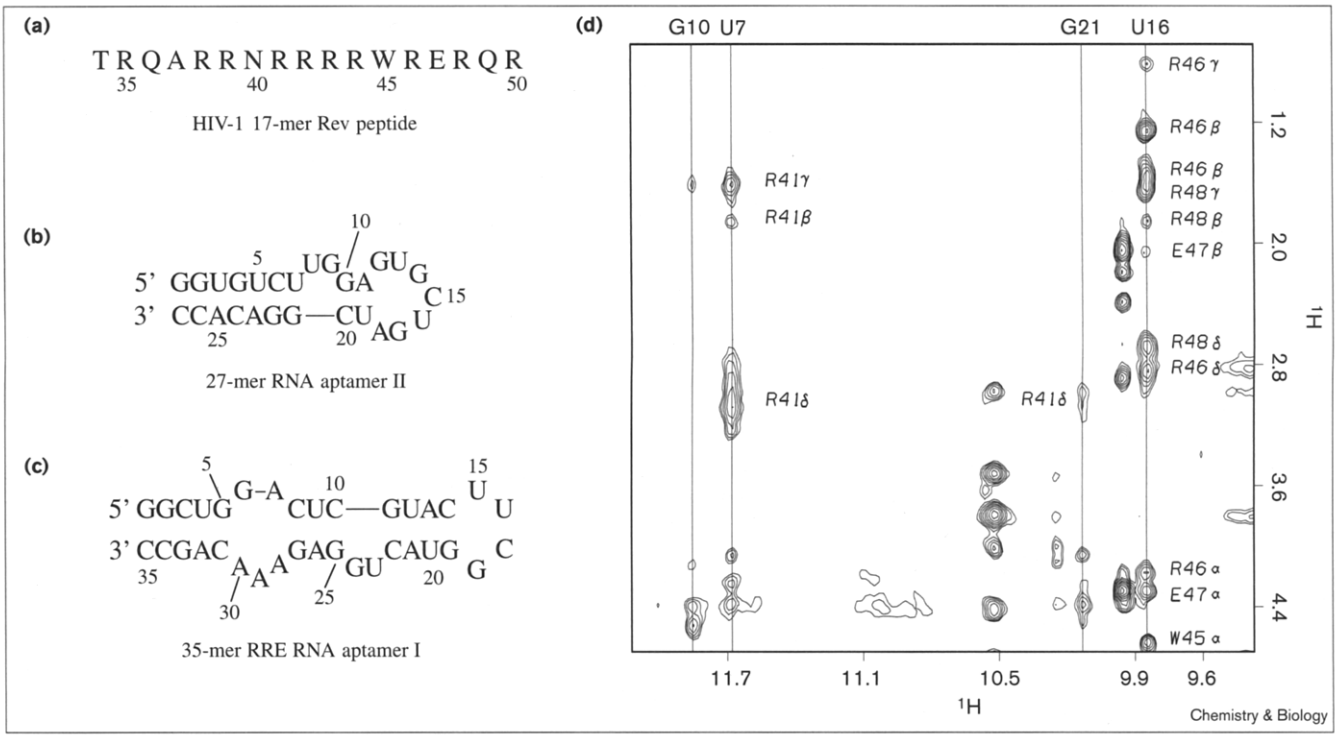
Introduction

The regulation of the lentiviral replication cycle is controlled by the RNA-binding Tat (*trans*-activator protein) and Rev (regulator of viral expression) proteins [1,2]. The HIV-1 Rev protein binds to the Rev response element (RRE) RNA located within the transcript of the *env* (envelope) gene [3–5] and mediates the export of intron-containing messenger RNAs that encode viral structural proteins to the cytoplasm. There is considerable interest in elucidating the structure of the complex between the HIV-1 Rev protein and its high affinity RRE RNA target as a prelude to the structure-based design of inhibitors that could interfere with this process, which is critical for the maintenance of the viral life cycle.

The HIV-1 Rev–RRE system has been approached biochemically by investigating its modular interacting components [6,7]. We know that an arginine-rich basic peptide segment of HIV-1 Rev extending from residues Thr34–Arg50 (Figure 1a) binds to the stem–loop IIB domain of the RRE RNA with nanomolar affinity and high specificity [8]. Solution structures are available for this complex [9] and a related complex [10] of the basic Rev peptide with an RNA aptamer [11] (Figure 1c) that is a higher-affinity binding variant of stem–loop IIB RRE RNA. The Rev peptide undergoes an adaptive transition to an α-helical conformation [12] and penetrates deeply into the major groove of the RNA centered around a zippered up asymmetric internal bubble in both complexes [9,10].

Two distinct families of RNA aptamers have been identified that target the arginine-rich HIV-1 Rev peptide (Figure 1a) with high specificity and nanomolar affinity. One of these, designated 35-mer RNA aptamer I (family I), has a secondary structure [11] that contains an asymmetric internal loop binding site (Figure 1c) analogous to the stem–loop IIB of the RRE RNA. The other family, designated 27-mer RNA aptamer II (family II), has a very different secondary structure [13] composed of a seven-residue hairpin loop and a two-base bulge (Figure 1b). The availability of two RNA aptamers that are targeted by a common HIV-1 Rev peptide should allows us to address the following fundamental question related to adaptive structural transitions of arginine-rich peptides [14] targeted to binding pockets associated with RNA tertiary structure. Does the bound Rev peptide adopt similar conformations on both RNA aptamer targets or does it adopt different folds defined

Figure 1



(a) Sequence and numbering of HIV-1 17-mer Rev peptide (using single-letter amino acid code). Sequence, secondary structure and numbering of **(b)** the 27-mer RNA aptamer II used in this study and **(c)** the 35-mer RRE RNA aptamer I studied previously [10]. **(d)** An expanded NOESY contour plot outlining NOEs between imino protons (9.6 to 12.0 ppm) and amino acid Hα and sidechain protons (0.6–4.8 ppm) of the bound RNA in the 17-mer Rev peptide–27-mer RNA aptamer II complex in H_2O buffer, pH 6.2 at 10°C. The labeled cross peaks involve intermolecular NOEs between the imino protons of G10, U7, G21 and U16 of the bound RNA aptamer and the Hα, Hβ, Hγ and Hδ protons of Trp45, Arg46, Glu47 and Arg48 of the bound Rev peptide.

by potentially distinct binding-pocket architectures associated with the two classes of RNA aptamer targets?

We report below on the solution structure of the 17-mer HIV-1 Rev peptide bound to the 27-mer RNA aptamer II elucidated using a combination of nuclear magnetic resonance (NMR) and computational approaches. This structure is compared with the structure of the 17-mer HIV-1 Rev peptide bound to the 35-mer RNA aptamer I reported previously [10]. Our studies establish that the same HIV-1 Rev peptide adopts an extended conformation in the complex with an RNA aptamer of class II type (Figure 2b; this study), whereas it has an α-helical conformation [9,10,12] in the complex with an RNA aptamer of class I type (Figure 2a) [10]. These studies establish that the same peptide can undergo adaptive structural transitions to distinct bound conformations dependent on the binding-pocket architecture of the RNA target.

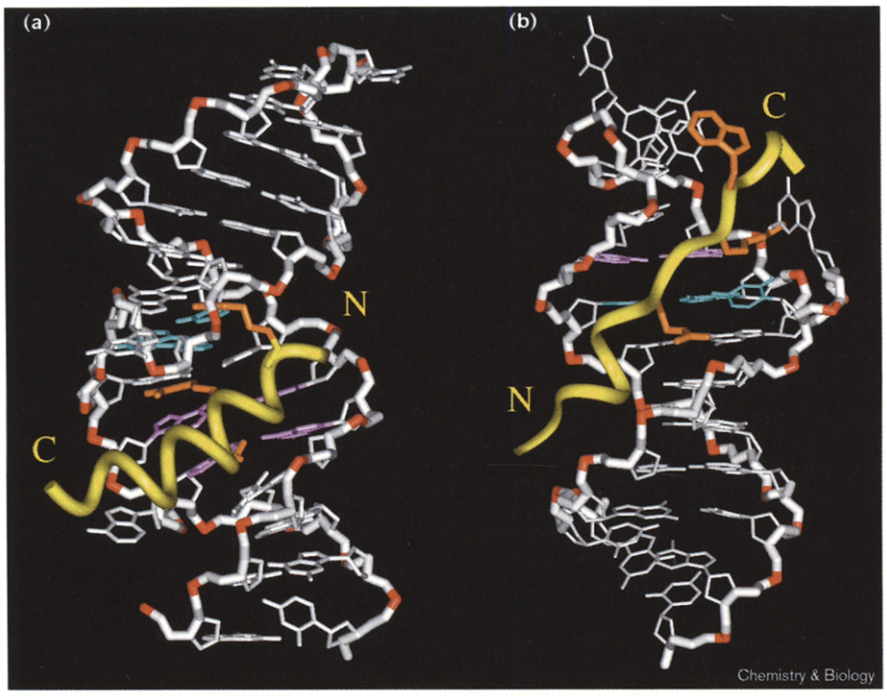
Results

Complex formation

The solution structure of the 17-mer HIV-1 Rev peptide complexed to the 27-mer RNA aptamer II (Figure 1b) was characterized using multidimensional heteronuclear NMR spectroscopy [15–17]. We have recorded the exchangeable proton spectra (8.5–14.5 ppm) of the free RNA aptamer on gradual addition of one equivalent of Rev peptide in H_2O buffer, pH 6.2 at 10°C. Imino proton resonances from both the free RNA aptamer and the complex were observed on substoichiometric addition of the Rev peptide, consistent with slow exchange between the free and bound forms and reflective of the high affinity associated with complex formation. The observation of a unique set of imino proton resonances of the bound RNA between 10 and 14 ppm and a unique set of amide proton resonances of the bound Rev peptide between 8.5 and 10.0 ppm is indicative of a predominant peptide–RNA conformation on complex formation.

We also generated complexes of the 17-mer Rev peptide with two additional RNA aptamers in the same class II family [13] and recorded their NMR spectra. The secondary folds of 29-mer RNA aptamer II and 36-mer RNA aptamer II differed from the 27-mer RNA aptamer II in the number of base pairs separating the common two-base bulge and 7-mer hairpin loop segments. Most importantly, they differed in the residues at positions 13 and 15 of the

Figure 2



A comparison of representative refined structures of the complexes of the HIV-1 17-mer Rev peptide with **(a)** the 35-mer RRE RNA aptamer I (Figure 1c) reported previously [10] and **(b)** the 27-mer RNA aptamer II (Figure 1b) reported in this study. The peptide backbones are represented by yellow ribbons with key sidechains in orange. The RNAs are white, backbone phosphorus atoms are red, base mismatches are pink and base triples are cyan. The peptide backbone adopts an α-helical fold in (a) and is extended in (b).

RNA hairpin loop and were therefore of considerable help in the spectral assignments of these residues, and other residues in overlap regions of the two-dimensional data sets.

The 17-mer HIV-1 Rev peptide contains ten arginine residues, and one challenge was to identify and distinguish between these amino acids with their long sidechains in the spectra of the complex. We therefore also generated and recorded NMR spectra for complexes of the 27-mer RNA aptamer II with analogs of the 17-mer Rev peptide containing amino-acid substitutions and/or deletions. These included a 15-mer peptide lacking the last two residues and Arg39→Ala (R39A) and Arg44→Asn (R44N) substitutions, and an 11-mer peptide lacking the first six residues and Arg42→Ala (R42A) and Arg43→Ala (R43A) substitutions. Both these mutant and truncated Rev peptides bound the 27-mer RNA aptamer II under slow-exchange conditions with spectra that corresponded to a predominant conformation for the complex.

Structure determination

The solution structure of the HIV-1 17-mer Rev peptide–27-mer RNA aptamer II complex was characterized using multidimensional NMR spectroscopy. Resonances from the bound RNA were assigned from the sample containing uniformly $^{13}C,^{15}N$-labeled 27-mer RNA, and augmented where necessary with complexes generated with unlabeled 29-mer and 36-mer RNA aptamer II analogs. Resonances from the bound Rev peptide were assigned from samples containing unlabeled and uniformly $^{13}C,^{15}N$-labeled 17-mer peptide, selectively $^{13}C,^{15}N$-labeled (at Arg41 and Arg45) 17-mer peptide, and augmented where necessary with complexes generated with unlabeled 15-mer and 11-mer peptide analogs.

We identified 49 intermolecular nuclear Overhauser effects (NOEs) between the 17-mer Rev peptide and the 27-mer RNA aptamer II in the complex (Table 1). Several of these intermolecular NOEs involve RNA–aptamer-exchangeable imino protons and examples of these between the imino protons of G10, U7, G21 and U16 and nonexchangeable arginine sidechain protons are shown in an expanded NOESY contour plot of the complex in H_2O buffer at 10°C (Figure 1d). The majority of these NOEs span the Arg39–Arg48 segment of the bound Rev peptide and the U7–C20 segment of the bound RNA aptamer. The largest number of intermolecular NOEs involve the peptide sidechains of Arg41 and Trp45 and the RNA base and sugar protons of U16 in the complex (Table 1). Of specific interest are the intermolecular NOEs between the sidechain of Arg41 and the RNA U7 and G10 residues and between the sidechains of the Trp45–Arg46–Glu47–Arg48 segment and the RNA U16 residue in the complex (Table 1). We did not detect intermolecular NOEs involving assigned arginine sidechain NHε protons in the complex, possibly because of the increased line width of these exchangeable resonances (Table 1). We were unable to assign any arginine sidechain $NH_2\eta$ protons because of increased line widths and resonance overlap, and were not, therefore, in a position to monitor

Table 1

Intermolecular NOEs in the 17-mer Rev peptide–27-mer RNA aptamer complex.

Amino acid		Nucleotide
Thr34	γCH_3	U3 (H5, H6)
Ala37	βCH_3	U5 (H5, H6)
Arg38	δCH_2	U5 (H5)
Arg39	βCH_2, γCH_2	U19 (H5)
	δCH_2	U19 (H5), C20 (H5)
Asn40	βCH_2	U19 (H5)
	δNH_2	G5 (H8), G6 (H8)
Arg41	βCH_2	U7(NH3, H5)
	γCH_2	G10 (NH1), U7 (NH3, H5), U19 (NH3)
	δCH_2	G10 (NH1), U7 (NH3, H5), U8 (H5)
Arg44	δCH_2	U8(H1′)
Trp45	$CH\delta_1$	U16 (NH3)
	$CH\varepsilon_3$	U16 (NH3, H1′, H2″, H3′)
	$CH\zeta_2$	U16 (NH3, H5)
	$CH\zeta_3$	U16 (H5, H1′, H2″, H3′)
Arg46	NH, βCH_2, γCH_2	U16 (NH3)
	δCH_2	U16 (NH3, H1′)
Glu47	NH	U16 (NH3)
Arg48	NH	U16 (NH3)
	βCH_2, δCH_2,	U16 (NH3), U13 (H5)

intermolecular NOEs involving these exchangeable protons in the complex.

The solution structure of the 17-mer Rev peptide–27-mer-RNA aptamer II complex was solved using molecular dynamics computations guided by NOE distance restraints. We generated 60 starting structures with both the peptide and RNA in randomized alignments and positioned far apart (separated by 100 Å). The helical stem extending from G1–C6 and from G22–C27 outside the core-binding region was restricted to an A-helical conformation by dihedral angle and hydrogen-bond restraints during the computations. In addition, hydrogen-bonding restraints were used to restrain the experimentally identified sheared G12•A18 mismatch alignment within the core-binding region in the complex. The protocol outlined in the Materials and methods section involved initial torsional space dynamics at 20,000 K followed by cartesian space dynamics at 300 K. A subset of 22 refined structures of the complex were identified based on low NOE violation energies and exhibit a pairwise root mean square deviation (rmsd) value of 2.79 Å ± 0.47 for the central core-binding region of the complex (RNA U7–G21 and peptide Asn40–Arg48; Table 2). This high rmsd value reflects our inability to precisely define the majority of the sidechains (in contrast to the better defined backbone) of the bound peptide and the G9, U13 and C15 looped out residues of the bound RNA aptamer in the core-binding region of the complex. The pairwise rmsd

Table 2

Restraints and refinement statistics for the HIV-1 Rev peptide–RNA aptamer complex.

NMR restraints in complex	
RNA (G1–C27)	
RNA distance restraints	204
Peptide (Thr34–Arg50)	
Peptide distance restraints	107
Peptide–RNA	
Intermolecular distance restraints	49
Structure statistics in complex	
NOE violations	
Number > 0.2 Å	2.7 ± 0.9
Maximum violations (Å)	0.38 ± 0.01
rmsd of violations (Å)	0.039 ± 0.002
Deviations from the ideal covalent geometry	
Bond length (Å)	0.016 ± 0.0001
Bond angle (°)	2.25 ± 0.031
Impropers (°)	0.63 ± 0.08
Pairwise rmsd (Å) among 22 refined structures (all heavy atoms)	
Core complex: RNA (U7–G21) and peptide (Asn40–Arg48, backbone and sidechains)	2.79 ± 0.47
Core complex: RNA (U7–U8, G10–G12, G14 and U16–G21) and peptide (Asn40–Arg48, backbone only)	1.55 ± 0.38

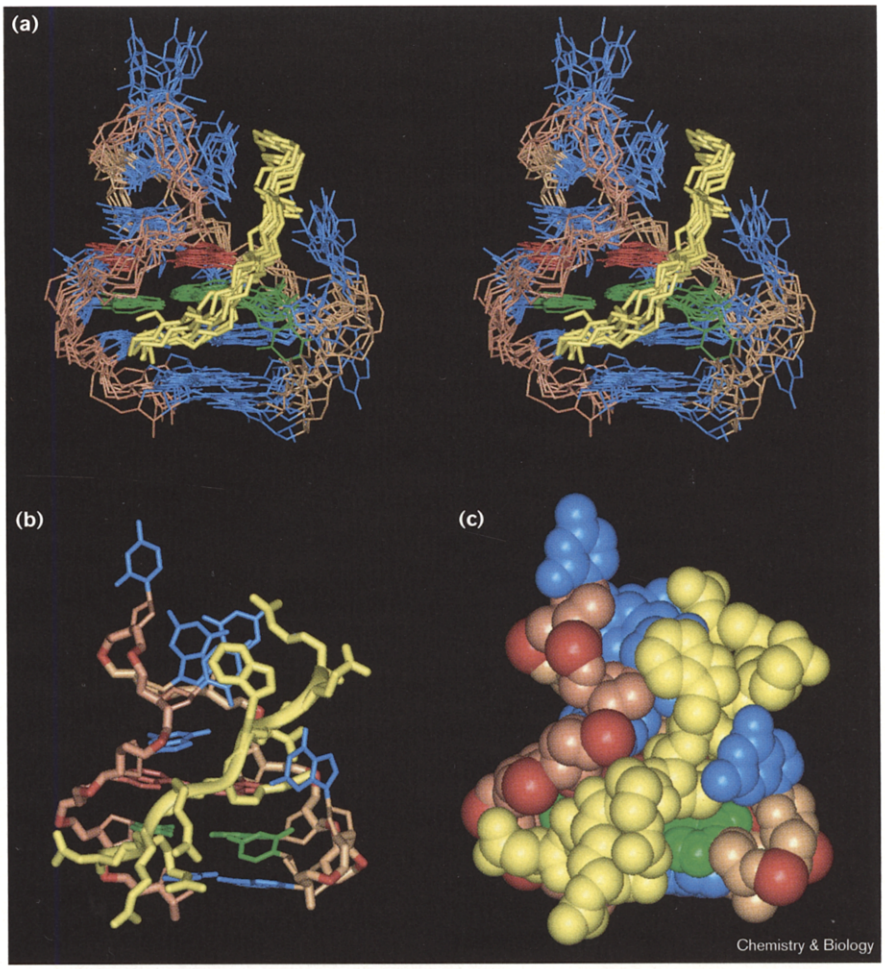
value of the complex decreases to 1.55 Å ± 0.38 for the core-binding region comprising well-defined U7–U8, G10–G12, G14 and U16–G21 RNA residues and backbone atoms only for the Asn40–Arg48 segment of the peptide in the complex (Table 2).

Structure analysis

A stereo view of eight superpositioned distance-refined structures of the core-binding region of the HIV-1 Rev peptide–RNA aptamer II complex looking normal to the helix axis and into the major groove is shown in Figure 3a. Stick and space-filling views of a representative refined structure of the core-binding region of the complex including peptide sidechains are shown in Figure 3b,c. The Asn40–Arg48 segment of the Rev peptide binds in an extended conformation to the major groove of the bulge-containing stem segment and attached hairpin-loop segment extending from U7–G21 of the RNA aptamer. The intermolecular contacts in the refined structures of the HIV-1 Rev peptide–RNA aptamer II complex, along with their counterparts in the refined structures of the HIV-1 Rev peptide–RNA aptamer I complex published previously [10], are summarized schematically in Figure 4.

Eight superpositioned and one representative refined structure of the seven-residue hairpin loop extending from G12–A18 of the bound RNA aptamer in the

Figure 3



(a) A stereo view of eight superpositioned refined structures of the 17-mer Rev peptide–27-mer RNA aptamer II complex. The peptide backbones without sidechains from residues Asn40–Arg48 are shown in yellow. The RNA aptamers are shown from residues U7–G21. The backbone is orange, the base and sugar rings are blue except for the sheared G12•A18 mismatch, which is red, and the U8•(A11•U19) triple, which is green. **(b)** Stick and **(c)** space-filling views of one representative structure of the 17-mer Rev peptide–27-mer RNA aptamer complex. The sidechains of the bound peptide are included in this view. The components of the complex and the color coding are the same as in (a).

complex are shown in Figure 5a,b, respectively. These views emphasize which residues of the bound RNA hairpin loop are well defined in the refined structures of the complex. Eight superpositioned and one representative refined structure of the Trp45–Arg48 segment of the bound Rev peptide in the complex are shown in Figure 5c,d. These views emphasize that the peptide backbone and Trp45 sidechain are better defined than the sidechains of Arg46, Glu47 and Arg48, amongst the refined structures of the complex.

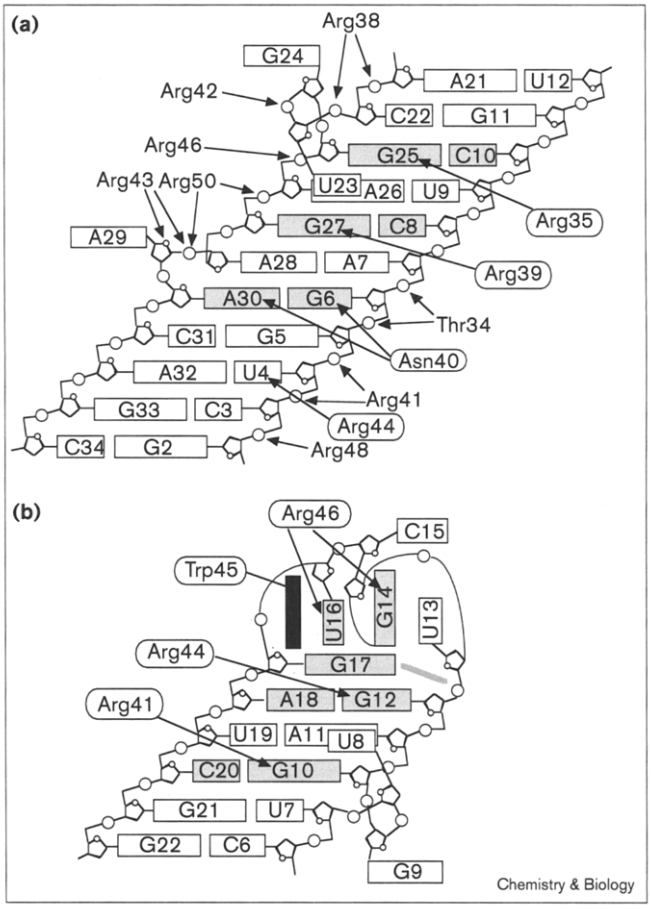
Several patterns of intermolecular interactions can be identified at the interface between the Rev peptide and the RNA aptamer within the core-binding segment of the complex. The orientations of the Trp45 and Arg48 sidechains of the Rev peptide relative to the G14 and U16 bases at the tip of the RNA hairpin loop in the complex are shown in Figure 6a,b. The interaction between the Watson–Crick edge of U16 at the tip of the RNA hairpin loop and the peptide backbone of residues Arg46 and Arg48 in the complex are shown in Figure 6c,d. Also shown is the orientation of the sidechain of Arg46 relative to the RNA backbone on either side of G12. The alignment of the Arg41 sidechain of the Rev peptide relative to the major groove edge of G10 and their positioning next to the buckled U8•(A11•U19) base triple in the complex are shown in Figure 6e,f.

Discussion

Quality of refined structures

We have not been able to determine a high-resolution structure of the entire HIV-1 17-mer Rev peptide–27-mer RNA aptamer II complex because of the less than optimal quality of the NMR data sets. Our structure determination was handicapped by severe overlap of Hα and sidechain proton resonances of the bound Rev peptide that could only be partially overcome with two-dimensional and three-dimensional experiments on samples of the complex labeled with uniformly ^{15}N-labeled and ^{13}C,^{15}N-labeled peptide. Several of the peptide protons in the complex were also broader than others, further complicating the spectral analysis. These limitations were reflected in our

Figure 4



Schematic summaries of structural features and intermolecular contacts (indicated by arrows) in the solution structures of **(a)** the 35-mer RRE RNA aptamer I (Figure 1c) reported previously [10] and **(b)** the 27-mer RNA aptamer II (Figure 1b) reported in this study. Bases are represented by rectangles, sugars by pentagons and phosphate groups by small circles. The amino acids involved in specificity-determining intermolecular interactions are enclosed in ovals.

inability to define precisely the majority of the peptide sidechains in the core-binding segment of the complex. There was also severe overlap of RNA sugar resonances of the stem segment (G1–C6 and G22–C27) in the complex. This required the inclusion of torsional and hydrogen-bonding restraints to define this helical segment located outside the core-binding region of the complex during the computations. Our structural analysis and subsequent discussion of the core-binding segment of the HIV-1 Rev peptide–RNA aptamer complex will, therefore, focus solely on those aspects of the structure that are unambiguously defined and merit further consideration.

Alignment of the Rev peptide within the RNA aptamer binding pocket

The 17-mer Rev peptide binds within the RNA major groove on complex formation, with the binding site spanning the helical and hairpin loop segments of the 27-mer RNA aptamer II. The carboxyl terminus of the Rev peptide is directed towards the hairpin loop with the best defined interface spanning the Asn40–Arg48 segment of the 17-mer peptide and U7–G21 segment of the 27-mer RNA aptamer II (Figure 3a). The RNA and peptide backbone are better defined than the peptide sidechains within this core-binding segment of the complex. The peptide in an extended conformation is embedded within the walls of the major groove of the stem and adjacent hairpin loop (Figure 3b), with several sequence and structure specific intermolecular contacts within the core-binding segment (Figure 4b) contributing to the specificity and affinity of complex formation.

Adaptive conformational transitions on complex formation

Both the Rev peptide and the RNA aptamer undergo adaptive conformational transitions on complex formation. This is readily apparent from changes in the NMR spectra of the exchangeable proton region where new imino resonances are observed between 9.5 and 10.5 ppm after the free RNA aptamer is converted to a complex with the Rev peptide. We also observe distinct amide proton chemical shifts between 8.5 and 10.0 ppm for the bound Rev peptide on complex formation.

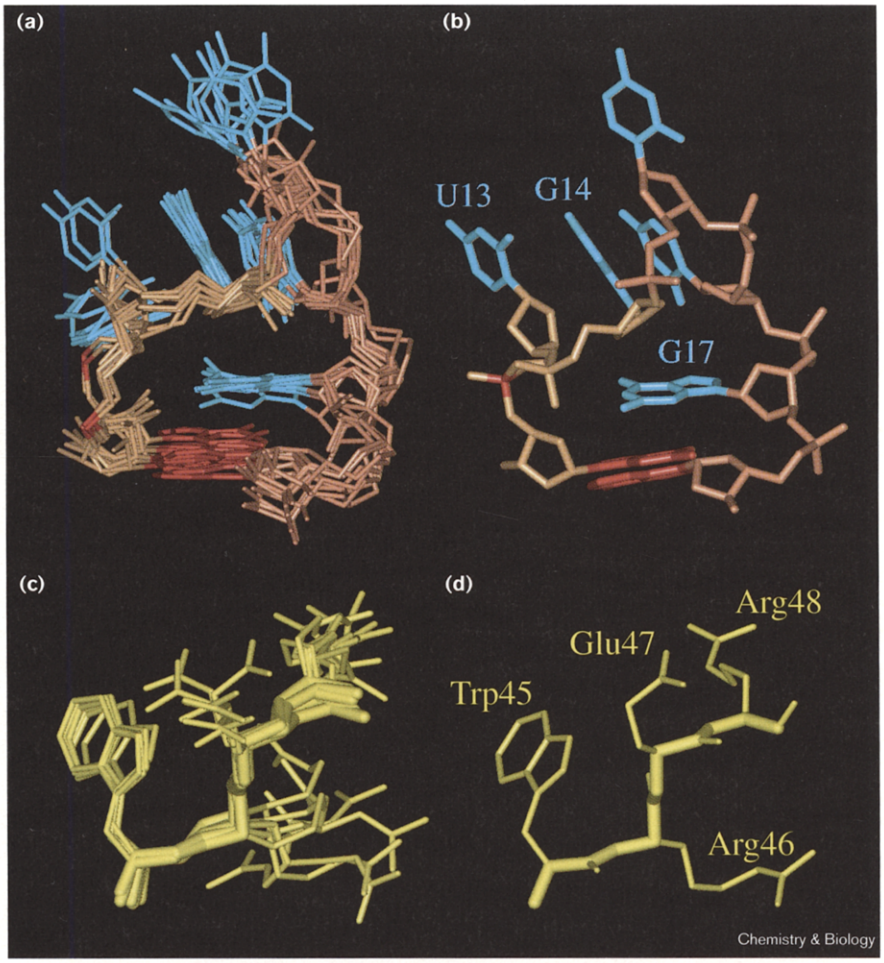
The free Rev peptide has been shown to be predominantly unstructured (≈10% α helicity by circular dichroism [12]) in solution and probably rapidly interconverts between an ensemble of unstructured states. By contrast, the same Rev peptide, when bound to a class II RNA aptamer, adopts a unique extended conformation. The distinct extended conformation reflects the linear threading of the peptide through the RNA recognition pocket.

Structural features of the bulge residues in the bound RNA

The U8 and G9 bulge residues are separated from the seven-residue hairpin loop by two Watson–Crick base pairs in the secondary fold of the RNA aptamer II (Figure 1b). These two bulge residues make a critical contribution to the generation of the peptide-binding pocket on the RNA. G9 adopts an alignment with its purine ring approximately parallel to the helix axis. G9 acts as a flap and makes an important contribution to anchoring the peptide within the walls of the major groove (Figure 3b). By contrast, the U8 bulge base participates in a U8•(A11•U19) base triple alignment on complex formation. The resulting stacking of the U8•(A11•U19) base triple over the G10•C20 base pair is consistent with the observed nonsequential NOE between H8 proton of G10 and the H1′ proton of U8 in the complex. The defined positioning of the U8 and G9 bulge residues is accompanied by an S-shaped fold for the U7-U8-G9-G10 segment of the RNA backbone, which is accompanied by mutual stacking of the junctional base pairs of the adjacent stem segments on complex formation.

Figure 5

Views of **(a,c)** eight superpositioned refined structures and **(b,d)** one representative refined structure of particular segments of the 17-mer Rev peptide–27-mer RNA aptamer II complex. (a,b) The bound RNA segment from G12–A18. This view emphasizes the potential intermolecular hydrogen-bonding interaction between donor hydrogens along the Watson–Crick edge of G17 and the phosphate oxygens at the G12–U13 step (phosphorus in red). Note that G14 is partially sandwiched between U13 and U16 within the loop segment of the complex. (c,d) The bound peptide segment from Trp45–Arg48. Certain amino acid sidechains (Trp45) are better defined than others (Arg46).



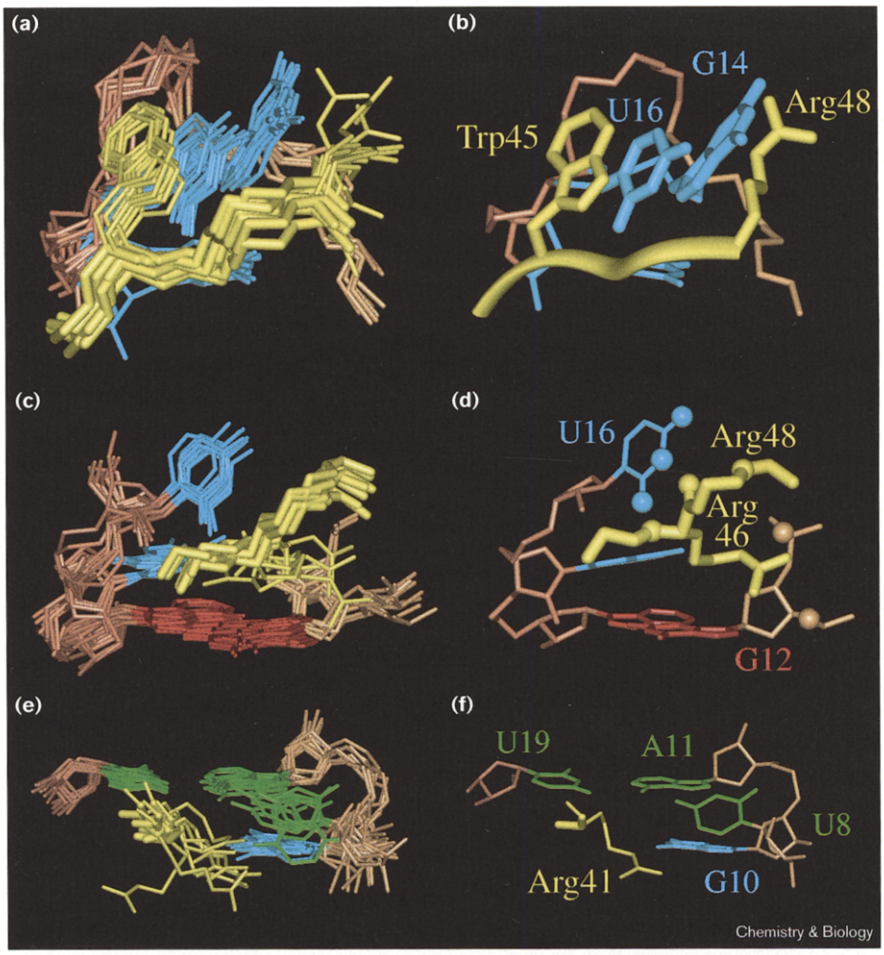
U8•(A11•U19) base triple in bound RNA

The U8 is positioned in the major groove and is inclined relative to the Watson–Crick A11•U19 pair on formation of the U8•(A11•U19) base triple amongst the distance refined structures of the complex (Figure 6e). Our inability to precisely define the position of U8 originates in the availability of limited markers on this residue because the base H5 and H6 protons are broad and we have been unable to assign the imino proton of U8 in the spectrum of the complex. It is interesting that well-defined approximately planar U•(A•U) triples have been identified for the bovine immunodeficiency virus (BIV) Tat peptide–TAR RNA [18] and HIV-1 Rev peptide–RNA aptamer I [10] complexes, where the third-strand uracil imino proton could be assigned and its alignment along the Hoogsteen edge of the adenine could be defined by an NOE between its imino proton and the H8 proton of the adenine [10,18]. By contrast, buckled U•(A•U) triples have been proposed for argininamide–HIV-1 TAR [19,20], argininamide–HIV-2 TAR [21] and HIV-1 Tat peptide–TAR RNA [19,20] complexes, where the third strand uracil imino proton could not be identified in the spectrum of the complex.

Structural features of the hairpin loop in the bound RNA

The seven-residue hairpin loops of all three RNA aptamers in the class II family (27-mer, Figure 1b; 29-mer and 36-mer) adopt a common unique conformation on complex formation with the bound peptide. The hairpin-loop fold is well defined except for two looped-out pyrimidines amongst the refined structures of the 17-mer Rev peptide–27-mer RNA aptamer II complex (Figure 5a). The loop is closed by a conserved sheared G12•A18 mismatch pair that was identified on the basis of an NOE between the imino proton of G12 and the H8 proton of A18, a feature characteristic of sheared G•A mismatch alignments [22,23]. There is extensive stacking between the sheared G12•A18 mismatch pair and the junctional Watson–Crick A11•U19 base pair in the refined structures of the complex. Hydrogen-bonding restraints for the sheared G12•A18 mismatch pair were incorporated during the molecular dynamics computations, but their removal at the end of the computations followed by further refinement did not affect the structure of the bound hairpin loop or the number of NOE violations within this segment. The G17 base is stacked on the sheared

Figure 6



Views of **(a,c,e)** eight superpositioned refined structures and **(b,d,f)** one representative refined structure with emphasis on intermolecular contacts in the 17-mer Rev peptide–27-mer RNA aptamer II complex. (a,b) Intermolecular contacts involving the sidechains of Trp45 and Arg48 (yellow) of the Rev peptide and the G14 and U16 bases (cyan) of the RNA aptamer. (c,d) Intermolecular contacts between the peptide backbone (yellow) and the Watson–Crick edge of U16 (cyan) and between the sidechain of Arg46 (yellow) and the phosphate oxygens (orange) at the A11–G12–U13 step of the RNA backbone. The balls identify donor and acceptor atoms involved in potential intermolecular hydrogen-bonding interactions. (e,f) Intermolecular contacts between the G10 base (cyan) and the sidechain of Arg41 (yellow) positioned next to the U8•(A11•U19) base triple (green).

G12•A18 mismatch pair and its Watson–Crick edge is directed towards the phosphate oxygens of the G12–U13 step across the loop (Figure 5a,b). This alignment is defined by the distribution of intramolecular NOEs observed between the slowly exchanging imino proton of G17 and the sugar protons of the G12-U13-G14 segment across the hairpin loop in the complex. The G14 and U16 hairpin loop bases are better defined than their U13 and C15 loop counterparts amongst the refined structures of the complex (Figure 5a). The G14 base adopts a *syn* alignment consistent with the strong NOE between the H8 proton and its own H1′ proton in short mixing time NOESY data of the complex. One such example is the nonsequential NOE between the H8 protons of G14 and G18 in the NOESY spectrum of the complex. The G14 and U16 bases are stacked on each other, consistent with the observed nonsequential NOE between the H8 proton of G14 and the H1′ proton of U16 in the NOESY spectrum of the complex. The stacked G14 and U16 bases are aligned approximately orthogonal to the plane of G17 in the complex (Figures 5a,b). Such a relative alignment explains the upfield ring current shift of the H8 (6.64 ppm) and H1′ (3.86 ppm) protons of G14, which are positioned over the aromatic ring of G17 in the complex. The well-defined alignment of the G14 and U16 bases reflects their participation in intermolecular packing contacts with peptide residues in the complex (Figures 6a,b). By contrast, U13 and C15, which play no role in molecular recognition of the bound Rev peptide, are poorly defined amongst refined structures of the complex. It is not surprising, therefore, that U13 in 27-mer RNA aptamer II (Figure 1b) can be replaced by A13 in the 36-mer RNA aptamer II and that C15 in 27-mer RNA aptamer II (Figure 1b) can be replaced by G15 in the 29-mer RNA aptamer II. The residues 13 and 15 are the only two non-conserved positions within the seven-residue hairpin loops of the three class II RNA aptamers that bind Rev peptide with comparable affinity.

Structural features of the bound Rev peptide

The 17-mer Rev peptide adopts an extended conformation along its entire length when bound to the class II

RNA aptamers. This result is consistent with the strong intramolecular NOEs between Hα(i) and NH(i + 1) protons and weak NOEs between adjacent amide protons in the complex. Further, no continuous NH(i) to NH(i + 2), NH(i) to NH(i + 3) and Hα(i) to NH(i + 3) NOEs characteristic of an α-helical peptide conformation were observed for the bound Rev peptide in the complex. The Asn40–Arg48 residues represent the best defined backbone segment of the bound Rev peptide (Figure 3a) in the complex. This is a consequence of the majority of the intermolecular NOEs originating from this segment of the bound peptide as listed in Table 1. The amino acid sidechains are poorly defined, except for Trp45 (Figure 5c,d), amongst the refined structures of the complex. The line broadening observed to different degrees for protons of several peptide sidechains may reflect contributions associated with lack of intramolecular hydrogen bonding within the bound peptide, coupled with local conformational flexibility of particular sidechains of the bound peptide in the complex.

Peptide–RNA interface in the complex

The NOE fingerprint patterns characteristic of specific complex formation are consistent with the truncated 15-mer Rev (lacking Gln49 and Arg50, and R39A and R44N mutations) and 11-mer Rev (lacking Thr34–Arg39, and R42A and R43A mutations) peptides binding to the same target site on the class II RNA aptamers, as does the 17-mer Rev peptide. This restricts the minimal core-binding domain of Rev to the sequence spanning the Asn40–Arg48 segment of the peptide, whereas weak intermolecular NOEs approximately position the amino terminus of the peptide within the major groove of the longer stem of the RNA aptamer. Arg42 and Arg43 within the core binding region, and Arg39 adjacent to it, do not appear to be critical for site-specific recognition because replacement by alanine residues does not adversely affect complex formation. The same conclusion holds following replacement of the longer sidechain of Arg44 by the shorter sidechain of asparagine.

The peptide is anchored within the binding pocket of the RNA aptamer through two sets of intermolecular interactions associated with the core-binding segment of the complex. The Trp45–Arg48 segment of the Rev peptide interacts with the hairpin-loop segment of the RNA aptamer (Figure 6a), whereas the sidechain of Arg41 is positioned along the major groove edge of the Watson–Crick G10•C20 base pair and adjacent U8•(A11•U19) base triple in the complex (Figure 6c). These intermolecular interactions in the complex are discussed in further detail below.

Rev Trp45–U16 RNA stacking interaction

The mutual alignment of the aromatic rings of Trp45 and U16 in the distance-refined structures of the complex (Figure 6a) are defined by a large number of intermolecular NOEs between these two residues in the complex (Table 1). The aromatic rings of Trp45 and U16 stack on each other (Figure 6b) resulting in the stabilization of the bound conformation of the RNA hairpin loop and anchoring of the carboxyl terminus of the bound Rev peptide in the RNA binding pocket. This stacking explains the upfield ring current shifted position of the H6 proton of U16 (6.80 ppm) in the complex. A related tryptophan–base intermolecular van der Waals stacking interaction has been reported in the λ N-peptide–boxB RNA complex [24–26]. In this latter case, a tryptophan aromatic ring stacks over an adenine ring and in the process extends the characteristic stacking associated with the 3′ residues of a GNRA (where N is any nucleotide and R is a purine) loop.

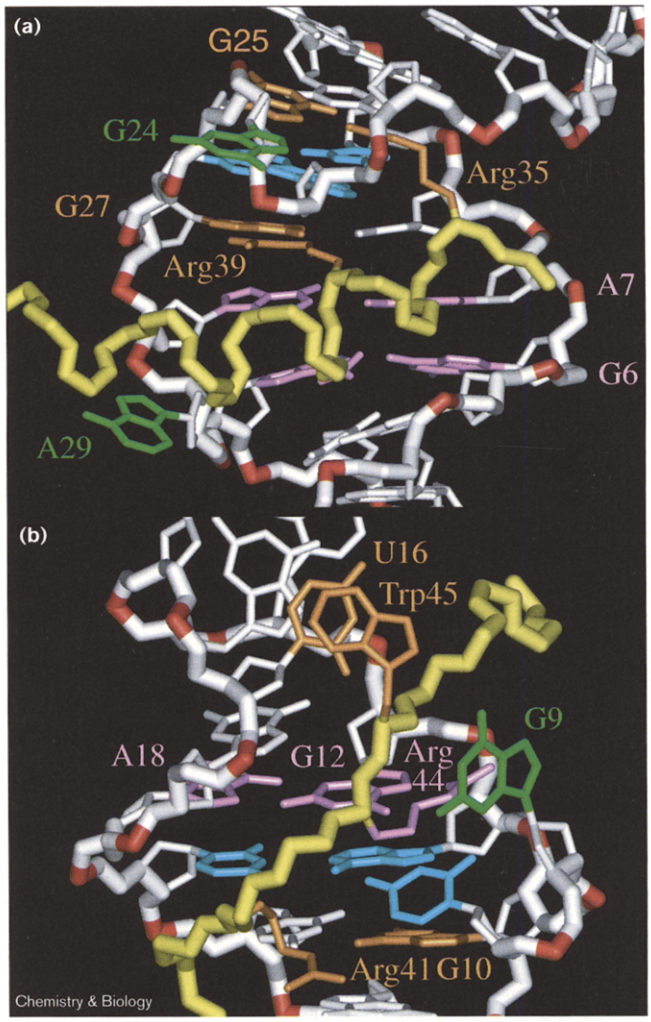
U16 Watson–Crick edge–peptide backbone hydrogen-bonding interaction

The base and sugar protons of U16 exhibit a large number of intermolecular NOEs to the backbone and/or sidechain protons of Trp45, Arg46, Glu47 and Arg48 in the complex (Table 1). Specifically, the imino proton of U16 exhibits NOEs to the backbone amide protons of Arg46, Glu47 and Arg48 (Table 1). The relative alignment of the Watson–Crick edge of U16 to the peptide backbone between Trp45 and Arg48 is well defined amongst the distance refined structures of the complex (Figure 6c). This alignment is stabilized by potential intermolecular hydrogen bonds between the amide proton of Arg46 and the O^2 carbonyl of U16 and between the amide carbonyl of Arg46 and the N^3 of U16 in the complex (Figure 6d). U16 is therefore anchored through stacking with Trp45 and by hydrogen bonding with the peptide backbone in the complex.

Arginine–guanine and arginine–backbone interactions

Our inability to assign the arginine sidechain ηNH2 protons, as well as to identify intermolecular NOEs between assigned arginine sidechain εNH protons and RNA aptamer protons in the NMR data sets on the complex, has resulted in poorly defined orientations of the arginine sidechain guanidinium groups relative to the RNA amongst the refined structures of the complex. This leaves us with qualitative conclusions about the orientation of the arginine sidechains based on available intermolecular NOEs between their nonexchangeable methylene protons and RNA aptamer protons in the complex (Table 1). The majority of these intermolecular NOEs involve Arg41, Arg46 and Arg48 (Table 1), and these residues will be the primary focus of our discussion. The sidechain of Arg41 is directed towards the major groove edge of G10 of the Watson–Crick G10•C20 pair (Figures 6e,f), whereas the sidechain of Arg44 is directed towards the major groove edge of G12 of the sheared G12•A18 mismatch pair amongst the majority of the refined structures of the complex. There are many examples in the literature of the guanidinium group of arginine aligning through hydrogen-bond formation with the major

Figure 7



Closeup comparative views of the key intermolecular interactions in representative refined structures of the complexes of the HIV-1 17-mer Rev peptide with **(a)** the 35-mer RRE RNA aptamer I (Figure 1c) reported previously [10] and **(b)** the 27-mer RNA aptamer II (Figure 1b) reported in this study. The peptide backbone is yellow and the RNA aptamer backbones are white except for phosphorus atoms, which are red. The base triples are cyan, base mismatches are pink and looped out bases are green. Key intermolecular contacts are shown in orange or pink and labeled by residue.

groove edge of guanine in peptide–RNA complexes (reviewed in [27]), including one such interaction with the major groove edge of guanine in a sheared G•A mismatch pair in the P22 N peptide–boxB RNA complex [17]. Indeed, the NOE patterns involving the nonexchangeable protons of U8 and G10 in the 17-mer Rev peptide–27-mer RNA aptamer II complex are similar to those previously reported for arginine-fork alignments involving arginine–guanine–U•(A•U) triple elements in amino acid–RNA [19–21] and peptide–RNA [10,17,18] complexes. By contrast, the guanidinium group of Arg46 is directed towards adjacent backbone phosphates flanking the G12 residue in the majority of the refined structures of the complex (Figures 6c,d).

Comparison of Rev peptide conformations bound to class I and II RNA aptamers

The solution structure of the 17-mer Rev peptide bound to the 35-mer RRE RNA aptamer I (class I, Figure 1c) reported previously [10] is shown in Figure 2a. An expanded view of the core-binding site is shown in Figure 7a. The Rev peptide binds in an α-helical conformation [9,10,12] within the major groove centered about adjacent G•A and A•A mismatches. Key intermolecular contacts involve hydrogen bonding between the guanidinium groups of Arg35 and Arg39 and the major groove edges of nonadjacent guanines involved in G•C pair formation and between the amide sidechain of Asn40 and both bases in the G•A mismatch pair (see also [28]). The complex is stabilized through formation of a U•(A•U) base triple, which serves as a platform for aligning the nonpolar sidechains of Arg35 and Arg39 involved in arginine–guanine pairing. The Rev peptide is bound deep in the major groove, which is widened by the presence of two looped out bases, one of which originates from the asymmetric internal bubble and the other from the two base bulge in the secondary structure of 35-mer RRE RNA aptamer I (Figure 1c). The intermolecular contacts in this complex are summarized schematically in Figure 4a.

The corresponding solution structure of the 17-mer Rev peptide bound to the 27-mer RNA aptamer II (Figure 1b) determined in this study is shown in Figure 2a and an expanded view of the core-binding site shown in Figure 7b. The Rev peptide binds in an extended conformation in the major groove, with the core-binding element centered about the hairpin-loop and adjacent short stem and two base-bulge segment. The binding-pocket architecture is defined, in part, through formation of a U•(A•U) triple, the looping out of one base from the two base-bulge which forms a flap over a portion of the bound peptide and a distinct hairpin-loop conformation closed by a sheared G•A mismatch pair. Key intermolecular contacts (summarized in Figure 4b) involve stacking of the Trp45 sidechain against a hairpin loop pyrimidine residue and arginine sidechain–guanine major groove edge interactions involving Arg41 and Arg44. The extended alignment of the bound peptide chain permits adjacent Arg44 and Trp45 residues to form intermolecular contacts with nonadjacent G12 (through hydrogen bonding) and U16 (stacking) RNA aptamer residues in the complex (Figure 7b). Similarly, an extended bound peptide conformation permits the NH and CO groups of Arg46 and the NH group of Arg48 to form intermolecular contacts with the Watson–Crick edge of U16 of the RNA aptamer in the complex (Figure 6b).

These studies establish that the Rev-peptide-binding pockets adopt very different architectures in the complexes

formed with class I RRE RNA aptamer I (Figure 1c) and class II RNA aptamer II (Figure 1b). The arginine-rich Rev peptide binds in an α-helical conformation to class I RRE RNA aptamer I (Figures 2a,7a) whereas it binds in an extended conformation to class II RNA aptamer II (Figures 2b,7b). Further, the bound peptides adopt opposite directionalities in these two complexes (Figure 2). The key specificity-determining amino acids are Arg35, Arg39 and Asn40 in the complex with class I RRE RNA aptamer I (Figure 7a) and Arg41, Arg44 and Trp45 in the complex with class II RNA aptamer II (Figure 7b). There are no intermolecular peptide-backbone–RNA interactions in the complex between the bound α-helical Rev peptide and the class I RRE RNA aptamer I target [10], in contrast to the observation of such interactions in the complex between the bound extended Rev peptide and the class II RNA aptamer II target (this study). These structural studies establish that short arginine-rich peptides undergo adaptive folding transitions, with the diversity of their bound conformations and the nature of their intermolecular contacts correlated with the unique architectures of major-groove-binding pockets of distinct RNA scaffolds.

Comparison of peptide–RNA and protein–RNA complexes

These conclusions regarding peptide–RNA complexes can be contrasted with related structural studies on protein–RNA complexes. An interesting example is the bacteriophage MS2 coat protein–RNA system, for which crystal structures are available for the coat protein dimer with the wild-type RNA [29] and two RNA aptamer complexes [30,31]. In this system, it is the protein dimer that retains its three-dimensional fold in all three complexes, whereas while the RNA components undergo adaptive structural transitions on complex formation. A critical feature of the MS2 coat protein–RNA system is that specific recognition nucleotides on the RNA are inserted into conserved hydrophobic binding pockets on the protein surface and anchored into position by intermolecular hydrogen bonding interactions. Similar observations have been made for the complex of *Escherichia coli* glutaminyl tRNA synthetase complexed to tRNA(Gln) and ATP [32], and recently for the complex of the HIV-2 nucleocapsid protein bound to SL3 ψ-RNA [33].

In essence, the two systems under comparison are complementary in that in peptide–RNA complexes it is the RNA tertiary structure and its binding pockets for minimal elements of peptide secondary structure that govern complex formation, whereas in protein–RNA complexes it is the protein tertiary structure and its binding pockets for minimal elements of RNA secondary structure that govern complex formation.

Implications of adaptive binding

Our results confirm that RNA aptamers can strongly induce conformational changes in peptide epitopes, just as antipeptide antibodies can undergo adaptive binding with their epitopes. These results have implications for both the biology of protein–RNA interactions and for drug discovery. To the extent that different RNA molecules can bind different conformations of the same amino-acid sequence, it might be possible for a single regulatory protein to interact with quite different RNA targets. Similarly, adaptive binding might allow the same protein–RNA interface to exist in multiple forms. For example, it might be possible for ribosomal protein arginine-rich motifs (ARMs) to remain associated with the translation apparatus during translocation. Finally, to the extent that strong interactions between the sequence of a RNA target and the chemical moieties present on a peptide or drug can drive binding irrespective of backbone conformations, it might prove possible to design drugs that can bind to multiple, related RNA targets. For example, a RNA-binding compound with sufficient affinity for a target sequence might still bind even if distal mutations cause the target to be presented in a different structural context. If this hypothesis continues to hold true, it might even be possible to design RNA-binding compounds that to some extent anticipate the evolution of resistant variants of viral binding elements. Indeed, based on the current example, it might be expected that even distantly related variants of the Rev-binding element could nonetheless be decoyed by an appropriate pharmaceutical peptide.

Significance

The solution structures of a limited number of arginine-rich basic peptide–RNA complexes have been reported to date. The majority of the peptides adopt an α-helical conformation (in complexes of the HIV-1 regulator of viral expression, Rev, with its RNA response element, RRE [9,12,34] and with a 'class I' RRE aptamer [10], as well as in the complexes of N peptide and boxB RNA involved in bacteriophage antitermination in the P22 [17] and λ [24–26] systems). In the complex of the BIV *trans*-activator protein (Tat) with its RNA response element, TAR, Tat adopts a β-sheet conformation [18,35]. The peptides are primarily unstructured when free in solution and undergo adaptive structural transitions to minimal elements of protein secondary structure on complex formation [27,36].

Here we show that in a 17-mer peptide fragment of Rev, when bound to a class II RNA aptamer, adopts an extended peptide conformation bound within the major groove of its RNA target site. This is in sharp contrast to the α-helical conformation the same peptide adopts when bound to the class I 35-mer RRE RNA aptamer I. These results unequivocally demonstrate that an unfolded peptide can utilize different folded conformations to recognize RNA partners with distinct peptide-binding pockets (this concept is discussed in [37]). This is achieved by using different amino-acid sidechains in the recognition

process and establishes the diversity of intermolecular alignments associated with peptide–RNA recognition.

Our results experimentally establish that unstructured peptides can target multiple RNA partners and, through adaptive binding, adopt distinct bound secondary folds at the peptide–RNA interface. Adaptive reorganization involving one or both partners on peptide–RNA complex formation could generate new surface architectures for the sequential addition of further components characteristic of multimeric systems.

Materials and methods

Sample preparation

The HIV-1 17-mer Rev peptide (Figure 1a) with a carboxamide protecting group at the carboxyl terminus and its 15-mer and 11-mer analogs were chemically synthesized and purified using high-performance liquid chromatography (HPLC) in the Sloan-Kettering Microchemistry Core Facility. The ^{15}N-labeled or $^{13}C,^{15}N$-labeled 17-mer Rev peptides were obtained by cloning the synthetic gene into a T7 expression plasmid, followed by over expressing the gene in *E. coli* grown in a medium supplemented either with $^{15}NH_4Cl$ or with $^{15}NH_4Cl$ and ^{13}C-glucose as sole nitrogen and carbon sources, respectively. To facilitate purification, the peptide was expressed as a histidine (His) tag fusion protein. The expressed protein formed inclusion bodies, which allowed the protein to be easily isolated from soluble proteins, and the His tag allowed further purification on a nickel chelating column. The desired peptide was then cleaved from the fusion protein with cyanogen bromide and purified using a C18 (Vydac) column by reverse-phase HPLC. The Rev-peptide-binding RNA aptamers were synthesized by *in vitro* transcription from a DNA template with commercially available NTPs (nucleoside triphosphates) or ($^{13}C,^{15}N$)-labeled NTPs using T7 RNA polymerase and subsequently purified using gel electrophoresis. The peptide–RNA complexes were generated by gradual addition of Rev peptide to the RNA aptamers with the stoichiometry monitored by following the intensity of the free and bound imino protons of the RNA aptamer (in slow exchange) in H_2O solution. The concentration of the Rev peptide–RNA aptamer complexes ranged between 0.5 and 2 mM.

NMR data collection and processing

NMR spectra of the Rev peptide–RNA aptamer complexes in 10 mM sodium phosphate buffer at pH 6.2 were acquired on Varian Inova 500 MHz and 600 MHz NMR spectrometers. Two-dimensional data sets were processed with VNMR (Varian) and analyzed with FELIX (Biosym/MSI). Three-dimensional data sets were processed with NMRpipe [38] and analyzed with PIPP [39]. Resonance assignments and NOE intensities for the Rev peptide and RNA aptamer components of the complexes were obtained from the following NMR experiments: two-dimensional NOESY (H_2O and D_2O), two-dimensional DQ-COSY and two-dimensional TOCSY (H_2O and D_2O) data sets were recorded on unlabeled samples of the complexes containing various combinations of Rev peptides and RNA aptamers. Two-dimensional ^1H-^{15}N HSQC (H_2O), two-dimensional ^1H-^{13}C HSQC (D_2O), two-dimensional ^{15}N-edited HMQC-NOESY (H_2O), three-dimensional HCCH-COSY (D_2O), three-dimensional HCCH-TOCSY (D_2O) and three-dimensional ^{13}C-edited NOESY–HMQC (D_2O) data sets were recorded on a sample of the complex containing uniformly $^{13}C,^{15}N$-labeled 27-mer RNA aptamer-II and unlabeled 17-mer Rev peptide. Two-dimensional ^1H-^{15}N HSQC (H_2O), three-dimensional ^{15}N-edited HMQC-NOESY (H_2O) and three-dimensional ^{15}N-edited HMQC–TOCSY (H_2O) were acquired on a sample of the complex with uniformly ^{15}N-labeled 17-mer Rev peptide and unlabeled 27-mer RNA aptamer II. Two-dimensional ^1H–^{15}N HSQC (H_2O) and two-dimensional ^1H–^{13}C HMQC-NOESY (H_2O) were collected on a sample of the complex with selectively $^{15}N,^{13}C$ labeled 17-mer Rev peptide (labeled at Arg41 and Arg46) and unlabeled 27-mer RNA aptamer II. Two-dimensional ^1H–^{15}N HSQC (H_2O), two-dimensional ^1H–^{13}C HSQC (H_2O), three-dimensional CBCA(CO)NH (H_2O), three-dimensional HNCACB (H_2O) and three-dimensional C(CO)NH (H_2O) were recorded on a sample of the complex with uniformly $^{15}N,^{13}C$ labeled 17-mer Rev peptide and unlabeled 27-mer RNA aptamer-II. Proton chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), whereas carbon and nitrogen chemical shifts were referenced to DSS and ammonia, respectively.

Distance restraints

Interproton distance restraints were obtained from NOE cross-peak volumes in two-dimensional NOESY and three-dimensional NOESY–HMQC data sets with the corresponding distance bounds calibrated using pyrimidine H5–H6 as a reference distance. The NOESY data sets were collected at various mixing time ranging from 80 ms to 300 ms. The distance bounds for restraints obtained from three-dimensional NOESY-HMQC (120 ms mixing time) spectra in D_2O or H_2O buffer and two-dimensional NOESY (100 and 150 ms mixing time) in H_2O or D_2O buffer were categorized as strong (2.5 Å ± 0.7), medium (3.5 Å ± 1.0) or weak (4.5 Å ± 1.5). The corresponding distance bounds were increased for restraints obtained from NOESY (mixing time > 150 ms) spectra in D_2O buffer and those involving more rapidly exchanging protons from NOESY spectra in H_2O buffer, which were categorized as strong (2.7 Å ± 1.0), medium (3.5 Å ± 1.5), and weak (4.5 Å ± 2.0). Larger distance bounds were also used for NOEs involving methyl-group protons.

Torsion-angle restraints of the endocyclic torsion angles within ribose sugars of RNA were obtained from an analysis of COSY experiments. The furanose rings with strong H1′–H2′ cross peaks were restricted to C2′-*endo* (P = 144–180°) sugar puckers, whereas those rings that showed no H1′–H2′ cross peaks were restricted to C3′-*endo* (P = 0–36°) sugar puckers. These sugar pucker pseudorotation angles were next translated to corresponding values and ranges for the sugar endocyclic torsion angles.

Molecular dynamics protocol

A molecular dynamics-simulated annealing protocol driven by NOE distance and dihedral restraints (X-PLOR package, version 3.8, [40]) was used to solve the solution structure of the HIV-1 Rev peptide–RNA aptamer complex. The refinement procedure include 60 simulated annealing trials starting from 60 random-chain conformations for both peptide and RNA chains with the components placed 100 Å apart in space. The dynamics protocol was undertaken in two stages. The first stage involved high-temperature molecular dynamics in torsional space. The molecules were equilibrated at 20,000 K (15,000 steps over 1.5 ps) and then cooled slowly to 1,000 K (30,000 steps over 15 ps). Nonbonded interactions were restricted to repulsive force field potentials and these first stage computations were guided by NOE distance and dihedral restraints. The second stage involved lower temperature dynamics in cartesian space using Lennard–Jones potentials for van der Waals interactions and guided by NOE distance and dihedral restraints. The structures were slowly cooled to 300 K (30,000 steps over 15 ps) and minimized until the gradient of energy was less than 0.1 kcal mol^{-1}. The helical stem extending from G1–C6 and from G22–C27 located outside the core-binding region was restrained to an A-form helix through incorporation of torsion and hydrogen bonding restraints during both high- and low-temperature dynamics. In addition, G12 and A18 were restricted to the experimentally identified sheared G12•A18 mismatch alignment during the computations. The force constants for NOE distance restraints were maintained at 30 kcal mol^{-1} and for dihedral restraints at 25 kcal rad^{-2}. A subset of 22 refined structures (≈15 kcal mol^{-1} variation of total energy within the subset) were selected based on lowest NOE violation energy and were separated by a gap of ≈250 kcal mol^{-1} of total energy from the rest of the ensemble.

Structure analysis

The RNA helical parameters were analyzed using the CURVES program [41]. Color figures were prepared using INSIGHT II (Molecular Simulations, Inc) program.

Accession numbers

The coordinates of the HIV-1 17-mer Rev peptide–27-mer RNA aptamer II complex (accession number: 484d) have been deposited with the Protein Data Bank.

Supplementary material

Supplementary material including NMR resonance assignments is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

This research was funded by NIH grant CA-49982 to D.J.P.

References

1. Cullen, B.R. & Malim, M.H. (1991). The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators. *Trends Biochem. Sci.* **16**, 346-350.
2. Gait, M.J. & Karn, J. (1993). RNA recognition by the human immunodeficiency virus Tat and Rev proteins. *Trends Biochem. Sci.* **18**, 255-259.
3. Zapp, M.L., Hope, T.J., Parslow, T.G. & Green, M.R. (1991). Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. *Proc. Natl Acad. Sci. USA* **88**, 7734-7738.
4. Heapy, S., Finch, J.T., Gait, M.J., Karn, J. & Singh, M. (1991). Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich 'bubble' located within the rev-response region of viral mRNAs. *Proc. Natl Acad. Sci. USA* **88**, 7366-7370.
5. Tiley, L.S., Malim, M.H., Tewary, H.K., Stockley, P.G. & Cullen, B.R. (1992). Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type I Rev protein. *Proc. Natl Acad. Sci. USA* **89**, 758-762.
6. Frankel, A.D. (1994). Using peptides to study RNA–protein recognition. In *RNA–Protein Interactions.* (Nagai, K. & Mattaj, I.W., eds), pp. 221-247, IRL Press, New York.
7. Karn, J., Gait, M.J., Churcher, M.J., Mann, D.A., Mikaelian, I. & Pritchard, C. (1994). Control of HIV gene expression by RNA-binding proteins Tat and Rev. In *RNA-Protein Interactions.* (Nagai, K. & Mattaj, I.W., eds), pp. 193-220, IRL Press, New York.
8. Kjems, J, Calnan, BJ, Frankel, AD & Sharp, PA (1992). Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* **11**, 1119-1129.
9. Battiste, J.L., *et al.*, & Williamson, J.R. (1996). α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273**, 1547-1551.
10. Ye, X., Gorin, A., Ellington, A. & Patel, D.J. (1996). Deep penetration of an α helix into a widened RNA major groove in the HIV-1 Rev peptide-RNA aptamer complex. *Nat. Struct. Biol.* **3**, 1026-1033.
11. Giver, L., Bartel, D., Zapp, M., Pawul, A., Green, M. & Ellington, A.D. (1993). Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Res.* **23**, 5509-5516.
12. Tan, R., Chen, L., Buettner, J.A., Hudson, D. & Frankel, A.D. (1993). RNA recognition by an isolated α helix. *Cell* **73**, 1031-1040.
13. Xu, W. & Ellington, A.D. (1996). Anti-peptide aptamers recognize amino acid sequence and bind a peptide epitope. *Proc. Natl Acad. Sci. USA* **93**, 7475-7480.
14. Lazinski, D., Grzadzielska, E. & Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine rich motif. *Cell* **59**, 207-218.
15. Pardi, A. (1995). Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol.* **261**, 350-380.
16. Jiang, F., Fiala, R., Live, D., Kumar, R.A. & Patel, D.J. (1996). RNA folding topology and intermolecular contacts in the AMP–RNA aptamer complex. *Biochemistry* **35**, 13250-13266.
17. Cai, Z., *et al.*, & Patel, D.J. (1998). Solution structure of P22 transcriptional antitermination N-peptide–boxB RNA complex. *Nat. Struct. Biol.* **5**, 203-212.
18. Ye, X., Kumar, R.A. & Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat-peptide–TAR RNA complex. *Chem. Biol.* **2**, 827-840.
19. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. & Williamson, J.R. (1992). Conformation of the Tar-RNA–arginine complex by NMR spectroscopy. *Science* **257**, 76-80.
20. Puglisi, J.D., Chen, L., Frankel, A.D. & Williamson, J.R. (1993). Role of RNA structure in arginine recognition of Tar RNA. *Proc. Natl Acad. Sci. USA* **90**, 3680-3684.
21. Brodsky, A.S. & Williamson, J.R. (1997). Solution structure of the HIV-2 TAR–argininamide complex. *J. Mol. Biol.* **267**, 624-639.
22. Li, Y., Zon, G. & Wilson, W.D. (1991). NMR and molecular evidence for a G•A mismatch base pair in a purine-rich DNA duplex. *Proc. Natl Acad. Sci. USA* **88**, 26-30.
23. Jucker, F.M., Heus, H.A., Yip, P.F., Moors, E.H. & Pardi, A. (1996). A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* **264**, 968-980.
24. Su, L., *et al.*, & Weiss, M.A. (1997). An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. *Genes. Dev.* **11**, 2214-2226.
25. Su, L., Radek, J.T., Hallenga, K, Hermanto, P, Labeots, LA & Weiss, M.A. (1997). RNA recognition by a bent α-helix regulates transcriptional antitermination in phage λ. *Biochemistry* **36**, 12722-12732.
26. Legault, P., Li, J., Mogridge, J., Kay, L.E. & Greenblatt, J. (1998). NMR structure of the bacteriophage λ N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289-299.
27. Patel, D.J. (1999). Adaptive recognition in RNA complexes with peptide and protein modules. *Curr. Opin. Struct. Biol.* **9**, 74-87.
28. Jain, C. & Belasco, J.G. (1996). A structural model for the HIV-1 Rev-RRE complex deduced from altered-specificity Rev variants isolated from a rapid genetic strategy. *Cell* **87**, 115-125.
29. Valegard, K., *et al.*, & Liljas, L. (1997). The three dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein–RNA interactions. *J. Mol. Biol.* **270**, 724-738.
30. Convery, M.A., *et al.*, & Stockley, P.G. (1998). Crystal structure of an RNA aptamer-protein complex at 2.8 Å resolution. *Nat. Struct. Biol.* **5**, 133-139.
31. Rowsell, S. & Stockley, P.G. (1998). Crystal structures of a series of RNA aptamers complexed to the same protein target. *Nat. Struct. Biol.* **5**, 970-975.
32. Rould, M.A., Perona, J.J., Soll, D. & Steitz, T.A. (1989). Structure of *E. coli* glutaminyl tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. *Science* **246**, 1135-1142.
33. deGuzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N. & Summers, M.F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 ψ-RNA recognition element. *Science* **279**, 384-388.
34. Ellington, A.D., Leclerc, F. & Cedergen, R. (1996). An RNA groove. *Nat. Struct. Biol.* **3**, 981-984.
35. Puglisi, J.D., Chen, L., Blanchard, S. & Frankel A.D. (1995). Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **270**, 1200-1203.
36. Grate, D. & Wilson, C. (1997). Role REVersal: understanding how RRE RNA binds its peptide ligand. *Structure* **5**, 7-11.
37. Frankel, A.D. & Smith C.A. (1998). Induced folding in peptide–RNA recognition: more than a simple handshake. *Cell* **92**, 149-151.
38. Delaglio, F., Grzesiek, S., Vuister, G., Zu, G., Pfieiffer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293.
39. Garrett, D.S., Powers, R., Gronenborn, A.M. & Clore, G.M. (1991). A common sense approach to peak picking in two, three and four dimensional spectra using automatic computer analysis of contour programs. *J. Magn. Reson.* **95**, 214-220.
40. Brunger, A.T. (1992). XPLOR User Manual (Version 3.1), Yale University, New Haven, CT, USA.
41. Lavery, R & Sklenar, H (1988). The definition of generalized helicoidal parameters and axis of curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**, 63-91.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.