11550 *J. Am. Chem. Soc.* **1994,** *116,* 11550−11553

Conformational Free Energies from Simulation: Stochastic Dynamics/Monte Carlo Simulations of a Homologous Series of Gellman's Diamides

D. Quentin McDonald and W. Clark Still*

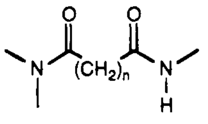
Contribution from the Department of Chemistry, Columbia University, New York, New York 10027

Received August 10, 1994⊗

Abstract: We report simulations using the recently described Monte Carlo/stochastic dynamics (MC/SD) simulation method to examine the conformational thermodynamics for a homologous series of diamides. The simulations use the AMBER* force field and the GB/SA model for $CHCl_3$ and are able to reproduce experimental observations of the degree of hydrogen bonding in the series. By carrying out simulations at different temperatures, we are able to obtain thermodynamic parameters for the equilibrium between hydrogen-bonded and non-hydrogen-bonded states. These parameters are found to be in general agreement with experiment. We also discuss the importance of the gas phase force field parameters, the solvation model, and the simulation methodology on the calculated results. One of the more surprising findings is that converged results are sensitive not only to the relative potential energies of conformational minima, but also to the heights of the barriers between conformations.

Introduction

In recent years, Gellman and co-workers have reported a number of experimental studies of the conformational preferences of di- and triamides in organic solution. One particularly interesting set of compounds described by Gellman in this context is the homologous series of diamides **1**−**5**.[1] Each member of this series has a single hydrogen bond acceptor and donor capable of forming an intramolecular hydrogen bond. Gellman studied the temperature dependence of the populations of the hydrogen-bonded and non-hydrogen-bonded states using variable temperature IR and ^{1}H NMR studies of these compounds in organic solvents. From these observations, the enthalpic and entropic parameters associated with the equilibrium between hydrogen-bonded and non-hydrogen-bonded states could be obtained. Such studies are important because they advance our meager knowledge of the conformational preferences of flexible molecules that are capable of forming intramolecular amide−amide hydrogen bonds. Because the Gellman experiments not only measure conformational populations but also provide thermodynamic parameters, these compounds provide a challenging test for computational molecular modeling. In this paper, we report simulations of diamides **1**−**5** using a molecular mechanics force field and free energy simulation methodology. We show that experiment can be well reproduced without *ad hoc* modifications to the calculation, provided that one deals with important issues involving force field parameterization and simulation convergence.



1 n=1
2 n=2
3 n=3
4 n=4
5 n=5

We recently described a preliminary study[2] of several amides including **1**−**5**. In that work, we estimated hydrogen-bonded populations by using stochastic dynamics (SD) at 200−300 K in $CHCl_3$ to generate representative, Boltzmann-weighted sets of conformations (ensembles) and then testing each such conformation in the sets for the presence of a hydrogen bond. While this approach worked well for evaluating the free energy based hydrogen-bonded populations of **1**, analogous calculations with **2**−**5** were problematic due to the more numerous minima and larger energy barriers separating them. In the work described here, we study diamides **2**−**5** using a newly developed simulation method that employs a mixed Monte Carlo/stochastic dynamics (MC/SD) algorithm.[3] This new method has been demonstrated to generate the same (Gibbs) ensemble as MC or SD alone but at a much more rapid rate in systems having multiple conformations. As we will demonstrate, MC/SD allows us to perform free energy simulations on **2**−**5** which are fully converged and thus provides precise theoretical measures of hydrogen-bonding populations.

Methods

Simulation Methodology. All calculations were performed using the MacroModel and BatchMin V4.5 molecular modeling programs. Simulations were performed using the recently described mixed Monte Carlo/stochastic dynamics (MC/SD) method or a stochastic variant (SD) of the velocity Verlet algorithm.[3,4] The total simulation time in all cases was 10 ns with a 1.5 fs time step for the SD part of the simulation algorithm. The first 50 ps of each simulation was taken as an equilibration period and discarded. The MC part of the MC/SD algorithm used random torsional rotations between ±60° and ±180° and was applied to all rotatable bonds except the peptide C−N bonds. The ratio of SD steps to MC steps was 1:1, and acceptance rates for the MC part of the simulation ranged from 0.5 to 5%. All SD and MC/SD simulations used a frictional coefficient (γ) of 2.5 ps^{-1}. Nonbonded cutoff distances were not used. Hydrogen-bonding populations were estimated by monitoring the geometry around the acceptor (O) and the donor (H) atoms during the simulation: a hydrogen bond was counted as present if the (N−)H⋯O(=C) distance was <2.5 Å, the N−H⋯O(=C) angle was >120° and the (N−)H⋯O=C angle was >90°.[5]

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1994.

[1] Gellman, S. H.; Dado, G. P.; Liang, G.-B.; Adams, B. R. *J. Am. Chem. Soc.* **1991,** *113,* 1164.

[2] McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992,** *33,* 7747.

[3] Still, W. C.; Guarnieri, F. *J. Comput. Chem.* **1994,** *15,* 1302.

[4] van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988,** *1,* 173.

[5] Kurz, M.; Mierke, D. F.; Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1992,** *31,* 210.

0002-7863/94/1516-11550$04.50/0

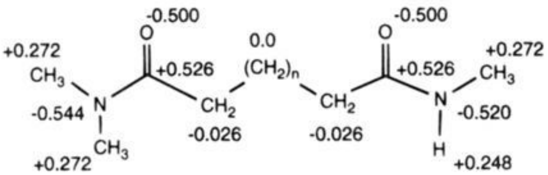


Figure 1. Atomic partial charges used in these calculations.

Table 1. Relative Conformational Energies of *n*-Butane (kcal/mol)

n-butane torsion angle	MM3	AMBER	AMBER*B
180° (anti)	0.0	0.0	0.0
(120°)	3.3	4.0	3.0
60° (gauche)	1.1	1.0	1.3
(0°)	4.8	7.0	4.9

Force Field. The molecular mechanics force field used in these calculations was united atom AMBER*. The AMBER* force field used here differs from authentic AMBER[6] in the use of modified amide $O{=}C{-}C{-}C$ torsional parameters which we developed to reproduce *ab initio* conformational energies of propanamide. These modified torsional parameters were previously reported and are $V1/2 = 0.0$, $V2/2 = 1.44$, and $V3/2 = -0.41$ kcal/mol.[2] Other parameters and atomic partial charges were assigned by analogy with the standard united atom AMBER treatment of peptide backbone amides and the glutamine side chain (see Figure 1). The hydrogen bonding potential was the 6−12-Lennard-Jones treatment recently described by Kollman.[7] To investigate the sensitivity of our findings to parameter details, we modified a second set of torsional parameters to reproduce MM3[8] energies of the conformational minima and maxima of *n*-butane.[9] While authentic AMBER correctly reproduces the *anti−gauche* conformational energy difference for *n*-butane, it somewhat overestimates butane's barriers to rotation. To bring the energies of *n*-butane's minima and *maxima* into line, we modified the torsional parameters for methylene−methylene fragments, and the corresponding force field is designated here as AMBER*B. The modified torsional parameters for united atom$-CH_2-CH_2-$arrays (AMBER designation X−C2−C2−X) are $V1/2 = -0.10$, $V2/2 = 0.31$, and $V3/2 = 1.05$ kcal/mol. Using AMBER*B, the relative energies of *n*-butane's minima and maxima are well reproduced (see Table 1). The AMBER*B calculations described below for **2**−**5** employ these new parameters for all methylene−methylene torsions in conjunction with the above-described modified amide parameters. For modeling the effect of chloroform solvation, we used the GB/SA analytical continuum model.[10]

Evaluation of Simulation Convergence. Results obtained from molecular simulations are meaningful only if the simulations can be demonstrated to be converged, *i.e.*, give results which are invariant of starting conditions and simulation length. Convergence can be established by showing (1) the stability of average properties such as potential energy during the simulation, (2) that simulations beginning from significantly different initial conditions give the same results, and (3) in the case of molecules having symmetry, that all equivalent conformations (*e.g.*, +/− gauche torsions in **2**−**5**) are equally populated in the simulation. In our studies of **2**−**5**, we were able to satisfy all of these criteria for convergence. First, all of our 10 ns SD and MC/SD simulations showed stability in the cumulative average potential energy with fluctuations below 0.1% after half of the total simulation time. Such potential energy stability is necessary but insufficient for demonstrating convergence. Consequently, we conducted two other tests to establish simulation convergence.

One of the most convincing proofs of convergence is the demonstration that identical average properties can be obtained by beginning simulations from very different initial conditions. To test for starting condition independence, we performed SD and MC/SD simulations with diamide **4** starting from two different conformations, one internally hydrogen bonded (**A**) and one extended (**B**). In our tests, the average property monitored was the distance between the donor (N−)H and the acceptor (C=)O, and its average is graphed against simulation time in Figure 2. As shown, the average distance found using an MC/SD simulation reaches a stable value after about 1 ns of simulation time, and simulations starting with conformations **A** or **B** give the same average distance (~5.7 Å) to within <1%. In contrast, the average distance obtained with SD shows a strong dependence on the starting conformation, and averages from **A** and **B** still differ by 7% after 5 ns of simulation. Thus, the MC/SD simulation of **4** is better converged after 0.5 ns than the SD simulation is after 5 ns. This result is consistent with our previous comparisons of the MC/SD and SD simulation methods which established that SD crosses even moderate energetic barriers (*e.g.*, 3 kcal/mol) rather infrequently at room temperature.[3]

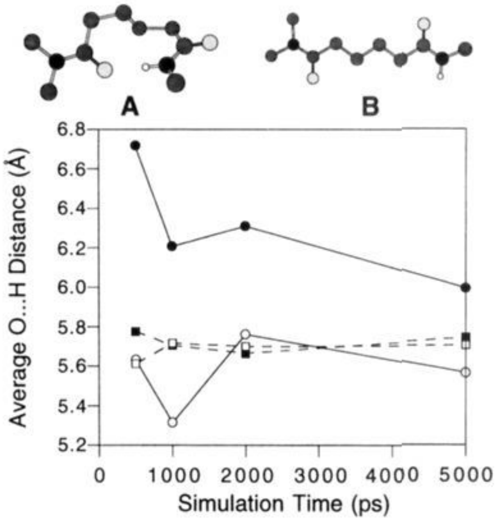


Figure 2. Average O···H distance of **4** as a function of simulation time for Monte Carlo/stochastic dynamics (MC/SD) and stochastic dynamics (SD) simulations at 300 K starting from hydrogen-bound (**A**) and extended (**B**) conformations: (filled circles) SD from **A**; (open circles) SD from **B**; (filled squares) MC/SD from **A**; (open squares) MC/SD from **B**.

We have also conducted convergence tests based on symmetry—the idea being that, at convergence, conformations which are enantiomeric or equivalent by symmetry must have equal populations. In the case of achiral **4**, there are five pairs of enantiomeric conformations which differ in only the sign of one C−C torsion angle. As a convergence test, we carried out a 10 ns, 300 K MC/SD simulation of **4** and monitored the values of each of these five torsions. As shown in Figure 3, the dihedral angle distributions of all five rotatable bonds show a high degree of symmetry around zero, indicating that equivalent conformations have been equally sampled during the simulation. In contrast, the dihedral distributions found using 10 ns, 300 K SD simulations were highly unsymmetrical (data not shown). Taken together, our tests establish that 10 ns MC/SD simulations of diamides **2**−**5** at 300 K give converged conformational populations to better than 1%. The same high degree of convergence using MC/SD was also found with these systems at 200 K, an achievement which would be difficult, if not problematic, using standard molecular dynamics or stochastic dynamics.

Results

Having parameterized our force field against *ab initio* quantum calculations and established that our MC/SD simulations are well converged, we carried out a series of 300 K MC/SD simulations to investigate the degree of hydrogen bonding as a function of the number of methylenes between the donor and acceptor groups in diamides **1**−**5**. We also performed simula-

---

[6] Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

[7] Ferguson, D. M.; Kollman, P. A. *J. Comput. Chem.* **1991**, *12*, 620.

[8] Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

[9] MM3 has been explicitly parameterized to reproduce high-level *ab initio* calculations on *n*-butane: Allinger, N. L.; Grev, R. S.; Yates, B. F.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1990**, *112*, 114.

[10] Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

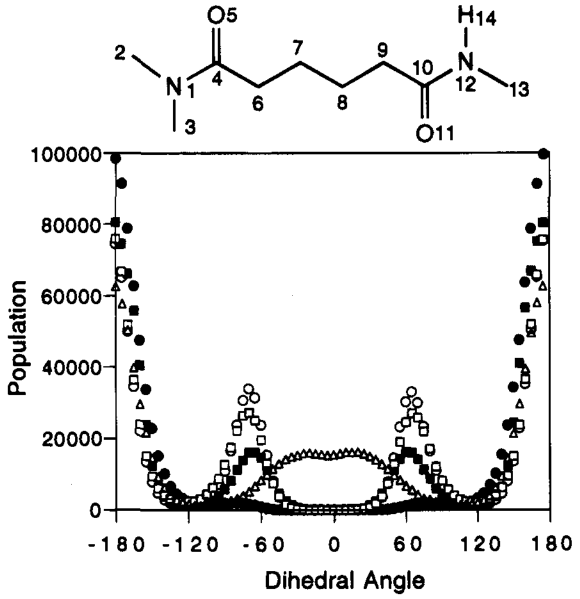


Figure 3. Dihedral angle distribution from 10 ns, 300 K MC/SD simulation of diamide **4**. Each point represents the population in a 5° range: (filled circles) angle 1−4−6−7; (open circles) angle 4−6−7−8; (filled squares) angle 6−7−8−9; (open squares) angle 7−8−9−10; (open triangles) angle 8−9−10−12.

Table 2. Hydrogen-Bonding Populations for Diamides **2**−**5** from 10 ns MC/SD Simulations at 300 and 200 K[a]

	population of hydrogen-bonded conformations (%)				
	AMBER*		AMBER*B		
diamide	300 K	200 K	300 K	200 K	experiment (CH_2Cl_2)[1]
2	38	53	40	54	more hydrogen bonding than **3**
3	3	2	12	18	little hydrogen bonding
4	20	58	18	45	little hydrogen bonding, except at low temperature
5	11	33	6	10	almost no hydrogen bonding

[a] All calculations employed GB/SA $CHCl_3$ AMBER* uses standard AMBER $-CH_2-CH_2-$ torsional parameters $V1/2 = 0.0$, $V2/2 = 0.0$, and $V3/2 = 2.0$ kcal/mol; AMBER*B uses modified $-CH_2-CH_2-$ torsional parameters $V1/2 = -0.10$, $V2/2 = 0.31$, and $V3/2 = 1.05$ kcal/mol.

tions at lower temperatures (*e.g.*, 200 K) to investigate the temperature dependence of hydrogen bonding. We previously noted that the barriers separating conformations of diamide **1** are sufficiently low to permit converged SD simulations around room temperature and reported that, consistent with experiment, 77% of **1**'s conformations are calculated to be hydrogen bonded at 300 K.[2] In that work, we were unable to carry out analogous SD calculations with **2**−**5** because of slow conformational interconversion. The convergence problem is now solved by the MC/SD simulation method with diamides **2**−**5**, and the converged results of our simulations with these molecules are summarized in Table 2.

Our simulations were carried out in GB/SA $CHCl_3$ using the two different force fields AMBER* and AMBER*B. The AMBER* calculations employ the original AMBER united-atom $-CH_2-CH_2-$ torsion parameters. The results with this parameter set are reasonably consistent with the experimental observations; however, diamide **3** is calculated to only be 3% hydrogen bonded at 300 K in $CHCl_3$. Such a small amount of hydrogen bonding would not be detected by the ^1H NMR and IR measurements employed and is inconsistent with the reported observation of small, but detectable amounts of hydrogen bonding.

The AMBER*B results utilized our modified $-CH_2-CH_2-$ torsional parameters which better reproduced MM3 energies of *n*-butane at its conformational minima and maxima (see Table 1). These AMBER*B simulations are also more consistent with the experimental findings that hydrogen bonding is preferred by **2** > **3** ≈ **4** > **5** in methylene chloride. Though experimental studies have not yet provided actual percentages of hydrogen bonding in these molecules, integrals of IR peaks of hydrogen-bonded and non-hydrogen-bonded forms of **2** and **3** are found in approximate ratios of 2:3 (**2**) and 2:7 (**3**), respectively—results that are generally consistent with our calculated populations for these molecules.[11] AMBER*B simulations also reproduce the large temperature coefficient for hydrogen bonding that is found experimentally with **4**. Both the AMBER* and AMBER*B results are more consistent with experiment than are simulations with authentic AMBER (torsional parameters from the glutamine side chain) which gave **3** as the most heavily (42%) hydrogen-bonded compound.

Though the results in Table 2 reproduce the hydrogen-bonding trends observed experimentally, the actual calculated percentages of hydrogen-bonding forms depend upon our definition of a hydrogen bond. As noted above, we defined a conformation as hydrogen bonding if it had (C=)O⋯H(−N) < 2.5 Å, (C=)O⋯H−N > 120°, and C=O⋯H(−N) > 90°. To test our results for sensitivity to the specifics of the hydrogen bond definition, we carried out additional simulations with other definitions. In particular, we conducted 10 ns, 300 K simulations with **2**−**4** using both increased and decreased stringencies of the hydrogen bond tests. These definition changes were effected by ±10% variations in the limits for the aforementioned distance and two angles. We found that the decreased stringency results ((C=)O⋯H(−N) < 2.75 Å, (C=)O⋯H−N > 108°, and C=O⋯H(−N) > 81°) gave hydrogen-bonding populations for **2**−**4** that were all larger by ~5% relative to the 300 K data in Table 2, and that the increased stringency simulations had hydrogen-bonding populations that were all smaller by ~30%. Thus, the hydrogen bond populations vary in the expected way with definition stringency. More importantly, the same hydrogen-bonding trend (**2** > **3** ≈ **4** > **5**) is found for all definitions.

For diamides **3** and **4** it is possible to extend the comparison with experiment beyond the qualitative comparison described above. These diamides have been analyzed as a two-state system having hydrogen-bonding and non-hydrogen-bonding states using eq 1 to obtain thermodynamic parameters which

$$\ln(K_{\mathrm{eq}}) = \left(\frac{-\Delta H^\circ}{R}\right)\left(\frac{1}{T}\right) + \frac{\Delta S^\circ}{R} \qquad (1)$$

quantify the enthalpic and entropic preference for intramolecular hydrogen bonding.[1] We have carried out a similar analysis using our MC/SD simulations to obtain hydrogen-bonded populations of **3** and **4** at temperatures between 200 and 300 K. In particular, we performed three new simulations with diamides **3** and **4** at temperatures of 225, 250 and 275 K. Using this data and results from Table 2, the fractions of hydrogen bonding at each temperature were converted to $\ln(K_{eq})$ and plotted against $1/T$. The resulting van't Hoff plot for diamide **4** is shown in Figure 4.

These van't Hoff plots display a number of interesting features. Firstly we find a clear linear relationship between the calculated $\ln(K_{eq})$ and $1/T$ for both the *in vacuo* and GB/SA $CHCl_3$ calculations as expected at convergence. As can be seen in Figure 4, the points calculated from our simulations performed using the GB/SA $CHCl_3$ solvation treatment are in close

---

(11) Raw IR data for **2** and **3** provided by Professor S. H. Gellman, University of Wisconsin.

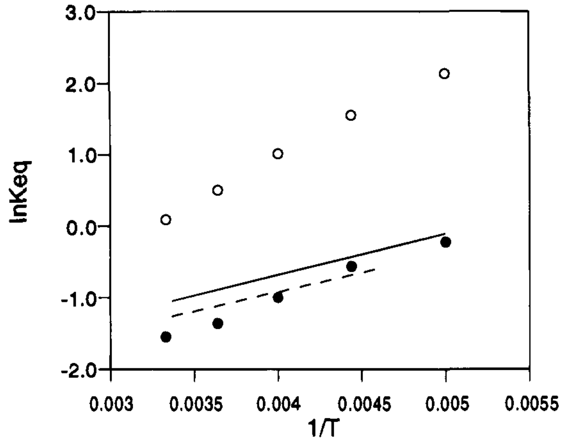


Figure 4. Van't Hoff plot constructed from hydrogen-bonded populations of **4** calculated during 10 s MC/SD AMBER* simulations and from experimental thermodynamic parameters: (filled circles) MC/SD simulation with AMBER*B and GA/SA $CHCl_3$; (open circles) MC/SD simulations with AMBER*B *in vacuo*; (dashed line) experiment by ^1H NMR; (solid line) experiment by IR.

Table 3. Thermodynamic Parameters for the Hydrogen-Bonded/Non-Hydrogen-Bonded Equilibrium Obtained from Experiment and 10 ns Mixed Mode Simulations (AMBER* and GB/SA $CHCl_3$)

	$\Delta H°$ (kcal/mol)			$\Delta S°$ (eu)		
diamide	^1H NMR	IR	simulation	^1H NMR	IR	simulation
3		−0.4	−0.62		−3.3	−6.1
4	−1.5	−1.4	−1.65	−7.5	−6.8	−8.6

agreement with values obtained from ^1H NMR and IR measurements of **4**. Similar results were obtained with **3**. In contrast, the calculated thermodynamic parameters are considerably different from experiment when the effects of solvation are omitted from the calculation. This finding supports the important role played by organic solvents in determining the balance between hydrogen-bonded and non-hydrogen-bonded states. It also underscores the importance of including a reasonable model for solvation in a calculation if quantitative accord with condensed phase experiments is desired.

The same calculated data can be used to derive thermodynamic parameters for the two-state process which can be compared directly with those obtained from experiment. These parameters are shown in Table 3 for diamides **3** and **4**.

As indicated in the table, there is generally good agreement between the calculated and experimentally measured thermodynamic parameters for both **3** and **4**. The calculated differences in enthalpy between the hydrogen-bonded and non-hydrogen-bonded states for the two diamides are in particularly good agreement with experiment. The calculated and measured entropic differences are in qualitative accord with experiment but are outside the estimated ±1 eu uncertainty for the measured values. Because of the high degree of convergence we are able to attain with the MC/SD simulation method, the discrepancies between calculation and experiment must reflect inaccuracies in the solvation treatment and/or the molecular mechanics potentials. Considering the simplicity of our continuum solvation treatment and the molecular mechanics model in general, we view the semiquantitative agreement between calculated and experimental thermodynamic parameters in this series of molecules as remarkably good.

Conclusion

Our results demonstrate that free energy simulations can reproduce the intramolecular hydrogen-bonding behavior of a simple diamide in an organic solvent—provided that important methodological issues are adequately addressed. The first issue is the molecular mechanics force field. We find that testing the *in vacuo* solute molecular mechanics conformational energies against high-level quantum mechanical calculations and molecular mechanics reparameterization as necessary to accurately reproduce the *ab initio* data is an essential first step in carrying out high-quality free energy simulations. The second issue is solvation. As we have found in this and other studies, accurate results cannot be obtained unless a reasonable solvation treatment is included. Though this point is widely appreciated when a highly polar solvent such as water is involved, it applies to organic solvents as well. The third issue is convergence. Unless a simulation directed toward the calculation of average properties can be shown to give the same results from significantly different initial conditions (*i.e.*, is converged), then the results cannot be taken as reliable or predictive. Achieving convergence in a free energy simulation is particularly difficult when the molecules involved have multiple conformational states. Even when the barriers to rotation are as low as 3 kcal/mol, convergence using molecular or stochastic dynamics methods can be very hard to achieve.[3] Similar concerns have been raised about the difficulty of achieving real convergence in other systems.[12] For the simple molecules described here, the mixed mode Monte Carlo/stochastic dynamics (MC/SD) simulation method works extremely well in generating fully converged data—an achievement that would be difficult by standard dynamics methods particularly in the case of the low-temperature (200 K) simulations. For multiconformational systems such as ours, both MC/SD and MC alone would seem to have distinct advantages over dynamics methods.

Finally, we would like to add that while conformational searching with calculation of Boltzmann populations is a popular approach to problems such as the one described here, such an approach makes a number of assumptions about the system harmonicity, entropy, and geometries of nonminimum energy conformations that may not be valid. Free energy simulations offer a substantially better alternative, at least in systems where convergence can be achieved. Given such a converged simulation and a reasonably accurate force field, the promise of molecular modeling—prediction—should be within grasp. The results described here provide further encouragement that useful results in hydrogen-bonding systems may be obtained with a conventional molecular mechanics force field and a continuum solvation treatment.

Acknowledgment. We thank Professor Sam Gellman of the University of Wisconsin for sharing his data on the hydrogen bonding of **2** and **3**. This work was supported by NSF Grant CHE92-08524.

---

[12] Mitchell, M. J.; McCammon, J. A. *J. Comput. Chem.* **1991**, *12*, 271. van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947. 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.