Nucleic Acids Research, 1999, Vol. 27, No. 4 **949–955**

A′-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC)

Yoshiyuki Tanaka[1,2,*], Satoshi Fujii[2,5], Hidekazu Hiroaki[2,3], Takeshi Sakata[2], Toshiki Tanaka[2,3], Seiichi Uesugi[2,4], Ken-ichi Tomita[2] and Yoshimasa Kyogoku[1]

[1]Division of Molecular Biophysics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565, Japan, [2]Faculty of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565, Japan, [3]Biomolecular Engineering Research Institute, Suita, Osaka 565, Japan, [4]Faculty of Engineering, Yokohama National University, Yokohama, Kanagawa 240, Japan and [5]University of Shizuoka, Shizuoka, Shizuoka 422, Japan

Received December 1, 1998; Accepted December 16, 1998 NDB no. AR0005

ABSTRACT

Here we demonstrate the presence of the A′-RNA conformation using the single crystal structure of a tridecamer: r(UGAGCUUCGGCUC). The average A′-RNA conformation deduced from X-ray fiber diffraction data had only been available previously, but now the presence of the A′-RNA conformation has been found in a single crystal structure for the first time. Statistical analysis showed that the A′-RNA conformation is distinguishable from the A-RNA conformation in a plot of the major groove width against the base pair inclination angle. The major groove of the A′-RNA conformation is wide enough to accommodate a protein or peptide while that of the A-RNA conformation is too narrow to do so. The presence of the A′-RNA conformation is significant for protein–RNA interaction.

INTRODUCTION

The sequence CUUCGG is known to form a thermally extraordinarily stable hairpin loop structure in solution (1–4). However, in a crystal, the tridecamer r(UGAGCUUCGGCUC) forms a double helix with an internal loop of four successive non-Watson–Crick base pairs instead of a monomeric hairpin loop structure, as previously reported (5,6). We have already reported the crystallization and crystal structure of the tridecamer using data at low resolution (7,8) and found that the local geometry of the non-Watson–Crick base pairs was similar to that in previous structures. In this study, we report a newly refined crystal structure based on new data obtained at higher resolution and examined it from the viewpoint of the conformational polymorphism of double helical RNA.

The major groove of nucleic acids has the potential to interact with proteins. As well as many transcriptional regulatory proteins, several proteins or peptides, such as Tat and Rev, bind to the major groove of the double helical regions of their target RNAs (TAR and RRE) (9,10). In such complexes, the conformation of the RNA duplex, especially the major groove width, is an important factor. The major groove of the A-RNA conformation is too narrow to accommodate a protein. In fact, it was observed that Tat and Rev increased the major groove widths of target RNAs upon binding. As in TAR and RRE, in the crystal structure of the 62 nt domain of *Escherichia coli* 5S rRNA, helix IV of a putative ribosomal protein (L25) binding site has a significantly wider major groove (11). Accordingly, it is crucial to determine the conformations of RNA duplexes themselves.

Numerous attempts have been made to clarify the conformational polymorphism of double helical polynucleotides by means of the X-ray fiber diffraction technique (12–16) and solid state ^{31}P NMR of RNA fibers (17). With respect to RNA, there are two major right-handed conformers; one is A-RNA, which has 11 nt in one helical pitch, and the other is A′-RNA, which has 12 nt in one helical pitch (16). However, the A′-RNA conformation is so rare in a single crystal that the conformational difference between A- and A′-RNAs has not been well studied. Furthermore, their conformations are so similar that they cannot be confidently discriminated from each other by means of root mean square deviation (r.m.s.d.). Accordingly, it is necessary to study their conformations in detail and to establish a novel method for classifying closely related but different conformations.

MATERIALS AND METHODS

Crystallization

The sample for crystallization was synthesized by means of the solid-phase phosphoramidite method and purified as described (2). The numbering of the bases and the base pair scheme in a crystal are presented in Figure 1a. Crystals were grown in a sitting drop vessel by means of the vapor diffusion technique at 10°C from a solution of 0.5 mM RNA tridecamer, r(UGAGCUUCGGCUC), 6.7 mM sodium cacodylate buffer (pH 6.5), 1.3 mM sodium phosphate, 13.3 mM NaCl, 3.3 mM $BaCl_2$, 3.7 mM spermidine·3HCl and 0.8% polyethylene glycol #400 (PEG400), against 4% PEG400 as a reservoir. The crystals belong to the monoclinic space group, C2, with one tridecamer per asymmetric unit and cell dimensions of a = 38.49 Å, b = 32.30 Å, c = 38.76 Å and β = 117.56°.

*To whom correspondence should be addressed at: Division of Molecular Biophysics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. Tel: +81 6 6879 8598; Fax: +81 6 6879 8599; Email: tanaka@protein.osaka-u.ac.jp

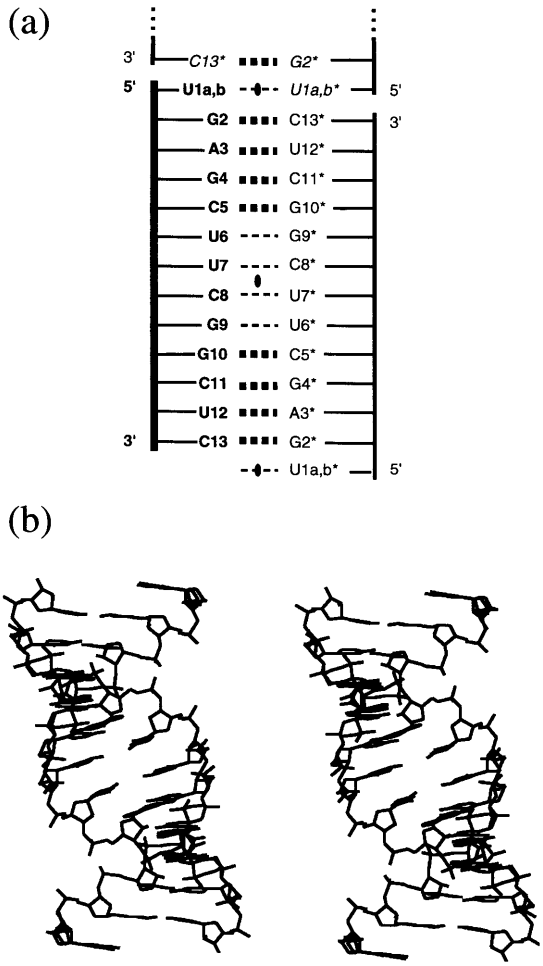

Figure 1. (**a**) The numbering of the bases and the base pair scheme are shown. A crystallographically independent molecule is presented in bold and symmetrically related molecules are presented in normal type with asterisks. The helices piled up along the *c*-axis are also presented in italics. Because of the disordered structure, U1 has two conformers, U1a and U1b, which are denoted as U1a,b for short. The bold and thin broken lines represent Watson–Crick base pairs and non-Watson–Crick base pairs, respectively. The black ovals represent the crystallographic two-fold axis. (**b**) Stereo view of the double-stranded form of the tridecamer.

Data collection

A single well-formed crystal was mounted and sealed in a glass capillary and then intensity data up to 1.8 Å resolution were collected. The numbers of observed reflections and reflections above 0.4σ(F) were 3556 and 3288, respectively. The completeness was 90.1 and 82.6%, respectively. The diffraction intensities were recorded in the θ/ω scan mode at a scan speed of 6°/min and a scan width of 1.2° with a Rigaku AFC5R diffractometer using graphite monochromated CuK_α radiation (λ = 1.5418 Å). The unit cell dimensions were refined with 2θ values of 20 reflections in the range of 16° < θ < 22°.

Structure determination

The crystal structure was solved by the molecular replacement method (18–21), using a single-stranded A-form dodecamer of r(GAGCUUCGGCUC) as the initial model. This gave a unique solution with reasonable packing at the R value of 40.0%. The resulting coordinates were subjected to rigid body refinements, refinements of individual atomic coordinates by the conjugate gradient method (22) and simulated annealing (23,24). Finally, 37 water molecules were introduced. The details are described in the supplementary material (Fig. S1). The final structure gave an R value of 19.8% for the working set of diffraction data [2942 reflections from 10.0 to 1.8 Å with |F| > 0.4σ(F)] and a free R value of 24.1% [318 reflections from 10.0 to 1.8 Å with |F| > 0.4σ(F)]. In all the above refinement stages, the free R values were used (25). As judged from a Luzzati plot (26), the estimated error of atomic coordinates is 0.20 Å. The deviations from the standard bond length and bond angle values are 0.005 Å and 0.90°, respectively. All structure calculations were performed with X-PLOR v.3.851 and improved topology and parameter files (27) were used throughout the refinements. QUANTA96 at the Research Center for Structural Biology was also used for graphics. Final coordinates were deposited in the Nucleic Acid Database (ID AR0005).

Calculation of structure parameters

Helical parameters, groove widths and torsion angles were calculated using the program CURVES (28,29). In Table 1, the mean *x*-displacement, inclination, rise, twist, minor groove width and major groove width for each duplex are listed, but other structure parameters are not. This is because the standard deviations of the omitted parameters for each duplex were much larger than the differences in the structure parameters between conformers, which means that the omitted parameters did not reflect differences in the conformation. The omitted parameters were insufficient for classification of the completely different conformations of the A- and B-form duplexes. They were insufficient for the classification of RNA duplexes as well. Caution must be taken regarding the parameter 'twist', because the standard deviation of 'twist' in each duplex was larger than the difference between conformers. It is, however, included in Table 1 because 'twist' is said to be an important parameter for the classification of nucleic acid conformations. The sequences, and their PDB and NDB ID codes, used for the calculation of the structure parameters are listed in the legend to Table 1. To calculate the major groove width precisely, oligomers of >10 bp were selected. However, oligomers including G-A pairs were eliminated because of their irregular backbone geometries. In the calculation of the structure parameters for the tridecamer, the 5′-terminal uridine at the dangling end was eliminated. Finally, the mean structure parameters of all the oligomers were averaged in terms of the structure parameters (AVERAGE) and then the standard deviation of each structure parameter was calculated (SD). These two parameters are structurally meaningless, but are required to calculate the normal distribution of each mean structure parameter for the following cluster analysis. The torsion angles of the oligomers are listed in Table 2.

Cluster analysis of RNA oligomers

The normal distribution of each mean structure parameter was calculated in terms of the structure parameters using AVERAGE and SD in Table 1 (supplementary material, Table S1). Then the pair-wise distances between the oligomers were calculated using the normal distributions of the structure parameters (supplementary

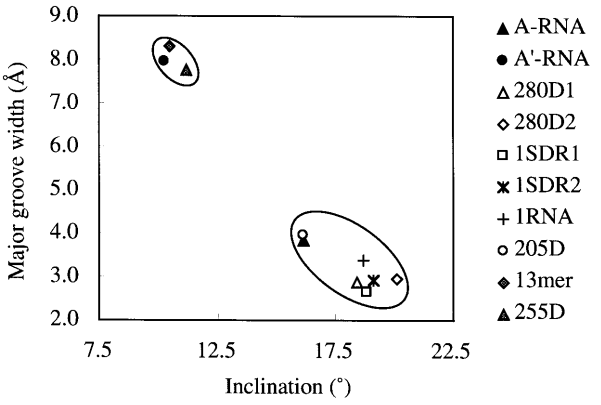

Figure 2. The mean inclination value of each oligomer duplex is plotted against the value of the mean major groove width of each oligomer duplex. Each point represents the corresponding oligomer duplex. The coordinates and names of the oligomers were taken from PDB, using its code. 1SDR and 280D include two duplexes per asymmetric unit, which are named 1SDR1 and 1SDR2 for 1SDR and 280D1 and 280D2 for 280D. The legend to Table 1 gives the names and sequences of the oligomers. A-RNA (fiber) and A′-RNA (fiber) represent the structures obtained from fiber diffraction data.

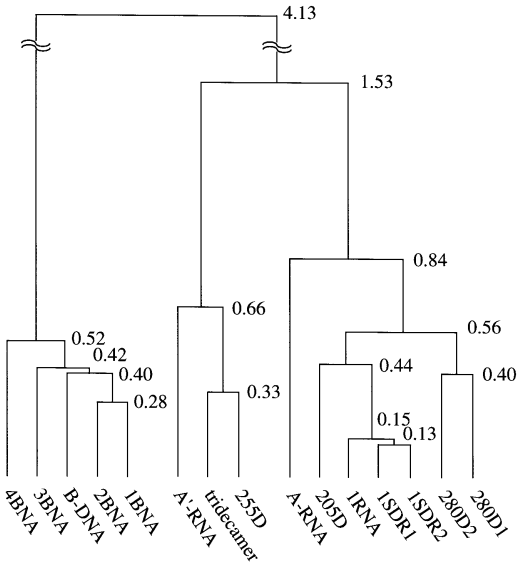

Figure 3. A dendrogram of the cluster analysis results. Oligomers are linked by the nearest neighbor method at distances which are listed in the supplementary material (Table S2) in bold. The legend to Table 1 gives the names and sequences of the oligomers.

material, Table S2). They were clustered by means of the nearest neighbor method and transformed into a dendrogram (Fig. 3). Details are given in the supplementary material (Tables S1 and S2).

RESULTS

Description of the structure

The crystal structure of the tridecamer r(UGAGCUUCGGCUC) contains one strand of the tridecamer and 37 water molecules per asymmetric unit. The 5′-terminal uridine residue adopts two conformations, with an occupancy of 0.5, because of the statistical disorder due to the asymmetric base pair around the crystallographic two-fold axis. The sequence, with the numbers of bases, and the interactions between the symmetrically related molecules are presented in Figure 1a.

Two strands of the tridecamer related by the crystallographic two-fold axis form a double-stranded structure (Fig. 1), as previously reported (5,6), instead of a monomeric hairpin loop structure. As a result, in the middle of the helix, four successive non-Watson–Crick base pairs (two G-U and two C-U pairs) are formed. A $2|F_o| - |F_c|$ map of the G-U pair is presented in the supplementary material (Fig. S1). Nevertheless, throughout the double helix, the right-handed structure is kept not only in the Watson–Crick base pair portion but also in the non-Watson–Crick base pair portion (Fig. 1b). The double helices are piled up along the crystallographic *c*-axis to form a pseudo-continuous double helix through the intermolecular U-U pair which has a similar arrangement to the Hoogsteen-like *trans* U-U pair (30). A detailed description of the U-U pair is given in the supplementary materials (Figs S2 and S3).

The arrangements of the non-Watson–Crick base pairs are similar to the previously reported structures, 255D (5) and 165D (6), which have the same core sequence, ‘CUUCGG’. The r.m.s.d. values of the core sequences in a double-stranded form between the tridecamer and the previous structures (5,6) were calculated. The r.m.s.d. between the tridecamer and 255D and that between the tridecamer and 165D were 1.12 and 1.40 Å, respectively, which indicates that the structure of the tridecamer is close to the previous structures and was refined correctly.

Evidence for the A′-form

The overall structure of the double helix was compared with those of the canonical A-RNA and A′-RNA conformers (11 residues/turn for the A-form and 12 residues/turn for the A′-form) (16). The r.m.s.d. of the coordinates from the tridecamer duplex to A-RNA and A′-RNA are 1.60 and 1.13 Å, respectively. This showed that the conformation of the tridecamer is rather similar to that of A′-RNA. However, the r.m.s.d. values were not very different from each other, so we could not determine the conformation of the tridecamer from these values.

To clarify the conformational properties in detail, we have listed in Table 1 the mean helical parameters, the major and minor groove widths of the tridecamer, together with those of canonical A-RNA, A′-RNA and B-DNA, and the single crystal structures of the oligonucleotides determined in previous works (5,31–37). All the coordinates were taken from PDB and their names represent the PDB codes (the legend to Table 1 gives the sequences and their NDB codes). 1SDR and 280D contain two double helices per asymmetric unit, which are named 1SDR1 and 1SDR2 for 1SDR and 280D1 and 280D2 for 280D. To calculate the major groove widths precisely, oligomers of >10 bp were selected. The parameters which are efficient for the classification of the conformations are listed in Table 1.

As far as the conformation of the RNA duplex is concerned, there are two remarkable features. (i) The major groove widths of the tridecamer and 255D are significantly greater than those of the others. (ii) The inclination angles of the tridecamer and 255D are much smaller than those of the others. When the major groove

Table 1. Mean helical parameters and major and minor groove widths

	X-displacement (Å)		Inclination (°)		Rise (Å)		Twist (°)		Minor groove width (Å)		Major groove width (Å)	
	mean	(sd)	mean	(sd)	mean	(sd)	mean	(sd)	mean	(sd)	mean	(sd)
A-RNA	-5.2		16.1		2.8		32.7		10.9		3.8	
A'-RNA	-5.3		10.3		3.0		30.0		10.8		8.0	
B-DNA	-0.6		-4.6		3.4		36.0		5.5		11.4	
1RNA	-4.9	(0.64)	18.7	(5.51)	2.6	(0.35)	33.5	(4.72)	10.1	(0.49)	3.4	(1.27)
205D	-4.9	(0.46)	16.1	(5.86)	2.6	(0.49)	33.9	(6.18)	9.4	(0.55)	3.9	(0.68)
1SDR1	-5.2	(0.43)	18.8	(2.44)	2.6	(0.24)	33.5	(4.36)	10.0	(0.50)	2.6	(0.47)
1SDR2	-5.1	(0.26)	19.1	(3.00)	2.6	(0.18)	33.6	(3.83)	10.0	(0.36)	2.9	(0.45)
280D1	-4.4	(0.84)	18.4	(3.17)	2.5	(0.37)	34.3	(6.07)	9.5	(1.07)	2.9	(0.87)
280D2	-4.7	(0.85)	20.1	(4.40)	2.4	(0.42)	33.7	(5.96)	9.6	(1.05)	3.0	(0.83)
tridecamer	-5.2	(1.08)	10.5	(2.36)	3.0	(0.17)	30.9	(8.02)	9.8	(0.84)	8.3	(0.51)
255D	-5.2	(1.08)	11.2	(2.41)	2.9	(0.18)	31.3	(8.81)	9.4	(0.78)	7.8	(0.55)
1BNA	-0.4	(0.28)	-1.0	(2.31)	3.4	(0.16)	35.8	(4.02)	5.1	(1.71)	11.4	(0.68)
2BNA	-0.3	(0.32)	-2.9	(2.68)	3.3	(0.25)	35.7	(4.98)	4.7	(1.88)	11.2	(0.70)
3BNA	-0.2	(0.31)	-3.7	(2.70)	3.4	(0.31)	36.4	(3.51)	4.8	(1.36)	11.3	(0.94)
4BNA	-0.1	(0.36)	-7.3	(2.80)	3.3	(0.32)	36.7	(4.15)	4.1	(0.61)	11.7	(1.24)
AVERAGE	-3.4		9.8		2.9		34.1		8.0		6.7	
SD	2.32		10.53		0.39		1.83		2.51		3.92	

Mean and (sd) indicate the mean structure parameter and standard deviation for each oligomer. AVERAGE and SD indicate the average value and its standard deviation of each structure parameter, respectively. The coordinates used here for the calculation of the structure parameters, except for those of the tridecamer, were taken from PDB (also NDB), the ID numbers being 1RNA (ARN035) for r(UUAUAUAUAUAUAA)$_2$ at 2.3 Å resolution (31), 205D (URL029) for r(GGACUUUGGUCC)$_2$ at 2.6 Å resolution (32), 1SDR (ARL062) for r(UAAGGAGGUGAU)·r(AUCACCUCCUUA) at 2.6 Å resolution (33), 280D (URL050) for r(GGCGCUUGCGUC)$_2$ at 2.4 Å resolution (34), 255D (ARL037) for r(GGACUUCGGUCC)$_2$ at 2.0 Å resolution (5), 1BNA (BDL001) for d(CGCGAATTCGCG)$_2$ at 2.3 Å resolution (35), 2BNA (BDL002) for d(CGCGAATTCGCG)$_2$ at 2.7 Å resolution (36), 3BNA (BDLB03) for d(CGCGAATT-^{Br}CGCG)$_2$ at 3.0 Å resolution (37) and 4BNA (BDLB04) for d(CGCGAATT^{Br}CGCG)$_2$ at 2.3 Å resolution (37). 1RNA, 255D, 205D, 1SDR, 280D and the tridecamer are RNA duplexes. 1BNA, 2BNA, 3BNA and 4BNA are B-DNA duplexes. A-RNA, A′-RNA and B-DNA represent the average structures deduced from X-ray fiber diffraction data.

widths were plotted against the inclinations, two clearly separated clusters appeared (Fig. 2). Interestingly, each cluster contains canonical A- and A′-RNAs, respectively. The standard deviation of each point along both axes is smaller than the distance between the clusters. From these data, we concluded that these two clusters must be assigned to the A-RNA and A′-RNA groups, respectively. This implies that the presence of the A′-RNA conformation was confirmed in the single crystal structures, i.e. the tridecamer and 255D belong to the A′-RNA conformation, while 1RNA, 1SDR1, 1SDR2, 205D, 280D1 and 280D2 belong to the A-RNA conformation.

Cluster analysis

We performed cluster analysis using all the structure parameters to reconfirm that the conformations of these oligomers can be divided into two groups. In cluster analysis in general, samples are classified based on the similarity of variables which represent the features of the samples. As mentioned previously, the mean structure parameters in Table 1 are variables which define the global conformational features. Therefore, cluster analysis can be performed using these parameters directly. However, they include distances and angles and also their values are dispersed. If the values in Table 1 are used for cluster analysis, the results might be affected only by a certain parameter which has the largest value. Accordingly, we calculated the normal distribution of each structure parameter before the cluster analysis and then calculated the distances, based on their normal distributions, between all combinations of the oligomers, A-RNA, A′-RNA and B-DNA (supplementary material, Tables S1 and S2). The oligomers, A-RNA, A′-RNA and B-DNA, were clustered by means of the nearest neighbor method and the resulting dendrogram is presented in Figure 3. The dendrogram shows the distances between clusters or duplexes. A short distance means that the compared structures are similar to each other.

The B-DNA group and all the RNA groups are separated by quite a great distance (4.13). As expected, the A-RNA and A′-RNA groups in Figure 2 were found to form different clusters with this method as well and they were separated by a distance of 1.53. On the other hand, the longest distance in each cluster was 0.84 and 0.66, respectively, which are much shorter than that between the A-RNA and A′-RNA groups. The differences in the inter- and intra-group distances are significant. This indicates that nucleic acid conformations can be classified by means of cluster analysis.

Comparison of torsion angles in A- and A′-RNAs

The mean torsion angles with their standard deviations for duplexes are listed in Table 2. In spite of the apparently different conformations of A- and A′-RNAs, the backbone torsion angles (α, β, γ, δ, ε and ζ) exhibited basically the same mean values for A- and A′-RNAs. The differences in the mean torsion angles were much smaller than the standard deviations. Only the χ angles of A- and A′-RNAs showed a difference which is comparable with their standard deviations. In conclusion, only the χ angle was

Table 2. Mean torsion angles

		χ	(°)	α	(°)	β	(°)	γ	(°)	δ	(°)	ε	(°)	ζ	(°)
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
A-RNA	280D1	199.8	9.12	289.7	40.72	177.1	15.41	54.7	29.94	82.0	2.96	208.0	8.92	286.3	8.46
	280D2	200.4	7.68	279.8	53.89	177.7	9.88	48.6	22.70	83.1	4.55	204.7	6.23	288.7	10.45
	1SDR1	198.0	6.17	287.4	11.19	168.6	10.29	63.1	7.13	75.6	5.58	206.3	10.36	289.1	8.30
	1SDR2	195.4	5.94	276.2	59.06	162.2	24.70	67.9	22.77	77.3	13.91	211.2	12.78	284.8	19.79
	1RNA	199.7	12.09	285.2	40.81	171.5	16.25	75.7	64.05	80.6	6.70	214.9	20.74	279.8	13.42
	205D	195.4	14.81	267.9	44.31	176.6	32.10	72.3	31.88	82.0	15.34	200.9	32.20	276.2	54.40
A'-RNA	255D	191.0	7.14	290.7	73.44	179.0	11.91	76.1	50.16	78.0	7.90	212.1	9.89	279.7	8.55
	tridecamer	193.5	6.81	280.5	43.45	178.5	9.92	66.2	37.65	80.5	3.36	208.2	7.53	286.1	7.53
AVE(A-RNA)		198.2		281.1		172.3		64.0		80.1		207.9		284.0	
SD(A-RNA)		9.76		43.48		19.88		34.70		9.53		17.60		24.70	
AVE(A'-RNA)		192.3		285.6		178.8		71.1		79.2		210.2		282.9	
SD(A'-RNA)		7.01		59.86		10.84		44.16		6.13		8.91		8.61	

Mean and sd indicate the mean torsion angle and standard deviation for each oligomer. AVE(A-RNA) and SD(A-RNA) indicate the average value and its standard deviation of each torsion angle for A-RNA, respectively. AVE(A′-RNA) and SD(A′-RNA) indicate the average value and its standard deviation of each torsion angle for A′-RNA, respectively.

slightly different between A- and A′-RNAs among all the torsion angles. The reason why backbone torsion angles are not affected by the difference in the conformation will be discussed later.

DISCUSSION

We showed that the A- and A′-RNA conformations could be classified as to the major groove widths and inclination angles (Fig. 2). In other words, the differences in the major groove width and the inclination angle between A- and A′-RNAs are clearer than that in the helical twist angle (Table 1), although the helical twist angle is still an important parameter for the classification of RNA duplexes. Then we further confirmed that the plot of the inclination versus the major groove width (Fig. 2) and the cluster analysis (Fig. 3) were basically unchanged even if other protocols were used for calculation of the structure parameters (38). Even at this stage, there remain two problems. One is how to correlate an inclination angle with a major groove width, and the other is why the backbone torsion angles of the two conformers are not different from each other (Table 2). Here we present a possible explanation which may solve the two problems at the same time. As shown in Figure 4a, if the inclination becomes smaller, the backbone of strand *a* moves downward and that of strand *b* upward. Consequently, the major groove width becomes larger vertically. As far as this model is concerned, the torsion angles do not need to change except for the χ angle. The change in the χ angle is expected to be approximately the same as the change in the inclination, because the vectors of the glycosyl bonds are nearly parallel to the rotation axis of the inclination angle. In fact, the difference in the inclination angles (8°) between A- and A′-RNAs is comparable with the difference in the χ angles (6°) between them (Table 2). Then we simulated the widening of the major groove (Fig. 4b). The vertical shift due to the inclination change is 2.6 Å, as indicated by a blue arrow (Fig. 4b). However, this value is less than the difference in the major groove width (4.5 Å) between A- and A′-RNAs and there should be a horizontal shift such as the decreased helical twist of A′-RNA relative to A-RNA (unwinding). The mean unwinding angle per base pair between A- and A′-RNAs is 2.7° and the nearest phosphorus atoms across the major groove are separated by 4 or 5 bp (Fig. 4b). Thus, the accumulated unwinding angle is ~14–16° and the resulting horizontal shift is ~2.4–2.8 Å (Fig. 4b). Finally, the major groove is widened by ~3.5–3.8 Å (Fig. 4b), which is approximately the same as the difference in the major groove width between A- and A′-RNAs.

We also demonstrated the presence of the A′-RNA conformation using the structure of a tridecamer in a single crystal. Then we further retrieved the occurrence of the A′-RNA conformation in the crystal structures of biologically active RNA molecules such as hammerhead ribozymes (39,40), the P4–P6 domains of group I intron (41,42) and the 62 nt domain of 5S rRNA (11). Interestingly, we found that helix IV of 5S rRNA adopted the A′-RNA conformation. It has characteristic features of the A′-RNA conformation, i.e. a wide major groove (8.7 Å) and a low inclination angle (9.5°). More interestingly, helix IV was suggested to be the binding site for the ribosomal protein L25, by the results of enzymatic probing (43–45). It could be thought that helix IV becomes ready for protein binding by taking on the A′-RNA conformation. Thus, the wide major groove of the A′-RNA conformation could be utilized for interaction with proteins, since the major groove of the A-RNA conformation is too narrow to accommodate a protein or peptide. Next we present space-filling models of the tridecamer, helix IV of 5S rRNA and canonical A- and A′-RNAs to determine their major groove widths (Fig. 5). As is apparent from Figure 4, the major groove of A-RNA indicated by X-ray fiber diffraction data (Fig. 5a) is quite narrow, while that of A′-RNA indicated by X-ray fiber diffraction data (Fig. 5b), the tridecamer (Fig. 5c) and helix IV of 5S rRNA (Fig. 5d) are wide enough to accommodate a protein or peptide. Thus, a difference in the major groove widths between A- and A′-RNA is evident.

Here we conclude that the crystal structure of the tridecamer belongs to the A′-RNA conformation and have revealed detailed structural features of the A′-RNA conformation at nearly atomic resolution. We also discussed the possible function of the

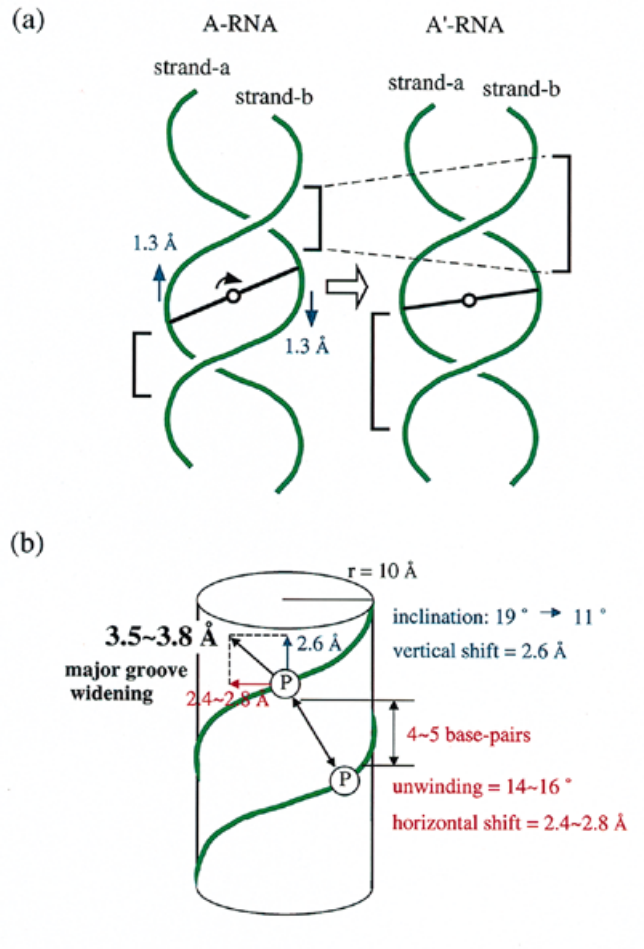

Figure 4. (**a**) Explanation of the correlation between the inclination angle and major groove width. A view from the major groove of the base pair (black bar). Green curved lines, black bars and open circles represent the backbones of nucleic acids, base pair planes and the rotation axis for inclination, respectively. Black and blue arrows indicate the rotational direction of inclination and vertical shifts of the strands, respectively. The vertical major groove width is presented as]. (**b**) The RNA duplex is presented as a cylinder and we assumed the radius of the cylinder to be 10 Å. Circles with the character P indicate the nearest phosphorus atoms across the major groove. The blue arrow and characters are the vertical shift and its value, respectively. The red arrow and characters are the horizontal shift and its value, respectively. The resulting major groove widening is indicated in black characters.

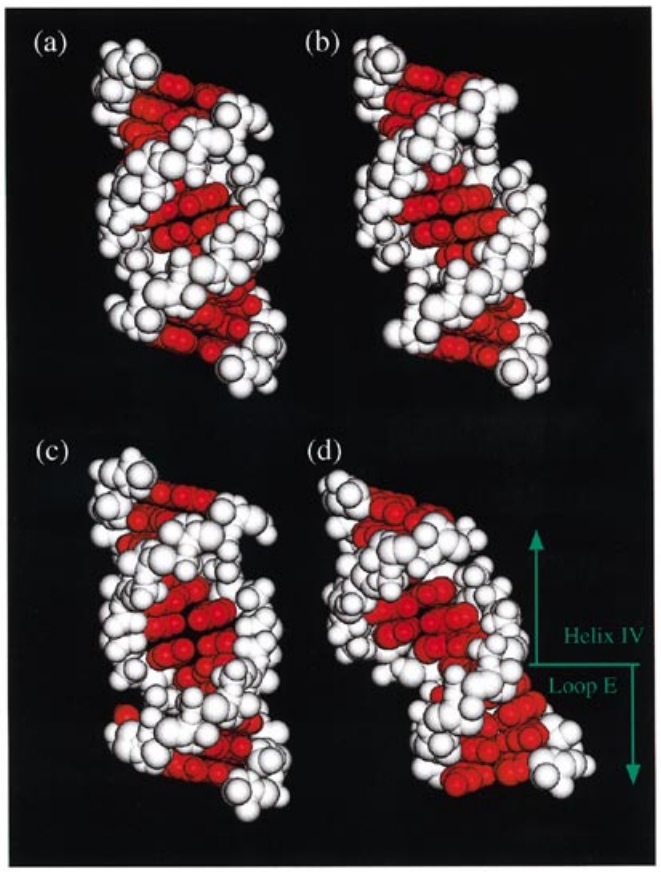

Figure 5. Space filling models of A-RNA deduced from X-ray fiber diffraction data (**a**), A′-RNA deduced from X-ray fiber diffraction data (**b**), the structure of the tridecamer (**c**) and the structure of helix IV with loop E in the 62 nt domain of *E.coli* 5S rRNA (URL065 from the Nucleic Acid DataBase, NDB) (11) (**d**). The helical axis for each duplex lies vertically. For (d), the helical axis was only calculated for helix IV because of the irregular structure of loop E.

A′-RNA conformation in a biological system. So far, it is reasonably safe to conclude that the A′-RNA conformation, which has a wide major groove, must be taken into consideration for RNA–protein interaction.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. Kazunari Taira, Dr De-Min Zhou, Dr Takeo Kohda, Dr Naruhisa Ota, Dr Masaki Warashina, Dr Tomoko Kuwabara, Mr Satoshi Fujita, Mr Ryuji Utsunomiya, Mr Masayuki Sano and Mr Satoru Sekiya (Tsukuba University) and Drs Toshio Yamazaki, Takashi S. Kodama, Chojiro Kojima, Koichi Uegaki, Hiroshi Matsuo, Eugene H. Morita, Mr Junichi Furui and Mr Mitsuaki Sugahara (Osaka University) for critical discussions and comments on the manuscript and also wish to thank Mr Tomokazu Hasegawa (Osaka University) for programing the drawing software. This work was supported by a Grant-in-Aid for Basic Scientific Research, Category B (no. 09480176), from the Ministry of Education, Science and Culture, Japan, and the Human Frontier Science Program.

See supplementary material available in NAR Online.

REFERENCES

1 Tuerk,C., Gauss,P., Thermes,C., Groebe,D.R., Gayle,M., Guild,N., Stormo,G., d'Aubenton-Carafa,Y., Uhlenbeck,O.C., Tinoco,I.,Jr, Brody,E.N. and Gold,L. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 1364–1368.
2 Sakata,T., Hiroaki,H., Oda,Y., Tanaka,T., Ikehara,M. and Uesugi,S. (1990) *Nucleic Acids Res.*, **18**, 3831–3839.
3 Cheong,C., Varani,G. and Tinoco,I.,Jr (1990) *Nature*, **346**, 680–682.
4 Antao,V.P. and Tinoco,I.,Jr (1992) *Nucleic Acids Res.*, **20**, 819–824.
5 Holbrook,S.R., Cheong,C., Tinoco,I.,Jr and Kim,S.-H. (1991) *Nature*, **353**, 579–581.
6 Cruse,W.B.T., Saludjian,P., Biala,E., Strazewski,P., Prangé,T. and Kennard,O. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 4160–4164.
7 Fujii,S., Tanaka,Y., Tomita,K.-I., Sakata,T., Hiroaki,H., Tanaka,T. and Uesugi,S. (1991) *Nucleic Acids Symp. Ser.*, **25**, 181–182.

8 Fujii,S., Tanaka,Y., Uesugi,S., Tanaka,T., Sakata,T. and Hiroaki,H. (1992) *Nucleic Acids Symp. Ser.*, **27**, 63–64.
9 Puglisi,J.D., Chen,L., Blanchard,S. and Frankel,A.D. (1995) *Science*, **270**, 1200–1203.
10 Battiste,J.L., Mao,H., Rao,N.S., Tan,R., Muhandiram,D.R., Kay,L.E., Frankel,A.D. and Williamson,J.R. (1996) *Science*, **273**, 1547–1551.
11 Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) *Cell*, **91**, 705–712.
12 Watson,D.J. and Crick,F.H.C. (1953) *Nature*, **171**, 737–738.
13 Langridge,R. and Gomatos,P.J. (1963) *Science*, **141**, 694–698.
14 Tomita,K.-I. and Rich,A. (1964) *Nature*, **201**, 1160–1163.
15 Fuller,W., Wilkins,M.H.F., Wilson,H.R., Hamilton,L.D. and Arnott,S. (1965) *J. Mol. Biol.*, **12**, 60–80.
16 Arnott,S., Hukins,D.W.L. and Dover,S.D. (1972) *Biochem. Biophys. Res. Commun.*, **48**, 1392–1399.
17 Shindo,H., Fujiwara,T., Akutsu,H., Matsumoto,U. and Kyogoku,Y. (1985) *Biochemistry*, **24**, 887–895.
18 Huber,R. (1965) *Acta Crysallogr.*, **A19**, 353–356.
19 Steigemann,W. (1974) Doctoral thesis, Technische Universität München.
20 Fuginaga,M. and Read,R.J. (1987) *J. Appl. Crystallogr.*, **20**, 517–521.
21 Brünger,A.T. (1990) *Acta Crystallogr.*, **A46**, 46–57.
22 Powell,M.J.D. (1977) *Math. Programming*, **12**, 241–251.
23 Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) *Science*, **235**, 458–460.
24 Brünger,A.T., Krukowski,A. and Erickson,J. (1990) *Acta Crystallogr.*, **A46**, 585–593.
25 Brünger A.T. (1992) *Nature*, **355**, 472–474.
26 Luzzati,P.V. (1952) *Acta Crystallogr.*, **5**, 802–810.
27 Perkinson,G., Vojtechovsky,J., Clowney,L., Brünger,A.T. and Berman,H.M. (1996) *Acta Crystallogr.*, **D52**, 57–64.
28 Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.
29 Stofer,E. and Lavery,R. (1993) *Biopolymers*, **34**, 337–346.
30 Wahl,M.C., Rao,S.T. and Sundaralingam,M. (1996) *Nature Struct. Biol.*, **3**, 24–31.
31 Dock-Bregeon,A.C., Chevrier,B., Podjarny,A., Johnson,J., de Baer,J.S., Gough,G.R., Gilham,P.T. and Moras,D. (1989) *J. Mol. Biol.*, **209**, 459–474.
32 Baeyens,K.J., De Bondt,H.L. and Holbrook,S.R. (1995) *Nature Struct. Biol.*, **2**, 56–62.
33 Schindelin,H., Zhang,M., Bald,R., Fürste,J.-P., Erdmann,V.A. and Heinemann,U. (1995) *J. Mol. Biol.*, **249**, 595–603.
34 Lietzke,S.E., Barnes,C.L., Berglund,J.A. and Kundrot,C.E. (1996) *Structure*, **4**, 917–930.
35 Drew,H.R., Wing,R.M., Takano,T., Broka,C., Tanaka,S., Itakura,K. and Dickerson,R.E. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 2179–2183.
36 Drew,H.R., Samson,S. and Dickerson,R.E. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 4040–4044.
37 Fratini,A.V., Kopka,M.L., Drew H.R. and Dickerson,R.E. (1982) *J. Biol. Chem.*, **257**, 14686–14707.
38 Tanaka,Y. (1998) Doctoral thesis, Graduate School of Science, Osaka University.
39 Pley,H.W., Flaherty,K.M. and Mckey,D.B. (1994) *Nature*, **372**, 68–74.
40 Scott,W.G., Finch,J.T. and Klug,A. (1995) *Cell*, **81**, 991–1002.
41 Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) *Science*, **273**, 1678–1685.
42 Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Szewczak,A.A., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) *Science*, **273**, 1696–1699.
43 Douthwaite,S., Christensen,A. and Garrett,R.A. (1982) *Biochemistry*, **21**, 2313–2320.
44 Huber,P.W. and Wool,I.G. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 322–326.
45 Toukifimpa,R., Romby,P., Rozier,C., Ehresmann,C., Ehresmann,B. and Mache,R. (1989) *Biochemistry*, **28**, 5840–5846.
46 Sussman,J.L., Seeman,N.,C., Kim,S.-H. and Berman,H.M. (1972) *J. Mol. Biol.*, **66**, 403–421.