Biophysical Chemistry 95 (2002) 273–282

Biophysical Chemistry

www.elsevier.com/locate/bpc

Hydrogen and hydration of DNA and RNA oligonucleotides

Muttaiya Sundaralingam*, Baocheng Pan

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

Received 26 January 2001; accepted 15 May 2001

Abstract

In this paper, hydrogen bonding interaction and hydration in crystal structures of both DNA and RNA oligonucleotides are discussed. Their roles in the formation and stabilization of oligonucleotides have been covered. Details of the Watson–Crick base pairs G•C and A•U in DNA and RNA are illustrated. The geometry of the wobble (mismatched) G•U base pairs and the *cis* and almost *trans* conformations of the mismatched U•U base pairs in RNA is described. The difference in hydration of the Watson–Crick base pairs G•C, A•U and the wobble G•U in different sequences of codon–anticodon interaction in double helical molecules are indicative of the effect of hydration. The hydration patterns of the phosphate, the 2′-hydroxyl groups, the water bridges linking the phosphate group, N7 (purine) and N4 of Cs or O4 of Us in the major groove, the water bridges between the 2′-hydroxyl group and N3 (purine) and O2 (pyrimidine) in the minor groove are discussed. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Hydrogen bonding; Hydration; X-ray crystal structures; Oligonucleotides; Stabilization

1. Introduction

Hydrogen bonding and hydration are two important aspects in nucleic acids. Both interactions are closely related to hydrogen bond donors and acceptors of bases, sugars and phosphates. The different patterns of hydrogen bond donors and acceptors will result in different patterns of hydrogen bonding and hydration patterns. Watson–Crick base pairs of hydrogen bonding patterns are crucial in the genetic information transfer from one generation to next in DNA. The four bases in DNA and RNA have different patterns of hydrogen bonding donors and acceptors and can have 31 different hydrogen bonding patterns [18,20]. As more and more crystal structures of oligonucleotides are available, it is expected that some new forms of hydrogen bond patterns will be found, such as C–H….O [25] and C–H….N hydrogen bonding [17].

The polynucleotides of DNA can adopt a B-form helix in high humidity and change to an A-form duplex in low humidity [10]. It has been proposed that the low humidity form of A-DNA has less hydration and may be considered to be more hydrophobic and result in the different packing modes for DNA [24]. Water molecules located in crystal structures can be classified as structural

---

*Corresponding author. Tel.: +1-614-292-2925; fax: +1-614-292-2524.
E-mail address: sundaral@chemistry.ohio-state.edu (M. Sundaralingam).

PII: S0301-4622(01)00262-9

water molecules. The structural water molecules are closely related to the structure and are important in stabilizing the crystal structure. They are usually located in the vicinity of the structure and are considered as part of the structure. High-resolution crystal structures can locate water molecules more accurately and the introduction of the *R*-free test makes them more reliable. Therefore, hydration of oligonucleotides can be now illustrated more convincingly as part of the crystal structures.

Base pairing patterns and stacking interactions have been reviewed in recent papers [5,18,19]. We here have discussed hydration of the oligonucleotides, particularly emphasizing the hydration of RNA oligomers.

2. Discussions

Base pairs G•C and A•U (T for DNA) are usually arranged in the Watson–Crick base pairing scheme where maximally three conventional hydrogen bonds are formed for G•C pair: N4–H(C)...O6(G), N1–H(G)… N3(C) and N2–H(G)…O2(C) and maximally two conventional hydrogen bonds and one C–H…O hydrogen bond for the A•U (T) pair: N6–H(A)…O4(U or T) and N3–H(T or U)…N1(A) and C2–H(A)...O2(U/T) (Fig. 1a,b). These conformations are involved in the *cis* hydrogen bonding pattern and is symmetrical to the C1′–C1′ vector of the two bases, with an angle of 43°. Fig. 1 shows the different hydrogen bond donors and acceptors in the major and minor grooves for the G•C and A•U base pairs. The differences are in the minor groove. These grooves are recognized during the interaction with proteins and other ligands.

The Watson–Crick base pairs G•C and A•U can be hydrated on both minor and major grooves. In the 1.38-Å RNA crystal structure r(GUAUGUA)dC [3], the Watson–Crick base pair A•U is hydrated by five water molecules with three in the major groove and two in the minor groove (Fig. 1d). The G•C pair in the 1.4-Å crystal structure of r(UUCGCG) [25] is hydrated by eight water molecules, five in the major groove and three in the minor groove (Fig. 1c). In the

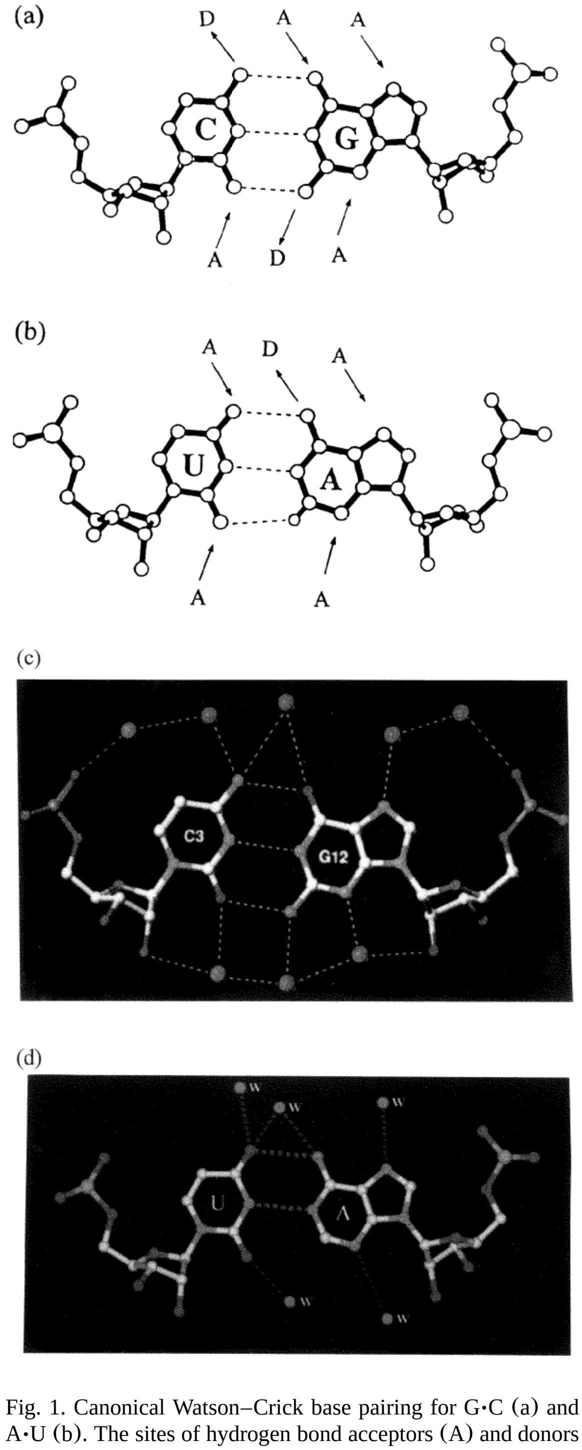

Fig. 1. Canonical Watson–Crick base pairing for G•C (a) and A•U (b). The sites of hydrogen bond acceptors (A) and donors (D) are showed with arrows. Hydration for the Watson–Crick base pairs the G•C in r(UUCGCG) (c), and the A•U in r(GUAUGUA)dC (d).

G•C base pair, there is one more water molecule in the minor groove than in the A•U pair, because of the G-based N2-amino group. In the major groove, the phosphate O1P groups on either sides of the double helix are linked to the base atoms N4(C) and N7(G) with two bridging water molecules. The result shows that the major groove in Watson–Crick G•C and A•T base pairs are more hydrated than the minor groove. It is noticed that the number of the water molecules directly hydrated with base atoms is the same as that of the hydrogen bond donors and acceptors for the base pairs. Similar hydration patterns have been observed for the G•C and A•T base pairs in the high resolution (0.74 Å) B-DNA decamer structure d(CCAGTACTGG) [13]. Because of the lack of 2′-hydroxyl groups in DNA, the water bridges between base atoms do not involve the sugar side.

The mismatched G•U base pairs, the most frequently occurring mismatch in RNA, are often involved in functional interactions, such as in RNA–protein recognition [12,15,16] and the splice site selection in group I introns [2,6,23]. They have been studied by X-ray crystallography in the case of isolated mismatches [14,21], tandem mismatches [3,4,9] and as a variation in a string of four mismatches whose end base pairs are G•U [1,8,11]. In every case, the G•U adopts the wobble conformation [7] where two hydrogen bonds are formed: N3–H(U)...O6(G) (C1′•••C1′ vector makes 57°) and N1–H(G)...O4(U) (C1′•••C1′ vector makes 43°) (Fig. 2a). This is asymmetrical to the C1′–C1′ vector of the two bases. The N2 of G protrudes into the minor groove and the O4 of U protrudes into the major groove. Sometimes the G•U bases can have bifurcated hydrogen bond possibilities.

The isolated G•U wobble pairs in the 2.5-Å crystal structure r(GG*G*CGC*U*CC) [21] are hydrated with two water molecules for the G2•U16 pair and three water molecules for the G11•U7 base pair (Fig. 2b,c). In both base pairs, a water molecule bridges N2(G), O2(U) and O2′(U) was noted in the minor groove side (Fig. 2b,c). In the 1.38-Å tandem, G•U wobble pairs in the crystal structure of r(GUAUGUA)dC [3] in motif I (5′–UG-3′/3′-GU-5′), there are four water molecules hydrating the G13•U4 pair and five water molecules hydrating the G5•U12 pair in the major groove (Fig. 3a,b). The water molecules are bound to the lone pair electrons of O6(G) and O4(U). It appears that the water-mediated hydrogen bonds in both G•U pairs stabilize the RNA duplex and partly compensate for the loss of the third hydrogen bond compared with the G•C base pair. In the 1.58-Å crystal structure r(GUGUGUA)dC, because of strand slippage, four G•U wobbles

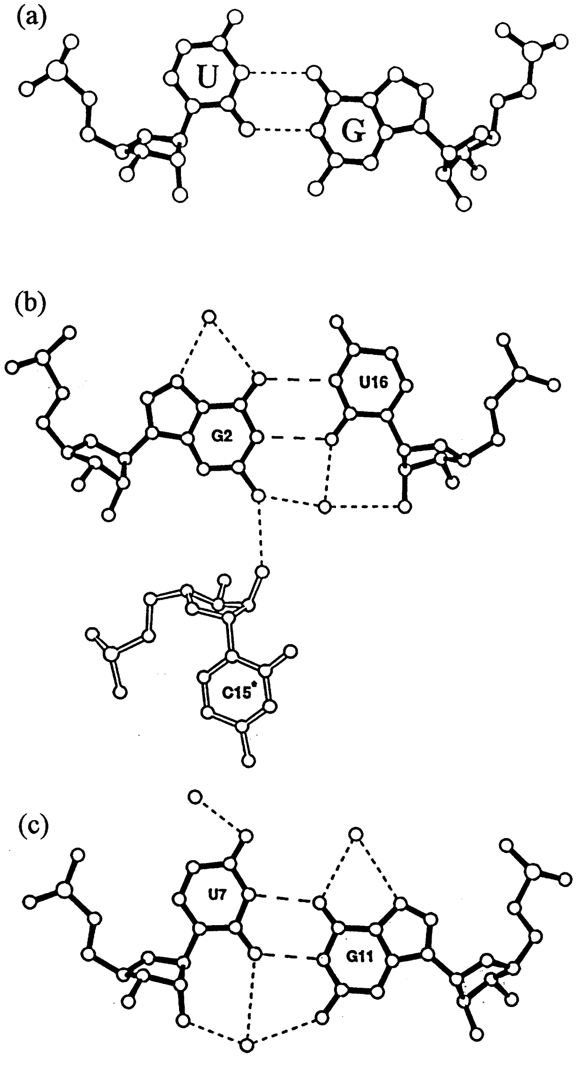

Fig. 2. (a) The wobble base pairing scheme for the G•U mismatch; (b) and (c) show the hydration of the isolated G•U base pairs in r(GGGCGCGCC).

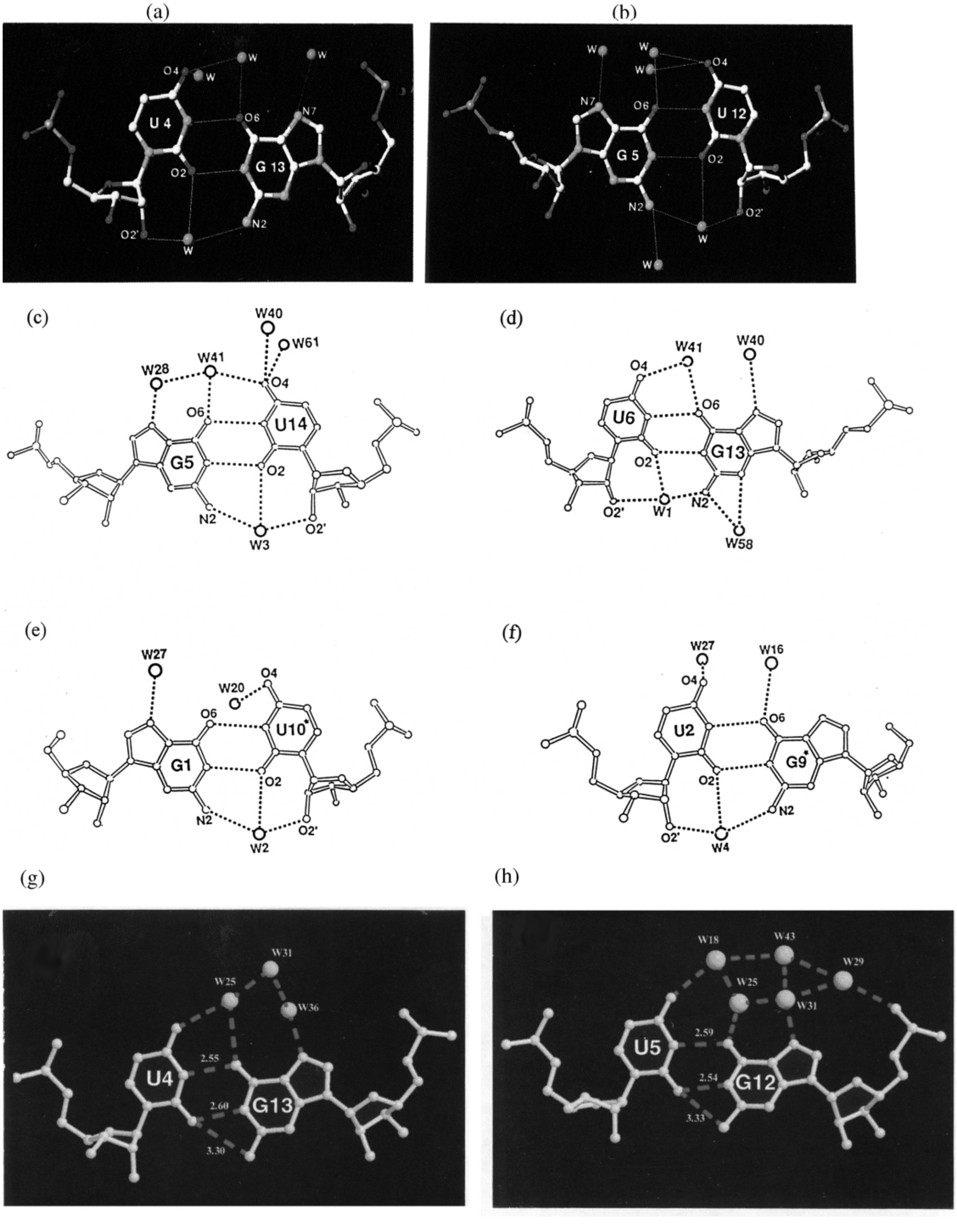

Fig. 3. Hydration of tandem G•U wobbles in crystal structures. Insets (a) and (b) show hydration of the motif I tandem G•U mismatched base pairs in r(GUAUGUA)dC; (c) and (d) show hydration of the motif II tandem G•U mismatched base pairs inside the duplex of r(GUGUGUA)dC while (e) and (f) are for the overhang wobbles in the structure of r(GUGUGUA)dC; (g) and (h) are the hydration for the motif III tandem G•U mismatched base pairs in r(GUAGGCAC)/r(GUGUUUAC).

result, with two, G5•U14 and G13•U6, inside the duplex, and two overhangs, G1•U10 and G9•U2 [4]. These tandem G•U wobbles belong to motif II (5′-GU-3′/3′-UG-5′). G5•U14 and G13•U6 are hydrated with five and four water molecules, respectively (Fig. 3c,d). G1•U10 and G9•U2 are hydrated with three each of water molecules (Fig. 3e,f). The G5•U14 and G13•U6 wobbles inside the duplex have similar hydration patterns in the minor groove, the water bridge in the major groove is broken (W11•••W14) (Fig. 3c,d). The wobbles inside the duplex are more hydrated than the outer overhang wobbles, which may account for the fact that the G•U wobbles inside the duplex are more thermodynamically stable than the overhanging G•U wobbles as illustrated by the B-factor. In the 1.9-Å crystal structure of r(GUG*UU*UAC)/r(GUA*GG*CAC) [9], the tandem G•U wobble pairs are in motif III (5′-GG-3′/3′-UU-5′). A bridge of three water molecules link the O4(U), O6(G) and N7(G) in the G13•U4 wobble pair while a bridge of five water molecules link O4(U), O6(G), N7(G) and O1P(G) in the G12•U5 wobble pair (Fig. 3g,h). It can be noticed that all the water molecules are in the major groove. The slippage of the uridine bases toward the minor groove results in the bifurcated hydrogen bonds of the bases: N1–H(G)...O2(U) and N2–H(G)...O2(U). In all, the G•U wobble pairs, a water molecule is always found connecting N2(G), O2(U) and O2′(U), except in Motif III in which the N2(G) forms hydrogen bonds with O2(U).

The U•U mismatch can adopt the *cis* conformation (Fig. 4a) [1,14] and the almost *trans* conformation (Fig. 4b) [25]. In the first *cis* conformation, it has a shortened C1′–C1′ distance of 9.1 Å, with hydrogen bonds N3–H(U2)–O4(U1) and N3–H(U1)–O2(U2) [1,14] (Fig. 4a). The O4 of one U2 is exposed to the major groove while O2 of U1 is exposed to the minor groove. It is asymmetric about the C1′–C1′ vector with the glycosyl bond angles of 64° and 48°, similar to the G•U wobble conformations. The overhang UU residues of the hexamer r(*UU*CGCG) adopts a different *cis* (almost *trans*) conformation, where the Watson–Crick sites of one of the Us face away from hydrogen bonding and form a novel C5–H(U2)...O4(U1) interaction and a conventional N3–H(U1)...O4(U2) hydrogen bond between the U bases [25] (Fig. 4b).

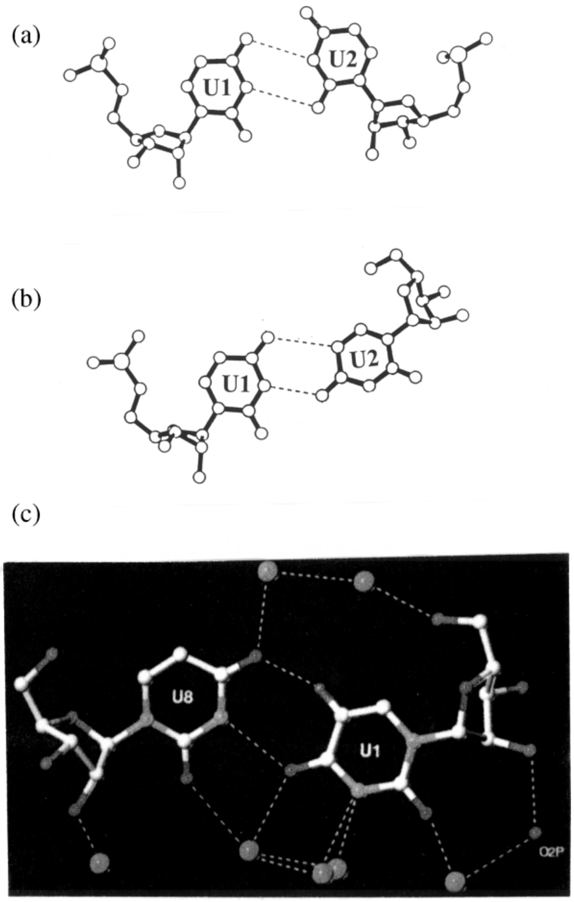

Fig. 4. (a) The *cis* conformation for the U•U mismatch; (b) the *trans* conformation for the U•U mismatch. (c) Hydration of the *trans* U•U base pair in r(UUCGCG).

In the structure, it is clearly seen that the overhang UU bases form the extreme *cis* (almost *trans*) U•U base pairs which is hydrated by six water molecules (Fig. 4c). The water bridging O4(U1) and O2(U8) forms a third pseudo-base pair hydrogen bond which is further bridged to N3(U1) through a disordered water molecule. The disordered water molecule is distributed between two sites, 2.1 Å apart, above and below the base plane with nearly equal proportions. The exocyclic O2 of U1 is connected through a water molecule to the anionic phosphate oxygen O2P(C5) of a

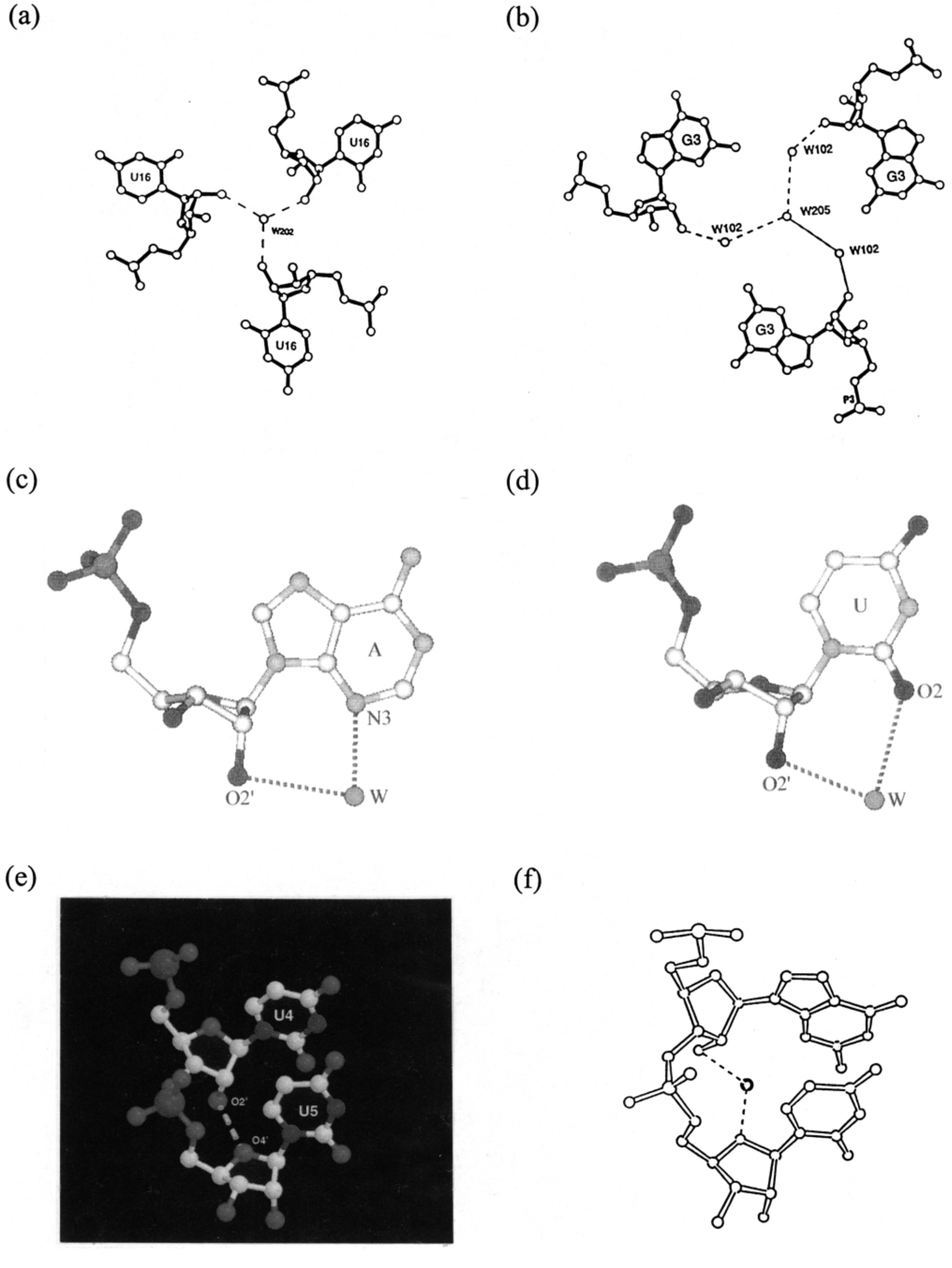

Fig. 5. Hydration of 2′-hydroxyl group. (a) and (b) are hydration of 2′-hydroxyl groups in r(GGGCGCUCC); (c) and (d) are hydration of 2′-hydroxyl group in the minor group. Direct interaction between O2′ and the adjacent O4′ in the 3′ side observed in r(GUAGGCAC)/r(GUGUUUAC) (e) and r(CGUAC)dG (f).

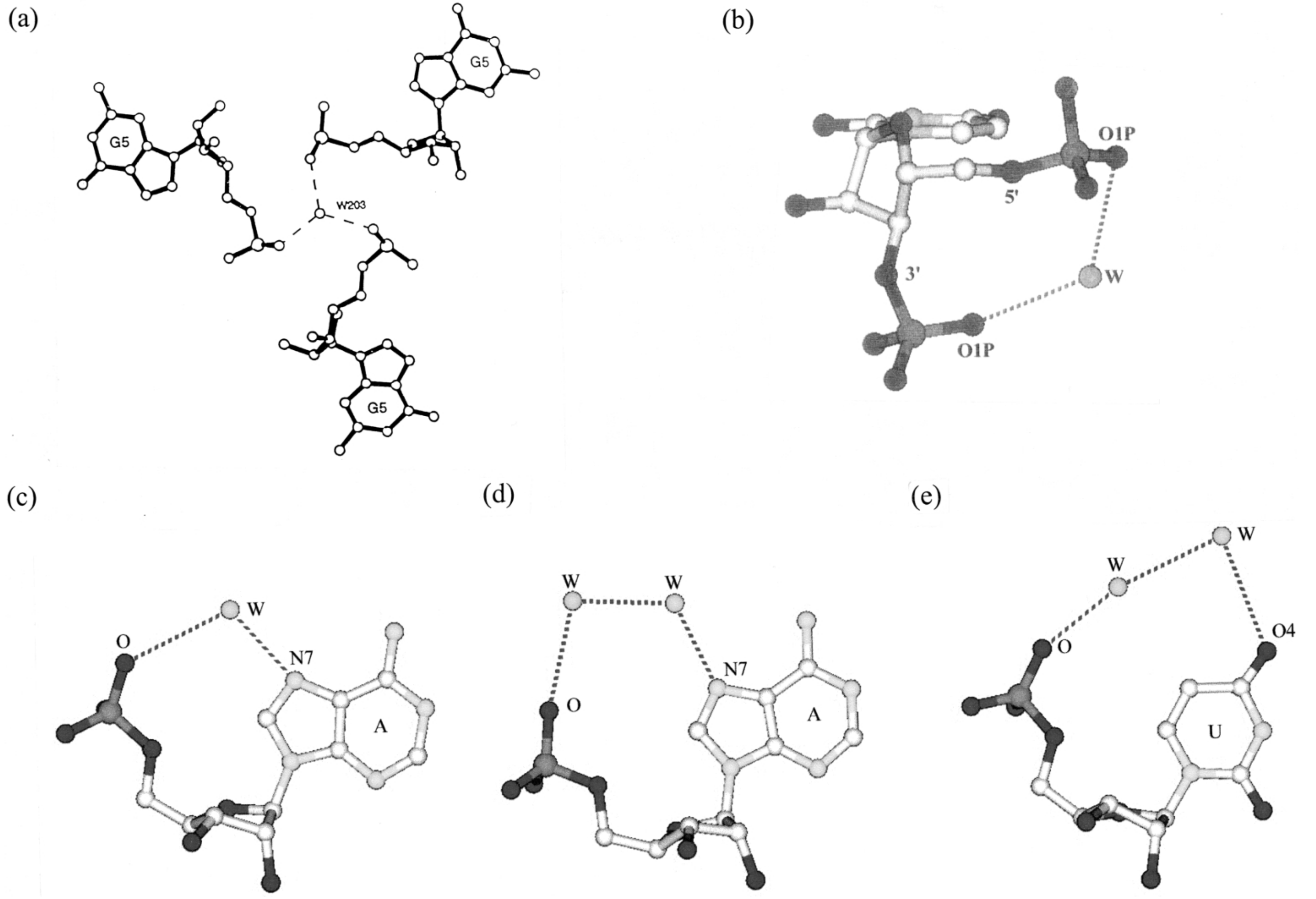

Fig. 6. Hydration of phosphate groups. (a) O2P is hydrated by a water molecule locating in the three axes in r(GGGCGCUCC); (b) water bridge between adjacent phosphate groups in r(GUAUACA)dC; (c), (d) and (e) are major group hydration in r(GUAUACA)dC.

neighboring duplex. This hydration scheme involving the exposed O4, N3 and O2 sites of U1 is crucial for additional stability of the base–base hydrogen bonds.

The 2′-hydroxyl group can be hydrated by a water molecules directly in the three-fold axis or mediated by another water molecule, as observed in the 2.5-Å crystal structure of r(GGGCGCUCC) [21] (Fig. 5a,b). 2′-hydroxyl group can also be hydrated in the minor groove and bridged to N3(A) and O2(U) as in r(GUAUACA)dC [22] (Fig. 5c,d). It can be linked by a water bridge connecting O2′ and O4′ of two adjacent residues on the 3′ side [9] (Fig. 5e) and [3] (Fig. 5f).

Phosphate groups in nucleic acids are usually highly hydrated because of their two anionic oxygen atoms, with O1P in the major groove side and O2P in the minor groove side. In the 2.5-Å crystal structure of r(G*G*GCGC*U*CC), the anionic oxygen O2P of phosphate G5 is hydrated by a water molecule located at the three-fold axis (Fig. 6a). The O1P of one molecule and the adjacent O1P of another can be directly linked by a water molecule because of the short intrastrand phosphate–phosphate distance (Fig. 6b). In addition, O1P can be bridged by one or two water molecules with N7(A) or O4(U) (Fig. 6c–e) as observed in the crystal structure of r(GUAUACA)dC [22].

In the high resolution (1.25 Å) A-DNA crystal structure of d(CCCGATCGGG) (Ramakrishnan, Sekharudu, Pan and Sundaralingam, unpublished results), the phosphate groups O1Ps are systematically hydrated by nine water molecules in the major groove. A bridging water molecule is usually found linked to the adjacent O1Ps. Such hydration pattern is not conserved for the O2Ps in the minor groove. This may be due to the abutting interactions in A-DNA or A-RNA (Fig. 7).

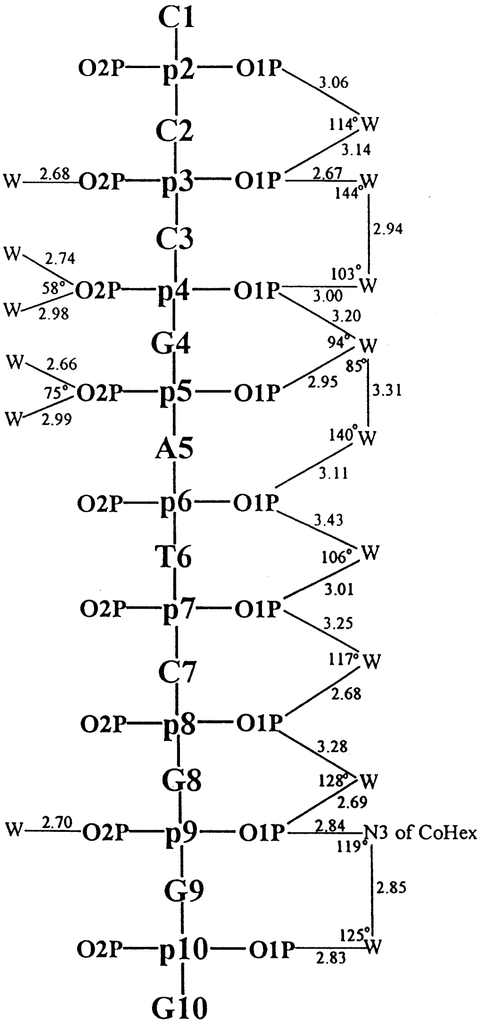

Fig. 7. Schematic diagram of hydration of phosphate groups in d(CCCGATCGGG).

In the high-resolution crystal structures (< 1 Å), we can see hydrogen atoms. We have recently done a high resolution (0.61 Å) tetraplex crystal structure at atomic resolution. Many of the hydrogen atoms in the structure r(cggggc)$_4$ can be visualized in the difference Fourier map ($Fo-Fc$) (Deng, Xiong and Sundaralingam, private communication). A composite drawing of the molecule is shown in Fig. 8 where the heavy atoms (density in blue) and hydrogen atoms (density in red) are indicated. The hydrogen atoms that are involved in the hydrogen bonding interaction can be localized as in cytosine, example N4–H1, the other N4–H2 is not seen; but both the amino N2–H1 and N2–H2 groups are seen in guanosine. It is very important to notice that the 5′-side of cytosine in the sequence 5′-r(cggggc)$_4$-3′ is protonated at N3 which is hydrogen-bonded to the oxygen atom in the phosphate group. We have visualized for the first time in an oligonucleotide crystal structure that an N3 atom of cytosine is protonated.

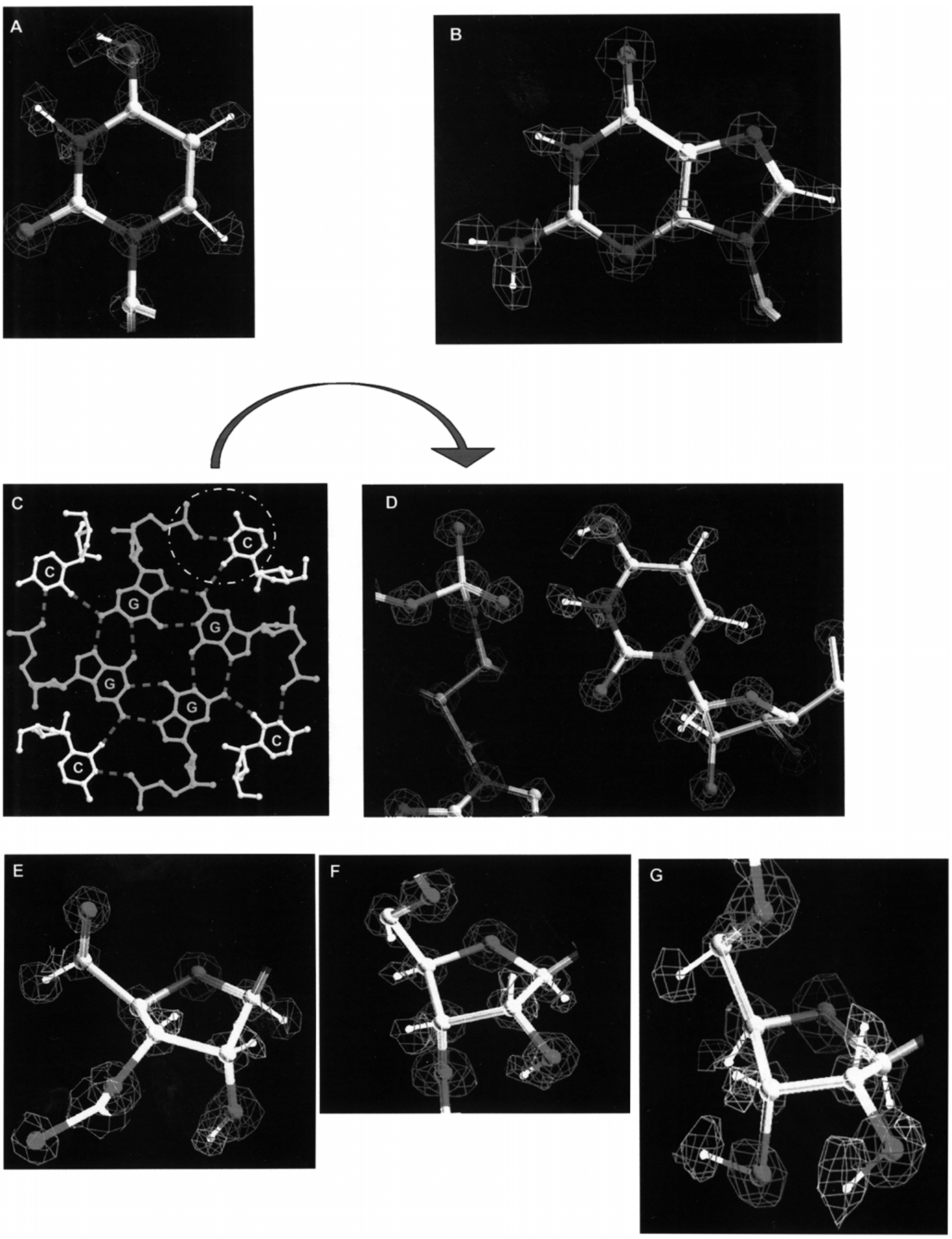

Fig. 8. Crystal structure in the tetraplex $r(cggggc)_4$ with hydrogen atoms. (A) Cytosine base with electron density map of ($2Fo-Fc$) (blue) for heavy atoms and difference Fourier map ($Fo-Fc$) (red) for hydrogen atoms. Hydrogen atoms are illustrated with small white balls. (B) Guanine base with hydrogen atoms illustrated. (C) The geometry of the octaplet in the structure of $r(cggggc)_4$. (D) The protonated cytosine and the phosphate group that are involved in hydrogen bonding interaction are shown with density maps. The difference Fourier map clearly shows that the N3 of cytosine is protonated. The visualized hydrogen atoms in the ribose of 5′-side cytosine (E), the residue inside the tetraplex (F) and 3′-side cytosine (G) are shown with hydrogen atoms in white balls and their difference Fourier maps in red.

In summary, the hydration of oligonucleotide double helices is an important interaction between the nucleic acids and the surroundings and it can stabilize the structures of oligonucleotide double helices. Hydration can stabilize the conformation of the double helices by bridging the phosphate groups and the base atoms in the major groove and hydrogen-bonding to the adjacent phosphate groups. Hydration can also link the 2′-hydroxyl group and the base atoms in the minor groove and stabilize the conformation. The conformation of mismatched base pairs and protein–DNA/RNA complexes are stabilized by hydrogen bonding interactions between water molecules and proteins, the atoms of the base that protrude into the major or the minor grooves. The different hydration for the Watson–Crick base pairs G•C and A•U/T results from their different patterns of hydrogen bond donors and acceptors. Thus, the two complementary base pairs are even more stabilized by hydration. The different hydration for the different motifs for the tandem wobble G•U mismatches may reflect that hydration plays a role in their stability. The hydration (Watson–Crick base pairs, wobble and mismatches) may be influenced by nucleic acid sequences. These results implicate that hydration of nucleic acids must be important in the genetic code recognition in codon–anticodon interactions.

Acknowledgments

We gratefully acknowledge support of this work by the grant GM-17378 from the National Institutes of Health and conferring an Ohio Eminent Scholar Award by the Board of Regents of the Ohio State University.

References

[1] K.J. Baeyens, H.L. DeBont, S.R. Holbrook, Natl. Struct. Biol. 2 (1995) 56.
[2] E.T. Barford, T.R. Cech, Mol. Cell. Biol. 9 (1989) 3657–3666.
[3] R. Biswas, M.C. Wahl, C. Ban, M. Sundaralingam, J. Mol. Biol. 267 (1997) 1149–1156.
[4] R. Biswas, M. Sundaralingam, J. Mol. Biol. 270 (1997) 511–519.
[5] T. Brown, W.N. Hunter, Biopolymers 44 (1997) 91–124.
[6] T.R. Cech, Sciences 236 (1987) 1532–1539.
[7] F.H. Crick, J. Mol. Biol. 19 (1966) 548.
[8] W.B.T. Cruse, P. Saludjian, E. Biala, P. Strazewski, T. Prange, O. Kennard, Proc. Natl. Acad. Sci. USA 91 (1994) 4160.
[9] J.P. Deng, M. Sundaralingam, Nucleic Acids Res. 28 (2000) 4376–4381.
[10] R.E. Franklin, R.G. Gosling, Nature 171 (1953) 740–741.
[11] S.R. Holbrook, C. Cheong, I. Tinoco, S.-H. Kim, Nature 353 (1991) 579.
[12] Y.M. Hou, P. Schimmel, Nature 333 (1988) 140–145.
[13] C.L. Kielkopf, S. Ding, P. Kuhn, D.C. Rees, J. Mol. Biol. 296 (2000) 787–801.
[14] S.E. Lietzke, C.L. Barnes, J.A. Berglund, C.E. Kundrot, Structure 4 (1996) 917.
[15] W.H. McClain, K. Floss, Sciences 240 (1988) 793–796.
[16] W.H. McClain, K. Floss, J. Mol. Biol. 202 (1988) 697–709.
[17] S.N. Mitra, R. Biswas, K. Shi, M. Sundaralingam, J. Biomol. Struct. Dyn. Conversation 11 (1999) 189–194.
[18] B. Pan, M. Sundaralingam, J. Quantum Chem. 75 (1999) 275–287.
[19] B. Pan, M. Sundaralingam, in: M. Vijayan, N. Yathindra, A.S. Kolaskar (Eds.), Perspectives in Structural Biology, Indian Academy of Sciences, Bangalore, 1999.
[20] W. Saenger, Principle of Nucleic Acid Structure, Spring-Verlag, New York, 1984.
[21] K. Shi, M. Wahl, M. Sundaralingam, Nucleic Acids Res. 27 (1999) 2196–2201.
[22] K. Shi, R. Biswas, S.N. Mitra, M. Sundaralingam, J. Mol. Biol. 299 (2000) 113–122.
[23] A.A. Strobel, T.R. Cech, Sciences 267 (1995) 675–679.
[24] M. Sundaralingam, R. Biswas, J. Biomol. Struct. Dyn. 15 (1997) 173–176.
[25] M.C. Wahl, S.T. Rao, M. Sundaralingam, Natl. Struct. Biol. 3 (1996) 24–31.