© 1998 Oxford University Press
Nucleic Acids Research, 1998, Vol. 26, No. 10 **2473–2480**

Crystal structures of B-DNA with incorporated 2′-deoxy-2′-fluoro-arabino-furanosyl thymines: implications of conformational preorganization for duplex stability

Imre Berger, Valentina Tereshko[1], Hisafumi Ikeda[2], Victor E. Marquez[2] and Martin Egli[1,*]

Institute for Molecular Biology and Biophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland, [1]Drug Discovery Program and Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School, Chicago, IL 60611-3008, USA and [2]Laboratory of Medicinal Chemistry, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received December 16, 1997; Revised and Accepted April 3, 1998 NDB nos BDLB84, BDLB85

ABSTRACT

The fundamental conformational states of right-handed double helical DNA, the A- and B-forms, are associated with distinct puckers of the sugar moieties. The furanose conformation itself is affected by the steric and electronic nature of the ring substituents. For example, a strongly electronegative substituent at the C2′ position, such as in the 2′-deoxy-2′-fluoro*ribo*furanosyl analogue, will drive the conformational equilibrium towards the C3′-*endo* type (north). Conversely, the 2′-deoxy-2′-fluoro*arabino*furanosyl modification with opposite stereochemistry at C2′ appears to have a preference for a C2′-*endo* type pucker (south). Incorporation of 2′-fluoroarabinofuranosyl thymines was previously shown to enhance the thermodynamic stability of B-DNA duplexes. We have determined the crystal structures of the B-DNA dodecamer duplexes $[d(CGCGAASSCGCG)]_2$ and $[d(CGCGAASTCGCG)]_2$ with incorporated 2′-deoxy-2′-fluoroarabinofuranosyl thymines S (south) at 1.55 Å resolution. In the crystal structures, all S residues adopt an O4′-*endo* conformation (east), well compatible with an overall B-form duplex geometry. In addition to the increased rigidity of S nucleosides, a clathrate-like ordered water structure around the 2′-fluorines may account for the observed larger thermodynamic stability of DNA duplexes containing 2′-deoxy-2′-fluoroarabino thymidines.

INTRODUCTION

Helicality and conformational flexibility of double-stranded DNA are consequences of the furanose moiety in the DNA backbone (1). Variations in the duplex geometry are chiefly the result of changes in the backbone torsion angle δ. Altered values of δ and concomitant changes in the sugar pucker give rise to the well known A-form and B-form double helix families. While both are right-handed, the average δ in A-form duplexes is ~80° and the average δ in B-form duplexes is ~120° (1). The relatively low energetic barriers between the conformational states of the deoxyribose sugars are consistent with the high plasticity of the DNA duplex (2). Among the factors that influence the DNA conformational equilibrium are relative humidity (3–5 and references therein), base sequence (6–11), interactions with proteins (e.g. refs 12–15) and the packing forces in crystal lattices (16–18). Moreover, while DNA displays geometrical variation, the overall conformation of the RNA duplex is restricted to the A-form as a result of the electronegative 2′-α-hydroxyl group (2,19,20). Chemical modification of the sugar thus provides an opportunity to influence its conformational properties and stabilize a particular puckering mode (21,22).

The two dominant furanose conformations in nucleosides and nucleotides are the C3′-*endo* (north) and C2′-*endo* (south) states. Changes in the conformation or pseudorotation phase angle *P* of the furanose are driven by the relative strengths of various gauche and anomeric effects (23). Thus, in 2′-deoxyribonucleosides, 5′-OH and 3′-OH prefer a *gauche* orientation with O4′ (as is the case in the south state), whereas in ribonucleosides the conformational equilibrium is affected by three more *gauche* effects, namely those between 2′-OH and 3′-OH, 2′-OH and O4′, and finally 2′-OH and nucleobase nitrogen atoms (see Fig. 1A for orientation). In addition, an anomeric effect exists between the furanose oxygen and the glycosidic bond. The preferred *antiperiplanar* orientation between the α lone pair of O4′ and the C1′-N bond, which furnishes an optimal overlap between the non-bonding and the antibonding σ* orbitals, shifts the conformational equilibrium towards north.

Some 250 chemical modifications affecting either the backbone, phosphate group, furanose sugar or nucleobase have been introduced in recent years as part of the search for antisense oligonucleotides with improved RNA affinity and nuclease resistance (24). Efforts to maximize the thermodynamic stability of a hybrid duplex between antisense strand and RNA target have

*To whom correspondence should be addressed: Tel: +1 312 503 0845; Fax: +1 312 503 0796; Email: m-egli@nwu.edu

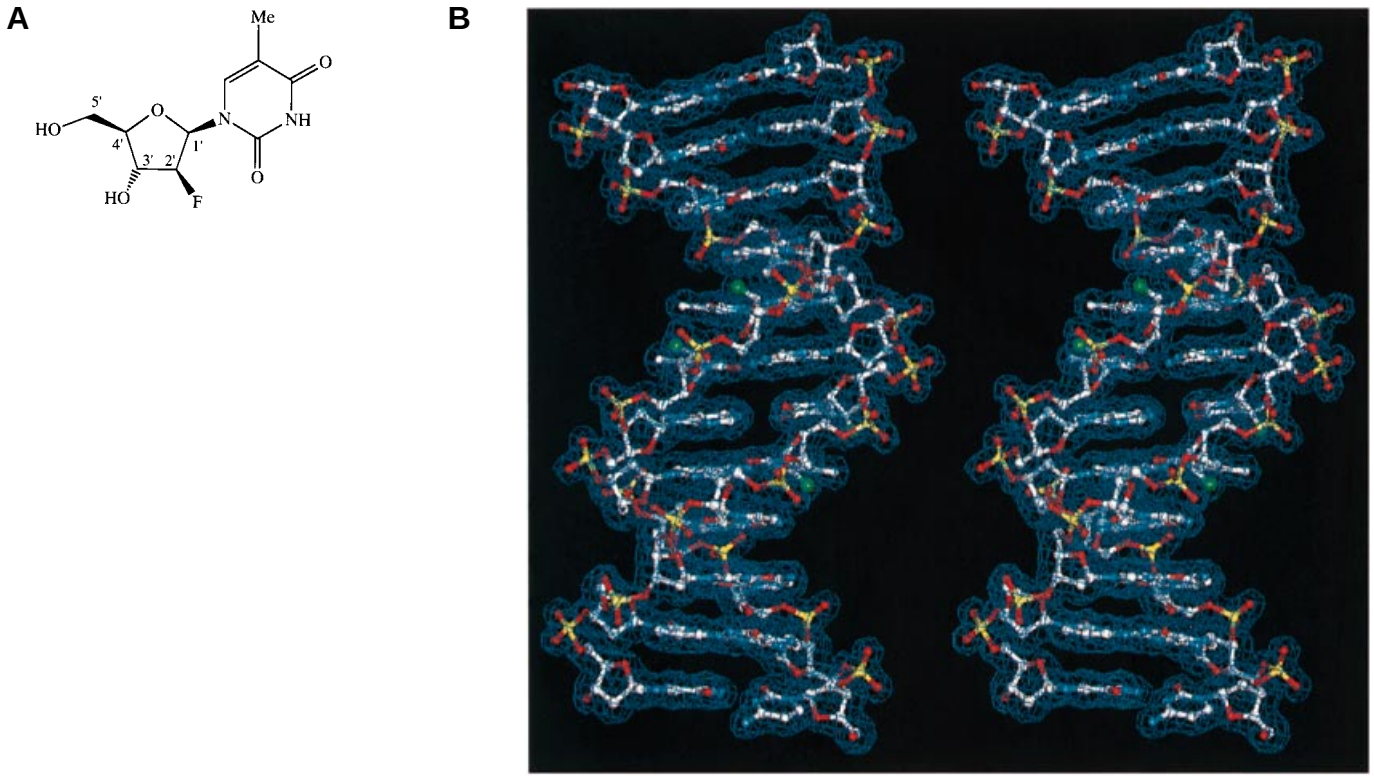

Figure 1. (**A**) Structure of 2′-deoxy-2′-fluoroarabinofuranosyl thymine. (**B**) The $[d(CGCGAASSCGCG)]_2$ structure in a ball-and-stick stereo representation in a $2F_o–F_c$ electron density envelope at 1.55 Å resolution (blue, 1σ contour). Carbon atoms are white, oxygen atoms red, nitrogen atoms blue and phosphorus atoms yellow. The four 2′-fluoro substituents are shown as larger spheres in green.

yielded several promising modifications. Many of these are 2′- or 2′-*O*-modifications that stabilize the C3′-*endo* conformation. Examples are the 2′-deoxy-2′-fluororibonucleosides (25–28), 2′-*O*-methylribonucleosides (29–31) and the 2′-*O*-methoxyethyl-ribonucleoside modification (32–34). Replacement of O3′ by a less electronegative substituent, such as in N3′→P5′ phosphoramidate DNA lacking a 2′-substituent, was also shown to stabilize an A-conformation (35,36). Thus, there are now many ways to stabilize a northern pucker in an oligonucleotide through chemical modification. Conversely, much less work appears to be have been invested in trying to drive the sugar conformational equilibrium towards south, with the aim to possibly stabilize an overall B-form duplex.

Individual nucleosides with a C2′-fluoro substituent in the arabino configuration favor a southern-type conformation (Fig. 1A) (37). Incorporation of 2′-deoxy-2′-fluoroarabinofuranosyl thymines into DNA oligonucleotides had a stabilizing effect relative to the native DNA duplex (38–40). On the other hand, incorporation of the same modification in combination with cytosine and uracil led to a slight destabilization of DNA duplexes (38,39). Recent work suggests that alternating stretches of 2′-deoxy-2′-fluororibo and 2′-deoxy-2′-fluoroarabino residues in a DNA duplex can introduce A–B junctions (40). In order to determine the conformation of 2′-deoxy-2′-fluoroarabino residues in the oligonucleotide context and to rationalize the origins of the higher stability of duplexes containing such residues, we have replaced either one or both thymidines in the Dickerson–Drew DNA dodecamer 5′-CGCGAATTCGCG by 2′-deoxy-2′-fluoroarabinofuranosyl thymines and determined the crystal structures of the corresponding DNA duplexes at high resolution. To avoid confusion, we have adopted the designation by Ikeda and coworkers (40) and we will refer to 2′-deoxy-2′-fluoroarabinofuranosyl thymine as S (for south). Initially, this letter was thought to represent the conformational preference of the nucleoside analogue (40). However, our work now demonstrates that these residues typically assume an O4′-*endo* pucker in B-DNA and that, perhaps, E (for east) would have been a more appropriate designation.

Here, we report the crystal structures of the duplexes $[d(CGCGAASSCGCG)]_2$ (SS dodecamer) and $[d(CGCGAASTCGCG)]_2$ (ST dodecamer) as well as the conformational properties of S nucleosides in a B-DNA environment. Detailed analysis of the structures suggests that the conformationally rigid S residues, compatible with a B-form conformation, as well as the ordered water structure in the vicinity of the 2′-fluorine atoms, are the main contributors to the elevated stability of DNA duplexes with S residues in place of T.

MATERIALS AND METHODS

Synthesis and purification

The SS and ST dodecamers were synthesized according to described procedures (40) in a 10 µmol scale by Oligos Etc. Inc., Wilsonville, OR. They were purified to >90% by reverse phase HPLC and after desalting and lyophilization, the modified oligonucleotides were obtained as white powders. The concentrations of stock solutions for both dodecamers were adjusted to 5 mM (single strand) in water as assessed by UV absorption.

Table 1. Selected crystal data and refinement parameters

Structure	[d(CGCGAASSCGCG)]$_2$[a]	[d(CGCGAASTCGCG)]$_2$[a]
space group	$P2_12_12_1$	
crystal system	orthorhombic	
unit cell dimensions (Å)	a = 24.93	a = 25.20
	b = 39.82	b = 39.71
	c = 65.62	c = 65.68
temperature (°C)	–170	
V_{asym} (Å^3)	16 286	16 431
V / basepair (Å^3)	1357	1369
strands per asym. unit	2	
non-hydrogen atoms	486 DNA and 4 F atoms	486 DNA and 2 F atoms
ions, solvent molecules	1 Mg^{2+}, 177 H_2O	1 Mg^{2+}, 186 H_2O
number of reflections	7894	8507
(8.0–1.55 Å, F ≥ 2σ[F])		
completeness (%) (8.0–1.65 Å)	89.8	94.4
completeness (%) (1.65–1.55 Å)	29.9	38.8
r.m.s deviation bonds (Å)	0.011	0.014
r.m.s. deviation angles (°)	1.88	2.18
r.m.s. deviation improper torsion angles (°)	1.79	1.93
final *R*-factor[b] (%)	20.8	21.8

[a]S = 2′-deoxy-2′-fluoroarabino-T.
[b]Using a multiscale procedure (10 bins) and all reflections with F ≥ 2σ (F) between 8.0 and 1.55 Å.

Crystallization and X-ray diffraction data collection

Crystals were grown at room temperature by the sitting drop vapor diffusion method. Droplets (20 µl) containing 1 mM oligonucleotide, 20 mM sodium cacodylate pH 7, 10 mM magnesium acetate and 3 mM spermine tetrahydrochloride were equilibrated against a reservoir of 25 ml 40% 2-methyl-2,4-pentanediol (MPD). Large hexagonal rods (SS-dodecamer) or rectangular plates with sharp edges (ST dodecamer) appeared after a week. For data collection, crystals were picked up with a nylon loop along with a small amount of mother liquor and were directly transferred into a cold nitrogen stream. Using an in-house R-axis II image plate system mounted on a Rigaku rotating anode X-ray generator, 3°-oscillation frames were collected at high and low resolution ranges for both modified dodecamers. Selected crystal data are summarized in Table 1. Data were processed and merged with the DENZO/SCALEPACK program package (41), and overall resolutions and reflection statistics are given in Table 1.

It is noteworthy that the resolutions of the SS and ST dodecamer data sets are well beyond those which have been reported for Dickerson–Drew type dodecamers over the last 15 years [the resolutions were retrieved from the Nucleic Acid Database (NDB; 42)]. Although one cannot exclude the possibility that the chemical modification may lead to more highly ordered and therefore better diffracting crystals, it is more likely that improvements in the synthesis and purification of oligonucleotides and the way crystallographic data are collected nowadays are responsible for the dramatic improvement in the resolutions that can now be achieved with dodecamer-type B-DNA crystals. This is fully consistent with the high resolution (1.4 Å) of the recently reported structure for the native Dickerson–Drew DNA dodecamer (43).

Structure determination and crystallographic refinement

The isomorphous high-resolution d(CGCGAATTCGCG) structure by Williams and coworkers (43) served as the starting model for the determination of the SS and ST structures [NDB ID BDL084]. The crystallographic refinements were carried out with program X-PLOR [Version 3.851 (44)]. The DNA-RNA-multi-endo nucleic acid parameter file was used, extended by a patch for the 2′-deoxy-2′-fluoroarabino thymines (45 and references therein). Rigid body refinement (five cycles) followed by positional refinement (25 cycles) of the unmodified DNA model with data between 8 and 2.5 Å resolution led to an *R*-factor of <30% in both cases. The 2′-fluorine atoms of the modified thymine residues were clearly visible in superimposed $2F_o–F_c$ sum and $F_o–F_c$ difference Fourier electron density maps. In addition, one hydrated magnesium ion per dodecamer duplex could be identified early on in electron density maps in both structures. Structured water molecules were located in superimposed sum and difference electron density maps and were included in further refinement cycles in groups of five. Selected refinement parameters are listed in Table 1 and a $2F_o–F_c$ sum electron density map around the final model for the SS dodecamer is depicted in Figure 1B.

Table 2. Comparison between backbone torsion angles, glycosidic torsion angles (in degrees) and sugar puckers for 2′-deoxy-2′-fluoroarabinothymidines and thymidines in the chemically modified and native B-DNA dodecamers, respectively[a,b]

Residue	α	β	γ	δ	ε	ξ	χ	*P*	pucker
[d(CGCGAASSCGCG)]$_2$[c]									
S7	–52.5	159.8	56.7	99.7	–178.2	–83.5	–133.4	103.1	O4′-*endo*
S8	–59.8	163.1	60.6	104.3	177.6	–94.0	–128.4	98.4	O4′-*endo*
S19	–67.0	169.6	67.1	99.4	–177.7	–90.3	–123.8	102.4	O4′-*endo*
S20	–51.9	156.9	55.7	95.7	176.3	–84.1	–127.9	95.7	O4′-*endo*
[d(CGCGAASTCGCG)]$_2$[c]									
S7	–57.5	161.6	59.3	92.6	–178.3	–80.2	–134.4	95.3	O4′-*endo*
S19	–60.9	171.3	60.9	96.8	179.2	–83.4	–136.5	96.3	O4′-*endo*
[d(CGCGAATTCGCG)]$_2$[d]									
T7	–49.0	172.9	48.6	112.7	–176.6	–95.6	–119.6	125.2	C1′-*exo*
T8	–54.1	168.3	53.4	114.0	171.5	–95.0	–119.9	126.9	C1′-*exo*
T19	–49.1	178.5	51.6	127.6	–172.8	–106.8	–115.9	143.0	C1′-*exo*
T20	–44.5	171.8	43.4	135.3	–164.1	–106.8	–110.1	151.4	C2′-*endo*
B-DNA									
global[e]	–41	136	38	139	–133	–157	–102	–	C2′-*endo*

[a]Calculated with program NEWHEL93, distributed by R.E.Dickerson.
[b]Backbone torsion angles are defined as O3′$_{n-1}$-P-α-O5′-β-C5′-γ-C4′-δ-C3′-ε -O3′-ξ-P-O5′$_{n+1}$, χ is the glycosidic torsion angle (O4′-C1′-N9-C4 for purines and O4′-C1′-N1-C2 for pyrimidines), and *P* is the pseudorotation phase angle (all in degrees).
[c]S = 2′-deoxy-2′-fluoroarabino-T.
[d]From the recent high-resolution structure of the native Dickerson–Drew dodecamer duplex (43).
[e]From ref. 2.

RESULTS AND DISCUSSION

Overall duplex geometries and conformation of 2′-deoxy-2′-fluoroarabino thymidine

The overall conformations of the two modified dodecamers deviate only minimally from one another and the rms deviation between SS and ST dodecamer is 0.24 Å. The rms deviations between SS dodecamer and ST dodecamer and the reference structure (NDB ID BDL084) are 0.47 and 0.39 Å, respectively. Therefore, the modified thymidines do not significantly alter the gobal geometry of the B-DNA duplex.

In the duplexes, all puckers of modified residues fall into the O4′-*endo* (east) range. The 2′-fluorines therefore shift the conformation in a northerly direction relative to the sugars in the unmodified reference structure. The average backbone, sugar and glycosidic torsion angles for all six S residues as well as a representative 2′-deoxy-2′-fluoroarabinofuranose conformation are shown in Figure 2. Backbone and glycosidic torsion angles as well as pseudorotation phase angles and pucker types for individual modified residues and the corresponding nucleosides in the reference structure are given in Table 2. If only considering the various gauche effects in the sugar moiety, one could have felt inclined to predict a more southern-type conformation for S residues. In such an arrangement, the torsion angles F2′-C2′-C1′-O4′ and F2′-C2′-C1′-N1 would display values of around –80° and +40°, respectively. By comparison, for residue S7 in the ST dodecamer structure (Fig. 2B), the F2′-C2′-C1′-O4′ torsion angle is –97° and the F2′-C2′-C1′-N1 torsion angle is +24°. In terms of anomeric effect alone, the conformation

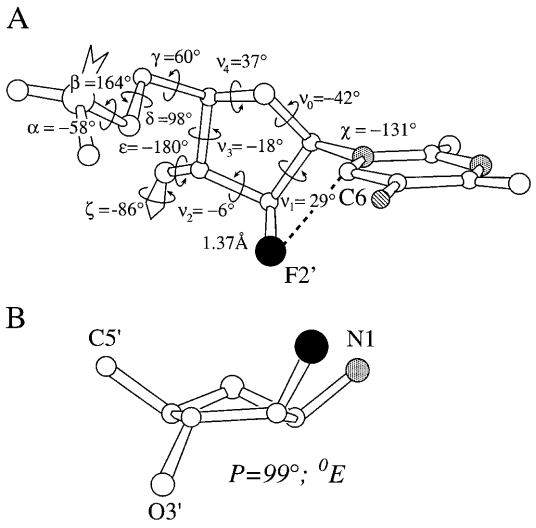

Figure 2. Geometry of the 2′-deoxy-2′-fluoroarabinofuranosyl moieties in the modified DNA dodecamer duplexes. (**A**) Average backbone and sugar torsion angles of the six modified residues in the SS and ST structures. The dashed line indicates the van der Waals contact between F2′ and C6 that is present in all modified residues in the two structures. (**B**) Drawing illustrating the 2′-deoxy-2′-fluoroarabinofuranosyl conformation of nucleoside S7 in the ST dodecamer duplex. Pseudorotation phase angle *P* and the sugar pucker represent the average value and range, respectively, observed in the six modified residues. Nitrogen atoms are stippled in grey, the C5-methyl carbon atom is dashed, and the 2′-fluorine atom is black. The C2′-F2′ bond length is 1.37 Å.

adopted by the modified thymines is favored over a pure C2′-*endo* conformation.

However, close inspection of the local environment of the 2′-fluorine atoms in a B-DNA environment suggests that steric rather than stereoelectronic reasons are responsible for the observed O4′-*endo* sugar conformations. The average intra-nucleoside distance between F2′ and C6 for the six modified residues is 2.93 Å (min. 2.79 Å, max. 3.12 Å). The corresponding average distance between F2′ and H6 (C-H bond length 1 Å) is 2.65 Å (min. 2.52 Å, max. 2.89 Å). These average values are comparable to the sum of the van der Waals radii for carbon (1.5 Å) and fluorine (1.35 Å), 2.85 Å, and for hydrogen (1.2 Å) and fluorine, 2.55 Å, respectively. Thus, for an O4′-*endo* conformation, the fluorine is in van der Waals contact with pyrimidine C6-H6. A southern shift would result in a clash between 2′-fluorine and base (Fig. 2A). For a C2′-*endo* conformation, the F2′....C6 distance would be ~2.7 Å and the F2′....H6 distance would be ~2.1 Å (C2′-F2′ bond length 1.35 Å; C6-H6 bond length 1.0 Å). This demonstrates that a C2′-*endo* pucker is prohibited because of an unfavorable 1–5 contact. As described in more detail below, the nucleoside cannot avoid this clash by altering the glycosidic angle (moving from *minus anticlinal* to *antiperiplanar*), since this would lead to an unfavorable inter-nucleoside contact between F2′ and the C5 methyl group of the 3′-adjacent modified thymidine. A further conclusion from this steric control of the sugar conformation in 2′-modified *ara* pyrimidines is that, in the case of modified purines, even the O4′-*endo* conformation would not prevent a repulsive interaction between F2′ and the purine C8-H8 bond. Therefore, it is likely that incorporation of 2′-deoxy-2′-fluoroarabinofuranosyl purine residues into DNA oligonucleotides would result in a destabilization of duplex formation.

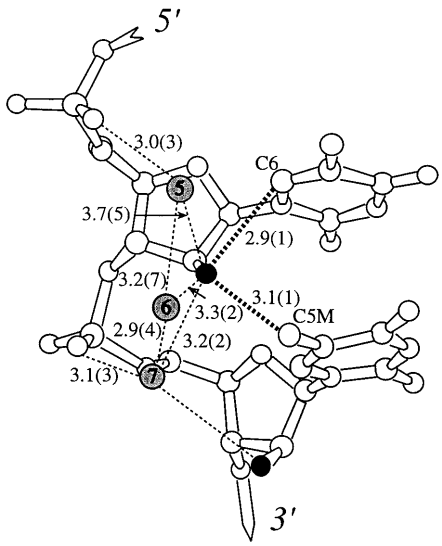

Figure 3. Ball-and-stick drawing illustrating the consensus water environment of the arabinose moiety of modified thymidines in the crystal structures of the SS and ST dodecamers. The view is into the border region of the major groove, 2′-fluorine atoms are black and water oxygen atoms are grey. Close contacts between F2′ of S7 (S19) and C6 of the same residue as well as the C5-methyl group of the 3′-adjacent residue S/T8 (S/T20) are indicated by thick dashed lines. Hydrogen bonds are drawn with thin dashed lines and the numbers correspond to the average distances (standard deviations in parentheses) observed in the four strands containing S residues. For more details, see Figure 4 legend.

Conformational properties of the SS and ST dinucleotide steps

The close resemblance between the SS and ST dodecamer structures and the one of the native DNA dodecamer shows that an O4′-*endo* conformation of sugars is compatible with a B-form geometry. The values for torsion angles of individual residues given in Table 2 are rather similar in all three duplexes and no dramatic local conformational perturbations appear necessary to accommodate the modified residues. In fact, residues G3 and C21 in the all-DNA reference duplex also assume eastern-type puckers. In the case of G3, the conformation is C4′-*exo* ($P = 65°$), and in the case of C21, the conformation is O4′-*endo* ($P = 94°$). In both SS and ST dodecamer, the puckers of residues 5′- and 3′-adjacent to the modified nucleosides appear virtually unaffected by the particular conformation displayed by the latter residues. None of the neighboring residues adopts an O4′-*endo* pucker. In the SS dodecamer, the sugar of residue C9 assumes a C2′-*endo* pucker and residue C21 displays a C1′-*exo* pucker. However, as pointed out above, that residue displays a similar conformation even in the all-DNA duplex. The thymidines 3′-adjacent to the S residues in the ST structure both assume C1′-*exo* puckers. Similarly, adenosines from the opposite strand which are base paired to S residues show no alterations in their puckers.

Rather than inducing a conformational discontinuity, the S residues are accommodated through subtle conformational changes in the neighboring residues with both modified dodecamers. This is consistent with earlier observations that the global conformations of DNA duplexes with incorporated mismatched base pairs were practically identical with those of the native duplexes (reviewed in ref. 6). The only X-ray crystallographic study of a DNA containing arabinofuranosyl residues was conducted with crystals of the left-handed Z-DNA duplex $[(araCdGaraCdGaraCdG)]_2$ (47). In that structure, the arabinoses of cytidines adopt C3′-*exo* pucker ($P = 190°$, south), closely related to the C2′-*endo* conformation normally displayed by the deoxyriboses of pyrimidines in crystal structures of Z-DNA hexamers. The overall conformation of the left-handed duplex changes only minimally as a result of the presence of the arabinofuranosyl residues. In spite of the fact that the sugar conformations of ara-C in Z-DNA and S in our SS and ST dodecamers are different, the small changes in the overall geometries caused by the arabino-type residues are a shared feature of the three structures.

In addition to the van der Waals contacts between 2′-fluorines and C6, the fluorines of residues S7 and S19 are forming close contacts to the C5 methyl group of the 3′-adjacent S (SS dodecamer) or T (ST dodecamer) residues. The average distance is 3.06 Å (min. 2.91 Å, max. 3.24 Å) and is thus about 0.1 Å longer than the average contact between 2′-fluorines and C6. An example of such a short contact is depicted in Figure 3. The conformational freedom of S residues is thus severely limited. In order to avoid a repulsive intra-nucleoside contact between fluorine and the 6-carbon, the sugar pucker is restricted to the O4′-*endo* range. It may initially appear possible to relieve a possible short contact and to expand the puckering range of the sugar by altering the glycosidic torsion angle χ. However, this would result in a short contact between the 2′-fluorine and the exocyclic methyl group of the 3′-adjacent T or S residue (Fig. 3, top and center). The values of glycosidic torsion angles for S and

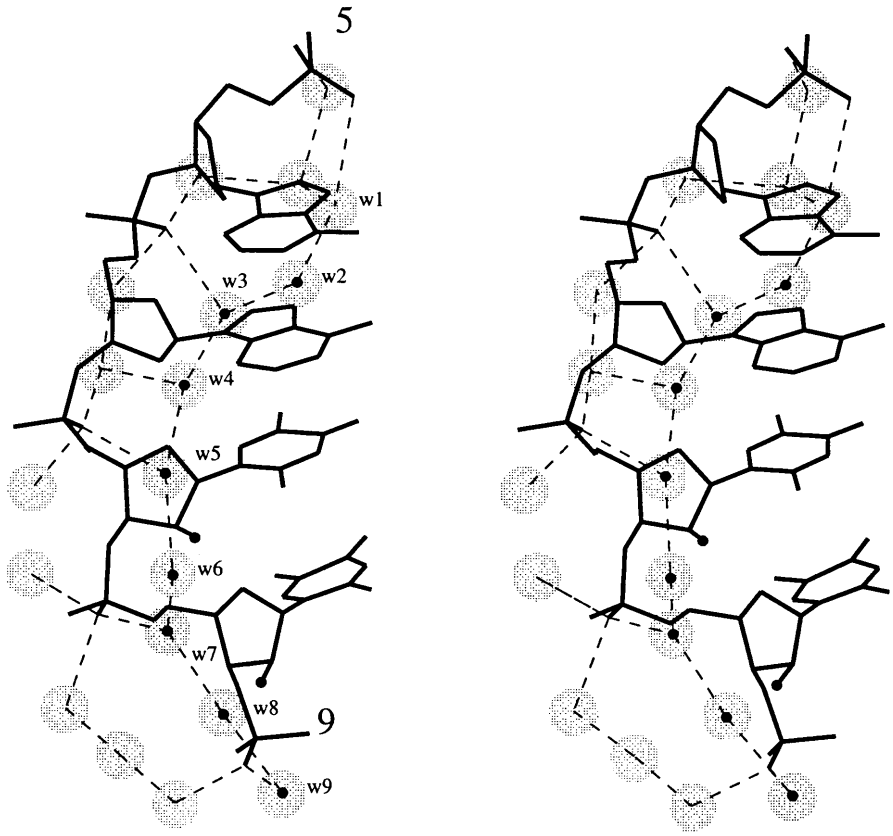

Figure 4. Stereo drawing depicting the consensus hydration pattern along the border of the major groove and around O2P oxygen atoms in the central portions (residues labeled) of the SS and ST dodecamers structures as well as their all-DNA counterpart. 2′-Fluorine atoms are black, water molecules are large spheres, stippled in grey, and hydrogen bonds are drawn with dashed lines. Water positions with a higher degree of conservation in the three crystal structures are highlighted with a black dot. The numbering of water molecules is consistent with Figure 3. Each of the 2′-fluorine atoms is in close contact with up to three water molecules. For residues 7 and 19 (second from bottom), these are water molecules W5, W6 and W7. In the case of residues 8 and 20 (bottom), the three corresponding water molecules are W7, W8 and W9. In Figure 3, these three consensus water molecules were arbitrarily numbered W5, W6 and W7.

T residues in the SS and ST dodecamer structures are slightly higher compared with those in the unmodified DNA (Table 2). This is consistent with our interpretation that opening up the χ torsion angle (toward *ap*) can relieve a possible clash between base-C6 and 2′-fluorine atom. One might have expected a shift towards south in the puckers of residues S7 and S19, as a change in their glycosidic torsion angle is not restricted by a potential short contact between the C5-methyl group and a fluorine atom from the 5′-adjacent residue (an unmodified A, Fig. 3). However, rotating a single base will produce propeller twisting in the base pair and lead to less optimal base stacking. Both changes are expected to negatively affect the enthalpic stabilization.

Initially only the SS dodecamer had been synthesized. While analyzing its crystal structure, the short contact between fluorine and the exocyclic methyl group of the 3′-adjacent S was noted and was interpreted as a possible weakly stabilizing interaction. Accordingly, the presence of the electronegative fluorine in the sugar moiety may lower the pK_a of the allylic protons and a close contact between methyl group and a putative acceptor would constitute a favorable interaction. In order to address this hypothesis, the structure of the ST dodecamer was also determined. In the ST dodecamer, absence of the fluorine in thymidine renders the pK_a of the allylic protons unchanged and if indeed the CH_3....F interaction were an attractive one, the distance between fluorine in S and the methyl group from the 3′-adjacent nucleotide would be expected to become longer as a result of the removal of fluorine in the former one. However, the observed slightly shorter CH_3....F distances for nucleotides S7/T8 and S19/T20 in the ST dodecamer (2.91 and 3.09 Å, respectively) relative to the corresponding distances for nucleotides S7/S8 and S19/S20 in the SS dodecamer (3.01 and 3.24 Å, respectively) clearly show that this is indeed a tolerated near van der Waals contact rather than a stabilizing electrostatic interaction. This finding provides good evidence for the interpretation that the uniform conformation of S residues in the modified duplexes is, although consistent with the expected gauche effects, mainly the result of an accommodation of the fluorine substituent without generating severely destabilizing, short intra- and internucleoside contacts.

Structured water around the 2′-fluorines

The high resolutions of the two modified DNA duplex structures provide a good opportunity to examine the arrangement of water molecules around fluorine atoms which are covalently linked to the DNA framework. Moreover, the even higher resolution of the reference structure allows for a meaningful assessment of changes in the hydration as a consequence of the chemically modified sugar moieties. We have limited our analysis of the DNA hydration in our crystal structures and its comparison to the situation in the structure of the native duplex to the major groove

borders, since this region is likely to be most affected by the presence of the 2′-fluorine atoms (Fig. 1B).

Organic fluorine is generally considered a poor hydrogen bond acceptor, but in cases where fluorine is part of an anion and therefore unusually electron rich as well as in situations where the donating water molecule is coordinated to a metal cation and thus a particularly strong donor, the existence of O-H....F hydrogen appears certain (48). DNA is a polyanion, but one should bear in mind that the negative charge is not located on the sugar moiety and conjugation to the covalently bound fluorine through the backbone is limited. The consensus water structure around S residues in the major groove along with the average F2′....O(W) distances are depicted in Figure 3. Figure 4 shows the consensus arrangement of water molecules along the rim of the major groove, including the phosphate backbone for the SS, ST and reference dodecamer structures.

Examination of the detailed geometry of water molecules around the 2′-fluorines reveals that each of them forms relatively close contacts to three water molecules (Fig. 3). Water molecules that are located at distances of between 3.0 and 3.5 Å from fluorines are usually engaged in a primary hydrogen bonding interaction to a phosphate group (Fig. 3). Typically, the latter hydrogen bond is then significantly shorter than the contact to the fluorine atom. In the case of residue S7 (SS dodecamer), water 5 (see Fig. 5 for orientation) is located at 2.76 Å from O2P of the same residue (W5....F2′[S7] = 3.37 Å), and water 7 is located at 2.98 Å from O2P of residue S8 (W7....F2′ [S7] = 3.08 Å). The distance between water 6, the central water molecule in Figure 5, and F2′ of residue S7 is 3.35 Å and is thus rather long. However, its main hydrogen bonding interactions are to water 5 (2.95 Å) and to water 7 (2.64 Å). The fluorine atom of residue S20 in the SS dodecamer structure forms three contacts to water molecules, all ≤3.1 Å. One is tempted to use the term hydrogen bond for describing these O-H....F2 contacts. Still, such relatively short contacts are also somewhat reminiscent of the clathrate-like water structure observed around the methyl groups of a leucine residue in the crystal structure of crambin (49) and of a valine residue in a peptide crystal (50). A comparison of the arrangements of water molecules around C2′ in the SS and ST dodecamer structures with those in the native DNA structure shows that the number of waters within a sphere centered at C2′ are similar for the 2′-deoxy-2′-fluoroarabinose and 2′-deoxyribose moieties (data not shown). In addition, the distribution of water molecules in this region within the major groove for the SS and ST duplexes as well as the reference duplex are rather similar (Fig. 4). However, it is likely that the presence of fluorine atoms leads to more highly ordered first-shell water molecules. The edges of thymidines (C6 and the C5-methyl group) form relatively hydrophobic patches along the borders of the major groove. The insertion of fluorines at such sites will trap water molecules between sugar and phosphate.

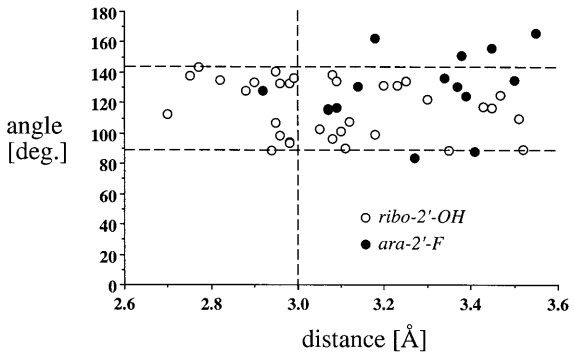

Figure 5. Comparison between the distributions of water molecules around the *ara* 2′-fluorine atoms in the SS and ST dodecamer duplex crystal structures (○) and around the *ribo* 2′-hydroxyl groups in the crystal structure of the RNA duplex $[r(CCCCGGGG)]_2$ (●; 51). The F2′....W and O2′....W distances are plotted against the C2′-F2′-W and C2′-O2′-W angles, respectively. Half the water molecules that form a hydrogen bond to an RNA 2′-hydroxyl group are located at 3 Å or less from the 2′-oxygen atom (this limit is indicated by the vertical dashed line). By comparison, despite the smaller van der Waals radius of fluorine relative to oxygen, only one water molecule lies less than 3 Å from a 2′-fluorine. Moreover, the distribution of water molecules around RNA 2′-hydroxyl groups displays a more pronounced clustering compared with the distribution of waters around 2′-fluorines. All C2′-O2′-W angles fall into a 50° range centered around 110° (upper and lower limits indicated by horizontal dashed lines), whereas almost half the C2′-F2′-W angles lie outside this range. Taken together, these findings argue strongly against fluorine covalently bound to DNA acting as an effective hydrogen bond acceptor.

Thus, none of the deoxyribose 2′-methylene groups in the native dodecamer displays contacts of less than 3.1 Å to three water molecules, as is the case for the 2′-fluorine of residue S20 in the SS dodecamer structure. However, the energetic consequences of such rearrangements in the first hydration shell are hard to assess.

In order to further characterize the potential of covalently bound fluorine to act as a hydrogen bond acceptor, we have compared the distributions of water molecules around the 2′-fluorines with those around the 2′-hydroxyl groups of RNA. The 2′-hydroxyl group is an excellent hydrogen bond donor and acceptor. Although oxygen has a slightly larger van der Waals radius than fluorine, the average distance between water molecules and ribose 2′-OH groups as observed in a recent high resolution RNA crystal structure (51) is clearly shorter than the one between water and fluorine in the SS and ST dodecamer structures (Fig. 5). In addition, the hydrogen bonds between water and 2′-OH display a narrower angle distribution compared with the interactions between water and fluorine. Together with the above results, our structures provide no evidence for the formation of strongly stabilizing hydrogen bonds by the fluorines in DNA. Nevertheless, the fluorines appear to have a definite ordering effect on the water structure on the DNA surface.

Table 3. UV melting temperatures (T_m)[a] and thermodynamic parameters[b] (40)

Oligonucleotide	T_m (°C)	ΔT_m (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal/mol)	$\Delta G°$ (kcal/mol)
d(CGCGAATTCGCG)	51.7	–	–46.8	–120.2	–9.6
d(CGCGAASSCGCG)[c]	59.1	7.4	–47.4	–119.3	–10.4
d(CGCGAASTCGCG)[c]	52.9	1.2	–49.7	–128.2	–9.9

[a]10 mM sodium phosphate pH 7, 0.1 mM EDTA, 100 mM NaCl.
[b]From $1/T_m$ versus log C_T plots, $\Delta G°$ at 37°C.
[c]S = 2′-deoxy-2′-fluoroarabino-T.

CONCLUSIONS

Replacement of thymidine by 2′-deoxy-2′-fluoroarabino thymidine in DNA oligonucleotides leads to a substantial gain in the thermodynamic stability of duplex formation (38–40). In the case of SS and ST dodecamer, the change in the UV melting temperatures relative to the native Dickerson–Drew sequence amounts to almost +2°C per modified residue (Table 3). The individual thermodynamic parameters reveal that the increased stabilities are due to favorable enthalpy and entropy terms. Our crystallographic analyses of two DNA dodecamers with incorporated modified thymidines show that, contrary to the anticipated south pucker, S nucleosides adopt an O4′-*endo* (east) pucker. Remarkably, O4′-*endo* residues can be accommodated by the B-DNA duplex without any drastic conformational perturbations in the overall geometry. It appears that it is mainly steric rather than stereoelectronic reasons that account for the observed conformation of the arabino moieties. The O4′-*endo* envelope constitutes an exception in B-DNA oligonucleotide crystal structures, but a pure C2′-*endo* (south) pucker would create a repulsive contact between F2′ and the 6-carbon of the base, provided the nucleoside maintains a glycosidic torsion angle that is compatible with a B-form geometry. Thus, the stability gain appears to be chiefly the result of a conformational preorganization of S nucleosides for the B-type duplex geometry. Despite an ordering effect on the water structure along the edges of the major groove due to the presence of the fluorine atoms, the capacity of the latter ones to act as hydrogen bond acceptors is clearly limited and cannot therefore account for the enthalpic stability gain. The SS and ST dodecamer structures constitute the first examples of crystal structures of oligonucleotides with covalently bound fluorine. It remains to be seen whether conformational preorganization, although now for an A-type duplex, is also the main reason for the superior thermodynamic stability displayed by oligo-2′-deoxy-2′-fluoro-*ribo*nucleotides.

ACKNOWLEDGEMENTS

We are grateful to Prof. Jack D.Dunitz, ETH-Zürich, for helpful discussions and Prof. Loren D.Williams, Georgia Institute of Technology, for sending us the coordinates of the 1.4 Å crystal structure of the native Dickerson–Drew B-DNA dodecamer prior to publication. This work was supported by NIH grant R01 GM-55237 (M.E.).

REFERENCES

1 Eschenmoser,A. and Dobler,M. (1992) *Helv. Chim. Acta* **75**, 218–259.
2 Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
3 Franklin,R.E. and Gosling,R.G. (1953) *Nature* **171**, 740–741.
4 Saenger,W., Hunter,W.N. and Kennard,O. (1986) *Nature* **324**, 385–388.
5 Harmouchi,M., Albiser,G. and Premilat,S. (1990) *Eur. Biophys. J.* **19**, 87–92.
6 Kennard,O. and Hunter,W.N. (1991) *Angew. Chem. Int. Ed. Engl.* **30**, 1254–1277.
7 Dickerson,R.E. (1992) In Lilley,D.M.J. and Dahlberg,J.E. (eds), *Methods in Enzymology*, Vol. 211, part A, *DNA Structures* Academic Press Inc., San Diego, CA, pp. 67–111.
8 Egli,M. (1994) In Bürgi,H.-B. and Dunitz,J.D. (eds), *Structure Correlation*, Vol. 2, VCH Publishers Inc., New York, NY, pp. 705–749.
9 Grzeskowiak,K. (1996) *Chem. Biol.* **3**, 785–790.
10 Hartmann,B. and Lavery,R. (1996) *Q. Rev. Biophys.* **29**, 309–368.
11 Subirana,J.A. and Faria,T. (1997) *Biophys. J.* **73**, 333–338.
12 Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) *Nature* **365**, 512–520.
13 Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993) *Nature* **365**, 520–527.
14 Guzikevich-Guerstein,G. and Shakked,Z. (1996) *Nature Struct. Biol.* **3**, 32–37.
15 Urpi,L., Tereshko,V., Malinina,L., Huynh-Dinh,T. and Subirana,J.A. (1996) *Nature Struct. Biol.* **3**, 325–328.
16 Shakked,Z., Guerstein-Guzikevich,G., Eisenstein,M., Frolow,F. and Rabinovich,D. (1989) *Nature* **342**, 456–460.
17 Jain,S. and Sundaralingam,M. (1989) *J. Biol. Chem.* **264**, 12780–12784.
18 Dickerson,R.E., Goodsell,D.S. and Neidle,S. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3579–3583.
19 Rich,A. (1995) In Chambers,D.A. (ed.), *DNA: The Double Helix.* Annals of the New York Academy of Sciences, Vol. 758, New York, NY, pp. 97–142.
20 Egli,M., Usman,N. and Rich,A. (1993) *Biochemistry* **32**, 3221–3237.
21 Griffey,R.H., Lesnik,E., Freier,S., Sanghvi,Y.S., Teng,K., Kawasaki,A., Guinosso,C., Wheeler,P., Mohan,V. and Cook,P.D. (1994) In Sanghvi,Y.S. and Cook,P.D. (eds.) *Carbohydrate Modifications in Antisense Research*. ACS Symposium Series Vol. 580, pp. 212–224.
22 Egli,M. (1996) *Angew. Chem. Int. Ed. Engl.* **35**, 1894–1909.
23 Plavec,J., Thibaudeau,C. and Chattopadhyaya,J. (1996) *Pure Appl. Chem.* **68**, 2137–2144.
24 Freier,S.M. and Altmann,K.-H. (1997) *Nucleic Acids Res.* **25**, 4429–4443.
25 Fazakerley,G.V., Uesugi,S., Izumi,A., Ikehara,M. and Guschlbauer,W. (1985) *FEBS Lett.* **182**, 365–369.
26 Williams,D.M., Benseler,F. and Eckstein,F. (1991) *Biochemistry* **30**, 4001–4009.
27 Kawasaki,A.M., Casper,M.D., Freier,S.M., Lesnik,A.E., Zounes,M.C., Cummins,L.L., Gonzalez,C. and Cook,P.D. (1993) *J. Med. Chem.* **36**, 831–841.
28 Monia,B.P., Lesnik,E.A., Gonzalez,C., Lima,W.F., McGee,D., Guinosso,C.J., Kawasaki,A.M., Cook,P.D. and Freier,S.M. (1993) *J. Biol. Chem.* **268**, 14514–14522.
29 Inoue,H., Hayase,Y., Imura,A., Iwai,S., Miura,K. and Ohtsuka,E. (1987) *Nucleic Acids Res.* **15**, 6131–6148.
30 Lesnik,E.A., Guinosso,C.J., Kawasaki,A.M., Sasmor,H., Zounes,M., Cummins,L.L., Ecker,D.J., Cook,P.D. and Freier,S.M. (1993) *Biochemistry* **32**, 7832–7838.
31 Lubini,P., Zürcher,W. and Egli,M. (1994) *Chem. Biol.* **1**, 39–45.
32 Martin,P. (1995) *Helv. Chim. Acta* **78**, 486–504.
33 Altmann,K.-H., Dean,N.M., Fabbro,D., Freier,S.M., Geiger,T., Häner,R., Hüsken,D., Martin,P., Monia,B.P., Müller,M., Natt,F., Nicklin,P., Phillips,J., Pieles,U., Sasmor,H. and Moser,H.E. (1996) *Chimia* **50**, 168–176.
34 Altmann,K.-H., Fabbro,D., Dean,N.M., Geiger,T., Monia,B.P., Müller,M. and Nicklin,P. (1996) *Biochem. Soc. Trans.* **24**, 630–637.
35 Gryaznov,S.M., Lloyd,D.H., Chen,J.-H., Schultz,R.G., DeDionisio,L.A., Ratmeyer,L. and Wilson,W.D. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5798–5802.
36 Tereshko,V., Gryaznov,S. and Egli,M. (1998) *J. Am. Chem. Soc.* **120**, 269–283.
37 Barchi,J.J.,Jr, Jeong,L.-S., Siddiqui,M.A. and Marquez,V.E. (1997) *Biochem. Biophys. Meth.* **34**, 11–29.
38 Rosenberg,I., Soler,J.F., Tocik,Z., Ren,W.-Y., Ciszewski,P.K., Pankiewicz,K.W., Spassova,M. and Watanabe,K.A. (1993) *Nucleosides Nucleotides* **12**, 381–401.
39 Kois,P., Tocik,Z., Spassova,M., Ren,W.-Y., Rosenberg,I., Soler,J.F. and Watanabe,K.A. (1993) *Nucleosides Nucleotides* **12**, 1093–1109.
40 Ikeda,H., Fernandez,R., Wilk,A., Barchi,J.J.,Jr and Marquez,V.E. (1998) *Nucleic Acids Res.* **26**, 2237–2244.
41 Otwinowski,Z. and Minor,W. (1997) *Methods Enzymol.* **276**, 307–326.
42 Berman,H.M., Olson,W.K., Beveridge,D.L., Westbrook,J., Gelbin,A., Demeny,T., Hsieh,S.-H., Srinivasan,A.R. and Schneider,B. (1992) *Biophys. J.* **63**, 751–759.
43 Shui,X., McFail-Isom,L., Hu,G.H. and Williams,L.D. (1998) *Biochemistry* **37**, in press.
44 Brünger,A.T. (1992) *X-PLOR, A System for X-ray Crystallography and NMR (Version 3.1)*. Yale University Press, New Haven, CT.
45 Parkinson,G., Vojtechovsky,J., Clowney,L., Brünger,A.T. and Berman,H.M. (1996) *Acta Crystallogr. D* **52**, 57–64.
46 Lavery,R. and Sklenar,H. (1989) *J. Biomol. Struct. Dyn.* **7**, 655–667.
47 Zhang,H., van der Marel,G.A., van Boom,J.H. and Wang,A.H.-J. (1992) *Biopolymers* **32**, 1559–1569.
48 Dunitz,J.D. and Taylor,R. (1997) *Chem. Eur. J.* **3**, 89–98.
49 Teeter,M.M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6014–6018.
50 Harlow,R.L. (1993) *J. Am. Chem. Soc.* **115**, 9838–9839.
51 Egli,M., Portmann,S. and Usman,N. (1996) *Biochemistry* **35**, 8489–8494.