Pergamon

Bioorganic & Medicinal Chemistry 7 (1999) 851–860

BIOORGANIC & MEDICINAL CHEMISTRY

Estimation of the Binding Affinities of FKBP12 Inhibitors Using a Linear Response Method

Michelle L. Lamb,† Julian Tirado-Rives and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, CT 06520-8107, USA

Received 15 September 1998; accepted 2 November 1998

Abstract—A series of non-immunosuppressive inhibitors of FK506 binding protein (FKBP12) are investigated using Monte Carlo statistical mechanics simulations. These small molecules may serve as scaffolds for chemical inducers of protein dimerization, and have recently been found to have FKBP12-dependent neurotrophic activity. A linear response model was developed for estimation of absolute binding free energies based on changes in electrostatic and van der Waals energies and solvent-accessible surface areas, which are accumulated during simulations of bound and unbound ligands. With average errors of 0.5 kcal/mol, this method provides a relatively rapid way to screen the binding of ligands while retaining the structural information content of more rigorous free energy calculations.

Introduction

The binding protein of the immunosuppressant natural product FK506 (Fig. 1) has been the target of extensive investigation by both biochemical and theoretical techniques during the last 10 years. Discovery of the *cis-trans* peptidyl-prolyl isomerase (PPIase or rotamase) activity of FKBP12 (MW = 12 kDa) led to dissection of the rotamase mechanism[1–4] and hopes for the rapid design of low molecular weight immunosuppressant molecules that inhibited this activity. The crystal structure of FK506 bound to FKBP12 was crucial to this endeavor, as it demonstrated that the peptidomimetic α-ketoamide and pipecolyl portions of the ligand were buried but that much of the macrocycle remained exposed to solvent.[5] However, in later studies rotamase inhibition was found to be insufficient for immunosuppression; the FK506–FKBP12 complex associates with the surface of calcineurin (CN), a serine/threonine phosphatase, and hinders binding of subsequent proteins in the T-cell signaling pathway.[6] The "effector" region of FK506, which contacts CN, is opposite the α-ketoamide-pipecolic acid moiety.[7,8] Thus, the PPIase inhibitors developed through structure-based design efforts (e.g. compounds **1**–**7** in Table 1) formed a set of potential scaffolds for immunosuppressive effector components.[9,10]

Many of these non-immunosuppressive FKBP12 ligands are also able to promote neuronal growth in vitro and in vivo through binding to FKBP12.[11–13] Dose-response studies of neurite outgrowth in chick dorsal root ganglia resulted in an ED_{50} of 0.058 nM for compound **10**, for example.[11] In addition, tethered dimers of molecules similar to **4** have been used to "chemically-induce" dimerization[14] of targeted cellular proteins that had been adapted to include FKBP12 domains.[15] Modification of targets in appropriate signaling pathways can result in apoptosis or the induction of transcription; the prospects of this technology for gene therapy are intriguing.

Consequently, we have used theoretical techniques to probe this class of small molecules (Table 1) to gain insight into the physical basis for differences in FKBP12-binding activity and to evaluate new methods for the calculation of protein–ligand binding affinities. In previous work,[16] *relative* free energies of binding ($\Delta\Delta G_b$) for compounds **2**, **4**, and **8**–**10** were calculated by the free energy perturbation (FEP) technique using a Metropolis Monte Carlo (MC) algorithm for configurational sampling.[17] The parameters and geometry of one ligand were transformed into those of another with these simulations, both in solution and while bound to the protein. At each step of the transformation, the system was brought to equilibrium and the free energy difference relative to the previous step was computed. The difference in binding free energy for the two ligands was then found from the difference in the total free energy change for each (bound and unbound) transformation. While theoretically rigorous and accurate, calculations of this type are computationally demanding

Key words: Monte Carlo; linear response; FKBP12; rotamase inhibitors.

*Corresponding author. Fax: +203-432-6299; e-mail: bill@adrik.chem.yale.edu

† Current address: Dept. of Pharmaceutical Chemistry, University of California, San Francisco, Box 0446, San Francisco, CA 94143-0446.

0968-0896/99/$ - see front matter
PII: S0968-0896(99)00015-2

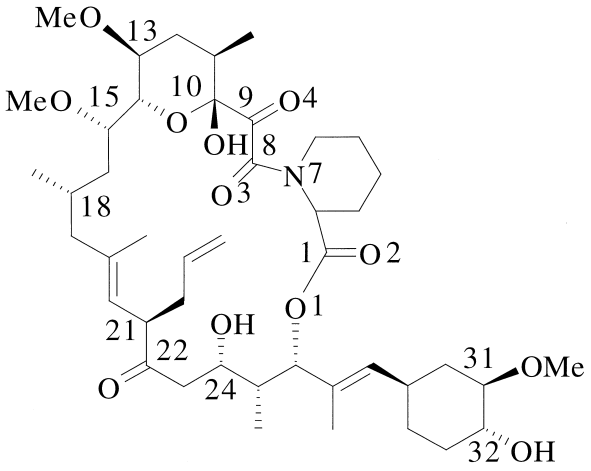


Figure 1. The immunosuppresent drug FK506.

and impractical for a large set of structurally diverse ligands.

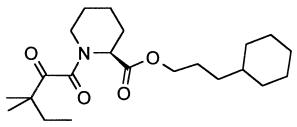
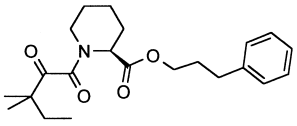
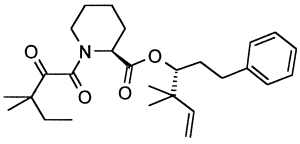
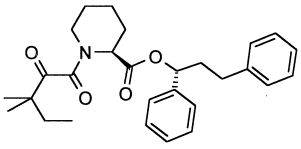
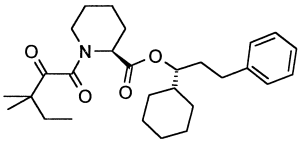
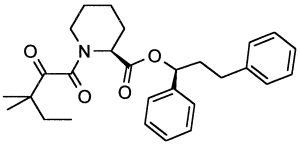
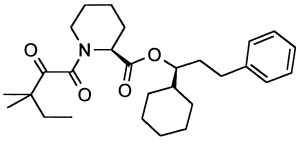
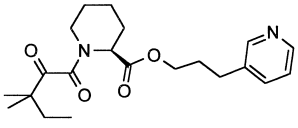
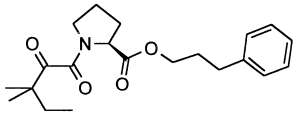
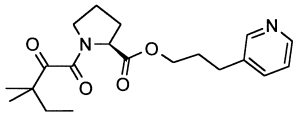
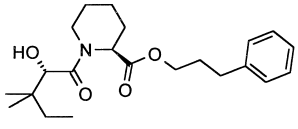
An attractive alternative has emerged in linear interaction energy or linear response techniques (LR), as recently reviewed.[18] Unlike most FEP calculations, *absolute* free energies of binding (ΔG_b) for protein–ligand systems are estimated, and simulations of non-physical states (intermediate steps) are not required. In other respects, the molecular dynamics (MD) or MC simulation protocols are the same.

As first proposed, the linear interaction energy method employs eq (1) for the estimation of binding free energies.[19]

$$\Delta G_b = \beta\langle \Delta E^{Coulomb} \rangle + \alpha\langle \Delta E^{Lennard-Jones} \rangle \qquad (1)$$

These terms represent the change in interaction between a ligand and its environment (solvent and/or protein) upon binding. The electrostatic energy differences are obtained from average Coulombic interaction energies between the ligand and solvent ($E^{Coulomb}_{l-water}$) and the ligand and protein ($E^{Coulomb}_{l-FKBP}$) accumulated during simulations of the bound and unbound ligands. The Lennard–Jones (van der Waals) energy differences are found in an analogous manner. The original value of $\beta = 0.5$ is consistent with analyses of the response of polar solutions to changes in electric fields, such as the charging of an ion in water. Observed linear correlations between the molecular size (surface area, chain length) and solvation free energies of hydrocarbons suggested van der Waals interactions might respond linearly as well. A scale factor $\alpha = 0.161$ was obtained empirically from fitting to experimental binding data for a small set of endothiapepsin inhibitors.[19] Recently, Åqvist and co-workers have advocated the assignment of one of four reduced values of β according to the charge-state or polarity of the ligand and have obtained an appropriate value for α by fitting to data for a larger set of protein–ligand pairs.[20–22] An extended linear response equation,[23,24] in which a cavitation term based on solvent-accessible surface areas (SASA) is added and all parameters are empirically determined, is given in eq (2) below. This model was first applied to the calculation of free energies of hydration for organic solutes; however, optimization of

Table 1. FKBP12 inhibitors[a]

	Compound	$K_{i,app}$, nM
1		186, 170[b]
2		110, 250[b]
3		250
4		10, 17[b]
5		7
6		300–600
7		300
8		165, 130[b]
9		42[b]
10		7.5[b]
11		2300[c]

[a]Rotamase inhibition data taken from refs 9 and 10 unless otherwise noted.
[b]Data from ref 11.
[c]K_i reported for a mixture of R and S isomers.

scale factors to $\beta = 0.131$, $\alpha = 0.131$ and $\gamma = 0.014$ yielded acceptable estimates of thrombin-binding affinity as well.[25] In addition, the extended linear response equation has potential for predictions of other properties important for pharmacological activity, such as the estimation of ligand lipophilicity.[24] It has been observed that the contribution of the ΔSASA term is often nearly constant, so the case in which the simple addition of a constant improves the fit has also been considered [eq (3)].[20,25]

$$\Delta G_b = \beta\langle \Delta E^{Coulomb} \rangle + \alpha\langle \Delta E^{Lennard-Jones} \rangle + \gamma\langle \Delta SASA \rangle \quad (2)$$

$$\Delta G_b = \beta\langle \Delta E^{Coulomb} \rangle + \alpha\langle \Delta E^{Lennard-Jones} \rangle + \gamma \quad (3)$$

In contrast to most proteins studied previously with this technique, FKBP12 has a distinctly hydrophobic binding pocket lined with aromatic residues, and only two intermolecular hydrogen bonds are observed in the crystal structure of inhibitor **4** with the protein.[9] Consequently, it was expected that the electrostatic behavior of these molecules might deviate from linear response and that the van der Waals contributions to binding might be larger than previously observed for other systems. To determine the suitability of the linear response approximation for binding to FKBP12, MC simulations of the bound and unbound states for all 11 inhibitors in Table 1 were performed.

Computational Details

The modeling strategy employed here was consistent with the FEP study described previously,[16] based on the **4**-FKBP12 crystal structure (1FKG)[9] and the OPLS force field.[26–28] (A full listing of parameters for these molecules may be found in ref 29) The first five inhibitors were easily built from **4**. Atoms of the 1-phenyl substituent were simply removed to obtain **2**, and for compound **1**, a united-atom cyclohexyl group[30] was positioned in the plane of the original 3-phenyl ring. The vinyl group of **3** was also represented with united-atom parameters and was positioned at the minimum of the $CH3-C-CH=CH_2$ torsional energy profile (180.0°) and aligned with an edge of the 1-phenyl ring of **4**.[29] One side of this ring was within 3.2 Å of His^{87} and Tyr^{82}, while the other was positioned more than 3.5 Å from Phe^{46} and Glu^{54}. Accordingly, the orientation which maximized hydrophobic contact between the protein and **3** was chosen. Inhibitor **5** incorporated the crystal structure orientations of the cyclohexyl and *tert*-pentyl groups from **5**-FKBP12 (1FKH).[9] In **1** and **5**, the cyclohexyl groups were treated as rigid units, as were all ligand and protein aromatic groups. It was thought that the conformational flexibility within the rings of these substituents would be less important to binding than the overall flexibility of the ligand, and thus sampling was focused accordingly. Starting geometries for the unbound ligand simulations were taken from these initial bound conformations.

All simulations were performed using the MCPRO program and Monte Carlo configurational sampling.[31] The ligands and protein–ligand complexes were solvated with 22 Å spheres of TIP4P water molecules. First, 1×10^6 (1 *M*) configurations of water-only equilibration with preferential sampling of molecules close to the inhibitor was performed, followed by 16 *M* configurations of sampling of the entire system with all solvent molecules sampled uniformly. Next, data was collected for 8 *M* configurations averaged in blocks of 2×10^5 configurations. To ensure convergence, averaging for all unbound ligands was extended to 16 *M* configurations.

To obtain protein-bound structures of **6** and **7**, the final conformations of **4** and **5** above were epimerized within the FKBP12 binding pocket via a slow perturbation protocol. Eight sequential double-wide windows with $\Delta\lambda = 10.0625$ were used. In each window, 4 *M* configurations were sampled to slowly transform between the two ligands in an energetically reasonable way. This procedure was repeated with the free ligands in solution. The simulations of ligands **8**–**11** were started from the final structures of FEP simulations reported previously.[16,29] In each case, energy components were averaged over 8 *M* (bound) and 16 *M* (unbound) configurations of the system.

As was mentioned above, the initial structure for **2** had been generated from the bound conformation of **4** with the 1-phenyl atoms removed. Following an FEP calculation from **4** to **2**, further sampling was carried out within the first windows of both a phenyl→pyridyl FEP[16] and a carbonyl→hydroxyl FEP.[29] The final conformations of these windows were then used to start two additional linear response MC simulations, **2a** and **2b**, respectively. These additional data points provide one gauge of precision.

Solvent-accessible surface areas for the ligands in both aqueous and protein environments were calculated using the SAVOL2 program.[32] This algorithm has been incorporated into MCPRO, and the necessary atomic radii are calculated from the corresponding OPLS σ parameters via 1/2 ($2^{1/6}\sigma$). Using the standard solvent radius for water, 1.4 Å, the SASAs of the ligands were calculated for the structure at the end of each block of MC configurations and were averaged.

Average energy and solvent-accessible surface area differences were fit to the experimental binding data (Table 1) to obtain linear response parameters α, β, and γ according to eqs (1)–(3). This procedure was performed with a Simplex-based algorithm. As inhibition data for the majority of the ligands in this set has been reported by Holt and co-workers,[9,10] in cases where two values have been measured the values from these authors were used for fitting.

Results and Discussion

Average intermolecular energy components and SASAs

Each ligand and protein–ligand complex was solvated with a sphere of explicit water molecules and sampled as described above. Average energy components and

ligand SASAs accumulated during the simulations are reported in Table 2. As has been observed with thrombin inhibitors, the ligand–solvent Coulombic energy components, $E_{l-water}^{\text{Coulomb}}$, fluctuate the most during the simulations.[25] However, both the magnitude of the energy contribution and the magnitude of the fluctuation are reduced compared to the positively charged or zwitterionic protease inhibitors, as expected for the neutral ligands of this set. All of the ligands in solution have undergone hydrophobic collapse relative to their bound conformations but to differing extents, as significant flexibility is observed in the propyl side chain. The Lennard–Jones interactions for the α-ketoamide molecules in solution scale generally with molecular size; **1** and **2** (and **8**–**10**) have the least favorable average energies, followed by compound **3**, and finally there is a small range of energies for **4**–**7**. The solvent accessible surface area of amide carbonyl O3 is generally ca. 10 Å^2 less than that of the keto (O4) or ester (O2) atoms, and usually only one water molecule is found to interact with this atom. This is consistent with molecular dynamics results for *trans*-FK506.[29] Often, O4 interacts strongly with one water molecule (H-O4 distance, ca. 1.8 Å), while a second molecule hovers 0.5 Å further away. Also, water molecules are found to interact with the faces of aromatic rings. In the case of **2**, an unusual long-lived solvation pattern is noted. The water molecule that hydrogen bonds to O3 further associates with one directed into the center of the less-accessible face of the phenyl ring (Fig. 2).

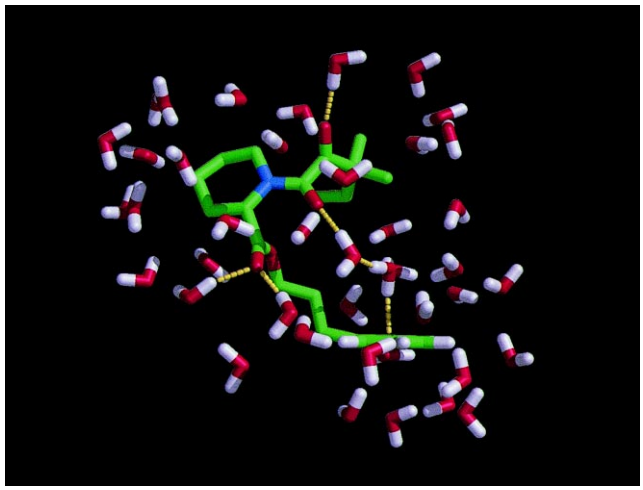


Figure 2. The hydrogen bonding network in **2**, with waters that bridge between the amide O3 and phenyl ring.

Typical of many FKBP12-ligand complexes,[33] the crystal structure of **4**-FKBP12 contains contacts between O4 and aromatic hydrogens of Tyr^{26}, Phe^{36}, and Phe^{99}, with the pipecolyl ring sitting over Trp^{59}.[9] The 3-phenylpropyl moiety binds in the solvent-exposed FK506-cyclohexyl groove of FKBP12 between Ile^{56} and Tyr^{82}, and these residues form hydrogen bonds with the ester and amide carbonyl oxygens of the ligand. The 1-phenyl substituent interacts with Phe^{46} and the tertiary pentyl group of the inhibitor. The two crystallographic intermolecular hydrogen bonds are maintained throughout all of the FKBP12-ligand simulations, and none of the ligands deviates significantly from the original orientation

Table 2. Average interaction energies (kcal/mol) and solvent-accessible surface areas of the inhibitors (Å^2) from the aqueous and FKBP12 MC simulations[a]

Compound	$E_{l-water}^{\text{Coulomb}}$	$E_{l-water}^{L-J}$	$E_{l-FKBP}^{\text{Coulomb}}$	E_{l-FKBP}^{L-J}	SASA[b]
1 aq	−23.16(0.32)	−35.96(0.15)			665.4(1.2)
1FKBP	−0.24(0.15)	−15.09(0.12)	−19.14(0.16)	−39.53(0.14)	202.1(1.6)
2 aq	−31.28(0.46)	−34.59(0.21)			653.9(1.9)
2a aq	−29.91(0.35)	−33.90(0.16)			629.3(2.3)
2b aq	−27.11(0.38)	−33.26(0.16)			620.6(2.3)
2FKBP	−5.15(0.31)	−13.74(0.09)	−19.73(0.21)	−40.29(0.13)	202.9(5.1)
2aFKBP	−2.14(0.19)	−13.99(0.12)	−21.56(0.16)	−38.01(0.17)	222.4(6.8)
2bFKBP	−1.62(0.19)	−13.95(0.14)	−20.97(0.20)	−37.92(0.21)	188.6(7.0)
3 aq	−27.68(0.36)	−36.52(0.16)			680.9(1.2)
3FBKP	−1.41(0.20)	−15.31(0.13)	−21.15(0.20)	−44.06(0.15)	221.2(3.0)
4 aq	−31.89(0.42)	−39.63(0.22)			715.1(1.4)
4FKBP	−5.30(0.38)	−17.84(0.11)	−21.25(0.15)	−42.73(0.19)	248.8(2.5)
5 aq	−28.43(0.50)	−40.12(0.19)			728.3(1.8)
5FKBP	−6.49(0.19)	−16.54(0.10)	−20.58(0.14)	−44.00(0.16)	244.9(2.1)
6 aq	−32.84(0.55)	−38.35(0.12)			712.2(1.6)
6FBKP	−8.00(0.19)	−18.10(0.13)	−17.26(0.15)	−38.46(0.20)	232.4(3.0)
7 aq	−29.64(0.34)	−38.48(0.15)			688.3(1.9)
7FKBP	−3.92(0.23)	−16.10(0.15)	−17.09(0.18)	−43.65(0.21)	282.3(7.1)
8 aq	−30.53(0.43)	−32.75(0.18)			618.4(1.8)
8FBKP	−13.51(0.48)	−12.03(0.15)	−18.03(0.16)	−38.80(0.19)	176.0(4.7)
9 aq	−29.88(0.50)	−31.67(0.17)			624.9(2.2)
9FBKP	−1.43(0.19)	−14.04(0.09)	−23.59(0.16)	−39.40(0.17)	179.3(6.4)
10 aq	−33.45(0.40)	−32.89(0.22)			630.8(2.2)
10FKBP	−11.95(0.31)	−12.82(0.10)	−20.02(0.18)	−38.77(0.15)	183.2(6.5)
11 aq	−42.57(0.43)	−31.24(0.21)			626.7(1.4)
11FKBP	−2.72(0.34)	−15.19(0.10)	−18.23(0.15)	−38.61(0.13)	185.5(4.4)

[a]The standard error of the means is given in parentheses.
[b]Calculated from structures saved every 2×10^5 configurations, $N=40$ (8 *M*) or 80 (16 *M*). (Standard deviations range from 10–50 Å^2.)

of **4**. In general, the aromatic rings of the inhibitors are oriented perpendicularly to the ring of Tyr[82], although a clearly T-shaped interaction is not dominant. Consistent with the hydrophobic binding pocket, the largest contribution to the protein–ligand interaction energy in all cases comes from the Lennard–Jones terms. In addition, the most favorable van der Waals interactions with FKBP12 are observed for the more hydrophobic **3**, **5**, and **7** compared to their more aromatic or smaller counterparts.

An estimate of the precision of the simulations was determined from the three simulations of **2** in solution. One model was derived directly from the **4**-FKBP12 crystal structure, while the others (**2a** and **2b**) were obtained through earlier FEP simulations from **4**. In solution, there is a 4 kcal/mol range in the $E^{\mathrm{Coulomb}}_{l-water}$ components and a 2 kcal/mol range in $E^{L-J}_{l-water}$. Variations among the average energy components reported for all simulations of **2**-FKBP12 are on the order of 2–3 kcal/mol. It must be noted that the terms from simulations **2a** and **2b** are more similar to each other than to those from the original **2** simulation.

As energy and SASA *differences* between the protein and aqueous environments are the quantities that are scaled to estimate binding affinity, these values are recorded in Table 3. Two of the lowest affinity inhibitors, **3** and **11** have the largest van der Waals energy differences between bound and unbound ligands. Ligands **6**, **7**, and **11** have the most unfavorable $\Delta E^{\mathrm{Coulomb}}$, while only **8** has a net favorable electrostatic interaction upon binding. Approximately 450 Å^2 of each ligand is buried upon binding to FKBP12. This represents 65% of the surface of inhibitor **4**, for example. The largest change is noted for inhibitors **5** and **6** (ca. 480 Å^2 buried), while the smallest difference is found for **7** (ca. 410 Å^2).

Optimization of the LR equations

The energies and surface areas of Table 2 were used first with previously determined scaling parameters to compute FKBP12-binding affinities. The original parameters of Åqvist[19] do not describe binding to this protein well at all; the free energies of binding are underestimated with an average unsigned error ($<|\mathrm{error}|>$) of 9.3 kcal/mol. Results with the thrombin-derived parameters are of more appropriate magnitude (−9.3 (**5**) to −6.3 (**11**) kcal/mol, $<|\mathrm{error}|> = 1.4$ kcal/mol), although both the set of atomic radii for SASAs and the all-atom representation of the thrombin inhibitors differed from the study presented here. However, to improve the correspondence with experiment for FKBP12, new values for α, β, and γ were found by fitting the average energy and surface area differences to the experimental binding free energies. The results for the various models investigated are summarized in Table 4.

The first step was to employ eq (1) with $\beta = 0.5$ and derive an appropriate value for α, as was done originally for endothiapepsin.[19] This model resulted in $\alpha = 0.626$ and yielded free energies which deviated significantly from experiment. In particular, the maximum unsigned error for model 1 was 4.4 kcal/mol, which diminishes its predictive value given that the range of experimental binding affinities is 3.4 kcal/mol. Furthermore, compound **3**, a poor inhibitor, was predicted to bind as well as the highest affinity ligand, **8**.

As expected, the linear response assumption for electrostatic energies, $\beta = 0.5$, does not appear to hold well for binding to FKBP12. When the value of β is set based on ligand composition[20] and α derived empirically with eq (1) (model 2) or both β and α treated as free parameters (model 3), average unsigned error and RMS to experiment were improved but the maximum errors are still greater than 2kcal/mol. With model 2, the α-ketoamide ligands required $\beta = 0.43$; **11** called for $\beta = 0.37$ due to its single hydroxyl substituent.[20]

Fitting the data with three parameters yielded an RMS deviation of 0.7 kcal/mol, whether or not the SASA difference was included explicitly (eq (2) versus eq (3)). Values of $\beta = 0.139$, $\alpha = 0.194$, and $\gamma = 0.0145$ (model 4) ranked the ligands in a qualitatively reasonable way

Table 3. Calculated energy and surface-area differences[a] with representative binding affinities

Compound	$\Delta E^{\mathrm{Coulomb}}$	ΔE^{L-J}	ΔSASA	ΔG_b, kcal/mol			
				model 2	model 4	model 6	expt[b]
1	3.8	−18.7	−463.3	−9.6	−9.8	9.7	−9.2(−9.2)
2	6.4	−19.4	−451.0	−8.9	−9.4	−9.4	−9.5(−9.0)
2a	6.2	−18.1	−406.9	−8.2	−8.5	−8.6	−9.5(−9.0)
2b	4.5	−18.6	−432.0	−9.2	−9.2	−9.3	−9.5(−9.0)
3	5.1	−22.7	−459.7	−11.5	−10.4	(−10.9)[c]	−9.0
4	5.3	−20.9	−466.3	−10.2	−10.1	−10.3	−10.9(−10.6)
5	1.4	−20.4	−483.4	−11.6	−10.8	−10.9	−11.1
6	7.6	−18.2	−479.8	−7.6	−9.4	−9.0	−8.7[d]
7	8.6	−21.3	−406.0	−9.0	−8.8	−9.3	−8.9
8	−1.0	−18.1	−442.4	−11.3	−10.1	−10.2	−9.2(−9.4)
9	4.8	−21.8	−445.6	−11.0	−10.0	−10.5	−10.1
10	1.5	−18.7	−447.6	−10.6	−9.9	−10.0	−11.1
11	21.6	−22.6	−441.2	−5.5	−7.8	−7.8	−7.7

[a]Units: kcal/mol and Å^2, respectively.
[b]References given in notes to Table 1, $\Delta G = RT ln K_i$.
[c]Compound **3** not included in the derivation of this model.
[d]Free energy estimated using $K_i = 450$ nM.

Table 4. Summary of the α, β, and γ parameters determined by fitting to the experimental ΔG_b data[a]

Model	eq	β	α	γ[b]	max. |error|[c]	< |error| >[c]	RMS to expt[c]
1	1	0.50[d]	0.626		4.39	1.24	1.75
2	1	0.43[d,e]	0.599		2.49	1.01	1.25
3	1	0.201	0.536		2.22	0.71	0.90
4	2	0.139	0.194	0.0145	1.39	0.57	0.72
		0.138 ± 0.016	*0.191 ± 0.006*	*0.0146 ± 0.0003*			
5	3	0.145	0.134	−7.75	1.12	0.55	0.69
6	2	0.176	0.348	0.0084	1.09	0.47	0.58
		0.174 ± 0.013	*0.348 ± 0.041*	*0.0085 ± 0.0002*			
7	3	0.180	0.328	−4.21	1.12	0.47	0.59

[a]Energy components and SASAs from MC simulations of **1**, **2**, **2a**, **2b**, and **3**–**11** were fit unless otherwise noted in the text. Cross-validation values with uncertainties given in italics.
[b]eq (2), kcal/mol·Å^2, eq (3), kcal/mol.
[c]Maximum unsigned error, average unsigned error, and RMS deviation to experiment in kcal/mol.
[d]Fixed value.
[e]β = 0.37 for hydroxyl-containing ligand **11** (see text).

(Table 3); again **3** was predicted to bind too well and was responsible for the largest deviation from experiment (1.4 kcal/mol). In performing the "leave-one-out" cross-validation of these scaling parameters, it became clear that compound **3** was the notable outlier. Thus, the fitting was repeated without including data for this compound, resulting in optimal parameters β = 0.176, α = 0.348, and γ = 0.0084 (model 6). The predicted order of affinities for the remaining ligands improved slightly, and there was a corresponding decrease in RMS deviation to 0.6 kcal/mol. The average unsigned error in this case improves by 0.1 kcal/mol to 0.5 kcal/mol. A plot of calculated versus experimental binding free energies based on model 6 is shown in Figure 3. Additional justification for setting aside compound **3** may be found when γ is a constant term in the linear response model. When **3** is included in the data set (model 5), the predicted binding energy results largely from the constant rather than the scaled energy terms (Table 4). However, when the fitting is performed without **3** (model 7), the value of γ (−4.21 kcal/mol) accurately reflects the nearly constant value of γ*ΔSASA found with model 6.

Overall, with or without **3**, the LR calculations are ordering the binding affinities well. In view of the statistical noise in the simulations and the experimental uncertainties, average errors of 0.5 kcal/mol are as good as can be hoped for.

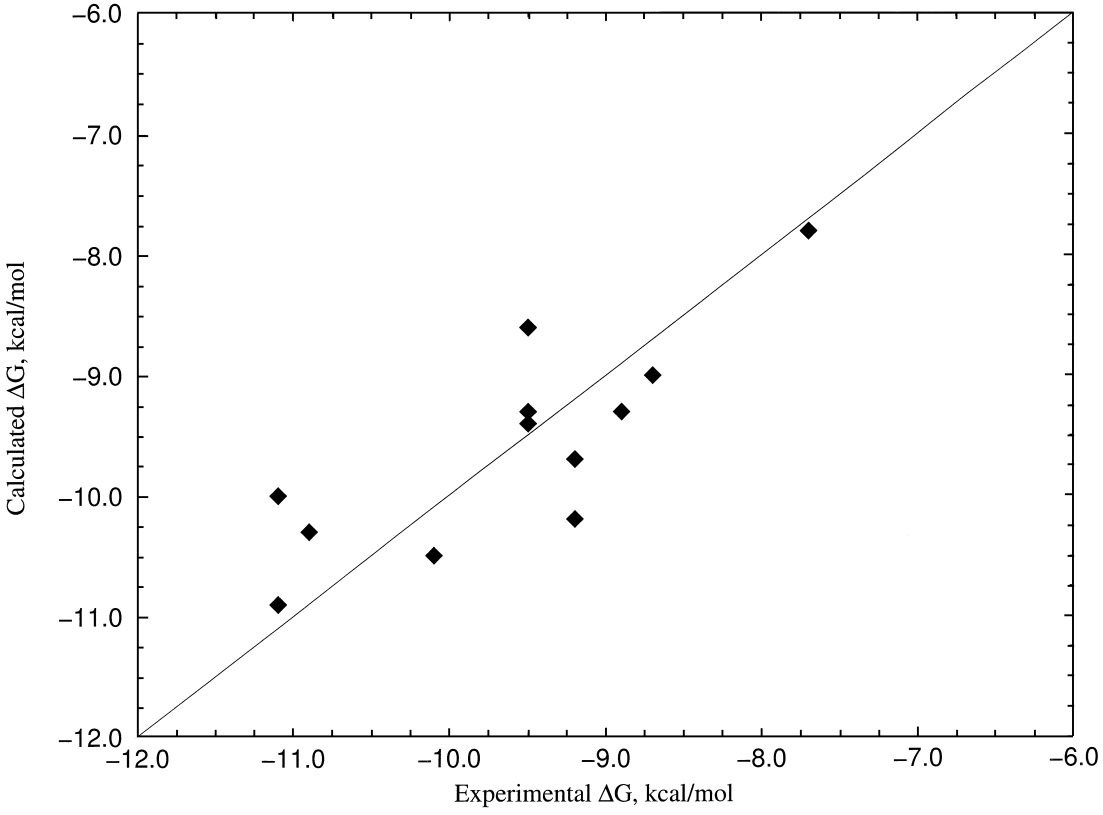


Figure 3. Calculated versus experimental ΔG_b plotted for all ligands. The binding free energies were estimated using optimal parameters β = 0.176, α = 0.348, and γ = 0.0084 and eq (2). Compound **3**, not included in this model, is not shown.

Structural observations

With reasonable reproduction of the observed binding free energies, we now turn to structures obtained during the simulations for possible insights into the origins of the predicted affinities. The benchmark example in our earlier work was the difference between the 3-phenylpropyl compound, **2**, and the (*R*)-1,3-diphenylpropyl, **4**. The reduced binding for **2** was attributed to a loss of hydrophobic contacts and specific aryl C−H...N,O interactions with the protein upon removal of the 1-phenyl substituent.[16] The relative free energy of binding computed with the LR method for any simulation of **2** and **4** ($\Delta\Delta G_b = 0.9$–1.8 kcal/mol) is in excellent agreement with the experimental and FEP-calculated values of 1.4 kcal/mol. Appropriately, the conformations of **2** and **4** bound to FKBP12 resemble those reported previously, although at the end of the simulations, the phenyl rings in these complexes are found nearer to Tyr^{82} than Glu^{54} (not shown). The mobility of the ligands within that region of the protein is seen throughout the MC simulations.

Among the final structures of the bound simulations of compound **2**, the two from the FEP-generated structures (**2a** and **2b**) are more similar to one another than to that of **2**, especially with regard to the dicarbonyl O=C−C=O dihedral angle, which is −108° in **2a** and **2b** and −125° in **2**. However, the electrostatic energy differences for **2** and **2a** are very similar (6.4 and 6.2 kcal/mol, respectively), while **2b** yields 4.5 kcal/mol. There is also variation in the ΔSASA terms for each model of **2** that parallels the trend in contribution to binding from ΔE^{L-J} for these ligands, in that **2** is most favorable, **2b** is intermediate, and **2a** is the least favorable. As a result, the predicted binding free energies for **2** and **2b** agree well with the experimental values (Table 3), while **2a** is estimated to bind 1 kcal/mol less well.

As one would expect, compound **1**, which contains no aromatic ring, has the least favorable average electrostatic interaction of all ligands in solution (−23.16 kcal/mol); however, its van der Waals interactions with solvent are comparable to those of **2**. The flexible propyl side chain allows the cyclohexyl ring of **1** to pack against the side chain of Ile^{56} (Fig. 4). It is also observed that much less water is found within 3 Å of the cyclohexyl group than the phenyl group, in both bound (Fig. 4) and unbound structures.

One of the best inhibitors, **5**, is notable for its packing among aromatic side chains of the protein. This is found both in its crystal structure with FKBP12[9] and in the structures modeled here. One H^{ε} of Tyr^{82} is 2.8 Å from the ligand phenyl ring center and this protein side chain fills the space between the *tert*-pentyl and phenyl groups (Fig. 5). Furthermore, the pipecolyl and cyclohexyl rings are parallel to one another and separated by Phe^{46}. The cyclohexyl moiety has more contact with Phe^{46} than does the analogous aromatic group of **4**. Without these contacts in solution, the cyclohexyl ring is free to rotate ca. 30° into a position still parallel to but displaced from the pipecolyl ring.

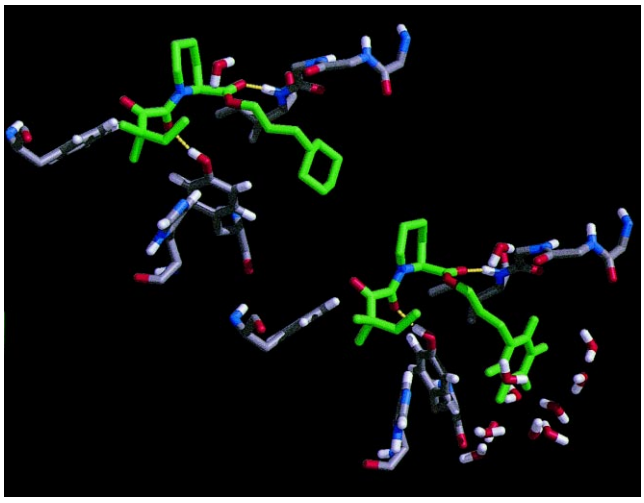


Figure 4. Final simulated structures of **1**-FKBP12 and **2**-FKBP12. Protein residues and solvent molecules within 3 Å of the ligands are displayed.

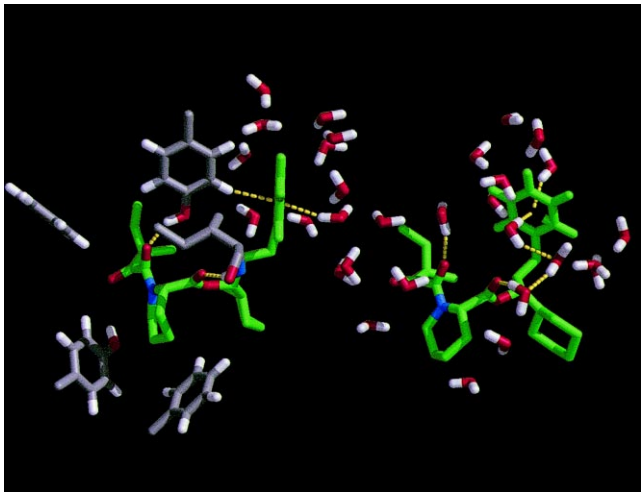


Figure 5. Snapshots from bound and unbound simulations of **5**.

The experimental and calculated binding affinities of **6** and **7** are much reduced from their (*R*)- counterparts, **4** and **5**; the scaled electrostatic and van der Waals energy differences from model 6 each contribute ca. 1 kcal/mol to the less favorable ΔG_b for these molecules. Thus, it appears that the bound conformations found for these ligands through epimerization of their counterparts (Fig. 6) are reasonable models. When bound to the protein, the two phenyl rings of **6** form a "parallel-stacked and displaced" structure with one another; and the distance between ring centers is 4.7 Å. This alignment of aromatic rings has also been noted previously for benzene dimer in solution[27] but is not seen in the simulations of **4**-FKBP12. In **7**-FKBP12, the $Tyr^{82}H^{\eta}$-O3 hydrogen bond is the shortest of all complexes, 1.7 Å, and the interaction between Ile^{56}H and O2 is the longest at 2.6 Å. Unlike the 1-phenyl group of **6**, the cyclohexyl ring in **7** interacts with Phe^{46} more closely than with its own *tert*-pentyl moiety.

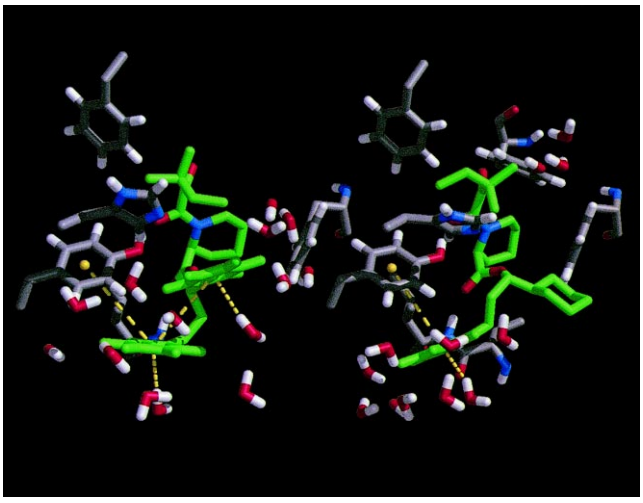


Figure 6. Final simulated structures of **6**-FKBP12 and **7**-FKBP12. The aromatic rings of **6** are nearly parallel, and water molecules point toward both ring centers. The distances between the ring centers are also highlighted.

In the simulations of the pyridyl ligands, **8** and **10**, differences in electrostatic interactions with solvent compared to their phenyl analogues (**2** and **9**) are comparable for both pipecolyl and prolyl molecules. The energy components confirm a more favorable interaction with water for pyridyl ligands and demonstrate continued interaction of the nitrogen atom with solvent upon binding to FKBP12, as was observed previously.[16] From Table 3 (model 6), the affinity of the pyridyl-containing ligand **8** is overestimated by the LR method, which results in a predicted $\Delta\Delta G_b$ relative to **2** of ca. −1 kcal/mol. In solution, the pyridyl ring of **8** has flipped 180° from its position at the end of the FEP simulation, which may contribute to the deviation from the free energy perturbation result of +0.8 kcal/mol for this pair of ligands.[16] There is uncertainty in the relative experimental $\Delta\Delta G_b$ values as well, as results from two laboratories differ as to which molecule has the higher affinity. However, the linear response estimate for the prolyl pair yields $\Delta\Delta G_b = 0.5$ kcal/mol for **10** relative to **9**, consistent with the value computed with the free energy perturbation method (0.9 kcal/mol). In this case, however, the available experimental binding data favors binding of **10** by 1 kcal/mol. The prediction of the affinity of prolyl- compared to pipecolyl-containing ligands is in qualitative accord with the experimental difference of −1.1 kcal/mol. For **9** and **2a**, the linear response calculations result in $\Delta\Delta G_b = -1.9$ kcal/mol, which may be further compared with the FEP result, $\Delta\Delta G_b = -1.7$ kcal/mol.

Reduction of the C9 carbonyl is correctly predicted to reduce binding by all three forms of the linear response equation shown in Table 3. With model 6, the predictions for the $\Delta\Delta G_b$ of **11** compared to **2** range from 0.8–1.6 kcal/mol (expt $\Delta\Delta G_b \approx 1.3$–1.8 kcal/mol). This ligand was originally of interest due to potentially unfavorable steric contacts between the hydroxyl group and the aromatic residues which surround O4, which would later be relieved by a compensating mutation in the protein. This "bump-hole" design could be used to avoid unproductive binding of ligand by wild-type FKBP12 in the cytoplasm in a chemically-induced dimerization scheme.[34,35] The electrostatic solute–solvent energy is 12–15 kcal/mol more favorable for the hydroxyl-containing **11** than its α-ketoamide parent, which is consistent with the preferential solvation of isopropanol compared with acetone[36] and the ability of alcohols to be incorporated into the hydrogen bonding networks of water.[37] The last structure in the unbound simulation has the hydroxyl hydrogen pointing towards the pipecolyl ring (H−O4−C9−C8 dihedral angle, −22°; H−O3 distance, 2.4 Å). It does make the expected two hydrogen bonds to water molecules, one as a donor, one as an acceptor. When bound to the protein, however there are no solvent molecules directly interacting with the ligand in the O4 binding pocket. The hydroxyl hydrogen of the ligand is positioned directly between and below the H^{δ} and H^{ε} of Phe[36]; these aromatic hydrogens interact with O4 (Fig. 7). The H−O4−C9−C8 dihedral angle is 54° in this conformation, but the H−O3 distance seen in solution is maintained through an H−O4−C9 bond angle of less than 109°. The increased charge of the hydroxyl oxygen ($q_O = -0.700$) compared to the carbonyl ($q_O = -0.424$) appears to draw the Tyr[26] hydroxyl group away from Asp[37] and toward O4. This could be expected to enhance binding of **11**, however, the overall change in electrostatic energy is most unfavorable when compared to results from simulations of **2** (Table 3). Clearly the preferential solvation of the unbound ligand is primarily responsible for the low predicted binding affinity of **11**.

All linear response models overestimated the binding affinity of **3**, a relatively poor rotamase inhibitor. Looking back at the average energy differences upon binding (Table 3), the Lennard–Jones contribution to binding for this molecule is more favorable than for all other α-ketoamides, while its other contributions are similar to compound **4**. In solution, the packing of the

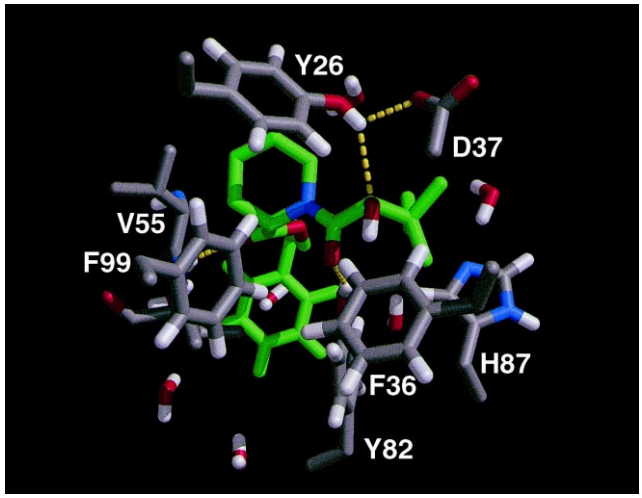


Figure 7. The final structure of the **11**-FKBP12 complex. The hydroxyl group of the ligand orients that of Tyr[26] away from the carboxylate of Asp[37]. A water molecule is also observed to interact with the back face of the ligand phenyl ring in this snapshot.

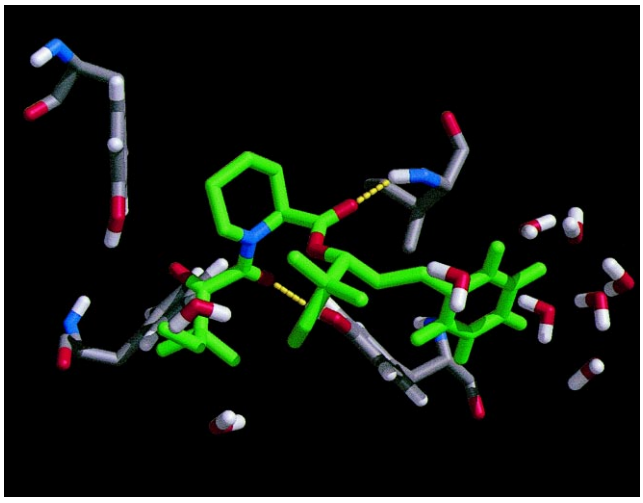


Figure 8. Final simulated structure of the **3**-FKBP12 complex.

allyl substituent against the *tert*-pentyl group is complementary and contributes to the compact core of the molecule. Of all the ligands, only the pipecolyl ring of **3** has not partially inverted compared to its bound conformation, as discussed in ref 16. The conformation of **3** bound to FKBP12 (Fig. 8) otherwise resembles its solution conformation. It may be noted that the homoprolyl C^{β} atom (C3) is separated from $Trp^{59}C^{\varepsilon 2}$ by 3.9 Å, more than for all other ligands. At close range this is an unfavorable contact,[16] so the distance from Trp^{59} may correlate with the most favorable E_{l-FKBP}^{L-J} for this system among the inhibitors (Table 2). In addition, the neutral CH_2 united-atom points in the same direction as does an aromatic hydrogen of **4** ($q_H = 0.115$), but this atom is also in close contact with the neutral atoms of the *tert*-pentyl portion of the ligand. It was thought that an all-atom representation of the vinyl group might improve the model, as interactions more similar to those of the 1-aryl ligands could be reproduced. However, repeating the simulations of **3** as suggested did not improve the predicted binding affinity (data not shown). It may be that the true bound conformation of this ligand has not yet been observed in the simulations. Free energy perturbation calculations could be used to investigate this anomaly further.

Conclusions

The linear response method is a valuable approach for the estimation of protein–ligand binding affinities from molecular simulations that include explicit water molecules. Only physically-relevant states are sampled, an attractive advantage over free energy perturbation calculations. As noted before, the scaling parameters for a linear response equation depend on the force field employed, and effects from these potentials are implicit in the derived LR model. In the applications investigated to date that employ the OPLS force fields, the optimal value of β is found to be reduced from the linear response assumption of 0.5.[23–25] For ligand-binding to FKBP12, the greatest variability has been observed in the average electrostatic energies accumulated during MC simulations, while the Lennard–Jones terms generally correlated with molecular size and packing, as expected. The present LR calculations for eleven ligands reasonably reproduce experimental free energies of binding with average errors of 0.5–0.6 kcal/mol. Notably, alcohol **11** was correctly predicted to have much lower affinity for the protein than the remaining compounds, and the most hydrophobic compound, **5**, was correctly predicted to be the strongest binder. With advances in inexpensive parallel computing environments,[38] it has become practical to carry out the multiple molecular simulations and the extensive sampling required for convergence of the intermolecular energy terms for this method. While the complementary technique of free energy perturbation theory is appropriate for detailed ligand optimization, the linear response method shows promise as a more rapid tool for initial screening of ligand binding affinities for structure-based drug design.

Acknowledgements

Gratitude is expressed to the National Institutes of Health for financial support.

References

1. Fischer, G. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1415–1436.
2. Rosen, M. K.; Schreiber, S. L. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 384–400.
3. Orozco, M.; Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1993**, *32*, 12864–12874.
4. Fischer, S.; Michnick, S.; Karplus, M. *Biochemistry* **1993**, *32*, 13830–13837.
5. Van Duyne, G.; Standaert, R.; Karplus, P.; Schreiber, S.; Clardy, J. *Science* **1991**, *252*, 839–842.
6. Liu, J.; Farmer Jr., J. D.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. *Cell* **1991**, *66*, 807–815.
7. Griffith, J. P.; Kim, J. L.; Kim, E. E.; Sintchak, M. D.; Thomson, J. A.; Fitzgibbon, M. J.; Fleming, M. A.; Caron, P. R.; Hsiao, K.; Navia, M. A. *Cell* **1995**, *82*, 507–522.
8. Kissinger, C. R.; Parge, H. E.; Knighton, D. R.; Lewis, C. T.; Pelletier, L. A.; Tempczyk, A.; Kalish, V. J.; Tucker, K. D.; Showalter, R. E.; Moomaw, E. W.; Gastinel, L. N.; Habuka, N.; Chen, X.; Maldonado, F.; Barker, J. E.; Bacquet, R.; Villafranca, J. E. *Nature* **1995**, *378*, 641–644.
9. Holt, D. A.; Luengo, J. I.; Yamashita, D. S.; Oh, H.-J.; Konialian, A. L.; Yen, H.-K.; Rozamus, L. W.; Brandt, M.; Bossard, M. J.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. *J. Am. Chem. Soc.* **1993**, *115*, 9925–9938.
10. Holt, D. A.; Konialian-Beck, A. L.; Oh, H.-J.; Yen, H.-K.; Rozamus, L. W.; Krog, A. J.; Erhard, K. F.; Ortiz, E.; Levy, M. A.; Brandt, M.; Bossard, M. J.; Luengo, J. I. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 315–320.
11. Hamilton, G.; Steiner, J. *Curr. Pharm. Des.* **1997**, *3*, 405–428.
12. Gold, B. G.; Zeleny-Pooley, M.; Wang, M. S.; Chaturvedi, P.; Armistead, D. M. *Exp. Neurol.* **1997**, *147*, 269–278.
13. Armistead, D. M.; Badia, M. C.; Deininger, D. D.; Duffy, J. P.; Saunders, J. O.; Tung, R. D.; Thomson, J. A.; DeCenzo, M. T.; Futer, O.; Livingston, D. J.; Murcko, M. A.; Yamashita, M. M.; Navia, M. A. *Acta Crystallogr.* **1995**, *D51*, 522–528.

14. Crabtree, G. R.; Schreiber, S. L. *Trends Biochem. Sci.* **1996**, *21*, 418–422.
15. Amara, J. F.; Clackson, T.; Rivera, V. M.; Guo, T.; Keenan, T.; Natesan, S.; Pollock, R.; Yang, W.; Courage, N. L.; Holt, D. A.; Gilman, M. *Proc. Natl. Acad. Sci. USA*. **1997**, *94*, 10618–10623.
16. Lamb, M. L.; Jorgensen, W. L. *J. Med. Chem.* **1998**, *41*, 3928–3939.
17. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
18. Lamb, M. L.; Jorgensen, W. L. *Curr. Opin. Chem. Biol.* **1997**, *1*, 449–457.
19. Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Engng.* **1994**, *7*, 385–391.
20. Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27–35.
21. Marelius, J.; Graffner-Nordberg, M.; Hansson, T.; Hallberg, A.; Åqvist, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 119–131.
22. Marelius, J.; Hansson, T.; Åqvist, J. *Int. J. Quantum Chem.* **1998**, *69*, 77–88.
23. Carlson, H. A.; Jorgensen, W. L. *J. Phys. Chem.* **1995**, *99*, 10667–10673.
24. McDonald, N. A.; Carlson, H. A.; Jorgensen, W. L. *J. Phys. Org. Chem.* **1997**, *10*, 563–576.
25. Jones-Hertzog, D. K.; Jorgensen, W. L. *J. Med. Chem.* **1997**, *40*, 1539–1549.
26. Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
27. Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768–4774.
28. Jorgensen, W. L.; McDonald, N. A. *J. Mol. Struct. (Theochem)* **1998**, *424*, 145–155.
29. Lamb, M. L. Ph.D. Dissertation, Yale University, May 1998.
30. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638–6646.
31. Jorgensen, W. L. *MCPRO, Version 1.5*, **1997**, Yale University, New Haven, CT.
32. Skell, J. M.; Pearlman, R. S. *SAVOL2*, **1988**, University of Texas, Austin, TX.
33. Babine, R. E.; Bender, S. C. *Chem. Rev.* **1997**, *97*, 1359–1472.
34. Belshaw, P. J.; Schoepfer, J. G.; Liu, K.-Q.; Morrison, K. L.; Schreiber, S. L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2129–2132.
35. Pierce, A. C.; Jorgensen, W. L. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1466–1469.
36. Cramer, C. J., Truhlar, D. G. *J. Comput-Aided Mol. Des.* **1992**, *6*, 629–666.
37. Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512–9521.
38. Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1996**, *117*, 1385–1386.