The hydration and unusual hydrogen bonding in the crystal structure of an RNA duplex containing alternating CG base pairs†‡

Dorota A. Adamiak,[a] Jan Milecki,[b] Ryszard W. Adamiak[a] and Wojciech Rypniewski*[a]

Received (in Montpellier, France) 27th October 2009, Accepted 10th December 2009
First published as an Advance Article on the web 2nd February 2010
DOI: 10.1039/b9nj00601j

The X-ray structures of two crystal forms of the RNA duplex $[CGCG(5\text{-}FC)G]_2$ are the first crystal structures for an RNA duplex containing alternating CG base pairs with unmodified 2′-hydroxyl groups. The triclinic and the rhombohedral crystal forms (PDB ids 3jxq, 3jxr) differ in the molecular packing, the amount of ordered structure and in molecular details, especially the hydration. The P1 data extend to 1.45 Å resolution, while the R32 data reach 1.25 Å. In both structures the RNA oligomers take the form of a right-handed half-turn with conformational parameters close to the canonical A-RNA. The duplexes stack coaxially in the crystal lattice. The geometry of C:G *vs.* 5-FC:G pairing is similar in terms of H-bond geometry. In CG steps, inter-strand guanines are parallel while cytosines are not parallel. In steps GC this motif is reversed. Both structures contain ordered water molecules, forming characteristic hydration network in the major and the minor grooves. In addition to water, the triclinic structure contains four magnesium cations, one of them forming an inner complex with a phosphate. Two types of unusual interactions are observed. In the major groove, the 5-fluorocytidine residues form C–F···H–O–H hydrogen bonds with water molecules. In the minor groove, intermolecular contacts include C–H hydrogen bonds (C1′–H1···O2′ and C4′–H4′···O4′) between two ribose rings.

Introduction

Double-stranded RNA helices exist principally in the right-handed A-form. RNA duplexes containing alternating pyrimidine-purine steps *e.g.* poly(CG)[1], $(CGCGCG)_2$[2] or $(CGCG)_2$[3] can be induced to undergo helicity reversal with the formation of *Z*-RNA, which distinguishes them from other RNA duplexes. The process is promoted not only by high salt, such as 6 M $NaClO_4$, but also by proteins containing Zα domain *e.g.* ADAR1.[4] The mechanism of salt-induced A to Z reversal has not been fully understood and in order to resolve this question one needs a detailed knowledge about the structure and hydration of both right- and left-handed RNA helices.

In view of this we have earlier analysed, using NMR, the right-handed structures of $(CGCGCG)_2$ and 2′-O–Me$(CGCGCG)_2$ in low-salt solution (150 mM NaCl).[5] The structure of the $(CGCGCG)_2$ appeared to be overwound, with an average of 9.7 bp per turn, and a unique stacking pattern and helical parameters, thus showing a geometry considerably deviating from a canonical A-RNA. At that time we were unable to crystallise $(CGCGCG)_2$ but we determined two crystal structures of 2′-*O*-Me$(CGCGCG)_2$. In both cases, the structures were similar to parent $(CGCGCG)_2$ determined in solution and revealed a characteristic hydration pattern of a single row of water molecules in the minor groove, directly stabilising unique base stacking interactions.[6a,b] More recently, we reported a high resolution NMR structure of the $(CGCGCG)_2$ in the left-handed form induced by high salt (6.0 M $NaClO_4$).[7] Unfortunately, the NMR method used cannot reveal the hydration pattern. Until now, the hydration pattern of a *Z*-RNA is available for the low salt X-ray structure of the left-handed $dUr(CG)_3$ helix, complexed with the Zα domain of the ADAR1 enzyme.[8]

To date there is no X-ray structure of right-handed RNA duplex containing alternating CG base pairs and bearing unmodified 2′-hydroxyl function, which would allow an inspection of the hydration of both CG and GC steps. Here, we present the X-ray structures of two crystal forms of $[CGCG(5\text{-}FC)G]_2$. The duplex containing 5-fluorocytidine residue (abbreviated 5-FC) was obtained for ^{19}F NMR study and, in contrast to $(CGCGCG)_2$, can be crystallised.

Results

Molecular packing and the refined models

In the triclinic structure the asymmetric unit (the unit cell) contains two RNA duplexes stacked end-to-end (chains A + B and K + L). This arrangement is repeated in the crystal lattice to form parallel semi-continuous columns (Fig. 1a). In the

[a] *Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland. E-mail: wojtekr@ibch.poznan.pl*
[b] *Faculty of Chemistry, Adam Mickiewicz University, Poznań, Poland*
† Dedication: This paper is dedicated to Professor Wojciech J. Stec on the occasion of his 70th birthday anniversary. This article is part of a themed issue on Biophosphates.
‡ Electronic supplementary information (ESI) available: Supplementary tables. RSCB PDB codes: 3jxq, 3jxr. For ESI see DOI: 10.1039/b9nj00601j

rhombohedral structure the asymmetric unit contains one complete duplex (chains A + B) and one strand of a second duplex. The other strand is symmetry-related to the first by a crystallographic two-fold axis located between the strands (chains K + K*, where * denotes a symmetry-related chain). The duplexes in the rhombohedral structure also stack end-to-end and form semi-infinite columns in the crystal lattice (Fig. 1b). The duplex A + B is tightly packed against its symmetry mates and well ordered, whereas the duplex K + K* is surrounded largely by solvent space and poorly ordered. This accounts for the poorer R-factor statistics of the rhombohedral structure, despite its nominally higher resolution, compared with the triclinic structure (Table 1). The triclinic structure also contains a relatively larger amount of ordered solvent molecules, which surround evenly both the RNA duplexes. In addition to water molecules, four magnesium cations have been identified in the triclinic structure, based on the regularity and distances within their coordination spheres. The positions of the fluorine atoms modifying the C5 cytidine residues were clearly visible in electron density maps in both structures.

The overall structure and conformational parameters

In both crystal forms, self-complementary hexamers CGCG(5-FC)G form approximately one half-turns of an RNA helix. The helices belong to the A-type RNA.[9]

Sugar, glycosidic bond and backbone conformation. In both duplexes, most of the conformational parameters describing the phosphor-sugar backbone are typical of the canonical A-helix.[9] The sugars are in stabilised C3′-*endo* pucker with the average P value 12°, except for G2 in chain A of the rhombohedral structure, which has been modelled in the C2′-*exo* conformation ($P = -1°$). All residues are characterised by high pucker amplitudes (average $\Phi = 44°$) and by *anti* glycosidic bond angles. The latter and the α, β, γ, δ, ε, ζ backbone torsion angles (Supplementary Table 1‡) fall in the range typical for the A-family of right-handed helices.[9] The γ torsion angle ranges from 38° to 90° describing their (+) *gauche* (+*sc*) conformations, except G4, chain L of the triclinic structure, which has $\gamma = 177°$ (−) *gauche* (*ap*). The β positive values (159–179°) and ε (−135 to −162°) torsion angles indicate a favoured *trans* conformation of corresponding bonds. In two instances β takes a negative value: for G4, chain L triclinic, which is correlated with the γ angle mentioned above.

Base pairs and stacking geometry. Watson–Crick base pairing is observed throughout the duplex in both crystal forms. The geometry of C:G *vs.* 5-FC:G pairing is similar in terms of H-bond lengths (in the range 2.8–2.9 Å). Every C:G or 5-FC:G pair has a characteristic negative propeller twist ranging −8 to −18°. Analysis of helical parameters reveals that they fall in the range typical of A-RNA, differing considerably from the previously studied 2′-*O*-methylated analogues[6a,b] (Supplementary Table 2‡). For example, the rise parameter is on average 2.7 Å, higher by 0.3 Å than the methylated structures but similar to $(CGCGCG)_2$ studied in solution.[6a] The $[CGCG(5F\text{-}C)G]_2$ structure is less tightly

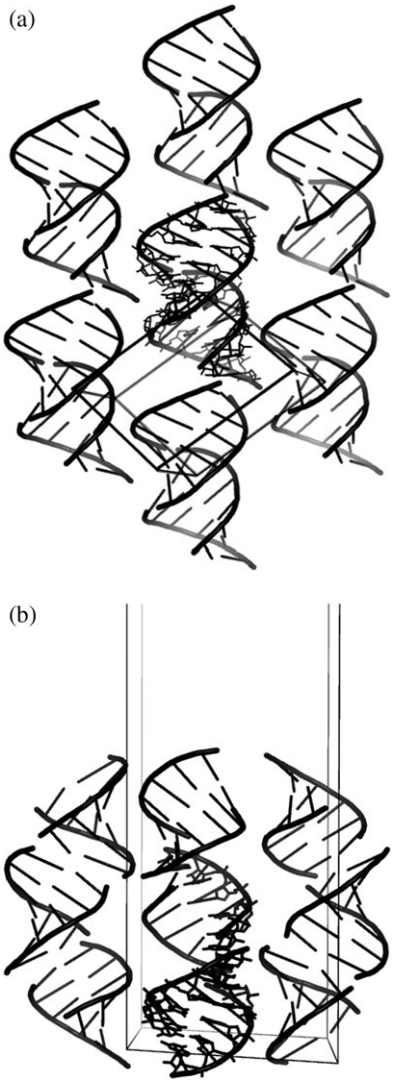

Fig. 1 Crystal packing of the $[CGCG(5\text{-}FC)G]_2$ duplexes in (a) the triclinic and (b) the rhombohedral crystal lattices. The content of the asymmetric units are drawn using the "sticks" model. The symmetry-related molecules are shown only in the cartoon mode.

wound (from 10.4 to 11.3 bp/turn) than the previously reported $(CGCGCG)_2$ in solution (9.7 bp/turn)[5] or 2′-*O*-Me$(CGCGCG)_2$ RNA crystal structures (9.9 and 10.2 bp/turn).[6a,b] A characteristic stacking pattern was observed for both the crystal structures. Within CG steps, inter-strand guanines are parallel while cytosines are not parallel. In GC steps this relationship is reversed. The arrangement is similar to the 2′-*O*-Me$(CGCGCG)_2$ structures[6a,b] and has not been observed in other RNA crystal structures. In GC steps considerable inter-strand stacking of guanines is observed (Fig. 2), which is a typical feature of A-RNA.[9]

Ions and hydration

Magnesium binding. Four ordered magnesium sites were located in the triclinic crystal lattice and none in the rhombohedral structure. Two Mg^{2+} ions are in the major groove, where they interact with ordered waters molecules associated with O6 and N7 sites of guanines. Two others stabilise close

Table 1 Summary of the X-ray data and model refinement for $[CGCG(5\text{-}FC)G]_2$

Crystal	Triclinic	Rhombohedral
Beam line	EMBL-BW7A	EMBL-X13
Wavelength/Å	1.100	0.802
Space group	P1	R32:H
Cell parameters	$a = 21.2$, $b = 26.5$, $c = 29.3$ Å, $\alpha = 97.6$, $\beta = 105.5$, $\gamma = 109.4°$	$a = b = 41.1$, $c = 143.3$ Å
Resolution range/Å	20.0–1.45 (1.47–1.45)[a]	25.0–1.25 (1.27–1.25)
Mosaicity (°)	1.2	0.4
Exposure time per image (min.)	1	1
R_{merge}[b]	0.041 (0.297)	0.043 (0.484)
$\langle I/\sigma(I)\rangle$	26 (2.7)	26 (4.2)
Completeness (%)	93.7 (80.0)	99.2 (99.8)
No. unique reflections	9441	13173
Overall multiplicity	3.1 (2.2)	8.6 (7.7)
Reflections $>3\sigma$ (%)	79 (45)	83 (58)
B-factor from Wilson plot (Å^2)	14	17
R_{work}	0.172	0.219
R_{free}[c]	0.205	0.276
Number of RNA atoms	510	381
Number of metal ions	4 Mg^{2+}	0
Number of water molecules	132	60
Solvent content (%)	40	45
r.m.s. deviation from ideal values		
Bond lengths/Å	0.011	0.018
Bond angles (°)	2.02	2.35
PDB code	3jxq	3jxr

[a] Values in brackets are for the highest resolution shell. [b] $R_{merge} = \Sigma_{hkl}\,\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle|/\Sigma_{hkl}\,\Sigma_i\,I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed individual and mean intensities of a reflection with indices hkl, respectively, Σ_i is the sum over the individual measurements of a reflection with indices hkl and Σ_{hkl} is the sum over all reflections. [c] R_{free} was calculated using 5% of the total reflections chosen randomly and omitted from the refinement.

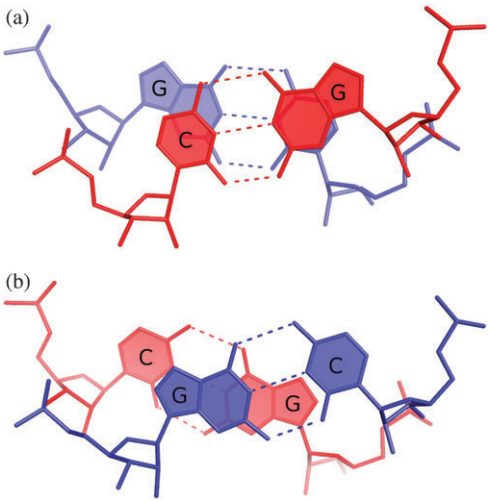

Fig. 2 Stacking pattern within CG (a) and GC steps (b) of the $[CGCG(5\text{-}FC)G]_2$ duplex.

contacts of phosphates from symmetry-related duplexes. One of the magnesium cations (Fig. 3) forms an inner complex with a phosphate oxygen atom (distance = 2.2 Å).

Major groove hydration. The major groove of $[CGCG(5\text{-}FC)G]_2$ is narrow but deep and the minor groove is broad and shallow. The grooves are hydrated in a regular way, with the majority of ordered waters located in the major groove. The patterns of hydration in the structures can be compared to the related crystal structure of $(CCCCGGGG)_2$; the first study in which the RNA duplex hydration scheme has been presented in detail.[10]

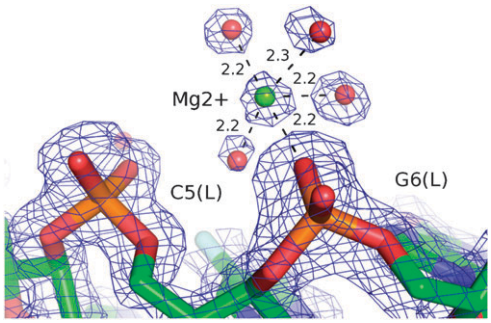

Fig. 3 Magnesium cation forming an inner complex with a phosphate group in the triclinic $[CGCG(5\text{-}FC)G]_2$ crystal structure.

The water molecules in the major groove form a regular pattern, similar to that observed in the 2′-*O*-Me$(CGCGCG)_2$ structures.[6a,b] As in other A-RNA crystal structures,[12] most of the intra-strand phosphate oxygens (O1P) are water-bridged with W-O1P distances ranging from 2.8 to 3.3 Å. Those bridges form one edge of a water network H-bonded (average distance 2.7–2.9 Å) with N^7 and O^6 of guanines as primary base sites. The N^4–H of cytosines are less strongly involved with an average distance 3.1 Å (Fig. 4a).

The C–F bonds in the 5-modified cytosines also participate in the scheme of the major groove hydration (Fig. 4b). Water oxygen atoms come as close as 2.9 Å to the fluorine atoms, indicating significant H-bonding interactions (C–F···H–O–H). The electronegative character of the fluorine atom is clearly seen in the $2F_o–F_c$ map, as a dislocation of electron density from the fluorine towards the carbon atom.

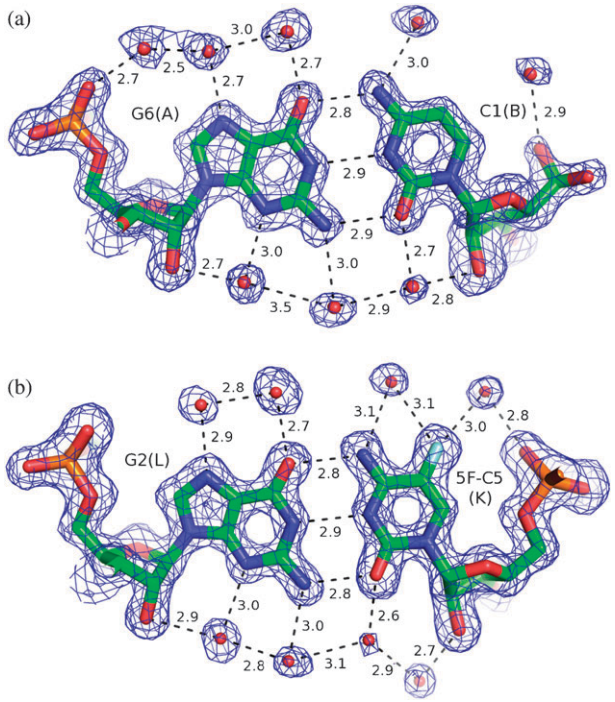

Fig. 4 Hydration of C:G pairs within the $[CGCG(5\text{-}FC)G]_2$ duplex from the triclinic crystal structure: (a) G6 (chain A):C1 (B) and (b) G2 (K): (5-FC)5 (chain L). The fluorine atom of the 5-fluorocytosine is shown in turquoise. The blue contours represent the $2F_o–F_c$ electron density map drawn at the 1σ level. Hydrogen bonds are indicated with dashed lines.

Minor groove hydration. The minor groove contains a dense network of water molecules (Fig. 5). Three types of interactions can be distinguished. The main features of hydration are pairs of water molecules that span the 2′-hydroxyl groups of ribose moieties from opposite strands. This has been recognised as a typical feature of A-RNA.[10–12] Additionally, the same waters are often involved in H-bonding of the base of the same nucleotide: O2 of cytosine and N3 of guanine. The third type of interactions involves bridging, by single water molecules, of carbonyl oxygen atoms from inter-strand parallel cytosines. Only one example has been observed of an analogous bridge between the N3 atoms of guanine residues.

A characteristic feature in both crystal structures is the presence in the minor groove of intrusions from neighbouring molecules, in contrast to the major groove, where only water and magnesium cations are observed. The contacts involve sugar–sugar C1′–H1′···O2′ and C4′–H4′···O4′ interactions. Two pairs of such interacting ribose rings are observed, involving two 5-fluorocytidines and two guanosines. The distance between the C and O atoms is 3.3–3.6 Å and their relative geometry indicates C–H hydrogen bonds (Fig. 6). Among other intruding groups are 2′-hydroxyls and phosphodiesters from neighbouring molecules (Fig. 5).

Discussion

Although the triclinic and the rhombohedral crystals differ in molecular packing, the overall structure of the RNA duplexes take similar forms of right-handed half-turns with conformational parameters close to the canonical A-RNA (Supplementary Tables 1 and 2‡), with end-to-end stacking of the molecules. The one instance of an unusual torsion angle of the sugar-phosphate backbone ($\gamma = 177^\circ$ for G4, chain L, Supplementary Table 1‡) is difficult to explain but clear in the electron density. The presence of the 5-fluorocytidine residues does not alter the typical C:G base pair geometry. This seems to be significant for RNA studies in solution, by ^{19}F NMR spectroscopy, using 5-fluorocytidine-labelled oligonucleotides.[13,14] Interestingly, the fluorine atoms of the 5-fluorocytidine residues form hydrogen bonds with water molecules, introducing additional elements to the hydration pattern, unavailable to unmodified cytidines (Fig. 4). To our knowledge, this is the first experimentally observed $C(sp^2)$-F···H–O bond between a modified nucleobase and solvent water. Hydration of C–F bonds was previously observed for DNA duplexes containing 2′-deoxyribo-2,4-difluorotoluene nucleoside analogue, an isostere of 2′-deoxythymidine.[15] The role of fluorine as a hydrogen bond acceptor in organic compounds has been examined, based on the available small-molecule crystallographic data, NMR and theoretical calculations.[16–18] Those studies indicated that such interactions occur rarely, having lower energies than classical H-bonds.

Three of the four ordered magnesium cations interact with water molecules without disrupting the hydration pattern of the RNA. The two Mg^{2+} ions in the major groove are associated with the water molecules that form the spine hydrating the Hoogsteen edge of guanine residues. The third magnesium cation is lodged between neighbouring duplexes, it is fully hydrated and spans three different phosphate groups *via* the waters from its coordination sphere. The fourth magnesium cation exceptionally sheds a water molecule from its hydration sphere and forms an inner complex with a phosphate oxygen atom.

In our previous studies we examined the structure of 2′-*O*-Me$(CGCGCG)_2$, especially its minor groove hydration.[6a,b] The present structure complements the study by showing the hydration of unmodified RNA duplex containing alternating CG base pairs. When the 2′-hydroxyl groups are methylated, the minor groove contains only a single row of ordered water molecules, linking and positioning pairs of nucleobases along the strand. This demonstrates how hydration supports the distinctive cross-strand parallel arrangement of the bases. When the 2′-hydroxyl groups are free, the minor groove is filled with ordered water molecules which, in addition to stabilising the relative positions of the bases, are engaged in spanning the two strands and stabilising the conformation within the nucleotide residues.

The question of the hydration scheme of right-handed RNA duplexes with $(CG)_n$ sequence is relevant in connected with the yet unknown mechanism of helicity reversal. The reversal can be induced by proteins containing Zα domain *e.g.* ADAR1, under low-salt conditions.[4] In order to promote this transformation in the absence of the protein, salt concentrations as high as 6 M $NaClO_4$ are necessary.[7] The interaction of salts with the polyanionic structure of nucleic acids and their water shells is the driving force in the helicity reversal of right-handed RNA or DNA $(CG)_n$ tracts. Salts at high concentration screen the repulsion between electronegative

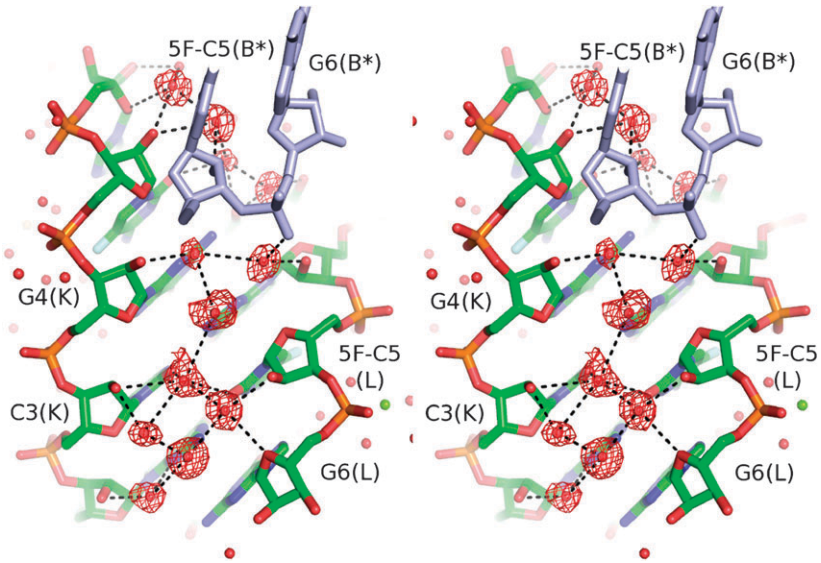

Fig. 5 A stereo view of a part of the hydration network of the minor groove in the triclinic crystal structure of the $[CGCG(5\text{-}FC)G]_2$ duplex. The red contours represent the F_o–F_c "omit" map drawn at the 3σ level. Hydrogen bonds are indicated with dashed lines. A fragment of an intruding, neighbouring duplex is shown in blue.

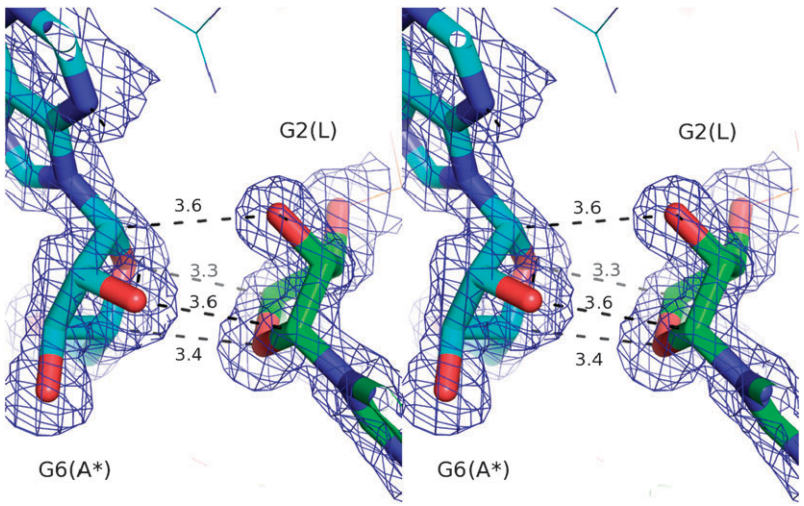

Fig. 6 A stereo view of the sugar–sugar interaction between neighbouring duplexes in the triclinic crystal form, involving C–H···O hydrogen bonds: a pair of C1′–H1···O2′ and a pair of C4′–H4′···O4′. The blue contours represent the $2F_o$–F_c electron density map drawn at the 1σ level. Hydrogen bonds are indicated with dashed lines.

phosphates in the left-handed form[19] and interacts with the water shell around RNA leading to its dehydration.[6b,20] In contrast to the DNA field, were various mechanisms have been proposed to describe B ⇌ Z DNA transition, there is no model of the RNA helicity reversal. Therefore, in order to evaluate the mechanism of A ⇌ Z transition under high salt conditions, information is necessary not only about atomic coordinates but also their hydration schemes on the atomic level, under different salt concentrations. This paper provides such data for the right-handed $(CG)_n$ RNA duplex, thus promoting further studies on the RNA helicity reversal mechanism.

The clarity of the hydration pattern in both grooves suggests that the bound water molecules play a structural role and can be considered part of the RNA structure. Nevertheless, the water molecules are readily displaced when macromolecules interact or associate with small ligands. The water molecules can be displaced by groups with similar H-bonding potential, *e.g.* ligand's hydroxyl groups, which preserves the basic H-bonding pattern.[6b,21] The present study also provides examples of interruption of the solvent's regular H-bonding scheme by neighbouring RNA chains (Fig. 5). The intrusions are stabilised by unusual C–H hydrogen bonds between pairs of interdigitated sugar moieties (Fig. 6). Although hydrogen bonding interactions involving C–H bonds are rarely observed (including the sugar ring oxygen atom[22]) and rather weak, the interface between the sugar rings features four of them. Therefore, their combined effect can be significant in stabilising the interaction between two RNA duplexes.

Experimental

Crystallization and data processing

The self-complementary hexamer CGCG(5-FC)G was prepared by an automated solid-phase synthesis using phosphoramidite chemistry, deprotected under standard conditions[23] and purified by Sephadex G10 filtration. The duplex was crystallised at 20 °C by the hanging drop/vapour diffusion method. Crystals of the triclinic form were prepared under the following conditions: The reservoir initially contained 0.5 ml of solution containing 2.0 M lithium sulfate, 10 mM magnesium chloride and 50 mM MES at pH 5.6. The crystallization drop was prepared by mixing 3 μl of RNA at 5 mg ml^{-1} and 3 μl of the reservoir solution. After several days, when no crystals appeared, 0.5 ml of MPD was added to the reservoir, giving the final MPD concentration 50% v/v. Crystals of the rhombohedral form were obtained under the following conditions: The reservoir initially contained 0.5 ml of solution containing 1.6 M ammonium sulfate, 10 mM magnesium chloride and 50 mM HEPES at pH 7.0. The crystallisation drop initially consisted of 3 μl of 5 mg ml^{-1} RNA solution and 3 μl of the reservoir solution. After several days, when no crystals appeared, 0.1 ml of MPD was added twice to the reservoir, at several days' interval, giving the final MPD concentration 28% v/v.

X-Ray diffraction data were obtained on the EMBL BW7A and X13 beam line at the DORIS storage ring, DESY, Hamburg. Prior to mounting, the crystals were transferred to cryoprotectant solutions similar to the reservoir solution but containing in addition 20% v/v glycerol. The crystals remained in the cryoprotectant solution for approximately 1 min, *i.e.* only the time that was needed to pick them again in the cryo-loop. The crystals were transferred in cryo-loops directly to the goniostat and vitrified in the stream of cold nitrogen gas. The diffraction images were recorded on a MAR CCD 165 mm detector. The diffraction intensities were integrated and scaled using the program suite Denzo/Scalepack.[24] The crystals and X-ray data are summarised in Table 1.

Structure solution and crystallographic refinement

Both crystal structures were solved by molecular replacement (MR) using the program Molrep.[25] The rhombohedral structure was solved first using as the search model the atomic coordinates of the 2′-*O*-Me$(CGCGCG)_2$ RNA (pdb id 1i7j).[6b] A solution obtained for one duplex in the asymmetric unit gave an R-factor of 0.55 and the correlation coefficient of 0.41, which was not outstanding, but was selected nevertheless because it gave reasonable base-pair interactions between symmetry-related duplexes in the crystal lattice. Electron density maps, $2F_o–F_c$ and $F_o–F_c$, calculated using phases derived from the initial MR model were initially poor but sufficient for refining and extending the model until it was complete. The triclinic structure was solved using as the search model the atomic coordinates of RNA from the rhombohedral structure. Although the asymmetric unit contained two RNA duplexes, a MR solution was obtained only for one duplex. The second duplex was built gradually, based on electron density maps, during the refinement. The atomic models were refined using the program Refmac5[26] from the CCP4 program suite.[27] Towards the end of the refinement of the rhombohedral structure, anisotropic atomic temperature factors were refined. It resulted in the decrease of both the R-factor and the R_{free} by approximately 0.025. Solvent molecules were inserted using the program ARP/wARP[28] with automatic determination of statistically significant density levels for inclusion of new water molecules. The program Coot was used for visualization of electron density maps and manual rebuilding of the atomic model.[29] Helical parameters were calculated using the 3DNA software.[30] All pictures were drawn in PyMOL v0.99rc6.[31]

Acknowledgements

This work was supported by the Ministry of Science and Higher Education (Poland, grant 7T09A 097 20 to RWA) and by the European Community Research Infrastructure Action under the F6P "Structuring the European Research Area Programme" (contract number RII3/CT/2004/5060008). The authors are grateful to M. Kluge for assistance with crystallization.

References

1 K. Hall, P. Cruz, I. Tinoco, Jr., T. M. Jovin and J. H. Van de Sande, *Nature*, 1984, **311**, 584–586.
2 R. W. Adamiak, A. Gałat and B. Skalski, *Biochim. Biophys. Acta, Gene Struct. Expression*, 1985, **825**, 345–352.
3 P. W. Davis, K. Hall, P. Cruz, I. Tinoco, Jr. and T. Neilson, *Nucleic Acids Res.*, 1986, **14**, 1279–1291.
4 B. A. Brown, 2nd, K. Lowenhaupt, C. M. Wilbert, E. B. Hanlon and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 13532–13536.
5 M. Popenda, E. Biala, J. Milecki and R. W. Adamiak, *Nucleic Acids Res.*, 1997, **25**, 4589–4598.
6 (*a*) D. A. Adamiak, J. Milecki, M. Popenda, R. W. Adamiak, Z. Dauter and W. Rypniewski, *Nucleic Acids Res.*, 1997, **25**, 4599–4607; (*b*) D. A. Adamiak, W. Rypniewski, J. Milecki and R. W. Adamiak, *Nucleic Acids Res.*, 2001, **29**, 4144–4153.
7 M. Popenda, J. Milecki and R. W. Adamiak, *Nucleic Acids Res.*, 2004, **32**, 4044–4054.
8 D. Placido, B. A. Brown, 2nd, K. Lowenhaupt, A. Rich and A. Athanasiadis, *Structure*, 2007, **15**, 395–404.
9 W. Saenger, *Principles of Nucleic Acids Structure*, 1984, Springer, Berlin, Heidelberg.
10 M. Egli, S. Portman and N. Usman, *Biochemistry*, 1996, **35**, 8489–8494.
11 P. Auffinger and Y. Hashem, *Bioinformatics*, 2007, **23**, 1035–1037.
12 P. Auffinger and E. Westhof, *J. Biomol. Struct. Dyn.*, 1998, **16**, 693–707.
13 A. Fischer, Z. Gdaniec, E. Biala, M. Lozynski, J. Milecki and R. W. Adamiak, *Nucleosides, Nucleotides Nucleic Acids*, 1996, **15**, 477–488.
14 B. Puffer, C. Kreutz, U. Rieder, M. O. Ebert, R. Konrat and R. Micura, *Nucleic Acids Res.*, 2009, **37**, 7728, DOI: 10.1093/nar/gkp862.
15 F. Li, P. S. Pallan, M. A. Maier, K. G. Rajeev, S. L. Mathieu, C. Kreutz, Y. Fan, J. Sanghvi, R. Micura, E. Rozners, M. Manoharan and M. Egli, *Nucleic Acids Res.*, 2007, **35**, 6424–6438.
16 L. Shimoni and J. P. Glusker, *Struct. Chem.*, 1994, **5**, 383–397.
17 J. A. K. Howard, V. J. Hoy, D. O'Hagan and G. T. Smith, *Tetrahedron*, 1996, **52**, 12613–12622.
18 J. D. Dunitz, *ChemBioChem*, 2004, **5**, 614–621.
19 A. Rich, A. Nordheim and A. H. Wang, *Annu. Rev. Biochem.*, 1984, **53**, 791–846.

20 S. N. Rao and P. A. Kollman, *J. Am. Chem. Soc.*, 1986, **108**, 3048–3053.
21 W. Rypniewski, M. Vallazza, M. Perbandt, S. Klussmann, L. J. DeLucas, C. Betzel and V. A. Erdmann, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 659–664.
22 I. Berger, M. Egli and R. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 12116–12121.
23 E. Biala, J. Milecki, A. Kowalewski, M. Popenda, W. Z. Antkowiak and R. W. Adamiak, *Acta Biochim. Pol.*, 1993, **40**, 521–530.
24 Z. Otwinowski and W. Minor, *Methods Enzymol.*, 1997, **276**, 307–326.
25 A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022–1025.
26 G. N. Murshudov, A. Vagin and E. J. Dodson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
27 Collaborative Computational Project, Number 4, *Acta Cryst.*, 1994, **D50**, pp. 760–763.
28 V. S. Lamzin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1993, **49**, 129–147.
29 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
30 W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X. J. Lu, S. Neidle and Z. Shakked, *et al.*, *J. Mol. Biol.*, 2001, **313**, 229–237.
31 W. L. De Lano, *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA, USA, 2002.