Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of an RNA dodecamer containing a fragment from SRP domain IV of *Escherichia coli*

Junpeng Deng, Yong Xiong, Baocheng Pan and Muttaiya Sundaralingam*

Departments of Chemistry and Biochemistry, The Ohio State University, 200 Johnston Laboratory, 176 West 19th Avenue, Columbus, Ohio 43210, USA

Correspondence e-mail: sundaral@chemistry.ohio-state.edu

Received 18 January 2003
Accepted 24 March 2003

NDB Reference: r(GCGUCAGGUC^{Br}CG)/r(CGGAAGCAG^{Br}CGC), AR0041.

PDB Reference: 1lnt, r1lntsf.

The crystal structure of an RNA dodecamer, r(GCGUCAGGUC^{Br}CG)/r(CGGAAGCAG^{Br}CGC), containing a fragment from the signal recognition particle (SRP) RNA (domain IV) of *Escherichia coli*, has been determined at 1.7 Å resolution with 21 666 independent reflections and an R_{work} and R_{free} of 20.1 and 22.5%, respectively. The structure exhibits a novel crystal packing pattern for RNA oligomer duplexes: one end of the duplex adopts the stacking interaction, while the other end adopts the abutting interaction in the minor groove. The symmetric loop of the SRP, r(CAGG)/r(AGCA), in the center of the dodecamer forms two different conformations of the A·C mismatch, a sheared G·G and a symmetrical G·A mismatch. These four mismatches present a unique surface for the abutting interaction. The involvement of the two A·C mismatches in the abutting interaction implies that these mismatches are the important sites for interaction with proteins. The conformation of the symmetric loop is greatly stabilized by hydrated metal ions, which display flexibility in adjusting their geometry and coordination in interaction with nucleic acids. Comparison with other crystal structures of fragments of 4.5S RNA indicates that the conformation of the symmetric loop is independent of the asymmetrical loop in domain IV.

1. Introduction

The signal recognition particle (SRP) is a ribonucleoprotein complex that directs translating ribosomes to the protein translocation apparatus of the endoplasmic reticulum in eukaryotes or the plasma membrane in prokaryotes. SRP RNA molecules can be divided into four structural domains (I, II, III and IV) and only domain IV is highly conserved in all homologues (Poritz *et al.*, 1988; Larsen & Zwieb, 1991). The bacterial SRP is composed of 4.5S RNA and Ffh protein. Domain IV in 4.5S RNA contains all the necessary components for specific recognition by Ffh protein (Lentzen *et al.*, 1996; Wood *et al.*, 1992). It consists of a GGAA tetraloop and symmetric and asymmetric internal loops (Lentzen *et al.*, 1996; Schmitz *et al.*, 1996; Schmitz, Behrens *et al.*, 1999; Schmitz, James *et al.*, 1999) (Fig. 1*a*). Biological studies have shown that Ffh protein binds primarily to the symmetric loop (Wood *et al.*, 1992; Selinger *et al.*, 1993; Althoff *et al.*, 1994; Zwieb *et al.*, 1996).

Recently, the crystal structures of domain IV (Jovine *et al.*, 2000) and its complex with the M domain of Ffh protein (Batey *et al.*, 2000) have been reported. Surprisingly, the RNA molecules in these two structures have essentially the same conformation. Based on this result, it has been proposed that the interaction between Ffh protein and 4.5S RNA involves rigid-body docking in the symmetry-loop region followed by

induced fit in the asymmetric loop (Jovine *et al.*, 2000). Previous studies have indicated that the asymmetric loop contributes significantly to the binding of the Ffh protein to the 4.5S SRP RNA (Schmitz *et al.*, 1996; Wood *et al.*, 1992). The asymmetric loop may function as a flexible hinge between the two adjacent double-helical segments and can change the orientation of helix c in domain IV when the M domain of Ffh protein binds to the symmetric loop (Schmitz, Behrens *et al.*, 1999; Jovine *et al.*, 2000). In the protein-bound domain IV, the 2′-hydroxyl group of A39 in the asymmetric loop is hydrogen bonded to the phosphate of A63 in the symmetric loop (Batey *et al.*, 2000), while such an interaction is missing in the free domain IV structure (Jovine *et al.*, 2000). In other words, the binding of the M domain to the symmetric loop results in a direct interaction between the symmetric loop and the asymmetric loop. Even though NMR structures have suggested that the asymmetric loop may not greatly influence the structure and stability of the symmetric loop (Schmitz *et al.*, 1996), there is not yet crystallographic evidence to show the structural correlation of these two important internal loops. We designed an RNA dodecamer containing helix *a*, the symmetric loop and part of helix *b* in order to study the conformation of the symmetric loop in the absence of the asymmetric loop (Fig. 1). In this paper, we report the crystal structure of the RNA dodecamer at 1.7 Å and study the conformation of the symmetric loop and its interaction with metal ions, hydration and stability.

2. Materials and methods

2.1. Synthesis and crystallization

The RNA dodecamer r(GCGUCAGGUC^{Br}CG)/r(CGGAAGCAG^{Br}CGC) was synthesized by the phosphoramidite method using an Applied Biosystem DNA synthesizer 391 and cleaved from the solid support using 3:1(*v*/*v*) ammonia/ethanol and incubated at room temperature overnight in the same solution. The 2′-hydroxyl groups were deprotected by triethylamine tris-hydrofluoride (TEA–3HF) for about 3 h at 328 K and purified using ion-exchange chromatography. The crystallization experiments were performed by the hanging-drop vapor-diffusion method at room temperature (293 K). The best crystals were obtained with 40 m*M* sodium cacodylate buffer pH 7.0, 20 m*M* magnesium chloride, 12.0 m*M* spermine tetrahydrochloride, 80 m*M* sodium chloride, 40 m*M* calcium chloride and 10%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD) equilabrated against 45% MPD in the reservoir. Crystals of dimensions 0.1 × 0.1 × 0.4 mm were obtained in about two weeks.

2.2. Data collection, structure solution and refinement

A set of three-wavelength MAD data was obtained from a single crystal using beamline 14BM-D at the Advanced Photon Source, Argonne National Laboratory after a fluorescence scan of the same crystal for determination of the bromine absorption-edge wavelength. The intensity data were integrated and scaled with *DENZO* and *SCALEPACK*. 2.0 Å resolution data were collected at both the absorption edge and at the peak sites of bromine. A total of 21 666 independent reflections (Friedel pairs unmerged), corresponding to 99.9% of the theoretically possible data, were collected to 1.60 Å resolution with an R_{merge} of 4.9% at the remote site. The crystal belonged to the tetragonal space group $P4_322$, with unit-cell parameters a = 35.61, b = 35.61, c = 133.96 Å. The asymmetric unit contains two RNA strands, with a volume per base pair of 1767 Å^3. The structure was solved by the MAD phasing method using the bromine anomalous signal. Phase refinement was successfully performed using *MLPHARE*/*CCP*4 (Otwinowski, 1991; Collaborative Computational Project, Number 4). The MAD-phased electron-density map could be clearly traced and the RNA chains and bases were fitted. The data set collected at the bromine remote site was used in the refinement. After simulated annealing, iterative positional and *B*-factor refinement and rebuilding of the model were carried out with *CNS* (Brünger *et al.*, 1998) and the graphics package *O*. The final R_{work} and R_{free} were 20.1 and 22.5%, respectively. The final model contains 512 RNA atoms, three Ca^{2+} ions, four Mg^{2+} ions and 111 water molecules. It should be mentioned that the Mg^{2+} ions are tetrahydrated $[Mg(H_2O)_4]^{2+}$ instead of the usual hexahydrated $[Mg(H_2O)_6]^{2+}$. Indeed, it is difficult to differentiate Mg^{2+} ions from Na^+ ions in the electron-density map because of the difference of only one electron between these two different ions. However, bond length and coordination are helpful in such identification. Na^+ ions usually have a bond length of about 2.4 Å and four or six ligands in coordination, while Mg^{2+} ions usually have bond lengths of about 2.1 Å and six ligands.

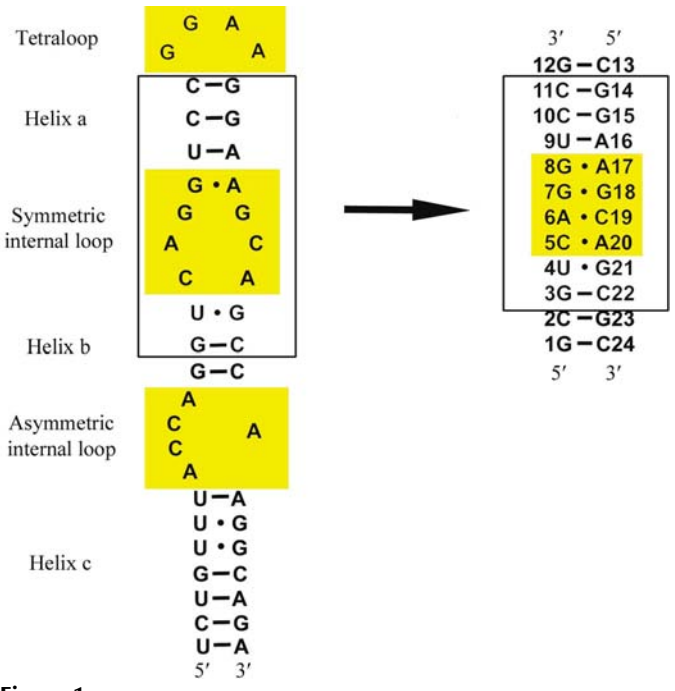

Figure 1
(*a*) Nucleotide sequence of SRP RNA domain IV from *E. coli*. The three conserved domains are highlighted in yellow. (*b*) The sequence of the present structure, with the native part of the SRP RNA domain IV from *E. coli* boxed.

Table 1
Crystal data and MAD phasing statistics for r(GCGUCAGGUC^{Br}CG)/r(CGGAAGCAG^{Br}CGC).

Values in parentheses are for the high-resolution shell (2.07–2.0 Å for the inflection and peak data sets and 1.66–1.60 Å for the remote data set).

	Inflection	Peak	Remote
Wavelength (Å)	0.9193	0.9196	0.8985
Space group	$P4_322$		
Unit-cell parameters (Å)	a = 35.61, b = 35.61, c = 133.96		
Resolution (Å)	20–2.0	20–2.0	20–1.6
Unique reflections	10840†	11382†	21666†
Redundancy	11.3 (7.5)	12.5 (8.8)	11.5 (5.8)
Completeness (%)	98 (99.8)	99(100)	99.9 (99.0)
R_{sym} (%)	3.5 (22.1)	3.7 (20.1)	4.9 (36.5)
$I/\sigma(I)$	50 (10.6)	54 (12.2)	46 (4.3)
Phasing statistics at 2.0 Å resolution			
Phasing power‡, isomorphous centric/acentric	—	0.18/0.27	0.32/0.46
R_{cullis}§			
Isomorphous centric/acentric	—	0.99/0.97	0.93/0.89
Anomalous	0.52	0.46	0.45
FOM			
Before *DM*	0.45		
After *DM*	0.76		

† Friedel pairs unmerged. ‡ Phasing power = F_H/lack of closure. § R_{cullis}: isomorphous = lack of closure/Δ_{iso}; anomalous = lack of closure/Δ_{ano}.

Table 2
Major-groove width of the dodecamer.

Phosphate group	Distance (Å)	Phosphate group
P(C2)	3.1	P(G18)
P(G3)	4.1	P(A17)
P(U4)	6.0	P(A16)
P(C5)	7.2	P(G15)
P(A6)	9.8	P(G14)

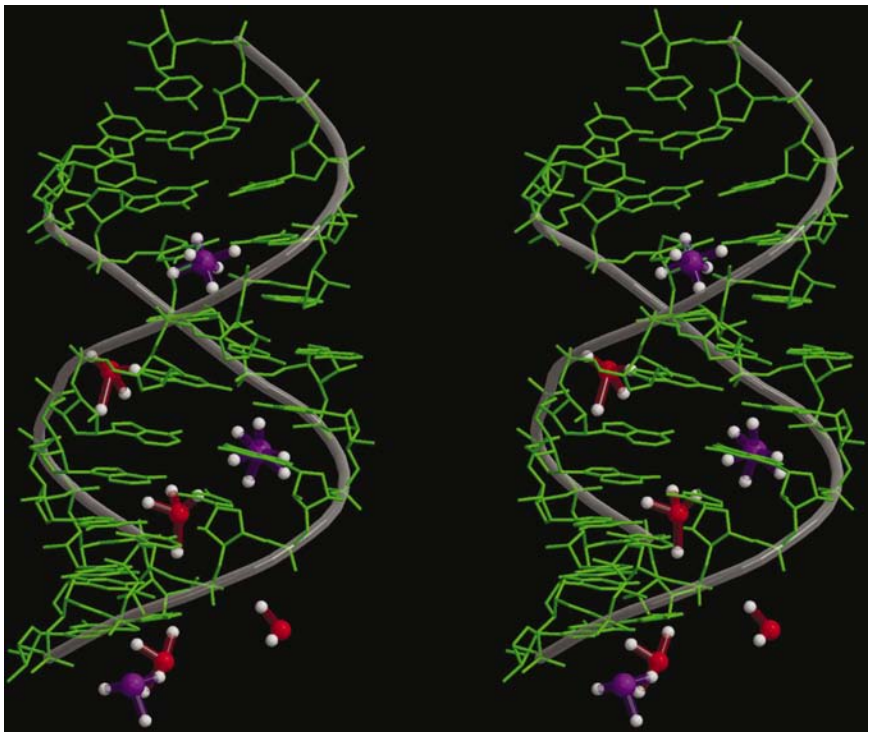

Figure 2
Stereoview of the duplex with four tetrahydrated Mg^{2+} ions (red) and three hexahydrated Ca^{2+} ions (purple).

In the present case, the average bond length for these cations is 2.2 Å. We assigned these cations as Mg^{2+} ions because of their bond lengths and our crystallization result, which implies the importance of Mg^{2+} ions. Information on the crystal, the bromo-derivative, the MAD phasing and the structure refinement are summarized in Table 1. The atomic coordinates and the structure factors have been deposited in the Nucleic Acid Database with NDB ID AR0041 and PDB code 1lnt (Berman *et al.*, 1992).

3. Results and discussion

3.1. Overall structure

The non-self-complementary RNA dodecamer r(GC**GU-BrC**G)/r(C**GGAAGCAG^{Br}C**GC) (the part of the sequence in bold is native in the 4.5S RNA in *E. coli*) crystallized as a right-handed A-RNA duplex with five mismatched base pairs in the center sandwiched by seven Watson–Crick base pairs (three at the 5′ end and four at the 3′ end) (Fig. 1*b*). There are two independent strands in the asymmetric unit. The structure is associated with three hexahydrated calcium ions and four tetrahydrated magnesium ions (Fig. 2). All the residues adopt the *anti* glycosyl conformation and C3′-*endo* sugar pucker except C5, which adopts the C4′-*exo* pucker. The major-groove width, defined as the shortest distance between the phosphate groups across the major groove less 5.8 Å, increases from one end of the duplex to the other (Table 2). Owing to the interactions O2P(A6)···N1(G18) (2.8 Å) and O2P(A6)···N2(G18) (2.9 Å), the phosphate group of A6 is pulled into the major groove, which results in major-groove widening (9.8 Å). This widening is much larger than the increase in groove width that is usually caused by mismatches (7.3 Å; Pan *et al.*, 1998, 1999). The structure exhibits a kink of 25° in the helical axis, in agreement with the previous observation in RNA duplex structures containing mismatches (Pan *et al.*, 1998, 1999; Jiang *et al.*, 1998).

Table 3 summarizes the torsion angles and helical parameters of the present dodecamer duplex calculated with the program *CURVES* (Lavery & Sklenar, 1989). Despite the non-Watson–Crick base pairings and irregularities in the helical conformation, the backbone and glycosyl torsion angles have similar values as the canonical RNA duplex. In other words, irregularities in helical conformation do not necessarily result in irregularities in the torsion angles of the helices. The dodecamer has the same average twist as and larger rise than the canonical RNA duplex. However, the twist angles at each step show great differences, especially for the symmetric loop (Table 3*b*). The very small twist (12°)

between base pairs U4:G21/C5:A20 and the large twist (65°) between C5:A20/A6:C19 result from the protrusion of C5 into the major groove. The similar abrupt trend in twists occurs between base pairs G8:A17/G7:G18 and G7:G18/A6:C19 arises from the protrusion of G18 into the major groove, which stacks well with C5.

Table 3
Torsion angles and helical parameters of the present structure.

(*a*) Torsion angles (°).

	P–O5′ α	O5′–C5′ β	C5′–C4′ γ	C4′–C3′ δ	C3′–O3′ ε	O3′– P ζ	C1′–N χ
G1			53	85	213	287	194
C2	295	173	47	82	207	288	198
G3	300	178	49	81	214	289	199
U4	298	168	48	82	216	304	203
C5	303	162	76	77	197	212	191
A6	300	163	51	77	233	300	179
G7	297	168	52	76	220	298	193
G8	290	172	53	82	204	298	190
U9	295	181	59	78	214	283	192
C10	307	167	56	81	204	292	194
C11	291	183	51	80	200	297	198
G12	293	187	51	80			207
C13			39	82	207	288	191
G14	293	176	52	80	208	292	189
G15	297	182	53	82	208	293	191
A16	291	174	62	73	195	282	193
A17	306	167	53	82	214	302	198
G18	299	169	58	185	83	257	202
C19	292	180	49	84	217	291	196
A20	290	174	58	80	228	290	189
G21	300	169	58	84	210	296	190
C22	288	178	48	81	200	292	201
G23	292	182	53	82	210	288	202
C24	303	176	44	79			208
Average	296	174	53	81	209	287	195
A-RNA	298	180	47	83	218	286	194

(*b*) Helical parameters.

	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (°)	Roll (°)	Twist† (°)
G1:C24						
	−0.1	−1.7	3.3	1	9	34
C2:G23						
	−0.1	−2.1	3.5	0	12	34
G3:C22						
	−0.2	−1.8	3.0	2	7	29
U4:G21						
	0.9	−2.8	3.8	2	7	12
C5:A20						
	−5.2	−0.6	2.8	0	10	65
A6:C19						
	4.5	−1.3	3.0	1	4	50
G7:G18						
	0.0	−2.1	3.4	−2	6	19
G8:A17						
	0.6	−2.5	3.3	−4	−2	27
U9:A16						
	0.4	−2.7	3.3	−3	1	28
C10:G15						
	−0.4	−2.6	3.5	−1	3	34
C11:G14						
	0.0	−2.2	3.4	2	6	32
G12:C13						
Average	0.1	−2.0	3.0	0	6	33

† Twist angles were measured directly from the structure.

3.2. Conformation and stability of the symmetric loop

The symmetric loop consists of five non-canonical base pairs. The G·U mismatch adopts the wobble conformation with two hydrogen bonds: N3(U4)···O6(G21) (2.7 Å) and N1(G21)···O2(U4) (3.0 Å) (Fig. 3*a*). The two tandem A·C mismatches adopt completely different conformations. There is one hydrogen bond in C5·A20: N6(A20)···O2(C5) (2.9 Å).

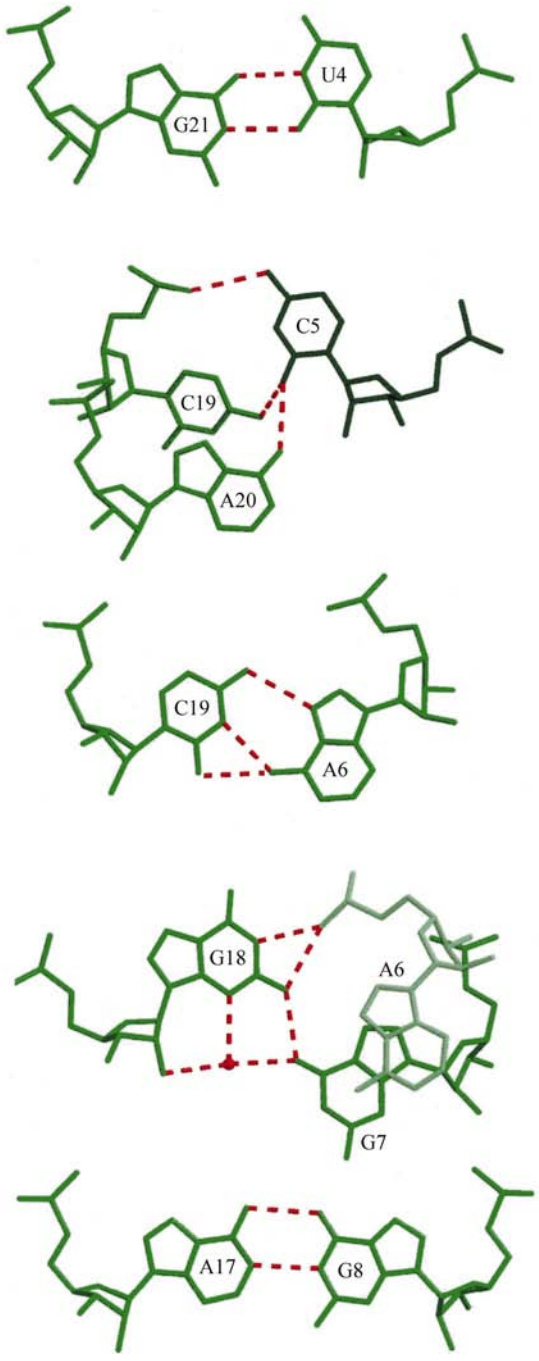

Figure 3
The five consecutive mismatches in the symmetric loop. (*a*) Wobble U4·G21; (*b*) C5·A20 with O2 (C5), forming bifurcated hydrogen bonds with N6 (A20) and N4 (C19); (*c*) reverse-Hoogsteen A6·C19; (*d*) G7·G18 with N2(G7), forming bifurcated hydrogen bonds with O6(G18) and O1P(A17); (*e*) G8·A17 in the symmetric conformation.

Figure 4
Stereoview of the crystal packing in the present structure. The central duplex (green) stacks with the duplex on its 5′ side (yellow) and abuts into the minor groove of the duplex on its 3′ side (blue). Another duplex (magenta) abuts into the central duplex (green).

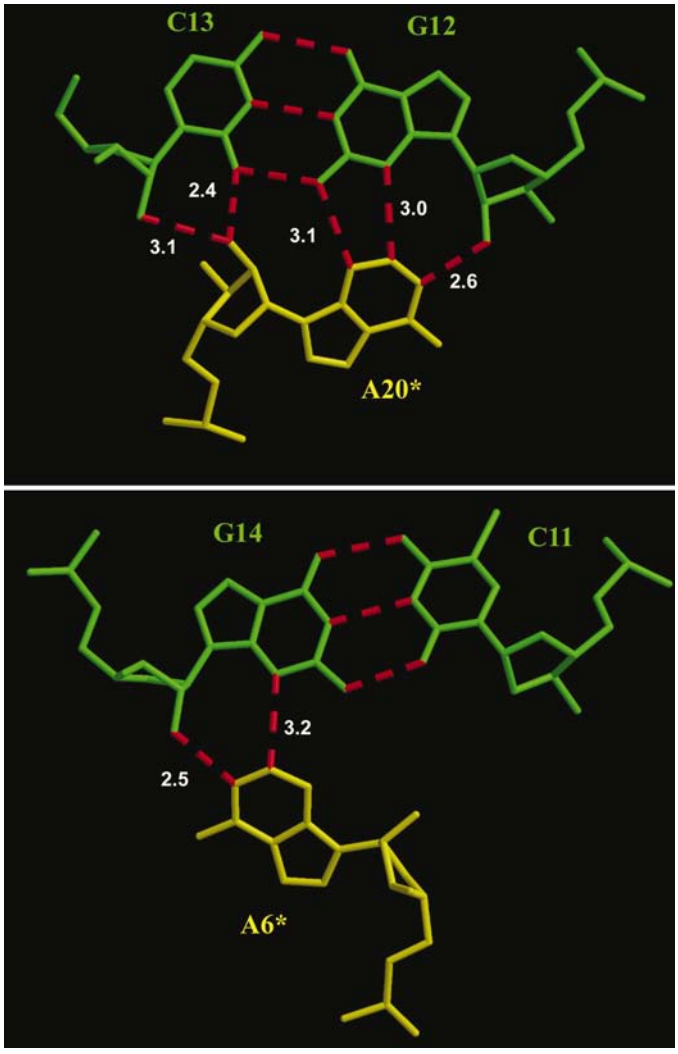

Figure 5
Two base triplets in the minor groove as a result of the abutting interaction. All values are in Å. * denotes bases from the symmetry-related duplex.

Because it lies approximately halfway between the planes of A20 and the A21·C19 base pair, C5 also forms two hydrogen bonds with C19: N4(C19)···O2 (C5) (3.0 Å) and N4(C5)···O2P(C19) (3.5 Å) (Fig. 3*b*). The A6·C19 base pair adopts the reverse-Hoogsteen base pairing with three hydrogen bonds: N6(A6)···O2(C19) (3.1 Å), N6(A6)···N3(C19) (2.9 Å) and N4(C19)···N7(A6) (3.1 Å) (Fig. 3*c*). In such a conformation, the tandem A·C base pairs expose N1 and N3 atoms of both A6 and A20 to the minor groove for possible interaction with proteins and other ligands. In the present structure, these groups are involved in the abutting crystal packing interaction (see below). There is one hydrogen bond in the sheared G·G mismatch N2(G18)···O6(G7) (2.9 Å). G18 also has hydrogen-bonding interactions with the phosphate group of A6: N1(G18)···O2P(A6) (2.8 Å) and N2(G18)···O2P(A6) (2.9 Å) (Fig. 3*d*). G7 protrudes into the minor groove and its Watson–Crick face is exposed and hydrated with four water molecules, while G18 protrudes into major groove and its Hoogsteen face is exposed in the major groove and hydrated with four water molecules. There are two hydrogen bonds for the G·A mismatch, N1(A17)···N1(G8) (2.8 Å) and N6(A17)···O6(G8) (2.9 Å) (Fig. 3*e*).

In addition to the base–base hydrogen-bonding interaction, the symmetric loop has extensive hydrogen-bonding interactions with water molecules and 2′-hydroxyl groups. The base atoms of the five mismatches are hydrated with 21 water molecules, two hexahydrated calcium ions $[Ca(H_2O)_6]^{2+}$ and one tetrahydrated magnesium ion $[Mg(H_2O)_4]^{2+}$. The symmetric loop also has nine inter-duplex interactions, of which five are involved with the 2′-hydroxyl group. These interactions compensate for the loss of energy of base mismatching and stabilize the conformation of the symmetric loop.

Comparison with the crystal structure of domain IV of SRP RNA molecule containing both symmetric and asymmetric loops (Jovine *et al.*, 2000) showed that the symmetric loop has essentially the same conformation as observed in the present dodecamer, with an r.m.s.d. of less than 0.6 Å for the superposition of all the atoms in the symmetric loop. All the base pairings in the symmetric loop adopt the same conformation in these two structures. However, C5·A20, A6·C19 and G7·G18 have slightly longer bond lengths (by 0.2 Å on average) than those in the symmetric loop in domain IV, while G8·A17 has almost the same bond length. The water bridge in the G7·G18 was missing in the G·G mismatch in the domain IV structure.

3.3. Crystal packing

The patterns of crystal packing are closely related to the helical form of oligonucleotide duplexes. A-form RNA, B-form and Z-form DNA duplexes usually adopt the stacking

pattern, while A-form DNA duplexes always adopt the abutting pattern. In the present structure, two packing patterns are observed in the same duplex. The 5′-terminal base pair, G1·C24, stacks with its symmetry-related base pair in the head-to-tail fashion with a twist angle of 17°, while the 3′-terminal base pair, G12·C13, abuts into the minor groove of a symmetry-related duplex (Fig. 4). To our knowledge, this is the first observation of such crystal packing in RNA oligomers. The abutting pattern in the present structure results in two A*(G·C) base triplets in the minor groove (Fig. 5). It is clear from the figure that the 2′-hydroxyl groups play an important role in stabilizing the conformation. It should be noted that the C—H···N interactions C2(A20*)—H···N3(G12) and C2(A6*)—H···N3(G14) are involved in the formation of the base triplets. These two conformations are different from that observed in the abutting pattern in A-DNA duplex, in which only one hydrogen bond is involved: N2(G)···N7(A) (3.2 Å; Nunn & Neidle, 1996). However, similar base triplets have been observed in hammerhead ribozyme (Pley *et al.*, 1994), the P4–P6 domain of *Tetrahymena thermophilla* intron (Cate *et al.*, 1996), viral pseudoknot (Su *et al.*, 1999) and overhang RNA oligomers (Mitra *et al.*, 2000; Shi *et al.*, 2000).

As demonstrated by oligonucleotide crystals, A-form DNA crystal structures usually adopt the abutting interaction (Wahl & Sundaralingam, 1997), which is mainly attributed to the formation of hydrophobic platform for A-form helices (Sundaralingam & Biswas, 1997). Even though they have similar helical conformation as A-DNA duplexes, A-RNA duplexes usually adopt the stacking pattern in crystal packing. This may be because of the fact that the hydrophilic 2′-hydroxyl groups disrupt the hydrophobic platform generated by the A-form helix. Until now, all A-RNA duplexes have adopted the head-to-tail (5′,3′/5′,3′) stacking pattern, generating pseudo-continuous helical columns, or the head-to-head (5′,5′/3′,3′) stacking pattern, which usually results in super-helical conformation (Portmann *et al.*, 1995; Shi *et al.*, 2000; Mitra *et al.*, 2000). The present structure shows that the combination of head-to-tail stacking and minor-groove abutting can also exist in RNA oligomer duplexes. In the abutting crystal packing, 2′-hydroxyl groups are involved in four hydrogen bonds (Fig. 5). This result indicates that although the hydrophobic platform do not favor the abutting crystal packing for RNA duplexes, strong interactions involving the 2′-hydroxyl groups can compensate for the loss in stability and the abutting crystal packing may prevail.

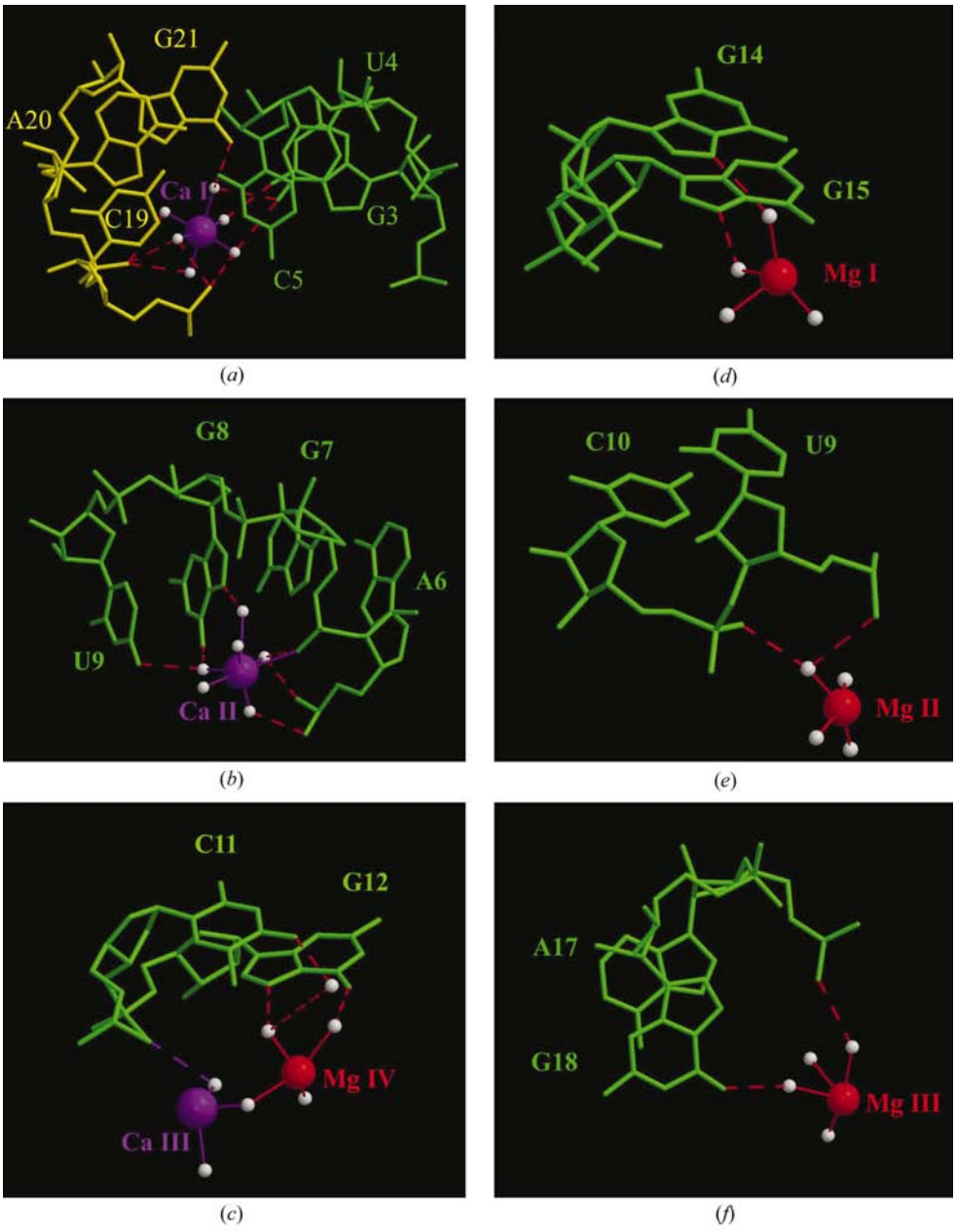

Figure 6
Interactions of the hydrated metal ions with the RNA duplex. Interactions with the bases from symmetry-related duplex are not shown.

3.4. Hydration and metal ions

111 water molecules, three calcium ions and four magnesium ions have been located in this structure. Phosphate groups are hydrated with 29 water molecules, the major groove with 56 water

Table 4
Hydrogen-bonding interactions between hydrated metal ions and RNA.

Hydrated metal ion	RNA atom	Distance (Å)
$[Ca(H_2O)_6]^{2+}$ (I)	O6 (G3)	3.0
$[Ca(H_2O)_6]^{2+}$ (I)	O4 (U4)	3.0
$[Ca(H_2O)_6]^{2+}$ (I)	O4 (U4)	3.2
$[Ca(H_2O)_6]^{2+}$ (I)	O2 (C5)	3.1
$[Ca(H_2O)_6]^{2+}$ (I)	O2P (C19)	2.8
$[Ca(H_2O)_6]^{2+}$ (I)	O2P (C19)	3.1
$[Ca(H_2O)_6]^{2+}$ (I)	O2P (A20)	3.0
$[Ca(H_2O)_6]^{2+}$ (I)	O2P (A20)	3.0
$[Ca(H_2O)_6]^{2+}$ (I)	O2P (A20)	3.1
$[Ca(H_2O)_6]^{2+}$ (I)	O6 (G21)	2.5
$[Ca(H_2O)_6]^{2+}$ (I)	N7 (G21)	3.2
$[Ca(H_2O)_6]^{2+}$ (II)	O1P (A6)	2.9
$[Ca(H_2O)_6]^{2+}$ (II)	O2P (A6)	3.0
$[Ca(H_2O)_6]^{2+}$ (II)	O2P (G7)	2.9
$[Ca(H_2O)_6]^{2+}$ (II)	O2P (G7)	2.9
$[Ca(H_2O)_6]^{2+}$ (II)	O2P (G8)	3.2
$[Ca(H_2O)_6]^{2+}$ (II)	O6 (G8)	2.6
$[Ca(H_2O)_6]^{2+}$ (II)	N7 (G8)	2.3
$[Ca(H_2O)_6]^{2+}$ (II)	O4 (U9)	2.9
$[Ca(H_2O)_6]^{2+}$ (III)	O2P (C11)	3.0
$[Ca(H_2O)_6]^{2+}$ (III)	O2P (C11)	3.1
$[Ca(H_2O)_6]^{2+}$ (III)	O2P (G12)	3.0
$[Ca(H_2O)_6]^{2+}$ (III)	O2P (C12)	3.0
$[Ca(H_2O)_6]^{2+}$ (III)	O2P (G12)	3.2
$[Mg(H_2O)_4]^{2+}$ (I)	N7 (G14)	3.0
$[Mg(H_2O)_4]^{2+}$ (I)	N7 (G15)	2.6
$[Mg(H_2O)_4]^{2+}$ (II)	O1P (U9)	3.1
$[Mg(H_2O)_4]^{2+}$ (II)	O2P (C10)	2.6
$[Mg(H_2O)_4]^{2+}$ (III)	O2P (A17)	3.1
$[Mg(H_2O)_4]^{2+}$ (III)	O6 (G18)	2.9
$[Mg(H_2O)_4]^{2+}$ (VI)	O1P (C11†)	2.9
$[Mg(H_2O)_4]^{2+}$ (VI)	O2P (C11†)	3.1
$[Mg(H_2O)_4]^{2+}$ (VI)	O6 (G12)	3.0
$[Mg(H_2O)_4]^{2+}$ (VI)	N7 (G12)	2.8

† Symmetry-related duplex.

molecules and the minor groove with 19 water molecules. The minor groove is less hydrated because of the abutting interaction of the crystal packing. Even though there are no obvious hydration patterns present, some features can be observed. Bridges of one and two water molecules are found between adjacent phosphate groups. The O6 and N7 atoms of guanine, the N6 and N7 atoms of adenine, the N4 of cytosine and the O4 of uridine are mostly hydrated by water molecules. Mismatches in the symmetric loop are heavily hydrated, implying the importance of water molecules in stabilizing the conformation of mismatches. For instance, C5 is hydrated with two water molecules at N4 and one water molecule at N3. In the minor groove, water molecules are mostly located in the vicinity of mismatches.

One important feature of the water molecules in the present structure is the hydration of metal ions. Only one direct hydrogen bond to a base atom is observed for seven metal ions. All the metal ions strongly interact with water molecules, with average bond lengths of 2.2 Å for Mg^{2+} and 2.2 Å for Ca^{2+}. The strong hydrogen bonds between metal ions and water molecules make them act as a complex in the interaction with RNA duplex. Calcium ions are hexahydrated, $[Ca(H_2O)_6]^{2+}$, and magnesium ions are tetrahydrated, $[Mg(H_2O)_4]^{2+}$. These hydrated metal ions form cluster centers that bridge adjacent phosphate groups, adjacent bases, base atoms and phosphate groups (Fig. 6). All of these interactions are listed in Table 4. $[Ca(H_2O)_6]^{2+}$ (I) and (II) bind to each edge of the symmetric loop and have strong interactions with the symmetric loop. $[Ca(H_2O)_6]^{2+}$ (I) has three interactions with the G·U wobble in the major groove and binds to C5 in A20·C5 mismatch (Fig. 6*a*). On the other side of the hydrated metal ion, it forms five interactions with the phosphate groups of C19 and A20. These interactions greatly stabilize the conformation of A20·C5, which only has one base–base hydrogen bond. $[Ca(H_2O)_6]^{2+}$ (I) also binds to three adjacent bases G3-U4-C5. $[Ca(H_2O)_6]^{2+}$ (II) has three interactions with the G8·A17 mismatch (Fig. 6*b*). On the other side, it forms two strong hydrogen bonds (2.9 and 3.0 Å) with the phosphate oxygen of A6, which binds to N1(G18) (2.8 Å) and N2(G18) (2.9 Å). $[Ca(H_2O)_6]^{2+}$ (II) makes a network of interactions between the adjacent mismatches G7·G18 and G8·A17. It binds to two adjacent bases G8-U9. At the 3′-terminal position of the duplex, $[Ca(H_2O)_6]^{2+}$ (III) and $[Mg(H_2O)_4]^{2+}$ (IV) form a complex $[Ca(H_2O)_6]^{2+}\cdot[Mg(H_2O)_4]^{2+}$ by sharing a common water molecule (Fig. 6*c*). This complex, extending to about 9 Å in size, interacts with G12·C13, the phosphate O atoms of C11 and G12 and the symmetry-related phosphate O atoms of C11* and G12*. Obviously, this complex stabilizes the crystal packing of the crystal structure. $[Mg(H_2O)_4]^{2+}$ (I) interacts with the G14-G15 step at N7, while $[Mg(H_2O)_4]^{2+}$ (II) sits at a twofold symmetry axis and interacts with the adjacent phosphate O atoms of U9 and C10 and their symmetry-related counterparts (Figs. 6*d* and 6*e*). $[Mg(H_2O)_4]^{2+}$ (III) interacts with O6(G18) and the phosphate of A17 (Fig. 6*f*). Of 34 interactions between hydrated metal ions and the RNA duplex, 19 are involved in the interaction with the symmetric loop (Table 4). Clearly, these hydrated ions are important in stabilizing the conformation of the symmetric loop.

The three hexahydrated calcium ions in the present structure display different binding sites in RNA duplex, the GG step, GU step and GUC step in the major groove, and two or three adjacent phosphate groups. One hexahydrated calcium ion can be involved in as many as three different binding sites, as shown by $[Ca(H_2O)_6]^{2+}$ (I). The four $[Mg(H_2O)_4]^{2+}$ ions exhibit binding sites in the GG step, adjacent phosphate groups and guanine and phosphate groups. Phosphate groups are the preferred binding sites for these two kinds of hydrated metal ions because of the attraction between the opposite electrical charges of the hydrated metal ions and phosphate groups. O6 and N7 of guanine are the atoms that the hydrated metal ion usually binds.

It is helpful to compare metal-ion bindings of the SRP RNA molecules in different situations. In the free form of the 4.5S RNA, no metal ions bind in the major groove of the symmetric loop (Jovine *et al.*, 2000). Two hexahydrated magnesium ions bind in the major groove with one at each end of the symmetric loop in the protein-bound 4.5S RNA (Batey *et al.*, 2000, 2001). In the present structure, two hexahydrated calcium ions are located in nearly the same position, interacting with the symmetric loop by mediation of water molecules. In other words, magnesium ions and calcium ions can bind equally well to the symmetric loop. This result is

consistent with the structural and energetic analysis, which indicates that the major-groove binding site only has low selectivity for a specific type of monovalent or multivalent ions (Batey & Doudna, 2002).

3.5. Biological implication

4.5S RNA and Ffh protein are the two components of bacterial SRP. Ffh protein primarily binds to the symmetric loop in domain IV of 4.5S RNA. Crystallographic studies have showed that the conformation of the domain IV of 4.5S RNA does not change when it binds to the M domain of Ffh protein (Batey *et al.*, 2000; Jovine *et al.*, 2000). Based on this fact, it was suggested that the interaction between Ffh protein and 4.5S RNA could be divided into two steps: rigid docking of the symmetric loop region, followed by induced ordering of the asymmetric loop (Jovine *et al.*, 2000). The symmetric loop in the present structure was almost identical to the corresponding part of the two structures above, as indicated by the r.m.s.d. of less than 0.6 Å. This similarity indicates that the symmetric loop adopts its structure independently of the asymmetric loop and the GGAA tetraloop. We suggest from this result that the induced ordering of the asymmetric loop will not affect the conformation of the symmetric loop.

Metal ions are very important in stabilizing the three-dimensional structures of RNA molecules. Magnesium ions have been shown to stabilize 4.5S RNA (Bourgaize *et al.*, 1984; Lentzen *et al.*, 1996; Schmitz *et al.*, 1996) and have been observed in the crystal structures of SRP and 4.5S RNA (Batey *et al.*, 2000; Jovine *et al.*, 2000), which can form a complex with water molecules as pentahydrated (Gao *et al.*, 1999) and hexahydated complexes (Batey *et al.*, 2000; Egli *et al.*, 1998; Robinson *et al.*, 2000). The present structure shows that magnesium can also form tetrahydrated complexes. This result indicates that the hydrated metal ions are flexible in adjusting their geometry and coordination when they interact with nucleic acids. These kinds of metal ions are an important motif in interactions with nucleic acids because they are flexible and extend over a large interaction range.

We gratefully acknowledge the National Institutes of Health for grant GM-17378 and an Ohio Eminent Scholar Endowment for supporting this work. We also thank BioCARs for the use of beamline BMD14 for collecting data at the Advanced Photon Source, Argonne National Laboratory supported by the US Department of Energy, Basic Energy Sciences, Office of Science under Contract No. W-31-109-Eng-38.

References

Althoff, S. M., Selinger, D. & Wise, J. A. (1994). *Nucleic Acids Res.* **22**, 1933–1947.
Batey, R. T. & Doudna, J. A. (2002). *Biochemistry*, **41**, 11703–11710.
Batey, R. T., Rambo, R. P., Lucast, L., Rha, B. & Doudna, J. A. (2000). *Science*, **287**, 1232–1239.
Batey, R. T., Sagar, M. B. & Doudna, J. A. (2001). *J. Mol. Biol.* **307**, 229–246.
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.
Bourgaize, D. B., Farrell, C., Langley, K. H. & Fournier, M. J. (1984). *Nucleic Acids Res.* **12**, 2019–2034.
Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cate, J. H., Gooding, A. R., Podell, E. R., Zhou, K., Golden, B., Kundrot, C., Cech, T. & Doudna, J. A. (1996). *Science*, **273**, 1678–1685.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Egli, M., Tereshko, V., Teplova, M., Minasov, G., Joachimiak, A., Sanishvili, R., Weeks, C. M., Miller, R., Maier, M. A., An, H., Cook, P. D. & Manharan, M. (1998). *Biopolymers*, **48**, 234–252.
Gao, Y.-G., Robinson, H., Sanishvilli, R., Joachimiak, A. & Wang, A. H.-J. (1999). *Biochemistry*, **38**, 16452–16460.
Jiang, S. B., Hung, L.-W., Chi, Y.-I., Holbrook, E., Carter, R. & Holbrook, S. R. (1998). *Biochemistry*, **37**, 11726–11731.
Jovine, L., Hainzl, T., Oubridge, C., Scott, W. G., Li, J., Sixma, T. K., Wonacott, A., Skarzynski, T. & Nagai, K. (2000). *Structure*, **8**, 527–540.
Larsen, N. & Zwieb, C. (1991). *Nucleic Acids Res.* **19**, 209–215.
Lavery, R. & Sklenar, H. (1989). *J. Biomol. Struct. Dyn.* **4**, 655–667.
Lentzen, G., Moine, H., Ehresmann, C., Ehresmann, B. & Wintermeyer, W. (1996). *RNA*, **2**, 244–253.
Mitra, S. N., Biswas, R., Shi, K. & Sundaralingam, M. (2000). *J. Biomol. Struct. Dyn.* **11**, 189–194.
Nunn, C. M. & Neidle, S. (1996). *J. Mol. Biol.* **256**, 340–351.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Pan, B., Mitra, S. N. & Sundaralingam, M. (1998). *J. Mol. Biol.* **283**, 977–984.
Pan, B., Mitra, S. N. & Sundaralingam, M. (1999). *Biochemistry*, **38**, 2826–2831.
Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). *Nature (London)*, **372**, 68–74.
Poritz, M. A., Strub, K. & Walter, P. (1988). *Cell*, **55**, 4–6.
Portmann, S., Usman, N. & Egli, M. (1995). *Biochemistry*, **34**, 7569–7575.
Robinson, H., Gao, Y.-G., Sanishvili, R., Joachimiak, A. & Wang, A. H.-J. (2000). *Nucleic Acids Res.* **28**, 1760–1766.
Schmitz, U., Behrens, S, Freymann, D. M., Keenan, R. J., Lukavsky, P., Walter, P. & James, T. L. (1999). *RNA*, **5**, 1419–1429.
Schmitz, U., Freymann, D. M., James, T. L., Keenan, R. J., Vinayak, R. & Walter, P. (1996). *RNA*, **2**, 1213–1227.
Schmitz, U., James, L., Lukavsky, P. & Walter, P. (1999). *Nature Struct. Biol.* **6**, 634–638.
Selinger, D., Brennwald, P., Liao, X. & Wise, J. A. (1993). *Mol. Cell. Biol.* **13**, 1353–1362.
Shi, K., Biswas, R., Mitra, S. N. & Sundaralingam, M. (2000). *J. Mol. Biol.* **299**, 113–122.
Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). *Nature Struct. Biol.* **6**, 285–292.
Sundaralingam, M. & Biswas, R. (1997). *J. Biomol. Struct. Dynam.* **15**, 173–176.
Wahl, M. C. & Sundaralingam, M. (1997). *Biopolymers*, **44**, 45–63.
Wood, H., Luirink, J. & Tollervey, D. (1992). *Nucleic Acids Res.* **20**, 5919–5925.
Zwieb, C., Müller, F. & Larsen, N. (1996). *Fold. Des.* **1**, 315–324.