Localized Influence of 2′-Hydroxyl Groups and Helix Geometry on Protein Recognition in the RNA Major Groove†

Stephen G. Landt, Alicia R. Tipton, and Alan D. Frankel*

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143-2280

Received September 26, 2004; Revised Manuscript Received February 10, 2005

ABSTRACT: The local geometry of a DNA helix can influence protein recognition, but the sequence-specific features that contribute to helix structure are not fully understood, and even less is known about how RNA helix geometry may affect protein recognition. To begin to understand how local or global helix structure may influence binding in an RNA model system, we generated a series of DNA analogues of HIV and BIV TAR RNAs in which ribose sugars were systematically substituted in and around the known binding sites for argininamide and a BIV Tat arginine-rich peptide, respectively, and measured their corresponding binding affinities. For each TAR interaction, binding occurs in the RNA major groove with high specificity, whereas binding to the all-DNA analogue is weak and nonspecific. Relatively few substitutions are needed to convert either DNA analogue of TAR into a high-affinity binder, with the ribose requirements being restricted largely to regions that directly contact the ligand. Substitutions at individual positions show up to 70-fold differences in binding affinity, even at adjacent base pairs, while two base pairs at the core of the BIV Tat peptide−RNA interface are largely unaffected by deoxyribose substitution. These results suggest that the helix geometries and unique conformational features required for binding are established locally and are relatively insulated from effects more than one base pair away. It seems plausible that arginine-rich peptides are able to adapt to a mosaic helical architecture in which segments as small as single base steps may be considered as modular recognition units.

Sequence-specific recognition of nucleic acids is commonly achieved by the complementary positioning of hydrogen bonding groups of the protein side chains and backbone with those of the bases, base pairs, or base pair steps in the major groove of a double helix (*1*, *2*). Thus, the form of the double helix (A, B, or other) can influence recognition by altering the relative positioning of base and backbone functional groups, as can the presence of local helical irregularities, bends, or nonhelical structures. Indeed, analyses of DNA−protein cocrystal structures have shown that many protein binding sites diverge significantly from the canonical B-form geometry, sometimes toward the A-form, and often in ways that increase complementarity between the major groove and a protein α-helix (*3*, *4*). The value of recognizing a sequence in an "A-like" geometry was further suggested by the observation that a zinc finger protein binds more tightly to its target sequence when presented in the context of an RNA/DNA hybrid than as DNA (*5*). While some proteins may recognize sequences within limited stretches of an A-form helix, an extended A-form duplex presents a deep and narrow major groove that is too small to accommodate a protein α-helix or *β*-sheet segment (*6*, *7*). Thus, recognition of base pairs within an RNA major groove typically is accompanied by local interruptions of helical structure, incorporating adjacent bulged or looped nucleotides that widen the groove sufficiently for protein access (*8*).

One type of RNA-binding domain that recognizes the RNA major groove is the arginine-rich motif (ARM)[1], a 10−20 amino acid segment enriched in arginine that typically binds RNA with high affinity and specificity as an isolated peptide (*9*). ARM peptides can adopt α-helical, *β*-hairpin, or extended conformations and often require RNA to assume their bound structures. It is likely that these peptides must conform to the structure of the deep major groove in order to create a sufficiently large interface for specific, high-affinity binding (*9*−*11*). Among the most well-characterized ARM peptide−RNA interactions are the Tat−TAR complexes from the human and bovine immunodeficiency viruses (HIV and BIV). The viral Tat proteins are essential transcription elongation factors that bind to their respective TAR RNA hairpins located at the 5′-end of the nascent transcripts (*12*).

In the HIV Tat−TAR complex, an ARM peptide comprised of amino acids 49−57 binds TAR with high affinity, primarily using one arginine (at position 52) to make specific contacts to the RNA (*8*, *13*−*15*). The free amino acid arginine, or even the guanidinium group alone, also binds TAR specifically, and NMR structures have shown that arginine and Tat peptides induce similar conformational

† This work was supported by a Howard Hughes Medical Institutes predoctoral fellowship (S.G.L.) and by NIH Grants GM47478 and AI29135 (A.D.F.).

* To whom correspondence should be addressed: Alan Frankel, Department of Biochemistry and Biophysics, UCSF, 600 16th Street, San Francisco, CA 94143-2280. Phone, 415-476-9994; fax, 415-514-4112; e-mail, frankel@cgl.ucsf.edu.

[1] Abbreviations: ARM, arginine-rich motif; HIV, human immunodeficiency virus; BIV, bovine immunodeficiency virus; CD, circular dichroism.

10.1021/bi047916t CCC: $30.25
Published on Web 04/06/2005

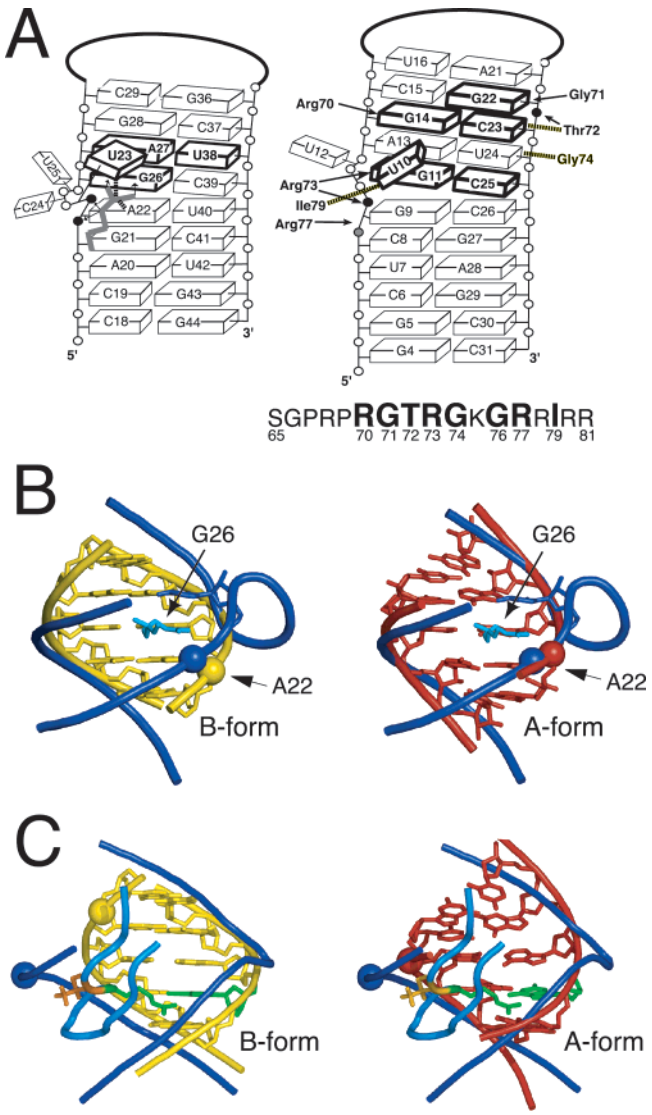

FIGURE 1: Models of HIV TAR−arginine and BIV TAR−Tat peptide complexes. (A) Schematic drawings of the HIV (left) and BIV (right) TAR complexes. Nucleotides in the RNAs and amino acids in the BIV Tat peptide important for binding (*26*, *27*) are highlighted. Phosphates whose ethylation strongly interferes with binding are indicated by black dots, and one with a moderate effect is in gray (*15*, *27*). Hydrogen bonding and electrostatic interactions are indicated by arrows, and van der Waals interactions are indicated by dashed lines. (B) Overlapped models of B-form (left, yellow) and A-form (right, red) HIV TAR helices on the HIV TAR−argininamide complex, with argininamide (light blue) and the extrahelical U23 residue (blue) shown in a stick representation and the RNA backbone as ribbon (blue) (*18*). Positions of the A22 phosphates are shown by spheres and marked by an arrow, and the location of the G26 base is indicated. (C) Overlapped models of B-form (left, yellow) and A-form (right, red) BIV TAR helices on the BIV TAR−Tat peptide complex, with the RNA (blue) and peptide (light blue) backbone shown as ribbons (*29*). The Arg73−G11 hydrogen bonding interaction is shown in green, and the Thr72−C23 phosphate interaction is shown in orange. Positions of the C23 phosphates are shown by spheres.

changes in the RNA. Upon binding, the helices surrounding a three-nucleotide bulge coaxially stack, and the bulged U23 residue is positioned in the major groove of the upper stem to form a base triple with the A27−U38 pair (Figure 1A) (*16*−*21*). Arginine, which appears to be presented from an extended peptide chain, stacks between A22 and U23, forms hydrogen bonds with the O6 and N7 atoms of G26 in the major groove, and makes electrostatic or hydrogen bond contacts with the phosphates of A22 and U23. This arginine-binding motif has been observed repeatedly in in vitro selection experiments with arginine or ARM peptides (*22*−*25*), suggesting that it represents a highly favorable major groove interaction.

Despite significant similarities between the HIV and BIV hairpin sequences, including conservation of the arginine-binding site, the BIV Tat−TAR interaction is far more extensive (Figure 1A) (*26*−*28*). The BIV interaction involves eight critical peptide residues, all contained within residues 70−79 of the full-length protein and a larger RNA surface (*26*, *27*). The peptide is unstructured on its own, but NMR structures of the complex demonstrate that it adopts a β-hairpin configuration upon binding in the major groove (*29*, *30*). A base triple similar to that in HIV TAR is observed, with arginine 73 making analogous contacts, although the A13−U24 base pair of the triple can be mutated to U−A with little effect on binding affinity (*27*). A critical isoleucine at position 79 appears to buttress the bulged U10 base by stacking against the hydrophobic C5−C6 surface, and additional contacts between arginine 70 and G14, glycine 71 and G22, threonine 72 and the G22/C23 step, and potentially between arginine 77 and G9 at the top of the lower stem all contribute to the high-affinity interaction.

The structures of DNA−protein complexes provide evidence that helical conformation plays an important role in protein recognition (*3*), but direct experimental evidence is generally lacking, and even less is known about how the form of the helix may contribute to RNA binding. To assess the roles of helix geometry in ARM peptide recognition, we systematically introduced ribose groups into DNA analogues of HIV and BIV TAR, which do not bind argininamide or the BIV Tat peptide, respectively, and identified positions critical for binding. Because 2′-hydroxyl groups are the major determinant of A-form helix geometry and project into the minor groove where they are inaccessible to ARM peptides bound in the major groove (*6*), we reasoned that such DNA/RNA hybrids might reveal which portions of the binding sites are particularly sensitive to helical conformation. We show that the ribose requirements for both interactions are highly localized, with some regions that are sensitive to 2′-hydroxyl substitution being located immediately adjacent to largely insensitive regions. Thus, some essential peptide−RNA contacts appear to require the A-form geometry, whereas others may adapt to variant helical forms, perhaps reflecting the conformational plasticity that is characteristic of ARM peptides.

MATERIALS AND METHODS

Modeling. Canonical A- and B-form helices with the sequences of HIV and BIV TAR were generated by the Nucleic Acid Builder software package (*31*). Using InsightII (Accerlys) or PyMOL (DeLano Scientific), these helices were superimposed on the NMR-derived model of the HIV TAR−argininamide complex (*18*), and the average NMR structure of the BIV Tat−TAR complex (1MNB) (*29*) by overlapping all nitrogen atoms of the four bases at the junction between the upper and lower TAR stems (the A22−U40 and G26−C39 base pairs for HIV TAR and the G9−C26 and G11−C25 base pairs for BIV TAR).

Preparation and Characterization of RNA, DNA and DNA/RNA Hybrids. RNA was prepared by in vitro transcription, purified, and labeled as described (*27*). DNA and DNA/RNA hybrids were synthesized on an Applied Biosystems model 394 oligonucleotide synthesizer using phosphoramidites from Glen Research (Sterling, VA) and cleaved and deprotected according to manufacturer's protocols. Briefly, hybrids were deprotected overnight in ethanolic ammonium hydroxide and evaporated to dryness. Following a 24 h incubation in triethylamine trihydrofluoride to remove the silyl protecting groups, oligonucleotides were precipitated once with 0.3 M sodium acetate, pH 6.0, and 3 vol of 1-butanol, and once with ethanol. Hybrids were purified on 15% polyacrylamide/urea gels and stored at −80° in deionized water. DNA and DNA/RNA hybrids were end-labeled with ^{32}P as described (*27*), and the positions of ribonucleotide substitutions were confirmed by hydrolysis in 33 mM sodium bicarbonate, pH 9.0, for 15 min at 90 °C followed by analysis on 20% polyacrylamide/urea gels.

L-Arginine-Affinity Chromatography. For HIV TAR analogue binding experiments, L-arginine agarose columns (1 mL of 8.4 μmol/mL packed resin in 10 mL disposable columns, Sigma Chemical Company) were prewashed with 5 mL of 1 M NaCl/1× binding buffer (10 mM Tris·HCl, pH 7.2, and 0.2 mM EDTA), followed by 5 mL of 100 mM NaCl/1× binding buffer. A mixture of ^{32}P-5′-end-labeled RNA or hybrid molecule (∼300 000 cpm) and 10 μg of TAR DNA was loaded onto the column at 4 °C in 100 mM NaCl/1× binding buffer, and oligonucleotides were eluted with a 100−500 mM NaCl gradient (100 mL) at 4 °C. Fractions (1.2 mL) were collected and analyzed on 15% polyacrylamide/urea gels to determine elution positions. Dissociation constants (K_ds) were estimated using a standard curve (see Figure 8 in ref *23*) relating the K_ds of a series of HIV TAR mutants measured by isocratic elution from an arginine column, under conditions similar to those described above, to the concentration of NaCl required for elution (*23*, *32*). For calibration, HIV TAR RNA elutes at 165 mM NaCl, whereas the DNA analogue elutes at 79 mM NaCl. Although column elutions were performed at 100 mM NaCl, it is unlikely that any weak interactions were missed because all hybrids eluted as single peaks later than the column void volume, indicating some degree of interaction with the matrix.

BIV Tat Peptide Affinity Chromatography. An 18-amino acid peptide corresponding to residues 65−81 of BIV Tat preceded by a cysteine (Cys-BTat, CSGPRPRGTRGKGRRIRR) was synthesized, purified as described (*26*), and quantified by reactivity of the thiol group to Ellman's reagent (*33*). The BTat affinity resin was generated by first activating 1 mL of packed λ-amino-hexyl agarose (4% agarose, epoxy-activated with a 12 atom spacer, Sigma) with freshly prepared 2.5 mM Sulfo-SMCC (Pierce) in 5 mL coupling buffer (50 mM sodium phosphate, pH 7.4) for 30 min at room temperature. Activated agarose was washed 3 times with 30 mL of coupling buffer and resuspended in 2 mL coupling buffer containing 50 μg Cys-BTat. The peptide−agarose mixture was incubated for 2 h at room temperature, and unreacted resin was blocked by incubating with 20 μL of 500 mM dithiothreitol for an additional 30 min. Resin was washed 3 times and resuspended in 1 mL coupling buffer.

For BIV TAR analogue binding experiments, a BTat column (333 μL packed resin in a 10 mL disposable column) was washed with 5 column vol of 900 mM NaCl/binding buffer (10 mM Tris-Cl, pH 7.4, 1 mM EDTA, and 0.005% Triton X-100) followed by 5 column vol of 100 mM NaCl/binding buffer. ^{32}P-5′-end-labeled RNA or hybrid (>100 000 cpm) was heated to 85 °C in 1 mL of 100 mM NaCl/binding buffer and slow cooled to 4 °C, and a total of 50 μg yeast tRNA (Invitrogen) was added. The mixture was loaded onto the column, and oligonucleotides were eluted with a 100−900 mM NaCl/binding buffer gradient (80 mL) at 4 °C. Elution volumes were determined by scintillation counting, and results from at least three independent experiments were averaged. K_ds were determined using a standard curve relating K_ds measured by fluorescence anisotropy (see below) to the concentration of NaCl required for elution. A tight linear relationship was obtained for all hybrids with measurable K_ds from anisotropy experiments.

Fluorescence Anisotropy. Cys-BTat was labeled with fluorescein at its N-terminus by incubating 50 μM peptide with 500 μM 5-(iodoacetamido)fluorescein in 20 mM sodium phosphate, pH, 8.0, and 2 mM EDTA for 2 h at room temperature in the dark. Labeled peptide was purified by C_4 reverse-phase HPLC as described for the unlabeled peptide (*26*) and was quantified by fluorescein absorbance at 475 nm. To measure binding, fluorescein-labeled Cys-BTat (2.5 nM) was incubated with varying concentrations of DNA, RNA, or DNA/RNA hybrids (0.5−8092 nM) in 30 mM Hepes, pH 7.5, 100 mM KCl, 40 mM NaCl, 10 mM ammonium acetate, 10 mM guanidinium, 2 mM $MgCl_2$, 0.5 mM EDTA, and 0.001% Nonidet P-40 for 30 min at room temperature using 20 μL reactions in 384-well plates. Fluorescence anisotropy was measured in a LJL Biosystems Criterion fluorometer using a fluorescein filter set (excitation at 485 nm and emission at 530 nm) and a *G*-factor of 0.8. Each point was measured three times, and values were averaged from three independent experiments. K_ds were determined by fitting data to a single-site binding equation using Kaleidagraph software.

Circular Dichroism (CD) Spectroscopy. CD spectra were measured using an Aviv model 62DS spectropolarimeter. Samples (50 μg/mL of oligonucleotide) were prepared in 10 mM sodium phosphate, pH 7.5, and 100 mM NaCl and maintained at 4 °C. Spectra were recorded from 320 to 210 nm using a 1 cm path length cuvette. The signal was averaged for 5 s at each wavelength, and scans were repeated 3 times and averaged. For HIV TAR and analogues, 10 mM argininamide was added, and for BIV TAR and analogues, a stoichiometric amount of peptide (5 μM) was added. To estimate the fraction of A- and B-form helical content in the HIV TAR hybrids, we determined the CD values for each molecule at 266 and 284 nm, which correspond to the peaks observed with the all-RNA and all-DNA versions of HIV TAR, respectively. We estimated the percentage of the two helical forms by comparing the ratio of values at these wavelengths to those of the RNA and DNA versions (considered to be 100% A- and B-form content, respectively), assuming a linear relationship between this ratio and the A- and B-form content, and neglecting any possible contribution from the nonhelical regions. We performed similar calculations for the BIV TAR hybrids using CD values at 267 and

279 nm, which correspond to the peaks observed with the all-RNA and all-DNA versions of BIV TAR, respectively.

Native Gel Analyses and Peptide-Binding Gel Shift Assays. BIV TAR oligonucleotides were heated to 85 °C and slowly cooled to 4 °C (in 40 mM Hepes, pH 7.5, 100 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 50 μg/mL of yeast tRNA, and 10% glycerol), and mobilities were characterized on native 20% polyacrylamide 1× TBE gels run at 600 V for 20−24 h at 4 °C. RNA-binding gel shifts with BTat 65−81 peptide were performed under similar conditions, with peptide at a saturating concentration (2 μM) incubated with oligonucleotides for 30 min at 4 °C. We did not perform extensive analyses of HIV TAR hybrids using native gels because the differences suggested by CD were relatively small and because argininamide-bound complexes are too short-lived to be detected on gels.

RESULTS

Arginine Binding to HIV TAR DNA/RNA Hybrids. Early studies with the HIV Tat protein and Tat-derived ARM peptides showed that a DNA version of HIV TAR did not bind the protein specifically even though no 2′-hydroxyl group was directly involved in binding (*13*, *14*, *34*), leading to the suggestion that ARM−RNA interactions may require an A-form helical geometry for recognition. Indeed, NMR studies of HIV TAR and BIV TAR described below show that all stems form continuous A-form helices (*18*, *20*, *21*, *29*, *30*). To better visualize how helix geometry might influence binding, we generated models of HIV TAR, based on an NMR complex with argininamide, in which the upper and lower stems were replaced by idealized A- or B-form helices (Figure 1B). It is apparent that the distances between the O6 and N7 atoms of G26 and the phosphate of A22 in HIV TAR, which make the most critical contacts to arginine, are lengthened in the B-form model and that they cannot simultaneously form hydrogen bonds to the guanidinium group. It is less obvious whether differences between the major groove widths or between helical twists at individual base steps might also affect interactions either with the guanidinium group or the aliphatic portion of the arginine side chain, and it is entirely unclear how localized any important differences might be.

One way to effectively change helical conformation is to alter the 2′-position of the sugar ring, using a hydroxyl group to favor the C3′-endo sugar pucker and resulting A-form geometry (*6*). Indeed, some studies have suggested that introducing a single 2′-hydroxyl into a DNA helix can drive its conformation toward the A-form, although solution structures generally suggest that these effects are limited to the vicinity of the substitution (*35*−*39*). Because the 2′-hydroxyl groups in an A-form helix project into the minor groove, they cannot directly contact arginine or ARM peptides bound in the major groove, and therefore, any effects on recognition are most likely attributable to changes in local or global helical conformation. To identify regions in HIV TAR where the helical form is most critical for recognition, we created a series of hybrids in which 2′-hydroxyl groups were systematically introduced into an all-DNA version of HIV TAR and measured their ability to rescue arginine binding, as monitored by salt-dependent elution from an L-arginine agarose column (*23*, *32*). Strikingly, ribose

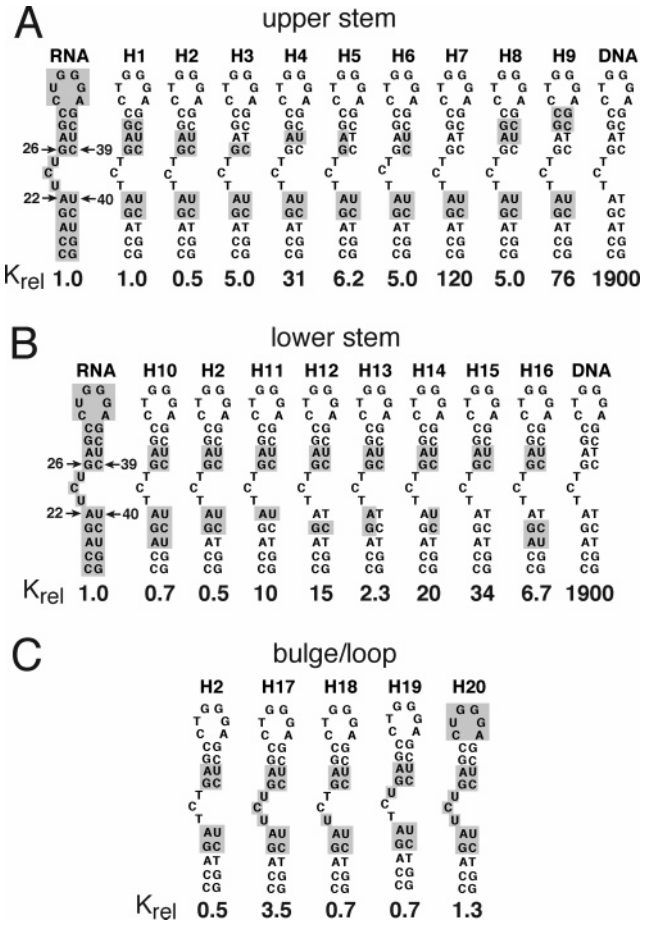

FIGURE 2: Binding affinities of arginine for HIV TAR hybrids. Positions of ribose substitution are highlighted. K_{rel} is the dissociation constant of each hybrid relative to that of HIV TAR RNA, as determined by affinity chromatography. Substitutions in (A) the upper stem, (B) the lower stem, and (C) the bulge and loop are shown. Numbered arrows by the RNA hairpin in A and B indicate the nucleotide numbering scheme.

substitution of the two base pairs immediately above and immediately below the bulge was sufficient to generate a hybrid (H2, Figure 2) with an affinity for arginine ~2-fold higher than the all-RNA version of HIV TAR.

The affinities of a variety of analogues in the upper stem (Figure 2A) indicate the particular importance of riboses in the base pair immediately above the bulge. Removing ribose groups from both upper stem base pairs in the H2 context (H7) reduces binding affinity ~250-fold relative to H2, while retaining ribose only at the A27−U38 pair (H4) still reduces affinity ~60-fold. A hybrid with ribose only at the G26−C39 base pair (H3) has an affinity just 10-fold lower than that of H2. Much of the affinity of H4 can be restored by adding riboses at an additional adjacent base pair (H8), perhaps suggesting a cooperative propagation of A-form helix structure toward the binding site. Substitutions made only at positions in the 5′- or 3′-strands (H5 and H6) showed similar affinity to substitutions at G26−C39, indicating that at least one 2′-hydroxyl group positioned at the junction of the two stems is important but that no single position is critical. This is consistent with the effects of single deoxyribose substitutions in HIV TAR RNA on Tat peptide binding (*40*).

Analogues in the lower stem (Figure 2B) indicate that riboses at the two base pairs below the bulge contribute

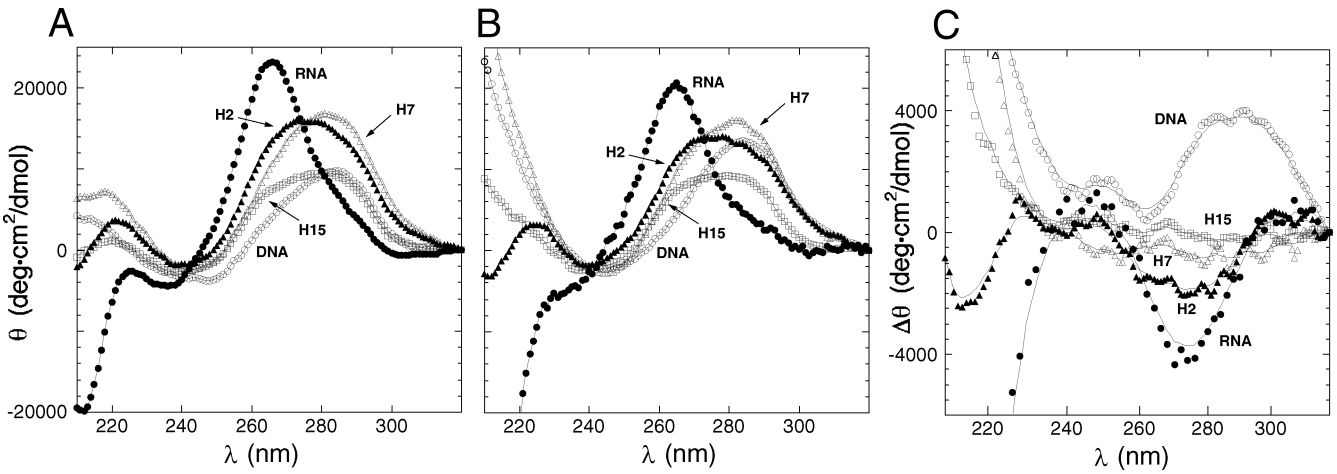

FIGURE 3: CD spectra of HIV TAR RNA (●), HIV TAR DNA (○), H2 (▲), H7 (△), and H15 (□) in the (A) absence or (B) presence of 10 mM argininamide. (C) Difference spectra calculated by subtracting the spectra in the absence of argininamide from the corresponding spectra in the presence of argininamide.

approximately equally to binding. Removing ribose groups from both base pairs in the H2 context (H15) reduces binding affinity ~70-fold relative to H2, whereas either single base pair substitution (at A22−U40 or G21−C41, H11 or H12) restores binding by only ~2−3-fold. There is a marked strand asymmetry at these base pairs, with riboses being much more important on the 5′-strand (compare H13 and H14). Previous work has shown that deoxyribose substitutions at either G21 or A22 have only a small effect on Tat peptide binding to the RNA form of TAR (*40*), but the effects in the DNA context shown here are more prominent, likely reflecting the importance of local A-form geometry in the stacked strand surrounding the binding site. The local nature of the effect is especially apparent in the lower stem as substituting riboses at an additional adjacent base pair shows less evidence of helix propagation (compare H12 and H16).

The extensive set of HIV TAR hybrids permits some analysis of how ribose substitutions at particular base pairs are affected by the ribose/deoxyribose composition of neighboring base pairs. To facilitate pairwise comparisons, we converted differences in binding affinities to free energy differences and identified cases where the $\Delta\Delta G$ values between hybrids with neighboring ribose or deoxyribose base pairs were significant. In the upper stem, addition of riboses at the G28−C37 base pair is ~1.4 kcal/mol (an affinity difference of ~12-fold) less favorable when the G26−C39 pair already has been ribose-substituted (compare H1 and H2 with H4 and H8), further supporting an indirect role for the G28−C37 conformation in recognition of the two adjacent lower pairs. In contrast, $\Delta\Delta G$ values of only ~0.6 kcal/mol are observed when the G26−C39 or A27−U38 pairs at the core of the binding site are ribose-substituted in the context of different neighboring substitutions (compare H2 and H4 with H3 and H7; H2 and H3 with H4 and H7). In the lower stem, ribose substitutions at both the A22−U40 and G21−C41 base pairs show substantial context dependence, with a ~1.2 kcal/mol greater benefit for each ribose-substituted base pair when the other of these two pairs also contains ribose (compare H2 and H12 with H11 and H15; H2 and H11 with H12 and H15), consistent with the apparent cooperative nature of the G21 and A22 conformations mentioned above.

The effects of some ribose substitutions in the nonhelical regions (Figure 2C) were unexpected, in particular the significant preference for deoxyriboses in the bulge (compare H2 and H17). There is NMR evidence that U23, which participates in base triple formation and is the only bulge residue critical for arginine binding, may exist in a C2′endo conformation when bound to argininamide (*18*, *40*, *41*), but the observed effect on binding is not specific to U23 (compare H18 and H19). In the context of an all-ribose bulge, it also is slightly favorable to have riboses in the loop (compare H17 and H20), resulting in an affinity similar to the all-RNA version of HIV TAR.

CD of HIV TAR DNA/RNA Hybrids. The binding data demonstrate the importance of 2′-hydroxyls at the base pairs immediately flanking the HIV TAR bulge. To assess whether these ribose substitutions result in localized A-form geometries or whether the helix propagates further along the stems, we estimated the relative A- and B-form helical content by CD, which gives characteristic spectra for the two helical forms (*42*). The CD spectrum of the all-RNA version of HIV TAR is characteristic of the A-form geometry, with strong positive ellipticity near 265 nm and much weaker negative ellipticity near 240 nm, whereas the all-DNA version shows the characteristics of a B-form spectrum, with a smaller and broader positive peak near 280 nm and negative ellipticity near 250 nm (Figure 3A). DNA/RNA hybrids, containing the important ribose base pairs in the lower (H7) or upper (H15) stems, or the tight binding H2 hybrid, containing substitutions in both stems, produced intermediate spectra. On the basis of the ratio of signals at 266 nm (A-form) and 284 nm (B-form), we estimate an 8% A-form content for H7, 14% for H15, and 19% for H2, consistent with localized induction of A-form geometry near the arginine-binding region and little propagation beyond the sites of ribose substitution.

We next asked whether arginine binding might lead to propagation of the A-form geometry beyond the binding region. CD spectra recorded in the presence of 10 mM argininamide (Figure 3B) are not grossly different from those taken in its absence, with estimated A-form contents of 7% for H7, 13% for H15, and 14% for H2. However, the CD difference spectra of the tight-binding H2 hybrid shows a

significant decrease in signal near 280 nm (Figure 3C), which is very similar to that observed with TAR RNA (*16*) and indicates a change in base stacking that is not seen with the poor-binding H7 or H15 hybrids. An opposite change in the CD signal is observed with the all-DNA analogue, possibly suggesting some nonspecific interactions. The results further suggest that the effects of the 2′-hydroxyl groups on the helix structure remain largely localized to the sites of substitution.

BIV Tat Peptide Binding to BIV TAR DNA/RNA Hybrids. While the argininamide−HIV TAR interaction highlights the importance of localized helix geometry, the analysis of hybrids based on the more extensive BIV Tat−TAR/ARM−RNA interface addresses additional features of recognition that may be relevant to more complex systems. This complex also involves a specific arginine−RNA interaction that is virtually identical to that in the HIV complex, but it is presented in the context of a β-hairpin peptide along with seven other required amino acids that form an interface comparable in surface area to those seen in DNA−protein complexes (*43*, *44*). To visualize the possible consequences of altering the helix geometry on BIV TAR recognition, as for the HIV TAR interaction, we generated models, based on an NMR complex, in which the upper and lower stems surrounding the two single-nucleotide bulges (Figure 1A) were replaced by idealized A- or B-form helices (Figure 1C). By considering just two essential interactions, the Arg73−G11 contact that is analogous to the arginine−G26 interaction in HIV TAR and the Thr72−C23 phosphate contact, it is apparent that a major reorientation of the β-hairpin would be required to maintain important contacts at both ends of the upper stem in a B-form conformation. The wider major groove of a B-form helix would poorly conform to the width of the β-hairpin, and numerous interactions to RNA base and backbone atoms, with their positions altered by differences in base pair tilt and displacement from the helix axis, would be disrupted. Given the diverse and flexible nature of ARM peptides and the preponderance of long, flexible arginine side chains (*9*, *45*), it is of particular interest to know to what extent a peptide such as BIV Tat can adapt to either localized or extensive changes in helical geometry.

BIV Tat peptide binding affinities were measured for a series of BIV TAR DNA/RNA hybrids using affinity chromatography and fluorescence anisotropy based assays, which are in good agreement with one another. The all-DNA version of BIV TAR binds >1000-fold more weakly than the RNA version, whereas an analogue with three ribose base pairs in the lower stem and riboses in the entire upper stem, where the peptide binds (B1), has a similar affinity to that of the RNA (Figure 4A). In examining the upper stem requirements (Figure 4A), we found that removing all 2′-hydroxyls (B7) reduced binding to the level of DNA, while substituting riboses only at base pairs G11−C25 and A13−U24 to generate an analogue (B10) similar to the tight arginine-binding H2 HIV TAR hybrid created a hybrid that still bound with ∼100-fold lower affinity than the RNA. However, adding just one more ribose base pair at the top of the stem (B9) brought affinity to within a factor of 8 of the RNA, and adding one additional ribose pair to the lower stem (B3) brought affinity to within a factor of 2 of the RNA. Surprisingly, adding riboses to base pairs G14−C23 and C15−G22, both of which are involved in specific peptide contacts, had less of an effect than substitution of the upper

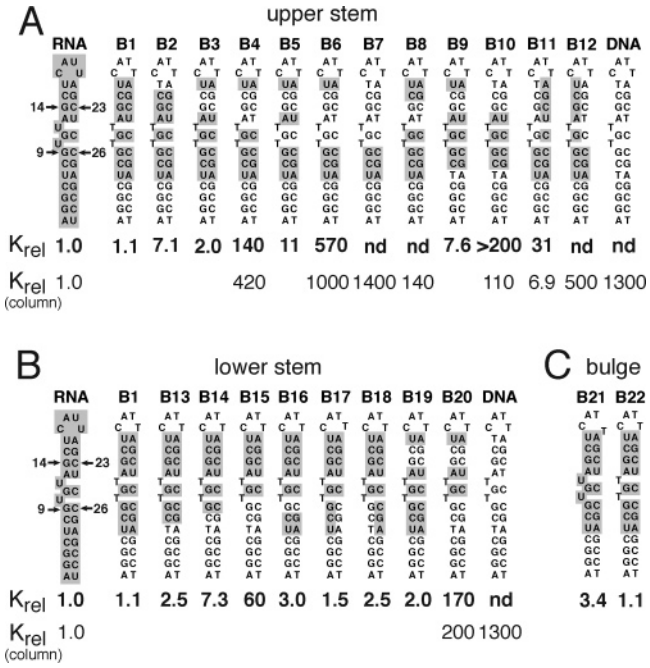

FIGURE 4: Binding affinities of BIV Tat peptides for BIV TAR hybrids. Positions of ribose substitution are highlighted. K_{rel} is the dissociation constant of each hybrid relative to that of BIV TAR RNA, as determined by fluorescence anisotropy. For hybrids in which peptide binding could not be measured accurately by anisotropy (nd), or for some selected hybrids, K_{rel} was estimated by affinity chromatography [K_{rel} (column)]. Substitutions in (A) the upper stem, (B) the lower stem, and (C) the bulge are shown. Numbered arrows by the RNA hairpin in A and B indicate the nucleotide numbering scheme.

base pair (compare B2 and B3). Thus, despite the very large differences in the interaction surfaces, the 2′-hydroxyl requirements for BIV TAR are nearly identical to those of HIV TAR, with some additional benefit derived from riboses at the top base pair and at one additional base pair in the lower stem. At both of these locations, mutagenesis has indicated a role in binding that is independent of specific contacts (*43*) and may be related to helical stability, which should be reduced in the context of a DNA/RNA hybrid (*46*, *47*).

Despite the general similarity in ribose requirements for binding to HIV and BIV TARs, the two show rather different local sensitivities to ribose substitution in the upper stem. Most notably, the G26−C39 base pair that hydrogen bonds to arginine in HIV TAR is ∼6-fold more sensitive to substitution than the A27−U38 base pair that participates in the base triple (compare H3 and H4 in Figure 2A), whereas the converse is true for the corresponding G11−C25 and A13−U24 base pairs in BIV TAR (compare B4 and B5 in Figure 4A). Peptide binding to BIV TAR shows a marked strand asymmetry, with a strong preference for riboses in the 3′-strand (Figure 4A, B11 and B12) versus little preference in HIV TAR (Figure 2A, H5 and H6). Additionally, a comparison of $\Delta\Delta G$ values between related hybrids, as described above for HIV TAR, reveals almost no context dependence for the G11−C25 and A13−U24 ribose substitutions. The observed differences of ∼0.2 kcal/mol (compare B3 and B5 to B4 and B6; B3 and B4 to B5 and B6) are much less than seen with the HIV TAR upper stem hybrids (discussed above).

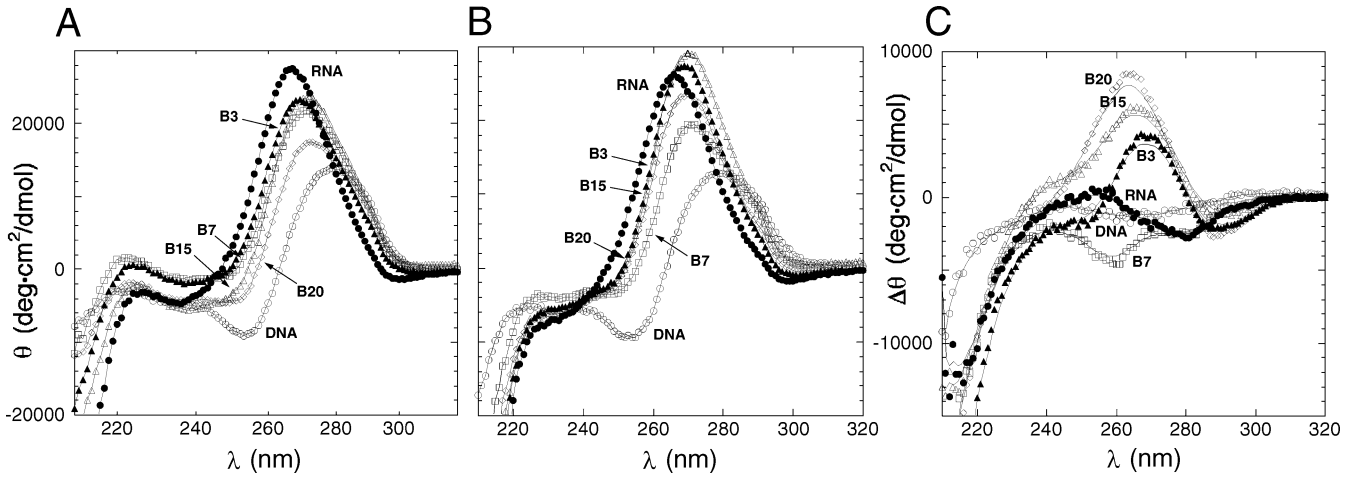

FIGURE 5: CD spectra of BIV TAR RNA (●), BIV TAR DNA (○), B15 (△), B3 (▲), B20 (◇), and B7 (□) in the (A) absence and (B) presence of stoichiometric BIV Tat peptide. (C) Difference spectra calculated by subtracting the spectra in the absence of peptide from the corresponding spectra in the presence of peptide.

In the BIV TAR lower stem, removing all 2′-hydroxyls is less detrimental than removing them from the upper stem, but still reduces binding by ~60-fold (Figure 4B, B1 and B15). Ribose substitutions at two of the three top lower stem base pairs restore most of the binding affinity (Figure 4B, B13, B14, and B16). There is no apparent strand preference for 2′-hydroxyls in the lower stem of BIV TAR (Figure 4B, B17 and B18), unlike the preference for riboses in the 5′-strand in HIV TAR (Figure 2B, H13 and H14). The context dependence of the G9−C26 and U7−A28/C8−G27 pairs also is smaller than in HIV TAR, with $\Delta\Delta G$ values of ~0.7 kcal/mol (compare B14 and B15 with B1 and B16, and compare B1 and B14 with B15 and B16), probably reflecting less conformational coupling in the lower stem. Thus, as with the upper stem, the overall regions affected by ribose substitution in the lower stem are similar in the two TAR cases but the local sensitivities are quite different.

In the bulge region of BIV TAR, the presence of 2′-hydroxyls is slightly unfavorable for binding (Figure 4C, B21 and B22), as also observed with HIV TAR (Figure 2C). This is consistent with NMR experiments on both complexes which indicate that the bulge nucleotides exist primarily in the C2′ endo conformation (*18*, *20*, *21*, *29*, *30*).

CD of BIV TAR DNA/RNA Hybrids. It is interesting that the G14−C23 and C15−G22 base pairs in the upper stem, whose identities are critical for BIV Tat peptide binding, bind similarly in the presence or absence of 2′-hydroxyl groups, whereas the U16−A21 pair at the top of the stem clearly prefers the 2′-hydroxyl substitution, allowing the minimally substituted B3 hybrid (Figure 4A) to bind with nearly the same affinity as BIV TAR RNA. We used CD to assess whether the entire upper stem might be cooperatively switched to the A-form geometry in this hybrid, either with or without BIV Tat bound, or whether some localized non A-form helix in the middle of the binding site might be accommodated by the peptide.

Like HIV TAR, the spectra of the RNA and DNA versions of BIV TAR are characteristic of A- and B-form structure (Figure 5A). The B20 hybrid, with three ribose pairs in the upper stem, has an estimated 35% A-form content; the B15 hybrid, with an all-ribose upper stem, has 47% A-form content, and the tight-binding B3 hybrid, with additional substitutions in the lower stem, has 65% A-form content. Thus, it appears that the A-form helix does not propagate across the G14−C23 and C15−G22 base pairs, since substitution with ribonucleotides increases the A-form content proportionally despite having flanking ribose base pairs in B20. The B7 hybrid, which contains three ribose pairs in the lower stem and none in the upper stem, has an unexpectedly high A-form content of 47%. We suspect that some propagation of A-form structure may occur in the lower stem or at the junction of the stems, which are better stacked in unbound BIV TAR than in HIV TAR due to differences in the bulge configurations (*29*, *30*).

Upon BIV Tat peptide binding, a significant shift is observed in the spectra of the two upper stem hybrids (B20 and B15), with an increased signal near 265 nm (Figure 5B,C), suggestive of a switch to a more A-form-like conformation. The A-form content of hybrid B20 is estimated to increase from 35% to 55%, and the A-form content of hybrid B15 is estimated to increase from 47% to 60%, perhaps reflecting propagation of helix structure into the lower stem driven by the energy of binding. Little change is seen with the tight-binding B3 hybrid, with a calculated change in A-form content from 65% to 66%, suggesting that any required A-form structure, which does not seem to include the G14−C23 and C15−G22 base pairs, may be preformed in this hybrid. The small change in calculated A-form values for B3 upon binding is consistent with gel mobility measurements described below, although we note some change in the difference spectrum (Figure 5C) that may reflect the CD characteristics of an A−B junction rather than A-form content per se. In contrast to the upper stem hybrids, BIV TAR RNA and DNA show little change in the CD spectra in the presence of peptide, whereas B7, with substitutions only in the lower stem, shows a slight decrease in signal near 265 nm (Figure 5B,C).

Native Gel Analyses of BIV TAR DNA/RNA Hybrids. Oligonucleotide mobility on native gels has been shown to be proportional to the fraction of B-form helix and thus can provide some indication of helical form (*42*, *48*, *49*). The relative mobilities of the BIV TAR hybrids generally correlate well with the level of ribose substitution in both the upper (Figure 6A) and lower (Figure 6B) stem hybrids, supporting the premise that the effects of ribose substitution are relatively additive and localized, with little propagation

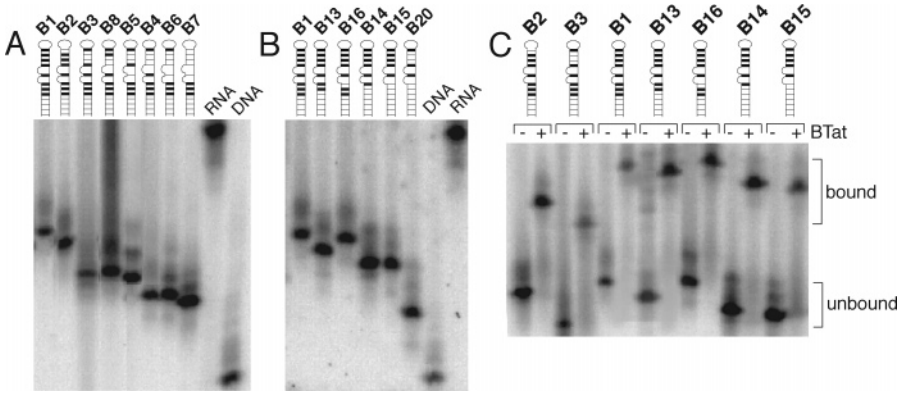

FIGURE 6: Native gel electrophoresis of BIV TAR hybrids in (A) the upper stem and (B) the lower stem. (C) Gel mobility shifts in the absence (−) and presence (+) of saturating (2 μM) BIV Tat peptide for the indicated hybrids. The positions of BIV Tat-bound complexes and unbound hybrids are indicated in (C). Schematic drawings above each lane mark the locations of ribose substitutions with bold lines. The additional, minor bands migrating with slower mobilities probably reflect some degree of conformational heterogeneity but are unlikely to affect the qualitative interpretation of the mobility results.

of helical conformation. The similar mobilities of B4 and B6, which differ by one ribose substitution in the upper stem, or the slightly different mobilities of hybrids with similar levels of substitution (compare B4 and B5 in Figure 6A, and compare B13 and B16 in Figure 6B), however, suggest some propagation at the lower stem−bulge boundary, consistent with the CD results. The mobilities of hybrids with and without ribose substitution at the G14−C23 and C15−G22 base pairs are substantially different (compare B1 and B3 in Figure 6A, and compare B15 and B20 in Figure 6B), consistent with the degree of chemical substitution and a localized difference in helix geometry.

Relative mobility shifts of BIV Tat peptide−hybrid complexes (Figure 6C) are consistent with conformational changes in the lower stem, as inferred from CD, and with a localized helix geometry. All hybrids containing five ribose pairs in the upper stem (B1, B13, B16, B14, and B15) form complexes of low mobility and thus are mostly A-like. Interestingly, the B2 hybrid, which lacks riboses at the top base pair of the upper stem, forms a complex with faster mobility than the other hybrids containing all-ribose upper stems, even though unbound B2 migrates similarly to B13 and slower than B14 and B15 in their unbound states. Thus, peptide binding may drive formation of A-form structure in the lower stem when it is only partially ribose-substituted (B13, B14, and B16) or even completely unsubstituted (B15). The greater mobility of B2 also suggests that, among the hybrids shown in Figure 6C, the degree of upper stem substitution is the primary determinant of mobility differences and, by inference, differences in the amount of A-form helix. The tight-binding B3 hybrid forms the fastest mobility complex, suggesting that the upper stem and, presumably, the deoxyribose G14−C23 and C15−G22 pairs are not entirely A-form.

DISCUSSION

We have defined the minimal 2′-hydroxyl group requirements for specific recognition of HIV TAR by arginine and BIV TAR by a BIV Tat peptide. In both cases, we hypothesize that the hydroxyl groups create an appropriate A-form geometry for major groove binding and thus serve an indirect role in recognition. We base our interpretations on the assumption that the ribose substitutions favor the C3′ endo conformation and are sufficient to alter the local helix geometry, although, clearly, the surrounding context also can influence the conformation. Surprisingly, the helical requirements appear highly localized to the binding regions, and the effects of ribose substitution on helix structure seem largely restricted to the altered base pairs. Furthermore, in the case of BIV TAR, it seems that two essential base pairs in the heart of the binding site need not be in an A-form conformation and we surmise that the BIV Tat ARM peptide is able to adapt to a more B-like conformation. Given the localized nature of the effects, we interpret the data with respect to the possible consequences of changing helix geometry on individual contacts observed in the NMR structures of the two complexes.

HIV TAR−Arginine Interactions. Ribose sugars at four base pair positions in HIV TAR (Figure 2A, H2) are both necessary and sufficient for arginine binding, and the requirement for each can be rationalized based on structural data. First, the NMR structures of TAR−argininamide complexes (*18*, *20*, *21*) position the guanidinium group of arginine to hydrogen bond to the G26 base and simultaneously to make electrostatic contacts to the A22 and U23 phosphates (Figure 1A). In the A-form geometry, the phosphates are substantially displaced along the *z*-axis relative to the center of a base pair (z_p) (*7*). This would be expected to position the phosphate of A22, and potentially of U23, near the upper stem and in closer proximity to the G26 base than would be expected in a B-form helix (Figure 1B). Consistent with this suggestion, removing 2′-hydroxyl groups from the A22−U40 base pair in the H2 context (Figure 2B, H12) reduces arginine binding affinity by ∼30-fold. Second, removing hydroxyl groups from the G26−C39 base pair (Figure 2A, H4), which reduces affinity by 60-fold, is likely to reposition G26 relative to the backbone phosphates and thereby disrupt the guanidinium binding interaction.

Third, it has been proposed that an overtwisting of the helix at the A22−G26 step helps create a pocket between A22 and U23 in which arginine can stack (*19*), and the greater twist of the B-form geometry may be expected to alter the base overlap and stacking interactions in this region. Removal of the 2′-hydroxyl groups from G21 and A22 in the lower stem decreases binding by 40-fold (Figure 2B,

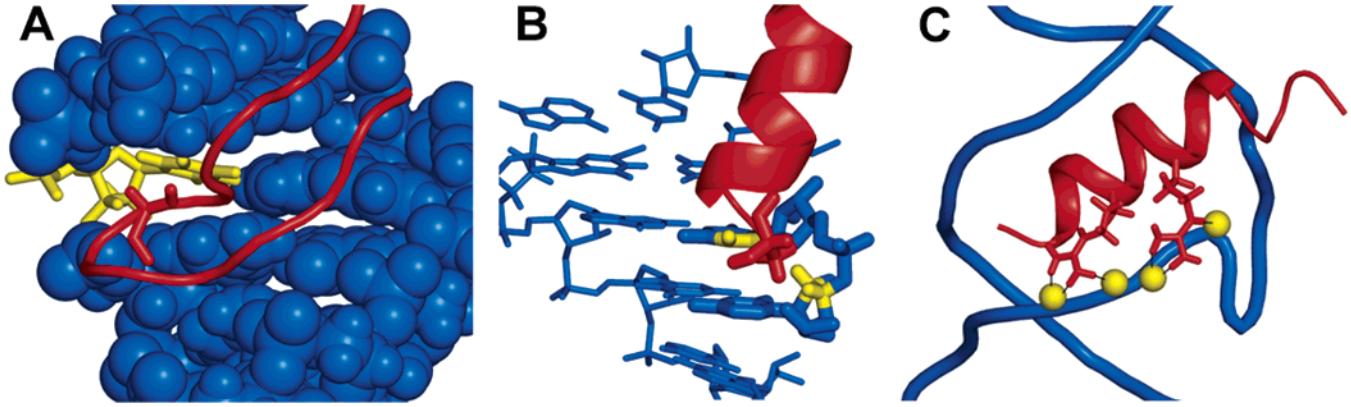

FIGURE 7: Examples of ARM interactions sensitive to helical conformation. (A) In the BIV TAR−Tat peptide complex (1MNB) (*29*), the pocket formed by the U24 nucleotide (yellow) accommodates Gly74 and the turn of the β-hairpin peptide (red). (B) In the bacteriophage λ boxB RNA−N peptide complex (1FQF) (*55*), Ala3 of the peptide α-helix (red) occupies the hydrophobic pocket created by the C2′ and C3′ atoms of cytosine4 and the C5 and C6 atoms of cytosine5 (yellow). (C) In a Rev peptide−aptamer complex (1ULL) (*24*), bivalent hydrogen bonds are formed between nonbridging oxygen atoms at consecutive phosphates (yellow) and arginine residues (red). Arg38 interacts with the oxygens of C22 and U23, and Arg42 interacts with the oxygens of U23 and G24.

H14), but only a small effect is observed when they are removed from the bases on the opposite strand (Figure 2B, H13). The G21 ribose may indirectly affect binding via stacking with A22, since deoxyribose substitution of the G21−C41 base pair (Figure 2B, H11) also reduces binding affinity by 20-fold. Thus, the G21/A22/arginine/U23 stacking network may be energetically coupled and may respond cooperatively to changes in sugar composition, reminiscent, on a local scale, of the cooperative changes in helix geometry proposed to take place in ethanol (*35*, *36*).

Fourth, a base triple between U23 and the A27−U38 base pair in the upper stem (Figure 1A) is formed upon arginine binding (*18*, *21*). Removing hydroxyl groups from the A27−U38 pair (Figure 2A, H3) reduces binding affinity by 10-fold, perhaps by repositioning the base pair and disrupting the triple. The moderate effect on binding is consistent with other results using base triple mutants (*18*, *32*, *50*). Thus, the requirement for riboses at each of the four base pairs in HIV TAR has a plausible structural interpretation, supporting the view that the main effects of each ribose substitution are highly localized within or adjacent to the site of substitution.

BIV TAR−Tat Peptide Interactions. This study reveals several new features of the BIV TAR−Tat interaction, described below, and also provides an interesting comparison to the HIV complex. The arginine-binding module of HIV TAR is conserved in the BIV complex and is very similar in structural detail, with Arg73 used to contact the G11 base and phosphate backbone in the BIV case (Figure 1A) (*28*−*30*). The ribose requirements at the arginine-binding site also appear to be similar, with the BIV interaction requiring 2′-hydroxyls at the two base pairs immediately flanking the U10 bulge in both the upper and lower stems (Figure 4). The sequences of the lower stems differ, and studies of BIV TAR mutants showed that its sequence is important for binding, probably due to subtle stacking interactions (*28*). Correspondingly, the BIV TAR−peptide interaction is enhanced by ribose substitution at one additional base pair in the lower stem (Figure 4, B3 and B9). The only other difference in ribose requirements between the two complexes is the need for ribose at the top base pair of the upper stem in BIV TAR, as discussed below.

The striking similarity in ribose requirements between HIV and BIV TAR seems remarkable given the large difference in the scale of the binding surfaces between arginine and the BIV peptide (Figure 1B,C). However, there are quantitative differences that likely reflect other constraints on the peptide complex. For example, the relative importance of the two upper stem base pairs is reversed: deoxy substitution of the A13−U24 base pair in BIV TAR reduces affinity by 140-fold and substitution of G11−C25 reduces affinity by 10-fold (Figure 4A, B4 and B5), whereas the converse is true for HIV TAR (Figure 2A, H3 and H4). In addition, there are different strand preferences for the two complexes, with 2′-hydroxyl groups being especially important on the 3′-strand of the upper stem of BIV TAR (Figure 4A, B11 and B12) and on the 5′-strand of the lower stem of HIV TAR (Figure 2B, H13 and H14). The preference for riboses in the 3′-strand of the BIV site, as well as the requirement for riboses at the A13−U24 base pair, which is surprisingly insensitive to mutation, can be explained by the positioning of Gly74 between the base and ribose of U24 (*27*, *29*, *30*). Gly74 is an essential residue that is deeply buried within the RNA−peptide interface and held rigidly to form the turn of the β-hairpin (Figure 7A) (*26*). A C2′ endo conformation at U24, in concert with the 5-methyl group of thymine, would substantially reduce the size of the Gly74 binding pocket, forcing a reorientation of the β-hairpin within the groove to maintain other essential RNA contacts. Furthermore, the overtwisting of the A-form helix in the arginine-binding module increases the depth of the pocket, probably further enhancing binding specificity. Thus, the β-turn appears particularly well-adapted to the A-form RNA helix geometry.

Our data suggest that some A-form helix may propagate into the lower stem upon peptide binding. It is observed that Arg77 interacts with the G9−C26 base pair at the top of the lower stem, which appears to stabilize an extensive stacking network involving G9/Arg73/U10/Arg70, with Ile79 stacking against the C5/C6 surface of U10 (*29*, *30*). Thus, the energy of peptide binding to the 5′-strand of the upper stem may help extend the A-form helix to the lower stem. In contrast, HIV TAR forms a less extensive G21/A22/Arg/U23 stacking arrangement, and we observe that ribose substitutions in the 5′-strand of the lower stem result in tighter binding than

substitutions in the 3′-strand (H13 and H14, Figure 2B), an asymmetry not observed with BIV TAR. Thus, HIV TAR may require the A-form geometry of the lower stem prior to binding to help stabilize its stacking network.

Perhaps the most striking result is that 2′-hydroxyl groups are not needed at two essential base pairs (G14−C23 and C15−G22) in the core of the BIV TAR binding site, located just above the arginine-binding module (Figure 1A). CD and native gel data suggest that these pairs are not converted to an A-form geometry despite flanking ribose pairs in the B3 hybrid (see Figure 4A). In the BIV complex, Arg70 hydrogen bonds to the O6 and N7 groups of G14, Gly71 hydrogen bonds to G22, and Thr72 may hydrogen bond to C23 and stack on the ribose. Because Arg70 does not make additional interactions to the phosphates or form other stacking interactions seen in the arginine-binding module, it might be expected to be largely insensitive to changes in helix geometry. However, the Gly71 and Thr72 interactions, which both contribute substantially to binding affinity, might be expected to be sensitive to helix geometry as both are largely buried on the inner strand of the β-hairpin and Thr72 contacts both a base and the backbone (*26*). In these cases, the adaptability of the ARM and the flexible nature of the arginine side chains may allow substantial adjustments to accommodate changes in helix geometry (*9*−*11*, *45*).

Finally, the sensitivity of BIV TAR to deoxy substitutions in the top base pair of the upper stem (Figure 4A, B10) may reflect fraying of the B-form stem induced by the immediately adjacent loop. Mutagenesis of BIV TAR RNA has shown that having a Watson−Crick base pair at this position is important, but that its identity does not matter (*28*). It is generally the case that DNA stems are less thermostable than the corresponding RNA stems (*42*), consistent with the larger loss of affinity of hybrid B10 observed in a fluorescence anisotropy assay at 25 °C than in an affinity chromatography assay at 4 °C (Figure 4A). Furthermore, the loss of affinity can largely be suppressed by ribose substitutions at G14−C23 and C15−G22 (Figure 4A, B2 and B10) despite the lack of a requirement for 2′-hydroxyls at those pairs (see above).

Implications for Other RNA−Protein Interactions. Both the requirements for specific A-form geometries and the mosaic character of a helical binding site undoubtedly apply to other RNA−protein complexes. The arginine-binding module discussed here is found in two aptamers that bind the HIV Rev peptide and in the HTLV RxRE RNA element, and both likely will show geometric requirements similar to those of HIV and BIV TAR (*22*, *24*, *51*, *52*). The bacteriophage λ and P22 N peptide−boxB RNA and RSG-1.2 peptide−RRE RNA complexes all utilize an ARM in which alanine is seen to interact in a pocket defined by the 2′- and 3′-carbons of the ribose sugar and the C5 and C6 carbons of a pyrimidine base (Figure 7B) (*53*−*56*). The confluence of hydrophobic groups is a consequence of the C3′ endo ribose conformation and is the same pocket occupied by Gly74 of the BIV Tat peptide (Figure 7A), suggesting that this may be a relatively common recognition feature in the RNA major groove. Another interaction likely to depend on helix geometry uses the arginine guanidinium group to make electrostatic contacts and/or hydrogen bonds to the oxygens of consecutive phosphates (Figure 7C). These types of multivalent "arginine fork" interactions (*15*), which can distinguish between helix geometries because the distance between neighboring phosphate oxygens is ∼1.4 Å shorter in the A-form than in the B-form (*6*), have been observed in several ARM complexes (*24*, *54*, *57*). On the other hand, it can be inferred that interactions made only to a base, such as an arginine−guanine interaction, may allow peptides or proteins to utilize particularly flexible side chains and electrostatic complementarity to adapt to different, perhaps locally unique, helix geometries (*45*, *58*). For this reason, the ARM may be especially well-suited to recognize a variety of helical contexts, including B-form helices found in structured single-stranded DNA elements (S.G.L. and A.D.F., in preparation).

Studies of other RNA−protein interactions using 2′-deoxy substitutions further highlight the importance of helix geometry. In the *E. coli* $tRNA^{Ala}$−synthetase complex, where recognition occurs in the minor groove, and the MS2 coat protein−operator complex, where stem and loop sequences are recognized, removal of 2′-hydroxyl groups believed to directly hydrogen bond to the protein reduced binding affinity by as much as 20-fold (*59*, *60*). However, DNA analogues containing only the critical ribose groups still showed 20-fold reduced affinity, implying that helical conformation probably also is important. The double-stranded RNA-binding domain (dsRBD) provides the most extreme example of helix-specific recognition. Each modular dsRBD recognizes about one helical turn of RNA and makes a set of hydrogen bonding interactions to 2′-hydroxyls and phosphate oxygens that precisely span the major, and perhaps the minor, groove of an A-form helix (*61*, *62*). The dsRBDs do not bind hybrids composed of one RNA and one DNA strand, and a small number of deoxyribose substitutions can significantly affect binding (*63*). Thus, as exemplified by these cases and the TAR complexes, indirect readout of helical conformation can be as important energetically for high-affinity binding as base-specific interactions or recognition of other tertiary features of RNA structure. It will be interesting to test the importance of localized helical changes in other complexes and to determine what types of RNA or even DNA conformational features can be effectively insulated from neighboring regions.

ACKNOWLEDGMENT

We thank Alex Ramirez for help with anisotropy assays, Valerie Calabro for help with modeling, members of the Frankel laboratory for helpful suggestions, and Valerie Calabro, Chandreyee Das, Matthew Daugherty, Yong-beom Lim, and Dennis Wang for comments on the manuscript.

REFERENCES

1. Cheng, A. C., Chen, W. W., Fuhrmann, C. N., and Frankel, A. D. (2003) Recognition of nucleic acid bases and base-pairs by hydrogen bonding to amino acid side-chains, *J. Mol. Biol. 327*, 781−796.
2. Garvie, C. W., and Wolberger, C. (2001) Recognition of specific DNA sequences, *Mol. Cell 8*, 937−946.
3. Nekludova, L., and Pabo, C. O. (1994) Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes, *Proc. Natl. Acad. Sci. U.S.A. 91*, 6948−6952.
4. Suzuki, M., and Yagi, N. (1996) An in-the-groove view of DNA structures in complexes with proteins, *J. Mol. Biol. 255*, 677−687.

5. Shi, Y., and Berg, J. M. (1995) Specific DNA-RNA hybrid binding by zinc finger proteins, *Science 268*, 282−284.
6. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
7. Lu, X. J., Shakked, Z., and Olson, W. K. (2000) A-form conformational motifs in ligand-bound DNA structures, *J. Mol. Biol. 300*, 819−840.
8. Weeks, K. M., and Crothers, D. M. (1993) Major groove accessibility of RNA, *Science 261*, 1574−1577.
9. Frankel, A. D. (2000) Fitting peptides into the RNA world, *Curr. Opin. Struct. Biol. 10*, 332−340.
10. Patel, D. J. (1999) Adaptive recognition in RNA complexes with peptides and protein modules, *Curr. Opin. Struct. Biol. 9*, 74−87.
11. Narayana, N., and Weiss, M. A. (1999) RNA recognition by arginine-rich peptide motifs, *Biopolymers 48*, 167−180.
12. Xie, B., Wainberg, M. A., and Frankel, A. D. (2003) Replication of human immunodeficiency viruses engineered with heterologous Tat-transactivation response element interactions, *J. Virol. 77*, 1984−1991.
13. Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, J. G., Gait, M. J., and Karn, J. (1993) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region, *J. Mol. Biol. 230*, 90−110.
14. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A., and Valerio, R. (1989) Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro, *Proc. Natl. Acad. Sci. U.S.A. 86*, 6925−6929.
15. Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., and Frankel, A. D. (1991) Arginine-mediated RNA recognition: the arginine fork, *Science 252*, 1167−1171.
16. Tan, R., and Frankel, A. D. (1992) Circular dichroism studies suggest that TAR RNA changes conformation upon specific binding of arginine or guanidine, *Biochemistry 31*, 10288−10294.
17. Tao, J., and Frankel, A. D. (1992) Specific binding of arginine to TAR RNA, *Proc. Natl. Acad. Sci. U.S.A. 89*, 2723−2726.
18. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy, *Science 257*, 76−80.
19. Brodsky, A. S., and Williamson, J. R. (1997) Solution structure of the HIV-2 TAR-argininamide complex, *J. Mol. Biol. 267*, 624−639.
20. Aboul-ela, F., Karn, J., and Varani, G. (1995) The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein, *J. Mol. Biol. 253*, 313−332.
21. Long, K. S., and Crothers, D. M. (1999) Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA, *Biochemistry 38*, 10059−10069.
22. Baskerville, S., Zapp, M., and Ellington, A. D. (1999) Anti-Rex aptamers as mimics of the Rex-binding element, *J. Virol. 73*, 4962−4971.
23. Tao, J., and Frankel, A. D. (1996) Arginine-binding RNAs resembling TAR identified by in vitro selection, *Biochemistry 35*, 2229−2238.
24. Ye, X., Gorin, A., Ellington, A. D., and Patel, D. J. (1996) Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex, *Nat. Struct. Biol. 3*, 1026−1033.
25. Ellington, A. D., Leclerc, F., and Cedergren, R. (1996) An RNA groove, *Nat. Struct. Biol. 3*, 981−984.
26. Chen, L., and Frankel, A. D. (1995) A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA, *Proc. Natl. Acad. Sci. U.S.A. 92*, 5077−5081.
27. Chen, L., and Frankel, A. D. (1994) An RNA-binding peptide from bovine immunodeficiency virus Tat protein recognizes an unusual RNA structure, *Biochemistry 33*, 2708−2715.
28. Smith, C. A., Crotty, S., Harada, Y., and Frankel, A. D. (1998) Altering the context of an RNA bulge switches the binding specificities of two viral Tat proteins, *Biochemistry 37*, 10808−10814.
29. Puglisi, J. D., Chen, L., Blanchard, S., and Frankel, A. D. (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex, *Science 270*, 1200−1203.
30. Ye, X., Kumar, R. A., and Patel, D. J. (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex, *Chem. Biol. 2*, 827−840.
31. Macke, T., and Case, D. A. (1998) in *Molecular Modeling of Nucleic Acids* (Leontes, N. B., and SantaLucia, J., Jr., Eds.) pp 379−393, American Chemical Society, Washington, DC.
32. Tao, J., Chen, L., and Frankel, A. D. (1997) Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction, *Biochemistry 36*, 3491−3495.
33. Darby, N., and Creighton, T. E. (1995) Disulfide bonds in protein folding and stability, *Methods Mol. Biol. 40*, 219−252.
34. Barnett, R. W., Delling, U., Kuperman, R., Sonenberg, N., and Sumner-Smith, M. (1993) Rotational symmetry in ribonucleotide strand requirements for binding of HIV-1 Tat protein to TAR RNA, *Nucleic Acids Res. 21*, 151−154.
35. Ban, C., Ramakrishnan, B., and Sundaralingam, M. (1994) A single 2′-hydroxyl group converts B-DNA to A-DNA. Crystal structure of the DNA-RNA chimeric decamer duplex d(CCGGC)r(G)d(CCGG) with a novel intermolecular G-C base-paired quadruplet, *J. Mol. Biol. 236*, 275−285.
36. Ivanov, V. I., Minchenkova, L. E., Minyat, E. E., Frank-Kamenetskii, M. D., and Schyolkina, A. K. (1974) The B to A transition of DNA in solution, *J. Mol. Biol. 87*, 817−833.
37. Ivanov, V. I., and Krylov, D. (1992) A-DNA in solution as studied by diverse approaches, *Methods Enzymol. 211*, 111−127.
38. Nishizaki, T., Iwai, S., Ohkubo, T., Kojima, C., Nakamura, H., Kyogoku, Y., and Ohtsuka, E. (1996) Solution structures of DNA duplexes containing a DNA·RNA hybrid region, d(GG)r(AGAU)d(GAC)·d(GTCATCTCC) and d(GGAGA)r(UGAC)·d(GTCATCTCC), *Biochemistry 35*, 4016−4025.
39. Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., and Reid, B. R. (1993) The DNA strand in DNA·RNA hybrid duplexes is neither B-form nor A-form in solution, *Biochemistry 32*, 4207−4215.
40. Hamy, F., Asseline, U., Grasby, J., Iwai, S., Pritchard, C., Slim, G., Butler, P. J., Karn, J., and Gait, M. J. (1993) Hydrogen-bonding contacts in the major groove are required for human immunodeficiency virus type-1 tat protein recognition of TAR RNA, *J. Mol. Biol. 230*, 111−123.
41. Sumner-Smith, M., Roy, S., Barnett, R., Reid, L. S., Kuperman, R., Delling, U., and Sonenberg, N. (1991) Critical chemical features in trans-acting-responsive RNA are required for interaction with human immunodeficiency virus type 1 Tat protein, *J. Virol. 65*, 5196−5202.
42. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) Sequence specific thermodynamic and structural properties for DNA·RNA duplexes, *Biochemistry 33*, 5298−5304.
43. Frankel, A. D., and Smith, C. A. (1998) Induced folding in RNA-protein recognition: more than a simple molecular handshake, *Cell 92*, 149−151.
44. Nadassy, K., Wodak, S. J., and Janin, J. (1999) Structural features of protein-nucleic acid recognition sites, *Biochemistry 38*, 1999−2017.
45. Wilkinson, T. A., Botuyan, M. V., Kaplan, B. E., Rossi, J. J., and Chen, Y. (2000) Arginine side-chain dynamics in the HIV-1 rev-RRE complex, *J. Mol. Biol. 303*, 515−529.
46. Soto, A. M., Gmeiner, W. H., and Marky, L. A. (2002) Energetic and conformational contributions to the stability of Okazaki fragments, *Biochemistry 41*, 6842−6849.
47. Gyi, J. I., Conn, G. L., Lane, A. N., and Brown, T. (1996) Comparison of the thermodynamic stabilities and solution conformations of DNA·RNA hybrids containing purine-rich and pyrimidine-rich strands with DNA and RNA duplexes, *Biochemistry 35*, 12538−12548.
48. Bhattacharyya, A., Murchie, A. I., and Lilley, D. M. (1990) RNA bulges and the helical periodicity of double-stranded RNA, *Nature 343*, 484−487.
49. Roberts, R. W., and Crothers, D. M. (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition, *Science 258*, 1463−1466.
50. Weeks, K. M., and Crothers, D. M. (1991) RNA recognition by Tat-derived peptides: interaction in the major groove?, *Cell 66*, 577−588.
51. Ye, X., Gorin, A., Frederick, R., Hu, W., Majumdar, A., Xu, W., McLendon, G., Ellington, A., and Patel, D. J. (1999) RNA architecture dictates the conformations of a bound peptide, *Chem. Biol. 6*, 657−669.

52. Jiang, F., Gorin, A., Hu, W., Majumdar, A., Baskerville, S., Xu, W., Ellington, A., and Patel, D. J. (1999) Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples, *Struct. Fold. Des. 7*, 1461−1472.
53. Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A., and Patel, D. J. (1998) Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex, *Nat. Struct. Biol. 5*, 203-212.
54. Gosser, Y., Hermann, T., Majumdar, A., Hu, W., Frederick, R., Jiang, F., Xu, W., and Patel, D. J. (2001) Peptide-triggered conformational switch in HIV-1 RRE RNA complexes, *Nat. Struct. Biol. 8*, 146−150.
55. 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.
56. Zhang, Q., Harada, K., Cho, H. S., Frankel, A. D., and Wemmer, D. E. (2001) Structural characterization of the complex of the Rev response element RNA with a selected peptide, *Chem. Biol. 8*, 511−520.
57. Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D., and Williamson, J. R. (1996) Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex, *Science 273*, 1547−1551.
58. Garcia-Garcia, C., and Draper, D. E. (2003) Electrostatic interactions in a peptide- RNA complex, *J. Mol. Biol. 331*, 75−88.
59. Musier-Forsyth, K., and Schimmel, P. (1992) Functional contacts of a transfer RNA synthetase with 2′-hydroxyl groups in the RNA minor groove, *Nature 357*, 513−515.
60. Baidya, N., and Uhlenbeck, O. C. (1995) The role of 2′-hydroxyl groups in an RNA−protein interaction, *Biochemistry 34*, 12363−12368.
61. Ramos, A., Grunert, S., Adams, J., Micklem, D. R., Proctor, M. R., Freund, S., Bycroft, M., St. Johnston, D., and Varani, G. (2000) RNA recognition by a Staufen double-stranded RNA-binding domain, *EMBO J. 19*, 997−1009.
62. Ryter, J. M., and Schultz, S. C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA, *EMBO J. 17*, 7505−7513.
63. Bevilacqua, P. C., and Cech, T. R. (1996) Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR, *Biochemistry 35*, 9983−9994.

BI047916T