Simulated Annealing of Chemical Potential: A General Procedure for Locating Bound Waters. Application to the Study of the Differential Hydration Propensities of the Major and Minor Grooves of DNA

Frank Guarnieri* and Mihaly Mezei

Department of Physiology and Biophysics
Mount Sinai School of Medicine
New York, New York 10029

Received May 3, 1996
Revised Manuscript Received July 15, 1996

A wide variety of experimental[1−9] and computational[10−19] studies are focused on the crucial role that water plays in DNA and protein architecture, and in many DNA and protein functions.[20−22] Chalikian et al.[4] point out that despite the successes of these efforts, many ambiguities about the role of water exist because different methodologies measure different properties, and thus have generally led to different conclusions. Therefore, various aspects of DNA and protein hydration remain unsettled. In this communication we demonstrate how the method of *simulated annealing of chemical potential* allows bulk waters to be distinguished from bound waters, and how differentially bound waters may be distinguished from each other based on their relative chemical potentials. This is illustrated by showing that it takes more free energy to desolvate the minor groove than the major groove of a charged DNA dodecamer.

Grand canonical ensemble simulations are generally performed by placing a molecule in a periodic simulation cell, setting a parameter B in such a way as to achieve an experimentally determined density, sampling potential hydration positions around the molecule by inserting and deleting water molecules from the simulation cell using a technique such as cavity-bias,[23,24] and accepting or rejecting the attempt based on a Metropolis Monte Carlo[25] criteria using a grand canonical ensemble probability function.[26] The parameter B is related to the excess chemical potential μ' as follows: $B = \mu'/kT + \ln\langle N\rangle$, where k is Boltzmann's constant, T is the absolute temperature, and $\langle N\rangle$ is the mean number of molecules. In the method of *simulated annealing of chemical potential*, the simulation is started with a large initial B-value so that a higher percentage of water insertion attempts are accepted. This causes the simulation cell to be flooded with water molecules. After this grand canonical ensemble simulation at high excess chemical potential is equilibrated, subsequent simulations are carried out at successively lower B-values. This causes a gradual removal of the bulk water molecules from the simulation cell. As the chemical potential is further "annealed", a point is reached at which water molecules do not readily leave the cell. At this point of the simulation, we have encountered the waters that are strongly influenced by the DNA, the so-called "bound water molecules". As the excess chemical potential is again lowered, ultimately some of these bound waters start to leave the cell. Since chemical potential is a free energy, this *simulated annealing of chemical potential* yields a numerical estimate of the differential free energy of binding of the different bound water molecules. It must be emphasized that our utilization of the term "annealing" applies strictly to the value of the chemical potential and that the *temperature is kept constant at 298 K* in all the simulations. For all simulations the DNA was held fixed, water molecules were added and deleted throughout all parts of the cell, extensive canonical Monte Carlo is performed between accepted grand canonical Monte Carlo steps, and periodic boundary conditions were used.

As an illustration of the method, we performed *simulated annealing of chemical potential* on the d(CGCGAATTCGCG)$_2$ dodecamer of Dickerson,[27] starting with $B = 1.0$ down to -26 in 37 increments performing 2 000 000 cavity-biased grand canonical ensemble Monte Carlo steps at each B-value. Figure 1 shows four final water configurations around the DNA at four successively lower B-values. Figure 1A, the final configuration of the simulation with $B = -6$, has 1120 water molecules. Figure 1B, the final configuration of the simulation with $B = -8$, has 533 water molecules. Figure 1C, the final configuration of the simulation with $B = -9$, has 390 water molecules. Figure 1D, the final configuration of the simulation with $B = -11$, has 215 water molecules. The most salient feature of this progression is the differential hydration of the major and minor groove of the DNA. Figure 1A shows the DNA essentially uniformly solvated. Figure 1B clearly shows that upon lowering of the chemical potential by 2 B-units, a majority of the nonbulk extracted waters come from the major groove, while the minor groove remains almost unaffected. Annealing the chemical potential further (Figure 1C) still leaves the minor groove well hydrated while the major groove is almost stripped. Lowering B even further (Figure 1D) results in the removal of almost all water molecules from both the major and minor groove. Quantitation of the hydration of the DNA as a function of chemical potential was computed by proximity analysis[28,29] with the results shown in Table 1. For $B = -6$, the first hydration shell (defined by the position of the first minimum of the radial distribution function) of the major and minor groove has a comparable density (0.012 and 0.013, respectively), while the second hydration shell of the minor groove has twice the density of the major groove. For $B = -8$ the hydration difference

(1) Eisenstein, M.; Shakked, Z. *J. Mol. Biol.* **1995**, *248*, 662−678.
(2) Brandes, R.; Rupprecht, A.; Kearns, D. R. *Biophys. J.* **1989**, *56*, 683−691.
(3) Lipscomb, L. A.; Peek, M. E.; Zhou, F. X.; Bertrand, J. A.; VanDerveer, D.; Williams, L. D. *Biochemistry* **1994**, *33*, 3649−3659.
(4) Chalikian, T. V.; Sarvazyan, A. P.; Plum, G. E.; Breslauer, K. J. *Biochemistry* **1994**, *33*, 2394−2401.
(5) Brunne, R. M.; Liepinsh, E.; Otting, G.; Wuthrich, K.; van Gunsteren, W. F. *J. Mol. Biol.* **1993**, *231*, 1040−1048.
(6) Dotsch, V.; Wider, G. *J. Am. Chem. Soc.* **1995**, *117*, 6064−6070.
(7) Denisov, V. P.; Halle, B. *J. Am. Chem. Soc.* **1994**, *116*, 10324−10325.
(8) Yao, J.; Bruschweiler, R.; Dyson, H. J.; Wright, P. E. *J. Am. Chem. Soc.* **1994**, *116*, 12051−12052.
(9) Otting, G.; Liepinsh, E. *Acc. Chem. Res.* **1995**, *28*, 171−177.
(10) Steinbach, P. J.; Brooks, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9135−9139.
(11) Veal, J. M.; Gao, X.; Brown, F. K. *J. Am. Chem. Soc.* **1993**, *115*, 7139−7145.
(12) McConnell, K. J.; Nirmala, R.; Young, M. A.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 4461−4462.
(13) Yang, L.; Pettitt, M. *J. Phys. Chem.* **1996**, *100*, 2564−2566.
(14) Miaskiewicz, K.; Osman, R.; Weinstein, H. *J. Am. Chem. Soc.* **1993**, *115*, 1526−1537.
(15) Forester, T. R.; McDonald, I. R. *Mol. Phys.* **1991**, *72*, 643−660.
(16) Beveridge, D. L.; Ravishanker, G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 246−255.
(17) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(18) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.
(19) Gu, W.; Schoenborn, B. P. *Proteins: Struct., Funct., Genet.* **1995**, *22*, 20−26.
(20) van Gunsteren, W. F.; Luque, F. J.; Timms, D.; Torda, A. E. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 847−863.
(21) Buckin, V. A.; Kankiya, B. I.; Rentzeperis, D.; Marky, L. A. *J. Am. Chem. Soc.* **1994**, *116*, 9423−9429.
(22) Berman, H. M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 345−350.
(23) Mezei, M. *Mol. Phys.* **1987**, *61*, 565−582.
(24) Resat, H.; Mezei, M. *J. Am. Chem. Soc.* **1994**, *116*, 7451−7452.
(25) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.
(26) Tolman, R. In *The Principles of Statistical Mechanics*; Dover Press: New York, 1971.
(27) Dickerson, R. E. *J. Mol. Biol.* **1981**, *149*, 761−786.
(28) Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102*, 4287.
(29) The effects of different partial charges on proximity analysis are described in: Mezei, M. *Mol. Simul.* **1988**, *1*, 327−332.

S0002-7863(96)01482-5 CCC: $12.00 © 1996 American Chemical Society

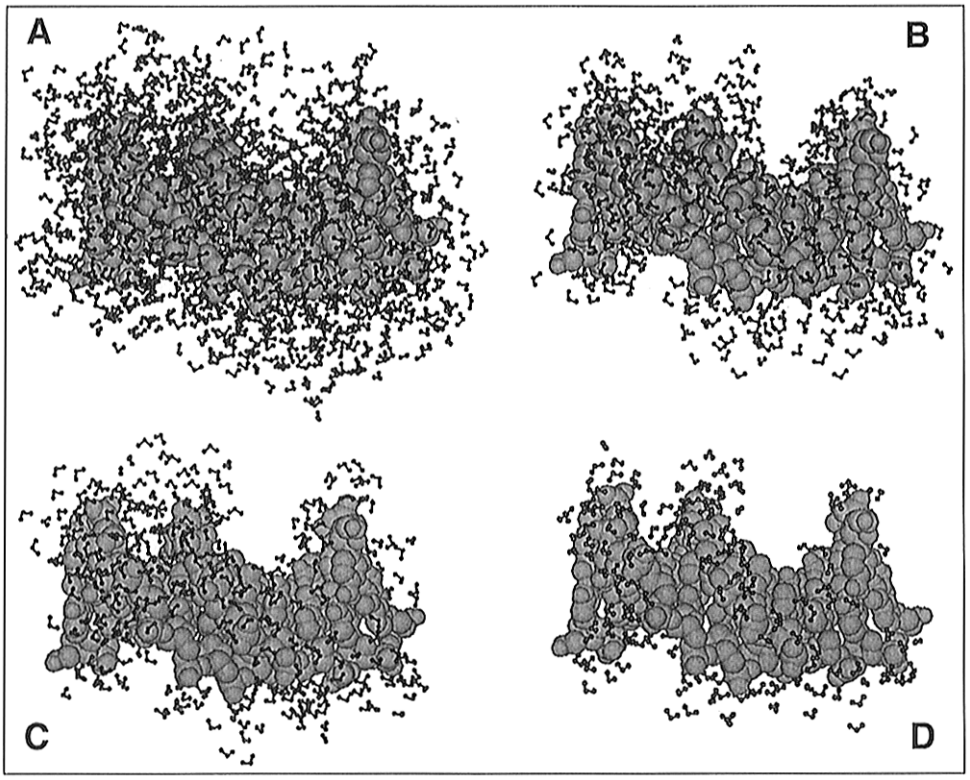

Figure 1. Final configurations of the simulations with $B = -6, -8, -9, -11$ containing 1120, 533, 390, and 215 water molecules, respectively.

Table 1. Hydration of the Major and Minor Groove of a DNA Dodecamer[a]

	first hydration shell				first and second hydration shell			
	minor groove		major groove		minor groove		major groove	
B	no. of waters	density	no. of waters	density	no. of waters	density	no. of waters	density
−6	7.27	0.013	13.23	0.012	21.3	0.021	41.7	0.011
−8	5.4	0.010	5.06	0.004	14.6	0.015	11.8	0.003
−9	4.08	0.007	4.36	0.004	11.5	0.011	9.7	0.003
−11	1.04	0.002	2.11	0.002	3.9	0.004	4.2	0.001

[a] Proximity analysis of the hydration patterns of the first and second solvation shells of a DNA dodecamer. The first block of data corresponding to the first four columns is the first hydration shell of the major and minor groove as a function of *B*. The second block of data corresponding to the second four columns is the first plus second hydration shell of the major and minor groove as a function of *B*. The column header no. of waters is the average number of waters over 4 000 000 steps. The volume of the first hydration shell of the major and minor groove is 1133 and 563 Å^3, respectively, and the volume of the first plus second hydration shell of the major and minor groove is 3779 and 1006 Å^3, respectively. The column header "density" is the number of waters divided by the volume.

becomes quite pronounced with the minor groove first and second shell hydration density being 2.5 fold and 5 fold higher than the major groove, respectively. For $B = -11$ the major and minor groove hydration density again becomes equal because at this value of the excess chemical potential both grooves are essentially stripped bare.

Illustrating the differential hydration propensities of the major and minor groove of DNA is computationally undemanding (3 days of CPU time to run one annealing schedule and 3 days of CPU time to run one proximity analysis[30] on an SGI Power Challenge) using *simulated annealing of chemical potential* because only a rather coarse "cooling" schedule of the chemical potential is required. Since the chemical potential is a free energy, a very fine cooling schedule may be used to estimate quantitatively the hydration free energy difference of two different functional groups or even two different atoms of the DNA. Two atoms that desolvate at the same *B*-value have similar solvation free energy, or alternatively, require a finer cooling schedule to resolve the differences. It should be noted that the model system used here consisted of ionic DNA with 22 negative charges and no sodium counterions. It is not known how the presence of sodium ions will affect the results. However, our findings about the preferential hydration of the minor groove corresponds very well to results from X-ray crystallographic[31−36] and NMR[37,38] studies. Possible reasons for the stronger binding of water molecules in the minor groove may include the following: the high density of the charged rows of phosphate groups, steric constraints, and specific water−water, water−DNA interactions. We are presently investigating the mechanisms responsible for the different hydration propensities of the major and minor grooves and several alternatives for incorporating sodium ion effects into grand canonical ensemble simulations.

Acknowledgment. This work was supported in part by a grant from the National Institute on Drug Abuse (Grant 5T32DA07135 to F.G.). The authors would like to thank Professor Harel Weinstein and Nina Pastor for helpful suggestions.

JA961482A

(30) Calculations of volume elements can be somewhat CPU intensive. The effects of volume element calculations on proximity analysis are described in Mezei, M.; Beveridge, D. L. In *Methods in Enzymology*; Packer, L., Ed.; Academic Press Inc.: New York, 1986; pp 21−47.

(31) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556.
(32) Quintana, J. R.; Grzeskowiak, K.; Yanagi, K.; Dickerson, R. E. *J. Mol. Biol.* **1992**, *225*, 379−395.
(33) Heinemann, U.; Alings, C. *J. Mol. Biol.* **1989**, *210*, 369−381.
(34) Heinemann, U.; Hahn, M. *J. Mol. Biol.* **1992**, *267*, 7332−7341.
(35) Baikalov, I.; Grzeskowiak, K.; Yanagi, K.; Qunitana, J.; Dickerson, R. E. *J. Mol. Biol.* **1993**, *231*, 768−784.
(36) Edwards, K. J.; Brown, D. G.; Spink, N.; Neidle, S. *J. Mol. Biol.* **1992**, *226*, 1161−1173.
(37) Liepinsh, E.; Otting, G.; Wuthrich, K. *Nucleic Acids Res.* **1992**, *20*, 6549−6553.
(38) Kubinec, M. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 8739−8740.