Comparison of the X-ray Crystal, NMR Solution, and Molecular Dynamics Calculated Structures of a Tandem G·A Mismatched Oligonucleotide Duplex

Edward P. Nikonowicz and David G. Gorenstein*

Contribution from the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907. *Received April 10, 1992*

Abstract: We have compared structures derived from the NOESY distance restrained molecular dynamics (MD) and unrestrained MD calculations of a tandem G·A mismatched decamer deoxynucleotide duplex d(CCAAGATTGG)$_2$ with the X-ray crystal structure. The importance of using restrained MD in water (with counterions and a box of water molecules containing the G·A mismatched decamer) or in the gas phase is evaluated. While structures derived starting from model-built coordinates and crystal coordinates using restrained MD in both water and gas-phase environments do obey the set of input restraints, global changes in the DNA observed in the gas phase occur faster than in the aqueous medium. Thus, the use of structures generated with minimal gas-phase simulation as input for further solution-phase refinement is quite reasonable given the results observed here. Although a number of significant differences exist, many features of the crystal structure are found in the NMR solution structure. Both NMR spectroscopy and X-ray crystallography describe the rather large propeller twist at the mismatch site *and* the bifurcated hydrogen-bonding pattern of the mismatched guanine residue. However, the crystal structure maintains a straight rod conformation in order to allow adjacent molecules in each unit cell to stack and form the crystal. Through wedging of the adjacent thymidine residues, and thus providing opposing kinks at either end, the duplex in the crystal is able to recover its global cylindrical shape. In the aqueous environment, however, no forces are present which would cause such kinking and destacking of the thymidine–thymidine residues. Thus, an overall curvature or bend in the duplex is observed. Finally, unrestrained molecular dynamics (in a 100-ps trajectory in the gas or solution phase) is unable to adequately locate either the crystal or solution NMR conformation. The structures determined from 100 ps of free dynamics depend upon the initial starting structure.

The past decade has seen a tremendous increase in the application of both computational[1-6] and multidimensional NMR techniques[7-11] for the structural determination of biomolecules. Molecular dynamics (MD) calculations now provide a powerful tool to describe the motions of molecules in solution and to assess the relative stability of various conformations.[12-14] The growing interest in such computational methods has been spurred in part by the availability of high-resolution NMR structural data. Typically, solution structures are obtained by evaluation of interproton distances from 2D-NMR nuclear Overhauser effect spectroscopy (NOESY).[7] Refinement methods are varied but have included distance geometry[15] and restrained molecular mechanics or dynamics.[16,17]

Although NMR can produce accurate short interproton distances, homonuclear ^{1}H–^{1}H NOESY is incapable of defining long-range distances, i.e., from end to end of a duplex DNA helix. Further, while NMR spectroscopy has been largely successful in defining the overall conformation of DNA as well as some sequence-specific variations in the local conformation of DNA,[18-24] several studies show disagreement between structures derived from X-ray crystallography and NMR-derived solution conformations.[24-27] The discrepancy between solution NMR structures and crystallographic X-ray structures raises the question of whether the sequence-specific structural variations observed in the X-ray crystallographic studies are the result of crystal packing forces. Indeed, variations in the local conformation may be found in different crystalline forms of the same duplex.[28,29] Alternatively, the differences between solution and crystal structures may simply reflect the intrinsic inability of NMR to accurately define these subtle structural details. For example, dynamic averaging is a potential problem in NOESY distance-restrained MD calculations.

(1) Post, C. B.; Brooks, B. R.; Karplus, M.; Dobson, D. M.; Artymiuk, P. J.; Cheetham, J. C.; Phillips, D. C. *J. Mol. Biol.* **1986**, *190*, 455–479.
(2) (a) Seibel, G. L.; Singh, U. C.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6537–6540. (b) Rao, S.; Kollman, P. *Biopolymers* **1990**, *29*, 517–532.
(3) Withka, J. M.; Swaminathan, S.; Beveridge, D. L.; Bolton, P. H. *J. Am. Chem. Soc.* **1991**, *113*, 5041–5049.
(4) Han, K. H.; Syi, J. L.; Brooks, B. R.; Ferretti, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 2818.
(5) de Vlieg, J.; Berendsen, H. J. C.; van Gunsteren, W. F. *Proteins: Struct. Funct. Genet.* **1989**, *6*, 104–127.
(6) Zielinski, T. J.; Shibata, M. *Biopolymers* **1990**, *29*, 1027–1044.
(7) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, NY, 1986.
(8) Gorenstein, D. G. In *DNA Structures, Methods in Enzymology*; Lilley, D. M., Dahlberg, J. E., Eds.; Academic Press, Inc.: Orlando, FL, 1992; Vol. 211, pp 254–286.
(9) Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F.; Kaptein, R. *Biochimie* **1985**, *67*, 707.
(10) Nilsson, L.; Clore, G. M.; Gronenborn, A. M.; Brünger, A. T.; Karplus, M. *J. Mol. Biol.* **1986**, *188*, 455–475.
(11) Gorenstein, D. G.; Meadows, R. P.; Metz, J. T.; Nikonowicz, E. P.; Post, C. P. In *Advances in Biophysical Chemistry*; Bush, C. A., Ed.; JAI Press: Greenwich, 1990; Vol. 1, pp 47–124.
(12) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(13) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287–303.
(14) Van Gunsteren, W. F.; Berendsen, H.; Geurtsen, R.; Zwinderman, H. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 287–302.
(15) Havel, T. F.; Kuntz, I. D.; Crippen, G. M. *Bull. Math. Biol.* **1983**, *45*, 665–720.
(16) Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179–182.
(17) Clore, G. M.; Gronenborn, A. M.; Brunger, A. T.; Karplus, M. *J. Mol. Biol.* **1985**, *186*, 435–455.
(18) Lefevre, J.-F.; Lane, A. N.; Jardetzky, O. *Biochemistry* **1987**, *26*, 5076–5090.
(19) Nilges, M.; Clore, G. M.; Gronenborn, A. M.; Piel, N.; McLaughlin, L. W. *Biochemistry* **1987**, *26*, 3734–3744.
(20) Gorenstein, D. G.; Schroeder, S. A.; Fu, J. M.; Metz, J. T.; Roongta, V. A.; Jones, C. R. *Biochemistry* **1988**, *27*, 7223–7237.
(21) Assa-Munt, N.; Kearns, D. R. *Biochemistry* **1984**, *23*, 791.
(22) Clore, G. M.; Gronenborn, A. M.; Moss, D. S.; Tickle, I. J. *J. Mol. Biol.* **1985**, *185*, 219–226.
(23) Patel, D. J.; Shapiro, L.; Hare, D. *Annu. Rev. Biophys. Chem.* **1987**, *16*, 423–454.
(24) Rinkel, L. J.; van der Marel, G. A.; van Boom, J. H.; Altona, C. *Eur. J. Biochem.* **1987**, *163*, 275–286.
(25) Nikonowicz, E. P.; Meadows, R. P.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 4193–4204.
(26) Sklenář, V.; Bax, A. *J. Am. Chem. Soc.* **1987**, *109*, 7525–7526.
(27) Joshua-Tor, L.; Rabinovich, D.; Hope, H.; Frolow, F.; Appella, E.; Sussman, J. L. *Nature* **1988**, *334*, 82–84.
(28) Jain, S.; Sundaralingam, M. *J. Biol. Chem.* **1989**, *264*, 12780–12784.
(29) Shakked, Z.; Guerstein-Guzikevich, G.; Eisenstein, M.; Frolow, F.; Rabinovich, D. *Nature* **1989**, *342*, 456–460.

0002-7863/92/1514-7494$03.00/0

Similarly, computational methods alone are incapable of adequately searching conformational space for the "correct" structure, due to the computationally time-demanding nature of MD calculations. For example, even with nanosecond MD simulations, only a limited number of alternative macromolecular structures can be explored within a reasonable time frame using supercomputer resources. Thus, the best solution structures will be obtained by a combined NMR and computational method. However, the final structures will depend critically upon the quality of the NMR restraints and the quality of the molecular mechanics/dynamics force field and calculations.

It has been suggested that the accuracy of the interproton distances obtained from the NOESY spectra has been limited primarily by failure to properly account for spin diffusion (indirect magnetization transfer from multiple spins). The errors in the interproton distances obtained from analysis of NOESY data when not explicitly accounting for indirect transfer of magnetization has been well documented.[30-33] The distance errors resulting from inadequate data treatment may be as large as 1.3 Å. Recently, in order to obtain more accurate distances, the Bloch equations of magnetization have been solved numerically[18,34] or by a complete relaxation matrix approach[11,25,30,33,35-38] including a "hybrid matrix" approach.[37-41] These methods take into account the effects of spin diffusion, which allows for the measurement of interproton distances with a higher degree of precision and accuracy. Several studies have addressed the question of the accuracy and precision of distance geometry and NOESY distance-restrained molecular dynamics to define duplex structures in solution.[10,33,42-46] The basic conclusions were that the complete relaxation method was indeed superior to the two-spin approximation methodology and that back-calculation of the NOESY volume matrix could more accurately reproduce the correct cross-relaxation rates and distances. It was also demonstrated that the hybrid matrix/restrained MD methodology can even provide direct information on the time-averaged conformation of duplex oligonucleotides, including a variety of helical parameters and backbone torsional angles.[46,47]

Molecular dynamics and NMR restrained MD calculations on biomacromolecules have been largely performed in the gas phase. For proteins, gas-phase MD refinement is not of significant concern because tertiary folding of the protein provides numerous long-range NOESY distance restraints. However, in nucleic acids, factors such as hydration and the presence of counterions are of greater concern due to the flexible nature of the molecules, the highly charged surfaces of nucleic acids, and the relatively high surface to volume ratios.[3] The absence of tertiary interactions in duplex DNA also restricts the refinement to include only intranucleotide and immediate neighbor nucleotide distance restraints. Just how important is the molecular dynamics force field in accurately reproducing the NMR-determined solution conformation of nucleic acids?[49] We address this issue by using a more realistic force field (DNA plus counterions in water) than that in a gas-phase restrained MD calculation.

Free, unrestrained gas-phase MD simulations generally fail to adequately reproduce solution or crystal structures. MD simulations which explicitly include water molecules and counterions appear to provide a more realistic model of the structure and dynamics of DNA.[2-6,50,51] One of the early studies on DNA in water with sodium counterions was performed on the model-built pentamer sequence d(CGAGA)·d(TCGCG) which contains a central G·A mismatch.[51] In this study, unrestrained molecular dynamics was run for a period of 90 ps. The authors found that the hydrogen bonding was poorer than in the nonmismatched, standard G·C pentamer[52] due to effects of the mismatch being propagated throughout the sequence. More recent molecular dynamics studies of DNA have concentrated on sequences containing the canonical Watson–Crick base-pairing schemes.[3,5]

Both X-ray crystallography and NMR spectroscopy have been used to probe the structural anomalies of mismatched base pairs.[23,53-56] Although each of the three most general conformations of duplex DNA (A-, B-, and Z-DNA) are represented in the mismatch structures observed by X-ray crystallography, the actual conformation adopted by the duplex may in part be the result of crystallization techniques and the crystallization medium. High-resolution X-ray crystal structures of base pair mismatch containing structures include two G·A mismatched DNA fragments, a dodecamer d(CGC*G*AATT*A*GCG)$_2$[54] and a decamer d(CCAA*G*ATTGG)$_2$.[55] The mismatched adenosine residues were found to be in a syn orientation[56] or an anti orientation,[55] respectively.

Comparison of the structures of the tandem G·A mismatched decamer d(CCAA*G*ATTGG)$_2$ in solution using ^{1}H–^{1}H NOESY restrained molecular dynamics and free unrestrained MD methods is presented in this paper. Interproton distance restraints derived using the hybrid relaxation matrix methodology and gas-phase dynamics[25] were incorporated into restrained molecular dynamics simulations using a bath of water and sodium counterions. Comparisons are made between structures calculated in vacuo[25] and in water with counterions to assess the use of strictly gas-phase methods for structural refinement. In addition the water-calculated structures and the structures refined in vacuo are compared to the structure determined in the crystalline state.

Experimental Procedure

Synthesis and NMR-Derived Distance Restraints. Details of the synthesis and NMR assignments of the decamer are given in ref 57. ^{1}H–^{1}H

(30) Keepers, J. W.; James, T. L. *J. Magn. Reson.* **1984**, *57*, 404–426.
(31) Dobson, C. M.; Olejniczak, E. T.; Poulsen, J. F. M.; Ratcliffe, R. G. *J. Magn. Reson.* **1982**, *48*, 97–110.
(32) Olejniczak, E. T.; Gampe, R. T.; Fesik, S. W. *J. Magn. Reson.* **1986**, *67*, 28–41.
(33) Post, C. B.; Meadows, R. P.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1990**, *112*, 6796–6803.
(34) Banks, K.; Hare, D.; Reid, B. *Biochemistry* **1989**, *28*, 6996–7010.
(35) Bothner-by, A. A.; Noggle, J. H. *J. Am. Chem. Soc.* **1979**, *101*, 5152–5155.
(36) Baleja, J. D.; Moult, J.; Sykes, B. D. *J. Magn. Reson.* **1990**, *87*, 375–384.
(37) Nikonowicz, E. P.; Meadows, R.; Post, C.; Jones, C.; Gorenstein, D. G. *Bull. Magn. Reson.* **1989**, *11*, 226–229.
(38) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299–311.
(39) Boelens, R.; Koning, T. M. G.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *J. Magn. Reson.* **1989**, *82*, 290–308.
(40) Nikonowicz, E. P.; Meadows, R. P.; Fagan, P.; Gorenstein, D. G. *Biochemistry* **1991**, *30*, 1323–1334.
(41) Borgias, B. A.; Gochin, M.; Kerwood, D. J.; James, T. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **1990**, *22*, 83–100.
(42) Broido, M. S.; James, T. L.; Zon, G.; Keepers, J. W. *Eur. J. Biochem.* **1985**, *150*, 117–128.
(43) Gronenborn, A.; Clore, G. *Biochemistry* **1989**, *28*, 5978–5984.
(44) Pardi, A.; Hare, D. R.; Wang, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8785–8789.
(45) Meadows, R. P.; Post, C. B.; Kaluarachchi, K.; Gorenstein, D. G. *Bull. Magn. Reson.* **1991**, *13*, 22–48.
(46) Kaluarachchi, K.; Meadows, R. P.; Gorenstein, D. G. *Biochemistry* **1991**, *30*, 8785–8797.
(47) Powers, R.; Jones, C. R.; Gorenstein, D. G. *J. Biomol. Struct. Dyn.* **1990**, *8*, 253–294.
(48) Oshiro, C. M.; Thomason, J.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1049–1064.
(49) Gorenborn, A.; Clore, G. *Biochemistry* **1989**, *28*, 5978–5984.
(50) Laaksonen, A.; Nilsson, L.; Jönsson, B.; Teleman, O. *Chem. Phys.* **1989**, *129*, 175–183.
(51) Rao, S. N.; Singh, C.; Kollman, P. *Isr. J. Chem.* **1986**, *27*, 189.
(52) Singh, U. C.; Weiner, S. J.; Kollman, P. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 755.
(53) (a) Kennard, O.; Hunter, W. N. *Q. Rev. Biophys.* **1989**, *22*, 327. (b) Patel, D. J.; Kozlowski, S. A.; Ikuta, S.; Itakura, K. *Biochemistry* **1984**, *23*, 3207–3217. (c) Patel, D. J.; Kozlowski, S. A.; Ikuta, S.; Itakura, K. *Biochemistry* **1984**, *23*, 3218–3226. (d) Kan, L.; Chandrasegaran, S.; Pulford, S. M.; Miller, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 4263–4265. (e) Kouchakdjian, M.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *J. Mol. Biol.* **1988**, *202*, 139–155. (f) Dodgson, J. B.; Wells, R. D. *Biochemistry* **1977**, *16*, 2367–2374.
(54) Brown, T.; Hunter, W. N.; Kneale, G.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2402–2406.
(55) Prive, G. G.; Heinemann, U.; Chandrasegaran, S.; Kan, L.; Kopka, M. L.; Dickerson, R. E. *Science* **1987**, *238*, 498–504.
(56) Kennard, O.; Cruse, W. B. T.; Nachman, J.; Prange, T.; Shakked, Z.; Robinovich, D. *J. Biomol. Struct. Dynam.* **1986**, *3*, 623–647.

STRUCTURE IDENTIFICATION

Initial Structures:

CTL : Crystal Structure

MDL : Model B DNA (A-G /Anti-Anti)

MA : MORASS/MD Refined In Vacuo A-DNA

MD : MORASS/MD Refined In Vacuo B-DNA

Molecular Dynamics Methods

U : Unrestrained

R : Restrained

G : Gas Phase 100 ps MD

W : Water Phase 100 ps MD

R/U : 100 ps Restrained, followed by 25ps Unrestrained MD

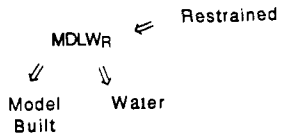

Figure 1. Identification of various decamer structures.

NOESY distance restraints were determined using the program MORASS as described in ref 40. The program MORASS (MORASS: multiple Overhauser relaxation analysis and simulation)[33,58] was used to calculate volume and rate matrices using the complete relaxation matrix approach[11,30,32] and also to implement the hybrid matrix methodology.[37,38] The refinement process has been monitored using the following percentage rms derivation in the calculated and experimental NOESY volumes ($\%rms_{vol}$):

$$\%rms_{vol} = \left(\frac{1}{N} \sum_{ij} \left(\frac{v_{ij}^{theoret} - v_{ij}^{exptl}}{v_{ij}} \right)^2 \right)^{1/2} 100\% \qquad (1)$$

where $v_{ij}^{theoret}$ and v_{ij}^{exptl} represent the theoretical and experimental volume matrix elements, respectively. rms_{vol}^{exptl} is the error between calculated and experimental volumes determined with the experimental volume in the denominator of eq 1 and $rms_{vol}^{theoret}$ contains the calculated volume in the denominator.[25]

NOESY Distance-Restrained Molecular Dynamics Calculations of the Duplex. Energy minimization and restrained molecular dynamics calculations were carried out either in vacuo or in a water bath containing sodium counterions using AMBER3[13] on a Silicon Graphics IRIS 4DS. MIDAS Plus software from T. E. Ferrin and R. Langridge, UCSF, was used for all modeling. Details of the gas-phase calculations are reported in ref 40.

Water MD calculations were performed by placing a Na^+ ion at each phosphate bisector 3.0 Å from the phosphorus atom and immersing the solute in a box of TIP3P water molecules.[59] After H_2O molecules > 7 Å from any solute atom were removed, 150 steps of steepest descent constant volume minimization and 7 ps of MD equilibration were carried out holding the solute atoms in a belly. Next, restrained and unrestrained molecular dynamics simulations of 100 ps using constant pressure were performed on each of the structures. Two of the restrained structures, $MA_{R/U}$ and $MB_{R/U}$, were allowed to undergo an additional 25 ps of unrestrained MD after the first 100 ps of restrained MD.

Final structures were determined by averaging the Cartesian coordinates of the last 20 ps of structures (either 106–125 ps or 81–100 ps) and subjecting the structures to 40 cycles of unrestrained gas-phase minimization using a distance-dependent dielectric to relieve bond length and bond angle distortions caused by the averaging process. All simulations were run under constant pressure at 1.0 bar (0.987 atm) and constant temperature at 300 K except from 10 to 20 ps where the temperature was increased to 360 K and from 35 to 40 ps when the temperature was increased to 340 K (see Figure 1). A temperature coupling parameter of $\tau = 0.2$ and a pressure relaxation time of 0.4 ps^{-1} were used; the coordinates were stored every 50 steps. The SHAKE algorithm with a tolerance of 0.0004 Å was used to restrain all bonds, and the time step for the integration was set to 2 fs. The cutoff distance for nonbonded pairs was set to 8.5 Å and updated every 50 steps, and a constant dielectric of $\epsilon = 1.0$ was used.

The starting structures used in the water dynamics simulations were as follows: Arnott model-built B-DNA with the A6 residue in an anti conformation (MDL), the crystal structure (CTL; from ref 55), and two structures previously derived from gas-phase NMR distance-restrained MD[40] which began as A-type (MA) or B-type (MB) helices.

Results and Discussion

Convergence of the Structures. The application of NMR methods to the area of macromolecular structure determination has become closely linked to various computational methods for refining the solution data. Implementation of the complete relaxation matrix approach for the analysis of NOESY data has allowed us and others[25,36–39,60] to extract more accurate interproton distances than are obtainable via the "two-spin" methodology. However, the validity of structures simulated in vacuo *may* be questioned since the use of a distance-dependent dielectric constant represents an approximation of the very important electrostatic energy terms. A particularly important question for us has been the origin of the apparent "bending" observed in the gas phase refined structures of duplexes containing mispairs or bulges relative to structures solved by X-ray crystallography.

Since it may be argued that an in vacuo environment does not adequately reflect the forces encountered in a true solvated environment, we have attempted to address this issue. Eight different water (W) simulations were considered, including both unrestrained and restrained simulations using previously derived NMR NOESY distance restraints (see Figure 1 for a summary of the structure labels). The distance restraints used in the simulations were those derived from the fifth MORASS/MD iteration starting from the A-DNA $A6_{anti}$ model.[40] Three gas-phase (G) simulations were also carried out.

The crystal coordinates (CTL) and those of a model duplex built using B-type DNA geometry (MDL) were used as starting structures for several of the simulations. The remaining restrained MD runs were performed using the crystal structure (CTL_R) and the MORASS/NOESY distance-restrained structures (M) generated after 36 ps of gas-phase dynamics, MA_R and MB_R.[40] The two MORASS-refined structures (MA and MB) initially began as model built A-DNA (A) and B-DNA (B) structures, with the mismatch G5 and A6 placed in an anti–anti orientation.

Both restrained (R) and unrestrained (U) simulations were performed on these four initial starting structures, CTL, MDL, MA, and MB. Finally, in order to remove any unfavorable bond geometries and to ensure that the water (W) refined structures MAW_R and MBW_R were in fact restrained to an energy minimum, restraints were removed after 100 ps in water and the simulation was continued for a further 25 ps producing structures $MAW_{R/U}$ and $MBW_{R/U}$. Note that three of the Na^+ ions rapidly became displaced from the phosphate bisectors, where they were originally placed, during the final 20 ps of the unrestrained portion of simulation on $MAW_{R/U}$. Comparison of the Cartesian coordinates of averaged snapshots from each run indicates that the structures MAW_R and $MAW_{R/U}$ and the structures MBW_R and $MBW_{R/U}$ are nearly identical to each other, (~0.6-Å rms difference of atomic Cartesian coordinates).

Convergence of the potential energy of the system (all structures) as well as convergence of the restraint energy (restrained MD structures) was used to define the portion of the MD run from which the coordinate sets to be used for averaging were taken. Figure 2 parts A and B depict the potential and restraint energy trajectories for the distance-restrained MD calculation of the crystal structure in water. Note that regions A and C of Figure 2A denote the portions of the MD run during which the temperature was increased to 360 K and to 340 K, respectively. The system appears to reach equilibrium after about 60 ps, which is consistent with other restrained MD simulations.[61] The potential

(57) Nikonowicz, E. P.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 8845–8858.

(58) Meadows, R.; Post, C. B.; Gorenstein, D. G. MORASS Program, 1989.

(59) Jorgensen, W.; Chandrasekhar, J.; Madra, J.; Impey, M.; Klein, R. *J. Chem. Phys.* **1983**, *79*, 926.

(60) Nikonowicz, E. P.; Roongta, V.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1989**, *28*, 8714–8725.

Table I. Compilation of the Energies and %Rms_{vol} Variations of the Final Water and Gas Phase Derived Structures

	energy, kcal/mol					
struct	E_{tot}[a] vacuo	E_{tot}[b]	$E_{tot/DNA}$[c] DNA + Na^+ + H_2O	$E_{electro}$[d]	E_{vdW}[e] DNA	%rms_{vol}[f] / %$R_{theoret}$/ %R_{exptl}
MAW_R	−726					45/52
MBW_R	−721	148 ± 25	−5525 ± 46	−144 ± 16	−272 ± 10	43/54
$MAW_{R/U}$	−729					48/58
$MBW_{R/U}$	−723	5 ± 23	−5550 ± 52	−193 ± 18	−295 ± 18	45/51
$CTLW_R$	−702	307 ± 180	−5433 ± 136	−128 ± 21	−198 ± 86	63/72
$CTLW_U$	−756	17 ± 30	−5486 ± 90	−247 ± 14	−234 ± 62	172/76
$MDLW_R$	−736					44/57
$MDLW_U$	−764	−60 ± 84	−5418 ± 86	−236 ± 26	−298 ± 19	168/104
MAG_R	−740	na[g]	na	na	na	42/45
MBG_R	−736	na	na	na	na	41/51
$CTLG_R$	−746	na	na	na	na	46/53
CTL	−748	na	na	na	na	191/75

[a] Total energy (kcal/mol) for minimizations done in vacuo using a distance-dependent dielectric. No restraints were applied during the minimization. [b] Total energy (kcal/mol) involving DNA–DNA interactions only. The decamer has been minimized in water in the presence of counterions (the DNA–water and DNA–Na^+ contributions have been subtracted out). Value reported is the average of five structures ± standard deviation observed for those five structures. [c] Total energy is the summation of the DNA and the solvent interaction energy and the DNA and the counterion interaction energy. Note that the numbers of solvent molecules contained in the four structural families were as follows: MAW 2510, MBW 2210, CTLW 2033, and MDLW 2449. [d] Electrostatic energy term for the DNA–DNA interactions only. [e] Non-bonded van der Waals energy for DNA–DNA interactions only. The van der Waals energies for DNA–solvent and DNA–counterion have been subtracted out. [f] Rms deviation between the theoretical NOESY volumes NOESY volumes calculated from MORASS using these structures and the experimentally measured volumes. [g] Not applicable = na.

energy profiles of the other simulations are comparable.

Evaluation of the Energies of the Structures and Their Fit to the Experimental NMR Data. All structures were minimized in vacuo in order that a direct comparison of the energies could be made between the structures derived in water and those calculated in the gas phase. MAG_R and MBG_R are the final structures derived from gas phase (G) NOESY distance-restrained/MORASS (M) refined structures starting from the MA or MB initial model, respectively.[40] $CTLG_R$ is the structure derived by averaging the final 20 ps of a 100-ps restrained MD run in the gas phase starting from the crystal (CTL) coordinates. The energies are reported in Table I.

A comparison of in vacuo energies of the minimized structures determined using the distance-dependent dielectric (second column, Table I) indicates that the water-derived structures MAW_R and MBW_R are highest in energy (20–30 kcal/mol) compared to the unrestrained crystal ($CTLW_U$) or model-built ($MDLW_U$) structures and would thus appear to be less energetically favorable. However, a closer inspection as to the origin of this energy difference shows that nearly all of the difference is attributable to the nonbond electrostatic term (data not shown). Thus examining the energies in a solvated environment with counterions present might be a more valid evaluation of the energy terms contributing to the structure.

Fifteen snapshots from the water simulation $MBW_{R/U}$ were minimized through 100 cycles via the steepest descent method. The energies of the snapshot structures were decomposed and analyzed to determine how each of the individual energy terms varied for a single run. Five snapshots from each of the other water structures were minimized as above and the decomposed energies for each structure averaged and used as the basis for comparison. Table I lists the total energy for the DNA alone ($E_{tot/DNA}$; solvent molecules and counterions are excluded from the energy evaluation) as well as the DNA + Na^+ + H_2O total energy of interaction (E_{tot}). Because the number of water molecules differs in the various simulations, comparison of this latter value is meaningless. Also reported are the averages for the electrostatic energy term ($E_{electro}$) and nonbonded van der Waals energy term (E_{vdW}) of the various structures (again the counterion and solvent contributions have been removed).

We have also used methods to assess the consistency of the calculated structures with the NMR data as previously described.[25,37,60] Back-calculation of the theoretical NOESY volume matrix using the different minimized (40 cycles of in vacuo

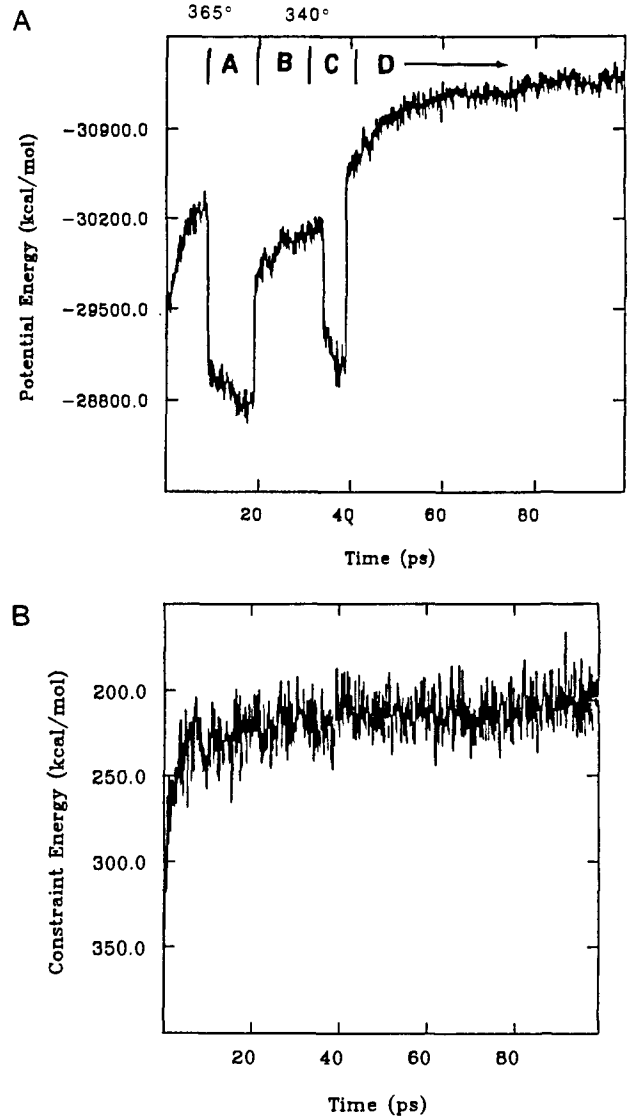

Figure 2. (A) Plot of the potential energy of the system during the 100 ps restrained molecular dynamics in water of the structure CTL. 300 K except in regions A (360 K) and C (340 K). (B) Plot of the restraint energy of the system during the 100 ps restrained molecular dynamics in water of the structure CTL.

(61) Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 2773–2781.

minimization) 20 ps averaged structures and comparison to the experimental volumes (Table I) clearly indicate that aqueous, unrestrained MD alone starting from model-built duplex ($MDLW_U$) or aqueous or gas-phase unrestrained MD beginning from the crystal coordinates ($CTLW_U$ and $CTLG_U$) is not capable of locating the NMR-defined minima within the time span of the simulation. Using $\%R_{theoret}$ (a measure of the percentage errors in the theoretically calculated NOESY volumes relative to the experimental data; Table I) the unrestrained MD leads to structures with errors in the NMR volumes of 172%, 168%, and 191% for $MDLW_U$, $CTLW_U$, and $CTLG_U$, respectively. With one exception, the $\%R_{theoret}$ for all of the restrained MD-calculated structures varies between 41% and 48%. Because of the inverse sixth power dependence to the NOESY volumes, this error is comparable to distance errors of ca. 8%, which is within the experimental error of the NOESY data.[25,37,60] Similar (although smaller) variation is noted for the $\%R_{exptl}$ criterion of quality of fit of the structure to the solution NMR data. Restrained MD refinement in water starting from the crystal structure ($CTLW_R$) did not give quite as good of a reproduction of the NOESY spectra ($\%R_{theoret} = 63\%$) as all of the other refined structures, even that starting from model-built DNA ($MDLW_R$). Note that the $\%R_{theoret}$ of the starting crystal coordinates is 191%, and presumably a longer refinement would have led to a converged refinement value of ca. 50%. Thus the theoretical NOESY spectra calculated from structures derived using NOESY distance restraints are consistent with the experimental NOESY spectra. These latter results are expected since the restraining distances were derived from the NMR data.

Comparison of Structural Features. The crystal structure of the $d(CCAAGATTGG)_2$ mismatched decamer has an average minor groove width of 7.2 Å (the P–P distance across the minor groove minus 5.2 Å to include the van der Waals surface of the phosphate group[55]). All of the structures calculated using either gas-phase or water MD maintain a narrower minor groove width, ranging about 4–5 Å on average. This P–P difference appears to result from the different accomodation of the mismatched G and A nucleotides in the duplex. The variation in the width of the minor groove within each NMR structure is also quite large, ~3–4 Å. Although less pronounced, the unrestrained structures, $MDLW_U$ and $CTLW_U$, show a similar intrastructure variation in minor groove width. An additional feature which is common to all restrained structures is that the widest portion of the groove is not symmetric across the mismatch site, i.e., the width of the minor groove on either side of the mismatch is not equal. Indeed, the greatest intrastructure variation of the minor groove width originates from the P–P distance on either side of the mismatch. This feature is also observed in structures $MDLW_U$ and $CTLW_U$, but again is less pronounced.

The helix twist parameter describes the relative rotation about a central axis from one base pair to the next. Figure 3 compares the helical twist of the crystal structure vs sequence with several other calculated structures. Following the practice of the crystallographers,[62] and required by the molecular two-fold axis of symmetry, the helix twist has been averaged for the two strands of the solution structures and is thus symmetric about the center position. As can be seen, the helix twist between the two structures MAW_R and MBW_R is maintained through the unrestrained portion of the MD simulations for $MAW_{R/U}$ and $MBW_{R/U}$ (Figure 3A). As shown in Figure 3 parts A and C the same trend for these helical twist variations obtained from restrained MD in the gas phase (MAG_R and MBG_R) is maintained through the water simulations.

Importantly, the crystal structure and the structures produced by unrestrained molecular dynamics show an opposite change in the helix twist at the $A^4pG^5 - G^5pA^6$ and $G^5pA^6 - A^6pT^7$ steps than that observed for the NMR restrained structures. This suggests that MD alone is *not* capable of reproducing helical trends which are observed in solution [at least in the 100-ps time span

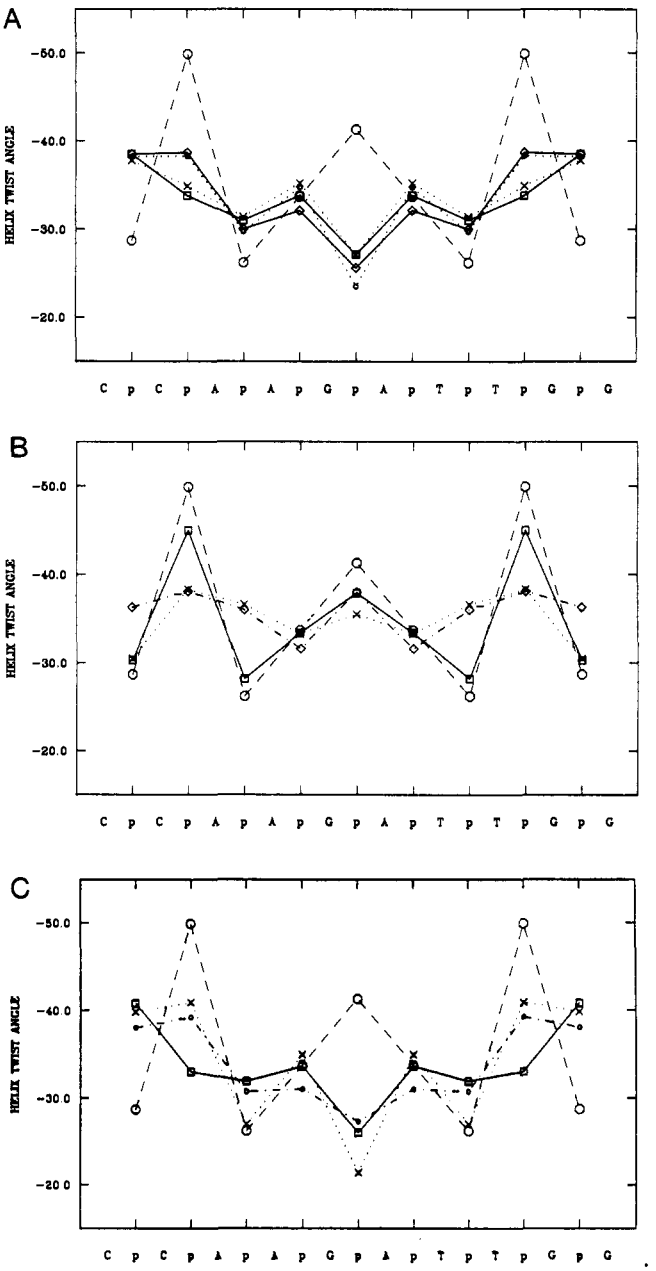

Figure 3. (A) Plot showing the helix twist as a function of sequence for the structures CTL (○), MAW_R (□), $MAW_{R/U}$ (×), MBW_R (◇) and $MBW_{R/U}$ (○). Note that the helix twist has been averaged for the two strands of the solution structures and is thus symmetric. (B) Plot showing the helix twist as a function of sequence for the structures CTL (○), $CTLW_U$ (□), $MDLW_U$ (×) and $MDLG_U$ (◇). (C) Plot showing the helix twist as a function of sequence for the structures CTL (○), MAG_R (□), MBG_R (×) and $CTLG_R$ (○).

sampled in water ($CTLW_U$ and $MDLW_U$) or 80 ps in the gas phase ($MDLG_U$)]. Simulations have indicated that NMR distance-restrained MD is capable of reproducing these sequence-specific helical twist and propeller twist (see below) variations;[46] Distance geometry gives comparable reproduction.[44]

The results obtained from the free dynamics also appear to be dependent upon the starting structure. Looking at the helix twist plots (Figure 3) unrestrained dynamics begins to move the crystal structure toward the solution conformation ($CTLW_U$). Similarly starting from a uniform 36° of helix twist throughout the model-built structures, we begin to see variations in helix twist that generally follow the pattern of the crystal structure (Figure 3B). Free dynamics in both the gas phase ($MDLG_U$) and in water ($MDLW_U$) produces very similar results. One exception is the difference between the helix twist values found for the CpA base step in the restrained MD calculations starting from the A-DNA

(62) Fratini, A. V.; Kopka, M. L.; Drew, H. R.; Dickerson, R. E. *J. Biol. Chem.* **1982**, *257*, 14686–14707.

Table II. Atomic Cartesian Coordinate Rms Deviations[a] between Full Structures (All Atoms)

struct[b]	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A		2.8	0.9	1.2	1.7	2.6	2.8	1.9	3.7	3.8	1.9	4.2	3.4	3.3
B			3.0	3.1	2.5	1.1	1.3	1.8	2.3	2.3	1.8	2.5	1.7	1.5
C				0.7	1.6	2.7	2.8	1.9	3.9	4.0	1.9	4.3	3.5	3.4
D					1.7	2.8	2.9	1.9	4.0	4.0	2.0	4.7	3.6	3.4
E						2.4	2.6	1.7	2.9	3.2	1.4	3.5	2.9	2.7
F							0.7	1.5	2.7	2.8	1.7	3.0	2.3	2.0
G								1.5	3.0	3.0	1.9	3.2	2.4	2.2
H									3.0	3.2	1.3	3.5	2.7	2.5
I										1.6	2.5	2.0	2.0	1.9
J											2.6	1.0	1.2	1.7
K												3.0	2.2	2.0
L													1.9	2.0
M														1.2

[a] In angstroms. [b] Where the letters A–N represent the minimized full structures defined below: A = MA, B = MB, C = MAW_R, D = $MAW_{R/U}$, F = MBW_R, G = $MBW_{R/U}$, H = MGB_R, I = $CTLW_R$, J = $CTLW_U$, K = $CTLG_R$, L = CTL, M = $MDLW_U$, and N = $MDLW_R$.

vs the B-DNA initial models (Figure 3 parts A and C). The smaller helical twist value for the MA structures at this base step is found for both the gas phase and solution phase restrained structures. This presumably reflects equally probable structures based upon the limited distance restraints in this portion of the molecule.

Most importantly, however, note the relatively large helical twist in the central GA step for both the crystal and the unrestrained structures, which is *opposite* to the underwound GA base step observed in the NMR-refined solution conformation (Figure 3A). It is satisfying to see that the unrestrained gas phase dynamics structure ($MDLG_U$) beginning from the model-built B-DNA coordinates produces the same helix twist trend as that produced by performing the calculation in aqueous environment ($MDLW_U$). Clearly however, unrestrained dynamics in the gas or water phase fails to adequately locate either the NMR solution or X-ray crystal structures.

Rms (Å) comparison of the structures using the full set of all atom Cartesian coordinates indicates in general a rather poor relation of one structure to another (with the exception of the gas phase derived structures, Table II). However, when the CG base pair at one terminus of MAW and MBW is not considered, the rms deviation between the two structures is reduced to 2.2 Å. Although no true fraying at the ends is observed, end effects such as increased flexibility which are present for gas-phase structures also appear to exist in the water-calculated structures. Such flexibility is further supported by the weak intensity observed for the terminal imino protons of the H_2O NMR spectrum (E. Nikonowicz and D. Gorenstein, unpublished results).

Importantly, the structures which were subjected to 100 ps of restrained MD and then a further 25 ps of free dynamics (MBW_R vs $MBW_{R/U}$ or MAW_R vs $MAW_{R/U}$) show little variation between the restrained and unrestrained conformations, Figure 3A and Figure 4B (rms difference in total Cartesian coordinates 0.7 Å, Table II). The difference that does exist can be attributed to the relief of bond length and bond angle strain (Table I) once the restraining forces are removed. It should be recalled that the data observed by NMR is dynamic in nature, and as a consequence, some of the NOESY-derived interproton restraints cannot simultaneously be obeyed within the defined limits. Note that the % error region of the flatwell restraint term (the region in which a restraint distance for a given proton pair may vary without an energy penalty being imposed) was set to ±5% of the restraint distance and was bounded by force constant barriers on either side of 40 kcal/mol Å^2.

A comparison of the rms deviation between the total coordinates for some of the structures is shown in Table II. Clearly the crystal (CTL) structure differs most (generally 3–3.5 Å) from any of the restrained or unrestrained MD structures, whether starting from the model-built or initial MORASS-derived structures. However, starting from the crystal structure itself, unrestrained MD in water leads to a structure that differs by only 1.0 Å from this initial structure. Similarly, all of the unrestrained MD structures generally differ the least from their initial starting structure. Thus, either 100 ps is too short of a simulation to locate either the "correct" crystal or solution structure or the free MD methodology itself is incapable of locating the structure(s). Note also that gas or water phase refined structures are quite comparable to each other although the differences are initial starting structure dependent.

Bending of the Solution Derived Structures. The most prominent features distinguishing the solution and crystal structures are the bending and underwinding observed in the solution structures relative to the crystalline state (Figure 4). The bending appears to be centered about the site of the G_{anti}·A_{anti} mismatch, in contrast to the rodlike structures observed in the crystal structure. The considerable bend observed in the NMR distance-restrained solution structures is seen in both water and gas phase calculated structures starting from any model (A-DNA, B-DNA, and Crystal). This bending is not observed after 100 ps of unrestrained MD starting from the crystal structure ($CTLW_U$) or from the model-built DNA structure ($MDLW_U$); however, these molecules are slightly underwound compared to their parent structures. This latter point is reflected by the differences in the widths of the major and minor grooves. When restraints are applied to either the crystal structure coordinates or the model-built B-DNA coordinates in the water-phase simulation, a small degree of bending is observed (Figures 4 and 5). When the aqueous simulation is carried out for an additional 150 ps for a total of 250 ps (results not shown), the change in the degree of bending observed is still rather small relative to the gas phase derived structures. Levitt[63] had previously observed a 26-ps periodicity for the bending motion of a DNA dodecamer during in vacuo simulations. When considering the effects of viscosity upon bending, Levitt postulated that the bending motion observed in water would be significantly damped. Thus, it is quite reasonable that the extent of bending observed after ~30 ps of gas-phase MD for the decamer structures would not be expected to be reached until >250 ps in the aqueous environment.

The structure resulting from 100 ps of restrained water dynamics applied to the crystal coordinates is anomalous in that one of the mismatched G·A basepairs is not base paired. This is not uncommon in unrestrained MD of duplex oligonucleotides.[2,3] Instead, the bases are stacked upon one another (Figure 5B). Accommodation of the G·A mismatch in this fashion allows the structure to remain relatively straight; however, a significant energy penalty is imposed both in terms of electrostatic interactions and restraint energy (Table I). Note that this anomalous feature is probably responsible for the unusually high $\%rms_{vol}$ values shown in Table I for this structure. Although the imino protons of the central G·A mispairs are observed in the H_2O spectrum (indicating H-bonding), hydrogen bond restraints were not imposed upon the central four residues for the MD simulations. This was done in order to allow adequate flexibility in this region so as not to predispose these residues to a specific conformation.

It had previously been noted in the crystal structure[55] that the guanine residue of the G·A mismatch was hydrogen bonded to

(63) Levitt, M. *Cold Spring Harbor Symp. Quant. Biol.* **1983**, *47*, 251–262.

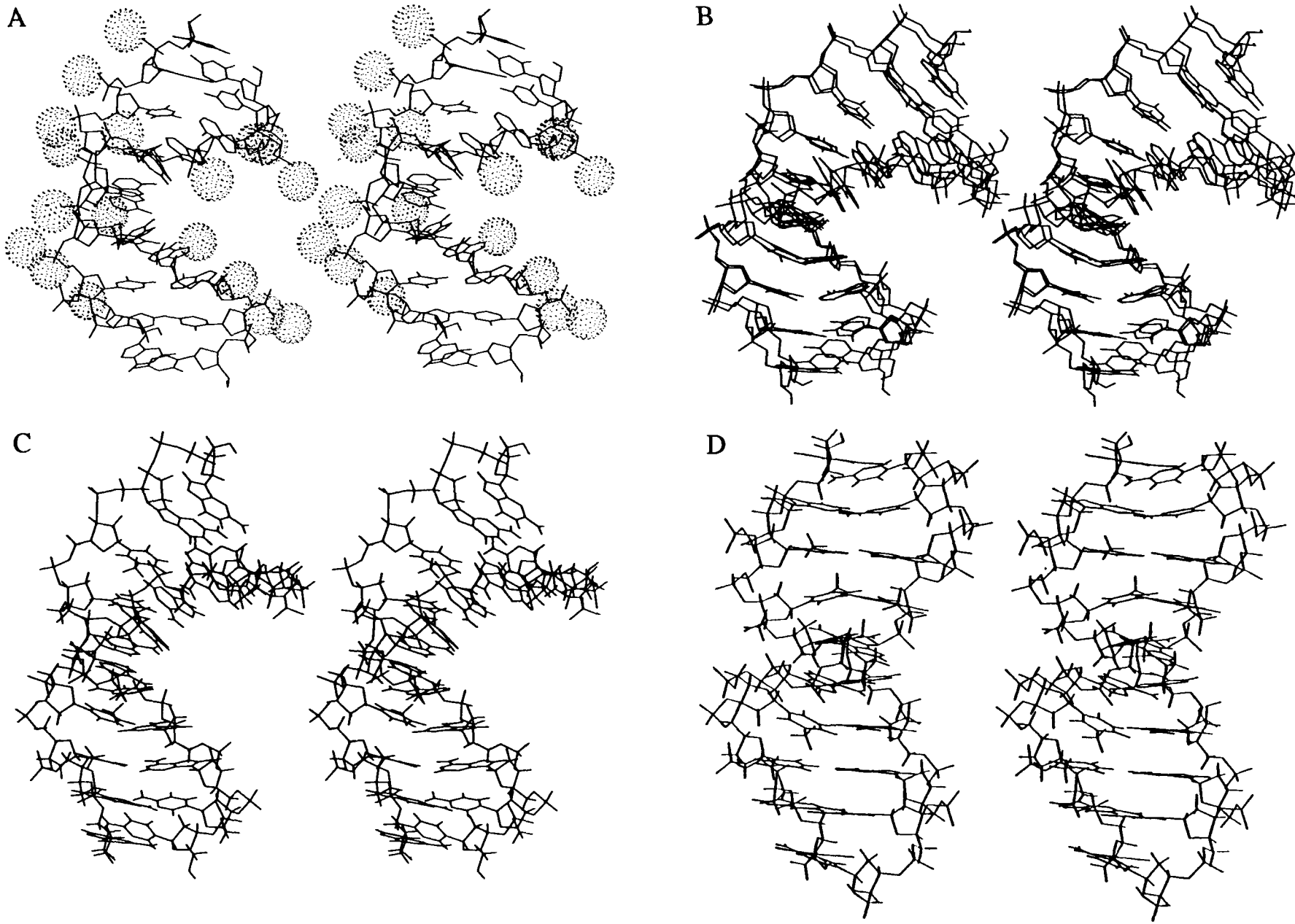

Figure 4. (A) Stereoview of the final water phase calculated structure starting from model-built B-DNA (MBW_R), using NMR derived restraints; Na^+ ions (stipled balls) are also shown. (B) Stereoview overlay of the structures MAW_R and $MAW_{R/U}$. (C) Stereoview of the final structure derived in the gas phase using NOESY restraints and starting from A-DNA (MAG_R). (D) Stereoview of the structure produced after 100 ps of unrestrained MD starting from the crystal structure ($CTLW_U$).

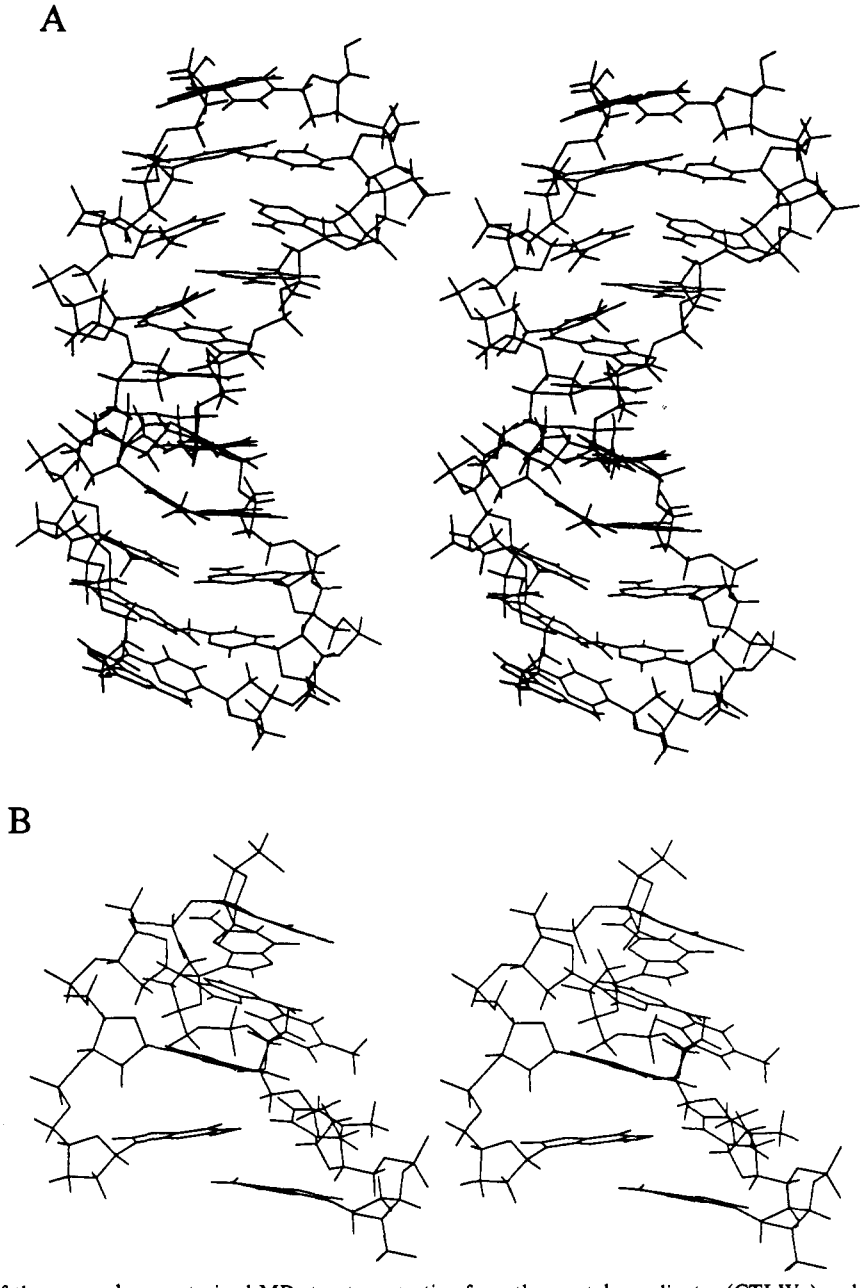

Figure 5. (A) Stereoview of the water phase restrained MD structure starting from the crystal coordinates ($CTLW_R$) and (B) an expansion around the mismatch site showing the unpaired GA mismatch.

both the adenine on the opposite strand and the thymine adjacent to the mismatched adenine residue. The NOESY distance-restrained structures derived in the gas phase and in the water phase have reproduced this same bifurcated hydrogen-bonding scheme. Further, this pattern is maintained without any apparent gross disturbances to the hydrogen-bonding scheme of the flanking A·T base pairs (Figures 6 and 7). However, the bifurcated hydrogen bonding of the mismatched A6 residue is introduced at the cost of rather large propeller twists, ~25°, in the crystalline and $CTLW_U$ structures and ~36° in the gas-phase and water-phase NMR-derived structures. Figure 8 shows a plot of the variation of the average (strand 1 and strand 2) propeller twist as a function of sequence for the structures CTL, MAW_R, $MBW_{R/U}$, and $CTLG_R$. The consequences resulting from different modes of accommodation of the unusually twisted basepair in the crystalline state and in the phase-calculated structures are observed through the overall bending, or straightness, of the duplex structures. In the crystalline state, the adjacent thymidine residues on each strand appear to unstack and form a wedge. This wedging, in essence, produces three helical axes, two defined by the three residues terminal on each end and one in the middle defined by the four central residues. The segregated nature of the central four residues was first described by Dickerson and co-workers.[55] The thymidine residues are situated in the sequence such that they are approximately 165° out of phase (as measured by helical twist) relative to each other. Thus, a reverse in the direction of the bend in the helical axis negates the comparable bend at the other thymidine–thymidine wedge. As a result, a straight structure is produced in the crystal. This gross cylindrical structure then allows the duplexes to stack end-to-end, as is usually observed for DNA crystals. Stacking of base pairs appears to contribute 9–13 kcal/mol to the stabilization of the duplex (through analysis of the duplex interaction energies in the MDANAL module of AMBER). Because of the wedging between the TpT bases, the stacking energy at this base step is considerably less than this (ca. 6 kcal/mol). Intermolecular end-to-end stacking of the helices in the crystal may therefore contribute to the overall straightness of the duplex in the solid state.

In the aqueous environment, however, crystal packing forces that cause the duplex to accommodate the unusually large propeller twisting of the G·A mismatch in a different fashion are not present. The gas- and water-phase MORASS/MD-refined structures

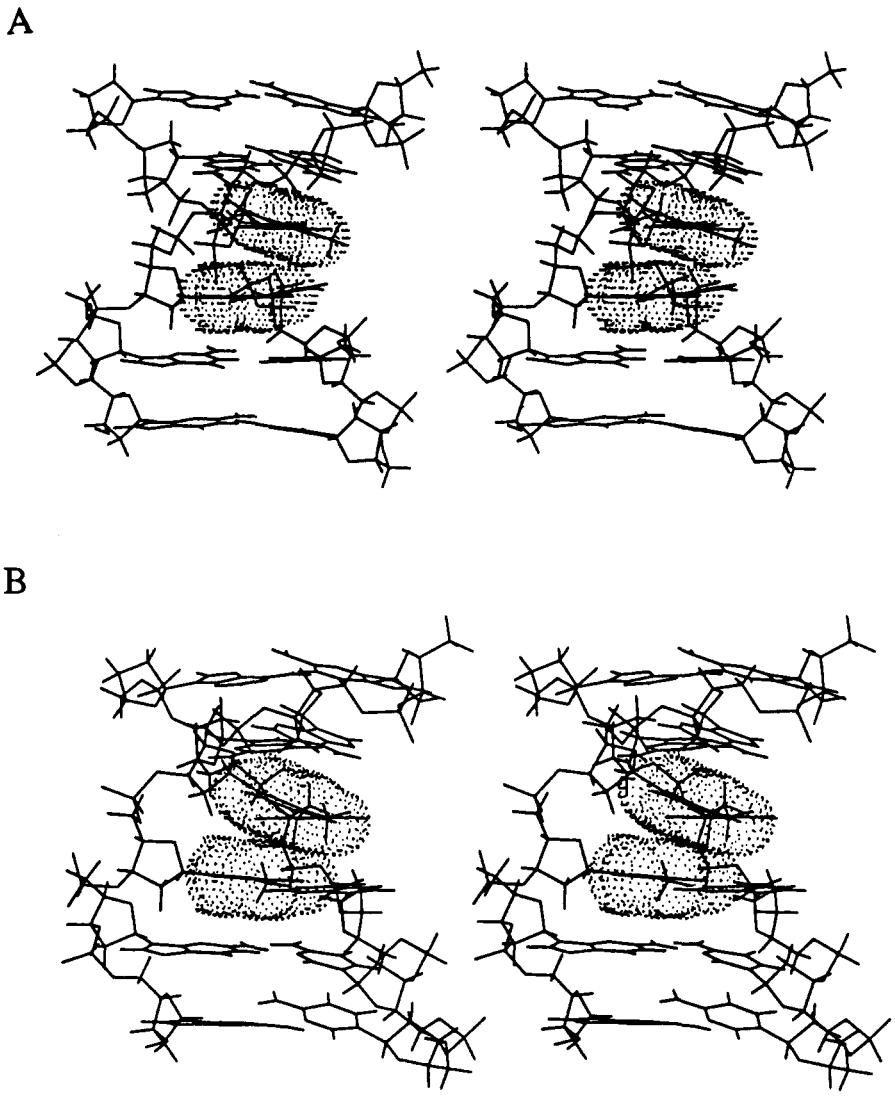

Figure 6. Stereoviews of residues 5–10 for the crystalline structure (A) and the final water dynamics structure $CTLW_U$ (B). T7 and T8 are depicted with their van der Waals surface.

are intramolecularly stacked from base-to-base; there is no wedging observed at the T^7pT^8 base steps at either end. Examination of the local helix axes at each base step in the water- and gas-phase structures shows little trend except that the local helix axis does not recover a collinearity with the overall helix axis as is the case with the crystal structure. It should be emphasized that the bend observed in *all* of the bent decamer structures from restrained dynamics is in the same direction, acting to "close around" the major groove. It is well-known that stacking interactions provide much of the stability for duplex DNA. Although the stacking energy for a TT base step is not as great as those for other base steps, the lack of two such interactions in a tandemly mismatched decamer sequence would result in significant destabilization. Thus, in a solution or solutionlike (i.e., in vacuo with a distance-dependent dielectric) environment, a duplex in which all bases are stacked and the structure is bent would be favored relative to a structure which was straight, but partially unstacked (Figures 6 and 7). Note that the van der Waals energies of the solution structures are lower than that of the crystal structure in agreement with a more favorable stacking interaction.

Further support for the above observations comes from the experimental plots of the NOESY data and the theoretical NOESY spectra calculated from the crystal structure coordinates (see ref 57). Given the wedging effect of the adjacent thymidine residues in the crystal structure, one would predict weak NOE interactions between these two residues. Comparison of the theoretical and experimental NOESY plots of the base–H1′/H3′ and base–H2′,H2″ regions does indeed show a discrepency in the intensities of the interresidue crosspeaks between T7 and T8. Comparison with the theoretically calculated spectra from the gas-phase structures does *not* display this same difference as observed using the crystal coordinates. Thus, the experimental NOESY data is consistent with a "normal" base-stacking scheme and does not support a wedged conformation.

Conclusions

The data presented here show that both crystallography and NMR provide important, but different, structural information about nucleic acids. Duplex DNA in crystals is subject to forces which are in fact appreciable and which significantly influence the structure adopted by the molecule even though the crystals are hydrated. In the present study, both NMR spectroscopy and X-ray crystallography have described the rather large propeller twist at the mismatch site *and* the bifurcated hydrogen-bonding pattern of the mismatched guanine residue. However, it would appear to be necessary for the crystal structure to maintain a straight rod conformation in order to allow adjacent molecules in each unit cell to stack and form the crystal. Through wedging of the adjacent thymidine residues, and thus providing opposing kinks at either end, the duplex is able to recover its global collinear cylindrical shape. In the aqueous environment, however, no forces are present which would cause such kinking and destacking of the thymidine–thymidine residues. Thus, an overall curvature or bend in the duplex is observed. A similar phenomenon is likely to be the explanation for the discrepencies between the NMR and crystal structures of an adenosine bulge tridecamer.[25] In order

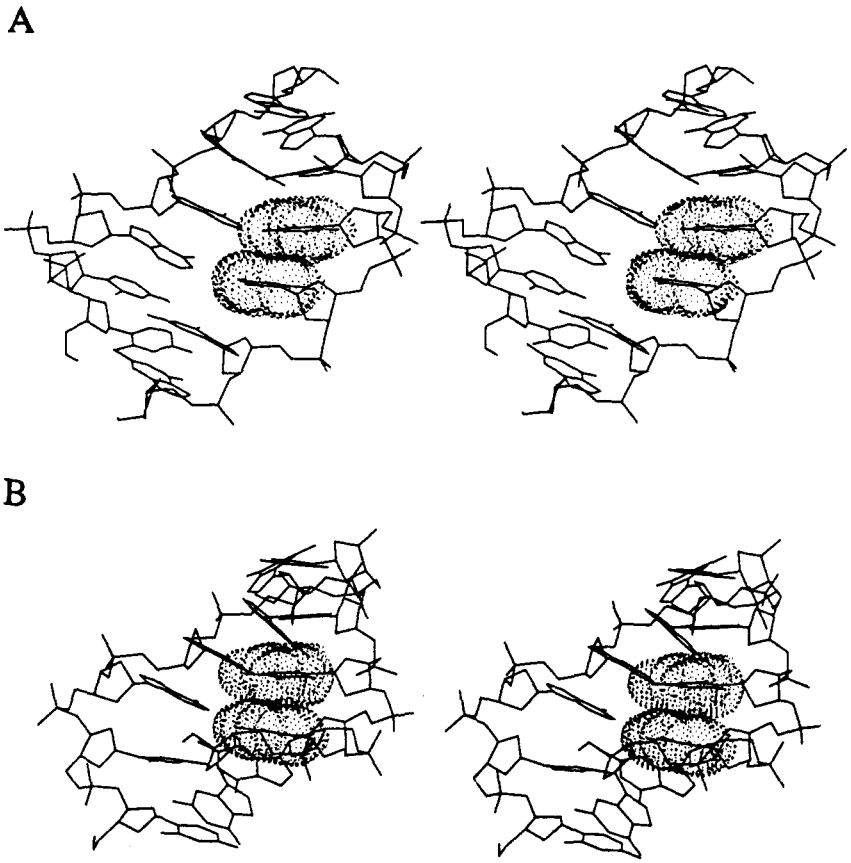

Figure 7. Stereoviews of residues 5–10 for the restrained structures MAW_R (A) and MBW_R (B). T7 and T8 are again depicted with their van der Waals surface.

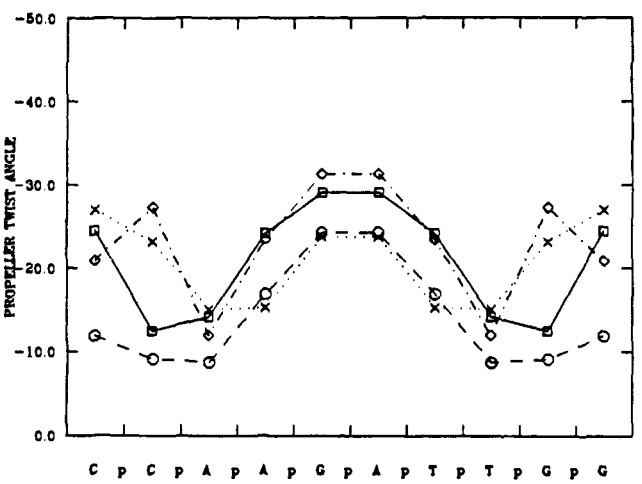

Figure 8. Plot of average propeller twist vs sequence for representative structures: CTL (○), MAW_R (□), $MBW_{R/U}$ (◇) and $CTLG_R$ (×).

for the tridecamer to stack end-to-end in the crystal, it must be straight. Wedging in of the bulge base forces the structure to kink, and consequently the duplex adopts in solution a bent geometry which is not favored for stacking and crystallization.

Various computational methods are currently used to simultaneously search for structures which are of minimal energy and structures which satisfy a set of NOESY-derived interproton distances. One such method, molecular dynamics, is capable of simulating an aqueous environment while performing the calculations in vacuo. The advantage to these gas-phase simulations is the speed with which they may be carried out. However, the distance-dependent dielectric can only approximate a general aqueous environment; the direct influence of counterions and water molecules cannot be accounted for. Thus, it may be argued that it is preferrable to carry out structural investigations in a bath of water where the solute is surrounded by counterions. The structures derived using restrained MD in water and starting from either model-built coordinates or crystal coordinates do obey the set of input restraints; unfortunately, the time frame in which these calculations were performed, appears to be too short to observe more global changes in the DNA. Global changes such as bending or curving in an aqueous medium would occur minimally on a time scale 3–4 times longer than that used here due to the viscosity of the aqueous medium. The differences observed between the gas-phase "preconditioned" structures and those derived after 100 and 125 ps of MD in water being considered, it appears that the bend observed in the gas-phase calculation is *not* restricted from forming in the aqueous phase. Thus, while an explicit aqueous medium might be preferred over a gas-phase environment, the in vacuo simulation appears to perform comparably to the fully solvated system. The use of structures generated with minimal gas-phase simulation as input for further solution-phase refinement is quite reasonable given the results observed here.

Finally it should be noted that unrestrained molecular dynamics (in a 100-ps trajectory) is unable to adequately locate either the crystal or solution NMR conformations. The structures determined from 100 ps of free dynamics depend upon the initial starting structure. Perhaps with considerably longer simulations (>1 ns) it may be possible to locate the "correct" structure. Alternatively, these results may suggest that currently, free molecular dynamics calculations even in an aqueous medium simulation are inherently limited in defining DNA structures.

Acknowledgment. This work was supported by NIH (AI27744), the Purdue University Biochemical Magnetic Resonance Laboratory which is supported by the NSF Biological Facilities Center on Biomolecular NMR, Structure and Design at Purdue (Grants BBS 8614177 and DIR-9000360 from the Division of Biological Instrumentation), and the NIH Designated AIDS Center at Purdue (AI727713).