Can Configuration Entropy Losses Be Predicted from the Binding Affinities of Hydrogen-Bonded Complexes with Varying Numbers of Single Bonds?

Kristy L. Mardis*

Pacific Lutheran University, Department of Chemistry, Tacoma, Washington 98447-0003

Received: September 1, 2005; In Final Form: October 20, 2005

The design of new supramolecular complexes often depends on reducing entropic contributions to improve binding. However, few studies provide reliable values for the cost of entropic contributions. Here, the binding affinities of a series of six α,ω-diamides to α,ω-dicarboxylates are calculated using a predominant states method and an implicit solvent model based upon finite difference solutions of the Poisson−Boltzmann equation. The calculations are able to reproduce the observed increase in binding free energy as the number of single bonds increases. However, calculations show that the increase in binding free energy is not due to an increase in entropy. Instead, the increase is due to the changing ability of the α,ω-diamides to form internal hydrogen bonds that must be disrupted to bind to the dicarboxylate receptors. This suggests that interpreting experimental free-energy trends to give rotational entropy contributions may be problematic.

I. Introduction

The construction of supramolecular complexes requires a detailed understanding of the energetics of molecular recognition. In the design phase, structures that reduce the conformational entropy of the precursor units are often preferred because the restriction of ligand torsional motions upon binding is entropically unfavorable.[1] However, modifications to reduce the entropy may actually reduce the binding affinities because hydrogen bonds are disrupted.[2] The charge interactions are no longer optimally situated and must pay a steric strain cost to align properly. The magnitude of the configurational entropy loss is thus an important parameter in design decisions. Experimentally, entropic changes associated with binding are difficult to measure because of the residual motion of the ligand in the complex and uncertainty concerning the entropy of the ligand in solution. Estimates from thermodynamic data at 300 K for the freezing of an internal torsion range from 0.3 to 1.6 kcal/mol.[3−8] This range results in uncertainty about the necessity of avoiding single bonds in molecular design.

By fitting the experimentally determined free energy to a linear function of the number of rotatable bonds, Eblinger and Schneider[3] suggest that the entropic disadvantage is relatively small for the complexes shown in Figure 1. Although such a fit is suggestive, it does not specifically determine if the change in free energy is due to changes in configurational entropy.

A more direct measure of the entropy is needed to assess the magnitude of configurational entropy loss upon binding for carbon−carbon single bonds. This work presents free-energy calculations that are able to calculate the role of configurational entropy directly in the binding of the same six α,ω-diamide/α,ω-dicarboxylate complexes. This series illustrates the complications that can arise in obtaining entropic loss values from Δ*G* data for even carefully designed systems.

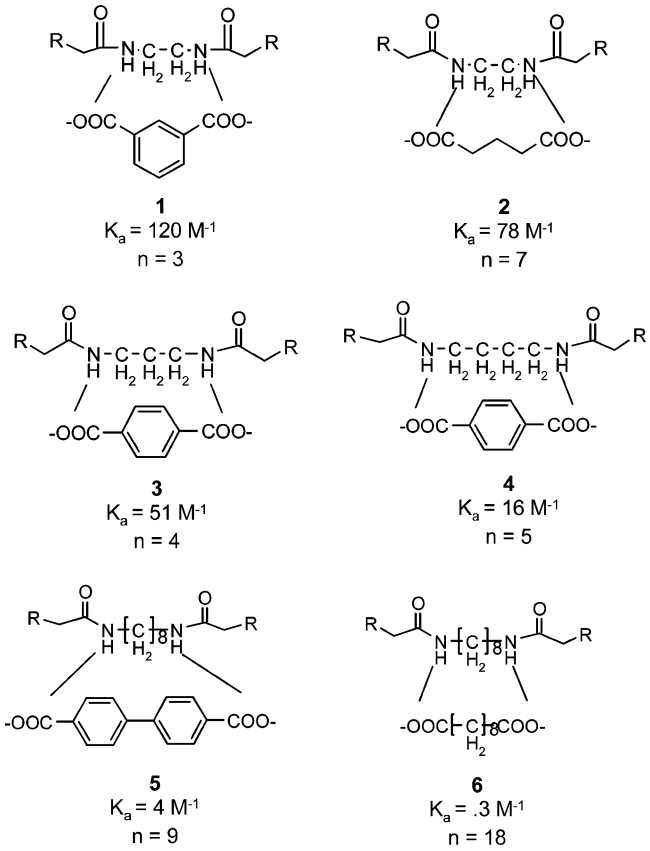

Figure 1. Structures of complexes studied in this work, experimental binding affinities (K_a) from ref 1, and the number of rotatable bonds (*n*) as defined in ref 3. R = adamantyl group.

II. Methods

A. Theory. The "mining minima" (MM) algorithm is used to calculate the free energies of binding. Because details of the algorithm have been described in previous work,[9,10] only a brief overview will be presented here.

* Corresponding author. E-mail: kmardis@csu.edu. Current Address: Department of Chemistry and Physics, Chicago State University, Chicago, IL 60628.

10.1021/jp054964u CCC: $33.50
Published on Web 12/18/2005

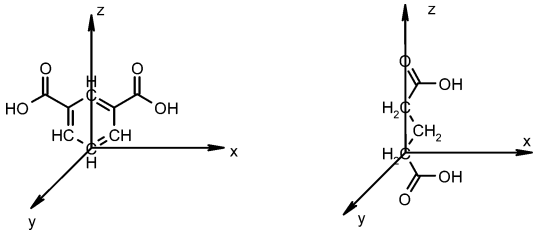

Figure 2. Coordinate system.

Ignoring symmetry numbers, which will cancel in the classical approximation to statistical thermodynamics, the standard chemical potential of molecular species X in a hypothetical ideal dilute solution equals[9,11]

$$\mu_X^\circ = -RT \ln Z_X - RT \ln\left(\frac{8\pi^2}{C^\circ}\right) \quad (1)$$

where C° is the standard concentration and Z_X is the configuration integral over the internal coordinates of the molecule. The configuration integral Z_X in turn can be written in the form[12]

$$Z_X = \int e^{-(U_X+W_X)/RT}\, d\vec{r}_{\text{int}} \quad (2)$$

The Boltzmann factor is given in terms of the vacuum potential energy, U_X, and the solvation energy, W_X, both of which depend on molecular conformation. The integral extends over the full range of the internal coordinates, $\vec{r}_{\text{int}}$, defined previously.[12]

For the noncovalent association of two molecules, $L + R \rightarrow L{:}R$, the standard free energy of binding can then be written in terms of the configuration integrals, as follows:

$$\Delta G_B^\circ \equiv \mu_{RL}^\circ - \mu_R^\circ - \mu_L^\circ = -RT \ln \frac{Z_{L:R}}{Z_L Z_R} + RT \ln \frac{8\pi^2}{C^\circ} \quad (3)$$

Here $Z_{L:R}$, Z_L, and Z_R are the configuration integrals of the complex and the isolated molecules, respectively.

The MM algorithm calculates Z_X by identifying a low-energy conformation, i, and evaluating the corresponding configuration integral, $Z_{X,i}$, for that energy well via a Monte Carlo method. The total configurational integral is then the sum over energy minima:

$$Z_X \approx \sum_i Z_{X,i} \quad (4)$$

The search for new energy wells continues until, for five successive minima, either (a) contributions to the free-energy drop to a fractional change of $<10^{-6}$ or (b) the cumulative average potential energy of the wells changes by less than 0.001 kcal/mol.[13]

B. Sampling. The conformational distributions of the "hard" bond and angle degrees of freedom are assumed to not vary greatly upon binding. This allows us to treat them as rigid and integrate eq 2 only over dihedral angles and, for the complexes, the relative rotational and translational degrees of freedom.[12]

The coordinates specifying the position of the diamide relative to the dicarboxylate are shown in Figure 2. For the benzyl carboxylates, the origin is located on one of the benzyl carbons. The z axis is defined by the line between this origin and the benzyl carbon directly across the ring. The xz plane is defined by the z axis and the benzyl carbon attached to the carboxylate substituent group. For the alkyl carboxylates, the origin is located on the alkyl carbon bonded to the carboxylate carbon. The z axis is defined by the line between the origin and the alkyl carbon bonded to the second carboxylate carbon. The xz plane is defined by the z axis and the alkyl carbon bonded to the carbon at the origin. The position of the diamide in this axis system is defined by the position of the centermost carbon in the alkyl chain.

The number of rotatable dihedral angles varied from 2 (isophthalate and terephthalate) to 15 (2-adamantan-1-yl-*N*[3-(2-adamantan-1-yl-acetylamino)-octyl]-acetamide). Starting structures were obtained by performing an initial conformational search for each individual molecule followed by a gas-phase minimization of the lowest energy structure. Because bond lengths and angles were not varied in the sampling, and because some of the diamides formed an intramolecular hydrogen bond, there was the possibility of bias. The bond lengths and angles of the groups involved in this might influence the strength of the intermolecular hydrogen bonds in the complex. In cases in which the lowest energy structure contained a hydrogen bond, a second structure, without an intramolecular hydrogen bond was chosen as a second starting structure. All binding calculations were run with both and the results of both runs combined by Boltzmann weighting. Finally, in light of our initial results, starting structures were constructed by performing a gas-phase minimization of the two binding partners as a complex. These bond lengths and angles were used as starting coordinates in binding calculations, which will be referred to below as the strain-minimized calculations.

C. Energy Model. The energy in the configurational integrals is computed as the sum of the potential, $U(r)$, and the solvation, $W(r)$, energy.[14] Here, the potential energy is computed with the CHARMm 22 force field[15] using the all-atom representation. Atom-type parameters were assigned with the program Vega.[16] Atomic charges, shown in Table 1, were assigned in accordance with CHARMm charges from the Quanta 98 implementation of CHARMm (set I) or in accordance with Charmm22 templates (set II). All of the terms in the CHARMm force field were included. The carboxylic acid groups are treated as fully ionized to match the experimental work in which the α,ω-dicarboxylates were prepared as tetrabutylammonium salts and then dissolved in chloroform.

During conformational sampling, the electrostatic portion of the solvation energy is estimated with the generalized Born model (GB).[17,18] The solvent dielectric constant is set to that of chloroform (4.8 at 25 °C). The cavity radius of each atom is set to the mean of its CHARMm σ parameter and the solvent probe radius (2.5 Å). The nonpolar component of the solvation energy, added after the sampling, is estimated as proportional to the solvent-accessible surface area using

$$\Delta G_{\text{np}} = aA_{\text{SA}} + b \quad (5)$$

where A_{SA} is the solvent accessible surface area. A value of -16.4 cal/mol/Å^2 was used for the proportionality constant, a. However, there is considerable uncertainty in both the value and the meaning of the offset.[13] Therefore, the Results section presents relative binding free energies that do not depend on the value of the offset.

After the sampling phase, the electrostatic solvation energies are adjusted to optimize the agreement with finite difference solutions of the Poisson−Boltzmann (PB) model. This is accomplished by recalculating the electrostatic solvation energy with both GB and PB for each distinct local minimum found during the conformational sampling. The deviation of GB relative to PB is then subtracted from the original conformation free energy of the minimum. Previous calculations[19] have

TABLE 1: Atomic Charge Parameters in Units of Electrons (1 Electron = 1.602×10^{-19} C)

diamide				dicarboxylate			
		charge set				charge set	
chemical group	atom	I	II	chemical group	atom	I	II
$-CH_2NH$	−C	−0.05	−0.02	aromatic $-C-CO_2^-$	−C−	−0.15	−0.10
	alkyl H	0.05	0.09		C	0.35	0.62
	N	−0.40	−0.47		O	−0.60	−0.76
	amide H	0.25	0.31	aromatic −CH	C	0.13	−0.115
>C═O	C	0.60	0.51		H	−0.13	0.115
	O	−0.55	−0.51	alkyl $-CH_2-CO_2^-$	−C−	−0.15	−0.28
chain $-CH_2^-$	C	−0.10	−0.18		H	0.00	0.09
	H	0.05	0.09		C	0.35	0.62
admantyl $-CH_2^-$	C	−0.10	−0.18		O	−0.60	−0.76
	H	0.05	0.09	$-CH_2^-$	C	−0.10	−0.18
admantyl >CH−	C	−0.05	−0.09		H	0.05	0.09
	H	0.05	0.09	aromatic >C−C<	C	0.00	0.00
admanyl >C<	C	0.00	0.00				

suggested that this correction improves the agreement of computation with experiment when partially desolvated systems are studied.

III. Results and Discussion

A. Free Energies. Figure 3 plots the calculated (using charge set I) and experimental binding free energies for each complex. As mentioned in the computational details section, each data series uses the binding of complex **1** as a baseline. The values for all other complexes are actually their relative difference in binding compared to complex **1**. This avoids difficulties with the nonpolar solvation offset value. The calculations reproduce the experimental trend of generally increasing free energy as the number of single bonds increases. Furthermore, the calculations reproduce the deviation for complex **2** seen experimentally, which has 7 single bonds, but binds better than either complex **3** or **4** with 4 and 5 single bonds, respectively.

B. Role of Entropy in Binding. A central goal of this work is to determine the role entropy plays on the increase of free energy as the number of single bonds increases. This section therefore examines the variation of entropic and enthalpic energy components as a function of the number of single bonds. As shown Figure 4, neither the trend in potential nor entropic energy tracks the trend in free energy. Entropic energy initially decreases (becomes more favorable toward binding) as the number of single bonds increases from three to four (complexes **1** and **3**). There is another drop in entropy between complex **2** (7 rotatable bonds) and complex **5** (9 rotatable bonds). In fact, the potential and entropic energy seem to compensate for each other, one rising as the other falls and vice versa. However, if one plots ΔH versus $T\Delta\Delta S$, as shown in Figure 5, then the standard enthalpy−entropy compensation that results in a relationship between ΔH and $T\Delta\Delta S$ with a slope greater than 1 does not hold completely, especially for complex **6**.[20] It is possible that enthalpy−entropy compensation requires the inclusion of explicit solvation effects that are not completely modeled here or that the range of ΔH and $T\Delta\Delta S$ values in this work is too small to identify the trend.

It is interesting that if the deviations are ignored and a linear fit is imposed on the entropic trend, then the slope, 0.29 kcal/mol/rotatable bond ($R = 0.71$), is remarkably similar to the value obtained by fitting the experimental free-energy curve (0.28 kcal/mol/rotatable bond, $R = 0.91$). Both values are similar to the value of 0.31 kcal/mol/rotatable bond obtained by Eblinger and Schneider.[3] An investigation of the bound molecular structures

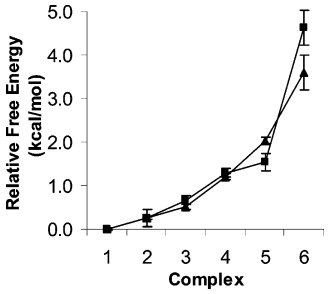

Figure 3. Experimental[3] (▲) and calculated (■) relative binding free energies for each complex. All values are shown relative to the binding free energy of complex **1**. Error bars for the experimental data are taken from ref 3. Calculated error bars are obtained by repeating the calculations (1) with different random numbers and (2) doubling the number of minima found.

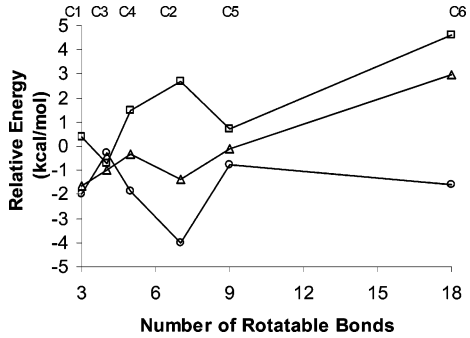

Figure 4. Change in free energy (△), potential energy (○), and $-T\Delta\Delta S$ (□) upon binding as a function of rotatable bonds. Following the convention of ref 3, the aromatic carbon−carboxyl carbon bonds are not considered rotatable bonds. The negative of the entropic energy is plotted to maintain the standard of a more negative value indicating more favorable binding. For reference, the points are labeled by complex number along the top of the graph. The ΔG values should be interpreted as relative values because no attempt was made to correct for the nonpolar solvation offset (see the Experimental Section).

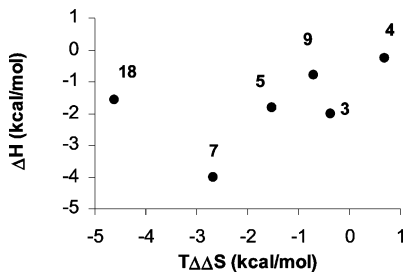

Figure 5. Plot of the calculated enthalpy of binding against the calculated entropy of binding for the six complexes. Each point is labeled by the number of rotatable bonds.

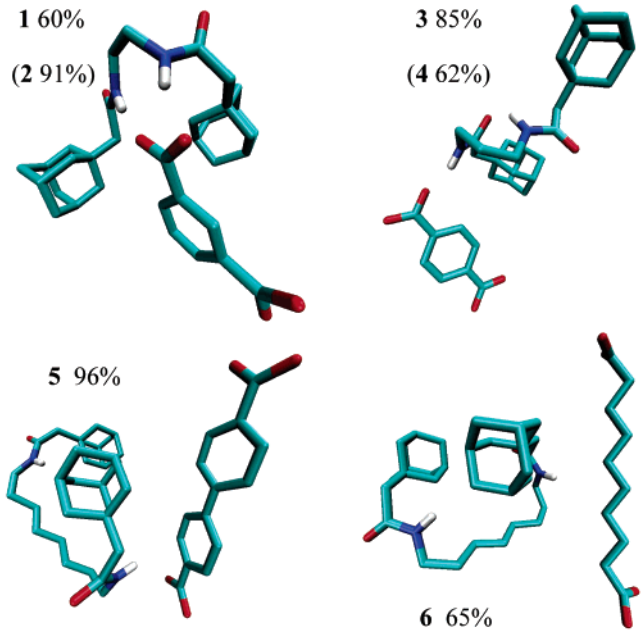

Figure 6. Most probable bound structures for complexes **1**, **3**, **5**, and **6**. The bound structures of complex **2** resemble those of complex **1**, and those of complex **4** resemble those of complex **3**. The total probability for each structure is shown as a percentage.

TABLE 2: Percentage of Conformers Having Various Hydrogen Bond Types[a]

	hydrogen bond type					diamide
complex	2[b] inter-	1 inter & 1 intra	1 intra-	1 inter-	none	intra- before binding
1	60	<1	32	<1	5	62
2	91	5	<1	3	<1	62
3	2	85	9	1		76
4	<1	62	11	21	4	39
5	<1	<1	<1	2	96	0
6	9	<1		23	65	0

[a] The last column provides information on how likely the diamide portion of the complex is to have an intramolecular hydrogen bond before binding. [b] The two intermolecular hydrogen bonds in complex **6** are to two different carboxylate acceptors. All others involve two donors bound to one carboxylate acceptor.

illustrates why using any of these values as a measure of the entropic cost of adding a rotatable bond is problematic.

C. Structure-Based Analysis. Ideally, the molecules would form complexes such as those shown in Figure 1. In this "bridged" conformation, the two amide hydrogens form hydrogen bonds with separate carboxyl groups. In this binding scheme it is easy to see why increasing the number of $-CH_2-$ groups would increase the entropic cost of binding. However, this is not the preferred binding conformation for any of the complexes.

As shown in Figure 6, for all complexes, the most probable binding conformation involves only one of the carboxylate groups. The breakdown of binding conformations for each complex is shown in Table 2. Instead of forming a hydrogen bond with the second carboxylate group, the second amide hydrogen binds to the same carboxylate group (**1** and **2**), forms an internal hydrogen bond with the amide oxygen (**3** and **4**), or is exposed to solvent (**5** and **6**). In fact, only complex **6** has a significant population (8%) bound in the bridged conformation.

The data in Figures 3−5 was obtained using charge set I. The lack of bridged conformations was surprising, so several additional calculations were performed. First, the calculations for complexes **1** and **4** (chosen as representative systems) were repeated using charge set I, but without the solvation component (a gas-phase calculation). In both cases, less than 0.1% of the resulting conformers were bridged. This suggests that error in the solvation model cannot account for the lack of bridged structures. Next, the calculations were repeated using charge set II for complex **1**. Again, no bridged conformations were found. In fact, the structures resulting from charge set II were much more likely to form intermolecular hydrogen bonds or to fold into structures similar to those of complex **6** in Figure 6 that maximized the van der Waals attractions. Although it is possible that the CHARMm force field is not capable of modeling these structures, this sort of error should have made it difficult to reproduce the free energy of binding trend.

Finally, to determine if the bridged structures could only be obtained with a very specific set of bond distances and angles not present in the original structures, complexes of the form of Figure 1 were constructed and minimized as a unit using CHARMm. These *strain-minimized* structures should have bond angles and distances pre-set to form the bridged structures. The mining minima calculations were repeated for complex **1**. Even when starting from these preoptimized for binding structures, the predominant configurations resembled those in Figure 6, not the bridged structures of Figure 1.

The primary difficulty in correlating the free-energy trend with the entropic loss trend is the tendency of the diamides to form intramolecular hydrogen bonds. Calculations suggest that prior to binding, the diamide involved in complexes **1** and **2** has an intramolecular hydrogen bond 62% of the time, while the unbound diamide in complex **3** is even more likely (76%) to exist in an intramolecular hydrogen bonded state. The *n*-butyl diamide (complex **4**) is less likely (37%) to have an intramolecular hydrogen bond, but there remains a significant fraction of the structures with intramolecular hydrogen bonds that must be disrupted to form a bridged complex. Only the *n*-octyl diamide (complex **5** and **6**) does not show evidence of intramolecular hydrogen bonding when in solution alone. Significantly, it is only this diamide that forms bridged complexes with the carboxylate acceptor.

D. Explanation of Binding-Energy Trend. The bound structures suggest that the increasingly weak binding is not due to the increasing entropic cost of restricting longer and longer carbon chains. In fact, potential energy changes alone can predict the trend between complexes **4**, **2**, and **5**. For these complexes, an analysis of the individual energy components indicates that the trend in binding affinity is primarily driven by complex **2**'s ability to form better (tighter and shorter) hydrogen bonds. In fact, as shown in Figure 6, the bound structure of complex **2** forms two hydrogen bonds as compared to the one in complex **4** and **5**. Nevertheless, entropy does play an important role in predicting the binding trend. As mentioned above (Table 2), the changing probability of maintaining the intramolecular amide hydrogen bond influences the binding. Comparing complex **1** and **2**, **2** binds more tightly despite its greater number of rotatable bonds. However, as shown in Table 2, complex **2** rarely maintains an intramolecular hydrogen bond upon binding. An analysis of the energy components of the individual binding conformations indicate that this is due to the higher entropic cost of maintaining a closed eight-membered ring in the case of complex **2** as compared to the seven-membered ring for complex **1**. This alone might suggest that **2** has a lower entropic cost to binding. However, the entropic cost of restricting the rotational and translational motions to form two amide−carboxylate bonds seems to be higher than the entropic gain

from opening the ring. A similar argument can be made for explaining the relative binding strengths of complexes **1** and **3**. Figure 4 indicates that despite the increased number of rotatable bonds complex **3** suffers less entropic loss. Upon analysis of the data in Table 2, it may be that the increased likelihood of the diamide in **3** to form intramolecular hydrogen bonds pre-organizes it for binding. Complex **3** has already "paid" the entropic cost to form a hydrogen bond. For complexes **5** and **6**, the poor binding is due to the fact that both bind primarily through van der Waals interactions rather than hydrogen bonds. These results suggest that complex **6** pays a larger entropic penalty because it loses more translational and rotation entropy when the two molecules align to maximize stacking-type contacts. In summary, because the entropic loss cannot be ascribed to a single event, but rather is due to the interplay of losing or maintaining intramolecular hydrogen bonds, forming one or more intermolecular hydrogen bonds, and the restriction of translational and rotational motions to optimize van der Waals attractions, no simple explanation for the binding trend is possible.

IV. Conclusions

Both calculation and experiment find that adding single bonds results in decreased binding affinity between the diamides and dicarboxylates. However, the calculations presented here indicate that this is not due to the increased flexibility causing a steady increase in the entropic loss upon binding. The addition of single bonds does influence the binding, but primarily by changing the probability of the formation of intramolecular hydrogen bonds. Experimental work on organic ion pairs,[21] which lack the ability to form intramolecular hydrogen bonds, still suggests that entropic losses may be even smaller than values in current use. Preliminary calculations suggest that this system may provide a more robust value for entropic loss. Thus, future work will focus on complexes that lack the ability to form intramolecular hydrogen bonds.

Acknowledgment. Funding for this work from the Research Corporation (Cottrell College Science Award no. CC5582) and Pacific Lutheran University (Regency Advancement Award) is gratefully acknowledged.

References and Notes

(1) Bohm, H.-J. *J. Comput-Aided Mol. Des.* **1994**, *8*, 243−256.

(2) Pochapsky, T. C. *Enantiomer* **1999**, *4*, 437−444.

(3) Eblinger, F.; Schneider, H.-J. *Angew. Chem., Int. Ed.* **1998**, *37*, 826−829.

(4) Page, M. I.; Jencks, W. P. *Proc. Nat. Acad. Sci. U.S.A.* **1971**, *68*, 1678−1683.

(5) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690−10697.

(6) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. G. *Biochemistry* **1998**, *37*, 16839−16845.

(7) Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. *J. Am. Chem. Soc.* **1991**, *113*, 7020−7030.

(8) Ajay; Murcko, M. A. *J. Med. Chem.* **1995**, *38*, 4953−4967.

(9) Head, M. S.; Given, J. A.; Gilson, M. K. *J. Phys. Chem. A* **1997**, *101*, 1609−1618.

(10) Luo, R.; David, L.; Hung, H.; Devaney, J.; Gilson, M. K. *J. Phys. Chem. B* **1999**, *103*, 727−736.

(11) In *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems*; Hill, T. L., Ed.; Springer-Verlag: New York, 1985.

(12) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys J.* **1997**, *72*, 1047−69.

(13) Luo, R.; Head, M. S.; Given, J. A.; Gilson, M. K. *Biophys. Chem.* **1999**, *78*, 183−93.

(14) Gilson, M. K.; Given, J. A.; Head, M. S. *Chem. Biol.* **1997**, *4*, 87−92.

(15) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(16) Pedretti, A.; Villa, L.; Vistoli, G. *J. C. A. M. D.* **2004**, *18*, 167−173.

(17) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(18) Qui, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.

(19) Mardis, K. L.; Luo, R.; Gilson, M. K. *J. Mol. Biol.* **2001**, *309*, 507−17.

(20) Calderone, C. T.; Williams, D. H. *J. Am. Chem. Soc.* **2001**, *123*, 6262−6267.

(21) Hossain, M. A. S.; Schneider, H. J. *Chem.—Eur. J.* **1999**, *5*, 1284−1290.