Molecular Recognition: Effect of Rotational Isomers on Host-Guest Binding

William R. Cannon,[†,‡] Jeffry D. Madura,[§] Randolph P. Thummel,[‡] and J. Andrew McCammon*[,†,‡]

Contribution from the Departments of Biochemistry and Chemistry, University of Houston, Houston, Texas 77204, and Department of Chemistry, University of South Alabama, Mobile, Alabama 36688. *Received May 29, 1992*

Abstract: Monte Carlo simulations have been used to study the relative binding of dimethylurea and imidazolidone to a synthetic host molecule in $CHCl_3$. The thermodynamic cycle-perturbation method was used to calculate the relative free energy of binding, which was compared with experimental data from NMR binding studies. Special techniques have been used to properly account for the different rotational isomeric states of dimethylurea in the thermodynamic averages. The computed relative free energy of binding $\Delta\Delta G$ = 3.6 kcal mol^{-1} favors the binding of imidazolidone and compares reasonably well with the experimental value of 3.1 kcal mol^{-1}.

Introduction

Interest in synthetic host-guest chemistry has increased greatly during the last ten years. Synthetic receptors and their guests are being used, among others, as enzyme mimics[1] and as models for drug-receptor[2] and DNA-protein complexes.[3] A goal of this work has been to gain an understanding of the molecular recognition process so that novel systems of scientific, medical, and economic importance can be rationally designed. The recognition process involves bringing two solute molecules together, desolvation, imposing any necessary conformational changes, and forming the electrostatic and steric interactions observed in the complex. Comprehensive reviews of the molecular recognition process and design principles have appeared in the literature.[1,4-7]

Experimental work by several groups has led to elucidation of important principles involved in molecular recognition. For instance, the importance of π-stacking in host-guest interactions has been demonstrated by Rebek and co-workers[8] and by Zimmerman and co-workers.[9] Several other groups have also been active in utilizing hydrogen bonds in complexation studies.[10-12] In particular, Chang and Hamilton[2] have designed a receptor that binds barbituate derivatives by hydrogen bond interactions with dissociation constants in the micromolar range.

In recent years molecular dynamics (MD) and Monte Carlo (MC) simulations have evolved into powerful methods for studying complicated molecular systems.[13-15] These techniques have become valuable tools in interpreting experimental results and predicting structural and energetic properties of complexes. For example, Jorgensen and Pranata recently demonstrated the importance of secondary hydrogen bond interactions in DNA complexes.[16] These observations were used by Jeong et al.[17,18] to explain differences in imide-imide, imide-lactam, and lactam-lactam host-guest chemistry.

The major theoretical tool for MD and MC studies of host-guest complexes is the thermodynamic cycle-perturbation method,[19,20] which allows calculations of relative free energies of binding. Such calculations have successfully addressed binding involving electrostatic interactions of ions,[21,22] hydrogen bond interactions,[23-25] and base stacking interactions.[24,26,27] However, when the guest or host molecules are capable of existing in multiple isomeric forms due to bond rotations, sampling of all relevant conformations can be a problem. For example, *N,N'*-dimethylurea can exist in three different rotational isomeric states: **1**, with both methyl groups syn to the carbonyl oxygen (s,s); **2**, with one methyl group anti and the other syn to the carbonyl oxygen (a,s); and **3**, with both methyl groups anti to the carbonyl oxygen (a,a).

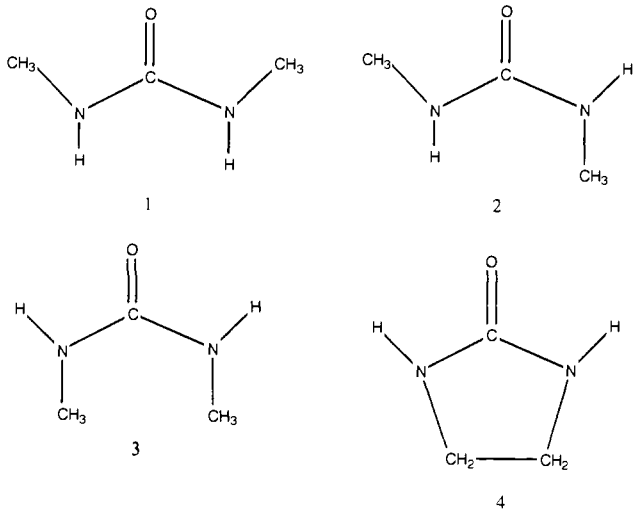
In such a case, uncertainty may arise as to which species are actually bound and what the reorganization of conformations upon binding may contribute to the relative free energy of binding. Recently, Hegde et al.[28,29] reported on such a system in which the binding of *N,N'*-dimethylurea and imidazolidone (**4**) to a cavity-shaped host (Figure 1) was examined. NMR titrations of the host in 1:1 $CDCl_3/CD_2Cl_2$ at 291 K were carried out wherein the addition of incremental amounts of the guest led to a progressive downfield shift of the host NH resonance (corrected for a small dilution effect). Analysis of these data led to the determination of association constants of 10 M^{-1} for *N,N'*-dimethylurea and 2240 M^{-1} for imidazolidone,[30] leading to an approximate $\Delta\Delta G$ of binding of 3.1 kcal mol^{-1} in favor of imidazolidone.

In this paper, MC free energy simulations are used to determine the relative binding constants of dimethylurea and imidazolidone

[†] Department of Biochemistry, University of Houston.
[‡] Department of Chemistry, University of Houston.
[§] Department of Chemistry, University of South Alabama.

(1) Cram, D. J. *Angew Chem., Int. Ed. Engl.* **1988**, *27*, 1009.
(2) Chang, S.-K.; Van Engen, D.; Fan, E.; Hamilton, A. D. *J. Am. Chem. Soc.* **1991**, *113*, 7640.
(3) Zimmerman, S. C.; Wu, W.; Zeng, Z. *J. Am. Chem. Soc.* **1991**, *113*, 196.
(4) Cram, D. J. *Nature* **1992**, *356*, 29.
(5) Schneider, H.-J. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 1417.
(6) Rebek, J., Jr. *Science* **1987**, *235*, 1478.
(7) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 89.
(8) Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1090.
(9) Zimmerman, S. C.; VanZyl, C. M. *J. Am. Chem. Soc.* **1987**, *109*, 7894.
(10) Neder, K. M.; Whitlock, H. W., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 9412.
(11) Goodman, M. S.; Rose, S. D. *J. Am. Chem. Soc.* **1991**, *113*, 9380.
(12) Adrian, J. C., Jr.; Wilcox, C. S. *J. Am. Chem. Soc.* **1989**, *111*, 8055.
(13) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: New York, 1987.
(14) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Wiley: New York, 1988.
(15) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.
(16) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.
(17) Jeong, K.-S.; Tjivikua, T.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 3215.
(18) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 201.
(19) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281.
(20) Straatsma, T. P.; McCammon, J. A. *Ann. Rev. Phys. Chem.* **1992**, *43*, 407.
(21) Mazor, M.; McCammon, J. A.; Lybrand, T. P. *J. Am. Chem. Soc.* **1990**, *112*, 4411.
(22) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 5716.
(23) Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770.
(24) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 7269.
(25) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 209.
(26) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768.
(27) Rao, S. N.; Kollman, P. A. *Proc. Natl. Acad. Sci., U.S.A.* **1987**, *84*, 5735.
(28) Hegde, V.; Madhukar, P.; Madura, J.; Thummel, R. P. *J. Am. Chem. Soc.* **1990**, *112*, 4549.
(29) Hegde, V.; Hung, C.-Y.; Madhukar, P.; Cunningham, R.; Höpfner, T.; Thummel, R. P. *J. Am. Chem. Soc.*, previous paper in this issue.
(30) Recalculated K_a values in ref 29 differ slightly from those reported in ref 28.

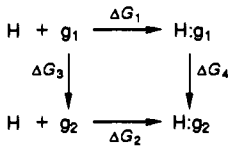
0002-7863/93/1515-879$04.00/0



with the host. Using the method of Straatsma and McCammon,[31] we address the problems associated with rotational isomers of dimethylurea. OPLS potentials are used when available and OPLS-derived potentials are used elsewhere. Partial charges for the host and guest are derived from ab initio calculations. Internal degrees of freedom for the partially saturated rings on the host molecule are incorporated by the use of a Fourier series to reproduce the quantum mechanical potential energy surface.

Theory

The relative free energy of binding of two guest molecules g_1 and g_2 to a host molecule H can be computed using the thermodynamic cycle shown in eq 1.



$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3 \quad (1)$$

The free energy changes ΔG_1 and ΔG_2 have been experimentally determined by NMR titration methods.[28] Computationally it is easier to obtain $\Delta\Delta G$ from calculations of ΔG_3 and ΔG_4.[20]

The free energy difference between two states can be written as

$$\Delta G = G_1 - G_0 = -k_B T \ln \frac{Q_1}{Q_0} \quad (2)$$

where Q_0 and Q_1 are the isothermal–isobaric partition functions for states 0 and 1, respectively. In MD or MC simulations, ΔG can be conveniently given by

$$\Delta G = -k_B T \ln \langle e^{-\Delta E/k_B T} \rangle_0 \quad (3)$$

Here the brackets represent the ensemble average taken over state 0 and ΔE is the configuration energy difference between state 1 and state 0. In the simulation, the configurational energy difference is composed of several inter- and intramolecular contributions

$$\Delta E = E^1_{\text{intra1,5}} + E^1_{\text{dihedral}} + E^1_{\text{solute–solute}} + E^1_{\text{solute–solvent}} - E^0_{\text{intra1,5}} - E^0_{\text{dihedral}} - E^0_{\text{solute–solute}} - E^0_{\text{solute–solvent}}$$

$E_{\text{intra1,5}}$, $E_{\text{solute–solute}}$, and $E_{\text{solute–solvent}}$ are the nonbonded Lennard-Jones and Coulombic contributions; E_{dihedral} is the energy for rotation about a bond including the 1,4-interactions.

In order to calculate free energy changes of flexible molecules accurately, the available phase space of the system must be adequately sampled. In this study, controlled sampling of the rotational isomers[31] is used to ensure that phase space is adequately explored. The free energy difference between a system that is free to move throughout its phase space and the system that is confined to a subspace (i.e., one of the rotational isomers) is[31]

$$G - G_j = -k_B T \ln \sum_{i=1}^{n} e^{-(G_i - G_j)/k_B T} \quad (4)$$

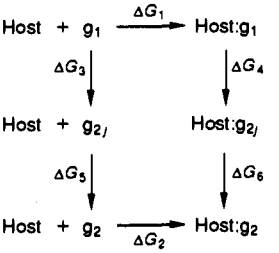
Here G represents the free energy of the molecule of interest, G_j represents the free energy of rotational isomer j, and n is the total number of subspace regions associated with distinguishable rotational isomers.

In the case of dimethylurea, the barrier for rotation of a secondary amide about the carbon–nitrogen bond is roughly 20–22 kcal mol^{-1}. An energy barrier of this height will not be crossed using standard simulation techniques. But the free energy difference between dimethylurea which is free to explore its entire phase space and a dimethylurea molecule confined to the phase space of the a,s rotational isomer is given by

$$G^{\text{dmu}} - G^{\text{dmu}}_{\text{a,s}} = -k_B T \ln \left(e^{-(G^{\text{dmu}}_{\text{s,s}} - G^{\text{dmu}}_{\text{a,s}})/k_B T} + e^{-(G^{\text{dmu}}_{\text{a,s}} - G^{\text{dmu}}_{\text{a,s}})/k_B T} + e^{-(G^{\text{dmu}}_{\text{a,a}} - G^{\text{dmu}}_{\text{a,s}})/k_B T}\right) \quad (5)$$

Due to the presence of saturated ethylene bridges, the host molecule is also capable of existing in three rotational isomeric states: gauche+/gauche+, gauche+/gauche−, and gauche−/gauche+. However, as discussed below, the energy barrier between the three states is sufficiently low that they are adequately sampled during a conventional simulation.

As a consequence of the rotational isomers, the thermodynamic cycle is modified. To calculate the difference in the free energy of binding of g_1 and g_2 for the case in which several rotational isomers of g_2 exist, the thermodynamic cycle is modified so that g_1 is first perturbed to one of the rotational isomers g_{2j} of g_2, as shown in eq 6.



$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_4 + \Delta G_6 - \Delta G_3 - \Delta G_5 \quad (6)$$

The rotational isomer g_{2j} is then perturbed to the full ensemble of rotational isomers of g_2.

Molecular Models

The structure of the host molecule shown in Figure 1 was obtained by optimization with MOPAC 5.0[32] using the AM1 Hamiltonian and is in close agreement with the crystal structure of the host–imidazolidone complex. The atomic charges were calculated for the optimized MOPAC structure using the CHELP[33] program. CHELP computes these charges by fitting the electrostatic potential generated from an ab initio wavefunction, STO-3G in this case, to a point charge model. Optimally,

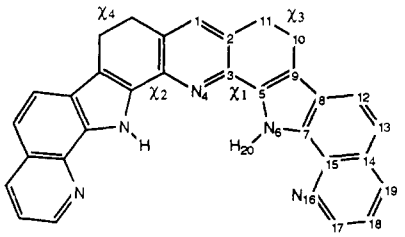


Figure 1. Structure of host with variable dihedral angles and atom-numbering scheme shown.

(31) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1989**, *90*, 3300.

(32) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1.

(33) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.

Table I. Atomic Charges Derived from MOPAC 5.0, GAUSSIAN 90, and CHELP Calculations for Host Molecule[a]

atom	MOPAC	G90	CHELP	σ	ϵ
H1C1	0.044	0.004	−0.032	3.750	0.110
C2	−0.956	0.002	−0.030	3.750	0.110
C3	0.054	0.083	0.233	3.750	0.110
N4	−0.150	−0.267	−0.301	3.250	0.170
C5	−0.016	0.095	−0.138	3.750	0.110
N6	−0.130	−0.315	−0.010	3.250	0.170
C7	0.008	0.110	−0.176	3.750	0.110
C8	−0.085	−0.028	0.044	3.750	0.110
C9	−0.113	−0.029	0.008	3.750	0.110
C10	0.100	0.017	−0.020	3.905	0.118
C11	0.074	0.026	0.083	3.905	0.118
C12	0.069	0.004	−0.042	3.750	0.110
C13	−0.015	−0.016	−0.020	3.750	0.110
C14	−0.061	−0.007	−0.097	3.750	0.110
C15	0.014	0.074	0.437	3.750	0.110
N16	−0.142	−0.258	−0.489	3.250	0.170
C17	0.099	0.110	0.283	3.750	0.110
C18	−0.042	−0.008	−0.118	3.750	0.110
C19	0.057	0.021	0.072	3.750	0.110
H20	0.278	0.252	0.148	0.000	0.000

[a] Lennard-Jones parameters are given by σ and ϵ. Numbering scheme is shown in Figure 1.

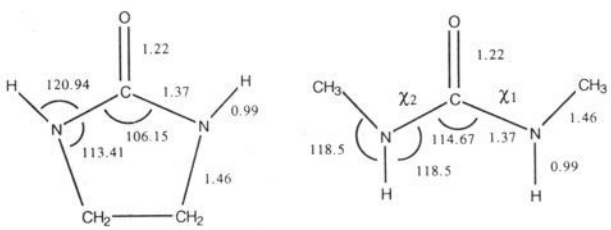


Figure 2. Structures of guests. *N,N'*-dimethylurea has two variable dihedral angles, χ_1 and χ_2, and imidazolidone has none in the models used.

a higher level basis set such as 6-31G(d) would be preferable, but the time and cost of such a calculation for 37 heavy atoms was judged to be prohibitive.

A comparison of the MOPAC Mulliken[34] charges, ab initio STO-3G Mulliken charges, and CHELP charges is provided in Table I. The CHELP charges were used in the simulations.

Optimized structures for the guest molecules dimethylurea and imidazolidone were determined at the 3-21G basis set level using the program GAUSSIAN 90.[35] The results for the three rotameric states of dimethylurea show that the a,s rotamer was 0.5 kcal mol^{-1} more stable than the s,s rotamer and 8.2 kcal mol^{-1} more stable than the a,a rotamer.

For simulation purposes, dimethylurea was modeled using an approximation in which the bond lengths and bond angles do not change from one isomeric state to another but dihedral angles are allowed to change. The amount of internal freedom in the imidazolidone is small, and it was modeled as a rigid planar structure for use in the simulations. Dimethylurea and imidazolidone geometries are summarized in Figure 2. Charges and Lennard-Jones parameters for the guest carbonyl carbon and oxygen are from the analogous OPLS urea atoms. The guest amide nitrogen, hydrogen, and methyl group were obtained from the respective OPLS amino acid terminal amide atoms with the nitrogen and methyl charges scaled so that there is no overall charge on the molecule. The methyl carbons on dimethylurea are changed to imidazolidone carbons by changing the Lennard-Jones parameters to those for the OPLS lysine ϵ-carbon. When combined with the rigid-rotor conformational analysis (below), the charges result in a net dipole of 5.11, which is in excellent agreement with the experimental value in benzene of 5.1.[36] Charges and Lennard-Jones parameters are summarized in Table II.

Table II. Charges and Lennard-Jones Parameters for Dimethylurea (dmu) and Imidazolidone[a]

molecule	atom	charge	σ	ϵ
imidazolidone,dmu	O	−0.390	2.96	0.210
imidazolidone,dmu	C	0.142	3.75	0.105
imidazolidone,dmu	N	−0.4660	3.25	0.170
imidazolidone,dmu	H	0.370	0.0	0.0
dmu	CH_3	0.2200	3.775	0.170
imidazolidone	CH_2	0.2200	3.905	0.118

[a] Source is OPLS parameters and OPLS-derived parameters as described in the text.

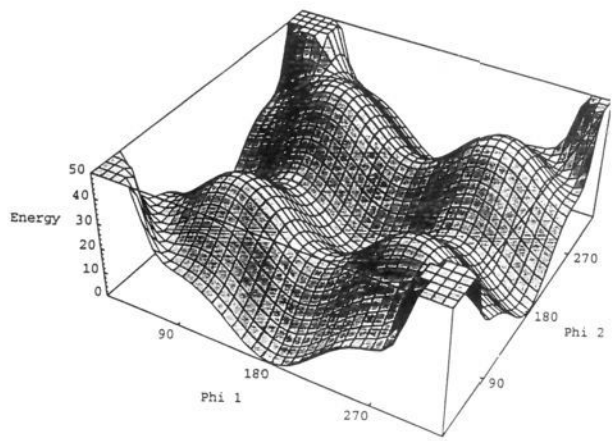


Figure 3. Rigid-rotor plot of ϕ_1 and ϕ_2 for dimethylurea. Energy is in kcal mol^{-1}. Minima at (0°,180°), (180°,0°), and (180°,180°) correspond to the a,s, s,a, and s,s isomers of dimethylurea, respectively.

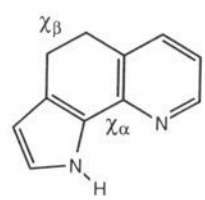


Figure 4. Analogue of host molecule used for ab initio studies of ring motions. χ_α and χ_β denote dihedral angles.

The energy surface obtained by varying the two dihedral angles of dimethylurea is shown in Figure 3. In this model the amide groups are locally planar and not allowed to pyramidalize. As can be seen, energy minima exist at values of ϕ_1 and ϕ_2 corresponding to the s,s and a,s rotamers. As ϕ_1 and ϕ_2 approach 0° from either direction, the energy increases markedly.

Since the inclusion of internal degrees of freedom for a closed ring is difficult due to correlated ring motions,[37-39] modeling the motion of the ethylene bridge of the host molecule was achieved by fitting Fourier terms to an adiabatic energy map of a small analogue of the host (Figure 4).

The analogue comprises the central pyridine ring and the adjacent six-member and pyrrole rings. Shown are the two dihedral angles that can deviate significantly from planarity, χ_α and χ_β. Because of correlated ring motions, these angles are not independent of one another and the motion of χ_β can be described for simulation purposes by using a pseudodihedral function on χ_α. In this model, the two saturated hydrocarbons are rigid but move independently of each other and χ_α follows the energy surface of the ring found from STO-3G ab initio calculations.

(34) Mulliken, R. S. *J. Chem. Phys.* **1962**, *36*, 3428.
(35) Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M. A.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A. *GAUSSIAN 90*; Gaussian, Inc.: Pittsburgh, PA, 1990.
(36) Devoto, G. *Gazz. Chim. Ital.* **1933**, *63*, 495.
(37) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304.
(38) Chandrasekhar, J.; Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5073.
(39) Chandrasekhar, J.; Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5080.

Table III. Fourier Coefficients for Rotation about Dihedral Angles[a]

molecule	dihedral	bonds	Fourier coefficients V_0	V_1	V_2	V_3
host	χ_1,χ_2	N—C—C—N	−0.119	−3769.738	7354.247	3772.150
dimethylurea	ϕ_1,ϕ_2	N—C(=O)—N(H)—C	0.000	2.800	21.200	0.000

[a] Coefficients for host were derived as described in text; dimethylurea coefficients are from the BOSS program. Units for the Fourier coefficients are in kcal mol^{-1}.

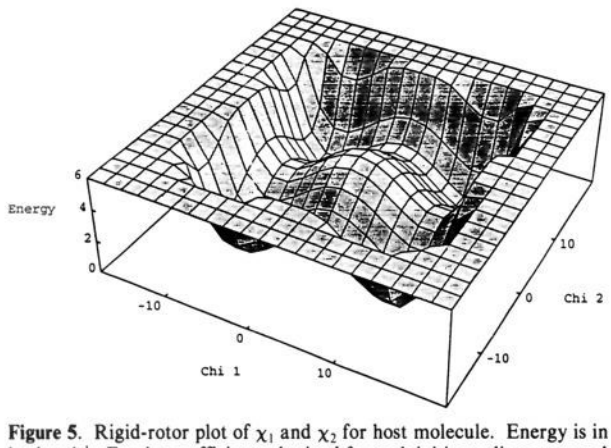


Figure 5. Rigid-rotor plot of χ_1 and χ_2 for host molecule. Energy is in kcal mol^{-1}. Fourier coefficients obtained from ab initio studies were used to generate an energy surface with the correct minima and maxima as described in the text.

Ab initio calculations of the analogue found energy minima corresponding to two gauche conformations of the ethylene bridge at approximately 10° and −10°; mode following was used to follow the one imaginary eigenvalue of the Hessian matrix back to a transition state[40,41] that corresponded to a flat C_s molecule that had the ethylene bridge in an eclipsed form. The difference in energy between either minima and the transition state was approximately 2 kcal mol^{-1}.

In the simulations the dihedral angle potential energy including nonbonded 1,4-interactions was represented by

$$V(\phi) = V_0 + \tfrac{1}{2}V_1(1 + \cos\phi) + \tfrac{1}{2}V_2(1 - \cos 2\phi) + \tfrac{1}{2}V_3(1 + \cos 3\phi) \quad (7)$$

Using the method of Chung-Phillips,[42,43] the energy surface obtained from the ab initio calculations was used to fit Fourier coefficients for use in eq 7. The bond between the two aliphatic hydrocarbons of the analogue was included by using a high-energy penalty via the Fourier coefficients for any value of the N–C–C–N dihedral angle (χ_α) that is much greater than 15°, which would correspond to an unreasonable stretching of the CH_2–CH_2 bond. Also included in the Fourier coefficients are the 1,5 and greater nonbonded neighbor interactions in the analogue. Table III lists the Fourier coefficients obtained and also those for dimethylurea which are from the OPLS force field. Although the coefficients found from the fitting are quite high, they correctly reproduce the energy surface as can be seen in the resulting rigid-rotor map (Figure 5), where χ_1 and χ_2 are the two N–C–C–N dihedrals of the host molecule. As can be seen, this energy surface is characterized by local minima at $\chi_1 = 10°$, $\chi_2 = 10°$; $\chi_1 = 10°$, $\chi_2 = -10°$; $\chi_1 = -10°$, $\chi_2 = 10°$; and $\chi_1 = -10°$, $\chi_2 = -10°$.

Monte Carlo Simulations

Monte Carlo simulations were carried out with Jorgensen's MC program BOSS 2.8,[44] which was modified to include intramolecular energies in the calculation of free energy differences. Also, the program was modified so that dihedral potentials could be perturbed simultaneously with bonds and bond angles. Simulations of the guest molecules were carried out in a box of chloroform 33.5 Å on a side that included 263 chloroform molecules, while the host and guest system was simulated in a box 38.5 Å on a side that included 417 chloroform molecules. The cutoff distance used was 12 Å in all cases, and preferential sampling was employed. Dihedral angles were allowed to change as much as 15°, and an attempt to move a particular dihedral angle was made about every 50 MC moves. The isothermal–isobaric simulation was carried out at 298 K and 1 atm. The simulation was broken into a window of single-wide sampling at each end point and four windows of double-wide sampling at intermediate steps in the simulation. Independent forward and reverse simulations were carried out in which equilibration was for 1 million MC moves, and each of the 60 MC averaging observations consisted of 50 000 random moves for a total of 3 million moves. In addition to the calculation of thermodynamic values in the forward and reverse simulations, the data collected were used to calculate overall thermodynanic values as if the two simulations were one long simulation with a break in the Markov chain using eq 8.

$$\Delta G = -RT \ln\left(\frac{1}{N}\sum_{i=\text{forward}}^{\text{reverse}}\sum_{j=1}^{60}\sum_{k=1}^{50\,000} e^{-\Delta E_{ijk}/k_B T}\right) \quad (8)$$

Here the index k is used for labeling an MC observation (a block of 50 000 moves), j is used to sum each of the 60 observations for a forward or reverse simulation, i sums the forward and reverse results, and $N = 6 \times 10^6$. This leads to adequate sampling of phase space and allows for the calculation of overall statistical errors.

Table IV. Free Energy Changes from Simulation in $CHCl_3$ at 25 °C and 1 atm

perturbation	ΔG, kcal mol^{-1}
a,s → s,s (in vacuo)	0.47 ± 0.00
a,s → s,s	0.18 ± 0.07
a,s → imidazolidone	−0.54 ± 0.06
host: a,s → host: s,s	6.94 ± 0.18
host: s,s → host: a,s	−5.19 ± 0.18
host: a,s → host: imidazolidone	−4.24 ± 0.15

Results and Discussion

Initial simulations of dimethylurea were carried out in vacuo and in solution ($CHCl_3$) in order to study the relative stabilities of the various rotamers and to get an estimate of the free energy of solvation. In vacuo the *anti,anti*-dimethylurea isomer was not stable and immediately isomerized to the a,s isomer; the solvated species behaved much the same, isomerizing within several thousand MC moves. As can be seen in Figure 3, the low stability of the a,a isomer is due to an internal energy surface that has a maximum in the region of the a,a isomer (χ_1, χ_2 near 0°). Although the shape of the energy surface is dictated by the potentials used (Table II) and the fact that dimethylurea is modeled with rigid bond angles, the low stability is consistent with the ab initio calculations that showed that the a,a isomer was 7.7 kcal mol^{-1} less stable than the s,s isomer. In contrast, both the a,s and s,s isomers were stable and did not undergo isomerization. The a,s isomer is slightly lower in free energy relative to the s,s isomer in vacuo and in solution, with $\Delta G = -0.47 \pm 0.00$ and -0.18 ± 0.07, respectively, as shown in Table IV. The result is a small relative shift of the a,s isomer to the s,s isomer upon transfer of dimethylurea to solution: $\Delta\Delta G_{\text{solvation}} = -0.28$ kcal mol^{-1} in favor of the s,s isomer. The solute–solvent energy difference between the s,s and a,s rotamers actually favors the s,s form by ca. 0.5

(40) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385.
(41) Schlegel, H. B. *Ab Initio Methods in Quantum Chemistry, Part 1*; Wiley: New York, 1987.
(42) Chung-Phillips, A. *J. Chem. Phys.* **1988**, *88*, 1764.
(43) Chung-Phillips, A. *J. Comput. Chem.* **1989**, *10*, 733.
(44) Jorgensen, W. L. BOSS Version 2.8. Yale University, New Haven, CT, 1989.

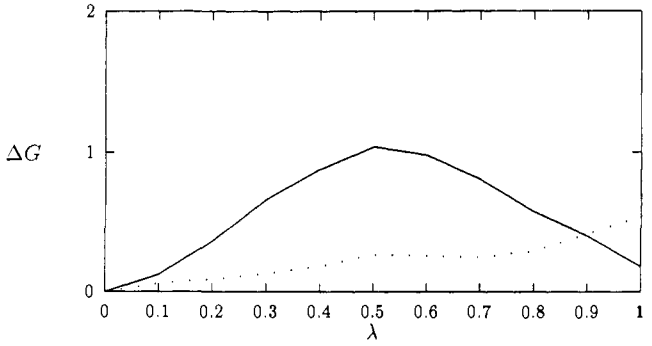


Figure 6. Perturbation profile of free energy (kcal mol^{-1}) versus λ for guest molecules in solution: a,s → s,s (dotted line), a,s → imidazolidone (solid line).

kcal mol^{-1}. The exact details of these energy orderings will vary with details of the molecular mechanics model; e.g., the use of flexible bond angles would lead to slightly different results. However, these variations are not expected to alter the final $\Delta\Delta G$ of binding significantly.

Using eq 5, the free energy change in solution going from the a,s rotamer of dimethylurea to all rotamers is then $\Delta G_5 = -0.33$ kcal mol^{-1} in the thermodynamic cycle of eq 6.

Results for the solvated imidazolidone to *anti,syn*-dimethylurea simulation show that $\Delta G_3 = 0.54 \pm 0.06$ kcal mol^{-1} in favor of the imidazolidone. Thus, the solvation free energy difference in solution between dimethylurea and imidazolidone is $\Delta G_{solvation} = \Delta G_3 + \Delta G_5 = 0.21$ kcal mol^{-1} in favor of the imidazolidone.

Figure 6 shows the free energy change for the solvated, uncomplexed guest molecules as a function of the simulation variable λ. The simulations were well behaved with standard deviations for the incremental changes in free energy not exceeding ±0.03 kcal mol^{-1}. The simulations appeared to converge within about half of the six million total MC moves.

Initial structures for the host–guest complex in solution were determined by in vacuo potential of mean force calculations along the host–guest intermolecular axis. At the lowest energy position, simulated annealing techniques were employed to find a single structure to be used as an initial configuration for the solvated complex.

In the optimized structure of the imidazolidone–host complex (Figure 7), the imidazolidone lies in the cavity of the host molecule and is essentially coplanar with it. Hydrogen bonding between the imidazolidone carbonyl oxygen and host pyrrole hydrogens and between the imidazolidone amide hydrogens and host pyridine nitrogens is seen. In this structure, the imidazolidone oxygen lies 3.0 Å from the central pyridine nitrogen and 2.1 Å from either pyrrole amide hydrogen.

The initial *anti,syn*-dimethylurea–host structure is similar to the initial structure for the imidazolidone–host complex (Figure 7), except that the guest molecule is twisted out of the plane of the host as a result of steric interactions of the guest methyl group syn to the oxygen with the host. The dimethylurea carbonyl oxygen again lies 3.0 Å from the central pyridine nitrogen of the host; the pyrrole amide hydrogens lie 2.1 and 2.6 Å from the carbonyl oxygen. As a consequence of the loss of the second guest amide hydrogen bond interaction, there are now at least two binding modes for *anti,syn*-dimethylurea; the second mode involving rotation of *anti,syn*-dimethylurea by 180° and interaction between the dimethylurea amide hydrogen and the other host pyridine nitrogen.

In contrast, the *syn,syn*-dimethylurea binds nearly perpendicularly to the host molecule due to steric interactions involving the two methyl groups (Figure 7). The pyridine nitrogen and guest oxygen distance is unaltered at 3.0 Å, but the pyrrole amide hydrogen to carbonyl oxygen distance is now symmetrical at 2.1 Å.

When the *anti,anti*-dimethylurea isomer is bound as a guest molecule in the cavity of the host, one of its dihedral angles again rotates 180°, transforming the a,a isomer into the a,s isomer. Even

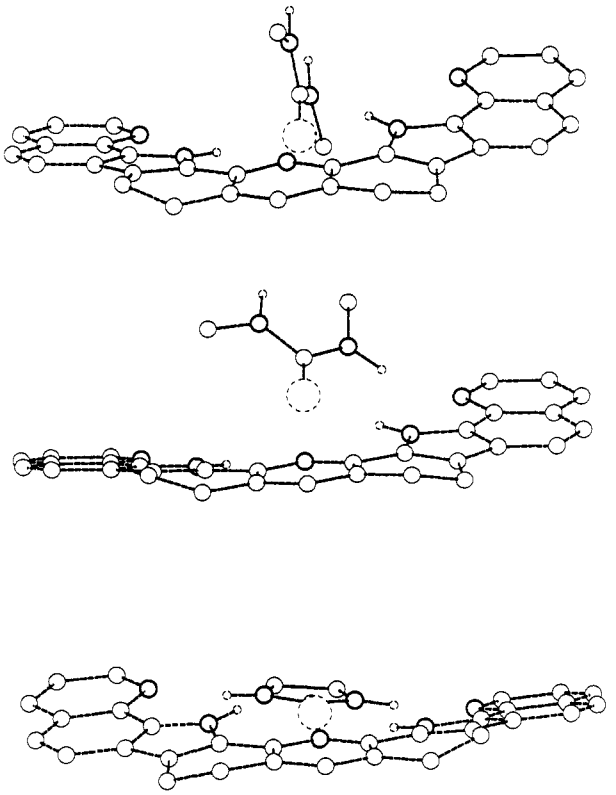


Figure 7. Initial structures of host–guest complexes: *syn,syn*-dimethylurea (top); *anti,syn*-dimethylurea (middle); imidazolidone (bottom).

though the formation of an additional hydrogen bond will favor the population of the a,a state relative to its in vacuo or solvated counterpart, the interaction energy between the two methyl groups dominates. This syn–anti isomerization occurs for the host–guest system both in vacuo and in solution.

Binding to the host by dimethylurea induces a population shift from the s,s to the a,s rotamer of the guest. The complex gains an additional hydrogen bond by this conformational change and permits the a,s rotamer to move toward the approximate plane of the host. In solution, the free energy lost by the a,s → s,s isomerization is 6.94 ± 0.18 kcal mol^{-1} (Table IV). The reverse simulation s,s → a,s surprisingly shows a free energy change of only −5.19 ± 0.18 kcal mol^{-1}. However, analysis of the simulations shows that a third, less favorable binding mode of *anti,syn*-dimethylurea to the host has appeared in which this rotamer binds in a perpendicular fashion to the host, in contrast to the structures found by simulated annealing. The carbonyl oxygen still binds the two host pyrrole hydrogens, however the dimethylurea amide hydrogen that is syn to the carbonyl oxygen now interacts favorably with the central pyridine nitrogen as shown in Figure 8. The dimethylurea syn methyl group appears to be in a stable van der Waals complex with the terminal pyridine rings of the host.

Again the simulations are well behaved, as can be seen in Figure 9, with standard deviations for the incremental changes in free energy not exceeding ±0.07 kcal mol^{-1}. The free energy change for λ = 0.00 → 0.10 and λ = 0.10 → 0.00 was somewhat large, ca. 2.0 kcal mol^{-1}, indicating a possible need to decrease the increment in λ. However, when the steps λ = 0.00 → 0.05 and λ = 0.10 → 0.05 are used, the free energy change is identical to that for λ = 0.00 → 0.10. Convergence of individual simulations is quite good (within approximately three million MC moves).

The apparent hysteresis (6.94 vs 5.19 kcal mol^{-1}) is actually not a hysteresis at all, since the two values of the free energy obtained are for two different processes: 6.94 kcal mol^{-1} is the work for the process of changing *anti,syn*-dimethylurea bound to the host in the manner of Figure 7 to the bound *syn,syn*-dimethylurea, while 5.19 kcal mol^{-1} is the work for the process of changing *anti,syn*-dimethylurea bound to the host in the manner

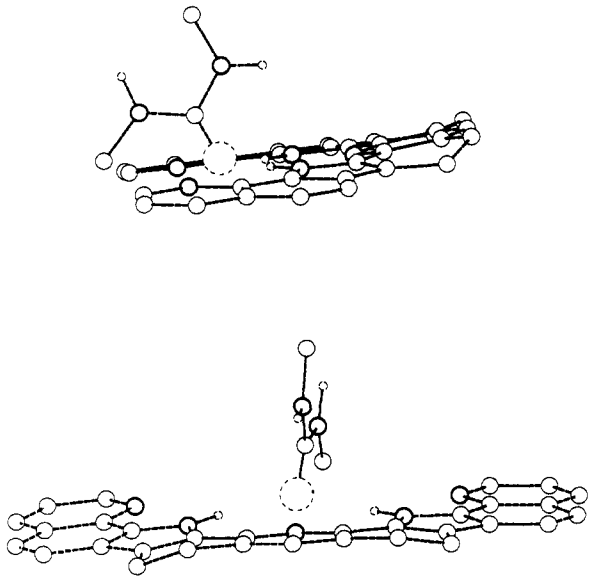


Figure 8. Side and front view of the third binding mode of *anti,syn*-dimethylurea to the host found in the simulation *syn,syn*- → *anti,syn*-dimethylurea, as described in the text.

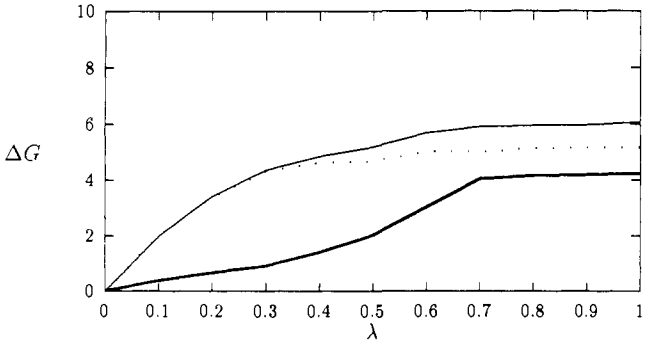


Figure 9. Perturbation profile of free energy (kcal mol^{-1}) versus λ for guest molecules while complexed to the host: s,s ← a,s (upper solid line), s,s → a,s (dotted line), imidazolidone → a,s (bold-face lower solid line).

of Figure 8 to the bound *syn,syn*-dimethylurea. Contribution of the new binding mode is calculated in analogy to rotational isomers using eq 4. This gives $\Delta G_6 = -0.43$ kcal mol^{-1} in the thermodynamic cycle of eq 6. The only significant contribution to the free energy change here comes from the two-fold degeneracy of the binding mode shown in the center of Figure 7. The binding mode shown in Figure 8 has a negligible population because of its higher free energy.

Simulation results for the complexed perturbation *anti,syn*-dimethylurea to imidazolidone involve the addition of a fourth hydrogen bond and the loss of the steric interaction due to the methyl group syn to the carbonyl oxygen. This results in the imidazolidone relaxing to a roughly coplanar configuration with the host molecule, in agreement with the energy-minimized complex and the X-ray structure of the complex. The resulting free energy change is $\Delta G_4 = 4.24 \pm 0.15$ kcal mol^{-1} (Table IV). The simulations are well behaved (Figure 9) and convergent, with standard deviations for the incremental changes in free energy not exceeding ±0.09 kcal mol^{-1}.

When rotational isomers of dimethylurea are taken into account, one determines $\Delta G_{binding} = \Delta G_4 + \Delta G_6 = 3.81$ kcal mol^{-1} in favor of imidazolidone over dimethylurea. Combining the binding results with the solvation results then gives a $\Delta\Delta G_{binding}$ of 3.60 kcal mol^{-1} in favor of imidazolidone, which compares with the experimentally determined value of 3.1 kcal mol^{-1}.[29] In light of previous results for host–guest systems,[25] this may seem to be a fair discrepancy between the experimental and simulation results. Much of the discrepancy can be attributed to the level of development of the potential parameters. In particular, the use of the STO-3G basis set in the estimation of partial charges for the host may not be as accurate as desired. Optimally, a higher level basis set such as 6-31G(d) would be preferable, but the time and expense of such a calculation for 37 heavy atoms would be prohibitive. Although the potential parameters for the guest were derived from OPLS parameters, they were not tested as rigorously as OPLS parameters. From a qualitative perspective, however, the results here are in agreement with experiment.

Concluding Remarks

Consideration of the possible isomers of the guest is important for a proper treatment of the system considered here. The free dimethylurea guests are almost as likely to be in the s,s conformation as the a,s conformation. Organization of the guests into the a,s conformation required for binding costs 0.3 kcal mol^{-1}, enough to reduce the binding constant of dimethylurea relative to the preorganized imidazolidone by ca. 60%. The a,a conformation of dimethylurea would be more complementary to the host than the a,s conformation but is excluded because of intramolecular steric strain in this model that is not adequately compensated by the improved host–guest interactions. The primary reason that imidazolidone binds to the host more strongly than dimethylurea is that its cyclic structure eliminates this steric strain factor and allows the guest to form the optimal hydrogen bonds to the host.

Acknowledgment. We thank Professor W. L. Jorgensen for providing the BOSS software. W.R.C. thanks Tjerk Straatsma for insightful discussions. This work was supported in part by grants from the NSF, the Robert A. Welch Foundation, and SDSC. W.R.C. is supported in part by the NIH Houston Area Molecular Biophysics Training Program. J.A.M. is the recipient of the George H. Hitchings Award from the Burroughs Wellcome Fund.