# Molecular Dynamics Simulation of a Zif268-DNA Complex in Water. Spatial Patterns and Fluctuations Sensed from a Nanosecond Trajectory

**G. Roxström, I. Velázquez, M. Paulino,† and O. Tapia***

*Department of Physical Chemistry, Uppsala University, Box 532, S-85121 Uppsala, Sweden*

*Received: June 27, 1997; In Final Form: September 19, 1997*

In this paper, with the help of a nanosecond-long molecular dynamics trajectory, we show that a fully charged protein−DNA system (Zif268-DNA) in a water environment can be simulated with a general force field (GROMOS) if solvation and counterion effects are appropriately represented. A model exhibiting ionic-atmosphere effects on mobile counterions (Tapia, O.; Velázquez, I. *J. Am. Chem. Soc.* **1997**, *119*, 5934−5938) was implemented. The root-mean-square deviations (rmsds) with respect to X-ray structure for the full complex, the protein, and the 12-base pair consensus sequence were 2.0, 1.95, and 1.35 Å, respectively, while the counterions displayed an rmsd from the initially equilibrated position of 1.2 Å. The mean-square fluctuation with respect to the average structure correlated with temperature factors for the protein and DNA; the agreement in trend is good. The results show that GROMOS87 force field with an appropriate representation of coion (ion atmosphere) effects on the counterions, and corrections for hydrophobicity default in the water to solute $C_{12}(\mathrm{Ow,Ow})^{1/2}$ parameter, previously used for proteins (Daura X.; Oliva, B.; Querol, E.; Avilés, F. X.; Tapia, O. *Proteins* **1996**, *25*, 89−103), are sufficient to obtain a fairly realistic simulation of this protein−DNA complex.

## Introduction

Molecular dynamics simulations of protein−DNA complexes may help in understanding dynamic aspects of the specificity and recognition mechanisms. This is an important problem in theoretical physical chemistry having much relevance in molecular biology. However, the highly charged nature of DNA and some bias in the force fields[1] have prevented simulations from leading to stable structures having a nanosecond lifetime.[2] The problem was compounded with difficulties in representing ionic atmosphere effects on the simulated counterion subsystem.[3] These factors have a bearing on the spatial fluctuation pattern of counterions contributing to the stabilization of the polyelectrolyte. In earlier work by other authors, screening of the phosphate charges and ad hoc hydrogen-bonded potential functions between base pairs have customarily been used to ease such instabilities.[4,5] Since an important part of the DNA−protein recognition and stability factors may depend on the Coulomb interaction, it is desirable to represent the fully charged system and eliminate any bias in the force field. Following this approach, GROMOS87 was recently used in a successful simulation of a fully charged 10-base pair (10 bp) DNA[6] where corrections to the water−solute interactions and a symmetrized counterion model were sufficient to render a molecule stable for more than 1 ns. This work extended the range of previous DNA simulations carried out by van Gunsteren and co-workers.[6] In a different spirit, recent versions of reparametrized AMBER and CHARMM packages have been successfully used to run nanosecond range trajectories with DNA in water by using state-of-the-art parametrizations and Ewald techniques to cope with the electrostatic problem.[8−13] However, no trajectory in the nanosecond time scale has been reported yet for fully charged protein−DNA complexes. This situation is overcome in this paper. The method proposed here to handle a protein−DNA complex is an alternative to previous ones in the sense that no Ewald summation technique is used for the electrostatics, and counterions related to phosphates not screened by the protein are incorporated with initial positions derived from the free DNA simulation.[6] All atoms are treated on the same basis with GROMOS87 force field;[14] a significant element is the use of a solute−water modified model that represents solvation effects of hydrophobic atoms.[6,16,17] The zinc finger complex Zif268-DNA[15] is used now to examine the hypothesis that the protein and the DNA can be treated with a unique force field avoiding ad hoc potential terms.

Zinc fingers constitute important eukaryotic DNA-binding domains, being present in many transcription factors.[18,19] The specific contacts determining the complex stability have been directly guessed from the crystal structure.[15,20] Interestingly, crystals are obtained with a consensus 10 bp only if dangling complementary bases are built at the 5′-ends of the double helix.[20] In this manner, the DNA may form long fibers that apparently prompt for crystallization. This experimental trick may or may not affect the details of the protein−DNA complex fluctuations. The molecular dynamics study of the crystal motif without dangling residues amounts to considering the fingers with their corresponding base pair subsites as a model functional unit.

In this paper, with the help of a nanosecond long molecular dynamics trajectory, we show that a protein−DNA fully charged system (Zif268-DNA) in a water environment can be simulated with a general force field previously used to represent proteins, if solvation and counterion effects are appropriately represented. Attention is focused on the complex and the protein behavior. While the results indicate that it is not necessary to reparametrize

† Permanent address: Department of Quantum Chemistry, Facultad de Química, Universidad de la República, 11800 Montevideo, Uruguay.

* To whom correspondence should be addressed.

S1089-5647(97)02093-2 CCC: $15.00 
Published on Web 02/14/1998

GROMOS in order to obtain meaningful simulations, they also point to the necessity for a more careful representation of surrounding medium effects in general.

## Method, Model, and Computational Details

**Method.** Hydrophobic effects are among the most important interactions relevant to the stability of protein-ligand complexes.[21−23] In recent investigations, Daura et al.[17] tested a scaling of the repulsive oxygen−solute term of the otherwise extensively used SPC model of water and standard GROMOS87 force field (the parameters are consistent now with the recently delivered GROMOS96[24]). The early model of water solvated everything to an extent that, for instance, methane was more stable in water than in nonpolar, reputedly hydrophobic, solvent. The water−solute parameter $C_{12}(\mathrm{Ow,Ow})^{1/2}$ is responsible for the repulsive interactions in the Lennard-Jones potential, and several authors have found an increased stability of proteins in water measured as persistence time of the corresponding fold if the value of this parameter is set to 793 kcal mol$^{-1}$ Å$^{12}$.[16,17] It is apparent that this solvent modified potential simulates the hydrophobicity aspect in a better way than the old one.[6,16]

The change of one parameter in a multiparameter problem with large cross correlations cannot be made by ignoring the interdependences. Such statement is not pertinent in GROMOS. To see this, note that the nonbonded van der Waals parameters $C_{12}(i,j)$ and $C_6(i,j)$, entering into the potential function $V(r_{ij}) = C_{12}(i,j)/r_{ij}^{12} - C_6(i,j)/r_{ij}^6$, are defined and stored in GROMOS force field in a matrix form; namely, there may exist different parameters for an atom engaged in different pair interactions. This implies that a change of a single pair parameter *does not* change any other pair. If one were to use geometric mean algorithms to calculate pair interactions such as $C_{12}(i,j) = C_{12}(i,i)^{1/2}C_{12}(j,j)^{1/2}$, with one parameter for each atom, then changing a single atom parameter would change a whole range of pair parameters.

We note that, with the above description on how GROMOS handles the Lennard-Jones coefficients, the water−water parameter in the SPC model is not changed. Similarly, the parameters entering in the description of protein−protein interactions are not changed by the correction of the solvation parameter. The testing of the ability of the standard parameter set to describe the DNA−protein complex in water will provide new physical insights in the schemes used up to now to treat these systems.

**Model System.** Zif268 protein is also known as Krox-24, NGFI-A, and Egrl. The molecular model of the complex containing the three zinc finger from Zif268 was derived from the X-ray structure of Pavletich and Pabo, solved at 2.1 Å resolution and refined to an *R* factor of 18.3%. The zinc finger domain from Zif268 has 85 residues. In the tertiary structure derived from the X-ray data, the three tandemly repeated zinc fingers are arranged in a semicircular (C-shaped) structure that fits into the major groove of DNA. Each zinc finger domain consists of an antiparallel $\beta$ sheet and an $\alpha$ helix, held together by a zinc ion and by a set of hydrophobic residues. Finger 1 spans over residues 1−28 and the linker between finger 1 and 2 extends over residues 29−32 (linker 1−2). Finger 2 consists of residues 33−56, the linker 2−3 is defined by residues 57−60, and finger 3 consists of residues 63−85.

For each finger, the zinc ion was constrained to its coordination sphere as usual in this type of simulation (no attempt is made at simulating the diffusion to or from this binding site). Equilibrium bond distances and force constants were taken from ab initio calculations[25] carried out at a Hartree−Fock level of theory. Zn is bonded to two S and two N atoms from Cys and His, respectively. We assumed zinc to have a charge of +2, with two deprotonated Cys residues neutralizing the complex.

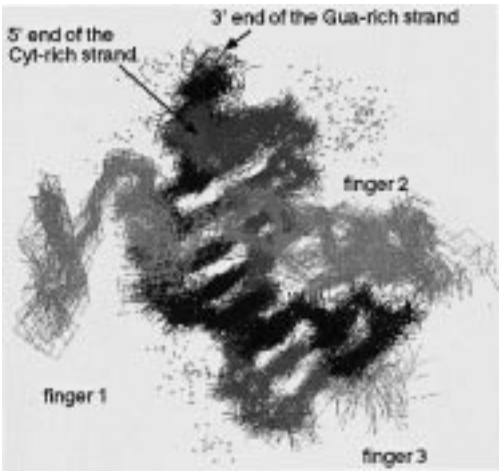


**Figure 1.** 55 frames of atom-distance display representing the 1 ns trajectory. Plot made with TOM-mdFRODO (for mdFRODO package, see: Nilsson, O. *J. Mol. Graph.* **1990**, *8*, 192). The time history fills the three-dimensional space producing a clear picture of the DNA and the three-finger structure. The water molecules are not plotted to help the viewing. Fingers 1 and 2 and fingers 2 and 3 are bridged by a short stretch on amino acid referred to as linkers (cf. Figure 2c). They connect the $\alpha$-helix to the corresponding $\beta$-strand.

The dangling bases appearing in the PDB file were taken away. The X-ray coordinates were then thoroughly minimized after embedding the system in a box of 5060 SPC-water molecules.

A zero net charge model was constructed with 454 DNA and 916 protein atoms including polar hydrogens for both molecules. Eight fully charged phosphate groups that were not screened by protein contacts were compensated with eight $Na^+$ ions. An additional four $Cl^-$ ions were distributed with PROION[14,24] to neutralize the global protein system.

**Computing Details.** The box length is 70.2 Å with octahedral periodic boundary conditions. The electrostatic interaction was calculated with cutoff radius of 1.2 nm. The van der Waals interactions were evaluated with a cutoff radius of 0.8 nm. The time step was of 0.002 ps. The number of pairs was updated every 10 steps. Coulomb interactions between 0.8 and 1.2 nm were updated each 20 fs. The cutoff radii were applied to geometric centres of neutral atomic groups or charged groups in the protein. SHAKE was used for constraining bond distances. A weak coupling to a Berendsen bath was used ($\tau = 0.1$). The solute optimization procedure started with 1000 steps of steepest descent energy minimization and continued with 500 steps of conjugate gradient energy minimization. The electrostatic factor, represented by the salt effects not included in the motif (mostly affecting DNA)[26] was represented as follows:[6] The equilibrated positions of counterions found for the DNA simulation were used in starting the calculations, and after the minimization procedure all counterions were weakly constrained to these relaxed positions (force constant of 0.58 kcal mol$^{-1}$ Å$^{-2}$).

A 1.1 ns run was performed with no restraints on the solute. The global system converged to a constant temperature of 307 K.

In Figure 1 are plotted the atom fluctuations for the 1 ns trajectory. One can see a convergent structure of the whole

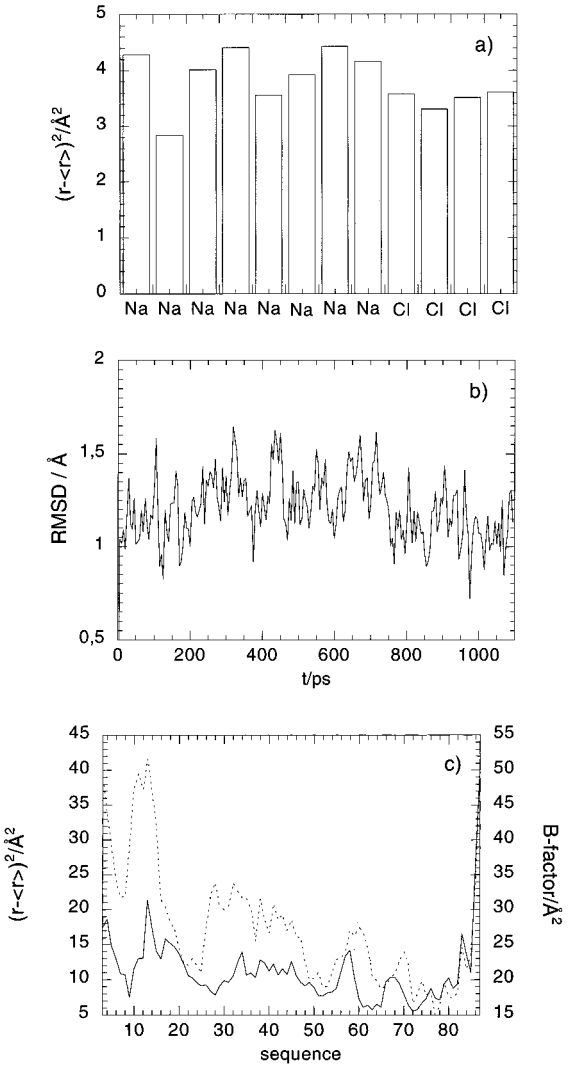


**Figure 2.** (a) Mean-square fluctuations of the counterions. (b) Root-mean-square deviation from the intitial position for all counterions. The average displacement is l.l Å. (c) Crystallographic temperature factors (broken lines) of the α-carbons for Zif268 obtained from the PDB file pdb1aay.ent and mean-square fluctuations (continuous line) from the 1 ns trajectory. The calculation represents the global fluctuation around the average structure without special least square fittings. The dotted regions represent $\beta$-sheets. The shaded zones are α-helices.

complex. This picture gives a first impression of the quality of the present model.

## Counterion Behavior

In Figure 2a the mean square deviations from the average position are plotted for all counterions. This "temperature factor" indicates that the ions are not tightly bound to a given position.

In Figure 2b is plotted the rmsd from the initial position taken as an average over all cations as a function of time. The ions, on average, move about 1 Å away from their initial positions. They actually fluctuate around their average positions with relatively large amplitudes.

The energy required to move an atom 1 Å away from its constrained location is about 0.3 kcal mol$^{-1}$ per constraint. This is a reasonable amount of energy to be payed for having a relatively stable and fairly diffuse ion distribution.

There is then a tradeoff between the solute forces attracting the counterions and the "cage" forces simulating the external ionic atmosphere effect. Neither the counterions nor the solute become frozen. In this sense, the counterion model may be considered as reasonably good so that the spatial patterns and fluctuations of the DNA-Zif268 may be an acceptable representation of this system in solution.

## Zif268-DNA Structural and Dynamical Characteristics

The quality of the counterion fluctuation patterns suggests that the model used to mimick ion atmosphere (coions) effects may produce a reasonable description of the complex.

The mark of the fluctuation pattern can be measured by comparing to experimental B-factors. In Figure 2c are plotted the mean-square fluctuations of the α carbons obtained from the 1 ns trajectory and the B-factors reported by Elrod-Eriksson et al. for the 1.6 Å resolution structure.[20] The trends agree quite well. Note that the B-factors reflect large atom fluctuations in average. Our calculations yield a similar picture. If one performs a least-squares fitting of the data to the average structure, then agreement between both data sets is very good in most regions, except in linkers which have larger fluctuations in the simulation. Such results are not unexpected as the one-molecule simulation does not include crystal contacts interactions. All in all, the fluctuation pattern (without numerical filtering) can be considered as reliable when compared to the experimental results.

The interesting point is that the fluctuation patterns actually define an excellent average structure, as suggested by Figure 1. All the simulation indicators are consistent with an almost equilibrated trajectory.

The rmsd for the DNA and Zif268 taken separately show their structures behaving somewhat differently. DNA is extremely stable; although it sometimes presents fluctuations attaining the region of 1.5 Å, its average rmsd is 1.35 Å. This result reflects, of course, the tight interactions between DNA with the protein since the free DNA simulation shows a larger rmsd. The protein, on the other hand, shows a change of conformation that became apparent beyond the 850 ps point.

The Zif268-DNA principle moments of inertia (PMI) as functions of time are displayed in Figure 3. The global structure presents a moderately cylindric mass distribution. This is hardly surprising as DNA imposes its form on the complex. The shape is conserved all along the trajectory. The almost coincident two largest PMI display an in-phase fluctuation. They too signal a uniform mass distribution perpendicular to the DNA axis.

The time series (cf. Figure 3b) formed with the PMIs for the protein alone show a stable global shape. Note that the two higher PMIs fluctuate in phase. The mass distribution is inhomogeneous in space in agreement with its C-shape form.

**Individual Fingers.** In Figure 4 are displayed the PMI time series for each finger. The effect of DNA on the individual fingers is partly reflected by a change in the PMI patterns. After less than 100 ps of equilibration, finger 2 presents a fairly stationary fluctuation. The size of the PMI reflects the shape of an elongated narrow parallelepiped: The α-helix stacks on top the $\beta$-sheet so that the thin dimension coincides with the α-helix's width. This shape is fairly stable along the 1 ns trajectory. Finger 3 possesses fewer atoms. The shape leans towards the shape of finger 2, especially after 600 ps.

Finger 1 presents a completely different behavior. It starts from a different conformation of the $\beta$-sheet when compared to fingers 2 and 3. During the first 400 ps it presents a relatively stable spatial fluctuation pattern. A regime of internal change of conformation is detected thereafter. The PMI converges toward a shape that looks like the one obtained for finger 2.

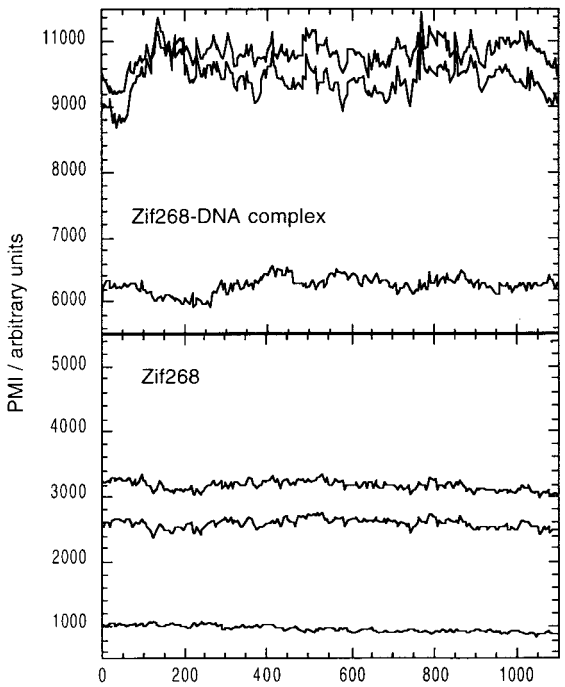


**Figure 3.** Principle moments of inertia as a function of time calculated for the Zif268-DNA complex and the protein alone. Time in picoseconds.

The internal structure of finger 1 in the crystal seems to be off equilibrium with respect to the present simulation in water. Finger 2 looks as having a well equilibrated internal structure. Finger 3 looks more similar to finger 2 but it also presents some structural reorganization as the trajectory proceeds.

For linkers, the atom-distance representation (not shown) indicates the emergence of a well defined structure. The linker has then a certain degree of "rigid body" behavior giving it the appearance of an arm. The end points are flexible.

## Discussion

MD simulations of the zinc finger domain from the Zif268 protein, in a complex with a consensus binding DNA sequence, yields a globally stable system with small yet significant internal readjustments. The results confirm the quality of the standard GROMOS87 force field with a scaled solute−water interaction (not altering the SPC water−water interactions) to simulate a fully charged DNA−protein system. The simulations of DNA alone, with the same force field, showed the important role played by the solvent and the symmetry of the counterion distribution.[6] In the present work, this feature has been retained when appropriate. The results of rmsd and B-factors illustrate the quality of the present simulation scheme.

The counterions, when treated as fully free ions around the protein−DNA complex, may approach to very short distances, producing structural perturbations. The effect would then be artifactual since in a solution the ionic atmosphere,[21,26] external to the simulated motif, would produce a force opposing the implosion. Recently, Ewald techniques have been used to represent such external (to the motif) effects.[3,8,27,28] The result is a sort of lowering (buffering as it were) of the electrostatic interactions, yielding more stable counterion fluctuation patterns. We obtain useful results with an alternative model that avoids freezing the ion positions so that the internal structure may fluctuate within reasonable bounds. The matching in trends of theoretical mean square fluctuation to experimental B-factors is satisfactory. It is worth noticing that the experimental values of B-factors for the protein are large. This may signal the large flexibility retained by Zif268 in the crystal.

The internal structure of the three fingers, while stable as such, have subtle but definite differences among themselves. Finger 1 presents a clear dissimilarity with respect to the other two. The persistence time of the DNA−protein contacts is shorter for this finger. The other two display recognition interactions whose persistence is unabated by thermal fluctuations. According to the present simulation, finger 2 and finger 3 might be the most conserved elements in the recognition of DNA. One may conjecture that the special form of DNA in the crystal may have molded finger 1 into a conformation that is not permanently used in solution as the co-crystals are obtained by forming long DNA fibers. A longer trajectory may help sense further changes and/or recurrences. For this type of system, a trajectory of 1000 ps may still be too short to sense subtle dynamical changes.

A further advantage of an equal basis representation of the protein and DNA atoms is to avoid ad hoc potential energy functions to represent hydrogen bonding. Base flipping, as found in DNA methyltransferase,[29−31] can be more naturally simulated in the present context. One of the advantages of the GROMOS force field, in so far as H-bonding is concerned, resides in its simple electrostatic representation. Further testing of the present model with other systems would be necessary to fully validate its usefulness.

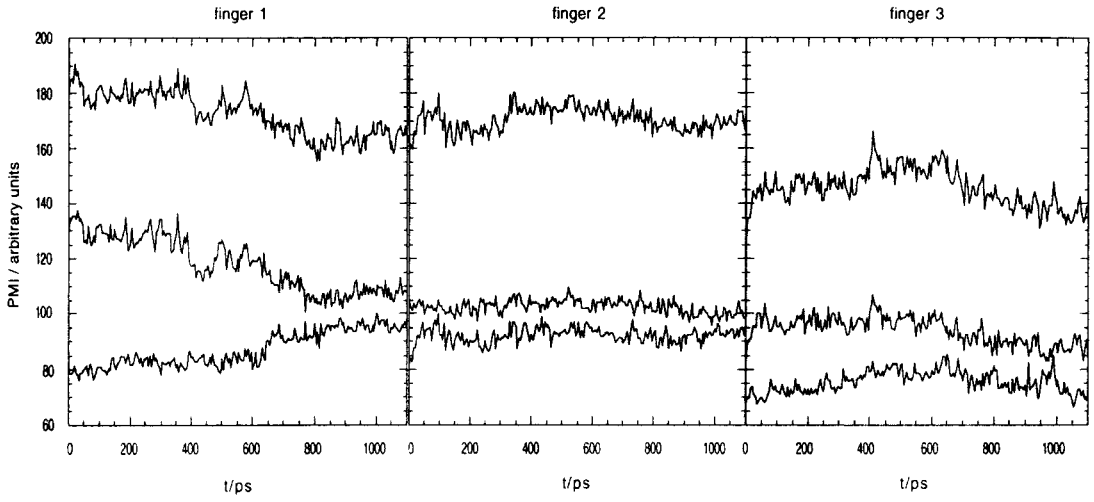


**Figure 4.** Time series for the principle moments of inertia calculated for the three fingers of Zif268. Time in picoseconds.

**Acknowledgment.** The authors are indebted to Claudio Scazocchio (Paris) for introducing them to the subject of DNA−protein interactions. M. Paulino is most grateful to Alan Mark for initiating her into the mysteries of GROMOS. The authors are most grateful for the constructive comments by the referees. O.T. is grateful to the NFR for financial support.

## References and Notes

(1) Harvey, S. C. *Proteins* **1989**, *5*, 78−92.

(2) Nilsson, L. In *Encyclopedia of Computational Chemistry*; Scheleyer, P. von R., Allinger, N. L., Clark, T., Gasteiger, P. A., Kollman, P. A., Schaefer, H. F., III, Eds.; Wiley: New York, 1997.

(3) Louise-May, S.; Auffinger, P.; Westhof, E. *Current Op. in Struct. Biol.* **1996**, *6*, 289−298.

(4) van Gunsteren, W. F.; Berendsen, H. J. C.; Geurtsen, R. G.; Zwinderman, H. R. L. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 287−303.

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

(6) Tapia, O.; Velázquez, I. *J. Am. Chem. Soc.* **1997**, *119*, 5934−5938.

(7) de Vlieg, J.; Berendsen, H. J. C.; van Gunsteren, W. F. *Proteins* **1989**, *6*, 104−127.

(8) Miaskiewicz, K.; Osman, R.; Weinstein, H. *J. Am. Chem. Soc.* **1993**, *115*, 1526−1537.

(9) McConnell, K. J.; Nirmala, R.; Young, M. A.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 4461−4462.

(10) Cheatham, I. T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.

(11) York, D. M.; Yand, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.

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

(13) Yang, L.; Pettit, B. M. *J. Phys. Chem.* **1996**, *100*, 2564−2566.

(14) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual;* BIOMOS B.V.: Nijenborgh 16, Groningen, The Netherlands, 1987.

(15) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809−817.

(16) Daura, X.; Oliva, B.; Querol, E.; Avilés, F. X.; Tapia, O. *Proteins* **1996**, *25*, 89−103.

(17) Daura, X.; Hünenberger, P. H.; Mark, A. E.; Querol, E.; Avilés, F. X.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1996**, *118*, 6285−6294. Also see: Mark, A. E.; van Helden, S. P.; Smith, P. E.; Janssen, L. H. M.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 6293−6302.

(18) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, *61*, 1053−1095.

(19) Goldman, A.; Glumoff, T. In *Nucleic Acids in Chemistry and Biology*; Blackburn, G. M., Gait, M. J., Eds.; Oxford University Press: Oxford, 1996.

(20) Elrod-Eriksson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171−1180.

(21) Richards, E. G. *An Introduction to Physical Properties of Large Molecules in Solution*; Cambridge University Press: Cambridge, 1980.

(22) Saenger, W. *Principles of Nucleic Acid Structure;* Springer-Verlag: New York, 1984.

(23) Israelachvili, J.; Wennerström, H. *Science* **1996**, *379*, 219−225.

(24) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. In *Groningen Molecular Simulation (GROMOS) Library Manual*; BIOMOS B.V.: Zürich, 1996; pp II36.

(25) Tapia, O.; Cardenas, R.; Andres, J.; Krechl, J.; Campillo, M.; Colonna-Cesari, F. *Int. J. Quantum Chem.* **1991**, *39*, 767−786.

(26) Vorontsov-Velyaminov, P. N.; Lyubartsev, A. P. *Mol. Simul.* **1992**, *9*, 285−306.

(27) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1996**, *105*, 4289−4293.

(28) Weerasinghe, S.; Smith, P. E.; Mohan, V.; Cheng, Y.-K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 2147−2158.

(29) Cheng, X.; Blumenthal, R. M. *Structure* **1996**, *6*, 639−645.

(30) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *Nature* **1996**, *384*, 87−92.

(31) Vassylyev, D. G.; Morikawa, K. *Structure* **1996**, *4*, 1381−1385.