Acta Cryst. (1998). D**54**, 570–576

1.76 Å Structure of a Pyrimidine Start Alternating A-RNA Hexamer r(CGUAC)dG

Roopa Biswas, Shome Nath Mitra and Muttaiya Sundaralingam*

Biological Macromolecular Structure Center, Departments of Chemistry, Biochemistry and The Ohio State Biochemistry Program, The Ohio State University, 100 West 18th Avenue, Columbus, OH 43210, USA. E-mail: sunda@biot.mps.ohio-state.edu

(*Received 15 May 1997; accepted 3 October 1997*)

Abstract

The crystal structure of the alternating RNA r(CGUAC)dG with a 3′-terminal deoxy G residue has been determined at 1.76 Å resolution. The crystal belongs to the orthorhombic space group $C222_1$, unit-cell dimensions a = 29.53, b = 44.61 and c = 94.18 Å, with two independent duplexes (I and II) per asymmetric unit. The structure was solved by the molecular-replacement method. The final R factor was 18.8% using 4757 reflections in the resolution range 8.0–1.76 Å. The model contains a total of 496 atoms and 85 solvent molecules. The two duplexes form the repeating unit and stack in the usual head-to-tail (5′,3′/5′,3′) fashion into a pseudocontinuous helical column. Almost all of the 2′-hydroxyl groups are engaged in the three modes of water-mediated interactions to the base N3/O2 atoms, the sugar O4′ atoms and the backbone phosphates. Thus, the 2′-hydroxyl group of RNA is probably contributing to the stability of the duplexes.

1. Introduction

More than two decades ago tRNA was the only RNA crystal structure known (Schimmel *et al.*, 1979). During the past decade other RNA crystal structures have been unravelled: the hammerhead ribozyme (Pley *et al.*, 1994; Scott *et al.*, 1995, 1996), the P4–P6 domain of group I intron (Cate *et al.*, 1996) after the model of the entire intron was proposed by Michel & Westhof (1990), complexes of RNA with proteins (Rould *et al.*, 1989; Ruff *et al.*, 1991; Biou *et al.*, 1994; Arnez & Steitz, 1994; Oubridge *et al.*, 1994; Valegard *et al.*, 1994) and duplex RNA (Dock-Bregeon *et al.*, 1988, 1989; Holbrook *et al.*, 1991; Leonard *et al.*, 1994; Betzel *et al.*, 1994; Cruse *et al.*, 1994; Portmann *et al.*, 1995; Schindelin *et al.*, 1995; Baeyens *et al.*, 1995; Wahl, Ban *et al.*, 1996). Complex RNA's, like tRNA, rRNA, ribozyme, have double helical stems, loops and other tertiary structures, while DNA's are commonly composed of duplexes. Here, we report the crystal structure of the pyrimidine–purine alternating RNA hexamer duplex, r(CGUAC)dG. The all-DNA analog d(CGTACG) of this hexamer sequence has been determined with daunomycin intercalating between the terminal dCpG steps (Wang *et al.*, 1987). It was hoped that the RNA analog with a terminal dG would also intercalate daunomycin at the rCpdG/rCprG steps.

2. Experimental

2.1. *Synthesis and purification*

The oligonucleotide RNA hexamer was synthesized using solid-phase phosphoramidite chemistry on an Applied Biosystems (ABI) nucleic acid synthesizer (Model 391). The oligomers were cleaved from the solid support using a 3:1 ammonium hydroxide/ethanol solution and deprotected overnight by incubation at 328 K in the same solution (Wahl, Ramakrishnan *et al.*, 1996). The resulting material was lyophilized and the 2′-hydroxyl groups were deprotected by incubating with 0.6 ml of tetrabutylammonium fluoride in tetrahydrofuran (THF) at room temperature for 24 h. Next a 0.6 ml volume of 0.1 *M* triethylamine acetate (TEAA) was added to the RNA, the mixture lyophilized and precipitated with ammonium acetate/ethanol. The precipitate was dissolved in 400 µl of buffer (0.05 *M* ammonium acetate, pH 7.0, 20% acetonitrile) and further purified by ion-exchange fast protein liquid chromatography using a gradient of 0–1 *M* lithium chloride in 20% acetonitrile, followed by lyophilization and ethanol precipitation. A 4 m*M* stock solution of the oligonucleotide single strand was prepared in water to carry out crystallization trials.

2.2. *Crystallization and data collection*

The hanging-drop vapor-diffusion method was used for the crystallization. Each drop contained 2 m*M* RNA (single-stranded concentration), 2.5 m*M* magnesium acetate, 0.5 m*M* spermine, 10 m*M* sodium cacodylate buffer (pH 6.5) and 0.2 m*M* daunomycin hydrochloride (DMA) against a reservoir containing 50% MPD at 291 K. Crystals appeared within 2 d and continued to grow for a week. The lack of red color in the crystals indicated the absence of daunomycin. A single crystal of

ISSN 0907-4449

dimensions 0.2 × 0.15 × 0.05 mm was mounted in a glass capillary. The crystal belongs to the orthorhombic space group $C222_1$ with unit-cell dimensions a = 29.53, b = 44.61 and c = 94.18 Å. The asymmetric unit contains two RNA duplexes with a volume of 1292 Å^3 per base pair. The X-ray data were collected to 1.76 Å resolution on the in-house R-axis IIc imaging plate with Cu $K\alpha$ radiation and a Rigaku generator operating at 50 kV and 100 mA. The data were processed using the *R-axis* software from Molecular Structure Corporation. The 4757 reflections, 76% complete, ($F \geq 2\sigma F$) in the resolution range 8.0–1.76 Å, with an $R_{\rm merge}$ of 6.7%, were used for refinement. The data completeness in the resolution range 8.0–2.0 Å is 80 and 65% in the range 2.0–1.76 Å.

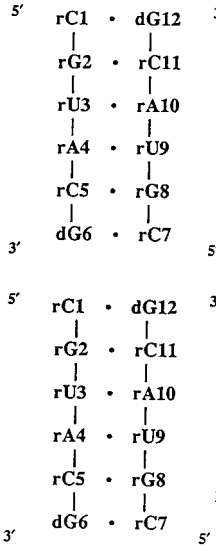
Table 1. *Crystal data and refinement statistics*

Space group	$C222_1$ (Orthorhombic)
Unit-cell dimensions (Å)	
a	29.53
b	44.61
c	94.18
Molecules per asymmetric unit	Two duplexes
Resolution range (Å)	8.0–1.76
No. of reflections ($F \geq 2.0\sigma F$)	4757
Data completeness (%)	76
Final *R* value (%)	18.8
Final model	
Nucleic acid atoms	496
Water O atoms	85
R.m.s. deviation from ideal geometry	
Parameter file	param_nd.dna
Bond lengths (Å)	0.007
Bond angles (°)	1.4

2.3. *Structure solution and refinement*

The structure of the oligonucleotide hexamer, r(CGUAC)dG, was solved by the method of molecular replacement using the program *AMoRe* (Navaza, 1994). The starting atomic coordinates were from the hexamer stretch r(UAUAUA) from the octamer sequence r(G**UAUAUA**dC (Wahl, Ban *et al.*, 1996). For the rotation search the data between 8.0 and 2.0 Å was used with a box size of 22 × 22 × 45 Å and a search radius of 10 Å. All the top 25 peaks from the rotation-function calculation were used for the translation search in *AMoRe* and subsequently for rigid-body refinement using the *FITTING* program. The top peak gave a correlation coefficient of 64.7% for one duplex, yielding an *R* factor of 51%. The second duplex, which again corresponded to the highest peak, was obtained by fixing the first duplex and performing another translation search, followed by rigid-body refinement as before. The two highest peaks gave a correlation coefficient of 75.2% and *R* factor of 47%. Refinement of the structure was continued using the program package *X-PLOR* (Brünger, 1992). After a few cycles of rigid-body and positional refinement the *R* factor dropped to 34.2% ($R_{\rm free}$ 40.8%). At this point a $2F_o - F_c$ map was calculated, displayed on *CHAIN* (Sack & Quiocho, 1992). The correct bases corresponding to the sequence r(CGUAC)dG were incorporated and the refinement continued. Positional refinement dropped the *R* factor to 30.4% ($R_{\rm free}$ 35.8%). This was followed by simulated annealing (Brünger, 1988) to remove any model bias by heating to 3000 K and slow cooling to 300 K at sampling intervals of 0.5 fs. Refinement of the individual temperature factors including all data to 1.76 Å gave an *R* factor of 25.8% ($R_{\rm free}$ 29.9%). The drug daunomycin was not found in the molecular structure. A total of 85 water molecules were picked from the electron-density maps in steps (Biswas & Sundaralingam, 1997). Positional and temperature-factor refinement of the structure including the water molecules resulted in an *R* factor of 18.8% ($R_{\rm free}$ 23.2%). The final model contains 496 atoms for the two duplexes and 85 water molecules.

The r.m.s. deviation from ideal geometry of the bond lengths and bond angles are 0.007 Å and 1.4°, respectively. The refinement statistics are listed in Table 1. The atomic coordinates and structure factors have been deposited with NDB (Berman *et al.*, 1992).†



Base Steps	Distance (Å) Molecule I	Molecule II
P2-P3	5.54	5.61
P3-P4	5.55	5.87
P4-P5	5.64	5.50
P5-P6	5.97	6.00
P8-P9	5.64	5.57
P9-P10	5.81	5.67
P10-P11	5.64	5.57
P11-P12	5.70	6.15

Fig. 1. (*a*) Schematics of the stacking of the two independent duplexes. The intrastrand phosphate–phosphate separations are also listed.

† Atomic coordinates and structure factors have been deposited with the Nucleic Acid Database (Reference: ARF0108). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0755). At the request of the authors, the atomic coordinates will remain privileged until 1st December 1998 and the structure factors will remain privileged until 1st December 1998. A list of deposited data is given at the end of this issue.

3. Results and discussions

3.1. *The molecular structure*

The asymmetric unit contains two right-handed A-RNA hexamers, r(CGUAC)dG with 9.79 base pairs per turn. The duplexes are packed in the usual head-to-tail fashion (5′,3′/5′,3′) (Figs. 1 and 2) and form the repeating unit of an infinite pseudocontinuous helix. The superposition of molecule I over molecule II gives an r.m.s. of 0.35 Å. The bases are all *anti* and the puckering of the ribo furanose sugars are C3′-*endo,* characteristic of A-RNA/A-DNA duplexes. All the backbone C4′—C5′ (γ) bonds assume the favorable *gauche*$^+$ conformation and the intrastrand phosphate–phosphate distance for the two molecules ranges from 5.50 to 6.15 Å, average 5.71 Å (Fig. 1).

3.2. *Alternating helical parameters*

The helical parameters of the pyrimidine–purine alternating duplex (Table 2), calculated with the program *NUPARM* (Bhattacharyya & Bansal, 1990), exhibit interesting alternation which are particularly pronounced for the roll and buckle. The twist and the roll angles in the alternating A-RNA duplex are higher for the 5′pyrimidine-3′purine (5′Y-3′R) steps and lower for the 5′purine-3′pyrimidine (5′R-3′Y) steps to avoid steric clashes between the base pairs (Dock-Bregeon *et al.*, 1989). Similar trends have also been observed for DNA (Calladine, 1982; Dickerson, 1983). The twist angles at the CpG steps (average 34.5°) are slightly higher than those at the UpA (average 30.3°) steps but no such trend is observed for the corresponding roll angles. The buckle and propeller twist are also alternating and exhibit somewhat lower values for the 5′-Y·R-3′ base pairs and higher values for the 5′-R·Y-3′ pairs.

3.3. *Hydration*

3.3.1. *Minor and major grooves*. Out of a total of 85 water molecules 34 are involved in the hydration of the minor and major grooves. The minor-groove N2 of guanine and O2 of cytosine of the terminal base pairs, dG12·rC1 in molecule I and dG6·rC7 in molecule II, are linked through a water molecule besides the Watson–Crick hydrogen bond. A cross-strand water bridge occurs between the minor-groove atoms in adjacent bases, U9(O2) and C5(O2) of molecule II. In the major groove both single and double water bridges are observed. In molecule I, the N7 of G8 is linked to O6 *via* two water molecules and similar water bridge occurs between the N7 and N6 of A4. Also intrastrand hydrogen bonding between adjacent bases of C5(N4) and dG6(O6) occur through a single water molecule, while a double water molecule bridges U3(O4) and A4(N7) in molecule I. In molecule II similar interactions occur, U9(O4)···W···A10(N6) and G2(N7)···W···W···U3(O4). A water molecule is commonly found to bridge the purine N7 to the O1P both in molecule I G2, A4, dG6, dG12 and molecule II dG6, A10. The base hydration of the RNA duplex is listed in Table 3.

3.3.2. *2′-Hydroxyl groups*. The 2′-hydroxyl group confers additional hydration and stability to the duplex. The 2′-hydroxyl group hydrogen bonds through a water molecule in three modes of interactions: to the base N3/O2, the sugar ring O4′ and the backbone ester O3′. In the first mode, a water molecule bridges the backbone

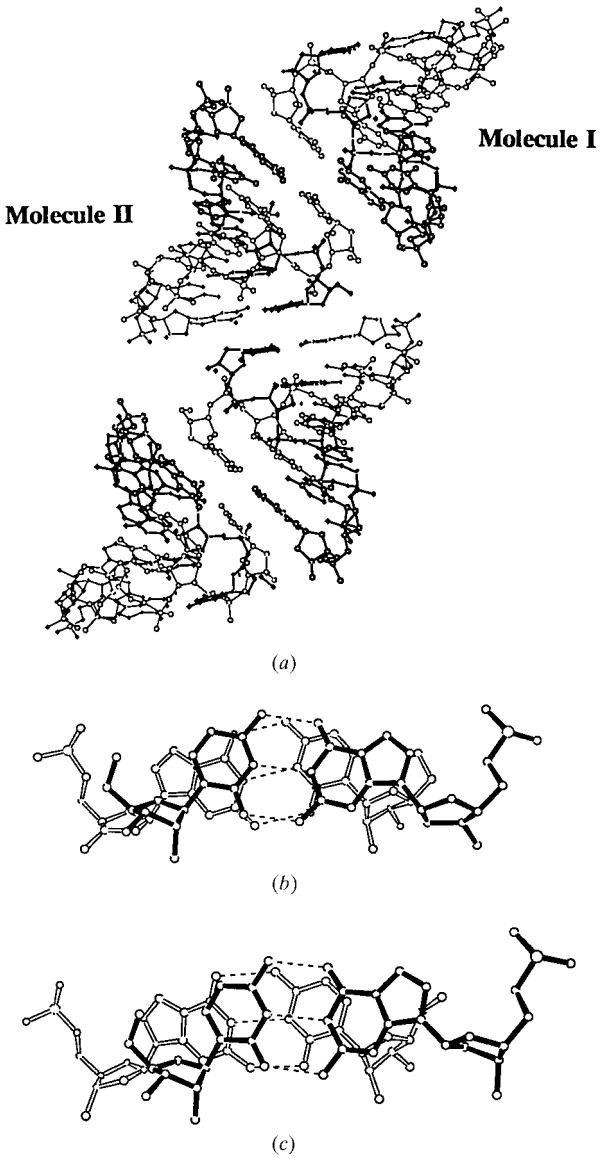


Fig. 2. (*a*) Pseudocontinuous stacking of four duplexes into a helical column. Base stacking at the two junctions of molecule I and II (*b*) intramolecular and (*c*) intermolecular stacking.

Table 2. *Helical parameters of r(CGUAC)dG*

Base steps	Twist (°)	Roll (°)	Propeller twist (°)	Buckle (°)	Dx (Å)
Duplex I					
C1·G12			−9.97	−0.59	−0.57
	35.09	18.48			
G2·C11			−16.18	2.63	−0.18
	32.65	8.21			
U3·A10			−18.35	−0.81	0.13
	30.71	16.82			
A4·U9			−8.45	3.44	0.54
	31.75	7.49			
C5·G8			−16.48	−4.55	0.39
	35.09	14.23			
dG6·C7			−11.58	4.20	
Duplex II					
C1·dG12			−13.08	−8.11	0.01
	36.00	18.64			
G2·C11			−18.91	0.10	−0.39
	30.98	12.03			
U3·A10			−13.33	−3.48	−0.23

Table 3. *Groove hydration of r(CGUAC)dG duplexes*

	Major-groove sites			Minor-groove sites		
Residue	Atom type	Water	Distance (Å)	Atom type	Water	Distance (Å)
Duplex I						
C(1)				O2	W37	3.39
G(2)	O6	W14	3.23			
	N7	W24	3.40			
	O1P		3.42			
U(3)	O4	W14	3.28	O2	W4	2.77
		W42	2.98		W66	3.48
A(4)	N7	W5	2.77	N3	W74	2.75
	N6	W42	3.50	O2′		3.41
C(5)	N4	W68	3.32			
dG(6)	N7	W10	2.48			
	O6	W68	2.63			
C(7)				O2	W41	2.49
				O2′		2.85
G(8)	O6	W27	3.04	N3	W46	3.24
	N7	W36	2.65	O2′		2.82
A(10)	N7	W7	2.90	N3	W1	2.92
				O2′		2.73
C(11)				O2	W16	3.00
				O2′		2.48
dG(12)	N7	W29	2.66	N2	W37	2.67
Duplex II						
G(2)	N7	W56	2.82	N3	W21	2.83
	O6	W31	2.62	O2′		2.61
U(3)	O4	W31	3.33	O2	W9	2.79
					W51	3.11
A(4)	N7	W28	2.49	N3	W62	2.92
C(5)				O2	W75	2.79
dG(6)	N7	W49	2.49	N2	W55	3.29
C(7)				O2		3.39
				O2′	W47	3.11
				O2		3.13
G(8)				N3	W8	2.91
				O2′		2.88
U(9)	O2	W54	2.98	O2	W75	2.94
A(10)	N7	W11	2.80	N3	W2	2.94
	N6	W54	3.48	O2′		2.88
		W31	3.28			
C(11)	N4	W11	3.49	O2	W9	2.62

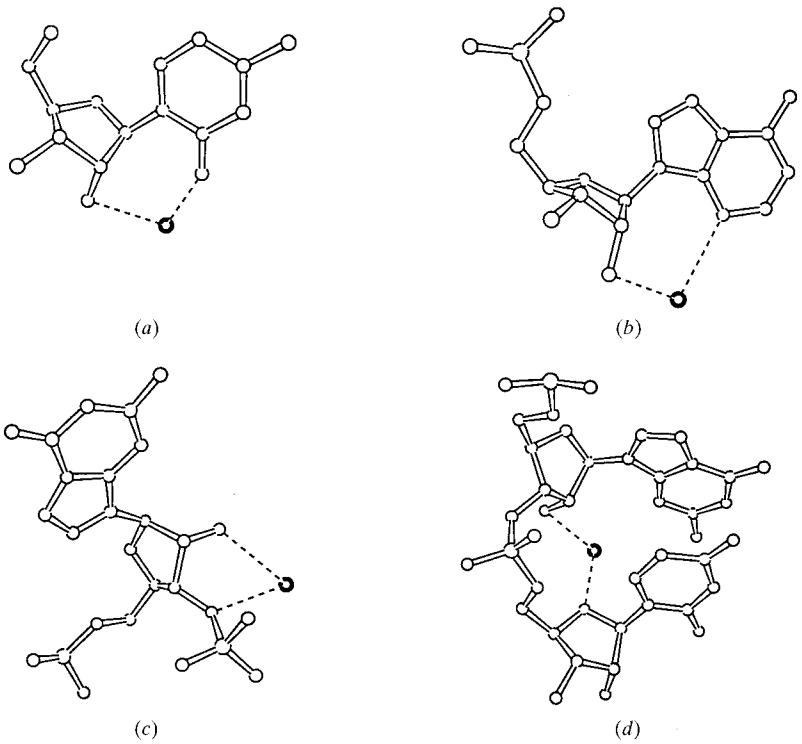


Fig. 3. Three modes of hydration of the 2′-hydroxyl group: with the base (*a*) cytosine O2 and (*b*) adenine N3, (*c*) with O3′ ester O atom and (*d*) with an adjacent sugar ring O4′ *via* water molecules.

O2′ to the bases, purine N3 or to the pyrimidine O2 (Figs. 3*a* and 3*b*), characteristic of A-RNA's (Portmann *et al.*, 1995; Wahl, Ban *et al.*, 1996). The majority of the purine bases, *viz.* G8, A4, A10 of molecule I, and G2, G8, A10 of molecule II, and some of the pyrimidine bases, *viz.* C7, C11 of molecule I and C7 of molecule II, display the above hydration pattern. It is interesting that the O2′'s of U3 and C11 of molecule II are bridged by two water molecules to the O2's of the respective bases. In the second mode, the O2′ hydroxyl group interacts with the adjacent sugar ring O4′ on the 3′ side through a water molecule (Fig. 3*c*) between G8–U9 (molecule I), U3–A4, C5–dG6, C7–G8 and C11–dG12 (molecule II). The third mode of interaction of the O2′ is with the phosphate ester O atoms, O3′, of the same residue (Fig. 3*d*) as in G2, C7, G8 of molecule I and A4 of molecule II.

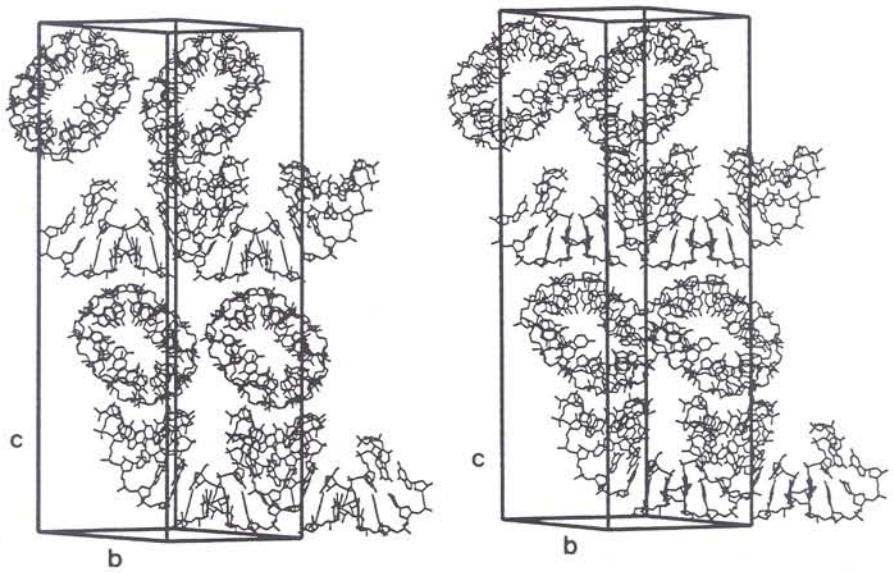
3.3.3. *Backbone phosphates*. The compressed RNA backbone allows the phosphate anionic O atoms, O1Ps, of P4–P5, P5–P6, P8–P9 of molecule I and P4–P5 of molecule II, to be bridged by a water molecule. Such O1P–O1P water bridges are commonly found in A-type duplexes, as in A-DNA (Saenger *et al.*, 1986; Wahl & Sundaralingam, 1996) and A-RNA (Portmann *et al.*, 1995; Wahl, Ban *et al.*, 1996). However, in one instance a water molecule is found to bridge the O1P of P10 to the O2P of P11, in molecule II. Another common mode of water molecule bridges is between O1P and the ester O5′ as in U3, U9 and A4 of molecule I and U3, U9, G2 and A10 of molecule II.

3.4. *Intermolecular interactions involving the 2′-OH group*

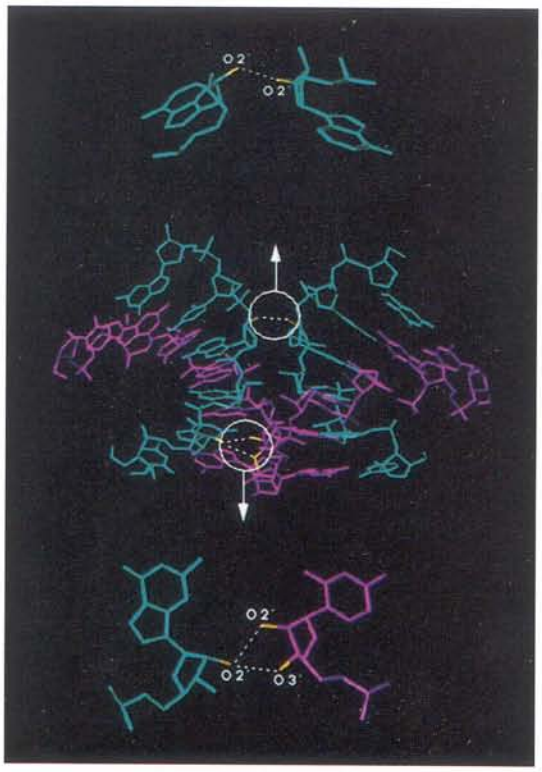
The tight packing of the duplexes in the unit cell (Fig. 4*a*) leads to a large number of intermolecular contacts involving the 2′-hydroxyl group. These are mainly direct lattice contacts and sometimes are water mediated. Direct lattice interactions occur between the 2′-hydroxyl groups of A4 and A4*† in molecule I and similarly between U9 (molecule I) and G8* (molecule II) and also between U3 and C11* (molecule II) (Fig. 4*b*). These O2′—O2′ interactions are different from the 'ribose-zipper' mode of interaction (Cate *et al.*, 1996). Intermolecular direct contacts also occur between O2′ and O4′, *viz.* O2′(C11) interacting with O4′(U3*) in molecule I and O2′(U3) interacting with O4′(C11*)in molecule II. An unusual interaction occurs between O2′ of G2 with O2′ and O3′ of C11* (Fig. 4*b*). This is similar to the intramolecular water bridge, mentioned earlier, between O2′ and O3′.

This illustrates how the 2′-hydroxyl groups are responsible for the stability of the RNA molecule.

† Asterisks symmetry-related molecules.



(*a*)



(*b*)

Fig. 4. (*a*) Stereoview of the packing in the unit cell. (*b*) Adjacent columns interact through the O2′: O2′–O2′ interactions (top) and O2′ interacting with O2′ and O3′ ester.

We thank the National Institutes of Health grants GM-17378, GM-45947, The Ohio State University and the Ohio Regents for an Eminent Scholar Endowment. We are grateful to the Ohio Board of Regents Invest-

ment Fund for the partial support to purchase the R-axis IIc instrument.

References

Arnez, J. G. & Steitz, T. A. (1994). *Biochemistry*, **33**, 7560–7567.

Baeyens, K. J., De Bondt, H. L. & Holbrook, S. R. (1995). *Nature Struct. Biol.*, **2**, 56- 62.

Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.

Betzel, C., Lorenz, S., Forste, J. P., Bald, R., Zhang, M., Schneider, T. R., Wilson, K. S. & Erdmann, V. A. (1994). *FEBS Lett.* **351**, 159–164.

Bhattacharyya, D. & Bansal, M. (1990). *J. Biomol. Struct. Dyn.* **8**, 539–572.

Biou, V., Yaremchuk, A., Tukalo, M. & Cusack, S. (1994). *Science*, **263**, 1404–1410.

Biswas, R. & Sundaralingam, M. (1997). *J. Mol. Biol.* **270**, 511–519.

Brünger, A. T. (1988) *Crystallographic Refinement by Simulated Annealing. Crystallographic Computing 4*: *Techniques and New Technologies*, edited by N. W. Issacs & M. R. Taylor, pp. 126–140. Oxford: Clarendon Press.

Brünger, A. T. (1992). *X-PLOR. A system for X-ray crystallography and NMR*. Yale University Press, New Haven, CT, USA.

Calladine, C. R. (1982). *J. Mol. Biol*. 161, 343–352.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B., Kundrot, C., Cech, T. & Doudna, J. (1996). *Science*, **273**, 1678–1685.

Cruse, W., Saludjian, P., Biala, E., Strazewski, P., Prange, T. & Kennard, O. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 416–4164.

Dickerson, R. E. (1983). *J. Mol. Biol.* **166**, 419–441.

Dock-Bregeon, A. C., Cheerier, B., Podjarny, A., Johnson, J., deBear, J. S., Gough, G. R., Gilham, P. T. & Moras, D. (1989). *J. Mol. Biol.* **209**, 459–474.

Dock-Bregeon, A. C., Cheerier, B., Podjarny, A., Moras, D., deBear, J. S., Gough, G. R., Gilham, P. T. & Johnson, J. (1988). *Nature (London)*, **335**, 375–378.

Holbrook, S. R., Cheong, C., Tinoco, I. Jr & Kim, S. H. (1991). *Nature (London)*, **353**, 579–581.

Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T. & Hunter, W. N. (1994). *Structure*, **2**, 483–494.

Michel, F. & Westhof, E. (1990). *J. Mol. Biol.* **216**, 585–610.

Navaza, J. (1994). *Acta Cryst*. A**50**, 157–160.

Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994). *Nature (London)*, **372**, 432- 438.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). *Nature (London)*, **372**, 68–74.

Portmann, S., Usman, N. & Egli, M. (1995). *Biochemistry*, **34**, 7569–7575.

Rould, M. A., Perona, J. J., Soll, D. & Steitz, T. A. (1989). *Science*, **246**, 1135–1142.

Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C. & Moras, D. (1991). *Science*, **252**, 1682–1689.

Sack, J. & Quiocho, F. A. (1992). *CHAIN Crystallographic modelling program*, Baylor College of Medicine, Houston, Texas, USA.

Saenger, W., Hunter, W. N. & Kennard, O. (1986). *Nature (London)*, **324**, 385–388.

Schimmel, P. R., Söll, D. & Abelson, J. N. (1979). *Transfer RNA: Structure, Properties and Recognition*, pp. 1–577. New York: Cold Spring Harbor Laboratory Press.

Schindelin, H., Zhang, M., Bald, R., Forste, J. P., Erdmann, V. A. & Heinemann, U. (1995). *J. Mol. Biol.* **249**, 595–603.

Scott, W. G., Finch, J. T. & Klug, A. (1995). *Cell*, **81**, 991–1002.

Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L. & Klug, A. (1996). *Science*, **274**, 2065- 2069.

Valegard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). *Nature (London)*, **371**, 623–626.

Wahl, M. C., Ban, C., Sekharudu, C., Ramakrishnan, B. & Sundaralingam, M. (1996). *Acta Cryst*. D**52**, 655–667.

Wahl, M. C., Ramakrishnan, B., Ban, C., Chen, X. & Sundaralingam, M. (1996). *Acta Cryst*. D**52**, 668–675.

Wahl, M. C. & Sundaralingam, M. (1996). *Bioploymers*, **44**, 45–63.

Wang, A. H.-J., Ughetto, G., Quigley, G. J. & Rich, A. (1987). *Biochemistry*, **26**, 1152–1163.