The EMBO Journal Vol.17 No.9 pp.2709–2718, 1998

Hydration and recognition of methylated CpG steps in DNA

Claudine Mayer-Jung, Dino Moras and Youri Timsit[1,2]

Laboratoire de Biologie Structurale IGBMC CNRS/INSERM/ULP, Parc d'Innovation, 1, rue Laurent Fries, Illkirch 67404, France

[1]Present address: IBPC 13, rue Pierre et Marie Curie, Paris 75005, France

[2]Corresponding author

The analysis of the hydration pattern around methylated CpG steps in three high resolution (1.7, 2.15 and 2.2 Å) crystal structures of A-DNA decamers reveals that the methyl groups of cytosine residues are well hydrated. In comparing the native structure with two structurally distinct forms of the decamer d(CCGCCGGCGG) fully methylated at its CpG steps, this study shows also that in certain structural and sequence contexts, the methylated cytosine base can be more hydrated that the unmodified one. These water molecules seem to be stabilized in front of the methyl group through the formation C–H•••O interactions. In addition, these structures provide the first observation of magnesium cations bound to the major groove of A-DNA and reveal two distinct modes of metal binding in methylated and native duplexes. These findings suggest that methylated cytosine bases could be recognized by protein or DNA polar residues through their tightly bound water molecules.

Keywords: A-DNA/CpG/cytosine methylation/hydration

Introduction

DNA methylation at CpG steps plays an essential role in the regulation of the genetic functions of eukaryotic cells. It is now established that CpG methylation is strictly required for normal embryonic development in mice (Li *et al.*, 1992). Methylation exerts its action at various levels of the control of gene expression, from the stable inactivation of entire chromosomes to the fine tuning of spatial and temporal expression of tissue-specific genes (Cedar, 1988; Tate and Bird, 1993). Recent studies have revealed some indications about the mode of action of cytosine methylation: DNA methylation can directly prevent the binding of transcription factors or act through the intermediate of methylated CpG-binding proteins (Tate and Bird, 1993). Another mode of action on gene repression is the induction of the formation of inactive higher order structures of chromatin (Lewis and Bird, 1991). However, the detailed mechanisms by which the addition of one methyl group on cytosine residues influences the DNA structure and function are poorly understood. A complex and somewhat paradoxical picture has emerged from the structural studies on methylated DNA. Although it is well established that methylation of CpG steps can affect DNA structure strongly in promoting helical transitions from B to Z (Behe and Felsenfeld, 1981; Fujii *et al.*, 1982) or B to A (Frederick *et al.*, 1987; Mooers *et al.*, 1995; Tippin *et al.*, 1997), the comparison of native and methylated crystal structures adopting the same helical form revealed only minor differences (Fujii *et al.*, 1982; Frederick *et al.*, 1987; Heineman and Hahn, 1992). Moreover, recent biochemical experiments also indicated that methylation does not influence DNA flexibility (Hodges-Garcia and Hagerman, 1995).

In this study, the question of how cytosine methylation influences the binding of water molecules around modified CpG steps is addressed in the context of recent findings establishing the role of DNA hydration in intermolecular recognition (reviewed in Shwabe, 1997) and in the stabilization of secondary and tertiary structures of DNA (reviewed in Westhof and Beveridge, 1990; Berman, 1994). A current view is that the methyl group of cytosine consists of a hydrophobic protuberance in the major groove acting mainly in preventing the binding of water molecules (Frederick *et al.*, 1987; Ho *et al.*, 1988; Tippin *et al.*, 1997). In a recent work, however, it has been shown that cytosine methylation does not prevent DNA–DNA recognition by groove–backbone interaction. By contrast, the two methyl groups of the modified CpG step of one helix were found in an appropriate geometry for clamping the phosphate group of another helix, establishing C–H···O hydrogen bonds with its anionic oxygen atoms (Mayer-Jung *et al.*, 1997). In showing that methyl groups can interact with polar atoms, these finding prompted us to examine further the hydration of methyl groups in our new crystallographic structures.

Here, the hydration patterns of the high resolution structures of the native and methylated palindromic A-DNA decamer d(CCGCCGGCGG) were compared. The sequence d(Cm^{5}CGCm^{5}CGGm^{5}CGG) methylated at the three CpG steps was crystallized in two space groups corresponding to distinct helical conformations. These structures provide experimental data on water molecules bound around methylated CpG steps in different sequence and structural contexts. In contrast to previous reports, we found that the methylation does not interfere with the major groove hydration of 5-methylcytosines. In some cases, the methyl groups provide new hydration sites and promote the binding of a larger number of water molecules than the corresponding unmodified sequences. Furthermore, despite quite different conformations, the two methylated structures share a quasi-identical magnesium cation-binding site. These results are analyzed in terms of the influence of the sequence and structures on the hydration pattern.

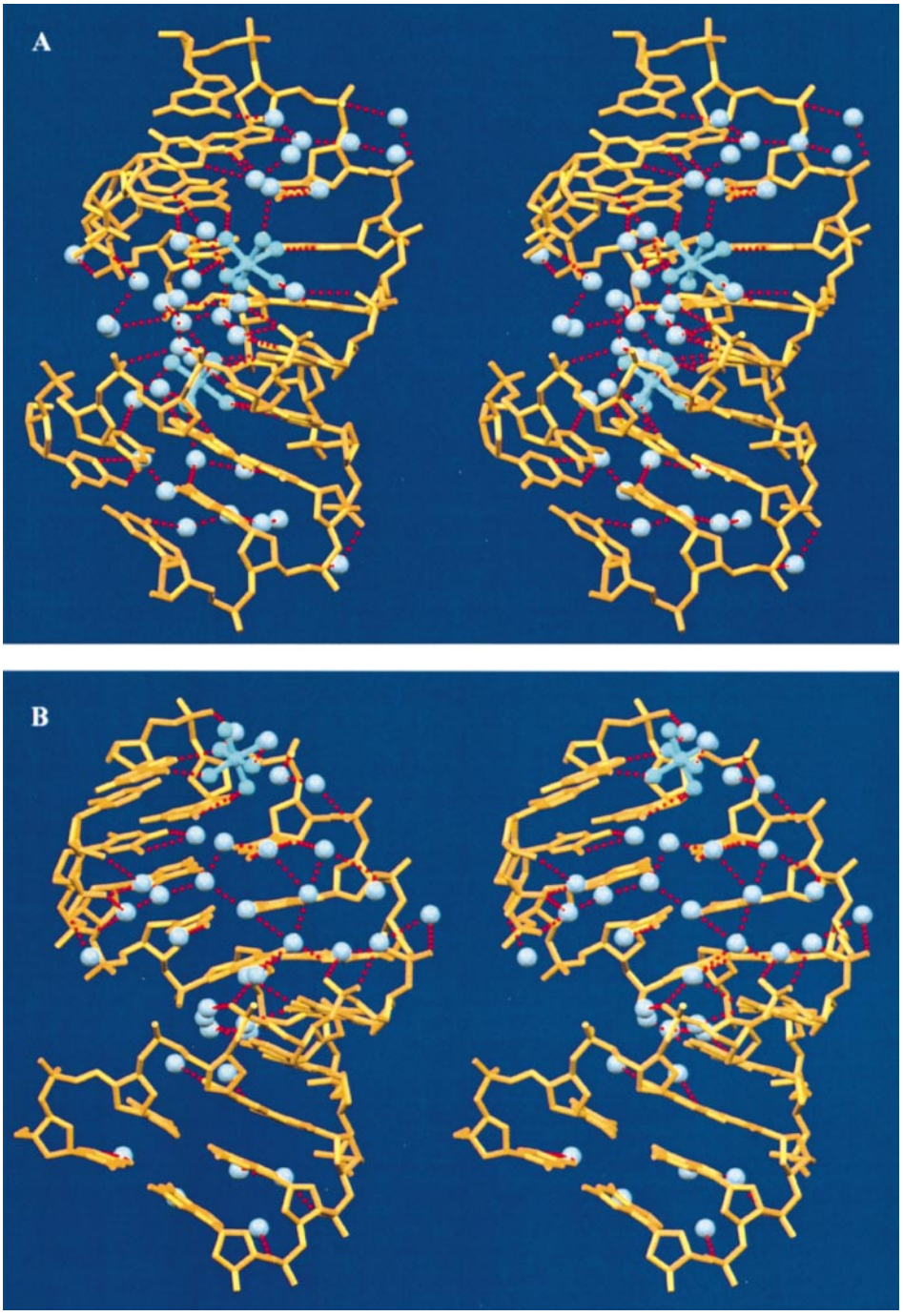
Results

Overall hydration features

The native and methylated sequences exhibit extremely different structural features. The structure of the native

decamer which crystallizes in $P6_522$, a new packing for an A-DNA decamer, is very regular and closely resembles fiber A-DNA (r.m.s. = 0.71 Å for 404 atoms) (Table I; Figure 1A). By contrast, the methylated decamer crystallizes in two different space groups, the orthorhombic $P2_12_12_1$ and the hexagonal $P6_1$, corresponding to two distinct conformations. Both forms are observed within the same crystallization drop. The orthorhombic form which is characterized by pronounced alterations and irregularities represents one of the most distorted A-DNA duplexes observed to date (Table I; Figure 1B). The structure exhibits large variations in the helical parameters and a severe bending of the helical axis (18°) between base 5 and 6, as reflected by the large roll angle. In $P6_1$, the methylated duplex displays a more regular and extended conformation (Table I, Figure 1C). The average r.m.s. deviation between the two methylated structures is 2.2 Å. A detailed comparison of the structures will be reported elsewhere.

With a well hydrated major groove and a poorly hydrated minor groove, the overall hydration pattern of the three structures resembles those of other A-DNA duplexes (Table I). However, the three decamers are less hydrated than other A-DNA duplexes in which an average of 10 water molecules are reported per base pair (Tippin and Sundaralingam, 1997b). Although their crystal forms have different volumes per base pair and water content, the two methylated decamers share a similar degree of hydration, with an average of nine solvent molecules per base pair (Table I). They also have in common a quasi-



identical magnesium ion-binding site located at the duplex end. In the slightly less hydrated native structure, two symmetry-related magnesium ions were found in the center of the major groove.

The views shown in Figure 1 illustrate the differences in the geometry and hydration pattern of the three structures. This structure becomes much less regular in passing from the native form to the most distorted methylated orthorhombic form. The structural alterations correlate with a significant increase in the helical diameter and a greater variability in the distances between the amino groups (N4–N4) at the CpG steps (Tables I and III). As a consequence, while the water molecules of the major groove are densely packed into the central channel of the native duplex, they appear more sparse in the modified structures (Figure1A–C).

Hydration of methylated CpG steps

In agreement with other works (Schneider *et al.*, 1993; Eisenstein and Shakked, 1995), two major groove hydration sites were found in the plane of the unmodified cytosine bases. As depicted on Figure 2A, one water molecule (wA) interacts with the N4 amino group while the other (wB) is linked to the anionic oxygen atom of

Table I. Overall structural characteristics and hydration of the three A-DNA native and methylated double helices

	Methylated *nae* $Cm^5CGCm^5CGGm^5CGG$		*naeA* CCGCCGGCGG
	Orthorhombic	Hexagonal	
Minor groove			
width (Å)	9.7	9.7	10.1
depth (Å)	0.8	1.2	0.9
Major groove			
width (Å)	4.5	8.3	2.7
depth (Å)	11.8	10.1	9.6
Helix diameter (Å)	20.1	19.8	18.5
P_n–P_{n+10} distances (Å)			
$n = 2$	5.1	9.1	4.7
$n = 3$	5.4	9.8	4.0
V/bp ($Å^3$/bp)	1270	1983	1429
No. of hexacoordinated Mg ions	1	1	2
No. of water molecules			
Total	79	80	64
First hydration shell	63	67	40
Major groove	19 (24%)	20 (25%)	16 (25%)
$N_{solvent}$/bp	8.8	8.9	7.8

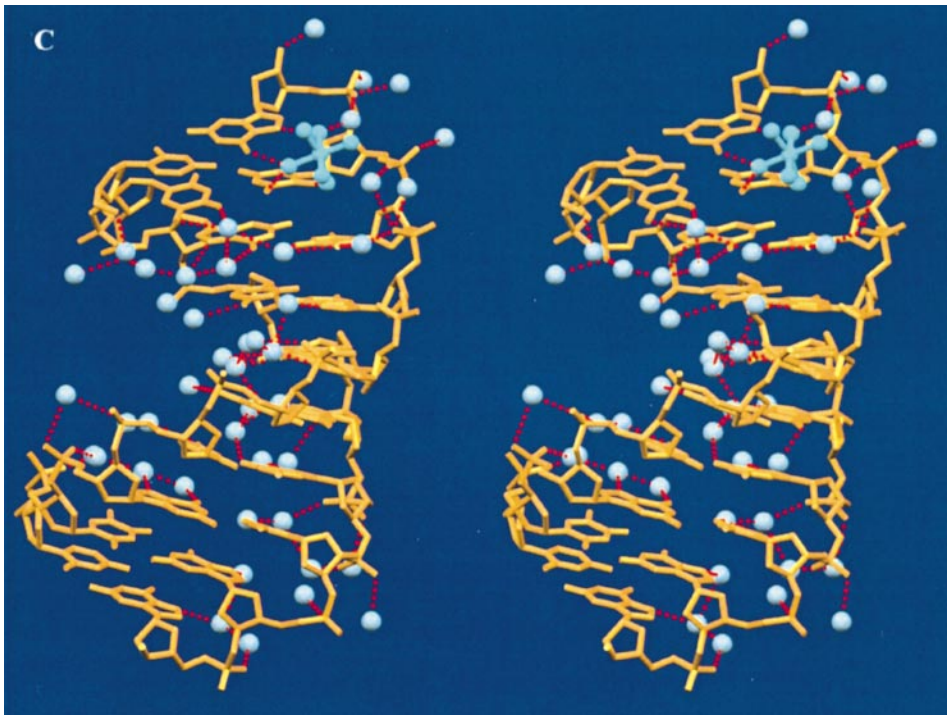


Fig. 1. Comparison of the overall major groove and backbone hydration patterns of the native and methylated A-DNA decamers. Frontal stereo view of the native decamer (**A**), the methylated orthorhombic decamer (**B**) and the methylated hexagonal decamer (**C**). Water molecules are indicated in light blue and hexacoordinated magnesium ions in cyan. Hydrogen bonds are represented by red dashed lines.

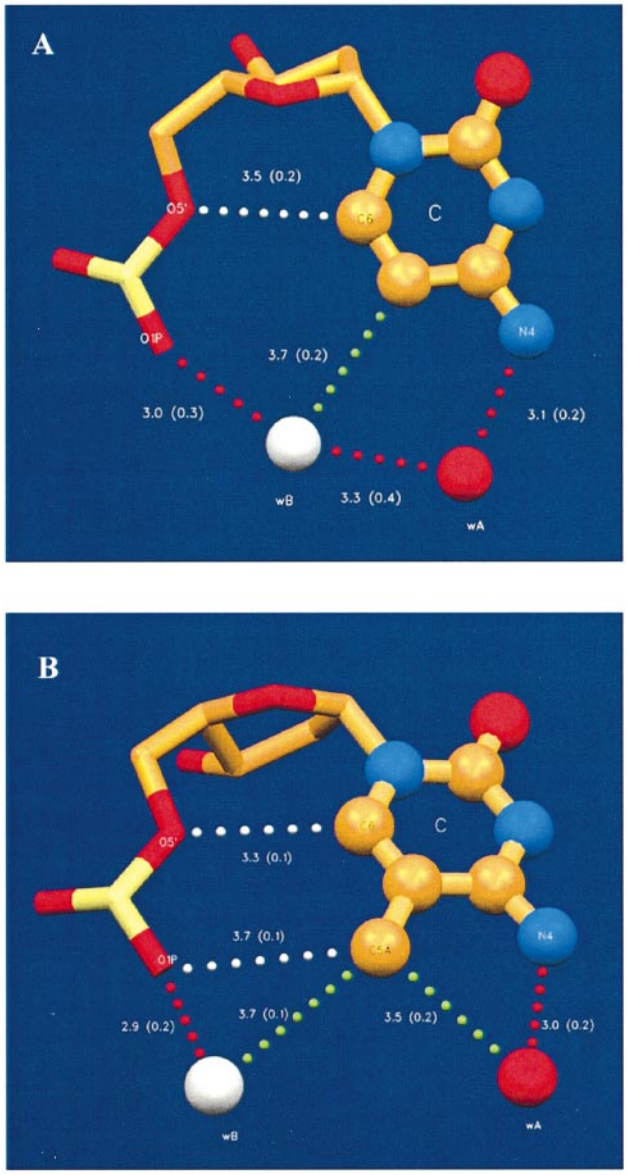


Fig. 2. Comparison of the geometry of hydration of a cytosine base (**A**) and a 5-methylated cytosine base (**B**). The water molecule interacting with the N4 amino group (wA) is represented in red while the other (wB) linked to the anionic oxygen atom of the phosphate is indicated in white. Conventional hydrogen bonds are represented by dashed red lines, C–H···O hydrogen bonds between the methyl group (C5A) and the water molecules are represented by dashed green lines, and other C–H···O hydrogen bonds are in white.

the phosphate group. This latter molecule is placed at a correct distance to form a C–H···O bond with the C5–H hydrogen of the cytosine base. In the methylated structures, this pattern is mainly conserved, with slight changes in the distances and angles (Table II, Figure 2B). This indicates that, in most cases, the methyl group does not prevent the hydration of the modified cytosine base. On the contrary, the methyl group appears to stabilize the binding of the water molecules wA and wB further through the formation of additional C–H···O hydrogen bonds. As indicated in Table II and Figure 2B, the majority of the distances C5A–wA and C5A–wB are between 3.4 and 4 Å and are thus compatible with this kind of interaction (Wahl and Sundaralingam, 1997). As observed in the Z-DNA structure of the methylated d(CGCGCG) hexamer (Fujii *et al.*, 1982), the methyl group has displaced the wA water molecule towards the N4 amino group and opened up the angle C5–N4–wA. The major features of the hydration geometry of the CpG steps in the three structures are described in Tables II and III.

In addition, new major groove hydration sites were found to be correlated with the presence of the methyl group of some m^5CpG sequences (Figure 3A and B). The water molecules were observed just in front of the methyl group at a distance too great for interacting with other donor or acceptor atoms of the DNA molecule. As indicated in Figure 3A and B, these water molecules are referred to wA′ or wB′ depending on their proximity to the wA and wB sites, respectively. For example, in the hexagonal form, the wA′ and wB′ water molecules are located in front of the methyl group of m^5C5 and m^5C18 residues, respectively (Figure 3A and B). These additional solvent molecules are also stabilized through the formation of hydrogen bond networks with their neighboring wA and wB solvent molecules. A particular helical conformation also contributes to organize solvent networks in front of the methylated bases: due to the high twist angle at step 2–3, the two wA water molecules bound to the N4 amino groups of m^5C2 and m^5C18 interact with each other (Figure 3A). Since in all cases the geometry of hydration is suitable for establishing C–H···O hydrogen bonds between the water molecules and the methyl group, it can be concluded that C–H···O bonds and solvent–solvent interactions are mainly responsible for the stabilization of these water molecules. At these two well-hydrated m5CpG steps 2–3 and 5–6, the double helix exhibits large helical twist angles and moderately high positive roll values (Table III). However, these deviations from regularity cannot be attributed to the methylation and/or the change in hydration since the native duplex shares similar helical parameters at the corresponding steps, with a different hydration pattern. Note that the geometry of step 2–3 is reminiscent of that observed in the B-DNA dodecamer methylated at its central CpG step where the anionic oxygen atoms of the phosphate group of a symmetry-related molecule occupy a similar position in front of the alternate N4 amino groups (Figure 3A; Figure 3 of Mayer-Jung *et al.*, 1997).

The step C5–G6 of the orthorhombic structure is characterized by a very high positive roll, rise and propeller twist, and a severe kink of the helical axis (Figure 1 and Table III). The methyl group of the m^5C5 methylcytosine base seems to participate in the stabilization of the distorted conformation by binding two additional water molecules. On both sides of the cytosine base plane they form a solvent network which bridges the adjacent residues C4 and G7 (Figure 3C). The methyl group points into a cone of three polar atoms including the two water molecules and the anionic oxygen atom of the phosphate group of the m^5C5 residue. This geometry is reminiscent of that observed around consecutive methylated adenine and thymine bases and a magnesium cation in the trigonal structure of a B-DNA decamer (see Figure 10 of Baikalov *et al.*, 1993).

Influence of cytosine methylation on magnesium binding

In the native decamer, two symmetry-related hexacoordinated magnesium ions bind to both DNA chains across the

Table II. Geometry of hydration (distances and angles) of the unmodified and methylated CpG steps. Cytosine residues showing this geometry are also indicated

Atom 1	Atom 2	Distance (Å)[a]	Angle (°) C4–N4–wA	Residues
Cytosine				
DNA–water				
N4	wA	3.1 (0.2)	122 (9)	ortho C(1), C(4),
O1P	wB	3.0 (0.3)		C(11), C(14)
C5	wB	3.7 (0.2)		hexa C(4), C(14)
DNA–DNA				*naeA* C(1), C(2), C(4),
C6	O5′	3.5 (0.2)		C(5), C(8)
Water–water				
wA	wB	3.3 (0.4)		
Methylated cytosine				
DNA–water				
N4	wA	3.0 (0.2)	139 (5)	ortho C(2), C(8),
O1P	wB	2.9 (0.2)		C(15), C(18)
C5A	wA	3.5 (0.2)		hexa C(2), C(15)
C5A	wB	3.7 (0.1)		C(8), C(12)
DNA–DNA				
C5A	O1P	3.7 (0.1)		
C6	O5′	3.3 (0.1)		

[a]Mean values of observed distances; mean value deviation in parentheses.

Table III. Helical parameters and hydration pattern of the CpG steps of the native and methylated structures

	Step CpG	Twist (°)	Rise (Å)	Roll (°)	Propeller twist (°)	N4–N4 distance (Å)	C5A–C5A distance (Å)	Hydration site
Ortho	C(2)G(3)	24	2.5	3	–5, –3	6.1	10.8	AB/A
	C(5)G(6)	30	3.0	18	–14, –15	4.9	9.9	BB′/A
	C(8)G(9)	32	2.1	8	–9, –4	5.4	10.3	A/–
Hexa	C(2)G(3)	34	2.8	9	–12, –12	5.1	9.7	AB/ABB′
	C(5)G(6)	30	2.9	11	–15, –6	5.7	10.4	AA′B/A
	C(8)G(9)	25	3.5	11	–12, –6	6.0	11.0	AB/B
naeA	C(2)G(3)	29	2.4	9	–13, –7	5.4	–	A/A
	C(5)G(6)	29	2.6	11	–10, –10	5.3	–	AB/AB
	C(8)G(9)	29	2.4	9	–7, –13	5.4	–	A/A
Methylated *ras*	C(6)G(7)	40	3.4	6	–10, –12	3.7	6.9	–

major groove. Four water molecules of their hydration sphere interact directly with the N7 and O6 atoms of the guanine bases G6, G7 and G13 on one side, and G16, G17 and G3 on the other side of the 2-fold axis (Figure 4A). In the methylated decamers, the magnesium-binding site is displaced to the duplex extremity. In both structures, the cation bridges the two adjacent guanine bases G19 and G20, using three of the water molecules of its coordination sphere. Although the two methylated forms exhibit quite distinct conformations, the binding site and the geometry of interaction are strikingly conserved (Figure 4B and C). A noticeable difference from the native structure is that the magnesium cation establishes direct contacts only with a single DNA strand. It is difficult, however, to explain why methylation has caused the displacement of the magnesium-binding site, from the center of the major groove to the duplex end, since in the methylated decamers, the cation lies in close proximity to the methyl group of the m^5C2 base (Figure 4B and C). Moreover, the major groove geometry of the native structure remains suitable for receiving a hexacoordinated magnesium ion, even if the cytosine bases C5 and C15 are methylated.

Discussion

In this study, the hydration pattern of the native A-DNA decamer duplex d(CCGCCGGCGG) was compared with that of its methylated analogs crystallized in two distinct conformations. This analysis revealed that, in most cases, cytosine methylation does not prevent the binding of the solvent molecules which are normally observed in front of the unmodified base. This study has also shown that cytosine methylation correlates with the presence of new hydration sites. In these sites, which are absent at the corresponding steps of the native decamer duplex, the water molecules appear to interact exclusively with the methyl groups through C–H···O interactions, according to geometric criteria. It seems reasonable, therefore, to conclude that cytosine methylation could be responsible, in some cases, for a better major groove hydration of m^5CpG steps. These findings are in agreement with

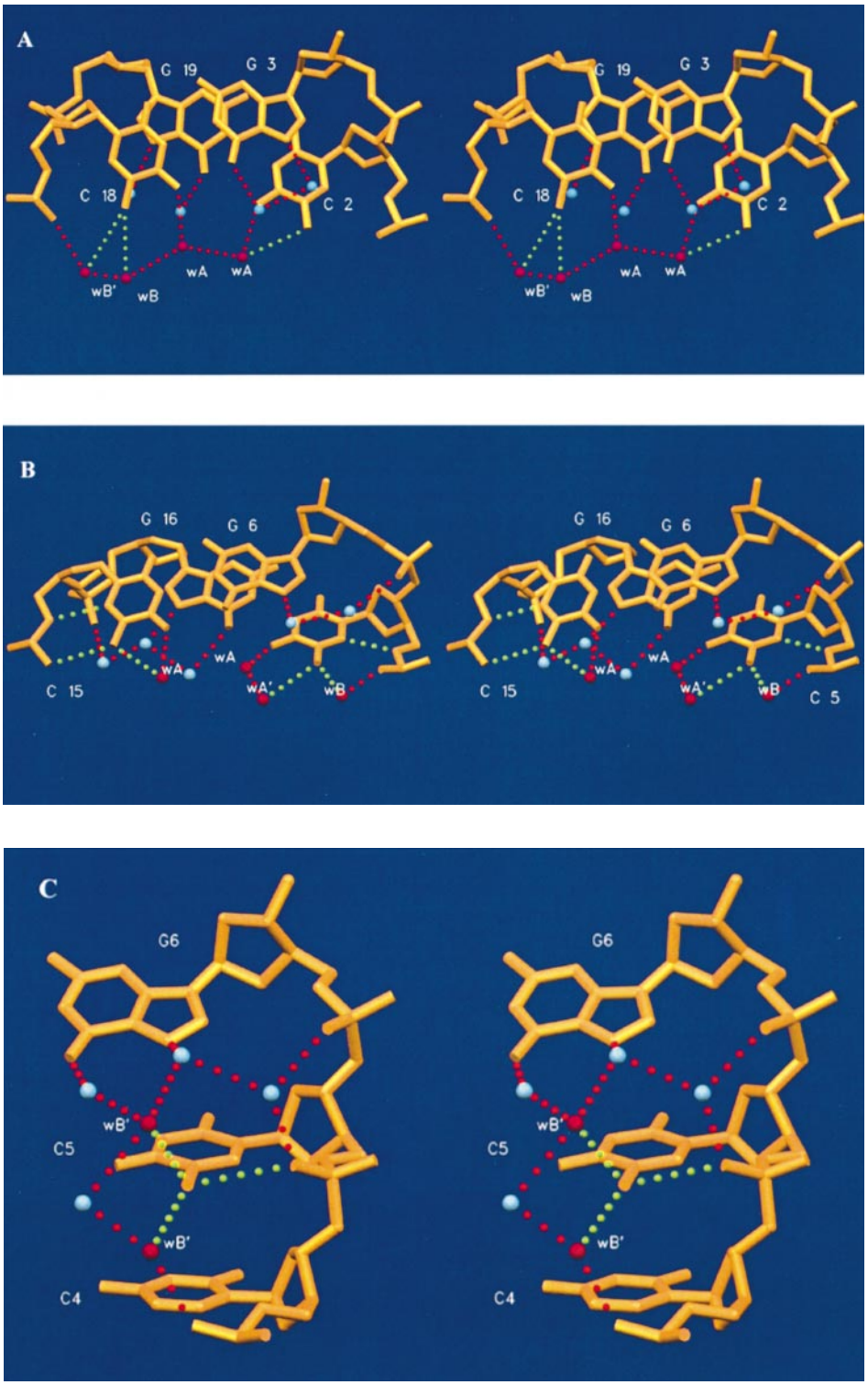


Fig. 3. New hydration sites in the methylated structures. The methylated CpG steps 2–3 (**A**) and 5–6 (**B**) of the hexagonal methylated decamer and 5–6 of the orthorhombic decamer (**C**). Additional water molecules are named wA′ and wB′ depending on their proximity to the wA and wB sites.

previous reports of well-hydrated methyl groups in nucleic acids (Drew and Dickerson, 1981; Westhof and Sundaralingam, 1986; Baikalov *et al.*, 1993). As previously suggested for the hydration of thymine residues (Drew and Dickerson, 1981), methyl groups can stabilize water molecules that interact with a phosphate oxygen. The existence of C–H···O interactions is now established (Taylor and Kennard, 1982) and their role in the stabilization of the structure of biological macromolecules was the subject of recent reviews (Derewenda *et al.*, 1995; Wahl and Sundaralingam, 1997). This kind of interaction was also found to be involved in the binding of water molecules in the crystal structures of small molecules (Steiner and Saenger, 1993). Thus, our study suggests a role for these interactions in the hydration of methylated bases in nucleic acids.

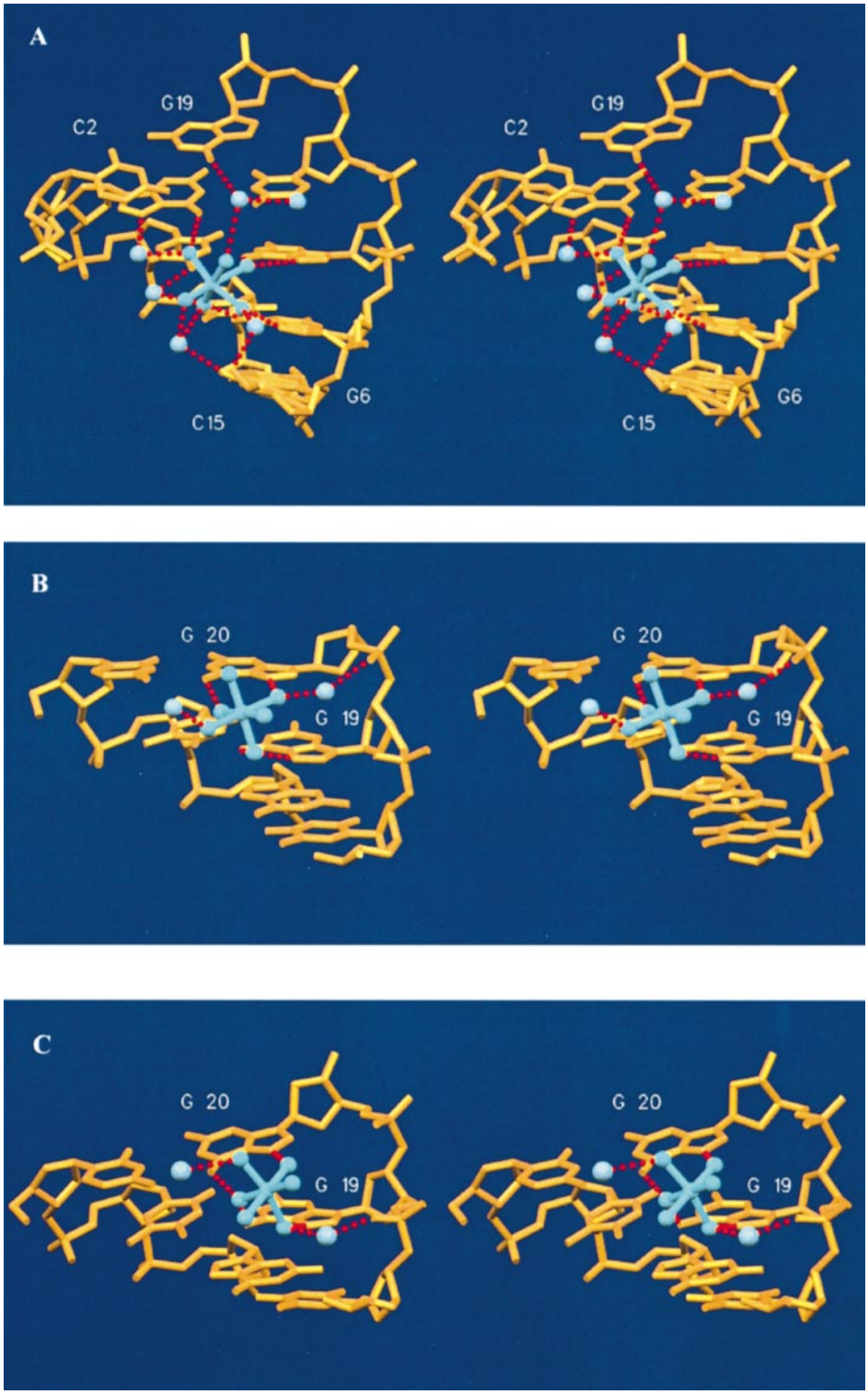


Fig. 4. Comparison of the magnesium-binding sites of the native (**A**) and methylated orthorhombic (**B**) and hexagonal (**C**) decamers. Water molecules are shown in light blue and hexacoordinated magnesium ions in cyan. Hydrogen bonds are represented by red dashed lines.

The finding that methylated cytosine residues can be hydrated is, however, in sharp contrast to previous crystallographic studies which report that the methyl groups create hydrophobic pockets in the floor of the major groove which hinder the approach of solvent molecules (Frederick *et al.*, 1987; Ho *et al.*, 1988; Tippin *et al.*, 1997). At the present stage of our work, this discrepancy is difficult to explain. Recent studies have indicated that unmodified CpG steps can also be extremely poorly hydrated in A-DNA and that their hydration can be influenced by the sequence and structure of the double helix (Eisenstein and Shakked, 1995; Tippin and Sundaralingam, 1997b). In a comparative study of the hydration patterns of 10 isomorphous tetragonal A-DNA octamers, it was

Table IV. Crystallization conditions, data collection and refinement statistics

	Methylated *nae* $Cm^5CGCm^5CGGm^5CGG$		*naeA* CCGCCGGCGG
	Orthorhombic	Hexagonal	
Crystallization conditions and data collection			
Crystallization conditions			
Temperature	20°C	4°C	16°C
pH	6	6	6
DNA (mM)	0.5	0.5	0.5
Mg (mM)	4	4	7
Mg/o	8	8	14
Spermine (mM)	0.4	0.4	0.5
MPD (%)	35	27	30
Space group	$P2_12_12_1$	$P6_1$	$P6_522$
Cell parameters (Å)	$a = 23.9$	$a = b = 54.9$	$a = b = 41.5$
	$b = 45.3$		
	$c = 46.9$	$c = 45.6$	$c = 57.5$
Resolution (Å)	1.7	2.15	2.2
No. of unique reflections	5804	4238	1510
R_{sym} (%)	3.8	4.4	5.6
Completeness (%)	94.4	97.2	88.9
Multiplicity	4.6	5.1	4.3
Refinement statistics			
No. of DNA atoms	410	410	202
No. of non-DNA atoms	80	81	33
No. of reflections (7.0–R_{max}) used in the refinement [$F_{obs} > 2\sigma(F_{obs})$]	4973	3827	1448
R-factor (%)	19.6	16.6	19.7
R.m.s.d. from ideal distances (Å)			
Bonds lengths	0.011	0.011	0.006
Angles	1.4	1.4	1.0
Torsion angles	27	27	26
Impropers	1.3	1.6	1.3

proposed that high twist angles, negative rolls and CG base pairs promote major groove hydration (Tippin and Sundaralingam, 1997b). It is possible that the hydration of methylated CpG steps also depends on the structural context and follows similar rules. The well-hydrated methylated CpG steps 2–3 and 5–6 of the hexagonal structure indeed display high twist angles, but positive rolls (Table III). However, in the methylated orthorhombic structure, step 2–3 is more hydrated than step 8–9 despite the lower value of its twist angle (Table III). Moreover, it appears clear that in the native structure, the CpG steps which are characterized by uniformly distributed high twist angles are less hydrated than the corresponding modified ones in the two methylated structures. This suggests that if an influence of the helical parameters (and mainly the twist angle) could contribute to overcome the 'negative' effect of the methyl group on hydration, other factors are responsible for the better hydration of the methylated CpG steps in our structures. It should be noted that these two modified decamers represent the most densely methylated structures analyzed to date. The high density of methyl groups pointing into the major groove could play a role in ordering water through a clathrate-like mechanism. In addition, each methylated form displays extreme structural features. While the orthorhombic form is the most distorted A-DNA structure analyzed to date, the very extended hexagonal form has a tendency to resemble to a B-DNA double helix. It could be suggested that these extreme features are also responsible for their unusual hydration properties.

In providing the first observation of hexacoordinated magnesium cations bound to the major groove of A-DNA duplexes, this study brings new insights into the influence of DNA methylation on magnesium binding. Two modes of magnesium binding were observed in the native and methylated structures. In the native structure, the hexacoordinated magnesium cations bridge the two DNA chains across the major groove at GGC sequences. This binding mode was observed previously in B-DNA in a similar sequence (Timsit and Moras, 1995, 1996), indicating that the determinants of magnesium–DNA recognition are conserved among the two right-handed DNA structural forms. By contrast, in the methylated decamers, magnesium ions only interact with a single DNA strand by bridging two consecutive guanine bases. Surprisingly, the two modified duplexes share a similar magnesium binding-site despite quite different conformations. Other cations such as cobalt hexammine have been shown to bind GpG steps in A-DNA in a closely related manner (Gao *et al.*, 1995; Nunn and Neidle, 1996). For the magnesium cation, similar binding modes were found in large RNA molecules and in the recent crystal structure of the 5S rRNA (Correl *et al.*, 1997). This study indicates that cytosine methylation has contributed indirectly to preventing the formation of interstrand magnesium bridges. Generalization of these findings would suggest that DNA methylation can interfere

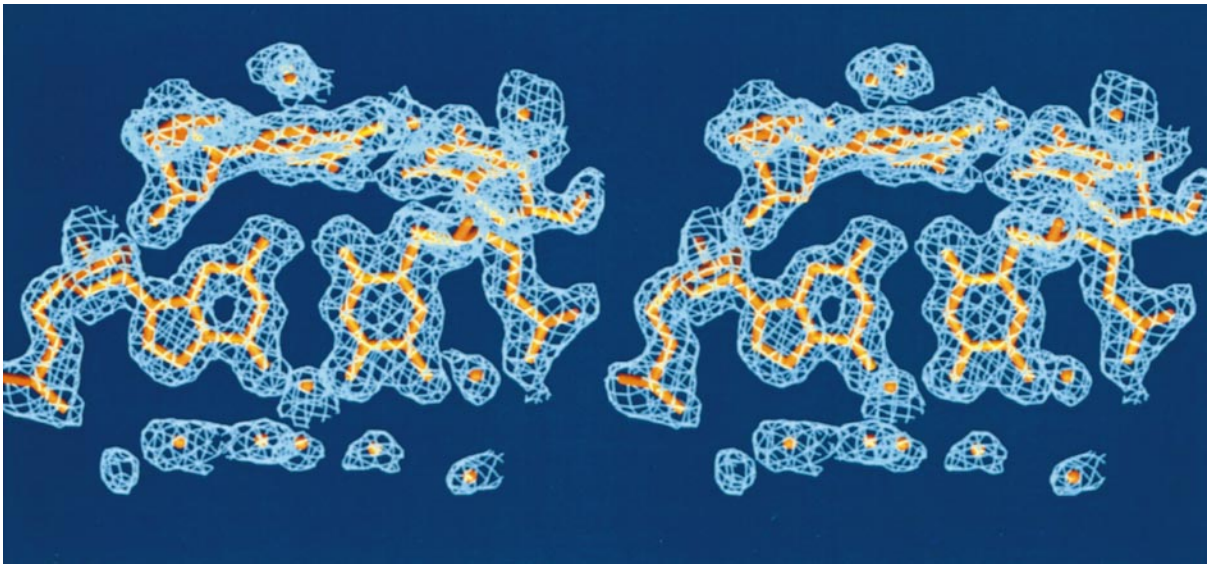


Fig. 5. Final $3F_o–2F_c$ map contoured at 1σ level using the 4973 reflections [$F_{obs} > 2\sigma(F_{obs})$] between 7 and 1.7 Å, superimposed on the final model. Part of the symmetry-related molecule is also shown.

with the ability of magnesium cations to stabilize the secondary structure of the double helix.

Conclusion

Our study has shown that, in certain structural and sequence contexts, methylated cytosine residues can be well hydrated, despite the 'hydrophobic' character of the methyl group. The careful analysis of the hydration geometry in three high resolution structures of A-DNA decamers has revealed that the methyl groups can indeed stabilize the binding of additional water molecules via the formation of C–H···O interactions. In the light of recent studies which have underlined the role of water molecules in protein–DNA recognition processes (Otwinowski *et al.*, 1988; Shakked *et al.*, 1994; Schwabe, 1997), this finding leads to the prediction that methylated cytosine residues could be recognized by polar residues, either directly or through the intermediate of tightly bound water molecules. This view is consistent with recent work which has shown the role of cytosine methylation in DNA–DNA recognition processes (Mayer-Jung *et al.*, 1997). In conclusion, this study suggests that the molecular mechanism through which cytosine methylation exerts its function involves more than the simple addition of a hydrophobic hindrance in the major groove of DNA.

Materials and methods

Crystallization and data collection

The native and methylated strands d(CCGCCGGCGG) and d(Cm^5-CGCm^5CGGm^5CGG) were synthesized by the phosphoramidite method and the crystallization was performed by the vapor diffusion technique. Single hexagonal crystals of the native duplex were grown at 16°C from hanging drops containing equal volumes of a 0.5 mM dodecamer duplex solution and a reservoir solution, 50 mM sodium cacodylate buffer (pH 6), 0.5 mM spermine, 7 mM magnesium acetate and 30% 2-methyl 2,4-pentanediol. Orthorhombic and hexagonal crystal forms of the methylated duplex were grown from the same drop under slightly different conditions (Table IV). The $P6_522$ hexagonal crystals of the native decamer diffracted to 2.2 Å resolution while the $P6_1$ hexagonal and $P2_12_12_1$ orthorhombic crystals of the methylated form diffracted at 2.15 and 1.7 Å resolution, respectively (Table IV). The diffraction data of the native duplex crystals were collected at 0°C using a Siemens area detector and Cu Kα from a rotating anode operated at 40 kV and 80 mA, and were processed using the XENGEN program package. The data sets of the methylated forms were collected at 0°C with a MarResearch imaging plate detector mounted on the W32 beamline of the LURE synchrotron (λ = 0.9 Å). The data were processed using the MarXDS program package (Kabsch, 1988).

Molecular replacement and refinement

Since isomorphous crystal structures were not available during this work, the structures of $P6_522$ and $P6_1$ hexagonal decamers were solved by molecular replacement using the program AMoRe in the range 9.0–4.0 Å resolution (Navazza, 1994), with fiber coordinates of A and B double helices as starting models. $P6_522$ represents a new space group for A-DNA decamers. The asymmetric unit consists of a single DNA strand. The search for the correct position was therefore restricted to a combination of a rotation–translation of the model around and along one of the two hexagonal unit cell 2-fold axes, with the helix oriented in such way that its dyad axis corresponds with the crystallographic 2-fold axis. While the search was performed in the two space groups $P6_122$ and $P6_522$, a clear solution appeared in $P6_522$ using the coordinates of a standard A-DNA decamer. The structure of $P6_1$ methylated decamer was solved using the coordinates of the octamer d(GGTATACC) (Shakked *et al.*, 1983) extended on both ends to form a decamer. The search procedure with other types of DNA models systematically failed. In the course of this study, another structure of an A-DNA decamer crystallized in $P6_1$ was reported (Tippin and Sundaralingam, 1997a). The orthorhombic methylated decamer was solved by using the isomorphous structure of d(CCCGGCCGGG) (Ramakrishnan and Sundaralingam, 1993) as the starting model.

The program X-PLOR (Brünger, 1992) was used for the refinement of the three structures, giving final *R*-factors of 19.7, 19.6 and 16.6% using the reflections between 7 Å and the limit of the resolution of the native, orthorhombic and hexagonal methylated decamers, respectively (Table IV; Figure 5). Solvent molecules were added at positive peaks over 3σ in difference Fourier ($F_o–F_c$) maps, provided they made sensible hydrogen bonds with the DNA molecules or other solvent atoms. Water molecules with a *B*-factor >40 were rejected. Addition of solvent was ended as soon as the agreement *R*-factors converged. A total of 33, 80 and 81 solvent molecules were placed in the native, orthorhombic and hexagonal methylated models, respectively (Table I). Magnesium ions were assigned based on the hexacoordinated geometry criteria of the observed solvent peaks. *B*-factors of all the solvent molecules, including those around the methyl groups of the methylated cytosines, are in the range 15–40 $Å^2$. The helical parameters of the three structures have been calculated using the program NewHelix93 (Dickerson, 1993) and their atomic coordinates have been deposited in the Nucleic Acid Database (Berman *et al.*, 1992).

Acknowledgements

We thank Zippi Shakked, Richard Lavery and Hugues Bertrand for helpful comments and discussions, and André Mitschler for his contribution to data collection.

References

Baikalov,I., Grzeskowiak,K., Yanagi,K., Quintana,J. and Dickerson,R. (1993) The crystal structure of the trigonal decamer CGTCG6meA-TCG: a B-DNA helix with 10.6 base pairs per turn. *J. Mol. Biol*., **231**, 768–784.

Behe,M. and Felsenfeld,G. (1981) Effects of methylation on a synthetic polynucleotide: the B–Z transition in poly(dG-m5dC)·polyd(dG-m5dC). *Proc. Natl Acad. Sci. USA*, **78**, 1619–1623.

Berman,H. (1994) Hydration of DNA: take 2. *Curr. Opin. Struct. Biol*., **4**, 345–350.

Berman,H., Olson,W., Beveridge,D., Westbrook,J., Gelbin,A., Demeny,T., Hsieh,S.-H., Srinivasan,A. and Schneider,B. (1992) The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.*, **63**, 751–759.

Brünger,A. (1992) *X-PLOR Version 3.0: A System for Crystallography and NMR*. Yale University Press.

Cedar,H. (1988) DNA methylation and gene activity. *Cell*, **53**, 3–4.

Correll,C., Freeborn,B., Moore,P. and Steitz,T. (1997) Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*, **91**, 705–712.

Derewenda,Z.,S., Lee,L. and Derwenda,U. (1995) The ocurrence of C–H···O hydrogen bonds in proteins. *J. Mol. Biol.*, **252**, 248–262.

Dickerson,R. (1993) NewHelix 1993.

Drew,R. and Dickerson,R. (1981) Structure of a B-DNA dodecamer. III Geometry of hydration. *J. Mol. Biol*., **151**, 535–556.

Eisenstein,M. and Shakked,Z. (1995) Hydration patterns and intermolecular interactions in A-DNA crystal structures. Implications for DNA recognition. *J. Mol. Biol*., **248**, 662–678.

Frederick,C., Saal,D., van der Marel,G., van Boom,J., Wang,A. and Rich,A. (1987) The crystal structure of d(GGm^{5}CCGGCC): the effect of methylation on A-DNA structure and stability. *Biopolymers*, **26**, S145–S160.

Fujii,S., Wang,A., van der Marel,G., van Boom,J. and Rich,A. (1982) Molecular structure of (m^{5}dC-dG)$_3$: the role of the methyl group on 5-methyl cytosine in stabilizing Z-DNA. *Nucleic Acids Res.*, **10**, 7879–7892.

Gao,Y.-G., Robinson,H., van Boom,J. and Wang,A. (1995) Influence of counter-ions on the crystal structures of DNA decamers: binding of $[Co(NH_3)_6]^{3+}$ and Ba^{2+} to A-DNA. *Biophys. J.*, **69**, 559–568.

Heineman,U. and Hahn,M. (1992) C-C-A-G-G-C-m5C-T-G-G helical fine structure, hydration and comparison with C-C-A-G-G-C-C-T-G-G. *J. Biol. Chem*., **267**, 7332–7341.

Ho,P., Quigley,G., Tilton,R. and Rich,A. (1988) Hydration of methylated and nonmethylated B-DNA and Z-DNA. *J. Phys. Chem*., **92**, 939–945.

Hodges-Garcia,Y. and Hagerman,P. (1995) Investigation of the influence of cytosine methylation on DNA flexibility. *J. Biol. Chem.*, **270**, 197–201.

Kabsch,W. (1988) Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr*., **21**, 67–71.

Lewis,J. and Bird,A. (1991) DNA methylation and chromatin structure. *FEBS Lett.*, **285**, 155–159.

Li,E., Bestor,T. and Jaenisch,R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.

Mayer-Jung,C., Moras,D. and Timsit,Y. (1997) Effect of cytosine methylation at CpG steps on DNA–DNA recognition. *J. Mol. Biol.*, **270**, 328–335.

Mooers,B., Schroth,R., Baxter,W. and Ho,P. (1995) Alternating and non-alternating dG–dC hexanucleotides crystallize as canonical A-DNA. *J. Mol. Biol*., **249**, 772–784.

Navazza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Nunn,C. and Neidle,S. (1996) The high resolution crystal structure of the DNA decamer d(AGGCATGCCT). *J. Mol. Biol*., **256**, 340–351.

Otwinoski,Z., Schevitz,R., Zhang,R.G., Lawson,C., Joachimiak,A., Marmostein,R., Luisi,B. and Sigler,P. (1988) Crystal structure of the trp repressor/operator complex at atomic resolution. *Nature*, **335**, 321–329.

Ramakrishnan,B. and Sundaralingam,M. (1993) High resolution structure of the A-DNA decamer d(CCCGGCCGGG). *J. Mol. Biol*., **231**, 431–444.

Schneider,B., Cohen,D. Schleifer,L. Srinivasan,A., Olson,W. and Berman,H. (1993) A systematic method for studying the spatial distribution of water molecules around nuclec acid bases. *Biophys. J*., **65**, 2291–2303.

Schwabe,J. (1997) The role of water in protein–DNA interactions. *Curr. Opin. Struct. Biol*., **7**, 126–134.

Shakked,Z., Rabinovich,D., Kennard,O., Cruse,W., Salisbury,S. and Viswamitra,M. (1983) Sequence dependent conformation of an A-DNA double helix: the crystal structure of the octamer d(GGTATACC). *J. Mol. Biol*., **166**, 183–201.

Shakked,Z., Guzikevich-Guerstein,G., Frolow,F., Rabinovich,D. Joachimiak,A. and Sigler,P. (1994) Determinants of repressor–operator recognition from the structure of the trp operator binding site. *Nature*, **368**, 469–473.

Steiner,T. and Sanger,W. (1993) Role of C–H···O hydrogen bonds in the coordination of water molecules. Analysis of neutron diffraction data. *J. Am. Chem. Soc*., **115**, 4540–4547.

Tate,P. and Bird,A. (1993) Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet. Dev.*, **3**, 226–231.

Taylor,R. and Kennard,O. (1982) Crystallographic evidence for the existence of C–H···O, C–H···N and C–H···Cl hydrogen bonds*J. Am. Chem. Soc*., **104**, 5063–5070.

Timsit,Y. and Moras,D. (1995) Self-fitting and self-modifying properties of the B-DNA molecule. *J. Mol. Biol*., **251**, 629–647.

Timsit,Y. and Moras,D. (1996) Cruciform structures and functions. *Q. Rev. Biophys*., **29**, 279–307.

Tippin,D. and Sundaralingam,M. (1997a) Nine polymorphic crystal structures of d(CCGGGCCCGG), d(CCGGGCCm5CGC), d(Cm5-CGGGCCm5CGG) and d(CCGGGCC(Br)5CGG) in three different conformations: effects of spermine binding and methylation on the bending and condensation of A-DNA. *J. Mol. Biol*., **267**, 1171–1185.

Tippin,D. and Sundaralingam,M. (1997b) Comparison of major groove hydration in isomorphous A-DNA octamers and dependence on base sequence and local helix geometry. *Biochemistry*, **36**, 536–543.

Tippin,D., Ramakrishnan,B. and Sundaralingam,M. (1997) Methylation of the Z-DNA decamer d(GC)5 potentiates the formation of A-DNA: crystal structure of d(Gm5CGm5CGCGCGC). *J. Mol. Biol*., **270**, 247–258.

Wahl,M. and Sundaralingam,M. (1997) C–H···O hydrogen bonding in biology. *Trends. Biochem. Sci*., **22**, 97–102.

Westhof,E. and Beveridge,D. (1990) Hydration of nucleic acids. *Water Sci. Rev.*, **5**, 24–123.

Westhof,E. and Sundaralingam,M. (1986) Restrained refinement of the monoclinic form of yeast phenylalanine transfer RNA. Temperature factors and dynamics, coordinated waters, and base-pair propeller twist angle. *Biochemistry*, **25**, 4868–4878.

Received January 21, 1998; revised and accepted March 3, 1998