Published online 26 December 2007 *Nucleic Acids Research, 2008, Vol. 36, No. 4* ***1227–1236***
doi:10.1093/nar/gkm1141

DNA mimicry by a high-affinity anti-NF-κB RNA aptamer

Nicholas J. Reiter[1], L. James Maher III[2] and Samuel E. Butcher[1,*]

[1]Department of Biochemistry, University of Wisconsin-Madison and [2]Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN, USA

Received September 20, 2007; Revised November 21, 2007; Accepted December 7, 2007

ABSTRACT

The binding of RNA molecules to proteins or other ligands can require extensive RNA folding to create an induced fit. Understanding the generality of this principle involves comparing structures of RNA before and after complex formation. Here we report the NMR solution structure of a 29-nt RNA aptamer whose crystal structure had previously been determined in complex with its transcription factor target, the $p50_2$ form of NF-κB. The RNA aptamer internal loop structure has pre-organized features that are also found in the complex, including non-canonical base pairing and cross-strand base stacking. Remarkably, the free RNA aptamer structure possesses a major groove that more closely resembles B-form DNA than RNA. Upon protein binding, changes in RNA structure include the kinking of the internal loop and distortion of the terminal tetraloop. Thus, complex formation involves both pre-formed and induced fit binding interactions. The high affinity of the NF-κB transcription factor for this RNA aptamer may largely be due to the structural pre-organization of the RNA that results in its ability to mimic DNA.

INTRODUCTION

An RNA aptamer that tightly binds transcription factor NF-κB

The subject of the present analysis is a small RNA aptamer (1,2) identified by *in vitro* selection (3,4) for affinity to the $p50_2$ form of mammalian transcription factor NF-κB. Because of its role in activating genes involved in inflammation, inhibition of apoptosis and HIV-1 activation, this dimeric transcription factor is of interest as a potential target for therapeutic inhibition (5,6). Previous work showed that the anti-NF-κB RNA aptamer bound with nanomolar affinity to NF-κB $p50_2$ in a manner that competed with DNA binding, that the binding interaction could be detected in the yeast three-hybrid system (7), and that aptamer variants with improved properties could be identified by yeast genetic selections (2). The X-ray crystal structure of a 29-nt version of the RNA was determined in complex with NF-κB $p50_2$ (8). The features of this complex have subsequently been reviewed and compared to the prior structure of NF-κB $p50_2$ bound to DNA (9,10).

The essential features of the RNA aptamer structure derived from its complex with NF-κB $p50_2$ are shown in Figure 1A. The RNA aptamer was predicted to fold as a stem-loop structure with an asymmetric internal loop, and a U–G wobble pair adjacent to a terminal GUAA tetraloop. Examination of the RNA aptamer within the NF-κB-aptamer complex (Figure 1B) validated secondary structure predictions, and revealed structural features deviating from A-form geometry. The latter included a large overall kink in the RNA and a complex pattern of base interactions within the asymmetric internal loop. These interactions included non-canonical pairing of A9–G22 and U6–C24, with stacking of unpaired bases U7, G8 and G23. Together with the U13–G18 wobble pair, these interactions provided a series of hydrogen bonding and van der Waals contacts with the DNA binding surface of NF-κB p50. In the crystal complex with NF-κB $p50_2$, one aptamer binds to each NF-κB subunit, requiring a large opening of the Rel homology domains in the protein dimer relative to the structure bound to DNA (Figure 1C and D). In fact, the aptamer structure mimics the major groove of the normal DNA binding sequence in such a manner that the interacting protein side chain conformations are largely preserved between the two complexes (8,10).

We have been interested in the degree to which the anti-NF-κB RNA aptamer structure is pre-formed before interaction with the target protein. Theoretically, maximum binding affinity can be achieved by the pre-formation of rigid complementary surfaces between two binding partners. This concept is inherent in pharmaceutical design and certain natural high-affinity interactions such as the binding of biotin by streptavidin (11). In contrast, many natural macromolecular binding

*To whom correspondence should be addressed. Tel: 608-263-4081; Fax: 608-262-3453; Email: butcher@biochem.wisc.edu

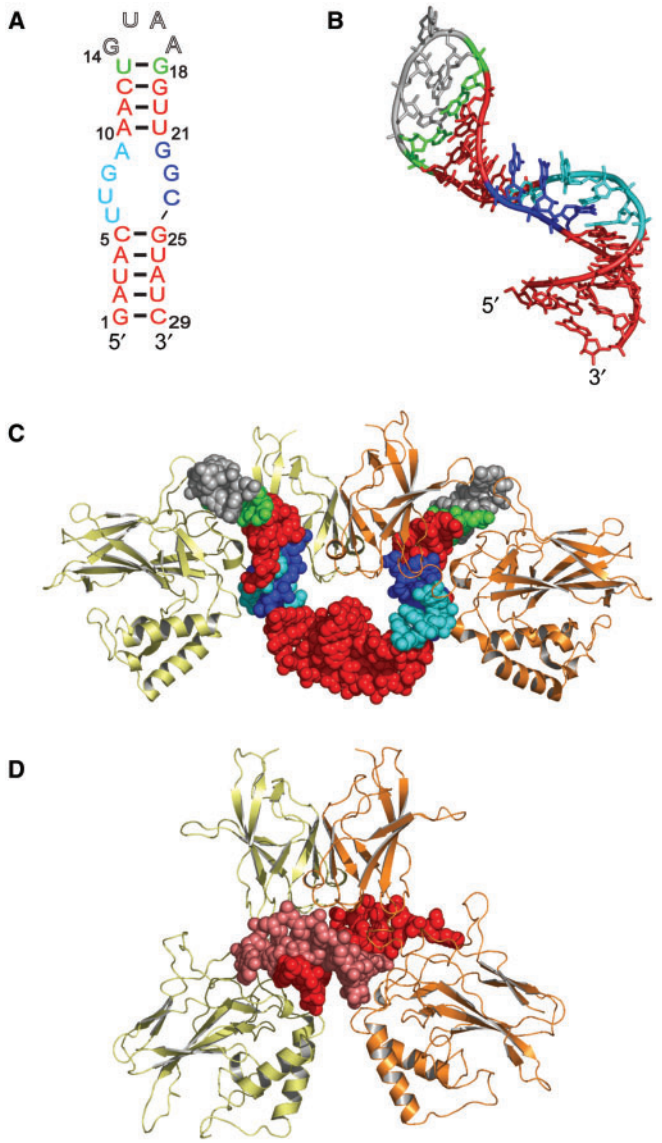

Figure 1. Anti-NF-κB RNA aptamer and crystal structure in complex with NF-κB $p50_2$. (**A**) Proposed secondary structure of the 29-nt anti-NF-κB RNA aptamer in solution. The RNA hairpin studied here is identical in sequence to the RNA aptamer co-crystallized with NF-κB $p50_2$ (with exception of the inversion of terminal nucleotides G1 and C29). Canonical RNA Watson–Crick base pairs are red, the 5′ region of the internal loop is cyan, the wobble pair is green, the GNRA-type tetraloop is gray, and the 3′ region of the internal loop is blue. (**B**) RNA aptamer structure extracted from the crystal structure and colored as in (A). (**C** and **D**) Comparison of the crystal complex of NF-κB $p50_2$ with the RNA aptamer (PDB ID 1OOA) or with bound DNA (PDB ID 1NFK). RNA binding requires substantial opening of the Rel homology domains (8,9). The two p50 subunits are shown in orange and yellow. Nucleic acids are shown as space filling models, where the RNA aptamer structures (C) are colored as in (A) and DNA structure (D) is colored red and pink.

interactions occur with intermediate affinity to accommodate the biological kinetics required for disassembly and regulation. For example, a common feature of protein–nucleic acid interactions is the mutual reorganization of both binding partners upon interaction. It has been argued that important energetic considerations are at play in such systems. In particular, organization of previously unstructured domains creates an unfavorable entropic contribution that reduces what could otherwise be unacceptably strong binding interactions between large molecular surfaces, while preserving specificity (12).

In order to understand the binding mechanism of an RNA–protein interaction, structural information is required for both the free components and the bound complex. There are a limited but growing number of cases where such detailed structural information is available (13–22). The present study shows that, even in the absence of its transcription factor target, a 29-nt anti-NF-κB RNA aptamer is pre-organized into a structure that mimics the natural DNA recognition site of the transcription factor.

MATERIALS AND METHODS

RNA synthesis and purification

The anti-NF-κB RNA aptamer was transcribed *in vitro* using purified His_6-tagged T7 RNA polymerase and synthetic DNA oligonucleotides (Integrated DNA Technologies), as previously described (23,24). RNA was purified by denaturing 15% polyacrylamide gel electrophoresis, identified by UV absorbance, excised from the gel, recovered by diffusion into 0.3 M sodium acetate and precipitated from ethanol. The RNA was further purified on an anion exchange column [MonoQ ($5 \times 1\,cm^2$)] and desalted on a gel filtration column [BioRad P6 ($5 \times 1\,cm^2$)] performed on a BioRad LC system. The purified RNA was lyophilized, resuspended in water and brought to pH 6.8 by the addition of 1 M NaOH. The sample was further lyophilized to enable preparation of a 99.99% D_2O solution for the acquisition of non-exchangeable proton data. $^{13}C/^{15}N$-labeled RNA was prepared using $^{13}C/^{15}N$-labeled rNTPs [Isotec Inc. (Sigma Aldrich) Miamisburg, OH]. All NMR samples were ~1 mM RNA and 50 mM NaCl.

NMR spectroscopy

NMR spectra were collected on 750 MHz Bruker Avance DMX or 900 MHz Varian Inova spectrometers at the National Magnetic Resonance Facility at Madison (NMRFAM). The spectrometers were equipped with a single *z*-axis gradient HCN cryoprobe and conventional room-temperature probe, respectively. Exchangeable resonances were assigned by reference to 2D NOESY spectra (50, 100 and 150 ms mixing time) in 90% H_2O/10% D_2O at 283 K using a 1-1 spin echo water suppression scheme (25). Non-exchangeable resonances were assigned by reference to 2D NOESY spectra (100, 200, 250 and 300 ms mixing times), 2D (^1H–^1H) TOCSY, 2D ^1H–^{13}C HSQC (aromatic and ribose optimized), 3D (^1H–^{13}C–^1H) TOCSY, 3D (^1H–^{13}C–^1H) COSY and 3D (^1H–^{13}C–^1H) NOESY-HMQC spectra of the RNA in 99.99% D_2O at 293 and 303 K, as previously described (26). The completeness of the ^1H chemical shift assignments was 97% for non-exchangeable and 67% for exchangeable resonances. For experiments in 99.99% D_2O, the residual HDO resonance was suppressed with a low power pre-saturation pulse. Hydrogen bonds for the ^{13}C, ^{15}N-labeled sample in 90% H_2O/10% D_2O at 298 K were detected using a $^2J_{NN}$ 2D HNN COSY experiment (27). Thermal

unfolding of the RNA aptamer was monitored by acquiring a series of 1D ^{1}H spectra at increasing temperatures (4, 8, 12, 18, 22, 26, 30, 35, 40, 45 and 50°C). Partial alignment of RNA for residual dipolar coupling (RDC) measurements was achieved by adding 17 mg/ml Pf1 filamentous bacteriophage (ASLA Ltd, Riga, Latvia) to the $^{13}C/^{15}N$-labeled samples (28). The $^{1}J_{C-H}$ couplings were measured under isotropic and anisotropic conditions using a *J*-modulated [^{13}C–^{1}H] CT-HSQC with a total of 20 and 10 *J*-modulated planes for aromatic and ribose optimized HSQC experiments, respectively (28,29).

Structure calculations

Structural constraints for the solution structure of the NF-κB RNA aptamer were categorized by the qualitative peak volumes of the NOE [strong (1.8–3.0 Å), medium (2.0–4.5 Å) or weak (3.0–6.0 Å)] obtained from 2D NOESY spectra (99.99% D_2O) with mixing times of 100, 200, 250 and 300 ms. Torsion angle restraints for residues in the lower (G1–C5, G25–C29) and upper (A10–G14, A17–U21) stem were constrained to A-form values (± 15°), which were consistent with NOESY, TOCSY, HNN-COSY and RDC data. Hydrogen bonds and weak planarity restraints (1 kcal mol^{-1} Å^{-2}) were enforced for the 8 Watson–Crick and 1 U–G wobble pair, since these base pairs were unambiguously identified by NOESY and HNN-COSY experiments. Two hydrogen bonds in the G–A pair within the GUAA tetraloop were also enforced, but only after structure calculations in the complete absence of non-experimental restraints indicated formation of a G–A pair in the majority of the accepted structures (based on acceptance criteria, *vide infra*). Aside from the two hydrogen bond restraints in the G–A pair in the tetraloop, no non-experimentally derived restraints (distance, dihedral or planarity) were used in the structure calculations for any of the internal loop or tetraloop nucleotides.

An extended structure generated in CNS 1.1 (30) was used to calculate 100 starting structures, as previously described (31). The 50 structures with lowest energy were refined using Xplor-NIH (32) with conformational database potentials (33). Use of the conformational database potentials did not alter the structure of the RNA loop regions, as determined by comparison of structures calculated in their absence, and were validated by an improved fit to the RDC data. PALES software (http://spin.niddk.nih.gov/bax/software/PALES) was used to estimate the values of the axial (D_a) and rhombic (R) components of the alignment tensor from the converged, low energy structures calculated by Xplor-NIH, as previously described (24). Optimal D_a/R-values were –25/0.14 for the 19 lower stem (G1–C5, G25–C29) RDC values. RDC values for the upper helix, internal loop and GUAA tetraloop could not be fit to the lower helix alignment tensor, and were therefore not used in the final structure calculations. The lowest-energy structures with no NOE violations (>0.5 Å), dihedral violations (> 5°) and RDC violations (>3 Hz) were accepted. Structures were analyzed in MOLMOL (34) and figures were created using PyMOL (www.pymol.org).

RESULTS

NMR analysis of the anti-NF-κB RNA aptamer

The 29-nt anti-NF-κB RNA aptamer sequence (Figure 1A) binds tightly to the $p50_2$ form of NF-κB, and competitively blocks NF-κB $p50_2$ binding to duplex DNA *in vitro* and *in vivo* (1,2,7). In the RNA–$p50_2$ complex (8), the RNA has a pronounced kink within its asymmetric internal loop (Figure 1B), which makes extensive interactions with protein residues.

Analysis of the free RNA aptamer under NMR conditions (1 mM RNA, 50 mM NaCl, pH 6.8) by native gel electrophoresis showed a single conformation indicative of a monomeric hairpin stem-loop structure (Supplementary Figure S1). Multi-dimensional TOCSY, NOESY and HSQC data exhibit disperse resonances throughout the RNA, suggesting that the molecule is well structured in a single hairpin conformation. **The** 1D ^{1}H NMR and 2D HNN COSY NMR spectra reveal that the RNA contains eight Watson–Crick base pairs, as expected (Figure 2A). It should be noted that the minor RNA conformational difference at the 5′ and 3′ termini is due to the presence of a mixture of n and $n + 1$-transcription product lengths. Additional resonances noted in Figure 2A likely correspond to alternative G1 and U28 imino protons and were not incorporated into the NMR structure calculations.

Overall, the NMR data are consistent with formation of three non-Watson–Crick pairs. The intense NOEs arising between the U13 and G18 imino protons are characteristic of a U-G wobble pair (data not shown), as is their lack of an N–H$\cdots$N correlation in the HNN COSY (Figure 2A). The sharp G imino proton at 10.3 ppm (Figure 2A) can be confidently assigned by NOESY to G14 and displays NOEs characteristic of the well-known sheared G-A pair conformation, as predicted for a GUAA tetraloop that belongs to the GNRA tetraloop family (35,36). In addition, the broad G imino proton resonance at 10.2 ppm displays internucleotide NOEs that allow assignment to G22, which is also consistent with a sheared G-A conformation. These internucleotide NOEs include: G22 to A9 (H1′–H2), G22 to G23 (H1–H8), G22 to A9 (H1–H8) and G22 to A9 (H1–H61). Based on 1D NMR spectra as a function of temperature, the relative peak intensities of the G14 and G22 imino protons seen in Figure 2A reflects the higher RNA stability observed in the tetraloop than in the internal loop region, respectively (data not shown). The identification of cross-strand adenine H2 to H1′ NOEs further confirmed the secondary structure across the RNA and gave additional evidence to support two sheared G-A pairs (A9–G22 and G14–A17).

Non-exchangeable base and sugar protons were assigned via sequential connectivities in NOESY data (Figure 2B). This ribose to base NOE pattern is indicative of a structured helix and is observed throughout the RNA with the exception of nt 15–17, whose NOE patterns are highly similar to the previously characterized GUAA tetraloop conformations (36). Structural features of the internal loop region (nt 6–9, 22–24) were revealed in the NMR data. G8 displays a strong intraresidue H1′–H8 NOE in addition to ribose H1′–H2 and H1′–H3′ scalar

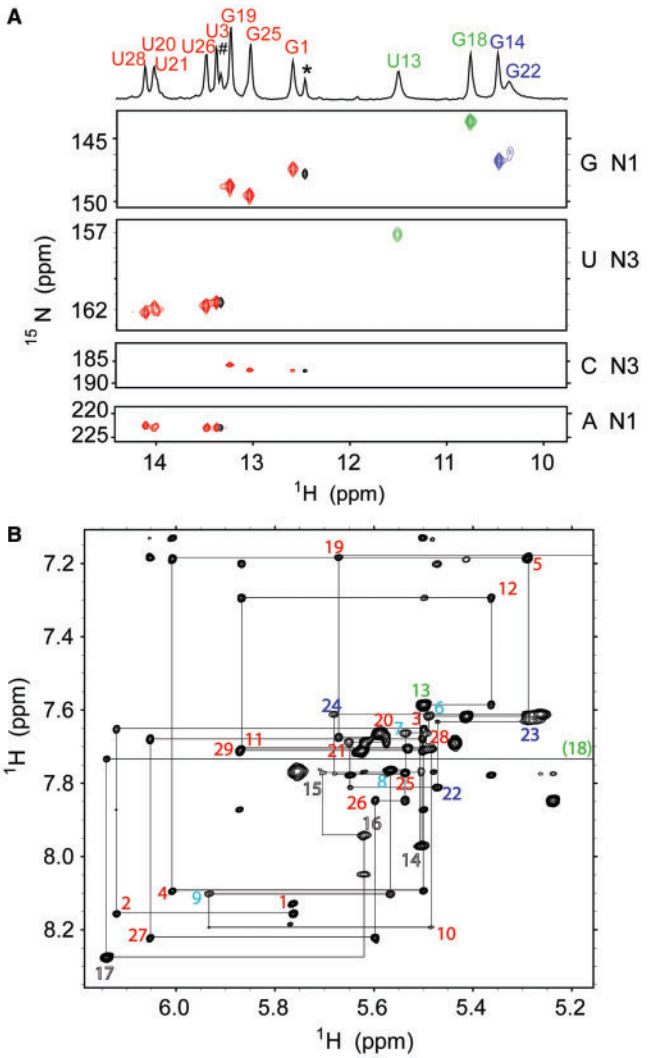

Figure 2. NMR evidence for a single, pre-organized RNA aptamer conformation. (**A**) Base pairing interactions of the RNA aptamer in solution. A 1D 1H NMR spectrum and resonance assignments for the imino protons (top). A 750 MHz 2D HNN COSY experiment showing cross-hydrogen bond scalar couplings ($^{2}J_{N,N}$) (bottom). Resonances and chemical shift assignments are colored as in Figure 1. The number and asterisk symbols correspond to a minor $n+1$ transcription product, which has different chemical shifts for the U28 (#) and G1 (*) imino protons, but does not otherwise interfere with analysis. (**B**) A 900 MHz 2D ^{1}H–^{1}H NOESY spectrum (250 ms mixing time) of the RNA aptamer. Intranucleotide H1′–H6/8 NOE assignments are indicated by number and a sequential internucleotide 'walk' is illustrated. Numbers are colored as in Figure 1. The parenthesis for nucleotide G18 denotes that its H1′ chemical shift is not in the region displayed.

couplings, indicating that it adopts a *syn* conformation and has a C2′ *endo* sugar pucker. Cross-strand NOEs, including G8 to G23 (H1′– H1′), G8 to G23 (H1′–H8), U7 to C24 (H5–H5) and U7 to C24 (H5–H5), were also observed in the internal loop.

Solution structure of the anti-NF-κB-RNA aptamer

Structures of the anti-NF-κB RNA aptamer were calculated with 541 NOE-derived distance restraints, 160 backbone torsion angle restraints for the helical regions, 25 hydrogen bond restraints based on NOESY and $^{2}J_{NN}$ COSY analysis, and 19 residual dipolar couplings incorporated at the refinement stage (Figure 2 and Table 1). The structural statistics and stereo superimposition of the 10 lowest-energy structures are shown (Table 1 and Figure 3A). The lower (nt 1–5, 25–29) and upper (nt 10–13, 18–21) helical regions adopt predominantly A-form geometries, and the GUAA fold has the same stacking and hydrogen bonding patterns as previously observed for GNRA tetraloop structures, despite the presence of an adjacent U13-G18 wobble pair (Figures 3 and 4A) (36,37). In contrast, the internal loop deviates from a typical A-form fold (Figures 3B and 4B). Unpaired pyrimidines 6 and 7 contain aromatic stacking interactions within the helix while C24 is stacked between the two uridines. Cross-strand stacking of G8 and G23 reduces the interphosphate cross-strand distance by 3 Å, down to 15 Å in comparison to a regular A-form strand width of 18.5 Å. The non-A-form helical geometry of the internal loop structure appears to be largely due to this cross-strand stack. Cross-strand stacking interactions involving G-A pairs are relatively common in RNA structures (38–41).

Table 1. Structure statistics of the energy-minimized NMR structures of the anti-NF-κB RNA aptamer

NOE-derived distance restraints	541
Intranucleotide	202
Internucleotide	339
Hydrogen bond constraints	25
Dihedral angle constraints	160
Residual dipolar couplings	19
RMSD for all heavy atoms to the mean coordinates (Å)	0.94
Internal loop (nt 6–9, 22–24)	0.56
Average NOE RMSD (Å)	0.059
Average RDC RMSD (Hz)	1.9

Comparison of the anti-NF-κB RNA aptamer alone and in complex with the p50 protein subunit illuminates the conformational changes induced upon protein binding, while also revealing a high degree of structural similarity (Figures 3B and 4). Although the lower helix and RNA secondary structure are predominantly unchanged between the free and bound states, regions within the GUAA tetraloop and internal loop are different. In contrast to the NMR structure and a 1.4 Å resolution GUAA RNA crystal structure (37), the GUAA tetraloop in the protein-bound state contains a flipped, unstacked conformation between U15 and A16 (Figure 4A). Multiple protein contacts with the phosphate backbone occur in the tetraloop to stabilize this alternative conformation, which may also be related to recently detected transient conformations of GNRA tetraloops (42). For the internal loop, the secondary structure features are nearly identical while the backbone fold is significantly different (Figures 3B and 4B). In both the free and bound structures, a sheared A9-G22 pair sets up the subsequent cross-strand stacking of G23 upon G8. Thus both the free and bound RNA structures contain a continuous G22–G23–G8 purine stack. The A9–G22–G23–G8 region has a 3.6 Å RMSD with

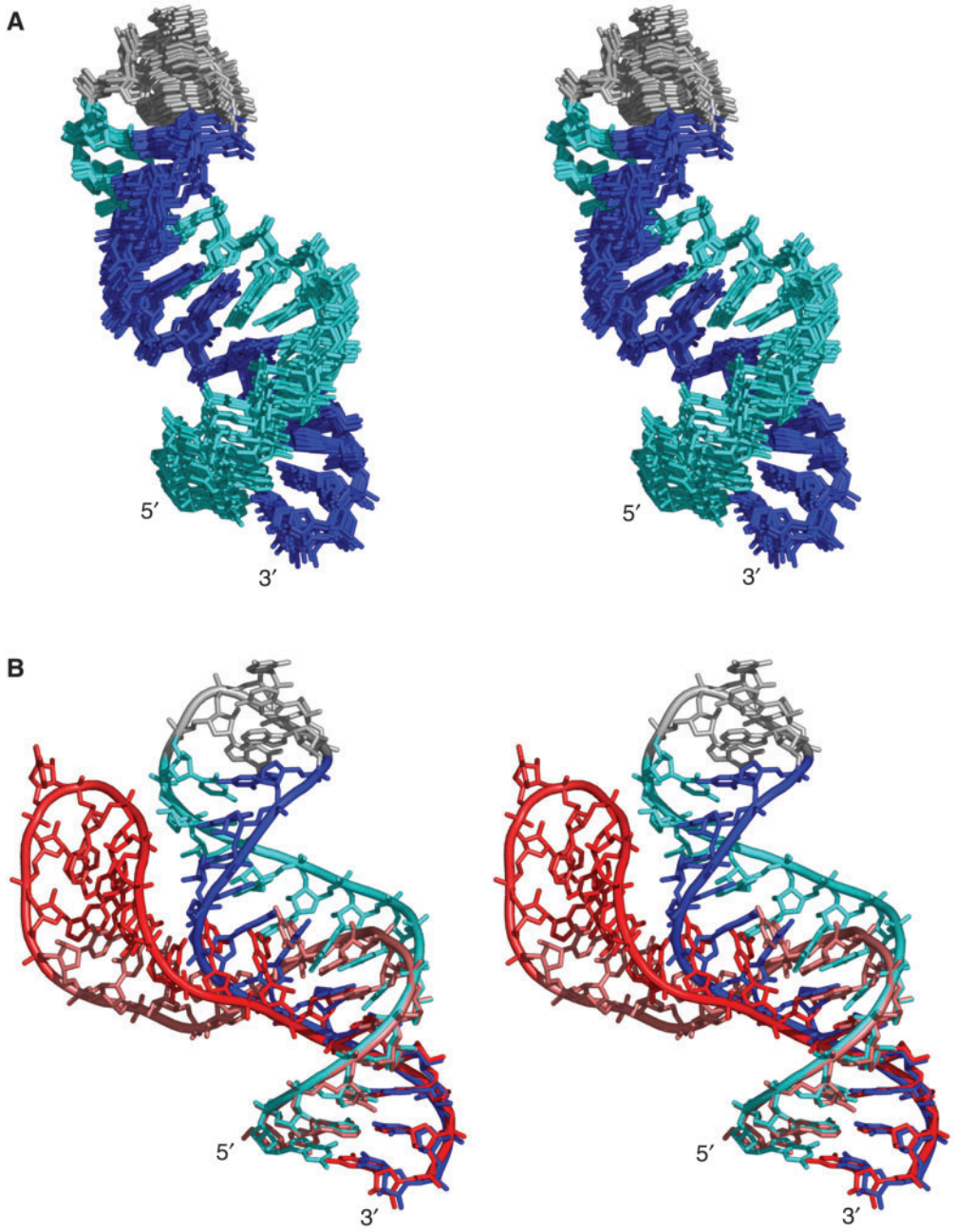

Figure 3. Global structure of the free and bound anti-NF-κB RNA conformations. (**A**) Stereo view of the superimposition over all non-hydrogen atoms of the 10 lowest energy solution structures (RMSD ~0.94 Å). The 5′ (nt 1–13) and 3′ (nt 18–29) termini of the RNA are colored cyan and blue, respectively, while the GUAA tetraloop is colored gray. (**B**) Stereo view of the superimposition over just the lower helix. The NMR structure is colored according to Figure 3A and the crystal structure is shown in red (nt 1–17) and pink (nt 18–29).

the crystal structure of the bound form (Figure 4B). The cross-strand purine stack is followed by a pyrimidine-rich region (U6, U7 and C24), which makes extensive backbone and base contacts when bound by the $p50_2$ homodimer (Figure 5). The free RNA is bent in the same direction as the bend observed for the bound form, but to a lesser degree. Measurement of the global axis curvature using CURVES 5.3 software (43) reveals $39.5 \pm 10.0^{\circ}$ and $97 \pm 8.0^{\circ}$ helical bend angles for the free (NMR) and protein bound (crystal) states of the RNA, respectively. It is likely that the extensive base and backbone protein contacts account for the higher degree of helical bending in the bound form, and significantly contribute to the irregular RNA fold observed in the co-crystal complex (Figure 5) (8).

A high degree of intrinsic flexibility of the free RNA structure may assist in the observed helical bending and induced fit. In molecules of this size, internal motions are coupled to overall rotational motions and are thus difficult to analyze. Nevertheless, significant variations in resonance intensities can report the net dynamics of local RNA motions relative to the applied magnetic field, even

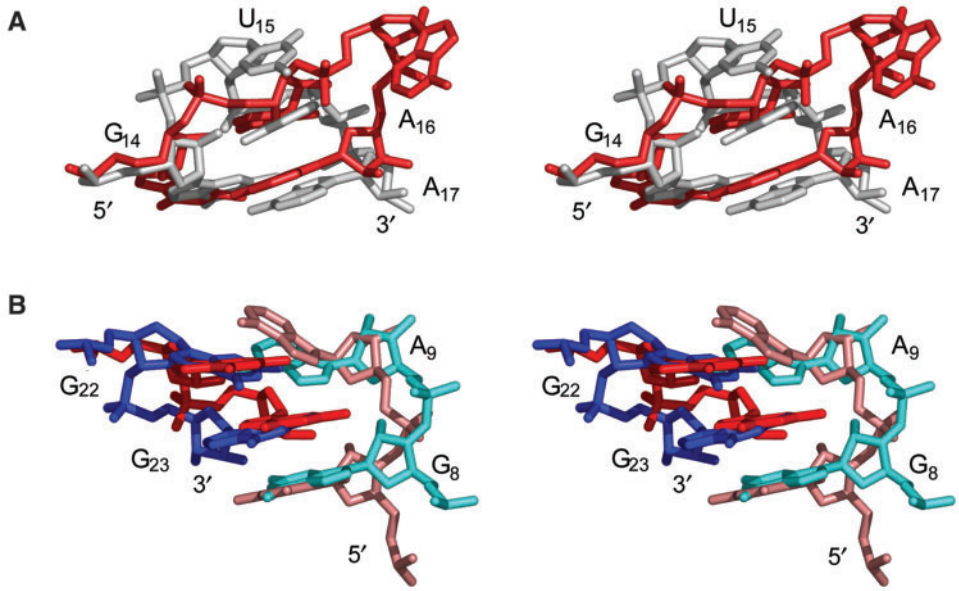

Figure 4. Comparison of the lowest energy NMR (free) and crystal (bound) structures of the GUAA tetraloop and internal loop regions. (**A**) Stereo view of the superimposed free (gray) and protein-bound (8) GUAA tetraloops. (**B**) Stereo view of the superimposed internal loop structures (nt 8,9, 22,23, RMSD = 3.6 Å). Color is as in Figure 3.

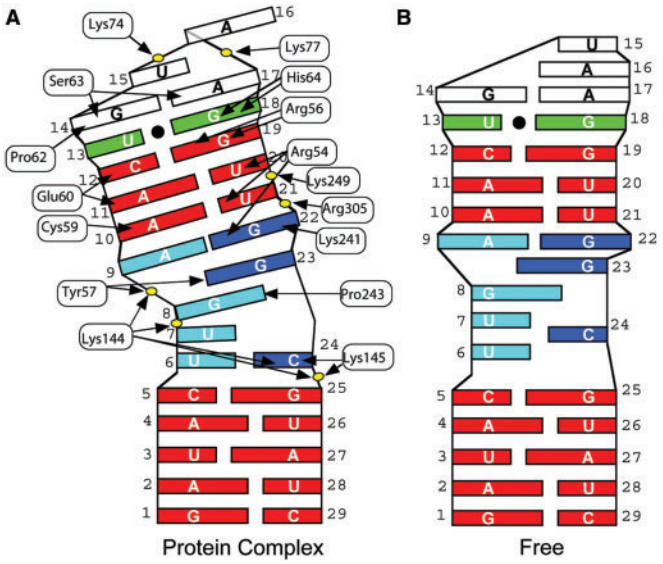

Figure 5. Comparative base stacking diagrams of anti-NF-κB RNA aptamer bound in the crystal complex (**A**) and free in solution (**B**). A summary of contacts between the RNA aptamer and the p50 monomer (based on PDB ID 1OOA) is shown in (A). Nucleotides are colored as in Figure 1 and phosphates in the RNA backbone contacted by p50 amino acid residues are represented as yellow ovals.

when internal and rotational motions are coupled (44). Resonances with higher intensities are indicative of internal motions that are faster (pico-nanosecond timescale) than the overall tumbling, whereas lower resonance intensities are indicative of slower (micro-millisecond timescale) internal motions. Quantification of the resonance intensities shows that the internal loop and tetraloop regions, in addition to the 3′ end of the molecule, experience fast timescale motions (Figure S2, Supplementary Data). In particular, nucleotides adjacent or within the internal loop (U7, U21, G22 and C24) and the tetraloop region (U13, A16 and A17) experience fast motions. Nucleotides near the 3′ end of the RNA (A27, U28 and C29) also show higher resonance intensities. Interestingly, G18 in the U–G wobble pair adjacent to the tetraloop was the only nucleotide to display slower timescale motions.

Analysis of the helical parameters in the free (NMR) and protein-bound (crystal) states of the anti-NF-κB RNA aptamer reveals striking similarities with both idealized B-form DNA and A-form RNA (Table 2). Statistical parameters, such as base pair slide, inclination, *x*-axis displacement, angular information [delta (δ°), chi (χ°), pseudorotation phase and ribose pucker] and groove distances/widths serve as reliable discriminators between B- and A-form geometries (45). While statistics for the helical regions of the NMR structure are more indicative of A-form RNA, substantial deviations favoring a B-form intermediate transition exist in the internal loop region. In fact, both NMR and crystal structures of the anti-NF-κB RNA aptamer contain an extended rise and closely imitate the major groove distance of idealized B-form geometry within this internal loop region. Interestingly, it is the expansion of the RNA major groove face, due to the guanine cross-strand stack, that mediates direct contacts in the RNA–$(p50)_2$ complex.

DISCUSSION

Pre-organized structure of the anti-NF-κB RNA aptamer

This study enables a comparison of the detailed structures of RNA and protein partners before and after complex formation. We show that the anti-NF-κB RNA aptamer is

Table 2. Helical parameters of the free (NMR) and $(p50)_2$-bound (crystal) forms of the anti-NF-κB RNA aptamer in comparison with a B-DNA or A-RNA duplex.[a]

	B-form DNA	RNA Aptamer Bound (crystal)[b]	RNA Aptamer free (NMR)[b]	A-form RNA
Rise/residue, (Å)	3.4	3.8	3.3	2.8
Slide (Å)	−0.76	−1.40 (−0.75)	−1.80 (−0.86)	−2.14
Incline (°)	−5.93	−11.90 (−12.50)	−1.34 (−5.70)	15.95
x-Displacement (Å)	−0.71	−0.86 (−0.14)	−2.40 (−0.86)	−5.30
Delta (δ°), Chi (χ°)	156.4, −97.9	84.2, −157.5 (91.6, −144.7)	86.1, −155.6 (94.3, −135.8)	83.5, −166.5
Pseudorotation (°)	191.6	29.8 (42.5)	46.8 (70.9)	13.4
Ribose sugar pucker conformation	C2′-endo	C3′-endo (C3′-endo)[c]	C3′-endo (C2′, O4′, and C3′-endo)[d]	C3′-endo
Major Groove[e]—distance (Å)	17.9	18.7 ± 3.5	19.1 ± 3.0	10.4
Minor Groove[f]—width, (Å)	5.90	11.10 (6.50)	11.30 (9.40)	11.05

[a]Idealized B-DNA [5′-GATACTTGAACGT-3′ and 5′-ACGTTCAAGTATC-3′] and A-RNA [5′-GAUACUUGAACGU-3′ and 5′-ACGUUCA AGUAUC-3′] models were made in Insight (Biosym). Helical parameters were generated by CURVES v5.3 (43). Unless noted, errors for all RNA aptamer parameters (including structures within the NMR ensemble) are within 20% of the average values reported.
[b]For the RNA aptamer, Watson–Crick helical regions (nt 2–5,10–13,19–21,25–28) are reported as the top number, whereas parameters for the internal loop region (nt 6–9 and 22–24), are given in parentheses.
[c]Ribose of G8 adopts a C4′-exo conformation in the crystal structure.
[d]For the free (NMR) RNA aptamer, ribose puckers left unrestrained in structure calculations gave rise to a mixture of sugar pucker conformations.
[e]Major groove distances between the A4–C24, C5–G23, and U6-G22 phosphorus atoms (or equivalent positions in idealized B-DNA and A-RNA) were measured in PyMol.
[f]Minor groove width values represent the minimal width with respect to the local helical axis, as defined by CURVES v5.3.

pre-structured, specifically through non-canonical base pairing and a cross-strand stacking arrangement within the asymmetric internal loop. This structure is likely crucial for engagement of the DNA-binding surface of the p50 protein. We find that RNA structural changes induced upon protein binding are (i) a hinge-like bending within the internal loop to change the orientation of the two Watson–Crick stems, and (ii) a protein-induced perturbation of the GUAA tetraloop, which otherwise adopts a canonical fold in the free RNA. Further, we propose that the cross-strand guanine stack within the major groove of the internal loop is a critical structural element for $p50_2$ recognition and specificity, assisting in the molecular mimicry of the κB DNA target sequence by expanding the major groove. A comparison of the overall fold and helical parameters between the lowest energy NMR structure, B-form DNA, and A-form RNA support this hypothesis, illustrating that the anti- NF-κB RNA aptamer closely imitates the exposed major groove face of the DNA (Table 2, Figure 6).

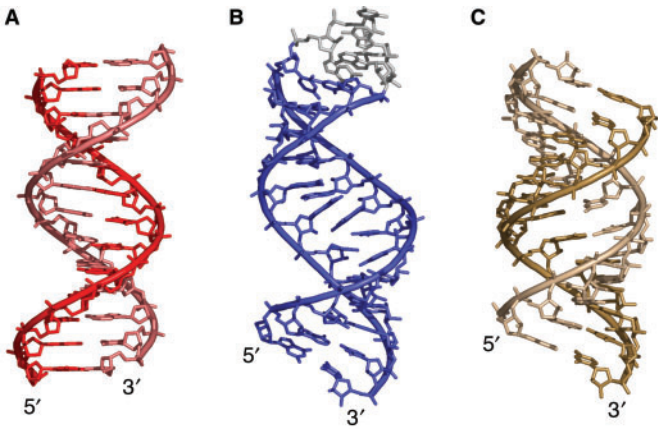

Figure 6. Comparison of B-form DNA, the NMR structure of the free NF-κB RNA aptamer and A-form RNA. (**A**) Idealized B-form duplex DNA is colored as in Figure 1D and contains the sequences 5′-GATACTTGAACGT-3′ and 5′-ACGTTCAAGTATC-3′. (**B**) Anti-NF-κB RNA aptamer NMR structure (blue), capped by the GUAA tetraloop (gray). (**C**) Idealized A-form duplex RNA is colored brown and tan and contains the sequence 5′-GAUACUUGAACGU-3′ and 5′-ACGUUCAAGUAUC-3′. The 26-nt DNA and RNA coordinates were generated using Insight (Biosym).

Precedence for pre-structured functional RNAs

Although an increasing number of cases in the literature investigate RNA folding in free and protein-bound states, there is a limited number of examples where both RNA conformations are structurally defined (13–22). How does the anti-p50 RNA aptamer compare with other RNAs with respect to folding rearrangement upon target binding? There is a general trend toward mutually induced fit between RNAs and their binding partners, but important examples exist where binding is not accompanied by major changes in RNA folding. In these examples, much or all of the interaction surface of the RNA molecule is preformed in the absence of the binding partner. The tRNA/tRNA synthetase complexes are examples where structured and pre-folded RNAs are observed (46–54). Upon protein binding, changes in secondary and tertiary tRNA structure are typically minor, and are limited to the flexible anticodon loop and the 3′ acceptor nucleotide. The overall geometry of the folded tRNA appears relatively rigid and autonomous, and this appears to also be true for a tRNA aptamer selected for increased synthetase affinity (55).

Other helical RNAs have also been found to be pre-structured prior to binding to their protein targets. Specifically, two examples (the SAM protein domain–SRE RNA and Rnt1p protein–snR47h RNA complex) reveal protein recognition through structured RNA pentaloops or tetraloops (17,18). In the case of the SAM domain–SRE RNA interaction, both the protein and the RNA are structured and complex formation occurs upon binding to the major groove of a pre-formed pentaloop

structure (17). For the Rnt1p–snR47h–RNA complex, the AGNN tetraloop conformation also retains the identical fold in the free and bound states (18). Thus, unlike the tetraloop conformational change that occurs in the anti-NF-κB aptamer, these two examples show rigid, pre-formed loop structures that do not undergo conformational re-organization upon protein binding.

Mutually induced fit in RNA–protein interactions

The concept of protein-induced RNA folding has been reviewed by Williamson (14), who highlighted structures of RNA binding partners determined both alone and in a complex. Cases include the ribosomal S15 and L11 proteins inducing large structural changes upon ribosomal RNAs, ribonucleoprotein U1A changing the folding of the 5′ untranslated region of its own mRNA upon binding, and a similar example for ribosomal protein L30 binding to its own mRNA. In each of these examples, both binding partners undergo conformational changes upon interaction, though the effect is most striking for the RNA. In the present case, protein contacts at the RNA/$p50_2$ interface are nearly identical to those observed in the κB-DNA/$p50_2$ interface, while only the overall N-terminal protein domain orientation has changed to accommodate aptamer binding (Figure 1C and D).

The increasing number of detailed structural studies for protein and RNA partners alone and in complex provides additional opportunity for comparison. In one example, structure determination of the free protein (56), free RNA (13) and complex (57) enabled a thorough analysis of nucleolin binding to a consensus RNA aptamer. The authors show that certain elements of RNA stem structure were detectable in the absence of protein, but a dynamic equilibrium of RNA loop structures was also observed. Nucleolin binding stabilized one particular RNA form, clearly confirming the existence of a mutually induced protein–RNA complex (13). Thus, like the anti-p50 RNA aptamer, the RNA reorganization upon protein complex formation might be characterized as moderate, with pre-formation of long-range RNA structure that is retained in the complex.

CONCLUSION

The interaction surface of p50 protein has been previously compared in its DNA and RNA complexes (8–10). The protein surface is essentially unchanged in the two complexes, exemplifying the DNA mimicry by the RNA aptamer. NMR evidence reveals that the free p50 RNA aptamer adopts a single structured conformation and displays features found in the complex, including non-canonical base pairing and cross-strand stacking of its internal loop bases. One implication of this pre-folded aptamer structure is that there should be a less unfavorable entropic contribution to the free energy change for the binding reaction relative to an RNA that only becomes folded upon protein binding. It is likely that this feature of the anti-NF-κB RNA aptamer contributes to its low nanomolar equilibrium dissociation constant for p50 binding (1,2). We propose that a pre-formed structural transition towards B-form helical geometry as well as helical bending induced by the internal loop play essential roles in high affinity binding to the $(p50)_2$ NF-κB transcription factor (Table 2, Figure 6). Parameters extracted for the helical and internal loop regions support this notion, revealing that an A- to B-form transition indeed occurs across this stem loop RNA. Additionally, $(p50)_2$ binding substantially alters the GUAA tetraloop structure through several backbone contacts (Figure 5A), indicating that this seemingly stable tetraloop conformation can also serve as a dynamic and mobile protein recognition element. Thus, the anti-NF-κB RNA aptamer imitates DNA, and both pre-formation of RNA structure and induced fit assist in formation of the protein–RNA complex.

Coordinates

Coordinates for the 29-nt anti-NF-κB-RNA aptamer have been deposited into the RCSB Protein Data Bank (accession code 2jwv) and NMR resonance assignments have been deposited into BioMagResBank (accession code 15538).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Marco Tonelli for assistance with NMR spectroscopy and Ryan Marcheschi for assistance with data analysis. This study made use of the National Magnetic Resonance Facility at Madison (www.nmrfam.wisc.edu), which is supported by National Institutes of Health grants P41RR02301 (Biomedical Research Technology Program, National Center for Research Resources) and P41GM66326 (National Institute of General Medical Sciences). Equipment in the facility was purchased with funds from the University of Wisconsin, the National Institutes of Health (P41GM 66326, P41RR02301, RR02781, RR08438), the National Science Foundation (DMB-8415048, BIR-9214394) and the US Department of Agriculture. This work was supported by the Mayo Foundation and NIH grant GM68128 to L.J.M., and a Milwaukee Foundation Shaw Scientist Award to S.E.B. Funding to pay the Open Access publication charges for this article was provided by NIH grant GM68128.

Conflict of interest statement. None declared.

REFERENCES

1. Lebruska,L.L. and Maher,L.J. 3rd. (1999) Selection and characterization of an RNA decoy for transcription factor NF-kappa B. *Biochemistry*, **38**, 3168–3174.
2. Cassiday,L.A. and Maher,L.J. 3rd. (2003) Yeast genetic selections to optimize RNA decoys for transcription factor NF-kappa B. *Proc. Natl Acad. Sci. USA*, **100**, 3930–3935.
3. Ellington,A.D. and Szostak,J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.

4. Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.
5. Darnell,J.E. Jr. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 740–749.
6. Yamamoto,Y. and Gaynor,R.B. (2001) Therapeutic potential of inhibition of the NF-kappaB pathway in the treatment of inflammation and cancer. *J. Clin. Invest.*, **107**, 135–142.
7. Cassiday,L.A. and Maher,L.J. 3rd. (2001) In vivo recognition of an RNA aptamer by its transcription factor target. *Biochemistry*, **40**, 2433–2438.
8. Huang,D.B., Vu,D., Cassiday,L.A., Zimmerman,J.M., Maher,L.J. 3rd and Ghosh,G. (2003) Crystal structure of NF-kappaB (p50)2 complexed to a high-affinity RNA aptamer. *Proc. Natl Acad. Sci. USA*, **100**, 9268–9273.
9. Ghosh,G., Duyne,G.V., Ghosh,S. and Sigler,P.B. (1995) Structure of NF-kB p50 homodimer bound to a kB site. *Nature*, **373**, 303–310.
10. Ghosh,G., Huang,D.B. and Huxford,T. (2004) Molecular mimicry of the NF-kappaB DNA target site by a selected RNA aptamer. *Curr. Opin. Struct. Biol.*, **14**, 21–27.
11. Hendrickson,W.A., Pahler,A., Smith,J.L., Satow,Y., Merritt,E.A. and Phizackerley,R.P. (1989) Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proc. Natl Acad. Sci. USA*, **86**, 2190–2194.
12. Spolar,R.S. and Record,M.T. Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science*, **263**, 777–784.
13. Bouvet,P., Allain,F.H., Finger,L.D., Dieckmann,T. and Feigon,J. (2001) Recognition of pre-formed and flexible elements of an RNA stem-loop by nucleolin. *J. Mol. Biol.*, **309**, 763–775.
14. Williamson,J.R. (2000) Induced fit in RNA-protein recognition. *Nat. Struct. Biol.*, **7**, 834–837.
15. Shajani,Z. and Varani,G. (2005) 13C NMR relaxation studies of RNA base and ribose nuclei reveal a complex pattern of motions in the RNA binding site for human U1A protein. *J. Mol. Biol.*, **349**, 699–715.
16. Walden,W.E., Selezneva,A.I., Dupuy,J., Volbeda,A., Fontecilla-Camps,J.C., Theil,E.C. and Volz,K. (2006) Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science*, **314**, 1903–1908.
17. Oberstrass,F.C., Lee,A., Stefl,R., Janis,M., Chanfreau,G. and Allain,F.H. (2006) Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. *Nat. Struct. Mol. Biol.*, **13**, 160–167.
18. Wu,H., Henras,A., Chanfreau,G. and Feigon,J. (2004) Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc. Natl Acad. Sci. USA*, **101**, 8307–8312.
19. Lu,M. and Steitz,T.A. (2000) Structure of Escherichia coli ribosomal protein L25 complexed with a 5S rRNA fragment at 1.8-A resolution. *Proc. Natl Acad. Sci. USA*, **97**, 2023–2028.
20. Stoldt,M., Wohnert,J., Ohlenschlager,O., Gorlach,M. and Brown,L.R. (1999) The NMR structure of the 5S rRNA E-domain-protein L25 complex shows preformed and induced recognition. *EMBO J.*, **18**, 6508–6521.
21. Jovine,L., Oubridge,C., Avis,J.M. and Nagai,K. (1996) Two structurally different RNA molecules are bound by the spliceosomal protein U1A using the same recognition strategy. *Structure*, **4**, 621–631.
22. Williams,D.J. and Hall,K.B. (1996) RNA hairpins with non-nucleotide spacers bind efficiently to the human U1A protein. *J. Mol. Biol.*, **257**, 265–275.
23. Milligan,J.F. and Uhlenbeck,O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **180**, 51–62.
24. Reiter,N.J., Blad,H., Abildgaard,F. and Butcher,S.E. (2004) Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry*, **43**, 13739–13747.
25. Sklenar,V., Brooks,B.R., Zon,G. and Bax,A. (1987) Absorption mode two-dimensional NOE spectroscopy of exchangeable protons in oligonucleotides. *FEBS Lett.*, **216**, 249–252.
26. Huppler,A., Nikstad,L.J., Allmann,A.M., Brow,D.A. and Butcher,S.E. (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat. Struct. Biol.*, **9**, 431–435.
27. Grzesiek,S., Cordier,F. and Dingley,A.J. (2001) Scalar couplings across hydrogen bonds. *Methods Enzymol.*, **338**, 111–133.
28. Hansen,M.R., Mueller,L. and Pardi,A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.*, **5**, 1065–1074.
29. Cutting,B., Tolman,J.R., Nanchen,S. and Bodenhausen,G. (2002) Accurate measurement of residual dipolar couplings in anisotropic phase. *J. Biomol. NMR*, **23**, 195–200.
30. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **54**, 905–921.
31. Reiter,N.J., Nikstad,L.J., Allmann,A.M., Johnson,R.J. and Butcher,S.E. (2003) Structure of the U6 RNA intramolecular stem-loop harboring an S(P)-phosphorothioate modification. *RNA*, **9**, 533–542.
32. Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, **160**, 65–73.
33. Clore,G.M. and Kuszewski,J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am. Chem. Soc.*, **125**, 1518–1525.
34. Koradi,R., Billeter,M. and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55, 29–32.
35. 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.
36. Jucker,F.M., Heus,H.A., Yip,P.F., Moors,E.H. and Pardi,A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.*, **264**, 968–980.
37. Correll,C.C. and Swinger,K. (2003) Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 A resolution. *RNA*, **9**, 355–363.
38. Hoffmann,B., Mitchell,G.T., Gendron,P., Major,F., Andersen,A.A., Collins,R.A. and Legault,P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc. Natl Acad. Sci. USA*, **100**, 7003–7008.
39. Schmitz,U., James,T.L., Lukavsky,P. and Walter,P. (1999) Structure of the most conserved internal loop in SRP RNA. *Nat. Struct. Biol.*, **6**, 634–638.
40. Wimberly,B.T., Brodersen,D.E., Clemons,W.M. Jr., Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
41. Hansen,J.L., Ippolito,J.A., Ban,N., Nissen,P., Moore,P.B. and Steitz,T.A. (2002) The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol. Cell*, **10**, 117–128.
42. Zhao,L. and Xia,T. (2007) Direct revelation of multiple conformations in RNA by femtosecond dynamics. *J. Am. Chem. Soc.*, **129**, 4118–4119.
43. Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.
44. Zhang,Q., Sun,X., Watt,E.D. and Al-Hashimi,H.M. (2006) Resolving the motional modes that code for RNA adaptation. *Science*, **311**, 653–656.
45. Vargason,J.M., Henderson,K. and Ho,P.S. (2001) A crystallographic map of the transition from B-DNA to A-DNA. *Proc. Natl Acad. Sci. USA*, **98**, 7265–7270.
46. Sankaranarayanan,R., Dock-Bregeon,A.C., Romby,P., Caillet,J., Springer,M., Rees,B., Ehresmann,C., Ehresmann,B. and Moras,D. (1999) The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell*, **97**, 371–381.
47. Shen,N., Guo,L., Yang,B., Jin,Y. and Ding,J. (2006) Structure of human tryptophanyl-tRNA synthetase in complex with tRNATrp reveals the molecular basis of tRNA recognition and specificity. *Nucleic Acids Res.*, **34**, 3246–3258.
48. Moor,N., Kotik-Kogan,O., Tworowski,D., Sukhanova,M. and Safro,M. (2006) The crystal structure of the ternary complex

of phenylalanyl-tRNA synthetase with tRNAPhe and a phenylalanyl-adenylate analogue reveals a conformational switch of the CCA end. *Biochemistry*, **45**, 10572–10583.
49. Silvian,L.F., Wang,J. and Steitz,T.A. (1999) Insights into editing from an ile-tRNA synthetase structure with tRNAile and mupirocin. *Science*, **285**, 1074–1077.
50. Briand,C., Poterszman,A., Eiler,S., Webster,G., Thierry,J. and Moras,D. (2000) An intermediate step in the recognition of tRNA(Asp) by aspartyl-tRNA synthetase. *J. Mol. Biol.*, **299**, 1051–1060.
51. Delagoutte,B., Moras,D. and Cavarelli,J. (2000) tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding. *EMBO J.*, **19**, 5599–5610.
52. Arnez,J.G. and Steitz,T.A. (1996) Crystal structures of three misacylating mutants of Escherichia coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP. *Biochemistry*, **35**, 14725–14733.
53. Biou,V., Yaremchuk,A., Tukalo,M. and Cusack,S. (1994) The 2.9 A crystal structure of T. thermophilus seryl-tRNA synthetase complexed with tRNA(Ser). *Science*, **263**, 1404–1410.
54. Rould,M.A., Perona,J.J., Soll,D. and Steitz,T.A. (1989) Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. *Science*, **246**, 1135–1142.
55. Bullock,T.L., Sherlin,L.D. and Perona,J.J. (2000) Tertiary core rearrangements in a tight binding transfer RNA aptamer. *Nat. Struct. Biol.*, **7**, 497–504.
56. Allain,F.H., Gilbert,D.E., Bouvet,P. and Feigon,J. (2000) Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target. *J. Mol. Biol.*, **303**, 227–241.
57. Allain,F.H., Bouvet,P., Dieckmann,T. and Feigon,J. (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin. *EMBO J.*, **19**, 6870–6881.