Bending and Adaptability to Proteins of the cAMP DNA-Responsive Element: Molecular Dynamics Contrasted with NMR

Sylvie Derreumaux and Serge Fermandjian

Département de Biologie et Pharmacologie Structurales, UMR 8532 Centre National de la Recherche Scientifique, Institut Gustave Roussy, 94800 Villejuif, France

ABSTRACT　DNA bending is assumed to play a crucial role during recognition of the cAMP-responsive element (CRE) by transcription factors. However, diverging results have been obtained for the bending direction of the unbound double helix. The refined NMR structures present a bend directed toward the minor groove, while biochemical methods conclude that there is a bend toward the major groove. The present 10-ns molecular dynamics (MD) simulation of d(GAGATGACGTCATCTC)$_2$, which contains the octamer CRE in its center, was carried out with AMBER in explicit water and counterions. It shows that CRE is a flexible segment, although it is bent slightly toward the major groove (7°-8°) on the average. The MD structure agrees with both the biochemical results and unrefined NMR data. The divergence with the NMR refined structures suggests an improper electrostatic parameterization in the refinement software. The malleability of the central CpG is certainly the major contribution to the curving of the whole CRE segment in both the unbound and bound states. Comparison with the crystal structure of CRE bound to GCN4 shows that the deformation induced by the protein is concentrated mainly on the CpG step, rendering the bound structure of CRE closer to the structure of the 12–0 tetradecanoylphorbol-β-acetate-responsive element.

INTRODUCTION

Recognition of DNA by bZIP (basic leucine zipper) dimers is a delicate process, involving subtle structural and dynamical variations that are generally related to dramatic biological effects. Some bZIP dimers recognize the cAMP-responsive element (CRE) in promoters of ubiquitary genes and the 12–O tetradecanoylphorbol-β-acetate (TPA)-responsive element (TRE) in protooncogene promoters. Proteins of the ATF-CREB subfamily bind to CRE and very weakly to TRE, whereas the yeast transcription factor GCN4 (general control protein 4) and the proteins belonging to the FOS-JUN subfamily bind to both CRE and TRE (Ryseck and Bravo, 1991; Hai and Curran, 1991; Hurst, 1994; Kim and Struhl, 1995). Such cross-bindings could be explained by the resemblances between CRE, d(TGACGTCA)$_2$, and TRE, d(TGACTCA)$_2$ (same TGA:TCA half-sites) and among the bZIP proteins (similar DNA binding sequences). Complex formation will further depend on the relative concentrations of the protein components, which can vary with the cell physiology. Some of them will be response productive and others nonproductive or even detrimental to cell life.

The specific recognition of DNA by proteins is highly dependent on the local and sublocal deformations occurring in the helix. According to Paolella et al. (1994), bending of DNA is the major structural factor influencing protein-CRE/TRE recognition. Actually, results of biochemical methods indicate a straight structure for TRE and a structure bent slightly (~10°) toward the major groove for CRE (Paolella et al., 1994, 1997; Sitlani and Crothers, 1998; Hockings et al., 1998; Sloan and Schepartz, 1998). Bending of CRE shortens both the distance and the difference in orientation between the half-sites TGA:TCA, seemingly reducing the structural difference between CRE and TRE and allowing the recognition of both CRE and TRE by the same bZIP dimers (Paolella et al., 1994; Keller et al., 1995). The DNA bending accompanying the complex formation varies as a function of the bZIP dimer. It could be responsible for a change in the dimer orientation in the transcription machinery, influencing the promotion of the gene transcription (Paolella et al., 1994; Leonard and Kerppola, 1998).

Our previous NMR refined molecular modeling argued in favor of a straight structure for TRE included in a 15-mer DNA and a ~30° bend toward the minor groove for CRE at the center of a 16mer DNA (Chaoui et al., 1999). A structural analysis of the NMR atomic coordinates of Conte et al. (1997) reported in the Protein Data Bank provided the same bend toward the minor groove.

The results from biochemical methods (Paolella et al., 1994, 1997; Sitlani and Crothers, 1998; Hockings et al., 1998; Sloan and Schepartz, 1998) and NMR restrained modeling, while they converge for TRE, are thus diverging for CRE. This motivates the present modeling study of CRE, in which no NMR constraints are applied. We believe this study to be more "realistic," as it resorts to a molecular dynamics (MD) simulation and includes neutralizing Na$^+$ and water explicitly. To our knowledge, it is the first MD study of the CRE DNA, allowing us to investigate its flexibility and its conformational substates. For the sake of comparison, the selected 16-mer DNA d(GAGATGACGT-CATCTC)$_2$ was kept the same as the one previously ana-

lyzed by NMR (Chaoui et al., 1999). A MD trajectory of 10 ns was carried out, as a previous study revealed that simulations longer than 5 ns were necessary to reach a full convergence of the trajectory for a system of that size (Young et al., 1997b). Calculations were performed with the AMBER software and the particle mesh Ewald implementation for the treatment of electrostatics. Previous studies have revealed a good agreement with experimental data, particularly regarding the sequence-specific structure and bending of DNA (Cheatham and Kollman, 1996, 1997; Young et al., 1997b; Young and Beveridge, 1998; de Souza and Ornstein, 1998; Sprous et al., 1999). The recently modified version of the force field of Cornell et al., parm98, was used (Cheatham et al., 1999).

The results indicate that the CRE DNA, although flexible, is on the average slightly curved toward the major groove (7°-8°). The inherent flexibility of CRE, mostly concentrated at the central CpG step, must allow it to easily adjust its curvature in the different DNA-protein complexes. In the cocrystal with GCN4 (Keller et al., 1995), the largest helical deformation involves the central CpG, which exhibits a correlated increase in the roll and a decrease in the twist, bringing the TGA:TCA half-sites closer to each other (distance and orientation) and attenuating the structural differences between CRE and TRE. MD results fit our NMR experimental data but disagree with the refined NMR structures (Conte et al., 1997; Chaoui et al., 1999), highlighting the impact of the electrostatic parameterization on the refinement of the DNA structure and, more particularly, of the axis curvature.

MATERIALS AND METHODS

Choice of the DNA sequence

The MD simulation was performed on the 16-mer duplex d(GAGAT-GACGTCATCTC)$_2$, which contains the CRE consensus sequence d(T-GACGTCA)$_2$ of interest at its center. Three factors led to the choice of this oligonucleotide: it has the minimum size needed to achieve stable interactions with the bZIP transcription factors (Talanian et al., 1992); the crystal structure of its complex with GCN4 has been determined (Keller et al., 1995); and its NMR refined structure has been reported recently (Chaoui et al., 1999). We used the following numbering for the DNA sequence: 5′-(G$_1$A$_2$G$_3$A$_4$T$_5$G$_6$A$_7$C$_8$G$_9$T$_{10}$C$_{11}$A$_{12}$T$_{13}$C$_{14}$T$_{15}$C$_{16}$: G$_{17}$A$_{18}$G$_{19}$A$_{20}$T$_{21}$-G$_{22}$A$_{23}$C$_{24}$G$_{25}$T$_{26}$C$_{27}$A$_{28}$T$_{29}$C$_{30}$T$_{31}$C$_{32}$)-3′.

Molecular dynamics: software and protocol

The MD simulation was carried out with the suite of programs AMBER 4.1 (Pearlman et al., 1995) and the modified version of the Cornell et al. force field (parm95), parm98 (Cornell et al., 1995; Cheatham et al., 1999). This version provides DNA structures that are in better agreement with experimental data. Water was simulated with the TIP3P model (Jorgensen et al., 1983). Periodic boundary conditions were simulated with a rectangular unit cell. The initial box was truncated to achieve a minimum distance of ~11 Å beyond all DNA atoms in all directions, resulting in a box size of 49 Å × 49 Å × 87 Å and 4752 water molecules. Thirty neutralizing Na cations were placed with a simple energy minimization algorithm (Leap module of

AMBER 5.0). The starting DNA conformation was the AMBER-generated canonical B-DNA double helix. The energy of the system was minimized for 2250 cycles, using a combination of steepest descent and conjugate gradient algorithms. During the energy minimization, the oligomer conformation and the counterion positions were maintained with harmonic restraints (100 kcal/mol Å^2), with a progressive diminution of the force constants for a final 500 cycles without restraints. The system was then heated to 303 K over 10 ps, the velocities were rescaled up to 150 K, and then the system was coupled to a heat bath with the weak coupling Berendsen algorithm (Berendsen et al., 1984). During heating, harmonic constraints (25 kcal/mol Å^2) were imposed on the atomic positions of the oligomer and of the sodium counterions. After an constant-mass, constant-volume, and constant-temperature (NVT) equilibration period of 5 ps at 303 K, the simulation was performed under constant-mass, constant-pressure, and constant-temperature (NPT) conditions (303 K, 1 atm.), using coupling constants of 0.2 ps. The restraints were relaxed over a period of 25 ps until a free system was achieved, and a final equilibration was carried out for 5 ps before the beginning of the 10-ns production phase. Then a Maxwell distribution of the velocities was assigned to the system, followed by 10 ps of free equilibration. This procedure was repeated before the beginning of the production phase. After 50 ps and 100 ps of production, a Maxwell distribution of the velocities was again assigned. The velocities were redistributed to attenuate the influence on the dynamics of the initial configuration of the system. All bond lengths involving hydrogen atoms were restrained using the SHAKE algorithm (tolerance = 0.0005) (Ryckaert et al., 1977), and a time step of 2 fs was used. Electrostatic interactions were calculated using the particle mesh Ewald summation method (Darden et al., 1983; Essmann et al., 1995), with a 10^{-6} Ewald convergence tolerance and a charge grid spacing of ~1 Å for the interpolation in the reciprocal space. A 9-Å residue-based cutoff was applied to the van der Waals and direct space Ewald interactions. Using a modified version of the Sander module of AMBER 4.1 (Young and Beveridge, 1998), we removed the translations and rotations of the DNA every 100 steps of MD, scaling the resulting solute atom velocities to conserve kinetic energy. The calculations were performed on a Silicon Graphics Origin200 workstation.

Structural analysis

CURVES version 5.2 was used to analyze the DNA structures stored every picosecond of MD (Lavery and Sklenar, 1988). CURVES fits a curved or a straight global helix axis that follows the double helix. It allows both global and local geometrical analysis, providing the values of the helical parameters and of the backbone dihedral angles. This work refers to the local parameters output by CURVES. To avoid end effects, only the 14 central base pairs were taken into account for the analysis. For the MD simulation (10 ns), each parameter was measured 10,000 times.

MD results were compared to the experimental NMR and x-ray data on unbound and bound CRE, respectively. For the 16-mer duplex d(GAGAT-GACGTCATCTC)$_2$ of interest, one NMR refined structure (Chaoui et al., 1999) and one crystal structure of its complex with the GCN4 transcription factor (2.2 Å) (Keller et al., 1995) (PDB ID: 2dgc) are available. Another NMR study concerns a different oligomer with CRE at its center, d(CAT-GTGACGTCACATG)$_2$ (Conte et al., 1997) (PDB ID: 1saa). Yet the loose NMR constraints used by these authors relative to the helical parameters (internucleotide distances) and phosphate positions (ϵ dihedral) result in a poor experimental assessment of these DNA morphological parameters. Thus comparisons were made with the NMR experimental structure of Chaoui et al. (1999), CRE_NMR; with the crystal structure (Keller et al., 1995), CRE_GCN4; and with the NMR data of Conte et al. (1997) for the sugar phases. Concerning the DNA bending, data of reference were provided by biochemical, NMR, and crystallography experiments. We calculated the local parameters and global bending describing the experimental structures, using the published PDB coordinates and CURVES 5.2. The canonical B-DNA and A-DNA conformations cited in the text are the ones generated by the AMBER software (Arnott et al., 1972).

RESULTS AND DISCUSSION

Molecular dynamics

A MD trajectory of 10 ns was simulated with the 16-mer duplex d(GAGATGACGTCATCTC)$_2$, using the AMBER software and the recent parm98 force field (Cheatham et al., 1999). According to Young et al. (1997b), simulations longer than 5 ns are necessary to reach a full convergence of the trajectory for a system with DNA of the above-mentioned size and containing explicit monovalent cations. The atom coordinates were saved and analyzed every picosecond. The results take into account the trajectory after 500 ps from its starting point (i.e., data collected through 9.5 ns).

In Fig. 1 a are plotted the root mean square deviations (rmsds) as a function of time, calculated for the coordinates of heavy atoms in the 14 central residues of the DNA, between the instantaneous structures extracted from the molecular dynamics trajectory (MD), CRE_MD, and the canonical B DNA conformation and A DNA conformation. One can see that after 500 ps, the DNA has globally relaxed from the starting conformation and follows its own dynamics.

The DNA conformation calculated with the parm98 force field looks like more the canonical B form than the canonical A form, with rmsd values of nearly 2.5 Å and 5.5 Å, respectively. In comparison, calculations performed with the parm95 force field (unpublished results) provide a DNA

conformation intermediate between the canonical B and A DNAs, with rmsd values of nearly 4.5 Å.

In the Appendix are reported the averaged values and the standard deviations over the MD trajectory (0.5–10 ns) of the interpair, pair-axis, base-base, and backbone structural parameters of the CRE oligomer. The DNA is globally stabilized in a B form, with averaged twist, rise, inclination, x displacement, and sugar phase values of 33.7°, 3.4 Å, 0.0°, −1.0 Å, and 147°, respectively, confirming that the new parm98 force field leads to DNA conformations that are more consistent with the experimental data compared to the parm95 force field (Young et al., 1997b; de Souza et al., 1998; Cheatham et al., 1997, 1999; Cieplak et al., 1997).

Backbone conformation

Sugar puckers. The average sugar phase angle in the CRE_MD structure is 147°. This value fits the B-DNA canonical value (162°) (Saenger, 1984) better than the value provided by the elder parm95 force field does (∼120°-125°) (Cheatham et al., 1999). The phase angle as a function of the nucleotide (Fig. 2) follows the well-known sequence dependence, with values higher at purines (∼160°) than at pyrimidines (∼135°) (Poncin et al., 1992). The transitions toward the Northern state (C3′-endo, ∼18°) are not numerous, and when they occur, they merely last more than 200 ps. Exceptions are the sugars of A12 (in the half-site TCA) and of C14 (outside the consensus sequence), which display transitions of nearly 500 ps between 1200 ps and 1700 ps and of 700 ps between 9000 ps and 9700 ps,
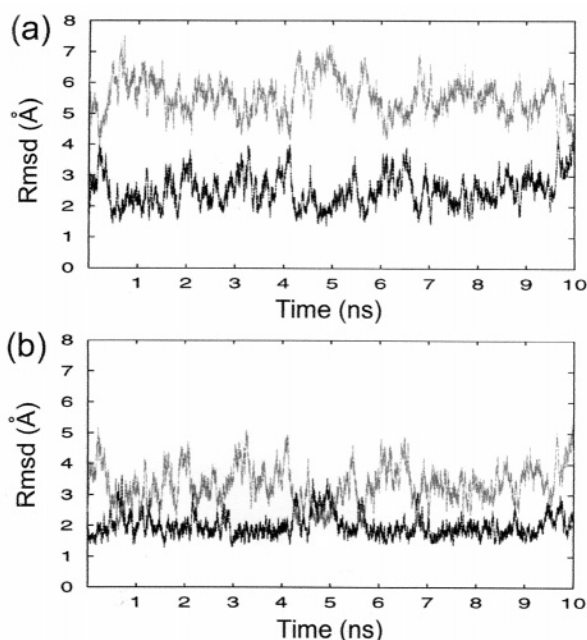

FIGURE 1 (a) Root mean square deviation (rmsd) between the MD structure of d(AGATGACGTCATCT)$_2$ (terminal residues are excluded from the calculation) as a function of the time and the canonical B DNA (*black line*) and A DNA (*gray line*) conformations. (b) Rmsd between the MD structure and the experimental x-rays corresponding to CRE bound to GCN4 (*black line*) (Keller et al., 1995) and NMR refined structure corresponding to unbound CRE (*gray line*) (Chaoui et al., 1999).
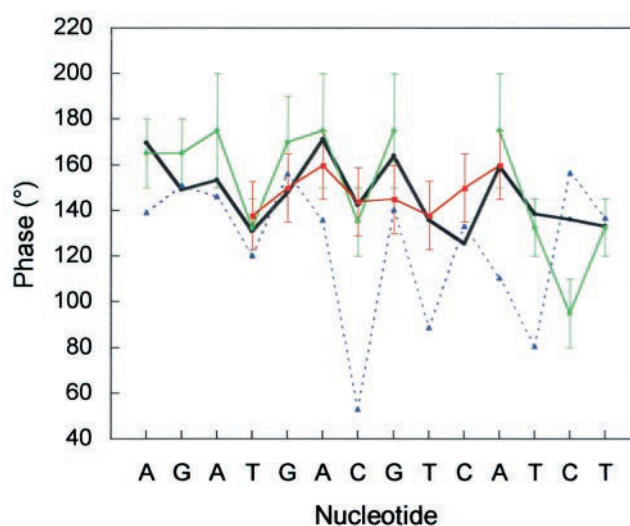

FIGURE 2 Sugar phase angles of (AGATGACGTCATCT)$_2$ as a function of the nucleotide: MD values averaged over 9.5 ns (*black line*); NMR data with error bars as measured by Chaoui et al. (1999) (*green line*) and by Conte et al. (1997) for the central CRE consensus sequence (*red line*); and values found in the cocrystal with GCN4 (Keller et al., 1995) (*dashed blue line*).

respectively. These two nucleotides are the only ones surrounded by pyrimidines.

Phosphate conformations. Numerous coupled variations of the dihedral angles ϵ and ζ associated with BI/BII transitions in the phosphate groups are observed during the dynamics, all over the DNA sequence (BI: ϵ in *trans* and ζ in *gauché*; BII: ϵ in *gauché* and ζ in *trans*) (Grzeskowiak et al., 1991; El antri et al., 1993; Gorenstein, 1994). The angle values are ~180° (280°) for ϵ and ~270° (150°) for ζ in the BI (BII) state. The difference $\epsilon - \zeta$ is nearly −90° for BI and 130° for BII and is convenient for following the transitions. The value of $\epsilon - \zeta$ averaged over time is proportional to the BII population percentage (Gorenstein, 1994). The BII population percentage averaged over the CRE_MD structure (21%) is in very good agreement with the percentage provided by NMR for solution DNAs and by x-ray for crystal DNAs (20–30%) (Gorenstein, 1994; Winger et al., 1998). Again the results are significantly improved compared to the results obtained with the parm95 force field (~10% of BII) (Cheatham et al., 1999).

The histogram of the BII population as a function of the nucleotide (Fig. 3 *a*) indicates that a period of 10 ns is sufficient to render the sequence dependency (similar variations along the two palindromic strands). Nevertheless, a significant difference (up to 20%) is found between the BII population of some symmetrical phosphates (for example, between G6pA7 and G22pA23). A longer MD simulation is needed to sample the phosphate conformations compared with the helical parameters (~10 ns versus ~2 ns). This is due to the potential energy barrier between the BI and the BII states and to the dependence of the phosphate conformations on the diffusion of water molecules (Winger et al., 1998).

The largest BII percentages are found within the CRE consensus segment, that is, the half-sites TGA:TCA and the central CpG linker: 70% at $G_6pA_7/G_{22}pA_{23}$, 45% at $C_8pG_9/C_{24}pG_{25}$ and $C_{11}pA_{12}/C_{27}pA_{28}$. With exception of $G_6pA_7/G_{22}pA_{23}$ and $T_5pG_6/T_{21}pG_{22}$, where BII rates are unexpectedly very high and low, respectively, the hierarchy of BII percentages (Fig. 3 *a*) is similar to YpR > RpR > YpY > RpY (where Y and R stay for pyrimidine and purine, respectively), obtained from statistical analysis of NMR and crystallographic data (El antri et al., 1993; Gorenstein, 1994; Winger et al., 1998; Bertrand et al., 1998).

Helical parameters

The twist angle of CRE_MD averaged over the whole sequence, 33.7°, is that expected for B-DNAs in solution (34°) (Ulyanov and James, 1995). The averaged rise ranges between 3.2 Å (A_4pT_5) and 3.8 Å (C_8pG_9), and the averaged twist varies from 27.2° (A_4pT_5) to 44.4° (C_8pG_9) (Appendix). The averaged inclination and x displacement vary slightly, with values ranging from −1.8° and 2.6°, and from −1.2 Å and −0.7 Å, respectively. The base-base buckle and
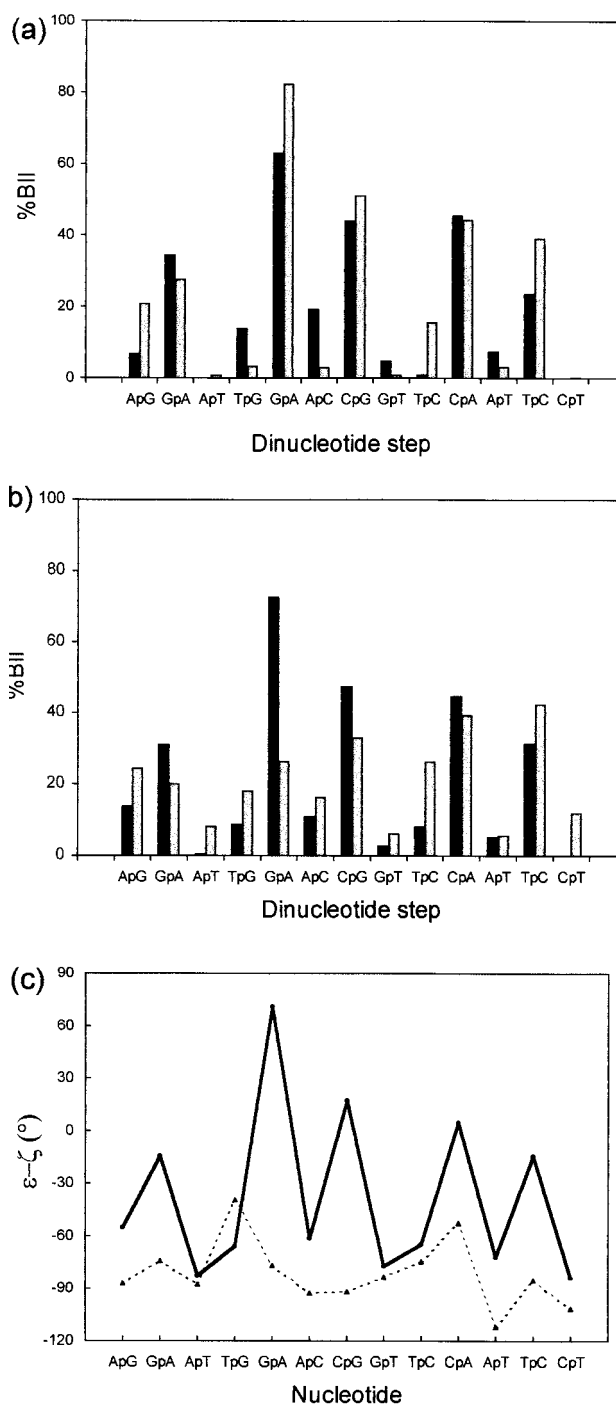

FIGURE 3 (*a*) Percentage population of the BII state in (AGATGACGTCATCT)$_2$ as a function of the dinucleotide step, for the first (*black histogram*) and the second (*gray histogram*) DNA strands simulated by the MD. (*b*) Comparison of the MD results averaged over the two DNA strands (*black histogram*) with the NMR data by Chaoui et al. (1999) (*gray histogram*). The NMR values have been calculated from the ^{31}P chemical shift measurements, using the relationship by Gorenstein (1994). (*c*) Difference between the ϵ and ζ dihedral angles of (AGATGACGTCATCT)$_2$ as a function of the dinucleotide step: MD values averaged over 9.5 ns (*solid line*) and values found in the cocrystal with GCN4 (Keller et al., 1995) (*dashed line*).

propeller twist parameters are submitted to important variations, with averaged values between $-13.9°$ (T_{15}:A_{18}) and $14.1°$ (A_2:T_{31}), and between $-16.0°$ (C_{14}:G_{19}) and $-3.7°$ (C_{11}:G_{22}), respectively. No disruption of hydrogen bonds is observed. In the phosphodiester backbone, most variable dihedral angles are ζ ($\Delta = 100°$), ϵ ($\Delta = 70°$), sugar phase ($\Delta = 50°$), and χ ($\Delta = 30°$). In contrast, the α, β, and γ angles are roughly constant, staying in g^-, t, and g^+ conformations, respectively, along the sequence. At last, both the minor and the major groove widths are typical of B-DNAs, with averaged values over CRE of 6.3 Å and 12.1 Å, respectively. Yet the grooves are narrower at the C_8pG_9 step in the center of CRE, with width values of nearly 5 Å (minor) and 11 Å (major).

DNA bending, roll, tilt

There is a sampling of crescent-shaped DNA structures with helix curvatures mostly concentrated on the central CpG. The DNA global curvature is defined by CURVES 5.2 as the angle between the local helical axes of the penultimate base pairs A_2:T_{31} and T_{15}:A_{18} (Lavery and Sklenar, 1988) (Fig. 4 *a, left*). The global curvature (Fig. 4 *b*) varies from $0°$ to $51°$ (14 bp), with an averaged value of $16.1°$ and a SD of $8.5°$. The bending direction is given by the Atip component of the kink angle (Fig. 4 *a, right*). The kink is defined as the angle between the two optimal linear axes for the DNA half-sides, fragments A_2:T_{31} to C_8:G_{25} and G_9:C_{24} to T_{15}:A_{18}, and the Atip is the component of the kink angle in the plane perpendicular to the long base-pair axis of the central base-pair step. The Atip component is negative or positive for a bending of the oligomer toward the minor or major groove, respectively. The variation of the Atip as a function of time (Fig. 4 *c*) shows that the CRE curvature oscillates between the major and the minor grooves, with an intermediate straight conformation, revealing no apparent anisotropy (examples of structures are shown in Fig. 5, *left* and *right*). The averaged Atip value reported as a function of the simulation length (Fig. 4 *d*) reveals that a trajectory of more than 5 ns is necessary to get a convergence of the DNA global bending. The distribution of the Atip values calculated over the entire trajectory (0.5–10 ns) is roughly Gaussian (data not shown), indicating that a trajectory of 10 ns is sufficient to sample the bending of a DNA 16-mer.

An understanding of why and how bending of CRE occurs requires examination of local perturbations in the double helix. Typical parameters responsible for DNA bending include the roll and, to a lesser extent, the tilt (EMBO Workshop, 1989; Young et al., 1995; Suzuki and Yagi, 1995; Dickerson, 1998). The roll is positive or negative for pinching toward the major or the minor groove of the DNA, respectively. The tilt is positive or negative for a pinching toward the second or the first strand backbone of the DNA, respectively. The MD structure shows loci of curvature, mainly in the central consensus sequence. The
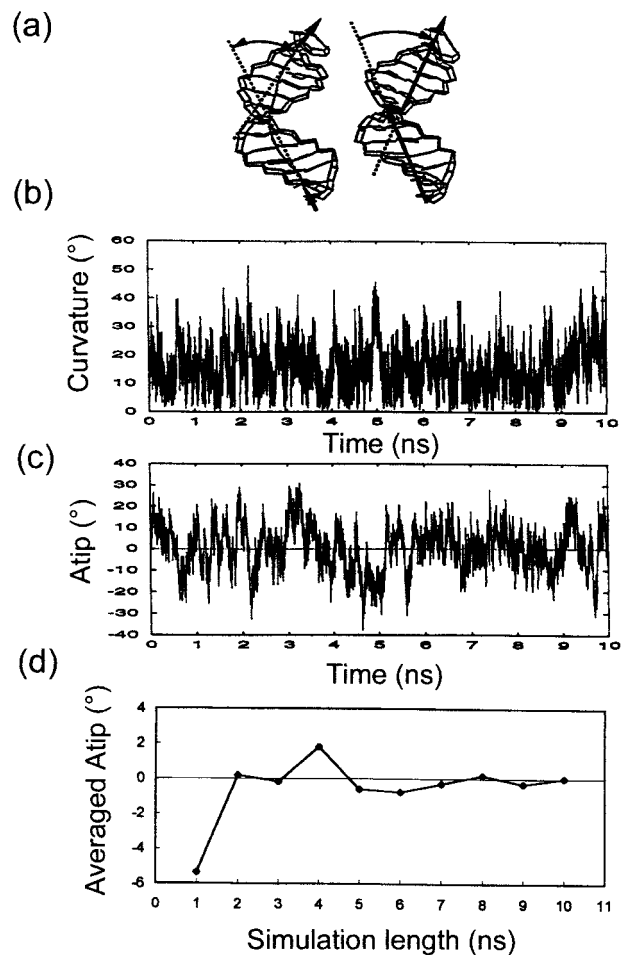

FIGURE 4 (*a*) Crescent-shaped conformations of CRE with representation of the curvature (*left*) and the Atip (*right*) structural parameters, as defined in the text. These are related to the global bending of the DNA oligomer and are calculated by CURVES. (*b* and *c*) Curvature and Atip values, respectively, for (AGATGACGTCATCT)$_2$ as a function of the MD simulation time. (*d*) Averaged Atip value as a function of the simulation length.

pyrimidine-purine steps T_5pG_6, C_8pG_9, and $C_{11}pA_{12}$ display averaged positive roll values of $8.9°$, $5.7°$, and $8.6°$, respectively, and high standard deviation values of $6.9°$, $7.3°$, and $6.8°$, respectively (Fig. 6 *a*). The major tilt deformations are found at the T_5pG_6 and G_6pA_7 steps ($2.4° \pm 5.5°$ and $-3.2° \pm 4.5°$), at the symmetrical $T_{10}pC_{11}$ and $C_{11}pA_{12}$ steps ($4.4° \pm 4.7°$ and $-3.4° \pm 5.6°$), and at C_8pG_9 ($1.0° \pm 5.7°$) (Fig. 6 *b*). The CpG and TpG/CpA steps appear as the most malleable steps, which conforms to the crystallographic (Young et al., 1995; Suzuki and Yagi, 1995; Dickerson, 1998) and NMR (Ulyanov and James, 1995; El antri et al., 1993; Lefebvre et al., 1995) data; the effect is related to a lower stacking energy in these steps (Bertrand et al., 1998).

The small tilts at T_5pG_6 + G_6pA_7 and $T_{10}pC_{11}$ + $C_{11}pA_{12}$ are in phase and additive (Dickerson, 1998) with the roll at C_8pG_9, resulting in a total kink angle of $7.5°$ for
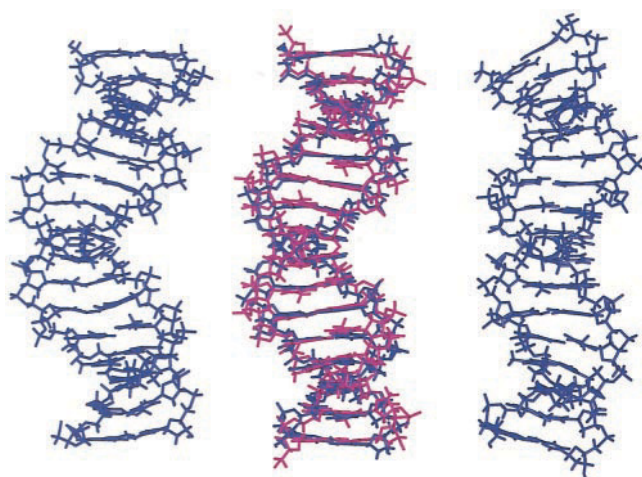
FIGURE 5 Various crescent-shaped conformations of (GAGAT-GACGTCATCTC)$_2$. (*Left*) Helix curved toward the major groove, at a MD simulation time of 3049 ps. (*Center*) Averaged MD conformation (9.5 ns) (*blue*) superimposed over the x-ray structure (*pink*) of the CRE 16-mer bound to GCN4 (Keller et al., 1995). Both DNAs are slightly curved toward the major groove. (*Right*) Conformation bent toward the minor groove, at a MD simulation time of 5050 ps.

the CRE consensus segment. Thus the CRE DNA segment is globally curved on the average by 7°-8° toward the major groove at its center, i.e., toward the dimer bZIP protein partner, with a global "C" form. The positive rolls (4.2°) at the A$_2$pG$_3$ and C$_{14}$pT$_{15}$ steps ending the DNA duplex are roughly in antiphase with the positive roll at the central C$_8$pG$_9$, as these steps are separated by 6 bp. As a consequence the Atip is reduced, explaining the apparent isotropy of the Atip curve (Fig. 4 *c*). In the binding half-sites, the high positive rolls at the T$_5$pG$_6$ and C$_{11}$pA$_{12}$ steps create local bends toward the major groove. Globally, the molecule is S-shaped in the plane perpendicular to the Atip one, with each half-site precurved toward one monomer of the bZIP protein.

The plot of rolls averaged over the MD structures and classified as a function of their Atip value (Atip > 10°, < −10°, or ≅ 0°) reveals a strong correlation between the roll of the central CpG and the Atip (Fig. 7). The structure of CpG modulates the overall bending: an increase in roll changes a global bending toward the minor groove (Atip < −10°) to a global bending toward the major groove (Atip > 10°); a null roll is related to a global bending toward the minor groove because of the presence of positive rolls at the A$_2$pG$_3$ and C$_{14}$pT$_{15}$ steps 6 bp from CpG. Regarding the tilts, no clear correlation with the Atip was found.

Comparison with experimental data

CRE bending: MD versus biochemical methods

The consensus CRE segment in CRE_MD (7–8°) presents an averaged curvature toward the major groove very similar
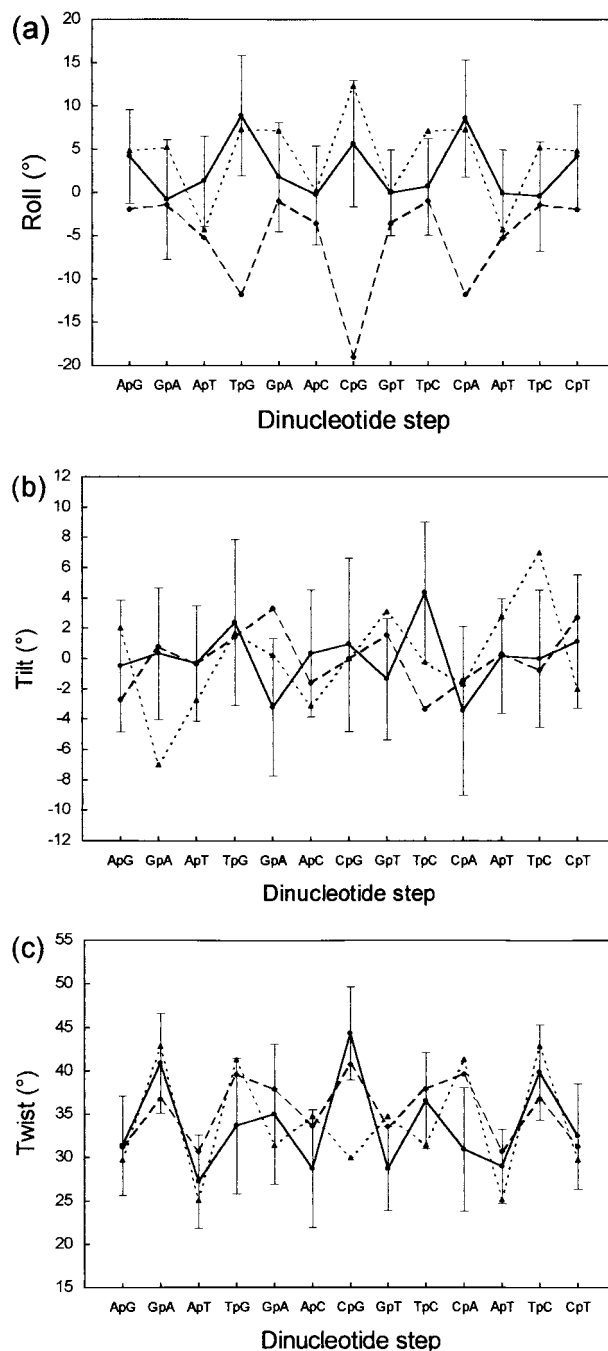
FIGURE 6 (*a–c*) Rolls, tilts, and twists, respectively, in (AGAT-GACGTCATCT)$_2$ as a function of the dinucleotide step: MD averaged values ± 1 SD (9.5 ns) (*solid line*), and values in the NMR refined structure by Chaoui et al. (1999) (*long dashed line*) and in the structure cocrystallized with GCN4 (Keller et al., 1995) (*short dashed line*).

to that obtained (nearly 10°) by phasing analysis (Paolella et al., 1994), cyclization kinetics (Hockings et al., 1998), and ligation ladder (Sloan and Schepartz, 1998). It can be noted, however, that the DNA bending is not due to the sole rolls

FIGURE 7   Rolls in $(AGATGACGTCATCT)_2$ as a function of the sequence, averaged over the MD conformations, classified as a function of their Atip value: Atip $> 10°$ (*solid line*), Atip $< 10°$ and $> -10°$ (*long dashed line*), and Atip $< -10°$ (*short dashed line*).
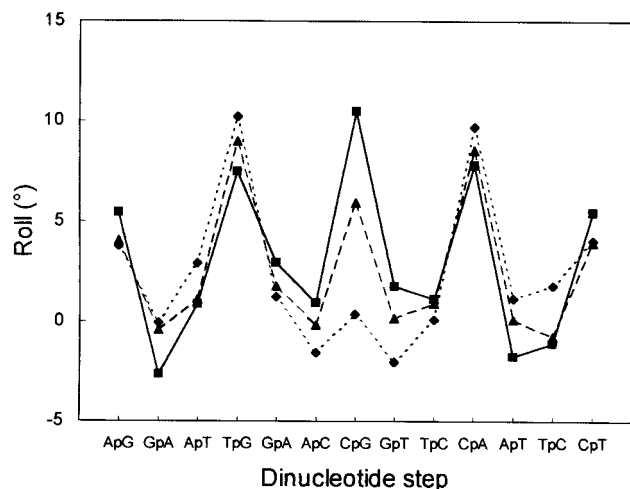
at the TpG/CpA steps, as suggested by Sloan and Schepartz (1998), but rather to a favorable combination of roll and tilts at the central CpG and the TGA half-sites, respectively.

CRE structure and bending: MD versus NMR

The CRE_NMR structure (Chaoui et al., 1999) has been calculated with the JUMNA 10.0 molecular mechanics software (Lavery, 1988; Lavery et al., 1995), using the JUMNA force field, Flex (Lavery et al., 1984, 1986a,b). A set of different initial structures corresponding to local energy minima were generated by JUMNA. These were minimized with the NMR restraints on the phases, the differences $\epsilon - \zeta$, and the ^1H-^1H distances, introduced step by step (Chaoui et al., 1999). All of the initial structures converged toward a final unique structure, CRE_NMR, with a curvature directed toward the minor groove. When the force field Flex was replaced by parm95 (Cornell et al., 1995), one obtained a final conformation with a rmsd of 1.0 Å to CRE_NMR (Chaoui et al., 1999).

In Fig. 1 *b* are plotted the rmsds as a function of time between the instantaneous MD structures, CRE_MD, and the available experimental structures. CRE_MD is clearly closer to CRE_GCN4 (rmsds ~2 Å) than to CRE_NMR (rmsds ~3.5 Å). CRE_NMR is found to be very different from ⟨CRE_MD⟩ and CRE_GCN4 (rmsd = 3.6 Å).

Backbone conformation. The sugar phase angles in CRE_MD were compared to the NMR data of Chaoui et al. (1999). There is a good agreement between calculation and experiments, with the same phase angle variation as a function of the nucleotide sequence. The phase angles averaged over the MD trajectory stay within the NMR uncertainties, except that of the C_{14} sugar (Fig. 2). The agreement between

the current MD data and the NMR data of Conte et al. (1997) related to the CRE segment is also good. Sugars at C_{14}/C_{30} show a preference for the Eastern state (O4′-endo, ~90°) in CRE_NMR and for the C1′-exo state (~126°) in CRE_MD. Actually, during the MD trajectory, these two sugars keep the C2′-endo conformation most of the time, with phase angles of nearly 150°. As they transit from time to time into the Northern conformation, their averaged conformation lies in the range of C1′-exo puckers. The Eastern state provided by NMR could better correspond to an average of Southern and Northern states. Whatever the real conformation adopted by residues C_{14}/C_{30}, there is a disagreement between the MD calculations and the NMR measurements.

Regarding the phosphate conformations, the BII populations predicted by MD simulation are close to those estimated from NMR ^{31}P chemical shifts, using the Gorenstein (1994) relationship (Fig. 3 *b*). For the average over the sequence, both NMR and MD provide 21% of BII population, whereas for the BII population at each step, the absolute difference does not exceed 10%. Exceptions are G_6pA_7/ $G_{22}pA_{23}$ and their complementary steps $T_{10}pC_{11}/T_{26}pC_{27}$. For the first steps, relative BII populations of 70% and only 25% are predicted by MD and NMR, respectively, while for the second steps MD predicts a smaller BII population (-20%). Globally, the sequence dependence of the BII population is found to be more pronounced by MD calculations than by NMR measurements (1 SD = 21% versus 12%).

Helical parameters and bending propensity. The current MD simulation disagrees with the recently reported NMR restrained molecular modelings of both Conte et al. (1997) and Chaoui et al. (1999). These show a curvature of 30° (16-mer) toward the minor groove and rolls negative at CpG ($-19°/-10°$) and null or negative at TpG:CpA ($0°/-12°$). Actually, the plot of rolls as a function of dinucleotide steps for CRE_NMR (Chaoui et al., 1999) is the opposite of that of CRE_MD (Fig. 6 *a*), but this also proves to be true for the tilts at G_6pA_7 and $T_{10}pC_{11}$. It seems that tilts adapt themselves to smooth the bending caused by rolls. In addition, the averaged twist in CRE_MD is smaller compared with that found in CRE_NMR (33.7° versus 35.4°). Yet the profiles of the twist sequence dependences are very similar (Fig. 6 *c*), with the greatest differences being observed at CpG and T_5pG_6 and $C_{11}pA_{12}$. The values of CRE_NMR are within the averaged values $\pm$ 1 SD of CRE_MD.

Influence of software and protocol on NMR structure determination. The curvature of CRE predicted by MD is similar to that derived from biochemical data but differs from that provided by NMR refinement methods.

DNA conformations resolved by NMR refinement depend on different factors such as measurements and methodology. The latter includes the modeling software and the refinement protocol. The helical parameters are determined by the internucleotide distances obtained from NOE, data which are refined through back-calculations of theoretical
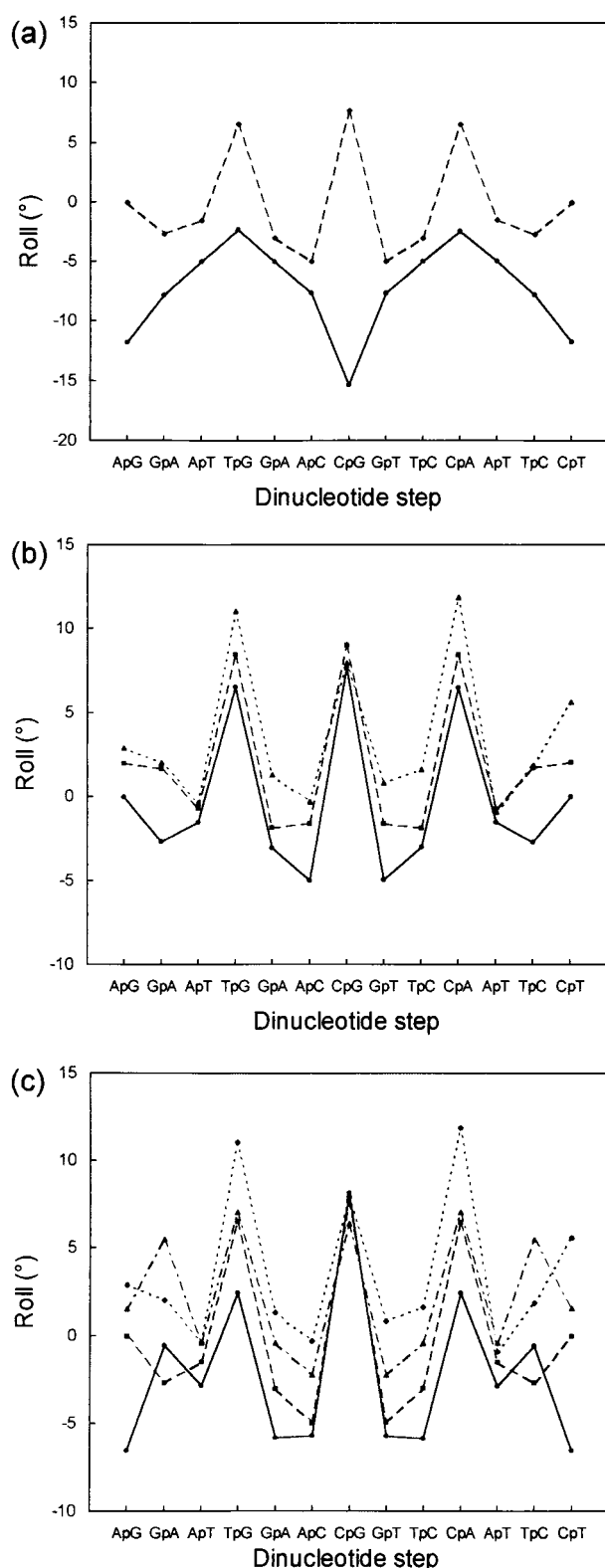
NOESY spectra, using the NOE Simulate routine of IN-SIGHT II (MSI) (Boelens et al., 1989; Lefebvre et al., 1995, 1996; Chaoui et al., 1999). Moreover, the bending can be strongly influenced by the force field. Its impact can be tested by comparing, for instance, the rolls in the minimal energy structures provided by JUMNA, using Flex and parm95. Conditions are the same as the ones used for the CRE refinement by Chaoui et al. (1999) (dielectric function and screening of the phosphate charges), but the NMR constraints are excluded. One finds that rolls are modulated by the force field (Fig. 8 a), especially at the CpG and TpG/CpA steps, where JUMNA-Flex provides negative rolls and JUMNA-parm95 provides positive ones. Replacing parm95 with parm98 (force-field parameters modified for three dihedral angles, no change of the partial charges) in JUMNA does not significantly alter the rolls: they differ by less than 2° at each step of the CRE sequence. As in most softwares dedicated to the NMR refinement of molecular structures, the DNA environment in JUMNA is simulated in an implicit way: the dielectric effect of water is modeled by a sigmoid function with a slope value of 0.16 and a plateau of 78, and the screening effects of counterions on phosphates are taken into account by reducing the charge of the phosphate groups to a value of $-0.5e$ (Lavery et al., 1995). Using different slope values of the dielectric function (0.16 and 0.356) and varying the apparent charge of the phosphates ($-0.5e$, $0.0e$, and $-1.0e$) for the minima provided by JUMNA-parm95 shows that the roll values increase parallel to the dielectric function and when the phosphate charge tends to zero. A MD simulation with the same parm95 force field (unpublished results) with an explicit consideration of the solvent provides a similar curve, but with more positive values of rolls (Fig. 8, b and c). However, we cannot conclude that it is possible to find an implicit electrostatic parameterization that would correctly simulate, whatever the DNA sequence, the influence of the solvent on the DNA structure. Actually, structure alterations caused by specific localization of water molecules or hydrated ions (Diekmann, 1987; Rouzina and Bloomfield, 1998; Lyubartsev and Laaksonen, 1998; Winger et al., 1998; McFail-Isom et al., 1998; Shui et al., 1998) cannot be simulated by varying the apparent phosphate charge or by using a dielectric function.

From the above results one can conclude that the electrostatic parameterization of the software exerts a significant effect on DNA bending, through both the force field

FIGURE 8   Rolls in (AGATGACGTCATCT)$_2$ for the conformations with minimal energy calculated by JUMNA with different physical parameters. (*a*) Rolls as a function of the force field used: Flex (*solid line*) or parm95 (*dashed line*). (*b*) Rolls with parm95 as a function of the dielectric model: sigmoid function with a slope of 0.16 (*solid line*) or 0.356 (*long dashed line*). Results were obtained by a MD simulation in explicit water with AMBER and the parm95 force field (*short dashed line*). (*c*) Rolls with parm95 as a function of the counterion model: apparent phosphate charges of $-1.0e$ (*solid line*), $-0.5e$ (*long dashed line*), or $0.0e$ (*mixed dashed line*). Results were obtained by a MD simulation in explicit solvent and Na$^+$ counterions with AMBER and the parm95 force field (*short dashed line*).

(higher polarity of the atomic charges for parm95/parm98 than for Flex) and the solvent model. As far as CRE is concerned, a major factor could be the electronegativity of the major groove along the consensus d(TGACGTCA)$_2$ helix. Actually, CRE displays a succession of intrastrand electronegative pockets, at 5′-T$_5$pG$_6$-3′, 3′-T$_{26}$pG$_{25}$-5′, 5′-G$_9$pT$_{10}$-3′, and 3′-G$_{22}$pT$_{21}$-5′, connected one to another by interstrand electronegative pockets 5′-G$_6$pA$_7$-3′, 5′-C$_8$pG$_9$-3′, and 5′-T$_{10}$pC$_{11}$-3′ (Young et al., 1997a). The succession of thymine O4 and guanine O6 and N7 atoms creates a continuous negative channel along the major groove, which could push away subsequent base pairs and induce negative (or less positive) rolls. The repelling of base pairs would in this case decrease as the dielectric effect increases.

The NMR restraints introduced successively on the sugar phases and on the $\epsilon - \zeta$ differences did not significantly modify the roll values. In contrast, the restraints on ^{1}H-^{1}H distances, particularly the H6/8-H6/8 distances, have both "pulled" the rolls of CpG and TpG/CpA toward strong negative values and bent CRE toward the minor groove. This occurs whatever the force field used, Flex or parm95 (Fig. 9) (Chaoui et al., 1999). Although the interbase H6/8-H6/8 distance between subsequent bases is related to the roll value, it is not sufficient to determine precisely the roll because, for the same H6/8-H6/8 distance, different positive or negative rolls are possible, depending on the twist value. Actually, both the H6/8-H6/8 distances and the H3′-H6/8 distances (which are not always amenable to 2D NMR experiments for signal overlapping reasons) are needed to determine the rolls univocally (Lefebvre et al., 1997). For instance, in the CRE NMR restrained structures obtained with JUMNA-Flex and JUMNA-parm95 by Chaoui et al. (1999), the roll values of the central CpG are different, although these structures present similar H6/8-H6/8 distances (Fig. 10). On the other hand, the roll values in the CRE structures determined by Conte et al. (1997) and Chaoui et al. (1999) (with JUMNA-parm95) are similar, although the H6/8-H6/8 distances are very different. In these structures the restraints on the H3′-H6/8 distances strongly related to the rolls are severely missed (Lefebvre et al., 1997). Perhaps, whatever the starting structure, the software reacts to the adjunction of the H6/8-H6/8 distance restraints by "pushing" the CRE conformation toward a region of the conformational space that is "more convenient" for it (i.e., abnormally negative rolls but correct twists).

Another obvious fault of the refinement procedure that could influence the rolls concerns the internal molecular motions, so that the refinement procedure provides erroneous internucleotide ^{1}H-^{1}H distances, especially at the most flexible steps (TpG:CpA and CpG).

The disagreement on the bending of CRE_NMR and CRE_MD could also result from the solvent conditions, which are different in the NMR and MD experiments, with regard to both the ionic strength (MD $\cong$ 0.5; NMR $\cong$ 0.1) and the nature of the ions (MD: Na$^+$; NMR: buffer + purification + impurities → Na$^+$, Li$^+$, K$^+$, and perhaps
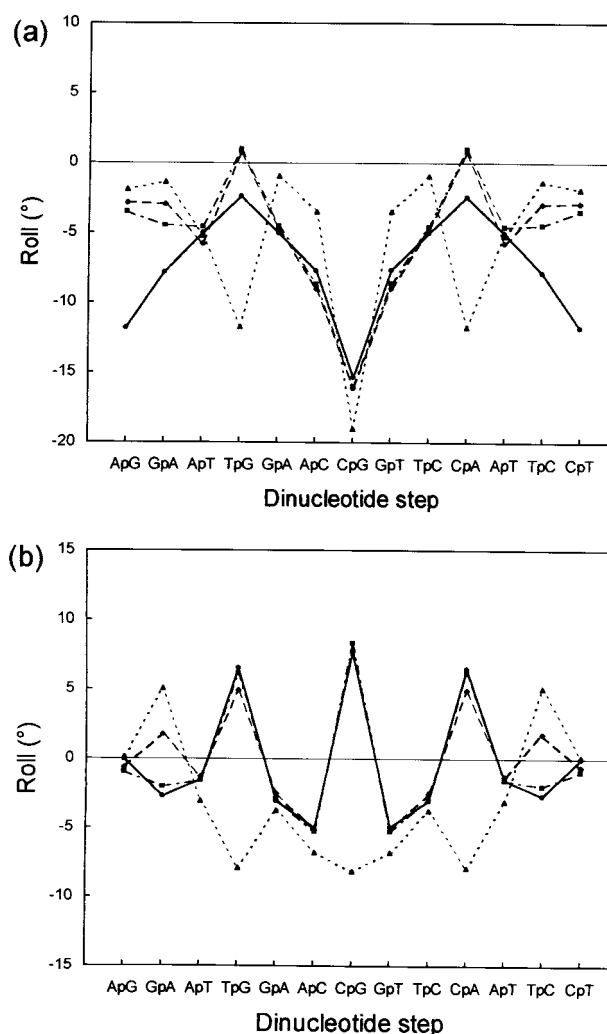

FIGURE 9   Rolls for the (AGATGACGTCATCT)$_2$ conformations with minimal energy calculated by JUMNA, showing the effects of the NMR restraints introduced successively during the refinement of the structure, with the Flex (*a*) and the parm95 (*b*) force fields. *Solid line*: without NMR restraints; *long dashed line*: restraints on the sugar phase angles; *mixed dashed line*: addition of the restraints on $\epsilon - \zeta$; *short dashed line*: addition of the restraints on the ^{1}H-^{1}H distances.

also traces of divalent cations like Mn^{2+}, Mg^{2+}, Zn^{2+}). Different monovalent ions could induce different local DNA deformations, as crystallographic (Tereshko et al., 1999) and MD (Lyubartsev and Laaksonen, 1998) studies performed with various monovalent cations argue that they bind to different specific DNA sites (for instance, Na$^+$ preferentially binds to the minor groove and Li$^+$ binds to the phosphate oxygens). In a recent MD study, Young and Beveridge (1998) have shown that the nature and the concentration of monovalent cations can have a significant impact on the DNA curvature: replacing neutralizing Na$^+$ with K$^+$ and KCl doubles the bending of a 25-bp DNA segment containing an A-tract. Moreover, gel retardation (Diekmann, 1987), theoretical investigations (Rouzina and
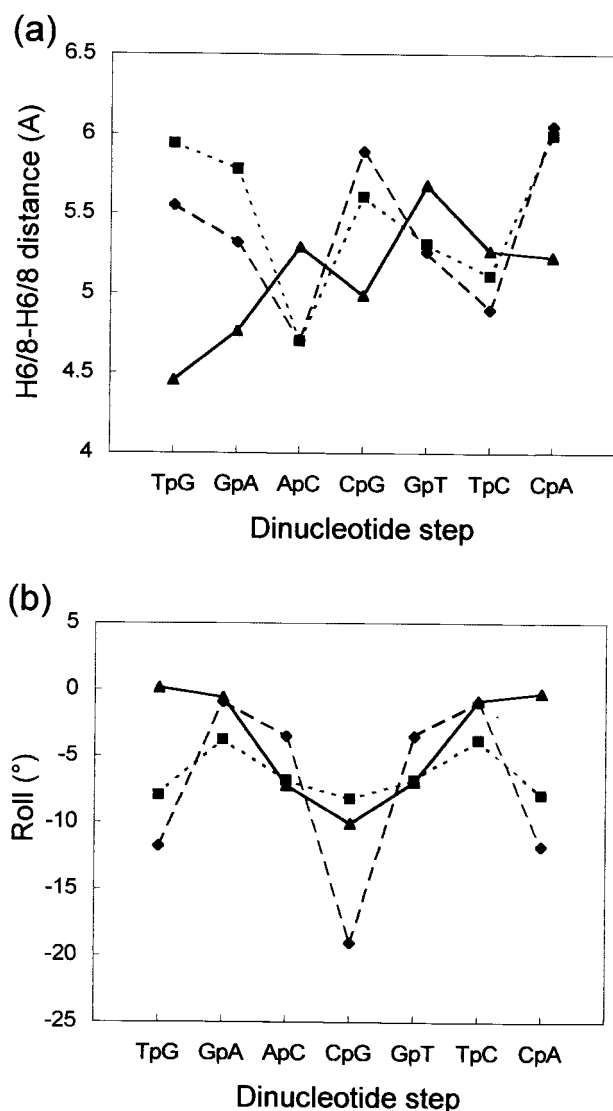
FIGURE 10 H6/8-H6/8 distances (*a*) and rolls (*b*) in the CRE consensus sequence (TGACGTCA)$_2$ as a function of the sequence, for the NMR refined structures obtained by Conte et al. (1997) (*solid line*) and by Chaoui et al. (1999) with the Flex (*long dashed line*) and the parm95 (*short dashed line*) force fields.

Bloomfield, 1998), and crystallographic data (McFail-Isom et al., 1998) suggest that divalent cations undergo sequence-specific interactions with DNA that have a particular impact on the DNA bending. Thus the nonequivalent ionic conditions between the NMR and MD studies could in principle contribute to the appearance of diverging results regarding the bending of CRE. As the MD results agree with the NMR data regarding the conformation of the phosphodiester backbone (sugar phases and phosphate conformations), this would mean that the diverging ionic conditions would only influence the helical parameters, including the rolls. But it is interesting to note that the biochemical data on the bending of CRE support the MD results, although the experimental

conditions diverge from the MD conditions. For instance, the measurements of cyclization kinetics (Hockings et al., 1998) have been made in the presence of KCl (50 mM) and MgCl$_2$ (5 mM), and the ligation ladder experiments could have been performed in the presence of Mg^{2+} impurities (Sloan and Schepartz, 1998).

In the end, the influence of the phosphorus NMR restraints has also to be considered. Such restraints could force the steps involved in a BI/BII equilibrium to adopt time-averaged $\epsilon - \zeta$ values devoid of physical sense. A forced backbone may strongly alter the shape of the helix because of the relation that exists between $\epsilon - \zeta$ and the helical parameters, including the rolls that are strongly involved in the DNA curvature (El antri et al., 1993; Hartmann et al., 1993; Gorenstein, 1994; Lefebvre et al., 1996; Bertrand et al., 1998). Actually, one could generate different initial structures with each phosphate either in the BI or in the BII conformation, making a combinatorial search with all of the steps susceptible to adopting both BI and BII states. One could then perform an energy minimization without constraint on the ϵ and ζ angles on the different initial structures and obtain final structures with different patterns of BI/BII phosphates. Finally, one could compare the averaged $\epsilon - \zeta$ values to the NMR data (Lefebvre et al., 1996; Tisné et al., 1998).

CRE in the complex with GCN4

The rmsd between ⟨CRE_MD⟩ (averaged MD structure without further minimization) and CRE_GCN4 is only 1.3 Å (Table 1 and Fig. 5, *center*). This is lower compared to the averaged rmsd of the instantaneous MD structures (~2 Å) (Fig. 1 *b*), probably because, at a first approximation, it does not take into account the vibrational motions of the DNA molecule. The good agreement between the simulated CRE structure and the crystal structure of CRE bound to GCN4 indicates that the binding of the protein does not significantly affect the DNA, in agreement with the biochemical results (Paolella et al., 1994, 1997; Hockings et al., 1998;

TABLE 1 Rmsds (in Å) between the averaged MD, experimental, and canonical structures

	⟨CRE_MD⟩	B DNA	A DNA	CRE_NMR	CRE_GCN4
⟨CRE_MD⟩		1.5	4.8	2.2	1.4
B-DNA	2.2		5.4	1.6	2.1
A-DNA	6.1	7.5		6.0	4.3
CRE_NMR	3.1	2.1	8.4		2.7
CRE_GCN4	1.3	2.7	5.8	3.6	

Root mean square deviations (rmsd) between the structure of (AGATGACGTCATCT)$_2$ averaged over the MD trajectory (9.5 ns), ⟨CRE_MD⟩; the canonical B DNA and A DNA conformations; the NMR refined structure by Chaoui et al. (1999), CRE_NMR; and the conformation of the DNA in the cocrystal with GCN4 (Keller et al., 1995), CRE_GCN4. The upper right part of the table gives the rmsd values calculated for the CRE consensus sequence, i.e., the central octamer (TGACGTCA)$_2$; and the lower left part of the table gives the rmsd values for the whole 14-mer.

Sloan and Schepartz, 1998). The rmsd for the central octamer is higher compared with that for the entire 14-mer (1.4 Å versus 1.3 Å), translating a larger DNA deformation at the binding sequence. Actually, in CRE_GCN4 the octamer helix has drifted away from the canonical B form toward the canonical A form: the rmsds from the B DNA (A DNA) are 1.5 Å (4.8 Å) and 2.1 Å (4.3 Å) for $\langle CRE\_MD \rangle$ and CRE_GCN4, respectively. Four sugars have smaller phase angles in CRE_GCN4 compared with the free oligomer (Fig. 2). For the other sugars, the low resolution of the crystal structure (2.2 Å) makes the precise determination of phase angles difficult, preventing a comparison with the MD results. In the same way, in CRE_GCN4 the protein blocks all of the phosphates, except at the TpG:CpA steps, in the BI conformation ($\epsilon - \zeta \approx -90°$) (Fig. 3 c), a characteristic of the A DNA helix. Such a preference of phosphate groups for the BI conformation has already been observed in DNA-protein complexes (Gorenstein, 1994). Note that previous mechanical studies have shown that the phosphate conformations correlate with the twists: the twist angle increases while the phosphate tends toward BII (Hartmann et al., 1993; Bertrand et al., 1998). This correlation is roughly observed in CRE_MD, but not at all in the crystal structure (Figs. 3 c and 6 c).

The crescent structure of CRE_MD resembles that displayed by CRE_GCN4, with a global curvature toward the major groove of 16.1° in the former and 15° in the latter. Rolls at the CpG and the TpG/CpA steps are positive, with, however, a higher value for CpG (12.3° versus 5.7°) and a slightly smaller value for TpG/CpA (7.3° versus 8.9° and 8.6°) (Fig. 6 a). A significant difference consists of rolls at G_6pA_7 and $T_{10}pC_{11}$ that are positive in the crystal and almost null in CRE_MD (7.1° versus 1.8° and 0.8°). However, the roll values of CRE_GCN4 stay within the averaged MD values $\pm$ 1 SD, suggesting that the small DNA bending deformation induced by GCN4 is accessible through thermal energy. The straight conformation adopted by CRE in the complex with BP1 (Paolella et al., 1994) could also be reached with little consumption of energy, as a null roll at CpG stays within the MD averaged value $\pm$ 1 SD. Rolls with averaged values and SD higher at the CpG and TpG/CpA steps reflect the malleability of the helix at both the linker and the half-site positions. The roll differences between CRE_MD and CRE_GCN4 are caused by a drift of the kinks from TpG/CpA to CpG induced by the protein, thus confirming the gel results (Sloan and Schepartz, 1998).

Differences in tilts are observed only for G_3pA_4 and the $T_{13}pC_{14}$ steps (Fig. 6 b). The twist angle of CRE_MD averaged over the whole sequence (33.7°) is similar to that found in CRE_GCN4 (33.8°). However, although similar for the residues lying outside the CRE consensus octamer, the twists are very different inside CRE, especially at the CpG step (29.9° in CRE_GCN4 versus 44.4° in CRE_MD). Such a twist variation at CpG (correlated to the roll) is not accessible through thermal energy (Fig. 6 c) and must be induced by work done during complexation.

The comparison of CRE with TRE seemed interesting, as both are stable partners of GCN4. As far as the distance between the half-sites TGA/TCA and their relative orientations are concerned, the structural difference between CRE and TRE is reduced upon protein binding, as TRE keeps a straight conformation (Ellenberger et al., 1992; Keller et al., 1995). The fact that a greater positive roll is mechanically correlated to a smaller twist (Hartmann et al., 1993; Bertrand et al., 1998) suggests that the coupled change in the central CpG roll and twist during the formation of the complex of CRE with GCN4 follows a general mechanism in DNA that allows a reduction of the structural difference between promoters with half-sites that diverge by the spacer length.

CONCLUSION

A long molecular dynamics simulation of 10 ns was made on a 16mer DNA sequence containing the CRE consensus sequence. Results were compared with those provided by NMR and biochemical methods. The B-DNA conformation presented by CRE oscillates between a helix curved toward the major groove and a helix curved toward the minor groove, through an intermediate straight form. However, as a final result, the averaged helix is bent slightly toward the major groove. Although this finding is consistent with the biochemical results, it disagrees with the refined NMR structures, where the bending is directed toward the minor groove. This reversion is imputed to an improper electrostatic parameterization in the NMR structure refinement softwares, inasmuch as the NMR data agree with the MD simulation results. In contrast, the MD structure appears to be very similar to the crystal structure of CRE bound to GCN4. In particular, the rolls (often involved in the DNA curvature) are roughly conserved in the complex. These parameters could be the structural determinants of the CRE recognition by GCN4. A large part of the helix deformation induced by GCN4 is concentrated at the central CpG step that links the TGA:TCA half-sites. CpG undergoes a concerted increase in roll and a decrease in twist that reduces both the distance between the TGA:TCA half-sites and their difference in orientation. As a result, the structure of CRE bound to GCN4 is more like the structure of TRE, which remains straight in its complex with GCN4.

Thus the current MD simulation reveals the high flexibility of the CRE consensus sequence, reflecting its ability to accommodate itself to various bZIP proteins. The impact of the central CpG dinucleotide on the conformation of CRE is clearly identified: the inherent bending and malleability of CpG are the major contributions to the curving and flexibility of CRE. The absence of CpG at the center of TRE could be the structural explanation for the observation by biochemical methods that TRE is unable to bind to such varied bZIP partners and in such numbers as CRE does.

APPENDIX

	Rise	Twist	Roll	Tilt	Shift	Slide
ApG	3.61 (0.36)	31.42 (5.75)	4.21 (5.41)	-0.47 (4.35)	0.08 (0.72)	-1.04 (0.61)
GpA	3.62 (0.32)	40.89 (5.73)	-0.77 (6.95)	0.33 (4.36)	-0.04 (0.61)	-0.81 (0.71)
ApT	3.19 (0.27)	27.25 (5.40)	1.35 (5.21)	-0.29 (3.81)	-0.10 (0.56)	-1.00 (0.45)
TpG	3.43 (0.34)	33.67 (7.83)	8.92 (6.94)	2.41 (5.49)	0.56 (0.86)	-0.28 (0.64)
GpA	3.31 (0.29)	35.03 (8.10)	1.85 (6.28)	-3.18 (4.54)	-0.70 (0.75)	-0.16 (0.55)
ApC	3.26 (0.30)	28.69 (6.83)	-0.26 (5.76)	0.36 (4.18)	0.26 (0.65)	-0.76 (0.44)
CpG	3.81 (0.36)	44.36 (5.34)	5.71 (7.32)	0.95 (5.72)	0.19 (0.77)	-0.50 (0.57)
GpT	3.28 (0.28)	28.70 (4.84)	0.02 (4.96)	-1.33 (4.00)	-0.43 (0.59)	-0.91 (0.45)
TpC	3.31 (0.28)	36.56 (5.58)	0.76 (5.57)	4.40 (4.68)	0.88 (0.67)	-0.23 (0.59)
CpA	3.38 (0.36)	30.96 (7.14)	8.61 (6.79)	-3.43 (5.56)	-0.88 (0.84)	-0.35 (0.56)
ApT	3.21 (0.26)	28.92 (4.30)	-0.09 (5.09)	0.18 (3.79)	0.06 (0.60)	-1.09 (0.39)
TpC	3.59 (0.32)	39.81 (5.51)	-0.38 (6.33)	0.01 (4.53)	0.07 (0.64)	-0.92 (0.67)
CpT	3.59 (0.35)	32.44 (6.10)	4.20 (6.05)	1.15 (4.41)	0.08 (0.81)	-0.95 (0.65)

	x-disp	y-disp	Inclination	Tip	Buckle	Propeller	Opening	Shear	Stretch	Stagger
A:T	-1.20 (0.56)	0.43 (0.39)	-1.38 (5.94)	-4.12 (5.18)	14.10 (10.14)	-13.61 (7.97)	3.56 (5.52)	0.21 (0.28)	0.19 (0.15)	0.22 (0.49)
G:C	-1.13 (0.60)	0.29 (0.38)	-1.81 (5.52)	-1.43 (4.48)	6.91 (10.45)	-15.71 (8.27)	1.65 (3.61)	0.06 (0.33)	0.13 (0.13)	-0.08 (0.40)
A:T	-1.07 (0.56)	0.27 (0.34)	-1.19 (4.89)	-1.52 (4.56)	-6.95 (10.34)	-12.03 (8.67)	3.26 (5.32)	0.10 (0.28)	0.16 (0.14)	-0.19 (0.42)
T:A	-1.08 (0.53)	0.03 (0.32)	0.16 (4.66)	-1.79 (4.69)	1.06 (10.11)	-10.48 (8.11)	4.06 (5.48)	-0.08 (0.29)	0.16 (0.14)	-0.10 (0.41)
G:C	-0.73 (0.51)	0.11 (0.29)	2.61 (4.40)	1.27 (4.62)	2.08 (11.27)	-5.98 (9.45)	1.14 (3.40)	-0.02 (0.32)	0.13 (0.12)	-0.13 (0.38)
A:T	-1.01 (0.48)	0.23 (0.29)	1.56 (4.33)	0.41 (4.24)	5.45 (11.41)	-10.68 (9.08)	1.56 (5.15)	0.21 (0.28)	0.15 (0.13)	0.05 (0.41)
C:G	-0.89 (0.44)	0.00 (0.28)	1.20 (4.32)	-1.61 (4.52)	10.41 (10.56)	-11.69 (8.91)	1.72 (3.47)	0.05 (0.31)	0.15 (0.13)	-0.10 (0.38)
G:C	-0.82 (0.48)	0.01 (0.28)	0.82 (4.26)	0.37 (4.38)	-11.39 (11.31)	-11.90 (8.66)	1.76 (3.50)	-0.06 (0.32)	0.15 (0.13)	-0.13 (0.39)
T:A	-1.06 (0.44)	-0.25 (0.30)	0.11 (4.33)	-0.57 (3.97)	-7.45 (10.30)	-12.69 (7.57)	2.55 (5.34)	-0.22 (0.28)	0.15 (0.13)	0.10 (0.43)
C:G	-0.70 (0.51)	-0.10 (0.28)	1.64 (4.44)	-0.91 (4.31)	-4.27 (11.74)	-3.75 (9.45)	1.15 (3.38)	0.01 (0.32)	0.13 (0.12)	-0.14 (0.40)
A:T	-1.18 (0.48)	-0.03 (0.33)	-0.61 (4.45)	2.89 (4.59)	-0.51 (9.49)	-10.24 (7.78)	3.47 (5.37)	0.10 (0.29)	0.16 (0.14)	-0.08 (0.41)
T:A	-1.16 (0.50)	-0.28 (0.34)	-1.36 (4.88)	1.77 (4.65)	6.21 (10.33)	-13.23 (8.12)	2.62 (5.35)	-0.13 (0.28)	0.16 (0.14)	-0.16 (0.41)
C:G	-1.20 (0.55)	-0.35 (0.38)	-1.29 (5.75)	1.38 (4.31)	-6.74 (10.60)	-15.96 (8.08)	1.84 (3.65)	-0.04 (0.33)	0.13 (0.14)	-0.08 (0.40)
T:A	-1.21 (0.53)	-0.46 (0.39)	-0.43 (6.26)	3.56 (5.36)	-13.90 (10.38)	-13.04 (8.87)	3.96 (5.95)	-0.20 (0.28)	0.18 (0.16)	0.18 (0.50)

Strand 1	Phase	Amplitude	Chi	Alpha	Beta	Gamma	Delta	Epsilon	Zeta
A	174.17 (23.76)	36.16 (6.15)	258.08 (12.82)	-70.39 (11.68)	171.48 (11.83)	48.98 (14.21)	143.28 (9.75)	174.17 (23.76)	36.16 (6.15)
G	147.26 (21.14)	43.33 (6.32)	252.42 (17.56)	-70.33 (13.52)	174.23 (17.71)	49.49 (10.43)	134.97 (13.01)	214.17 (46.89)	221.15 (54.49)
A	156.56 (39.56)	35.87 (6.10)	243.41 (15.69)	-66.65 (10.62)	174.46 (10.83)	46.92 (12.57)	133.26 (17.88)	185.69 (9.49)	269.24 (11.08)
T	130.82 (29.26)	41.24 (6.33)	248.48 (18.03)	-73.19 (12.73)	171.44 (12.25)	51.62 (10.44)	123.00 (18.36)	201.98 (28.96)	257.17 (42.60)
G	151.09 (26.27)	42.56 (7.25)	263.63 (15.75)	-76.69 (14.62)	164.86 (16.29)	45.72 (12.21)	136.90 (12.77)	241.11 (44.44)	191.92 (57.66)
A	167.87 (30.49)	37.33 (6.99)	256.34 (14.36)	-70.06 (11.65)	164.27 (14.98)	43.28 (13.10)	141.31 (12.64)	204.36 (36.65)	247.95 (50.45)
C	143.68 (17.71)	44.15 (6.36)	257.07 (17.28)	-73.40 (14.28)	168.73 (17.33)	49.28 (10.59)	133.91 (12.95)	223.47 (44.68)	213.52 (54.86)
G	159.67 (36.44)	36.54 (6.45)	244.22 (15.60)	-65.39 (10.77)	175.65 (12.48)	45.89 (12.73)	135.38 (17.01)	186.93 (20.96)	260.11 (28.92)
T	134.93 (27.13)	38.96 (6.05)	242.44 (14.68)	-67.55 (10.44)	169.18 (10.83)	53.54 (9.82)	123.13 (16.64)	188.38 (10.93)	268.52 (15.36)
C	124.46 (43.59)	41.42 (7.01)	252.22 (22.57)	-73.35 (14.87)	167.23 (16.23)	52.42 (10.14)	123.59 (22.96)	220.15 (40.83)	217.82 (56.55)
A	156.33 (44.96)	36.18 (6.79)	247.39 (17.85)	-67.20 (11.86)	171.26 (12.49)	48.35 (14.18)	135.19 (17.67)	191.49 (26.10)	259.49 (36.80)
T	138.59 (23.56)	41.80 (6.67)	248.36 (17.22)	-67.46 (11.91)	166.72 (18.75)	53.14 (10.71)	128.14 (16.37)	203.89 (39.21)	235.17 (49.16)
C	132.71 (43.68)	38.05 (6.44)	235.73 (14.84)	-64.80 (9.91)	177.01 (11.15)	53.15 (10.37)	123.85 (20.11)	185.28 (9.31)	269.94 (11.90)
T	132.21 (31.38)	38.88 (6.31)	239.30 (15.26)			53.37 (10.18)	122.16 (17.93)		
Strand 2									
A	164.85 (32.67)	37.90 (7.07)	259.41 (16.05)	-72.38 (13.77)	168.67 (13.77)	49.82 (13.48)	140.53 (12.90)	203.13 (36.35)	243.90 (49.53)
G	150.96 (21.72)	42.17 (6.43)	251.32 (17.50)	-70.28 (13.50)	174.09 (16.88)	49.04 (11.64)	135.91 (13.24)	207.93 (43.36)	229.73 (51.71)
A	150.11 (40.58)	36.63 (6.23)	241.48 (15.32)	-65.26 (10.03)	175.38 (11.31)	48.41 (11.83)	130.42 (19.23)	184.86 (11.98)	267.48 (15.16)
T	130.83 (31.73)	40.55 (6.05)	245.89 (15.72)	-73.47 (11.90)	171.66 (12.02)	52.42 (10.10)	122.70 (18.02)	195.11 (15.77)	271.80 (23.52)
G	144.38 (18.52)	44.67 (6.07)	266.18 (16.12)	-79.62 (14.75)	162.87 (15.38)	46.17 (10.84)	136.11 (10.70)	258.85 (36.98)	166.37 (47.61)
A	174.28 (24.76)	35.96 (5.74)	253.52 (12.67)	-67.90 (10.48)	169.11 (10.71)	41.84 (13.11)	142.25 (12.03)	188.48 (17.03)	268.05 (22.42)
C	140.84 (18.21)	45.07 (5.94)	257.91 (17.20)	-73.61 (14.74)	169.34 (17.24)	49.97 (10.06)	132.99 (14.00)	228.84 (45.02)	204.23 (51.59)
G	168.01 (30.47)	35.49 (5.89)	244.63 (13.92)	-66.56 (10.97)	176.24 (11.54)	44.66 (13.27)	138.88 (14.17)	184.76 (11.89)	266.74 (14.30)
T	136.83 (27.50)	39.98 (6.66)	248.94 (17.96)	-69.96 (11.44)	164.27 (15.73)	51.80 (10.17)	126.23 (17.03)	202.17 (32.09)	251.72 (45.83)
C	126.56 (43.09)	41.81 (7.24)	253.35 (22.61)	-73.82 (14.72)	168.46 (15.38)	50.75 (10.85)	124.94 (22.75)	223.27 (42.89)	216.49 (59.12)
A	162.34 (34.40)	36.04 (6.23)	249.95 (15.00)	-67.67 (11.35)	171.59 (10.95)	48.01 (13.90)	136.60 (15.68)	188.09 (18.71)	264.60 (25.02)
T	138.32 (21.56)	43.35 (6.67)	253.24 (18.62)	-70.29 (13.10)	162.46 (19.90)	51.74 (10.63)	129.86 (15.72)	218.89 (45.51)	217.11 (56.86)
C	139.37 (36.92)	37.87 (6.35)	236.62 (13.93)	-65.18 (10.03)	178.10 (11.42)	52.11 (10.72)	126.60 (18.24)	185.26 (9.38)	268.79 (11.04)
T	133.65 (28.26)	39.81 (6.19)	240.17 (15.98)			52.41 (10.16)	123.09 (17.38)		

REFERENCES

Arnott, S., and D. W. L. Hukins. 1972. Optimized parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun.* 47:1504–1509.

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

Bertrand, H.-O., T. Ha-Duong, S. Fermandjian, and B. Hartmann. 1998. Flexibility of the B-DNA backbone: effects of local and neighbouring sequences on pyrimidine-purine steps. *Nucleic Acids Res.* 26:1261–1267.

Boelens, R., T. M. G. Köning, G. A. van der Marel, J. H. van Boom, and R. Kaptein. 1989. Iterative procedure for structure determination from proton-proton NOEs using a full relaxation matrix approach. Application to a DNA octamer. *J. Magn. Reson.* 82:290–308.

Chaoui, M., S. Derreumaux, O. Mauffret, and S. Fermandjian. 1999. An intrinsic curvature towards the minor groove in the cAMP-responsive element DNA found by combined NMR and molecular modelling studies. *Eur. J. Biochem.* 259:877–886.

Cheatham, T. E., P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16:845–862.

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

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

Cieplak, P., T. E. Cheatham, and P. A. Kollman. 1997. Molecular dynamics simulations find that 3′ phosphoramidate modified DNA duplexes undergo a B to A transition and normal DNA duplexes an A to B transition. *J. Am. Chem. Soc.* 119:6722–6730.

Conte, M. R., A. N. Lane, and G. Bloomberg. 1997. Solution structure of the ATF-2 recognition site and its interaction with the ATF-2 peptide. *Nucleic Acids Res.* 25:3808–3815.

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

Darden, T., D. York, and L. Pedersen. 1983. Particle mesh Ewald: an $N·\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

de Souza, O. N., and R. L. Ornstein. 1998. Inherent DNA curvature and flexibility correlate with TATA box functionality. *Biopolymers.* 46: 403–415.

Dickerson, R. E. 1998. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* 26:1906–1926.

Diekmann, S. 1987. Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. *Nucleic Acids Res.* 15:247–265.

El Antri, S., P. Bittoun, O. Mauffret, M. Monnot, O. Convert, E. Lescot, and S. Fermandjian. 1993. Effect of distortions in the phosphate backbone conformation of six related octanucleotide duplexes on CD and ^{31}P NMR spectra. *Biochemistry.* 32:7079–7088.

Ellenberger, T. E., C. J. Brandl, K. Struhl, and S. C. Harrison. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. *Cell.* 71:1223–1237.

EMBO Workshop. 1989. EMBO workshop. *EMBO J.* 8:1–4.

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

Gorenstein, D. G. 1994. Conformation and dynamics of DNA and protein-DNA complexes by ^{31}P NMR. *Chem. Rev.* 94:1315–1338.

Grzeskowiak, K., K. Yanagi, G. G. Privé, and R. E. Dickerson. 1991. The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G. *J. Biol. Chem.* 266:8861–8883.

Hai, T., and T. Curran. 1991. Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. *Proc. Natl. Acad. Sci. USA.* 88:3720–3724.

Hartmann, B., D. Piazzola, and R. Lavery. 1993. BI-BII transitions in B-DNA. *Nucleic Acids Res.* 21:561–568.

Hockings, S. C., J. D. Kahn, and D. M. Crothers. 1998. Characterization of the ATF/CREB site and its complex with GCN4. *Proc. Natl. Acad. Sci. USA.* 95:1410–1415.

Hurst, H. C. 1994. Transcription factors 1: bZIP proteins. *Protein Profile.* 1:123–168.

Jorgensen, W. L., J. Chandrasekhar, and J. D. Madura. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Keller, W., P. König, and T. J. Richmond. 1995. Crystal structure of a bZIP/DNA complex at 2.2 Å: determinants of DNA specific recognition. *J. Mol. Biol.* 254:657–667.

Kim, J., and K. Struhl. 1995. Determinants of half-site spacing preferences that distinguish AP-1 and ATF/CREB bZIP domains. *Nucleic Acids Res.* 23:2531–2537.

Lavery, R. 1988. Junctions and bends in nucleic acids: a new theoretical modeling approach. *In* Structure and Expression, Vol. 3, DNA Bending and Curvature. W. K. Olson, M. H. Sarma, and M. Sundaralingam, editors. Adenine Press, New York. 191–211.

Lavery, R., I. Parker, and J. Kendrick. 1986a. A general approach to the optimization of the conformation of ring molecules with an application to valinomycin. *J. Biomol. Struct. Dyn.* 4:443–461.

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

Lavery, R., H. Sklenar, K. Zakrzewska, and B. Pullman. 1986b. The flexibility of the nucleic acids. II. The calculation of internal energy and applications to mononucleotide repeat DNA. *J. Biomol. Struct. Dyn.* 3:989–1014.

Lavery, R., K. Zakrzewska, and A. Pullman. 1984. Optimized monopole expansions for the representation of the electrostatic properties of the nucleic acids. *J. Comp. Chem.* 5:363–373.

Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimisation of nucleic acids). *Comp. Phys. Comm.* 91:135–158.

Lefebvre, A., S. Fermandjian, and B. Hartmann. 1997. Sensitivity of NMR internucleotide distances to B-DNA conformation: underlying mechanics. *Nucleic Acids Res.* 25:3855–3862.

Lefebvre, A., O. Mauffret, B. Hartmann, E. Lescot, and S. Fermandjian. 1995. The structural behaviour of the CpG step in two related oligonucleotides reflects its malleability in solution. *Biochemistry.* 34: 12019–12028.

Lefebvre, A., O. Mauffret., E. Lescot, B. Hartmann, and S. Fermandjian. 1996. Solution structure of the CpG containing d(CTTCGAAG)$_2$ oligonucleotide: NMR data and energy calculations are compatible with a BI/BII equilibrium at CpG. *Biochemistry.* 35:12560–12569.

Leonard, D. A., and T. K. Kerppola. 1998. DNA bending determines Fos-Jun heterodimer orientation. *Nature Struct. Biol.* 5:877–881.

Lyubartsev, A. P., and A. Laaksonen. 1998. Molecular dynamics simulations of DNA in solution with different counter-ions. *J. Biomol. Struct. Dyn.* 16:579–592.

McFail-Isom, L., X. Shui, and L. D. Williams. 1998. Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base π systems. *Biochemistry.* 37:17105–17111.

Paolella, D. N., Y. Liu, M. A. Fabian, and A. Schepartz. 1997. Electrostatic mechanism for DNA bending by bZIP proteins. *Biochemistry.* 36: 10033–10038.

Paolella, D. N., C. R. Palmer, and A. Schepartz. 1994. DNA targets for certain bZIP proteins distinguished by an intrinsic bend. *Science.* 264: 1130–1133.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* 91:1–41.

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

Rouzina, I., and V. A. Bloomfield. 1998. DNA bending by small, mobile multivalent cations. *Biophys. J.* 74:3152–3164.

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

Ryseck, R.-P., and R. Bravo. 1991. c-JUN, JUN B, and JUN D differ in their binding affinities to AP-1 and CRE consensus sequences: effect of FOS proteins. *Oncogene.* 6:533–542.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer Verlag, New York.

Shui, X., C. C. Sines, L. McFail-Isom, D. VanDerveer, and L. D. Williams. 1998. Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry.* 37:16877–16887.

Sitlani, A., and D. M. Crothers. 1998. DNA-binding domains of Fos and Jun do not induce DNA curvature: an investigation with solution and gel methods. *Proc. Natl. Acad. Sci. USA.* 95:1404–1409.

Sloan, L. S., and A. Schepartz. 1998. Sequence determinants of the intrinsic bend in the cyclic AMP response element. *Biochemistry.* 37:7113–7118.

Sprous, D., M. A. Young, and D. L. Beveridge. 1999. Molecular dynamics studies of axis bending in d(G$_5$-(GA$_4$T$_4$C)$_2$-C$_5$) and d(G$_5$-(GT$_4$A$_4$C)$_2$-C$_5$): effects of sequence polarity on DNA curvature. *J. Mol. Biol.* 285:1623–1632.

Suzuki, M., and N. Yagi. 1995. Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res.* 23:2083–2091.

Talanian, R. V., J. C. McKnight, R. Rutkowski, and P. S. Kim. 1992. Minimum length of a sequence-specific DNA binding peptide. *Biochemistry.* 31:6871–6875.

Tereshko, V., G. Minasov, and M. Egli. 1999. A "hydrat-ion" spine in a B-DNA minor groove. *J. Am. Chem. Soc.* 121:3590–3595.

Tisné, C., E. Hantz, B. Hartmann, and M. Delepierre. 1998. Solution structure of a non-palindromic 16 base-pair DNA related to the HIV-1 κB site: evidence for BI-BII equilibrium inducing a global dynamic curvature of the duplex. *J. Mol. Biol.* 279:127–142.

Ulyanov, N. B., and T. L. James. 1995. Statistical analysis of DNA duplex structural features. *Methods Enzymol.* 261:90–115.

Winger, R. H., K. R. Liedl, S. Rüdisser, A. Pichler, A. Hallbrucker, and E. Mayer. 1998. B-DNA's BI → BII conformer substate dynamics is coupled with water migration. *J. Phys. Chem. B.* 102:8934–8940.

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

Young, M. A., B. Jayaram, and D. L. Beveridge. 1997a. Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.* 119:59–69.

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

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