

The crystal structure of an RNA oligomer incorporating tandem adenosine–inosine mismatches

Richard J. Carter, Katrien J. Baeyens¹, John SantaLucia², Douglas H. Turner³ and Stephen R. Holbrook*

Structural Biology Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720, USA, ¹Laboratory of Analytical Chemistry and Medicinal Physicochemistry, KU Leuven, Belgium, ²Department of Chemistry, Wayne State University, Detroit, MI 48202, USA and ³Department of Chemistry, University of Rochester, Rochester, NY 14627, USA

Received May 27, 1997; Revised and Accepted August 29, 1997

ABSTRACT

The X-ray crystallographic structure of the RNA duplex $[r(\text{CGCAIGCG})_2]$ has been refined to 2.5 Å. It shows a symmetric internal loop of two non-Watson–Crick base pairs which form in the middle of the duplex. The tandem A–I/I–A pairs are related by a crystallographic two-fold axis. Both A(anti)–I(anti) mismatches are in a head-to-head conformation forming hydrogen bonds using the Watson–Crick positions. The octamer duplexes stack above one another in the cell forming a pseudo-infinite helix throughout the crystal. A hydrated calcium ion bridges between the 3'-terminal of one molecule and the backbone of another. The tandem A–I mismatches are incorporated with only minor distortion to the backbone. This is in contrast to the large helical perturbations often produced by sheared G–A pairs in RNA oligonucleotides.

INTRODUCTION

Internal loops are one of the most common elements of RNA secondary structure. In fact, as shown by several crystallographic and NMR analyses of RNA oligomers, symmetric internal loops are not 'loops' at all, but runs of non-Watson–Crick base pairs within a continuous double helix (1–9). The RNA double helix in these structures is distorted to varying degrees depending on the sequence of the internal loop.

The RNA octamer $r(\text{CGCAIGCG})$ has been crystallized and its three-dimensional structure determined at 2.5 Å resolution. Non-canonical base pairing between the central nucleotides leads to the formation of tandem A–I, I–A base pairs within a double helix.

Assuming that at least two hydrogen bonds must form to produce a stable base pair, there are five possible pairings between adenosine and guanosine (10). These are shown in Figure 1. Type I base pairs are in the head-to-head conformation and form hydrogen bonds between the hydrogen of N6(A) and O6(G) and between the N1(A) and the hydrogen of N1(G). Type III pairs form hydrogen bonds between N6(A) and O6(G) and between N7(A) and N1(G). Type V pairs form bonds between N6(A) and O6(G) and between N1(A) and N7(G). The sheared type II base pair and type IV base pair both induce large backbone distortions in the

duplex. The structure of an A–I pair which lacks the N2 amino group must conform to one of either type I, III or V motifs.

Inosine has chemical properties very similar to guanosine although the missing N2 amino group gives inosine a versatility for base pairing with other bases in an A-form duplex (11). This is utilized by some tRNA molecules which have inosine at the 5'-end of the anticodon where it appears to pair with adenosine, uridine or cytidine without destabilizing the duplex (12) and also in the middle position of an anticodon to pair with adenosine (13). Only a few crystal structures of inosine containing duplexes have been solved. These include the structures of the deoxynucleotides I(anti)–A(syn) (14), an I–T wobble base pair (15) and an I–C Watson–Crick base pair (16,17).

The crystallographic structural analysis of the RNA octamer described below allows comparison between tandem A–I/I–A pairs, A–G/G–A pairs (9) which occur rarely in nature, and G–A/A–G tandem pairs which occur with much greater frequency in biological systems (18).

MATERIALS AND METHODS

Crystallization and data collection

The RNA oligomer of sequence CGCAIGCG was synthesized and purified as described previously (19). Crystals were grown in hanging drops at room temperature from a solution containing 1.5 µl of 1.5 mM $r(\text{CGCAIGCG})$ and 1.5 µl of crystallization buffer consisting of 0.2 M calcium chloride, 0.1 M HEPES pH 7.5 and 28% polyethyleneglycol (PEG) 400. The drops were equilibrated against crystallization buffer and bipyramidal needles appeared after 24 h.

The crystals were mounted in thin walled quartz capillaries for X-ray data collection on a Rigaku R-axis IIC diffractometer using Cu K α radiation ($\lambda = 1.5418$ Å). Data were measured at 24 °C using ω scans with a scan width of 2°. The data belonged to either space group P6₁22 or P6₅22 with dimensions of $a = b = 39.05$ Å, $c = 58.85$ Å, $\alpha = \beta = 90^\circ$ and $\gamma = 120^\circ$. A total of 5952 reflections were measured which reduced to 1014 unique reflections. This represented 91.8% of a complete data set. The R_{merge} for this data was 7.9%. The refinement statistics are shown in Table 1.

*To whom correspondence should be addressed. Tel: +1 510 486 4304; Fax: +1 510 486 6059; Email: srholbrook@lbl.gov

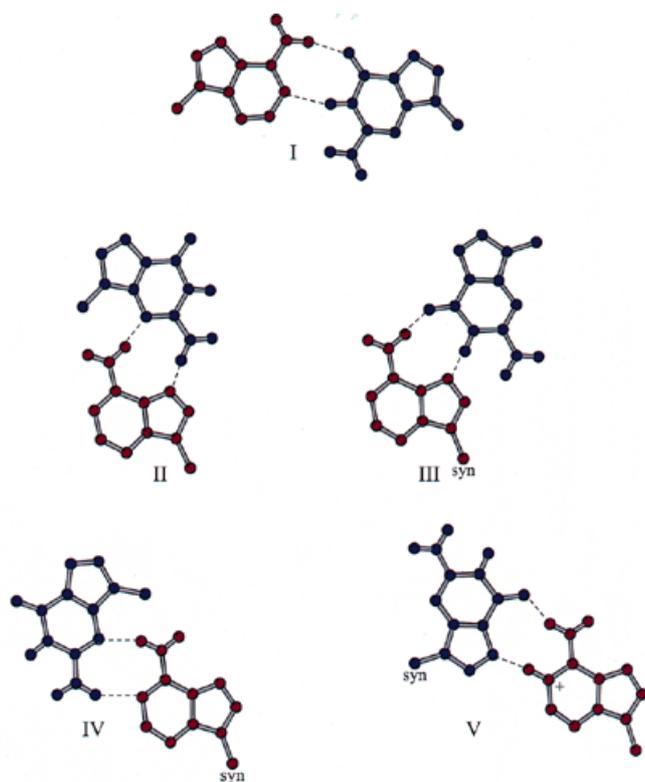


Figure 1. The five possible A-G mismatches (adapted from Gautheret *et al.*, 10). The adenosines are coloured red, the guanosines are coloured blue and the bases are *anti* except where noted.

Structure solution and refinement

The structure was determined by molecular replacement (MR) using X-PLOR (20). A number of search models were tried including a canonical A-form RNA duplex in both space group $P6_122$ and $P6_522$ but produced no satisfactory solution. A successful solution was found using the three base pair fragment CGC:GCG from the decamer structure of Leonard *et al.* (21) in space group $P6_122$. The top peak in the direct PC search followed by translation gave a solution with a high translation function and a good packing function. Rigid body refinement of the model reduced the R factor from 46.4% to 41.8%. The electron density maps generated at this stage showed the positions of the missing

adenosines and inosines in the duplex. The structure was then subjected to cycles of rigid body, positional and simulated annealing refinement with X-PLOR using data having $F \geq 3\sigma$ in the resolution range 12–2.5 Å. This resulted in a final R factor of 24.6% and R_{free} of 26.6%. The refined model consists of a single RNA octamer, six water molecules and one calcium molecule per asymmetric unit.

The geometry of the model is good with root mean square (rms) deviation from ideal bond lengths of 0.013 Å for the ribose/base and 0.017 Å for the backbone. The rms deviation from the canonical values for the angles was 1.863°. The van der Waals distances for both inter- and intra-molecular contacts are good with no evidence for clashing.

Table 1. Refinement statistics

Resolution	25.0–2.52 Å
Number of measurements	5952
Number of unique reflections	1014
R_{merge}	7.9%
Completeness	91.8%
Space group	$P6_122$
Dimensions	39.05 Å, 39.05 Å, 58.85 Å, 90°, 90°, 120°
Number of non-hydrogen RNA atoms (per unique strand)	199
Number of solvent molecules (per unique strand)	6
R factor (%) ($F \geq 3\sigma$)	24.6%
R_{free}	26.6%
rms Deviations from ideality	
Bond length	0.011 Å
Bond angle	1.70°

RESULTS

Duplex structure

The octamer r(CGCAIGCG) forms a self-complementary duplex in the crystal (AI duplex) with a global helical rise of 2.76 Å per base pair and helical twist of 32.30° per base pair which places it in the A-form family (22). The pseudorotation angles indicate that the ribose moieties in the structure belong to the C3'-endo family (Table 2) except for the adenosines which have a C2'-exo ribose pucker.

Table 2. Sugar–phosphate backbone and glycosyl torsion angles (°)

Residue	χ	α	β	γ	δ	ξ	ζ	Pseudo-rotation
C1	-174.7	-70.4	169.1	60.7	80.9	-153.0	-71.7	10.2
G2	-177.8	-59.4	-160.5	58.7	79.1	-152.3	-61.4	9.2
C3	-149.6	-175.0	152.4	36.5	81.3	-161.8	-60.7	6.6
A4	-165.3	-54.6	157.1	155.0	93.1	-118.8	-69.5	336.6
I5	-163.9	-60.5	-176.2	49.1	74.5	-164.3	-69.1	22.6
G6	-173.4	-94.2	-153.5	45.3	81.5	-177.9	-42.0	17.2
C7	-163.6	176.0	-176.2	58.9	81.4	-178.4	-62.7	35.1
G8	-179.8	–	–	156.5	88.8	–	–	20.4
A-form RNA	-158	-68	178	54	82	-153	-71	18

Note: As the structure is symmetrical, the parameters for the remaining bases (C9–G16) are identical to those above. Entries in bold face differ from canonical A-form by >15°.

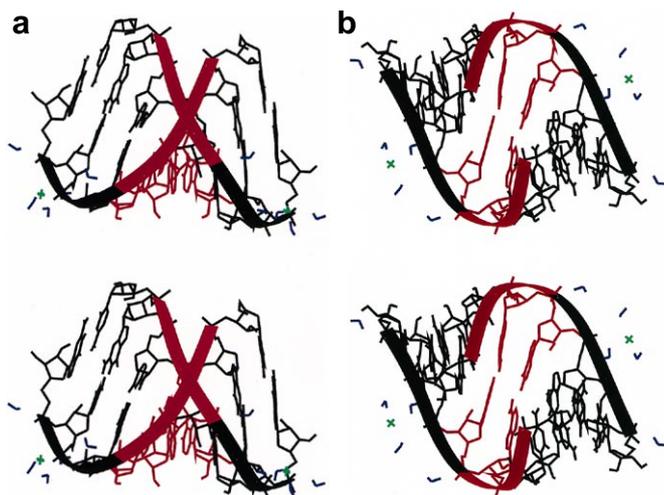


Figure 2. (a) A ribbon representation of $[r(CGCAIGCG)]_2$. The calcium atoms are represented by crosses at either end of the duplex. The mismatched region is shown in a different colour. (b) The same representation rotated by 90° .

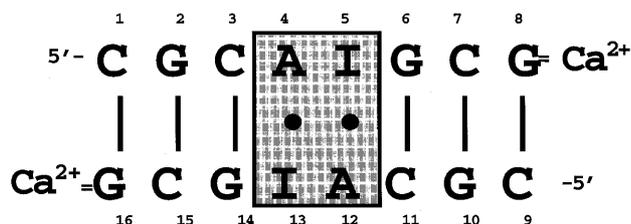


Figure 3. Schematic view of the $[r(CGCAIGCG)]_2$ showing the two non-Watson-Crick base pairs (boxed). A crystallographic two-fold axis around the A-I/A mismatches relates the two halves of the duplex.

From Figure 2 it can be seen that the helix has a deep, narrow, major groove and a wide, shallow, minor groove. The average width of the minor groove is 11.1 Å which is consistent with the minor groove width of A-form RNA. However, the groove width is increased in the region containing the tandem A-I mismatch. Here it has a width of up to 12.06 Å. The separations between adjacent phosphorus atoms on the same strand range from 5.36 to 6.24 Å with an average distance of 5.87 Å which is the same as the native A-form.

The AI duplex consists of Watson-Crick base paired ends flanking tandem adenosine-inosine mismatches (Fig. 3). In the crystal, duplexes stack above one another in a head-to-head fashion with the 3'-terminal of one molecule joined via a water bridge to the 3'-terminal of the next (further discussed in water structure). This stacking forms a discontinuous pseudo-infinite helix throughout the crystal (Fig. 4). At the junction of adjacent helices, there is a distance of 5.25 Å from the phosphate of residue G8 to the phosphate of a symmetry related G8 residue.

Analysis of the A-I duplex using the program CURVES (23) indicated that the helical axis is curved by 18.6° from linear. The largest region of irregularity is around the mismatch. The A-I base pairs are offset by 1.32 Å from a linear axis. The kink angles between the two A-I base pairs and between the A-I base pair and its neighboring Watson-Crick base pair are 5.3° . This causes an

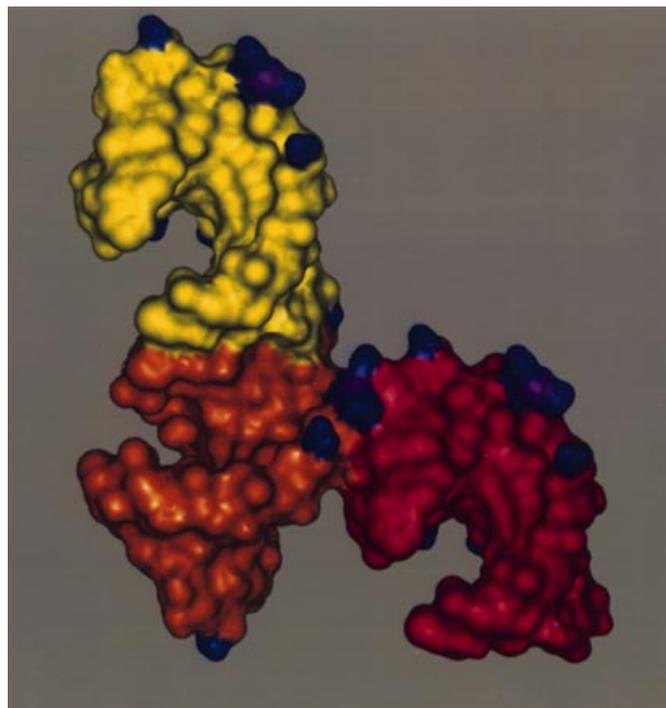


Figure 4. Duplexes stacked upon one another in a head-to-head fashion and a duplex from the next layer interacting at the interface between the first duplex. (The two stacked duplexes are shown in yellow and orange. The duplex from the next layer is red.) The hydrated calciums (purple with waters shown in blue) lie at the interface of two duplexes but only bond to one member of the pseudo-helix and then to the backbone of the adjacent duplex.

expansion of the minor groove (12.06 Å at its largest) and the diameter of the helix to increase to 19.2 Å (c.f. 17.4 Å in standard A-form). Table 4 also shows that the roll and tilt values are highest around the tandem mispairs. No measurement of the major groove could be made as there is not a full turn of the helix.

We have observed the calculations of helical distortion determined by the CURVES algorithm to be dependent upon the definition of the non-canonical base pairs in the internal loop. Therefore, we have used a second approach to measure the helical deformation due to the internal loop. By the superposition of canonical A-form helices onto the Watson-Crick ends of the octamer duplex, the angle and displacement between the helical axes can be easily calculated. Using this method, the two axes surrounding the internal loop in the AI duplex are nearly colinear (translational displacement 0.1 Å and interhelical angle 0.2°), in contrast to other internal loop structures which show high bending or displacement (1,3).

Packing

The duplexes stack one above another in the crystal as discussed above and shown in Figure 4. The duplexes also form in layers beside one another related by 60° to the adjacent layer. The layers lie such that the junction between two helices in the first layer interacts with the backbone of the next duplex at residue G6 (Fig. 4). This junction contains a calcium ion and its neighboring waters (discussed later). The packing is mainly stabilized by the hydrated calcium and interactions at the 3'-terminus. There is also

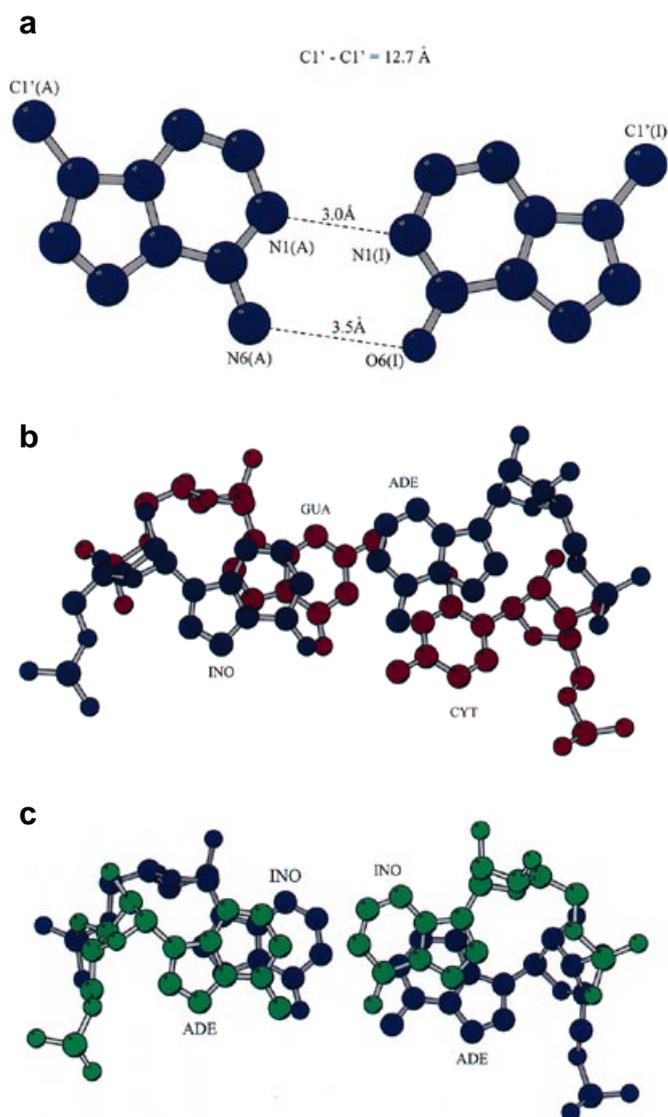


Figure 5. (a) The A-I base pair showing the hydrogen bonds. (b) The stacking of the A-I base pair on the preceding C-G base pair. (c) The stacking of tandem A-I base pairs. Both A-I mismatches are in the head-to-head imino hydrogen bonded conformation.

a hydrogen bond between the O2' of residue G2 and the O2' of residue I5 of a molecule related by crystallographic symmetry.

Conformation of the A-I base pairing

The A-I base pairs in the structure of $[r(\text{CGCAIGCG})]_2$ adopt an A(anti)-I(anti) type I conformation, as shown in Figure 5. The two hydrogen bond distances are 3.0 Å for N1(A)-N1(I) and 3.4 Å for N6(A)-O6(I). The propeller twists of the A-I base pairs are -4.3° (Table 3). The C1'(A)-C1'(I) distance is 12.67 Å compared with 10.64 Å, the average in the base paired region. An increase in the backbone torsion angles β , γ and δ and a decrease in ζ around the adenosine allow an expansion at the mismatch which allows the larger A-I base pairs to be incorporated into the duplex. The glycosyl angles are in the anti range (Watson-Crick $52-62^\circ$), with values of 54.8° for the inosine and 48.4° for the adenosine.

Table 3. Geometric parameters for base pairs in $[r(\text{CGCAIGCG})]_2$

Base pair	Propeller twist($^\circ$)	C1'-C1' (Å)	Inclination ($^\circ$)	Offset
C1-G16	0.79	10.51	5.0	0.00
G2-C15	-11.91	10.56	4.8	0.46
C3-G14	-7.37	10.84	7.5	1.07
A4-I13	-4.29	12.67	18.6	1.32

Table 4. Helical parameters [definitions as for CURVES (23)]

Step	Rise	Angle	Slide	Roll	Tilt	Twist
C1-G2	3.26	2.29	-2.34	-0.18	-0.39	32.60
G2-C3	2.78	0.59	-3.00	1.97	4.69	32.98
C3-A4	2.80	5.34	-1.85	15.15	8.45	31.60
A4-I5	1.72	5.25	-1.85	13.68	-0.02	24.61
Average	2.77	3.08	-2.32	6.81	0.01	31.29
A-form	2.8	-	-1.5	-0.4	13.0	32.7

There is only minor distortion to the backbone caused by the tandem A-I base pairs. However, the increased helical diameter necessary to incorporate the purine-purine A-I pair is accompanied by a reduced rise per residue of only 1.7 Å for the A-I-I-A step and, as can be seen from Figure 5 and Table 4, a low helical twist angle (24.6°) between the A-I-I-A base pairs. This allows for maximum stacking of the tandem mismatches. It is also worth noting the pyrimidine preceding the A-I mismatch (C3, C11) is unstacked on the 3' side (Fig. 5b) leaving it available for interaction with other molecules.

Metal binding site and water structure

Figure 6 shows the site of the hydrated calcium ion. One calcium ion is associated with each strand of the duplex. It is bonded to the O1P of residue G6 via a water molecule and with the O2' and O3' atoms of the 3'-terminal residue G8 from a symmetry related molecule. The calcium is hydrated with four waters surrounding it at a distance of 2.6-2.7 Å. As it lies between helices, there is sufficient room to accommodate the relatively bulky hydrated calcium without causing any large distortion to the helical nature of the duplex. Two symmetry related hydrated calcium atoms are only 6.0 Å apart. Two of the four water molecules associated with the calcium also interact with backbone atoms of the duplex. Water 1001 is 2.6 Å from O1P (residue G6) and water 1003 is also associated with O3' (residue G8) at a distance of 2.6 Å. Water 1005 acts as a bridge between two adjacent molecules in a helix, linking the 3'-residues across the phosphates. The water is ~ 2.3 Å from each O1P (residue G8) and 3.2 Å from each O2P (residue G8). Finally, water 1006 lies within 2.7 Å of O1P (residue I5). There are indications that there are more water molecules associated with the structure, most likely with the backbone atoms, especially the phosphate oxygens. These lie on crystallographically special positions (sites of crystal symmetry elements), however, and have not been refined. It is notable that no waters were observed in the region of the A-I mismatches. The waters around the calcium are well ordered with an average B

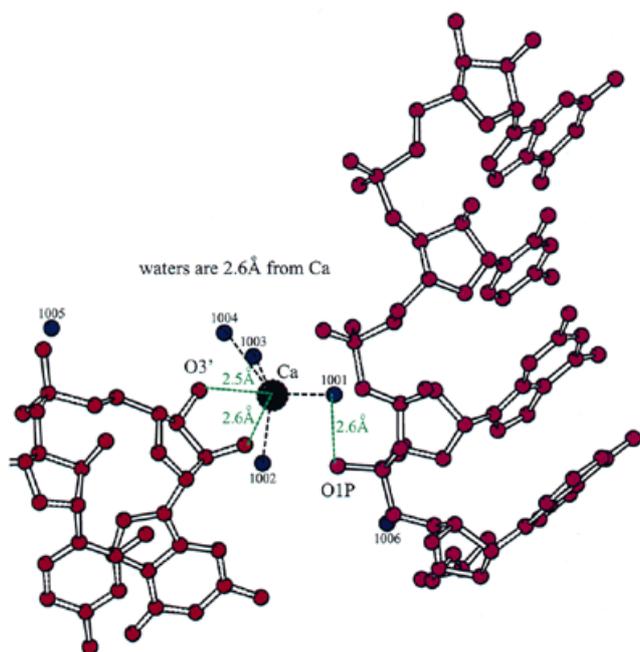


Figure 6. The calcium binding pocket. The calcium (black sphere) is surrounded by four water molecules (blue) and binds via a water to the O1P of a guanosine residue and to the ribose oxygens of the 3'-terminal of an adjacent duplex.

factor of 27.5 \AA^2 . The water molecule which lies near O1P (residue I5) is not as well ordered and has a B factor of 48.2 \AA^2 .

Analysis of the B factors for the molecule show that the duplex is relatively rigid. There is greatest mobility at the 5'-terminus of each strand, whilst the 3'-terminus is relatively less mobile. This is probably due to the stabilizing effect of the hydrated calcium bound to the 3'-terminus. The adenosine ribose ring and the inosine base have above average B factors, but the density fit is good in these regions.

DISCUSSION

Tandem A-G and A-I base pairs are among the most thermodynamically stable of all possible combinations of tandem mismatches (after G-U pairs) (19,24). Yet despite their physical stability, tandem A-G/G-A mismatches are not seen in ribosomal RNA (18,25). It has been shown by NMR studies (6,9,18) that G-A/A-G and A-G/G-A tandem mismatches have different structures, when flanked by a 5' C-G base pair. These studies indicate that A-G/G-A tandem pairs adopt an A-like conformation, leading to the suggestion (18) that A-G/G-A is rarely found in nature because its similarity to A-form RNA does not induce a unique, recognizable structural change in the duplex. Since the N2 amino group is missing from inosine as compared to guanosine, the A-I base pair can only assume conformations of type I, III or V. In this structure, both A-I base pairs conform to type I, the imino-hydrogen bonded head-to-head conformation. This is the same conformation seen for the A-G/G-A motif in the NMR structure of the octamer $[r(\text{GGCAGGCC})]_2$ (9). Interestingly, the A-I/I-A motif is $\sim 1 \text{ kcal/mol}$ more stable than the

A-G/G-A at 37°C (19,26). This is consistent with the structural observation that the guanosine N2 amino moiety is not involved in base pairing (9) or easily available for hydrogen bonding with solvent. The greater stability of A-I pairs may also arise from differences in pKa's for I (8.8) and G (9.2) (19). It should be noted, however, that the thermodynamics of the duplexes have only been characterized in the absence of divalent cations and thus may not represent *in vivo* conditions.

Several examples of A-I and A-G base pairs have been observed in crystal and NMR studies of DNA and RNA oligomers as summarized below.

Two A-I mismatches have been observed in crystals of DNA oligomers. The structure of $[\text{d}(\text{CGCIAATTAGCG})]_2$ (14) showed that the purine-purine mismatch was in the type III A(syn)-I(anti) conformation. The mismatch was incorporated into the B-DNA helix with little distortion of either local or global conformation. The structure of $[\text{d}(\text{CGCAAATTIGCG})]_2$, containing two adenosine-inosine pairs (27) showed the mismatches in the A(anti)-I(syn) type V conformation. Again, there was little or no perturbation of the helix. Leonard *et al.* (27) used ultraviolet melting studies to show that pH was crucial for the conformation of the mismatch. Below pH 6.5 the base pairs adopt the protonated AH⁺(anti)-I(syn) conformation, whilst in the pH range of 6.5-8.0 the base pairs are in the non-protonated A(anti)-I(anti) conformation. Crystallization at or below pH 6.5 preferentially selects the anti-syn base pairs. Our crystals were grown at pH 7.5 and, in agreement with these findings, are in the A(anti)-I(anti) conformation.

The crystal structure of an RNA dodecamer $[\text{r}(\text{CGCGAAUUAGCG})]_2$ duplex incorporating two G(anti)-A(anti) type I mismatches (21) shows a number of features also observed in the A-I duplex. In both cases, the C1'-C1' distances at the purine-purine pairs are increased by $\sim 2 \text{ \AA}$ from regular A-form RNA. There are also decreases in the twist and rise around the mismatches. The N1(G)-N1(A) distance compares favourably with the A-I structure, but the N6(A)-O6(G) distance of 3.0 \AA is smaller than the corresponding distance in A-I (3.4 \AA). The G-A mismatch is incorporated into the duplex with little change to the structure of the phosphodiester backbone. The A-I/I-A structure also shows that the mismatches are accommodated by small distortions to the backbone (Table 2).

A functionally and structurally important single G-A mismatch has been observed in the NMR structure of the Rev responsive element of HIV-1 (RRE) (28). This mispair has a similar C1'-C1' distance to that of A-I: 12.9 \AA in RRE, 12.7 \AA for A-I. The G-A mismatch in RRE helps expand the major groove to allow the Rev protein to bind to the RNA and regulate splicing and transport of mRNA from the nucleus (29).

The structure of the double helical RNA octamer $[\text{r}(\text{GGCAGGCC})]_2$ has been solved by NMR (9) and shown to incorporate tandem type I A-G mismatches with similar N1-N1 hydrogen bond lengths, C1'-C1' distances, and angles between the glycosyl bonds as those observed in the tandem A-I pairs. They both also have a decrease in the rise (1.7 \AA for A-I and 2.1 \AA for A-G) and the twist (24° for A-I and 27° for A-G) between the mismatches. Again, the N6-O6 bond is shorter in A-G than in A-I: 3.1 and 3.4 \AA , respectively. The A-I/I-A and the A-G/G-A mismatches and their flanking base pairs superimpose within an rms deviation of 0.84 \AA . Overall, both the A-I/I-A duplex and the A-G/G-A structure are very similar to standard A-form RNA.

Sheared (type II) G–A mismatches have been observed in the crystal structures of hammerhead catalytic RNA (30,31) and an RNA dodecamer duplex [r(GGCCGAAAGGCC)]₂ (3). A divalent metal (i.e. Mn²⁺, Mg²⁺) binding site was identified in the pocket between the preceding Watson–Crick C–G and the G–A base pairs. The hydrogen bonding in this structure occurs between N3(G)–N6(A), N2(G)–N7(A) and O2'(G)–N6(A). This type of sheared G–A pair has also been observed by NMR in an octameric RNA in the absence of a divalent cation (6).

The hydrated calcium ions in our structure lie in the junction between helices, bonding directly to the 3'-terminal O2' and O3' of base G8 and indirectly via a water molecule to the O1P of G6 in a neighbouring duplex (Fig. 6). In DNA structures (32,33), hydrated calcium binds indirectly through its coordinated waters. The calcium lies in the minor groove and forms ligands using seven water molecules, each at a distance of 2.4 Å. In the A–I duplex there are four water molecules associated with each calcium ion.

Summary

There are many possible structures for A–I and A–G base pairs (22). We have determined the crystal structure of an RNA octamer duplex with an internal loop of tandem A–I/I–A type I base pairs. This agrees with an NMR structure containing tandem A–G/G–A mismatches which also form type I A–G base pairs (9). The type I A–I base pairs do not significantly distort the double helix from A-form. This explains the presence of inosine in the wobble position of some transfer RNA anticodon loops, which can base pair with A, C or U in the codon, without distorting the helix.

Tandem A–G/G–A pairs have a similar thermodynamic stability to G–A/A–G pairs (18), but assume different structures. The absence of A–G/G–A motifs in biological RNAs may thus arise from their similarity to A-form conformation which does not present a unique recognition site for binding proteins or other ligands, compared with G–A/A–G sequences which assume a sheared conformation, distorting the helix and, when preceded by a C–G pair, creating a divalent cation binding site.

The crystal structure also reveals a novel calcium ion binding site with direct ligation to the 3'-terminus of double helical RNA.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health Grant, NIGMS, GM 49215. Facilities and equipment were provided through support of the Office of Energy Research, Office of Health and Environmental Research and the Health Effects Research Division of the US Department of Energy.

REFERENCES

- Holbrook, S. R., Cheong, C., Tinoco, I., Jr and Kim, S.-H. (1991) *Nature*, **353**, 579–581.
- Baeyens, K. J., DeBondt, H. L. and Holbrook, S. R. (1995) *Nature Struct. Biol.*, **2**, 56–62.
- Baeyens, K. J., De Bondt, H. L., Pardi, A. and Holbrook, S. R. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 12851–12855.
- Lietzke, S. E., Barnes, C. L., Berglund, J. A. and Kundrot, C. E. (1996) *Structure*, **4**, 917–930.
- Wimberly, B. (1994) *Nature Struct. Biol.*, **1**, 820–827.
- SantaLucia, J., Jr and Turner, D. H. (1993) *Biochemistry*, **32**, 12612–12623.
- 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.
- Wu, M. and Turner, D. H. (1996) *Biochemistry*, **35**, 9677–9689.
- Wu, M., SantaLucia, J., Jr and Turner, D. H. (1997) *Biochemistry*, **36**, 4449–4460.
- Gautheret, D., Konings, D. and Gutell, R. R. (1994) *J. Mol. Biol.*, **242**, 1–8.
- Martin, F. H., Castro, M. N., Aboul-ela, F. and Tinoco, I., Jr (1985) *Nucleic Acids Res.*, **13**, 8927–8938.
- Felsenfeld, G. and Miles, H. T. (1967) *Annu. Rev. Biochem.*, **31**, 407–448.
- Davis, D. B., Anderson, P. and Sparking, P. F. (1973) *J. Mol. Biol.*, **76**, 223–232.
- Corfield, P. W. R., Hunter, W. N., Brown, T., Robinson, P. and Kennard, O. (1987) *Nucleic Acids Res.*, **15**, 7935–7949.
- Cruse, W. B. T., Aymani, J., Kennard, O., Brown, T., Jack, A. G. C. and Leonard, G. A. (1989) *Nucleic Acids Res.*, **17**, 55–72.
- Kumar, V. D., Harrison, R. W., Andrews, L. and Weber, I. T. (1992) *Biochemistry*, **31**, 1541–1550.
- Xuan, J.-C. and Weber, I. T. (1992) *Nucleic Acids Res.*, **20**, 5457–5464.
- SantaLucia, J., Jr, Kierzek, R. and Turner, D. H. (1990) *Biochemistry*, **29**, 8813–8819.
- SantaLucia, J., Jr, Kierzek, R. and Turner, D. H. (1991) *J. Am. Chem. Soc.*, **113**, 4313–4322.
- Brunger, A. T. (1992) X-PLOR (Version 3.1) A system for X-ray Crystallography and NMR, Yale University Press, New Haven, CT.
- Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T. and Hunter, W. N. (1994) *Curr. Biol.*, **2**, 483–494.
- Saenger, W. (1984) In Cantor, C. R. (Ed.) *Principles of Nucleic Acid Structure*. Springer advanced texts in chemistry. Springer-Verlag, New York, NY.
- Ravishankar, G., Swaminathan, S., Beveridge, D. L., Lavery, R. and Sklenar, H. (1989) *J. Biomol. Struct. Dynamic.*, **6**, 669–699.
- Wu, M., McDowell, J. A. and Turner, D. H. (1995) *Biochemistry*, **34**, 3204–3211.
- Gutell, R. R., Weiser, B., Woese, C. R. and Noller, H. F. (1985) *Prog. Nucleic Acid Res. Mol. Biol.*, **32**, 155–216.
- SantaLucia, J., Jr, Kierzek, R. and Turner, D. H. (1991) *Biochemistry*, **30**, 8242–8251.
- Leonard, G. A., Booth, E. D., Hunter, W. N. and Brown, T. (1992) *Nucleic Acids Res.*, **20**, 4753–4759.
- 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.
- Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. and Cullen, B. R. (1989) *Nature*, **338**, 254–257.
- Pley, H. W., Flaherty, K. M. and McKay, D. B. (1994) *Nature*, **372**, 68–74.
- Scott, W. G., Finch, J. T. and Klug, A. (1995) *Cell*, **81**, 991–1002.
- Yuan, H., Quintana, J. and Dickerson, R. E. (1992) *Biochemistry*, **31**, 8009–8021.
- Grzeskowiak, K., Goodsell, D. S., Kaczor-Grzeskowiak, M., Cascio, D. and Dickerson, R. E. (1993) *Biochemistry*, **32**, 8923–8931.