HRMS m/z for $C_{46}H_{64}N_4O_6$ (M^+) calcd 768.4825, found 768.4825.

2,6-Naphthalenediamide Tricyclic Dilactams 21. A solution of **20a** (0.70 g, 0.91 mmol) in EtOH (60 mL) and THF (60 mL) was stirred with $NaBH_4$ (1.40 g) at room temperature for 24 h. The solution was poured into water (200 mL), and the aqueous solution was extracted with $CHCl_3$ (2 × 200 mL). The organic layer was washed with brine and dried over anhydrous Na_2SO_4. The solvent was removed in vacuo, the residue was taken up in CH_2Cl_2 (50 mL), a catalytic amount of *p*-TsOH was added, and the mixture was stirred for 3 h at room temperature. The solution was washed with saturated $NaHCO_3$ and brine and then dried with anhydrous Na_2SO_4. The solvent was evaporated in vacuo, and the resulting mixture of two diastereomers was separated by flash chromatography (16% $EtOAcCH_2Cl_2$) to yield a less polar diastereomer **21** (0.30 g, 45%) and a more polar diastereomer **21** (0.26 g, 39%) as white solids. Less polar diastereomer (meso) **21**: mp 298-300 °C; IR (NaCl, neat) 3220, 2958, 2931, 2871, 1676, 1440, 1376, 1259, 1084, 743 cm^{-1}; ^1H NMR ($CDCl_3$, 250 MHz) δ 7.89 (d, 2 H, J = 8.9 Hz), 7.70 (d, 2 H, J = 1.8 Hz), 7.31 (dd, 2 H, J_1 = 8.9 Hz, J_2 = 1.8 Hz), 5.88 (d, 2 H, NH, J = 3.1 Hz), 4.70 (d, 2 H, J = 3.1 Hz), 2.22 (d, 2 H, J = 14.1 Hz), 2.09-0.80 (m, 50 H); HRMS m/z for $C_{46}H_{64}N_4O_4$ (M^+) calcd 736.4927, found 736.4927. More polar diastereomer (racemic) **21**: mp > 300 °C; IR (NaCl, neat) 3345, 2957, 2929, 2872, 1693, 1646, 1457, 1383, 1235, cm^{-1}; ^1H NMR ($CDCl_3$, 250 MHz) δ 7.89 (d, 2 H, J = 8.9 Hz), 7.70 (d, 2 H, J = 1.8 Hz), 7.31 (dd, 2 H, J_1 = 8.9 Hz, J_2 = 1.8 Hz), 5.88 (d, 2 H, NH, J = 3.1 Hz), 4.68 (d, 2 H, J = 3.1 Hz), 2.22 (d, 2 H, J = 14.1 Hz), 2.09-0.80 (m, 50 H): HRMS m/z for $C_{46}H_{64}N_4O_4$ (M^+) calcd 736.4927, found 736.4927.

2,6-Naphthalenediamide Dilactams 22a and 22b. Racemic tricyclic dilactam **21** (0.20 g, 0.27 mmol) was dissolved in CF_3CO_2H (3 mL) and Et_3SiH (0.2 mL). The mixture was stirred overnight at room temperature. The mixture was concentrated in vacuo and the residue was taken up in CH_2Cl_2 (15 mL), the organic phase was washed with saturated $NaHCO_3$, brine, then dried with anhydrous Na_2SO_4. The crude product was purified by flash chromatography (80% $EtOAc/CH_2Cl_2$) to yield pure **22** (racemic, 0.13 g, 64%) as a white solid. mp 174-176 °C; IR (NaCl, neat) 3432, 3197, 2957, 2917, 2871, 1684, 1662, 1585, 1525, 1457, 1393, 1269, 1156 cm^{-1}; ^1H NMR ($CDCl_3$, 250 MHz) δ 8.00 (d, 2 H, J = 1.7 Hz), 7.71 (d, 2 H, J = 8.7 Hz), 7.60 (s, 2 H, NH), 7.46 (dd, 2 H, J_1 = 8.7 Hz, J_2 = 1.7 Hz), 5.34 (s, 2 H, NH), 3.17 (d, 2 H, J = 11.9 Hz), 2.99 (d, 2 H, J = 11.9 Hz), 2.89 (d, 2 H, J = 13.6 Hz), 2.24 (d, 2 H, J = 15.1 Hz), 2.00 (m, 2 H), 1.85 (d, 2 H, J = 12.2 Hz), 1.61-0.88 (m, 46 H); HRMS m/z for $C_{46}H_{68}N_4O_4$ (M^+) calcd 740.5240, found 740.5240.

Racemic **22** was chromatographed on a Pirkle column (L-3,5-dinitrophenylglycine, Regis Chem. Co.) to yield a less polar enantiomer **22b** ($[\alpha]_D$ = -77.6°, c 1.13 in CH_2Cl_2) and a more polar enantiomer **22a** ($[\alpha]_D$ = +77.4°, c 1.31 in CH_2Cl_2).

2,6-Naphthalenediamide Dilactam 23a (meso). The preparation of **23** is the same as that described for **22a** and **22b** except that the less polar diastereomer **21** (0.20 g, 0.27 mmol) was used instead of the more polar diastereomer **21**: a white solid (0.12 g, 60%); mp > 300 °C; IR (NaCl, neat) 3438, 3174, 2956, 2845, 1675, 1635, 1605, 1539, 1456, 1283 cm^{-1}; ^1H NMR ($CDCl_3$, 250 MHz) δ 8.00 (d, 2 H, J = 1.7 Hz), 7.71 (d, 2 H, J = 8.7 Hz), 7.60 (s, 2 H, NH), 7.46 (dd, 2 H, J_1 =8.7 Hz, J_2 = 1.7 Hz), 5.34 (s, 2 H, NH), 3.17 (d, 2 H, J = 11.9 Hz), 2.99 (d, 2 H, J = 11.9 Hz), 2.89 (d, 2 H, J = 13.6 Hz), 2.24 (d, 2 H, J = 15.1 Hz), 2.00 (m, 2 H), 1.85 (d, 2 H, J = 12.2 Hz), 1.61-0.88 (m, 46 H); HRMS m/z for $C_{46}H_{68}N_4O_4$ (M^+) calcd 740.5240, found 740.5240.

***N*-Butylbarbituric Acid 32.**[20] A solution of diethyl butylmalonate (1.00 g, 4.62 mmol), urea (0.28 g, 4.63 mmol), and sodium ethoxide (9.25 mmol) in absolute ethanol (40 mL) was refluxed for 6 h under N_2 atmosphere. The mixture was concentrated in vacuo, and the residue was triturated with 1 N HCl (100 mL) and filtered to yield a white solid. The mother liquid was saturated with NaCl and extracted with $CHCl_3$ (2 × 100 mL). The organic phase was dried with anhydrous Na_2SO_4 and concentrated in vacuo. Combined crude product was recrystallized from hot $CHCl_3$ to yield a white solid (0.53 g, 63%): mp 211-212 °C; IR (NaCl, neat) 3227, 2960, 2926, 2863, 1684, 1419, 1337, 1209 cm^{-1}; ^1H NMR ($CDCl_3$, 250 MHz) δ 6.05 (s, 2 H, NH), 3.99 (d, 2 H, J = 9.5 Hz), 1.93-1.60 (m, 6 H), 1.00 (d, 6 H, J = 6.2 Hz), 0.95 (d, 6 H, J = 6.2 Hz); HRMS m/z for $C_8H_{12}N_2O_3$ (M^+) calcd 350.1994, found 350.1994.

Acknowledgment. We thank the National Institutes of Health for support of this work and Professor Jorgensen for his advice and calculations. G.D. expresses his gratitude to the Natural Sciences and Engineering Research Council of Canada for a postdoctoral fellowship. M.F. is grateful to the D.A.A.D. for a NATO Science Fellowship.

Chemical Chameleons: Hydrogen Bonding with Imides and Lactams in Chloroform

William L. Jorgensen*,† and Daniel L. Severance

Contribution from the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907. *Received June 7, 1990*

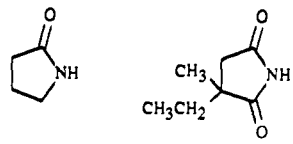
Abstract: Monte Carlo statistical mechanics simulations have been used to elucidate the origin of the novel variations in complex formation observed for imides and lactams. The OPLS potential functions were employed in conjunction with statistical perturbation theory to calculate relative association constants for the dimerization of succinimide and butyrolactam as well as for their cross complex in chloroform. The solution environment significantly dampens the gas-phase preference for the hydrogen bonding with butyrolactam. Consistent with recent experimental results for related intramolecular associations, the symmetry-corrected K_a ratios for lactam-lactam over imide-lactam and imide-lactam over imide-imide are both computed to be about 3, while the differences in optimal gas-phase interactions are nearly 2 kcal/mol. However, the remarkable observations of stronger association of imides rather than lactams with adenines also emerges from similar simulations for complexes of succinimide and butyrolactam with 9-methyladenine. The symmetry-corrected K_a ratios favoring the imide are now ca. 2 and 6 for the Watson-Crick and Hoogsteen orientations. The origin of these variations is shown to arise in a straightforward way from secondary electrostatic interactions.

Introduction

Amido groups are the most common sites for molecular recognition through hydrogen bonding in natural systems. Consequently, it is important to understand the energetic and structural details of such interactions which can often be subtle. A case in point is the notable variation in preferences for association with imides and lactams that has been observed. The self-association of lactams is greater than for related imides. Typical association constants, K_a, in carbon tetrachloride are 100-300 M^{-1} for lactams and 10-60 M^{-1} for imides including butyrolactam and 2-ethyl-2-methylsuccinimide.[1-3] It should be realized that the imide

† Address correspondence to the Department of Chemistry, Yale University, 225 Prospect Street, New Haven, CT 06511.

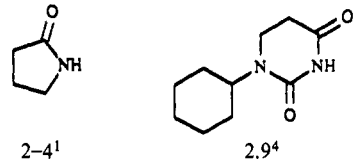
0002-7863/91/1513-209$02.50/0



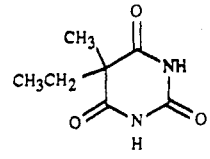
K_a (CCl_4)	145[1]	59[3]
K_a^{corr}	145	15

dimerization has a factor of 4 statistical advantage owing to the four essentially degenerate ways the doubly hydrogen-bonded complex can be formed. Thus, in comparing the strengths of association for a single amido site, the corrected K_a's for imides are 3–15 M^{-1} in CCl_4. These results suggest at first glance that the enhanced N–H acidity of imides does not fully offset the effects of lower oxygen basicity for hydrogen bond formation; however, this notion is challenged below.

In progressing to the more polar solvent, chloroform, the K_a's for lactams and imides are leveled to below 10 M^{-1}.[1-6] Lactam association still appears to be favored after the statistical factor is considered as for butyrolactam and 1-cyclohexyl-5,6-dihydrouracil. Even barbiturates, where the statistical factor is now 16, show K_a's in chloroform of only 2–8 as in the case of barbital below.[5]

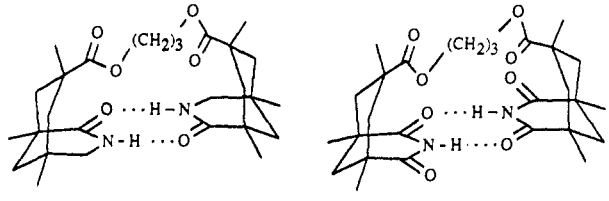


K_a ($CHCl_3$)	2–4[1]	2.9[4]
K_a^{corr}	2–4	0.7



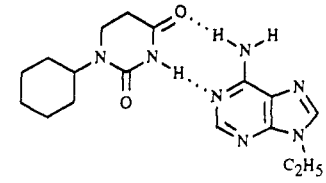
K_a ($CHCl_3$)	7[5]
K_a^{corr}	0.4

The stronger association for lactams than imides in chloroform has recently been clearly established by Rebek and co-workers.[7,8] They used cleverly constrained systems, developed from Kemp's triacid, in which the association is intramolecular. The results are shown below for the compounds with three methylene groups linking the esters.[7] The corresponding imide–lactam has a K_a of 8 M^{-1}, so the corrected enhancement in binding is ca. 3 for each step, imide–imide to imide–lactam to lactam–lactam.



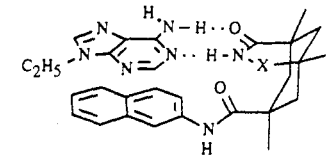
K_a ($CHCl_3$)	24	5.2
K_a^{corr}	24	2.6

Given these observations, it would seem that imides are intrinsically poorer for hydrogen bonding than lactams. Remarkably, this is not the case when comparisons are made with adenine derivatives as the hydrogen-bonding partner. For example, Kyogoku et al. did not observe association between valerolactam



K_a ($CHCl_3$)	30[5]

and 9-ethyladenine (9-Et-A) in chloroform, while complexes with K_a's of 30–120 M^{-1} were formed for 1-cyclohexyl-5,6-dihydrouracil and various barbiturates with 9-Et-A.[5] Furthermore, Rebek and co-workers obtained a K_a of 220 M^{-1} for the illustrated naphthyl imide, while the corresponding lactam exhibits a K_a of only 15.[6,8]

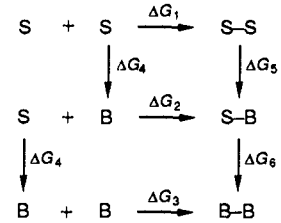


	X = CO	X = CH_2
K_a ($CHCl_3$)	220	15
K_a^{corr}	110	15

The present computational study was undertaken to elucidate these variations in complexation for imides and lactams. Previously, we have developed intermolecular potential functions for chloroform,[9] *cis*-amides (lactams),[10] imides,[11] and nucleotide bases,[12] and we have contributed to the development and application of methodology for computing relative and absolute association constants in solution.[10-14] Combination of these elements could then be used to address the present problem. First, it is established that the computations with the present potential functions do reproduce the observed association patterns in chloroform. Then, the origin of the variations can be analyzed with reasonable confidence. The systems that have been chosen for study compare the self-association and cross-association for butyrolactam (B) and succinimide (S) and the association of S and B with 9-methyladenine (A).

Computational Details

Monte Carlo Simulations. Statistical mechanics calculations were carried out initially to compute the relative free energies of binding of succinimide with itself versus succinimide with butyrolactam versus butyrolactam with itself. The appropriate thermodynamic cycle is shown in eq 1. The relative free energies of binding, $\Delta G_1-\Delta G_2$ and $\Delta G_2-\Delta G_3$,



(1)

are readily obtained[15] by computing ΔG_4, ΔG_5, and ΔG_6 since $\Delta G_1-\Delta G_2$

(1) Krikorian, S. E. *J. Phys. Chem.* **1982**, *86*, 1875.
(2) Kulevsky, N.; Reineke, W. *J. Phys. Chem.* **1968**, *72*, 3339.
(3) Hine, J.; Hahn, S.; Hwang. J. *J. Org. Chem.* **1988**, *53*, 884.
(4) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 250.
(5) Kyogoku, Y.; Lord, R. C.; Rich, A. *Nature* **1968**, *218*, 69.
(6) (a) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1082. (b) Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1090.
(7) Jeong, K.-S.; Tjivikua, T.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 3215.
(8) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J. Jr. *J. Am. Chem. Soc.* Preceding paper in this issue.
(9) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683.
(10) Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212.
(11) Jorgensen, W. L.; Nguyen, T. B.; Boudon, S. *J. Am. Chem. Soc.* **1989**, *111*, 755. Pranata, J.; Jorgensen, W. L. Unpublished results.
(12) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008. Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* Submitted for publication.
(13) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
(14) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 7269.

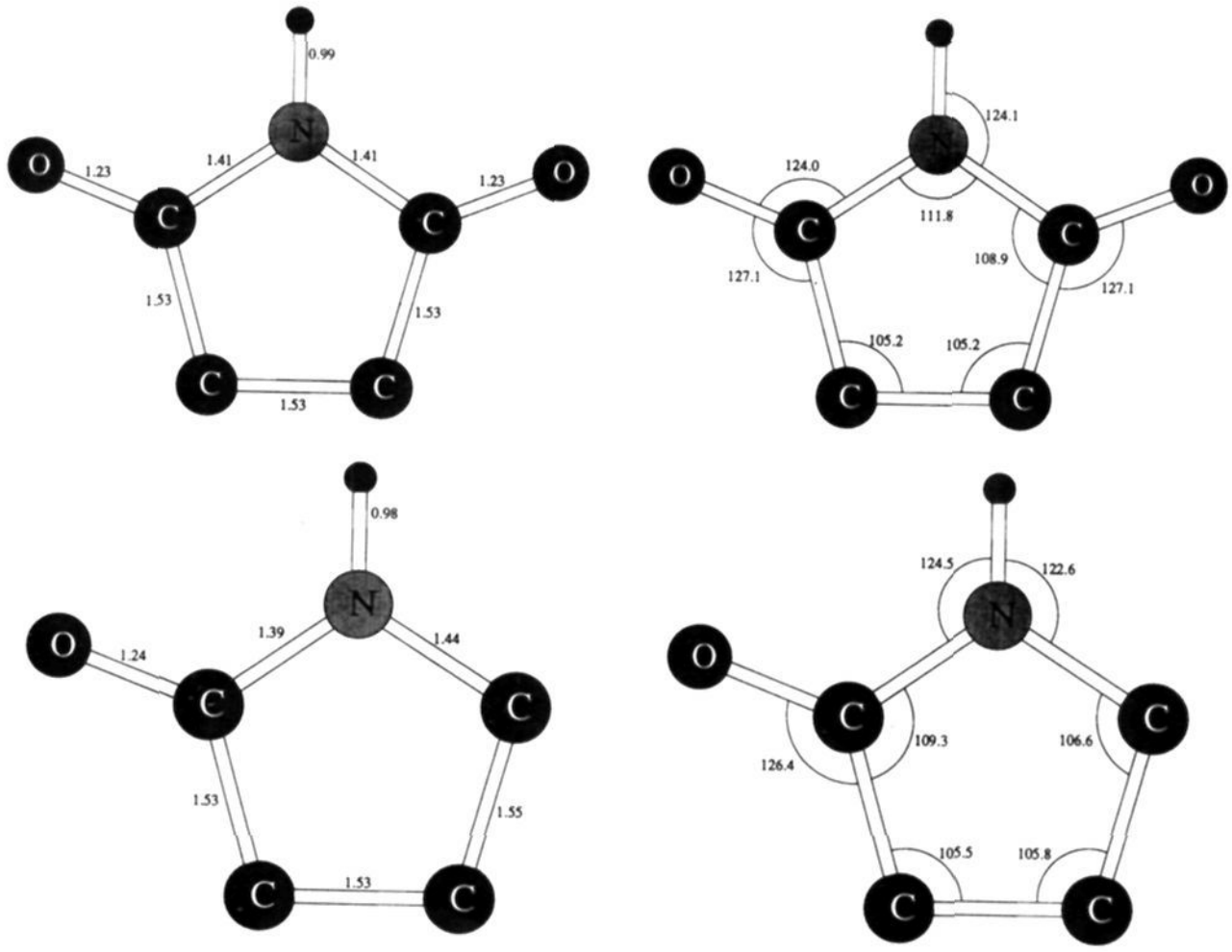


Figure 1. Bond lengths and bond angles for succinimide (top) and butyrolactam (bottom) from AM1 optimizations.

$= \Delta G_4 - \Delta G_5$ and $\Delta G_2 - \Delta G_3 = \Delta G_4 - \Delta G_6$. Each required mutation then involves converting a succinimide molecule to butyrolactam by changing one carbonyl to a methylene group.

In the second series of simulations the binding of succinimide and butyrolactam to 9-methyladenine was compared according to eq 2. Now, $\Delta G_7 - \Delta G_8 = \Delta G_4 - \Delta G_9$ so the only new interconversion is for the complexes with A. However, simulations were carried out for both the Watson–Crick and Hoogsteen forms of the complexes.

$$\begin{array}{ccccc} S & + & A & \xrightarrow{\Delta G_7} & S\text{–}A \\ \downarrow \Delta G_4 & & & & \downarrow \Delta G_9 \\ B & + & A & \xrightarrow{\Delta G_8} & B\text{–}A \end{array} \qquad (2)$$

The free energy changes were computed by using statistical perturbation theory[16] with procedures that are now well-established.[11-14,17] The succinimide to butyrolactam mutations were performed gradually over a series of six Monte Carlo simulations with double-wide sampling[17] which provided 12 incremental free energy changes. All simulations used periodic boundary conditions, Metropolis and preferential sampling, and the isothermal isobaric ensemble at 25 °C and 1 atm.[18] The systems for the transformations in eq 1 consisted of the one or two solutes plus 125 chloroform molecules in a cubic box ca. 26 Å on a side. A larger, rectangular periodic cell (ca. 26 × 26 × 39 Å) with 185 chloroform molecules was used for the mutations of the complexes with 9-methyladenine.

In all cases, the intersolute interactions were always included, while solvent–solvent interactions were spherically truncated at a C–C separation of 12 Å with quadratic feathering over the last 0.5 Å. A solute–solvent interaction was included with the same feathering if the distance between the carbon of chloroform and essentially any non-hydrogen atom in the solute was less than 12 Å. Each simulation consisted of 1.1×10^6 configurations of equilibration followed by 3.0×10^6 configurations of averaging. The solute and solvent molecules were independently translated and rotated at random such that an overall acceptance rate of ca. 40% was obtained for new configurations. No intramolecular degrees of freedom were included in these simulations, i.e., the geometries of the solvent molecules were kept constant. The structures for succinimide and butyrolactam were obtained from full AM1 optimizations[19] and are both planar, as summarized in Figure 1. The geometry for 9-methyladenine is the same as used previously.[12,14] It is also planar and was taken from results of an optimization for adenine via ab initio calculations with the 3-21G basis set.[20] A standard methyl group was attached to N9 with a C–N bond length of 1.45 Å. For chloroform, the experimental geometry was adopted with r(C–Cl) = 1.758 Å and ∠Cl–C–Cl = 111.3°.[9,21]

All computations were carried out with the BOSS program, version 2.8, on SUN/4 and Silicon Graphics 4D computers in our laboratory.

Table I. OPLS Parameters for Imides and Lactams

atom	q, e⁻	σ, Å	ϵ, kcal/mol
	Succinimide		
N	−0.49	3.250	0.170
H (N)	0.37	0.000	0.000
C (O)	0.42	3.750	0.105
O	−0.42	2.960	0.210
CH_2	0.06	3.905	0.118
	Butyrolactam		
N	−0.55	3.250	0.170
H (N)	0.35	0.000	0.000
C (O)	0.53	3.750	0.105
O	−0.53	2.960	0.210
CH_2 (N)	0.20	3.800	0.118
CH_2	0.00	3.905	0.118

Intermolecular Potential Functions. The intermolecular interactions are represented in the Coulomb plus Lennard–Jones format of eq 3.

$$\Delta E_{ab} = \sum_i \sum_j (q_i q_j e^2 / r_{ij} + A_{ij}/r_{ij}^{12} - C_{ij}/r_{ij}^6) \qquad (3)$$

Thus, the interaction between two molecules a and b involves summing over all pairings of the interaction sites i in a and j in b.[22] The A and C parameters are related to Lennard–Jones σ's and ϵ's 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 partial charges and Lennard–Jones parameters used here were developed previously.[9-12] Specifically, the OPLS four-site model with a united atom CH group was adopted for chloroform.[9] The charges for

(15) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281.
(16) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(17) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.
(18) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087. Owicki, J. C. *ACS Symp. Ser.* **1978**, *86*, 159.
(19) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(20) Aida, M. *J. Comput. Chem.* **1988**, *9*, 362.
(21) Jen, M.; Lide, D. R. *J. Chem. Phys.* **1962**, *36*, 2525.
(22) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

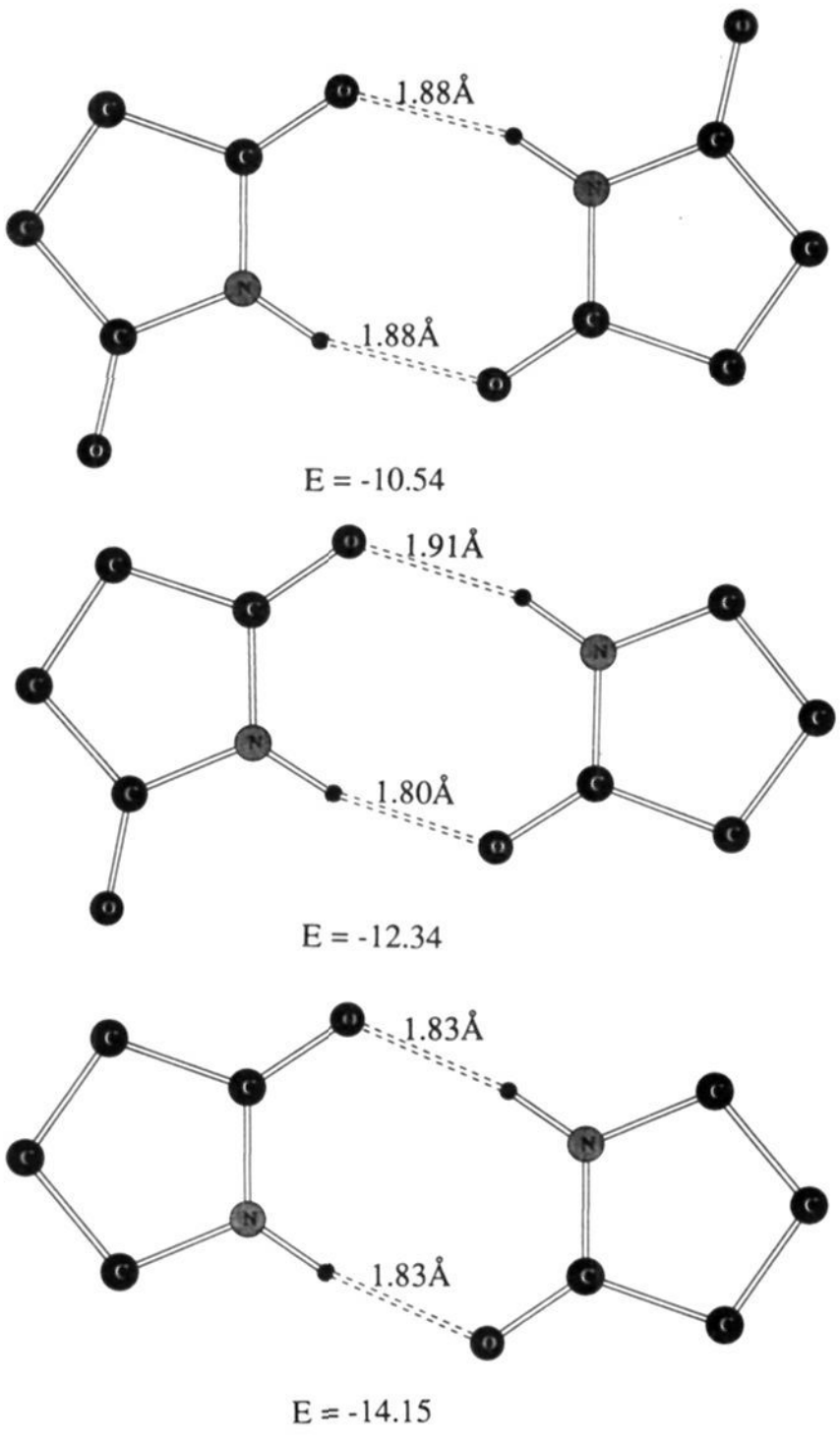


Figure 2. Geometries and interaction energies (kcal/mol) for the lowest energy complexes of succinimide and butyrolactam.

imides and *cis*-amides were obtained by fitting to the results of ab initio calculations with the 6-31G(d) basis set for the geometries and interaction energies of hydrogen-bonded complexes of a water molecule with (*E*)-*N*-methylformamide and (*E*,*E*)-formimide.[10,11] The resulting parameters for succinimide and butyrolactam are summarized in Table I and incorporate standard Lennard–Jones parameters for amides.[23] It should be noted that the methylene groups in these molecules are represented as united atoms centered on carbon, while the remaining atoms are explicit. Also, though the charges were developed independently, it is found that the acidic hydrogen is more positive and the oxygens less negative in the imide than in the amide consistent with traditional ideas from resonance structures. Finally, for 9-methyladenine, the parameters have been described in detail previously.[12] An all-atom representation is used except for a united atom methyl group on N9. These parameters were also derived by fitting to 6-31G(d) results for gas-phase complexes with water, so there is a consistent origin for the solute parameters. Furthermore, the appropriateness of the present intermolecular potential functions is supported by the favorable results on diverse thermodynamic quantities from the previous liquid-phase simulations.[9-14]

Results and Discussion

Complexes with Succinimide and Butyrolactam. To begin, optimizations were carried out with the present OPLS potential functions for the dimers of succinimide and butyrolactam and for the cross complex in isolation. The same results for the lowest energy pairs were obtained either from Simplex optimization or Monte Carlo searches near 0 K with the BOSS program. The resultant structures and interaction energies are shown in Figure

(23) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569.

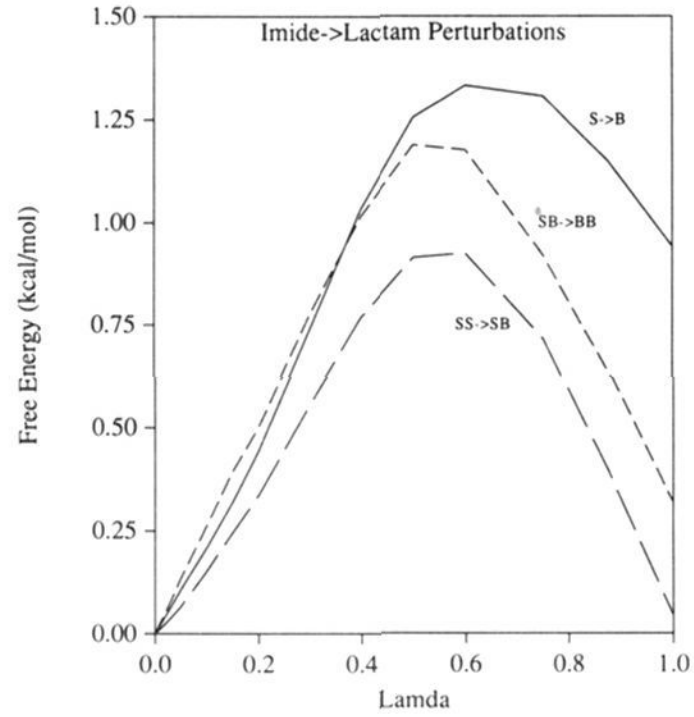


Figure 3. Computed free energy changes vs λ for the imide to lactam perturbations S → B, SS → SB, and SB → BB.

Table II. Computed Free Energy Changes in Chloroform at 25 °C

conversion[a]	ΔG, kcal/mol
S → B	0.94 ± 0.06
S–S → S–B	0.04 ± 0.07
S–B → B–B	0.32 ± 0.09
S–A $\xrightarrow{wc}$ B–A	1.26 ± 0.08
S–A $\xrightarrow{h}$ B–A	1.99 ± 0.09

[a] S = succinimide, B = butyrolactam, A = 9-methyladenine, wc = Watson–Crick orientation, h = Hoogsteen orientation.

2. The hydrogen bonding becomes more favorable in increments of 1.8 kcal/mol as each imide is replaced by a lactam. The overall preference of 3.6 kcal/mol for the lactam dimer over the imide dimer is striking. Though the trend is consistent with the experimental data reviewed in the introduction, the corrected K_a ratios in solution of ca. 3–10 correspond to free energy differences of only 0.7–1.4 kcal/mol. This damping should emerge from inclusion of the solvent environment and configurational averaging in the Monte Carlo simulations.

The progress of the three mutations for eq 1, S → B, SS → SB, and SB → BB, is shown in Figure 3 as a function of the coupling parameter λ which goes from 0 for succinimide to 1 for butyrolactam. The initial geometries for the complexes were doubly hydrogen bonded and oscillated around these forms throughout the simulations. Stereoviews of the last configurations from the simulations of the imide and lactam dimers are illustrative in Figure 4. Naturally, the oxygen not involved in the hydrogen bonding was chosen for mutation; the C=O bond was contracted to 0.5 Å, and all charges, Lennard–Jones parameters, and other geometrical parameters were scaled linearly with λ as succinimide was converted to butyrolactam. The calculations were all well-behaved owing to the length of the averaging runs and the use of 12 increments for the modest mutation. This is apparent in the smoothness of the curves in Figure 3. Also, standard deviations for the incremental free energy changes were computed from separate averages over batches of 2×10^5 configurations and did not exceed 0.04 kcal/mol in any case.

Quantitatively, the overall free energy changes from the S → B, SS → SB, and SB → BB transformations were computed to be $\Delta G_4 = 0.94 \pm 0.06$, $\Delta G_5 = 0.04 \pm 0.06$, and $\Delta G_6 = 0.32 \pm 0.09$ kcal/mol, as summarized in Table II. Combination gives a predicted stronger complexation for succinimide with butyrolactam than with itself of 0.90 ± 0.09 kcal/mol corresponding to $\Delta G_1 - \Delta G_2 = \Delta G_4 - \Delta G_5$. This translates to a corrected K_a ratio of 4.5, while the observable ratio would be half this value owing to the four ways of forming the SS complex as compared to two ways for the SB complex. Similarly, the dimerization of butyrolactam is predicted to be favored over formation of the cross

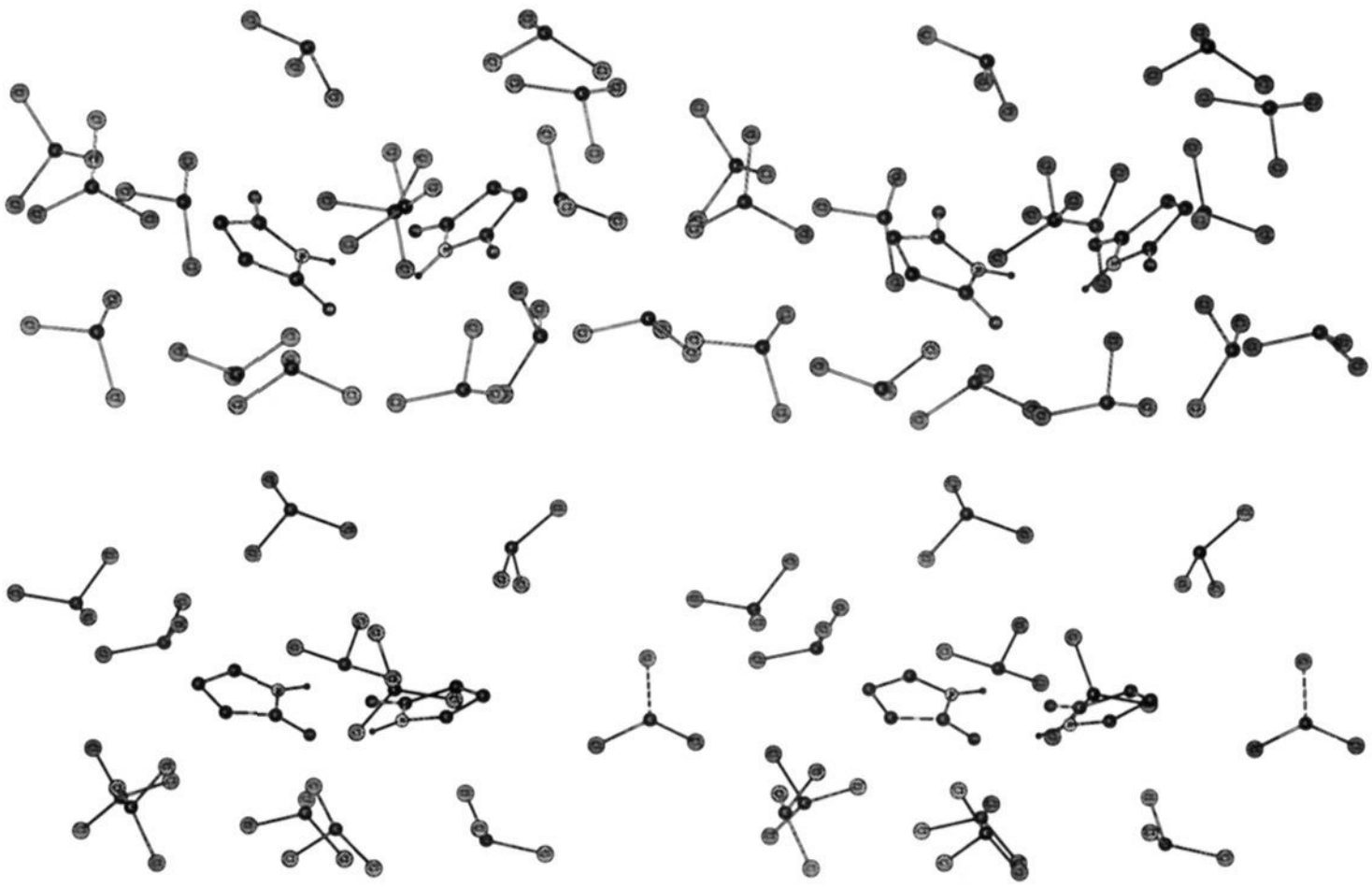
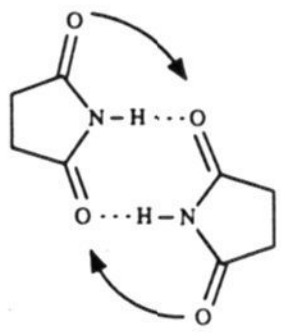


Figure 4. Stereoviews of the last configurations from the simulations of the succinimide (top) and butyrolactam dimers (bottom) in chloroform. The hydrogens in the methylene and methine groups are implicit. Only solvent molecules with an atom within 4 Å of any solute atom are shown.

Chart I



complex by $\Delta G_2 - \Delta G_3 = \Delta G_4 - \Delta G_6 = 0.62 \pm 0.11$ kcal/mol. This gives a corrected K_a ratio of 2.9 or an observable value of 1.4. Overall, these results are in excellent agreement with the available experimental data in chloroform. In particular, it should be recalled that the study by Rebek and co-workers for the formation of the intramolecular complexes with the trimethylene diesters yielded corrected K_a ratios of 3 for both imide–lactam over imide–imide and lactam–lactam over imide–lactam.[7] Thus, the present potential functions appear to represent acceptably the constituent intermolecular interactions, and the simulations properly account for the damping of the gas-phase energetic preferences. The extent of the damping is quantified here and is notable for a solvent like chloroform with a dielectric constant of only 4.8.

The obvious question at this point is why is the intrinsic attraction between imides weaker than for amides? In our recent analysis of variations in complexation for triply hydrogen-bonded systems, the importance of secondary interactions between adjacent hydrogen-bonding sites was pointed out.[12] This notion was applied by Rebek and co-workers to explain the extra destabilization of the interactions with imides as arising from the secondary electrostatic interactions between the non-hydrogen-bonding (spectator) oxygen in one molecule and the hydrogen-bonding oxygen in the other (Chart I).[7,8] The proposal can be tested by decomposing the interaction energies from the OPLS potential functions. First, the variation in the primary interactions can be assessed by computing the interaction energy between just the hydrogen-bonding HNCO fragments for the three optimized complexes in Figure 2. The results are −8.77, −8.72, and −7.36 kcal/mol for the imide dimer, imide–lactam complex, and lactam dimer, respectively. Thus, the contribution from the primary hydrogen-bonding interactions is actually wearker for the lactam dimerization.

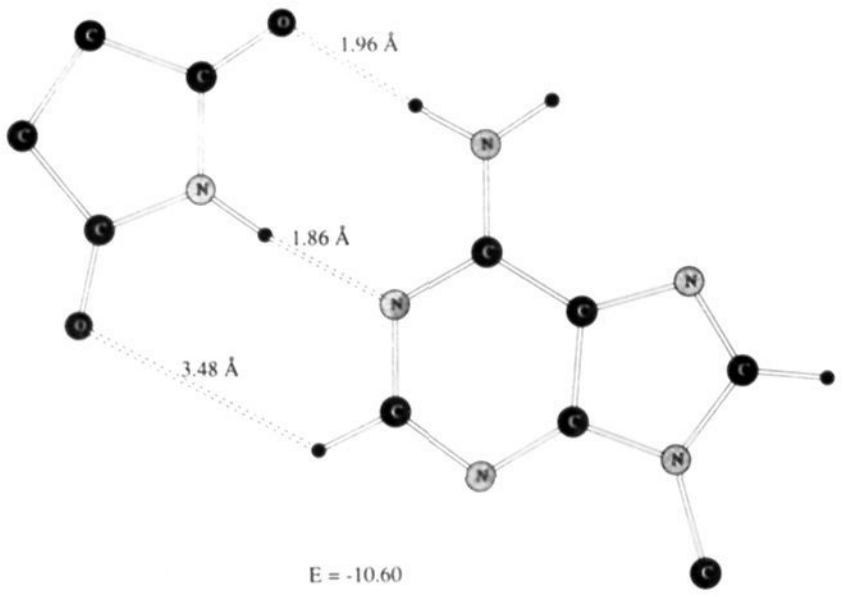
Further analysis then shows that the interactions with the spectator oxygen are indeed largely responsible for the destabilization with imides. This is apparent from decomposing the energy into the contribution from each atom. Namely, eq 3 can be rewritten as eq 4 where the i and j indices still refer to the atoms

$$\Delta E_{ab} = \sum_i e_i = \sum_j e_j$$

$$e_i = \sum_j (q_i q_j e^2/r_{ij} + A_{ij}/r_{ij}^{12} - C_{ij}/r_{ij}^{6}) \qquad (4)$$

in molecules a and b. Thus, e_i is the total contribution to the intermolecular energy owing to interactions with atom i. For the imide dimer, e_i is −1.98 and +1.84 kcal/mol for the spectator carbonyl carbon and oxygen. However, for the lactam dimer, e_i is −2.13 for the corresponding methylene group adjacent to nitrogen, and the destabilizing oxygen is absent. Furthermore, interaction with the non-hydrogen-bonding oxygen also helps cause the e_i for the hydrogen-bonding oxygen to be only −6.93 kcal/mol in the imide dimer versus −7.79 kcal/mol for the lactam dimer. Though there are numerous individual atom–atom contributions to these values, the most destabilizing interaction for the spectator oxygen is with the hydrogen-bonding oxygen in the other molecule (+13.6 kcal/mol), followed by its interaction with the other molecule's nitrogen (+11.5 kcal/mol). And, of course, the hydrogen-bonding oxygen in the reference imide also suffers an equally destabilizing interaction with the other molecule's spectator oxygen. Thus, it is clear that the present potential functions attribute the destabilization of the hydrogen bonding in these complexes by imides chiefly to the interactions with the non-hydrogen-bonding oxygen and not to variations in the interactions between the primary sites. It should be noted that the e_i values presented here are highly dominated by the Coulombic contributions.

Complexes with 9-Methyladenine. The results of the geometry optimizations for the isolated complexes with 9-Me-A in the Watson–Crick and Hoogsteen orientations are illustrated in Figures 5 and 6, respectively. Remarkably, the OPLS potential functions predict the optimal interactions to be slightly more favorable for the imide than the lactam in the Watson–Crick form and significantly more favorable in the Hoogsteen form. The only one of these optimized complexes that is not coplanar is for bu-



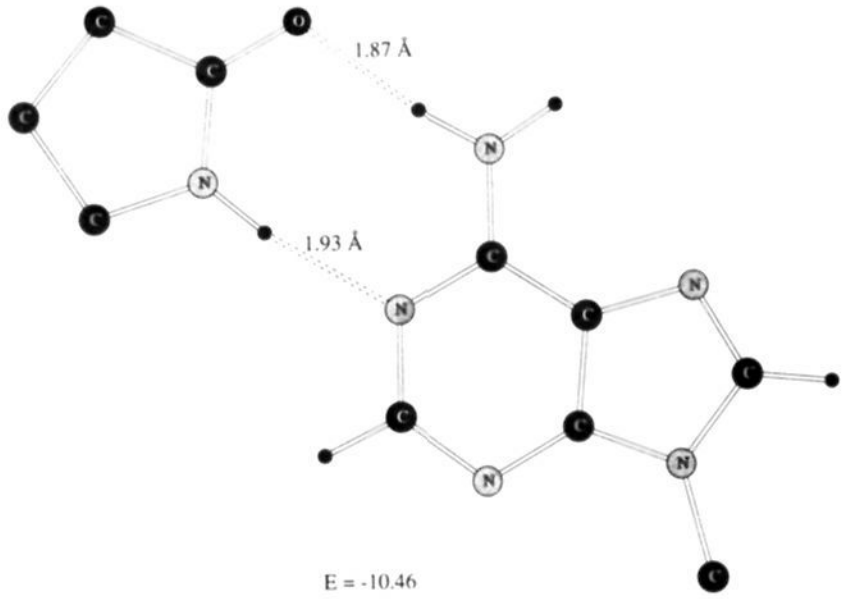


Figure 5. Geometries and interaction energies (kcal/mol) for the lowest energy complexes of succinimide and butyrolactam with 9-Me-A in the Watson–Crick orientation.

tyrolactam in the Hoogsteen orientation. However, when this complex is constrained to be coplanar and reoptimized, the interaction energy rises by only 0.14 kcal/mol to −9.13 kcal/mol. These preliminary energetic results are again consistent with the experimentally observed trend of stronger binding of imides than lactams to adenine derivatives.[5,6,8] The interaction energies of ca. −10.5 kcal/mol in the Watson–Crick orientation are also the same as for complexes of 9-Me-A with 1-methylthymine from the OPLS potential functions.[12]

The Monte Carlo simulations were subsequently carried out to assess the actual preferences in chloroform. The SA → BA mutations corresponding to ΔG_9 in eq 2 were well-behaved and yielded the curves for free energy versus the coupling parameter λ in Figure 7. Two sets of simulations were run for the complexes in the Watson–Crick and Hoogsteen orientations. The systems remained near these doubly hydrogen-bonded structures throughout the simulations, as reflected in the stereoviews of the last configurations for the Hoogsteen forms in Figure 8.

As indicated in Figure 7 and Table II, the mutation of uncomplexed succinimide to butyrolactam is now less endoergic than the same mutation in the complexes. Thus, the binding of the imide is preferred over the lactam in both orientations. For the Watson–Crick mode, $\Delta G_7 - \Delta G_8 = \Delta G_4 - \Delta G_9 = -0.32 \pm 0.10$ kcal/mol, and the difference is -1.05 ± 0.11 kcal/mol for the Hoogsteen mode. These values transform to corrected K_a ratios of 1.7 and 5.9 favoring complexation of succinimide over butyrolactam with 9-methyladenine. A predicted observable K_a ratio is not directly accessible from these results since the relative contributions of the Watson–Crick and Hoogsteen forms is not known. However, the imide has a statistical factor of 2 advantage over the lactam. Also, on the basis of the results of the gas-phase optimizations in Figure 5 and 6 and the consistent, higher corrected K_a ratio in the Hoogsteen mode, it is likely that the two orientations contribute about equally to the complexation of 9-Me-A with the

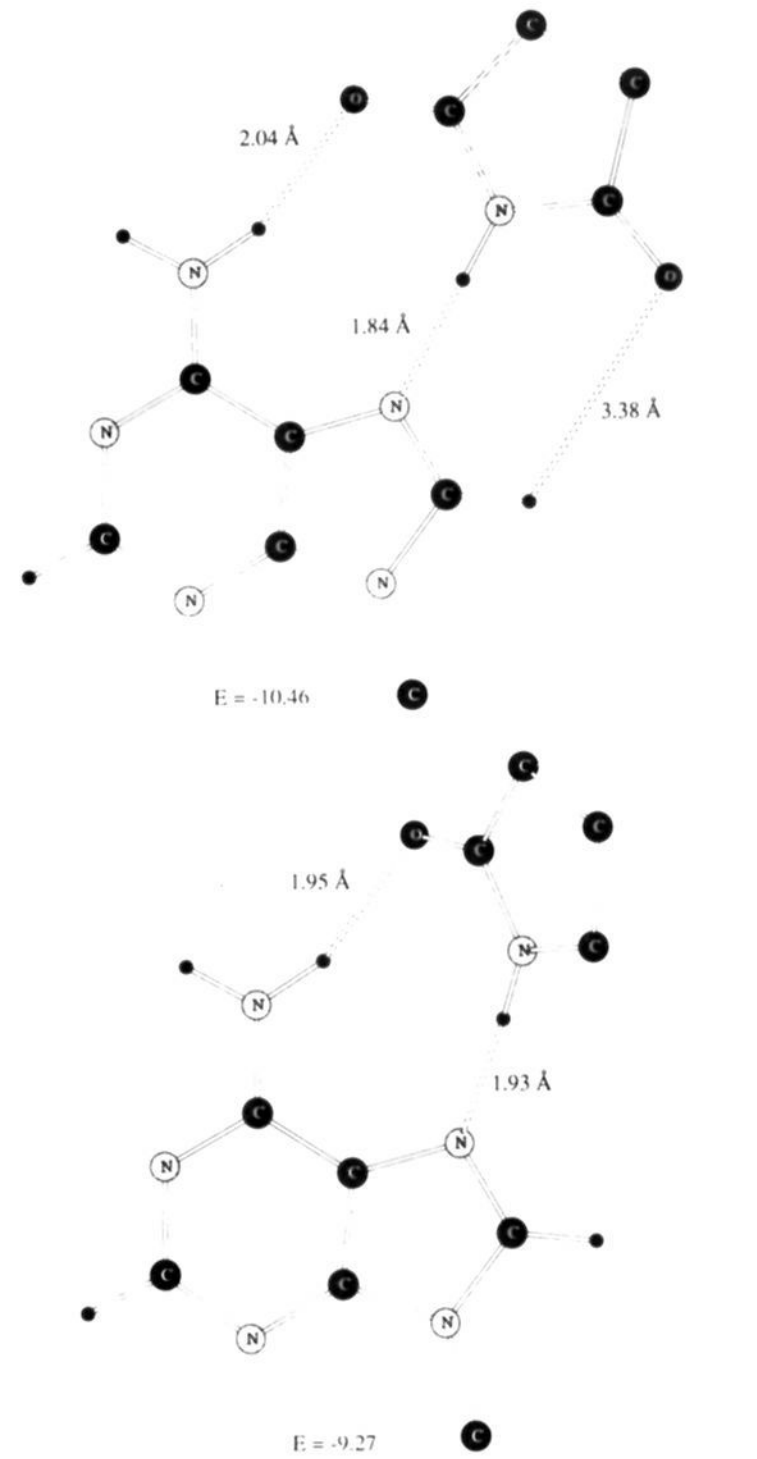


Figure 6. Geometries and interaction energies (kcal/mol) for the lowest energy complexes of succinimide and butyrolactam with 9-Me-A in the Hoogsteen orientation.

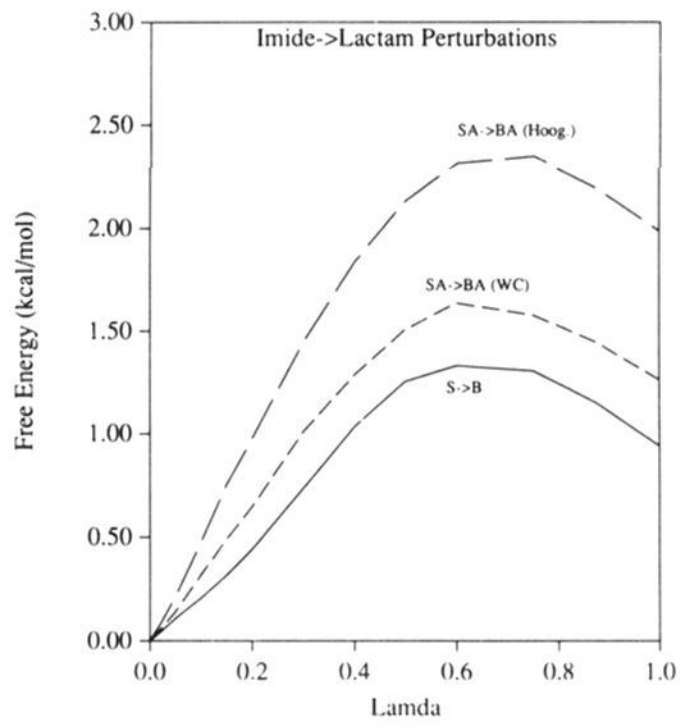


Figure 7. Computed free energy changes vs λ for the imide to lactam perturbations S → B SA → BA in the Watson–Crick orientation, and SA → BA in the Hoogsteen orientation.

imide, while the contribution from the Hoogsteen form with the lactam is small. Combined with the intrinsic statistical factor, this would give a net factor of about 4 favoring complexation of

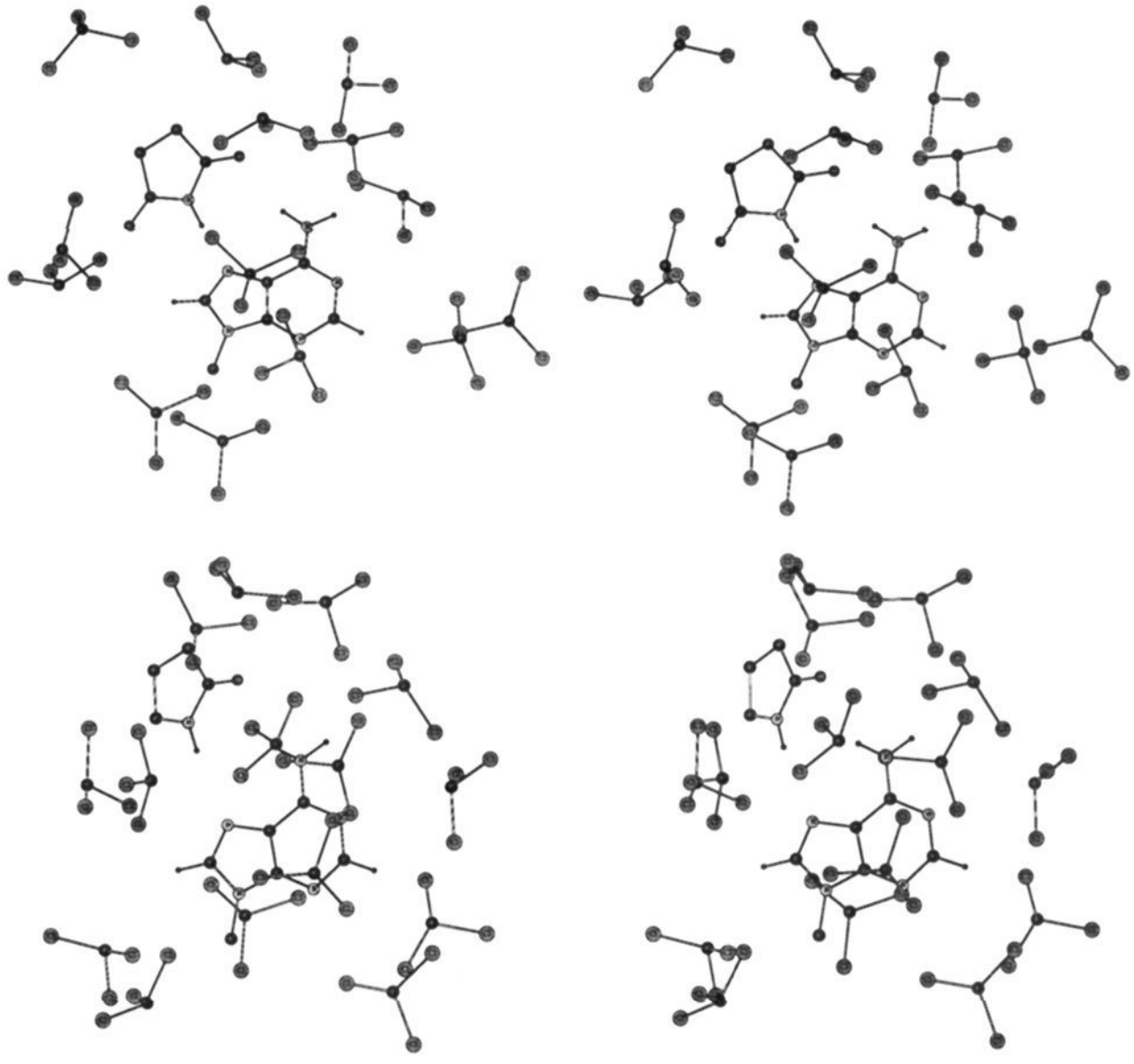


Figure 8. Stereoviews of the last configurations from the simulations of the complexes of succinimide (top) and butyrolactam (bottom) with 9-Me-A in the Hoogsteen orientation. Solvent molecules were selected as in Figure 4.

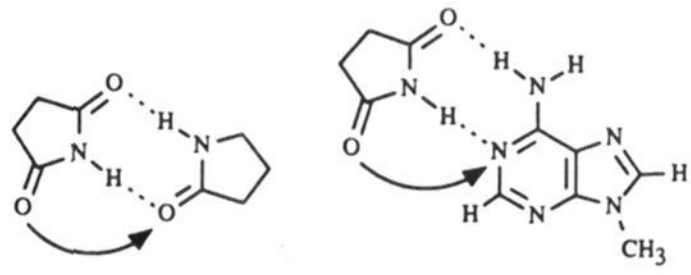
the imide and a rough predicted value of $4 \times 1.7 = 7$ for the observable K_a ratio.

The corrected K_a ratio of 1.7 is smaller than Rebek's value of $110/15 = 7$ for the binding of 9-Et-A by the naphthyl imide and lactam hosts.[6,8] However, the qualitative trend is correct and the systems under comparison are significantly different. If the Hoogsteen mode was also ineffective for Rebek's lactam, the corrected ratio of 7 would become 3.5. In addition, as Rebek and co-workers have pointed out,[6a] the binding with their naphthyl imide appears from NOE measurements to be enhanced by bifurcated hydrogen bonding of the adenine N–H to both the imide carbonyl oxygen and the oxygen of the amide in the naphthyl linkage. Presumably in the lactam, the latter oxygen would be rotated away from the lactam oxygen, and the bifurcated interaction would not occur. There are interesting subtleties in these systems, though it is clear that there is concurrence on the stronger binding of imides than comparable lactams to adenine derivatives.

The final issue is trying to understand the energetic reversals that are highlighted in Figures 2, 5, and 6. Hydrogen bonding with the lactam is preferred by 1.8 kcal/mol over the imide in Figure 3, while hydrogen bonding of 9-Me-A is stronger by 0.1–1.2 kcal/mol for the imide than the lactam in Figures 5 and 6. For the complexes with 9-Me-A, the interactions between the HNCO unit of the imide or lactam and the $N1C6NH_2$ unit in the Watson–Crick form or the $N7C5C6NH_2$ unit in the Hoogsteen form can be considered. The results for the optimized structures are −9.39 and −8.34 kcal/mol for the imide and lactam in the Watson–Crick form and −10.22 and −9.74 kcal/mol for them in the Hoogsteen form. As presented above, these interactions between the primary hydrogen-bonding sites are also favored for the imide dimer over the lactam dimer in Figure 2 by 1.4 kcal/mol. Consequently, variations in the interactions among the primary sites are not the source of the reversals.

However, consideration of the e_i (eq 4) values for the mutated groups is revealing. In the Watson–Crick orientations, e_i for the spectator carbonyl carbon and oxygen in the imide is −2.67 and +1.18 which total to −1.49 kcal/mol, while the corresponding methylene group in the lactam has $e_i = -1.45$ kcal/mol. These values can be compared with the results for the imide and lactam interacting with the lactam in Figure 2 instead of 9-Me-A. For the imide, $e_i = -4.32$ and +4.14 kcal/mol for the spectator carbon and oxygen which total to −0.18 kcal/mol, while $e_i = -2.13$ kcal/mol for the methylene group on nitrogen in the lactam. Thus, the differences in the e_i values for these atoms find the imide-to-lactam conversion to be more favorable by about 2 kcal/mol when the binding partner is butyrolactam rather than 9-Me-A. It was shown above that the destabilization for the SS and SB complexes could be attributed primarily to the Coulombic interaction of the spectator oxygen and the hydrogen-bonded oxygen in the other molecule. An analogous interaction exists between the spectator oxygen and N1 of 9-Me-A as illustrated in Chart II. However, the destabilization in this case is offset by the stabilization from the interaction of the oxygen with the highly charged $(+0.42e)$[12] C2H unit in 9-Me-A. For comparison, the CH_2 group adjacent to the CO in butyrolactam has no net charge in the OPLS model; even if it were somewhat positively charged, the geometrical differences between the two types of complexes force the methylene group away from the spectator oxygen, while the HC2–N1 dipole is well-aligned in an antiparallel fashion with the spectator C=O dipole in the complex with 9-Me-A.

Chart II



Similarly, for the Hoogsteen complexes, the HC8 unit is highly charged $(+0.40e)$ and functions like the C2H unit in the Watson–Crick orientation. The spectator C=O dipole of the imide

is now aligned in an antiparallel fashion with the HC8N7 dipole which overcomes the added Coulombic repulsion between the spectator oxygen and N7 (Figure 6). The e_i values in this case are actually negative for both the spectator carbon and oxygen (-1.39 and -0.84 kcal/mol) as well as for the methylene group α to the nitrogen in the Hoogsteen complex for butyrolactam (-1.23 kcal/mol). Thus, the net interactions with the spectator carbonyl group in the imide are now attractive by 2.2 kcal/mol, whereas they contributed only 0.2 kcal/mol for the succinimide-butyrolactam complex (Figure 2).

In summary, the interesting reversal in the preferences for complexation with imides and lactams can be attributed to predominantly Coulombic secondary interactions. For complexation with another imide or lactam, the repulsion between the spectator oxygen of the imide and the hydrogen-bonding oxygen of the other molecule dominates to make the attraction weaker for the imide than lactam. However, for complexation with adenine derivatives, favorable electrostatic interactions for the HC2N1 and HC8N7 units in the Watson-Crick and Hoogsteen orientations with the spectator carbonyl group enhance the complexation for an imide over a lactam. These analyses and the appropriateness of the intermolecular potential functions upon which they are based are supported by the accord between the present results for relative free energies of binding in chloroform and related experimental data. It is gratifying that the observed differences in binding, which are not large, can be probed with adequate precision by current methods of condensed-phase theory.

Acknowledgment. Gratitude is expressed to Professor Julius Rebek, Jr., for helpful discussions and data prior to publication, to Dr. Julianto Pranata for computational assistance, and to the National Science Foundation for support of this research.

Synthesis and Enzymatic Evaluation of Two Conformationally Restricted Trisaccharide Analogues as Substrates for *N*-Acetylglucosaminyltransferase V

Ingvar Lindh and Ole Hindsgaul*

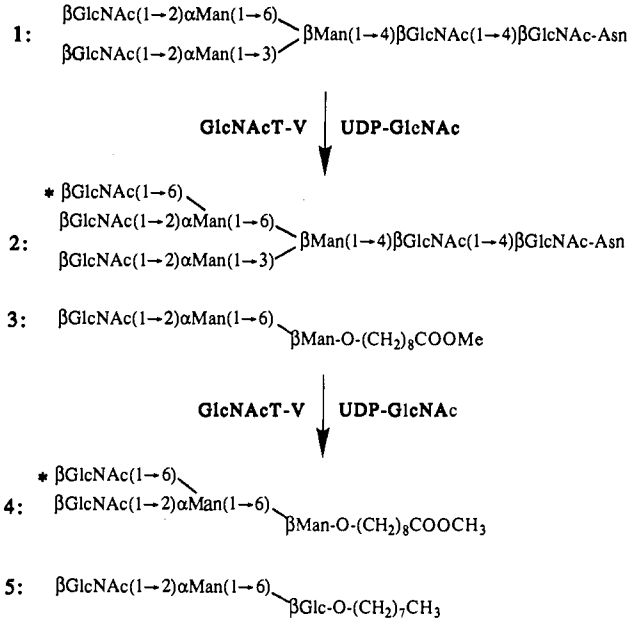
Contribution from the Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada. *Received June 5, 1990*

Abstract: The synthetic trisaccharide βGlcNAc(1→2)αMan(1→6)βGlc-O$(CH_2)_7CH_3$ (**5**) is an acceptor substrate for the enzyme *N*-acetylglucosaminyltransferase V (GlcNAcT-V) which adds an additional βGlcNAc residue to the 6-position of the central αMan unit. Two trisaccharide analogues of **5** have been chemically synthesized where the possibility for rotation about the C5-C6 bond of the βGlc residue has been eliminated by linking O4 and C6 with an ethylene bridge. The resulting fused-ring trisaccharides **8** and **9** represent conformationally restricted models for the "gt" and "gg" conformations, respectively, of asparagine-linked oligosaccharides. Both **5** and its conformationally restricted gg analogue **9** showed similar acceptor properties toward the enzyme from hamster kidney, while the gt trisaccharide **8** was an extremely poor substrate. These results demonstrate that GlcNAcT-V preferentially recognizes the carbohydrate chains of asparagine-linked glycoproteins in their gg conformations.

UDP-GlcNAc:α-mannoside $\beta(1\rightarrow6)$-*N*-acetylglucosaminyltransferase (GlcNAcT-V,EC 2.4.1.155) is a key enzyme involved in the biosynthesis of highly branched asparagine-linked oligosaccharides.[1,2] This enzyme has become the subject of intense investigation following observations that GlcNAcT-V activity increases when cells are transformed by both tumor viruses[3,4] and oncogenes.[5] Specific increases in the activity of this enzyme have also recently been shown to correlate with the metastatic potential of human and rodent tumor cells.[6,7]

Biosynthetically, GlcNAcT-V transfers an *N*-acetyl-D-glucosamine (GlcNAc) residue from uridine 5′-diphospho-GlcNAc (UDP-GlcNAc) to oligosaccharide acceptors having the minimum heptasaccharide sequence **1** (Scheme I), converting it to the octasaccharide **2**.[1,2] We have previously shown that the much simpler synthetic trisaccharide **3**, a partial structure of **1**, is also an effective substrate for the enzyme yielding the expected tetrasaccharide **4**.[8] The enzyme further tolerates the substitution of the βMan residue in **3** by a βGlc residue since trisaccharide **5** was also found to be an excellent acceptor.[9] The aliphatic aglycons in **3** and **5**

Scheme I. Glycosylation Reactions Catalyzed by GlcNAcT-V



were incorporated into these structures in order to facilitate the enzyme assay procedures.[10]

(1) Cummings, R. D.; Trowbridge, I. S. *J. Biol. Chem.* **1982**, *257*, 13421-13427.
(2) Schachter, H. *Biochem. Cell Biol.* **1986**, *64*, 163-181.
(3) Yamshita, K.; Tachibana, Y.; Ohkura, T.; Kobata, A. *J. Biol. Chem.* **1985**, *260*, 3963-3969.
(4) Arango, J.; Pierce, M. *J. Cell. Biochem.* **1988**, *37*, 225-231.
(5) Dennis, J. W.; Kosh, K.; Bryce, D.-M.; Breitman, M. L. *Oncogene* **1989**, *4*, 853-860.
(6) Dennis, J. W.; Laferte, S.; Waghorne, C.; Breitman, M. L.; Kerbel, R. S. *Science* **1987**, *236*, 582-585.
(7) Dennis, J. W. *Cancer Surveys* **1988**, *7*, 573-594.
(8) Hindsgaul, O.; Tahir, S. H.; Srivastava, O. P.; Pierce, M. *Carbohydr. Res.* **1988**, *173*, 263-272.
(9) Srivastava, O. P.; Hindsgaul, O.; Shoreibah, M.; Pierce, M. *Carbohydr. Res.* **1988**, *179*, 137-161.

0002-7863/91/1513-216$02.50/0