Design of Protein–Ligand Binding Based on the Molecular-Mechanics Energy Model

F. Edward Boas and Pehr B. Harbury*

Department of Biochemistry, Stanford University School of Medicine, Beckman B437, 279 Campus Drive West, Stanford, CA 94305-5307, USA

Received 8 October 2007; received in revised form 25 March 2008; accepted 1 April 2008
Available online 8 April 2008

While the molecular-mechanics field has standardized on a few potential energy functions, computational protein design efforts are based on potentials that are unique to individual laboratories. Here we show that a standard molecular-mechanics potential energy function without any modifications can be used to engineer protein–ligand binding. A molecular-mechanics potential is used to reconstruct the coordinates of various binding sites with an average root-mean-square error of 0.61 Å and to reproduce known ligand-induced side-chain conformational shifts. Within a series of 34 mutants, the calculation can always distinguish between weak ($K_d > 1$ mM) and tight ($K_d < 10$ μM) binding sequences. Starting from partial coordinates of the ribose-binding protein lacking the ligand and the 10 primary contact residues, the molecular-mechanics potential is used to redesign a ribose-binding site. Out of a search space of $2 \times 10^{12}$ sequences, the calculation selects a point mutant of the native protein as the top solution (experimental $K_d = 17$ μM) and the native protein as the second best solution (experimental $K_d = 210$ nM). The quality of the predictions depends on the accuracy of the generalized Born electrostatics model, treatment of protonation equilibria, high-resolution rotamer sampling, a final local energy minimization step, and explicit modeling of the bound, unbound, and unfolded states. The application of unmodified molecular-mechanics potentials to protein design links two fields in a mutually beneficial way. Design provides a new avenue for testing molecular-mechanics energy functions, and future improvements in these energy functions will presumably lead to more accurate design results.

Edited by B. Honig

Keywords: protein design; generalized Born; force field; dissociation constant; structure prediction

Introduction

Computer-aided design of a ligand-binding site is similar to solving a three-dimensional jigsaw puzzle: it involves fitting together the right pieces (amino acid mutations) to create a properly shaped and functionalized pocket for a ligand. The inputs to the design procedure are the crystal structure of a scaffold protein, a ligand structure, and a set of amino acid positions that will be mutated to create the binding site. The orientations of candidate jigsaw puzzle pieces are determined by modeling the conformations that the ligand and the surrounding amino acids can adopt, so as to identify the lowest-energy arrangement. The design procedure searches through thousands of candidate sequences for one that optimizes the computed binding free energy of the ligand with the protein. The whole process depends heavily on the potential energy function (PEF)—a mathematical expression embodying the physical laws that govern the protein–ligand and solvent system.

Over the past 30 years, PEFs have played a central role in the molecular-mechanics field. This field has converged onto a small set of standard PEFs that have been extensively tested. Identification and
correction of the limitations of these energy models are areas of active research.2–4 Modern molecular-mechanics potential energy functions (MM-PEFs) treat proteins as a collection of atoms with partial charges and van der Waals parameters, connected by springs that maintain bond lengths and angles. The parameters are derived from quantum calculations and from experimental data on a wide range of systems.5 MM-PEFs have been used to calculate binding constants,6–10 protein folding kinetics,11 protonation equilibria,12 and active-site coordinates.8,13,14

Perhaps surprisingly, standard MM-PEFs are not used for protein design.15 The reason is that computing energies using MM-PEFs requires significant computer time and is very sensitive to detailed atom positions, necessitating fine conformational sampling. When thousands of different sequences must be evaluated, the computation time per sequence becomes critical. In order to accelerate calculations, design algorithms typically use simplified PEFs with various ad hoc energy terms13,16–28 (these PEFs have also been used to predict binding constants9,30 and to predict active-site coordinates28). Water is treated in a simplified way: for example, by inserting a distance-dependent dielectric constant into Coulomb’s law and by applying a surface-area-based solvation energy.16,17 The van der Waals interaction is frequently smoothed, so that it is less sensitive to spatial position and thus can be optimized with coarse sampling.16–18 Rather than explicitly modeling reference states, such as the unfolded state, the reference states are typically treated implicitly by modifying the PEF.16,18 Statistical terms derived by counting how frequently different residues and functional groups interact in crystal structures are included as well.16–18 Relative weights for the various energy terms are adjusted empirically so as to match experimental data.18,19 Similar approximations were used in the early days of molecular-mechanics calculations, but were replaced as better models and increased computational power became available.

There are several motivations for trying to identify a single standardized energy function that is practically useful for protein design. First, design results produced in different laboratories using a standardized PEF could be compared, and those results would collectively address where the energy model had failed and how to improve it. Second, the practice of computational protein design would be simplified if PEF development were not required. Finally, a PEF that had been broadly validated might be expected to generalize better to new design problems than would a customized PEF.

One reasonable choice for a universal energy function would be an MM-PEF. MM-PEFs are the most broadly tested PEFs,5 and a direct correspondence exists between them and more rigorous quantum mechanical treatments of matter.5 A large group of scientists are working on MM-PEFs, and the advances they make would be directly applicable to design. Here, we test whether protein–ligand binding sites can be successfully designed based on a standard MM-PEF that does not include any heuristic corrections. We first describe how we directly apply an MM-PEF to the protein design problem and then detail various tests applied to the ribose-binding protein (RBP).

Results

Design scheme

Using the genetic algorithm,32 we search through thousands of sequences to find one sequence that maximizes the calculated protein–ligand dissociation energy without destabilizing the protein by >5 kcal/mol. To evaluate dissociation and unfolding energies, the bound, unbound, and unfolded states are modeled, and their calculated energies are differentiated. For each state, we use a mean field rotamer-reppacking algorithm to find the atomic coordinates that minimize the energy. As part of the rotamer repacking, titratable residues are allowed to protonate or deprotonate, depending on the local energetics. Good structural sampling is achieved by using extremely large rotamer libraries (≥5449 rotamers per position) and several thousand ligand poses that sample the translational, rotational, and internal degrees of freedom of the ligand. The optimal structure generated by rotamer repacking is then subjected to gradient-based energy minimization. The energies of each state are evaluated with the unmodified CHARMM22 MM-PEF33 and the generalized Born solvation formalism34 developed by Lee et al.35 The design procedure is outlined in Fig. 1. To evaluate the approach, we apply three tests: structural prediction, energetic prediction, and prediction of a binding-site sequence.

Structural prediction

For structural predictions, we started with crystal structures and discarded the coordinates of the ligand and all contacting side chains. These coordinates were then predicted in the context of the rest of the protein. No native coordinates for the ligand or side chains were included as rotamers. We first explored the effect of sampling resolution by predicting the structure of RBP bound to ribose using four rotamer libraries of increasing size (Fig. 2). With fewer than 5449 rotamers per position, the calculated energy of the predicted structure is less favorable than the calculated energy of the crystal structure, indicating that the crystal structure conformation is missed due to inadequate sampling resolution. At 5449 rotamers per position, the predicted structure has the same energy as the energy-minimized crystal structure, and the coordinates differ by a root-mean-square (RMS) error of 0.148 Å. This level of accuracy exceeds the experimental error in the crystallographic coordinates. This apparently surprising result likely occurs because the fixed portion of the crystallographic
coordinates constrains the possible solutions at the modeled positions. However, this constraint alone is not sufficient for specifying the binding-site sequence and geometry (see below).

Using high-resolution rotamer libraries (either 5449 or 6028 rotamers per position), side chains in the binding sites of five different structures were predicted with an average RMS error of 0.61 Å (Figs. 3 and 4). The number of predicted residues ranged from 9 to 23. The error was generally larger for surface residues and when more positions were predicted.

For the RBP-ribose calculations, we restricted the ligand poses to be within 1.8 Å RMS of the native pose, resulting in the 4639 poses shown in Fig. 3. Similarly, 4111 arabinose poses within 1.0 Å RMS of the native pose were used for the arabinose-binding protein (ABP)-arabinose calculations (Fig. 3). Although we would have preferred to perform the calculations without this filter, it was necessary to reduce the number of ligand poses to a manageable number (the precalculated interaction energy matrices had to be smaller than 2 GB to fit into memory).

We explicitly model the bound and unbound states, providing predictions of side-chain conformational shifts upon binding. The predicted changes match the crystal structures in 70% of the residues with the largest conformational shifts (Fig. 5). Single-state design algorithms ignore such conformational shifts, in contrast to a multistate design framework.22 Note that we did not predict the backbone shift upon binding [4.1 Å RMS for RBP and 0.8 Å RMS for vascular endothelial growth factor (VEGF)] because the bound and unbound backbone coordinates were used as inputs to the calculation.

The calculation predicts that one aspartic acid and one glutamic acid in the binding site of ABP are protonated (Supplementary Table 5). If these residues are not allowed to protonate, the structural prediction is degraded (Supplementary Fig. 7).

**Energetic prediction**

To test whether the energy function can properly rank the binding affinities of different binding-site sequences, we first computed ligand binding energies for the native ABP and RBP sequences and for 1000 scrambled sequences. As expected, none of the scrambled sequences has better predicted stability and dissociation energy than the native (Fig. 6a). Next, we calculated the relative binding energies of 34 mutants of ABP for which dissociation energies have been measured. Two sequences were predicted...
to destabilize the protein by $>10$ kcal/mol relative to native ABP and presumably to adopt alternative backbone conformations. The binding energies of the remaining sequences are predicted with a correlation coefficient of $r^2 = 0.57$ (Fig. 6b, Supplementary Table 6). The predictions were performed without any adjustable parameters. As each calculation required about 1 min of central processing unit time on a Pentium processor, the approach is fast enough for design applications. The data set includes single, double, and triple point mutants of wild-type ABP, and covers a wide range of mutation types (hydrophobic to hydrophobic, hydrophobic to polar/charged, polar/charged to hydrophobic, and polar/charged to polar/charged).

Within the data set, the calculation can always distinguish between weak ($K_d > 1$ mM) and tight ($K_d < 10$ μM) binding sequences. However, the absolute dissociation energies are not predicted correctly. One important possible source of error is that there is no published crystal structure of unbound ABP. We model the unbound protein backbone conformation based on the crystal structure of bound ABP. In reality, the unbound protein likely exists in an open conformation with better-solvated binding-site residues. An incorrect unbound state might explain the 21.2-kcal/mol offset in calculated dissociation energies. The slope of the regression line is $N$, which is likely due to modes of structural relaxation (such as backbone motions) that were not modeled. The resulting clashes will exaggerate any energy differences between sequences. Another possibility is that we

Fig. 2. Higher rotamer resolution improves structural predictions for the RBP-binding site [Protein Data Bank (PDB) code 2DRI]. Δ Energy is the difference in potential energy between the calculated structure and the crystal structure, after both have been subjected to local energy minimization. RMS error is the root-mean-square deviation between the calculated coordinates and crystallographic coordinates of the repacked atoms, comprising the ligand and 10 active-site side chains. The phenylalanine rotamers from each rotamer library are shown to illustrate the sampling resolution. The lowest-resolution rotamer library shown is the Lovell penultimate rotamer library, with protonation states added for His, Asp, and Glu. The other rotamer libraries were derived by clustering side-chain conformations in high-resolution crystal structures from the PDB (see Supplementary Information).

<table>
<thead>
<tr>
<th>Ligand poses</th>
<th>Side chain rotamers</th>
<th>crystal structure / predicted structure</th>
<th>RMS error</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABP-arabinose</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>RBP-ribose</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Fig. 3. Prediction of binding-site coordinates. Starting from crystal structures stripped of the ligand and the contacting residues, the active site was reconstructed by finding the lowest-energy arrangement of the ligand and side chains. For ABP-arabinose (PDB code 6ABP), the coordinates of the arabinose and 15 contacting residues (10, 14, 16, 17, 64, 89, 90, 108, 145, 147, 151, 204, 205, 232, and 259) were predicted using 6028 rotamers per position and 4111 ligand poses. For RBP-ribose (PDB code 2DRI), the coordinates of ribose and 10 contacting residues (13, 15, 16, 89, 90, 141, 164, 190, 215, and 235) were predicted using 5449 rotamers per position and 4639 ligand poses.
are not adequately modeling entropy losses upon binding. 39

**Binding site design**

The final and most stringent test of the molecular-mechanics energy model was a redesign of the binding site in RBP (Fig. 7). We discarded the ligand coordinates, and the sequence and coordinates of the 10 residues contacting the ligand. The total size of the sequence space searched was $17^{10} = 2.0 \times 10^{12}$ (Gly, Pro, and Cys were not allowed). The calculation was initiated from a population of random sequences. After the evaluation of 8888 sequences, the energy function identified a point mutant (N13L) of native RBP as the tightest binding sequence. After 8964 sequences, it picked native RBP as the second tightest binding sequence. Evaluation of an additional 8879 sequences did not yield any further improvement. The entire process was repeated with a different random initial sequence population, and the same optimal sequences were selected. During

![Predicted structure](image1)

Fig. 4. Prediction of binding site-coordinates for bevacizumab–VEGF (1BJ1), unbound VEGF (2VPF), and unbound RBP (1URP). For bevacizumab–VEGF, the following 23 residues were repacked using 6028 rotamers per position: V17, V21, W48, W79, W81, W82, W83, W91, W93, H28, H30, H31, H32, H54, H55, H99, H101, H102, H103, H105, H106, H107, and H108. V and W: VEGF chains; H and L: antibody heavy and light chains. For unbound VEGF and RBP, the same set of residues was predicted as the bound structure.

![Predicted structure](image2)

Fig. 5. Prediction of side-chain conformational shifts in RBP upon binding ribose, or in VEGF upon binding bevacizumab. The five largest experimentally observed conformational shifts are shown for each protein. The residues were superimposed by aligning the backbone amide nitrogen, $\alpha$ carbon, and carbonyl carbon. * Correct predictions, where the unbound and bound predictions are closest to the unbound and bound crystallographic coordinates, respectively.
Fig. 6. Predicting dissociation energies. (a) Calculated stability and dissociation energy distinguish the native sequence (x) from 1000 scrambled sequences (◆) for ABP and RBP. Sequences predicted to be >10 kcal/mol destabilized relative to the native are shown in gray. (b) Predicting relative dissociation energies of mutants. The graph shows data on mutants of ABP binding to arabinose. Experimental data are from Declerck and Abelson37 and from measurements reported in Supplementary Table 6. An experimental dissociation energy of zero means that there was no detectable binding. Calculations were performed using 6ABP as the scaffold structure for both the bound state and the unbound state, with 6028 rotamers per position. Coordinates of the 15 primary ligand contacts and of residues 20 and 235 were optimized. The circled points are predicted to be destabilized by >10 kcal/mol relative to the native.

Fig. 7. Redesigning the ribose-binding site in RBP. Positions identical with the native are highlighted in yellow. The figure shows the best sequence as a function of the number of sequences considered, using as the criterion either the mean field dissociation energy (blue trajectories) or, alternatively, the dissociation energy calculated using minimized structures (red trajectories). Purple sequence numbers indicate sequences identified by both criteria. All sequences with a mean field dissociation energy of >30 kcal/mol (corresponding to −7.5 kcal/mol relative to the native sequence; dashed line) were locally energy minimized to generate the red trajectory. Sequence 8871 is the top sequence when ranked by mean field dissociation energy (corresponding to Table 1b) and sequence 8888 is the top sequence when ranked by minimized dissociation energy (corresponding to Table 1d). The native sequence was found out of a possible 2×1012 sequences after 8964 sequence evaluations. Dissociation and unfolding energies are reported in kilocalories per mole relative to the native sequence. The number of protein–ligand hydrogen bonds was determined using bndlst.40 Shape complementarity (ranging from 0=perfectly noncomplementary surfaces to 1=perfectly complementary surfaces) was calculated using sc.41 Backbone coordinates for the bound state are from 2DRI, and backbone coordinates for the unbound state are from 1URP.
the course of the design, stability was achieved first, then hydrogen bonding, and, finally, shape complementarity. The same pattern has been seen experimentally in the affinity maturation of antibodies against lysozyme.42

We experimentally tested the three top sequences from four different RBP-ribose redesign calculations to determine which aspects of the design algorithm were essential (Table 1). Decreasing the rotamer resolution (row a), omitting the final continuous minimization step (row b), or using a less accurate electrostatics model (row c) produces sequences that bind very weakly. Only when we use a high-resolution rotamer library, a final continuous minimization step, and accurate electrostatics does the design algorithm predict sequences that bind well (row d).

Discussion

This article reports the first successful redesign of an entire binding site based on an unmodified MM-PEF. This is a stringent test of the energy function because the native sequence and a point mutant are distinguished from \( 2.0 \times 10^{12} \) alternative sequences. Good hydrogen bonds and steric complementarity were picked out directly by the energy function, without energy terms or selection criteria that specifically required these features. Given that the underlying physics is the same for the design of new proteins and for the simulation of known proteins, it is satisfying to see that the same energy models can be used as well.

We tested a number of simplifications that are commonly used in protein design calculations and found that they all resulted in less successful predictions. For example, low sampling resolution or an inaccurate solvation model led to sequences that lacked critical hydrogen bonds. Scaling down the electrostatic energy (which is frequently performed to compensate for a crude electrostatics model) reduced the accuracy of the energetic predictions. Eliminating the unfolded state resulted in unstable designed proteins. Softening the van der Waals interaction allowed atoms to pack together more closely, making hydrogen bonds and salt bridges appear artificially strong (Supplementary Fig. 8) and resulting in the burial of charged and polar functional groups.

An important conclusion from this work is that MM-PEFs must be paired with an accurate continuum solvent model and with protonation equilibria in order to correctly redesign a polar binding site. Individual polar protein–ligand interactions can exhibit energies of up to 100 kcal/mol (the Coulomb energy between unit charges separated by 3.3 Å). These energies are almost exactly counterbalanced by interactions with water in the unbound protein. Thus, small errors in the solvation energy grossly alter the design predictions. Finite-difference algorithms are generally considered as the most accurate methods for solving the Poisson–Boltzmann differential equation that defines the continuum solvent model, but they are currently too slow for protein design. Very accurate generalized Born approaches have been developed over the last few years and produce solvation energies that differ from the finite-difference result by only 2% (Supplementary Information). We have shown that this level of accuracy is both necessary and sufficient for protein design calculations.

### Table 1

<table>
<thead>
<tr>
<th>Design calculation</th>
<th>Rotamers per position</th>
<th>Local minimization</th>
<th>Solvent treatment</th>
<th>Rank</th>
<th>Number of residues identical to native</th>
<th>( K_d ) (experimental), ( \mu M )</th>
<th>Sequence (10 primary contacts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) 2800</td>
<td>yes</td>
<td>Lee</td>
<td></td>
<td>1</td>
<td>3</td>
<td>210,000±80,000*</td>
<td>NMLMLMMMNAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>8800±400</td>
<td>NMLMLMMMNAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>83,000±32,000*</td>
<td>NMLMLMMMNAN</td>
</tr>
<tr>
<td>(b) 5449</td>
<td>no</td>
<td>Lee</td>
<td></td>
<td>1</td>
<td>8</td>
<td>12,000±900*</td>
<td>NFFDFRFPPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>9</td>
<td>48,000±13,000*</td>
<td>NFFDFRFPPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>8</td>
<td>84,000±10,000*</td>
<td>NFFDFRFPPSS</td>
</tr>
<tr>
<td>(c) 5449</td>
<td>yes</td>
<td>Qiu</td>
<td></td>
<td>1</td>
<td>6</td>
<td>99,000±2000*</td>
<td>NYFDFDFPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>6</td>
<td>84,000±2000*</td>
<td>NYFDFDFPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>7</td>
<td>13,000±1000*</td>
<td>NYFDFDFPSS</td>
</tr>
<tr>
<td>(d) 5449</td>
<td>yes</td>
<td>Lee</td>
<td></td>
<td>1</td>
<td>9</td>
<td>19±8</td>
<td>LFDFDFPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>10</td>
<td>0.3±0.09*</td>
<td>LFDFDFPSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>9</td>
<td>80±28</td>
<td>LFDFDFPSS</td>
</tr>
<tr>
<td>Native</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>0.3±0.07*</td>
<td>NFDHDFSD</td>
<td></td>
</tr>
</tbody>
</table>

Multiple design calculations (a–d) were performed using different sampling resolutions and solvent models. The top three sequences from each calculation and their experimentally measured binding constants are shown. Parts of the sequence identical with the native sequence are highlighted in yellow. (a) Design calculation using a lower-resolution rotamer library. (b) Design calculation without a gradient-based local minimization step. (c) Design calculation using a less accurate generalized Born solvent treatment.43 (d) Design calculation using a high-resolution rotamer library, gradient-based local minimization, and an accurate generalized Born solvation model.42 Sequences are ranked by calculated dissociation energy, allowing 5 kcal/mol destabilization relative to the native sequence for 5449 rotamers per position and 20 kcal/mol destabilization for 2800 rotamers per position. The native sequence was not within the top 100 sequences for design calculation (a), (b), or (c). \( K_d \) measured using the solid-phase radioligand-binding assay. \( K_d \) measured using the centrifugal concentrator assay. The reported error is the standard deviation of three measurements.
The results in this article suggest that the protein design and molecular-mechanics fields can work together on the same PEFs and that future developments in MM-PEFs will be immediately applicable to protein design. Currently, there are active efforts to develop polarizable PEFs that more accurately reproduce the physical characteristics of small molecules and hybrid quantum mechanical/molecular mechanical PEFs that model charge transfer and changes in covalent bonding. It will be exciting to see how these improved energy models will impact the protein design problem.

Materials and Methods

Calculations

Protein structures were predicted using a rotamer-based mean field algorithm. The energy was calculated as the sum of the CHARMM22 molecular-mechanics energy, a generalized Born surface-area solvation energy using a microscopic surface tension of 0.0072 kcal/mol/Å², and deprotonation energy. The most probable mean field structure was then locally minimized using the L-BFGS optimization algorithm to obtain the final structure and energy. The unfolded protein energy was calculated by assuming that the protein backbone adopts an ensemble of random walk conformations in water (see Slovic et al., Zhou, and Supplementary Information). The stability of the protein was calculated as the energy difference between the unfolded protein and the folded unbound protein, and the dissociation energy was calculated as the energy difference between the uncomplexed protein–ligand system and the complexed protein–ligand system. All calculations were performed at 25 °C, pH 7.0, and 100 mM monovalent salt. RBPs were designed using a genetic algorithm that optimized the calculated ribose dissociation energy, given a 5 kcal/mol limit on protein destabilization. The genetic algorithm was initialized with a population of random sequences. Calculations were performed using CNSsolve, TINKER, and custom code written in C++, and run on a Pentium-based Linux cluster.

Protein purification and constructs

RBP without a periplasmic signal peptide was cloned into the Ncol/Xhol sites of pET28a (EMD Biosciences), generating a derivative with a C-terminal His tag. Mutants were made by Kunkel mutagenesis or by QuikChange (Stratagene). Protein was expressed in BL21 DE3 Escherichia coli (Novagen) with 1 mM IPTG for 4 and hybrid quantum mechanical/molecular mechanical PEFs that model charge transfer and changes in covalent bonding. It will be exciting to see how these improved energy models will impact the protein design problem.

Solid-phase radioligand-binding assay

A solid-phase radioligand-binding assay was used to detect binding with $K_d$ values in the high-millimolar range. Nickel-NTA agarose slurry (Qiagen) was washed and resuspended in buffer [20 mM potassium phosphate (pH 7.0) and 100 mM NaCl] to form a 50% (vol/vol) slurry. Twenty microliters of the slurry was mixed with 5 nmol of His$_{r}$-tagged protein and 1.0 μCi of radioligand in a final buffer volume of 50 μl. Following a 30 min equilibration, the mixture was transferred to 0.45 μm centrifugal filter units (no. UFC30HSV5; Millipore) and centrifuged at 12,000 g for 2 min to remove unbound ligand. The resin was washed three times by the addition of 500 μl of 50% ethanol and centrifuged at 12,000 g for 2 min. The bound ligand was eluted with 250 μl of guanidinium HCl and quantified by scintillation counting. Radioligand eluted from a no-protein control was included to account for nonspecific binding to the resin, and a control of 0.5 μCi of radioligand was used to determine counting efficiency. Dissociation constants were calculated as $K_d = \frac{2P}{r} \frac{2L}{r+1}$, where $r$ is the fraction of protein bound to radioligand, $P$ is the initial protein concentration, and $L$ is the initial radioligand concentration. We chose conditions where $P > K_d$ and $r$ fell between 1.2 and 20. The analysis depends on the assumption that water and the ligand cross the membrane at equal rates. This assumption was tested by centrifuging a ribose solution across the membrane in the absence of protein; the specific activities of the retentate and the filtrate were identical to within 4%.

Acknowledgements

We thank J.D. King, R.S. Fenn, and S.R. Pfeffer for help with protein purification, and R.B. Altman, R.L. Baldwin, A.T. Brunger, R.S. Fenn, J.J. Havranek, D. Herschlag, M. Levitt, L.L. Looger, W.L. Martin, V.S. Pande, E.L. Raffauff, T.J. Wandless, R.M. Weisinger, and S.J. Wrenn for many helpful suggestions. This work was supported by National Institutes of Health grants GM068126-01 and 1-DP1-0D000429 to P.B.H. F.E.B. was partially supported by a training grant from the National Institute of General Medical Sciences (5T32 GM07365-28).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.04.001
References


