

# The 1.3 Å resolution structure of the RNA tridecamer r(GCGUUUGAAACGC): Metal ion binding correlates with base unstacking and groove contraction

YOURI TIMSIT<sup>1</sup> and SOPHIE BOMBARD<sup>1,2</sup>

<sup>1</sup>Laboratoire de Cristallographie et de Biochimie Théorique, Institut de Biologie Physico-Chimique, UPR 9080, CNRS, 75005 Paris, France

<sup>2</sup>Laboratoire de Chimie et de Biochimie Pharmacologiques et Toxicologiques, Université René Descartes, UMR8601 CNRS, Paris, France

## ABSTRACT

Metal ions play a key role in RNA folding and activity. Elucidating the rules that govern the binding of metal ions is therefore an essential step for better understanding the RNA functions. High-resolution data are a prerequisite for a detailed structural analysis of ion binding on RNA and, in particular, the observation of monovalent cations. Here, the high-resolution crystal structures of the tridecamer duplex r(GCGUUUGAAACGC) crystallized under different conditions provides new structural insights on ion binding on GAAA/UUU sequences that exhibit both unusual structural and functional properties in RNA. The present study extends the repertory of RNA ion binding sites in showing that the two first bases of UUU triplets constitute a specific site for sodium ions. A striking asymmetric pattern of metal ion binding in the two equivalent halves of the palindromic sequence demonstrates that sequence and its environment act together to bind metal ions. A highly ionophilic half that binds six metal ions allows, for the first time, the observation of a disodium cluster in RNA. The comparison of the equivalent halves of the duplex provides experimental evidences that ion binding correlates with structural alterations and groove contraction.

**Keywords:** X-ray crystallography; asymmetry; electrostatic; monovalent cations

## INTRODUCTION

Metal ions play an essential role in RNA folding and activity (Pyle 2002). Two modes of metal ion binding contribute to stabilize RNA secondary and tertiary structures (Draper et al. 2005). In diffuse binding, the ions interact with RNA through long-range electrostatic interactions. Site bound metal ions are, on the other hand, bound directly via RNA. The two modes of ion binding both contribute to stabilize the structure of RNA. In particular, divalent cations are essential components of RNA tertiary structure and are required for both stability and catalytic activity. RNA folds can form a specific pocket for their tight binding. Crystallographic studies have shown that magnesium ions have a marked preference for binding Hoogsteen edges of guanine (Jack et al. 1977; Scott et al. 1995; Ennifar et al. 2003). In addition, some specific

sequences such as the tandem wobble G.U base pairs in RNA also constitute magnesium binding sites (Cate and Doudna 1996). On the other hand, monovalent metal ions stabilize specific RNA motifs (Draper and Misra 1998) and can support folding of catalytically active tertiary structure at high concentrations (Murray et al. 1998). Certain RNA motifs and pockets were found to specifically chelate potassium ions (Basu et al. 1998; Batey et al. 2000; Conn et al. 2002). Crystallographic studies of DNA oligonucleotides have also demonstrated that monovalent cations play an essential role in DNA structure and compaction (McFail-Isom et al. 1999). However, locating monovalent cation in nucleic acid structures requires accurate crystallographic data and great care (Tereshko et al. 2001). Due to the lack of high-resolution data, the role of RNA sequence on the binding of monovalent ions and in particular sodium ions that are more difficult to identify crystallographically is sparsely documented. In addition, how sequence and other factors act together to influence the specific binding of metal ions is still not well understood. Indeed, a comparative study of ion binding in two similar RNA sequences has noted that “subtle changes that are not immediately rationalized” affect ion binding in RNA (Ennifar et al. 2003).

**Reprint requests to:** Youri Timsit, Laboratoire de Cristallographie et de Biochimie Théorique, Institut de Biologie Physico-Chimique, UPR 9080, CNRS, 13, rue Pierre et Marie Curie, 75005 Paris, France; e-mail: Timsit@ibpc.fr; fax: 0033158415026.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.730207>.

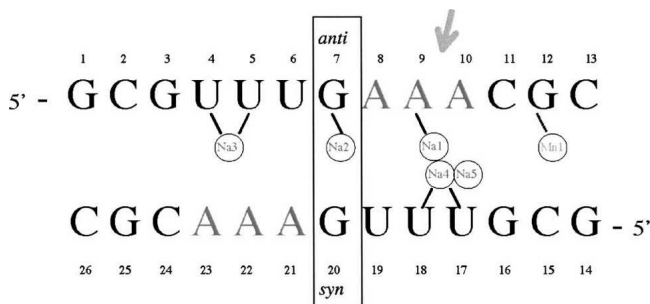
On the basis of the high-resolution crystallographic structures of the native and brominated palindromic RNA tridecamer r(GCGUUUGAAACGC) bearing the GAAA tetranucleotide, we attempt to answer the following questions: (1) Are there specific sites along the sequence for the binding of monovalent cations in RNA? (2) How do sequence and electrostatic environment act together to bind ions specifically? GAAA sequences in RNA are of particular interest since they exhibit unusual functional and structural features. First, they constitute a minimal  $Mn^{2+}$  ribozyme when paired to the UUU triplet (Dange et al. 1990; Kazakov and Altman 1992; Kuo and Herrin 2000). In catalyzing the specific cleavage between the GpA step, they represent the smallest ribozyme observed to date and may constitute an interesting model for active prebiotic molecules. However, structural information about ion binding in this sequence and the catalytic mechanism still remains to be elucidated. GAAA sequences are also part of the polypurine tracts (PPT) of retroviruses that are resistant to degradation by retroviral RNase H (Telesnitsky and Goff 1997). Crystallographic studies have shown that in their RNA/DNA hybrid form, they adopt an unusual conformation that prevents them from being hydrolyzed by enzymes (Safiranos et al. 2001; Kopka et al. 2003). Interestingly, GAAA tetranucleotide is very rarely observed within a regular double in RNA structures available in the structural data banks (Berman et al. 2000). However, they are frequently found in GNRA tetraloops, and they also constitute a conserved part of the hammerhead ribozyme where they adopt an irregular structure containing noncanonical base pairs (Scott et al. 1995; Martick and Scott 2006).

The present study provides detailed structural information on the hydration of GAAA/UUU sequences and the localization of divalent and monovalent cations in four isomorphous structures obtained in different crystallization conditions. The high-resolution crystallographic data have allowed the unambiguous identification of sodium ions specifically bound through their inner sphere to groups lining the major groove of RNA. A marked asymmetry in ion binding is observed in the two equivalent halves of the duplex with respect to the palindromic axis. The highly ionophilic half consisting of six base pairs that binds five metal ions is reminiscent of the loop E of the 5S ribosomal RNA (Correl et al. 1997). The equivalent half only chelates one sodium ion. Comparison of their hydration in the context of the crystal packing environment shows how the sequence and the electrostatic environment influence the specific binding of the metal ions. This study also shows that GAAA tetranucleotides inserted in a regular RNA duplex present an unusual structural plasticity that correlates with metal ion binding. These results support previous structural studies that observed unusual conformations in GAAA sequences of RNA/DNA hybrids (Safiranos et al. 2001; Kopka et al. 2003) and show that the structural alterations correlate with cation binding.

## RESULTS AND DISCUSSION

### Overall structure

In order to analyze the influence of metal ions on RNA structure, four isomorphous structures of the tridecamer r(GCGUUUGAAACGC) (Fig. 1) were crystallized with different relative amounts of divalent cations. The Br1 and Br2 brominated duplexes grew in crystallization drops containing both magnesium and manganese ions (Table 1). In contrast, the brominated BrMn and the native NatMn structures were crystallized with manganese ion alone as a divalent cation (0.5 mM and 3 mM, respectively). In addition, a concentration of 100 mM of sodium ions (NaCl) was used for the crystallization of the four duplexes. The four duplex structures do not display any catalytic activity. The RNA tridecamers Br1, Br2, BrMn, and NatMn adopt a similar geometry. The two palindromic halves that contain the GAAA tetranucleotide surround a central well-ordered  $G_{anti}:G_{syn}$  mispair. With an RMSD value of 0.089 Å, the two brominated structures Br1 and Br2 obtained in similar crystallization conditions are identical. The brominated BrMn and native NatMn structures crystallized at lower concentrations of divalent ions differ slightly from Br1 (Table 1, S1). However, since the four structures exhibit almost identical structural features, Br1 has been used as a representative model for the following structural analysis. Most of the helical values are typical for an A-form helix (Table 2). However, a marked unstacking of the adenines at step A9/A10 is observed (Fig. 2A; Table 2). The two complementary uracils U17 and U18 are also unstacked (Fig. 2B). On the contrary, the equivalent base pairs at step A22/A23 are well stacked. Another particularity of this structure is illustrated by the high values of the B-factor of the phosphate groups of the two central GpA steps (Fig. 3). It is rare to observe higher values of B-factors in the backbone of a residue located in the middle of a double helix duplex. This unusual feature may therefore reflect an intrinsic flexibility of the GpA step. The GpA step corresponds to the cleavable step in the  $Mn^{2+}$  ribozymes (Dange et al. 1990; Kazakov and Altman 1992; Kuo and Herrin 2000). A higher



**FIGURE 1.** Sequence of the tridecamer. In the brominated Br1, Br2, and BrMn structures, the U6 and U19 uracils are replaced with 5-bromo-uracils. The ions are schematically represented within circles. The arrow indicates the destacking between A9 and A10.

TABLE 1. Statistic of data collection, phasing, and refinement

	Br1	Br2	BrMn	NatMn
Sequence	r(GCGUU <sup>Br</sup> UGAAACGC)	r(GCGUU <sup>Br</sup> UGAAACGC)	r(GCGUU <sup>Br</sup> UGAAACGC)	r(GCGUUUGAAACGC)
PDB code	2R1S	2R20	2R21	2R22
Divalent cations	3 mM MnCl <sub>2</sub> 2 mM MgCl <sub>2</sub>	5 mM MnCl <sub>2</sub> 5 mM MgCl <sub>2</sub>	0.5 mM MnCl <sub>2</sub>	3 mM MnCl <sub>2</sub>
Data collection and phasing				
Cell parameters				
<i>a, b, c</i> (Å)	25.38, 25.67, 29.92	25.28, 25.80, 29.99	25.25, 25.65, 29.72	25.28, 25.50, 29.60
$\alpha, \beta, \gamma$ (°)	98.64, 105.43, 100.67	98.62, 105.58, 100.92	98.13, 106.15, 100.83	97.97, 106.69, 101.62
Resol (Å)	1.4	1.3	1.6	1.5
Peak				
$\lambda = 0.919487$				
$N_{\text{meas}}/N_{\text{ref}}$	35,101/11,142	47,954/13,060	16,234/8117	17,032/8516
<i>I</i> / $\sigma$ <i>I</i>	19.4	9.1	4.2	7.8
<i>I</i> <sub>mean</sub>	5227	22,981	5119	45,551
<i>R</i> <sub>sym</sub> (%)	2.3	4.9	7.7	6.4
Cullis <i>R</i> <sub>fact</sub>	0.75			
Phasing power	1.13			
FOM init/FOM final	0.64/0.87			
Edge $\lambda = 0.919704$				
$N_{\text{meas}}/N_{\text{ref}}$	16,568/8284			
<i>I</i> / $\sigma$ <i>I</i>	13.6			
<i>I</i> <sub>mean</sub>	2958			
<i>R</i> <sub>sym</sub> (%)	4.0			
Cullis <i>R</i> <sub>fact</sub>	0.77			
Phasing Power	1.16			
Remote1 $\lambda = 0.88558$				
$N_{\text{meas}}/N_{\text{ref}}$	16,570/8285			
<i>I</i> / $\sigma$ <i>I</i>	6.0			
<i>I</i> <sub>mean</sub>	4794			
<i>R</i> <sub>sym</sub> (%)	7.7			
Cullis <i>R</i> <sub>fact</sub>	0.915			
Phasing power	0.60			
Remote2 $\lambda = 0.979776$				
$N_{\text{meas}}/N_{\text{ref}}$	18,068/7055			
<i>I</i> / $\sigma$ <i>I</i>	27.0			
<i>I</i> <sub>mean</sub>	19,999			
<i>R</i> <sub>sym</sub> (%)	2.2			
Refinement				
<i>N</i> <sub>ref</sub>	11,018	12,692	7596	8452
<i>R</i> <sub>fact</sub> / <i>R</i> <sub>free</sub>	15.8/18.5	18.1/19.8	16.9/19.1	17.8/20.1
<i>N</i> atoms	552	552	552	550
<i>N</i> solvent	182	171	165	150
Metal ions	6	6	5	5
RMSD versus BR1		0.089	0.183	0.347

B-factor in the center of the duplex may be also related to the presence of the G7:G20 Hoogsteen base pair (Fig. 2C). This study reveals that the GAAA sequence presents an important structural plasticity when inserted within a regular RNA duplex. Previous studies have shown that a similar sequence within the polypurine tract of retrovirus exhibit important structural alterations in RNA/DNA hybrid duplexes (Safiranos et al. 2001; Kopka et al. 2003).

#### The G7<sub>anti</sub>:G20<sub>syn</sub> mispair

In the center of the duplex, the G7 and G20 form a Hoogsteen-type asymmetric base pair stabilized by

G7(N2)–G20(N7) and G7(N1)–G20(O6) hydrogen bonds (Fig. 2C). A water molecule bridges the two O6 carbonyl groups and stabilizes the interaction. The G20 adopts a *syn* conformation with a glycosidic torsion angle  $\chi$  of 5.7°. A hydrogen bond between the G20 phosphate oxygen and the G20(N2) amino nitrogen also stabilizes the nucleotide conformation. The G<sub>syn</sub>:G<sub>anti</sub> mispair is rare in RNA. It has been observed in the HIV Rev response element (RRE) (Hung et al. 2000; Ippolito and Steitz 2000) in which it is required for specific recognition by the protein Rev. The C1'–C1' distance is 11.4 Å, a value close to that observed in the two RRE structures. However, in the present structure, the sugar pucker remains in C3'-endo while it

**TABLE 2.** Selected helical parameters

	Rise (Å)	Tilt (°)	Roll (°)	Twist (°)	X-disp (Å)	Y-disp (Å)	Incl (°)	Tip (°)	Overlap (strand 1) (Å <sup>2</sup> )	Overlap (strand 2) (Å <sup>2</sup> )
1 GC/GC	3.27	-0.79	-0.30	34.85	-2.51	0.60	-0.51	1.32	4.86	7.06
2 CG/CG	3.30	0.89	13.76	30.45	-4.99	0.63	24.67	-1.59	0.00	0.00
3 GU/AC	3.26	0.96	4.75	30.07	-4.21	-1.32	9.09	-1.84	7.32	3.15
4 UU/AA	3.15	-0.11	5.21	28.54	-4.47	-1.73	10.46	0.21	0.66	0.63
5 UU/AA	3.39	0.06	7.51	31.88	-4.85	-1.61	13.44	-0.10	1.19	2.54
6 UG/GA	3.02	6.53	6.86	—	-2.83	-0.57	5.31	-5.06	0.92	0.00
7 GA/UG	3.02	0.71	2.80	—	-2.24	-3.00	-6.85	1.73	2.31	3.86
8 AA/UU	3.34	-0.22	5.19	30.52	-4.41	-1.76	9.77	0.42	3.25	1.73
9 AA/UU	3.44	-2.87	11.25	32.43	-5.01	0.99	19.38	4.95	0.13	0.17
10 AC/GU	3.26	-0.68	5.13	30.85	-4.55	0.91	9.55	1.26	3.48	6.55
11 CG/CG	3.19	1.35	8.09	28.56	-5.23	-1.05	16.00	-2.67	0.06	0.00
12 GC/GC	3.24	1.21	-1.70	33.67	-3.22	-0.31	-2.94	-2.09	5.89	4.40
Average	3.24	0.59	5.71	30.71	-4.04	-0.42	8.95	-0.29		
SD	0.13	2.20	4.34	22.18	1.06	1.39	9.16	2.58		

The helical parameters are calculated with the program X3DNA (Lu and Olson 2003).

adopts the C2'-endo, C4'-exo, or C1'-exo forms in the RRE structures.

### Ion binding to RNA duplex

Monovalent and divalent cations have been unambiguously identified in the high-quality electron density maps (Fig. 2) of the four structures according to geometric and crystallographic criteria (Table 3).

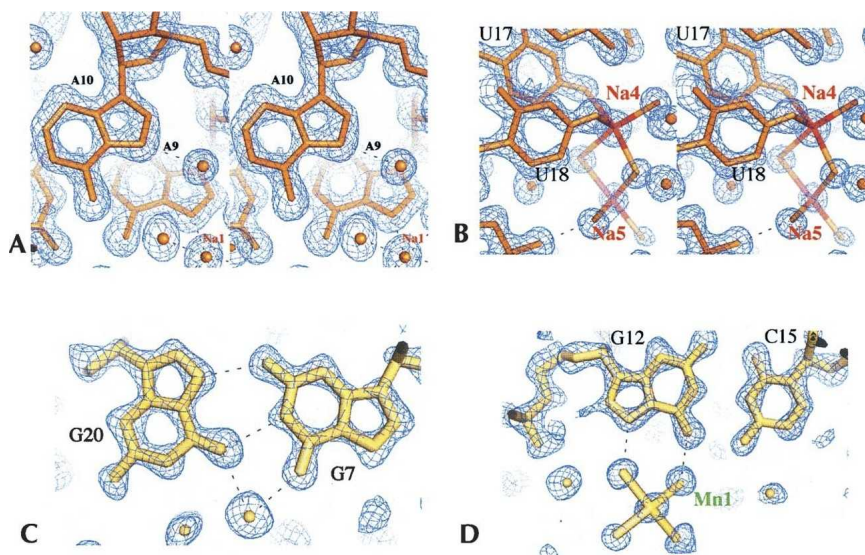
#### Divalent cations

A divalent cation has been found at G12 in all four structures. The average distance of 2.2 Å between the metal ion and the water molecules of its coordination sphere indicates that it could be either a magnesium or a manganese ion. As described in Table 3 and Figure 2D, this metal ion interacts with the O6 and N7 atoms of the guanine G12 through two water molecules of its coordination sphere. Such a binding mode is common in A-form DNA or RNA duplexes (Mayer-Jung et al. 1998). In the BrMn and the NatMn structures whose crystal grew only in the presence of Mn<sup>2+</sup>, this site must correspond to a hexahydrated manganese ion.

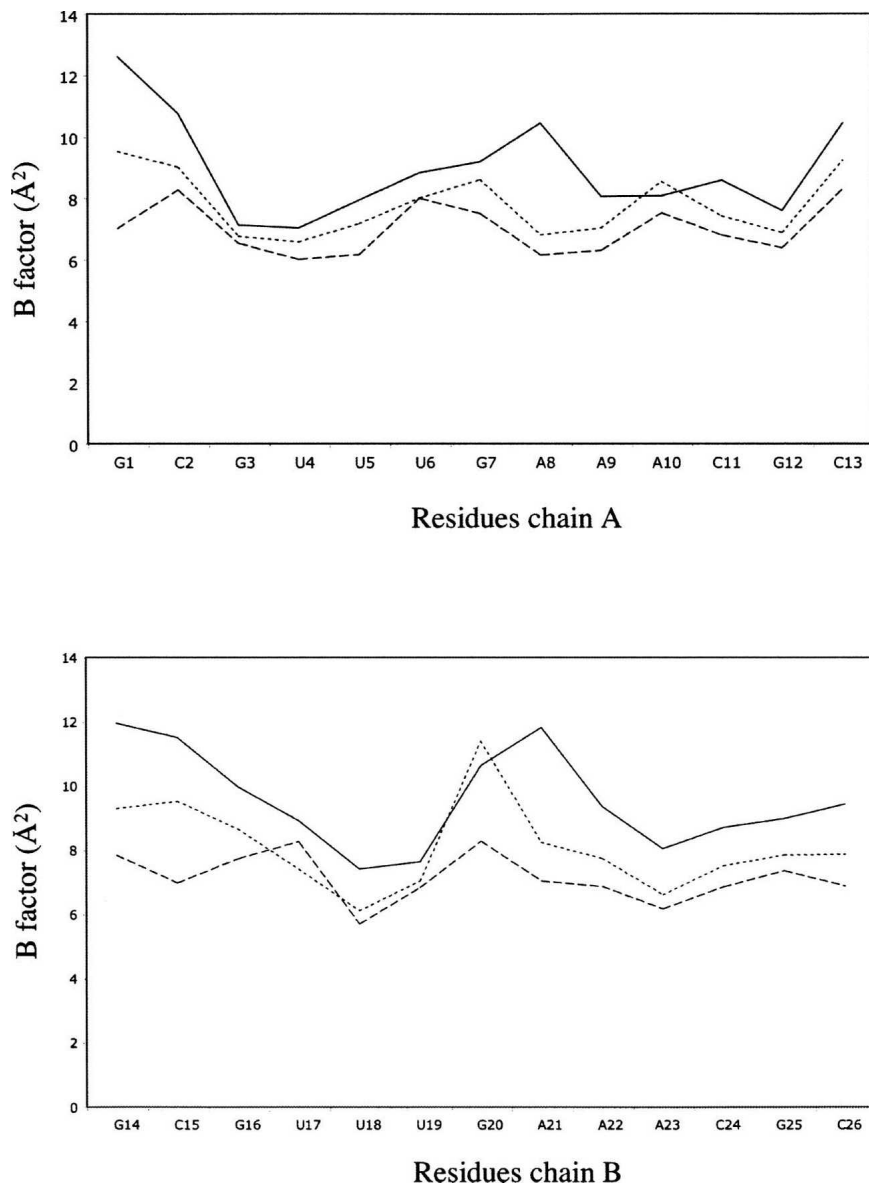
#### Monovalent cations

A total of five monovalent cations assigned to sodium ions have been located in the electron density maps.

These ions have been independently found in the four refined structures except Na5, which is not observed in either BrMn or NatMn. The metal ions common to the four structures (Na1–Na4) chelate the N7 atoms of purines or O4 carbonyl groups of uracils in the major groove. The fifth one (Na5), which is involved in a bimetallic cluster, binds anionic oxygen atoms through its inner coordination sphere (Fig. 2B). Two modes of direct coordination of sodium ions are observed. In the first one, sodium ions



**FIGURE 2.** (A) Stereo view of the unstacking at the A9/A10 step. Electron density map 2fo-fc at 1.4 Å of resolution (Br1) and contoured at 1.4σ. (B) Stereo view of the disodium bound at the unstacked U17/U18 step. Electron density map 2fo-fc at 1.4 Å of resolution (Br1) and contoured at 1.4σ. (C) The Hoogsteen pair G7–G20. Electron density map 2fo-fc at 1.3 Å (Br2) of resolution and contoured at 1.5σ. (D) Binding of the divalent cation at G12. Electron density map 2fo-fc at 1.3 Å (Br2) of resolution and contoured at 1.5σ.



**FIGURE 3.** Plot of the B-factors of strand 1 (A) and strand 2 (B) of the Br1 structure. Average B-factors of phosphate groups (uninterrupted line), sugar (dotted line), and bases (dashed line) are represented.

directly coordinate the N7 of purines. Na1, which is located at the unstacked step A9/A10, binds directly to N7 of A9 and indirectly to N6 of A10. Its hydration sphere also interacts with the phosphate group of A9. In a similar manner, Na2 binds directly to N7 of G7. The second binding mode involves the coordination of the O4 groups of the consecutive uracils. Figure 4, A and B, shows that sodium ions Na3 and Na4 chelate the carbonyl groups of the two uracils U4/U5 and their equivalent step U17/U18, with respect to the palindromic axis. It has been shown previously that monovalent metal ions coordinate RNA atoms and stabilize RNA motifs such as the AA platform motif of the RNA tetraloop receptor (Basu et al. 1998),

noncanonical base pairs of the RNA in the signal recognition particle (Batey et al. 2000), and an RNA pocket in a ribosomal RNA fragment (Conn et al. 2002). Monovalent ions also stabilize RNA tetrad in chelating carbonyl groups of guanines and uracils (Pan et al. 2003). Analysis of ions around 23S RNA has shown that monovalent ions modeled as  $\text{Na}^+$  bind to the electronegative edge of guanines in the context of G-U wobble base pairs (Klein et al. 2004). Interestingly, the monovalent ions described in these studies as well as the sodium ions reported in the MeRNA database (Stefan et al. 2006) display larger distances between the  $\text{Na}^+$  ions and their ligands (between 2.8 and 3.2 Å). These distances are significantly larger than the ones observed in our structure and in other high-resolution structures of nucleotides (Kennard et al. 1971). It is therefore possible that the longer distances that are observed at a lower resolution reflect the presence of hybrid binding sites with partial occupancies that combine water and sodium ions (McFail-Isom et al. 1999). By showing that the two first bases of the UUU triplet constitute a specific site for a sodium ion, our study adds a new sequence to the repertory of RNA ion binding sites. One should note that in both equivalent sites, the metal ions bind the first two bases of the triplet. It is likely that the electronegative potential generated contiguous carbonyl groups in the GUUU tetranucleotide favors ion binding. Similarly, it has been observed that tandem G-U wobble base pairs in group I intron structure constitute specific divalent binding sites due to their high electronegative potential (Cate and Doudna 1996; Keel et al. 2007).

#### Disodium cluster

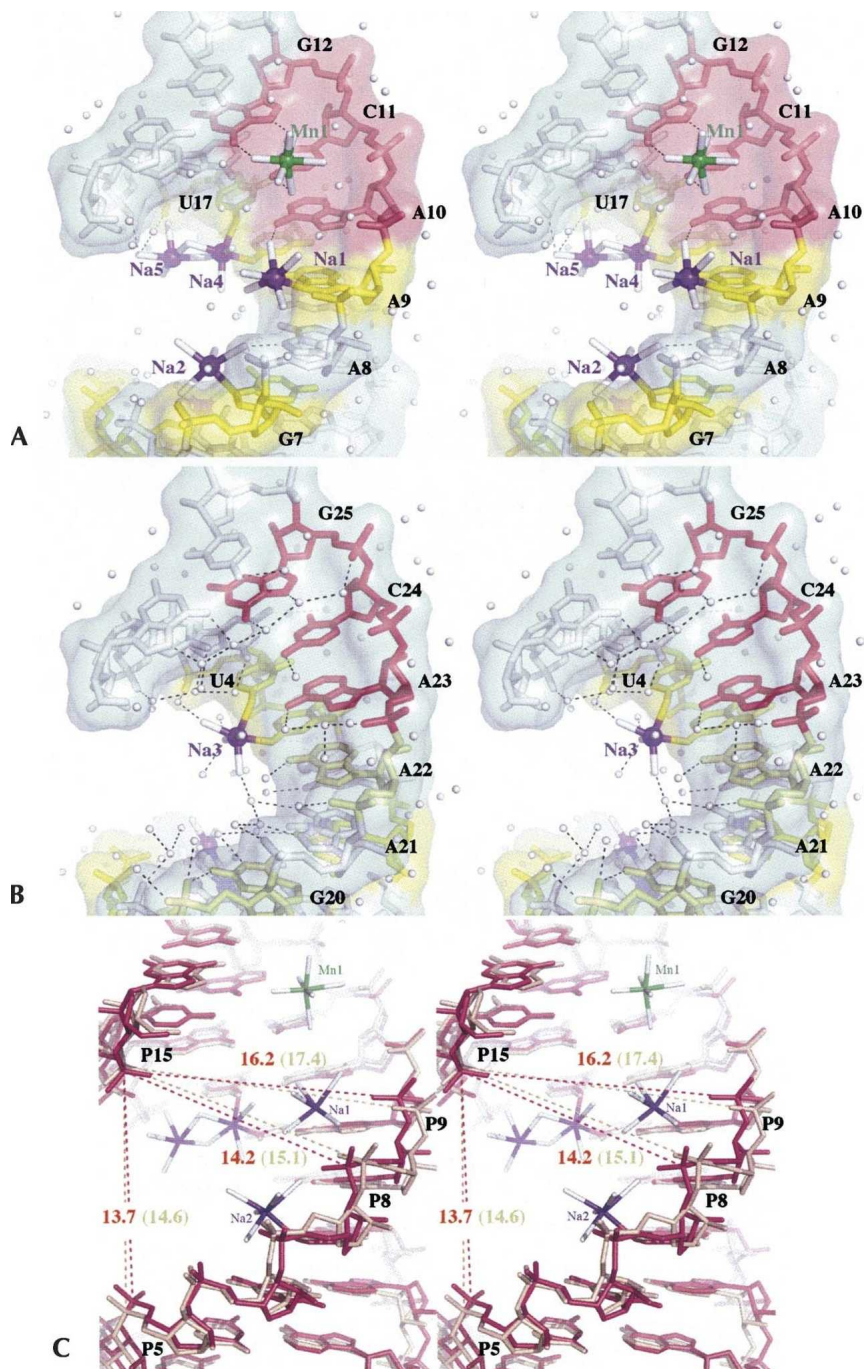
Our structure also provides the first example of a disodium bound to an RNA molecule. The two sodium ions Na4 and Na5 share two water molecules of their hydration sphere (Fig. 2B). Na4 coordinates the O4 carbonyl groups of U17 and U18 and Na5 interacts through its outer sphere with the anionic oxygen atoms of G16 and G17. Since the occupancies of both sodium cations are close to unity, it is

**TABLE 3.** Geometry of the cation coordination in the Br1 structure

	Element	Inner sphere ligand	Distance (Å)	Occupancy	B-factor	Outer sphere RNA ligand
Divalent 1 Mn1	Mg <sup>2+</sup> /Mn <sup>2+</sup>	W1	2.31	1.0	19.5	
		W2	2.45	0.85	23.7	
		W3	2.40	0.93	19.41	N4 (C11)
		W4	2.03	0.81	32.2	
		W5	1.99	0.99	16.28	O6 (G12)
		W6	2.31	0.77	31.63	N7 (G12)
Monovalent 1 Na1	Na <sup>+</sup>	N7 (A9)	2.57	1.0	17.35	
		W2	2.27	1.0	6.47	
		W3	2.59	1.0	24.82	
		W4	2.48	0.98	17.20	O2P (A9)
		W5	2.59	0.98	14.40	N6 (A10)
		W6	2.42	1.0	34.66	N6 (A9)
Monovalent 2 Na2	Na <sup>+</sup>	N7(G7)	2.49	0.97	20.36	
		W2	2.5	1.0	8.12	
		W3	2.6	1.0	12.51	O2P (G7)
		W4	2.5	1.0	17.49	O2P (U6)
		W5	2.74	1.0	20.22	O2P (A8)
		W6	2.55	0.92	23.59	O2P (G7)
Monovalent 3 Na3	Na <sup>+</sup>	O4 (U4)	2.52	1.0	12.48	
		O4 (U5)	2.44	1.0	7.17	
		W3	2.51	1.0	6.87	
		W4	2.31	1.0	13.64	
		W5	2.44	1.0	17.73	
		W6	2.39	1.0	16.05	N6 (A23)
Monovalent 4 Na4 (cluster)	Na <sup>+</sup>	O4 (U17)	2.47	0.97	11.07	
		O4 (U18)	2.39	1.0	10.3	
		W3	2.44	1.0	5.77	
		W4	2.40	1.0	15.12	
		W5	2.41	1.0	13.64	N6 (A9)
		W6	2.52	1.0	14.78	
Monovalent 5 Na5 (cluster)	Na <sup>+</sup>	W1	2.34	1.0	25.79	
		W2	2.44	1.0	14.78	
		W3	2.40	1.0	13.12	
		W4	2.42	1.0	13.67	O2P (G16)
		W5	2.32	0.80	27.65	O2P (G17)
					35.50	O2P (G16)

likely that they simultaneously occupy this binding site. The geometry of the disodium coordination is close to that observed in the structure of ATP (Kennard et al. 1971). However, the binding environment is different. In the ATP structure, each sodium ion of the cluster is coordinated in a

symmetric manner. The coordination sphere of a Na<sup>+</sup> ion is composed of five anionic oxygen atoms and a N7 atom. The two sodium ions share two anionic oxygens and are surrounded by a more electronegative potential than in the present structure. Disodium that shares two water



**FIGURE 4.** Ion binding sites in the Br1 tridecamer. (A) Ionophilic half (G7–C13). (B) Hydrophilic half (G20–C26) (C) Superimposition of the ionophilic half (dark pink) and the hydrophilic half (gold). The distances between the phosphorus atoms are reported in ångströms. Distances between the corresponding phosphorus atoms of the hydrophilic half are depicted in parentheses.

molecules is observed in the structure of the salt  $\text{Na}_2\text{H}_2\text{P}_2\text{O}_6 \cdot 6\text{H}_2\text{O}$  (Collin and Willis 1971). Disodium that shares three ligands is observed in the structure of monosodium inosine 5'-phosphate (Rao and Sundaralingam 1969) and of sodium guanylyl-3', 5'-cytidine nonahydrate (Rosenberg

et al. 1976). In contrast to the present structure, the binuclear clusters are coordinated by a symmetric environment. Thus, the present structure shows that disodium can also exist in an asymmetric environment. Interestingly, the formation and the binding of a disodium ion has been observed in a recent molecular dynamics study of a DNA dodecamer (Várnai and Zakrzewska 2004). During the simulation, two sodium ions that share two water molecules of their inner coordination spheres bind to the anionic oxygen of a phosphate group (P. Várnai, pers. comm.). Other bimetallic clusters have been observed in nucleic acids. In the Z-DNA duplex  $d(m^5\text{CGTAm}^5\text{CG})$ , a mixture of magnesium and sodium ions was identified in a polynuclear cluster (Wang et al. 1984). In the structure of the loop E of 5S RNA, a magnesium binuclear cluster has been observed (Correl et al. 1997).

It is probable that the electronegative potential of O6 and O4 carbonyl groups of the GUUU sequence along an RNA strand has contributed to the binding of a bimetallic cluster. The disodium that bridges the consecutive uracil bases to their phosphate groups may also be a factor to rigidify the RNA structure. This would explain why the two structures, BrMn and NatMn, that lack Na5 display higher thermal B-factors than Br1 and Br2.

#### *Asymmetry of ion binding*

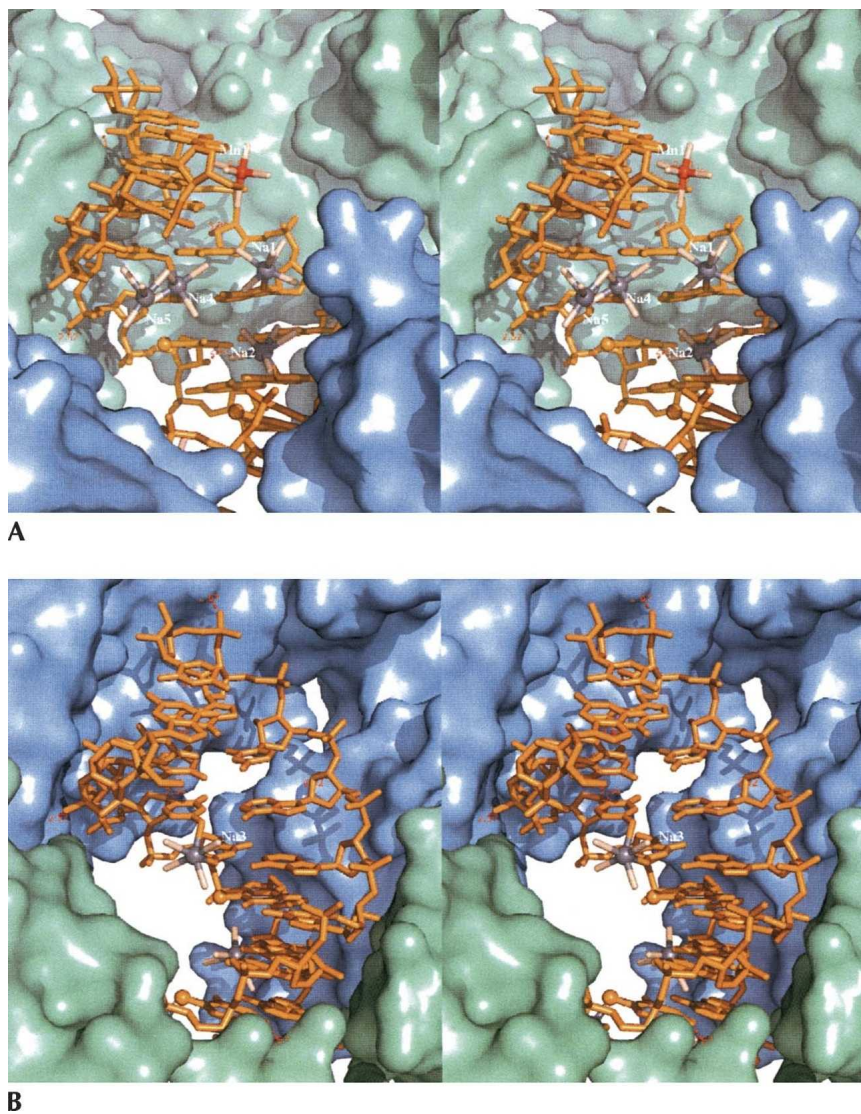
Despite their equivalent sequences, the two halves of the duplex display an asymmetric pattern of metal ion binding. In the ionophilic half, G7–C13, that binds five cations, a metal ion is found every two base pairs (Fig. 4A). The base pair A9:U18 exhibits the unusual property of being able to bind three metal ions. Interestingly, this correlates with the unstacking of the adenines at the A9/A10 step (Fig. 2A). Although this region adopts the geometry of a standard A-form RNA, its ionophilic character is reminiscent of the deep groove binding sites seen in the 5S rRNA loop E motif (Correl et al. 1997; Auffinger et al. 2004). However, despite a similar sequence, the equivalent half (G20–G26) only chelates one

sodium ion at the step U4/U5. This spectacular asymmetry provides the experimental evidence that factors other than the base sequence influence ion binding in RNA. This contrasts with the observation of symmetric ion binding sites within the 5S rRNA loop E motif (Auffinger et al. 2004). Two structural features seem to correlate with the ionophilic character of the half G7–C13/G14–G20 part of the tridecamer. First, the geometry of the ionophilic half is slightly more compact than the hydrophilic half. Figure 4C shows that the groove is narrower in the ionophilic half than in its symmetric counterpart. The contraction of the ionophilic groove is irregular and is indicated by the shorter distances between the phosphate groups of U5, A8, and A9 and that of C15, on the opposite strand. It is likely that the higher concentration of the negative charges around the metal ion binding sites in the ionophilic half has increased the residence time and favored the inner sphere binding of the metal ions on the RNA. However, it is difficult to determine if the helix contraction is the cause or the consequence of its ionophilic character. The second structural feature associated with the ion binding is the marked unstacking of adenines A9 and A10.

In space group P1 the two halves of the duplex are different by definition. Although the two halves are both surrounded by three equivalent RNA duplexes (S5, S6), they have a different electrostatic environment. The ionophilic one is locked between three molecules arranged in a more compact manner (Fig. 5). This analysis demonstrates that the electrostatic environment plays an important role for the binding of metal ions to RNA bases. This provides experimental support for the current view that the occupancies of metal binding sites in RNA are strongly influenced by negative electrostatic potential that is created by the folded RNA structures (Hermann and Westhof 1998; Banato et al. 2003). This study also reveals that the UUU triplet constitutes a strong binding site for monovalent cation that is less influenced by the environment.

## MATERIALS AND METHODS

The brominated  $r(\text{GCGUU}^{\text{Br}^-}\text{UGAAACGC})$  and native  $r(\text{GCGUUUGAAACGC})$  tridecamer were synthesized and puri-



**FIGURE 5.** Crystal packing environment around the ionophilic half (green) (A) and around the hydrophilic half (blue) (B).

fied by Dharmacon. Crystals were grown by the vapor diffusion method at 21°C. Hanging drops were made by mixing 2  $\mu\text{L}$  of RNA solution (0.5 mM) to an equivalent volume of well solution containing 50 mM Tris (pH 7), 100 mM NaCl, various concentrations of  $\text{MnCl}_2$  or  $\text{MgCl}_2$  (Table 1) and 30% of 2-methyl-2,4-pentane-diol (MPD). The data sets were collected at B30A beamline (ESRF). Crystals of the brominated molecule were used to collect data at three wavelengths for multiple-wavelength anomalous dispersion (MAD) experiments. Triclinic crystals diffracted up to 1.3 Å resolution. Data were processed with MOSFLM and scaled with SCALA of the CCP4 package (CCP4 1994). The heavy atom search, phasing, and solvent flattening were performed with CNS (Brunger et al. 1998). Two peaks were found by anomalous Patterson map analysis in all the crystals of the brominated tridecamer. However, the first phasing attempts using the data sets collected from a single crystal (Br2) diffracting up to 1.3 Å failed due to X-ray-induced debromination during data collection



(Ennifar et al. 2002). Growing larger crystals and collecting data at the different wavelengths from distinct crystals helped us to overcome this obstacle (Br1). The first experimental electron density maps obtained after solvent flattening showed that the asymmetric unit is composed of an RNA duplex. Subsequent cycles of building and refinement were performed with the program O (Jones et al. 1991) and CNS. The isomorphous crystal structures of the brominated (BrMn) and the native (NatMn) tridecamer grown only in the presence of manganese ions were refined independently. The statistics of phasing and refinement are reported in Table 1. Figures were made with the program Pymol (<http://www.pymol.org>). The coordinates and the structure factors are deposited in the PDB (Table 1). The helical parameters were analyzed with the program X3DNA (Lu and Olson 2003).

### Supporting information

Supporting information is available at [Timsit@ibpc.fr](mailto:Timsit@ibpc.fr).

### ACKNOWLEDGMENTS

We thank Peter Várnai for fruitful discussions and the careful reading of the manuscript. We are grateful to Philippe Dumas for helpful advice concerning bromine photolysis during data collection. We thank the people of the beamline BM30A team (European Synchrotron Radiation Facility, Grenoble, France).

### NOTE ADDED IN PROOF

The influence of sodium ions on RNA structure has been recently illustrated by the paper of Casiano-Negroni et al. (2007).

Received July 12, 2007; accepted August 28, 2007.

### REFERENCES

Auffinger, P., Bielecki, L., and Westhof, E. 2004. Symmetric K<sup>+</sup> and Mg<sup>2+</sup> ion-binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. *J. Mol. Biol.* **335**: 555–571.

Banatao, D.R., Altman, R.B., and Klein, T.E. 2003. Microenvironment analysis and identification of magnesium binding sites in RNA. *Nucleic Acids Res.* **31**: 4450–4460. doi: 10.1093/nar/gkg471.

Basu, S., Rambo, R.P., Strauss-Soukup, J., Cate, J.H., Ferré-D'Amaré, A.R., Strobel, S.A., and Doudna, J.A. 1998. A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nat. Struct. Biol.* **5**: 986–992.

Batey, R.T., Rambo, R.P., Lucast, L., Rha, B., and Doudna, J.A. 2000. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* **287**: 1232–1239.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The protein data bank. *Nucleic Acids Res.* doi: 10.1093/nar/28.1.235. **28**: 235–242.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 905–921.

Casiano-Negroni, A., Sun, X., and Al-Hashimi, H.M. 2007. Probing Na<sup>+</sup>-induced changes in the HIV-1 TAR conformational dynamics using NMR residual dipolar coupling; new insights into the role of the counterions and electrostatic interactions in adaptative recognition. *Biochemistry* **46**: 6525–6535.

Cate, J.H. and Doudna, J. 1996. Metal-binding sites in the major groove of a large ribozyme domain. *Structure* **4**: 1221–1229.

Collaborative Computing Project 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 760–763.

Collin, R.L. and Willis, M. 1971. The crystal structure of disodium dihydrogen hypophosphate hexahydrate (Na<sub>2</sub>H<sub>2</sub>P<sub>2</sub>O<sub>6</sub>•6H<sub>2</sub>O) and disodium dihydrogen pyrophosphate hexahydrate (Na<sub>2</sub>H<sub>2</sub>P<sub>2</sub>O<sub>7</sub>•6H<sub>2</sub>O). *Acta Crystallogr. B* **27**: 291–302.

Conn, G.L., Gittis, A.G., Lattman, E.E., Misra, V.K., and Draper, D.E. 2002. A compact RNA tertiary structure contains a buried backbone-K<sup>+</sup> complex. *J. Mol. Biol.* **318**: 963–973.

Correl, C.C., Freeborn, B., Moore, P.B., and Steitz, T.A. 1997. Metals, motifs and recognition in the crystal structure of a 5S rRNA domain. *Cell* **91**: 705–712.

Dange, V., Van Atta, R.B., and Hecht, S.M. 1990. A Mn<sup>2+</sup>-dependent ribozyme. *Science* **248**: 585–588.

Draper, D.E. and Misra, V.K. 1998. RNA shows its metal. *Nat. Struct. Biol.* **5**: 927–930.

Draper, D.E., Grilley, D., and Soto, A.M. 2005. Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **34**: 221–243.

Ennifar, E., Carpentier, P., Ferrer, J.-L., Walter, P., and Dumas, P. 2002. X-ray induced debromination of nucleic acids at the Br K absorption edge and implication for MAD phasing. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 1262–1268.

Ennifar, E., Walter, P., and Dumas, P. 2003. A crystallographic study of the binding of 13 metal ions to two related RNA duplexes. *Nucleic Acids Res.* **10**: 2671–2682. doi: 10.1093/nar/gkg350.

Hermann, T. and Westhof, E. 1998. Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations. *Structure* **6**: 1303–1314.

Hung, L.-W., Holbrook, E.L., and Holbrook, S.R. 2000. The crystal structure of the Rev binding element of HIV-1 reveals novel base pairing and conformational variability. *Proc. Natl. Acad. Sci.* **97**: 5107–5112.

Ippolito, J.A. and Steitz, T. 2000. The structure of the HIV-1 RRE high affinity Rev binding site at 1.6 Å resolution. *J. Mol. Biol.* **295**: 711–717.

Jack, A., Ladner, J.E., Rhodes, D., Brown, R.S., and Klug, A. 1977. A crystallographic study of metal-binding to yeast phenylalanine transfer RNA. *J. Mol. Biol.* **111**: 315–328.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110–119.

Kazakov, S. and Altman, S. 1992. A trinucleotide can promote metal-ion-dependent specific cleavage of RNA. *Proc. Natl. Acad. Sci.* **89**: 7939–7943.

Keel, A., Rambo, R., Batey, R., and Kieft, J. 2007. A general strategy to solve the phase problem in RNA crystallography. *Structure* **15**: 761–772.

Kennard, O., Isaacs, N.W., Motherwell, W.D.S., Coppola, J.C., Wampler, D.L., Larson, A.C., and Watson, D.G. 1971. The crystal and molecular structure of adenosine triphosphate. *Proc. R. Soc. Lond. A* **325**: 401–436.

Klein, D., Moore, P.B., and Steitz, T.A. 2004. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10**: 1366–1379.

Kopka, M.L., Lavelle, L., Han, G.W., Ng, H.-L., and Dickerson, R.E. 2003. An unusual sugar conformation in the structure of an RNA/DNA decamer of the polypurine tract may affect recognition by RNase H. *J. Mol. Biol.* **334**: 653–665.

Kuo, T.C. and Herrin, D.L. 2000. Quantitative studies of Mn<sup>2+</sup>-promoted specific and nonspecific cleavages of a large RNA: Mn<sup>2+</sup>-GAAA ribozymes and the evolution of small ribozymes. *Nucleic Acids Res.* **28**: 4197–4206. doi: 10.1093/nar/28.21.4197.

Lu, X.-J. and Olson, W.K. 2003. 3DNA: A software for the analysis, the rebuilding and the visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**: 5108–5121. doi: 10.1093/nar/gkg680.

- Martick, M. and Scott, W.G. 2006. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* **126**: 309–320.
- Mayer-Jung, C., Moras, D., and Timsit, Y. 1998. Hydration and recognition of methylated CpG steps in DNA. *EMBO J.* **17**: 2709–2718.
- McFail-Isom, L., Sines, C.C., and Williams, L.D. 1999. DNA structure: Cations in charge? *Curr. Opin. Struct. Biol.* **9**: 298–304.
- Murray, J.B., Seyhan, A.A., Walter, N.G., Burke, J.M., and Scott, W.G. 1998. The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem. Biol.* **5**: 587–595.
- Pan, B., Xiong, Y., Shi, K., Deng, J., and Sundaralingam, M. 2003. Crystal structure of an RNA purine-rich tetraplex containing adenine tetrads: Implications for specific binding in RNA tetraplexes. *Structure* **11**: 815–823.
- Pyle, A.M. 2002. Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.* **7**: 679–690.
- Rao, S.T. and Sundaralingam, M. 1969. Stereochemistry of nucleic acids and their constituents. V. The crystal and molecular structure of a hydrated monosodium inosine 5'-phosphate. A commonly occurring unusual nucleotide in the anticodons of tRNA. *J. Am. Chem. Soc.* **91**: 1210–1217.
- Rosenberg, J.M., Seeman, N., Day, R.O., and Rich, A. 1976. RNA double-helical fragments at atomic resolution: II. The crystal structure of sodium guanylyl-3',5'-cytidine nonahydrate. *J. Mol. Biol.* **104**: 145–167.
- Safranios, S.G., Das, K., Tantillo, C., Clark, A.D., Ding, J., Whitcomb, J.M., Boyer, P.L., Hughes, S.H., and Arnold, E. 2001. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**: 1449–1461.
- Scott, W.G., Finch, J.T., and Klug, A. 1995. The crystal structure of an all-RNA hammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. *Cell* **81**: 991–1002.
- Stefan, L.R., Zhang, R., Levitan, A.G., Hendrix, D., Brenner, S.E., and Holbrook, S. 2006. MeRNA: A database of metal ion binding sites in RNA structures. *Nucleic Acids Res.* **34**: D131–D134. doi: 10.1093/nar/gkj058.
- Telesnitsky, A. and Goff, S.P. 1997. Reverse transcriptase and the generation of retroviral DNA. In *Retroviruses* (eds. J.M. Coffin et al.), pp. 121–160. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Tereshko, V., Wilds, C.J., Minasov, G., Prakash, T.P., Maier, M.A., Howard, A., Wawrzak, Z., Manoharan, M., and Egli, M. 2001. Detection of alkali metal ions in DNA crystals using the state-of-the-art X-ray diffraction experiments. *Nucleic Acids Res.* **29**: 1208–1215. doi: 10.1093/nar/29.5.1208.
- Várnai, P. and Zakrzewska, C. 2004. DNA and its counterions: A molecular dynamics study. *Nucleic Acids Res.* **32**: 4269–4280. doi: 10.1093/nar/gkh765.
- Wang, H.-J., Hakoshima, T., van der Marel, G., van Boom, J.H., and Rich, A. 1984. AT base pairs are less stable than GC base pairs in Z-DNA: The crystal structure of d(m<sup>5</sup>CGTAm<sup>5</sup>CG). *Cell* **37**: 321–331.