A 5-Nanosecond Molecular Dynamics Trajectory for B-DNA: Analysis of Structure, Motions, and Solvation

Matthew A. Young, G. Ravishanker, and D. L. Beveridge
Chemistry Department and Molecular Biophysics Program, Wesleyan University, Middletown, Connecticut 06459 USA

ABSTRACT We report the results of four new molecular dynamics (MD) simulations on the DNA duplex of sequence d(CGCGAATTCGCG)$_2$, including explicit consideration of solvent water, and a sufficient number of Na^+ counterions to provide electroneutrality to the system. Our simulations are configured particularly to characterize the latest MD models of DNA, and to provide a basis for examining the sensitivity of MD results to the treatment of boundary conditions, electrostatics, initial placement of solvent, and run lengths. The trajectories employ the AMBER 4.1 force field. The simulations use particle mesh Ewald summation for boundary conditions, and range in length from 500 ps to 5.0 ns. Analysis of the results is carried out by means of time series for conformationalm, helicoidal parameters, newly developed indices of DNA axis bending, and groove widths. The results support a dynamically stable model of B-DNA for d(CGCGAATTCGCG)$_2$ over the entire length of the trajectory. The MD results are compared with corresponding crystallographic and NMR studies on the d(CGCGAATTCGCG)$_2$ duplex, and placed in the context of observed behavior of B-DNA by comparisons with the complete crystallographic data base of B-form structures. The calculated distributions of mobile solvent molecules, both water and counterions, are displayed. The calculated solvent structure of the primary solvation shell is compared with the location of ordered solvent positions in the corresponding crystal structure. The results indicate that ordered solvent positions in crystals are roughly twice as structured as bulk water. Detailed analysis of the solvent dynamics reveals evidence of the incorporation of ions in the primary solvation of the minor groove B-form DNA. The idea of localized complexation of otherwise mobile counterions in electronegative pockets in the grooves of DNA helices introduces an additional source of sequence-dependent effects on local conformational, helicoidal, and morphological structure, and may have important implications for understanding the functional energetics and specificity of the interactions of DNA and RNA with regulatory proteins, pharmaceutical agents, and other ligands.

INTRODUCTION

Issues of molecular motion and flexibility are expected to be a crucial component in developing a fundamental understanding of the relationship between DNA structure and function in molecular biophysics and genetics. Diverse biophysical methods can be applied, but no one experimental technique is capable of full elucidation of the dynamical structure of DNA in a solution environment. Molecular dynamics (MD) simulation can, in principle, provide a complete theoretical description of DNA structure and motions, and is thus a valuable independent means of developing models and interpreting experimental data. Furthermore, MD force fields now play a key supplementary role in the determination of structures by iterative refinement procedures in crystallography and NMR spectroscopy. However, although MD is a well-defined theoretical methodology, approximations in the underlying force field, effects of simulation protocols, and the intractability of run lengths approaching those of many phenomena of experimental interest currently limit the applicability of the results.

A number of MD simulations on B-form DNA oligonucleotides using various force fields have been reported over the period from 1982 to the present (Beveridge and Ravishanker, 1994; Beveridge et al., 1993). The results to date have provided considerable information on the ability of MD to simulate DNA accurately, as well as details of the dynamical structure of various model DNA systems. Advances in computer power now permit MD to be performed on increasingly realistic models of DNA systems, explicitly including solvent water and counterions, and to be extended from the picosecond well into the nanosecond time frame. Recently, a new version of the AMBER force field has been proposed (Cornell et al., 1995), with enhancements and specific improvements in the parameters for nucleic acids. In addition, a refined treatment of the long-range region of interaction potentials, the particle mesh Ewald (PME) method (Cheatham et al., 1995; York et al., 1994), has been proposed, to deal with a long-standing and well-known limitation in MD protocol.

This laboratory has been engaged for some time in the characterization of MD models of DNA, investigating their capabilities and limitations. We present here a series of extended MD simulations of the DNA duplex of sequence d(CGCGAATTCGCG)$_2$, particularly configured to independently characterize the AMBER 4.1 model of B-form DNA, and to simultaneously examine the sensitivity of MD results to the treatment of boundary conditions and electrostatics, initial placement of solvent, and run lengths. A

Received for publication 23 December 1996 and in final form 5 August 1997.

Address reprint requests to Dr. David L. Beveridge, Department of Chemistry, Wesleyan University, Hall-Atwater and Shanklin Laboratories, Middletown, CT 06457-0280. Tel.: 860-685-2575; Fax: 860-685-2211; E-mail: bever@rose.chem.wesleyan.edu.

0006-3495/97/11/2313/24 $2.00

number of new analysis tools are introduced here to critically analyze the results. A description of the calculated structure of the hydration density and counterion atmosphere is included. The results are compared with corresponding crystallographic and NMR studies on the d(CGCGAATTCGCG)$_2$ duplex and placed in the context of the observed behavior of B-DNA via comparisons with results from the current crystallographic database of B-form structures (Berman et al., 1992).

The calculations described herein also provide details on the dynamics of axis bending and groove widths in B-DNA, and some new ideas about the role of solvent in the structure of the DNA double helix. In a study preliminary to this article (Young et al., 1997), analysis of results from 1.5 ns of MD led us to suggest a reexamination of the solvent structure in the minor groove of d(CGCGAATTCGCG)$_2$, previously interpreted as a the "spine of hydration." The current set of MDs, extended now to 5.0 ns, provides further theoretical evidence that counterions may intrude on the minor groove spine of hydration in B-form DNA and that solvation of the minor groove is best described in terms of the fractional occupancies of counterions as well as water. The idea of localized complexation of otherwise mobile counterions in specific electronegative pockets in the grooves of DNA helices introduces an indirect means by which sequence can effect local conformational, helicoidal, and morphological structure in DNA, and may have important implications for understanding the functional energetics and specificity of the interactions of DNA and RNA with regulatory proteins, pharmaceutical agents, and other ligands.

BACKGROUND

The oligonucleotide duplex of sequence d(CGCGAATTCGCG)$_2$ was the first to crystallize in the B-form of DNA (Dickerson and Drew, 1981; Drew and Dickerson, 1981; Drew et al., 1981; Wing et al., 1980). The structure of this sequence has been the subject of a considerable number of subsequent crystallographic (Dickerson et al., 1985; Holbrook et al., 1985; Kopka et al., 1985; Nunn et al., 1993; Teng et al., 1988; Westhof, 1987), NMR (Alam et al., 1991; Hare et al., 1983; Lane et al., 1991; Liepinsh et al., 1992; Moe and Russu, 1990, 1992; Nerdal et al., 1989; Ott and Eckstein, 1985; Patel et al., 1982; Pelton and Wemmer, 1988; Withka et al., 1992) and Raman spectroscopic investigations (Wang et al., 1987; Weidlich et al., 1990), and now serves the field as a benchmark system for further experimental and theoretical (Beveridge et al., 1990; Falsafi and Reich, 1993; McConnell et al., 1994; Miaskiewicz et al., 1993; Prevost et al., 1993; Rao and Kollman, 1990; Ravishanker et al., 1989; Srinivasan et al., 1990; Subramanian and Beveridge, 1989; Subramanian et al., 1988, 1990; Swaminathan et al., 1991; Withka et al., 1992) studies. The central 6 bp of this sequence are the recognition site for the *Eco*RI restriction endonuclease, characterized at the molecular level in the first protein-DNA crystal structure (Frederick et al., 1984). The native crystal structure of d(CGCGAATTCGCG)$_2$, the "*Eco*RI dodecamer," displayed sequence-dependent irregularities in certain helicoidal parameters, particularly base pair roll and propeller twist. These results led to the hypothesis of "indirect readout," i.e., the idea that the conformational fine structure of DNA sequence plays a role in molecular recognition processes (Dickerson, 1983; Steitz, 1990).

At the level of helix morphology, the native crystalline form of the *Eco*RI dodecamer exhibited axis bending in the vicinity of the junctions between the AATT tract and the CGCG flanking sequences. Although the sequence is palindromic, the symmetry of the native DNA structure is not maintained in the crystalline environment, as it is when complexed with the *Eco*RI enzyme (Grable et al., 1984), suggesting an influence of packing effects (Dickerson et al., 1991; Narayana et al., 1991). The extent and nature of axis bending are different at the two AATT/CGCG junctions, which have come to be viewed as sites with a particular intrinsic propensity toward bending. The crystal structure also exhibits a relative narrowing of the minor groove in the AATT. Packing effects imposed on DNA helices in a crystalline environment may affect groove widths as well as bending and fine structure, and the effects of crystal packing remain to be fully disentangled from structural features intrinsic to the DNA (Dickerson et al., 1987, 1991; DiGabriele et al., 1989; Jain and Sundaralingam, 1989; Shakked et al., 1989). Recent progress on this issue has been reported (Berman et al., 1997). A prominent sequence of ordered solvent peaks in the electron density map of the minor groove region of the AATT was observed in the original crystal structure of the *Eco*RI dodecamer, and this feature, termed the "spine of hydration," has been discussed as a particular stabilizing feature in B-form DNA. A number of ordered water molecules were also assigned to positions proximal to the atoms of the major groove of the sequence, but no particular evidence of water spines or networks was reported.

The NMR structures for the *Eco*RI dodecamer currently support a conventional B-form structure in solution (Lane et al., 1991), although an early study reported evidence for an axis deformation at the central ApT step of the recognition site (Nerdal et al., 1989), similar to that observed in the crystal structure of the DNA complex with *Eco*RI endonuclease. Raman spectra and circular dichroism likewise support a B-form structure for this sequence in solution (Wang et al., 1987; Weidlich et al., 1990). NMR studies have subsequently been focused on the elucidation of the hydration, and support for the idea of the spine of hydration in the minor groove has been obtained (Kubinec and Wemmer, 1992; Liepinsh et al., 1992). The extensive experimental data available for the *Eco*RI dodecamer in crystal and solution, and the crystal structure of the 94 examples of B-DNA of various sequences in the Nucleic Acids Data Base (Berman et al., 1992) serve as a basis for characterization of the theoretical models.

MD studies of DNA oligonucleotides in general have recently been surveyed elsewhere (Beveridge and Ravishanker, 1994; Beveridge et al., 1993), and we confine our attention here to issues particularly relevant to the *Eco*RI sequence. A series of progressively longer MDs on the *Eco*RI dodecamer were reported from this laboratory, based on the GROMOS force field (van Gunsteren and Berendsen, 1986). An initial study revealed a tendency to base pair opening in GROMOS DNA on the picosecond time scale (Swaminathan et al., 1991). A subsequent analysis revealed that the interaction energy of Watson-Crick hydrogen bonds as described by GROMOS 86 is underestimated, based on experimental binding energies of prototypes and corresponding estimates from *ab initio* quantum chemistry (Gould and Kollman, 1993). The interactions were adjusted with a Tung-Harvey (Tung et al., 1984) potential, and the MD model with this modification resulted in a more B-like dynamical structure (McConnell et al., 1994).

DNA studies based on the AMBER 3.0 force field (Weiner et al., 1984, 1986) were characterized in a model treating the solvent effects by means of a distance-dependent dielectric screening function (Rao and Kollman, 1990; Srinivasan et al., 1990), as incorporated in MD on the d(CGCGAATTCGCG)$_2$ duplex. In more recent reports, the sequence plus solvent was treated explicitly and shown in a 140-ps trajectory to support a structure that was somewhat distorted with respect to canonical B-DNA (Miaskiewicz et al., 1993), but plausible. However, MD models of sequences expected to form B-DNA based on AMBER 3.0 showed a marked tendency to equilibrate in a region intermediate between that of canonical A- and B-forms and characterized by nucleotide base pairs, assuming a somewhat off-center alignment with respect to the helix axis and some degree of base pair inclination. Similar results were found with the CHARMM force field (Louise-May, 1994).

In the past year, revised parameterization of both AMBER and CHARMM force fields has been proposed, with improvements introduced specifically for the description of nucleic acids. Computational power has increased to the point that MD on macromolecular systems, explicitly including both water and counterions, is now readily feasible. The parameters of the earlier versions (first generation) of empirical force fields were developed based on a treatment of solvent via dielectric screening functions. Cornell et al. (1995) have developed a "second-generation" derivation of the series of empirical force fields, parameterized it in a manner consistent with explicit inclusion of solvent, and incorporated into the AMBER suite of programs (version 4.1). Net atomic charges are balanced with the commonly used effective pairwise potentials for water-water interactions such as TIP3P (Jorgensen, 1981), which have values for empirically derived charges that reflect many-body polarization effects in a mean-field sense. Modifications were also introduced in van der Waals parameters, with the latter optimized to reproduce liquid-state properties.

In addition to choice of force field, the run-time parameters necessary to produce reproducible, stable MD trajectories are also operational elements of a well-defined dynamical model. The principal issues in this area are 1) the treatment of long-range interactions (Loncharich and Brooks, 1989) and 2) the length of trajectories vis-à-vis stabilization and convergence. Long-range interactions are especially important in nucleic acids relative to proteins because of the proportionally larger number of charged groups. For reasons of both economy and tractability, long-range interactions in MD on macromolecules have traditionally been dealt with by various schemes in which pairwise potentials are truncated and then smoothed out with "switching functions" (Ravishanker et al., 1997), to avoid effective discontinuities in the evaluation of forces from derivatives. The degree to which artifacts may arise from this procedure has recently been characterized in simulations on aqueous NaCl and DNA (Auffinger and Beveridge, 1995). Another, more recent approach to this problem is applied through the use of force-shifting functions, as implemented in CHARMM (Loncharich and Brooks, 1989; Steinbach and Brooks, 1994). An alternative to either of these approaches, which in effect involves no truncation per se, is that of Ewald summation (Allen and Tildesly, 1987), which involves treating the system as a quasicrystal and forming lattice sums. Recently, a fast algorithm for Ewald sums was devised, called the "particle mesh Ewald (PME) method" (York et al., 1993, 1994), that makes this approach viable in MD. The initial round of PME simulations on both proteins and DNA demonstrated a higher degree of stability and less deviation from the starting structure (Cheatham et al., 1995). MD based on the AMBER 3.0 force field was carried out for the duplex d(CGCGAATTCGCG)$_2$ crystal (Duan et al., manuscript submitted for publication; York et al., 1995). Markedly improved stability in the MD was reported, and a heavy-atom root mean square deviation (RMSD) of 1.16 Å for the MD structure with the corresponding crystal structure was reported. Thus the method is currently attracting a considerable following, but whether the long-range electrostatics problem is completely solved by PME remains to be fully established. However, concerns have previously been expressed that the model implicit in PME may (Teleman and Wallqvist, 1990) or may not (Smith and Pettitt, 1996) introduce unphysical correlations when applied to aqueous solutions, and raise the possibility that PME could lead to an overestimation or physically incorrect description of dynamic stability (Smith and Pettitt, 1991). To further illuminate this issue, we compare in this article some MD results based on truncated potentials with those from PME.

A further issue in MD simulation is the choice of starting structure. In a previous study (Young et al., 1994), we found that GROMOS MD results for the d(CGCGAATTCGCG)$_2$ duplex (McConnell et al., 1994) and the d(CGCAAAAATGCG)$_2$ duplex (Young et al., 1995b) were essentially independent of whether the starting structure was canonical B-form, the crystal structure, or in the case of the *Eco*RI dodecamer (see also Kumar et al., 1994), the perturbed structure induced by protein binding in the endonuclease

complex. We now extend these considerations in the context of AMBER 4.1 to the sensitivity of the results to the initial placements chosen for mobile counterions. The various possibilities range from simply placing the ions at the bisectors of anionic phosphate groups (an ion-pair model), to locating them on the basis of an independent theoretical calculation based on electrostatics (Pearlman et al., 1995) or derived from a counterion Monte Carlo simulation (Jayaram et al., 1990b). Providing the system is ergodic, an MD simulation carried to sufficient length should produce results independent of this choice, but the length needed to attain this on any MD simulation of DNA including water and counterions has not yet been definitively established (Ravishanker et al., 1997). DNA, water, and counterions comprise a complex system, and the length scale for convergence of macromolecular MD including counterions and water together has not yet been characterized.

Convergence per se in MD computer simulation, in fact, can never be proved. (We note with interest that this issue predates both computers and simulation, being the Scottish philosopher David Hume's (1711–1776) objection to Sir Francis Bacon's (1561–1626) method of inductivism.) However, stability (oscillation of the dynamic structure around a mean) can be demonstrated, and lacking stability, an MD is definitely "unconverged." An MD trajectory may, in principle, be stable over an extended period of time and then legitimately transit to another structural form, if the underlying potential energy surface supports more than one thermally accessible local minimum ("substate"). Current practice is to run each trajectory as long as is practically affordable, with trajectory lengths at the nanosecond level now feasible for macromolecules in solution. In principle, all dynamic indices of structure should be executing oscillatory motion, or the calculation is not yet demonstrably stabilized, a criterion probably not yet achieved in any MD of a macromolecular system in solvent carried out to date. In our earlier reported nanosecond MD based on the GROMOS force field (McConnell et al., 1994), we characterized substates that emerged in a DNA simulation in going from the picosecond to nanosecond time scale (see also (Poncin et al., 1992)). It remains to be determined if this type of behavior persists in longer trajectories and is sensitive to force fields; we pursue this issue further herein.

In addition to the experimental structure determination specifically on the *Eco*RI dodecamer, it is of interest to examine the extent to which the calculated dynamical model behaves in a manner consistent with the entire class of B-form oligonucleotide structures with respect to conformational, helicoidal, and morphological indices of structure. The nature of the distributions of these indices observed in crystal structures can be gleaned from the NDB, and we have recently developed several techniques for analysis of structures applicable to detailed comparison of calculated and observed values. In particular, a comprehensive study of local bending across all available DNA crystal structures has recently been reported from this laboratory (Young et al., 1995a). Selected results of this study are used here to examine the extent to which computed dynamical models exhibit characteristics known for the class of B-form structures obtained from crystallography.

The structure of DNA is well known to be sensitive to water activity and salt (Franklin and Gosling, 1953; Ivanov et al., 1974; Wolf and Hanlon, 1975). An analysis of the behavior of counterions in continuum electrostatics calculations, Monte Carlo simulation, and MD on d(CGCGAATTCGCG)$_2$ has recently been reported from this Laboratory (Young et al., 1997). Detailed analysis of the solvent dynamics revealed evidence for the incorporation of ions in the primary solvation of the minor groove B-form DNA. We surveyed base pair steps of all possible compositions in the major and minor grooves of DNA, and indicated those with a high propensity for serving as electronegative pockets. A simple superimposition of electonegative potentials indicates that the steps ApT, TpA, and ApA (or TpT) of the minor groove and GpG (or CpC) of the major groove could serve as good pockets. A number of other steps provide possible pockets. A quantitative study of this issue based on continuum electrostatics as well as MD and MC simulation is in progress (Young, manuscript in preparation). The idea of localized complexation of otherwise mobile counterions in electronegative pockets in the grooves of DNA helices introduces an additional source of sequence-dependent effects on local conformational, helicoidal, and morphological structure, and may have important implications for understanding the functional energetics and specificity of the interactions of nucleic acids with regulatory proteins, pharmaceutical agents, and other ligands. Further analysis of the counterion structure in the model system is described in this article, based on the results of a considerably more extended trajectory.

The sensitivity of DNA structure to water activity was noticed in Franklin and Gosling's early experimentation on DNA fibers (Franklin and Gosling, 1953) and led to the proposition of A- and B-forms of DNA. Subsequently, it was shown that the position of the A- to B-form equilibrium transition could be irreducibly correlated with water activity (Ivanov et al., 1974; Saenger, 1983), with a preferential stability for the biologically prevalent B-form over A-DNA established at high water activity. At the molecular level, putative ordered forms have frequently been invoked to explain DNA sequence effects and axis bending (Westhof and Beveridge, 1989). An analysis of the ordered water structure in B-DNA crystal structures has recently been set forth by Schneider et al. (Schneider and Berman, 1995; Schneider et al., 1992, 1993). Because crystal structures reveal only "ordered" water structure (defined as that which can be assigned to residual electron density in a crystal structure determination), the full complement of water molecules hydrating a DNA is never experimentally observed. Thus little is actually known about DNA hydration at the molecular level beyond the ordered water structure, and, moreover, the nature of the causal role of hydration in conformational equilibria and free energy awaits a detailed explanation. The calculations described herein contain fur-

ther extensive details on the dynamical organization of counterions and water around DNA. DNA solvation is displayed here for the first time in terms of superimposed water and counterion densities, and we investigate two issues readily accessible to study: 1) the calculated distribution of counterions vis-à-vis conventional wisdom about counterion condensation phenomena in nucleic acids, and 2) the relationship of a robust theoretical description of first shell hydration to the information provided from crystallography about "ordered waters" and from physical measurements that are conventionally interpreted in terms of "hydration shells."

Recent reports of MD simulations from several other groups relate in various degrees to the study described herein. Cheatham and Kollman (1996) have carried out several nanosecond-length MD's on the decamer duplex sequence $d(CCAACGTTGG)_2$ starting at both A- and B-forms, and they find convergence on the B-form, starting from either canonical structure. A study of $d(CGCGAATTCGCG)_2$ based on the newly proposed version CHARMM Force field has been reported by Yang and Pettitt (1996) and indicates that the new CHARMM supports a model for the *Eco*RI dodecamer closer to the A-form than the B-form at 0.45 M NaCl. D. R. Langley has evidence that CHARMM 23 (version 6.1) (Mackerell et al., 1995) favors the A-form, even at minimal salt, and an extensive reparameterization reports improved results (Langley, 1996). Osman and co-workers (Miaskiewicz et al., 1996; Osman et al., 1991) have studied the effect of a model of thymine photodimerization of the $d(CGCGAATTCGCG)_2$ duplex bending. We present herein an independent characterization of AMBER 4.1 applied to the *Eco*RI dodecamer as a prototype B-form DNA, and investigate further the principal implications of the results, particularly in the realm of solvent effects, i.e., counterion condensation and hydration.

METHODS

A series of four molecular dynamics (MD) simulations were performed on the DNA oligonucleotide $d(CGCGAATTCGCG)_2$ duplex. The principal calculation on which this article is based is a fivefold extension of a trajectory used to elucidate the solvation of the minor groove (Young et al., 1997). Each strand of the duplex contains 11 PO_4^- anions, and each end is terminated with OH groups. The solvent was composed of a solvent bath of ~4000 water molecules and 22 Na^+ counterions, a sufficient number to provide electroneutrality to the system, modeling "minimal salt" ionic conditions. The simulations are thus configured for a condition of high water activity, an environment in which the B-form would be expected to predominate (Wolf and Hanlon, 1975). All MD calculations employed the AMBER 4.1 suite of programs (Pearlman et al., 1995), generously provided to us by Kollman and co-workers, the all-atom force field for the DNA and Na^+ (Cornell et al., 1995), and the TIP3P model for water (Jorgensen, 1981). The MD simulations were performed on the Cray C90 and the T3D parallel computer at the Pittsburgh Supercomputer Center and the CRAY J90 at FCRC of the National Institutes of Health. The starting DNA configuration for each case was the canonical form of the B-DNA double helix (Arnott et al., 1976).

Details of the four simulations are as follows:

1. MCI/PME: The initial placement of the counterions in this simulation was based on an independent Metropolis Monte Carlo calculation of the counterions with the CIMC program (Jayaram et al., 1990a). A CIMC of 22 Na^+ ions around a fixed canonical B-form DNA was carried out, with the effect of water introduced via a sigmoidal distance-dependent dielectric screening function (Hingerty et al., 1985). The minimum energy configuration from a CIMC run of 3.6×10^5 passes (~4 M two-particle moves) on the system was taken as the initial ion placement for the subsequent MD simulation. Notably, one Na^+ in the minimum energy structure was found to be located in the AATT region of the minor groove in a position and to interact favorably with the electronegative A-N7 and T-O_2 atoms. The ultimate disposition of this ion and behavior of the system in the vicinity of this site in the MD are thus a matter of considerable interest. Boundary conditions in the simulation were treated by the PME method. The overall length of the MCI/PME MD trajectory was 5.0 ns.

2. ESI/TRCO: The initial placement of the Na^+ counterions was based on iterative electrostatic (ES) calculations, using the CIONS method of the EDIT program in AMBER. The CION method determines the most favorable position for the first counterion site from the calculated electrostatic potential of the DNA. The process is iterated until all 22 are placed. The lowest energy counterion configuration with all particle interactions taken into consideration does not coincide with that of iterative sequential placement of the ions. In particular, note that the MCIONS protocol results in a cation residing in the minor groove, whereas CIONS does not. The MD trajectory in this simulation was created by using a twin range cutoff (TRCO) option in which the nonbonded pair list is updated every 10 steps for distances shorter than 7.5 Å and every 25 steps for distances between 7.5 and 11.5 Å.

3. ESI/PME: This MD was carried out with the ESIONS protocol as described in simulation 2 above, with PME boundary conditions used in place of the TRCO protocol. The trajectory length was 1000 ps.

4. BSI/PME: In this simulation, the ions were positioned 6 Å from the P atoms along the bisector of the backbone OPO groups. The PME treatment of boundary conditions was used. The length of the trajectory was 500 ps.

The DNA and ions in their initial configurations were hydrated in a rectangular prism by TIP3P water molecules. The box dimensions were truncated to achieve a minimum distance of ~12 Å beyond all DNA atoms in all directions, resulting in a box size of $\sim 45 \times 45 \times 70$ Å, solvating the DNA with ~4000 water molecules in each case. The effective concentration of the DNA sample (calculated by con-

verting the number of particles per box volume to moles/liter) within the periodic box is ~15 mM. Before MD, the various starting configurations were subjected to a series of energy minimizations to relieve any local atomic clashes. All PME calculations were carried out with a 9-Å cutoff for direct space nonbonded calculations and a 0.00001 convergence criterion for the Ewald portion of the nonbonded interactions.

Each MD was initiated with several rounds of semirestrained and eventually unrestrained minimizations of the entire system. Harmonic restraints of 25 kcal/mol·Å^2 were placed on DNA and ion atom positions during the first round of 500 steps of minimization. The restraints were relaxed on the ions more quickly than on the DNA atoms over the course of five subsequent 100-step minimizations. The final round of 500 steps of minimization involved all atoms of the systems. Heating and initial system equilibration was performed in a parallel semirestrained fashion as follows. First, 10 ps of heating was performed on the constant-volume system while the DNA and ion atom locations were restrained. This was followed by reducing the restraints on the ions more quickly than the DNA atoms up to 25 ps. Unrestrained constant-volume dynamics were continued for another 5 ps, at which time constant-temperature, constant-pressure MD (isothermal-isobaric ensemble) was initiated, utilizing the Berendsen algorithm for temperature bath coupling (Berendsen et al., 1984). A 2-fs time step was used for all simulations. The system energy was stable when coupled with a SHAKE constraint of 5×10^{-5}) Å on all covalent bonds involving hydrogen atoms. The MCI simulation became the most informational, as described below, and was extended to a run length of 5.0 ns. The main emphasis in the analysis of results is placed on this simulation, with the results of the other simulations discussed in comparison with the extended length MCI run. The other MDs were carried out for times ranging from 500 ps upward, with run lengths decided on the basis of the potential new information content to be acquired.

RESULTS

Dynamic structure

A series of snapshots extracted at periodic intervals from the MCI/PME MD is shown in Fig. 1. Overall, the duplex is seen to be intact over the course of the simulation. In overall physical appearance, the MD structure conforms closely to that expected of a B-form of DNA subject to thermal fluctuations. This is significant, because previous results from earlier versions of the AMBER force field (3.0) (Weiner et al., 1986) tended to stabilize around structures intermediate between the A- and B-forms of DNA. A view of the calculated dynamic structure of the Amber 4.1 DNA based on thermal ellipsoids is shown in Fig. 2. This representation shows clearly that the dynamic range of motion with the AATT tract is relatively restricted compared with that of the CGCG flanking sequences, and that a larger range of dynamic motion is exhibited in the region of the junctions between the AATT and CGCG tracks, particularly in the nucleotide bases. A dodecamer is short enough that the influence of end effects, including fraying at the ends and effects on structure propagated throughout the molecule, could be a problem (Olmsted et al., 1991). However, 12-mers are currently the longest of oligonucleotides to be studied extensively by crystallography in unliganded form. In the MCI/PME simulation, Watson-Crick base pairing is well maintained throughout the simulation. The nucleotide units at both the 5′ and 3′ ends of the sequence, as indicated by larger thermal parameters, exhibit a larger range of dynamical motion, but the effect does not appear to be propagated to a significant extent into the interior of the structure.

MD stability and convergence

Convergence and stability profiles for the 5.0-ns MCI/PME MD model are shown in Fig. 3. The root mean square (RMS) deviations of MD structure as a function of time with respect to canonical B DNA, canonical A DNA, and MD average structure are shown in Fig. 3 (*top*). The RMS deviation is computed from the mass-weighted mean square difference of all DNA atoms with respect to the starting structure (ions and water excluded). The results indicate that the MD structure transits quickly in the simulation to ~3.6-Å RMS deviation from the starting structure, and exhibits an irregular but reasonably stable oscillation centered about a distance of ~2.8-Å RMS deviation from the starting structure throughout the remaining course of the simulation. The all-atom RMSD of the average structure over the entire simulation is 2.71 Å from the canonical B-form, and 2.88 Å from the crystal structure of the sequence. The range of oscillation of the dynamical structure around the MD average is ~1.5 Å.

The corresponding two-dimensional (2D RMS) time series is displayed in matrix form in Fig. 3 (*bottom left*). Entries in the off-diagonal triangles of the matrix represent the RMS deviation of each structure in the simulation with every other. Shaded areas indicate regions of low RMS (<2.0 Å), with the intensity of the shading inversely proportional to RMS deviation. During the first 500 ps of the simulation, there are no cross-peaks in the 2D RMS diagram, indicating that the MD structure executes a concerted change away from the canonical form. Beyond 500 ps, considerable cross-peak activity is observed. The results in Fig. 3 (*bottom left*) indicate that equilibration, the period of the simulation before oscillation sets in, is achieved at ~500 ps. Some degree of structural correlation is implied over the remainder of the trajectory, indicative of a system oscillating in an broad energy basin. Within this basin, a series of blocks of ~500-ps length are observed as especially low-RMSD (<1.0 Å) regions. There is considerable indication of cross-peaks between the blocks, indicative of a dynamic oscillation of the structure among thermally accessible sub-

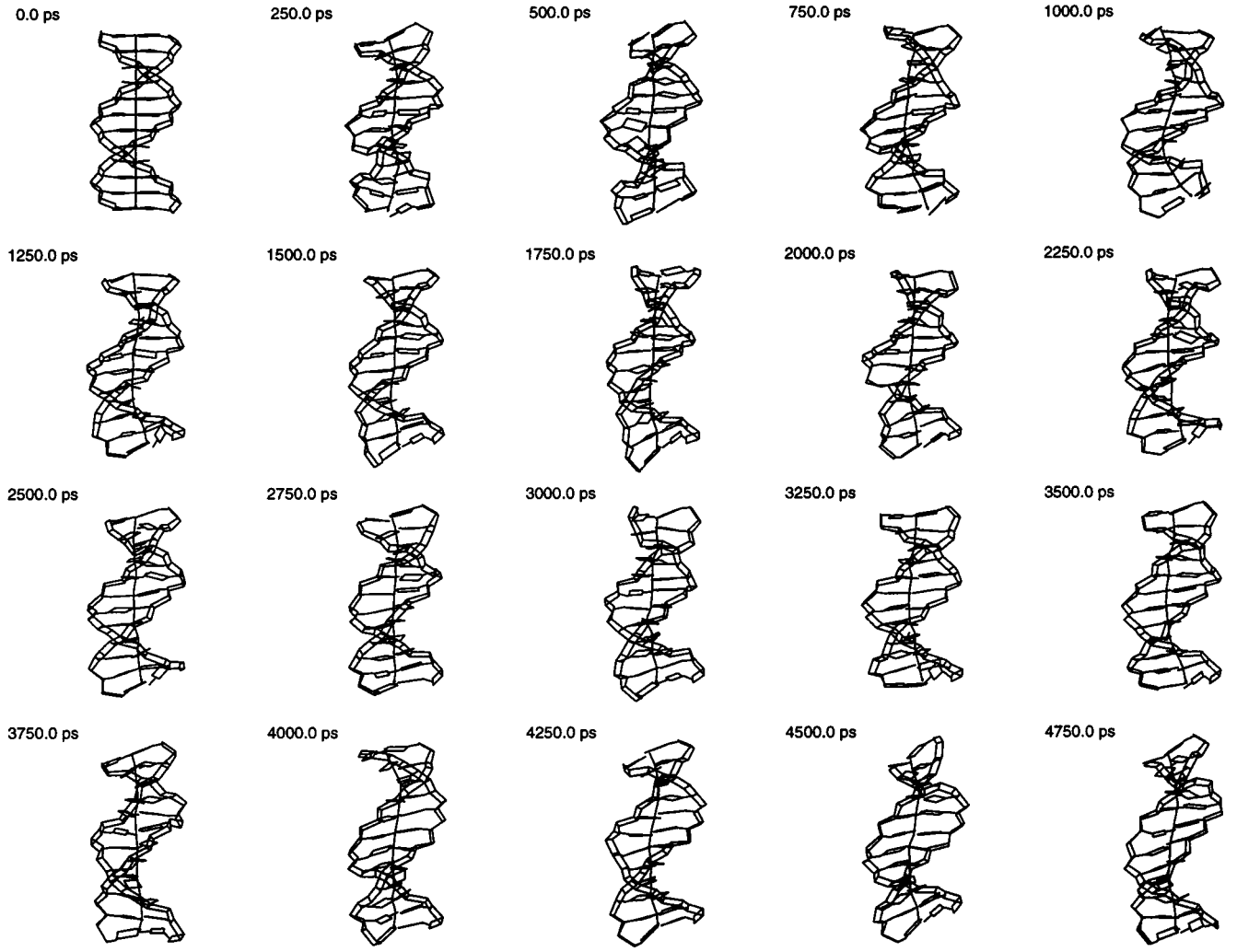


FIGURE 1 A series of 20 DNA schematic snapshots of from the 5.0-ns MCI/PME MD on the *Eco*RI dodecamer based on the MCI starting structure. Counterions and waters are removed for clarity.

states or at least an incipient tendency toward this behavior in the model system.

Conformational analysis

A composite view of the calculated conformational dynamics is shown in Fig. 4 (*top row*). A set of conformational plots is presented, one for each unique torsion angle in the DNA backbone (α, β, γ, δ, ϵ, ζ, χ) and sugar pucker P. These plots show a superposition of results for each step in the dodecamer, and thus generic rather than sequence-specific behavior. The shaded areas in these dials are the range of values of each torsion that have been deduced from independent model building considerations (Olson, 1982a,b) to be sterically unfavorable. The MD results strongly support the predicted conformationally allowed regions. The distribution of sugar puckers extends significantly from C2′ endo into the adjacent C1′ exo region throughout, with the C2′ endo of canonical B-DNA clearly predominating. Corresponding results from the available crystal structure data on A- and B-form oligonucleotide structures are shown in Fig. 4 (*middle* and *bottom*, respectively). Fig. 4 (*middle*) shows the superimposed results for conformational distributions from 94 B-form DNA crystal structures. The distributions of the MD calculated and crystal structure results in Fig. 4 (*middle*) are markedly similar and clearly distinct from the conformational distributions found in 47 A-form oligonucleotide crystals, shown in Fig. 4 (*bottom*). Note, however, that the MD results in Fig. 4 (*top*) come from a time series on the single sequence $d(CGCGAATTCGCG)_2$, whereas the results in Fig. 4 (*middle* and *bottom*) are a composite of time-averaged results on diverse sequences belonging to the B and A families, respectively. The results of Fig. 4 show that the MD model exhibits a dynamic range of values consistent with the distribution of B-form crystal structures, evidence that the MD model behaves reasonably well in describing B-form DNA.

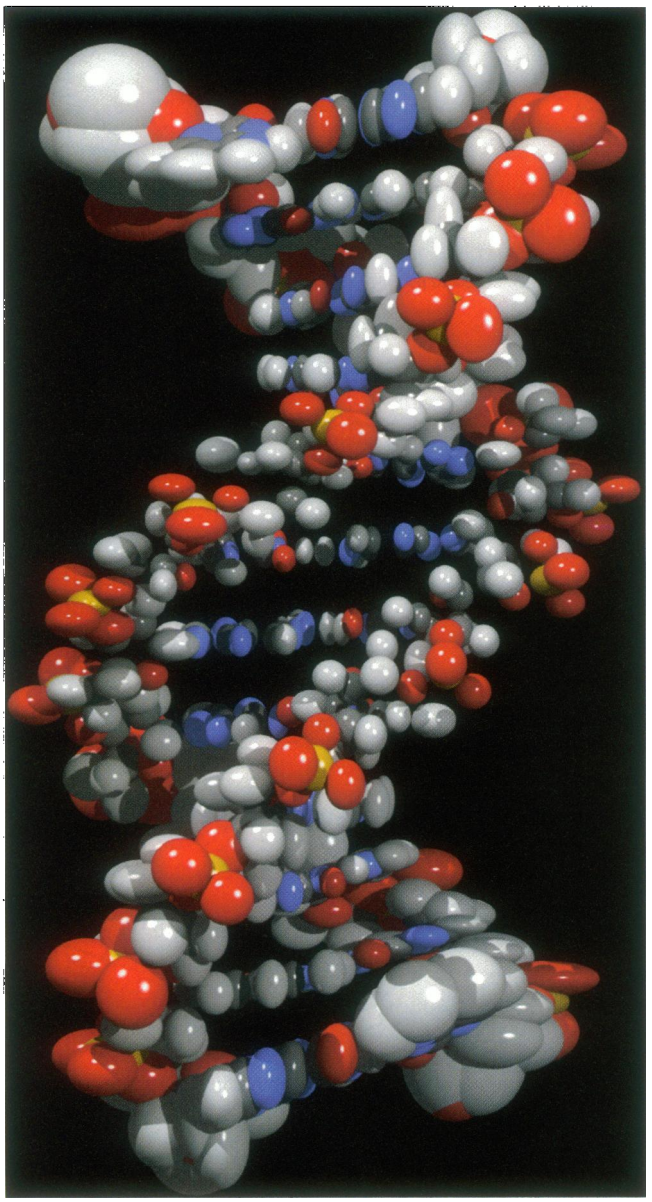


FIGURE 2 Dynamical structure of the 5.0-ns MCI/PME MD on the *Eco*RI dodecamer, depicted via thermal ellipsoids positioned on the average MD structure. The ellipsoids describe the extents of one standard deviation of atomic motions.

A series of more detailed figures, labeled W1–W7, have been made available as a supplement accessible to the interested reader on the Internet at the Biophysical Society web site. In summary, the results reveal the parameter δ of the furanose ring to be relatively stable over the time course of the simulation and oscillating freely in the sterically permitted region in the vicinity of (+) ac to (+) ap. (We use the terms *gauche* (*g*) and *trans* (*t*) to refer to broad regions of conformational space, and use the Klyne-Prelog notations such as synclinal (sc), anticlinal (ac), and antiperiplanar (ap) when more specific designations are called for.) In B_I to B_{II} transitions in DNA, the ϵ and ζ parameters are coupled and transit from (t, $g-$) to ($g-$, t). At ~3 ns, the MCI/PME MD shows distinct B_I to B_{II} transitions at the G2-C23 step on both strands simultaneously, and at G4. There is B_I to B_{II} activity at several other steps during the course of the trajectory, including C9, G10, and especially at T19, the center of the *Eco*RI recognition site. There is clearly more overall backbone flexibility at the ApT step, which is a point of considerable distortion of the DNA sequence in the *Eco*RI endonuclease crystal structure.

The phosphodiester torsions α and γ and the backbone angle β are relatively stable, except at the position noted above for B_I to B_{II} activity, T19. At this step there is also considerable correlated conformational activity between α and β, whereas γ assumes an uncharacteristic value of t rather that the usual g^+. The exocyclic parameter χ hovers throughout the trajectory in the region of $ap(-)$, displaced slightly from the canonical B value. Correlated transitions in α and γ, separated by β, define a "crankshaft motion " of the conformational dynamics, compensatory changes in two individual torsion angles that take place in such a way that the helix is conserved. Whereas crankshaft motions were observed in previous MD on DNA (Beveridge and Ravishanker, 1994), the current trajectory shows no examples of this motion and exhibits oscillatory motions about $\alpha = g-$, $\beta = g+$. One notable feature of the conformational dynamics is the behavior at the ApT step, in which α and γ tend to *trans* extended forms.

In the pseudorotation dynamics of the sugars, the $C_{2'}$ endo structure characteristic of B-form of DNA is favored. However, the sugar rings show considerable dynamical excursion into the $C_{1'}$ exo and $O_{4'}$ endo regions over the course of the simulation. There is transient repuckering to the $C_{3'}$ endo value characteristic of an A-DNA, but little stability of sugar puckers beyond the region between $C_{2'}$ endo and $O_{4'}$ endo. The sugar puckering remains dynamically active, even at the termination of the trajectory.

Helicoidal analysis

A set of 20 helicoidal parameters are involved and are divided into four groups: axis-base pair (4), intrabase pair (6), interbase pair (6), and axis-junction parameters (4). We omit consideration of the latter because this information is effectively communicated through other parameters in the analysis. For each MD structure analyzed, a global helical axis is obtained by using the CURVES procedure (Lavery and Sklenar, 1988), which minimizes a function that describes simultaneously the change in orientation between successive nucleotides and the nonlinearity of the helicoidal axis. The optimal helicoidal solution produces the overall helical axis that best fits the conformation considered, with structural irregularity distributed between changes in the axis-base pair parameters and axis curvature. The distributions of the axis-base pair, intrabase pair, and interbase pair helicoidal parameters are shown in Figs. 5, 6, and 7, respectively. Each of these plots includes the corresponding crystal structure data for B-form and A-form DNA in a manner

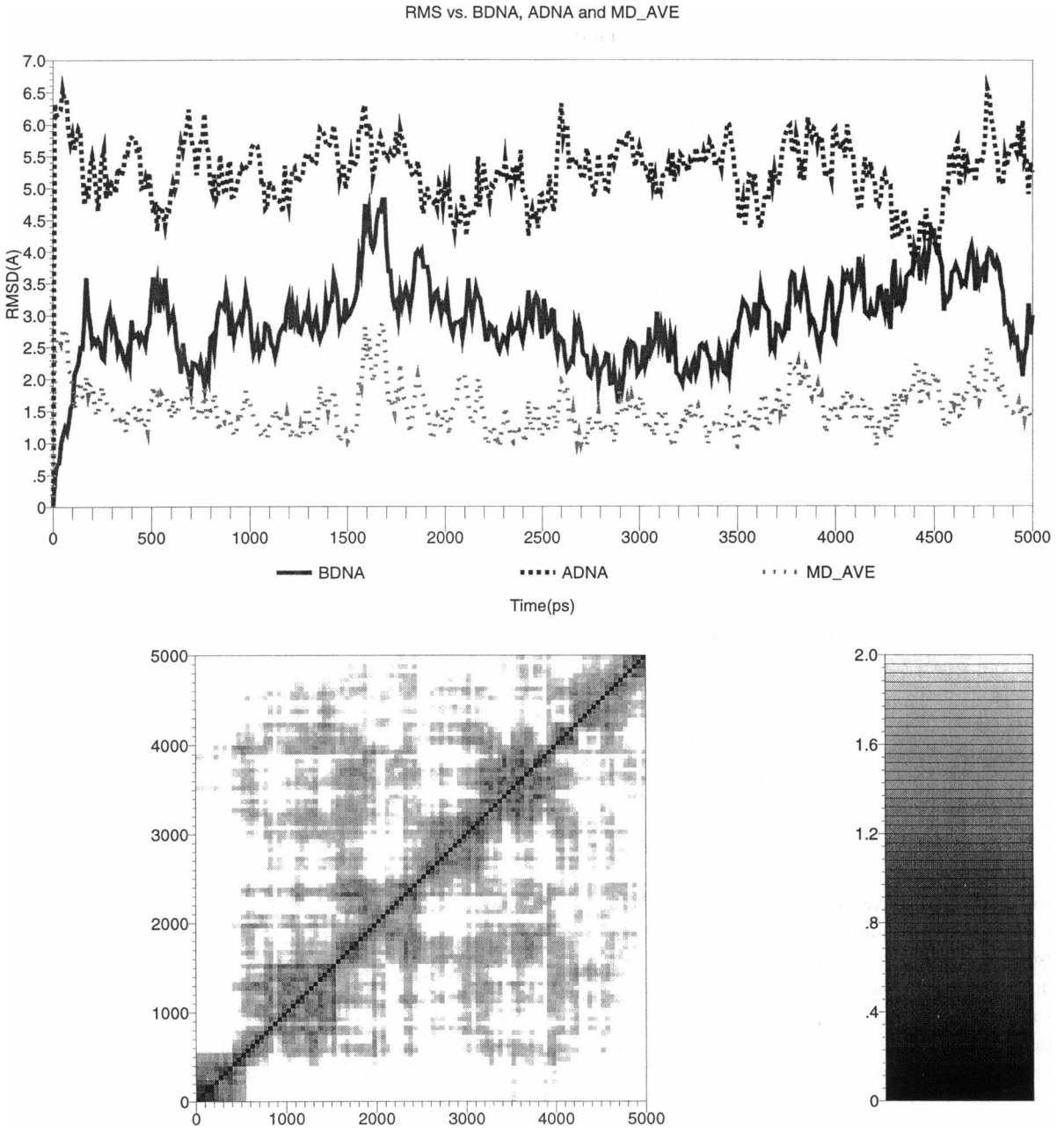


FIGURE 3 Mass-weighted root mean square deviation (RMSD) analysis of the *Eco*RI dodecamer as a function of time for the 5.0-ns MCI/PME/MD trajectory. (*Top*) (1-D RMS) for MD structures with respect to the canonical B-DNA starting structure (*middle trace*), canonical A-DNA (*top trace*), and the MD average for the dodecamer (*bottom trace*). (*Bottom*) RMSD of all MD structures with respect to one another (2-D RMS). The intensity of the shading is inversely proportional to the calculated RMS.

analogous to that in Fig. 4. The MD distributions of helicoidal parameters are in all cases similar in appearance to that from the results of B-form crystal structures, but the MD distributions generally tend to be broader. A minor discrepancy of note is the tendency of the calculated distribution of roll angles in the MD to slightly more positive values (compression of the major groove) than the composite results from the B-form crystal structures.

The subset of helicoidal parameters XDP, INC, PRP, RIS, and TWS are critical in distinguishing B-form from A-form DNA. For XDP, which is 0.0 Å in canonical B DNA, the MD average is −0.92 Å, as compared with −0.07 Å in B-form crystal structures and −3.18 Å in A form crystal structures and −5.2 Å in canonical A DNA. The distributions for B-form crystal structures values span the canonical B value, with the peak slightly below that of the canonical B-form value of 0.0 Å. The MD results for $d(CGCGAATTCGCG)_2$ show a similar tendency, distinct from the A-DNA results in Fig. 6 (*right column*), and confirm the B-form character of the MD model. For INC,

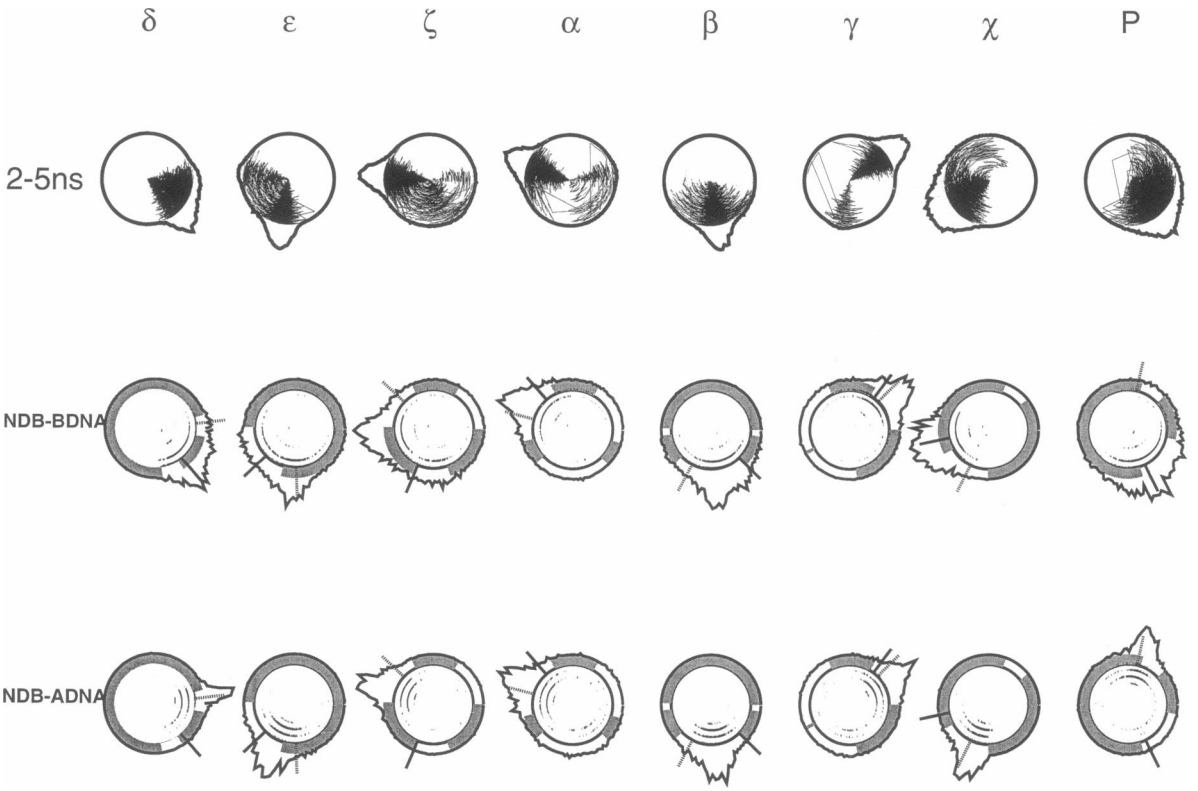


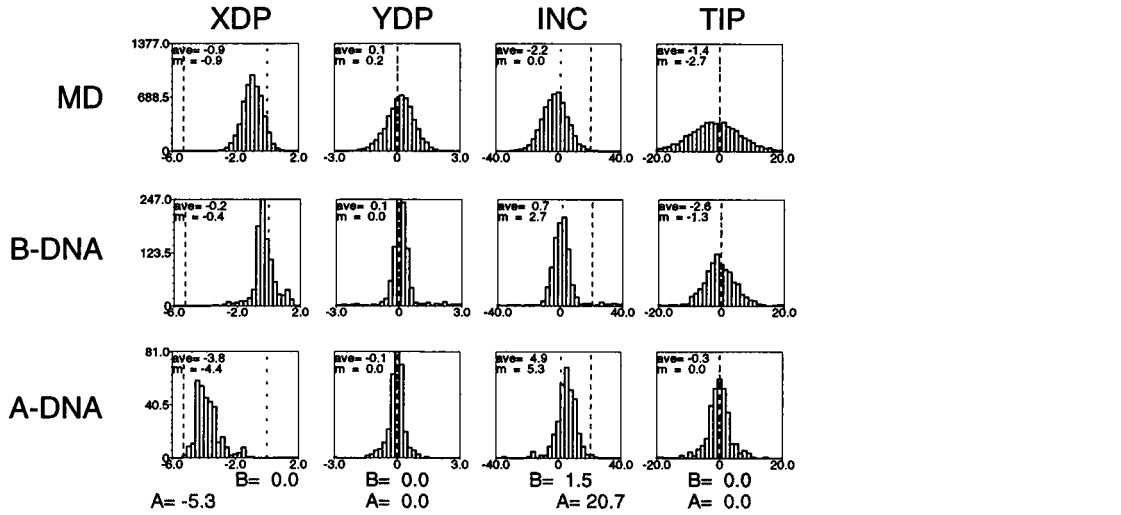
FIGURE 4 Comparison of MD calculated and crystallographically observed distributions of conformational angles and sugar puckers. (*Top row*) Conformational "dials" for the collected results from the 5.0-ns MCI/PME/MD trajectory. The outer ring provides a summation of the collected results from the various MD structures, presented as a function of angle, i.e., a representation of probability density. (*Bottom row*) Conformational dials for 94 B-form oligonucleotide crystal structures. The results for oligonucleotides of various lengths are presented on the radial coordinate of the dial as a function of length from dinucleotides (*inner radius*) to dodecamers (*outer radius*). The outer ring provides a summation of the collected results from the various crystal structures as a function of angle. Regions identified as sterically unfavorable (Olson, 1982a,b) are indicated by shading.

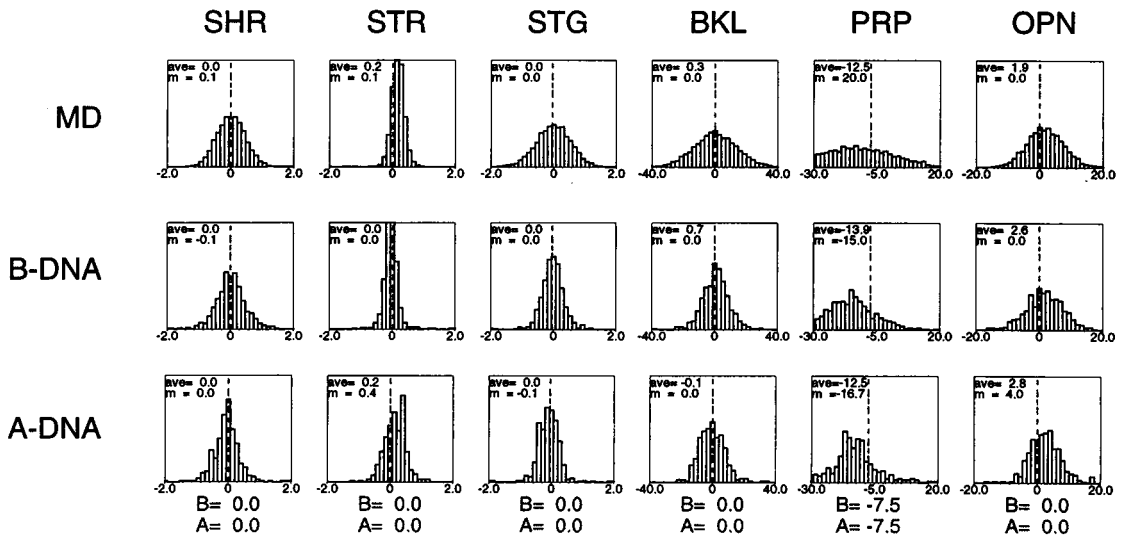
RIS, and PRP, the B-form MD, B-form crystal structures, and the canonical B values are closely aligned with each other and distinct from results from A-DNAs (Fig. 6, *right column*). The distributions of TWS in the MD values and the crystal forms show a close correspondence which spans that of the reference values of canonical B- and A-DNA and extends somewhat above and below the canonical values. The distribution peak, however, is in the vicinity of the canonical B and A values but slightly less than the 34° value corresponding to 10.5 bp per helical turn. The results indicate that a wide range of TWS angles is possible in both MD and crystallographic models of B-DNA and hence that they are not reliable diagnostic indexes of B versus A forms.

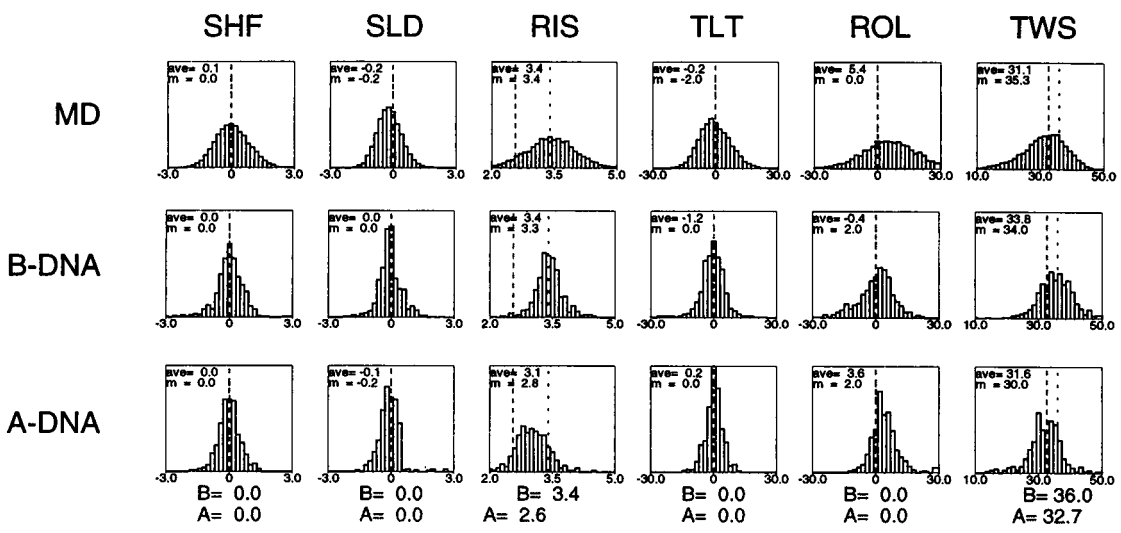
Detailed analysis of the time series of the helicoidal parameters for the dodecamer are also provided in the Supplement, Figs. W3–W5. In summary, on the basis of XDP, the dodecamer definitely remains in the B range over the entire MD trajectory and averages out to a value just under that of canonical B. The parameter YDP is stable, with incipient grand cycle behavior appearing toward one end of the sequence. The parameters INC and TIP specify the orientation of the base pairs with respect to the helical axis. The INC parameter tends to values that are in the vicinity of zero or slightly positive in the CG regions, and slightly more negative than that of canonical B in the AATT region. The parameter TIP is dynamically stable. The distance parameters shear (SHR), stretch (STR), and stagger (STG) are all relatively stable. A sequence dependence in the propeller (PRP) twist of the base pairs is observed in the simulation, with PRP values of the AATT region slightly but distinctly more negative, a probable consequence of the three Watson-Crick hydrogen bonds for GCs versus two for ATs. This is not, however, a pronounced effect. Shift (SHF), slide (SLD), and rise (RIS) are oscillatory and stable in cycles or grand cycles. The parameter ROL shows marked displacement at or near the junctions of the AATT and CGCG tracts, with compensatory behavior extending over two base pair steps. Base pair RIS in the dynamical structure remains in all cases at or near the value for canonical B DNA. The values of TWS remain essentially close to those of canonical B-DNA, although an incipient underwinding can be discerned, which is often correlated with increases in ROL magnitude, as observed in DNA crystal structures (Dickerson, 1991; Gorin et al., 1995). Sequence-dependent displacement in TWS is seen at G4, T7, and G10.

Helix morphology

The morphology of the DNA duplex can be described in terms of two essential characteristics: axis bending and


XDP
YDP
INC
TIP
MD
B-DNA
A-DNA
B= 0.0
A= -5.3
B= 0.0
A= 0.0
B= 1.5
A= 20.7
B= 0.0
A= 0.0


SHR
STR
STG
BKL
PRP
OPN
MD
B-DNA
A-DNA
B= 0.0
A= 0.0
B= 0.0
A= 0.0
B= 0.0
A= 0.0
B= 0.0
A= 0.0
B= -7.5
A= -7.5
B= 0.0
A= 0.0


SHF
SLD
RIS
TLT
ROL
TWS
MD
B-DNA
A-DNA
B= 0.0
A= 0.0
B= 0.0
A= 0.0
B= 3.4
A= 2.6
B= 0.0
A= 0.0
B= 0.0
A= 0.0
B= 36.0
A= 32.7

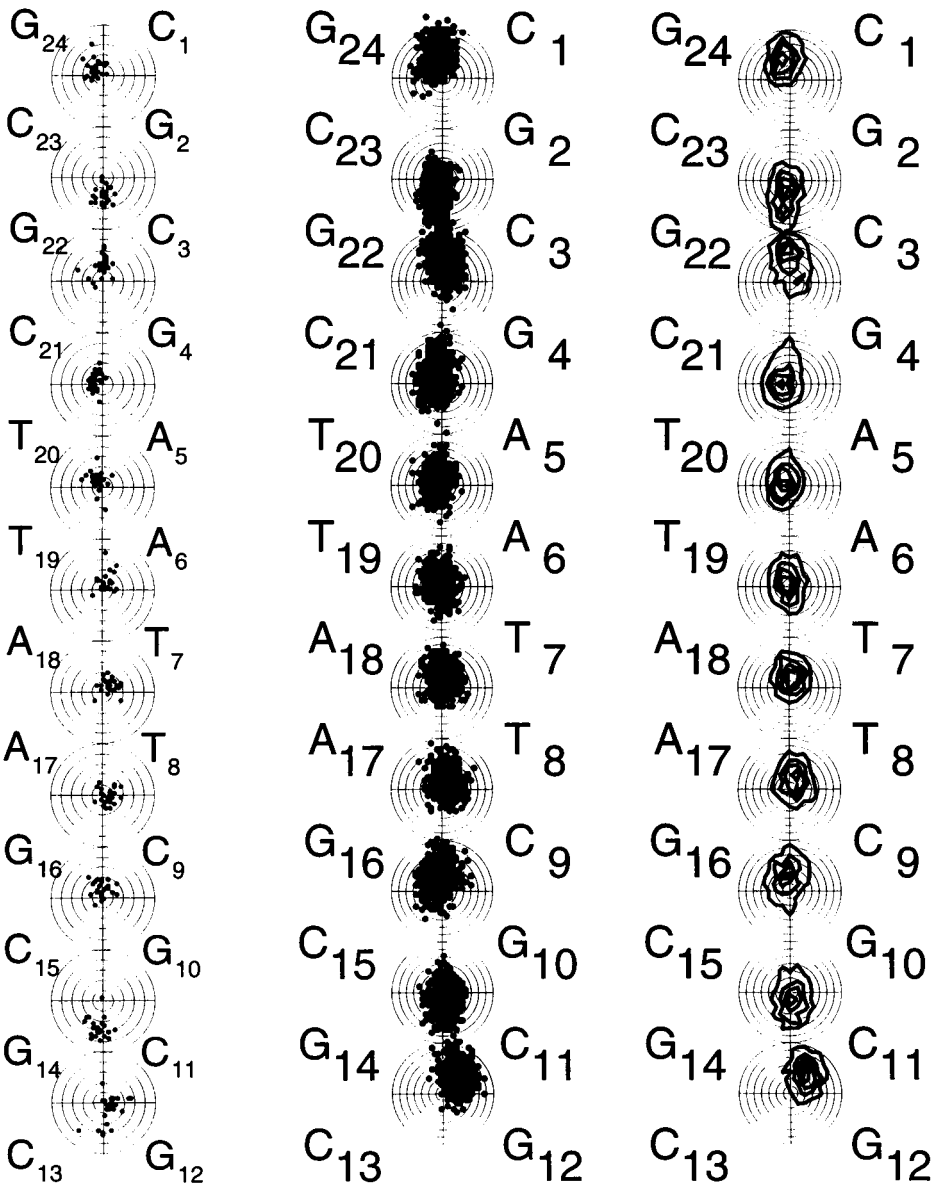


FIGURE 6 Analysis of local axis bending computed for the *Eco*RI dodecamer from the 5.0-ns MCI MD trajectory, presented in terms of "bending dials" (Young et al., 1995a). (*Left column*) Bending observed at each step in 25 reported crystal structures of the *Eco*RI dodecamer under various conditions of temperature and complexation. (*Middle column*) Bending at each base pair step, from a superposition of 500 snapshots taken at equally spaced intervals over the time course of the simulation. (*Right column*) Calculated bending presented in terms of probability density contour maps. At each step bending toward the major groove is plotted in the northern hemisphere, and bending toward the minor groove is plotted in the southern hemisphere.

groove widths. The analysis of axis bending in crystallographic DNA helices has been the subject of a recent study (Young et al., 1995a). The magnitude of bending θ and directionality ϕ at a given base pair step were computed from values of the inter-base pair helicoidal parameters ROL and TLT, and the results were displayed on a polar "bending dial," in which θ is the radial coordinate and ϕ is the angular coordinate. This same approach may be used to analyze the axis bending of a dynamical model as obtained from MD simulation (Young et al., 1994). Analysis of the axis bending in the AMBER 4.1 MD model is shown in Fig. 6 (*middle*) in a time series representation and in Fig. 6 (*right*) as probability contours. The loci of bending are at or near the junctions of the AATT and CGCG tracts, and support a "junction model" of axis bending as opposed to a "wedge model." The junction bend is composed primarily of displacements in roll toward the major groove at G2 and C3. The bending locus at the 3′ junction, the C9 and G10

FIGURE 5 (*Top*) Comparison of MD calculated and crystallographically observed distributions of axis-base pair parameters. *Top row*: Results from the 5.0-ns MCI/PME/MD. *Middle row*: Results from 92 B-form oligonucleotide crystal structures. *Bottom row*: Results from 42 A-form oligonucleotide crystal structures. (*Middle*) Comparison of MD calculated and crystallographically observed distributions of intra-base pair parameters. *Top row*: Results from the 5.0-ns MCI/PME/MD. *Middle row*: Results from 92 B-form oligonucleotide crystal structures. *Bottom row*: Results from 42 A-form oligonucleotide crystal structures. (*Bottom*) Comparison of MD calculated and crystallographically observed distributions of inter-base pair parameters. *Top row*: Results from the 5.0-ns MCI/PME/MD. *Middle row*: Results from 92 B-form oligonucleotide crystal structures. *Bottom row*: Results from 42 A-form oligonucleotide crystal structures. Experimental data obtained from the Nucleic Acids Data Base (Berman, 1992); individual citations provided previously (Young et al., 1995a).

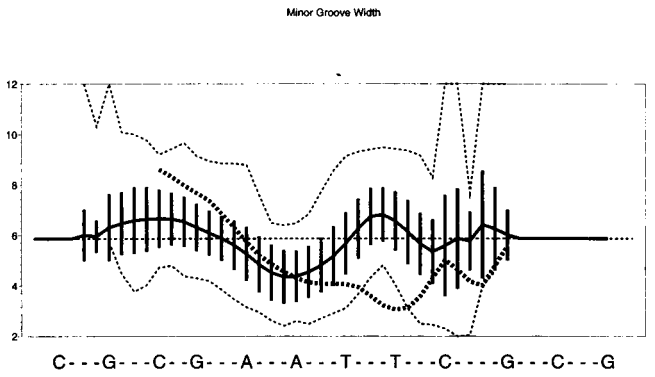


FIGURE 7 Minor groove width calculated by the Curves program, reporting the closest distance at each step between splines fit through the backbone atoms (Stofer and Lavery, 1994). The plots show the extents of instantaneous values (*thin dashed lines*), the average value (*thick solid line*), and one standard deviation over trajectory (*vertical bars*). Values for the native crystal structure of the dodecamer are plotted in the thick dashed line (end base regions could not be calculated).

steps, has a slightly higher component of tilt than the locus in the 5′ junction, at G2 and C3.

The bending dials for 25 crystal structure determinations of the sequence d(CGCGAATTCGCG)$_2$ under various conditions of temperature and complexation are shown in Fig. 6 (*left*). The crystallographic structures behave in a consistent manner and show bending toward the minor groove at the G2-C3 step, followed in the succeeding step (G3-C4) by a bend toward the major groove. A behavior similar to that at the 5′ junction is seen at the 3′ junction. The MD model exhibits bending in the same general region as that of the crystal structures, supporting the idea of a propensity toward bending in this region. The nature of the bending in the MD model is primarily toward the major groove, producing a discrepancy at the level of fine detail between calculated and observed values. The effect of crystal packing, which can be significant in DNA oligonucleotides, is not considered in the MD model, and thus the results would not be expected to correspond perfectly. The bending tends toward more palindromic symmetry in the MD than in the crystal structures, consistent with observed solution behavior.

The dynamic structure of the major and minor groove width groove as described by the AMBER 4.1 MCI/PME MD, along with corresponding results for crystal and canonical forms, is shown in Fig. 7. The presentation is based on the backbone spline-fitting algorithm developed by Lavery et al. (Stofer and Lavery, 1994) and available within Curves 5.1 (Lavery and Sklenar, 1996). The plot contrasts results from the crystal structure against both the average values from the MD as well as the dynamic range of values spanned throughout the entire trajectory. The AMBER 4.1 MD structures show a clear tendency toward minor groove narrowing in the region of the AATT tract, although it is not quite as pronounced as that of the crystalline form of the dodecamer. The dynamic range covered throughout the MD demonstrates that the helix is capable of undergoing significant minor groove width "breathing" within the 5.0-ns time frame of the simulation.

Results from other runs

The AMBER 4.1 ESI/TRCO simulation, after less than 250 ps, was found to result in a seriously destabilized double helix, underwound and elongated to the point of groove collapse. Thus this option does not result in a reasonable dynamic model of the B-form of DNA. The AMBER 4.1 ESI/PME simulation was carried out for 500 ps and resulted in a dynamic structure much improved over that of ESI/TRCO, indicating that the treatment of long-range electrostatic forces has a significant influence on the stability of the simulation, consistent with previously reported results carried out with this same force field and utilizing both truncated potentials and PME (Cheatham et al., 1995). The Amber 4.1 BSI/PME simulation was carried out to 1000 ps, at which time the structure and the behavior of the ions were determined to be consistent with those of MCI/PME and ESI/PME. Overall, the PME results give an improved description of the DNA as compared with truncated potentials we are currently considering or have previously analyzed. With regard to simulation stability, our results indicate that the general helicoidal and morphological characteristics of the dodecamer are not dependent on the initial ion placement in the slow heating/PME protocol utilized in this study.

DISCUSSION

Comparison of MD with experiment

The MCI/PME simulation for the AMBER 4.1 model for the d(CGCGAATTCGCG)$_2$ duplex supports a remarkably stable model of B-form DNA over the 5-ns MD trajectory. The resulting dynamic structure is closer to that of B-DNA than any of the previous models from the diverse force fields we have examined to date. On the experimental side, the B-form structure for this sequence in a high water activity environment is supported by crystallographic (Drew et al., 1981, 1982), NMR (Nerdal et al., 1989; Withka et al., 1992), and Raman spectroscopy (Benevides et al., 1988; Weidlich et al., 1990). Essential features of axis bending and sequence effects on groove widths, as discussed in the previous section, are adequately reproduced. Thus we are led to the opinion that the AMBER 4.1 force field supports an accurate model of B-form DNA under the conditions of this simulation.

The average MD structure from the 5-ns trajectory and the corresponding crystal structure are shown side by side in Fig. 8. The correspondence between the calculated and observed results is visually quite close. The mass-weighted heavy atom RMSD of the two structures is 2.50 Å, which is quite reasonable, especially considering that the MD was not performed on the crystallographic unit cell. The expected palindromic symmetry of the DNA sequence is more

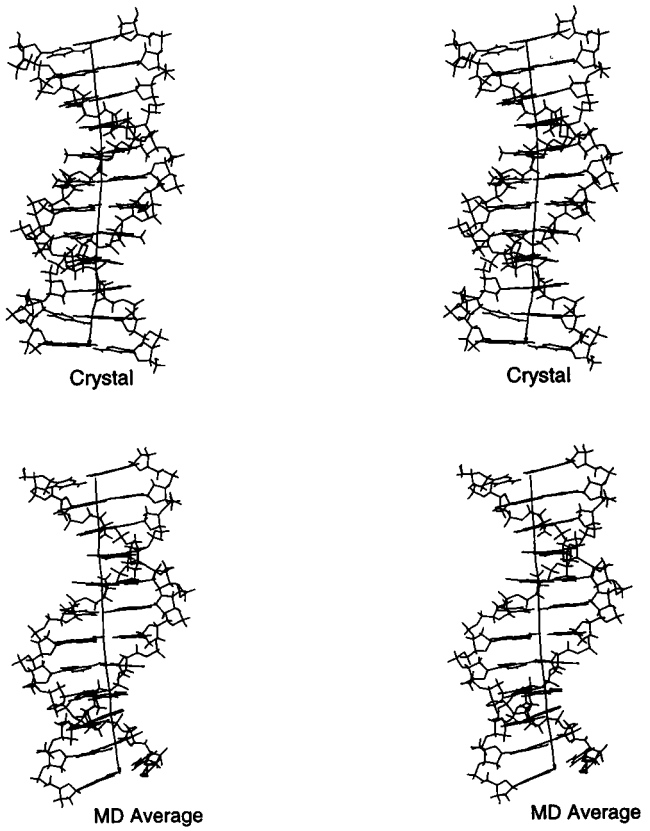


FIGURE 8 Comparison of MD calculated and observed structures for the *Eco*RI dodecamer. (*Top*) Average structure computed from the 5.0-ns MCI/PME/MD. (*Bottom*) Crystal structure (Drew et al., 1981).

strongly represented in the structure generated from the MD trajectory than is seen in the crystal structure. The RMS value between the top 6 bp and the bottom 6 bp of the MD average structure is 0.72 Å, versus 1.02 Å in the crystal structure. Examined more closely, the mass-weighted heavy atom RMS values between the MD-average structure and the crystal structure for the upper 6, middle 6, and lower 6 bp are 1.26, 1.46, and 1.87 Å, respectively. The extremely close agreement between the top half and the center of the molecule implies that the crystal structure of these base pairs is most likely closer to the solution structure than that of the bottom flanking base pairs, which are perhaps more directly influenced by packing contacts in the crystal (Dickerson et al., 1987, 1991; Narayana et al., 1991). Thus the even better agreement between the top 6 bp indicates that the top half of the crystal structure is expected to be a better prediction of solution behavior than the bottom half.

Comparison of data from our MD structure with the NMR study by Lane et al. (1991) shows close correspondence between the calculated and the observed pseudorotation phase angles, with sugars falling predominantly in the range of 130–180° (C2′ endo and O1′ exo), with a modest admixture of C3′ endo. Of the helicoidal parameters for which experimental data are reported, there is close agreement between calculated and observed values of SHF, SLD, and RIS. The MD model shows a tendency to be slightly underwound relative to the NMR results, with an average TWS of 32.5° from MD versus 36.8° reported by Lane et al.

Ion solvation and counterion condensation

All Na^+ counterions in all four of the reported dynamics trajectories were observed to execute some degree of diffusive motion, and no ions ultimately appeared to be "stuck" or exhibited other behavior indicative of ergodic problems. The general stability of the B-form DNA helix is maintained in all three of the simulations employing PME, indicating that the dynamic stability of the DNA is not dependent on starting ion configuration. A plot of the calculated DNA-Na^+ cylindrical distribution function $g(R)$ and running coordination number $N_C(R)$ for the MCI/PME MD trajectory is presented in Fig. 9. This plot shows inflection points at 7.5 Å and 21 Å, with the inflection at 21 Å defining the Manning radius for this model. The area under the curve indicates that 76% of the counterions are described by this model as "condensed," in good agreement with Manning's prediction (Manning, 1978). The presence of the inflection point at 8.5 Å in the MCI MD RDF indicates that counterions found within the groove region of the double helix in MC and PB calculations are also present to a significant extent the all-atom, explicit solvent, dynamic model from MD in which the DNA is permitted to execute thermal motions. A view of the calculated ion atmosphere of DNA from the MD is shown in Fig. 10.

An animation of the dynamic motion of the system as described by the MD was prepared with the program MPEG and is deposited on the BJ web page. The ion initially positioned in the ApT pocket in the MCI placement equilibrated between and roughly equidistant from the T-7 and T-19 O_2's of the central ApT step in the sequence. An MD snapshot of the solvated DNA from this segment of the MCI trajectory is presented in Fig. 11 and clearly shows the Na^+ ion in proximity to the two O_2 atoms from the central T-7 and T-19 bases on opposite strands of the duplex as well as the oxygen atoms of four water molecules.

At ~500 ps in the MCI trajectory, the Na^+ in the ApT pocket was observed to exchange positions with a nearby water molecule. The ion was subsequently displaced further away from the minor groove, in concert with the motions of a number of diffusing water molecules. An MD snapshot from this segment of the trajectory is shown in Fig. 12. The solvation in the minor groove region at this point corresponds to an intact spine of hydration. The Na^+ previously residing in the ApT pocket is now associated with the DNA backbone. This situation was observed to persist for the next 500 ps of the simulation. At ~1 ns into the trajectory, another Na^+ ion was observed to penetrate the minor groove region and situated itself in the second hydration shell of the DNA, and an Na^+ ion had also become situated in the 3′-CpG region, an intrastrand pocket.

At ~4.0 ns into the simulation, a Na^+ ion reentered the AATT region of the minor groove, situated itself at the TpT

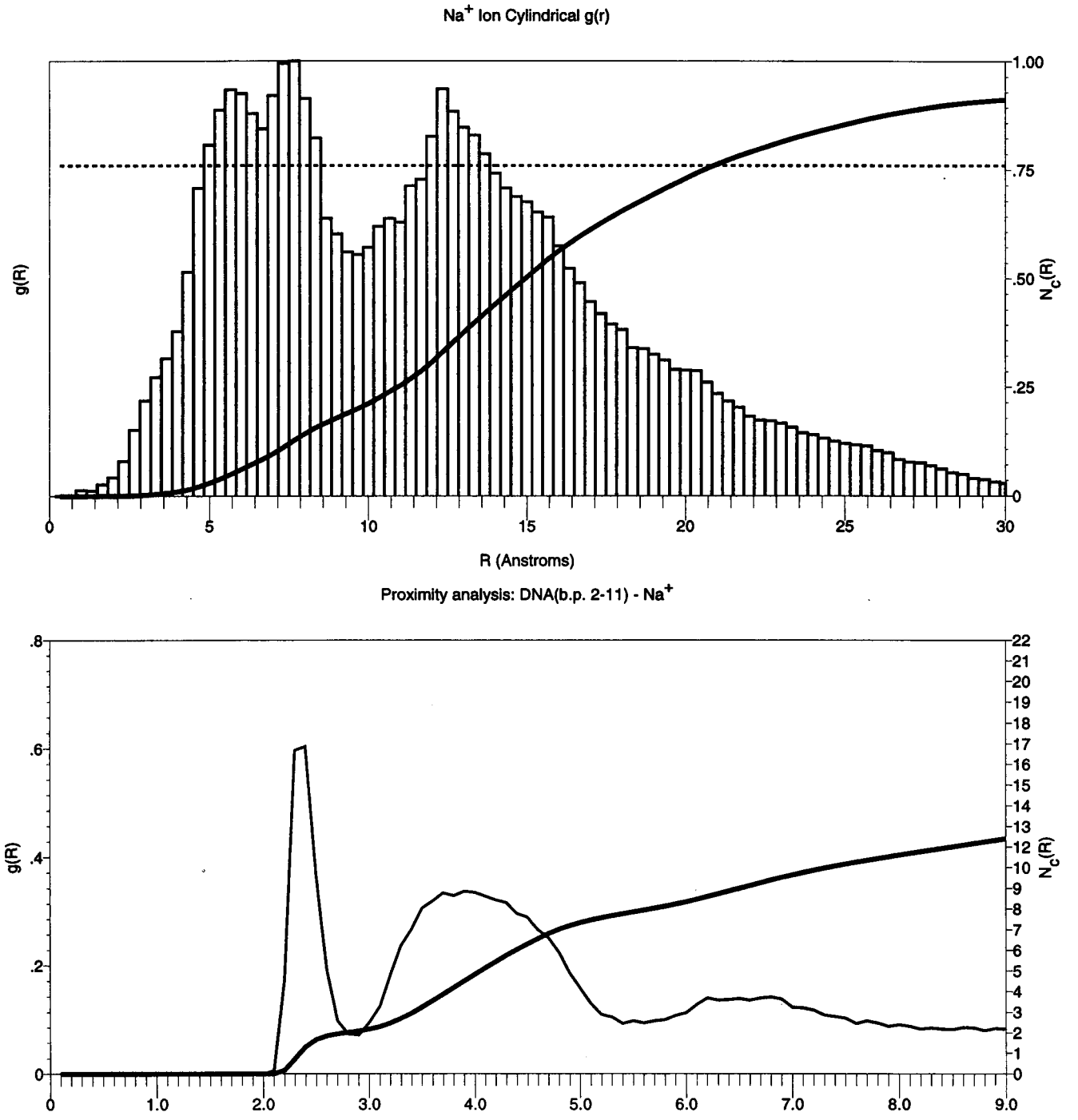


FIGURE 9 (*Top*) Calculated DNA-Na^+ cylindrical distribution function $g(r)$ (*histograms*) and running coordination number $N_c(r)$ (*solid line*) of the Na^+ ions around the DNA helical axis. (*Bottom*) Calculated DNA-Na^+ distribution function and running coordination number calculated via the proximity analysis method (Mehrotra and Beveridge, 1980) (to eliminate end effects, data are calculated for the central 10 bp only).

(ApA) step, and remained there for ~750 ps, at which point it exchanged positions with a water molecule. Further along in the minor groove, the Na^+ that became situated in the CpG step at 1 ns into the trajectory was found to reside for more than 1 ns. Obviously this means that trajectories of multiple nanoseconds are required to obtain a proper description of the dynamics of monovalent ions, and we make no claim that this aspect of the simulation is fully stabilized, even at 5.0 ns. Describing the dynamics and equilibrium properties of DNA systems containing multivalent ions, which is also of considerable interest (Bloomfield, 1991), will require even longer trajectories.

Referring back to Fig. 10, there are two other regions in which high counterion density is indicated: the CpGpA

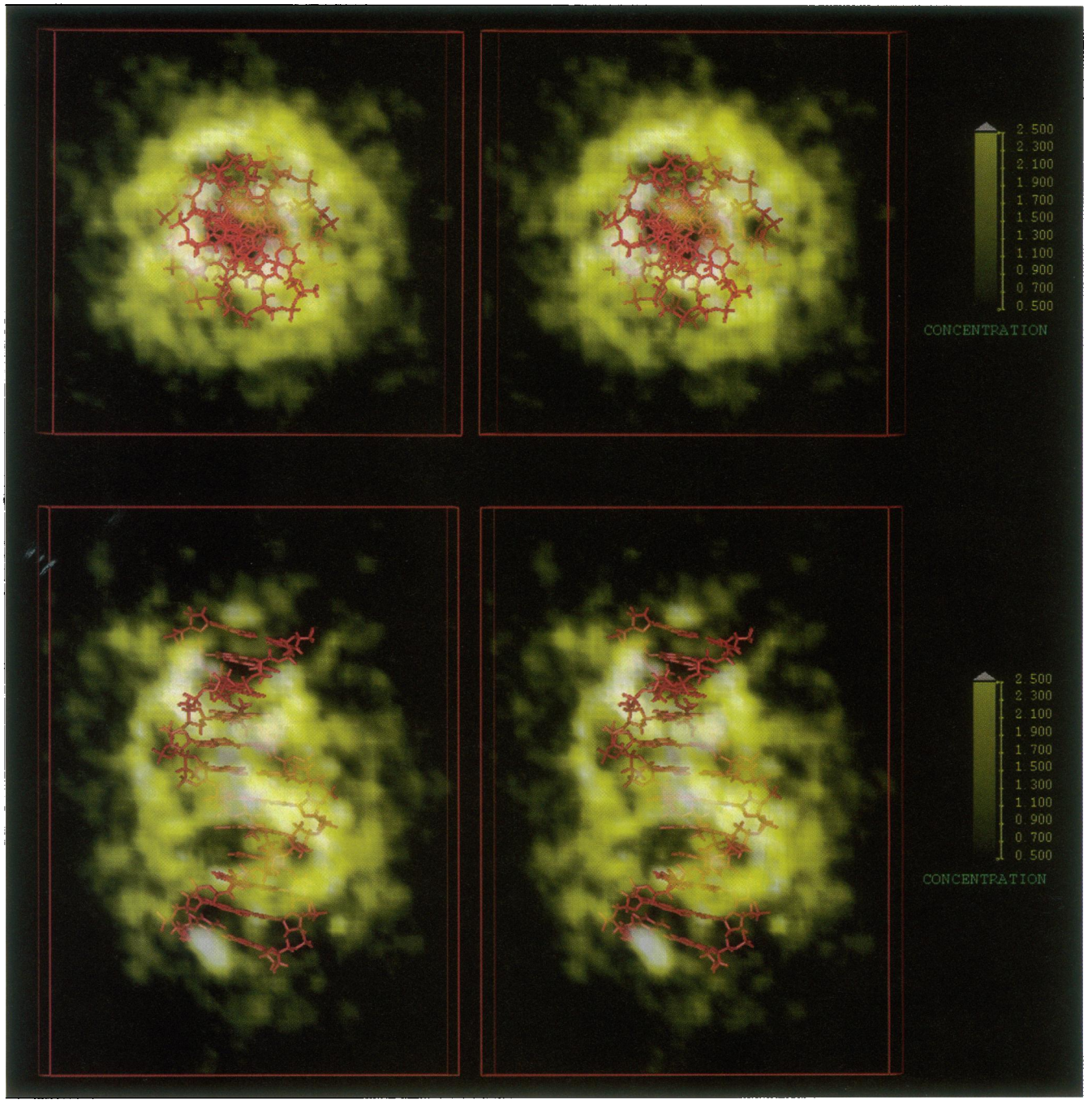


FIGURE 10 Two stereo views of the distribution of Na^+ ions over the 5.0-ns trajectory superimposed on the average DNA conformation. Atomic densities are calculated on a fixed 1.0-Å^3 resolution grid, and grid voxels are shaded on a density spectrum to indicate the local Na^+ densities. Densities shown range from 500 mM (3×10^{-4}) ions/Å^3) (*transparent*) to 2.5 M (1.5×10^{-3}) ions/Å^3) (*white*). (*Top*) View down the helical axis. (*Bottom*) View into the major groove.

region of the minor groove at one end of the sequence, and the GpA step of the major groove. Analysis of the trajectory reveals that there is considerable dynamic motion underlying the disposition of ions in the minor groove of the CGCC flanking sequence, and frequent ion-water exchanges. The fact that the net ion density is observed at only one of the ends of a self-complementary sequence suggests that the dynamic structure of the monovalent counterions, even though well stabilized, is not yet totally converged, even over the course of a 5-ns trajectory. We note in passing the apparent tendency of ions to cluster at positions in the flanking sequences that are sites of bending or to demonstrate a propensity toward bending.

The results of this extended simulation indicate that in the AMBER 4.1 MD model, counterions do intrude into the minor groove spine of hydration in a B-form DNA double

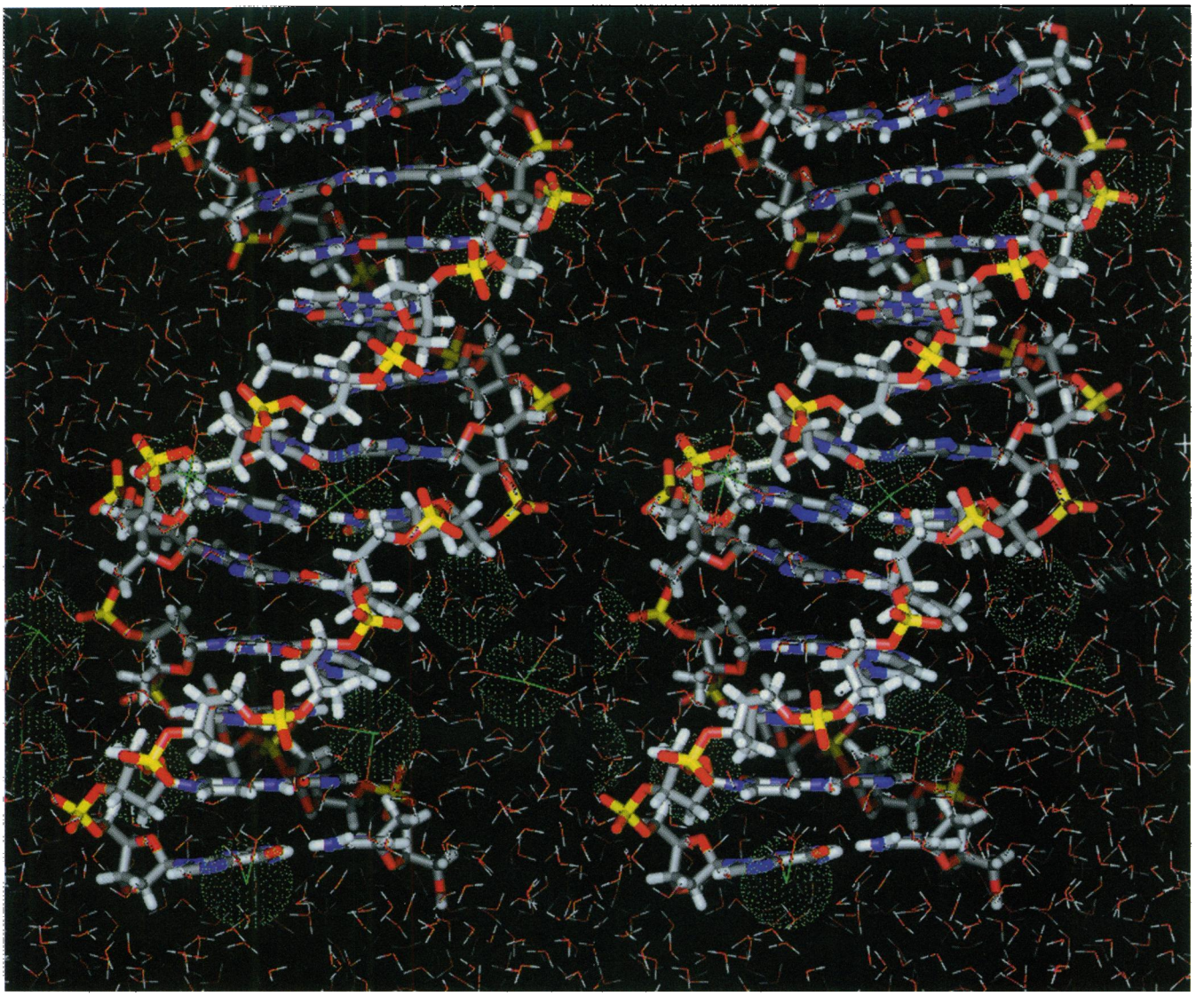


FIGURE 11 Stereo view of a structure obtained from the MCI/PME/MD trajectory, showing the structural detail of the incorporation of a Na^+ into the first solvation shell of the minor groove. The structure is representative of the first 500 ps of the trajectory.

helix at sequence-specific positions of electronegative pockets in the electrostatic potential. The AATT region of the minor groove manifests this property, and the effect can be further enhanced by propitious local geometry. These results are consistent with earlier calculations of the electrostatic potential of B-DNA based on a point charge model by Lavery and Pullman (1985), the nonlinear PB calculations of Jayaram and Honig (Jayaram et al., 1989), and follow-up PB calculations on $d(CGCGAATTCGCG)_2$ (Young et al., 1997), carried out to further elucidate the sequence effects issue.

None of the previous crystallographic studies of oligonucleotides relevant to the minor groove spine were geared to distinguishing water molecules from sodium ions. The previous report most relevant to this results is the crystal structure of ApU (Seeman et al., 1976), in which a sodium ion was found to be complexed to the carbonyl groups of successive uracils that protrude into the proto-minor groove. The coordination of the sodium ion was noted to pull the two bases closer together, resulting in a locally anomalous propeller twist angle. The Na^+ coordination we found in the MCI MD simulation on $d(CGCGAATTCGCG)_2$ shows essentially close coincidence with the structure seen in prototype in the crystal structure of ApU, and thus experiment and theory are to this extent mutually supportive of the idea that counterions may be involved in the minor groove solvation in B-DNA. The MD further predicts that the nature of these structures is not static but dynamic, with significant exchange of ions and water at favorable sites in the minor groove spine to be expected. A recent report by Hud and Feigon (1997) of localization of divalent metal ions in the minor groove of DNA A-tracts based on NMR spectroscopy has just appeared.

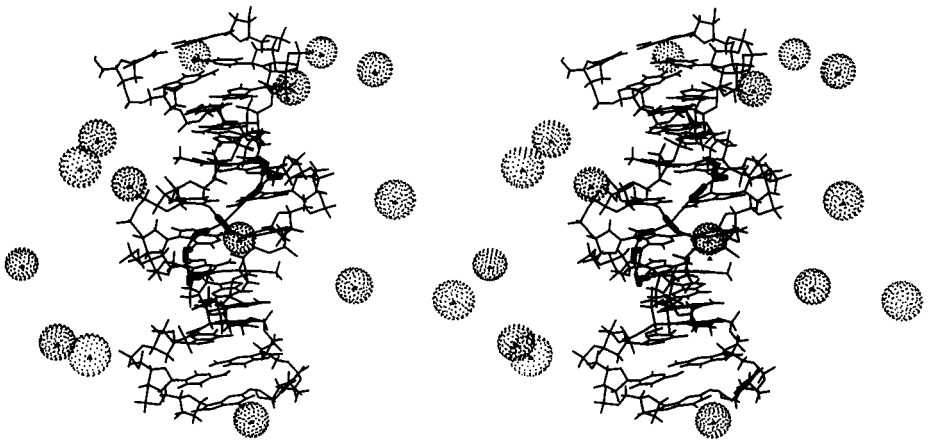


FIGURE 12 Stereo view of a structure obtained from the MCI/PME/MD trajectory, showing the structural detail of an intact "spine of hydration" the first solvation shell of the minor groove. The structure is representative of the second 500 ps of the trajectory.

Hydration

A computer graphics rendering of the calculated first shell solvation density is shown in Fig. 13. The total solvation density is composed of contributions due to water molecules and counterions intertwined. In the minor groove there is considerable localization of both water molecules and Na^+ ions at various sites on or close to the surface of the DNA relative to the densities found in the bulk solvent. Close inspection of the sites constituting the previously identified minor groove "spine of hydration" shows that this feature is well reproduced in the MD, but the solvation of the minor groove must be described in terms of fractional occupation by water and Na^+ ions, some of which have residence times of 500 ps or greater. The view into the major groove, Fig. 14, also shows localized ion and water density, but the residence time of ions is 100 ps or less. We find no particular evidence for spines or filaments in the first shell of major groove waters, but we do find a fully connected network featuring a distribution of interaction energies. This result is consistent with the larger configurational volume of the major groove, which imposes no geometrical restrictions on the motions of counterions and solvent water.

The boundary of the first solvation shell of the DNA as described by the simulation may be defined in terms of inflection points in molecular distribution functions for solvent molecules, as defined by the proximity criterion, an approach analogous to that used to describe the shell structure of liquids (Mehrotra and Beveridge, 1980). The number of waters in the first shell of the DNA (defined as waters with oxygens located less than 3.0 Å from the DNA surface), eliminating end effects, is found to average out at 22.2 waters per nucleotide base pair. This number is consistent with estimates from spectroscopic measurements of 11–12 waters per nucleotide (Saenger, 1983; Wolf and Hanlon, 1975).

A question of considerable interest to crystallographers is the relationship of the observed or "ordered" waters to the structure of the hydration in general. In d(CGCGAATTCGCG)$_2$, comparing the MD calculated first shell hydration numbers with the number of ordered water sites identified in the crystal structure indicates that if the MD figure is accurate, we calculate that ~18% of the water in the primary solvation shell (<3.0 Å from the DNA) was ordered enough to be assigned in the crystallographic refinement of solvent (NDB structure 2BNA; Drew et al., 1982). The total number of waters peaks identified in the 16K crystal structure of the dodecamer was reported to be 27% of the total number of waters in the asymmetrical unit (Kopka et al., 1983), but not all of these waters are part of the primary solvation layer. To pursue this issue further, we examined the extent to which density fluctuations in the calculated solvation can be identified with ordered water positions in the major and minor grooves of the crystal. To accomplish this, we examined the calculated solvation density by varying the display threshold with respect to a reference of bulk water. The maximum correspondence was found at a threshold density value twice that of the bulk, indicating that ordered solvent sites observed in crystal are twice as ordered as bulk water. A comparison of crystallographically ordered water sites and calculated above-threshold water density in the model solution structure is shown in Fig. 15. The correspondence is quite close overall, and the agreement appears to us to be quite reasonable, considering that the calculated solution structure and the refined crystal structure of the DNA are not identical.

A more detailed analysis of the calculated hydration and counterion structure and motion based on quasicomponent distribution functions and the proximity criterion (Mehrotra and Beveridge, 1980; Mezei and Beveridge, 1986) will be provided in a later paper. We have performed an extensive visual examination of the calculated solvation and find the water networks to be robust, with no convincing evidence of distinct substructures such as filaments, or spines in regions other than those, such as the minor groove, where DNA geometry dictates. Thus the solvation seems to involve a continuous distribution of molecular structural and energetic environments, analogous to that of the preference in liquid water for a continuum model over a mixture model (Beveridge et al., 1983; Rahman and Stillinger, 1971; Swaminathan and Beveridge, 1978). Overall, the MD sup-

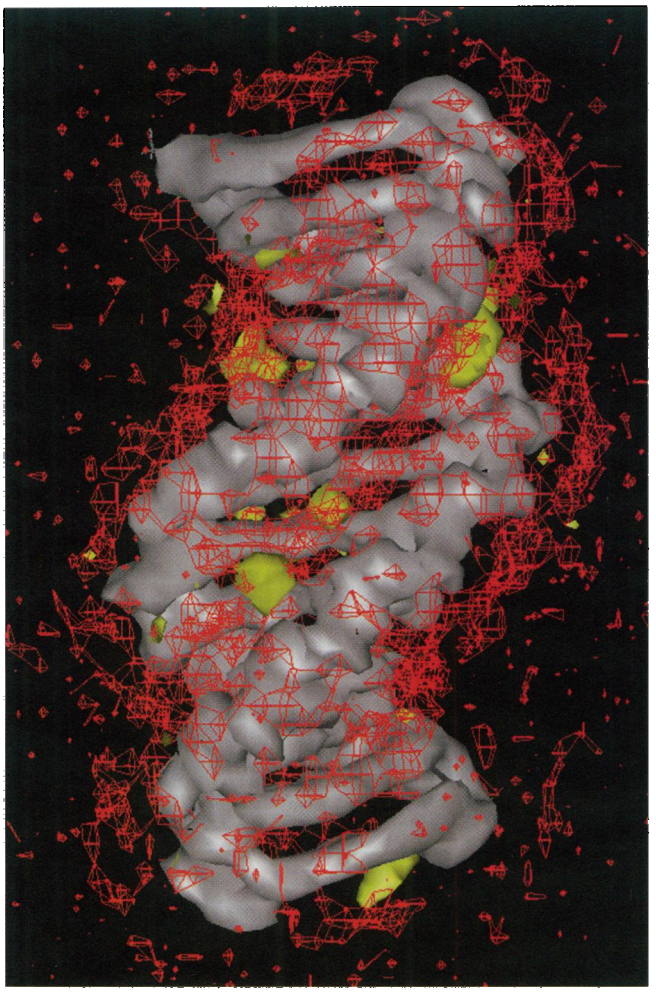


FIGURE 13 Isoprobability surfaces for the DNA atoms (*solid gray*), the water oxygens (*red*), and the Na^+ ions (*green*). Densities were calculated on a fixed 1.0-Å^3 grid after each snapshot was superimposed on the reference frame of the DNA, and are normalized over the entire 5.0-ns trajectory. The surfaces shown are 0.03 atoms/Å^3 for the DNA, 0.045 atoms/Å^3 for the waters, and 0.0045 atoms/Å^3 for the ions. Bulk water has a density of 0.0326 waters/Å^3 (55 M), so the borders of the "first shell" waters immediately surrounding the DNA here are ~50% more ordered than bulk water.

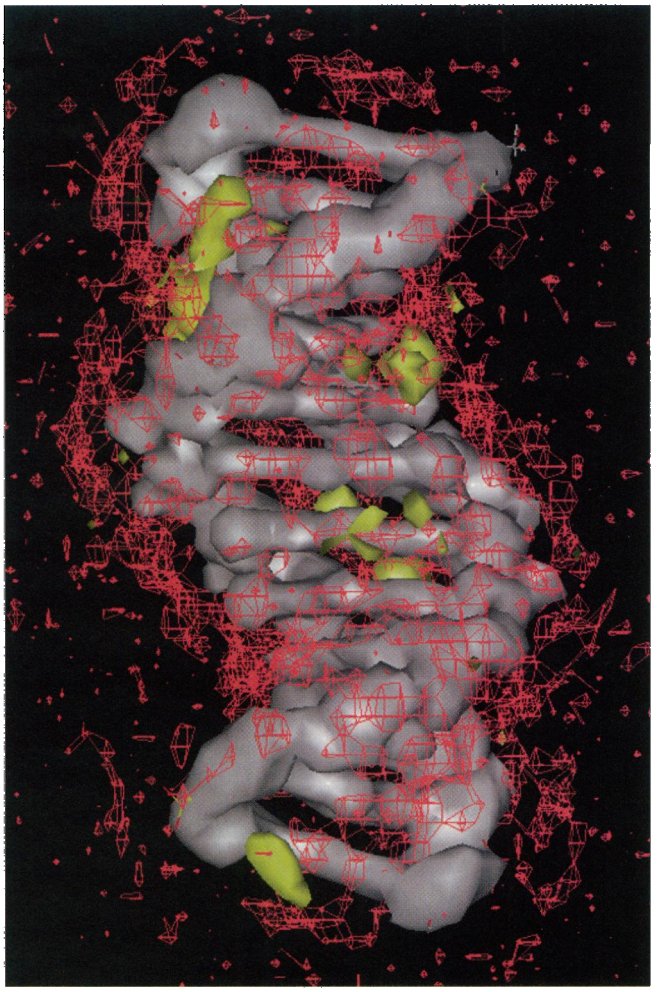


FIGURE 14 Same isoprobability surfaces as in Fig. 10, viewed in the major groove.

ports the idea that water is an integral part of helical as well as nonhelical structures of nucleic acids, and that the development of an adequate treatment of hydration and solvent effects in general is essential for correct modeling and simulation of nucleic acids.

The implications of the results are of potentially considerable interest in the structural biology of DNA and RNA sequences. The presence of ions in sequence-specific pockets in the minor groove would be expected to have the net effect of mitigating electrostatic repulsions among the various anionic phosphates in the region. This could, in turn, influence the helix morphology, i.e., axis bending and groove widths. We have examined the effect of ion binding in the ApT pocket on groove width by analyzing a DNA structure form by MD, in which an Na^+ ion occupies the ApT pocket (cf. Fig. 7), and a form in which it has exchanged with water in this site (Fig. 8). The Na^+ in the AT pocket shows a noticeable effect on the minor groove local to the region of ion association, leading evidence that ion complexation plays a significant role in the narrowing of the minor groove in AT-rich DNA sequences.

SUMMARY AND CONCLUSIONS

The results of four new molecular dynamics (MD) simulations on the DNA duplex of sequence d(CGCGAATTCGCG)$_2$, including explicit consideration of solvent water and a sufficient number of Na^+ counterions to provide electroneutrality to the system, indicate that the new AMBER 4.1 force field with a PME treatment of long-range interactions leads to an accurate description of the dynamic structure of B DNA. To support this claim, the MD results are compared with corresponding crystallographic and NMR studies on the d(CGCGAATTCGCG)$_2$ duplex, and placed in the context of observed behavior of B-DNA by comparisons with the complete crystallographic database of

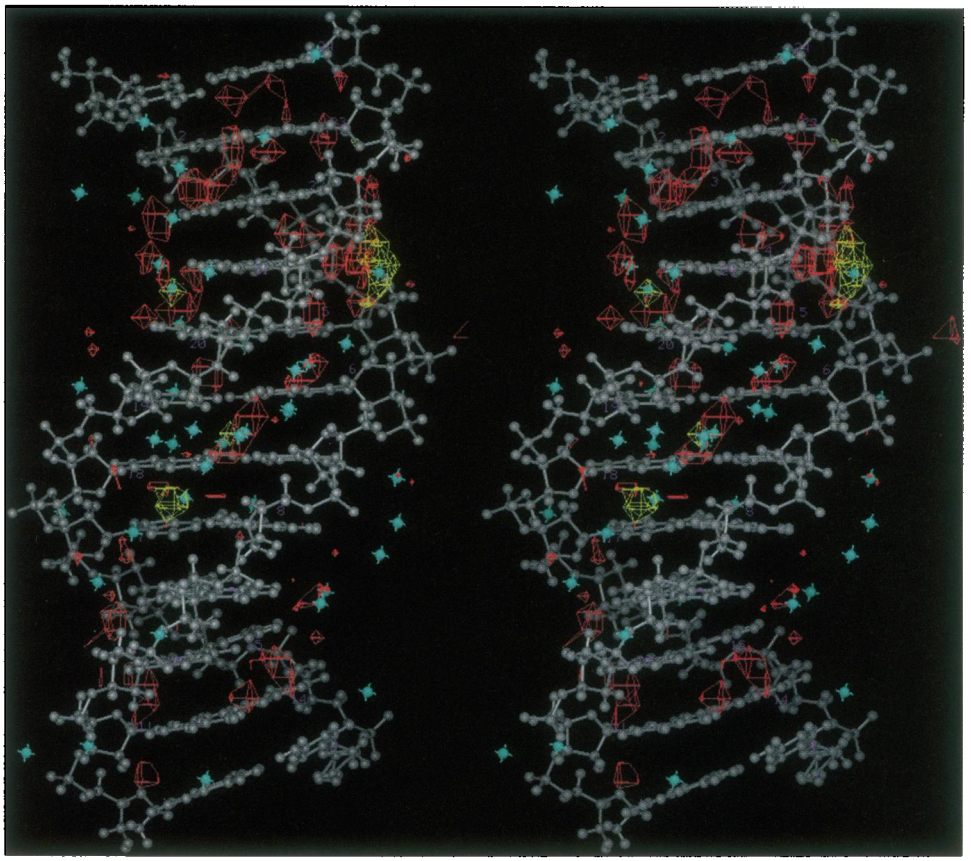


FIGURE 15 Red spheres indicate ordered water sites located less than 3.0 Å from the DNA in the native CGCGAATTCGCG crystal structure. The average structure from the MD was then superimposed on the crystal structure to compare ordered solvent sites in the MD with refined water positions in the crystal. The occupancy threshold for the calculated waters was adjusted until a reasonable match with the crystallographic sites was achieved. This occupancy turned out to be ~0.06 waters/Å^3, approximately twice as high as the 0.0326 water/Å^3 value of bulk water, and 1.5 times as high as the "first shell" waters identified in Figs. 14.

B-form structures. Note, however, that we cannot eliminate the possibility that the AMBER 4.1 dynamic model overestimates the stability of B-DNA; further studies of this issue are in progress.

The analysis of the distribution of mobile counterions in the simulation shows that counterions may intrude on the minor groove spine of hydration in B-form DNA and may subsequently influence the environmental structure and thermodynamics in a sequence-dependent manner. The idea of localized complexation of otherwise mobile counterions in electronegative pockets in the grooves of DNA helices introduces a heretofore mostly unappreciated source of sequence-dependent effects on local conformational, helicoidal, and morphological structure, and may have important implications for understanding the functional energetics and specificity of the interactions of DNA and RNA with regulatory proteins, pharmaceutical agents, and other ligands. If direct ionic interactions in the minor and major grooves are a significant element of understanding the sequence-dependent fine structure of nucleic acid helices, then modeling the effect of ions in modes previously utilized by ourselves and others, including simple distance-dependent dielectrics, reduced charges on phosphates, salt-dependent interphosphate screening functions, or artificially large hydrated counterions (solvatons), will not suffice to describe the physics of the system adequately. Explicit consideration of hydration and ionic solvation and the delicate balance of numerous local attractions and repulsions in the system at temperature will thus be crucial to providing an accurate model of the system.

The accuracy of MD on DNA must still be considered an evolving story, but new and improved force fields, the PME method for long-range interactions, and increased supercomputer power have demonstrably advanced the field. The indication of this study is that accurate, lengthy all-atom MD explicitly including water and counterions can achieve an accurate description of the DNA double helix in solution, and produce ideas about the nature of dynamic structure and solvation that will be useful to consider in further experimental and theoretical studies.

CAPTIONS FOR SUPPLEMENTARY FIGURES (AVAILABLE ON THE WWW)

(*Note: Contact the corresponding author for URLs.*)

Figure W-1. Time series analysis of all conformational parameters of the "left-hand" strand of the *Eco*RI dodecamer computed from the MCI MD trajectory, and presented as conformational "dials" (Ravishanker and Beveridge, 1997). The definition of the parameters follows IUPAC convention.

Figure W-2. Time series analysis of all conformational parameters of the "right-hand" strand of the *Eco*RI dodecamer computed from the 5.0 MCI/PME/MD trajectory, and presented as conformational "dials" (Ravishanker and Bev-

eridge, 1997). The definition of the parameters follows IUPAC convention.

Figure W-3. Time series analysis of the axis-base pair parameters for the *Eco*RI dodecamer computed from the 5.0 MCI-MD trajectory, presented via helicoidal "Windows" (Ravishanker and Beveridge, 1997). The definition of the helicoidal parameters follows convention (Dickerson et al., 1989); they were computed with respect to a global helical axis by using the CURVES procedure (Lavery and Sklenar, 1988).

Figure W-4. Time series analysis of the intra-base pair parameters for the *Eco*RI dodecamer computed from the 5.0 MCI/PME/MD trajectory, presented via helicoidal "Windows" (Ravishanker and Beveridge, 1997). The definition of the helicoidal parameters follows convention (Dickerson et al., 1989); they were computed based on a global helical axis by using the CURVES procedure (Lavery and Sklenar, 1988).

Figure W-5. Time series analysis of the inter-base pair parameters for the *Eco*RI dodecamer computed from the 5.0 MCI-MD trajectory, presented via helicoidal "Windows" (Ravishanker and Beveridge, 1997). The definition of the helicoidal parameters follows convention (Dickerson et al., 1989); they were computed based on a global helical axis by using the CURVES procedure (Lavery and Sklenar, 1988).

Figure W-6. Possible location of electronegative pockets in the minor groove of DNA as a function of sequence (schematic).

Figure W-7. Possible location of electronegative pockets in the major groove of DNA as a function of sequence (schematic).

The authors gratefully acknowledge discussions and advice on this project from Dr. S. C. Harvey of the University of Alabama at Birmingham, Dr. H. M. Berman of Rutgers University, and Dr. Richard Lavery of the Institut de Biologie Physico-chemique, Paris.

This research was supported primarily by grant GM-37909 from the National Institutes of Health. MAY is the recipient of an National Institute of General Medical Studies Traineeship in Molecular Biophysics, from grant GM-08271 to Wesleyan University. Computer graphics analyses were facilitated by a grant from the Dreyfus Foundation and use of the computer programs MD Tool Chest (Ravishanker and Beveridge, 1997) and UCSF Midas-Plus (Ferrin et al., 1988; Huang et al., 1991). Cray time was provided for these projects by the Pittsburgh Supercomputer Center, and the Frederick Biomedical Supercomputer Center of the National Cancer Institute, National Institutes of Health.

REFERENCES

Alam, T. M., J. Orban, and G. P. Drobny. 1991. Deuterium NMR investigation of backbone dynamics in the synthetic oligonucleotide [d(CGCGAATTCGCG)]2. *Biochemistry.* 30:9229–9237.

Allen, M. P., and D. J. Tildesly. 1987. Computer Simulation of Liquids. Clarenden Press, Oxford.

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. CRC Press, Boca Raton, FL. 411–422.

Auffinger, P., and D. L. Beveridge. 1995. A simple test for evaluating the truncation effects in simulations of systems involving charged groups. *Chem. Phys. Lett.* 234:413.

Benevides, J. M., A. H.-J. Wang, G. A. Va, J. H. V. B. Der Marel, and G. J. Thoma, Jr. 1988. Crystal and solution structures of the B-DNA dodecamer d(CGCAAATTTGCG) probed by Raman spectroscopy: heterogeneity in the crystal does not persist in the solution structure. *Biochemistry.* 27:931–938.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, and A. DiNola. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Berman, H. M., A. Gelbin, and J. Westbrook. 1996. Nucleic acid crystallography: a view from the nucleic acids data base. *Prog. Mol. Biol. Biophys.* 66:255–298.

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

Beveridge, D. L., M. Mezei, P. K. Mehrotra, F. T. Marchese, G. Ravishanker, T. R. Vasu, and S. Swaminathan. 1983. Monte Carlo computer simulation studies of the equilibrium properties and structure of liquid water. *In* Molecular-Based Study of Fluids. American Chemical Society, Washington, DC. 297–351.

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

Beveridge, D. L., P. Subramanian, B. Jayaram, S. Swaminathan, and G. Ravishanker. 1990. Molecular simulation studies on the d(CGCGAATTCGCG) duplex: hydration, ion atmosphere, structure and dynamics. *In* Structure and Methods III. DNA AND RNA. Adenine Press, New York. 79–112.

Beveridge, D. L., S. Swaminathan, G. Ravishanker, J. M. Withka, J. Srinivasan, C. Prevost, S. Louise-May, D. R. Langley, F. M. DiCapua, and P. H. Bolton. 1993. Molecular dynamics simulations on the hydration, structure and motions of DNA oligomers. *In* Water and Biological Molecules. Macmillan Press, London. 165–225.

Bloomfield, V. A. 1991. Condensation of DNA by multivalent cations: considerations on mechanism. *Biopolymers.* 31:1471–1481.

Cheatham, T. E., III, and P. A. Kollman. 1996. Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* 259:434–444.

Cheatham, T. E., III, J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman. 1995. Molecular dynamics simulations 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.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., 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.

Dickerson, R. E. 1983. The DNA helix and how it is read. *Sci. Am.* 249:94–111.

Dickerson, R. E. 1991. DNA structure from A to Z. *Methods Enzymol.* 211:67–111.

Dickerson, R. E., M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. M. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C. S. Tung, A. H. J. Wang, and V. B. Zhurkin. 1989. Definitions and nomenclature of nucleic acid structural parameters. *EMBO J.* 8:1–4.

Dickerson, R. E., and H. R. Drew. 1981. Structure of a B DNA dodecamer. II. Influence of base sequence on helix structure. *J. Mol. Biol.* 149: 761–786.

Dickerson, R. E., D. S. Goodsell, M. L. Kopka, and P. E. Pjura. 1987. The effect of crystal packing on oligonucleotide double helix structure. *J. Biomol. Struct. Dyn.* 5:557–579.

Dickerson, R. E., K. Grzeskowiak, M. Grzeskowiak, M. L. Kopka, T. Larson, A. Lipanov, G. G. Prive, J. Quintana, P. Schultze, K. Yanagi, H. Yuan, and H.-C. Yoon. 1991. Polymophism, packing, resolution, and reliablity in single-crystal DNA oligomer analyses. *Nucleosides Nucleotides.* 10:3–24.

Dickerson, R. E., M. L. Kopka, and P. Pjura. 1985. Base sequence, helix geometry, hydration and helix stability in B-DNA. *Biol. Macromol. Assemblies.* 2:37–126.

DiGabriele, A. D., M. R. Sanderson, and T. A. Steitz. 1989. Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract. *Proc. Natl. Acad. Sci. USA.* 86:1816–1820.

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

Drew, H. R., S. Samson, and R. E. Dickerson. 1982. Structure of a B-DNA dodecamer at 16 K. *Proc. Natl. Acad. Sci. USA.* 79:4040–4044.

Drew, H. R., R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itikura, and R. E. Dickerson. 1981. Structure of a B DNA dodecamer. I. Conformation and dynamics. *Proc. Natl. Acad. Sci. USA.* 78:2179–2983.

Falsafi, S., and N. O. Reich. 1993. Molecular dynamics simulations of B-DNA: an analysis of the role of initial molecular configuration, randomly assigned velocity distribution, long integration times, and nonconstrained termini. *Biopolymers.* 33:459–473.

Ferrin, T. E., C. C. Huang, L. E. Jarvis, and R. Langridge. 1988. The MIDAS display system. *J. Mol. Graph.* 6:13–27.

Franklin, R. E., and R. G. Gosling. 1953. The structure of sodium thymonucleate fibers. I. The influence of water content. *Acta Crystallogr.* 6:673–677.

Frederick, C. A., J. Grable, M. Melia, C. Samudzi, L. Jen-Jacobson, B.-C. Wang, P. Greene, H. W. Boyer, and J. M. Rosenberg. 1984. Kinked DNA in crystalline complex with *Eco*RI endonuclease. *Nature.* 309: 327–330.

Gorin, A. A., V. B. Zhurkin, and W. K. Olson. 1995. B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* 247:34–48.

Gould, I., and P. A. Kollman. 1993. Ab initio and molecular mechanics calculations on the hydrogen bond strengths in guanine-cytosine and adenine-thymine base pairs. *J. Am. Chem. Soc.* 116:2493–2494.

Grable, J., C. A. Frederick, C. Samudzi, L. Jen Jacobson, D. Lesser, P. Greene, H. W. Boyer, K. Itakura, and J. M. Rosenberg. 1984. Two-fold symmetry of crystalline DNA-*Eco*RI endonuclease recognition complexes. *J. Biomol. Struct. Dyn.* 1:1149–1160.

Hare, D. R., D. E. Wemmer, S. H. Chou, G. Drobny, and B. R. Ried. 1983. Assignment of the non-exchangeable proton resonances of d(CGCGAATTCGCG) using 2D-NMR methods. *J. Mol. Biol.* 171:319–336.

Hingerty, B. E., R. H. Ritchie, T. L. Ferrel, and J. E. Turner. 1985. Dielectric effects in biopolymers: the theory of ionic saturation revisited. *Biopolymers.* 24:427–439.

Holbrook, S. R., R. E. Dickerson, and S. H. Kim. 1985. Anisotropic thermal parameter refinement of the DNA dodecamer CGCGAATTCGCG by the segmented rigid body method. *Acta. Crystallogr. Sect. B.* 41:255–262.

Huang, C. C., E. F. Pettersen, T. E. Klein, T. E. Ferrin, and R. Langridge. 1991. Conic: a fast renderer for space-filling molecules with shadows. *J. Mol. Graph.* 9:230–236.

Hud, N. V., and J. Feigon. 1997. Localization of divalent metal ions in the minor groove of DNA A-tracts. *J. Am. Chem. Soc.* 119:5756–5757.

Ivanov, V. I., L. E. Minchenkova, E. E. Minyat, M. D. Frank-Kamentskii, and A. K. Schyolkina. 1974. The B to A transition of DNA in solution. *J. Mol. Biol.* 87:817–833.

Jain, S., and M. Sundaralingam. 1989. Effect of crystal packing environment on conformation of the DNA duplex. *J. Biol. Chem.* 264: 12780–12784.

Jayaram, B., K. A. Sharp, and B. Honig. 1989. The electrostatic potential of B-DNA. *Biopolymers.* 28:975–993.

Jayaram, B., S. Swaminathan, and D. L. Beveridge. 1990a. Monte Carlo simulation studies of the structure of the counteranion atmosphere of B-DNA. Variations on the primitive dielectric model. *Macromolecules.* 23:3156–3165.

Jayaram, B., S. Swaminathan, D. L. Beveridge, K. Sharp, and B. Honig. 1990b. Monte Carlo simulation studies on the structure of the counterion atmosphere of B-DNA. Variations on the primitive dielectric model. *J. Phys. Chem.* 23:3156–3165.

Jorgensen, W. L. 1981. Transferable intermolecular potential functions for water, alcohols and ethers. Application to liquid water. *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 qualitative study. *J. Mol. Biol.* 163:129–146.

Kopka, M. L., P. Pjura, C. Yoon, D. Goodsell, and R. E. Dickerson. 1985. The binding of netropsin to double helical DNA of sequence CGCGAATTCGCG: single crystal x-ray structure analysis. *In* Structure and Motion: Membranes, Nucleic Acids and Proteins. Adenine Press, Guilderland, NY. 461ff.

Kubinec, M. G., and D. E. Wemmer. 1992. NMR evidence for DNA bound water in solution. *J. Am. Chem. Soc.* 114:8739–8740.

Kumar, S., Y. Duan, P. A. Kollman, and J. M. Rosenberg. 1994. Molecular dynamics simulations suggest that the *Eco*RI kink is an example of molecular strain. *J. Biomol. Struct. Dyn.* 12:487–525.

Lane, A., T. C. Jenkins, T. Brown, and S. Neidle. 1991. Interaction of Berenil with the *Eco*RI dodecamer d(CGCGAATTCGCG) in solution studied by NMR. *Biochemistry.* 30:1372–1385.

Langley, D. R. 1996. The BMS nucleic acid force field. Bristol-Myers Squibb Company, Wallingford, CT.

Lavery, R., and B. Pullman. 1985. The dependence of the surface electrostatic potential of B-DNA on environmental factors. *J. Biomol. Struct. Dyn.* 5:1021–32.

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

Lavery, R., and H. Sklenar. 1996. Curves 51—helical analysis of irregular nucleic acids. Laboratoire de Biochimie Theorique, CNRS URA 77, Institut de Biologie Physico-Chimique, Paris.

Liepinsh, E., G. Otting, and K. Wüthrich. 1992. NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution. *Nucleic Acids Res.* 20:6549–6553.

Loncharich, R. J., and B. R. Brooks. 1989. The effects of truncating long-range forces on protein dynamics. *Proteins.* 6:32–45.

Louise-May, S. 1994. Molecular Dynamics Investigations of the Local Structural Characteristics of DNA Oligonucleotides. Wesleyan University, Middletown, CT.

Mackerell, A. D., Jr., T. Wiorkiewicz-Kuizera, and M. Karplus. 1995. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* 117:11946–11975.

Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11:179–246.

McConnell, K. M., R. Nirmala, M. A. Young, G. Ravishanker, and D. L. Beveridge. 1994. A nanosecond molecular dynamics trajectory for a B DNA double helix: evidence for substates. *J. Am. Chem. Soc.* 116: 4461–4462.

Mehrotra, P. K., and D. L. Beveridge. 1980. Structural analysis of molecular solutions based on quasi-component distribution functions. Application to $[H_2CO]_{aq}$ at 25°C. *J. Am. Chem. Soc.* 102:4287–4294.

Mezei, M., and D. L. Beveridge. 1986. Structural chemistry of biomolecular hydration: the proximity criterion. *Methods Enzymol.* 127:21–47.

Miaskiewicz, K., J. Miller, M. Cooney, and R. Osman. 1996. Computational simulations of DNA distortions by a *syn, cis* cyclobutane thymine dimer lesion. *J. Am. Chem. Soc.* 118:9156–9163.

Miaskiewicz, K., R. Osman, and H. Weinstein. 1993. Molecular dynamics simulation of the hydrated d(CGCGAATTCGCG)2 dodecamer. *J. Am. Chem. Soc.* 115:1526–1537.

Moe, J. G., and I. M. Russu. 1990. Proton exchange and base-pair opening kinetics in 5′-d(CGCGAATTCGCG)-3′ and related dodecamers. *Nucleic Acids Res.* 18:821–827.

Moe, J. G., and I. M. Russu. 1992. Kinetics and energetics of base-pair opening in 5′-d(CGCGAATTCGCG)-3′ and a substituted dodecamer containing GT mismatches. *Biochemistry.* 31:8421–8428.

Narayana, N., S. L. Ginell, I. M. Russu, and H. M. Berman. 1991. Crystal and molecular structure of a DNA fragment: d(CGTGAATTCACG). *Biochemistry.* 30:4449–4455.

Nerdal, W., D. R. Hare, and B. R. Reid. 1989. Solution structure of the *Eco*RI DNA sequence: refinement of NMR-derived distance geometry structures by NOESY spectrum back-calculations. *Biochemistry.* 28: 10008–10021.

Nunn, C. M., T. C. Jenkins, and S. Neidle. 1993. Crystal structure of d(CGCGAATTCGCG) complexed with propamidine, a short-chain homologue of the drug pentamidine. *Biochemistry.* 32:13838–13843.

Olmsted, M. C., C. F. Anderson, and M. T. Record, Jr. 1991. Importance of oligoelectrolyte end effects for the thermodynamics of conformational transitions of nucleic acid oligomers: a grand canonical Monte Carlo analysis. *Biopolymers.* 31:1593–1604.

Olson, W. 1982a. Theoretical studies of nucleic acid conformation: potential energies, chain statistics, and model building. *In* Topics in Nucleic Acid Structure. MacMillan, London. 1–79.

Olson, W. K. 1982b. How flexible is the furanose ring? An updated potential energy estimate. *J. Am. Chem. Soc.* 104:278–286.

Osman, R., K. Miaskiewicz, and H. Weinstein. 1991. Structure-function relations in radiation damaged DNA. *In* Structure. Function Relations in Radiation Damaged DNA. Plenum Press, New York. 1–30.

Ott, J., and F. Eckstein. 1985. Phosphorous NMR spectral analysis of the dodecamer d(CGCGAATTCGCG). *Biochemistry.* 24:2530–2535.

Patel, D. J., S. A. Kozlowski, L. A. Marky, C. Broka, J. A. Rice, K. Itakura, and K. J. Breslauer. 1982. Premelting and melting transitions in the d(CGCGAATTCGCG) self-complementary duplex in solution. *Biochemistry.* 21:428–436.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, III, D. M. Fergusen, G. L. Seibel, U. C. Singh, P. Weiner, and P. Kollman. 1995. AMBER 41. University of California, San Francisco, CA.

Pelton, J. G., and D. E. Wemmer. 1988. Structural modeling of the distamycin A-d(CGCGAATTCGCG)2 complex using 2D NMR and molecular mechanics. *Biochemistry.* 27:8088–8096.

Poncin, M., B. Hartmann, and R. Lavery. 1992. Conformational sub-states in B-DNA. *J. Mol. Biol.* 226:775–794.

Prevost, C., S. Louise May, G. Ravishanker, R. Lavery, and D. L. Beveridge. 1993. Persistence analysis of the static and dynamical helix deformations of DNA oligonucleotides: application to the crystal structure and molecular dynamics simulation of d(CGCGAATTCGCG)2. *Biopolymers.* 33:335–350.

Rahman, A., and F. H. Stillinger. 1971. Molecular dynamics study of liquid water. *J. Chem. Phys.* 55:3336–3359.

Rao, S. N., and P. Kollman. 1990. Simulations of the B-DNA molecular dynamics of d(CGCGAATTCGCG)2 and d(GCGCGCGCGC)2: an analysis of the role of initial geometry and a comparison of united and all-atom models. *Biopolymers.* 29:517–532.

Ravishanker, G., P. Auffinger, D. R. Langley, B. Jayaram, M. A. Young, and D. L. Beveridge. 1997. Treatment of Counterions in Computer Simulations of DNA. *In* Reviews in Computational Chemistry. 11: 317–372.

Ravishanker, G., and D. L. Beveridge. 1997. MD Toolchest. Wesleyan University, Middletown, CT.

Ravishanker, G., S. Swaminathan, D. L. Beveridge, R. Lavery, and H. Sklenar. 1989. Conformational and helicoidal analysis of 30 PS of molecular dynamics on the d(CGCGAATTCGCG) double helix: "curves," dials and windows. *J. Biomol. Struct. Dyn.* 6:669–699.

Saenger, W. 1983. Principles of Nucleic Acid Structure. Springer Verlag, New York.

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

Schneider, B., D. Cohen, and H. M. Berman. 1992. Hydration of DNA bases: analysis of crystallographic data. *Biopolymers.* 32:725–750.

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

Seeman, N. C., J. M. Rosenberg, F. L. Suddath, J. J. P. Kim, and A. Rich. 1976. RNA double-helical fragments at atomic resolution. I. The crystal and molecular structure of sodium adenylyl-3′,5′-uridine hexahydrate. *J. Mol. Biol.* 104:109–144.

Shakked, Z., G. Guerstein Guzikevich, M. Eisenstein, F. Frolow, and D. Rabinovich. 1989. The conformation of the DNA double helix in the crystal is dependent on its environment. *Nature.* 342:456–460.

Smith, P. E., and B. M. Pettitt. 1991. Peptides in ionic solutions: a comparison of the Ewald and switching function techniques. *J. Chem. Phys.* 95:8430–8441.

Smith, P. E., and B. M. Pettitt. 1996. Ewald artifacts in liquid state molecular dynamics simulations. *J. Chem. Phys.* 105:4289–4293.

Srinivasan, J., J. M. Withka, and D. L. Beveridge. 1990. Molecular dynamics of an in vacuo model of duplex d(CGCGAATTCGCG) in the B-form based on the amber 30 force field. *Biophys. J.* 58:533–547.

Steinbach, P. J., and B. R. Brooks. 1994. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comp. Chem.* 15: 667–683.

Steitz, T. A. 1990. Structural studies of protein-nucleic acid interactions: the sources of sequence-specific binding. *Q. Rev. Biophys.* 23:205–280.

Stofer, E., and R. Lavery. 1994. Measuring the geometry of DNA grooves. *Biopolymers.* 34:337–346.

Subramanian, P. S., and D. L. Beveridge. 1989. A theoretical study of the aqueous hydration of canonical B d(CGCGAATTCGCG): Monte Carlo simulation and comparison with 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(GCGCTTAAGCGC): Monte Carlo computer simulation. *Proc. Natl. Acad. Sci. USA.* 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.

Swaminathan, S., and D. L. Beveridge. 1977. A theoretical study of the structure of liquid water based on quasicomponent distribution functions. *J. Am. Chem. Soc.* 99:8392–8398.

Swaminathan, S., G. Ravishanker, and D. L. Beveridge. 1991. Molecular dynamics of B-DNA including water and counterions: a 140 ps trajectory for d(CGCGAATTCGCG) based on the GROMOS force field. *J. Am. Chem. Soc.* 111:5027–5040.

Teleman, O., and A. Wallqvist. 1990. Ewald summation retards translational motion in molecular dynamics simulation in water. *Int. J. Quant. Chem. Quant. Chem. Symp.* 24:245–249.

Teng, M.-K., N. Usman, C. A. Frederick, and A. H.-J. Wang. 1988. The molecular structure of the complex of Hoechst 33258 and the DNA dodecamer d(CGCGAATTCGCG). *Nucleic Acids Res.* 16:2671–2690.

Tung, S. C., S. C. Harvey, and A. J. McCammon. 1984. Large amplitude bending motions in phenylalanine transfer RNA. *Biopolymers.* 23: 2173–2193.

van Gunsteren, W. F., and H. J. C. Berendsen. 1986. GROMOS Groningen Molecular Simulation Program. University of Groningen, Groningen, the Netherlands.

Wang, Y., G. A. Thomas, and W. Peticolas. 1987. Sequence dependent conformation of oligomeric DNAs in aqueous solution and in crystals. *J. Biomol. Struct. Dyn.* 5:249–274.

Weidlich, T., S. M. Lindsay, Q. Rui, A. Rupprecht, W. L. Peticolas, and G. A. Thomas. 1990. A Raman study of low frequency intrahelical modes in A-, B-, and C-DNA. *J. Biomol. Struct. Dyn.* 8:139–171.

Weiner, S. J., P. Kollman, D. A. Case, C. U. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. 1984. A new force field for molecular mechanical simulations of proteins and nucleic acids. *J. Am. Chem. Soc.* 106:765–784.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–238.

Westhof, E. 1987. Re-refinement of the B-dodecamer d(CGCGAATTCGCG) with a comparative analysis of the solvent in it and in the Z-hexamer d(^{5Br}CG^{5Br}CG^{5Br}CG). *J. Biomol. Struct. Dyn.* 5:581–600.

Westhof, E., and D. L. Beveridge. 1989. Hydration of nucleic acids. *In* Water Science Reviews: The Molecules of Life. Cambridge University Press, Cambridge. 24–136.

Wing, R. M., H. R. Drew, T. Takano, C. Broka, S. Tanaka, I. Itakura, and R. E. Dickerson. 1980. Crystal structure analysis of a complete turn of B-DNA. *Nature.* 287:755–758.

Withka, J. M., S. Swaminathan, J. Srinivasan, D. L. Beveridge, and P. H. Bolton. 1992. Toward a dynamical structure of DNA: comparison of theoretical and experimental NOE intensities. *Science.* 255:597–599.

Wolf, B., and S. Hanlon. 1975. Structural transitions of deoxyribonucleic acid in aqueous electrolyte solutions. II. The role of hydration. *Biochemistry.* 14:1661–1670.

Yang, Y., and B. M. Pettitt. 1996. B to A transition of DNA on the nanosecond timescale. *J. Phys. Chem.* 100:2564–2566.

York, D. M., T. A. Darden, and L. G. Pedersen. 1993. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J. Chem. Phys.* 99:8345–8348.

York, D. M., A. Wlodawer, L. G. Pedersen, and T. A. Darden. 1994. Atomic-level accuracy in simulations of large protein crystals. *Proc. Natl. Acad. Sci. USA.* 91:8715–8718.

York, D. M., W. Yang, H. Lee, T. Darden, and L. G. Pedersen. 1995. Toward the accurate modeling of DNA: the importance of long-range electrostatics. *J. Am. Chem. Soc.* 117:5001–5002.

Young, M. A., B. Jayaram, and D. L. Beveridge. 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.

Young, M. A., R. Nirmala, J. Srinivasan, K. J. McConnell, G. Ravishanker, D. L. Beveridge, and H. M. Berman. 1994. Analysis of helix bending in crystal structures and molecular dynamics simulations of DNA oligonucleotides. *In* Structural Biology: The State of the Art. Proceedings of the 8th Conversation. Adenine Press, Albany, NY. 197–214.

Young, M. A., G. Ravishanker, D. L. Beveridge, and H. M. Berman. 1995a. Analysis of local helix bending in crystal structures of DNA oligonucleotides and DNA-protein complexes. *Biophys. J.* 68: 2454–2468.

Young, M. A., J. Srinivasan, I. Goljer, S. Kumar, D. L. Beveridge, and P. H. Bolton. 1995b. Structure determination and analysis of local bending in an A-tract DNA duplex: comparison of results from crystallography, nuclear magnetic resonance, and molecular dynamics simulation on d(CGCAAAAATGCG). *Methods Enzymol.* 261:121–144.