



Improved convergence of binding affinities with free energy perturbation: Application to nonpeptide ligands with pp60^{src} SH2 domain

Daniel J. Price & William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, CT 06520-8107, U.S.A.

Accepted 17 May 2001

Key words: binding affinity, free energy perturbation, Monte Carlo, Src, SH2

Summary

Free Energy Perturbations (FEP) in the context of Monte Carlo (MC) simulations were conducted to predict the relative free energies of binding for a series of human Src SH2 domain ligands. Two procedures for disappearing atoms during a single-topology FEP are investigated and dramatic differences in free energy convergence behavior are seen. Comparison of these two protocols suggests that the coupling of the removal of angular constraints with the disappearance of an atom may significantly slow free energy convergence. The series of ligands under investigation here cover a range of modifications at the 3-position of 4-([4-(cyclohexyl methoxy)benzyl]amino)carbonyl) phenyl phosphate. Unlike any other compound in this study, the 3-amide analog can form two hydrogen bonds within the region of the perturbation, one to a backbone amide hydrogen and one to a highly coordinated water molecule. Agreement with experimental trends in binding affinity is seen, although the computed relative free energy of binding of the amido compound is underestimated. These results are reconciled by examination of the hydration energies of model systems, which predict primary amides as too hydrophilic.

Introduction

The potential of incorporating statistically rigorous calculations of relative free energies into the arsenal of computational drug-design tools is alluring. But while extremely accurate free energies have been achieved for small, organic systems, [1] determination of free energies in macromolecular systems has been hindered by poor convergence. The challenges to obtaining highly converged FEP results are twofold: (1) the still-elusive problem of poor sampling of the complicated energy landscapes inherent to biomolecular systems, and (2) disproportionate sampling of high-energy states for the perturbed system owing to dissimilarity of the reference and perturbed states. There have been many contributions in recent years to improving the sampling of low-energy perturbed states including shrinking bond lengths during the removal of an atom, the GASP procedure, dynamic windowing, and alternative (nonlinear) coupling pathways [2–6].

In light of the second point, the free energy convergence as a function of the positional constraints on the flexible dummy atoms at the beginning or end of an FEP are considered in this study. Dummy atoms are a necessary artifact of the inability to change the number of degrees of freedom in a perturbation, specifically when perturbing between functional groups of differing number of atoms. In most cases, there should be no torsional potential on the dummy atoms at the endpoint, as torsional distributions often are not well converged and different dihedral occupancies in the bound and unbound legs will hamper convergence of relative free energies of binding ($\Delta\Delta G_b$). As bonds and angles are treated harmonically here, it is necessary to have a non-zero force constant on one bond (k_r) and on at least one angle (k_θ) to each dummy atom in order to confine the dummy atom to a reasonable volume and to prevent numerical difficulties associated with linear angles, respectively [7]. There is some appeal to having only one angle with a $k_\theta > 0$ kcal/mol rad², as minimized structures are guaranteed to have the same absolute energy as the system

*To whom correspondence should be addressed. E-mail: william.jorgensen@yale.edu

without dummy atoms. On the other hand, the dummy atoms will sample a large number of geometries with only one angular constraint, most of which are high-energy states to the opposite endpoint [8]. As a prelude to a thermodynamic and structural analysis of Src SH2 domain binding with a series of peptidomimetic compounds, a single mutation, where one amide hydrogen is transformed to a carbon, and three dummy atoms grow out into hydrogens, is performed in forward and reverse under two cases of angular constraints in order to examine free energy convergence behavior.

For reasons of convergence, studies of protein-ligand binding with FEP have focused on systems that mimic a traditional medicinal chemistry scenario, that being the functionalization of a common template or scaffold [9]. Similarly, this study investigates a series of Src SH2 domain ligands of the form 4-([4-(cyclohexyl methoxy)benzyl]amino)carbonyl phenylphosphate (compounds **1–7** in Table 1), which were developed at Parke-Davis [10].

The Src protein mediates the prototypical signal transduction pathway whereby extracellular cytokines upregulate growth-related cellular processes. Any number of mitogens (EGF, PDGF, FGF, CSF-1, etc.) can stimulate homodimerization and subsequent *in-trans* autophosphorylation of their respective transmembrane receptors. Recognition of select phosphorylated tyrosines (Y^*) on the cytoplasmic side of these receptors by the SH2 domain of Src activates Src's own kinase domain, which in turn initiates a cascade of intracellular phosphorylation events. Src's numerous substrates allow for propagation of this proliferative signal to a broad range of cellular processes, including cell cycle progression, DNA synthesis, cytoskeletal rearrangement, and cell-cell adhesion [11]. Considering the cellular events modulated by Src's activity, it may not be surprising that elevated levels and activated forms of Src have been implicated in several cancers, particularly breast, colon, and pancreatic carcinomas, and have been identified in tissues (colon polyps and Barrett's esophageal epithelium) that are predisposed to dysplasia and malignant transformation [12–17].

The therapeutic potential of disrupting interactions between the SH2 domain of Src and its cognate receptor ligands, thereby disabling these pathways in hyperproliferative cells, is attractive. Reduction of Src levels in an ovarian cancer cell line expressing antisense Src RNA has been associated with regression in tumor size in mouse xenografts, but otherwise normal proliferative functions [18]. Furthermore, a reduction

in bone resorption was the only significant phenotypic difference between wild-type and homozygous null $src^{-/-}$ mice [19]. These results demonstrate the viability of such a therapeutic campaign, both in the sense that inhibition of Src can impact aberrant signalling pathways and that side effects are limited. In fact, affecting bone remodelling may be exploited in the treatment of osteoporosis. To date, small molecule inhibitors exhibit significant binding affinity but limited efficacy, presumably due to poor cellular penetration [20]. Research in neutral prodrug analogs and phosphate mimetics that prevent enzymatic dephosphorylation may extend the utility of these compounds [21–23].

The binding site of the SH2 domain has been well-characterized by numerous X-ray and NMR structures of the SH2 domain complexed with high affinity peptides and peptidomimetic compounds [10, 20, 24–28]. Peptides bind in an extended manner, anchored by a Y^* binding pocket that is characterized by ion-ion, hydrogen-bonding, cation- π , and CH_2 - π interactions. A shallow, largely hydrophobic cleft preferentially binds a β -branched, nonpolar residue (Ile or Val) three residues C-terminal to Y^* . An additional region of interaction between these two sites exists; three water molecules mediate indirect surface contacts to the $Y^* + 2$ glutamic acid in the consensus pentamer (Figure 1) [20].

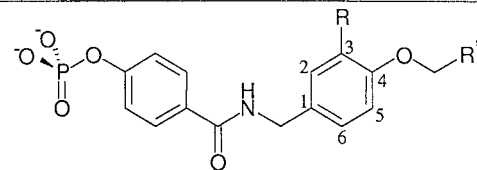











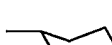



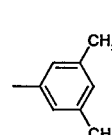
Experimental structures have not been reported for the present ligands with Src SH2 domain. However, an X-ray crystal structure of a related compound, **8**, reveals a hydrogen bond between the 3-amide carbonyl ($C=O$) and Lys $\beta D6$'s backbone amide nitrogen (HN) [10]. Furthermore, the 3-amide displaces a structural water molecule, thus improving the entropic contribution to binding. It is unclear whether an amide at this position is ideal, as protein contacts to the amide NH_2 are distant. This study aims to more comprehensively characterize the binding geometry of these flexible ligands and analyze the requirements for high affinity binding within this binding region using MC/FEP methodology.

Computational methods

Initial geometries

Protein coordinates are based on an X-ray crystal structure (PDB entry: 1SHD) [20] of the SH2 domain bound to a consensus pentamer (Y^*EEIE). The N- and

Table 1. Available experimental IC₅₀ data.

			
Compound	R	R'	IC ₅₀ (μM)
1			6.5 ^a
2			30 ^a
3			69 ^b
4			
5			
6			
7			> 100 ^a
8			5.6 ^a

^aRef. 10.^bLunney, E.A., personal communication.

C-termini were capped with acetyl and methylamine groups, respectively, to neutralize the backbone; there are a total of 107 residues including caps. Coordinates for 4 residues at the beginning of the βC strand, unresolved in the crystal structure, were taken from an average NMR structure of the same complex (PDB entry: 1HCS) [27]. The ligand with the highest experimental affinity, **1**, was modeled into the binding site based on the coordinates of the consensus peptide and on qualitative structural comments made on the unavailable crystal structure of the Src SH2 domain complexed with **8** [10]. Specifically, the phosphophenyl ring was overlaid on the phosphotyrosine in the available peptide complex structure, and the hydrogen bonds between the ligand's backbone HN and

His βD4 C=O and between the 3-amide C=O and Lys βD6 HN were manually placed before minimization. Twenty-four steps of conjugate gradient minimization were performed to alleviate poor steric contacts. The minimization was performed in a distance dependent dielectric ($\epsilon = r$) in order to dampen electrostatic interactions that would otherwise be overemphasized in a vacuum. The initial structures for the other complexes were derived by straightforward replacement of the amide group of [1], and identical coordinates were used for the protein backbone in all perturbations.

For the unbound ligands, structures were generated from 200 starting points using conformational searching with the GB/SA continuum solvation model [29]. Conformational searching is a utility in the

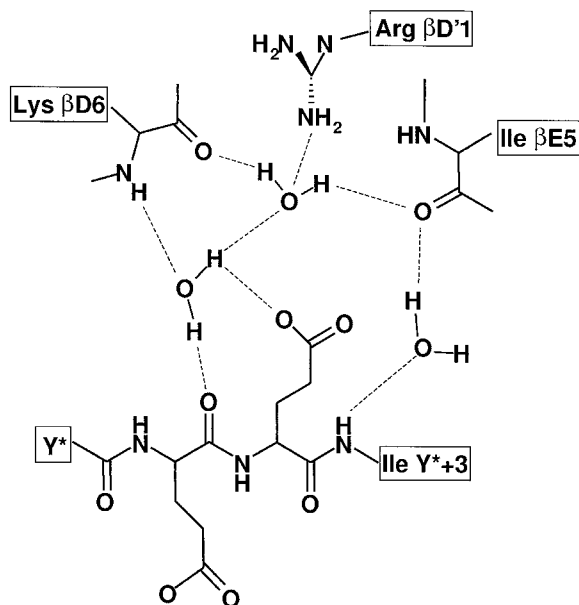


Figure 1. Schematic showing water coordination in the glutamic acid binding region from crystal structure of the SH2 domain with pentamer Ac-Y*EEIE (PDB entry: 1shd) [20].

BOSS program that entails generating a collection of starting structures by making random variations for specified degrees of freedom followed by energy minimization for each structure [30]. The lowest-energy structure was used as a starting point for the MC/FEP simulations.

Gas-phase relaxed potential energy scans of ligand 1

To further explore the conformational preferences for the ligands, torsional energy scans (repetitive energy minimizations where one dihedral angle is systematically varied over 360°) were performed in a distance-dependent dielectric on compound **1** and several of its derivatives. To improve efficiency, the geometry of the phosphophenyl group and the adjacent amide were not optimized. Every freely rotatable bond was independently scanned with the exception of those in the phosphate.

Free energy perturbations

FEP Protocols for “alchemical transformations” in MC simulation have been discussed previously in detail [31]. Briefly, the change in free energy for a perturbation from one state, A, to another, B, is described by Zwanzig’s equation:

$$\Delta G(A \rightarrow B) = -k_B T \ln \langle \exp[-(E_B - E_A) / k_B T] \rangle_A, \quad (1)$$

where $\langle \rangle_A$ denotes an ensemble average sampled in state A [32]. When states A and B are considerably different, as is usually the case when transforming one functional group into another, ΔG will converge slowly because the low-energy phase space for system B is rarely sampled. It is therefore necessary to perform the mutation in a series of smaller perturbations, or windows. The transformations under investigation here are performed in the single-topology paradigm, that is, the atomic, bond-stretching, angle-bending, and torsional parameters of the initial functional group (collectively denoted as ξ) are smoothly mapped into the parameters of the second group as a function of an imaginary reaction coordinate, λ , such that the parameters at any intermediate state between the endpoints $\lambda = 0$ and $\lambda = 1$ can be described by:

$$\xi_\lambda = \lambda \xi_1 + (1 - \lambda) \xi_0. \quad (2)$$

If the change in functionality involves a change in the number of atoms, real atoms must perturb to (or from) dummy atoms (DM) that have no charge or Lennard–Jones parameters. Dummy atoms have equilibrium bond lengths of 0.4 \AA , a value shorter than the hard shell Lennard–Jones radius of any connecting atom, thereby preventing convergence problems owing to cavitation at the endpoint [4].

All FEPs here involved 10 windows of double-wide sampling, where each window was equilibrated under the conditions discussed below. The total change in free energy of the transformation is the sum over the results for the series of steps. The difference between the ΔG calculated for the mutation in the binding site and the mutation when the ligand is free in solution equals the relative free energy of binding ($\Delta \Delta G_b$).

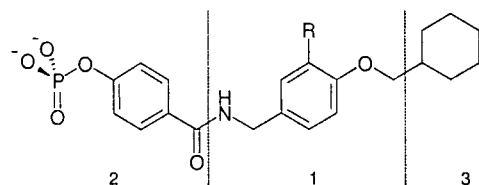
Simulation details

Subsequent to optimization, all bond lengths not directly involved in the perturbation were held rigid. All bond angles and dihedral angles in the ligand and protein side chains were variable, while all internal degrees of freedom of the protein backbone and of residues with more than 3 atoms 18 \AA or farther from the center of the ligand were held rigid. A sphere with a 20 \AA radius of pre-equilibrated TIP4P water [33] was centered on the ligand in the binding site.

A small half-harmonic potential (force constant of 1.5 kcal/mol-Å²) was applied to any water that leaves the sphere's radius, as determined by the distance from the center of the sphere to the water's oxygen, thus preventing evaporation. Total translations and rotations for the water molecules and ligand were sampled.

The protonation states of all histidines were assigned by inspection; two histidines (βD4 and BG6) are protonated, and the remaining histidine αB12) is the δ-tautomer. Charged protein residues 18 Å or farther from the ligand were neutralized so as to avoid deformation of the solvent sphere; neutralization of distant residues changed the net charge of the complex from +5 to +3. As the predicted ligand structure was considerably more extended when unbound, a 22-Å sphere of water was necessary to ensure complete solvation of the phosphate and cyclohexyl groups in the unbound simulations.

All simulations were carried out at 25 °C. To improve sampling of the ligands, they were divided into three artificial residues with integral charges:



Attempted Monte Carlo moves were performed for one water molecule or for one protein or ligand residue at a time using the Metropolis algorithm [34]. A residue-based cutoff of 10 Å was imposed [35]. Water-only equilibration was performed for 10 million (M) configurations where waters close to the ligand were preferentially sampled [36]. Subsequently, 10–130 M configurations of equilibration were performed with sampling of all variable degrees of freedom. Averaging periods were variable for reasons discussed below.

No translational or rotational moves were executed for the protein. Ranges for acceptable translations and rotations for the ligands were 0.01 Å and 0.2°, respectively, and were chosen to enhance ligand acceptance rates. Acceptable maximum dihedral moves ranged from 2° for ligand backbone motions to 15° for methyl rotations, with other maximum dihedral moves lying within this range depending on the number of atoms affected by the rotation. These choices resulted in acceptance ratios of ca. 0.4, 0.4, and 0.15 for attempted moves of water molecules, protein residues, and ligand residues, respectively. On a 500 MHz Pentium III, it

takes ca. 1.5 and 8.0 h to execute 1 M Monte Carlo configurations for the unbound ligand in water and for a protein-ligand complex in water, respectively.

Statistics

Reported uncertainties for changes in free energies are standard errors of the mean (*sem*) and are calculated based on accumulating separate averages for each 'batch' of 0.25 M or 0.5 M configurations. The *sem* of the Δ*G* is related to the standard deviation of the means by the number of batches, *N*, as follows:

$$sem = \sigma / \sqrt{N} \quad (3)$$

Structural correlations were also performed with snapshots recorded every 0.25 M or 0.5 M configurations. The Pearson product-moment or sample correlation coefficient, *r_{xy}*, is calculated as follows:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad (4)$$

and measures the extent to which two variables, *x* and *y*, are linearly related. The absolute value of *r_{xy}* ranges from zero to one where a value of one indicates an exact linear relationship, and the sign of *r_{xy}* corresponds to the direction of the slope. Structural correlations in this study are most likely underestimated because the snapshot from each batch may not be representative of all of the structures sampled in the batch from which the thermodynamic quantities are calculated.

Force field and software

All parameters are taken from the OPLS all-atom force field [1]. Previously unpublished parameters are available in the Supporting Information. All MC/FEP calculations were performed with MCPRO, version 1.6 or greater [37]. Torsional scans and gas-phase minimizations of representative model systems were done with BOSS, version 4.0 or greater [30]. *Ab initio* calculations were performed with Gaussian 95 [38].

Results

Effects of angular constraints on free energy convergence

The 1 → 2 and 2 → 1 mutations were performed where (1) the dummy angle force constants and equilibrium values were unchanged from its methyl counterpart, and (2) two angles per dummy atom (DM-H-DM) have their *k_θ* removed (Table 2). With the

Table 2. Bond-stretching and angle-bending parameters for dummy atoms in amide endpoint of forward and reverse mutations of **1** \rightarrow **2**^a.

Bond/Angle	Equilibrium value (r_{eq} or θ_{eq})	Case 1 – force constant (k_r or k_θ)	Case – force constant (k_r or k_θ)
DM-N	0.4	340.0	340.0
DM-N-H	109.5	35.0	35.0
DM-N-DM	107.8	33.0	0.0

^aTorsional parameters (V_1 , V_2 , V_3) for dihedral angles involving a dummy atom equal 0.

smaller number of constraints, the hysteresis in the $\Delta\Delta G_b$ is 0.9 kcal/mol as compared to a hysteresis of 0.2 kcal/mol achieved in a fraction of the simulation time in the system with the larger number of constraints (Table 3). In addition, the total hysteresis in the former case is not completely indicative of its convergence problems, as the individual bound and unbound legs had much larger hystereses (1.0 and 1.9 kcal/mol, respectively)

The majority of these hystereses arises from the windows closest to the amide endpoint ($\lambda = 0.05$ of **1** \rightarrow **2** and $\lambda = 0.95$ of **2** \rightarrow **1**) of both the bound (Figure 2a) and unbound legs. Fluctuations in ΔG in these windows are in the range of 1–3 kcal/mol, and correlate best ($r_{xy} \geq 0.99$) with the change in angle-bending energy, ΔE_{angle} , whose fluctuations are on the same order of magnitude (Figure 2b). Furthermore, these two quantities correlate strongly ($r_{xy} \geq 0.91$) to the square of the deviations in the DM-H-DM angles from their equilibrium values (see note in Computational Details). Windows farther from the amide endpoint have progressively smaller fluctuations in ΔE_{angle} ; windows $\lambda = 0.15$ and 0.25 have ΔE_{angle} fluctuations on the order of 0.8 and 0.6 kcal/mol, respectively, which correlate well with ΔG ($r_{xy} \geq 0.92$). Subsequent windows have fluctuations in ΔE_{angle} that fall within the other noise in the simulation.

Noise in ΔG and ΔE_{angle} is dramatically reduced in the simulations with the higher number of constraints (Figure 2c). The absolute value of these ΔG 's are expected to be different for reasons discussed below. For comparison, fluctuations in ΔE_{angle} in any window of the more constrained perturbations are smaller than 0.03 kcal/mol.

Relative binding affinities of ligands

When possible, perturbations were run with three angles defined with $k_\theta > 0$ kcal/mol rad² per dummy

atom (**1** \rightarrow **2**, **5** \rightarrow **6**, and **5** \rightarrow **7**), or with two angles otherwise (**1** \rightarrow **3**, **1** \rightarrow **4** and **4** \rightarrow **6**). In these perturbations, the ΔG at the windows closest to the endpoint with dummy atoms still correlate well with fluctuations in ΔE_{angle} ($r_{xy} \geq 0.88$), however the fluctuations in these perturbations are considerably smaller (< 1 kcal/mol) than with only one angle per dummy atom defined with a non-zero force constant. In the **1** \rightarrow **5** mutation, ΔG was poorly converged when each dummy atom had only 2 angles defined with $k_\theta > 0$ kcal/mol rad², owing in part to the fluctuations in ΔE_{angle} , but perhaps also related to the larger change in geometry than for the other mutations. A dummy atom constrained to more than two non-dummy atoms with angle or dihedral constraints produces an additional, nonphysical coupling between those real atoms. However, mutating **1** \rightarrow **5** with three angular constraints on the dummy atom that becomes the methyl hydrogen improves convergence significantly ($\Delta\Delta G_b$ reported in Table 4). Interpretation of this result requires the assumption that the contribution from this additional constraint cancels when considering the bound and unbound legs.

Free energies for all FEPs are listed in Table 4. The cycle of **1** \rightarrow **4** \rightarrow **6** \rightarrow **7** \rightarrow **5** \rightarrow **1** produces an acceptable $\Delta\Delta G$ of 1.1 kcal/mol, and the forward and reverse transformation **1** \rightarrow **2** \rightarrow **1** yields a $\Delta\Delta G$ of 0.2 kcal/mol. Though large hystereses for such cycles are indicative of poor convergence, small hystereses do not guarantee good convergence, if the equilibration and averaging periods are much shorter than the relaxation time of the system [39]. In practice, it is often difficult to gauge local relaxation in a large simulation, as the decay of the relevant energetic terms often falls within the noise of the simulation. Consequently, it is necessary to examine the running average of ΔG as a function of the averaging period to ascertain if initial high-energy states are biasing the result (Figure 3). Averaging periods were therefore deter-

Table 3. FEP results for forward and reverse mutations of **1** \rightarrow **2** for two cases of angular constraints.

Mutation	Number of angles to DM with $k_\theta > 0$	$\Delta\Delta G_{\text{bound}}$ (kcal/mol)	Simulation length - bound (M)	$\Delta\Delta G_{\text{unbound}}$ (kcal/mol)	Simulation length - unbound (M)	$\Delta\Delta G_b$ (kcal/mol)
1 \rightarrow 2	1	21.7 ± 0.2	12	22.0 ± 0.1	80	-0.3 ± 0.2
2 \rightarrow 1	1	-19.8 ± 0.2	10	-21.0 ± 0.1	40	1.2 ± 0.2
1 \rightarrow 2	3	18.3 ± 0.2	4	18.3 ± 0.1	20	0.0 ± 0.2
2 \rightarrow 1	3	-18.2 ± 0.2	4	-18.0 ± 0.2	20	-0.2 ± 0.3

Table 4. Results of FEPs

Mutation	$\Delta\Delta G_{\text{bound}}$ (kcal/mol)	Simulation length - bound (M)	$\Delta\Delta G_{\text{unbound}}$ (kcal/mol)	Simulation length - unbound (M)	$\Delta\Delta G_b$ calculated (kcal/mol)	$\Delta\Delta G_b$ expt. ^a (kcal/mol)
1 \rightarrow 2	18.3 ± 0.2	4	18.3 ± 0.1	20	0.0 ± 0.2	0.9
2 \rightarrow 1	-18.2 ± 0.2	4	-18.0 ± 0.1	20	-0.2 ± 0.3	-0.9
1 \rightarrow 3	26.8 ± 0.2	8	26.4 ± 0.2	20	0.4 ± 0.3	1.4
1 \rightarrow 4	7.3 ± 0.3	10	3.8 ± 0.1	30	3.5 ± 0.3	
1 \rightarrow 5	12.5 ± 0.2	4	9.7 ± 0.2	20	2.8 ± 0.3	
4 \rightarrow 6	18.5 ± 0.1	6	20.70 ± 0.06	40	-2.2 ± 0.1	
5 \rightarrow 7	4.6 ± 0.1	4	2.76 ± 0.07	20	2.4 ± 0.1	
6 \rightarrow 7	-1.58 ± 0.07	4	-3.75 ± 0.05	20	2.17 ± 0.3	
1 \rightarrow 5 \rightarrow 7 ^b					4.6 ± 0.3	> 1.6
1 \rightarrow 4 \rightarrow 6 \rightarrow 7 ^b					3.5 ± 0.3	> 1.6

^aExperimental $\Delta\Delta G_b$ was calculated based on the assumption that the ratio of IC₅₀ is equal to the ratio of dissociation constants.

^bIndividual mutations for **1** \rightarrow **7** were not explicitly performed, but are determined from the sum of several transformations.

mined by inspection on a case-by-case basis and are reported in Table 4.

The binding affinity of compound **1** is underestimated in this study; the computations predict **1** and **2** as approximately equipotent. Calculated relative free energies of hydration of acetamide and N-methyl acetamide (NMA) using these atomic parameters predict acetamide as too hydrophilic by ca. 1.7 kcal/mol [40]. Difficulties in reproducing experimental trends in amide solvation are not peculiar to these parameters. In fact, recent amide parameterizations have also been plagued by the anomalous trends in amide solvation as a function of increased alkyl substitution [41, 42]. Improved parameters that address these concerns are being prepared presently [40]. On the other hand, pure liquid simulations of acetamide and NMA with these parameters produced densities and heats of vaporization within 2% of experiment [1]. Taken together, these results suggest that a correction from improved parameterization will apply more to the unbound leg, thus improving agreement with experiment on the **1** \rightarrow **2** and **2** \rightarrow **1** mutations. However, the amide is

still partially solvated (Figure 4) in the complex, so 1.7 kcal/mol is an upper bound of such a correction.

The experimental IC₅₀ of the unsubstituted compound, **7**, was reported as a lower limit, as its affinity was beyond the sensitivity of the assay. Though the **1** \rightarrow **7** mutation was not explicitly performed, the relative free energy of binding of **7** can be determined by summing the $\Delta\Delta G_b$ of several mutations. The calculated relative free energies as determined by the unique pathways **1** \rightarrow **5** \rightarrow **7**, **1** \rightarrow **4** \rightarrow **6** \rightarrow **7**, and **1** \rightarrow **5** \rightarrow **6** \rightarrow **7** range from 3.5 to 4.6 kcal/mol, all in accord with the experimental findings. Overall, the predicted order of binding agrees with experiment, though the selectivity for **1**, **2**, and **3** is compressed. Of the compounds without experimental data, the amino, methyl, and unsubstituted derivatives (**4**, **5**, and **7**) are predicted to be poor binders, whereas the chloro analog, **6**, is intermediate between **3** and **7**.

General comments on binding site geometry

The conjugate gradient minimization of the protein-**1** complex did not deviate significantly from the starting crystal structure model. The *rmsd* to backbone atoms

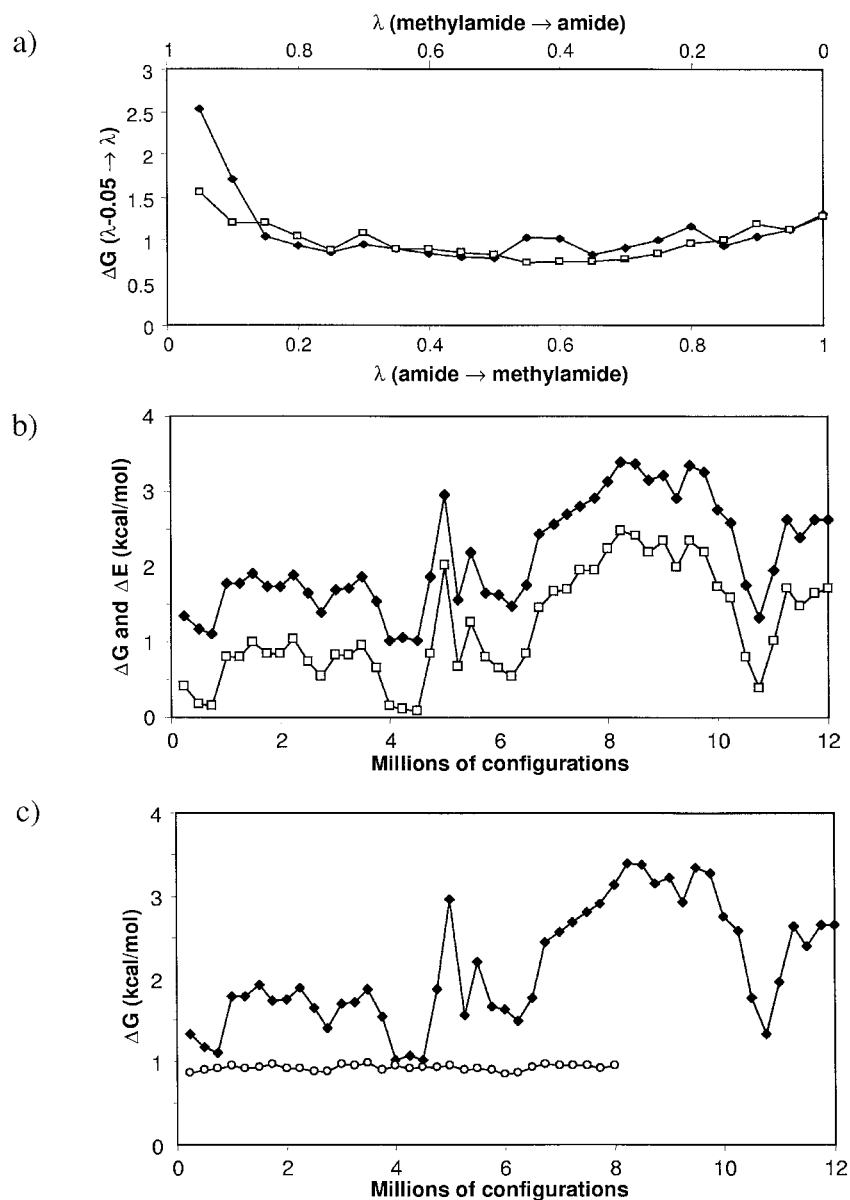


Figure 2. (a) ΔG for each window of $1 \rightarrow 2$ (closed diamonds) and $2 \rightarrow 1$ (open squares) mutations where dummy atoms have only one $k_{\theta} > 0$ kcal/mol rad². (b) Comparison of $\Delta G_{\lambda=0.0 \rightarrow \lambda=0.05}$ (closed diamonds) and $\Delta E_{\text{angle}, \lambda=0.0 \rightarrow \lambda=0.05}$ (open squares) for $1 \rightarrow 2$ mutation where dummy atoms have only one $k_{\theta} > 0$ kcal/mol rad². (c) Comparison of $\lambda = 0.05$ window of $1 \rightarrow 2$ mutation where dummy atoms have one (closed diamonds) or three (open circles) angles with $k_{\theta} > 0$ kcal/mol rad².

is 0.51 Å, and the *rmsd* for all carbon atoms is 1.03 Å. The critical phosphate and phosphophenyl ring interactions noted in reference 20 are readily apparent in these simulations, including the bidentate ion bridge to Arg β D5, CH/ π interactions with Lys β D6, sub-2-Å hydrogen bonds to Ser β B7 and Thr BC2, and both the amino-aromatic and ion-ion interactions with Arg α A2. Adjacent to the phosphate, the hydrogen bond

between the ligand's backbone HN and His β D4 C=O is nearly linear (ca. 165°).

Only two distances were reported from the crystal structure with **8**; the N-O distance for the His β D4 C=O ... ligand HN hydrogen bond is 2.9 Å, and the N-O distance of the Lys β D6 HN ... ligand C=O hydrogen bond is 3.0 Å [10]. Averaged over all near-amide windows, these distances with **1** were similar,

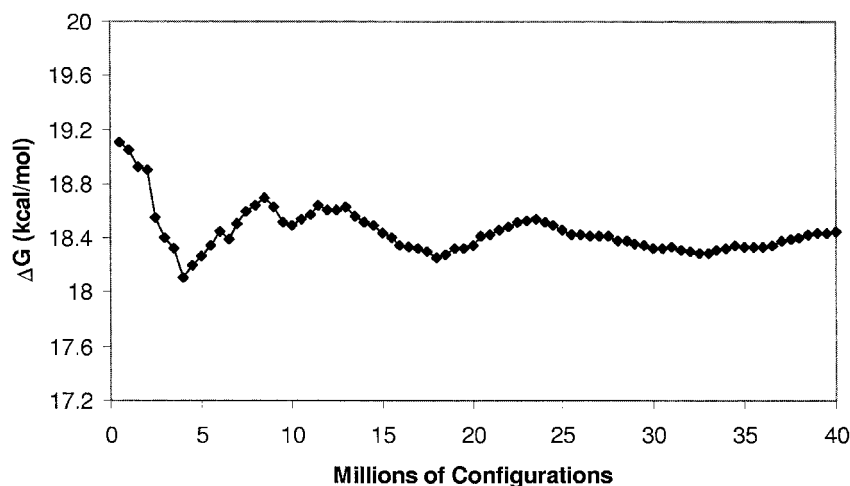


Figure 3. The running average of $\Delta G_{1 \rightarrow 2}$ (sum of all windows) for the unbound leg with 3 angles with $k_\theta > 0$ kcal/mol rad² per dummy atom as a function of simulation duration.

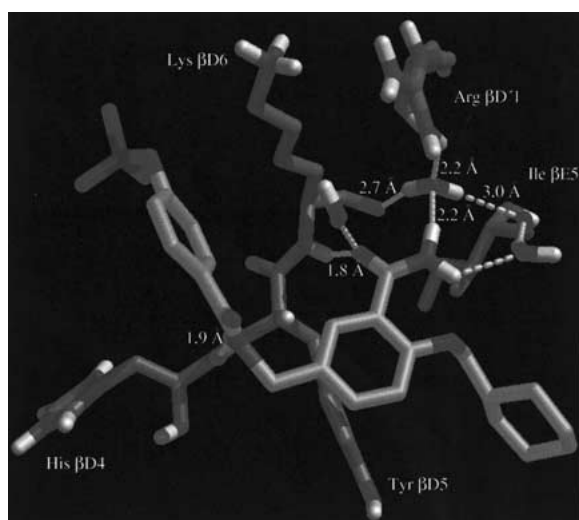


Figure 4. Compound **1** bound to residues in and near the glutamic acid binding region. Ligand carbon atoms are shown in gold; aliphatic hydrogens are not shown for clarity. Distances are averaged over the $\lambda = 0.05$ windows of **1** \rightarrow **2** \rightarrow **2**, **1** \rightarrow **3**, **1** \rightarrow **4**, **1** \rightarrow **5**, and the $\lambda = 0.95$ window of **2** \rightarrow **1**. Water molecules without distances exchanged during at least one equilibration or averaging period. Graphics prepared with MidasPlus [48].

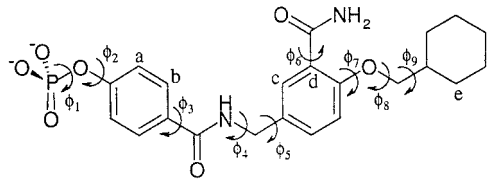
2.9 Å and 2.7 Å, respectively. Lunney and coworkers discussed a subtle change in the binding mode of **8**. They observed that Arg α A2 swings away from the phosphate, thus losing its cation/ π interactions with the phosphophenyl ring and allowing the ring to rotate. A minimization where χ_3 of Arg α A2 is manually repositioned from -75° in the peptide crystal structure to 180° results in Arg α A2 rotated away from

the phosphate and interacting with Glu α A6. However, the accompanying rotation of the phosphophenyl ring was not observed. It is unclear whether this interaction would be preferred in the presence of solvent, as both charged side chains could be otherwise solvated. As complete structural information was available for the $\chi_3 = -75^\circ$ conformer, and the rest of the structure of the complex is unaffected by the position of Arg α A2, the $\chi_3 = -75^\circ$ rotation was used for all FEPs.

The cyclohexyl methoxy moiety is highly mobile in the hydrophobic binding site, typically having an *rmsd* of at least a factor of two times that of the phosphophenyl group within any given simulation. The acceptance rates also reflect this higher mobility; the cyclohexyl methoxy ‘residue’ had an acceptance rate of 26% in this window as compared to 9% for the phosphophenyl moiety. The majority of this mobility stems from the three torsions between the cyclohexyl group and the central ring, each of which had broad distributions spanning ± 20 – 35° of their respective means.

When bound, several ligand dihedral angles do not occupy their respective solution-phase minima as determined by the GB/SA conformational search (Table 5). The bound conformation is about 6 kcal/mol higher in energy than the unbound global minimum including the GB/SA term. In particular, the phenyl ether torsion ϕ_7 is significantly out of the ring plane, and the benzyl amide torsion ϕ_5 is virtually coplanar. It should be noted that comparison of the forward and reverse perturbations of **1** \rightarrow **2** reveals that some dihedral distributions have not converged. In particular,

Table 5. Average dihedral values for **1** when bound versus dihedral values of GB/SA-minimized unbound **1**.



Dihedral	Connectivity	Average bound value ^a	Minimized unbound value
ϕ_1	O-P-O-C	49.1°	59.8°
ϕ_2	P-O-C-C _a	176.8°	0.3°
ϕ_3	C _b =C-C-N	352.9°	356.6°
ϕ_4	C-N-C-C	255.0°	265.3°
ϕ_5	N-C-C=C _c	2.6°, 342.0° ^b	90.2°
ϕ_6	N-C-C=C _c	163.4°	183.5°
ϕ_7	C _d =C-O-C	265.6°	177.2°
ϕ_8	C-O-C-C	250.0°	181.6°
ϕ_9	O-C-C-C _e ^c	301.9°	186.6°

^aDihedrals values for complex with **1** are an average of $\lambda = 0.05$ windows of **1** \rightarrow **2**, **1** \rightarrow **3**, **1** \rightarrow **4**, **1** \rightarrow **5** and $\lambda = 0.95$ window of **2** \rightarrow **1**, weighted by length of averaging period.

^bTwo averages are listed when distributions are bimodal.

^cThe hydrogen of the connecting carbon of the cyclohexyl ring is pointing towards the reader.

the benzyl amide torsions in $\lambda = 0.05$ of **1** \rightarrow **2** and $\lambda = 0.95$ of **2** \rightarrow **1** are largely non-overlapping, though both positions are distant from the preferred orthogonal geometry.

Glutamic acid binding region

Though water was not included in the minimizations, several water molecules in the glutamic binding region in the simulations occupy positions similar to water molecules resolved in the pentamer complex structure, minus one water molecule displaced from its Lys β D6 HN contact. Of the remaining bound water molecules in this region, one is highly coordinated between one Arg β D'1 guanidium hydrogen, Lys β D6 C=O, and Ile β E5 C=O, and was observed in every snapshot (recorded every 0.25–0.50 M configurations) of every simulation. Average hydrogen bond lengths to this water are shown in Figure 4. Compound **1** is the only ligand in this series that could hydrogen-bond to this ubiquitous water molecule. For the rest of the ligands, there is effectively a desolvation penalty for the water molecule, as it loses a hydrogen bond to another solvent molecule upon binding. In addition to the contact with the bridging water molecule, the C=O of **1** also

makes a hydrogen bond of average length 1.8 Å to Lys β D6 HN. The amide rotates out of the ring plane by ca. 16°, presumably to improve this contact. Another water molecule bridges the second amide hydrogen to Ile β E5 C=O.

In order to accommodate the methyl group of **2**, the bound water molecule translates further into the pocket by ca. 0.6 Å. The possibility that the methyl amide assumes a *cis* geometry was considered, as the amide hydrogen would be able to maintain its hydrogen bond to the bridging water molecule. A gas-phase torsional scan of the model compounds, benzamide and *o*-methoxy benzamide, reveals that trans-*cis* isomerization incurs an 5.9 kcal/mol penalty in internal energy for the parent model, and a 7.5 kcal/mol penalty with the *ortho*-substituent present. The former value was confirmed with *ab initio* calculations at the HF/6-31G*//HF/6-31G* level, resulting in a ΔE trans \rightarrow *cis*) of 5.5 kcal/mol. Therefore, it is unlikely that **2** will adopt a *cis* conformation.

Interestingly, compound **3** is preorganized to make the out-of-plane contact between its C=O and Lys β D6 HN. The methyl of the ketone interacts unfavorably with the *ortho*-ether substituent, causing the ketone to rotate out of the ring plane by ca. 44° in so-

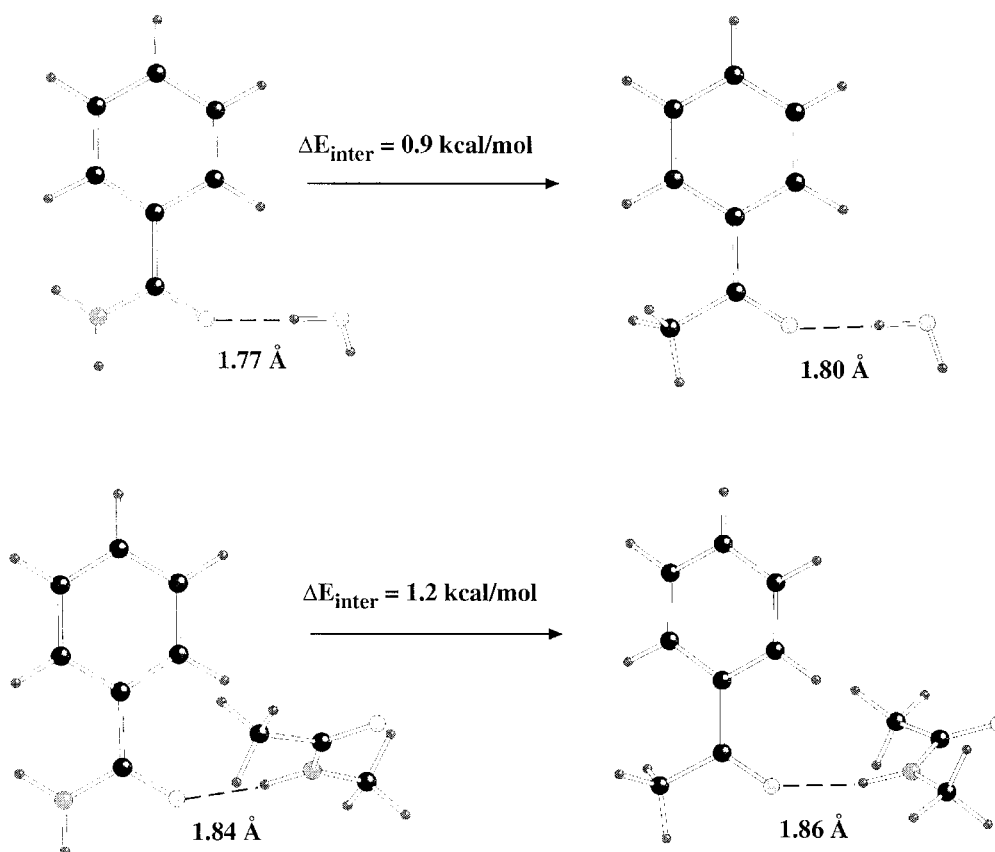


Figure 5. Relative interaction energies for complexes of benzamide and acetophenone with water (top) and with N-methylacetamide (bottom).

lution. In complex, the ketone rotates out-of-plane by an additional 10° , presumably from steric interactions with the ether methylene, which is positioned close to the ketone in the binding geometry. The hydrogen bond consequently overshoots linearity by ca. 12° . In addition, gas-phase bimolecular minimizations using the OPLS-AA force field reveal that the benzamide carbonyl oxygen is a better hydrogen bond acceptor than acetophenone, and relatively stronger when complexed with N-methyl benzamide versus water (Figure 5). The results imply that the amido compounds benefit more from replacing a solvent hydrogen bond with a hydrogen bond to the protein backbone than does the keto compound.

Two binding modes are observed for the amino compound, **4**. Transformation of **1** \rightarrow **4** swaps hydrogen-bonding capacity at the carbonyl oxygen position of **1**, potentially creating repulsive electrostatic interactions for **4**. In the later windows of this perturbation, the amine rotates, moving the nitrogen within 2.6 Å of Lys β D6 HN and turning the offending amine

hydrogen towards the solvent (Figure 6a). Alternatively, in the earlier windows of the **4** \rightarrow **6** mutation, the entire ligand moves farther towards the bridging water molecule, thereby lengthening the amine hydrogen...Lys β D6 HN distance to ca. 3.1 Å. This amine hydrogen is unable to hydrogen bond in this position, while the nitrogen makes a contact to a solvent molecule above the plane of the amine. Translation of the ligand also lengthens the backbone amide...His β D4 C=O hydrogen bond by ca. 0.3 Å (Figure 6b). In both cases, the second amine hydrogen forms a hydrogen bond to a water molecule that is, in turn, hydrogen bonding with Ile β E5. It is interesting to note that these binding modes do not interconvert within a single simulation.

Compounds **5**–**7** cannot hydrogen-bond to either the bridging water or the Lys β D6 HN. Furthermore, they block solvent access to these sites, which contributes to the desolvation penalty for the system. The possibility exists that these ligands are more mobile in the binding site in the absence of an anchoring hydro-

gen bond in the glutamic acid binding region. Indeed, comparison of snapshots at the endpoints of the **5** \rightarrow **6**, **5** \rightarrow **7** and **6** \rightarrow **7** mutations show average intersolute distances varying by as much as 0.7 Å. Disparity in the distribution of positions sampled in different simulations may contribute to the 1.1 kcal/mol hysteresis observed in the smaller cycle.

Unbound ligand structure

The GB/SA conformational search generated 63 minimized structures for **1** within 7 kcal/mol of the global minimum. The six lowest-energy structures were within 0.25 kcal/mol and consisted of combinations of the two orthogonal benzyl amide torsional minima and the three possible cyclohexyl torsional minima (*trans*, *gauche*[−], and *gauche*⁺). The geometries of the next eight structures (within 1 kcal/mol of global minimum) likely would be less populated in the presence of explicit solvent than the GB/SA minimizations indicate. Specifically, the carbonyl of the 3-amide points towards the 4-ether on the central ring, though out of the ring plane by $\pm 38^\circ$, thus limiting solvent access to both the carbonyl and ether oxygens.

Less than 10% of all of the unbound windows experienced a transition between conformers in the equilibrium or averaging periods. Clues to the effect that a change in conformation may have on the results of the perturbation can be gleaned from comparison of the **1** \rightarrow **2** and **2** \rightarrow **1** mutations. For example, the $\lambda = 0.75$ window of **1** \rightarrow **2** samples in region of the alkyl-ether torsional minimum (180°), while the $\lambda = 0.25$ window of **2** \rightarrow **1** samples in the *gauche*⁺ minimum (75°). However, the two free energies accumulated in each of the two windows agree to within 0.01 kcal/mol. On the other hand, the $\lambda = 0.35$ window of **2** \rightarrow **1** samples a high-energy, out-of-the-ring-plane conformations of the 3-amide while the $\lambda = 0.65$ window of **1** \rightarrow **2** samples a much more planar arrangement of the central ring and the 3-amide, and a second conformer of the cyclohexyl group. The changes in free energy between these two windows disagree by ca. 15% (0.15 kcal/mol), suggesting that one or both of these conformational changes may give new contributions to the ΔG .

As discussed above, comparison of the GB/SA minimized ligand structure to bound ligand structures reveals that the ligands are not preorganized to their binding geometries. Modification of the ligands may improve binding affinity by virtue of reducing the internal energy penalty for adopting the binding

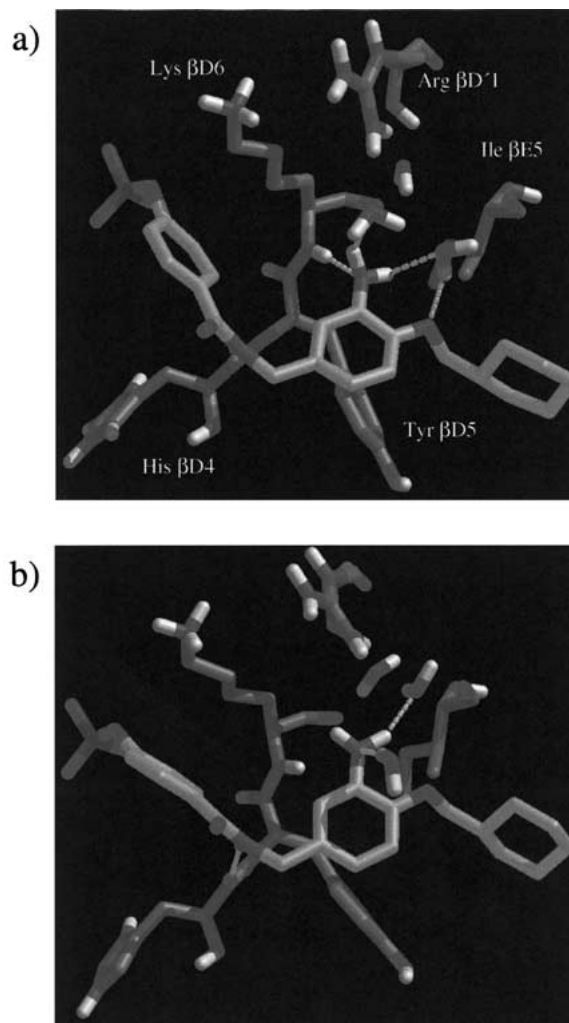


Figure 6. (a) Amino compound, **4**, from $\lambda = 0.95$ window of **1** \rightarrow **4** and (b) from $\lambda = 0.05$ window of **4** \rightarrow **6**. Ligand carbon's are gold; aliphatic hydrogens have been hidden for clarity. Graphics prepared with MidasPlus [48].

conformation. Indeed, torsional scans performed in a distance-dependent dielectric indicates that stereospecific methylation of the benzyl amide methylene may lower the relative energy of the binding conformation. Lunney et al. noted that a racemic mixture of compound **1** methylated at this position shows slightly improved binding (experimental $\Delta\Delta G_b = 0.09$ kcal/mol) [10]. The value of the benzyl amide torsion (ϕ_5 in Table 5) in bound, near-amide windows is bimodal; whereas pro(R) methylation is necessary to reduce the relative energy of one conformer, either pro(R) or pro(S) methylation can affect the other (Figure 7a). As pro(R) methylation would position this substituent into bulk solvent, binding would be likely

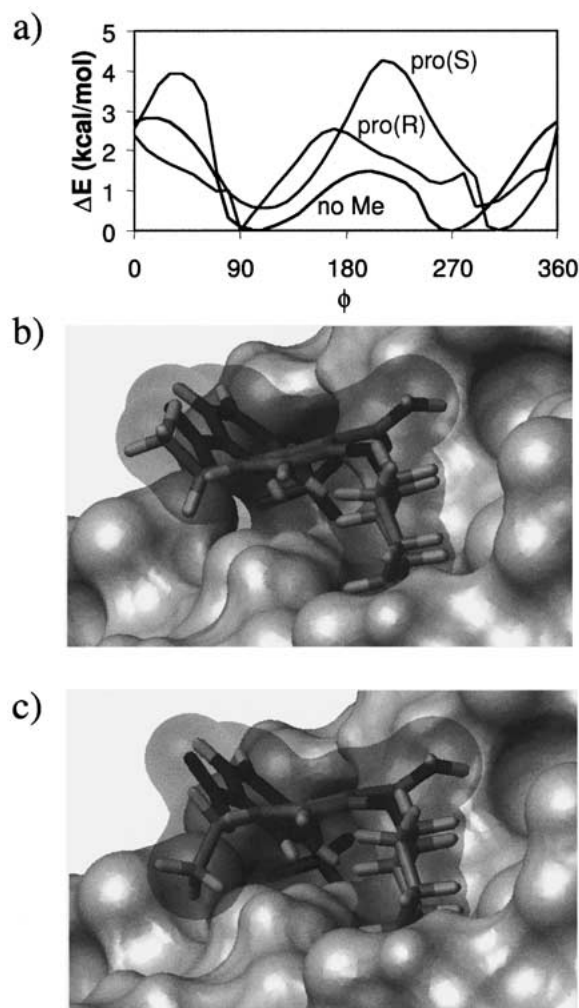


Figure 7. (a) Rotational scan of benzyl amide torsion (N-C-C=O) performed in a distance-dependent dielectric with or without stereospecific methylation at the benzyl carbon. (b) Compound **1** modeled with pro(R) methylation at benzyl amide methylene, and (c) with methylation at the pro(S) position. The methyl group is at the far left in the structures. Graphics prepared with MOLMOL [49].

unaffected (Figure 7b). Pro(S) methylation can also be accommodated, fitting into a small pocket between His β D4 and Tyr β D5 (Figure 7c). This analysis suggests that either enantiomer may confer some advantage to the ligand in adopting the binding conformation.

Similarly, methylation at the pro(S) position of the ether methylene may promote the binding geometry of the alkyl-ether torsion (ϕ_8 in Table 5). However, modeling of this modification suggests that without some rearrangements in the binding site, particularly the EF loop, it is unlikely that a methyl group can be accommodated in this position. Finally, a gas-

phase torsional scan of anisole reveals that there is a ca. 2 kcal/mol penalty for the phenyl ether to deviate from the ring-plane by 90° (ϕ_{17} in Table 5). In contrast, thioanisole prefers the perpendicular geometry to the planar conformation by ca. 1.8 kcal/mol. Unfortunately, modeling of the ether to thioether substitution shows that the longer carbon-sulfur bonds as compared to carbon-oxygen bonds will likely cause steric clashes in the hydrophobic binding pocket for the cyclohexyl moiety or a slightly smaller cyclopentyl group.

Discussion and conclusions

As discussed in the introduction, it seems attractive to limit the number of topological constraints on a disappearing atom, so as to have the endpoint of the perturbation most closely resemble the molecule without the dummy atoms. However, it is clear from the simulations performed here that the equienergetic cone of space available to the dummy atoms when constrained with only one angle and one bond is not sufficiently sampled within the duration of the simulations, and creates unreasonable fluctuations in the ΔG . Similar observations have been alluded to for the 'ideal gas atom' endpoint for dual-topology perturbations [8]. Increasing the positional constraints on dummy atoms essentially limits sampling of the angles to dummy atoms to the portion of angle-space that minimizes the dissimilarity between the reference and perturbed states. And while it may seem that not sampling these angular degrees of freedom may circumvent the problem all together, it has been shown that sampling of angles promotes torsional exchange, presumably by relaxing tight Lennard-Jones contacts [43]. In other words, having flexible angles may improve sampling of an arguably more relevant part of phase space.

There is a free energy difference at $T \neq 0$ K between a molecule or molecular fragment with and without flexible dummy atoms; consequently, the reader must take care in interpreting the ΔG of a mutation. As a result of having vibrational energy associated with a dummy atom, one cannot expect thermodynamic closure for a cycle of unbound or bound perturbations individually when the number of flexible atoms changes within the cycle. However, the ΔG associated with removing the energy attached to the vibrating dummy atoms is presumably very sim-

ilar in the bound and unbound simulations and will therefore cancel, giving a meaningful $\Delta\Delta G_b$ [44].

Ten years ago, Mitchell and McCammon postulated that the difficulty with free energy calculations in macromolecular systems is the large statistical error owing to the inadequate sampling of phase space [45]. While the previous discussion is relevant in that it addresses poorly sampled angle-space at the endpoint simulations during FEPs, the problem of torsional sampling is also apparent in this study. Even a decade later, with an increase in computer power roughly commensurate with Moore's law, a complete computational analysis of these ca. 70-atom ligands suffers from infrequent interconversion between conformers. The GB/SA continuum model was useful here in limiting the impact of poor sampling by identifying the global minimum of the ligands in solution, thus allowing the unbound simulations to begin sampling in the most relevant part of conformational space. Starting at a higher energy state would increase the likelihood that averaging would be performed in that geometry or while the system is in transition from high to low energy states, that is, while the system is not at equilibrium. Furthermore, accumulation of hysteresis in a thermodynamic cycle owing to the sampling of different minima in adjacent simulations makes evaluation of convergence within individual energetic wells difficult. On the other hand, occasional conformer exchange in the $1 \rightarrow 2$ and $2 \rightarrow 1$ mutations demonstrates that slightly different contributions to the ΔG may come from a more complete sampling of phase space, so true convergence of $\Delta\Delta G_b$ is not seen. Moreover, evaluation of predictions in ligand modifications that affect preorganization are beyond the scope of these simulations, as interconversion of the low-energy conformers is not frequent enough to determine the effect these modifications may have on altering conformer populations.

In the discussion of convergence, the reader should note that the *sem* of the ΔG in these simulations does not always give a strong indicator of convergence in a given simulation, or the uncertainty in ΔG . For example, $\Delta G_{\text{bound}}(1 \rightarrow 2) = 20.3 \pm 0.2$ kcal/mol after 6M configurations with only one angle per dummy with $k_\theta > 0$ kcal/mol rad², whereas the same simulation after 12 M configurations yields a $\Delta G = 21.7 \pm 0.2$ kcal/mol. The reason the *sem* is a poor indicator of uncertainty and convergence in this situation stems from the fact that the blocks are small relative to the frequency of the noise, so individual ΔG measurements (the mean from one batch) are not near the true

mean [46]. In these cases, the *sem* may be more meaningful in interpreting the rate of change in the overall mean.

Despite the challenges in achieving a high level of convergence in FEPs, the simulations predict the correct trends in binding affinity. Furthermore, the simulations with explicit solvent give valuable detail at the atomic level of the binding modes available to each ligand and have located a two-hydrogen-bond criterion distinguishing the ligands. The best binder, **1**, hydrogen-bonds to Lys β D6 HN and a highly-coordinated water molecule. The next tier of ligands, **2** and **3**, make only one of these hydrogen bonds, and the poorest binders (**4**, **5**, **7**) make neither of these contacts. Though the chloro compound, **6**, cannot hydrogen-bond in the traditional sense, it makes favorable electrostatic contacts with Lys β D6 HN, thus ranking in between the latter two groups.

The GB/SA conformational search identified at least six low-energy minima for the ligands in solution. Though the FEPs do not address the absolute ΔG_b for each ligand, the fact that there is a large entropic penalty contributing to the absolute ΔG_b that comes from constraining such a flexible ligand upon binding is expected from such an analysis. Furthermore, comparison of low-energy structures to the distributions obtained from the bound simulations has established what the ligands are not in a low-energy conformation when bound, further contributing an internal-energy penalty to $\Delta\Delta G_b$. Though several predictions can be made for modifications that lower the relative energy of the binding geometry, backbone motion in the bound simulations and frequent interconversion between conformers in the unbound simulations are necessary to confirm them.

Acknowledgements

Gratitude is expressed to Dr Julian Tirado-Rives and Melissa L.P. Price for valuable discussions. This research is supported by the National Institute of Health (GM32136).

References

1. Jorgensen, W.L., Maxwell, D.S. and Tirado-Rives, J., J. Am. Chem. Soc., 118 (1996) 11225.
2. Kong, X. and Brooks, C.L. III, J. Chem. Phys., 105 (1996) 2414.
3. Pearlman, D.A. and Kollman, P.A., J. Chem. Phys., 90 (1989) 2460.

4. Pearlman, D.A., Kollman, P.A., *J. Chem. Phys.*, 94 (1991) 4532. Mark, A.E., van Gunsteren, W.F. and Berendsen, H.J.C., *J. Chem. Phys.*, 94 (1991) 3808.
5. Severence, D.L., Essex, J.W. and Jorgensen, W.L., *J. Comput. Chem.*, 16 (1994) 311.
6. Axelson, P.H. and Li, D., *J. Comput. Chem.*, 19 (1998) 1278.
7. Boresch, S. and Karplus, M., *J. Phys. Chem. A*, 103 (1999) 103.
8. Essex, J.W. and Jorgensen, W.L., *J. Phys. Chem. B*, 101 (1997) 9663.
9. McCarrick, M.A. and Kollman, P.A., *J. Comput. Aid. Mol. Des.*, 13 (1999) 109.
10. Rao, B.G., Kim, E.E. and Murcko, M.A., *J. Comput. Aid. Mol. Des.*, 10 (1996) 23.
11. Reddy M.R., Varnery, M.D., Kalish, V., Viswanadhan, V.N. and Appelt, K., *J. Med. Chem.* 37 (1994) 1145.
12. Lunney, E.A., Para, K.S., Rubin, J.R., Humblet, C., Fergus, J.H., Marks, J.S. and Sawyer, T.K., *J. Am. Chem. Soc.*, 119 (1997) 12471.
13. Brown, M.T. and Cooper, J.A., *Biochim. Biophys. Acta*, 1287 (1996) 121.
14. Verbeek, B.S., Vroom, T.M., Adriaansen-Slot, S.S., Ottenhoff-Kalff, A.E., Geertzema, J.G.N., Hennipman, A. and Rijksen, G., *J. Pathol.*, 180 (1996) 383.
15. Irby, R.B., Mao, W., Coppola, D., Kang, J., Loubeau, J.M., Trudeau, W. Karl, R. Fujita, D.J., Jove, R. and Yeatman, T.J., *Nat. Genet.*, 21 (1999) 187.
16. Mao, W., Irby, R., Coppola, D., Fu, L., Wloch, M., Turner, J., Yu, H., Garcia, R., Jove, R. and Yeatman, T.J., *Oncogene*, 15 (1997) 3083.
17. Lutz, M.P., Eber, I.B.S., Flossman-Kast, B.B.M., Vogelmann, R., Lührs, H., Friess, H., Büchler, M.W. and Adler, G., *Biochem. Bioph. Res. Co.*, 243 (1998) 503.
18. Kumble, S., Omary, M.B., Cartwright, C.A. and Triadafilopoulos, G., *Gastroenterology*, 112 (1997) 348.
19. Cartwright, C.A., Meisler A.I. and Eckhart, W., *Proc. Natl. Acad. Sci. USA*, 87 (1990) 558.
20. Wiener, J.R., Nakano, K., Kruzlock, R.P., Bucana, C.D., Bast, R.C. and Gallick, G.E., *Clin. Cancer Res.*, 5 (1999) 2164.
21. Soriano, P., Montgomery, C., Geske, R. and Bradley, A., *Cell*, 64 (1991) 693.
22. Gilmer, T., Rodriguez, M., Jordan, S., Crosby, R., Alligood, K., Green, M., Kimery, M., Wagner, C., Kinder, D., Charifson, P., Hassell, A. M., Willard, D., Luther, M., Rusnak, D., Sternback, D.D., Mehrotra, M., Peel, M., Shampine, L., Davis, R., Robbins, J., Patel, I.R., Kassel, D., Burkhart, W., Moyer, M., Bradshaw, T. and Berman, J., *J. Biol. Chem.*, 269 (1994) 31711.
23. Stankovic, C.J., Surendran, N., Lunney, E.A., Plummer, M.A., Para, K.S., Shahripour, A., Fergus, J.H., Marks, J.S., Herrera, R., Hubbell, S.E., Humblet, C., Saltiel, A.R., Stewart, B.H. and Sawyer, T.K., *Bioorg. Chem. Med. Lett.*, 7 (1997) 1909.
24. Fu, J. and Castellano, A.L., *Bioorg. Chem. Med. Lett.*, 8 (1998) 2813.
25. Domchek, S.M., Auger, K.R., Chatterjee, S., Burke, T.R. and Shoelson, S.E., *Biochemistry*, 31 (1992) 9865.
26. Pacofsky, G.J., Lackey, K., Alligood, K.J., Berman, J., Charifson, P.S., Crosby, R.M., Dorsey Jr., G.F., Feldmen, P.L., Gilmer, T.M., Hummel, C.W., Jordan, S.R., Mohr, C., Shewchuk, L.M., Sternbach, D.D. and Rodriguez, M., *J. Med. Chem.*, 41 (1998) 1894.
27. Plummer, M.S., Holland, D.R., Shahripour, A., Lunney, E.A., Fergus, J.H., Marks, J.S., McConnell, P., Mueller, W.T. and Sawyer, T.K., *J. Med. Chem.*, 40 (1997) 3719.
28. Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. and Kuriyan, J., *Cell*, 72 (1993) 779.
29. Xu, R.X., Word, J.M., Davis, D.G., Rink, M.J., Willard, D.H.J. and Gampe, R.T.J., *Biochemistry*, 34 (1995) 2107.
30. Charifson, P.S., Shewchuk, L.M., Rocque, W., Hummel, C.W., Jordan, S.R., Mohr, C., Pacofsky, G.J., Peel, M.R., Rodriguez, M., Sternback, D.D. and Consler, T.G., *Biochemistry*, 36 (1997) 6283.
31. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., *J. Am. Chem. Soc.*, 112 (1990) 6127. Some parameters were adjusted for applicability to OPLS-AA (Price, M.L.P., Price, D.J., Jorgensen, W.L. unpublished results).
32. Jorgensen, W.L., BOSS, v.4.0, 4.1, 4.2, Yale Univ., New Haven, 1999.
33. Jorgensen, W.L., In Schleyer, P.v.R. (ed.) *Encyclopedia of Computational Chemistry*, vol. 2. John Wiley, New York, NY, 1998, 1061–1070.
34. Kollman, P.A., *Chem. Rev.*, 93 (1993) 2395.
35. Zwanzig, R.W., *J. Chem. Phys.*, 22 (1954) 1420.
36. Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D., *J. Chem. Phys.*, 79 (1983) 926.
37. Allen, M.P. and Tildesley, D.J., *Computer Simulation of Liquids*, Oxford Science, Oxford, 1987.
38. Essex, J.W., Severence, D.L., Tirado-Rives, J. and Jorgensen, W.L., *J. Phys. Chem. B*, 101 (1997) 9663.
39. Jorgensen, W.L., *J. Phys. Chem.*, 87 (1983) 5304.
40. Jorgensen, W.L., MCPRO, v.1.65, Yale Univ., New Haven, 1999.
41. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Gill, P.M., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T., Petersson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foreman, J.B., Cioslowski, J. and Stefanov, B.B., *Nanayakkar, Gaussian 94, Revision*
42. B.2, Gaussian, Inc., Pittsburgh, PA, 1995.
43. van Gunsteren, W.F., In van Gunsteren, W.F., Weiner, P.K. (eds.) *Computer Simulations of Biomolecular Systems: Theoretical and Experimental Applications*, ESCOM Science, Leiden, 1989, 53.
44. Pearlman, S. and Jorgensen, W.L., unpublished.
45. Morgantini, P. and Kollman, P.A., *J. Am. Chem. Soc.*, 117 (1995) 6057.
46. Ding, Y., Bernardo, D.N., Krogh-Jespersen, K. and Levy, R.M., *J. Phys. Chem.*, 99 (1995) 11575.
47. McCammon, J.A. and Karplus, M., *Biopolymers*, 19 (1980) 1375.
48. Eriksson, M.A. and Nilsson, L., *J. Mol. Biol.*, 253 (1995) 453.
49. Mitchell, M. and McCammon, J.A., *J. Comput. Chem.*, 12 (1991) 271.
50. Taylor, J., *An Introduction to Error Analysis*, University Science Books, Mill Valley, CA, 1982.
51. Ferrin, T.E., Huang, C.C., Jarvis, L.E. and Langridge, R., *J. Mol. Graphics*, 6 (1988) 13.
52. Koradi, R., Billeter, M. and Wüthrich, K., *J. Mol. Graphics*, 14 (1996) 51.