Intrusion of Counterions into the Spine of Hydration in the Minor Groove of B-DNA: Fractional Occupancy of Electronegative Pockets

Matthew A. Young,[†] B. Jayaram,[‡] and D. L. Beveridge*,[†]

Contribution from the Department of Chemistry and Molecular Biophysics Program, Wesleyan University, Middletown, Connecticut 06459, and Department of Chemistry, Indian Institute of Technology, Hauz Kauz, New Delhi 110016, India

Received February 12, 1996. Revised Manuscript Received August 27, 1996[⊗]

Abstract: A sequence of ordered solvent peaks in the electron density map of the minor groove region of ApT-rich tracts of the double helix is a characteristic of B-form DNA well established from crystallography. This feature, termed the "spine of hydration", has been discussed as a stabilizing feature of B-DNA, the structure of which is known to be sensitive to environmental effects. Nanosecond-range molecular dynamics simulations on the DNA duplex of sequence d(CGCGAATTCGCG) have been carried out, including explicit consideration of ∼4000 water molecules and 22 Na^+ counterions, and based on the new AMBER 4.1 force field with the particle mesh Ewald summation used in the treatment of long-range interactions. The calculations support a dynamical model of B-DNA closer to the B form than any previously reported. Analysis of the dynamical structure of the solvent revealed that, in over half of the trajectory, a Na^+ ion is found in the minor groove localized at the ApT step. This position, termed herein the "ApT pocket", was noted previously (Lavery, R.; Pullman, B. *J. Biomol. Struct. Dyn.* **1985**, *5*, 1021) to be of uniquely low negative electrostatic potential relative to other positions of the groove, a result supported by the location of a Na^+ ion in the crystal structure of the dApU miniduplex [Seeman, N.; et al. *J. Mol. Biol.* **1976**, *104*, 109) and by additional calculations described herein based on continuum electrostatics. The Na^+ ion in the ApT pocket interacts favorably with the thymine O2 atom on opposite strands of the duplex and is well articulated with the water molecules which constitute the remainder of the minor groove spine. This result indicates that counterions may intrude on the minor groove spine of hydration on B-form DNA and subsequently influence the environmental structure and thermodynamics in a sequence-dependent manner. The observed narrowing of the minor groove in the AATT region of the d(CGCGAATTCGCG) structure may be due to direct binding effects and also to indirect modulation of the electrostatic repulsions that occur when a counterion resides in the minor groove "AT pocket". 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 in understanding the functional energetics and specificity of the interactions of DNA and RNA with regulatory proteins, pharmaceutical agents, and other ligands.

Introduction

The structure of DNA is well-known to be sensitive to solvent effects,[1,2] and the energetics and thermodynamics of ligand binding to DNA involve important contributions from dehydration and counterion release.[3−5] No single experimental method or measurement can provide a comprehensive description of this aspect of the system, excepting crystallography in the rare cases in which all solvent in the system is ordered;[6] thus, theoretical studies may provide a valuable source of information on solvent. A series of new nanosecond-level molecular dynamics (MD) simulations on the DNA duplex of sequence d(CGCGAATTCGCG) are presented herein. The MD model includes explicit consideration of counterions and water and is based on a newly developed empirical force field[7] and an improved treatment of long-range interactions.[8,9] Analysis of the dynamical structure provides informative new details on the molecular arrangement of solvent. The results are found to be consistent with corresponding results from Monte Carlo (MC) simulation and nonlinear Poisson−Boltzmann (PB) calculations and suggest a possible reinterpretation of the crystallographically ordered solvent peaks in the minor groove of DNA known as the "spine of hydration".

Background

The preferential stability of the right-handed A and B forms of DNA is effected by an increase in water activity from 76 to 92% in a DNA fiber and from B to Z by an increase in salt concentration in systems of GC-rich sequences.[1] A phenom-

[†] Weslyan University.
[‡] Indian Institute of Technology.
[⊗] Abstract published in *Advance ACS Abstracts,* December 1, 1996.

(1) Saenger, W. *Principles of Nucleic Acid Structure.*; Springer-Verlag: New York, 1984.
(2) Westhof, E.; Beveridge, D. L. In *Hydration of Nucleic Acids.*; Westhof, E., Beveridge, D. L., Eds.; Cambridge University Press: Cambridge, England, 1989; Vol. 5, pp 24−136.
(3) Olmsted, M. C.; Bond, J. P.; Anderson, C. F.; Record, M. T., Jr. *Biophys. J* **1995**, *68*, 634−647.
(4) Misra, V. K.; Sharp, K. A.; Friedman, R. A.; Honig, B. *J. Mol. Biol* **1994**, *238*, 245.
(5) Record, M. T., Jr.; Anderson, C. F.; Lohman, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103−178.
(6) Berman, H. M. *Curr. Opin. Struct. Biol.* **1991**, *1*, 423−427.
(7) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(8) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193.
(9) York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345.

S0002-7863(96)00459-3 CCC: $14.00

enological model for counterions around DNA was proposed some 30 years ago based on the idea of counterion condensation (CC),[5,10] in which a stable fraction of counterions, essentially independent of salt concentration, remains associated in close proximity with the DNA. For a monovalent counterions around a line charge model of polyanionic DNA, CC theory predicts 76% of the ions to be within 10 Å of the DNA surface (17 Å from the effective helical axis) and the radius, i.e., the "Manning radius", defines the boundary of the CC shell.

The principal source of experimental data on solvation at the molecular level comes from x-ray crystallography.[11] The DNA dodecamer duplex d(CGCGAATTCGCG) was found to crystallize as a hydrate in a B form of DNA that corresponds closely to the Watson−Crick double helix, with a well-defined major and minor groove structure.[12,13] The crystallographic analysis assumed all solvent peaks to be full-occupancy water molecules. Although the fraction of water that turned out to be crystallographically ordered was only ∼25% of the total, an interesting and provocative feature was noted: a sequence of ordered solvent peaks in the electron density map of the minor groove region of AT-rich tracts. This element of structure has been termed the "spine of hydration".[14] The original spine of hydration was postulated to be specific to AT tracts, where the N3 atom of adenine and the O2 atom of thymine are readily available as hydrogen bond acceptors in interactions with water molecules. The N2 donor group on guanine was thought to disrupt the spine,[14] but hydration in the CpG region of the dodecamer was blocked in the crystal by a spermine ion and by helix−helix packing and thus could not be directly observed.

The minor groove spine has been observed in a number of subsequent B-form crystal structure determinations.[15,16] NMR experiments have also been reported to detect its presence.[17−19] Theoretical calculations of varying degrees of rigor have investigated the molecular nature of the spine[20−22] and other features of DNA hydration,[23] and the possibility that the spine may extend into GpC tracts was noted. An early MD study of a solvated B-DNA system was carried out on the sequence d(CGC).[24] Sequences rich in A−T base pairs are known to preferentially stabilize and possibly rigidify the B form of DNA, with the spine providing a possible explanation. A number of minor groove-binding drugs are also known to preferentially position themselves about A−T-rich sequences.[25−27] Netropsin has been shown to induce an A to B transition in DNA by binding in the minor groove,[28] which along with thermodynamic binding studies[29] gives credence to a stabilizing propensity for the minor groove spine.

Crystal structures of the dinucleotides dCpG and dApU were reported in 1976.[30,31] (Note that we adopt the notation CpG, ApU, or ApT whenever there is any chance of ambiguity between the designation of a base pair step and a Watson−Crick base pair.) These structures were of higher resolution than that described above for the d(CGCGAATTCGCG) dodecamer, and a number of ordered solvent peaks could be assigned. In dApU, a Na^+ ion was found to reside in the protominor groove region, whereas the Na^+ ions in dCpG were associated only with phosphates. Lavery and Pullman[32] carried out calculations of the electrostatic potential of the groove regions of DNA and found the elecrostatic potential in the minor groove to be markedly electronegative in the vicinity of AT steps. Thus, experimental and theoretical results collectively raised the possibility of the residence of cations in the minor groove spine in B-DNA. In a crystal structure refinement such as that of d(CGCGAATTCGCG) (∼2.3 Å), it is understandably difficult to reliably distinguish between electron density peaks due to water and those arising from monovalent cations.

Theoretical and computational studies of the ion atmosphere of nucleic acids include contributions from CC theory as described above,[10] continuum electrostatics in which the environment is modeled as a heterogeneous dielectric continuum,[33] and MD computer simulation[34] in which the solvent molecules may be considered explicitly. Due to the high dimensionality of the problem, MD studies are computationally intensive, requiring several hundred hours of (CRAY C90) supercomputer time for each nanosecond of trajectory. MC simulation[35] and PB calculations[36] are useful for the study of the ion atmosphere in models in which the DNA is held fixed.

We have recently investigated further the dynamical structure of DNA and particularly the minor groove solvation using improved levels of theoretical and computational methodology.[37] Theoretical studies of DNA in an aqueous solution environment are complicated by the highly charged nature of the system and the concomitant importance of long-range interactions, the inherent flexibility of the DNA molecule, and the difficulties of achieving numerical stabilization and reliability in a simulation on a complex mixed solvent system within currently realizable run lengths.[38,39] A recent proposed improvement in the force field for MD simulation[7] and the application of the

(10) Manning, G. S. *Q. Rev. Biophys.* **1978**, *11*, 179−246.
(11) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Kopka, M. L.; Pjura, P. E. *Cold Spring Harbor Symp.* **1983**, *47*, 13.
(12) Drew, H. R.; Wing, R. M.; T. Takano; C. Broka; S. Tanaka; Itikura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179−2983.
(13) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761−786.
(14) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556.
(15) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hseih, S. H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751−759.
(16) Schneider, B.; Cohen, D.; Berman, H. M. *Biopolymers* **1992**, *32*, 725−750.
(17) Liepinsh, E.; Otting, G.; Wüthrich, K. *Nucleic Acids Res.* **1992**, *20*, 6549−6553.
(18) Fukuzaki, M.; Umehara, T.; Kurita, D.; Shioya, S.; Haida, M.; Mashimo, S. *J. Phys. Chem.* **1992**, *96*, 10087−10089.
(19) Kubinec, M. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 8739−8740.
(20) Subramanian, P. S.; Swaminathan, S.; Beveridge, D. L. *J. Biomol. Struct. Dyn.* **1990**, *7*, 1161−1165.
(21) Subramanian, P. S.; Ravishanker, G.; Beveridge, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1836−1840.
(22) Chuprina, V. P. *Nucleic Acids Res.* **1987**, *15*, 293−311.
(23) Westhof, E. In *Structural Water of Nucleic Acids.*; Westhof, E., Ed.; Academic Press: New York, 1990; pp 11−18.
(24) Forester, T. R.; McDonald, I. R. *Mol. Phys.* **1991**, *72*, 643−660.
(25) Kopka, M. L.; Pjura, P.; Yoon, C.; D. Goodsell; Dickerson, R. E. In *Structure and Motion: Membranes, Nucleic Acids and Proteins*; Clementi, E., Coronsiu, G., Sarma, M. H., Sarma, R. H., Eds.; Adenine Press: Guilderland, NY, 1985; pp 461 ff.
(26) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376−1380.
(27) Pjura, P. E.; Grzeskowiak, K.; Dickerson, R. E. *J. Mol. Biol.* **1987**, *197*, 257−271.
(28) Ivanov, V. I.; Minchenkova, L. E.; Minyat, E. E.; Frank-Kamentskii, M. D.; Schyolkina, A. K. *J. Mol. Biol.* **1974**, *87*, 817.
(29) Marky, L. A.; Blumenfeld, K. S.; Breslauer, K. J. *Nucleic Acids Res.* **1983**, *11*, 2857−2870.
(30) Seeman, N. C.; Rosenberg, J. M.; Suddath, F. L.; Kim, J. J. P.; Rich, A. , *J. Mol. .Biol.* **1976**, *104*, 109−144.
(31) Rosenberg, J. M.; Seeman, N. C.; Day, R. O.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 145−167.
(32) Lavery, R.; Pullman, B. *J. Biomol. Struct. Dyn.* **1985**, *5*, 1021−1032.
(33) Sharp, K. A.; Honig, B. *Curr. Opin. Struct. Biol.* **1995**, *5*, 323−328.
(34) Jayaram, B.; Beveridge, D. L. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−94.
(35) Jayaram, B.; Swaminathan, S.; Beveridge, D. L. *Macromolecules* **1990**, *23*, 3156−3165.
(36) Jayaram, B.; Sharp, K. A.; Honig, B. *Biopolymers* **1989**, *28*, 975−993.
(37) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.*, submitted.

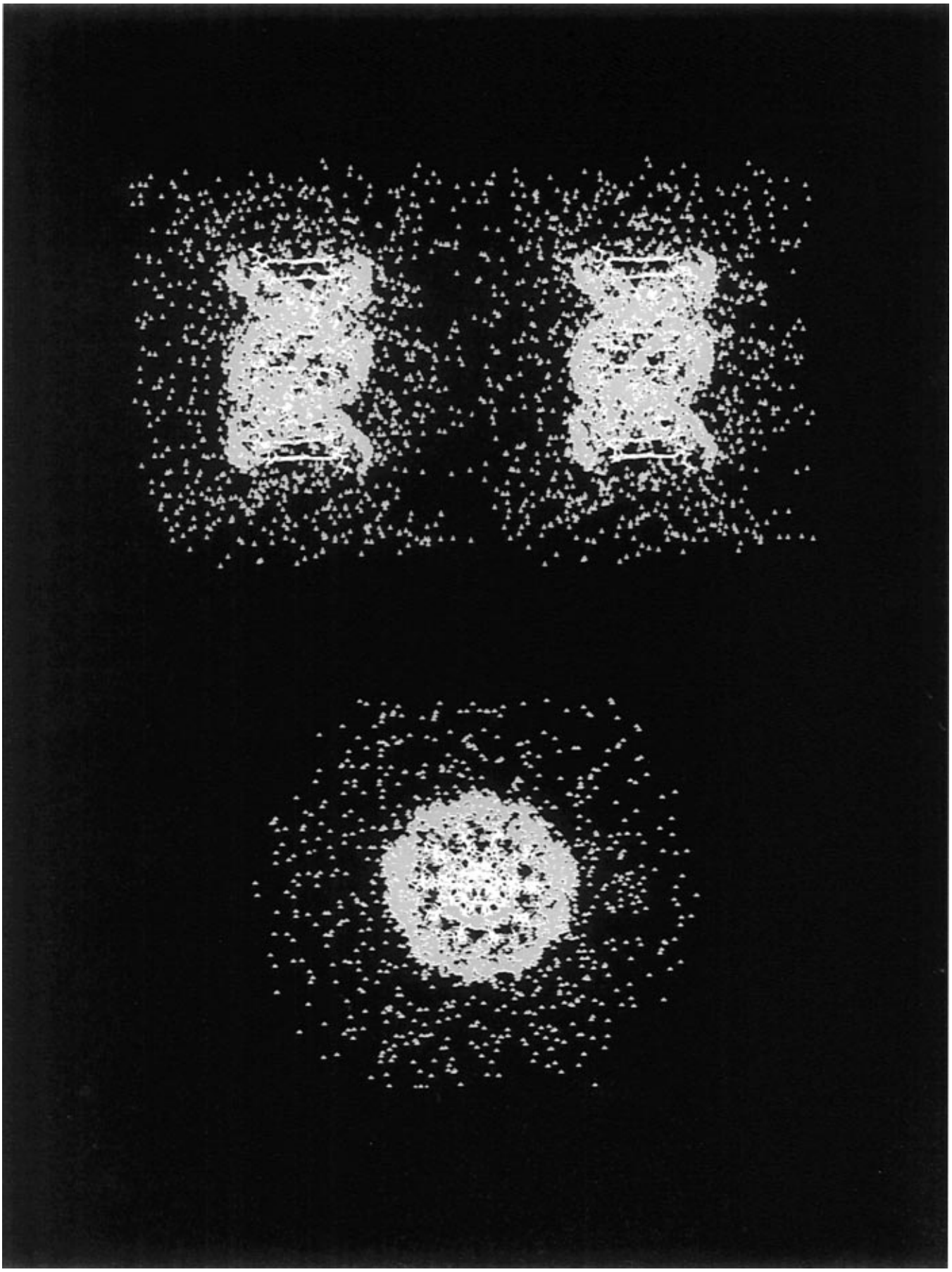

Figure 1. Calculated Na^+ counterion density around a canonical B form of the d(CGCGAATTCGCG) duplex, based on a superposition of 400 ion configurations obtained from a CIMC simulation on the DNA and 22 Na^+ ions in a dielectric continuum. Details of the method are described in ref 36.

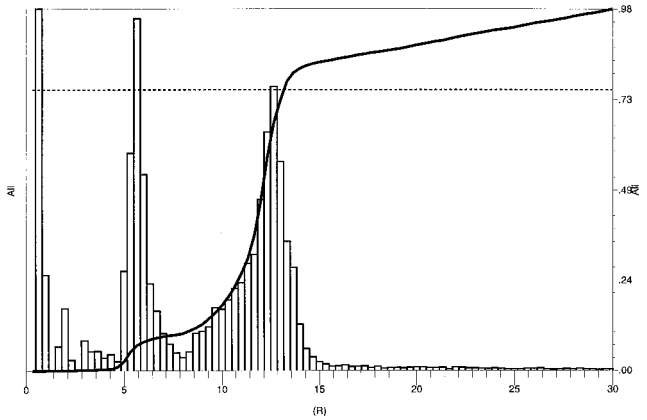

Figure 2. Calculated DNA−Na^+ radial distribution function $g(R)$ (histograms) and running coordination fraction $N_C(R)$ (solid line) from the CIMC simulation-generated configurations in Figure 1. The reported $g(R)$ is the volume-normalized density of counterions counted in concentric cylindrical shells of radius "R" measured from the center of the DNA helical axis. The $N_C(R)$ ion fraction is normalized relative to the total number of 22 ion configurations sampled and is not volume-normalized.

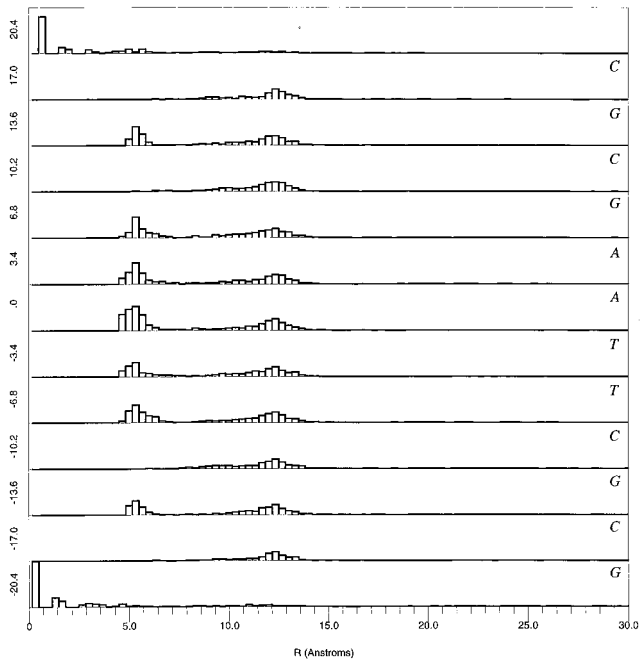

Figure 3. Calculated DNA−Na^+ $g(R)$ partitioned along the DNA helical axis into 3.4 Å tall regions indicative of single base pair steps in order to display sequence effects. The data in the individual panels of Figure 3 sums to the total $g(R)$ as presented in Figure 2.

particle mesh Ewald (PME) summation method to the long-range interaction problem[40] have resulted in more reliable MD simulations on DNA.[8,37] The new MD trajectories described in this article are based on this protocol and provide a basis for some new predictions and deeper insight into DNA solvation. MC and PB calculations, carried out initially to provide additional information on the electrostatic potential and starting structures for subsequent MDs, proved to each contribute useful additional perspectives on the solvation, as described below.

Monte Carlo Simulations. (a) Calculations. MC calculations of the ion atmosphere about a canonical B form of DNA phosphates with water treated as a dielectric continuum were described previously.[36,41] In the current project, we extended this computational model to the treatment of all-atom DNA as described by the charges and van der Waals in the new AMBER force field,[7] and carried out an MC Metropolis simulation on 22 mobile Na^+ counterions around a fixed canonical form of the d(CGCGAATTCGCG) duplex. The effect of solvent water on DNA−Na^+ and Na^+−Na^+ interactions was modeled with a sigmoidal dielectric screening function.[42] The calculations employed Born−von Karmann mass-conserving periodicity within a cylindrical cell of height 70 Å and a radius of 32 Å.

(b) Results. The MC calculated counterion density around the d(CGCGAATTCGCG) DNA helix is shown in Figure 1. This result provides graphic, independent support of the general idea of CC as anticipated almost 20 years ago by Manning.[10] Further analysis of the counterion distribution reveals significant new details. The DNA−Na^+ radial distribution function (RDF) $g(R)$ from the MC-generated ensemble and the corresponding running coordination number $N_C(R)$ are plotted in Figure 2. The MC RDF shows two inflection points, one at 6.5 Å and another at 14.9 Å from the axis of the DNA. The major inflection point at 14.9 Å defines the Manning radius for this model. Some 82% of the counterions are found to be condensed within this radius, as compared with Manning's value of 76% at 17 Å based on a line charge of DNA. Analysis of the inflection point in the MC at 5.5 Å , which is followed by a small plateau region extending to ∼9 Å, shows that it is produced by counterions positioned within the major and minor groove regions of the double helix. Beyond this point, the distribution corresponds to counterions distributed along the backbone of the DNA and beyond.

To elucidate sequence effects in the DNA−Na^+ distribution, the MC RDF was decomposed into contributions corresponding to each base pair. Figure 3 shows the resulting set of MC-calculated DNA−Na^+ RDFs, partitioned into 3.4 Å high cylinders along the long axis of the helix. Note particularly that the groove region (∼4.5−6.5 Å) of the central ApApTpT tract shows a notable counterion density. In the flanking sequences some density is indicated at the GpC steps and virtually none at the CpG steps. Proximity analysis[43] shows this density to be concentrated in the minor groove, consistent with results reported from other laboratories.[44] The lowest energy counterion configuration identified from the MC ensemble of structures shows one Na^+ ion located at the ApT step of the minor groove of the CGCGAATTCGCG sequence, in a condition of primary coordination with the carbonyl groups of the successive thymines on opposite strands of the duplex. This position, noted above to be an electronegative region, is referred to henceforth as the "ApT pocket".

Continuum Electrostatics. (a) Calculations. The general characteristics of nonlinear Poisson−Boltzmann calculations on DNA were described previously by Jayaram and Honig.[36] To investigate sequence effects, nonlinear PB calculations were

(38) Beveridge, D. L.; Swaminathan, S.; Ravishanker, G.; Withka, J. M.; Srinivasan, J.; Prevost, C.; Loiuse-May, S.; Langley, D. R.; Dicapua, F. M.; Bolton, P. H. Molecular Dynamics Simulations on the Hydration, Structure and Motions of DNA Oligomers. In *Water and Biological Macromolecules*; Beveridge, D. L., Swaminathan, S., Ravishanker, G., Withka, J. M., Srinivasan, J., Prevost, C., Loiuse-May, S., Langley, D. R., Dicapua, F. M., Bolton, P. H., Eds.; 1993.

(39) Beveridge, D. L.; Ravishanker, G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 246−255.

(40) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001.

(41) Jayaram, B.; Swaminathan, S.; Beveridge, D. L.; Sharp, K.; Honig, B. *J. Phys. Chem.* **1990**, *23*, 3156−3165.

(42) Hingerty, B. E.; Ritchie, R. H.; Ferrel, T. L.; Turner, J. E. *Biopolymers* **1985**, *24*, 427−439.

(43) Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102*, 4287−4294.

(44) Lamm, G.; Wong, L.; Pack, G. R. *Biopolymers* **1994**, *34*, 227−237.

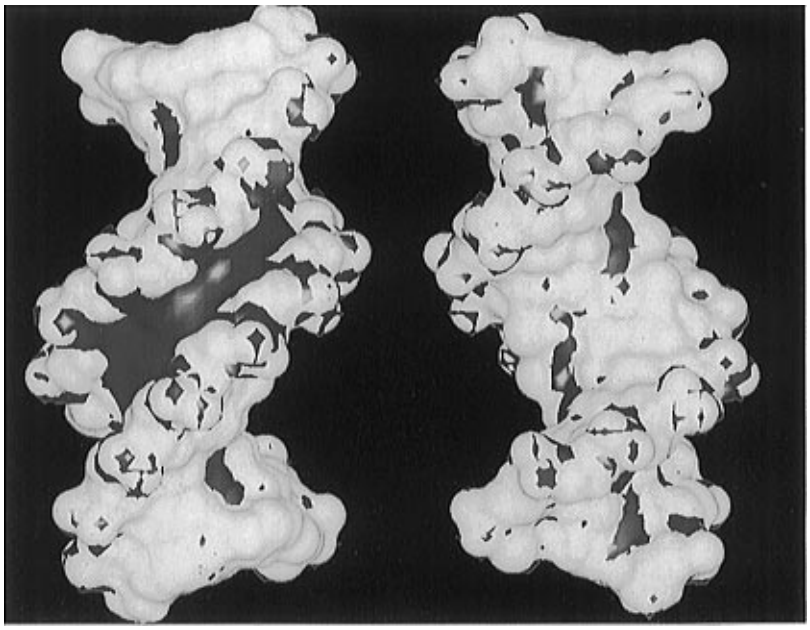

Figure 4. Calculated electrostatic potential of the d(CGCGAATTCGCG) duplex based on nonlinear Poisson−Boltzmann calculations performed using the DELPHI program. Shown in red is a plot of the −5 kT isopotential surface with the DNA viewed into the minor groove (left) and the major groove (right).

carried out on the d(CGCGAATTCGCG) duplex using the Cornell et al. charges and van der Waals parameters.[7] Solvent water was assigned dielectric constant of 80, and the dielectric was set at two within the DNA. A charge grid of 1 Å resolution was constructed within a box of dimensions 65 by 45 by 45 Å. Nonlinear Poisson−Boltzmann calculations were performed using the program Delphi[45,46] on the DNA−counterion system, at a salt concentration sufficient to provide electrical neutrality, 0.145 M. Results from these calculations can be sensitive to the assumed geometry.

(b) Results. The PB calculated electrostatic potential for the *Eco*RI dodecamer is shown in Figure 4. The central AATT region of the minor groove of the molecule (Figure 4a) was found to exhibit a large negative potential extending along the floor of the groove. The surrounding CpG steps in the flanking sequence show some evidence of electronegative patches; however, these patches are more localized at the base pair step level than was seen for the AATT region. The electrostatic potential of the major groove (Figure 4b) shows only small, localized electronegative patches in the CG region. The results indicate that the minor groove of the AATT region is favorable for occupation by counterions and are consistent with the results described above from the MC calculations. Together, the MC and PB results raise the distinct possibility that Na^+ counterions could be favorably situated in the AATT region of the minor groove of B-form DNA. Further studies based on molecular dynamics simulation were subsequently carried out to further explore this issue.

Molecular Dynamics Simulations. (a) Calculations. A series of MD simulations were performed on the DNA oligonucleotide d(CGCGAATTCGCG) duplex with 22 anionic phosphate groups using the AMBER 4.1 suite of programs.[47] Each strand was terminated with an OH group. The solvent was composed of a water bath containing 22 mobile Na^+ counterions. All calculations employed the Cornell et al. force field[7] for the DNA and Na^+ and the TIP3P model for water[48,49] as incorporated into AMBER 4.1. A series of four independent simulations were carried out to evaluate sensitivity of the current protocol to initial starting conditions and to assess trajectory stability and convergence. All four simulations begin with a canonical B form of the DNA derived from fiber diffraction.[50] The first three differ only in the initial placement of the 22 Na^+ ions, as shown in Figure 5. The placement for the first MD, denoted "MCI", (Figure 5a) is the lowest energy configuration of the 22 Na^+ ions about a fixed DNA as obtained from the MC calculation described above, using a modification of the "MCIONS" program[35] and henceforth referred to as MCI. The second MD was initiated from the placement "ESI" (Figure 5b) obtained from the AMBER CION program.[47] The CION program iteratively positions each of the Na^+ ions about the DNA at the point of next lowest electrostatic potential. Note the MCIONS calculation results in an Na^+ ion residing at the ApT step of the minor groove, whereas CION does not. There is also an ion residing at a GC step of the 5′ flanking sequence in the MCI structure. We observed that the lowest energy configuration with all particle interactions taken into consideration (as in MCIONS) does not coincide with that of iterative sequential placement of the ions (i.e., CION). The ion placement for a third simulation, denoted "BSI", was constructed by simply placing the ions 6 Å away from the phosphorus atoms along the O−P−O bisector of each phosphate group. The bisector configuration also served as the starting configuration for a fourth MD in which the DNA atoms were restrained to their canonical starting positions and only the ions and water molecules were allowed to move. This latter MD calculation provides for a direct comparison with the MC and PB calculations about the fixed DNA structure.

(45) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1988**, *9*, 327−335.
(46) Sharp, K., Honig, B., Eds. *DelPhi*; 2.3 ed.; Biosym Technologies Inc.: San Diego, CA, 1992.
(47) Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, Fergusen, D. M., Seibel, G. L., Singh, U. C., Weiner, P., Kollman, P., Eds. *AMBER 4.1*; 4.1 ed.; UCSF: San Francisco, CA, 1995.
(48) Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 341−345.
(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−936.
(50) Arnott, S., Campbell-Smith, P. J., Chandrasekaran, R., Eds. *Atomic Coordinates and Molecular Conformations for DNA-DNA, RNA-RNA, and DNA-RNA Helices*; CRC Press: Cleveland, OH, 1976; Vol. 2, pp 411−422.

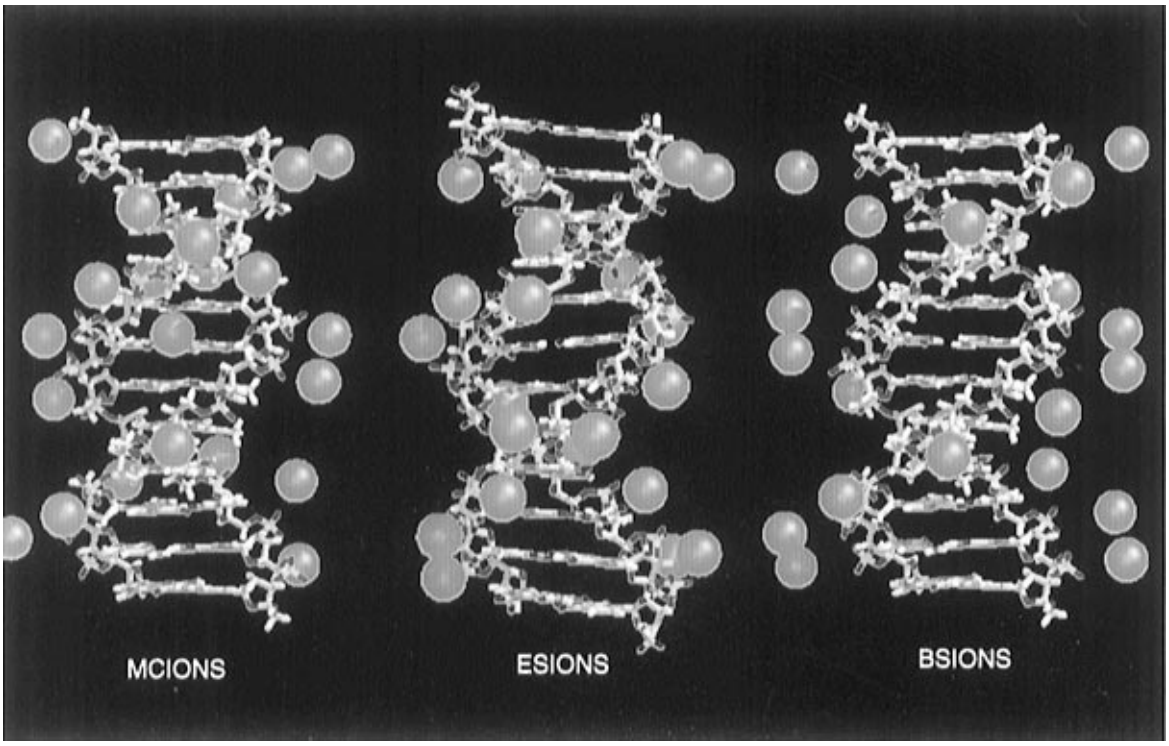

Figure 5. Initial positions of the 22 Na^+ ions around DNA prior to MD simulation: (a) MCI configuration; (b) ESI configuration; (c) BSI configuration.

Following placement of the counterions about a canonical B form of the DNA structure, each system was hydrated in a rectangular prism by TIP3P water molecules in a pre-equilibrated configuration. The box dimensions were truncated to achieve a minimum distance of approximately 12 Å beyond all DNA atoms in all directions, resulting in a box size of ~45 × 45 × 70 Å, solvating the DNA with ~4,000 water molecules in each case. The effective concentration of the DNA sample within the periodic box is ~15 mM. Prior to MD, the various starting configurations were subjected to a series of energy minimizations to relieve any local atomic clashes. Periodic boundary conditions were treated via the PME summation method[40] as implemented in AMBER 4.1, with a 9 Å cutoff for direct space nonbonded calculations and a 0.000 01 convergence criteria.

MD simulations were initiated with several rounds of semirestrained minimizations, followed by an unrestrained minimization 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 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. Next, the restraints on the ions were reduced 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.[51] A 2 fs time step was used for all simulations. The system energy was stable when coupled with a SHAKE constraint of 0.000 05 Å on all covalent bonds involving hydrogen atoms under the constant-temperature conditions maintained by the Berendsen algorithm. The MCI simulation became the most informational of the set, as described below, and was analyzed in detail after a run length of 1500 ps (1.5 ns). 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.

Table 1. Average and Mode for Selected DNA Helicoidal and Torsional Parameters over 1.5 ns of MD Compared Against the Fiber Diffraction Values[50] for A- and B-Form DNA

	1.5 ns MD		canonical values	
	av	mode	A-DNA	B-DNA
RIS	3.47	3.63	2.56	3.4
TLT	−0.67	−0.18	0	0
ROL	4.5	0.29	0	0
TWS	31.51	33.61	32.7	34
XDP	−0.88	−0.91	−5.43	0
INC	−1.23	−2.97	19.12	1.5
Pucker	133.66	143.98	13	154
Chi	244.53	240.66	206	258

(b) Results. A superposition of 20 DNA/counterion structures from the MCI MD trajectory at evenly spaced intervals throughout simulation is shown in Figure 6. Overall, the results as summarized in Table 1 indicate that the MD based on AMBER 4.1 and the Cornell et al. force field support a better dynamical model of B form DNA than any other simulation we have yet examined.[39] Similar demonstrations have been submitted by Kollman and co-workers based on nanosecond-length MD on DNA decamers.[8,52] We confirm independently that the Cornell et al. force field with PME appears to be a significant improvement over previous versions in the AMBER suite of programs and, as far as we can tell to date, other proposed MD protocols for nucleic acids as well. A detailed comparison of the results of the Cornell et al. force field with CHARMM on the d(CGATCG) duplex support this claim as

(51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(52) Cheatham, T. E., III; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.

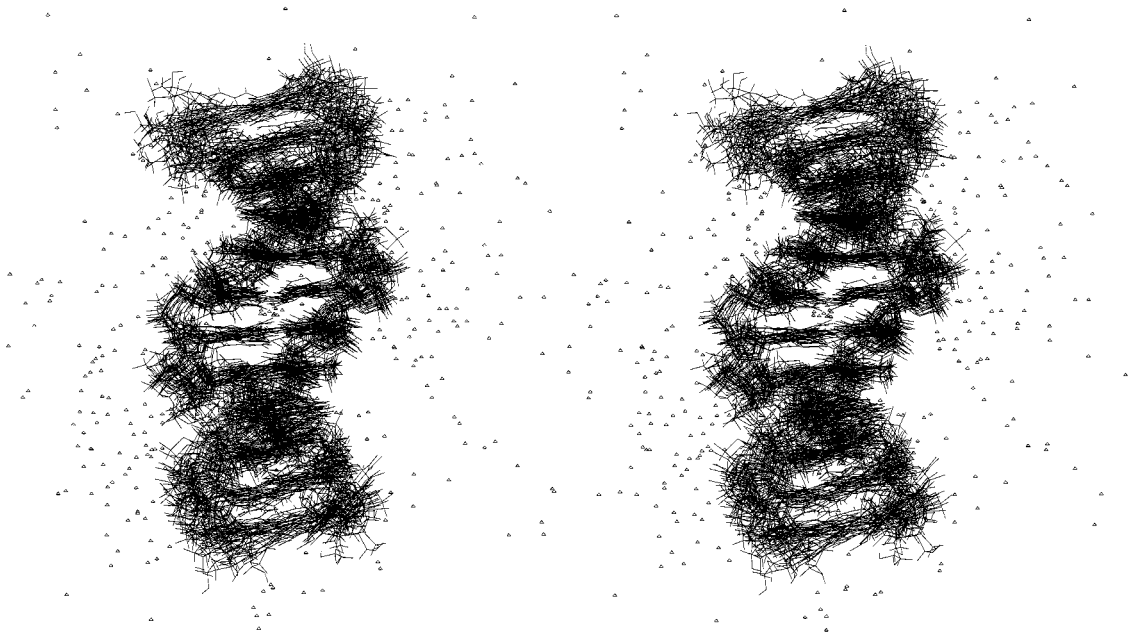

Figure 6. Calculated dynamical structure of the d(CGCGAATTCGCG) duplex based on 15 structures of a 1.5 ns MD trajectory including counterions and water. The inital DNA−Na^+ configuration of the MD was MCI (cf. Figure 5a).

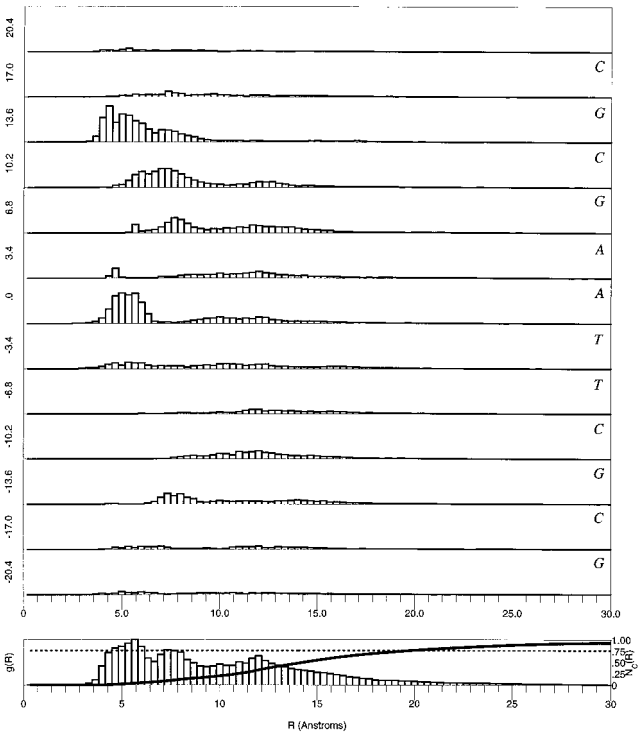

Figure 7. Calculated DNA-Na^+-$g(R)$ from the 1.5 ns MD trajectory shown in Figure 6, partitioned according to base pair step analogous to the analysis presented for the CIMC simulation in Figure 3.

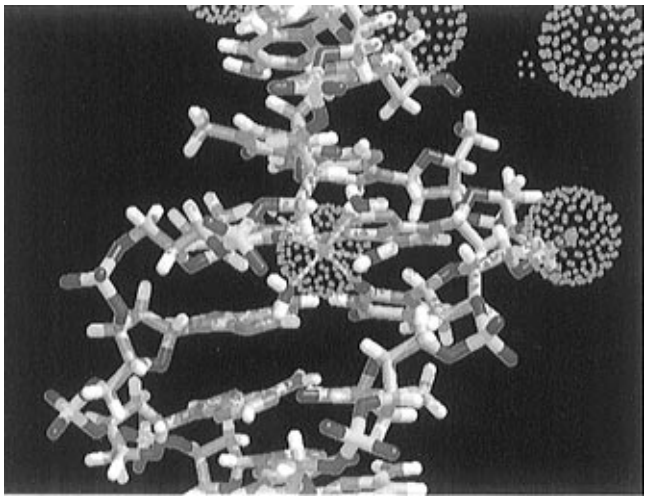

Figure 8. Structure obtained from the MCI MD trajectory on the d(CGCGAATTCGCG) system (see Figure 6 and text) showing the structural detail of the incorporation of a sodium ion in the first solvation shell of the minor groove. The structure is representative of those obtained in the first 500 ps segment of the trajectory.

well,[53] and further detailed characterization studies will appear shortly. A full description of the structure and motions of the d(CGCGAATTCGCG) duplex in this present round of simulations is being provided elsewhere.[37] In the present study we focus on the analysis of solvent structure and dynamics, particularly in the minor groove region of the duplex.

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. A plot of the calculated DNA−Na^+ $g(R)$ and $N_C(R)$ for the MCI MD trajectory is presented in Figure 7. This plot shows and inflection points at 7.5 and at 21 Å, the latter being not quite as prominent as in Figure 3 due to the internal motions of the DNA. The inflection at 21 Å defines the Manning radius for this model, and the area under the curve indicates that 76% of the counterions are described by this model as "condensed." The presence of the inflection point at 8.5 Å in the MD ROF indicates that counterions found within the groove region of the double helix in MC and PB calculations are also present to a significant extent in the all-atom, explicit solvent, dynamical model from MD in which the DNA is permitted to execute thermal motions.

An animation of the dynamical motion of the system as described by the MD was prepared using the program MPEG and can be examined on the website http://linus.chem.wesleyan.edu/∼mayoung. The ion initially positioned in the AT pocket in the MCI placement equilibrated between and roughly equidistant from the Thy7 and Thy19 O2's of the central ApT step

(53) Langley, D. R. Private communication.

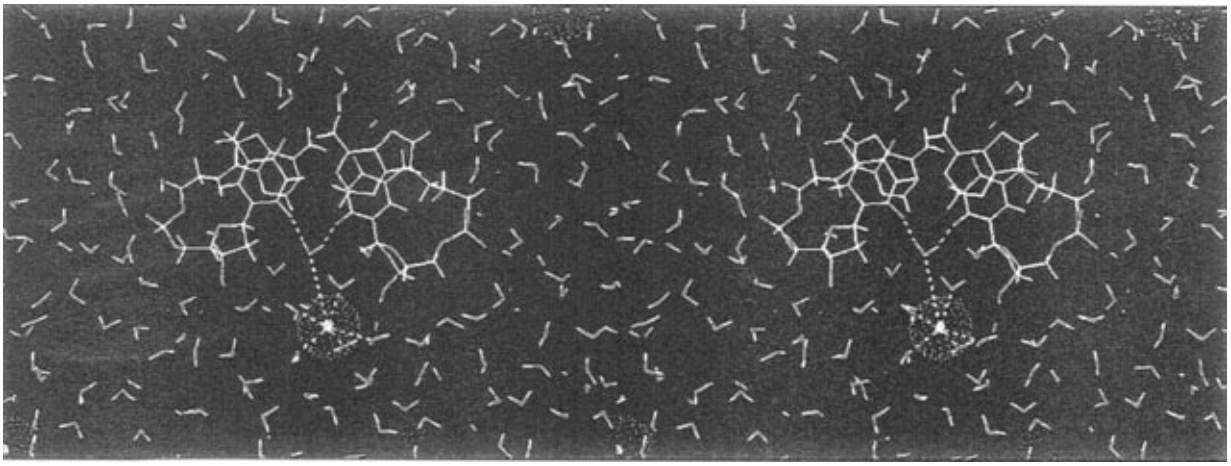

Figure 9. Structure obtained form the MCI MD trajectory (see Figure 6 and text) showing Na^+ ions residing in the second hydration shell of the minor groove "spine".

in the sequence. An MD snapshot of the solvated DNA from this segment of the MCI trajectory is presented in Figure 8 and shows clearly the Na^+ ion in proximity to the two O2 atoms from the central Thy7 and Thy19 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. 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 as shown in Figure 9. Near the end of this segment of the simulation, a population of Na^+ ions developed in the 3′-CG region as well. This point will be investigated further in future studies.

The corresponding results from the additional MD simulations on d(CGCGAATTCGCG) described in the preceding section permit us to examine the sensitivity of results to the choice of starting structure. MPEG animations of these trajectories are also accessible at http://linus.chem.wesleyan.edu/~mayoung. The ESI simulation began with ions placed according the iterative electrostatic protocol embodied in the CIONS routine of AMBER 4.1. Here the initial configuration features ions distributed asymmetrically along the sugar phosphate backbone. The MD was extended out to 500 ps. The dynamical structure of the DNA turned out to be much the same as that described above for the MCI simulation. The counterions diffused readily, and in one case, an Na^+ ion penetrated the minor groove in the vicinity of the AATT tract. It got as far as the second solvation shell but had not at 500 ps returned back into the ApT pocket.

The BSI MD was performed for two distinct cases: one in which the DNA structure was permitted a full range of dynamical motion, and one in which the atomic positions of the DNA were restrained to the canonical B form. The latter MD provides us with a control on the sensitivity of results to DNA dynamics, as well as a point of comparison of this work with the results of other computational methods and an earlier MD.[24] The BSI (mobile DNA) MD was carried out for 500 ps and did not result in any counterions in the grooves in this time frame. The BSI (fixed DNA), also carried out to 500 ps, showed Na^+ ions situated at two places in the first coordination shell of the AATT tract in the minor groove spine. The DNA−Na^+

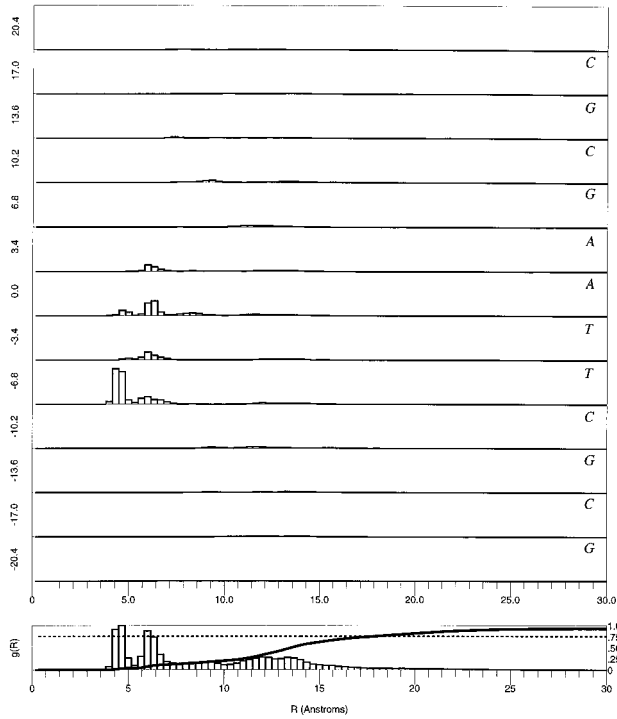

Figure 10. Calculated DNA−Na^+ $g(R)$ from the BSI MD simulation based on a fixed B-form d(CGCGAATTCGCG) duplex, partitioned according to base pair step (cf. Figures 3 and 7).

RDF for this MCI MD, broken down by sequence, is shown in Figure 10. An MD snapshot structure showing the positioning of the two ions in the minor groove is shown in Figure 11. The groove is maintained in a relatively wider state in the canonical B structure, providing more space for ions to enter. It appears that mutual Na^+:Na^+ electrostatic repulsions prevents either one from entering the most favorable region of the ApT pocket, and a compromise structure results. Note (cf. Figures 7 and 8), when the DNA is permitted dynamical motion, small but significant differences arise.

Discussion

The results of the MC, PB, and MD calculations as described above provide leading theoretical evidence of the possibility that counterions intrude into the minor groove spine of hydration in a B-form DNA double helix in solution. Furthermore, the MD indicates that this phenomenon occurs at sequence specific electronegative pockets in the electrostatic potential of the DNA.

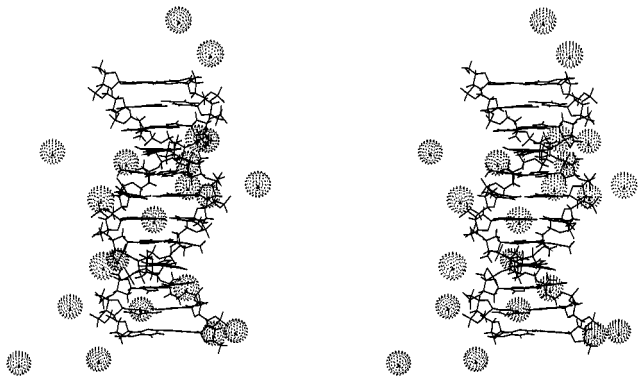

Figure 11. Structure obtained from the BSI MD based on fixed B-form d(CGCGAATTCGCG) duplex, in which two ions are found in the primary solvation shell of the minor groove.

In the case at hand, the functional groups of certain base pair steps, for example, the Thy−O2's of an A−T pair followed by a T−A, combine to provide a locally favorable electrostatic environment for counterions, an ApT pocket. In this section we will further consider the plausibility of this suggestion, some ideas for further experiments, and the possible implications of sequence-specific ionic solvation in the grooves on the structural biology of nucleic acids.

The force of the argument from the theoretical side comes from the combined results of MC simulations on a primitive model of d(CGCGAATTCGCG) and counterions as well as MD simulations on an all-atom model of the sequence including counterions and water. These results are consistent with earlier calculations of the electrostatic potential of B-DNA based on a point charge model by Lavery and Pullman, the nonlinear PB calculations of Jayaram et al.,[36] and the followup PB calculations on d(CGCGAATTCGCG) described in this article, which were carried out to further elucidate the sequence effects issue. None of the previous crystallographic or NMR studies of oligonucleotides relevant to the minor groove spine were geared toward distinguishing water molecules from sodium ions. Existing experimental data do not at present permit us to further confirm or deny this possibility. However, with the models provided by this study, further high-resolution analyses may be pursued and form the basis of a study currently being carried out in collaboration with Prof. H. M. Berman.

The previous report most relevant to the predictions herein is the crystal structure of ApU,[30] in which a sodium ion was found to be complexed to the carbonyl groups of successive uracils that protrude into the protominor groove. This structure was noted by Seeman et al. to be an "unexpected" and "most unusual" consequence of a 2-fold symmetric local field. 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) shows essentially close coincidence with the structure seen in prototype in the crystal structure of ApU. 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. Thus electrostatic pockets of the genre described herein must be treated in terms of fractional occupancy in the analysis and interpretation of crystallographic or NMR results.

The idea suggested by the results of this simulation, linked to earlier considerations of Seeman as described in his film *Deep Groove* (Seeman, N., private communication), is that the edges

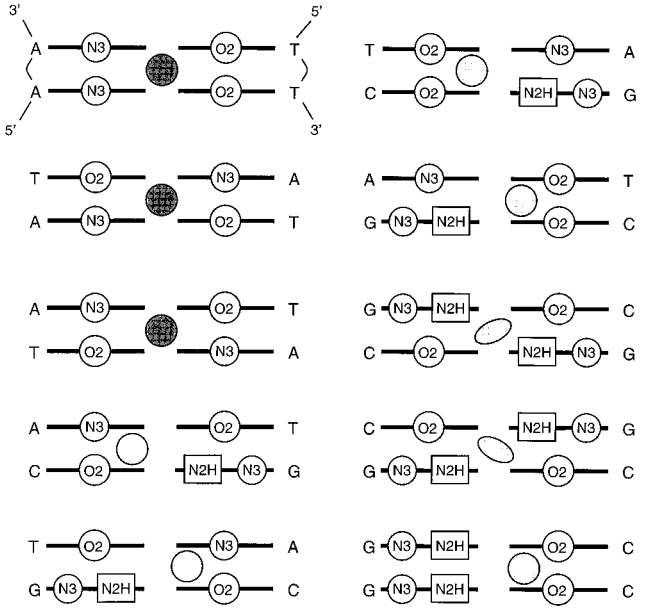

Figure 12. Schematic view of DNA base pair steps in the minor groove indicating possible locations of electrostatic pockets, indicated by shaded regions, with shading proportional to the putative efficacy of the site. On the base pair edges, electronegative atoms are indicated by circles, electropositive hydrogen atoms by squares, and hydrophobic methyl groups by triangles.

of the nucleic acid bases A, T (U in RNA), G, and C present a pattern of electronegative and electropositive sites at the base of the major groove and the minor groove of the duplex; at the base pair step level, the pattern motif is extended: the superposition of the electrostatic potentials of polar sites of the like charge can lead to regions in which the potential is markedly amplified. For example, the proximity of two or more negatively charged sites on successive base pairs in a step can lead to a localized region of enhanced electronegativity between the two bases of the step. This electronegative "pocket" may thus be especially favorable for some degree of occupancy by one of the mobile counterions. The counterion, with no intervening water to modulate its effect, can interact strongly with the DNA atoms and produce a perturbing influence on structure, leading to local axis bending and alteration of groove width due to attenuation of phosphate−phosphate anionic repulsions.

The overall idea can be described succinctly based on the diagrammatic representations of base pair steps shown in Figures 12 and 13. Figure 12 schematically depicts the steps where electronegative pockets may form as a consequence of polar groups lining the minor groove for DNA or RNA. The analysis suggests the presence of robust pockets at the ApT, TpA, and ApA steps, in which the bases on both strands of the duplex present a superposition of purely electronegative potential to the grooves. Further consideration of Figure 12 also suggests the possibility of intrastrand electronegative pockets at CpA, ApC, TpC, and CpC steps of the minor groove. However, these sites would be asymmetrically positioned in the groove compared with those of ApT, TpA, and ApA and give rise to smaller electronegative patches. Intrastrand pockets, also covering a lesser amount of configuration space than those of the ApT, TpA, and ApA steps, are suggested at CpG steps and CpC steps of the minor groove. Thus, all sites in the minor groove may exhibit some propensity for counterions, but the ApT, TpA, and ApA steps are the most likely to host ions since they arise from a superposition of both intrastrand and interstrand electronegative potentials and support larger, more extensive pockets. Furthermore, runs of ApT, TpA, and ApA steps in a sequence

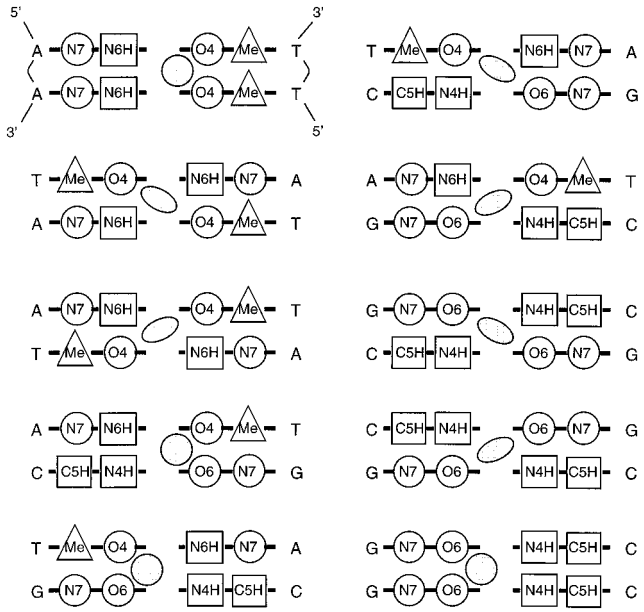

Figure 13. Diagrammatic view of possible electrostatic pockets of DNA base pair steps in the major groove. The same convention is used as in Figure 12.

produce an extended, contiguous pocket region with a correspondingly high propensity for ions.

In the major groove, Figure 13, there are no pockets comparable to those of the ApT, TpA, and ApA steps of the minor groove, and thus an overall lesser propensity for fractional occupancy by mobile cations is indicated. The most favorable region for ion occupancy in the major groove according to this theory would be the intrastrand pocket at the GpG step. Other intrastrand pockets, smaller in size, are indicated at TpT, TpG, and GpT of the major groove. Major groove pockets resulting from an interstrand superposition of potentials are possible at ApT, TpA, CpT, GpA, CpG, and GpC but also with lesser weight in configuration space. The steps of the major groove all present a mix of electronegative and electropositive potentials that, at least at this level of description, would tend to limit the extent of electronegative pockets and thus diminish the efficacy of attracting mobile cations to the major groove region. Thus, the most likely sites of electronegative pockets and fractional occupancy of mobile counterions in the dynamical structure of the system are the ApT, TpA, and ApA steps of the minor groove. This idea is essentially supported by the various theoretical calculations reported herein.

The implications of these results are of potentially considerable interest in the structural biology of DNA and also 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 from the MD in which an Na^+ ion occupies the ApT pocket (cf. Figure 7) and one in which it has exchanged with water in this site (Figure 8). The structure with a Na^+ in the AT pocket shows a noticeable narrowing of the minor groove local to the region of ion association. This may indicate that ion complexation plays a significant role in the narrowing of the minor groove in AT-rich DNA sequences, but further studies are required to investigate this hypothesis. We are proceeding to quantify the description shown schematically in Figures 12 and 13 with extensive electrostatic potential calculations using DelPhi and with further nanosecond-level MD studies that will permit theoretical estimates of the residence times of ions in the grooves and critical consideration and further analysis of the pocket hyothesis.

In concluding this section, we note that if direct ionic interactions in both 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 the various simplistic models previously utilized by ourselves and others, including simple distance-dependent dielectrics, reduced charges on phosphates,[54−56] salt-dependent interphosphate screening functions,[57] or artificially large hydrated counterions (solvatons)[58] will not suffice to describe the physics of the system adequately. Nonlinear Poisson−Boltzmann continuum electrostatics can accommodate this effect to some extent, but since they are base on assumed forms of the DNA, they may miss important structural consequences. All-atom MD including water and counterions explicitly can, however, provide a proper description of the system, if based on accurate potentials and carried out to sufficiently long run lengths.

Summary and Conclusions

The theoretical and computational studies described herein provide leading theoretical evidence for the idea of the participation of complexed ions in the primary coordination sphere of nucleotide bases in the minor groove of B-DNA. The circumstances in which this phenomenon may occur can be anticipated by screening structures based on continuum electrostatics. However, our study indicates that the nature of the structures is likely to be linked intimately with the temporal dynamic motions of the DNA. The differential effects of hydration vs ionic solvation, all of which reflect a detailed force balance of numerous local attractions and repulsions in the system at room temperature, will thus be crucial in providing an accurate model of the system. The accuracy of MD on DNA must be considered an evolving story, but new and improved force fields, the PME method for long-range interactions, and increased supercomputer power appear to have demonstrably improved the situation.

The plausibility of the idea of intrusion of ions into the spine of hydration was discussed in the context of available experimental data. The particular example developed in this article is the case of ions in the ApT pocket of the minor groove, but it is worth noting that this phenomenon may arise in other sequence-specific regions of both the major and minor grooves of nucleic acid helices and affect the functional energetics and specificity of ligand interactions. It has not escaped our attention that 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,

(54) McConnell, K. M.; Nirmala, R.; Young, M. A.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 4461−4462.
(55) Tidor, B.; Irikura, K. K.; Brooks, B. R.; Karplus, M. *J. Biomol. Struct. Dyn.* **1983**, *1*, 231−252.
(56) Levitt, M. *Cold Spring Harbor Symp. Quant. Biol.* **1983**, *47*, 251−262.
(57) Soumpasis, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5116−5120.
(58) Singh, U. C.; Weiner, S. J.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 755−759.
(59) Ravishanker, G., Ed. *Molecular Dynamics ToolChest*; 1.0 ed.; Wesleyan University: Middletown, CT, 1991.
(60) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *J. Mol. Graphics* **1988**, *6*, 13−27.
(61) Huang, C. C.; Pettersen, E. F.; Klein, T. E.; Ferrin, T. E.; Langridge, R. *J. Mol. Graphics* **1991**, *9*, 230−236.
(62) *Insight II*, 95.0 ed.; Biosym/MSI: San Diego, CA, 1995.

and morphological structure. Similar effects may be present in RNA helices and ligand interactions as well. Furthermore, this phenomenon may be the origin of weak salt dependencies in CC and counterion release on ligand binding to DNA. Extensions and variations of the idea of ApT pocket−ion binding phenomena may thus have important implications in understanding the functional energetics and specificity of the interactions of DNA and RNA with regulatory proteins, pharmaceutical agents, and other ligands. We plan to explore these issues further in subsequent nanosecond-length MD studies on longer oligonucleotide sequences.

Acknowledgment. This research was supported primarily by Grant GM-37909 from the National Institutes of Health. M.A.Y. is the recipient of an NIGMS 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 ToolChest,[59] UCSF Midas-Plus,[60,61] and Insight II.[62] Cray time was provided for these projects by the Pittsburgh Supercomputer Center and the Frederick Biomedical Supercomputer Center of the National Cancer Institute, NIH. The authors gratefully acknowledge discussions and advice on this project from Dr. G. Ravishanker, Prof. H. M. Berman, Dr. Richard Lavery, and Prof. Ned Seeman. We thank Amy Duffield for a critical reading of the final manuscript.

JA960459M