RNA Hydration: A Detailed Look[†,‡]

Martin Egli,*[,§] Stefan Portmann,[§,||] and Nassim Usman[⊥]

Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School, Chicago, Illinois 60611-3008, Organic Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich, Switzerland, and Department of Chemistry and Biochemistry, Ribozyme Pharmaceuticals Inc., 2950 Wilderness Place, Boulder, Colorado 80301

Received March 25, 1996; Revised Manuscript Received May 6, 1996[⊗]

ABSTRACT: The crystal structure of the RNA duplex $[r(CCCCGGGG)]_2$ has been refined to 1.46 Å resolution with room temperature synchrotron diffraction data. This represents the highest resolution reported to date for an all-RNA oligonucleotide and is well beyond the best resolution ever achieved with an A-form DNA duplex. The analysis of the ordered hydration around the octamer duplex reveals conserved regular arrangements of water molecules in both grooves. In the major groove, all located first shell water molecules can be fitted into a pattern that is repeated through all eight base pairs, involves half the phosphate oxygens, and joins the two strands. In the minor groove, roughly across its narrowest dimension, tandem water molecules link the 2′-hydroxyl groups of adjacent nucleotides in base-pair steps in a similarly regular fashion. The structure provides evidence for an important role of the 2′-hydroxyl groups in the thermodynamic stabilization of RNA, beyond their known functions of locking the sugar pucker and mediating 3′ → 5′ intrastrand O2′···O4′ hydrogen bonds. The ribose 2′-hydroxyls lay the foundation for the enthalpic stability of the RNA relative to the DNA duplex, both as a scaffold for the water network in the minor groove and through their extensive individual hydration.

Progress in the chemical synthesis and purification of RNA oligonucleotides (Wincott *et al.*, 1995, and publications cited therein) has led to the availability of highly pure material for crystallization and X-ray structure determination. The number of solved RNA crystal structures is thus rising steadily [reviewed in Wahl and Sundaralingam (1995)] (Baeyens *et al.*, 1995; Schindelin *et al.*, 1995; Ott *et al.*, 1996; Portmann *et al.*, 1996; Wahl *et al.*, 1996), and the next years promise progress in the elucidation of RNA and RNA−protein complex structures. We reported the crystallization (Egli *et al.*, 1995) and the structures of a hexagonal and a rhombohedral crystal form of the RNA octamer r(CCCCGGGG) or $r(C_4G_4)$ at resolutions of 2.8 and 1.8 Å, respectively (Portmann *et al.*, 1995). Crystals of the rhombohedral form diffract X-rays to exceptionally high resolution and allow a systematic study of RNA hydration. Here, we describe the regular network of water molecules surrounding the RNA duplex, on the basis of a refinement with X-ray synchrotron diffraction data to 1.46 Å. The quality of the data permits location of nearly all molecules in the first hydration sphere. Thanks to the repetition in the base sequence, the arrangement of water molecules is highly regular. The repeated water patterns in both grooves of the duplex and the fact that these patterns are practically uninterrupted even by interduplex contacts could lead one to consider the first hydration shell as part of the RNA. The 2′-OH in RNA is the only H-bond donor in the backbone. Since the O−H···O (lone pair) is a strongly directional interaction, the presence of the 2′-OH groups not only brings water molecules into regions that are hydrophobic in the DNA duplex but also provides a set of strongly directional arms on which to build the rest of the water molecule framework. For donor ability, DNA has to rely completely on the donor groups of the bases. The extensive hydration of the RNA minor groove in the present structure, aided by the ribose 2′-hydroxyl groups along its rim, and the regularity of the water distribution in both grooves suggest an important role of such structured water molecules in RNA−RNA and RNA−protein recognitions.

METHODS

Crystallization and Data Collection. Rhombohedral crystals of the RNA octamer $r(C_4G_4)$ can be grown from high gradients of ammonium sulfate within a pH range of 4.6−8.5 (Egli *et al.*, 1995). The crystal used for data collection was obtained at room temperature from a 10 μL hanging droplet, containing 1 mM RNA oligonucleotide (single stranded), 50 mM sodium acetate, pH 4.6, and 1 M ammonium sulfate, equilibrated against a 0.5 mL reservoir (100 mM sodium acetate, pH 4.6, 2 M ammonium sulfate). The crystal of size 0.4 × 0.4 × 0.4 mm, which is depicted in Figure 2d of a preliminary crystallization and X-ray diffraction report for the $r(C_4G_4)$ crystals (Egli *et al.*, 1995), was mounted in a thin-walled glass capillary with a droplet of mother liquor prior to transport. The space group is *R*32 and the cell constants are $a = b = 42.26$ Å and $c = 130.89$ Å (hexagonal setting). Data to 1.46 Å resolution were collected at room temperature on the X31 beamline of the European Molecular Biology Laboratory (EMBL) outstation at the German Electron Synchrotron (DESY) in Hamburg, Germany. The wavelength was 0.919 Å, the crystal-image

[†] This paper is dedicated to Professor Vladimir Prelog, ETH-Zürich, on the occasion of his 90th birthday. S.P. is supported by an ETH graduate fellowship.

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (filename 259D) and the Nucleic Acid Database (ARH074).

* To whom correspondence should be addressed.

[§] Northwestern University Medical School.

[||] Swiss Federal Institute of Technology.

[⊥] Ribozyme Pharmaceuticals Inc.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1996.

S0006-2960(96)00721-0 CCC: $12.00

Table 1: Completeness and Quality of the Crystallographic Data and Refinement Results[a]

shell limits (Å)	$I \leq 1\sigma(I)$ (%)	R_{merge} (%)[b]	completeness (%)[b]	*R*-factor (%)[c]
10.00−3.89	0.9	5.7	99.1	12.4
3.89−3.12	0.2	4.7	97.2	12.9
3.12−2.73	0.7	3.2	98.8	17.4
2.73−2.49	1.7	3.3	99.8	20.5
2.49−2.31	2.6	3.1	100.0	19.7
2.31−2.18	1.8	3.3	100.0	22.4
2.18−2.07	2.4	3.9	100.0	25.9
2.07−1.98	4.2	4.3	100.0	24.5
1.98−1.90	2.8	5.5	100.0	26.7
1.90−1.84	3.2	6.2	100.0	26.0
1.84−1.78	6.7	7.6	100.0	27.6
1.78−1.73	5.2	5.6	100.0	26.7
1.73−1.68	7.5	4.2	99.7	26.6
1.68−1.64	8.3	5.5	100.0	28.3
1.64−1.61	10.6	10.1	100.0	32.8
1.61−1.57	12.6	15.5	100.0	31.7
1.57−1.54	22.0	26.3	100.0	35.0
1.54−1.51	25.9	24.5	100.0	34.3
1.51−1.49	18.5	34.8	99.8	34.1
1.49−1.46	25.7	49.9	97.1	36.9
10.00−1.46	8.0	5.1	99.6	20.1

[a] The final rms deviations for bond lengths are 0.012 Å, and for bond angles they are 3.04°. [b] $R_{merge} = \sum(I - \langle I \rangle)^2/\sum I^2$, where $\langle I \rangle$ is the average intensity of a particular reflection. No σ limit was applied to determine R_{merge}, and the completeness refers to all unique 8040 data. [c] Based on a total of 7858 unique reflections with $I \geq 1\sigma(I)$.

plate distance was 120 mm, and 51 frames with a frame size of 1.5° were collected on a MAR research 180 mm image plate system. The 46 226 reflections were indexed with the program DENZO (Otwinowski, 1993) and merged to 8040 unique reflections with the program SCALEPACK (Minor, 1993). Details of the resolution, completeness, and quality of the data are given in Table 1. It is noteworthy that the current resolution limit does not represent the physical diffraction limit of the rhombohedral crystal from. We believe that data collection on beamlines with higher flux, such as the Advanced Photon Source at Argonne National Laboratory outside Chicago whose construction is currently being completed, in combination with flash-freezing, could ultimately yield diffraction data for the RNA octamer with a resolution limit near 1.2 Å.

Crystal Structure Refinement and Water Molecule Assignments. The coordinates of the room temperature 1.8 Å resolution structure of the rhombohedral crystal form minus the water molecules (Portmann *et al.*, 1995; PDB entry code 1RXB) were used for the initial refinement cycles. Refinement was performed with the program X-PLOR (Brünger, 1992a), using the standard RNA dictionaries. Positional refinement was applied several times, and the resolution of the data was continuously extended. Fourier electron density maps were displayed at 1σ ($2F_o - F_c$ sum) and 2.5σ ($F_o - F_c$ difference) levels on a Silicon Graphics Indigo2 graphics workstation, and water molecules were assigned to regions of superimposed sum and difference density. After placement of a total of 92 water molecules using 92% of the data at a 4σ (F_o) level to full resolution [8% of the reflections were set aside for calculation of R_{free} (Brünger, 1992b)], refinement was stopped. The population factor of all atoms was kept at 1.0, and the weighting factor of all energy terms in X-PLOR for treatment of contacts between water molecules was lowered to 0.2. The water shells surrounding the RNA double helix were then visually inspected for patterns within major and minor grooves and around backbone phosphate groups. The longitudinal hydration motifs in the major groove and the transversal bridging of strands across the minor groove were immediately apparent, although interrupted at a few places. During the refinement, the electron density at some assumed locations of water molecules was only minimal. This had been attributed to partial occupancy of these positions, and such water molecules were then rejected. In the sequel, the very regular nature of the water structure around the RNA prompted us to inspect the electron density maps once more in regions where the hydration motifs were apparently disrupted. Indeed, using the same map density display levels as before, significant amounts of difference density and in some cases sum density were located at every one of the suspected hydration sites and at several additional locations, often within a 3.5 Å range from the RNA. An additional 21 water molecules were placed in several rounds of refinement. The reflections previously used for the calculation of R_{free} (the value before the final cycle was 26.3%) were included in a final cycle. For water molecules, both the population parameters and the thermal *B*-factors were varied. The final *R*-factor was 20.1% including 338 RNA atoms and 113 water molecules, based on 7858 2σ (F_o) reflections. A resolution breakdown of the correspondence between measured and calculated structure factor amplitudes in the final structure is given in Table 1.

Analysis of Water Structure. Residues on one strand of the RNA are numbered 1−8, and residues on the other are numbered 9−16. Water molecules within 3.5 Å of RNA donor or acceptor functions were generally considered to constitute the first hydration shell. In the duplex, all potential hydrogen bond donors and acceptors, except the ribose O4′ and the backbone O3′ and O5′ atoms (Table 2), are engaged in at least one contact to a water molecule. In the deep major groove, the solvation is undisturbed, but in the wide and shallow minor groove water molecules with direct contacts to the RNA are occasionally replaced by phosphate and ribose 2′-oxygen atoms from neighboring duplexes. However, the overall hydration pattern appears practically undistorted as a consequence of interactions between duplexes. The patterns best defined involve water molecules located near the floor of both grooves. Near the periphery of the major groove the hydration is generally less well defined, even at this high resolution. At its narrowest points between sugars the minor groove can just be spanned by two water molecules, each contacting the bases as well. Water bridges involving more than two water molecules are not included in the subsequent discussion, as they are likely to be of less importance for stabilization and recognition.

RESULTS AND DISCUSSION

Hydration Sites and Overall Hydration. Comparison between the hydration of individual hydrophilic sites in the $[r(C_4G_4)]_2$ RNA duplex and the A-DNA duplex $[d(GG^{Br}U-A^{Br}UACC)]_2$ [whose hydration had been analyzed at 1.7 Å resolution, the highest reported for an A-DNA duplex (Kennard *et al.*, 1986)] reveals a dramatically increased hydration of the sugar−phosphate backbone in the RNA (Table 2). Table 2 shows that there are many more contacts

Table 2: Total and Average[a] Number of Hydrogen Bonds to Water Molecules by Hydrophilic Nucleic Acid Atoms in the Crystal Structures of the RNA Duplex $[r(C_4G_4)]_2$ and the A-DNA Duplex $[d(GG^{Br}UA^{Br}UACC)]_2$[b]

acceptor/donor atom	$[r(C_4G_4)]_2$	$[d(GG^{Br}UA^{Br}UACC)]_2$
entire duplex	170	110
phosphates	51	57
sugars	64	10
bases	55	43
minor groove	0.4	0.4
O2 (Py)	1.0/0.4	0.9/0.5
N2 (G)	1.1/0.5	0.5/0.3
N3 (Pu)	1.0/0.4	1.0/0.5
major groove	0.5	0.5
N4 (C)	1.1/0.5	1.0/0.4
O6 (G)	1.4/0.6	1.5/0.8
backbone	0.5	0.2
O2′	2.0/0.9	
O3′ chain	0.9/0.4	0.1/0.0
O3′ terminal	0.5	2.5
O4′	0.4/0.2	0.6/0.3
O5′ chain	0.6/0.3	0.2/0.1
O5′ terminal	1.5	1.0
anionic oxygens	0.8	0.9
O1P	2.3/1.0	1.9/1.0
O2P	1.4/0.6	1.4/0.7

[a] In the cases where hydration of individual atoms is analyzed, the first numbers are the ratios between all the water molecules within 3.5 Å [DNA 3.4 Å (Westhof, 1987)] from each acceptor or donor atom type in the RNA/DNA duplex and the number of occurrences of that particular atom type in the two strands. The second numbers are the first values normalized with regard to the largest relative affinity in each of the duplexes (2.3 for O1P in the RNA and 1.9 for O1P in the DNA). An analogous analysis was previously conducted with representative structures from the right- and left-handed DNA duplex families (Westhof, 1987), and the numbers for the A-DNA octamer duplex were taken from that publication. [b] Kennard *et al*., 1986; Westhof, 1987.

to waters by base atoms in the minor groove of the RNA relative to the DNA. The presence of the additional hydroxyl group in the sugar moiety not only confers more water on the RNA backbone portion, but also leads to an overall more ordered water structure in its minor groove. Further notable differences in the relative hydrations of RNA and DNA duplexes examined in Table 2 are the improved hydration of the bridging phosphate oxygens in the RNA backbone and the additional water contacts by the exocyclic $N(2)H_2$ amino group of guanosines in the RNA minor groove. In $[r(C_4G_4)]_2$, O1P phosphate oxygens are systematically bridged by water molecules along both strands (see below), an A-conformation-specific feature that was previously observed in portions of the backbone in crystal structures of A-form duplexes [e.g., Kennard *et al*. (1986), Saenger *et al.* (1986), Leonard *et al.* (1994), and Westhof (1987)]. Relative to B-DNA, intrastrand phosphates are more closely spaced in A-form duplexes, allowing a single water molecule to form tight hydrogen bonds to two adjacent phosphate groups. In the $[r(C_4G_4)]_2$ duplex, such water molecules lie close to the O5′ atom of a residue in O1P−W−O1P bridged nucleotides. On the one hand, these contacts may be considered to be secondary since the average W−O5′ distance of 3.37 Å (only including contacts below 3.5 Å) is above the average distance of 3.17 Å for the primary W−O1P hydrogen bond. On the other hand, this difference is small, and a more detailed analysis given in the major groove hydration section below shows that these waters are not located halfway between the two O1Ps but are consistently drawn toward the one from the 5′-residue. Thus, this pattern may represent a water molecule shared among two adjacent O1P oxygens and a 5′-oxygen. Closer examination of the apparently improved hydration of the 3′-oxygens in the RNA duplex reveals that water molecules within 3.5 Å are primarily bound to the ribose 2′-hydroxyl groups from the same nucleotide. Thus, the average distance between 3′-oxygens and waters is 3.25 Å, while that between these waters and the 2′-hydroxyl oxygens is 2.93 Å. These water molecules can be recognized as cluster b in Figure 1. In three instances, water molecules are shared between 3′-oxygens and O2P phosphate oxygens from 5′-adjacent residues. Again, the primary contact is to the phosphate oxygens, with a longer distance for W−O3′ (average distances 3.01 and 3.25 Å). A water molecule bonded to the O1P and O2P oxygens of the same phosphate was observed once in the RNA duplex, at residue G7. A similar explanation as in the cases of the O5′ and O3′ atoms can be invoked for the presence of additional water molecules in the vicinity of the exocyclic amino group of G residues in the minor groove. In two residues, this group donates in two hydrogen bonds, but one of the water molecules is clearly also bonded to the neighboring N3, and its distance to the latter is shorter. In other residues, only one water molecule contacts the guanine minor groove edge, and it is then positioned at nearly equal distances from the donor N2 and the acceptor N3. In yet other residues, each nitrogen contacts a different water molecule. In every case, the increased hydration of guanine N2 in the minor groove can be attributed at least in part to the systematic, O2′-mediated hydration of the adjacent N3.

O2′ Hydration. A total of 33 hydrogen bonds to water molecules by the 16 2′-hydroxyl groups per duplex are observed in the $[r(C_4G_4)]_2$ crystal structure. This places them on a similar level as the phosphate oxygens in terms of hydration (Table 2). Around certain 2′-hydroxyl groups three well ordered water molecules are found in the $[r(C_4G_4)]_2$ duplex. The discrepancy in the relative hydration levels of the O1P and O2P oxygens may be due to the fact that the former are directed into the shielded major groove, whereas the latter jut out into the open solvent channels. Contrary to water molecules around phosphates, those bonded to 2′-oxygens are involved in additional contacts to polar atoms (O3′, O4′, O2P). Indeed, they are embedded in a water network that spans the floor of the minor groove, the 2′-hydroxyls serving as a scaffold. A scatter plot for the positions of water molecules around 2′-oxygens is depicted in Figure 1 and reveals four water clusters, labeled a, b, c, and d. Cluster a consists of waters that bridge the 2′-oxygens to guanine N3 or cytosine O2. A second cluster b is formed by water molecules that link 2′-oxygens to backbone O3′ atoms (see preceding paragraph). Together with the C2′ atoms and the locations of water molecules within the a and b clusters, the two additional clusters c and d each represent the fourth vertex of an approximate tetrahedron around the O2′ positions. The presence of water molecules above and below the ribose hydroxyl groups (Figure 1), in addition to the waters bridging these groups to backbone and base sites, suggests that 3′−5′ intrastrand hydrogen bonds between ribose O2′ and O4′ are weak and may not account for a significant stabilization of the RNA A-conformation. As apparent in Figure 1, the distances between O2′ atoms and adjacent O4′ atoms are typically longer than the hydrogen-bonding cutoff of 3.5 Å; the average value is 3.68 Å, with

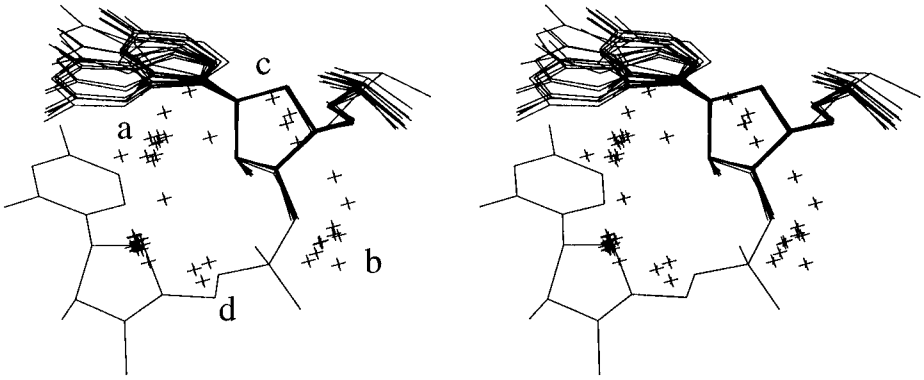

FIGURE 1: Distribution of water molecules (crosses represent oxygen positions) around the ribose 2′-hydroxyl groups within a distance range of 3.5 Å for 16 nucleotides (bold) in $[r(C_4G_4)]_2$. The four clusters are labeled a, b, c, and d. The superposed nucleotides (top residues) all adopt C3′-*endo* type puckers, and the 5′-adjacent residue was included to provide for a typical environment of the ribose. The locations of 5′-adjacent O4′ oxygens were included for comparison, independent of whether the corresponding O2′···O4′ contact was shorter than 3.5 Å or not, and are discernible as tightly clustered crosses around the O4′-position of the bottom residue.

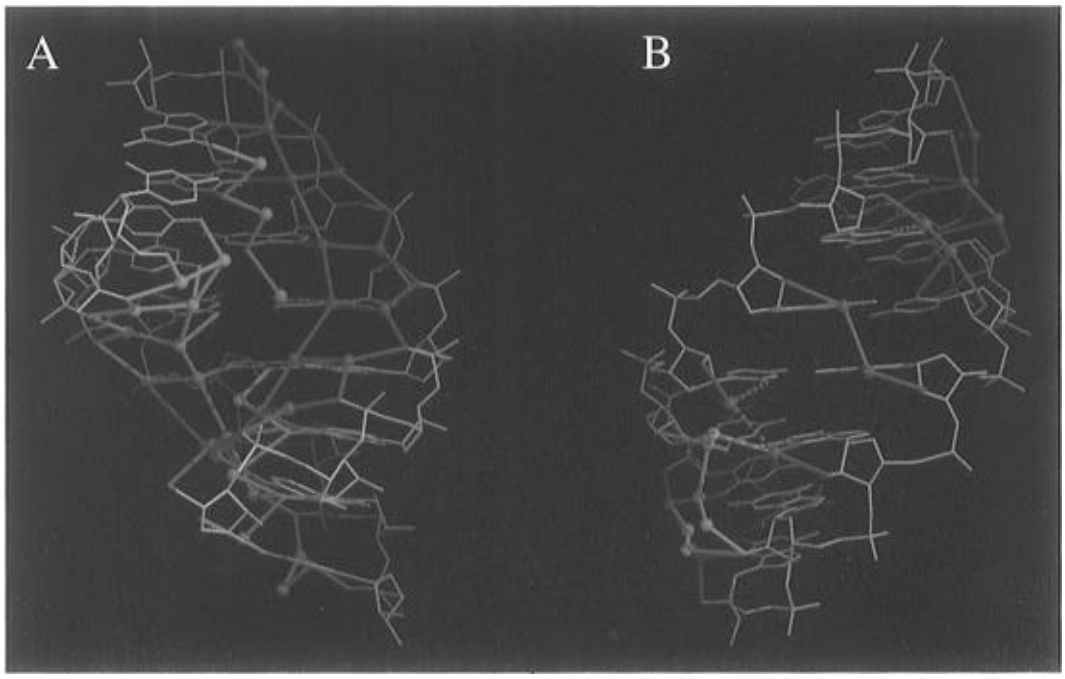

FIGURE 2: Overall views of the hydration patterns in the RNA (A) major and (B) minor grooves (related by a rotation of ca. 180° around the horizontal). RNA atoms are white, and bright portions are closer to the reader. In the major groove, water pentagons along strand 1 are red, along strand 2 they are green, and water bridges between O6 atoms of guanines are yellow and cyan in the top and bottom halves of the duplex, respectively. In the minor groove, water bridges between 2′-hydroxyl groups from opposite strands within base-pair steps are red (dotted for distances >3.5 Å). Bridges between 2′-hydroxyl groups within base-pair planes and sharing a water molecule with the former bridges are cyan. Two water molecules and a phosphate oxygen from an adjacent duplex link the hydroxyl groups of residues G7 and C11 and are colored yellow.

a minimum of 3.37 Å and a maximum of 3.89 Å. Evidently, water molecules are better acceptors than ribose O4′.

The Water Network in the Major Groove. In the $[r(C_4G_4)]_2$ duplex, cytidines form hydrogen bonds to two water molecules, and guanidines form hydrogen bonds to three water molecules in the major groove. The waters bonded to the O1P and O6 oxygens are shared among adjacent bases, while those bonded to N4 and N7 atoms of C and G residues, respectively, are associated with individual bases. Thus, water molecules bridging adjacent O1P oxygens and those bonded to N7 of G or N4 of C are located roughly within the planes defined by individual bases (Figure 2A). They are linked and form the joint edges of pentagons which wind down the outer portion of the major groove along both strands, with O1P atoms sitting at one corner in the pentagons. As shown in Figure 3, water molecules bridging O1P atoms are usually closer to the O1P of the 5′-residue, the one sitting between G5 and G6 being the only exception. At the duplex termini, O5′ atoms of residues C1 and C9[#] (from a stacked, symmetry-related duplex) replace the missing O1P atoms. The water ribbon smoothly follows the curvature of the RNA strands, resulting in puckering of the individual pentagons. The water molecules that bridge O1P atoms on the border of the groove are too far away from one another to form direct hydrogen bonds; those on the inner side of the pentagon ribbon and closer to the center of the groove are hydrogen bonded to one another (Figures 2A and 3). In the case of G residues, these water molecules are linked in turn to further waters that bridge O6 atoms of adjacent bases in the G4 stretches in both duplex halves (Figures 2A and 3). Interestingly, no ordered water molecule bridges the O6 atoms of G5 and G13 in the central base-

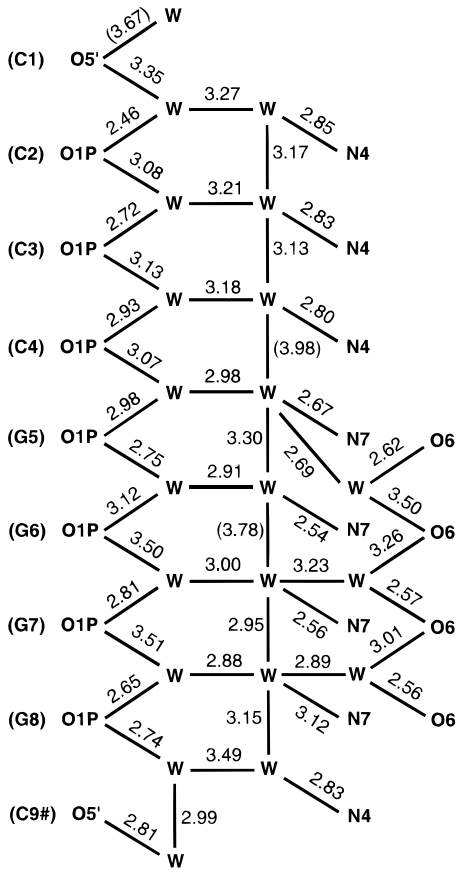

FIGURE 3: Schematic diagram of the longitudinal water structure in the major groove for strand 1 (residues C1 to G8; strand 2 very similar). Distances are in Ångstroms and are placed in parentheses where longer than 3.5 Å.

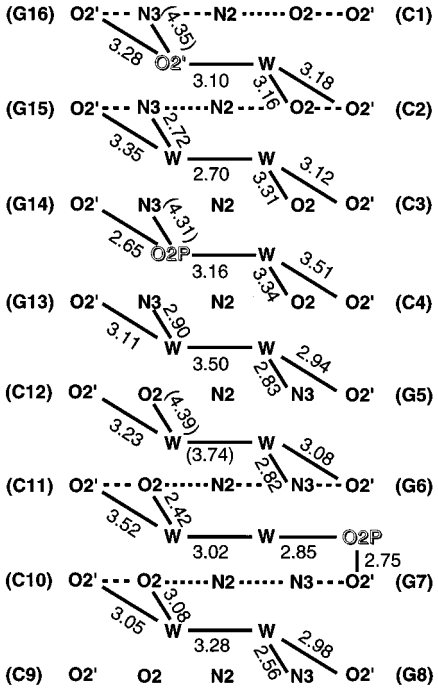

FIGURE 4: Schematic diagram of the transversal water structure in the minor groove, bridging 2′-hydroxyl groups of adjacent residues from opposite strands across the narrowest dimension of the groove. Distances are in Ångstroms and are placed in parentheses where longer than 3.5 Å. RNA atoms from neighboring duplexes replacing water molecules in the hydration motif are drawn with outline font. Water bridges across the groove and roughly within the planes of base pairs are included as dashed lines.

pair step, although the O···O distance of 3.4 Å is close to the separation between O6 pairs in stacked intrastrand guanines (Figure 2A). The waters bridging O6 atoms are systematically closer to one of the bases, lying more or less in the plane defined by that base, portrayed in Figure 2A (center of the major groove, with the yellow and the cyan water lying approximately in the planes defined by G5 and G13, respectively). The distances given in Figure 3 show the asymmetric positioning of the bridging water molecules. At the central step, there are thus two possibilities: linking O6 atoms directly or via bridging water molecules. This may result in positional disorder of a water molecule at that site, obscuring its presence in electron density maps. It is noteworthy that the central base-pair step in the crystal structure of the A-DNA $[d(C_4G_4)]_2$ features a water molecule that contacts both O6 atoms as well as two water molecules bound to the keto oxygens, suggesting bifurcated hydrogen bonds or orientational disorder of that water (Eisenstein & Shakked, 1995). However, no systematic bridging of O6 atoms was observed in that structure. The $[d(C_4G_4)]_2$ crystal structure displays an A′ conformation with enlarged distances between adjacent intrastrand phosphates (Haran *et al.*, 1986; Portmann *et al.*, 1995). The larger spacing between phosphates results in the loss of water molecules bridging O1P atoms on the border of the groove. The integral role of these water molecules in the major groove hydration network in the RNA duplex may explain the lack of a more systematic hydration of that groove in many A-DNA structures that exhibit higher backbone flexibility and a range of intrastrand phosphate distances. Thus, the enthalpic stability gain of the RNA duplex relative to the DNA duplex can be attributed in part directly to the ribose 2′-hydroxyl groups which lock the sugar−phosphate backbone in a conformation that allows waters to bridge adjacent phosphates.

The Water Network in the Minor Groove. In the $[r(C_4G_4)]_2$ duplex, roughly across the narrowest dimension of the minor groove, O2′ atoms from adjacent residues in base-pair steps are separated by 8.52 Å on average. This gap can just be spanned by two water molecules, each residing more or less in the base plane of the nucleotide contributing its O2′ and contacting either the O2 of cytidines or the N3 of guanidines (see center of the minor groove in Figure 2B). Occasionally, these water molecules are further removed from the floor of the minor groove. In the crystal structure, duplexes are stacked coaxially with their backbones inserted into the minor grooves of nearest neighbors and vice versa (Portmann *et al.*, 1995). Thus, hydrophilic atoms from such duplexes can replace water molecules within the hydration network in the minor groove (Figure 4). The average distance between O2′ atoms across strands and roughly within base-pair planes is 11.30 Å, considerably longer than the above. At base pairs G6·C11 and G7·C10, three water molecules connect the O2′ atoms from paired nucleotides (Figure 5). In the former base pair, each of these waters contacts the minor groove edges of the bases, while the water molecule bordering the backbone in the latter one is too far removed to form a hydrogen bond to the N3 atom (Figure 5, bottom). Thus, both prevailing hydration motifs in the minor groove are transversal rather than longitudinal as in the major groove, and they share a water molecule. In the two cases discussed above, these branch point waters are hydrogen bonded to cytosine O2 atoms, while at the opposite end of the duplex,

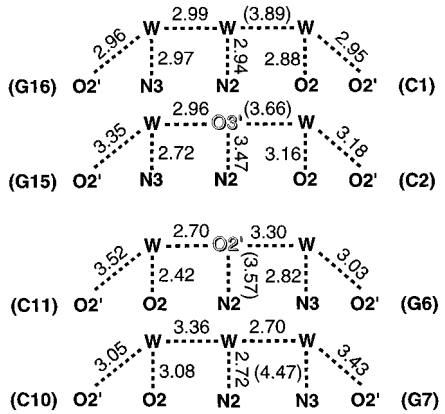

FIGURE 5: Schematic diagram of the second transversal hydration pattern in portions of the duplex, bridging 2′-hydroxyl groups of paired nucleotides approximately within the base-pair plane across the minor groove. Distances are in Ångstroms and are placed in parentheses where longer than 3.5 Å. RNA atoms from neighboring duplexes replacing water molecules in the hydration motif are drawn with outline font.

they are hydrogen bonded to guanine N3 atoms. However, the distances between O2′ atoms from opposite strands at these sites are somewhat too long to be effectively bridged by three water molecules (Figure 5, top).

Conclusions. We have shown that, in its crystal structure, the RNA duplex $[r(C_4G_4)]_2$ is much more extensively hydrated than the DNA duplex $[d(GG^{Br}UA^{Br}UACC)]_2$ (Kennard *et al.*, 1986) or the DNA duplex $[d(C_4G_4)]_2$ with identical sequence (Haran *et al.*, 1986; Eisenstein & Shakked, 1995). Indirectly, through maintaining an A-type backbone conformation with closely spaced phosphate groups, and directly, by linking backbone and base hydration across strands, the ribose 2′-hydroxyl groups propagate stable and conserved water networks in both grooves of the RNA duplex. Consistent with the higher number of waters in its first hydration shell, the higher thermodynamic stability of $[r(C_4G_4)]_2$ relative to $[d(C_4G_4)]_2$ [$\Delta T_m = 25.5$ °C, $\Delta\Delta G = -7.1$ kcal mol^{-1} (at 37 °C) (Egli, 1996)] is enthalpy driven ($\Delta\Delta H = -31.1$ kcal mol^{-1}), with the entropic term favoring the DNA duplex ($-T\Delta\Delta S = +24$ kcal mol^{-1}). Although it may appear to be preorganized for duplex formation because of reduced flexibility in the ribose moieties compared with deoxyriboses, RNA pays a large entropic price for the improved hydration of its backbones and grooves. The cyclic pentamer is a well established low-energy structural component of water (Liu *et al.*, 1996, and publications cited therein). It also constitutes one of the basic motifs in the solvation of nucleic acids (Neidle *et al.*, 1980; Kennard *et al.*, 1986; Bingman *et al.*, 1992) and proteins (Teeter, 1984; Baker *et al.*, 1985) in addition to the more common three- and four-membered rings (Lipscomb *et al.*, 1996). In place of the concatenated or fused water pentamers observed in these structures, the pentagons in the major groove of the RNA are formed by four waters and a phosphate oxygen, a pattern also seen at isolated locations in the crystal structure of the A-form DNA $[d(G_3CGC_3)]_2$ (Eisenstein & Shakked, 1995). The water structure in the major and minor grooves of $[r(C_4G_4)]_2$ lends strong support to the idea that the arrangement of water molecules around bases in oligonucleotides is systematic (Schneider *et al.*, 1993). Although the water network may appear to be complex, it can consist of simple repeated modules. In intact stem regions, RNA base composition should not influence the arrangement of tandem waters across the minor groove to a major extent as all combinations of standard base pairs can in principle provide an acceptor function on either border of that groove to stabilize the water molecules bound to O2′ atoms. Similarly, the pentameric hydration motif on the major groove borders may not only exist in GC-rich regions but may also prevail at A·U base pairs.

ACKNOWLEDGMENT

We are grateful to Dr. Zbyszek Dauter of EMBL for his help during data collection and processing and to Prof. Jack D. Dunitz, ETH-Zürich, for advice and suggestions on the manuscript.

REFERENCES

Baeyens, K. J., De Bondt, H. L., & Holbrook, S. R. (1995) *Nat. Struct. Biol. 2*, 56−62.
Baker, T., Dodson, E., Dodson, G., Hodgkin, D., & Hubbard, R. (1985) in *Crystallography in Molecular Biology*, pp 179−192, Plenum, New York.
Bingman, C., Li, X., Zon, G., & Sundaralingam, M. (1992) *Biochemistry 31*, 12803−12812.
Brünger, A. T. (1992a) *X-PLOR, A System for X-ray Crystallography and NMR* (*Version 3.1*), Yale University Press, New Haven, CT.
Brünger, A. T. (1992b) *Nature 355*, 472−475.
Egli, M. (1996) *Angew. Chem., Int. Ed. Engl. 35* (in press).
Egli, M., Portmann, S., Tracz, D., Workman, C., & Usman, N. (1995) *Acta Crystallogr. D51*, 1065−1070.
Eisenstein, M., & Shakked, Z. (1995) *J. Mol. Biol. 248*, 662−678.
Haran, T. E., Shakked, Z., Wang, A. H.-J., & Rich, A. (1986) *J. Biomol. Struct. Dyn. 5*, 199−217.
Kennard, O., Cruse, W. B., Nachman, J., Prange, T., Shakked, Z., & Rabinovich, D. (1986) *J. Biomol. Struct. Dyn. 3*, 623−647.
Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T., & Hunter, W. N. (1994) *Structure 2*, 483−494.
Lipscomb, L. A., Zhou, F. X., & Williams, L. D. (1996) *Biopolymers 38*, 177−181.
Liu, K., Cruzan, J. D., & Saykally, R. J. (1996) *Science 271*, 929−933.
Minor, W. (1993) *XDISPLAYF Program*, Purdue University, West Lafayette, IN.
Neidle, S., Berman, H., & Shieh, H. S. (1980) *Nature 288*, 129−133.
Ott, G., Dörfler, S., Sprinzl, M., Müller, U., & Heinemann, U. (1996) *Acta Crystallogr. D52* (in press).
Otwinowski, Z. (1993) Oscillation Data Reduction Program, in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N., & Bailey, S., Eds.) pp 56−62, SERC Daresbury Laboratory, England.
Portmann, S., Usman, N., & Egli, M. (1995) *Biochemistry 34*, 7569−7575.
Portmann, S., Grimm, S., Workman, C., Usman, N., & Egli, M. (1996) *Chem. Biol. 3*, 173−184.
Saenger, W., Hunter, W. N., & Kennard, O. (1986) *Nature 324*, 385−388.
Schindelin, H., Zhang, M., Bald, R., Fürste, J. P., Erdmann, V. A., & Heinemann, U. (1995) *J. Mol. Biol. 249*, 595−603.
Schneider, B., Cohen, D. M., Schleifer, L., Srinivasan, A. R., Olson, W. K., & Berman, H. M. (1993) *Biophys. J. 65*, 2291−2303.
Teeter, M. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6014−6018.
Wahl, M. C., & Sundaralingam, M. (1995) *Curr. Opin. Struct. Biol. 5*, 282−295.
Wahl, M. C., Rao, S. T., & Sundaralingam, M. (1996) *Nat. Struct. Biol. 3*, 24−31.
Westhof, E. (1987) *Int. J. Biol. Macromol. 9*, 186−192.
Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Trazc, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., & Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.

BI9607214