

The role of minor groove functional groups in DNA hydration

Kristen Kruger Woods, Tao Lan¹, Larry W. McLaughlin¹ and Loren Dean Williams*

School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA and
¹Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, MA 02467, USA

Received October 29, 2002; Revised December 11, 2002; Accepted December 24, 2002

ABSTRACT

Here we describe the crystal structure of modified [d(CGCGAATTCGCG)]₂ refined to 2.04 Å. The modification, which affects only the two thymines at the central ApT step, involves isosteric removal of the 2-keto oxygen atoms and substitution of the N1 nitrogen with carbon. The crystal structure reveals the ability of this modified thymine to effectively base pair with adenine in [d(CGCGAAtTCGCG)]₂. The structure also suggests that the minor groove 'spine of hydration' is destabilized but essentially intact.

INTRODUCTION

The stability and conformation of duplex DNA depend on interactions with surrounding water molecules (1) and cations (2). Conversely, affinity and mobility of water molecules (3) and cations (4,5) for DNA depends on surface functional groups of the DNA. The type, proximity and orientation of DNA functional groups vary with the sequence. Therefore, the interactions of water molecules and cations depend on the DNA sequence and conformation (6,7). In the work described here, we have determined some hydration effects resulting from chemical alteration of DNA minor groove functional groups.

The minor groove of A-tract DNA provides a unique chemical environment. The polarity and electronegative potential are high. The floor and walls of the groove are lined with hydrogen bond acceptors but are devoid of hydrogen bond donors. A stabilizing two-tier 'spine of hydration' was postulated from medium and low resolution X-ray diffraction data (8). High resolution data reveals a more organized and extensive hydration pattern (9). An interdependent ensemble of water molecules, four tiers deep by one molecule thick, extends along the minor groove, forming a 'fused hexagon motif'. These water molecules are rotationally and translationally restrained by the imbalance of hydrogen bond donors and the shape of the minor groove. The rotationally and translationally restrained water molecules are polarized, with aligned dipoles. Narrowing of the minor groove induces water polarization, increasing rotational and translation restraints.

It has recently become clear from molecular dynamics (MD) simulations that the geometry and hydration of the A-tract minor groove correlate with the contents of the groove (10–13). These correlations hold if one averages over time or monitors real-time thermal fluctuations. The groove is narrow, with a highly ordered and polarized monolayer of water molecules when cations are contained within the groove. These cations might be localized near the 'floor' of the groove, or near the 'lip', closer to phosphate groups. When the groove is especially narrow, cations in the lip region are stabilized by close proximity of phosphate groups and by interactions with polarized water molecules. When the groove is wide, with less-ordered water molecules and more remote phosphate groups, interactions with cations are less favorable.

In the study presented here, we use X-ray diffraction to examine a DNA dodecamer duplex [d(CGCGAAtTCGCG)]₂ (AAtT) containing two modified A-T base pairs. Functional groups normally present in the A-tract minor groove, the O2 keto oxygens of two T residues, have been removed without altering base pairing properties (Fig. 1). In the modified thymine, called *t*, a functional group that is thought to be a primary determinant of minor groove hydration has been removed. The significance of this keto oxygen and its impact on minor groove hydration is described. The importance of this functionality in helix stability and conformation has been examined previously (14,15). We have done a careful comparison of this modified structure with the high resolution structure of [d(CGCGAATTCGCG)]₂ (AATT) (9,16).

MATERIALS AND METHODS

Crystallization, data collection and reduction, and refinement

The modified dodecamer, dCGCGAAtTCGCG, was synthesized as described previously (14). The ammonium salt of reversed-phase HPLC-purified dodecamer was used for crystallization. Crystals were grown at 22°C from sitting drops containing 1.9 mM d(CGCGAAtTCGCG), 31 mM sodium cacodylate (pH 6.5), 19 mM MgCl₂, 6.9% 2-methyl-2,4-pentanediol (MPD) and 4.9 mM spermine tetrahydrochloride. The droplets were equilibrated by vapor diffusion against a reservoir of 40% MPD. Orthorhombic (*P*2₁2₁2₁) crystals appeared within several days. The crystal chosen for data collection measured 0.3 mm × 0.25 mm × 0.25 mm. The crystal was captured in a cryoloop and shock-cooled by direct

*To whom correspondence should be addressed. Tel: +1 404 894 9752; Fax: +1 404 894 7452; Email: loren.williams@chemistry.gatech.edu

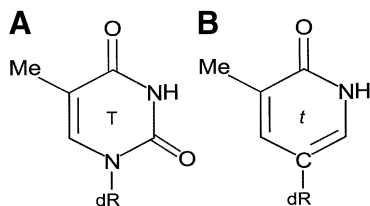


Figure 1. (A) Thymine and (B) *t*. The N1 atom of thymine is replaced with a carbon in *t* and the O2-carbonyl is absent.

transfer into a -110°C N_2 stream. X-ray intensity data were collected at low temperature (-110°C) with a Siemens SMART CCD detector and monochromated Mo $\text{K}\alpha$ radiation ($\lambda = 0.7107 \text{ \AA}$). Data were reduced using the program SAINT. Excluding solvent molecules, coordinates of a previous high resolution structure of $[\text{d}(\text{CGCGAATTCGCG})]_2$ (NDB entry bd1084) were used as the starting model (16). The model was annealed and refined using CNS (version 1.0) (17). The initial rounds of refinement retained the native model (with unmodified DNA). Negative difference peaks centered on the O2 atoms of the central thymine residues (Fig. 2) of the native DNA confirmed the presence of the modified residue in the DNA. After the modification was confirmed, the parameter file dna-rna-multi-endo.param was modified to include the modified thymine of $\text{d}(\text{CGCGAATTCGCG})$ and the refinement was completed by extending the model to include water molecules and ions. CNS was used to calculate Fourier maps. The model was refined against all data between 10 and 2.04 \AA resolution (4393 unique reflections) (Table 1). The final model contains 76 water molecules and one hydrated magnesium ion. The geometry of the final model was characterized with the program CURVES.

RESULTS

Modified thymine (*t*) forms Watson–Crick base pairs with adenine

Figure 2 shows the electron density map of one of the modified thymine–adenine base pairs. In the $t(7)$ -A(18) base pair, the

hydrogen bond distances are 3.02 \AA (N3–N1) and 3.14 \AA (O4–N6). The other modified base pair, $t(19)$ -A6, has slightly longer hydrogen bond distances of 3.13 \AA (N3–N1) and 3.37 \AA (O4–N6). These modified base pairs can be compared with the two corresponding A-T base pairs in the native dodecamer (16). The native T(7)-A(18) pair has hydrogen bond distances of 2.81 \AA (N3–N1) and 3.08 \AA (O4–N6), while the T(19)-A(6) hydrogen bonds are 2.82 \AA (N3–N1) and 3.19 \AA (O4–N6). The hydrogen bonding distances are slightly longer in the modified base pair than in the native A-T base pairs but are within reasonable hydrogen bonding limits. Therefore, the crystal structure illustrates the ability of *t* to form a standard Watson–Crick base pair in the dodecamer $[\text{d}(\text{CGCGAATTCGCG})]_2$.

The modified base *t* alters but does not disrupt first shell hydration in the minor groove

High resolution data collected from crystals of native $[\text{d}(\text{CGCGAATTCGCG})]_2$ reveal what appears to be a co-operative hydration structure within the A-tract minor groove. The native motif, observed at high resolution (9), is a four water molecule deep (tiers denoted by P, S, T and Q) by one water molecule thick ‘fused hexagon’ of hydration sites. The four sites in the primary tier of the monolayer, in contact with the floor of the groove, are denoted P1, P2, P3 and P4. These P sites interact with hydrogen bond acceptors including the N3 atoms of adenines and O2 atoms of thymines, which line the floor, as well as the O4′ atoms of deoxyriboses, which line the walls of the groove. To distinguish the native and modified dodecamers, hydration sites in the modified dodecamer are denoted here by P′, S′, T′ and Q′. The resolution of the data collected for the modified dodecamer is sufficient to allow analysis of the first two hydration tiers (P′, S′) but not the third or fourth tiers (T′, Q′). Similarly the data are not sufficient to observe or infer the sites of monovalent cations.

The P′1 site (of the modified dodecamer) is nearly superimposable on the P1 site (of the native dodecamer). P′1 and P1 both form hydrogen bonds with the O2 of C9 (P′1, 3.1 \AA ; P1, 2.7 \AA) and N3 of A(17) (P′1, 2.7 \AA ; P1, 2.9 \AA). Similarly, the P2 and P′2 sites are nearly superimposable.

Table 1. Data collection and refinement statistics

Unit cell dimensions	$\alpha = \beta = \gamma = 90^{\circ}$ $a = 25.170 \text{ \AA}$, $b = 41.053 \text{ \AA}$, $c = 65.723 \text{ \AA}$
DNA (asymmetric unit)	$[\text{d}(\text{CGCGAATTCGCG})]_2$
Space group	$P2_12_12_1$
Temperature of data collection (K)	163
Number of reflections	30 312
Number of unique reflections	4927
Number of reflections used in refinement ($F > 3\sigma F$)	4393 ^a
$R(\text{int})$ (%) ^b	8.4
RMS deviation of bonds from ideality (\AA)	0.0086
RMS deviation of angles from ideality (\AA)	1.22
Max. resolution of observed reflections (\AA)	1.97
Max. resolution of refinement (\AA)	2.04
Number of DNA atoms	484
Number of water molecules, excluding magnesium first shell	76
Number of magnesium ions plus coordinating water molecules	7
R -factor (%)	20.67 (22.34, including test set)
R -free (%)	28.15

^aThe data are 74% complete at the highest resolution shell (2.04 – 2.11 \AA , 378 reflections).

^b $\sum |F_o^2 - F_c^2| / \sum F_o^2$.

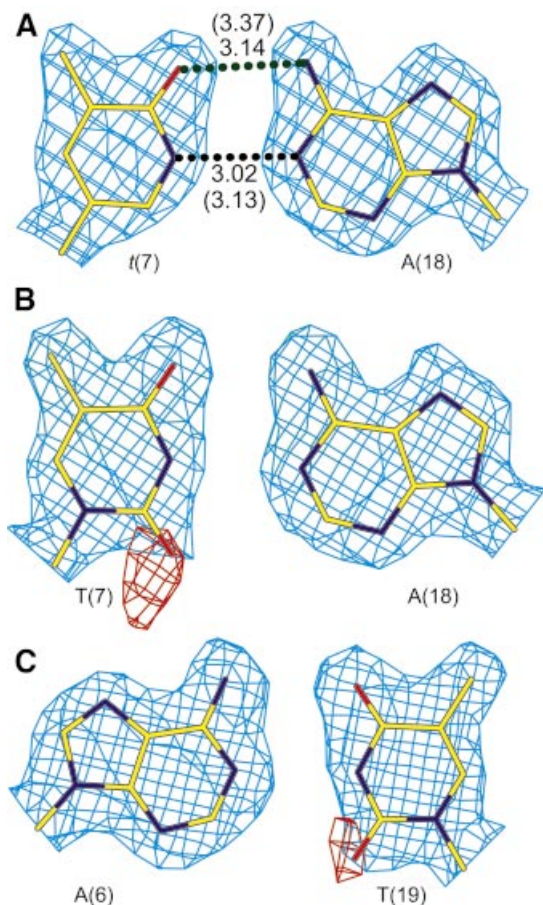


Figure 2. The modified base *t* forms Watson–Crick base pairs with A. (A) The blue net indicates the sum ($2F_o - F_c$) Fourier electron density surrounding the refined *t*(7)–A(18) base pair. The hydrogen bonding distances are indicated. The distances in parentheses are for the equivalent hydrogen bonds of the A(6)–*t*(19) base pair. Although the hydrogen bonding distances are slightly longer than observed for A–T base pairs in the same environment, they are still within normal range. (B) Negative difference ($F_c - F_o$) Fourier electron density map (red net), demonstrating that an A–T base pair does not fit the data. A native [d(CGCGAATTCGCG)]₂ dodecamer model was partially refined against data collected from modified [d(CGCGAATTCGCG)]₂. The negative difference density on the O2 atom of T7 indicates excess scattering by the model. (C) Negative difference density surrounding the O2 of T19, calculated in the same way as given above. Sum maps are blue and are contoured at 1.0σ . Difference maps are red and are contoured at -3.0σ . Distances are in Angstroms.

P'2 and P2 both form hydrogen bonds with the O2 of T8 (P'2, 2.9 Å; P2, 2.9 Å) and N3 of A(17) (P'2, 3.0 Å; P2, 3.0 Å).

In the native structure, the position of the P3 site is highly restrained, by coordination by two O2 atoms (T7 and T19), in addition to two O4' atoms (T8 and T20) and two solvent molecules in the S tier of the hydration monolayer. These interactions combine to restrict the P3 site to the position almost directly in the center of the ApT step, between the planes of the base pairs. The two O2 atoms that coordinate the P3 site in the native dodecamer are absent from the modified dodecamer, shifting the position of the P'3 site significantly. The P'3 site has dropped toward the floor of the groove, or rather the floor has dropped from beneath it, allowing formation of a CH–O hydrogen bond (3.2 Å from C to O) with the exposed C2 atom of *t*7. The P'3 site has also shifted

along the groove. The only hydrogen bond conserved by the P3 and P'3 sites is with the O4' of T(8). The P'3 site is 0.3 Å closer than the P3 site to this O4' (3.0 Å; 3.3 Å). Similarly, P'3 has shifted ~0.3 Å toward P'2 and away from P'4. In contrast to the P sites, the P' sites are unevenly spaced. The distance from P'2 to P'3 is 4.4 Å, while the distance from P'3 to P'4 is 5.5 Å.

The modified base *t* partially disrupts second shell hydration in the minor groove

In the native dodecamer each S site bridges two P sites. This hydration pattern is also observed, in part, in the modified dodecamer. Sites S'1, S'2, S'3 and S'5 correspond closely to sites S1, S2, S3 and S5, respectively. However, no electron density is evident at the S'4 site (Fig. 3A), which would span P'3 and P'4. Either the S'4 site is not occupied, or the potential well at that site is sufficiently broad that the position is not fixed within the ensemble of sites that constitute the crystal. In either interpretation, it appears that alteration of the P' layer, resulting in increased spacing between P'3 and P'4, in comparison with P3 and P4, is transmitted to the next hydration tier, partially disrupting the ordered hydration motif.

The program CURVES (18) was used to analyze the geometry of the modified dodecamer. The largest changes in the modified dodecamer, in comparison with the native structure, are localized in the A-tract and involve helical twist and roll. At base pair A(5)–T(20), the helical twist of AAtT is 8° less than that of AATT, while at A(6)–*t*(19) the helical twist is 6° more than the native. Roll is inversely affected. The roll at A(5)–T(20) of AAtT is 11° higher than that of AATT, while the roll at base pair A(6)–*t*(19) is 5° less than native. Propeller twist is also slightly affected. AAtT A(5)–T(20) is twisted 5° less than native, while A(6)–*t*(19) is twisted 5° more than native. Minor groove width is only slightly changed, with AAtT being 0.3 Å wider throughout the A-tract.

DISCUSSION

McLaughlin and coworkers previously assayed the role of minor groove functional groups on DNA conformation (15) and stability (14) in solution. They synthesized base analogs designed to retain Watson–Crick interactions with opposing bases while lacking minor groove functional groups. From the results of those solution studies they concluded that removal of 2-keto oxygen atoms from T within an A-tract decreases duplex stability and shifts the A/B equilibrium away from B-DNA toward A-DNA. Their results suggest that the effect is cooperative, and that removal of two keto groups from an A-tract sequence embedded in a dodecamer has a moderate effect, while removal of four keto groups has a much more dramatic effect on stability (14).

Here we use X-ray crystallography of a DNA dodecamer to confirm the Watson–Crick interaction of a T-analog that lacks the keto oxygen at the 2-position. Two keto oxygens have been removed from a 4 bp A-tract in the center of the dodecamer. The DNA conformation is not significantly altered by the modification. Minor groove hydration is largely intact in comparison with the control dodecamer. The positions of eight of nine water molecules in the first two hydration layers are conserved (Fig. 3). However, the structure clearly suggests

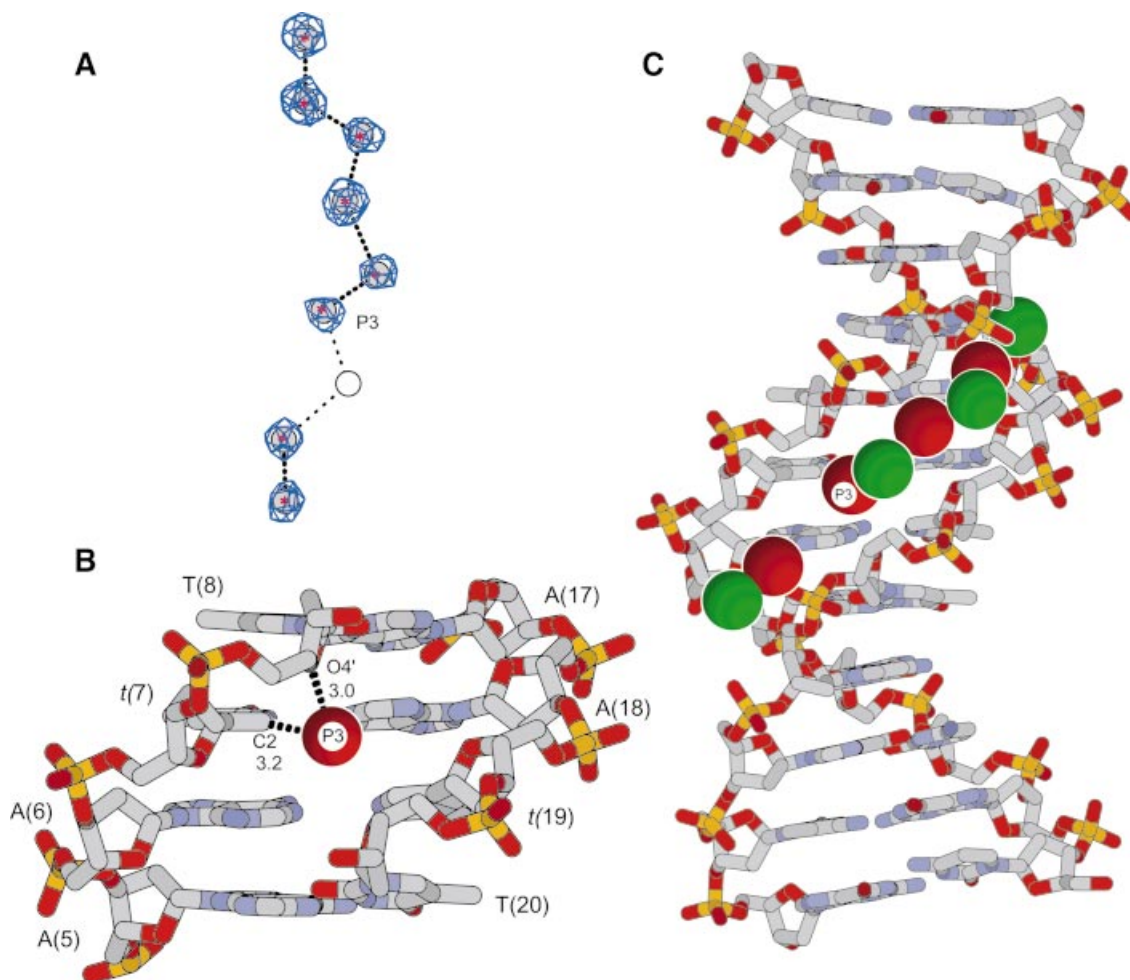


Figure 3. (A) Sum electron density surrounding the spine of hydration. No electron density is observed at one of the sites in the secondary layer, suggesting either disorder or very low occupancy. (B) View into the minor groove of the two modified bases, showing the *t*-A/A-*t* step. Coordination of the P3 site, by O4' and a C2. (C) The minor groove and the spine of hydration. The primary solvent layer is shown in red and the secondary layer is shown in green.

that the hydration scheme is destabilized. The geometric regularity of the hydration scheme is compromised. One hydration site appears vacant, probably because the locations of adjacent water molecules are shifted. An ApT step (the P3 site in the native structure) can contain monovalent cations (4,9,16,19–22). We believe cation occupancy would be less probable in the modified structure, which lacks coordinating carbonyl oxygens at the P3 site. However, the quality of the data here are not high enough to differentiate monovalent cations from water.

Solution studies of the modified versus the native dodecamer indicate that, based upon CD spectra (data not shown), both the native and the modified dodecamer exhibit transitions consistent with a B-form conformation. In comparison, an earlier study indicated that the introduction of two modified A-T base pairs, where both the N3-nitrogen of the A, and the O2-carbonyls of the T had been removed, resulted in a sequence that was not clearly A-form or B-form in character. Replacement of the entire (AATT)₂ core with similarly modified derivatives resulted in a spectra indicative of a largely A-like helical conformation. Solution studies of the doubly modified dodecamer (AAtt)₂ also suggest a shift away

from B-form conformation. The double modification, in which both thymine residues of d(CGCGAATTCGCG) have been substituted with the modified base *t*, lowers the duplex T_m sufficiently so that the DNA exists predominantly as the monomolecular hairpin. Attempts to crystallize the doubly modified dodecamer have not been successful, possibly because the spine of hydration and minor groove environment are severely altered.

The results described here explain previous solution results, which indicate that removing two keto groups in A-tracts is destabilizing but does not alter conformation. Similarly, these modifications destabilize, but do not disrupt, hydration. It has been suggested from X-ray structures that solvent sites within an A-tract minor groove are interdependent and that hydration is cooperative (9). Each occupied solvent site stabilizes those around it and each layer stabilizes the others. That interpretation is supported here. The solvent sites in the center of the A-tract are maintained by surrounding solvent sites, even though stabilizing interactions with the DNA are absent.

Water and associated monovalent metal ions located in the minor groove of A-tract sequences appear to be critical to drive duplex DNA to adopt the B-form conformation. Without

the minor groove functional groups of the A and T residues in the duplex, the equilibrium is shifted from the B to the A conformation.

ACKNOWLEDGEMENTS

The authors thank Nick Hud for helpful discussions. This work was supported by grants to L.W.M. (NSF MCB 0077667) and to L.D.W. (NSF MCB 9976498).

REFERENCES

- Franklin,R.E. and Gosling,R.G. (1953) Molecular configuration in sodium thymonucleate. *Nature*, **171**, 740–741.
- Anderson,C.F. and Record,M.T.,Jr (1995) Salt-nucleic acid interactions. *Annu. Rev. Phys. Chem.*, **46**, 657–700.
- Liepinsh,E., Leupin,W. and Otting,G. (1994) Hydration of DNA in aqueous-solution—NMR evidence for a kinetic destabilization of the minor-groove hydration of d-(TTAA)(2) versus d-(AAATT)(2) segments. *Nucleic Acids Res.*, **22**, 2249–2254.
- Hud,N.V., Sklenar,V. and Feigon,J. (1999) Localization of ammonium ions in the minor groove of DNA duplexes in solution and the origin of DNA A-tract bending. *J. Mol. Biol.*, **286**, 651–660.
- Denisov,V.P. and Halle,B. (2000) Sequence-specific binding of counterions to B-DNA. *Proc. Natl Acad. Sci. USA*, **97**, 629–633.
- Hud,N.V. and Polak,M. (2001) DNA–cation interactions: the major and minor grooves are flexible ionophores. *Curr. Opin. Struct. Biol.*, **11**, 293–301.
- McFail-Isom,L., Sines,C. and Williams,L.D. (1999) DNA structure: cations in control? *Curr. Opin. Struct. Biol.*, **9**, 298–304.
- Drew,H.R. and Dickerson,R.E. (1981) Structure of a B-DNA dodecamer. III. Geometry of hydration. *J. Mol. Biol.*, **151**, 535–556.
- Shui,X., Sines,C., McFail-Isom,L., VanDerveer,D. and Williams,L.D. (1998) Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry*, **37**, 16877–16887.
- Feig,M. and Pettitt,B.M. (1999) Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J.*, **77**, 1769–1781.
- Hamelberg,D., McFail-Isom,L., Williams,L.D. and Wilson,W.D. (2000) The flexible structure of DNA: ion dependence of minor-groove structure and dynamics. *J. Am. Chem. Soc.*, **122**, 10513–10520.
- Hamelberg,D., Williams,L.D. and Wilson,W.D. (2001) Influence of the dynamic positions of cations on the structure of the DNA minor groove: sequence-dependent effects. *J. Am. Chem. Soc.*, **123**, 7745–7755.
- Hamelberg,D., Williams,L.D. and Wilson,W.D. (2002) The effect of neutralized phosphate backbone on the minor groove structure of B-DNA. *Nucleic Acids Res.*, **30**, 3615–3623.
- Lan,T. and McLaughlin,L.W. (2000) Minor groove hydration is critical to the stability of DNA duplexes. *J. Am. Chem. Soc.*, **122**, 6512–6513.
- Lan,T. and McLaughlin,L.W. (2001) Minor groove functional groups are critical for the B-form conformation of duplex DNA. *Biochemistry*, **40**, 968–976.
- Shui,X., McFail-Isom,L., Hu,G.G. and Williams,L.D. (1998) The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry*, **37**, 8341–8355.
- Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S., Read,R.J., Rice,L.M., Simonson,T. and Warren,G.L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **54**, 905–921.
- Stofer,E. and Lavery,R. (1994) Measuring the geometry of DNA grooves. *Biopolymers*, **34**, 337–346.
- Howerton,S.B., Nagpal,A. and Williams,L.D. (2002) Surprising roles for electrostatic interactions in DNA-ligand complexes. *Biopolymers*, in press.
- Woods,K., McFail-Isom,L., Sines,C.C., Howerton,S.B., Stephens,R.K. and Williams,L.D. (2000) Monovalent cations sequester within the A-tract minor groove of [d(CGCGAATTCGCG)]₂. *J. Am. Chem. Soc.*, **122**, 1546–1547.
- Young,M.A., Jayaram,B. and Beveridge,D.L. (1997) Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.*, **119**, 59–69.
- Bartenev,V.N., Golovamov,Eu.I., Kapitonova,K.A., Mokulskii,M.A., Volkova,L.I. and Skuratovskii,I.Y. (1983) Structure of the B DNA cationic shell as revealed by an X-ray diffraction study of CsDNA. Sequence-specific cationic stabilization of B form DNA. *J. Mol. Biol.*, **169**, 217–234.