Structure, Vol. 11, 533–545, May, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00069-8

Structural Basis of the Highly Efficient Trapping of the HIV Tat Protein by an RNA Aptamer

Akimasa Matsugami,[1] Shin-ichiro Kobayashi,[1] Kiyoshi Ouhashi,[1] Seiichi Uesugi,[1] Rika Yamamoto,[2,4] Kazunari Taira,[2,4] Satoshi Nishikawa,[3] Penmetcha K.R. Kumar,[3] and Masato Katahira[1,*]

[1]Department of Environment and Natural Sciences
Graduate School of Environment and Information Sciences
Yokohama National University
79-7 Tokiwadai
Hodogaya-ku
Yokohama 240-8501
Japan

[2]Gene Discovery Research Center

[3]Institute of Molecular and Cell Biology
National Institute of Advanced Industrial Science and Technology
1-1-1 Higashi
Tsukuba 305-8565
Japan

[4]Department of Chemistry and Biotechnology
Graduate School of Engineering
University of Tokyo
Hongo
Tokyo 113-8656
Japan

Summary

An RNA aptamer containing two binding sites exhibits extremely high affinity to the HIV Tat protein. We have determined the structure of the aptamer complexed with two argininamide molecules. Two adjacent U:A:U base triples were formed, which widens the major groove to make space for the two argininamide molecules. The argininamide molecules bind to the G bases through hydrogen bonds. The binding is stabilized through stacking interactions. The structure of the aptamer complexed with a Tat-derived arginine-rich peptide was also characterized. It was suggested that the aptamer structure is similar for both complexes and that the aptamer interacts with two different arginine residues of the peptide simultaneously at the two binding sites, which could explain the high affinity to Tat.

Introduction

The Tat protein of the human immunodeficiency virus (HIV) plays a crucial role in viral replication [1–3]. Tat stimulates transcription of the viral genome through binding to the *trans*-activating region (TAR) located at the 5′ end of all pre-messenger RNA transcripts [4]. A small portion of TAR is responsible for transcriptional activation and binding of Tat. The sequences of the essential portions of the TARs of HIV type 1 (HIV-1) and type 2 (HIV-2) are shown in Figures 1A and 1B, respectively. The loop sequences are critical for the transcriptional activation but do not affect the binding of Tat to TAR. For the binding of Tat, a U residue (U23) in the bulge and two base pairs above the bulge (G26:C39 and A27:U38) are required [5–9].

Short peptides spanning the arginine-rich region of Tat bind specifically to TAR, recognizing the U23, G26, and A27 residues [5, 7–11]. This suggests that the essence of the TAR-Tat interactions can be examined by studying the TAR-Tat peptide interactions. In spite of the usefulness of TAR-Tat peptide complexes for biochemical studies, NMR spectra of the high quality required for detailed structural analysis could not be obtained for the complexes in the case of an HIV-1 system [12–14]. The high-resolution NMR structure of the TAR-Tat peptide complex has been determined in the case of a BIV system [15, 16].

The TAR-Tat interactions can be further simplified to those in a TAR-arginine complex [17]. A single arginine residue in a lysine-rich peptide binds specifically to TAR, recognizing U23, G26, and A27, as Tat does [11, 17, 18]. The NMR structures of TAR-argininamide complexes have been reported for HIV-1 [12, 13] and HIV-2 [19]. Additionally, the crystal structure of free HIV-1 TAR has been reported [20].

An RNA aptamer that binds Tat 100 times more strongly than the authentic TAR was isolated by means of the in vitro selection method [21]. The sequence of the aptamer implies the existence of two nearly symmetrical Tat binding sites (boxed regions) that are found in TAR RNAs (Figures 1C and 1D). The sequence of the hairpin at the top of the aptamer was originally UUUGC. In order to stabilize the structure, that sequence was replaced by a GAAA sequence, which is known to form a stable tetranucleotide loop structure [22]. The aptamer competed efficiently for the binding of Tat in the presence of a large excess of TAR of either HIV-1 or HIV-2 in vitro [23]. The aptamer specifically prevents Tat-dependent *trans*-activation both in vitro and in vivo [23]. Moreover, unlike TAR, the affinity of the aptamer to Tat does not depend on cellular proteins such as cyclin T1 [23]. Thus, the aptamer has the potential for use as a decoy molecule in gene therapy.

The dissociation constant (K_d) of the aptamer-Tat complex is 10^{-10} M. The K_d of the complex between the aptamer and the partial peptide spanning the arginine-rich region of Tat, RKKRRQRRRPPQG (RG peptide), is 10^{-9} M, the decrease in affinity being only one order. In order to elucidate the structural basis of the extremely high affinity of the aptamer to Tat, the structure of the aptamer complexed with argininamide, the simplest analog of Tat, has been determined by NMR. Unique structural features responsible for the high affinity have been found: the formation of two adjacent U:A:U base triples, the resultant widening of the major groove, the formation

*Correspondence: masakata@ynu.ac.jp

Key words: aptamer; decoy; HIV Tat; RNA-protein interaction; RNA structure

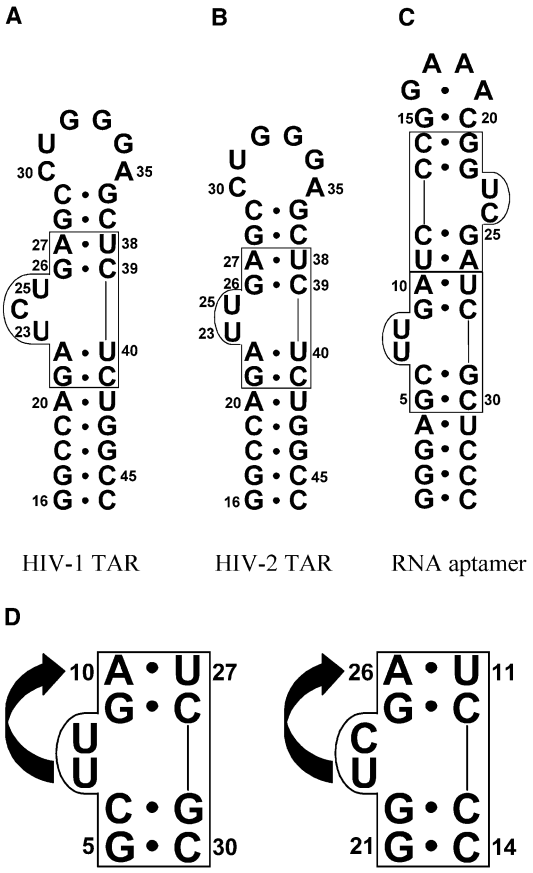

Figure 1. RNA Sequences

HIV-1 TAR (A); HIV-2 TAR (B); and RNA aptamer for the HIV Tat protein (C). Postulated Tat binding sites are boxed. Expected base pairs are indicated by dots.

(D) Comparison of the two binding sites, sites 1 (left) and 2 (right), of the RNA aptamer. The orientation of site 2 is inverted compared to that in (C) for clarity. The formation of U:A:U base triples in the complex is indicated by arrows (see text).

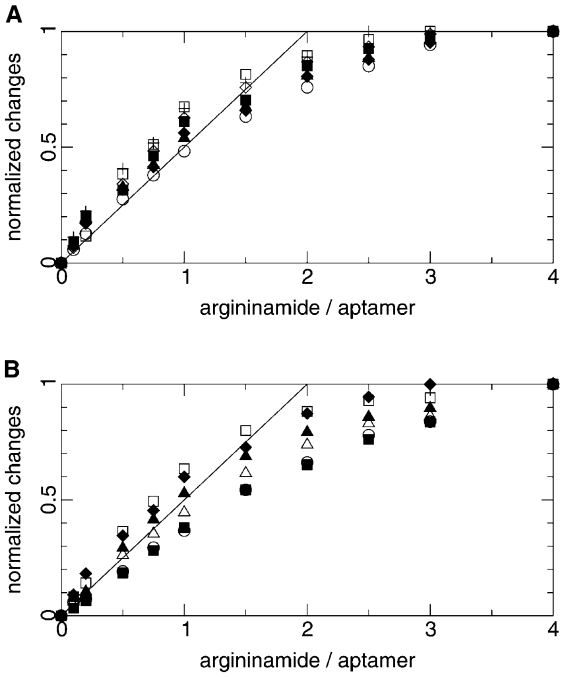

Figure 2. Normalized Changes of Chemical Shifts as a Function of the Molar Ratio of Argininamide to the Aptamer

Normalized changes for H5 (A) and H6 (B) at 35°C and pH 6.5. Open circles, C6; open triangles, U8; open squares, U11; open diamonds, C13; closed circles, C14; closed triangles, C24; closed squares, U27; closed diamonds, C30; and crosses, U31.

of hydrogen bonds between a G base and argininamide, and the stabilization of the binding through stacking interaction of a guanidinium group with bases. Structural characterization of the aptamer complexed with the RG peptide has also been carried out. Simultaneous interactions of the aptamer with two arginine residues of the RG peptide at two binding sites are strongly suggested. A combination of structural studies on the aptamer-argininamide and aptamer-RG peptide complexes provides a comprehensive explanation of the extremely high affinity of the aptamer to Tat.

Results and Discussion

Stoichiometry of the Aptamer-Argininamide Complex

The formation of a complex between the aptamer and argininamide was monitored by tracing the H5-H6 correlation peaks of the U and C residues of the aptamer in a series of TOCSY [24] spectra during titration. It turned out that the free and bound forms of the aptamer were in fast exchange on the NMR time scale. The normalized changes of the chemical shifts of either H5 or H6 resonances (see the next section for assignments) are plotted as a function of the molar ratio of argininamide to the aptamer in Figure 2. For reference, the line joining (0, 0) and (2, 1) is drawn in Figures 2A and 2B. When all data for both H5 (Figure 2A) and H6 (Figure 2B) are taken into account together, the initial slopes of the plots indicate that two argininamide molecules bind per aptamer. This is as expected, because the base sequence of the aptamer implies the existence of two binding sites.

Assignment of Resonances of the Aptamer-Argininamide Complex

Exchangeable ^{1}H and Attached ^{15}N Resonances of the Aptamer

All imino proton-nitrogen correlation peaks expected for the base pairs illustrated in Figure 1C, that is, three A:U, ten G:C, and one G:A base pairs, were observed on ^{1}H-^{15}N HSQC (Figure 3A). The correlation peaks originating from U and G residues were discriminated on the basis of their ^{15}N chemical sifts, 161–163 ppm for U and 145–148 ppm for G. Observation of an upfield-shifted G imino proton at 10.6 ppm is indicative of the formation of the expected sheared G16:A19 base pair in the G16A17A18A19 tetranucleotide loop. Thus, the imino proton resonance of G16 is a good starting point for the assignment. As shown in Figure 3B, imino proton resonances were sequentially assigned from G16 to U11

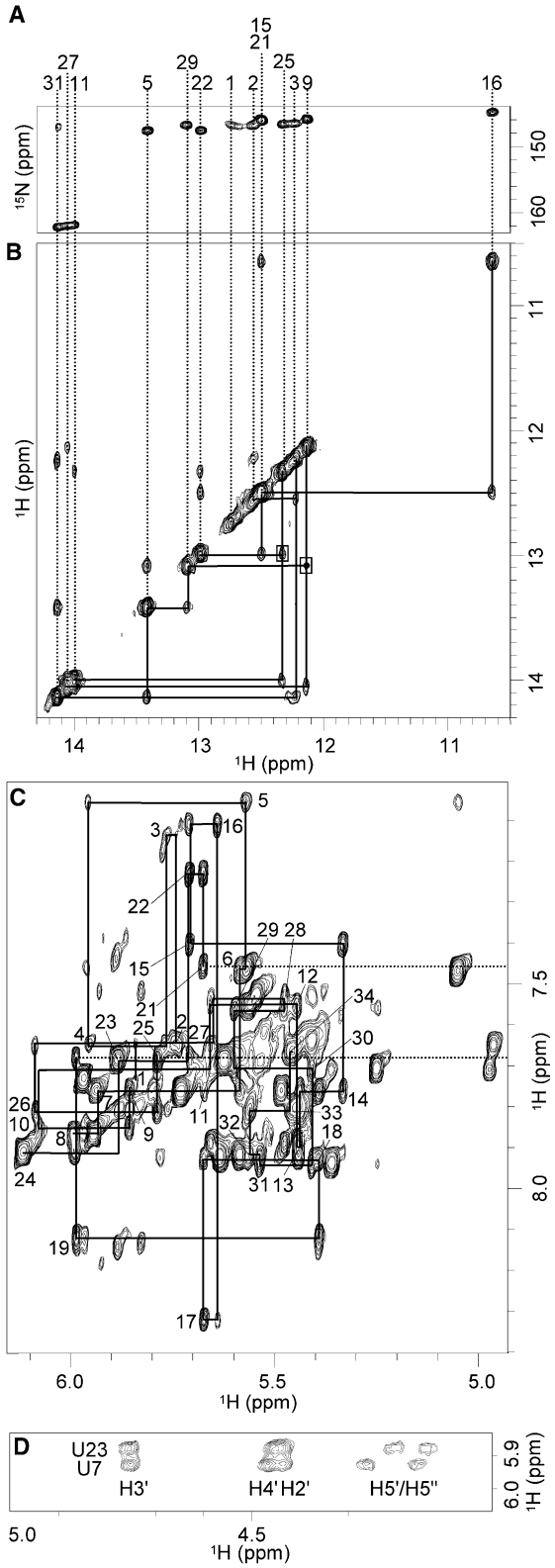

Figure 3. Assignments of Resonances

(A) The imino proton-nitrogen region of the ^{1}H-^{15}N HSQC spectrum of the aptamer-argininamide complex at 1°C and pH 6.5; the assignments are labeled with the residue numbers.

(B) The imino proton region of the NOESY spectrum of the complex with a mixing time of 200 ms at 1°C and pH 6.5; the sequential assignments are indicated. Cross peaks "across a bulge," G29-G9 and G22-G25, are boxed. The G29-G9 cross peak indicated by a dot becomes visible when the level of the plot is lowered.

(C) The H6/H8-H1′/H5 region of the NOESY spectrum of the complex with a mixing time of 200 ms at 35°C and pH 6.5. The lines show the H1′(i−1)-H6/H8(i)-H1′(i) connectivities; the intraresidue H6/H8-H1′ cross peaks are labeled with their residue numbers. The C20H1′-C20H6 and C20H1′-G21H8 cross peaks are outside of this region due to an extremely upfield-shifted C20H1′ resonance.

(D) The correlation from H1′ to H2′/H3′/H4′/H5′/H5″ for the U7 and U23 residues in the HCCH-TOCSY spectrum of the complex with a mixing time of 24.8 ms at 35°C and pH 6.5. The H1′-H2′ and H1′-H4′ peaks of U23 are overlapping.

for the upper half of the aptamer, the overlapping of the G15 and G21 resonances being resolved at higher temperature. Sequential assignments from U27 to G1 were also made for the lower half of the aptamer (Figure 3B).

Eleven pairs of amino proton-nitrogen correlation peaks observed on ^{1}H-^{15}N HSQC were interpreted as those of C residues due to their ^{15}N chemical shifts of 94–100 ppm. Ten of the eleven pairs were assigned to specific C residues on the basis of the NOESY [25] cross peaks between a G imino proton and the C amino protons for each G:C base pair. The remaining amino proton-nitrogen pair was assigned to C24 automatically. Two additional pairs of amino proton-nitrogen correlation peaks observed on ^{1}H-^{15}N HSQC were interpreted as those of G residues due to their ^{15}N chemical shifts of 73–75 ppm. These pairs were assigned to G9 and G25, respectively, on the basis of the intraresidue imino-amino NOESY cross peak for each G residue. Finally, another pair of amino proton-nitrogen correlation peaks observed on ^{1}H-^{15}N HSQC was interpreted as that of an A residue due to its ^{15}N chemical shift of 84 ppm. This was assigned to the overlapping of the amino proton-nitrogen correlation peaks of A10 and A26, because of the strong NOESY cross peaks between these amino protons and the imino protons of U11 and U27.

Nonexchangeable ^{1}H and Attached ^{13}C Resonances

Nonexchangeable ^{1}H resonances were assigned using standard methods [26–28]. Figure 3C shows expansion of the NOESY spectrum, which allows the sequential assignments of H1′ and H6/H8 through H1′(i−1)-H6/H8(i)-H1′(i) connectivities. When the sheared G:A base pair is formed in the GAAA tetranucleotide loop, H1′ of the residue located next to the loop is upfield shifted due to the ring current effect [22]. Observation of an upfield-shifted C20H1′ (3.65 ppm) is indicative of the expected G16:A19 sheared base pair. H2′, H3′, H4′, and H5′/5″ were assigned by means of HCCH-COSY [29] and HCCH-TOCSY [30] spectra (Figure 3D), the ^{13}C resonances of sugars, C1′, C2′, C3′, C4′, and C5′, being assigned at the same time. Unambiguous assignments of the H5′/H5″ methylene protons were made by changing the delay time of reverse INEPT [31]. The assignments were confirmed by successful tracing of the H2′(i−1)-H6/H8(i)-H2′(i) and H3′(i−1)-H6/H8(i)-H3′(i) connectivities (data not shown). The H3′(i−1)-P(i−1)-H4′/H5′/H5″(i) connectivities in the ^{1}H-^{31}P HetCor [32] spectrum provide some additional confirmation of the assignments (data not shown). AH2 was assigned from

the intraresidue correlation with AH8 in the HCCH-TOCSY [33, 34] spectrum. Observation of strong NOESY cross peaks between AH2 and UH3 for the A4:U31, A10:U27, and A26:U11 base pairs (data not shown) further confirmed the exchangeable and nonexchangeable ^{1}H assignments. The ^{13}C resonances of bases, C2, C5, C6, and C8, were assigned based on the ^{1}H-^{13}C HSQC spectrum. The appearance of the C5 resonances of U residues (102–106 ppm) in the lower field region compared to those of C residues (96–100 ppm) was consistent with the general tendency [28]. Finally, the cross peaks between UH3 and AN1, and those between GH1 and CN3 in the HNN-COSY [35, 36] spectrum, (data not shown) confirmed the formation of the A:U and G:C base pairs indicated in Figure 1, the assignments being confirmed as well.

Resonances of Argininamide

Hα, Hβ, Hγ, Hδ, Hϵ, and αNH$_2$ of argininamide in the complex were assigned based on TOCSY spectra obtained in ^{1}H$_2$O and ^{2}H$_2$O. An extra exchangeable proton resonance was found in the ^{13}C-, ^{15}N-filtered NOESY spectrum [37], which was assigned to the amide NH$_2$ on the basis of the NOESY cross peak to Hα. An Hη resonance was not detected. The resonances of the two argininamide molecules are almost identical, as expected from the nearly symmetrical base sequences of the two binding sites.

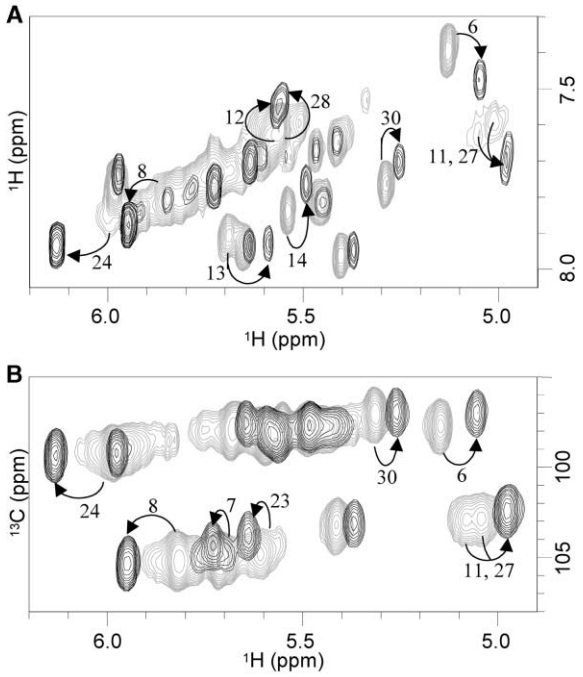

Figure 4. Chemical Shift Perturbations on Binding of Argininamide

H5-H6 correlation peaks in the TOCSY spectrum (A) and H5-C5 correlation peaks in the ^{1}H-^{13}C HSQC spectrum (B) of the free (gray) and argininamide-bound (black) aptamers at 25°C and pH 6.5; the assignments are labeled with the residue numbers.

Confirmation of the Existence of Two Binding Sites in the Aptamer

The existence of two binding sites, boxed in Figure 1C, was predicted from the sequence of the aptamer. Large chemical shift perturbations upon binding of argininamide of either the H5-H6 correlation peaks in the TOCSY spectrum (Figure 4A) or the H5-C5 correlation peaks in the ^{1}H-^{13}C HSQC spectrum (Figure 4B) were observed for the C30, C6, U7, U8, C28, and U27 residues, and also for the C14, C13, U23, C24, C12, and U11 ones. Large chemical shift perturbations of the H8-N7 correlation peaks in the ^{1}H-^{15}N HSQC spectrum were observed for the G9 and A10 residues, and also for the G25 and A26 residues (Figure 5B). The assignments of the aptamer in the free form were made on the basis of the assignments for the complex form by tracing back the correlation peaks during the course of the titration, and are utilized in these analyses. All of the perturbation data confirmed the existence of two binding sites. Hereafter, the lower and upper boxes in Figure 1C are named binding sites 1 and 2, respectively.

The normalized changes of the chemical shifts increased almost linearly up to the aptamer:argininamide molar ratio of 1:1, which indicates that almost all argininamide added bound to the aptamer up to the 1:1 ratio under the high aptamer concentration employed for NMR experiments. At the 1:1 ratio, the normalized changes of the chemical shifts were ca. 0.5 for residues of both sites 1 and 2 (Figure 2). Two explanations are possible for this result. The first explanation is that the binding of two argininamide molecules is completely independent and that two sites exhibit the similar affinity to argininamide. In this case, at the 1:1 ratio, four different states exist with nearly equal population: only site 1 of the aptamer occupied with argininamide, only site 2 occupied, both sites occupied, and both sites empty. If these states are in fast exchange on the NMR time scale, the experimental result can be explained. The second explanation is that the binding of two argininamide molecules is completely cooperative. In this case, at the 1:1 ratio, two states exist with nearly equal population, both sites occupied and both sites empty. Again, if these states are in fast exchange on the NMR time scale, the experimental result can be explained. The intermediate between the first and the second explanations is also possible. The binding of argininamide to the aptamer was monitored also by measuring a series of circular dichroism spectra during the titration. Hill plot analysis of these spectra did not indicate clear cooperativity on the binding of argininamide (data not shown). Therefore, the first explanation is favorable, although the existence of some cooperativity cannot be excluded completely. In any case, when the excess amount of argininamide is supplied at the 1:4 ratio, both sites are occupied for all aptamers in solution. The structural analysis was applied for this fully occupied form.

Formation of Two Adjacent U:A:U Base Triples in the Aptamer-Argininamide Complex

The proximity of the G29:C6 and G9:C28 base pairs was indicated by the G29H1-G9H1 (boxed in Figure 3B) and other NOESY cross peaks. The proximity of the G22:C13 and G25:C12 base pairs was also indicated by the G22H1-G25H1 (boxed in Figure 3B) and other NOESY cross peaks. These results imply the formation of two base triples, U7:A10:U27 and U23:A26:U11 (Figure 5A), when the base triple formation observed in the TARs of

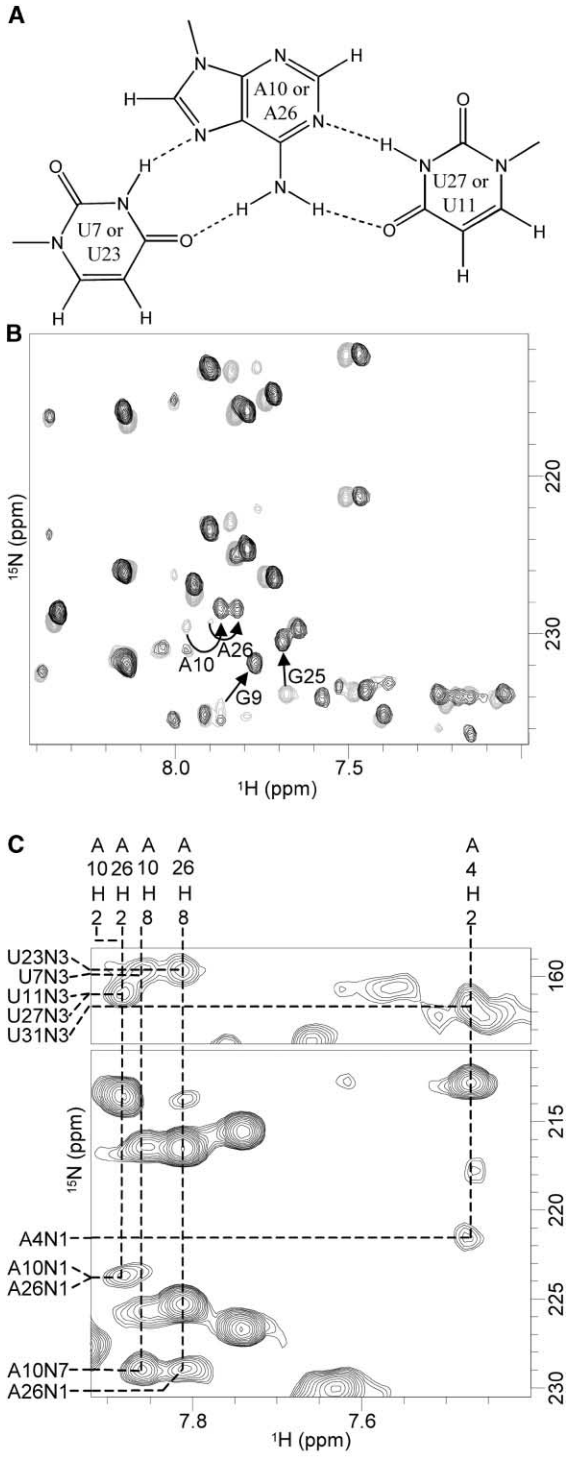

Figure 5. Formation of U:A:U Base Triples

(A) A scheme of the U:A:U base triple.

(B) AH2-AN1, AH2-AN3, and H8-N7 correlation peaks in the ^{1}H-^{15}N HSQC spectrum of the free (light) and argininamide-bound aptamer (dark) at 35°C and pH 6.5. Peaks exhibiting large changes in the ^{15}N chemical shift upon complex formation are labeled with their residue numbers.

(C) The HNN-COSY spectrum of the aptamer complexed with argininamide at 45°C and pH 6.5. The AH2-UN3 and AH2-AN1 correlation peaks for the Watson-Crick A:U base pairs, and the AH8-UN3 and AH8-AN7 correlation peaks for the Hoogsteen A:U base pairs are labeled.

HIV-1, HIV-2, and BIV is recalled [12, 14–16, 19]. When the U7:A10 and U23:A26 base pairs of the Hoogsteen type are formed in addition to Watson-Crick A10:U27 and A26-U11 base pairs, the restriction of rotation around a C6-N6 bond and/or the suppression of the exchange of amino protons with solvent water are expected for the amino moieties of A10 and A26. Thus, clear observation of amino proton-nitrogen correlation peaks in the ^{1}H-^{15}N HSQC spectrum for A10 and A26 residues, correlation peaks for the other A residues not being observed, is consistent with the formation of two base triples in the complex.

An upfield shift of the N7 resonance upon complex formation was observed for A10 and A26 (Figure 5B). The chemical shift perturbations of the N7 resonances of these A residues were larger than those of most other residues. This is consistent with the formation of the two U:A:U base triples, in which the N7 atoms of the two A residues are acceptors of the hydrogen bonds.

Direct evidence of the formation of the two U:A:U base triples was obtained by the detection of spin-spin coupling across the hydrogen bond, $^{2h}J_{NN}$. As imino proton resonances of either U7 or U23 were not observed, probably due to exchange with solvent water, magnetization originating from a nonexchangeable proton, AH8, was used to detect $^{2h}J_{NN}$ in an HNN-COSY spectrum [38, 39]. The A10H8-U7N3 and A26H8-U23N3 cross peaks in the HNN-COSY spectrum (Figure 5C) clearly confirm the formation of the Hoogsteen A10:U7 and A26:U23 base pairs, respectively. The assignments of UN3 were made based on intraresidue UH5-UN3 cross peaks in the ^{1}H-^{15}N HSQC spectrum (data not shown). The A4H2-U31N3, A10H2-U27N3, and A26H2-U11N3 cross peaks (Figure 5C) are consistent with Watson-Crick A4:U31, A10:U27 and A26:U11 base pairs, respectively. $^{2h}J_{NN}$ was calculated on the basis of the intensity ratio of the diagonal to cross peaks in the HNN-COSY spectrum [35]. The $^{2h}J_{NN}$ between AN1 and UN3 for the three Watson-Crick A:U base pairs was ca. 6 Hz. We have confirmed that the same value was obtained in another kind of HNN-COSY experiment in which the magnetization originating from UH3 was used [35, 36]. The value of ca. 6 Hz is close to those reported previously for the Watson-Crick A:U base pairs in other RNA sequences, 7-8 Hz [35, 40]. The $^{2h}J_{NN}$ between AN7 and UN3 for the two Hoogsteen A:U base pairs was ca. 4-5 Hz, which seems to be smaller than those reported previously for the Hoogsteen A:T base pairs of a DNA sequence, 6–8 Hz [41]. This may reflect that the Hoogsteen A:U base pairs in the two base triples dynamically open up for a certain period.

The formation of the U:A:U base triple for the HIV-1 TAR has been a matter of dispute [12–14]. The HNN-COSY result clearly demonstrates the formation of the two U:A:U base triples at least for the aptamer complexed with argininamide molecules.

The Mode of Binding of an Argininamide Molecule to a G Base

The chemical shift perturbations upon complex formation of the N7 resonances of the G9, A10, G25, and A26 residues were larger than those of other residues (Figure 5B). The large upfield shifts of the N7 resonances of the A10 and A26 residues were attributed to these N7 atoms

being acceptors of the hydrogen bonds of the Hoogsteen A:U base pairs in the complex. Similarly, the large upfield shifts of the N7 resonances of the G9 and G25 residues suggest that these N7 atoms are acceptors of the hydrogen bonds with argininamide molecules, as illustrated in Figure 6A. The NOESY cross peaks between Hβ/Hγ/Hδ of argininamide and the amino protons of the C12 and C28 residues (data not shown), and those between Hβ/Hγ/Hδ and the imino protons of the G9 and G25 residues (Figure 6B) are consistent with the formation of such hydrogen bonds. When either the G9 or G25 residue was replaced by a 7-deaza analog, the affinity of the aptamer to an arginine-rich partial peptide of the Tat protein decreased dramatically [23], which constitutes another line of support of the formation of the hydrogen bonds illustrated in Figure 6A.

The type of hydrogen bonds shown in Figure 6A is widely observed when a G:C base pair is recognized by an arginine residue [42, 43]. A similar type of hydrogen bonds is observed in HIV-2 TAR complexed with argininamide [19], and BIV TAR complexed with an arginine-rich peptide of BIV Tat [15], where the Hηs of an arginine residue are involved in the hydrogen bonds, whereas Hε is not. An alternative type of hydrogen bonds is GO6-Hε and GN7- Hη, as observed in BIV TAR complexed with an arginine-rich peptide of BIV Tat [16]. In this type, Hε is close to the amino protons of a C residue base paired to a G residue. The fact that an Hε-CNH_2 NOE was not observed in our case favors the type of hydrogen bonds shown in Figure 6A, in which Hε is far from the amino protons.

Intermolecular NOEs between the Aptamer and Argininamide

Intermolecular NOESY cross peaks between the imino protons of the aptamer and Hβ/Hγ/Hδ of argininamide were observed for the G9, U27, and G29 residues of site 1 (Figure 6B). Similar cross peaks were observed for G25, U11, and G22 of site 2 (Figure 6B). Intermolecular NOESY cross peaks between amino protons and Hβ/Hγ/Hδ were identified for the A10, C28, and C6 residues of site 1 in the ^{15}N-edited NOESY spectrum with the aid of separation in the ^{15}N dimension (data not shown). Similar cross peaks were identified for A26, C12, and C13 of site 2 (data not shown). Intermolecular NOESY cross peaks regarding Hβ/Hγ/Hδ were identified additionally for H6/H8 of the C6 and U7 residues of site 1 (Figure 6C), for H5 of the C6, U7, U27, and C28 residues of site 1 (Figure 6D), and for H2′ and H3′ of the C6 residue of site 1 (data not shown), with the aid of separation in the ^{13}C dimension in the ^{13}C-edited NOESY spectrum in some cases. Similar cross peaks were identified for H6/H8 of G22 and U23 of site 2 (Figure 6C), for H5 of U23, U11, and C12 of site 2 (Figure 6D), and for H2′ and H3′ of G22 of site 2 (data not shown). Intermolecular NOESY cross peaks with respect to α and amide NH_2 of the argininamide were observed for the imino protons of U27 of site 1 and U11 of site 2.

As mentioned above, intermolecular NOESY cross peaks were observed for the C6:G29 and C13:G22 moieties. It is impossible for one argininamide molecule to be close to the two separate moieties. It is reasonable to assume that two argininamide molecules bind to the aptamer, one molecule to binding site 1 and the other molecule to binding site 2. This is consistent with the

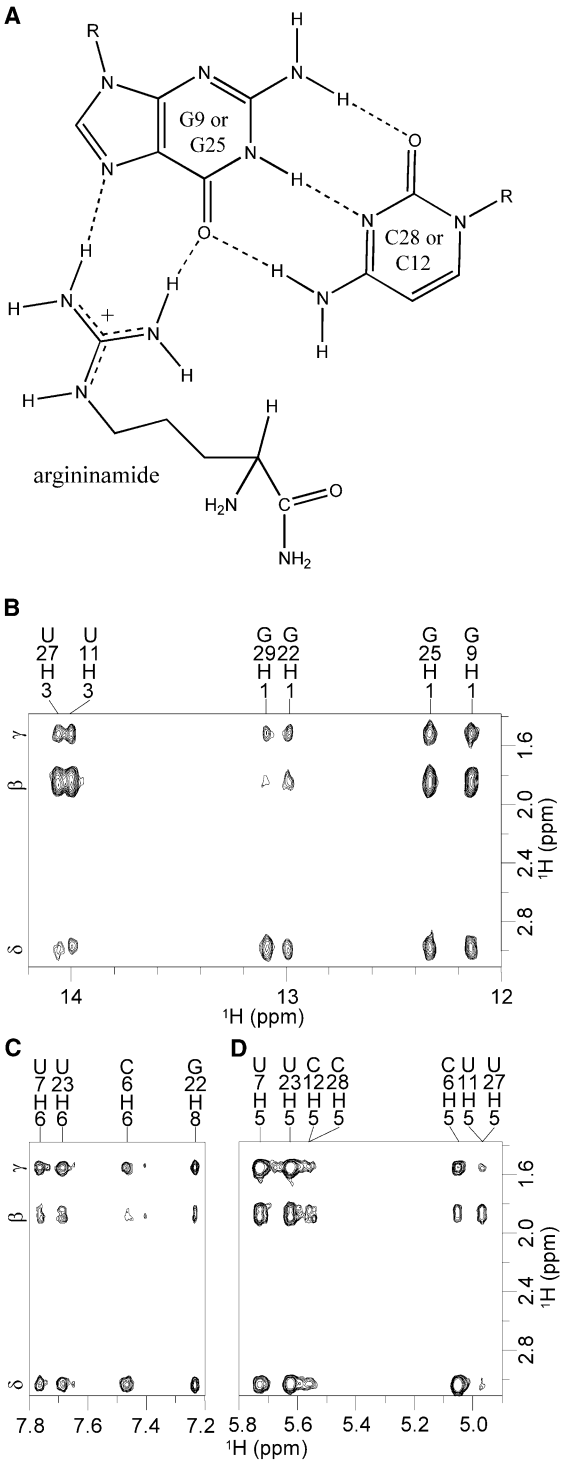

Figure 6. Intermolecular Contacts

A scheme of the binding of an argininamide molecule to the G base of a G:C base pair (A). The intermolecular NOESY cross peaks between Hβ/Hγ/Hδ of argininamide and imino protons of the aptamer in H_2O at 1°C (B), and those between Hβ/Hγ/Hδ and H6/H8 (C) and H5 (D) of the aptamer in ^2H_2O at 35°C.

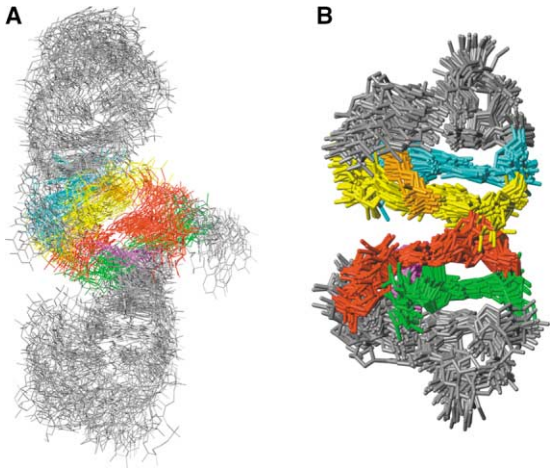

Figure 7. Superposition of the Twenty Final Structures of the Aptamer Complexed with Two Argininamide Molecules

Superposition with respect to all heavy atoms (A) and heavy atoms of the Tat binding sites (B). G9:C28, green; U7:A10:U27, red; U23:A26:U11, yellow; G25:C12, cyan; argininamide 1, magenta; argininamide 2, orange; and others, gray. In (B), U8 and C24 are not shown for clarity.

stoichiometry observed in the titration experiment by NMR. Hereafter, the argininamide molecules bound to sites 1 and 2 are named argininamide 1 and argininamide 2, respectively.

Mostly, only one set of resonances was observed for the two argininamide molecules. This indicated that the resonances of argininamide 1 and argininamide 2 are almost completely degenerated, as expected from the nearly symmetrical base sequences of the two binding sites. The degeneration requires the following treatment on interpretation of the intermolecular NOESY cross peaks. The intermolecular NOESY cross peaks due to the residues of binding sites 1 and 2 were regarded as those for argininamide 1 and argininamide 2, respectively. The structural calculation was carried out on the basis of this interpretation.

Structural Determination of the Aptamer-Argininamide Complex

As the formation of three Watson-Crick A:U, ten Watson-Crick G:C, one sheared G:A, and two Hoogsteen A:U base pairs was confirmed experimentally, corresponding distance constraints on the hydrogen bonds were included in the structural calculation. Distance constraints on hydrogen bonds between argininamide 1 and G9, and between argininamide 2 and G25 were also included on the basis of the experimental implications mentioned above. Additionally, rather moderate planarity constraints were also imposed. Structural calculations gave 20 final structures, which are superimposed in Figures 7A and 7B for the whole region and for the two Tat binding sites, respectively. The root-mean-square-deviations (rmsd) of the 20 final structures versus the mean structure for heavy atoms were 2.50 $\pm$ 0.76 Å for the whole region and 1.04 $\pm$ 0.25 Å for the two Tat binding sites with two argininamide molecules, excluding flipped-out and the least defined U8 and C24 residues. The structural statistics of the 20 final structures are presented in Table 1.

Structure of the Aptamer Complexed with Two Argininamide Molecules

Figures 8A and 8B show stereo views of a representative structure of the aptamer-argininamide complex with the lowest energy among the 20 final structures. The aptamer contains three stems, G1-C6:G29-C34, G9-C12:G25-C28, and C13-G15:C20-G22 (Figure 1C). The three stems are stacked on each other, which yields an almost continuous double helix (Figure 8A). As a whole, the structures of the two Tat binding sites are essentially symmetric, as expected from the base sequence. U7 and U23 of the bulges are in the major groove and form the two U:A:U base pairs adjacent to each other at the center of the aptamer, whereas U8 and C24 are flipped-out into the minor groove. The RNA duplex usually takes on an A-form structure and its major groove is narrow. The formation of the two base triples in the complex serves to widen the major groove to create space for accommodation of two argininamide molecules. The width of the major groove of the Tat binding sites of the aptamer, as defined by the nearest phosphorous-phosphorous distance, is larger than that of the A form by ca. 3–6 Å. In contrast, the width of the minor groove is smaller than that of the A form by ca. 2 Å.

Argininamide 1 and argininamide 2 form hydrogen bonds with the major groove edges of the G9 and G25 bases, respectively. The guanidinium group of argininamide 1 is stacked on the C6 base, lying where G9 would be positioned if it continued the A-form helix through coaxial stacking on C6 (Figure 8C). G9 is pushed aside due to over-twisting, while on the opposite strand, C28-G29 stacking is achieved, as it would in the A-form helix. Similarly, the guanidinium group of argininamide 2 is stacked on the G22 base, lying where G25 would be positioned (Figure 8D). G25 is pushed aside due to over-twisting, while on the opposite strand, C12-C13 stacking is achieved. A U7 base forming a Hoogsteen base pair in the base triple is stacked on the guanidinium group of argininamide 1 (Figure 8C). Thus, the guanidinium group of argininamide 1 is sandwiched between two bases, C6 and U7. In the same way, the guanidinium group of argininamide 2 is sandwiched between the G22 and U23 bases (Figure 8D). Similar features were observed for HIV-2 TAR complexed with an argininamide molecule [19]. The stacking interactions stabilize the binding of the two argininamide molecules.

Hη of argininamide 1 is located close to the oxygen atoms of the phosphate group between C6 and U7 (ca. 2.7–3.1 Å). Similarly, Hη of argininamide 2 is located close to the oxygen atoms of the phosphate group between G22 and U23 (ca. 2.4–2.9 Å). These results suggest the stabilization of the complex through the hydrogen bonds between argininamide molecules and these phosphate groups. The formation of the hydrogen bond with the phosphate group is reported for the BIV system [16].

Stoichiometry of the Aptamer-RG Peptide Complex

The formation of a complex between the aptamer and the RG peptide was monitored by means of H5-C5 corre-

Table 1. NMR Constraints and Structural Statistics for the Aptamer-Argininamide Complex

(A) NMR Constraints	
Distance constraints	440
Intraresidue	124
Sequential (i, i + 1)	132
Medium to long range ≥ (i, i + 2)	86
Intermolecular	98
Dihedral angle constraints	219
β	6
γ	4
δ	32
ϵ	17
ν_0–ν_4	160
Hydrogen bonding constraints	92
Planarity constraints	14
(B) Structural Statistics for Twenty Final Structures	
X-PLOR energies (kcal/mol)	
E_{total}	314 ± 5
E_{bond}	26 ± 1
E_{angle}	198 ± 2
$E_{improper}$	21 ± 1
E_{vdw}	15 ± 2
E_{noe}	22 ± 3
E_{cdih}	3 ± 1
Rmsd from idealized geometry	
Bond lengths (Å)	0.005 ± 0.00005
Bond angles (°)	0.77 ± 0.003
Impropers (°)	0.43 ± 0.02
NOE violations	
Number of violations greater than 0.3 Å	0 ± 0
Rmsd of violations (Å)	0.03 ± 0.002
Dihedral angle violations	
Number of violations greater than 4°	0 ± 0
Rmsd of violations (°)	0.38 ± 0.05
Rmsd of 20 final structures versus the mean structure for heavy atoms (Å)	
Whole region	2.50 ± 0.76
Two Tat binding sites with two argininamide molecules, excluding U8 and C24	1.04 ± 0.25

lation peaks of the aptamer in a series of ^{1}H-^{13}C HSQC spectra during titration. The free and bound forms of the aptamer were found to be in slow exchange on the NMR time scale. Fast exchange was observed in the case of the binding of argininamide. This difference is expected from the very strong binding of the RG peptide to the aptamer (dissociation constant of 10^{-9} M). The stoichiometry of the aptamer-RG peptide complex was determined by tracing the changes in the intensities of cross peaks during titration. The decrease in the intensities of several cross peaks representing the free form of the aptamer is plotted as a function of the molar ratio of the RG peptide to the aptamer in Figure 9A. The plot indicates that one RG peptide binds per aptamer. Similar stoichiometry was observed on plotting the increase in the intensities of the cross peaks representing the bound form of the aptamer (data not shown). This is consistent with the biochemical finding that another peptide, closely related to the RG peptide, binds to the aptamer in a one to one stoichiometry [23]. Frankel and coworkers designed RNAs in which the two Tat binding sites of BIV TAR are arranged in a symmetrical manner, as in our aptamer. They found the binding of two peptides of BIV Tat to some of these RNAs [44]. Thus, the difference in the stoichiometry between our aptamer and their RNAs is noteworthy.

The binding of one RG peptide per aptamer is in contrast to the binding of two argininamide molecules per aptamer. This difference will be discussed below.

Similarity between the RG Peptide-Bound and Argininamide-Bound Forms in the Aptamer Structure and Intermolecular Contacts

Figure 9B shows the chemical shift perturbations of the H5-C5 correlation peaks of the aptamer in the ^{1}H-^{13}C HSQC spectrum upon binding of the RG peptide. The pattern of the perturbations is very similar to that observed upon the binding of argininamide (Figure 4B). Similarity of the pattern of the perturbations was observed for other correlation peaks (data not shown). These results indicate that the structural changes and thus the resultant structure of the aptamer are quite similar for RG peptide-bound and argininamide-bound forms. The affinity between the aptamer and the RG peptide is almost comparable to that between the aptamer and Tat, the difference in dissociation constant being only one order. Therefore, it is expected that the structure of the aptamer for the Tat-bound form is also similar to that for the argininamide-bound form determined in this study.

It is notable that the aptamer-RG peptide complex gives NOESY cross peaks that are quite comparable to the intermolecular NOESY cross peaks observed for the aptamer-argininamide complex (Figures 9C–9E). The as-

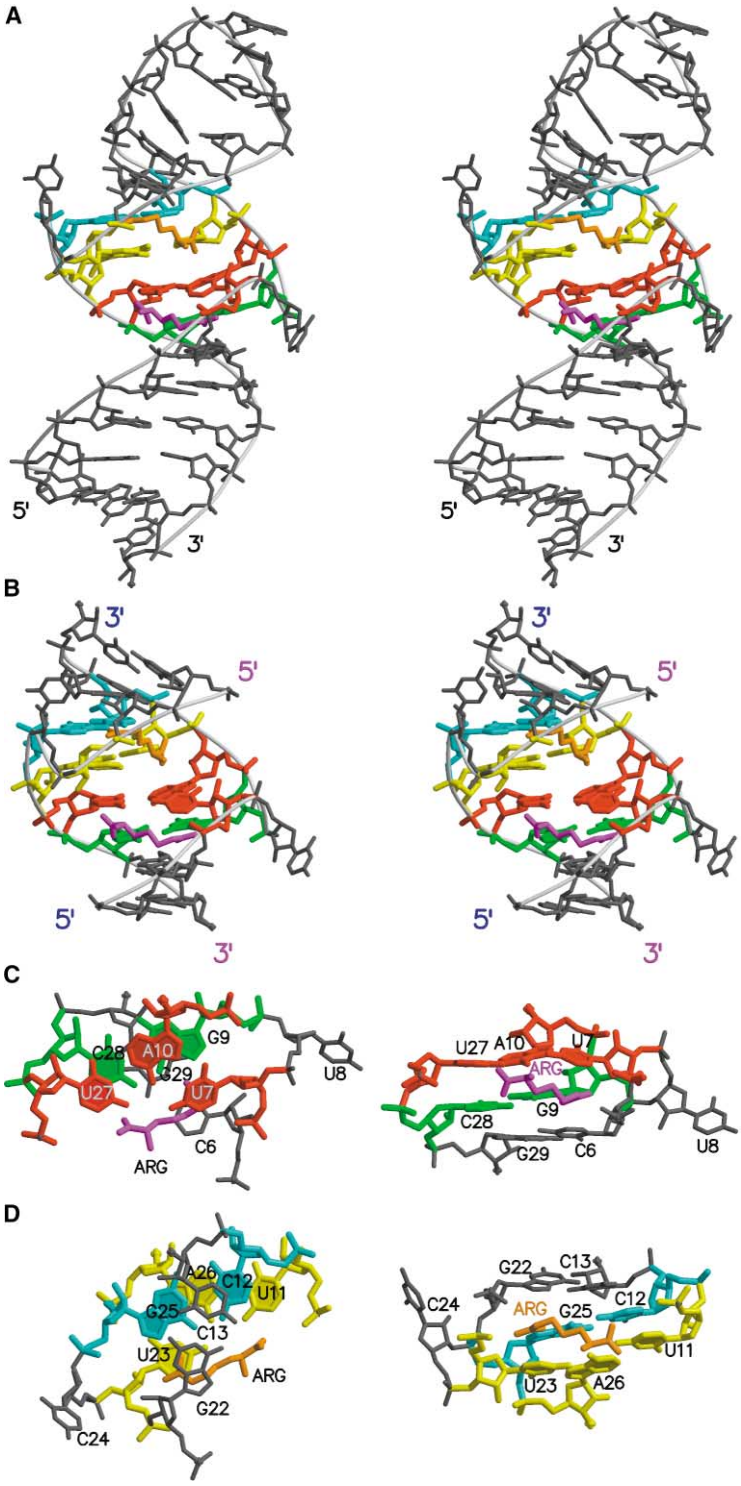

Figure 8. Views of a Representative Structure of the Aptamer Complexed with Two Argininamide Molecules

Stereo side views of whole region (A) and the Tat binding sites (B). The same coloring as in Figure 7 is applied. Phosphorus atoms are traced by thin gray lines. Top (left) and side (right) views of U7:A10:U27, G9:C28, U8, C6:G29, and argininamide 1 (C), and those of G22:C13, C24, G25:C12, U23:A26:U11, and argininamide 2 (D).

signments of the cross peaks for the aptamer-peptide complex were tentatively made by comparing the patterns of the cross peaks between the two complexes. The assignments are consistent with both partially accomplished sequential assignments for the aptamer-RG peptide complex and the results of analysis of the chemical shift perturbations for the aptamer-peptide complex, as shown in Figure 9B, for example. The chemical shifts of the cross peaks in Figures 9C–9E suggest strongly that the amino acid residues that give these cross peaks are arginine ones.

For the aptamer-peptide complex, intermolecular contacts with the arginine residues were observed for two separate moieties of the aptamer, the C6:G29 and G22:C13 base pairs, as in the case of the aptamer-argininamide complex. It is impossible that a single argi-

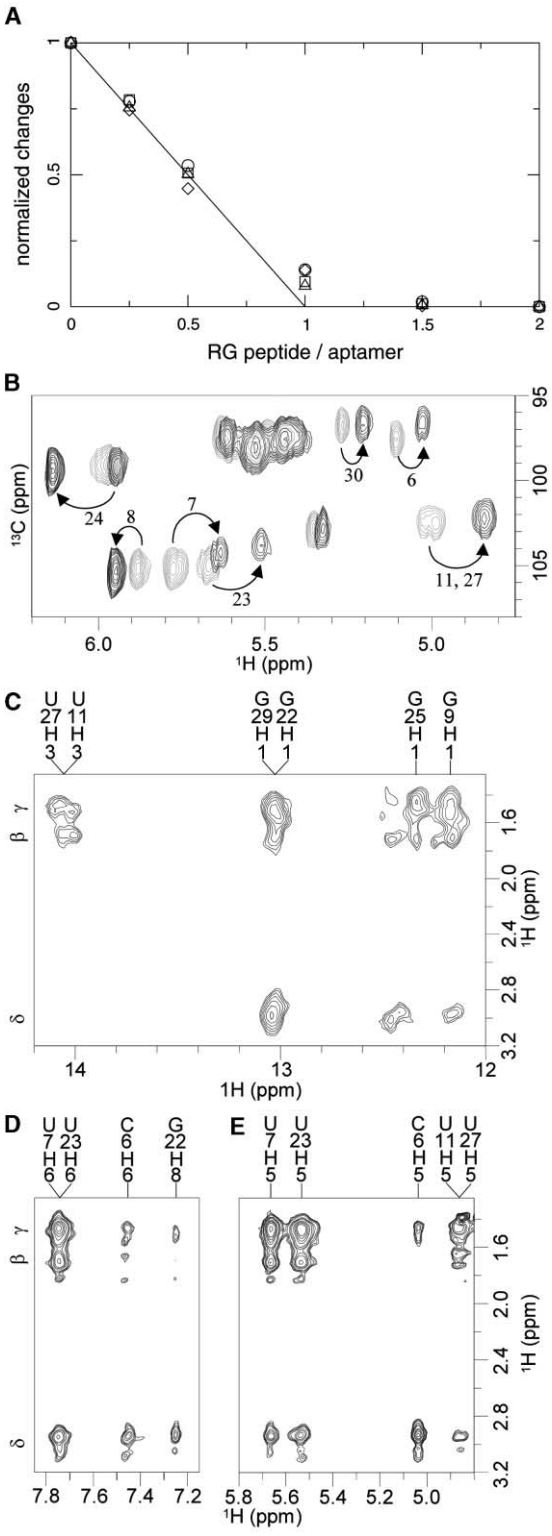

Figure 9. Interactions of the Aptamer with the RG Peptide

(A) Normalized changes of the intensities of the H5-H6 correlation peaks of the aptamer in the free form as a function of the molar ratio of the RG peptide to the aptamer at 35°C and pH 6.5. Circles, C6; squares, U11; diamonds, U27; and triangles, C30.

(B) H5-C5 correlation peaks in the ^{1}H-^{13}C HSQC spectra of the free (gray) and RG peptide-bound (black) aptamers in H_2O at 35°C and pH 6.5; the assignments are labeled with the residue numbers. The intermolecular NOESY cross peaks between Hβ/Hγ/Hδ of arginine residues of the RG peptide and imino protons of the aptamer in H_2O at 1°C (C), and those between Hβ/Hγ/Hδ and H6/H8 (D) and H5 (E) of the aptamer in $^{2}H_2O$ at 35°C are shown.

nine residue of the peptide interacts with these two separate moieties. It is thus supposed that two different arginine residues of the peptide interact with these moieties, respectively. The similarity in the patterns of the intermolecular contacts between the two complexes implies that the two arginine residues of the peptide interact with the aptamer in the same way as two argininamide molecules interact with the aptamer. This explains the experimental finding that one peptide molecule binds per aptamer, while two argininamide molecules bind per aptamer.

The Origin of the High Affinity of the Aptamer

The combination of structural studies on the aptamer-argininamide and aptamer-peptide complexes suggests the origin of the high affinity of the aptamer to Tat. The aptamer binds Tat 100 times more strongly than TAR. Simultaneous interactions of the aptamer at two sites, sites 1 and 2, with two different arginine residues of the arginine-rich region of Tat are supposed to be the primary reason for the high affinity. Double interactions could enhance the affinity by as much as the square of the binding constants, if each interaction is perfectly achieved, respectively [44]. Practically, it has been reported that by joining DNA binding domains covalently, the affinity to the multiple binding sites increases by two or three orders of magnitude [45, 46]. The 100-fold enhancement of the affinity observed for our aptamer is reasonable in this sense.

The ability of the aptamer to form two adjacent U:A:U base triples with the Hoogsteen base pairs in the central portion is crucial for achieving the simultaneous interactions, because the widening of the major groove on the formation of the two adjacent base triples is indispensable for creating the space to accommodate two bulky arginine residues and probably also several residues linking them. The major groove of the aptamer-argininamide complex is wider than that of the A form. The major groove may be able to become even wider when the aptamer needs to accommodate the two arginine and joining residues of Tat. The two adjacent U:A:U base triples at the center may serve to confer such deformability on the aptamer.

Hoogsteen base-paired U residues of the base triples contribute to stabilization of the complex through stacking interactions with the guanidinium groups of the arginine residues at the two sites, respectively. The importance of the formation of the Hoogsteen base pairs is consistent with the biochemical data revealing that the replacement of A10 and A26 by 7-deaza analogs, which cannot accept the Hoogsteen hydrogen donor, abolish the binding of the Tat peptide [23].

Future Prospects

When the symmetric arrangement of the two Tat binding sites of the aptamer is considered, it is assumed that

two arginine residues of a single peptide bind to the binding sites in a symmetric way. To achieve this, it may be required that the two arginine residues are separated by at least several residues along the primary peptide sequence and that the peptide forms a turn structure for these arginine residues to be arranged symmetrically.

The complex formation of the aptamer with a pentapeptide, RKKRR, was monitored by NMR. It turned out that the free and bound forms of the aptamer were in slow exchange on the NMR time scale and that one pentapeptide bound per aptamer (data not shown), as exactly observed for the binding of the RG peptide. The pentapeptide corresponds to the first 5 residues of the RG peptide and also corresponds to the residues 49–53 of Tat. The result suggests that either R49 and R52 or R49 and R53 may be the residues making simultaneous double interactions with the aptamer. It is notable that R52 and R53 of Tat are known to be critical for binding to TAR [18, 47, 48]. Structural determination of the aptamer-peptide complexes is currently being challenged in our laboratory.

As discussed previously, the RNA structure of the RG peptide-bound form and presumably that of the Tat-bound form also are close to the RNA structure of the argininamide-bound form determined in this study. Therefore it may be possible to design new RNAs that inhibit Tat or to develop RNA biosensors for Tat by mutating the current RNA aptamer in light of its elucidated structure, without having the detailed structural information on Tat.

Biological Implications

The Tat protein of HIV plays a critical role in viral replication. The development of effective inhibitors of Tat could lead to gene therapy. We isolated an RNA aptamer that binds to Tat 100 times more strongly than the authentic TAR RNA. The aptamer functions as a decoy molecule for Tat and thus inhibits the action of Tat both in vitro and in vivo. Structural basis of the highly efficient trapping of Tat by the aptamer was elucidated by this study.

The aptamer exhibits unique structure when it binds the analog of Tat. The formation of two adjacent U:A:U base triples was observed for the first time at the central portion of the aptamer, which widens the major groove to make space for the accommodation of the analog. The aptamer makes the interactions with two arginine residues of the analog at both sides of the central two base triples. The simultaneous double interactions of the aptamer could explain its extremely high affinity to Tat.

The elucidation of key elements of the aptamer to exhibit high affinity at atomic resolution greatly enhances the rational design of more effective inhibitors of Tat. Moreover, the strategy to have simultaneous double interactions with the target protein through two symmetrically arranged binding sites may be applicable to some other RNA-protein systems utilizing the arginine-rich motif in order to accomplish high affinity.

Experimental Procedures

Sample Preparation

Nonlabeled and ^{13}C- and ^{15}N-labeled RNA aptamers (34-mer) were synthesized in vitro by transcription, using T7 RNA polymerase, as described previously [49]. The lyophilized aptamer was dissolved in 10 mM sodium phosphate buffer (pH 6.7) containing 30 mM NaCl, 0.1 mM EDTA, and 3 mM NaN_3. DSS was used as an internal chemical shift reference. The concentration of the aptamer was 1.7 mM.

The aptamer-argininamide complex was prepared by the addition of a concentrated argininamide (Sigma) solution to the RNA aptamer solution step by step. The aptamer-peptide complexes were prepared in the same way.

NMR Spectroscopy

NMR spectra were recorded with a Bruker DRX600 spectrometer equipped with a quadruple-resonance probe with X, Y, and Z gradients. The following NMR spectra were recorded for the aptamer-argininamide complex (aptamer:argininamide = 1:4): NOESY, ^{13}C-, ^{15}N-filtered NOESY, TOCSY, ^{31}P-decoupled DQF-COSY, ^{1}H-^{31}P HetCor with suppression of ^{1}H-^{1}H passive coupling, spin echo difference constant time ^{1}H-^{13}C HSQC [50], ^{13}C-edited NOESY-HSQC [51], ^{15}N-edited NOESY-HSQC [52], HCCH-COSY, HCCH-TOCSY, and HNN-COSY spectra. Spectra were processed with Capp/Pipp/Stapp [53]. A similar set of spectra was obtained for the aptamer-peptide complex (aptamer:peptide = 1:1.5).

Distance Constraints

Interproton distances were calculated with the NOESY spectrum with a mixing time of 50 ms, as described previously [54]. Four hundred and forty distance constraints (124 intraresidue, 132 sequential, 86 medium- to long-range, and 98 intermolecular ones) were obtained for the aptamer-argininamide complex.

On the basis of experimental confirmation (see Results and Discussion), hydrogen bond constraints were included for three Watson-Crick A:U base pairs, ten Watson-Crick G:C base pairs, one sheared G:A base pair, and two Hoogsteen A:U base pairs. Hydrogen bond constraints between the aptamer and argininamide were also included.

Dihedral Angle Constraints

Dihedral angle constraints for the β torsion angle were derived from $^{3}J_{H5'\text{-}P}$ and $^{3}J_{H5''\text{-}P}$ couplings [28, 55]. The βs of the G3, A4, G5, G15, A19, and G21 residues were constrained to $180^\circ \pm 20^\circ$. For the other residues, β was left unconstrained.

Dihedral angle constraints for the γ torsion angles were derived from $^{3}J_{H4'\text{-}H5'}$ and $^{3}J_{H4'\text{-}H5''}$ couplings [28, 55]. The γs of the G2, C13, G16, and A17 residues were constrained to $60^\circ \pm 20^\circ$. For the other residues, γ was left unconstrained.

Dihedral angle constraints for the δ, and endocyclic ν_0, ν_1, ν_2, ν_3, and ν_4 torsion angles were derived from $^{3}J_{H1'\text{-}H2'}$ and $^{3}J_{H3'\text{-}H4'}$ couplings [28, 54, 55]. The δ and endocyclic ν_0-ν_4 torsion angles of all the residues, except for G1, G3, U7, U8, U23, and C24, were moderately constrained, leaving the sugar free to take on any conformation without an energy penalty between C2′-*exo* and C4′-*exo* including C3′-*endo* in a pseudorotation cycle. In the same way, those of the U7, U8, U23, and C24 residues were constrained between O4′-*endo* to C3′-*exo* including C1′-*exo* and C2′-*endo*. Those of the G1 and G3 residues were left unconstrained.

Dihedral angle constraints for the ϵ torsion angle were derived from $^{3}J_{C2'\text{-}P}$ and $^{3}J_{H3'\text{-}P}$ couplings [28, 50, 55]. The ϵs of G9, G16, A18, and A19 were constrained to $-145^\circ \pm 20^\circ$, those of A4, G5, C14, A17, C20, G21, G25, G29, C30, and U31 to $-170^\circ \pm 50^\circ$, and those of U7, U8 and U23 to $-100^\circ \pm 25^\circ$. For the other residues, ϵ was left unconstrained.

No dihedral angle constraints were used for the α and ζ torsion angles.

Structure Calculation

Structure calculations were carried out using distance and dihedral angle constraints using a simulated annealing protocol supplied with X-PLOR v. 3.8 [56] on an O_2 workstation (Silicon Graphics). Planarity constraints were imposed for the G9:C28 base pair and the guanidinium group of argininamide 1, and also for the G25:C12 base pair and the guanidinium group of argininamide 2, with the moderate weight of 4 kcal mol^{-1} $Å^{-2}$. For the other base pairs, planarity constraints with the rather moderate weight of 0.4 kcal mol^{-1} $Å^{-2}$ were imposed, including for the Hoogsteen A:U base

pairs. Twenty final structures were selected from 200 calculations on the basis of the criteria of the smallest residual energy values. None of them violated the distance constraints by more than 0.3 Å or the dihedral angle constraints by more than 4°. The structures were viewed with Insight II (MSI).

Acknowledgments

M.K. was supported by Grants-in-Aid for Scientific Research (no. 10179102, 12470487, and 14035219) and the Protein 3000 Project from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. R.Y., K.T., S.N., and P.K.R.K. were supported by a grant of R&D for Evolutionary Molecular Engineering (NEDO). We also wish to thank Dr. Y. Tanaka and Mr. T. Hori for technical assistance.

Received: December 10, 2002
Revised: February 28, 2003
Accepted: March 3, 2003
Published: May 6, 2003

References

1. Fisher, A.G., Feinberg, M.B., Josephs, S.F., Harper, M.E., Marselle, L.M., Reyes, G., Gonda, M.A., Aldovini, A., Debouk, C., Gallo, R.C., et al. (1986). The trans-activator gene of HTLV-III is essential for virus replication. Nature *320*, 367–371.
2. Frankel, A.D. (1992). Activation of HIV transcription by Tat. Curr. Opin. Genet. Dev. *2*, 293–298.
3. Jones, K.A., and Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. Annu. Rev. Biochem. *63*, 717–743.
4. Muesing, M.A., Smith, D.H., and Capon, D.J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell *48*, 691–701.
5. Cordingly, M.G., LeFemina, R.L., Callahan, P.L., Condra, J.H., Sardana, V.V., Graham, D.J., Nguyen, T.M., LeGrow, K., Gotlib, L., Schlabach, A.J., et al. (1990). Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive element of human immunodeficiency virus type 1 in vitro. Proc. Natl. Acad. Sci. USA *87*, 8985–8989.
6. Roy, S., Delling, U., Chen, C.H., Rosen, C.A., and Sonenberg, N. (1990). A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated *trans*-activation. Genes Dev. *4*, 1365–1373.
7. Weeks, K.M., and Crothers, D.M. (1991). RNA recognition by Tat-derived peptides: interaction in the major groove. Cell *66*, 577–588.
8. Churcher, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Butler, P.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.
9. Karn, J. (1999). Tackling Tat. J. Mol. Biol. *293*, 235–254.
10. Weeks, K.M., Ampe, C., Schultz, S.C., Steitz, T.A., and Crothers, D.M. (1990). Fragments of the HIV-1 Tat protein specifically bind TAR RNA. Science *249*, 1281–1285.
11. Long, K.S., and Crothers, D.M. (1995). Interaction of human immunodeficiency virus type 1 Tat-derived peptides with TAR RNA. Biochemistry *34*, 8885–8895.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. Tao, J., and Frankel, A.D. (1992). Specific binding on arginine to TAR RNA. Proc. Natl. Acad. Sci. USA *89*, 2723–2726.
18. Tao, J., and Frankel, A.D. (1993). Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins. Proc. Natl. Acad. Sci. USA *90*, 1571–1575.
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. Ippolito, J.A., and Steitz, T.A. (1998). A 1.3-Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. Proc. Natl. Acad. Sci. USA *95*, 9819–9824.
21. Yamamoto, R., Murakami, K., Taira, K., and Kumar, P.K.R. (1998). Isolation and characterization of an RNA that binds with high affinity to Tat protein of HIV-1 from a completely random pool of RNA. Gene Ther. Mol. Biol. *1*, 451–466.
22. Heus, H.A., and Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. Science *253*, 191–194.
23. Yamamoto, R., Katahira, M., Nishikawa, S., Baba, T., Taira, K., and Kumar, P.K.R. (2000). A novel RNA motif that binds efficiently and specifically to the Tat protein of HIV and inhibits the *trans*-activation by Tat of transcription in vitro and in vivo. Genes Cells *5*, 371–388.
24. Davis, D.G., and Bax, A. (1985). Assignment of complex ^{1}H NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. J. Am. Chem. Soc. *107*, 2820–2821.
25. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. J. Chem. Phys. *71*, 4546–4553.
26. Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids (New York: John Wiley and Sons).
27. Katahira, M., Kanagawa, M., Sato, H., Uesugi, S., and Fuji, S. (1994). Formation of sheared G:A base pairs in an RNA duplex modeled after ribozymes, as revealed by NMR. Nucleic Acids Res. *22*, 2752–2759.
28. Varani, G., Aboul-ela, F., and Allain, F.H.T. (1996). NMR investigation of RNA structure. Prog. in Nucleic Magn. Reson. Spect. *29*, 51–127.
29. Bax, A., Clore, M., Driscoll, C., Gronenborn, A.M., Ikura, M., and Kay, M. (1990). Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of ^{13}C-labeled proteins. J. Magn. Reson. *87*, 620–627.
30. Bax, A., Clore, M., and Gronenborn, A.M. (1990). ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. J. Magn. Reson. *88*, 425–431.
31. Pardi, A., and Nikonowicz, E.P. (1992). Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. J. Am. Chem. Soc. *114*, 9202–9203.
32. Williamson, D., and Bax, A. (1988). Resolution-enhanced correlation of ^{1}H and ^{31}P chemical shifts. J. Magn. Reson. *76*, 174–177.
33. Marion, J.P., Prestegard, J.H., and Crothers, D.M. (1994). Correlation of adenine H2/H8 resonances in uniformly ^{13}C labeled RNA by 2D HCCH-TOCSY: a new tool for ^{1}H assignment. J. Am. Chem. Soc. *116*, 2205–2206.
34. Legault, P.L., Farmer, B.T., Mueller, L., and Pardi, A. (1994). Through-bond correlation of adenine protons in a ^{13}C-labeled ribozyme. J. Am. Chem. Soc. *116*, 2203–2204.
35. Dingley, A.J., and Grzesiek, S. (1998). Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^{2}J_{NN}$ couplings. J. Am. Chem. Soc. *120*, 8293–8297.
36. Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M., and Wüthrich, K. (1998). NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. Proc. Natl. Acad. Sci. USA *95*, 14147–14151.
37. Ikura, M., and Bax, A. (1992). Isotope-filtered 2D NMR of a protein-peptide complex: study of a skeletal muscle myosin

light chain kinase fragment bound to calmodulin. J. Am. Chem. Soc. *114*, 2433–2440.
38. Majumdar, A., Kettani, A., Skripkin, E., and Patel, D.J. (1999). Observation of internucleotide NH N hydrogen bonds in the absence of directly detectable protons. J. Biomol. NMR *15*, 207–211.
39. Hennig, M., and Williamson, J.R. (2000). Detection of N-H N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances. Nucleic Acids Res. *28*, 1585–1593.
40. Wöhnert, J., Dingley, A.J., Stoldt, M., Görlach, M., Grzesiek, S., and Brown, L.R. (1999). Direct identification of NH…N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. Nucleic Acids Res. *27*, 3104–3110.
41. Dingley, A.J., Masse, J.E., Peterson, R.D., Barfield, M., Feigon, J., and Grzesiek, S. (1999). Internucleotide scalar couplings across hydrogen bonds in Watson-Crick and Hoogsteen base pairs of a DNA triplex. J. Am. Chem. Soc. *121*, 6019–6027.
42. Saenger, W. (1983). Principles of Nucleic Acid Structure (New York: Springer-Verlag).
43. Sinden, R.R. (1994). DNA Structure and Function (San Diego: Academic Press).
44. Campisi, D.M., Calabro, V., and Frankel, A.D. (2001). Structure-based design of a dimeric RNA-peptide complex. EMBO J. *20*, 178–186.
45. Robinson, C.R., and Sauer, R.T. (1996). Covalent attachment of Arc repressor subunits by a peptide linker enhances affinity for operator DNA. Biochemistry *35*, 109–116.
46. Kim, J.S., and Pabo, C.O. (1998). Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. Proc. Natl. Acad. Sci. USA *95*, 2812–2817.
47. 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.
48. Calnan, B.J., Biancalana, S., Hudson, D., and Frankel, A.D. (1991). Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. *5*, 201–210.
49. Yamamoto, R., Koseki, S., Ohkawa, J., Murakami, K., Nishikawa, S., Taira, K., and Kumar, P.K.R. (1997). Inhibition of transcription by the TAR RNA of HIV-1 in a nuclear extract of HeLa cells. Nucleic Acids Res. *25*, 3445–3450.
50. Legault, P., Jucker, F.M., and Pardi, A. (1995). Improved measurement of ^{13}C, ^{31}P J coupling constants in isotopically labeled RNA. FEBS Lett. *362*, 156–160.
51. Ikura, M., Kay, L.E., Tschudin, R., and Bax, A. (1990). Three-dimensional NOESY-HMQC spectroscopy of a C-13-labeled protein. J. Magn. Reson. *86*, 204–209.
52. Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A., and Bax, A. (1989). Three-dimensional heteronuclear NMR of ^{15}N-labeled proteins. J. Am. Chem. Soc. *111*, 1515–1517.
53. Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. (1991). A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. J. Magn. Reson. *95*, 214–220.
54. Matsugami, A., Ouhashi, K., Kanagawa, M., Liu, H., Kanagawa, S., Uesugi, S., and Katahira, M. (2001). An intramolecular quadruplex of $(GGA)_4$ triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. J. Mol. Biol. *313*, 255–269.
55. Smith, J.S., and Nikonowicz, E.P. (2000). Phosphorothioate substitution can substantially alter RNA conformation. Biochemistry *39*, 5642–5652.
56. Brünger, A.T. (1993). X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR (New Haven: Yale University Press).

Accession Numbers

The atomic coordinates have been deposited in the Protein Data Bank under ID code 1NBK.