NMR structures of double loops of an RNA aptamer against mammalian initiation factor 4A

Taiichi Sakamoto, Akihiro Oguro, Gota Kawai[1], Takashi Ohtsu and Yoshikazu Nakamura*

Department of Basic Medical Sciences, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan and [1]Department of Life and Environmental Sciences, Faculty of Engineering, Chiba Institute of Technology, Narashino-shi, Chiba 275-0016, Japan

ABSTRACT

A high affinity RNA aptamer (APT58, 58 nt long) against mammalian initiation factor 4A (eIF4A) requires nearly its entire nucleotide sequence for efficient binding. Since splitting either APT58 or eIF4A into two domains diminishes the affinity for each other, it is suggested that multiple interactions or a global interaction between the two molecules accounts for the high affinity. To understand the structural basis of APT58's global recognition of eIF4A, we determined the solution structure of two essential nucleotide loops (AUCGCA and ACAUAGA) within the aptamer using NMR spectroscopy. The AUCGCA loop is stabilized by a U-turn motif and contains a non-canonical A:A base pair (the single hydrogen bond mismatch: Hoogsteen/Sugar-edge). On the other hand, the ACAUAGA loop is stabilized by an AUA tri-nucleotide loop motif and contains the other type of A:A base pair (single hydrogen bond mismatch: Watson–Crick/Watson–Crick). Considering the known structural and functional properties of APT58, we propose that the AUCGCA loop is directly involved in the interaction with eIF4A, while the flexibility of the ACAUAGA loop is important to support this interaction. The Watson–Crick edges of C7 and C9 in the AUCGCA loop may directly interact with eIF4A.

INTRODUCTION

High affinity molecules selected against ligands or proteins are called aptamers. Nucleic acid (i.e. RNA or DNA) aptamers can be selected from combinatorial libraries by *in vitro* selection (referred to as SELEX for systematic evolution of ligands by exponential enrichment) of RNA or DNA molecules that bind to targets (1–3). Two important properties of aptamers are their high affinity and high specificity. Occasionally, aptamers acquire the potential to inhibit the biological function of their target molecules and therefore therapeutic or diagnostic applications of aptamers have been proposed (4). Some of these DNA or RNA aptamers have been examined for their 3D structures by X-ray crystallography or NMR spectroscopy (3). These structural analyses have been useful to generalize two types of aptamers. High affinity aptamers generated against proteins that are potentially nucleic acid-binding proteins are short-length. These aptamers bind to targets by the specific recognition of a limited number of contact sites, such as RRM, KH domain or dsRBDs, through precise stacking of flat moieties, hydrogen bonding and molecular shape complementation. On the other hand, aptamers that are selected to proteins of potentially no or weak affinity to nucleic acids are often large molecules necessary to achieve strong binding to targets (5,6). Structural information of aptamers will enhance our understanding of protein–RNA interaction at the atomic level in general, and facilitate effective design and improvement of RNA aptamers for therapeutic applications.

Mammalian translation initiation is a complex and highly regulated process involving multiple initiation factors (eIFs) (7–9). The initial association of mRNA with the small (40S) ribosomal subunit requires the participation of at least three initiation factors, eIF4A, eIF4B and eIF4F. eIF4F consists of three subunits (eIF4E, eIF4A and eIF4G), and binds to the cap structure [m^{7}GpppN, where N is any nucleotide (10)], which is present at the 5′ end of all cellular mRNAs, via the cap-binding protein subunit, eIF4E. eIF4E is the least abundant factor of all eIFs (11) and the recognition of the mRNA cap by eIF4E is the rate-limiting step of eukaryotic translation initiation. eIF4A is an RNA-dependent ATPase that unwinds the secondary structure present in the 5′ untranslated region of mRNAs (12,13). eIF4A is a prototype member of the DEAD-box RNA helicase

protein family. DEAD-box (and related DEXH-box) proteins contain several highly conserved amino acid sequence motifs (12–15) and are involved in a variety of biological processes involving RNA unwinding and/or rearrangement (16–18). eIF4A has a 'dumbbell'-like structure consisting of two compact domains connected by an extended, 11-residue (~18 Å long) linker (19–21). It is suggested that the inter-domain movement between the N- and C-terminal domains of eIF4A is either necessary for or coupled with ATP hydrolysis and the helicase action (21,22). This, along with changes in proteolytic digestion patterns (22,23), suggests that eIF4A undergoes a series of ligand-dependent conformational changes as it binds its substrates (RNA and ATP), hydrolyzes ATP and releases products.

The accurate control of translation initiation is important for cell growth. Increasing reports show a close relationship between aberrant expression or deregulation of initiation factors and malignant transformation of mammalian cells. For example, overexpression of eIF4E leads to deregulated cell growth (24) and malignant transformation of rodent and human cells (25). Deregulated phosphorylation of eIF4E or 4E-BP1 in Akt signaling leads to tumorigenesis by the activation of eIF4F complex (26–28). Moreover, the cellular level of eIFs, such as eIF4A and eIF4E, is a prognostic indicator of the clinical outcome of a variety of human cancers including breast cancer, head and neck squamous cell carcinoma (29,30) and melanoma (31). Consistent with these findings, antisense RNA against eIF4A has been shown to be effective, at least in part, to suppress proliferation of human melanoma cell lines (32). Therefore, a therapeutic approach that targets eIF4A and other eIFs might have potential for novel cancer therapy.

RNA aptamers that have been selected against eIF4A by SELEX are a useful molecular tool to facilitate the study of eIF4A helicase action and exhibit potential as a novel cancer therapeutic agent (5). One selected RNA aptamer, no. 21, was 87 nt long and was shown to efficiently inhibit ATP hydrolysis of eIF4A and cap-dependent *in vitro* translation. Aptamer no. 21 showed highest binding affinity to eIF4A among the selected aptamers, thus we chose no. 21 for NMR study. Surface plasmon resonance analysis suggests that the dissociation constant between aptamer no. 21 and eIF4A is 50 pM or less, one or two orders of magnitude lower (i.e. higher affinity) than that of an antibody (A. Oguro, T. Ohtsu and Y. Nakamura, unpublished data). Aptamer no. 21 consists of a 5′-domain and a 3′-domain connected by a short linker. Site-directed mutation and truncation experiments revealed that a 58 nt core structure of this aptamer is required for efficient recognition of eIF4A (5). Furthermore, the aptamer binds only to an intact eIF4A, and not to either domain when split at the linker of the 'dumbbell' structure (5). Therefore, the aptamer was suggested to 'staple' together the two domains of eIF4A, leading to inhibition of the inter-domain movement coupled to ATP hydrolysis and helicase action.

In this study, we examined the solution structure of APT58 (58 nt long), a variant of the aptamer no. 21 core structure that sustained 4 nt substitutions (G3→A, C6→U, G10→A and C12→U) for NMR analysis optimization while still retaining high affinity to eIF4A (Figure 1). Here, we present two NMR structures of a 15mer RNA (APT15) containing a hexanucleotide loop and a 16mer RNA (APT16) containing a
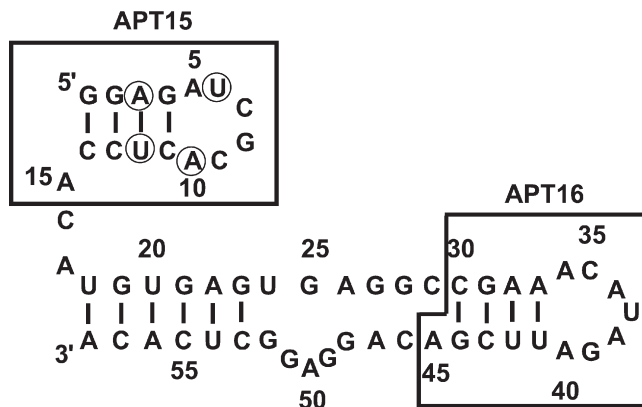

Figure 1. The secondary structure of APT58. Two hairpin derivatives used for NMR analysis, APT15 and APT16, are outlined. The four substituted residues are indicated by open circles: G3→A, C6→U, G10→A and C12→U.

hepta-nucleotide loop found in APT58 (see Figure 1). Mutational analysis has indicated that each loop of the aptamer is crucial for high affinity to eIF4A. By comparing the NMR spectra of the isolated hairpin with the 58mer aptamer, the structural similarity of isolated hairpin loops to the hairpin loops in the aptamer context was confirmed. Thus, the determination of the structures of these localized regions, APT15 and APT16, will enhance our understanding of the interaction between the aptamer and eIF4A.

MATERIALS AND METHODS

RNA synthesis and purification

APT15 was synthesized by *in vitro* transcription reaction using T7 RNA polymerase. APT16 was chemically synthesized by a phosphoramidite method, using an automatic DNA/RNA synthesizer, Expedite Model 8909 (PerSeptive Biosystem, Inc.). Ammonia and tetra-*n*-butylammonium fluoride deprotection, PAGE purification under denaturing conditions and extensive desalting by ultrafiltration (Centricon YM-3, Millipore Inc.) were carried out. A 20 residue RNA, APT20, that consists of APT16 and additional two G:C base pairs for *in vitro* transcription reaction using T7 RNA polymerase was designed for the heteronuclear NMR experiments. ^{13}C, ^{15}N-double-labeled APT20 was purchased from Nippon Sanso Corporation. The unlabeled APT15 was dissolved in 10 mM sodium phosphate buffer (pH 6.5). The unlabeled APT16 and the ^{13}C, ^{15}N-labeled APT20 were dissolved in 20 mM sodium phosphate buffer (pH 6.5) containing 50 mM NaCl. Sample concentrations were 0.5 mM for unlabeled APT15/APT16, and 1 mM for the ^{13}C, ^{15}N-labeled APT20, respectively. Finally, the samples were annealed by heating at 90°C for 5 min and snap-cooling on ice. To confirm the hairpin structure, the samples were subjected to native PAGE analysis before and after NMR experiments.

Analysis of NMR spectra

NMR spectra were measured on Bruker DRX-600 and DRX-500 spectrometers. Spectra were recorded at 7–15°C and NMR data at 10°C were used for structure calculation.

Resonance assignments of APT16 followed the well-established procedures including heteronuclear NMR experiments (33). 2D HNN-COSY experiments were employed to establish base-pairing schemes (34). HCCH-COSY and HCCH-TOCSY experiments were used to assign sugar spin systems (35), while through-backbone assignments were made with 2D HP-COSY (36) and HCP experiments (37). H2 protons of adenosines were assigned using HCCH-TOCSY and 2D HSQC experiments (38). Resonance assignments of APT15 followed the NOE-based procedures not including heteronuclear NMR experiments. NOE distance restraints from non-exchangeable protons were obtained from 2D NOESY experiments (50, 100 and 400 ms of mixing times) in D_2O (39). Exchangeable proton NOEs were determined by 2D NOESY in H_2O with mixing time of 150 ms using the jump-and-return scheme and the gradient pulses for water suppression. Dihedral restraints were obtained from 2D TOCSY (40), 2D DQF-COSY (41), 2D HP-COSY and HCP experiments, as described below.

NOE intensities from exchangeable protons were interpreted as distances of strong (0–3.5 Å) or weak (0–6 Å), while NOE intensities from non-exchangeable protons were interpreted as distances of strong (0–3 Å), medium (0–4 Å), weak (0–5 Å) or very weak (0–7 Å). In order to estimate the δ (C5′–C4′–C3′–O3′) dihedral angle, sugar pucker was analyzed using 2D TOCSY and 2D DQF-COSY spectra. A large $^3J_{H1'-H2'}$ coupling constant and a small $^3J_{H3'-H4'}$ coupling constant indicate the C2′-*endo* conformation ($\delta = 160° \pm 30°$), whereas a small $^3J_{H1'-H2'}$ coupling constant and a large $^3J_{H3'-H4'}$ coupling constant correspond to the C3′-*endo* conformation ($\delta = 85° \pm 30°$). Restraints for the C3′-*endo* conformation were used for G2-A5 and C11-C14, and no δ angle restraint was used for other nucleotides for APT15. Restraints for the C3′-*endo* conformation were used for G31–C35 and U41–G44, a restraint for the C2′-*endo* conformation was used for U37 and no δ angle restraint was used for other nucleotides for APT16. Hydrogen bonding restraints for Watson–Crick base pairs were introduced as distance restraints between protons and heavy atoms (1.8–2.5 Å).

Structure calculation

A set of 100 structures was calculated using a simulated annealing protocol with the InsightII/Discover package. The amber force field was used. A total of 155 NOE distance restraints, 45 dihedral restraints, 11 hydrogen bondings and 4 base planarity restraints were used for APT15. A total of 228 NOE distance restraints, 47 dihedral restraints, 13 hydrogen bondings and 5 base planarity restraints were used for APT16. The force constants were 50 kcal mol^{-1} Å^{-2} for distance restraints and 120 kcal mol^{-1} rad^{-2} for dihedral restraints. The randomized structures were heated to 1000 K during 5 ps. Distance and dihedral restraints were gradually scaled to full value during 15 ps of molecular dynamics, while maintaining a low value for interatomic repulsion, which was subsequently increased to full value during another 20 ps of dynamics. Then, the temperature was gradually scaled to 300 K during 10 ps. A final minimization step was performed, which included a Lennard-Jones potential, but no electrostatic terms. The structures were further refined through a second

simulated annealing protocol. After the structure was heated to 1000 K during 10 ps, all restraints and interatomic repulsion energy term were rescaled to 1/10 of full value. All restraints were gradually scaled to full value during 5 ps of molecular dynamics, while maintaining low value for interatomic repulsion, which was subsequently increased to full value during another 5 ps of dynamics. An additional 5 ps of dynamics were performed at 1000 K and the temperature was gradually scaled to 300 K during 10 ps. A final minimization step was performed, which included a Lennard-Jones potential, but no electrostatic terms. The 20 final structures that had the lowest total energy were chosen.

RESULTS AND DISCUSSION

Structure determination of APT15 and APT16

NMR signals were assigned using the standard method involving heteronuclear experiments (33). The assignment of imino proton signals confirmed the formation of the expected 4 bp in the stem including closing base pair, G4:C11 for APT15 and A33:U41 for APT16, respectively (Figure 2A). Furthermore, the imino proton signal of G39 was observed owing to the slow-exchange of the imino proton with solvent protons. NOE between amino protons of C35 and imino proton of G39 indicates a Watson–Crick base pair C35:G39. The signals at 10.68 and 11.18 p.p.m. in the spectrum of APT15 were assigned to G8 and U6, respectively. The only signal at 10.68 p.p.m. was observed and the signal at 11.18 p.p.m. was not observed in the mutant substituting the ACCGCA loop for the AUCGCA loop (data not shown). Furthermore, intra-residue NOE between the imino proton signal at 10.68 p.p.m. and amino protons suggested that the signal was derived from G8 (data not shown).

For APT15, all of the non-exchangeable base protons (H2, H5, H6 and H8), H1′ and H2′ as well as the H3′ and H4′ protons of G1–A10 (the loop residues) were assigned. For APT16, all of the non-exchangeable base protons, H1′, H2′ and H3′ as well as the H4′ protons of A34–U41 (the loop residues) were assigned. Because the length of APT15 was short, overlapping of NMR signals was avoided, which resulted in reliable assignments of the signals and sufficient information for the structural determination. Although several broadened signals (e.g. H5, H6 protons of C9 and C35 as well as the H1′ proton of G39) indicate conformational flexibility in the loop, NOE connectivity in the D_2O NOESY spectrum showed that the stem has the expected A-form geometry, and the hexa-nucleotide loop and hepta-nucleotide loop regions form well-defined structures (Figure 2B and C). Several unusual protons were observed in the ACAUAGA loop (A34–A40). The H1′ signal of G39 and the H5 signal of C35 were observed at 5.13 and 4.67 p.p.m., respectively. Sugar protons of U37 showed upfield chemical shifts (3.96 p.p.m. for H2′, 3.83 p.p.m. for H4′ and 3.50 p.p.m. for either H5′ or H5″). U6–C7 and A34–C35 phosphate groups have ^{31}P signals that are shifted upfield (-0.12 and -0.92 p.p.m., respectively) compared with the envelope of A-form phosphorus chemical shifts.

The absence of crosspeaks between H1′–H2′ in 2D TOCSY and 2D DQF-COSY experiments implied that the nucleotides in the stem, the A5 in the AUCGCA loop, as well as the A34
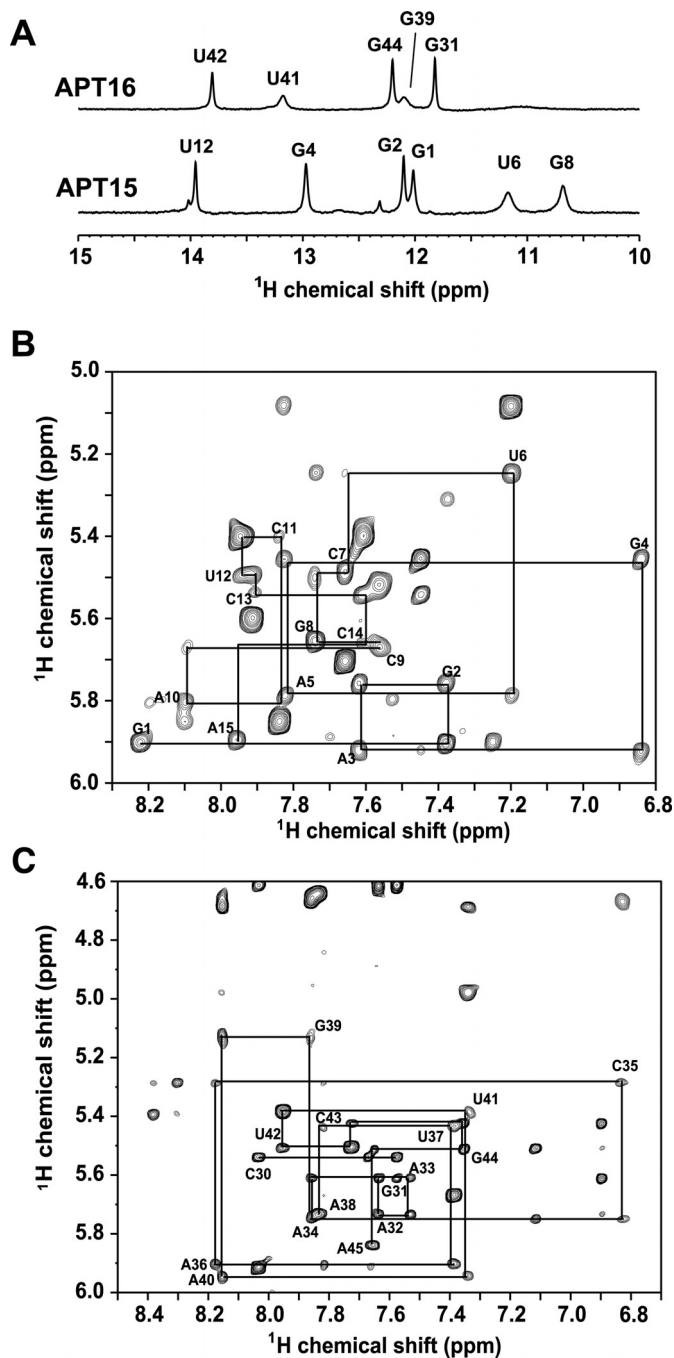
Figure 2. NMR spectra of APT15 and APT16. (**A**) Spectra of exchangeable protons at 10°C. Imino proton assignments are denoted with residue numbers. (**B**) 2D NOESY spectrum of the APT15 in D$_2$O. (**C**) 2D NOESY spectrum of the APT16 in D$_2$O. The NOESY spectra (mixing time = 400 ms) were recorded at 10°C showing crosspeaks between aromatic H6 and H8 protons and ribose H1′ protons. Sequential connectivities are indicated by lines and intra-residue NOEs are labeled by residues.

and C35 in the ACAUAGA loop are in the C3′-*endo* conformations. For the ACAUAGA loop, U37 was found predominantly in the C2′-*endo* conformation, A38 is in equilibrium between the C2′-*endo* and C3′-*endo* conformations, as estimated based on $^3J_{H1′-H2′}$ and $^3J_{H3′-H4′}$ as described previously (42). Sugar puckers of U6, C7, G8, C9 and A10 in the AUCGCA loop as well as A36, G39 and A40 in the

ACAUAGA loop could not be estimated because of overlapping crosspeaks between H3′–H4′. From HCP experiments, a pattern of large $^3J_{C2′-P}$ and weak $^3J_{C4′-P}$ was observed for A34 and U37, indicating ε dihedral angle (C4′–C3′–O3′–P) as −*gauche*. The opposite pattern (weak $^3J_{C2′-P}$ and large $^3J_{C4′-P}$) was observed for other nucleotides, indicating ε *trans*. Moreover, a strong NOE between the H1′ and the H8 base proton of A38 indicates the A38 base has a *syn* conformation rather than the usual A-form *anti* conformation.

The structures of APT15 and APT16 were determined using a standard methodology (33). For APT15, a total of 155 NOE and 45 dihedral angle restraints were obtained from NMR data. For APT16, 228 NOE and 46 dihedral angle restraints were obtained. Structures were calculated using restrained molecular dynamics in a simulated annealing protocol. A total of 20 structures for each of APT15 and APT16 converged to low total energy (Figure 3). The overall structures are well defined by the NMR data with a heavy atom root-mean-square deviation (r.m.s.d.) of 0.87 Å for the 20 structures of both APT15 and APT16. NMR restraints and structural statistics are summarized in Table 1. The final minimizations, which included electrostatic terms with a dielectric constant of 7, were also performed in this study (43,44). The refined structure with electrostatic terms was similar to those without electrostatic terms. The pairwise r.m.s.d. between the two (with and without electrostatic terms) minimized averaged structures were 0.88 for APT15 and 0.96 for APT16.

Structure of the AUCGCA loop

The local structure of the AUCGCA loop (G4–C11) is well defined in the calculated structure of APT15 (heavy atoms r.m.s.d. was 0.52 Å), which contains a non-canonical A:A base pair and a U-turn (Figure 3B). The A5:A10 base pair is categorized as a single hydrogen bond mismatch: Hoogsteen/Sugar-edge [(45), nomenclature proposed by N.B. Leontis and E. Westhof]. The N3 of A5 forms a hydrogen bond with the N6 amino group of A10 (Figure 4A), which is consistent with an NOE between H2 of A5 and H8 of A10. This type of A:A base pair has been observed in 16S rRNA (46,47), a group I intron (48), a hairpin ribozyme (49) and an AGAA tetra-nucleotide loop, which is the recognition site of *Saccharomyces cerevisiae* RNase III (Rnt1p) (50).

Three nucleotides, G4–U6, are continuously stacked, and a sharp turn in the phosphodiester backbone occurs between U6 and C7. The 5 nt, C7–C11, are continuously stacked. These are consistent with the NOE connectivity for H8(G4)–H8(A5), H8(A5)–H6(U6), H6(C7)–H8(G8), H8(G8)–H6(C9), H6(C9)–H8(A10) and H8(A10)–H6(C11). Furthermore, consistencies with a U-turn motif include (i) an upfield ^{31}P signal of the U6–C7 linkage, (ii) non-sequential NOE between H1′ of U6 and H8 of G8 and (iii) non-sequential NOE between H1′ of U6 and H5 of C9.

The U-turn is a common structural motif observed in several RNA structures and is primarily involved in the interaction with protein or RNA. Salient features of the U-turn motif include reversal of the phosphodiester backbone following the pivotal U-residue and two specific cross-loop hydrogen bonds. The calculated structure shows that the sharp foldback of the phosphodiester backbone is stabilized by the U-turn, in which the N3 of U6 forms a hydrogen bond with the phosphate
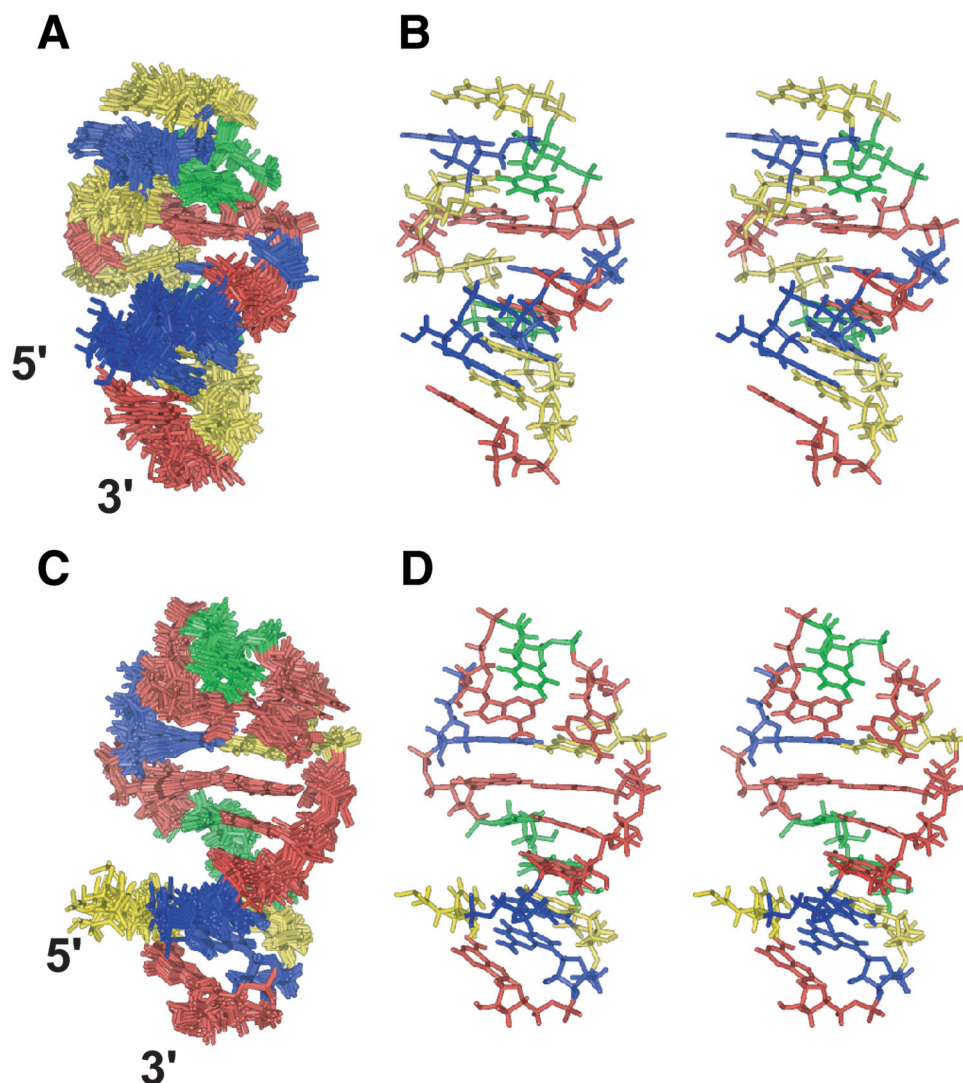
Figure 3. Tertiary structures of APT15 and APT16 solved by NMR. (**A**) A superposition of the final 20 structures of APT15. (**B**) Stereo view of the minimized average structure of APT15. (**C**) A superposition of the final 20 structures of APT16. (**D**) Stereo view of the minimized average structure of the APT16. G, A, U and C residues are colored in blue, red, green and yellow, respectively.

between G8 and C9, and the 2′ OH of U6 forms a hydrogen bond with the N7 of G8. The U-turn motif was first observed in the anticodon loop of tRNA[Phe] and functions to stabilize the stacked conformation of the anticodon (51), although the loop is a hepta-nucleotide loop and therefore different in length from the AUCGCA. The U-turn structures of the stem–loop IIa in yeast U2 snRNA (GUAACA loop) (52) and the A-rich loop in HIV-1 RNA (GUAAAA loop) (43) were determined previously. Although those folds form a U-turn and contain a sheared G:A pair (*Trans* Hoogsteen/Sugar-edge) instead of the A:A base pair of the AUCGCA loop, the overall folds are the same (Figure 5). Although the A:A base pair is unstable compared with a G:A pair and should destabilize the U-turn conformation, the U-turn structure is preserved in the AUCGCA loop. Since both the non-canonical base pairs belong to the Hoogsteen/Sugar-edge type and have the same geometry, it is expected that replacement of the G:A base pair with the A:A base pair would maintain the overall fold of U-turn motif.

Structure of the ACAUAGA loop

The local structure of the ACAUAGA loop (A33–U41) is well defined in the calculated structure of APT16 (heavy atoms r.m.s.d. was 0.88 Å), which contains a non-canonical A:A base pair, a Watson–Crick G:C base pair and an AUA tri-nucleotide loop (Figure 3D). The A34:A40 base pair is categorized to be a single hydrogen bond mismatch: Watson–Crick/Watson–Crick (Figure 4B). The N6 amino group of A34 forms a hydrogen bond with the N1 of A40, which is consistent with an NOE between the H2 of A34 and the H2 of A40. This type of A:A base pair has been observed in rRNA (46,47,53). The hepta-nucleotide loop domain is compact, in which the A38 base is inside the loop, while the A36 base is pushed out in solution. The A40, G39, A38 and U37 nucleotides are partially stacked in a consecutive fashion. The structure qualitatively explains the unusual chemical shifts observed in the loop. The H5 and H6 signals of C35 are shifted upfield slightly by the ring current of the adenosine base in

A36 while the H1′ signal of G39 is shifted upfield slightly by the ring current of the adenosine base in A40. The ribose ring of U37 is located on top of the A38 base. This is consistent with the observation that the sugar proton signals of U37 were shifted upfield.

The structure of the AUA in the ACAUAGA loop is similar to the AUA tri-nucleotide loop structure found in the stem–loop C of brome mosaic virus (BMV) RNA, which is recognized by the viral replicase (54–56) (Figure 6). As in the case with the BMV RNA, a strong NOE between the H1′ and the H8 base proton of the third adenosine residue in the AUA tri-nucleotide loop (A38 in APT16) was observed. Similar to the case of the BMV RNA, the calculated structure shows that the AUA tri-nucleotide loop is stabilized by the formation of a hydrogen bond between the N6 amino group of A36 and the phosphate between A34 and C35. Thus, APT16 is

suggested to consist of an AUA tri-nucleotide loop and a stem containing the A:A mismatch.

An RNA melting study showed that APT16 is thermodynamically unstable ($T_m = 46°C$) compared with the loop structure of the BMV RNA (13mer, GGUGCAUAGCACC, $T_m = 76°C$). The calculated structure revealed that A:A base pair widens the inter-strand distance (the distance between C4′ of A34 and that of A40 is 17.3 Å) compared with a Watson–Crick base pair in the A-form helical stem (about 15 Å). This distortion of the A-form helical structure results in destabilization of the stem–loop structure in APT16.

Functional relevance of local structural information of aptamer

This study solved the solution structure of two loops, AUCGCA (residues 5–10) and ACAUAGA (residues 34–40) that are essential domains for high affinity aptamer (APT58) to eIF4A. The original selection scheme was performed in the presence of 2.5 mM magnesium acetate, whereas the structures were solved in the absence of Mg^{2+} ions. Although the addition of Mg^{2+} ions to APT16 causes broadening of NMR signals, the chemical shift changes of the signals were scarcely observed (data not shown), suggesting that the folding of APT16 was not changed appreciably upon the addition of Mg^{2+} ions. On the other hand, the addition of Mg^{2+} ions to APT15 caused extreme broadening of NMR signals, and induced notable changes in C7 and G8 signals. These changes are, at least in part, explained by assuming that the addition of Mg^{2+} ions induces a kissing-loop type dimer conformation for APT15. Thus, the structures of APT16 and APT15 were determined in the absence of Mg^{2+} ions.

Substitution of guanosine residue for A36 did not diminish the binding affinity to eIF4A (5). Thus, we analyzed NMR spectra of the ACGUAGA loop variant, and confirmed that the conformation of the variant is similar to that of APT16 (data not shown). It is consistent with the fact that the conformation of the GUA tri-loop is very similar to the tri-loop conformation of the AUA tri-loop in BMV RNA (56). Either the substitution of a G:C base pair for the A:A base pair in the ACAUAGA loop (see Figure 6) or the substitution of the

Table 1. NMR restraints and structural statistics

	APT15	APT16
Number of experimental restraints		
Distance restraints	155	228
Dihedral restraints	45	46
Hydrogen bonding distance restraints	11	13
Base planarity restrains	4	5
r.m.s.d from restraints[a]		
Distance restraints (Å)	0.002 ± 0.001	0.012 ± 0.002
Dihedral restraints (°)	0.2 ± 0.1	0.5 ± 0.1
r.m.s.d. from idealized geometry		
Bonds (Å)	0.0026 ± 0.0002	0.0032 ± 0.0001
Angle (°)	0.58 ± 0.01	0.62 ± 0.01
Impropers (°)	0.443 ± 0.003	0.436 ± 0.002
Heavy-atoms r.m.s.d from mean structure (Å)		
All	0.87	0.87
Loop (G4–C11)	0.52	
Loop (A33–U41)		0.88

[a]The converged structures did not contain experimental distance violation of >0.2 Å or dihedral violation >5°.
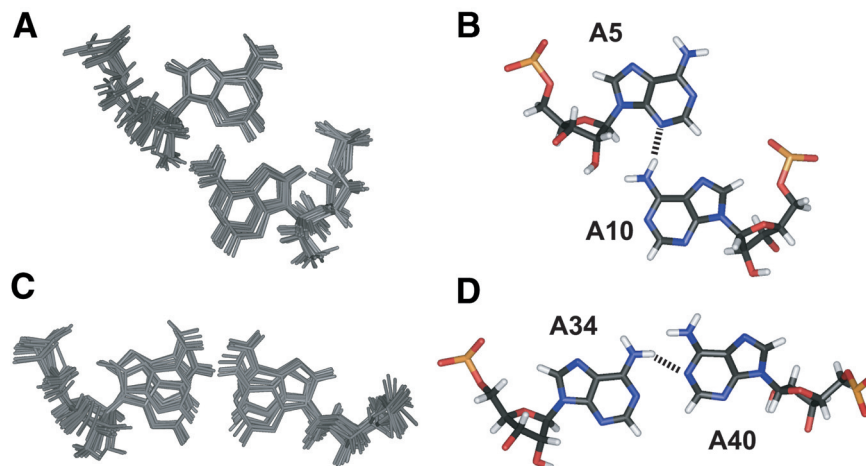

Figure 4. A local superposition of the final 20 structures (**A**) and schematic representation (**B**) of non-canonical A:A base pairs in APT15. A local superposition of the final 20 structures (**C**) and schematic representation (**D**) of non-canonical A:A base pairs in APT16.
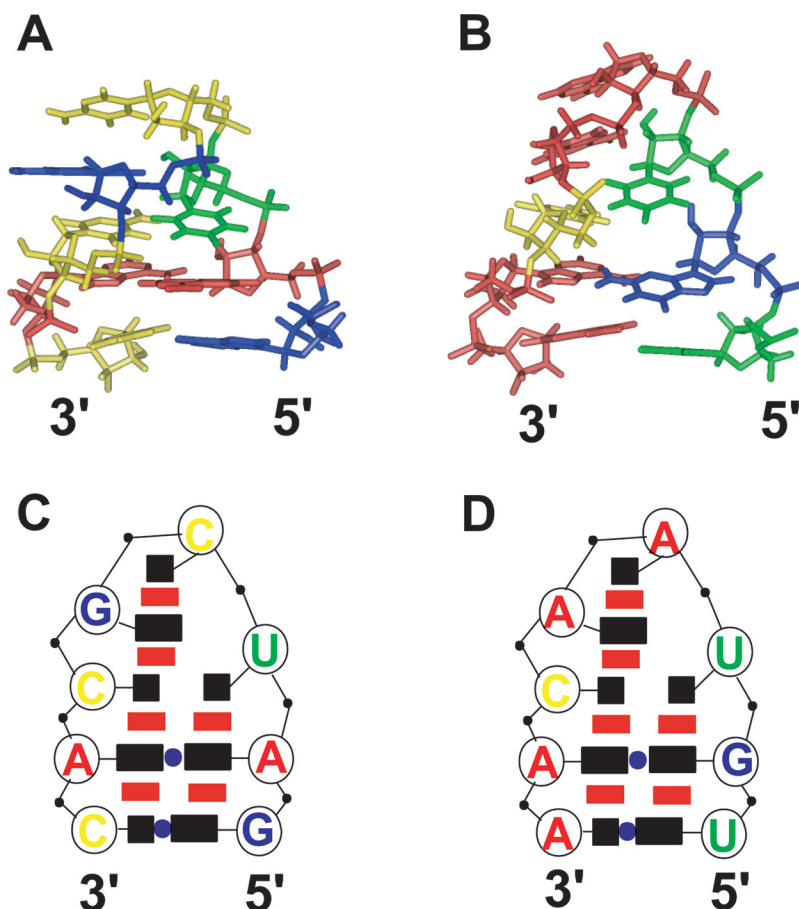
Figure 5. Structural comparison of the AUCGCA loop of APT15 and the stem–loop IIa in yeast U2 snRNA. (**A**) The AUCGCA loop structure in APT15. (**B**) The GUAACA loop structure in loop IIa of U2 snRNA (38). Schematic representations of the (**C**) AUCGCA loop structure and (**D**) GUAACA loop structure are shown. In the schematic representations, symbols are used as follows: black rectangle, *anti* base; open rectangle, *syn* base; red rectangle, stacking interaction; blue circle, hydrogen bonding interaction; open circle, C3′-*endo* ribose; open square, C2′-*endo* ribose.

entire loop for the stable UUCG tetra-nucleotide loop, weakly but significantly, diminished the binding affinity to eIF4A (data not shown). Therefore, we suggest that the flexible nature of the ACAUAGA loop may be important for aptamer binding to eIF4A, although the loop may not be directly involved in the interaction with eIF4A.

The APT15 was constructed for the NMR study by replacing the 4 nt in the original 15mer hairpin RNA. Initially, we tried to determine the original 15mer RNA (5′-GGGGACCGCGCCCCA-3′) derived from the aptamer no. 21. NMR and native PAGE analyses of the 15mer RNA suggested that the ACCGCG loop forms a kissing-loop type homodimer. To avoid difficulties in the structural determination due to multi-conformations of RNA molecule, we replaced G10 of the 15mer RNA with adenosine. The non-canonical A:A base pair was observed in APT15, whereas a similar base pair cannot be formed in the original 15mer RNA. Although the ACCGCG loop seems to be less stable due to a lack of the A:A base pair, stacking interaction between G10 and the contiguous residues may stabilize the loop conformation. In the U-turn motif, the uridine residue is thought to be important to stabilize the conformation. Thus, we tried to analyze NMR spectra of the loop variant (ACCGCA), which shows high affinity for eIF4A. Although the structure

of the ACCGCA loop could not be determined due to overlapping and broadening of NMR signals, NOE connectivity for H8(G8)–H6(C9), H6(C9)–H8(A10) and H8(A10)–H6(C11) indicated that 3 nt G8–C11 are continuously stacked in the variant (data not shown). Non-sequential NOE between H1′ of C6 and H8 of G8 was observed in the variant as shown in APT15. Furthermore, NOE signal between H2 of A5 and H8 of A10 indicated that the similar A:A base pair as APT15 is formed in the variant. Therefore, we assume that the ACCGCA loop forms a similar conformation to U-turn motif and that the U-turn like conformation could be important for eIF4A binding.

The U-turn structure of the AUCGCA loop results in solvent accessibility of the Watson–Crick edges of C7, G8 and C9, which should allow base specific interaction with eIF4A. In fact, substitutions of A for C7 and U for C9 diminish eIF4A binding; however, substitution of A for G8 preserves most of the binding affinity (5). Thus, the Watson–Crick edges of C7 and C9 are thought to be recognized by eIF4A. Furthermore, NMR analysis and native PAGE analysis of the loop variant (AUCGCG) suggested that the AUCGCG loop forms a kissing-loop type homodimer (data not shown) such that the Watson–Crick edges of C7, G8, C9 and G10 may allow the kissing loop–loop interaction. Similarly, the U-turn structure
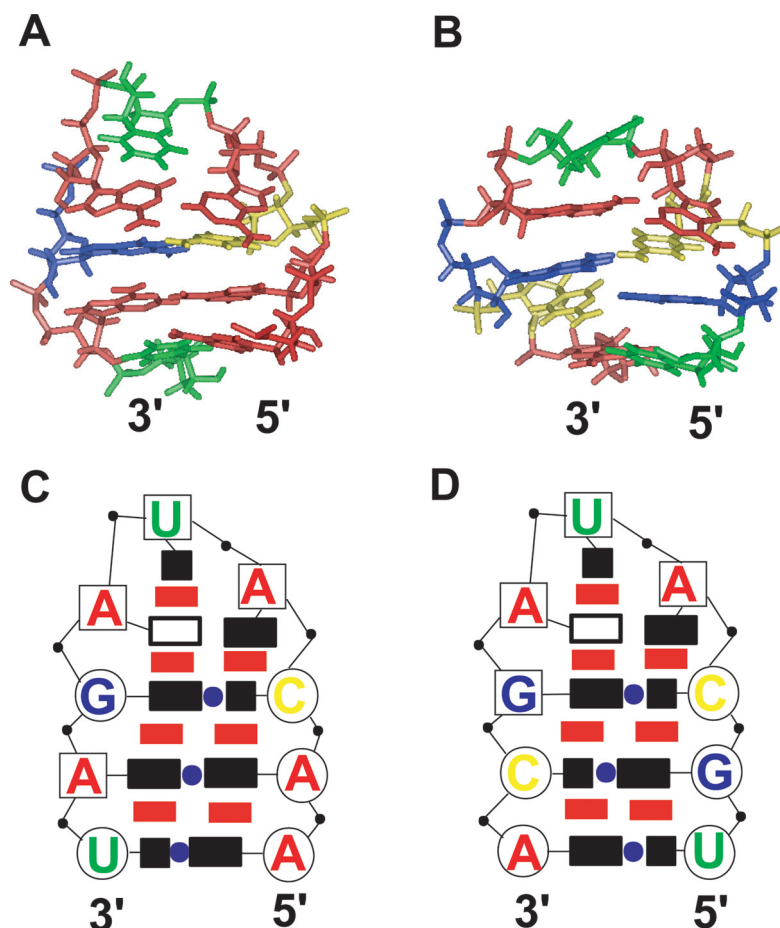
Figure 6. Structural comparison of the AUA tri-nucleotide loops in (**A**) APT16 and (**B**) loop C of BMV RNA (41). Schematic representation of the (**C**) ACAUAGA loop structure and (**D**) the AUA loop structure are shown. Symbols used are the same as Figure 5.

of the stem–loop IIa in U2 snRNA has been proposed to interact with a downstream sequence whose complementarity to the loop sequence is phylogenetically conserved (52). The U-turn structure of the A-rich loop was also suggested to form base pairs with four uridines in the anticodon loop of $tRNA_3^{Lys}$ to assist in tRNA binding to viral RNA, which is required for reverse transcription of viral RNA (43). Elucidation of detailed interactions between U-turn motifs and proteins or RNAs will enhance our understanding of the molecular and structural basis of protein–RNA interactions.

It is noteworthy that the minimum length of efficient aptamer binding to eIF4A is at least 58 nt. Aptamers selected to other initiation factors, such as eIF4E (6), eIF4G and eIF1A (A. Oguro, S. Miyakawa, T. Ohtsu and Y. Nakamura, manuscript in preparation), were also large aptamers over 70 nt long, with most of these nucleotides essential for high affinity binding to target proteins. Although these eIFs play critical roles in translation initiation, they do not contain any RNA-recognition motifs and do not exhibit strong binding to RNA. Therefore, it might be argued that selected aptamers to proteins possessing only weak, if any, affinity to RNA require the ability to capture the protein's global conformation. This may be the reason why RNA aptamers against these proteins are large. This is completely different from RNA aptamers against RNA recognition motifs, which acquire pinpoint affinity. In a more general context, this type of global recognition of target conformation is in sharp contrast to specific epitope (<10 amino acids) recognition by antibodies. The complete understanding of aptamer–eIF4A interactions requires structure determination of the aptamer–eIF4A complex at atomic resolution. Nevertheless, the present structural study of the local conformations of two functionally essential loops in APT58 will provide us with an important step toward the understanding of the global conformation recognition by RNA.

ACKNOWLEDGEMENTS

REFERENCES

1. Ellington,A.D. and Szostak,J.W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.
2. Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.
3. Hermann,T. and Patel,D.J. (2000) Adaptive recognition by nucleic acid aptamers. *Science*, **287**, 820–825.
4. Cerchia,L., Hamm,J., Libri,D., Tavitian,B. and de Franciscis,V. (2002) Nucleic acid aptamers in cancer medicine. *FEBS Lett.*, **528**, 12–16.
5. Oguro,A., Ohtsu,T., Svikin,Y.V., Sonenberg,N. and Nakamura,Y. (2003) RNA aptamers to initiation factor 4A helicase hinder cap-dependent translation by blocking ATP hydrolysis. *RNA*, **9**, 394–407.
6. Mochizuki,K., Oguro,A., Ohtsu,T., Sonenberg,N. and Nakamura,Y. (2005) High affinity RNA for mammalian initiation factor 4E interferes with mRNA-cap binding and inhibits translation. *RNA*, **11**, 77–89.
7. Hershey,J.W.B. and Merrick,W.C. (2000) Pathway and mechanism of initiation of protein synthesis. In Sonenberg,N., Hershey,J.W.B. and Mathews,M.B. (eds), *Translational Control of Gene Expression.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 33–88.
8. Gingras,A.-C., Raught,B. and Sonenberg,N. (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.*, **68**, 913–963.
9. Sonenberg,N. and Dever,T.E. (2003) Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.*, **13**, 56–63.
10. Shatkin,A. (1976) Capping of eukaryotic mRNAs. *Cell*, **9**, 645–653.
11. Duncan,R., Milburn,S.C. and Hershey,J.W.B. (1987) Regulated phosphorylation and low abundance of HeLa cell initiation factor 4F suggest a role in translational control. *J. Biol. Chem.*, **262**, 380–388.
12. Rozen,F., Edery,L., Meerovitch,K., Dever,T.E., Merrick,W.C. and Sonenberg,N. (1990) Bidirectional RNA helicase activity of eukaryotic translation initiation factors 4A and 4F. *Mol. Cell. Biol.*, **10**, 1134–1144.
13. Tanner,N.K. and Linder,P. (2001) DExD/H box RNA helicases: from genetic motors to specific dissociation functions. *Mol. Cell*, **8**, 251–262.
14. Linder,P. (2003) Yeast RNA helicases of the DEAD-box family involved in translation initiation. *Biol. Cell*, **95**, 157–167.
15. Tanner,N.K., Cordin,O., Banroques,J., Doere,M. and Linder,P. (2003) The Q motif. A newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell*, **11**, 127–138.
16. Venema,J. and Tollervey,D. (1995) Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. *Yeast*, **11**, 1629–1650.
17. Py,B., Higgins,C.F., Krisch,H.M. and Carpousis,A.J. (1996) A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature*, **381**, 169–172.
18. Staley,J.P. and Guthrie,C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, **92**, 315–326.
19. Benz,J., Trachsel,H. and Baumann,U. (1999) Crystal structure of the ATPase domain of translation initiation factor 4A from *Saccharomyces cerevisiae*—the prototype of the DEAD box protein family. *Structure Fold Des.*, **7**, 671–679.
20. Johnson,E.R. and McKay,D.B. (1999) Crystallographic structure of the amino terminal domain of yeast initiation factor 4A, a representative DEAD-box RNA helicase. *RNA*, **5**, 1526–1534.
21. Caruthers,J.M., Johnson,E.R. and McKay,D.B. (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl Acad. Sci. USA*, **97**, 13080–13085.
22. Lorsch,J.R. and Herschlag,D. (1998) The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry*, **37**, 2194–2206.
23. Lorsch,J.R. and Herschlag,D. (1998) The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry*, **37**, 2180–2193.
24. De Benedetti,A. and Rhoads,R.E. (1990) Overexpression of eukaryotic protein synthesis initiation factor 4E in HeLa cells results in aberrant growth and morphology. *Proc. Natl Acad. Sci.*, **87**, 8212–8216.
25. Lazaris-Karatzas,A., Montine,K.S. and Sonenberg,N. (1990) Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5′ cap. *Nature*, **345**, 544–547.
26. Ruggero,D., Montanaro,L., Ma,L., Xu,W., Londei,P., Cordon-Cardo,C. and Pandolfi,P. (2004) The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.*, **10**, 484–486.
27. Avdulov,S., Li,S., Michalek,V., Burrichter,D., Peterson,M., Periman,D.M., Manivel,J.C., Sonenberg,N., Yee,D., Bitterman,P.B. and Polunovsky,V.A. (2004) Activation of translation complex eIF4F is essential for the genesis and maintenance of the malignant phenotype in human mammary epithelial cells. *Cancer Cell*, **5**, 553–563.
28. Wendel,H.-G., de Stanchina,E., Fridman,J.S., Malina,A., Ray,S., Kogan,S., Cordon-Cargo,C., Pelletier,J. and Lowe,S.W. (2004) Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature*, **428**, 332–337.
29. Kerekatte,V., Smiley,K., Hu,B., Smith,A., Gelder,F. and De Benedetti,A. (1995) The proto-oncogene/translation factor eIF4E: a survey of its expression in breast carcinomas. *Int. J. Cancer*, **64**, 27–31.
30. Nathan,C.A., Liu,L., Li,B., Nandy,I., Abreo,F. and De Benedetti,A. (1997) Detection of the proto-oncogene eIF4E in surgical margins may predict recurrence in head and neck cancer. *Oncogene*, **15**, 579–584.
31. Eberle,J., Krasagakis,K. and Orfanos,D.E. (1997) Translation initiation factor eIF-4A1 mRNA is consistently overexpressed in human melanoma cells *in vitro*. *Int. J. Cancer*, **71**, 396–401.
32. Eberle,J., Fecker,L.F., Bittner,J.U., Orfanos,C.E. and Geilen,C.C. (2002) Decreased proliferation of human melanoma cell lines caused by antisense RNA against translation factor eIF-4A1. *Br. J. Cancer*, **86**, 1957–1962.
33. Allain,F.H.-T. and Varani,G. (1995) Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.*, **250**, 333–353.
34. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
35. Pardi,A. and Nikonowicz,E.P. (1992) Simple procedure for resonance assignment of the sugar protons in ^{13}C labeled RNA. *J. Am. Chem. Soc.*, **114**, 9202–9203.
36. Sklénar,V., Miyashiro,H., Zon,G., Miles,H.T. and Bax,A. (1986) Assignment of the ^{31}P and ^1H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett.*, **208**, 94–98.
37. Varani,G., Aboul-ela,F., Allain,F. and Gubser,C.C. (1995) Novel three-dimensional ^1H–^{13}C–^{31}P triple resonance experiments for sequential backbone correlations in nucleic acids. *J. Biomol. NMR*, **5**, 315–320.
38. Legault,P., Farmer,B.T.,II, , 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.
39. Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Phys. Chem.*, **71**, 4546–4553.
40. Bax,A. and Davis,D.G. (1985) MLEV-17-based 2D homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.*, **65**, 355–360.
41. Rance,M., Sørensen,O.W., Bodenhausen,G., Wagner,G., Ernst,R.R. and Wüthrich,K. (1983) Improved spectral resolution in COSY NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commu.*, **117**, 479–485.
42. Altona,C. and Sundaralingam,M. (1973) Conformational analysis of the sugar ring in nucleotides and nucleosides. Improved method for the Interpretation of proton magnetic resonance coupling constants. *J. Am. Chem. Soc.*, **95**, 2333–2344.
43. Puglisi,E.V. and Puglisi,J.D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. *Nature Struct. Biol.*, **5**, 1033–1036.
44. Lebars,I., Lamontagne,B., Yoshizawa,S., Elela,S.A. and Fourmy,D. (2001) Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. *EMBO J.*, **20**, 7250–7258.
45. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.
46. Carter,A.P., Clemons,W.M., Brodersen,D.E., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407**, 340–348.
47. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905–920.
48. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.

49. Butcher,S.E., Allain,F.H. and Feigon,J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature Struct. Biol.*, **6**, 212–216.

50. Wu,H., Yang,P.K., Butcher,S.E., Kang,S., Chanfreau,G. and Feigon,J. (2001) A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III. *EMBO J.*, **20**, 7240–7249.

51. Quigley,G.J. and Rich,A. (1976) Structural domains of transfer RNA molecules. *Science*, **194**, 796–806.

52. Stallings,S.C. and Moore,P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5**, 1173–1185.

53. Fourmy,D., Yoshizawa,S. and Puglisi,J.D. (1998) Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA. *J. Mol. Biol.*, **277**, 333–345.

54. Kim,C.H., Kao,C.C. and Tinoco,I.,Jr (2000) RNA motifs that determine specificity between a viral replicase and its promoter. *Nature Struct. Biol.*, **7**, 415–423.

55. Kim,C.H. and Kao,C.C. (2001) The mutant viral RNA promoter with an altered conformation retains efficient recognition by a viral RNA replicase through a solution-exposed adenine. *RNA*, **7**, 1476–1485.

56. Kim,C.H. and Tinoco,I.,Jr (2001) Structural and thermodynamic studies on mutant RNA motifs that impair the specificity between a viral replicase and its promoter. *J. Mol. Biol.*, **307**, 827–839.