Tetrahedron Vol. 47, No. 14/15, pp. 2491-2501, 1991
Printed in Great Britain

0040-4020/91 $3.00+.00

Monte Carlo Simulations Yield Absolute Free Energies of Binding for Guanine–Cytosine and Adenine–Uracil Base Pairs in Chloroform

Julianto Pranata and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06511

(Received in USA 17 September 1990)

Abstract: Monte Carlo simulations with statistical perturbation theory and the OPLS potential functions were employed to calculate the absolute free energies of binding between 9-methyladenine (A) and 1-methyluracil (U) and between 9-methylguanine (G) and 1-methylcytosine (C). The results for both A–U (ca. −3.6 kcal/mol) and G–C (ca. −7.9 kcal/mol) compare well with the values from experimentally determined association constants for these systems. Several thermodynamic cycles were also considered that demonstrated the high precision of the methodology.

Introduction

During the past decade the use of Monte Carlo and molecular dynamics simulations as tools for studying chemical processes in solution has grown rapidly.[1,2] A major reason for the surge of activity is the advancement that has been made in the accurate calculation of free energies in solution,[3,4] which facilitates direct comparisons between theory and experiment. Applications of such calculations are numerous and diverse, with examples including free energies of solvation,[5] relative pK_a values[6] and partition coefficients,[7] conformational equilibria,[8] energetics of ion pairs,[9] binding in host–guest systems[10] and in biomolecules,[11] and chemical reactions.[12] In this paper we report the use of free energy calculations in Monte Carlo simulations to quantify the association of nucleotide base pairs in chloroform.

The hydrogen bonded base pairs between adenine and thymine or uracil and between guanine and cytosine are central elements in the structure of nucleic acids and in the processes of replication, transcription, and translation of the genetic code. Consequently, numerous investigators have studied the association of nucleotide bases using both experimental and theoretical approaches The experimental work has used a variety of methods, including infrared[13] and NMR spectroscopy,[14] mass spectrometry,[15] calorimetry,[16] and phase solubility methods.[17] The theoretical studies include the applications of various levels of multipole approximations,[18] ab initio calculations,[16a,19] and, of course, Monte Carlo and molecular dynamics simulations.[20,21] In previous work, we computed the difference in free energies of binding for 9-methylguanine with 1-methylcytosine (G–C) and 9-methyladenine with 1-methyluracil (A–U).[22] These studies are now extended to the calculation of the absolute free energies of binding for these base pairs This provides important tests of our methodology and intermolecular potential functions.

Methodology

Both Monte Carlo and molecular dynamics simulations rely on statistical mechanical averages to obtain equilibrium thermodynamic properties of a system. The system typically consists of one or two solute molecules and several hundred solvent molecules in a cubic or rectangular box. Periodic boundary conditions are invoked, in which the central box is surrounded in all directions by its images, in order to avoid boundary effects and to allow representation of a bulk fluid by a manageable number of molecules. In molecular dynamics, the time evolution of the system is traced by numerical integration of Newton's equations of motion, and time averages are used to calculate thermodynamic properties.[2] In Monte Carlo simulations, the averages are performed over millions of instantaneous geometrical configurations that are selected by the Metropolis algorithm,[23] occasionally augmented by umbrella sampling[24] (in which an artificial biasing function is added to the potential energy) and preferential sampling[25] (in which solvent molecules near the solutes are moved more frequently than those farther away).

Of central importance to the simulations is the intermolecular potential functions that describe the interactions between the components of the system. The OPLS potential functions from our laboratory represent the intermolecular interaction energies by Coulombic and Lennard–Jones terms between sites centered on the nuclei (eq. 1).[26] Thus, the interaction energy between molecules a and b is given by the

$$\Delta E_{ab} = \sum_{i}^{\text{on } a} \sum_{j}^{\text{on } b} \left(q_i q_j e^2 / r_{ij} + A_{ij}/r_{ij}^{12} - C_{ij}/r_{ij}^{6} \right) \tag{1}$$

sum of the interactions between sites i on a and sites j on b. The Lennard–Jones parameters A and C are related to the more familiar σ and ϵ by $A_{ii} = 4\epsilon_i\sigma_i^{12}$ and $C_{ii} = 4\epsilon_i\sigma_i^{6}$, and the combining rules are $A_{ij} = (A_{ii}A_{jj})^{1/2}$ and $C_{ij} = (C_{ii}C_{jj})^{1/2}$.

The OPLS parameters have been optimized to reproduce experimental thermodynamic and structural results for ca. 40 pure organic liquids and water.[26] Additional parameters, e.g., for ions, were developed largely by fitting to the results of ab initio calculations on ion–solvent molecule complexes.[6,27] Parameters for the nucleotide bases were developed in an analogous manner.[22]

For the present purposes, it is important to note that the *difference* in free energies between two systems can be obtained via statistical perturbation theory (eq. 2),[28] as long as the two systems are similar. Thus,

$$G_j - G_i = -k_B T \ \ln\langle \exp[-(H_j - H_i)/k_B T] \rangle_i \tag{2}$$

ΔG between systems i and j is expressed as an average of a function of their energy difference. The averaging is for sampling based on system i, so j is treated as the perturbed system. For systems that differ too much to be treated by a single perturbation, multiple simulations can be run to connect i and j through intermediate points using a coupling parameter, λ (eq. 3). Features ξ of the system, including the geometry and potential functions, can then be interconverted as λ goes from 0 to 1. Typically λ represents

$$\xi(\lambda) = \xi_i + \lambda(\xi_j - \xi_i) \tag{3}$$

the linear admixture of i and j, so $\lambda = 0.3$, for example, represents an intermediate system with 70% i character and 30% j character.

The calculation of absolute free energies of binding (ΔG_b) using Monte Carlo or molecular dynamics simulations can be done in several ways. One can calculate the free energy profile or "potential of mean force" (pmf), $w(r)$, as a function of intersolute separation. In this case the perturbation variable λ can be the separation of the centers-of-mass. The association constant (K_a) is obtained by integration of the PMF to a cutoff limit c that defines association (eq. 4),[29] and ΔG_b is given by $-k_B T \ln K_a$.

$$K_a = 4\pi \int_0^c r^2 \exp(-w(r)/k_B T)\, dr \tag{4}$$

This approach has been applied to the association of ion pairs in water,[9d] a pair of N-methylacetamide molecules in water and chloroform,[30] two N-methylformamide molecules in water and CCl_4,[31] the A–T base pair in water,[21b] the uracil–2,6-diaminopyridine (U–DAP) base pair in chloroform,[22b] and a crown ether–potassium ion complex in water.[10c] One difficulty with this method is that eq. 2 only gives free energy differences, so the pmf must be zeroed by some other means. Another difficulty is that eq. 4 requires complete orientational averaging for the solutes.

Another route to ΔG_b comes from consideration of the following thermodynamic cycle. Clearly, $\Delta G_b =$

$$\begin{array}{ccccc} & & \overset{\Delta G_1}{\longrightarrow} & & \\ \mathrm{E} + \mathrm{S} & & & \mathrm{E\text{–}S} & \textit{gas phase} \\ \Delta G_2 \downarrow & \downarrow \Delta G_3 & & \downarrow \Delta G_4 & \\ \mathrm{E} + \mathrm{S} & & \underset{\Delta G_b}{\longrightarrow} & \mathrm{E\text{–}S} & \textit{solution} \end{array}$$

$\Delta G_1 + \Delta G_4 - \Delta G_2 - \Delta G_3$. So the calculation requires a knowledge of the gas-phase binding free energy (ΔG_1) and the free energies of solvation for the individual solutes (ΔG_2 and ΔG_3) and the complex (ΔG_4). ΔG_1 can come from gas-phase calculations, while the other ΔG's can come from the differences in free energy changes from pairs of simulations in which the solute or complex is made to vanish in the gas phase and in solution.[32] This procedure was used by Cieplak and Kollman to obtain ΔG_b's for the base pairs A–T and G–C in water.[21a]

A less arduous procedure can be gleaned from the following relationship.[33] ΔG_b is obtained as $\Delta G_2 - \Delta G_1$. All that is needed are two simulations in solution, one in which S disappears by itself and one in which it disappears while bound to E. Note also that if contributions from internal degrees of freedom

$$\begin{array}{lll} \mathrm{E} \longrightarrow \mathrm{E\text{–}S} & -\Delta G_1 \\ \mathrm{S} \longrightarrow 0 & \Delta G_2 \\ \hline \mathrm{E} + \mathrm{S} \longrightarrow \mathrm{E\text{–}S} & \Delta G_b \end{array}$$

are neglected, $-\Delta G_2$ is simply the free energy of solvation for S. This procedure is called the "double annihilation method". It has been used to calculated ΔG_b for two methane molecules in water,[33] two N-methylformamide molecules in water and CCl_4,[31] and the binding of adenine to a molecular tweezer in chloroform.[10d] It is the method we have used for the present study. The calculations still require substantial computational effort since the disappearence of a solute must be done gradually, using numerous simulations at different values of λ.

Calculation of relative free energies of binding ($\Delta\Delta G_b$) are considerably less demanding, because the changes involved are usually much less severe. Consideration of the thermodynamic cycle below leads

$$\begin{array}{ccccc} \mathrm{E_1} + \mathrm{S_1} & & \xrightarrow{\Delta G_1} & & \mathrm{E_1\text{–}S_1} \\ \Delta G_3 \downarrow & & \downarrow \Delta G_4 & & \downarrow \Delta G_5 \\ \mathrm{E_2} + \mathrm{S_2} & & \xrightarrow[\Delta G_2]{} & & \mathrm{E_2\text{–}S_2} \end{array}$$

to $\Delta\Delta G_b \equiv \Delta G_1 - \Delta G_2 = \Delta G_3 + \Delta G_4 - \Delta G_5$. Many calculations of this type have been reported, with applications to host–guest systems[10a,b,d] and to enzyme–substrate[11a] and drug–DNA binding.[11b] Our earlier calculation of the relative binding free energy between G–C vs A–U also falls into this category.[22]

Computational Details

The absolute free energy of binding for each base pair (A–U and G–C) was calculated in two ways. The first consists of a double annihilation of the pyrimidines (U and C), while the second features double annihilation of the purines (A and G). Thus, a total of eight perturbation calculations were carried out. The Monte Carlo simulations were performed in the isothermal–isobaric ensemble at 25 °C and 1 atm. Annihilation of the individual bases was performed with 125 chloroform molecules as the solvent in a cubic box with dimensions ca. 26 × 26 × 26 Å^3. For the annihilation of a base from a base pair, the number of solvent molecules was 185 and the cell dimensions were ca. 26 × 26 × 39 Å^3. Simulations for the complexes were started with the base pairs in the Watson–Crick orientation. The bases were allowed to move independently of one another.

Intermolecular interaction energies were calculated using the OPLS potential functions. For the nu-

cleotide bases these are in an all-atom format, i.e., all atoms are explicitly represented including the hydrogens bound to carbons.[22] The geometries for the individual bases were the same as in the prior studies.[22] For chloroform, a four-site model is used with the four sites corresponding to the three chlorine atoms and a carbon atom with an implicit hydrogen.[7] The solvent–solvent interactions were truncated at a C–C separation of 12 Å and the solute–solvent interactions were included if the chloroform carbon atom was within 12 Å of any solute atom. A switching function was used to feather quadratically the interactions to zero over the last 0.5 Å.

Each simulation consisted of 0.5 – 1.0 × 10^6 configurations of equilibration followed by averaging over 2.0 × 10^6 configurations. Averages were computed separately for blocks of 2.0 × 10^5 configurations to provide estimates of statistical uncertainties in the ΔG's. The configurations were generated by Metropolis[23] and preferential sampling[25] algorithms.

The simulations also feature "double-wide sampling" in which the perturbations are performed in both directions, i.e, towards smaller and larger values of λ simultaneously.[5a] This enables the computation of two incremental ΔG's from one simulation. Another feature is decoupling of the electrostatic and van der Waals interactions;[5c] each annihilation was performed by first gradually reducing the partial charges to zero without perturbing the Lennard-Jones parameters, followed by gradual reduction of the σ's and ϵ's while simultaneously shrinking the molecule to a single point. It has been found that without this decoupling, oppositely charged sites draw close which causes unacceptable fluctuations in the free energy increments.[10d]

The perturbation steps ($\Delta\lambda$'s) were not uniform, but were kept small enough so that the magnitudes of the incremental ΔG's do not exceed ca. 1 kcal/mol and their uncertainties ($\pm$ 1σ) do not exceed ca. 0 1 kcal/mol. This required 10 – 18 increments in the electrostatic part and 18 – 24 increments in the van der Waals part of the calculations

All simulations were performed with the BOSS program (version 2 8) on Silicon Graphics 4D and Sun workstations in our laboratory.

Results and Discussion

The free energy changes computed for the various annihilation processes are shown in Table I. In converting these ΔG's to free energies of binding, one must take into account the possibility of multiple binding conformations. For A–U, there are four doubly hydrogen bonded conformations that are essentially isoenergetic: the Watson–Crick and Hoogsteen forms and their reversed counterparts [22b] One can thus assume each of these four to be present in approximately equal amounts However, our simulations take into account only the Watson–Crick form. Thus, there is a correction factor of RT ln 4 to be added to $\Delta G_{(A\text{–}U\rightarrow A)}$ and $\Delta G_{(A\text{–}U\rightarrow U)}$.[34] This correction factor is included in the ΔG values shown in Table I. No such correction is necessary for G–C, because the triply hydrogen bonded Watson–Crick form is considerably more stable than any other possible binding conformation.[22b] It is therefore expected to be predominant.

Table I. Calculated Free Energy Differences (kcal/mol)

	Process	Electrostatic Part	Van der Waals Part	Total
1	U→0	2.2±0.1	11.3±0.2	13.6±0.2
2	A-U→A	4.0±0.2	12.0±0.3	16.8±0.3[a]
3	A→0	2.8±0.1	11.8±0.2	14.6±0.2
4	A-U→U	5.2±0.1	12.5±0.3	18.6±0.3[a]
5	C→0	4.7±0.1	10.5±0.2	15.2±0.3
6	G-C→G	12.5±0.2	10.6±0.6	23.1±0.4
7	G→0	5.3±0.1	13.1±0.3	18.3±0.3
8	G-C→C	13.2±0.2	12.9±0.3	26.1±0.3
9	A+U→A-U	(from **1** and **2**)		−3.2±0.4
10	A+U→A-U	(from **3** and **4**)		−4.0±0.4
11	G+C→G-C	(from **5** and **6**)		−7.9±0.5
12	G+C→G-C	(from **7** and **8**)		−7.8±0.4

[a]With RT ln 4 correction.

As seen in Table I, similar values of ΔG_b are obtained for both base pairs from the two double annihilations. We point out that the two calculations for each base pair are completely independent. The similarity of the results thus provides confidence in the methodology. Support for the potential functions comes from comparisons with experimental data. The reported association constants for A–U from infrared spectroscopic studies is 100 ± 20 M^{-1},[13] which translates into a ΔG_b of −2.7 kcal/mol, while the present results is −3 to −4 kcal/mol. For G–C, the association constant has been estimated to be in the range of 10^4 – 10^5 M^{-1},[13a] which translates into a ΔG_b of −5.5 to −6.8 kcal/mol. Considering that the experimental values are "rough estimates",[13a] the accord with our calculated ΔG_b's is quite good.

Decoupling of the electrostatic and van der Waals terms in the calculations allows analysis of the source of the interactions in base pairing in chloroform. The data in Table I clearly show that these interactions are mainly electrostatic in nature, e.g., for C → 0 and G–C → G the van der Waals terms are nearly identical, while the electrostatic terms differ by 7.8 kcal/mol. This is of course due to hydrogen bonding. In contrast, the solvation of the individual bases by chloroform is dominated by the van der Waals interactions. The absence of hydrogen bonding with the solvent allows the solutes to hydrogen bond with one another. This is in accord with traditional ideas about base pairing in a nonpolar medium.[35]

In the earlier study we reported the relative binding free energy between G–C and A–U to be −7.2 ± 0.3 kcal/mol.[22] This value came from three separate series of simulations in which G was converted to A, C to U, and G–C to A–U. Individual ΔG's for these processes are $\Delta G_{(G\to A)}$ = 3.7 ± 0.2 kcal/mol, $\Delta G_{(C\to U)}$ = 1.9 ± 0.2 kcal/mol, and $\Delta G_{(G\text{-}C\to A\text{-}U)}$ = 12.8 ± 0.2 kcal/mol. Compared to the present results, $\Delta\Delta G_b$ is too negative. The source of this discrepancy appears to be associated with the use of harmonic constraints for the G–C → A–U process.[22] When the calculation is repeated without the use of the constraints, the recomputed value of $\Delta G_{(G\text{-}C\to A\text{-}U)}$ is 10.3 ± 0.2 kcal/mol. With this new value, $\Delta\Delta G_b$ between G–C and

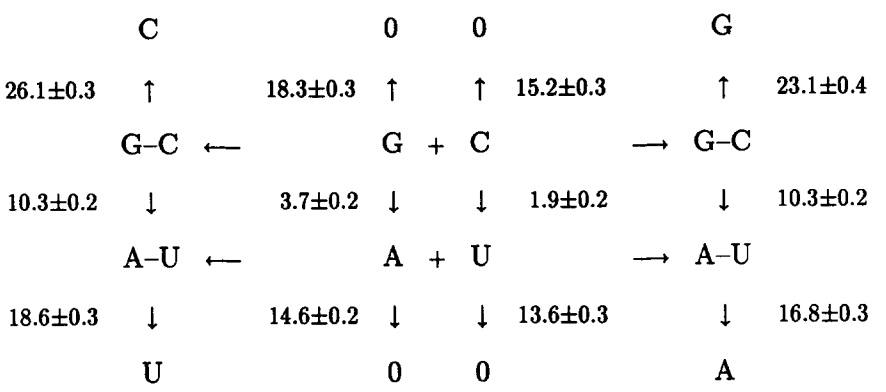

Figure 1. Free energy differences (kcal/mol) from the various perturbation calculations.

Table II. Thermodynamic Cycles from Figure 1

Cycle	ΔG (kcal/mol)
0→ U → C →0	−0.3
0→ A → G →0	0.0
A→A–U→G–C→G→A	−0.3
U→A–U→G–C→C→U	−0.9

A–U becomes −4.7 ± 0.3 kcal/mol, in excellent agreement with the present calculations.*

The free energy differences for the various processes involving the four nucleotide bases are collected in Figure 1. Presentation of the data in this manner allows easy identification of various thermodynamic cycles in addition to the ones involved in the ΔG_b calculations. Some of these are shown in Table II. The errors are quite small and within the statistical uncertainties of the calculations. Although the cycles are not all independent of one another, we point out that these data represent eleven independent free energy calculations that are nicely consistent with each other. We further note that the ΔG's involved are nontrivial; some have magnitudes greater than 20 kcal/mol. The closure of these thermodynamic cycles is an impressive validation of the present methodology. The large number of increments that have been used for the annihilations is critical to this success.

In an earlier calculation of this type a tendency was noted for the magnitudes of the free energy changes to be slightly smaller for exoergic perturbations than for endoergic ones.[10d] This trend has been

*One of the main points of the earlier paper was the similarity of ΔG_b's between A–U and U–DAP (1-methyluracil–2,6-diaminopyridine), which came from the similarities of $\Delta\Delta G_b$'s between G–C vs. A–U and G–C vs. U–DAP. The present result does not invalidate this point, because a recalculation of $\Delta G_{(G\text{–}C\rightarrow U\text{–}DAP)}$ without the harmonic constraints leads to a $\Delta\Delta G_b$ of −5.1 ± 0.7 kcal/mol between G–C and U–DAP. This is again similar to the new value of −4.7 kcal/mol for $\Delta\Delta G_b$ between G–C vs. A–U. Thus, the errors introduced by the use of the harmonic constraints apparently cancel out. We are currently investigating the source of the problems associated with the use of harmonic constraints in the earlier simulations.

Table III. Free Energy Changes (kcal/mol) for the van der Waals Part of the U→0 Process

λ_i	λ_j	$\Delta G_{i\to j}$	$\Delta G_{j\to i}$
0.00	0.05	0.785±0.094	−0.781±0.062
0.05	0.10	0.979±0.057	−0.694±0.092
0.10	0.15	0.764±0.116	−0.632±0.039
0.15	0.20	0.900±0.035	−0.663±0.079
0.20	0.25	0.708±0.068	−0.685±0.068
0.25	0.30	0.737±0.075	−0.648±0.033
0.30	0.35	0.735±0.054	−0.539±0.040
0.35	0.40	0.644±0.092	−0.496±0.075
0.40	0.45	0.450±0.079	−0.613±0.083
0.45	0.50	0.720±0.069	−0.509±0.052
0.50	0.55	0.661±0.043	−0.521±0.026
0.55	0.60	0.580±0.043	−0.546±0.025
0.60	0.65	0.647±0.072	−0.514±0.023
0.65	0.70	0.615±0.062	−0.528±0.019
0.70	0.75	0.637±0.020	−0.467±0.019
0.75	0.80	0.555±0.027	−0.465±0.010
0.80	0.80	0.903±0.023	−0.581±0.030
0.90	1.00	0.248±0.018	−0.154±0.012
Totals		12.268±0.272	−10.036±0.215

investigated further using the van der Waals part of the U → 0 process. Another set of simulations with double-wide sampling was run using values of λ which overlap those from the previous series. This allows the accumulation of free energy changes for processes exclusively in the forward (U → 0) and backward (0 → U) directions. The results are shown in Table III. The total ΔG from the second set of simulations with double wide sampling is 11.0 ± 0.3 kcal/mol, in agreement with the earlier value (11.3 kcal/mol; Table I). However, as seen in Table III, accumulation of free energy changes exclusively in the forward direction gives a ΔG of 12.3 kcal/mol, while in the backward direction it gives 10.0 kcal/mol. This is the same tendency observed earlier. The average, 11.2 kcal/mol, may be considered the best estimate of the true ΔG. It is apparent that double wide sampling also provides an excellent estimate of this value, as found and analyzed previously.[10d] Basically, even for small $\Delta\lambda$, the calculated ΔG's tend to be a little too positive. In the U → 0 calculation, the forward process has positive ΔG's, therefore the magnitudes of the calculated values will be too large. Conversely the magnitudes of ΔG's in the backward process, which are negative, will be too small. These systematic errors cancel out in a double-wide sampling calculation, in which half the perturbations are performed in the forward direction and half in the backward direction.[10d]

The individual errors can be minimized by using very small perturbation steps, so that the displacement of the perturbed system is insignificant. This point is confirmed by zooming in on the λ = 0.05 – 0.10 window of the U → 0 calculation (Table IV). Calculation of ΔG in this window with $\Delta\lambda$'s of 0.01 results in nearly identical values in the forward and backward directions Comparison with the values calculated with

Table IV. Free Energy Changes (kcal/mol) for the van der Waals Part of the U→0 Process (with small $\Delta\lambda$'s)

λ_i	λ_j	$\Delta G_{i\to j}$	$\Delta G_{j\to i}$
0.05	0.06	0.147±0.010	−0.151±0.008
0.06	0.07	0.156±0.010	−0.152±0.009
0.07	0.08	0.153±0.010	−0.138±0.012
0.08	0.09	0.135±0.014	−0.169±0.013
0.09	0.10	0.167±0.013	−0.147±0.010
Totals		0.758±0.026	−0.757±0.024

the larger $\Delta\lambda$'s (Table III) verifies the hypothesis: the magnitude of ΔG from the forward perturbation is too large and from the backward perturbation too small. Or, if their signs are considered, they are both too positive. This analysis shows the utility of the double-wide sampling procedure. Not only does it give twice the number of free energy increments, it also provides a better estimate of the accumulated ΔG's than simulations that are run in only one direction.

Conclusions

Monte Carlo simulations with statistical perturbation theory have been performed to calculate the absolute free energies of binding for G–C and A–U base pairs in chloroform. The quality of the methodology was demonstrated by consideration of various thermodynamic cycles which showed good consistency in the free energy changes. Favorable comparisons with experimental data provided additional validation of the procedures and intermolecular potential functions.

On the technical side, it is recommended that perturbation calculations of this type be carried out using the double-wide sampling procedure.[5a] Systematic errors present in unidirectional perturbation calculations tend to cancel in this procedure, leading to more accurate results.

Acknowledgments. Gratitude is expressed to the National Institutes of Health for support of this work.

References

1. Beveridge, D. L; Jorgensen, W. L., Eds. *Ann. N. Y. Acad. Sci.* **1986**, *482*.
2. McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University: Cambridge, 1987.
3. Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
4. Beveridge, D. L.; DiCapua, F. M. *Ann. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.
5. (a) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050. (b) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. *J. Am. Chem. Soc.* **1985**, *107*, 7793. (c) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564. (d) Singh, U. C.; Brown, F. K.: Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607.

6. Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, *109*, 6857. Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, *111*, 4190.

7. Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683.

8. Rosenberg, R. O.; Mikkilineni, R.; Berne, B. J. *J. Am. Chem. Soc.* **1982**, *104*, 7647. Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5757. Jorgensen, W. L.; Gao, J.; Ravimohan, C. *J. Phys. Chem.* **1985**, *89*,3470. Jorgensen, W. L.; Buckner, J. K. *J. Phys. Chem.* **1987**, *91*, 6083. Jorgensen, W. L.; Gao, J. *J Am. Chem. Soc.* **1988**, *110*, 4212. Mezei, M.; Mehrotra, P. K., Beveridge, D. L. *J. Am. Chem. Soc.* **1985**, *107*, 2239. Jayaram, B.; Mezei, M.; Beveridge, D. L. *J. Am. Chem. Soc.* **1988**, *110*, 1691.

9. (a) Berkowitz, M.; Karim, O. A.; McCammon, J. A.; Rossky, P. J. *Chem. Phys. Lett.* **1984**, *105*, 577. (b) Dang, L. X.; Pettitt, B. M. *J. Am. Chem. Soc.* **1987**, *109*, 5531. (c) Jorgensen, W. L.; Buckner, J. K.; Huston, S. E.; Rossky, P. J. *J. Am. Chem. Soc.* **1987**, *109*, 1891. (d) Buckner, J. K.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 2507. (e) Boudon, S.; Wippf, G.; Maigret, B. *J. Phys. Chem.* **1990**, *94*, 6056.

10. (a) Lybrand, T. P.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A* **1986**, *83*, 833. (b) Jorgensen, W. L.; Nguyen, T. B.; Boudon, S. *J. Am. Chem. Soc.* **1989**, *111*, 755. (c) Dang, L. X.; Kollman, P. A. *J Am Chem. Soc.* **1990**, *112*, 5716. (d) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.*, in press.

11. (a) Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830. Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574. Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551. Hirono, S.; Kollman, P. A. *J. Mol. Biol.* **1990**, 212, 197. (b) Gago, F.; Richards, W. G. *Mol. Pharmacol.* **1990**, *37*, 341.

12. Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 3049. Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154. Chandrasekhar, J., Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974. Madura, J. D.; Jorgensen, W. L. *J. Am Chem. Soc.* **1986**, *108*, 2517.

13. (a) Kyogoku, Y.; Lord, R. C ; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 250. (b) Kyogoku, Y.; Lord, R. C ; Rich, A. *Biochim. Biophys. Acta* **1969**, *179*, 10.

14. Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010. Iwahashi, H.; Kyogoku, Y. *J Am. Chem. Soc.* **1977**, *99*, 7761. Petersen, S. B.; Led, J. J. *J. Am. Chem. Soc.* **1981**, *103*, 5308. Williams, N. D.; Williams, L. D.; Shaw, B. R. *J. Am. Chem. Soc.* **1989**, *111*, 7205.

15. Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.

16. (a) Ohta, Y ; Tanaka, H,; Baba, Y.; Kagemoto, A.; Nishimoto, K. *J. Phys Chem.* **1986**, *90*, 4438. Ohta, Y., Nishimoto, K.; Tanaka, H ; Baba, Y.; Kagemoto, A. *Bull. Chem Soc. Jpn.* **1989**, *62*, 2441. (b) Williams, L. D.; Chawla, B.; Shaw, B. R. *Biopolymers* **1987**, *26*, 591.

17. Nakano, N. I.; Igarashi, S. J. *Biochemistry* **1970**, *9*, 577.

18. Pullman, B.; Claverie, P.; Caillet, J. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 904. Kudritskaya, Z. G ; Danilov, V. I. *J. Theoret. Biol.* **1976**, *59*, 303. Langlet, J.; Claverie, P., Caron, F.; Boeuve, J. C., *Int J. Quant. Chem.* **1981**, *19*, 299.

19. Aida, M. *J. Comput. Chem.* **1988**, *9*, 362.

20. Pohorille, A.; Pratt, L. R.; Burt, S. K.: MacElroy, R. D. *J. Biomol. Struc. Dyn.* **1984**, *1*, 1257. Pohorille, A.; Burt, S. K.; MacElroy, R. D. *J. Am. Chem. Soc.* **1984**, *106*, 402.

21. (a) Cieplak, P; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 3734. (b) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 503.

22. (a) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008. (b) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L., submitted for publication.

23. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

24. Patey, G. N.; Valleau, J. P. *J. Chem. Phys.* **1975**, *63*, 2334.

25. Owicki, J. C. *A.C.S. Symp. Ser.* **1978**, *86*, 159.

26. Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

27. Gao, J.; Garner, D. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 4784.

28. Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

29. Prue, J. E. *J. Chem. Educ.* **1969**, *46*, 12. Justice, M. C.; Justice, J. C. *J. Sol. Chem.* **1976**, *5*, 543.

30. Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770.

31. Sneddon, S. F.; Tobias, D. J.; Brooks, C. L., III. *J. Mol. Biol.* **1989**, *209*, 817

32. Jorgensen, W. L.; Blake, J. F.; Buckner, J. K. *Chem. Phys.* **1989**, *129*, 193.

33. Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742.

34. Bailey, W. F.; Monahan, A. S. *J. Chem. Educ.* **1978**, *55*, 489.

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