

JMBAvailable online at www.sciencedirect.com ScienceDirect

Predicting Absolute Ligand Binding Free Energies to a Simple Model Site

David L. Mobley^{1†}, Alan P. Graves^{2†}, John D. Chodera²
Andrea C. McReynolds¹, Brian K. Shoichet^{1*} and Ken A. Dill^{1*}

¹*Department of Pharmaceutical Chemistry, University of California at San Francisco San Francisco, CA 94143-2518 USA*

²*Graduate Group in Biophysics University of California at San Francisco, San Francisco CA 94143-2518, USA*

A central challenge in structure-based ligand design is the accurate prediction of binding free energies. Here we apply alchemical free energy calculations in explicit solvent to predict ligand binding in a model cavity in T4 lysozyme. Even in this simple site, there are challenges. We made systematic improvements, beginning with single poses from docking, then including multiple poses, additional protein conformational changes, and using an improved charge model. Computed absolute binding free energies had an RMS error of 1.9 kcal/mol relative to previously determined experimental values. In blind prospective tests, the methods correctly discriminated between several true ligands and decoys in a set of putative binders identified by docking. In these prospective tests, the RMS error in predicted binding free energies relative to those subsequently determined experimentally was only 0.6 kcal/mol. X-ray crystal structures of the new ligands bound in the cavity corresponded closely to predictions from the free energy calculations, but sometimes differed from those predicted by docking. Finally, we examined the impact of holding the protein rigid, as in docking, with a view to learning how approximations made in docking affect accuracy and how they may be improved.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: free energy calculation; docking; alchemical free energy; conformational change; isothermal titration calorimetry

*Corresponding authors

Introduction

A central problem in ligand discovery and design is the prediction of ligand–receptor binding free energies. Current methods cover a spectrum of physical rigor and computational cost. Among physics-based methods, physics-based docking and scoring are computationally the least expensive. In this approach, ligand orientations (poses) are

assigned scores, related to the intermolecular interaction energy, and ranked relative to other poses and other ligands.¹ A few scoring functions include an explicit or implicit estimate of the desolvation free energy of the receptor and ligand.² Receptor flexibility, strain energies,^{3,1} and various entropies are usually neglected, as is any reference to the unbound protein and ligand states. These approximations put the estimation of binding affinities well out of the reach of docking methods, although these methods can often correctly rank-order candidate molecules for testing.

At a higher level of theory are Molecular Mechanics – Generalized Born/Surface Area and Molecular Mechanics – Poisson-Boltzmann/Surface Area (MM-GBSA/PBSA) methods.^{4–6} These methods estimate the absolute free energies of bound and unbound reference states. Enthalpies are estimated using average energies from a molecular mechanics force field, and combined with an entropy estimate and a solvation free energy from an implicit solvent model. The difficulty is that the binding free energy

† D.L.M. and A.P.G. authors contributed equally to this work.

Present address: J. D. Chodera, Department of Chemistry, Stanford University, Stanford, CA 94305-5080, USA.

Abbreviation used: ITC, isothermal titration calorimetry; MM-GBSA/PBSA, Molecular Mechanics – Generalized Born/Surface Area and Molecular Mechanics – Poisson-Boltzmann/Surface Area.

E-mail addresses of the corresponding authors: shoichet@cgl.ucsf.edu; dill@maxwell.compbio.ucsf.edu

is a small difference between very large absolute free energies, requiring either very high accuracy in computing these large numbers or cancellation of errors. Thus while these approaches have had successes,^{5,6} they also have several drawbacks, such as sensitivity to details of the implicit solvent model used^{7,8} and to the method used for estimating the entropy term. Such methods perform poorly on some test sets.^{9,10}

At the highest level of rigor are various free energy methods, including the alchemical free energy calculations described below and potential of mean force-based methods^{11,12} (for recent reviews of free energy methods, see works by Rodinger and Pomès, Kofke, and Shirts *et al.*).^{13–15} Here we focus on alchemical free energy methods, which evaluate ratios of partition functions to estimate binding free energies and thus include entropic and other contributions neglected at lower levels of theory. These methods, combined with some theoretical developments first laid out in the mid-1990s^{16–18} and refined later,¹⁹ now allow absolute binding free energies to be computed rigorously and exactly, provided that the molecular mechanics force field used accurately describes the underlying physics, and that enough sampling can be performed that the estimates of the relevant thermodynamic averages are converged.²⁰

If, in principle, alchemical free energy calculations allow for the exact prediction of binding free energies, the requirements of accurate force fields and adequate sampling introduce an error into the computed free energies. This error is often difficult to isolate in the complicated environment of protein active sites. In such sites, failures of sampling or force fields are exacerbated by binding-induced conformational changes, titratable groups, metal ions, and ordered water molecules, among other complications. Furthermore, when sampling is inadequate, alchemical free energy methods can easily give biased results; for example, computed free energies are often sensitive to the choice of the initial receptor or ligand structure.^{20–25,58}

Here, to isolate sources of error, we study a highly simplified binding site using alchemical free energy methods and molecular dynamics. We focus on the binding of small aromatic ligands to the small, buried hydrophobic binding site in an engineered mutant of T4 lysozyme (the L99A site; **Figure 1**) that has been studied extensively experimentally,^{26–32} with docking methods,^{30,32} and in some previous computational free energy studies.^{18,19,23,33} Here we systematically evaluate the effect of various approximations on computed binding free energies. This model binding site provides a good starting point because it is simple and has been thoroughly characterized experimentally.

A second advantage of this model binding site is that it also provides an excellent opportunity for prospective predictions, since it is relatively easy to find new compounds that bind.³⁰ This is valuable because it can be far easier to suggest explanations for previous observations than to actually make new

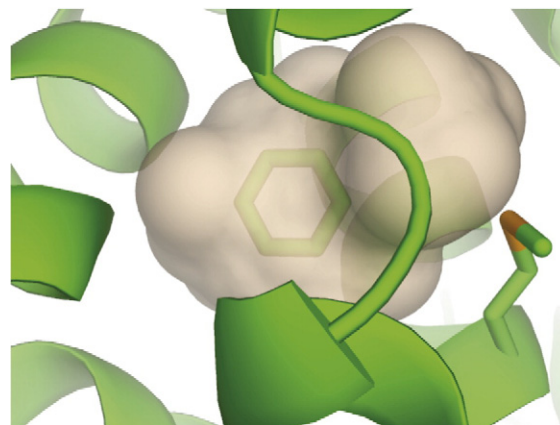


Figure 1. The model hydrophobic binding site in the L99A mutant of T4 lysozyme. The enclosed molecular surface of the cavity is shown (brown) as is the crystallographic geometry of a bound benzene ligand (green), within the context of the overall structure of T4 lysozyme (green ribbons). The side-chain of Met102 is also shown for reference.

predictions, and predictive ability provides a fundamental test for methods.

Results

Overview

Here we performed two sets of studies: retrospective, in which we studied binding of ligands with previously measured affinities, and prospective, in which we predicted, in a blind test, the binding modes and affinities of several previously uncharacterized small molecules. After making predictions, we tested them experimentally, using isothermal titration calorimetry (ITC) to measure affinities and X-ray crystallography to determine structures of the complexes.

Retrospective studies: comparison with previous experimental results

We first computed binding free energies for a test set of 13 small neutral compounds using alchemical free energy calculations, as described in Methods (**Table 1**). Of these, binding affinities for 11 had previously been measured by ITC,²⁹ and two had previously been determined not to bind more strongly than an affinity of 10 mM using a thermal denaturation assay.^{28,32}

Binding affinities are underestimated from single docking poses

We started with a simple approach. We used the best-scoring docking pose for each compound as a starting structure from which to simply compute binding free energies using our standard free energy calculation protocol discussed in Methods. We

Table 1. Calculated and experimental binding free energies for ligands of the apolar binding site considered here

Molecule	$\Delta G_{\text{exp}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{single}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{multiple}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{calc}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{calc}}^{\circ} - \Delta G_{\text{exp}}^{\circ}$ (kcal/mol)
2,3-Benzofuran	-5.46 ± 0.03	-2.77 ± 0.08	-3.45 ± 0.06	-3.53 ± 0.06	1.93 ± 0.07
Benzene	-5.19 ± 0.16	-3.05 ± 0.20	-4.53 ± 0.20	-4.56 ± 0.20	0.63 ± 0.26
Ethylbenzene	-5.76 ± 0.07	-4.95 ± 0.10	-5.36 ± 0.10	-6.36 ± 0.18	-0.60 ± 0.19
Indene	-5.13 ± 0.01	-0.63 ± 0.11	-1.56 ± 0.06	-1.75 ± 0.07	3.38 ± 0.07
Indole	-4.89 ± 0.06	0.06 ± 0.10	-0.24 ± 0.07	-0.42 ± 0.08	4.47 ± 0.10
Isobutylbenzene	-6.51 ± 0.06	-0.16 ± 0.15	-4.14 ± 0.12	-5.01 ± 0.20	1.50 ± 0.21
<i>n</i> -Butylbenzene	-6.70 ± 0.02	-4.03 ± 0.11	-4.44 ± 0.11	-4.87 ± 0.14	1.83 ± 0.14
<i>n</i> -Propylbenzene	-6.55 ± 0.02	-5.29 ± 0.10	-5.70 ± 0.10	-5.88 ± 0.11	0.67 ± 0.12
<i>o</i> -Xylene	-4.60 ± 0.06	-0.15 ± 0.10	-0.56 ± 0.10	-1.27 ± 0.18	3.33 ± 0.19
<i>p</i> -Xylene	-4.67 ± 0.06	-2.13 ± 0.09	-2.96 ± 0.09	-3.54 ± 0.17	1.13 ± 0.18
Toluene	-5.52 ± 0.06	-3.76 ± 0.09	-4.17 ± 0.09	-4.58 ± 0.12	0.94 ± 0.14
Phenol	> -2.74	-0.86 ± 0.09	-1.27 ± 0.09	-1.26 ± 0.09	N/A
2-Fluorobenzaldehyde	> -2.74	0.99 ± 0.25	-2.43 ± 0.10	-2.92 ± 0.14	N/A
<i>Statistics</i>					
RMS error		3.51 ± 0.04	2.55 ± 0.03	2.24 ± 0.04	
Correlation, <i>R</i>		0.51 ± 0.05	0.72 ± 0.05	0.72 ± 0.05	

Experimental values (denoted by $\Delta G_{\text{exp}}^{\circ}$) are according to Morton *et al.*,²⁹ except for 2-fluorobenzaldehyde and phenol, where no binding was observed (by a circular dichroism ΔT_m upshift assay^{28,32}), giving only a lower bound on the binding free energy. Calculated values shown are $\Delta G_{\text{single}}^{\circ}$, the free energy computed using only the single best-scoring docking orientation; $\Delta G_{\text{multiple}}^{\circ}$, the full computed binding free energy using all orientations considered; and $\Delta G_{\text{calc}}^{\circ}$, the computed binding free energy including multiple orientations and using the confine-and-release approach to account for protein conformational change at Val111. The final column is the difference from the experiment. At the bottom is the RMS error relative to the experiment across binders for each set of free energies, and the correlation coefficient, *R*, between calculated and experimental values. Experimental binding affinities were measured at 302 K; binding free energies were computed at 300 K. Calculated values used AM1-CM2 charges.

previously found that this single-pose approach often results in ligands remaining trapped in the vicinity of their starting orientation on simulation timescales.²⁰ Thus, with this approach, the free energy calculations effectively become an expensive rescoring of docking poses, including conformational averaging and entropic effects, but only for a single orientation. We present results using this approach as $\Delta G_{\text{single}}^{\circ}$ (Table 1). In keeping with the approximations typically made in docking, we consider only single orientations for these results, meaning that we also neglect symmetry-equivalent orientations for molecules like toluene, phenol, and benzene, which in reality also contribute to binding.^{17,19,20,23} The RMS error for $\Delta G_{\text{single}}^{\circ}$ relative to the experiment is 3.51 ± 0.04 kcal/mol, and the correlation coefficient (*R*) between computed free energies and experiment is 0.51 ± 0.05 . (These RMS and correlation calculations do not include the non-binders, since free energies of association for these are not known.) (Figure 2(a)). This approach underestimates all the binding affinities. This is likely due to undersampling, a failure to adequately sample the most optimal binding conformations. Previous work had suggested that this approach would fail in the case where the best docking pose is not the orientation that actually contributes most to binding,²⁰ so this outcome was expected.

Accounting for multiple potential bound orientations reduces the error in computed binding affinities

Next, we account for the presence of multiple potential ligand binding modes separated by kinetic barriers. We compute binding free energies of different possible binding modes separately and com-

bine their contributions to get an overall binding free energy²⁰ (see also Simulation methods). We refer to these free energies, which also include the contributions of orientations related by symmetry, as $\Delta G_{\text{multiple}}^{\circ}$ (Table 1). With this approach, the computed binding free energies are substantially closer to the experiment in several cases (and are never worse) than those computed using the single orientation approach, since occasionally the best-scoring pose is not the pose that contributes most to the binding free energy. This inclusion of these contributions reduces the RMS error in the computed free energies relative to the experiment, from 3.51 ± 0.04 kcal/mol to 2.55 ± 0.03 kcal/mol, and raises the correlation coefficient, *R*, from 0.51 ± 0.05 to 0.72 ± 0.05 .

The improvements with this approach come for several reasons. For three of the ligands (indene, indole, and 2,3-benzofuran), multiple orientations are within *kT* of one another and all contribute substantially. For isobutylbenzene, the best pose from docking is not in the orientation that contributes most to binding, so including multiple candidate orientations results in inclusion of the dominant orientation. For the remainder of the compounds, improvements come from inclusion of symmetry number corrections. These issues have been addressed in more detail in a work on a related binding site.²⁰ In general, it is extremely difficult to predict in advance whether multiple orientations may be relevant.

Accounting for more protein conformational change further improves computed binding free energies

The section above describes our treatment of relevant ligand orientations. However, the protein may also have relevant slow degrees of freedom

which can be difficult to sample.^{28,34} Here a key change is the reorientation of the Val111 side-chain observed in X-ray structures in response to ligands such as *n*-butylbenzene, isobutylbenzene, *o*-xylene, and *p*-xylene (Figure 3).²⁸ The energy barriers associated with this reorientation are large enough to prevent the side-chain from rotating on simulation timescales.³⁴ This leads to an apparent dependence of computed free energies on the initial protein structure used in simulations. For example, binding free energies that are computed from the *holo* protein structure are too negative (favorable) if the side-chain does not have time to reorient as the ligand is removed, because the protein strain energy (the energetic cost of deforming the protein on binding) is not properly accounted for. On the other hand, if the *apo* protein structure is used, as we did here, binding free energies are too positive (unfavorable), as the ligand sterically clashes to some degree with the protein.³⁴ This dependence on the starting structure

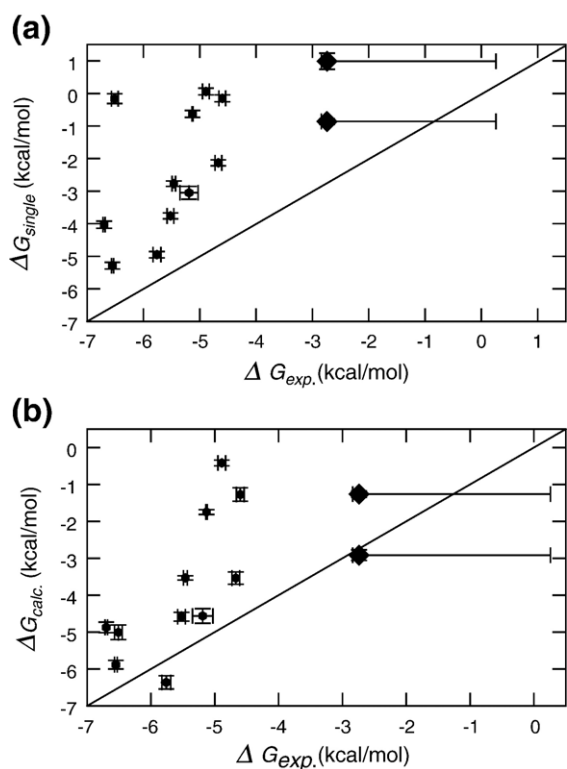


Figure 2. Calculated binding free energies compared with the experiment. Calculated and experimental binding free energies are shown with error bars; the calculated error bars represent one standard deviation. The two points shown as larger diamonds are the nonligands phenol and 2-fluorobenzaldehyde; for these, only a lower limit on the experimental binding free energy is known, as denoted by a large experimental error bar to the right. The diagonal $x=y$ line denotes perfect agreement with the experiment. (a) Calculated $\Delta G_{\text{single}}^{\circ}$, single-orientation binding free energies, including only the contribution from the single best docking orientation. (b) $\Delta G_{\text{calc}}^{\circ}$, binding free energies, including all relevant ligand orientations and contributions from releasing Val111 from its kinetic confinement.

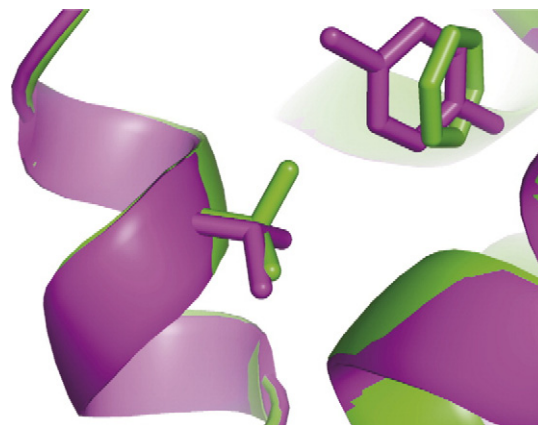


Figure 3. Val111 reorients on ligand binding. Val111 is observed to adopt a different side-chain rotamer from the *apo* crystallographic structure in cocrystal structures with several different ligands. Shown here is the benzene-bound structure, green, which is virtually identical with the *apo* structure of the protein. Also shown is the *p*-xylene-bound structure in magenta. The sticks at the left show the reorientation of the Val111 side-chain on binding to *p*-xylene by roughly 120° relative to the benzene-bound and *apo* structures.

is simply due to kinetic trapping of the protein in conformations near its starting conformation.

To overcome the kinetic trapping of Val111, we use a recently developed “confine-and-release” framework to obtain correct binding free energies that are independent of the starting structure.³⁴ Specifically, when the Val111 remains trapped, computed binding free energies are really “confined” binding free energies, with Val111 confined to a particular orientation, so we use umbrella sampling to compute the free energy of releasing the valine from its confinement in the bound and unbound states. Here this is accomplished by forcing the sampling of alternative orientations using a harmonic biasing potential, and recovering the free energy landscape for this degree of freedom.³⁴ We do this for all of the compounds considered here, although for many compounds, only the *apo* orientation of the valine is found to be relevant, as observed experimentally.²⁸ This is a rigorous way to account for kinetic trapping. The confine-and-release framework is a generalization (to protein degrees of freedom) of the biasing potential approaches applied previously to ligands in a number of studies.^{11,16,18–20,23,58}

The confine-and-release approach, which yields the total estimated binding free energy ΔG_{bind} , further improves the agreement of computed binding free energies with the experiment (Table 1 and Figure 2(b)). With this approach, the RMS error relative to the experiment further decreases from 2.55 ± 0.03 kcal/mol to 2.24 ± 0.04 kcal/mol, while the correlation with the experiment remains unchanged ($R=0.72 \pm 0.05$). There is significant improvement in the agreement with the experiment for a number of the ligands, especially isobutylbenzene, *p*-xylene, *n*-butylbenzene, *o*-xylene, and ethylbenzene. As mentioned above, for the first four of

these, Val111 reorientation on binding is observed in the cocrystal structure.²⁸ For ethylbenzene, the deposited structure does not show the Val111 rotated relative to the benzene-bound structure, but the electron density seems to allow the possibility of either orientation.²⁸ A prediction of this work is that, if the crystal structure with ethylbenzene can be solved at a higher resolution, then the Val111 side-chain should be observed to adopt a conformation similar to that in the *p*-xylene-bound structure.

The confine-and-release approach can be applied to a variety of protein conformational changes. Here we chose to apply it specifically to a single Val111 dihedral. This choice was motivated by the fact that Val111 was previously observed (experimentally) to reorient on ligand binding.²⁸ Also, previous computational work led us to believe that sampling of this degree of freedom could be very slow.²³ We therefore examined our initial simulations to look for Val111 reorientation and found the kinetic trapping described above.³⁴ This led us to apply the confine-and-release approach to this particular degree of freedom.

It is worth noting that this is not simply an issue of predicting a correct bound structure. Rather, using either the *apo* or *holo* structure leads to biased binding free energies (Tables 1 and 2) when the confine-and-release approach is not used. Only when we account for Val111 reorientation using confine-and-release do computed free energies become consistent between simulations beginning from *apo* and *holo* starting structures (see below).³⁴

Binding free energies from *holo* structures agree with those from *apo* after accounting for Val111 reorientation

After the confine-and-release calculations, there was still particularly poor agreement with the experiment for several ligands, especially *o*-xylene, indene, indole, isobutylbenzene, and 2,3-benzofuran (the first three of these are the only binders in Figure 2(b) with computed binding free energies worse than -2 kcal/mol). One possible explanation is inadequate sampling, perhaps due to additional protein conformational rearrangements that are not being sampled. For example, for indene, isobutylbenzene, and *o*-xylene, helix F, which forms one side of the

Table 2. Binding free energies calculated for selected ligands using their *holo* structures as a starting point

Molecule	$\Delta G_{\text{multiple}}^{\text{apo,holo}}$ (kcal/mol)	$\Delta G_{\text{calc}}^{\text{apo,holo}}$ (kcal/mol)	$\Delta G_{\text{calc}}^{\text{apo,holo}} - \Delta G_{\text{calc}}^{\text{apo}}$ (kcal/mol)
Indene	-1.44 ± 0.07	-1.64 ± 0.09	0.10 ± 0.11
<i>n</i> -Butylbenzene	-9.17 ± 0.13	-5.32 ± 0.22	-0.45 ± 0.26
Isobutylbenzene	-8.98 ± 0.13	-4.80 ± 0.21	0.20 ± 0.29
<i>p</i> -Xylene	-7.27 ± 0.09	-3.31 ± 0.20	0.23 ± 0.26
<i>o</i> -Xylene	-5.93 ± 0.12	-1.91 ± 0.21	-0.64 ± 0.28

Shown are the calculated binding free energies, beginning from the *holo* structures, with and without including any Val111 reorientation, and the difference from the calculated binding free energies using the *apo* structure as a starting point (including the Val111 reorientation) in Table 1.

cavity, shifts around 2 \AA on ligand binding, making the binding site larger.²⁸ Additionally, previous free energy calculations on the same system tended to overestimate binding free energies for some of these same compounds when beginning from the *holo* structures.²³

However, inspection of simulation trajectories suggests that this helix motion is being sampled. As a more quantitative test, we repeated the calculations for selected compounds beginning from the *holo* structures. If computed binding free energies are different starting from *apo* and *holo* structures, even after applying the confine-and-release approach for Val111, it would indicate inadequate sampling. While computed binding free energies are significantly different for calculations starting from the *apo* and *holo* structure before accounting for Val111 reorientation, the differences are essentially negligible (within uncertainty) when the confine-and-release approach is used to account for this change (Table 2) (the largest difference, using the confine-and-release approach, is for *o*-xylene, -0.64 ± 0.28 kcal/mol; since the uncertainty represents one standard deviation, this is still only a 2σ variation). This implies that the sampling of these conformational changes is probably sufficient and that the error lies elsewhere.

Free energies computed using the *holo* starting structures also show that the *holo* protein structure of several of these ligands is unfavorable by roughly 4 kcal/mol in the absence of bound ligand (Table 2).³⁴ This is presumably because of steric clashes within the protein and is the reason why only some of the ligands induce this conformational change on binding.

The AM1-BCC charge model further increases the accuracy of binding free energies

Next, we considered another possible source of error: the simulation parameters. There are different methods for assigning partial charges for small molecules.³⁵ In the work reported above, we used Austin Model 1 - Charge Model 2 (AM1-CM2)³⁶ partial atomic charges for the small molecules, as in docking studies. However, we found previously that Austin Model 1 - Bond Charge Correction (AM1-BCC) charges performed better than AM1-CM2 charges for hydration free energies, perhaps because they are more similar to the Hartree-Fock 6/31G* charges the force field was parameterized with.³⁵ Therefore, we tested the AM1-BCC charges here as well. Table 3 shows that AM1-BCC charges further reduce the RMS error between computed and experimental binding free energies from 2.24 ± 0.04 to 1.89 ± 0.04 kcal/mol, and the correlation coefficient, *R*, increases from 0.72 ± 0.05 to 0.79 ± 0.07 .

Alchemical methods are more accurate than docking

One major challenge for docking methods is to discriminate between binders and nonbinders. We

Table 3. Calculated and experimental binding free energies for ligands of the apolar binding site considered here, as in Table 1 except using AM1-BCC charges

Molecule	$\Delta G_{\text{exp}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{BCC}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{BCC}}^{\circ} - \Delta G_{\text{exp}}^{\circ}$ (kcal/mol)	$\Delta G_{\text{AM1-CM2}}^{\circ} - \Delta G_{\text{BCC}}^{\circ}$ (kcal/mol)
2,3-Benzofuran	-5.46 ± 0.03	-3.66 ± 0.06	1.80 ± 0.06	0.13 ± 0.08
Benzene	-5.19 ± 0.16	-3.95 ± 0.20	1.24 ± 0.26	-0.61 ± 0.28
Ethylbenzene	-5.76 ± 0.07	-5.82 ± 0.14	-0.06 ± 0.16	-0.54 ± 0.23
Indene	-5.13 ± 0.01	-1.63 ± 0.07	3.50 ± 0.07	-0.12 ± 0.09
Indole	-4.89 ± 0.06	-1.37 ± 0.10	3.52 ± 0.12	0.96 ± 0.13
Isobutylbenzene	-6.51 ± 0.06	-8.09 ± 0.18	-1.58 ± 0.19	3.09 ± 0.27
<i>n</i> -Butylbenzene	-6.70 ± 0.02	-5.70 ± 0.20	1.00 ± 0.20	0.83 ± 0.25
<i>n</i> -Propylbenzene	-6.55 ± 0.02	-5.44 ± 0.10	1.11 ± 0.11	-0.44 ± 0.16
<i>o</i> -Xylene	-4.60 ± 0.06	-3.23 ± 0.25	1.37 ± 0.25	1.96 ± 0.31
<i>p</i> -Xylene	-4.67 ± 0.06	-3.59 ± 0.12	1.08 ± 0.14	0.05 ± 0.21
Toluene	-5.52 ± 0.06	-4.07 ± 0.10	1.45 ± 0.12	-0.51 ± 0.16
Phenol	> -2.74	-1.07 ± 0.20	N/A	-0.19 ± 0.22
2-Fluorobenzaldehyde	> -2.74	-3.14 ± 0.13	N/A	0.22 ± 0.19
<i>Statistics</i>				
RMS error		1.89 ± 0.04		
Correlation, <i>R</i>		0.79 ± 0.07		

Shown are $\Delta G_{\text{BCC}}^{\circ}$ -full binding free energies done with AM1-BCC including contributions from multiple ligand orientations and any Val111 reorientation. These are equivalent to $\Delta G_{\text{calc}}^{\circ}$ from that table but done with AM1-BCC charges. Also shown are the differences between the AM1-BCC results and the experiment (next to last column), and between the AM1-BCC and AM1-CM2 results (last column). When the values in the last two columns have the same sign, AM1-BCC charges improved the agreement with the experiment. At the bottom is RMS error relative to the experiment across binders for each set of free energies, and the correlation coefficient, *R*, between calculated and experimental values.

have included two known nonbinders (with affinities worse than 10 mM) in the set of molecules examined here: phenol and 2-fluorobenzaldehyde. For these two compounds, computed binding free energies indicate only weak affinity: -2.9 ± 0.1 kcal/mol and weaker (more positive) (Table 1). A 10-mM detection threshold in binding affinity corresponds to a binding free energy of roughly -2.7 kcal/mol. Thus the computed binding free energies for these two compounds are at the detection limit, essentially consistent with the experimental observation that they are nonbinders.

Our free energy calculations are computationally expensive. Are the results any more accurate than those that can be obtained from molecular docking? As shown in Figure 4, DOCK scores for the ligands

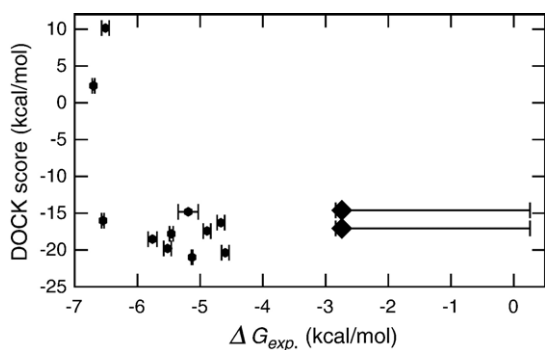


Figure 4. DOCK scores for the best-ranked pose for each molecule *versus* experimental binding free energies. The correlation coefficient (*R*) is -0.69 , meaning that compounds that DOCK predicts should bind strongly tend to bind weakly. Additionally, the two nonbinders have similar DOCK scores to a number of the binders.

studied here correlate poorly with experimental binding free energies. In fact they are anticorrelated ($R = -0.69$), the opposite of what one would like. Moreover, the two nonbinders have DOCK scores similar to those for the majority of the true ligands (and much more favorable scores than several ligands), hence it is impossible to discriminate between binders and nonbinders. In fairness to docking, it is worth noting that these nonbinders were included in our test set because they have proven challenging for docking to discriminate from the binders. Also, the first goal of docking is to separate likely from unlikely ligands, and it does seem to be performing remarkably well in this binding site, where about 80% of the top 100 docking hits would probably bind. Additionally, we find that docking also performs quite well in this site at generating sterically reasonable potential bound orientations. That said, the free energy calculations give substantially better affinity estimates and correlations than docking does, and are better at recognizing nonbinders.

The docking results discussed here used the benzene-bound protein structure, which is virtually identical with the *apo* structure. However, docking to alternate protein conformations seems not to result in significant improvements in the quality of docking results, except when many different crystallographic protein conformations are used.³¹

A large source of error in docking is the rigid-protein approximation

Docking typically treats proteins as rigid. How big is the error introduced by this assumption? To test this, we held the protein rigid and repeated our free

energy calculations, including the effects of ligand symmetries and multiple ligand orientations (and using AM1-CM2 charges). This led to essentially zero correlation ($R = -0.05 \pm 0.09$) between computed free energies and experimental values, and an RMS error of 19.78 ± 0.06 kcal/mol (Figure 5(a)).

As a simple improvement on this rigid-protein approximation, we also allowed the protein to relax to a different structure for each ligand. First, we minimized the entire protein in the presence of each ligand, using the initial docking geometry, and subsequently held the protein rigid during the simulations. This resulted in a correlation (R) of 0.82 ± 0.09 and an RMS error of 4.92 ± 0.07 kcal/mol relative to the experiment (Figure 5(b)). Second, we minimized only a region of the protein around the binding site (Binding free energies to a rigid protein) in the presence of each ligand, before holding the protein fixed during the simulations. This resulted in a correlation (R) of 0.32 ± 0.08 and an RMS error of 4.06 ± 0.06 kcal/mol relative to the experiment (Figure 5(c)).

Overall, it appears that keeping the protein rigid while estimating binding free energies is detrimental to binding free energy estimation, even if minimization is performed separately for each ligand.

Prospective studies: predictions and experimental tests

Distinguishing binders from nonbinders

We also performed a blind test of these free energy methods. We selected five small molecules that were among the top-scoring molecules from a docking screen of a compound library (using protocols described previously).³² We calculated binding free energies in the same manner as above, and compared the resulting dissociation constants with the detection threshold of 10 mM for the experimental thermal denaturation assay. We predicted that four of the molecules would bind and one would not[‡] (the 10 mM threshold fell just between affinities for two of the compounds). We then tested these predictions experimentally using the upshift in thermal denaturation, in which the melting temperatures of the protein in the presence and absence of the ligand were compared.²⁶ All molecules were tested in their neutral forms, using either circular dichroism or fluorescence to monitor the transition from the folded to the unfolded state, and resulting T_m values were compared to that of the *apo* protein. All melts occurred reversibly in a manner consistent with two-state unfolding. 1,2-Dichlorobenzene, *n*-methylaniline, 1-methylpyrrole, and 1,2-benzenedithiol increased the T_m significantly, between 1.0 and 2.9 °C (Table 4). Conversely, thieno[2,3-*c*]pyridine was not observed to increase the T_m , even at a 2.5-mM concentration, consistent with the predic-

[‡] Initial predictions were made with AM1-CM2 charges, but AM1-BCC charges were tried later and led to the same outcome.

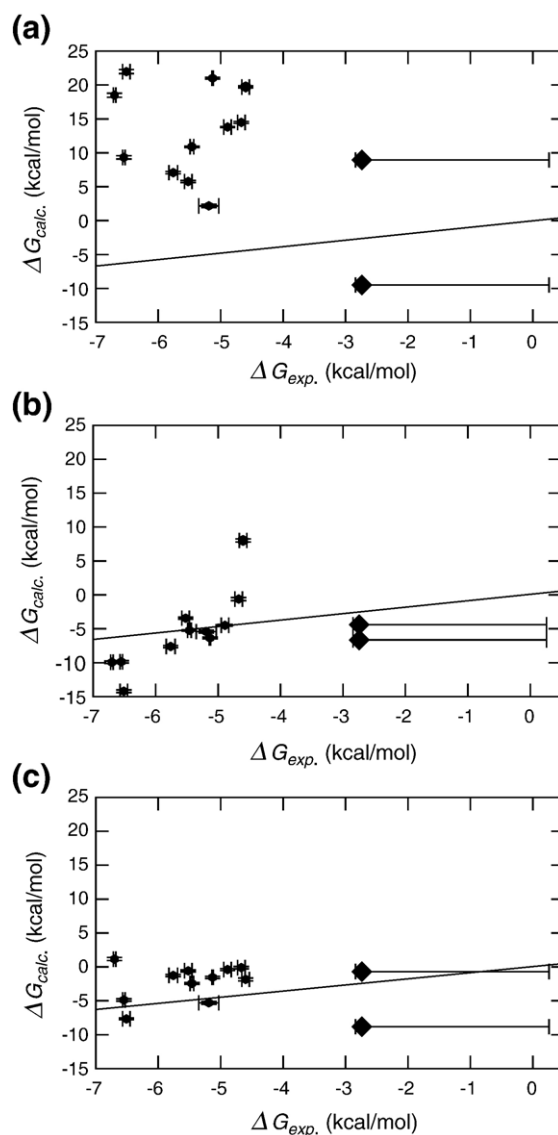


Figure 5. Comparison of calculated and experimental binding free energies with the protein held rigid. (a) Binding free energies with the protein completely rigid. The RMS error relative to the experiment is 19.78 ± 0.06 kcal/mol and the correlation coefficient (R) is -0.05 ± 0.09 . (b) Binding free energies with the whole protein minimized separately for each ligand. The RMS error relative to the experiment is 4.92 ± 0.07 kcal/mol and the correlation coefficient (R) is 0.82 ± 0.09 . (c) Binding free energies with only the binding site minimized for each ligand. The RMS error relative to the experiment is 4.06 ± 0.06 kcal/mol and the correlation coefficient (R) is 0.32 ± 0.08 . The $x=y$ indicates perfect agreement with the experiment.

tions of the free energy calculations. In contrast, docking had predicted that all five would bind, but the free energy methods correctly identified the nonbinder (thieno[2,3-*c*]pyridine) (Table 4; Figure 6).

Predicting bound orientations was successful

We then obtained crystal structures (Table 5) to determine how well these free energy calculations

Table 4. Novel compounds for which predictions were made and later tested experimentally

Molecule	DOCK score (kcal/mol)	Alchemical prediction ^a	ΔT_m (°C)	Experiment (kcal/mol)	$\Delta G_{\text{calc}}^{\circ}$ ^b	$\Delta G_{\text{exp}}^{\circ}$ ^c (kcal/mol)
1,2-Dichlorobenzene	-19.99	Binder	2.90	Binder	-5.66±0.15	-6.37
<i>n</i> -Methylaniline	-17.29	Binder	1.00	Binder	-5.37±0.11	-4.7
1-Methylpyrrole	-15.27	Binder	2.20	Binder	-4.32±0.08	-4.44
1,2-Benzenedithiol	-18.51	Binder	2.50	Binder	-2.79±0.13	N.D.
Thieno[2,3- <i>c</i>]pyridine	-18.81	Nonbinder	-0.40	Nonbinder	-2.56±0.07	N.D.

DOCK scores suggested that all five should bind. Alchemical free energy calculations were initially used to predict whether or not these molecules would bind, then ΔT_m values were determined experimentally to test these predictions. Following this, final binding free energy predictions ($\Delta G_{\text{calc}}^{\circ}$) were tested experimentally with ITC; results are as shown ($\Delta G_{\text{exp}}^{\circ}$). The RMS difference between predicted ΔG° and experiment for the three compounds tested with ITC is 0.57 kcal/mol. N.D. indicates that the measurements were not done.

^a Initial predictions were made using AM1-CM2 charges, but the outcome was unchanged with AM1-BCC charges.

^b Before doing ITC, predictions were refined using AM1-BCC charges, which testing had indicated gave higher accuracy.

^c Calorimetry was done at 283 K.

could predict the bound ligand conformations. We soaked three of the ligands into L99A lysozyme protein crystals. The crystals diffracted between 1.7 and 2.07 Å on a home source. In all three structures, initial $F_o - F_c$ electron density unambiguously identified the orientation of the ligand in the site; for dichlorobenzene, two orientations of the ligand were apparent.

In parallel, we predicted dominant bound orientations for each of these ligands (Simulation methods). Then we determined the structures for these three ligands, and compared with our predicted structures, the best predicted DOCK poses, and the electron density (Figure 7).

For 1-methylpyrrole, the X-ray, docking, and molecular dynamics poses were quite similar (Figure 7). The particular molecular dynamics snapshot that was selected as representative appears slightly twisted relative to the other two structures, but this is simply due to the arbitrary selection of a single molecular dynamics conformation from an ensemble. The X-ray pose falls well within the range of structures sampled by the simulation from which this snapshot was chosen (and the underlying electron density is consistent with a range of structures seen in simulation). The RMSD between the docking pose and X-ray structure is 0.39 Å and that

between the free energy snapshot and X-ray is 0.94 Å.

For *n*-methylaniline, free energy methods predicted two orientations with essentially equal occupancy probabilities (Figure 7). The X-ray and docking poses match well with one of these orientations, but there is no evidence in the electron density for the second orientation. The two orientations differ only by rotation around the C1-N bond. The RMSD between the docking pose and X-ray is 0.63 Å, and that between the lower RMSD free energy snapshot and X-ray is 0.69 Å (the higher RMSD orientation is 1.29 Å away from X-ray).

For 1,2-dichlorobenzene, two separate orientations were observed in the crystal structure, differing by a rotation of around 60° in the plane of the aromatic ring. DOCK failed to properly identify either of these orientations as the best-scoring pose, whereas our free energy methods picked out one but indicated that the second was energetically unfavorable. In particular, the ligand would occasionally transition into this alternate orientation, but it would typically remain only transiently. We further tested this by conducting a separate set of calculations where the ligand was restrained to remain in the alternate orientation, but, with our parameter set, it appears substantially less favorable for binding than the orientation predicted to dominate using our free energy methods. This could be a force field failure, inadequate sampling, or a difference between experimental and simulation conditions. The docking pose is 2.8–2.9 Å away from the X-ray orientations, while the free energy snapshot is only 0.77 Å away from the most similar X-ray orientation.

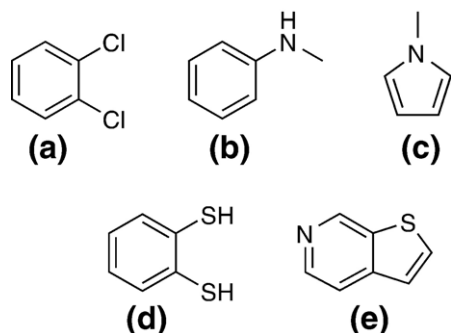


Figure 6. Five compounds for which binding predictions were made. (a) 1,2-Dichlorobenzene; (b) *n*-methylaniline; (c) 1-methylpyrrole; (d) 1,2-benzenedithiol; and (e) thieno[2,3-*c*]pyridine.

Predicting binding affinities

Next, we predicted binding free energies and then measured them by ITC. Since the AM1-BCC charge model worked best in retrospective studies, we used these charges for binding free energy predictions. Ligand titrations gave easily modeled curves using a low *c*-value protocol (Figure 8).³⁷ Not only did binding free energy calculations give the correct rank ordering of binders, but computed binding free energies

Table 5. X-ray data collection and refinement

	Ligands bound to L99A		
	1,2-Dichlorobenzene	<i>n</i> -Methylaniline	1-Methylpyrrole
Cell dimensions			
<i>a</i> = <i>b</i> (Å)	60.2	59.9	60.2
<i>c</i> (Å)	97.0	96.1	96.4
Resolution (Å)	1.70 (1.76) ^a	2.07 (2.14) ^a	1.94 (2.01) ^a
Reflections	18,469 (2263) ^a	12,102 (1169) ^a	15,355 (1484) ^a
<i>R</i> _{merge} (%)	9.8 (64.5) ^a	13.5 (52.8) ^a	11.2 (56.9) ^a
Completeness (%)	99.8 (99.9) ^a	95.0 (92.8) ^a	99.4 (98.4) ^a
<i>I</i> / <i>σ</i> (<i>I</i>)	9.8 (2.3) ^a	7.4 (2.3) ^a	8.7 (2.1) ^a
<i>R</i> -factor (%)	19.8	21.9	19.1
<i>R</i> _{free} (%)	23.5	23.4	25.8
Resolution range (Å)	50.0–1.70	50.0–2.07	50.0–1.94
Δ _{bondlengths} (Å)	0.008	0.007	0.009
Δ _{bondangles} (°)	1.004	0.916	1.074
PDB Code	20TY	20T2	20U0

^a Values in parentheses are for the highest resolution shell.

for these compounds were also remarkably accurate (an RMS error of 0.57 ± 0.09 kcal/mol, Table 4).

Discussion

Accuracy of free energy calculations in retrospective and prospective studies

Alchemical free energy calculations using molecular dynamics can be used to compute fairly accurate binding free energies of ligands in the T4 lysozyme L99A binding site, with an RMS error of 1.89 ± 0.04 kcal/mol in retrospective tests. This is a much higher accuracy than docking, where scores were inversely correlated with the experiment, at least among the top-scoring ligands here. Admittedly, the docking program, DOCK, was never designed to predict binding affinities and performs remarkably well at ranking likely ligands highly in large libraries.³² Also, in these calculations, we are comparing with previously known results. A more rigorous test is to compare genuinely new predictions on untested candidate ligands with a subsequent experiment.

Therefore, in a blind test, we predicted affinities and binding orientations for five previously uncharacterized compounds predicted by DOCK to bind, then tested these predictions experimentally. With alchemical free energy calculations, we correctly recognized the one nonbinder, accurately predicted ligand-bound orientations, and quantitatively predicted binding free energies. In each of these areas, free energy methods agreed better with the experiment than did docking. Free energy methods, unlike docking, also correctly ranked the ligand binding affinities in these prospective tests. Thus it appears that alchemical free energy methods can be truly predictive and can rescue docking failures.

The approach described here, including the retrospective study of previously published data, required no knowledge of the bound structure of the protein and ligand, and used the *apo* protein struc-

ture. Previous work on this same binding site has required an accurate bound structure of the complex of interest as a starting point.²³

Essential ingredients for accuracy

This present study is limited to a simplified model binding site, where many complications of other binding sites are absent. The L99A cavity studied here has no interface with bulk water, no ordered water molecules to displace, and is small, and the dominant interactions appear to be mainly nonpolar. Nevertheless, the use of this system has allowed us to be systematic in isolating and solving various sampling problems. We identified several keys to obtaining accurate binding free energies: first, multiple potential ligand orientations must be included; one cannot rely on the single top docking pose to be the dominant ligand orientation. There can be large energetic barriers between different ligand orientations, which make timescales for orientational change long compared to simulation times. Second, even seemingly small protein conformational changes on ligand binding, such as the reorientation of a single side-chain, Val111, can be difficult to sample correctly in free energy calculations, yet it is essential to include these conformational changes to get correct binding free energies.

On this second point, it is interesting to note that previous computational work on this same binding site gave free energies that were more than 2 kcal/mol too negative for four of 10 known binders in calculations initiated from the *holo* structures.²³ We performed similar free energy calculations for these compounds and found that our calculations overestimated binding affinities for the same ligands, despite the fact that we use a different force field and different parameters, but only when we failed to account for the free energy associated with Val111 reorientation. Indeed, in the previous work, results were sensitive to the starting orientation of the Val111 side-chain for at least one ligand, indicating that this side-chain degree of freedom was

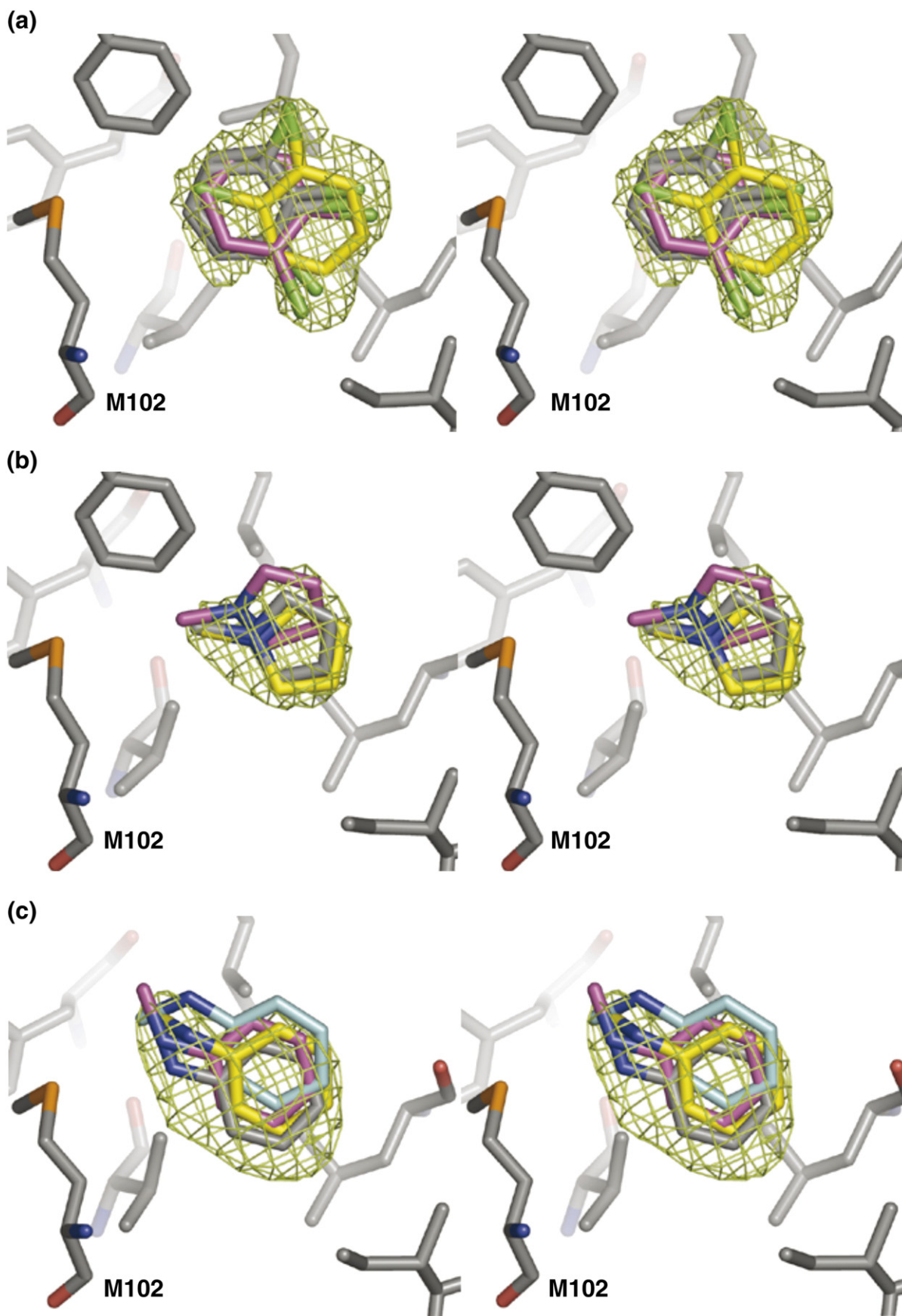


Figure 7 (legend on next page)

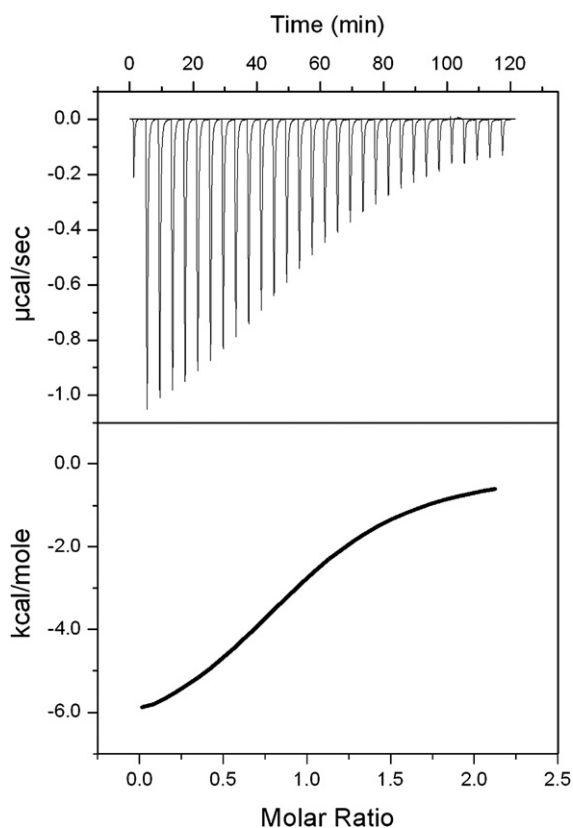


Figure 8. Representative ITC data. Data and fit for T4 lysozyme L99A (0.063 mM) titrated with 1,2-dichlorobenzene (~ 0.6 mM). An initial injection of 2.5 μ l was followed by 29 injections of 10 μ l of the ligand solution made every 2.5 min into the 1.4 ml reaction cell. After subtraction of blank runs, titrations were fit as described under Experimental methods to obtain the results in Table 4.

inadequately sampled.²³ Our results indicate that if the free energy associated with this reorientation is not included, then it can bias computed binding free energies by as much as 4 kcal/mol.

Given that these small conformational changes can contribute so substantially to overall binding free energies, it will likely prove essential to develop improved methods for computing free energies associated with larger-scale protein conformational changes, such as loop motions on ligand binding.

Lessons for docking

To address the rigid protein assumption typically used in docking, we also tried free energy calculations with the protein held rigid. Using a rigid *apo* structure for all ligands resulted in very large errors

(RMS error of 20 kcal/mol; zero correlation). Minimizing the protein separately in the presence of each ligand worked better, but RMS errors remained high (above 4 kcal/mol) and the approach lacked the ability to recognize nonbinders. Apparently, for this binding site, holding the protein rigid cannot easily produce binding affinities that agree quantitatively (even within a 4 kcal/mol RMS error) with the experiment, but strategies involving the minimization of the protein can provide some improvement over treating it as completely rigid. But for high accuracy, it may be necessary to include not only protein conformational change, but also a correct accounting for the free energy costs associated with these protein conformational changes, which can be substantial, even at the single side-chain level. Lastly, our results indicated that higher quality charges can lead to substantial improvements in binding affinities; thus the AM1-BCC charges that performed best here may also be a better choice for docking.

Conclusions

Overall, our results indicate that free energy methods are reaching the point where they can be useful when used predictively. However, in the relatively simple system examined here, this reasonably high level of accuracy depends on carefully accounting for the presence of multiple potential ligand-bound orientations and the possibility of protein conformational changes on ligand binding. In principle, extremely long molecular dynamics simulations can handle both of these, but in practice, the computational cost of such simulations is often prohibitive. For now, treating both problems requires deliberate sampling of the relevant degrees of freedom, and so some preknowledge of these degrees of freedom. We suspect that challenges observed in this model binding site will be found in biologically relevant binding sites as well. Despite these limitations, alchemical free energy methods hold great promise, both in predictive power and in guiding the improvement of more approximate physics-based methods.

Methods

Simulation methods

Overview

We begin with the benzene-bound crystal structure of the T4 lysozyme L99A mutant^{28,29} (which is virtually

Figure 7. Predicted and experimental ligand orientations. Stereo images comparing the experimental and predicted poses for three ligands bound to L99A. (a) The two observed configurations of 1,2-dichlorobenzene, structure determined to 1.70 Å resolution. (b) 1-Methylpyrrole, structure determined to 1.94 Å resolution and (c) *n*-methylaniline, structure determined to 2.07 Å resolution. The crystallographic carbon atoms of protein residue M102 and each ligand are colored grey. The carbon atoms of the docking predictions are colored yellow, and the carbon atoms of the free energy predictions are colored magenta. The carbon atoms of the second free energy prediction in (c) are colored cyan. The $F_o - F_c$ density maps are contoured at 3σ (green mesh).

identical with the *apo* structure, 1L90, from which it has an RMSD of 0.14 Å), dock our ligands into the binding site, and determine which poses or orientations are kinetically stable and distinct. We then begin free energy calculations from stable, distinct orientations, as described previously,²⁰ to compute binding free energies. Finally, for each ligand, we combine contributions from these different orientations to rigorously estimate binding free energies.

System preparation

Except where indicated otherwise, the initial protein structure for molecular dynamics simulations and free energy calculations is the benzene-bound structure of the L99A lysozyme mutant, which is essentially identical with the *apo* structure. This was prepared with the GROMACS^{38,39} 3.3 utility PDB2GMX with default protonation states, using a GROMACS port⁴⁰ of the AMBER 96⁴¹ force field. Since the cavity that makes up the binding site is completely hydrophobic without any nearby titratable groups, these protonation states present no difficulties. Following preparation, the protein was placed in a dodecahedral simulation box and surrounded by roughly 6000 water molecules which were pre-equilibrated for 1 ns with the protein held fixed prior to the equilibration of the full system, which is discussed below.

Docking

We used DOCK 3.5.54 to fit the molecules of interest into the protein structure (Tables 1 and 4). We retained all of the generated poses (numbering in the thousands) and scores of the molecules, then sorted these by score. We then began with the best-scoring pose and worked toward the worst, retaining every pose that was different by more than 2 Å RMSD from a better scoring pose, to generate a set of the best-scoring, distinct ligand orientations. This typically resulted in 10–40 distinct poses, of which we retained only the group of top-scoring poses (typically 3–8).

Identifying candidate orientations

From these poses, we generated general AMBER force field⁴² parameters for each ligand using ANTECHAMBER version 1.2.4,⁴³ and AM1-CM2 charges⁴⁴ as discussed previously.²⁰ These charges were employed in docking studies on the same system,³² and we sought to separate parameter differences from methodology differences as much as possible. We also present results in this work where we use AM1-BCC^{45,46} charges computed with ANTECHAMBER.

From the resulting small-molecule amber topology and coordinate files, we generated GROMACS topology and coordinate files using the amb2gmx conversion utility[§]. These ligand topologies and coordinates were then merged with those for the presolvated protein system prior to simulation.

To further reduce the number of ligand orientations we consider in free energy calculations, we initiated separate 1-ns molecular dynamics simulations from all candidate orientations to identify those that are kinetically distinct.²⁰ We only retained one orientation of each set of orienta-

tions that interconvert easily within simulation timescales. This typically resulted in 1–4 kinetically distinct orientations which were used for the calculations presented in Results.

Choosing restrained orientations

We then chose reference orientations for restraining the ligand in the binding site relative to the protein for subsequent free energy calculations. These are defined by picking a specific value in each of six relative protein–ligand degrees of freedom to which to restrain the ligand.²⁰ These values were chosen as the most probable value of each degree of freedom as determined from histograms computed during the 1-ns simulations, although in principle this choice is arbitrary.^{19,20} The degrees of freedom used are as described previously.^{19,20}

Binding free energy calculations

We carried out independent binding free energy calculations for each kinetically distinct orientation. Using the orientational decomposition procedure described previously,²⁰ we combined the effective binding free energies of each orientation into an overall binding free energy ($\Delta G_{\text{multiple}}^{\circ}$). We also computed binding free energies that would have resulted had we only considered a single potential bound orientation and neglected symmetry number corrections, as done in docking ($\Delta G_{\text{single}}^{\circ}$).

Binding free energy calculations were carried out in GROMACS 3.3 (with several crucial bugfixes described previously²⁰) using the Bennett acceptance ratio^{47,48} method to estimate free energy differences. To calculate absolute binding free energies, we used a thermodynamic cycle developed previously.^{17,19,20} In this cycle, we begin with the ligand bound to the protein, then restrain the ligand harmonically to a reference orientation within the binding site. We then annihilate the ligand's partial charges, then decouple its Lennard–Jones interactions with the rest of the system. The final ligand state is equivalent to a noninteracting ligand with no electrostatics, restrained, in vacuum or water. We then analytically calculate the free energy of removing the restraints and compute the free energy of restoring first the Lennard–Jones and then the electrostatic interactions in water. This entire process forms a thermodynamic cycle that transfers the ligand from the binding site to bulk water in the standard state. If all of the component calculations are converged, this rigorously provides a measurement of the absolute binding free energy, ΔG° , for the force field and solvent model used^{||}. As part of each of the steps in the cycle, independent free energy calculations were conducted at a number of intermediate alchemical states (denoted by the parameter λ) which were the same as those described in our previous work.²⁰

Following these binding free energy calculations and the predictions discussed below, we also computed binding free energies for the set of small molecules using AM1-BCC charges. To do this, we computed the free energy of changing AM1-CM2 to AM1-BCC charges in water for each compound, and then repeated the restraining and charging calculations for the compound in the

§ <http://folding.stanford.edu/ffamber>

|| The sum of the components around the cycle is actually the negative of ΔG° .

protein for each orientation. Since the Lennard–Jones decoupling is done with the compound’s electrostatics already turned off, it was unnecessary to repeat these calculations.

Simulation protocols

For all of the simulations discussed here (at each λ value), equilibration was performed as follows. First, velocities were assigned from a Maxwell–Boltzmann distribution at 300 K and the system was subjected to 10 ps of isothermal molecular dynamics. This was followed by 100 ps of isothermal-isobaric dynamics with pressure regulated by the Berendsen weak-coupling scheme⁴⁹ as discussed previously.²⁰ Following this, the simulation cell size was fixed and production simulations were run with isothermal dynamics, using the Langevin integrator for temperature control with a friction coefficient of 1 ps^{-1} . Production simulations were 1 ns in length for simulations of the complex (at each alchemical intermediate state, or λ value) and 5 ns for the ligand in water, except where noted otherwise.

All remaining protocols are as discussed previously,²⁰ with several exceptions: first, PME parameters were modified from those used previously to increase accuracy. Here, we used a PME spline order of 6, a relative tolerance of 10^{-6} , and a Fourier spacing as close as possible to 1.0 \AA . Additionally, we applied a long-range van der Waals correction (in addition to the analytical correction employed previously²⁰) to correct for the effect of truncating the long-range dispersive interactions at a finite cutoff. These interactions are everywhere attractive and can contribute significantly to binding free energies due to the fact that proteins have a higher density of attractive sites than water.⁵⁰ While this issue will be discussed in detail elsewhere,⁵⁰ the approach used here was, briefly, to run as usual the set of simulations where the ligand Lennard–Jones interactions are decoupled (that is, the ligand–environment Lennard–Jones interactions are turned off). These simulations were then reprocessed with long (24 \AA) cutoffs for Lennard–Jones interactions, and the weighted histogram analysis method⁵¹ was used to reweight the data from the simulations conducted with the short cutoff in order to estimate what the decoupling free energy would have been had we run with the longer cutoff. This is a relatively small correction ($0.2\text{--}0.8 \text{ kcal/mol}$) in the direction of increased binding affinity. This correction would be larger had not an approximate analytical dispersion correction already been included in the original runs by using the GROMACS correction option ENERPRES, and tends to be larger for larger ligands.⁵⁰

To aid the convergence of the calculations for benzene, which has a very high symmetry number, we used an approach employed previously²³ and restricted benzene to stay within a single symmetric orientation during our free energy calculations, then simply included the effect of symmetry as a symmetry number correction to the binding free energy.²⁰

Confine-and-release for Val111

For some ligands, a valine side-chain (Val111) in the binding cavity is observed experimentally to change rotameric states on ligand binding. This conformational change is not typically sampled (or not well sampled) during our molecular dynamics simulations (discussed in Results). Neglect of this change leads to an underestimate of binding affinities for those ligands when the *apo* protein structure is used, and an overestimate when the *holo*

protein structure is used. In the former case, the protein typically remains trapped in a conformation where the valine interferes with ligand binding; in the latter, when the ligand is removed from the binding site, the protein remains kinetically trapped with the side-chain in an unfavorable orientation, leading to neglect of protein strain energy in the free energy calculation. One way to properly account for the presence of these multiple conformations is to use the “confine-and-release” strategy.³⁴ The basic idea is to compute the binding free energy of the ligand to the protein, with the protein conformation confined (either kinetically or with artificial restraints) to a particular region of configuration space, then to compute the free energy of releasing the protein from confinement in the bound and unbound states. This provides a rigorous approach for computing binding free energies which include contributions from these conformational changes.³⁴

We apply the confine-and-release approach to compute the binding free energies of all the compounds considered here. To do this, we first compute the binding free energy with the valine side-chain kinetically trapped in the orientation from the *apo* structure (which we check by monitoring the dihedral angle throughout all of our simulations). In some cases, the valine actually manages to briefly escape from its kinetic trap at one or several λ values in the alchemical part of the calculation; we discard any simulation snapshots where it had done so from our data analysis in order to apply the confine-and-release approach. Once we have confined binding free energies, computed with the side-chain trapped, we use umbrella sampling and weighted histogram analysis method⁵¹ to compute the potential of mean force for rotating the valine side-chain in the bound and unbound states. From this, we compute the free energy of releasing the protein from confinement, and thus the binding free energies. We present our results with and without the confine-and-release approach, which provides a rigorous way to account for inadequate sampling.

Simulation details for the umbrella sampling calculations are as described previously.³⁴

Since free energy calculations were conducted with two charge sets, two sets of umbrella sampling calculations for Val111 were carried out for each ligand: one where the ligand had AM1-BCC charges, and one where it had AM1-CM2 charges. This was important, since the details of the ligand electrostatics can influence the free energy landscape associated with this reorientation.

Predictions

To predict whether untested compounds would bind, we selected five small molecules predicted by docking (using protocols described previously)³² to be binders. We followed the same protocols described above: docking to the binding site, retaining a number of different kinetically distinct starting orientations, running separate binding free energy calculations for each of these, and then combining them to get a total binding free energy. We also applied the confine-and-release approach to account for any reorientation of the Val111 side-chain on binding of these molecules. From computed binding free energies, we calculated dissociation constants, and then predict that those compounds with dissociation constants less than 10 mM (the experimental detection threshold for the thermal upshift assay) should bind. These predictions, and those of bound structures below, were made with AM1-CM2 charges, as AM1-BCC charges were only examined later.

Predicting bound structures of these unknown ligands is challenging, since our method is intended to provide an accurate estimate of the binding free energy which includes contributions from a variety of different ligand and protein structures, and neglects any effects of the crystal environment which are present in X-ray structures, as well as other differences. Here we attempted to identify the dominant bound structure by identifying which kinetically distinct ligand binding orientation contributes most favorably to the total binding free energy; to do this, we calculated occupancy probabilities of the different dominant orientations from their estimated binding free energies. The predicted orientation was the orientation with the highest probability of occupancy. Then we predicted a bound structure by taking a representative snapshot from a simulation where the ligand remains, without restraints, stably within the region of configuration space corresponding to that orientation. Our predicted bound orientations, then, were single snapshots from molecular dynamics simulations. For one ligand, we predicted that two orientations would have nearly comparable occupancy probabilities, so we predicted that both would be observed.

After these predictions, we continued retrospective studies and found that the AM1-BCC charge model gave more accurate binding free energies. Therefore, we used the AM1-BCC charge set to make predictions for binding free energies prior to measuring these calorimetrically.

Binding free energies to a rigid protein

To compare our methodology more closely with docking, we repeated the free energy calculations using essentially the same protocols, but with the protein held rigid in its prepared starting structure, as is often done in docking. To do so, we used the GROMACS option of defining frozen groups that are held fixed during dynamics. Because there are so many fewer degrees of freedom when the protein is completely rigid, convergence was more rapid, and production simulations required only 100 ps at each λ value. Protocols were otherwise the same, and these calculations used AM1-CM2 charges.

We considered several choices for the rigid protein structure. First, we held the protein rigid in its starting structure. Second, motivated by testing approaches that could easily be applied in docking and scoring, we minimized the entire protein in the presence of each ligand individually in vacuum, and used each of these structures for the appropriate ligand. The RMSD to the starting prepared structure is typically around 0.5 Å with this approach. Finally, we modified this second protocol to allow only residues near the binding site (residues 78, 84, 85, 87, 88, 91, 98, 99, 100, 102, 103, 106, 111, 118, 121, 133, and 153) to move during minimization. With this protocol, changes in structure were very minor (often less than 0.01 Å RMSD (the RMSD reported is for the protein as a whole)).

Minimization protocols were as discussed above and previously,²⁰ except the order of minimization was reversed (steepest descents followed by limited memory Broyden-Fletcher-Goldfarb-Shanno) and, since minimization was done in vacuum, cutoff electrostatics was used instead of PME, using a cutoff of 11 Å.

Error analysis

Calculated uncertainties reported here are one standard deviation of the mean over 40 block bootstrap trials, where the block length is taken to be equal to the autocorrelation time, as described previously.²⁰

Experimental methods

Binding detection by upshift of thermal denaturation temperature

To detect binding, L99A protein was denatured reversibly by temperature in the presence and absence of the putative ligand. Molecules that bind preferentially to the folded cavity-containing protein should stabilize it relative to the *apo* protein, raising its temperature of melting.²⁸ Thermal denaturation experiments were carried out in a Jasco J-715 spectropolarimeter with a Jasco PTC-348WI Peltier-effect in-cell temperature control device and in-cell stirring. Each compound was screened in its neutral form. 1,2-Benzenedithiol was assayed in a pH 3 buffer containing 25 mM KCl, 2.9 mM phosphoric acid, and 17 mM KH₂PO₄. Compounds 1,2-dichlorobenzene and 1-methylpyrrole were screened in a pH 5.4 buffer containing 100 mM sodium chloride, 8.6 mM sodium acetate, and 1.6 mM acetic acid. Compounds thieno[2,3-*c*]pyridine and *n*-methylaniline were screened in a buffer composed of 50 mM potassium chloride (pH 6.8) and 38% (v/v) ethylene glycol. All buffers are as previously described.²⁸ Thermal denaturation of the protein in the presence of compounds 1,2-dichlorobenzene, thieno[2,3-*c*]pyridine, and *n*-methylaniline was monitored by circular dichroism between 223 and 234 nm (although the 223 nm wavelength is the ideal wavelength for measuring the helical signal of T4 lysozyme, the higher wavelengths, which were less affected by absorbance from some of the compounds, can be used to monitor the edge of the helical signal). For 1,2-benzenedithiol and 1-methylpyrrole, which have high absorbance in the far UV region, thermal denaturation was measured by the intensity of the integrated fluorescence emission for all wavelengths above 300 nm, exciting at 290 nm. Thermal melts were performed at a temperature ramp rate of 2 K/min. A least-squares fit of the two-state transition model was performed with the program EXAM⁵² to calculate T_m and van't Hoff ΔH values for the thermal denaturations. The ΔC_p was set to 1.94 kcal mol⁻¹ K⁻¹.

Thermal denaturation of *apo* T4 lysozyme L99A was carried out with 0.02–0.04 mg/ml protein in the same buffer conditions described above. Compounds were included at concentrations as high as 10 mM. Each denaturation experiment was performed at least three times.

Isothermal titration calorimetry

Quantitative estimates of association for ligand binding to L99A T4 lysozyme were obtained by ITC using a Microcal VP-ITC calorimeter⁵³ operated at 10 °C with a reference power of 10 μ cal/s, a stirring speed of 300 rpm, and a data collection interval of 4 min per injection. An initial injection of 2 μ l of ligand was followed by an additional 29 injections of 10 μ l totaling 292 μ l. These were added to 0.05 to 0.13 mM protein in the 1.4266 ml sample cell. The concentration of small molecule ligands in the syringe was adjusted such that the final molar ratio of ligand to protein was at least twofold by the end of the titrations. Protein concentrations were determined by molar absorptivity at 280 nm in 0.5 M NaCl, 0.1 M sodium phosphate buffer (pH 6.8). Ligand concentrations were determined from the volume of material added to a known volume of buffer. Baseline mixing heats were estimated by injection of ligand into buffer. Reaction heat profiles were fit to the single

binding site model using the ITC worksheet of ORIGIN version 7.0.

Protein preparation and crystallography

T4 lysozyme mutant L99A was overexpressed and purified, and crystals were grown as described previously.⁵⁴ The crystals belong to space group $P3_22_1$. Crystals were soaked overnight to four days in crystallization buffer containing as much as 50 mM compound. In addition, drops of neat compound were added to the cover slip surrounding the drop containing the crystal. After soaking, the crystals were cryoprotected with a 50:50 Paraton-N (Hampton Research, Aliso Viejo, CA), mineral oil mix. X-ray data were collected at 110 K with an in-house Raxis IV detector. Reflections were indexed, integrated, and scaled using the HKL package.⁵⁵ The complex structures were refined using REFMAC5.⁵⁶ For model building and water placement, we used Coot.⁵⁷

Protein Data Bank accession numbers

The benzene-bound crystal structure of the T4 lysozyme L99A mutant has been deposited in the Protein Data Bank (PDB) with accession number 181L. The X-ray crystal structures of 1,2-dichlorobenzene, *n*-methylaniline, and 1-methylpyrrole have been deposited in the PDB with accession numbers 2OTY, 2OTZ, and 2OU0, respectively. The bound structures of indene, *n*-butylbenzene, isobutylbenzene, *p*-xylene, and *o*-xylene have accession numbers 183L, 186L, 184L, 187L, and 188L, respectively.

Acknowledgements

We thank Sarah Boyce for help with protein preparation. We thank Benoît Roux (University of Chicago), Bill Swope (IBM Almaden), Jed Pitera (IBM Almaden), Guha Jayachandran (Stanford University), and Vijay Pande (Stanford University) for helpful discussions, and the reviewers for insightful comments on the manuscript. JDC was supported in part by HHMI and IBM predoctoral fellowships. DLM and KAD acknowledge NIH grant GM63592, Anteon Corporation grant USAF-5408-04-SC-0008, and a UCSF Sandler Award. BKS acknowledges NIH grant GM59957. This work was performed in part with the UCSF QB3 Shared Computing Facility.

References

- Halperin, I., Ma, B., Wolfson, H. & Nussinov, R. (2002). Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins: Struct. Funct. Genet.* **47**, 409–443.
- Shoichet, B. K., Leach, A. R. & Kuntz, I. D. (1999). Ligand solvation in molecular docking. *Proteins: Struct. Funct. Genet.* **34**, 4–16.
- Huang, N., Kalyanaraman, C., Bernacki, K. & Jacobson, M. P. (2006). Molecular mechanics methods for predicting protein–ligand binding. *Phys. Chem. Chem. Phys.* **8**, 5166–5177.
- Srinivasan, J., Cheatam, T. E., III, Cieplak, P., Kollman, P. A. & Case, D. A. (1998). Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **120**, 9401–9409.
- Kuhn, B. & Kollman, P. A. (2000). Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J. Med. Chem.* **43**, 3786–3791.
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L. *et al.* (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897.
- Allen, T. W., Andersen, O. S. & Roux, B. (2004). On the importance of atomic fluctuations, protein flexibility, and solvent in ion permeation. *J. Gen. Physiol.* **124**, 679–690.
- Lee, M. S. & Olson, M. A. (2006). Calculation of absolute protein–ligand binding affinity using path and endpoint approaches. *Biophys. J.* **90**, 864–877.
- Pearlman, D. A. (2005). Evaluating the molecular mechanics Poisson–Boltzmann surface area free energy method using a congeneric series of ligands to p38 map kinase. *J. Med. Chem.* **48**, 7796–7807.
- Steinbrecher, T., Case, D. A. & Labahn, A. (2006). A multistep approach to structure-based drug design: studying ligand binding at the human neutrophil elastase. *J. Med. Chem.* **49**, 1837–1844.
- Woo, H.-J. & Roux, B. (2005). Calculation of absolute protein–ligand binding free energy from computer simulations. *Proc. Natl Acad. Sci. USA*, **102**, 6825–6830.
- Zhang, Y. & McCammon, J. A. (2006). Potentials of mean force for acetylcholine unbinding from the alpha7 nicotinic acetylcholine receptor ligand-binding domain. *J. Am. Chem. Soc.* **128**, 3019–3026.
- Rodinger, T. & Pomès, R. (2005). Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr. Opin. Struct. Biol.* **15**, 164–170.
- Kofke, D. A. (2005). Free energy methods in molecular simulation. *Fluid Phase Equilib.* **228–229**, 41–48.
- Shirts, M. R., Mobley, D. L. & Chodera, J. D. Alchemical free energy calculations: ready for prime time? *Ann. Rep. Comp. Chem.* In the press.
- Roux, B., Nina, M., Pomès, R. & Smith, J. C. (1996). Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys. J.* **71**, 670–681.
- Gilson, M. K., Given, J. A., Bush, B. L. & McCammon, J. A. (1997). The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **72**, 1047–1069.
- Hermans, J. & Wang, L. (1997). Inclusion of the loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *J. Am. Chem. Soc.* **119**, 2707–2714.
- Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. (2003). Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem., B*, **107**, 9535–9551.
- Mobley, D. L., Chodera, J. D. & Dill, K. A. (2006). On the use of orientational restraints and symmetry

- corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902.
21. Wang, J., Deng, Y. & Roux, B. (2006). Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* **91**, 2798–2814.
 22. Fujitani, H., Tanida, Y., Ito, M., Jayachandran, G., Snow, C. D., Shirts, M. R. *et al.* (2005). Direct calculation of the binding free energies of FKBP ligands. *J. Chem. Phys.* **123**, 084108.
 23. Deng, Y. & Roux, B. (2006). Calculation of standard binding free energies: aromatic molecules in the internal cavity of the T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* **2**, 1255–1273.
 24. van den Bosch, M., Swart, M., Snijderst, J. G., Berendsen, H. J. C., Mark, A. E., Oostenbrink, C. *et al.* (2005). Calculation of the redox potential of the protein azurin and some mutants. *Chem. Bio. Chem.* **6**, 738–746.
 25. Zhou, Y., Oostenbrink, C., Jongejan, A., Hagen, W. R., de Leeuw, S. W. & Longejan, J. A. (2006). Computational study of ground-state chiral induction in small peptides: comparison of the relative stability of selected amino acid dimers and oligomers in homo-chiral and heterochiral combinations. *J. Comp. Chem.* **27**, 857–867.
 26. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*, **255**, 178–183.
 27. Eriksson, A. E., Baase, W. A. & Matthews, B. W. (1993). Similar hydrophobic replacements of leu99 and phe153 within the core of t4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol.* **229**, 747–769.
 28. Morton, A. & Matthews, B. W. (1995). Specificity of ligand binding in a buried non-polar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. *Biochemistry*, **34**, 8576–8588.
 29. Morton, A., Baase, W. A. & Matthews, B. W. (1995). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry*, **34**, 8564–8575.
 30. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339–355.
 31. Wei, B. Q., Weaver, L. H., Ferrari, A. M., Matthews, B. W. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **337**, 1161–1182.
 32. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for docking. *J. Med. Chem.* **48**, 3714–3728.
 33. Mann, G. & Hermans, J. (2000). Modeling protein–small molecule interactions: structure and thermodynamics of noble gases binding in a cavity in mutant phage T4 lysozyme L99A. *J. Mol. Biol.* **302**, 979–989.
 34. Mobley, D. L., Chodera, J. D. & Dill, K. A. Confine and release: obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.* 2007, <http://dx.doi.org/10.1021/ct700032n>.
 35. Mobley, D. L., Dumont, E., Chodera, J. D. & Dill, K. A. (2007). Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. *J. Phys. Chem., B.* **111**, 2242–2254.
 36. Chambers, C. C., Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. (1996). Model for aqueous solvation based on class iv atomic charges and first solvation-shell effects. *J. Phys. Chem., A*, **100**, 16385–16398.
 37. Turnbull, W. B. & Daranas, A. H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125**, 14859–14866.
 38. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 306–317.
 39. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718.
 40. Sorin, E. J. & Pande, V. S. (2005). Exploring the helix–coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* **88**, 2472–2493.
 41. Kollman, P. A. (1996). Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc. Chem. Res.* **29**, 461–469.
 42. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general AMBER force field. *J. Comp. Chem.* **25**, 1157–1174.
 43. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **26**, 247–260.
 44. Li, J., Zhu, T., Cramer, C. J. & Truhlar, D. G. (1998). New class IV charge model for extracting accurate partial charges from wave functions. *J. Phys. Chem., A*, **102**, 1820–1831.
 45. Jakalian, A., Jack, D. B. & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. am1-bcc model: li. Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641.
 46. Jakalian, A., Bush, B. L., Jack, D. B. & Bayly, C. I. (2000). Fast, efficient generation of high-quality atomic charges. am1-bcc model: I. Method. *J. Comput. Chem.* **21**, 132–146.
 47. Bennett, C. H. (1976). Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **22**, 245–268.
 48. Shirts, M. R. & Pande, V. S. (2005). Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett Acceptance Ratio, and thermodynamic integration. *J. Chem. Phys.* **122**, 144107-1–144107-16.
 49. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3584–3690.
 50. Shirts, M. R., Mobley, D. L., Chodera, J. D. & Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *J. Phys. Chem., B*. Submitted.
 51. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules: I. The method. *J. Comp. Chem.* **13**, 1011–1021.
 52. Kirchhoff, W. (1993). *EXAM: A Two-State Thermodynamic Analysis Program*. NIST, Gaithersburg, Maryland.
 53. Plotnikov, V. V., Brandts, J. F. & Lin, L. N. (1997). A new ultrasensitive scanning calorimeter. *Anal. Biochem.* **250**, 237–244.
 54. Poteete, A. R., Dao-Pin, S., Nicholson, H. & Matthews, B. W. (1991). Second-site revertants of an inactive t4 lysozyme mutant restore activity by restructuring the active site cleft. *Biochemistry*, **30**, 1425–1432.

-
55. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
56. Murshudov, G. N., Vagin, A. A. & Dobson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., D Biol. Crystallogr.* **53**, 240–255.
57. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., D Biol. Crystallogr.* **60**, 2126–2132.
58. Shirts, M. R. Calculating Precise and Accurate Free Energies in Biomolecular Simulations. Ph.D. Dissertation, Stanford University, 2004. ProQuest location: <http://wwwlib.umi.com/dissertations/fullcit/3153076>.

Edited by M. Levitt

(Received 10 April 2007; received in revised form 29 May 2007; accepted 1 June 2007)
Available online 8 June 2007