Crystal Structure of an RNA 16-mer Duplex R(GCAGAGUUAAAUCUGC)$_2$ with Nonadjacent G(Syn)·A$^+$(Anti) Mispairs†

Baocheng Pan, Shome Nath Mitra, and Muttaiya Sundaralingam*

Biological Macromolecular Structure Center, Departments of Chemistry, Biochemistry, and Biophysics Program, The Ohio State University, 1060 Carmack Road, Columbus, Ohio 43210-1002

Received September 2, 1998; Revised Manuscript Received December 22, 1998

ABSTRACT: G·A mispairs are one of the most common noncanonical structural motifs of RNA. The 1.9 Å resolution crystal structure of the RNA 16-mer r(GCAGAGUUAAAUCUGC)$_2$ has been determined with two isolated or nonadjacent G·A mispairs. The molecule crystallizes with one duplex in the asymmetric unit in space group *R*3 and unit cell dimensions $a = b = c = 49.24$ Å and $\alpha = \beta = \gamma = 51.2°$. It is the longest known oligonucleotide duplex at this resolution and isomorphous to the 16-mer duplex with the C·A$^+$ mispairs [Pan, et al., (1998) J. Mol. *Biol. 283*, 977−984]. The C·A$^+$ mispair behaves like a wobble pair while the G·A$^+$ does not. The G·A mispairs are protonated at N1 of the adenines as in the C·A$^+$ mispairs, and two hydrogen bonds in the G(syn)·A$^+$(anti) conformation are formed. The syn guanine is stabilized by an intranucleotide hydrogen bond between the 2-amino and the 5′-phosphate groups. The G(syn)·A$^+$(anti) conformation can provide a different surface for recognition in the grooves compared to other G·A hydrogen bonding schemes. The major groove is widened between the two mispairs allowing access to ligands. One of the 3-fold axes is occupied by a sodium ion and a water molecule, while a second is occupied by another water molecule.

Base-pair mismatches are found in both DNA and RNA, but they are corrected by repair enzymes in DNA (*1*) and serve as recognition sites in RNA. G·A mispairs can adopt at least five possible conformations (*2*, *3*). Crystallographic studies so far have shown that G·A mispairs in DNA adopt three conformations: G(anti)·A(anti) (*4*); G(anti)·A(syn) (*5*, *6*); G(syn)·A(anti) (*7*). NMR studies on DNA have shown that the G·A conformations depend on the flanking sequences (*8*). Nonadjacent and tandem G·A mispairs are found in ribosomal RNAs (*9*). Various possible conformations of G·A mispairs in RNA and their effects on helical regularity have attracted crystal structural studies. A variety of G(anti)·A(anti) conformations have already been found in the crystal structures of r(CGC*G*AAUU*A*GCG) (*10*), r(GGCC*G*AA*A*GGCC) (*11*), and the 62-nucleotide 5S rRNA fragment (*12*) (italicized bases are mismatched). We have designed the 16-mer r(GCAGA*G*UUAA*A*UCUGC) with two isolated or nonadjacent G6·A27/A11·G22 mispairs to study their conformation(s) and helix distortions.

MATERIALS AND METHODS

Synthesis, Crystallization, and Data Collection. The 16-mer RNA r(GCAGA*G*UUAA*A*UCUGC) (GA duplex) was synthesized by the phosphoramidite method and purified by the usual procedure (*13*). The crystallization was carried out by the hanging-drop vapor diffusion method at room temperature (293 K). The best crystals were obtained with 1 mM RNA (single-stranded concentration), in the presence of 50 mM sodium HEPES buffer[1] (pH 7.0), 50 mM KCl, 5 mM $CaCl_2$, and 5% PEG400 equilibrated with 30% PEG400 in the reservoir. A crystal of dimensions 0.3 mm × 0.3 mm × 0.2 mm was mounted in a thin-walled glass capillary with some mother liquor at one end and sealed for intensity data collection. The 1.90 Å resolution intensity data comprising 5521 independent reflections ($F \geq 2\sigma(F)$) were collected at room temperature on our in-house R-AXIS IIc imaging plate and a 50 kV/100 mA graphite-monochromated Cu Kα X-ray beam. The R_{merge} was 7.3%. The data set was 81.0% complete, and the last resolution bin (1.99−1.90 Å) had 52.5% of the data. The data were processed using the software Version 2.1 from the Molecular Structure Corp. The crystal belonged to the rhombohedral space group *R*3 with unit cell constants $a = b = c = 49.25$ Å and $\alpha = \beta = \gamma = 51.2°$. The asymmetric unit contains one duplex with a volume/per base pair of 1390 Å^3. The structure was solved and refined in the hexagonal setting with cell constants $a = b = 42.53$ Å, $c = 128.06$ Å, and $\gamma = 120.0°$ (Table 1).

Structure Solution and Refinement. The *a* and the *b* axes for the GA duplex are the same as the 16-mer CA duplex r(GCAGACUUAAAUCUGC)$_2$ recently solved by us (*13*) while the *c*-axis is about 3% longer. The coordinates of the CA duplex (NDB ID No. AR0002) were used to start the refinement of the GA duplex by X-PLOR (*14*). A total of

† We gratefully thank the National Institutes of Health for Grant GM-17378 and the Board of Regents of Ohio for an Eminent Scholar endowment in support of this work. We also acknowledge the Hayes Investment Fund Program for the partial support toward the purchase of an R-axis II c imaging plate system.

* To whom correspondence should be addressed. Phone: (614) 292-2925. Fax: (614) 292-2524. E-mail: sundaral@chemistry.ohio-state.edu.

[1] Abbreviations: NMR, nuclear magnetic resonance; FPLC, fast performance liquid chromatography; HEPES − 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid, PEG − poly(ethylene glycol).

10.1021/bi982122y CCC: $18.00
Published on Web 02/13/1999

Table 1: Crystal Data and Refinement Statistics for the 16-mer RNA r(GCAGAGUUAAAUCUGC)

cryst system	rhombohedral
space group	*R*3
cell params (hexagonal setting)	
$a = b$ (Å)	42.53
c (Å)	128.06
γ (deg)	120
V/base pair (Å^3)	1390
resolution (Å)	1.90
no. of unique reflns ($\geq 2\sigma(F)$)	5521
data completeness (%): overall (1.99−1.90 Å)	81.0 (52.5)
R_{sym} (%)	7.3
resolution ranged used (Å)	8.0−1.9
no. of reflns used ($\geq 2\sigma(F)$)	5433
R_{work}/R_{free} (%)	19.7/29.6
est coordinate error (Luzzati plot) (Å) 0.29	
rms dev from target values	
param file	param_nd.dna
bond length (Å)	0.005
bond angle (deg)	1.0
dihedral angle (deg)	12.6
improper angle (deg)	1.3

5433 unique reflections ($F \geq 2\sigma(F)$) between 8.0 and 1.9 Å resolution, representing 80.8% data, were used in the structure refinement. R_{free} was calculated with 10% of randomly selected reflections (*15*). The rigid body refinement dropped the R_{work}/R_{free} to 0.342/0.410 which further dropped to 0.278/0.329 on positional refinement. At this stage, omit $|F_o| - |F_c|$ maps were calculated by removing C6 and C22, one base at a time. The maps clearly indicated that both C's were substituted by G's in the syn conformation (Figure 1a). Further refinement by Powell conjugate gradient energy minimization followed by simulated annealing dropped R_{work}/R_{free} to 0.248/0.315. The 55 water molecules and 1 sodium ion were located at various steps of refinement from the $|F_o| - |F_c|$ difference density maps at or above 3σ level which simultaneously satisfied 2|Fo| − |Fc| maps at the 1σ level. Only those sites were accepted as water molecules which were at hydrogen-bonding distances to the RNA molecule or neighboring water molecules. Individual *B*-factor and positional refinements dropped the R_{work}/R_{free} to 19.7%/29.6%. The model contains 678 nucleic acid atoms, 55 water molecules, and a sodium ion. The root-mean-square deviation of the model from the ideal geometry was 0.005 Å in bond lengths and 0.9° in bond angles (Table 1). The atomic coordinates and structure factors have been deposited with the Nucleic Acid Database (NDB ID No. AR0006) (*16*).

RESULTS AND DISCUSSION

Global Conformation. The hexadecamer r(GCAGAGUUAAAUCUGC) crystallized as a right-handed A-form duplex with two nonadjacent G·A mispairs separated by four Watson−Crick base pairs. The helical axis changes its direction at or near the two mispairs producing a curved helical axis (Figure 2). The helical axis makes an angle of 21° at the two sites where the axis changes its path from a straight helix. The program CURVES was used for the calculation of the helical axis (*17*). Comparison with fiber A-RNA indicates that the 16-mer has an end-to-end bending angle of only 3.8°. The helical axis is displaced from linearity by 1.28 Å at the A11·G22 mispair and 1.30 Å at the other mispair G6·A27.

Table 2a summarizes the torsion angles for the duplex. The three residues G6, G22, and U30 have the trans, trans

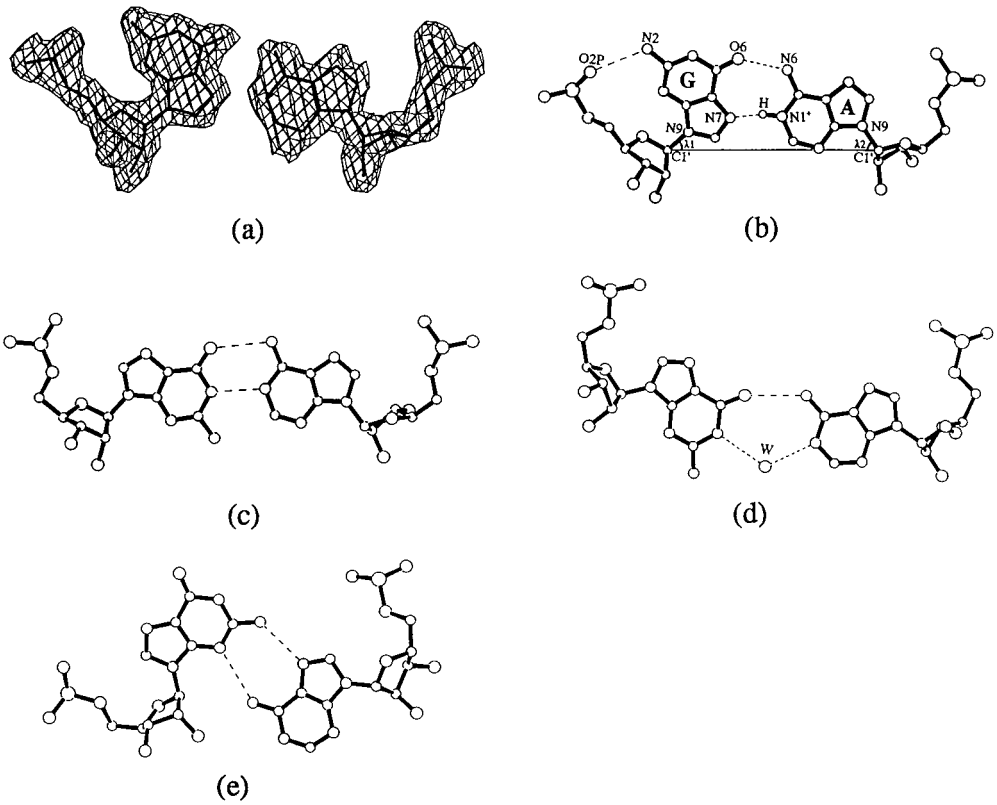

FIGURE 1: (a) The $|F_o| - |F_c|$ omit map showing the G(syn)·A$^+$(anti) in the present structure. (b) Stick drawing of the mismatch conformation. (c) G(anti)·A(anti) conformation in r(CGCGAAUUAGCG) (Leonard et al. (1994)). (d) Water-mediated G(anti)·A(anti) conformation in r(CCGAUGGUAGUG)/r(GCGAGAGUAGGC) (Correll et al. (1997)). (e) "Sheared" G(anti)·A(anti) conformations in r(GGCCGAAAGGCC) (Baeyens et al. (1996)) and r(CCGAUGGUAGUG)/r(GCGAGAGUAGGC) (Correll et al. (1997)).

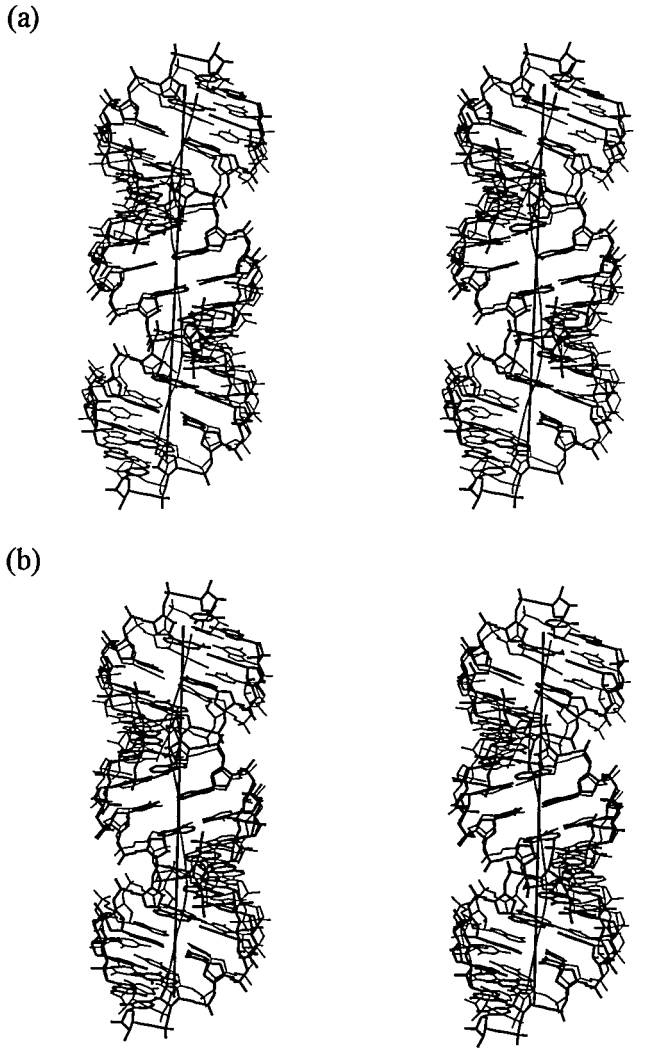

FIGURE 2: Stereoviews of the 16-mers with their curved helical axes, superimposed on the fiber A-RNA linear helix, showing the reversal of directions for the axis at the mispair sites: (a) GA duplex; (b) CA duplex.

conformation for α and γ compared to the more favorable g^-, g^+ conformation (*18*) for the other 27 residues. For the trans, trans conformation, α is considerably smaller than 180° while γ is near 180°; β is above 180° in strand 1 while near 180° in strand 2. It is to be noted that the trans, trans conformation occurs only in one strand at the site where the helical axis reverses its direction. These backbone conformations extend the intra phosphorus−phosphorus distances to 6.61 Å for P6−P7, 6.60 Å for P22−P23 and 6.61 Å for P30−P31 compared to the average of all the P−P distances, 5.74 ± 0.23 Å. Superposition with the A-RNA fiber shows that the divergence of the phosphorus atoms of these three residues are significantly higher (root-mean-square deviations of 2.3−2.5 Å) than the rest of the structure. The means of all the torsion angles (α, β, γ, δ, ϵ, and ζ) excluding G6, G22, and U30 are very close to the A-RNA structure.

Table 2b summarizes the helical parameters calculated with the program NEWHEL92 (*19*). The average twist angle is 32.4°, close to the fiber A-RNA. The largest twist angles of 36 and 37° of the 16-mer structure are at the penultimate steps. The average rise is 2.68 Å, and higher values are observed next to the mispairs. The roll angles for the successive base pair steps vary significantly, and the largest

Table 2: (a) Torsion Angles (deg) and (b) Helical Parameters of r(GCAGAGUUAAAUCUGC)A)

(a) Torsion Angles

	α	β	γ	δ	ϵ	ζ	χ
G1			71.8	87.4	228.2	277.6	184.5
C2	299.0	159.6	54.1	80.2	211.9	285.7	197.1
A3	294.8	172.5	49.7	77.2	207.2	283.6	196.4
G4	303.5	175.6	50.9	80.6	209.1	285.2	199.1
A5	305.0	179.3	45.0	80.8	199.6	284.1	205.5
G6	143.0	195.5	187.0	81.0	225.9	289.6	0.9
U7	295.9	175.8	50.8	81.8	205.9	288.7	195.6
U8	295.2	176.8	53.3	81.0	206.8	280.4	201.4
A9	302.7	170.2	46.7	80.4	209.0	279.5	201.3
A10	310.6	163.6	49.2	80.8	211.7	279.8	199.9
A11	302.3	164.4	54.1	81.8	212.1	280.8	205.6
U12	294.4	175.3	50.1	81.8	216.8	286.3	205.6
C13	294.0	171.4	53.3	81.2	199.2	297.3	204.4
U14	280.2	178.8	63.5	76.2	192.4	291.1	190.5
G15	283.6	184.9	59.3	83.9	211.5	287.0	196.5
C16	303.9	170.3	53.8	79.6			206.4
G17			67.2	85.2	218.6	285.8	186.8
C18	293.2	168.5	53.6	80.5	208.5	288.5	200.4
A19	297.4	177.3	45.1	78.9	210.8	279.6	197.6
G20	314.4	169.5	44.4	78.7	200.9	294.6	198.5
A21	291.6	189.3	50.3	84.7	193.9	290.5	207.3
G22	149.2	187.6	181.9	83.2	222.0	290.1	6.9
U23	292.0	175.2	50.6	81.4	207.5	285.3	194.9
U24	298.8	171.9	56.6	81.7	207.7	280.4	197.8
A25	301.2	171.3	46.3	80.3	202.7	287.5	207.1
A26	303.5	170.7	51.6	80.3	213.5	279.3	199.2
A27	297.4	161.6	59.5	82.5	214.2	277.8	205.1
U28	293.4	177.7	50.0	82.7	218.1	279.8	202.7
C29	309.9	167.0	43.8	78.8	185.5	297.7	209.0
U30	160.3	170.6	177.4	88.7	216.6	293.1	187.5
G31	291.3	179.5	54.9	83.0	214.2	283.7	197.6
C32	308.7	164.3	50.5	78.6			206.0
av	283.7	173.9	64.9	81.4	209.4	285.7	187.4
SD	45.7	8.2	38.8	2.6	9.5	5.7	48.5

(b) Helical Parameters

	twist (deg)	rise (Å)	roll (deg)	tilt (deg)	propeller (deg)	slide (Å)
G1·C32					−17.1	
	33.6	2.38	−10.16	0.35		−2.87
C2·G31					−15.9	
	36.4	2.24	−5.59	−4.13		−3.21
A3·U30					−14.0	
	30.8	2.60	1.82	−3.78		−3.33
G4·C29					−13.6	
	30.5	3.29	9.12	−2.97		−3.00
A5·U28					−12.5	
	32.8	3.02	−4.16	5.17		−3.18
G6·A27					−13.9	
	32.7	2.70	8.01	1.08		−3.15
U7·A26					−10.0	
	29.9	2.56	−3.00	2.22		−2.99
U8·A25					−9.8	
	33.1	2.73	−6.95	−0.62		−2.68
A9·U24					−3.2	
	29.7	2.67	−1.93	−1.73		−2.95
A10·U23					−11.1	
	33.2	2.68	11.92	−2.63		−3.10
A11·G22					−12.8	
	32.6	2.96	−7.52	−2.74		−3.10
U12·A21					−16.5	
	30.5	3.27	9.93	0.96		−3.07
C13·G20					−10.9	
	29.8	2.62	0.92	4.74		−3.29
U14·A19					−9.3	
	37.5	2.18	−4.77	3.75		−3.16
G15·C18					−15.9	
	33.0	2.36	−11.62	0.49		−2.98
C16·G17					−11.2	
av	32.4	2.68	−0.93	0.01	−12.4	−3.07
SD	2.3	0.34	7.59	3.03	3.5	0.17

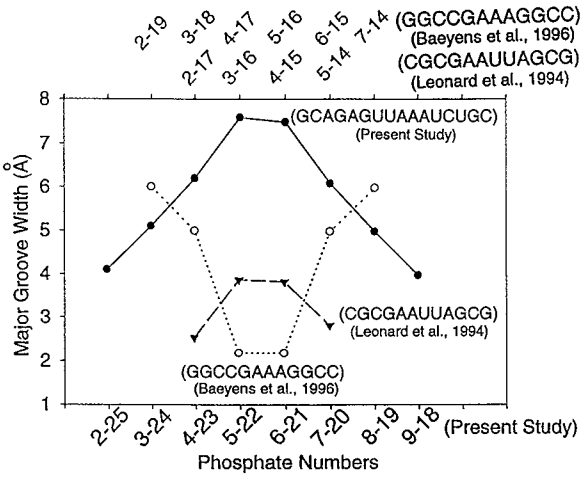

FIGURE 3: Major groove widths (P−P distances across the groove less 5.8 Å) in the present study compared to r(CGCGAAUUAGCG) (Leonard et al. (1994)) and r(GGCCGAAACCGG) (Baeyens et al. (1996)).

differences between adjacent steps are seen around the mispairs. In addition to the trans, trans conformation for the two guanines, the change in roll angles, which are maximum at either end of the mispairs, results in the deviation of the course of helical axis from linearity (*20*). The molecule is mainly composed of A·U base pairs at the middle. The propeller twist of the central A·U base pair is constricted, and the G·C pairs at the ends are higher. The average tilt and slide are 0° and −3.07 Å, respectively.

G(syn)·A$^+$(anti) Mispairs. The glycosyl conformations for the mismatched base pairs are G(syn)·A(anti) (Figure 1b) with two hydrogen bonds, N6(A)···O6(G) (2.68/2.75 Å) and N1(A)···N7(G) (2.57/2.50 Å). The shorter distances are observed for the adenines protonated at N1. The C1′−C1′ distances for the two G·A$^+$ mispairs are 10.54/10.42 Å and the $\lambda 1$ and $\lambda 2$ are 41.4°/42.9° and 56.0°/56.2°, respectively (Figure 1b). The nucleotides can display both anti and syn glycosyl conformations, even though the anti is more frequent. Further, guanine is more prone to the syn conformation than any other nucleotides (*21*). Theoretical studies of guanosine 5′-monophosphate and guanosine 3′,5′-diphosphate showed that the glycosyl angle (χ) and the exocyclic C4′−C5′ torsion angle, γ, are correlated; a syn G favors trans γ (present structure), and an anti G favors g^+ (*22*, *23*). In the C3′-endo sugar puckering the syn guanine is additionally stabilized by a direct hydrogen bond between the N2 amino group of the base and the 5′-phosphate (Figure 4). While in the C2′-endo sugar puckered B-DNA, a water molecule bridges this hydrogen bond (*7*).

Crystallographic studies on DNA oligonucleotides show that the conformations of the G·A mispairs are pH- and sequence-dependent (*4*−*7*). While crystallographic studies on RNA oligonucleotides do not show a pH dependence (Table 3), it seems there is a correlation between the sequence and the type of conformation that G·A may adopt in RNA. The sequence 5′-P-*G*-R-3′ (where P is a pyrimidine and R a purine) appears to give G(anti) while the present sequence 5′-R-*G*-P-3′, the only such sequence with G·A mispair studied so far where the flanking bases are switched, adopts a G(syn) conformation. Evidence is still lacking for G·A mismatches in RNA for A(syn). Further structural data might help to determine if the sequence actually plays a role on

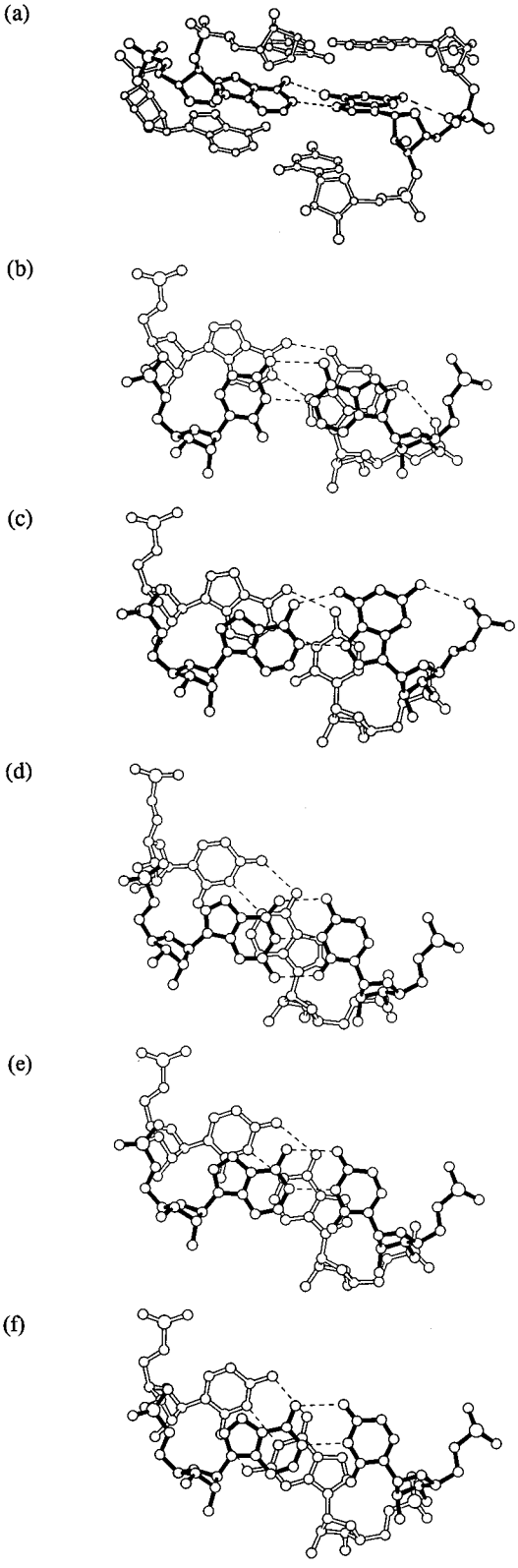

FIGURE 4: (a) Conformational features of the G(syn)·A$^+$(anti) mispairs (in dark lines), with the flanking Watson−Crick base pairs. The syn G makes the backbone torsion angle to the trans, trans conformation, and the syn conformation is further stabilized by a base−backbone hydrogen bond between the N2 amino group and the 5′-phosphate anionic oxygen atom. Base stacking occurs between b) A5·U28 (filled bonds) and G6·A27 (open bonds), (c) G6·A27 (filled bonds), and U7·A26 (open bonds), (d) C2·G31 (filled bonds) and A3·U30 (open bonds), (e) U8·A25 (filled bonds), and A9·U24 (open bonds) and (f) U14·A19 (filled bonds) and G15·C18 (open bonds).

Table 3: Comparison of Crystal Structures with G·A Mispairs in RNA

sequence	conformation	C1′−C1′ dist for G·A (Å)	pH	refs
CGC*G*AAUU*A*GCG	G(anti)·A(anti)	12.5	6.5	Leonard et al. (1994)
GGCC*G*AA*A*GGCC	G(anti)·A(anti) (sheared)	9.5	7.5	Baeyens et al. (1996)
5′CC*G***A**UGGU*A*GUG3′ [a] 3′CGGAUGAG**A***G*CG5′	G(anti)·A(anti) (sheared)	9.6	6.0	Correll et al. (1997)
5′CCGAU*G*GUAGUG3′ 3′CGGAUG*A*GAGCG5′	G(anti)·A(anti) (water-mediated)	14.8	6.0	Correll et al. (1997)
GCAGA*G*UUAA*A*UCUGC	G(syn)·A^+(anti)	10.5	7.0	this work

[a] The bold **A**'s are involved in reverse Hoogsteen A·U pairing.

the types of conformation of the G·A mismatches. This information can help in the prediction of the conformations of unknown RNA systems.

Recognizable Site and Major Groove Widening. The duplex has three interstrand purine−purine stacks at step 2 (C2·G31/A3·U30), step 8 (U8·A25/A9·U24), and step 14 (U14·A19/G15·C17) which locally unstack the two pyrimidines in the major groove (Figure 4). Stacking of the mispair G6·A27 on the adjacent U7·A26 base pair exposes the N2 amino group of G6 in the major groove for interaction (Figure 4d). Similar observation is made for the mispair A11·G22. Thus, the unstacking produces a potential recognizable surface for interaction.

The two mispairs expand the major groove to 7.6 Å at the center (Figure 3) while the minor groove widths remain unaffected, ranging from 9.6 to 10.6 Å. The unstacked G6 and G22 bases provide interaction sites for ligands while the expanded major groove might allow their entry. It may be noted that in the structure with G(anti)·A(anti) conformation the major groove is unaffected (*10*) while in the sheared G(anti)·A(anti) structure the major groove width is significantly narrowed to 2.2 Å (*11*). Therefore, the polymorphic G·A mispair can affect the grooves of the global structure differently indicating that the interacting proteins or ligands may affect the local conformation at the mispair sites and distort the helix.

Solvents, Metal Ion, and Crystal Packing. Of the 55 water molecules in the structure, 50 are in the first shell directly interacting with the RNA duplex, 12 in the major groove, 7 in the minor groove, and 31 hydrating the anionic phosphates of the backbone. There are some short water bridges between adjacent phosphates, water bridges between phosphates, and N7 atoms of G and A and N6 atoms of A. One water molecule on the 3-fold axis is tetrahedrally coordinated, making direct contacts with the phosphates of U30 (Figure 5a). Thus, U30 adopts the trans, trans conformation for the α and γ torsions. A water molecule on the other 3-fold axis makes hydrogen bonds with 2′-OH of U12, O4′ of C13, and O2 of C13* (symmetry related to C13) through a bridging water molecule (Figure 5b). A sodium ion located on the 3-fold axis is octahedrally coordinated to a water molecule and phosphate of U14 at distances of 2.5 and 2.6 Å, respectively (Figure 5c). The molecules are packed in the head-to-tail fashion to form endless helical columns resembling the packing of the RNA octamer r$(CCCCGGGG)_2$ in a rhombohedral unit cell (*24*). The crystal packing and structural features of the CA duplex are nearly the same as in this structure, but the sodium ion and the water molecules

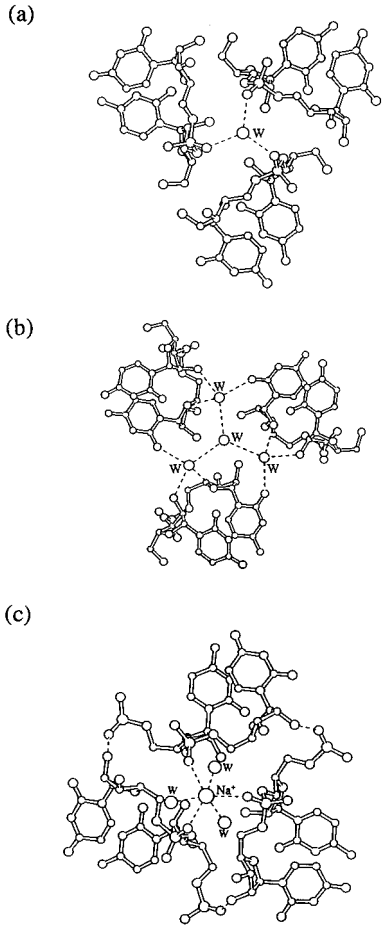

FIGURE 5: Interactions of the water molecules and the sodium ion located on the 3-fold axis: (a) One of the water molecules directly hydrogen bonding with O2P of the phosphate groups; (b) the other water molecule interacting with the RNA duplex through water molecules; (c) the Na^+ ion in an octahedral coordination interacting with water and O2P atoms.

on the 3-fold axis found in the present structure could not be located in the former, perhaps due to the limited resolution (2.5 Å).

Comparison of the GA and CA Duplexes. The present structure is isomorphous to the 16-mer CA duplex (*13*). The two sequences differ in that C6 and C22 in the CA duplex have been replaced by G6 and G22 in the GA duplex. However, the replaced guanines are syn while the cytidines are anti giving, therefore, different geometries for the mismatches. Both 16-mers crystallized in the same crystal lattice with N1 of adenine in the mispairs protonated. The overall conformations are very similar; the root-mean-square deviation on superposition of the common atoms is 0.60 Å.

Another difference in the GA duplex concerns the backbone conformations of G6 and G22 which are trans, trans (α, γ) while those of C6 and C22 in the CA duplex are g^-,g^+. The only other residue in the trans, trans conformation in the GA duplex is U30, while only A10 in the CA duplex is trans, trans. However, in both structures the helical axes follow similar course by changing their directions near the mismatches, which might be related to a systematic variation of the roll angles around the mispairs. The other major difference in the GA and CA duplexes is in the twist angles between the flanking Watson−Crick base pairs and the mispairs. In the CA duplex the twist angles are different; on one side of the mispairs it is higher (36°, 37°) than the average of 32°, while on the other side it is lower (27°, 26°). In the GA duplex they are close to the average (Table 2). The two structures display almost similar base stacking patterns, but at the mismatches, GA duplex in particular, the preference is for intrastrand purine−purine overlap. The difference is manifested in the base stacking of the mispairs with the 3′-side Watson−Crick pairs. In the wobble C·A^+ pair, the cytosine rotates toward the major groove and the A^+ translates toward the minor groove (*13*) and stacks with the flanking 3′-side Watson−Crick pair, the adenine stacks over adenine with slight overlap, and the two pyrimidines are totally unstacked (Figure 3c in ref 13). In the GA duplex, the mismatched pairs stack with slightly better overlap of the A on one side and the G and U on the other side are unstacked. It is also interesting to note that the two significantly different mispairs having varied effects at the local conformation around them still impart very similar global changes in the helix, e.g., opening up the major groove and local perturbations in the helical axes. The positions of the G·A^+ and C·A^+ mispairs in the duplexes may be a determinant in having similar effects on the groove widths.

Biological Significance of G·A Mispairs. G·A mispairs are commonly found in the secondary structural motifs of biological RNA molecules including tRNAs (*25*), ribosomal RNA (*9*, *26*, *27*), hammerhead ribozyme (*28*), and group I intron from *Tetrahymena thermophilla* (*29*). The appendant atoms of the mispairs afford tertiary interactions (*30*−*32*), serve as recognition sites for proteins (*12*, *33*), and also stabilize three-dimensional structures, for example, the GAAA tetraloop in ribozymes (*28*, *34*). These functions may be achieved by the different conformations of the G·A mispairs. This study shows that G bases can adopt the syn conformation and indicates that the G·A mispairs in RNA are polymorphic as in DNA. Stacking of the G·A mispairs with the flanking Watson−Crick base pair at the 3′-sides leaves the syn guanines unstacked so that the exocyclic 2-amino group may serve as the potential recognition site for proteins in the major groove. The widening of the major groove by the mispairs may produce accessibility to interacting proteins or any other ligands.

REFERENCES

1. Modrich, P. (1987) *Annu. Rev. Biochem. 56*, 435−466.
2. Gautheret, D., Konings, D., and Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
3. Pan, B., and Sundaralingam, M. (1999) *Int. J. Quantum Chem.* (in press).
4. Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., and Dickerson, R. (1987) *Science 238*, 498−504.
5. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2402−2406.
6. Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., and Tickle, I. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693−6697.
7. Brown, T., Leonard, G. A., Booth, E. D., and Chambers, J. (1989) *J. Mol. Biol. 207*, 455−457.
8. Cheng, J.-W., Chou, S.-H., and Reid, B. R. (1992) *J. Mol. Biol. 228*, 1037−1041.
9. Gutell, R. R., Larsen, N., and Woese, C. R. (1994) *Microbiol. Rev. 58*, 10−26.
10. Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T., and Hunter, W. N. (1994) *Structure 2*, 483−494.
11. Baeyens, K. J., De Bondt, H. L., Pardi, A., and Holbrook, S. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12851−12855.
12. Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) *Cell 91*, 705−712.
13. Pan, B., Mitra, S. N., and Sundaralingam, M. (1998) *J. Mol. Biol. 283*, 977−984.
14. Brunger, A. T. (1992) *XPLOR Manual, Version 3.0*, Yale University, New Haven, CT.
15. Kleywegt, G. J., and Brunger, A. T. (1996) *Structure 4*, 897−904.
16. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R., and Schneider, B. (1992) *Biophys. J. 63*, 751−759.
17. Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 4*, 655−667.
18. Sundaralingam, M. (1973) in *Conformation of Biological Molecules and Polymers, Proceedings of Jerusalem Symposium* (Bergmann, E., and Pullman, B., Eds.) Vol. 5, pp 417−456, Jerusalem Academic Press.
19. Fratini, A. V., Kopka, M. L., Drew, H. R., and Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686−14707.
20. Calladine, C. R., and Drew, H. R. (1984) *J. Mol. Biol. 178*, 773−782.
21. Guschlbauer, W. (1972) in *The Purines-Theory and Experiment, Proceedings of Jerusalem Symposium* (Bergmann, E. D., and Pullman, B., Eds.) Vol. 4, pp 297−310, Jerusalem Academic Press.
22. Yathindra, N., and Sundaralingam, M. (1973) *Biopolymers 12*, 2261−2277.
23. Pullman, B., and Saran, A. (1976) *Prog. Nucleic Acid Res. Mol. Biol. 18*, 215−322.
24. Portmann, S., Usman, N., and Egli, M. (1995) *Biochemistry 34*, 7569−7575.
25. Kim, S.-H. (1976) *Prog. Nucl. Acid Res. Mol. Biol. 17*, 181−216.
26. Traub, W., and Sussman, J. L. (1982) *Nucleic Acids Res. 10*, 2701−2708
27. Noller, H. F. (1984) *Annu. Res. Biochem. 53*, 119−162.
28. Pley, H. W., Flaherty, K. M., and McKay, D. (1994) *Nature 372*, 111−113.
29. 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.
30. Michel., F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
31. Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol. 236*, 49−63.
32. SantaLucia, J., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
33. Zwieb, C. (1992) *J. Biol. Chem. 267*, 15650−15656.
34. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) *Science 256*, 217−219.

BI982122Y