# Structure and Hydration of *Bam*HI DNA Recognition Site: A Molecular Dynamics Investigation

T. Castrignanò,* G. Chillemi,† and A. Desideri*

*INFM and Department of Biology, University of Rome "Tor Vergata," 00133 Rome; and †CASPUR, Supercomputing Center for University and Research, c/o Università "La Sapienza," 00185 Rome, Italy

ABSTRACT The results of a 3-ns molecular dynamics simulation of the dodecamer duplex d(TATGGATCCATA)$_2$ recognized by the *Bam*HI endonuclease are presented here. The DNA has been simulated as a flexible molecule using an AMBER force field and the Ewald summation method, which eliminates the undesired effects of truncation and permits evaluation of the full effects of electrostatic forces. The starting B conformation evolves toward a configuration quite close to that observed through x-ray diffraction in its complex with *Bam*HI. This configuration is fairly stable and the Watson-Crick hydrogen bonds are well maintained over the simulation trajectory. Hydration analysis indicates a preferential hydration for the phosphate rather than for the ester oxygens. Hydration shells in both the major and minor groove were observed. In both grooves the C-G pairs were found to be more hydrated than A-T pairs. The "spine of hydration" in the minor groove was clear. Water residence times are longer in the minor groove than in the major groove, although relatively short in both cases. No special long values are observed for sites where water molecules were observed by x-ray diffraction, indicating that water molecules having a high probability of being located in a specific site are also fast-exchanging.

## INTRODUCTION

Water plays an important role in modulating the structural and functional properties of biological macromolecules, and its interaction with both proteins and nucleic acids is the subject of an intense investigation from both an experimental and a simulative approach. The surface hydration of proteins depends on the geometry and quality of surface groups: waters within polar cavities exchange more slowly and can be detected by x-ray diffraction, while nonpolar cavities do not usually contain long-lasting water molecules (Schwabe, 1997; Kuhn et al., 1992; Luise et al., 2000). In the case of nucleic acids it has been shown that their hydration is directly related to their conformation (Westhof, 1988, 1993; Rentzeperis et al., 1993; Chalikian et al., 1994; Jayaram and Beveridge, 1996; Cheatham and Kollman, 1997a). Information on the water position around a DNA sequence is provided by x-ray diffraction, which reveals the favored average position occupied by a water molecule. The observation of highly localized water molecules forming a "spine of hydration" in the minor groove of the B-DNA dodecamer (CGCGAATTCGCG) (Drew and Dickerson, 1981; Kopka et al., 1983; Subramanian et al., 1988, 1990) was successfully reproduced using static DNA models (Subramanian and Beveridge, 1989; Berman, 1994) and later by using a dynamic model for several DNA sequences (Duan et al., 1997; Young et al., 1997; Feig and Pettitt, 1999).

A network of water molecules bridging the side-chain protein atoms and the DNA bases have also been observed in the x-ray diffraction pattern of protein-DNA systems, such as the complex between trp repressor and trp operator or between the *Bam*HI restriction enzyme and its specific DNA sequence (Shakked et al., 1994; Newman et al., 1994, 1995). These and other related works have raised the question of whether interfacial water molecules play an important role in mediating sequence-specific recognition and in enhancing the protein affinity for specific DNA sequences (Schwabe, 1997).

An accurate structural and dynamical description of DNA in aqueous solution, helping to understand the molecular basis of protein-DNA and drug-DNA recognition, can be obtained through molecular dynamics simulations. Realistic simulations of DNA molecules are currently possible because of three methodological improvements: the inclusion of hydration and ionic patterns in their environment, the introduction of Ewald summation technique (Cheatham et al., 1995, 1997; Young and Beveridge, 1998) to account for the undesired effects of the truncation method, and the use of an accurate force field. The dynamical properties of the water molecules around the trp operator have been recently studied by MD and NMR spectroscopy (Bonvin et al., 1998; Sunnerhagen et al., 1998). Both the simulative and experimental study indicated that water molecules observed by x-ray diffraction are not characterized by long residence time.

Here we make use of the most recent Amber force field (Cornell et al., 1995) and of the particle mesh method (Darden et al., 1993) to investigate through MD simulation the dynamical properties of water around the DNA duplex interacting with the *Bam*HI endonuclease. The results indicate that hydration is characterized by relatively short water residence times at any DNA site. However, evidence is

*Received for publication 24 February 2000 and in final form 5 June 2000.*

Address reprint requests to A. Desideri, Department of Biology, University of Rome, "Tor Vergata," Via della Ricerca Scientifica, 00133 Rome, Italy. Tel.: +39-06-72594376; Fax: +39-06-72594326; E-mail: desideri@uniroma2.it.



0006-3495/00/09/1263/10 $2.00

presented for a spine of hydration in the DNA minor groove, indicating that sites having a high probability of being hydrated do not necessarily trap water molecules for a long time.

## COMPUTATIONAL METHODS

### System setup

The canonical B-DNA structure (Arnott et al., 1976) was generated for the DNA sequence d(TATGGATCCATA)$_2$ using the nucgen module included in the AMBER 5.0 software package. The B conformation was chosen because x-ray diffraction has shown that in the DNA-*Bam*HI endonuclease complex (Newman et al., 1995) the DNA duplex retains a B-like conformation.

The DNA macromolecule was immersed in a rectangular box filled with TIP3P (Jorgensen et al., 1981) imposing a minimum solute-wall box distance of 10 Å. The system was neutralized with $Na^+$ cations using the AMBER leap module. The total system is composed of 762 DNA atoms, 22 $Na^+$ counterions, and 3376 water molecules, giving a total of 10,912 atoms.

### Dynamics simulation protocol

The system was modeled using the AMBER95 all-atom force field (Cornell et al., 1995) with periodic boundary conditions. A cutoff radius of 9 Å was used for nonbonded interactions, updating the neighbor pair list every 10 steps. The electrostatic interactions were calculated with the Particle Mesh Ewald method (Darden et al., 1993; Essmann et al., 1995). The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain all bond lengths involving hydrogen. Optimization and relaxation of solvent and ions were initially performed while keeping the solute atoms constrained to their initial position with decreasing force constants of 500, 25, 15, and 5 kcal/mol Å$^2$. Thereafter the system was minimized without any constraints, and warmed up. A 3-ns simulation was carried out at constant temperature (Berendsen et al., 1984) of 300 K and at a constant pressure of one atmosphere with a 2-fs time step. Pressure and temperature coupling constants were 0.5 ps.

### Analysis

During the production run, from 0.5 to 3 ns, the atom coordinates were saved every 0.1 ps for analysis. The AMBER carnal module was used to check some structural properties (root-mean-square deviation (rmsd), hydrogen bond). The hydrogen bond criterion was a maximum donor-acceptor distance of 3.5 Å and a minimum donor-proton-acceptor angle of 120°. The mean and maximum water residence times and the coordination number were calculated using the algorithm described by Garcia and Stiller (1993), considering a radius sphere of 3.5 Å around DNA sites and a time resolution of 0.1 or 10 ps, respectively. The solvent-accessible surface area of DNA sites was evaluated with the program NACCESS (Hubbard and Thornton, 1993), using a default probe size of 1.4 Å. DNA geometry parameters were calculated with CURVES (Lavery and Sklenar, 1988, 1989).

## RESULTS AND DISCUSSION

### Evaluation of the energetic contributions

A molecular dynamic simulation of the DNA oligonucleotide d(TATGGATCCATA)$_2$, entirely solvated and neutralized with $Na^+$ ions, was performed for 3 ns using the AMBER95 all-atom force field. Fig. 1, *A* and *B* show the electrostatic and van der Waals energies versus simulation time of DNA-DNA, DNA-ions, and DNA-water interaction, respectively. Both the electrostatic and van der Waals potential energies for the DNA-DNA, DNA-water, and DNA-ions interactions were all fairly stable over the 3 ns of simulation. The equilibration of the system is dominated by the DNA-solvent and DNA-ions electrostatic interactions, which show relatively large fluctuations, greater than those observed in a previous MD simulation, carried out with the Gromos force field for a DNA sequence corresponding to the Trp operator (Bonvin et al., 1998). This may be due to the fact that the partial charges attributed to the atoms in the Gromos force field are lower than those attributed in AMBER, which was used in this work (Cornell et al., 1995). The DNA-DNA electrostatic interaction is almost negligible over all the trajectory. The van der Waals potential energies are negligible for what concerns the DNA-ion interaction, while the DNA-

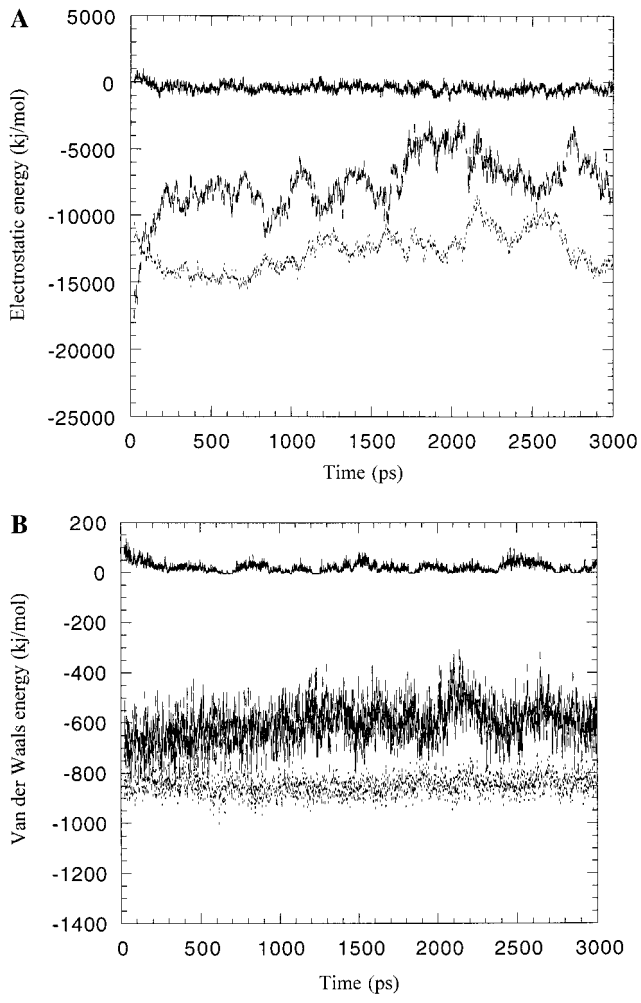


FIGURE 1 Non-bonded energies as a function of simulation time of the DNA duplex interacting with *Bam*HI. (*A*) Electrostatic energies; (*B*) van der Waals energies.

water and mainly the DNA-DNA interactions have an important role for the internal stability of the double helix.

Because of the above consideration the system can be considered fully equilibrated after 300–400 ps. To have a reliable sampling of our system, the analysis concerning its structural and hydration properties have been carried out for the last 2.5 ns of our trajectory, i.e., the time range between 0.5 and 3.0 ns.

## Structural analysis

The stability of our system during the MD trajectory has been checked by monitoring specific structural parameters as a function of time. The rmsd calculated for all atoms from the B-DNA, the A-DNA, the average molecular dynamics, and the crystallographic x-ray structure (PDB entry 2bamhi), excluding the first two and the last two basepairs of the DNA tract, are reported in Fig. 2, *A* and *B*, respectively. The deviation from B-DNA increases with time, whereas it decreases from A-DNA (Fig. 2 *A*), although over all the trajectory the simulated structure remains closer to that of B-DNA. The rmsd from the average molecular dynamics structure (Fig. 2 *B*, *bottom curve*) oscillates around 1 Å, indicating that the DNA settles in a well-defined and stable configuration during the simulation. The rmsd from the crystallographic x-ray structure of the DNA

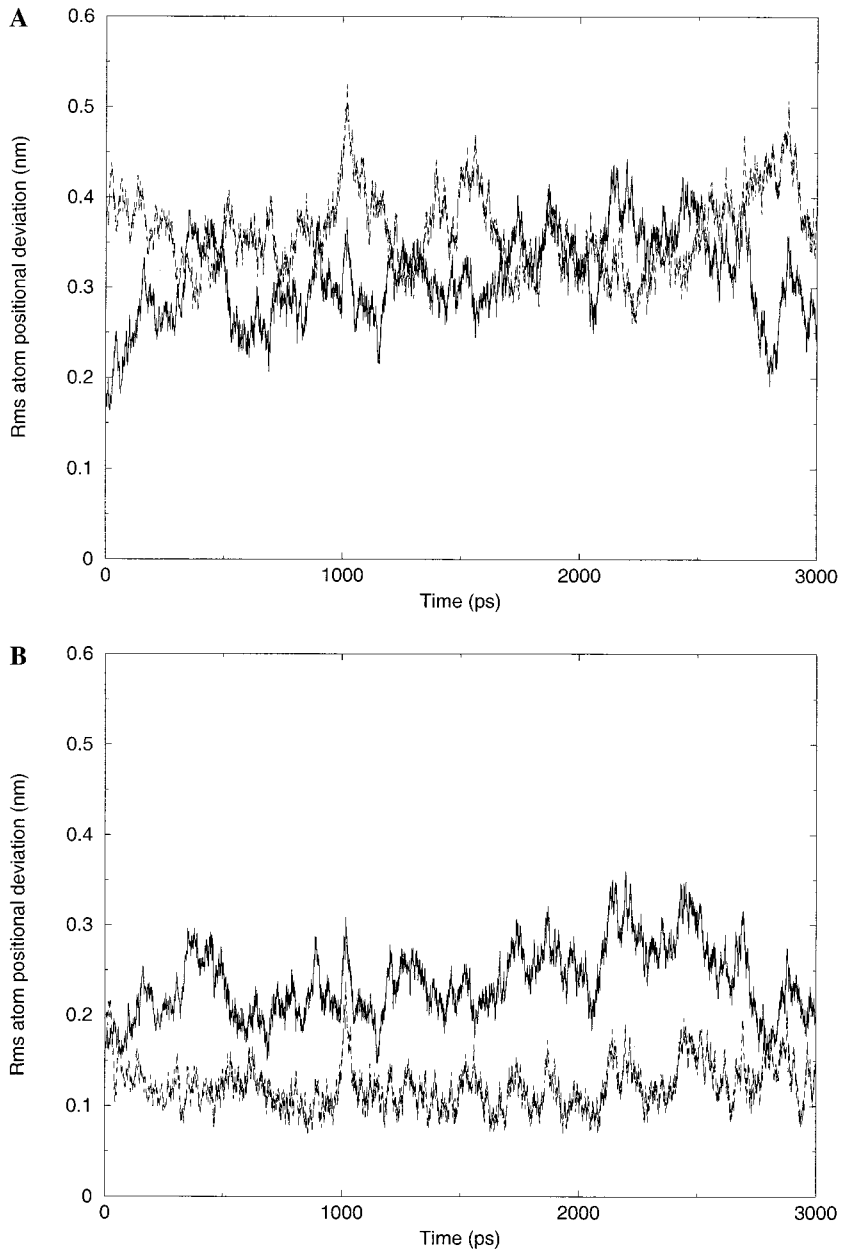


FIGURE 2 Root-mean-square atom positional deviations as a function of simulation time. (*A*) From A-DNA (*top curve*), from B-DNA (*bottom curve*); (*B*) from the average MD structure of the DNA in solution (*bottom curve*) and from the x-ray diffraction structure (*top curve*). The structures were superimposed and the rms deviations calculated for all atoms excluding the first two and the last two basepairs.

segment interacting with the *Bam*HI endonuclease (Fig. 2 *B*, *top curve*) is less than that from the B-DNA structure. Because the starting structure of the simulation is the B-DNA, the system spontaneously evolves toward the crystallographic structure during the thermalization and oscillates around this configuration. This result is in agreement with the x-ray evidence that in the DNA-*Bam*HI complex the DNA structure remains quite rigid, without any major bends or kinks, while the protein undergoes a series of conformational changes (Newman et al., 1995). Up to now the 3D structure of the DNA segment in absence of the *Bam*HI endonuclease is not available, preventing any comparison with our simulation.

A further confirmation of the stability of the simulated structure comes from the analysis of the average number of the interstrand Watson-Crick hydrogen bonds (*black bars*) and of the whole hydrogen bonds (*gray bars*) plotted in Fig. 3. The number of genuine Watson-Crick hydrogen bonds is strictly close to two, and to three for the A-T and C-G basepairs, respectively. A lower number is only observed for the last two basepairs at one end of the helix, which leads to a subtle structural deformation at the tail of the DNA helix. Comparison of our data with the ones previously extracted from an MD simulation for a DNA duplex belonging to the trp operator fragment (Bonvin et al., 1998) indicates that our structure displays a higher ability to maintain the inter-strand hydrogen bonds. This may be due either to a higher intrinsic stability of our DNA segment or to the avoidance of truncation effect in the treatment of the electrostatic interactions because of the use in our simulation of the particle mesh Ewald method (Essman et al., 1995), which has been shown to be well suited for the treatment of electrostatics in the case of strongly charged systems like DNA.

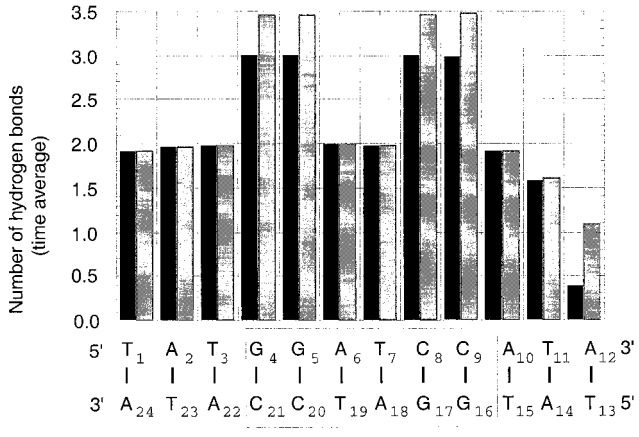


FIGURE 3 Average number of inter-strand basepairing hydrogen bonds (*gray bars*) as a function of the DNA sequence calculated from the last 2.5-ns trajectory. The fractions of the Watson-Crick hydrogen bond are indicated in black. The criterion for the occurrence of a hydrogen bond is defined in Computational Methods. The box indicates central basepairs involved in *Bam*HI recognition.

## Hydration analysis

The trajectory has been analyzed to work out information on the degree of hydration around the various groups forming the DNA sequence. Before performing the analysis, the presence of time-dependent artifacts (Sagui and Darden, 1999; Harvey et al., 1998) due to the use of a relatively long nonbond update time, i.e., 20 fs, was checked, evaluating the water dynamics throughout the trajectory. The water kinetic energy was fairly constant during the overall simulation, its variation being $<5\%$, allowing us to exclude the presence of any artifact. In the analysis we have evaluated the average and maximum residence time and the coordination number of water molecules around specific sites. These data are reported for the backbone and the major and minor groove, respectively, and are discussed in comparison with other structural an dynamical values obtained in previous work.

## Backbone

The coordination number and the average and maximum residence times of water molecules around the phosphate and ribose oxygen atoms evaluated from the last 2.5 ns of simulation are reported in Table 1. The data indicate that of the four phosphate oxygens, the two partially charged phosphate oxygens, O1P and O2P, are highly hydrated, their average water coordination number ranging between 3.5 and 3.8. However, the coordination number around the ester oxygens, O3′ and O5′, oscillates around 1. An interesting observation is that the oxygen phosphate and the ester oxygen coordination number does not depend on the presence of specific bases because these values are fairly constant all along the DNA sequence. Such a finding is in agreement with the analysis carried out by Schneider et al. (1998) on 59 crystallographic structures of DNA. In that case it is reported that each charged phosphate oxygen (O1P and O2P) has three water-ordered molecules in its first coordination sphere. The lower coordination number observed by Schneider is likely due to the fact that in his work the selection has been made on water molecules hydrogen-bonded to the oxygens, so that the considered constraints (water-O1P and -O2P distance 2.8 Å and hydrogen bond angle $>125°$) are more strict than in our case, where we have considered all the water molecules having an oxygen phosphate–oxygen-water distance $\leq 3.5$ Å. However, in both analyses it is found that water has a high affinity for the phosphate oxygens and that ester oxygens are hydrated only marginally.

The two types of oxygen atoms behave differently also as far as the water residence time is concerned. Phosphate oxygen atoms, in fact, have average and maximum residence times that are about twice those displayed by the ester oxygen atoms (see Table 1). This could be attributed to the difference in partial charge, which is greater for the phos-

**TABLE 1 Average coordination number, average and maximum water residence time around phosphate and ribose oxygen atoms**

| | O1P | | | O2P | | | O3′ | | | O5′ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bases | $n_{wat}$* | $t$ (ps)† | $t_{max}$ (ps)‡ | $n_{wat}$ | $t$ (ps) | $t_{max}$ (ps) | $n_{wat}$ | $t$ (ps) | $t_{max}$ (ps) | $n_{wat}$ | $t$ (ps) | $t_{max}$ (ps) |
| $G_4$ | 3.8 | 7 | 127 | 3.6 | 9 | 208 | 1.3 | 3 | 51 | 1.0 | 2 | 83 |
| $G_5$ | 3.8 | 7 | 139 | 3.6 | 8 | 133 | 1.3 | 2 | 81 | 1.2 | 2 | 86 |
| $A_6$ | 3.8 | 8 | 107 | 3.6 | 8 | 242 | 1.3 | 2 | 82 | 1.1 | 2 | 116 |
| $T_7$ | 3.8 | 7 | 173 | 3.6 | 7 | 167 | 1.2 | 2 | 66 | 1.1 | 3 | 95 |
| $C_8$ | 3.8 | 8 | 129 | 3.6 | 9 | 178 | 1.3 | 4 | 77 | 1.1 | 2 | 74 |
| $C_9$ | 3.8 | 6 | 101 | 3.7 | 8 | 174 | 1.4 | 3 | 120 | 1.1 | 2 | 152 |
| $G_{16}$ | 3.8 | 8 | 172 | 3.6 | 12 | 147 | 1.3 | 3 | 83 | 0.9 | 3 | 61 |
| $G_{17}$ | 3.7 | 9 | 135 | 3.5 | 11 | 203 | 1.3 | 3 | 63 | 1.4 | 3 | 96 |
| $A_{18}$ | 3.8 | 7 | 105 | 3.6 | 8 | 249 | 1.3 | 2 | 95 | 1.0 | 2 | 115 |
| $T_{19}$ | 3.8 | 7 | 145 | 3.5 | 12 | 325 | 1.3 | 2 | 52 | 1.0 | 3 | 83 |
| $C_{20}$ | 3.8 | 7 | 139 | 3.7 | 14 | 117 | 1.3 | 2 | 71 | 1.1 | 2 | 69 |
| $C_{21}$ | 3.8 | 9 | 127 | 3.7 | 10 | 140 | 1.3 | 2 | 83 | 1.1 | 2 | 78 |

*Average water coordination number.
†Average water residence time.
‡Maximum water residence time.

phate (−0.776) than for the ester (−0.509) oxygen atoms, resulting in a long-lasting water interaction for the first class of oxygens. A comparable result is also found in an MD simulation of the trp operator (Bonvin et al., 1998) indicating that phosphate group hydration is independent of DNA sequence.

## Major and minor groove

The average and maximum water residence times and the coordination number of water molecules around the various sites of the bases belonging to the major and minor groove are reported in Tables 2 and 3, respectively. In the tables the average solvent accessibility surface of the various sites evaluated from the MD trajectory is also reported in comparison with the solvent accessibility surface worked out from the DNA structure as observed in the x-ray diffraction of the *Bam*HI-DNA complex (Newman et al., 1995), once the DNA sequence and the protein are separated. The two values are similar for all the major groove sites and some deviation is only observed in the minor groove around the Gua-5 base. This result is confirmed by the average values of the minor groove width taken over 500-3000 ps, reported in Fig. 4 *A*, in comparison to the values extracted from the DNA structure of the *Bam*HI-DNA complex. The largest difference is observed in correspondence of the Gua-5 base. The average twist, obtained from the MD simulation, is found to have values lower than those observed in the x-ray structure (Fig. 4 *B*), suggesting that the amber force field underestimates this quantity, as reported in other MD simulations (Cheatham and Kollman, 1997b; Young et al., 1997). However, these results, together with the trend of the structural parameters of Fig. 2 confirm, that MD is sampling conformations quite close to the x-ray structure.

In the major groove each pyrimidine has one hydration site; in particular, at least one water molecule is present during all the trajectory in the first coordination sphere of the N4 of cytosine and O4 tymine, respectively (Table 2). The average water residence time is slightly shorter for water molecules residing in the proximity of a cytosine than of a tymine, however in both cases water is fast-exchanging with the bulk solvent, indicating that these sites, although highly hydrated, do not trap water molecules for a long time. In the case of purines, they both display two hydration sites in correspondence to the N7 and N6 atoms of the

**TABLE 2 Hydration and water accessibility parameters in the major groove**

| Bases | Atom type | $n_{wat}$ | $t$ (ps) | $t_{max}$ (ps) | $S$ (Å$^2$) MD* | $S$ (Å$^2$)† X-ray |
|---|---|---|---|---|---|---|
| $G_4$ | N7 (acceptor) | 1.4 | 7 | 104 | 12 | 11 |
| | O6 (acceptor) | 1.4 | 7 | 137 | 6 | 6 |
| $G_5$ | N7 (acceptor) | 1.7 | 3 | 119 | 15 | 15 |
| | O6 (acceptor) | 1.7 | 5 | 127 | 11 | 13 |
| $A_6$ | N7 (acceptor) | 1.3 | 10 | 162 | 10 | 6 |
| | N6 (donor) | 1.4 | 2 | 81 | 8 | 8 |
| $T_7$ | O4 (acceptor) | 1.4 | 9 | 112 | 13 | 9 |
| $C_8$ | N4 (donor) | 1.6 | 3 | 105 | 14 | 13 |
| $C_9$ | N4 (donor) | 1.5 | 2 | 96 | 12 | 13 |
| $G_{16}$ | N7 (acceptor) | 1.6 | 3 | 93 | 13 | 13 |
| | O6 (acceptor) | 1.6 | 4 | 135 | 8 | 8 |
| $G_{17}$ | N7 (acceptor) | 1.8 | 2 | 270 | 14 | 13 |
| | O6 (acceptor) | 1.7 | 7 | 149 | 11 | 11 |
| $A_{18}$ | N7 (acceptor) | 1.3 | 7 | 245 | 12 | 10 |
| | N6 (donor) | 1.6 | 2 | 337 | 11 | 15 |
| $T_{19}$ | O4 (acceptor) | 1.3 | 13 | 287 | 10 | 5 |
| $C_{20}$ | N4 (donor) | 1.6 | 3 | 123 | 12 | 12 |
| $C_{21}$ | N4 (donor) | 1.5 | 2 | 128 | 12 | 12 |

$n_{wat}$, $t$ (ps), and $t_{max}$ (ps) have the same definition as in Table 1.
*Solvent average accessibility evaluated from the MD trajectory.
†Solvent accessibility from x-ray.

**TABLE 3 Hydration and water accessibility parameters in the minor groove**

| Bases | Atom type | $n_{wat}$ | $t$ (ps) | $t_{max}$ (ps) | $S$ (Å$^2$) MD | $S$ (Å$^2$) X-ray |
|---|---|---|---|---|---|---|
| $G_4$ | N2 (donor) | 1.8 | 4 | 246 | 12 | 13 |
| | N3 (acceptor) | 1.1 | 4 | 205 | 5 | 6 |
| $G_5$ | N2 (donor) | 1.6 | 2 | 208 | 8 | 1 |
| | N3 (acceptor) | 1.0 | 10 | 344 | 3 | 1 |
| $A_6$ | N3 (acceptor) | 1.0 | 9 | 444 | 5 | 5 |
| $T_7$ | O2 (acceptor) | 1.3 | 16 | 351 | 6 | 5 |
| $C_8$ | O2 (acceptor) | 1.1 | 8 | 378 | 6 | 3 |
| $C_9$ | O2 (acceptor) | 1.1 | 11 | 429 | 6 | 3 |
| $G_{16}$ | N2 (donor) | 1.7 | 2 | 233 | 13 | 10 |
| | N3 (acceptor) | 1.1 | 4 | 216 | 5 | 5 |
| $G_{17}$ | N2 (donor) | 1.5 | 5 | 215 | 7 | 3 |
| | N3 (acceptor) | 0.9 | 5 | 290 | 2 | 1 |
| $A_{18}$ | N3 (acceptor) | 1.0 | 9 | 271 | 4 | 3 |
| $T_{19}$ | O2 (acceptor) | 1.4 | 19 | 256 | 10 | 11 |
| $C_{20}$ | O2 (acceptor) | 1.1 | 12 | 249 | 6 | 2 |
| $C_{21}$ | O2 (acceptor) | 1.1 | 17 | 486 | 6 | 4 |

See Table 2 for definitions.

adenine and N7 and O6 atoms of the guanine base, respectively. The average coordination number is around 1 for both the N7 and N6 atoms of the adenine bases, while it is higher for the guanine atoms. The water molecules detected in the major groove are not shared between atoms belonging to different strands or between atoms belonging to the same base except for the Ade-6 and Ade-18, where one water molecule is shared for 2% and 20% of the simulation trajectory, respectively, between the N7 and N6 atoms. This effect slightly decreases the coordination number of the N7 and N6 atoms of the adenine bases and makes the C-G pairs more hydrated than the A-T basepairs, as may be observed by summing the average coordination number of the CG and the AT reported in Table 2. A higher hydration of the C-G basepairs when compared to the AT pairs was found in a recent long MD simulation (10 ns) of the DNA fragment d(C5T5)(A5G5) (Feig and Pettitt, 1999). This result is also in agreement with the hydration values evaluated by molar volume and compressibility measurement as a function of base composition (Chalikian et al., 1994) and by free energy calculations (Elcock and McCammon, 1995). The water residence times of the purines are also characterized by relatively short values (Table 2).

In the minor groove (Table 3) the average water coordination number for each base is close to one except for the guanine base, which displays a non-shared water molecule around both the N2 and N3 atoms, again confirming the higher degree of hydration of the C-G than the A-T basepairs. The average water residence time is short and comparable to that observed in the major groove. Some differences between major and minor grooves are observed at the level of the maximum water residence time, which is longer in the minor than in the major groove. The number of water molecules with maximum residence time greater than 80 ps,

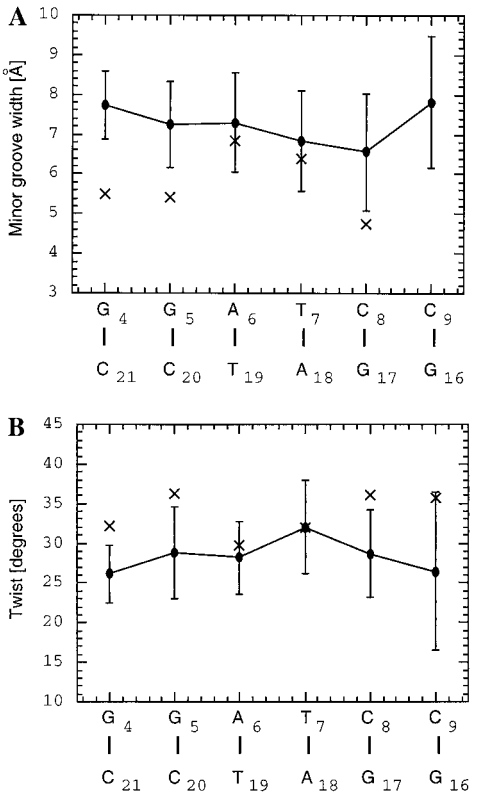


FIGURE 4 Minor groove average width (*A*) and average twist (*B*) for the six central basepairs calculated from the last 2.5-ns trajectory (●) compared with the values evaluated from the crystal structure (×). Error bars are single standard deviation about the mean. The two parameters were averaged from values extracted every 50 ps.

reported in Fig. 5 for the six central basepairs, is higher in the minor than in the major groove, indicating that water is trapped for a longer time in the minor groove. As a matter of fact, the minor groove shows the presence of localized

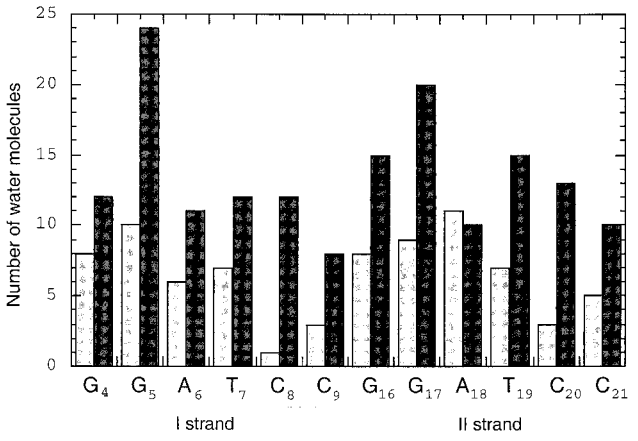


FIGURE 5 Number of water molecules residing on each base having a maximum residence time >80 ps in the major (*gray bars*) and minor (*black bars*) grooves, respectively.

water molecules. In fact, the 5′-$A_6T_7C_8$-3′ strand shares water molecules with the opposite 5′-$A_{18}T_{19}C_{20}$-3′ bases for a significant percentage of time. In Fig. 6 a scheme of the minor groove, where base atoms sharing water molecules are connected by a line, has been reported. There are pairs of atoms, such as $C_8$ $O_2$ and $A_{18}$ N3, $T_7$ $O_2$ and $T_{19}$ $O_2$, $A_6$ N3 and $C_{20}$ $O_2$, which share a water molecule for 24%, 44%, and 28% of the total simulation time, respectively. This implies that a water molecule has a high probability of being found in a position bridging two atoms belonging to the two strands despite the fact that the average water residence time is short. These positions are local minima for water molecules and create a spine of hydration, as already observed in the minor groove of other DNA sequences (Duan et al., 1997; Feig and Pettitt, 1999). Fig. 5 also shows that there are water molecules in the minor groove that may be shared for a different percentage of time by various atoms, sometimes even by four sites, although for a relatively short percentage of the simulation time, i.e., 5%. Such shared water molecules are not observed in the major groove, indicating that water localization occurs only in the minor groove. However, despite this localization, the average water residence times are relatively short also in the minor groove, although slightly longer than in the major

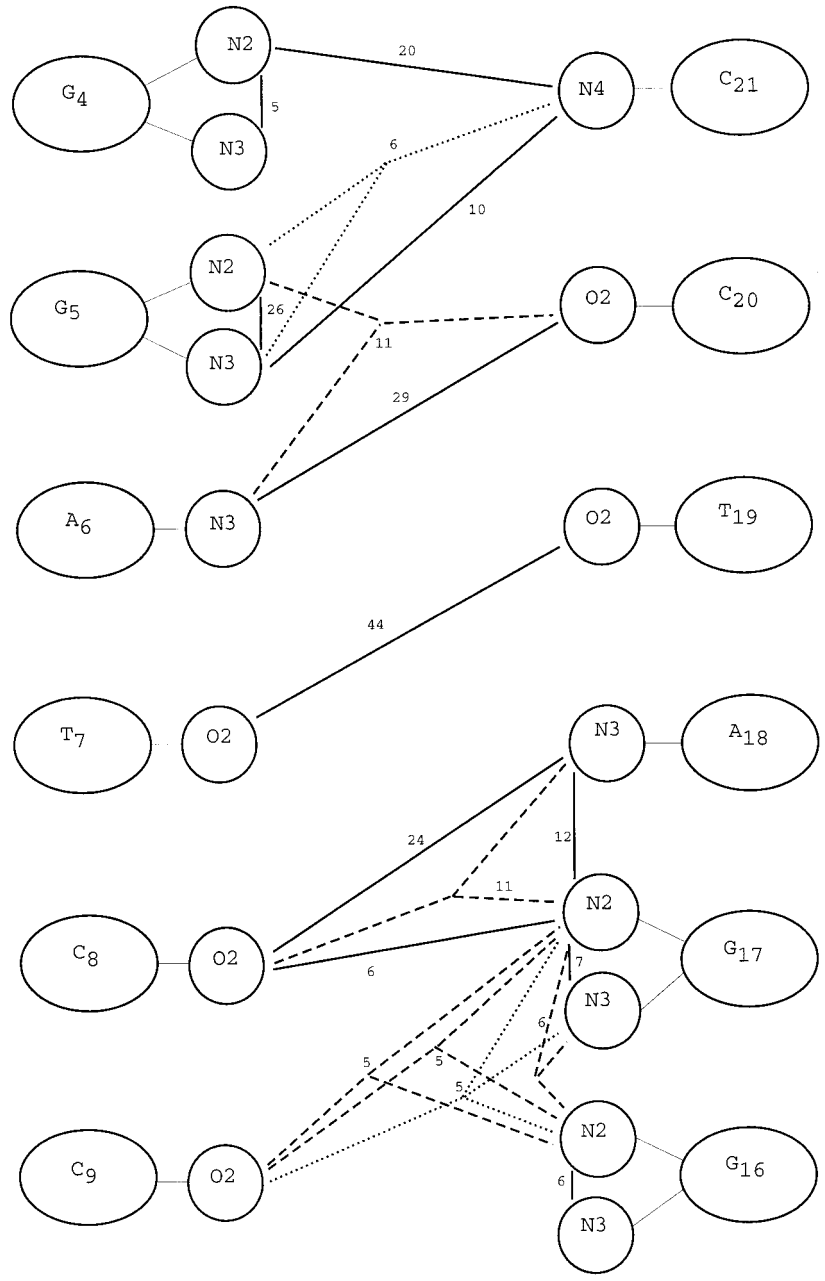


FIGURE 6 Connectivity of the hydration sites in the six central basepairs of the DNA duplex recognized by the *Bam*HI endonuclease. Atoms sharing a water molecule, having a distance <3.5 Å from each atom, are connected by lines. The reported value represents the percentage of time over the whole simulation time in which water molecules are shared by the connected atoms. Lines are solid, dashed, and dotted for water molecules shared by two, three, and four sites, respectively. The values represent the percentage of simulation time in which water molecules are shared between the various sites.

groove (see Tables 2 and 3), confirming that the static and dynamic properties of water are not necessarily correlated (Levitt and Park, 1993). In fact, a localized water molecule, i.e., a water molecule revealed by x-ray diffraction, has a favored average position where the water is almost always present, although likely fast-exchanging with bulk water. This behavior is better clarified in Fig. 7, where the time history of water molecules in the first coordination sphere of $T_7$ $O_2$ and $T_{19}$ $O_2$ are represented. Each water molecule is represented by a different color and the fast changes of colors indicate that the two sites are almost continuously hydrated, but by many different water molecules. The figure also represents as a function of time the water molecules that are shared by the above-mentioned atoms and it confirms that although water molecules have a high probability of bridging the two sites during the trajectory, each water molecule resides in this position for a relatively short time.

The average residence times of water molecules within 3.5 Å of nonexchangeable DNA protons were also calculated using a time resolution of 10 ps, i.e., counting water molecules leaving and entering the 3.5 Å radius within 10 ps as being continuously within the distance cutoff. The values shown in Table 4 are three to four times lower than those reported in a previous MD study carried out on a different DNA sequence using the same resolution time (Bonvin et al., 1998). Such a difference can, at least in part, be due to the use in our study of the TIP3P water model, which has a diffusion constant higher (Van der Spoel et al., 1998) than the SPC model used in the Bonvin work. Decreasing the time resolution to 0.1 ps reduces the evaluated residence time by one order of magnitude (data not shown) indicating that resolution time is the parameter that must be calibrated when comparing experimental and calculated water residence times.

**TABLE 4 Average residence times in picoseconds of water molecules within 0.35 nm of DNA hydrogen atoms**

| Base | H1′ | H2′ | H2″ | H3′ | H4′ | H5′ | H5″ | H6 | C-H5 | T-C5H3 | A-H2 | H8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $G_4$ | 46 | 20 | 20 | 26 | 24 | 24 | 20 | | | | | 25 |
| $G_5$ | 85 | 48 | 16 | 45 | 34 | 24 | 23 | | | | | 21 |
| $A_6$ | 65 | 21 | 38 | 30 | 37 | 37 | 19 | | | | 43 | 35 |
| $T_7$ | 20 | 19 | 17 | 21 | 15 | 26 | 25 | 39 | | 9 | | |
| $C_8$ | 27 | 14 | 23 | 31 | 14 | 14 | 17 | 31 | 22 | | | |
| $C_9$ | 47 | 21 | 13 | 24 | 20 | 30 | 40 | 24 | 23 | | | |
| $G_{16}$ | 31 | 22 | 36 | 29 | 57 | 16 | 24 | | | | | 28 |
| $G_{17}$ | 37 | 24 | 24 | 30 | 16 | 16 | 17 | | | | | 21 |
| $A_{18}$ | 38 | 23 | 18 | 27 | 50 | 16 | 16 | | | | 19 | 29 |
| $T_{19}$ | 15 | 24 | 13 | 45 | 21 | 31 | 26 | 28 | | 9 | | |
| $C_{20}$ | 30 | 19 | 16 | 38 | 16 | 28 | 33 | 28 | 26 | | | |
| $C_{21}$ | 71 | 16 | 26 | 29 | 26 | 18 | 22 | 18 | 28 | | | |

A time resolution of 10 ps was used for calculating the residence times.

Upon considering the importance of the choice of the resolution time the calculated residence time values are comparable to the ones experimentally measured through NMR on various B-DNA sequences (Sunnerhagen et al., 1998; Phan et al., 1999). In Phan's work it is also found that the residence times of the major groove were shorter than those in the minor groove as found in our simulation, suggesting that this trend is a general feature of the grooves that likely depends on their geometrical properties and not on the specific DNA sequence.

The x-ray diffraction of the *Bam*HI-DNA complex have indicated that Gua-5 interacts with the protein through water-mediated hydrogen bonds (Newman et al., 1995). We don't find any long-lasting water molecule in this or any other position; however, we find that Gua-5, together with Gua-17, is the most hydrated base. A similar result has been obtained in the MD simulation of the trp operator (Bonvin

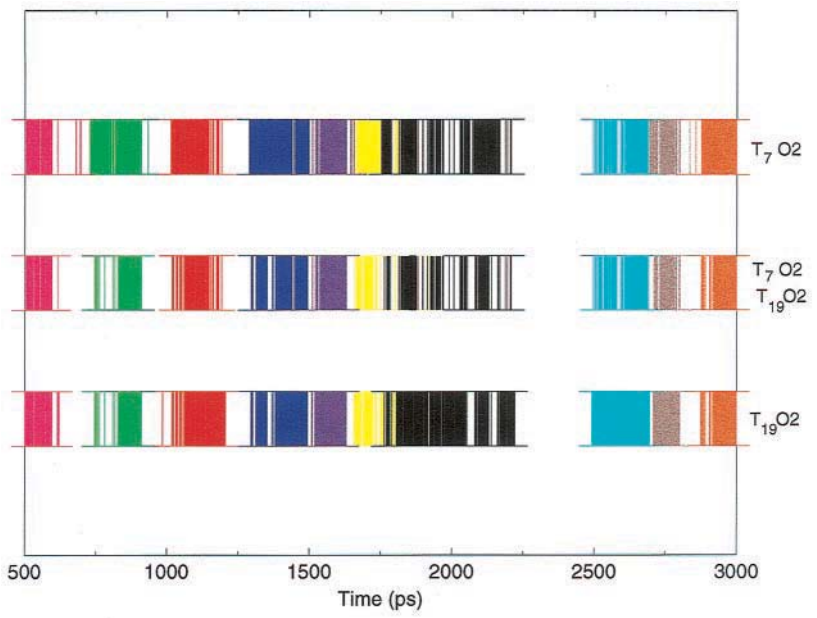


FIGURE 7 Hydration time history of the $O_2$ $T_7$ and $O_2$ $T_{19}$ atoms. Each color represents a different water molecule at a distance <3.5 Å from the $O_2$ $T_7$ atom (*top line*) and from the $O_2$ $T_{19}$ atom (*bottom line*). The middle line represents the fraction of time in which water molecules are shared between these two atoms.

et al., 1998) where no specifically long residence time values were observed in any site corresponding to the water crystallographic sites.

## CONCLUSIONS

The structure and hydration of the DNA sequence interacting with *Bam*HI has been investigated by means of a 3.0-ns molecular dynamics simulation in water. The use of the particle mesh Ewald method allowed an accurate treatment of the electrostatics and the generation of a stable MD trajectory. The solvated structure in the MD simulation reaches a stable configuration close to the configuration observed through x-ray diffraction in its complex with the *Bam*HI endonuclease. This is in line with the evidence that, in the DNA-*Bam*HI complex, the DNA remains quite rigid, while the protein undergoes conformational change to fit with the DNA sequence (Newman et al., 1995)

Hydration analysis indicates a preferential hydration for phosphate oxygen atoms than for ester oxygen atoms, as already observed in other static and dynamical analysis (Falk et al., 1962; De Oliviera Neto, 1986; Schneider et al., 1998; Bonvin et al., 1998). Water residence times are relatively short for both the backbone and the major and minor groove sites, being slightly shorter in the major than in the minor groove, as also experimentally reported in other DNA duplexes (Phan et al., 1999). A spine of hydration is found in the minor groove along the central CTA sequence confirming that it is not a unique feature of the AATT sequence (Duan et al., 1997; Feig and Pettitt, 1999).

No special long water residence time values were observed for sites where crystallographic waters have been detected, as also found in the case of the trp operator. These results indicate that the sites where water molecules are observed by x-ray diffraction correspond to sites where water molecules have a high probability of being found, but these molecules are usually fast-exchanging with the bulk solvent.

We thank Dr. A. Bonvin (Utrecht University) and Dr. N. Sanna (CASPUR, Roma) for useful discussions.

## REFERENCES

Arnott, S., P. J. Campbell-Smith, and R. Chandrasekaran. 1976. Atomic coordinates and molecular conformations for DNA-DNA, RNA- RNA, and DNA- RNA helices. *In* CRC Handbook of Biochemistry and Molecular Biology, Vol. 2.CRC Press, Cleveland, OH. 411–422.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gusteren, A. Di Nola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Comput. Phys.* 81:3684–3690.

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

Beveridge, D. L., and G. Ravinshanker. 1994. Molecular dynamics study of DNA. *Curr. Opin. Struct. Biol.* 4:246–255.

Bonvin, A. M., M. Sunnerhagen, G. Otting, and W. F. van Gusteren. 1998. Water molecules in DNA recognition II. A molecular dynamics view of the structure and hydration of the trp operator. *J. Mol. Biol.* 282: 859–873.

Chalikian, T. V., A. P. Sarvazyan, and E. G. Plum. 1994. Influence of base composition, base sequence, and duplex structure on DNA hydration: apparent molar volumes and apparent adiabatic compressibilities of synthetic and natural DNA duplexes at 25°C. *Biochemistry.* 33: 2394–2401.

Cheatham, T. E., and P. A. Kollman. 1997a. Insight into the stabilization of a DNA by specific ion association: spontaneous B-DNA to A-DNA transitions observed in molecular dynamics simulations of d(ACCCGCGGGT)$_2$ in the presence of hexammine cobalt (III). *Structure.* 15:1297–1311.

Cheatham, T. E., and P. A. Kollman. 1997b. Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA and DNA:RNA hybrid duplexes. *J. Am. Chem. Soc.* 119:4805–4825.

Cheatham, T. E., J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman. 1995. Molecular dynamics simulation on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA and proteins. *J. Am. Chem. Soc.* 117:4193–4194.

Cheatham, T. E., J. L. Miller, T. I. Spector, P. Cieplak, and P. A. Kollman. 1997. Molecular dynamics simulations on nucleic acid systems using the Cornell et al. force field and particle mesh Ewald electrostatics. *In* Molecular Modeling of Nucleic Acids (N. B. Leontis, editor. American Chemical Society, Washington, DC. 285–302.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, M. Kenneth, J. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N.log(n) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

De Oliviera Neto, M. 1986. Rapid location of the preferred interaction sites between small polar molecules and macromolecules. II. Binding of water to a model segment of B-DNA. *J. Comput. Chem.* 7:629–639.

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

Duan, Y., P. Wilkosz, M. Crowley, and J. M. Rosenberg. 1997. Molecular dynamics simulation study of DNA dodecamer d(CGWAATTWGC) in solution: conformation and hydration. *J. Mol. Biol.* 272:553–572.

Elcock, A. H., and J. A. McCammon. 1995. Sequence dependent hydration of DNA: theoretical results. *J. Am. Chem. Soc.* 117:10161–10162.

Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

Falk, M., K. A. Hartman, and R. C. Lord. 1962. Hydration of deoxyribonucleic acid. I. A gravimetric study. *J. Am. Chem. Soc.* 84:3843–3846.

Feig, M., and B. M. Pettitt. 1999. Modeling high-resolution hydration patterns in correlation with DNA sequence and conformation. *J. Mol. Biol.* 286:1075–1095.

Garcia, A. E., and L. Stiller. 1993. Computation of the mean residence time of water in the hydration shells of biomolecules. *J. Comp. Chem.* 14:1396–1406.

Harvey, S. C., R. K.-Z. Tan, and T. E. Cheatham. 1998. The flying ice cube: velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comput. Chem.* 19:726–740.

Hubbard, S. J., and J. M. Thornton. 1993. "NACCESS." Computer Program, Department of Biochemistry and Molecular Biology, University College, London.

Jayaram, B., and D. L. Beveridge. 1996. Modeling DNA in aqueous solutions: Theoretical and computer simulation studies on the ion atmosphere of DNA. *Annu. Rev. Biophys. Biomol. Struct.* 25:367–394.

Jorgensen, W. L. 1981. Transferable intermolecular potential functions for water alcohols and ethers. Application to liquid waters. *J. Am. Chem. Soc.* 103:335–340.

Kopka, M. L., A. V. Fratini, H. R. Drew, and R. E. Dickerson. 1983. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* 163:129–146.

Kuhn, L. A., M. A. Siani, M. E. Pique, C. L. Fisher, E. D. Getzoff, and J. A. Tainer. 1992. The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. *J. Mol. Biol.* 228:13–22.

Lavery, R., and H. Sklenar. 1988. The definition of the generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 4:655–667.

Levitt, M., and B. H. Park. 1993. Water: now you see it, now you don't. *Structure.* 1:223–226.

Luise, A., M. Falconi, and A. Desideri. 2000. Molecular dynamics simulation of solvated azurin: influence of surface solvent accessibility on water residence times. *Proteins*. 39:56–67.

Newman, M., T. Strzelecka, L. F. Dorner, I. Schildkraut, and A. K. Aggarwal. 1994. Structure of restriction endonuclease *Bam*HI and its relationship to *Eco*RI. *Nature.* 368:660–664.

Newman, M., T. Strzelecka, L. F. Dorner, I. Schildkraut, and A. K. Aggarwal. 1995. Structure of *Bam*HI endonuclease bound to DNA: partial folding and unfolding on DNA binding. *Science.* 269:656–663.

Phan, A. T., J.-L. Leroy, and M. Gueron. 1999. Determination of the residence time of water molecules hydrating B′-DNA and B-DNA, by one-dimensional zero-enhancement nuclear Overhauser effect spectroscopy. *J. Mol. Biol.* 286:505–519.

Rentzeperis, D., D. W. Kupke, and L. A. Marky. 1993. Volume changes correlate with entropies and enthalpies in the formation of nucleic acid homoduplexes: differential hydration of A and B conformations. *Biopolymers.* 33:117–125.

Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23: 327–341.

Sagui, C., and T. A. Darden. 1999. Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.* 28:155–179.

Schneider, B., and H. M. Berman. 1995. Hydration of the DNA bases is local. *Biophys. J.* 69:2661–2669.

Schneider, B., K. Patel, and H. M. Berman. 1998. Hydration of the phosphate group in double-helical DNA. *Biophys. J.* 75:2422–2434.

Schwabe, J. W. R. 1997. The role of water in protein-DNA interactions. *Curr. Opin. Struct. Biol.* 7:126–134.

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

Subramanian, P. S., and D. L. Beveridge. 1989. A theoretical study of the aqueous hydration of the canonical B d(CGCGAATTCGCG): Monte Carlo simulation and comparison crystallographic ordered water sites. *J. Biomol. Struct. Dyn.* 6:1093–1122.

Subramanian, P. S., G. Ravishanker, and D. L. Beveridge. 1988. Theoretical considerations on the "spine of hydration" in the minor groove of d(CGCGAATTCGCG)* d(CGCGAATTCGCG): Monte Carlo computer simulation. *Proc. Natl. Acad. Sci. U.S.A.* 85:1836–1840.

Subramanian, P. S., S. Swaminathan, and D. L. Beveridge. 1990. Theoretical account of the spine of hydration in the minor groove of duplex d(CGCGAATTCGCG). *J. Biomol. Struct. Dyn.* 7:1161–1165.

Sunnerhagen, M., V. P. Denisov, K. Venu, A. M. J. J. Bonvin, J. Carey, B. Halle, and G. Otting. 1998. Water molecules in DNA recognition I: hydration lifetimes in the trp operator in solution measured by NMR spectroscopy. *J. Mol. Biol.* 282:847–858.

Van der Spoel, D., P. J. Van Maaren, and H. J. C. Berendsen. 1998. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* 108:10220–10230.

Young, A. Y., and D. L. Beveridge. 1998. Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.* 281:675–687.

Young, A. Y., G. Ravishanker, and D. L. Beveridge. 1997. A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J.* 73:2313–2336.

Westhof, E. 1988. Water: an integral part of nucleic acid structure. *Annu. Rev. Biophys. Chem.* 17:125–144.

Westhof, E. 1993. Structural water bridges in nucleic acid structure. *In* Water and Biological Macromolecules. CRC Press, Boca Raton, FL. 226–243.