JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

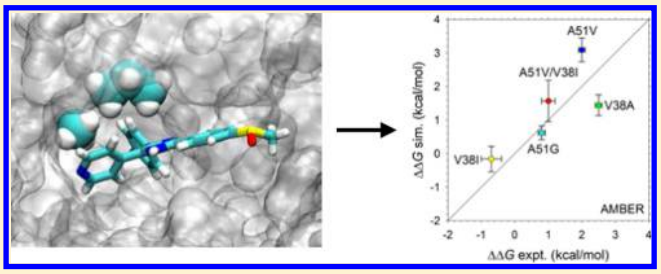
Accurate Calculation of Mutational Effects on the Thermodynamics of Inhibitor Binding to p38α MAP Kinase: A Combined Computational and Experimental Study

Shun Zhu, Sue M. Travis, and Adrian H. Elcock*

Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242, United States

Supporting Information

ABSTRACT: A major current challenge for drug design efforts that are focused on protein kinases is the development of drug resistance caused by spontaneous mutations in the kinase catalytic domain. The ubiquity of this problem means that it would be advantageous to develop fast, effective computational methods that could be used to determine the effects of potential resistance-causing mutations before they arise in a clinical setting. With this long-term goal in mind, we have conducted a combined experimental and computational study of the thermodynamic effects of active-site mutations on a well-characterized and high-affinity interaction between a protein kinase and a small-molecule inhibitor. Specifically, we developed a fluorescence-based assay to measure the binding free energy of the small-molecule inhibitor, SB203580, to the p38α MAP kinase and used it measure the inhibitor's affinity for five different kinase mutants involving two residues (Val38 and Ala51) that contact the inhibitor in the crystal structure of the inhibitor−kinase complex. We then conducted long, explicit-solvent thermodynamic integration (TI) simulations in an attempt to reproduce the experimental relative binding affinities of the inhibitor for the five mutants; in total, a combined simulation time of 18.5 μs was obtained. Two widely used force fields—OPLS-AA/L and Amber ff99SB-ILDN—were tested in the TI simulations. Both force fields produced excellent agreement with experiment for three of the five mutants; however, simulations performed with the OPLS-AA/L force field produced qualitatively incorrect results for the constructs that contained an A51V mutation. Interestingly, the discrepancies with the OPLS-AA/L force field could be rectified by the imposition of position restraints on the atoms of the protein backbone and the inhibitor without destroying the agreement for other mutations; however, the ability to reproduce the experiment was dependent on the strength of the restraints' force constant. Imposition of position restraints in corresponding simulations that used the Amber ff99SB-ILDN force field had little effect on their ability to match the experiment. Overall, the study shows that both force fields can work well for predicting the effects of active-site mutations on small molecule binding affinities and demonstrates how a direct combination of experiment and computation can be a powerful strategy for developing an understanding of protein−inhibitor interactions.



INTRODUCTION

Protein kinases are of significant current interest as drug targets, because aberrantly regulated kinase activities are associated with a wide variety of human diseases, ranging from rheumatoid arthritis to cancer.[1−3] While there have been significant successes in developing small-molecule inhibitors to target selected kinases—most notably, the development of imatinib to treat chronic myelogenous leukemia[4]—it has been recognized for many years that a major challenge to overcome with such inhibitors is the development of drug resistance.[5,6] Resistance can occur via several mechanisms,[7,8] but the most important in the case of kinases seems to be spontaneous mutations in the ATP-binding domain that diminish inhibitor binding while preserving the kinase's catalytic function.[9,10] Given the apparent ubiquity of resistance as a problem in the development of kinase inhibitors, it would be helpful to have methods that enable potential resistance-causing mutations to be predicted or identified before they are encountered clinically.

One way to achieve this is to develop sophisticated experimental genetic screens.[11] An alternative, and the basis for the present study, is to use computational methods to explicitly calculate the binding thermodynamics of an inhibitor with potential mutants of the kinase target: if such calculations could be performed accurately and rapidly enough, it would be possible, in principle, to pre-emptively screen all likely mutations that might occur within the active site of a kinase and to begin design efforts to deal with those mutations predicted to cause resistance. A key requirement of such an approach, of course, is the ability to accurately compute the thermodynamic effects of active-site mutations on the binding affinities of small-molecule inhibitors for proteins. While there are many approaches that might be used for this purpose, the most accurate predictions of ligand−protein binding affinities are, in principle, to be obtained from thermodynamic

Received: February 8, 2013
Published: May 28, 2013

dx.doi.org/10.1021/ct400104x | J. Chem. Theory Comput. 2013, 9, 3151−3164

integration (TI) or free energy perturbation (FEP) techniques that explicitly compute absolute or relative free energies of binding; several excellent reviews of the use of these methods to compute ligand−protein binding affinities have recently been published.[12−14] Methods such as TI have the advantage of allowing binding thermodynamics to be computed from molecular dynamics (MD) simulations that explicitly include solvent and that directly account for contributions made by changes in conformational flexibilities of the inhibitor and the protein. While these advantages are to be weighed against the significant computational expense associated with such calculations, the continuing increase in computer power means that calculations can now be performed for much longer periods of time than was originally possible.[15,16] One purpose of the present study, therefore, is to explore the extent to which explicit-solvent free-energy calculations, combined with MD simulations and using modern computing resources, can accurately predict the experimentally observed effects of single-residue mutations on the affinity of an inhibitor for a protein kinase.

As a well-characterized model system for conducting a direct comparison of simulation and experiment, we have chosen to study the interaction between the p38α MAP kinase and the competitive small-molecule inhibitor SB203580, which binds with nanomolar affinity in the kinase's ATP-binding pocket.[17,18] Since many resistance-causing mutations involve relatively conservative changes in amino acid, we have selected mutations with the deliberate intention of causing only modest perturbations of the ATP-binding pocket. Our goal in each case has been (a) to present the inhibitor with mild challenges to binding without abolishing binding completely, (b) to experimentally measure the inhibitor's binding affinity for the mutant kinase, and (c) to then attempt to reproduce the measured binding affinity using free-energy calculations. Five mutants of p38α were selected for study (Figure 1): A51V, V38I, A51G, V38A, and A51V/V38I. The first two of these mutations explore the effects of decreasing the size of the binding pocket (by increasing the size of side chains that contact the inhibitor), the next two explore the effects of increasing the size of the binding pocket, and the final mutation explores the combined actions of the first two mutations.

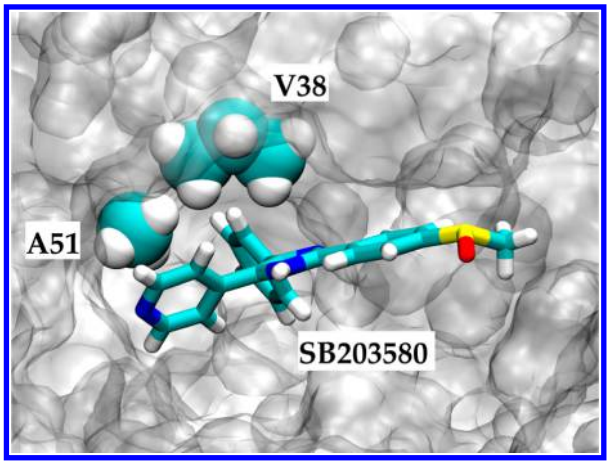


Figure 1. Close-up view of the p38α−SB203580 binding site; p38α kinase (in gray) is shown in molecular surface representation (with a probe radius of 1.2 Å). The inhibitor SB203580 (in cyan) is shown in stick representation. The two mutation sites, Ala51 and Val38, are highlighted using space-fill representation.

Although all of these mutations have been selected primarily as vehicles for exploring the effects of modest structural changes on the binding affinity of a kinase inhibitor, it is perhaps worth noting that (a) the A51V mutation has been found in somatic mutations associated with stomach cancer,[19] and (b) the so-called glycine-rich loop of the kinase, of which V38 is a part, is one of the most common locations for resistance-causing mutations to occur in the BCR-ABL kinase.[10]

This paper is organized as follows. In the Experimental Section, we outline (a) the development of an experimental assay that utilizes changes in the fluorescence intensity of the SB203580 inhibitor to directly measure protein-inhibitor binding affinities, and (b) the simulation protocols used to compute the binding free energy of the inhibitor for each mutant kinase, relative to that of the wild-type protein. In the Results section, we compare the experimental relative binding free energies of each mutant with calculated relative binding free energies obtained using combined molecular dynamics simulations and thermodynamic integration calculations (MD/TI). We compare results obtained with two different simulation force fields—OPLS-AA/L[20] and Amber ff99SB-ILDN[21,22]—and show that, although both force fields perform well for most of the mutants, simulations performed with the former force field produces surprisingly poor estimates for mutants that contain the A51V mutation. We then show that these apparent deficiencies can be corrected by applying harmonic restraints to the positions of atoms of the protein backbone and the inhibitor; however, we also show that the results obtained with this approach are sensitive to the strength of the harmonic restraints' force constant. Overall, the results indicate that, under the right circumstances, MD/TI simulations can compute the effects of active-site mutations on protein-inhibitor interaction thermodynamics with reasonable accuracy. If similar accuracy can also be achieved in calculations that focus instead on the binding of the kinase's true substrate (ATP), it should be possible to use MD/TI simulations to predict potential resistance-causing mutations in protein kinases prior to their occurrence in the clinic.

EXPERIMENTAL SECTION

Mutagenesis, Expression, and Purification of p38α Proteins. The bacterial vector pET21a-6xHIS-rTEV-p38α encoding wild-type p38α (GenBank ID: CAQ52036.1) was a generous gift from Professor Ernesto J. Fuentes (University of Iowa). The mutants A51V, A51G, V38A, V38I, and A51V/V38I were made using the QuikChange site-directed Mutagenesis Kit (Stratagene) and primers were synthesized by Integrated DNA Technologies. All mutant vectors were sequenced to confirm the DNA sequence (DNA core facility, College of Medicine, University of Iowa). All p38α proteins were expressed in *E. coli.* strain Rosetta 2 (Novagen). Bacterial cultures were grown at 37 °C until the optical density (OD) reached 0.6 before being induced with 1 mM (final concentration) IPTG and incubated at 30 °C for 6 h. Following centrifugation at 6000 rpm for 25 min (Beckman Coulter, Model J2-21M), pellets were resuspended in buffer containing 50 mM Tris base, 500 mM NaCl, 25 mM imidazole, 5% glycerol, and 1 mM DTT at pH 8.0. Cells were then lysed by sonication (Misonix Sonicator, Model W-385) using three 2-min rounds of sonication, with each round consisting of 40 pulses (pulse cycle: on for 1 s, then off for 2 s). The lysed cells were then centrifuged at 14 000 rpm for 30 min (Beckman Coulter, Model J2-21M). The supernatant was loaded onto a HiTrap Ni^{2+} affinity column (GE

Healthcare) and elution was performed using a linear gradient of imidazole (25 to 300 mM); p38α eluted at ~220 mM imidazole. Finally, the pooled fractions were loaded onto a Superdex 75 (GE Healthcare) gel filtration column and protein was eluted with buffer containing 5 mM HEPES, 150 mM NaCl, 20 mM $MgCl_2 \cdot 6H_2O$, and 2 mM TCEP at pH 7.4.

Fluorescence Binding Assay. We developed a fluorescence-based assay to experimentally measure the binding affinities of wild-type and mutant p38α proteins to the inhibitor SB203580. The assay exploits the fact that the inhibitor's intrinsic fluorescence increases significantly in intensity and "blue shifts" from 420 nm to 394 nm upon binding to the protein (Figure 2a). Discontinuous titration experiments performed with a range of inhibitor and protein concentrations can therefore be used as a sensitive measure of the binding affinity.

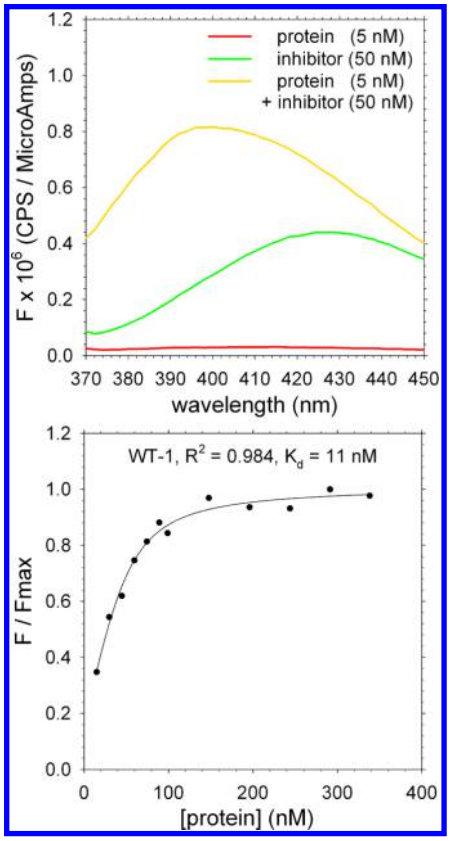


Figure 2. Fluorescence binding assay to measure the binding affinity of wild-type p38α kinase to the inhibitor SB203580. (a) Fluorescence spectrum of the inhibitor, protein, and a solution containing both the inhibitor and the protein; excitation was performed at 320 nm. (b) Typical titration of wild-type protein with the inhibitor; dots represent experimental data and lines represent fits to the quadratic function (see the Experimental Section).

For each mutant, titrations were performed in 12 separate vials, each with a different concentration of protein and a constant concentration (50 nM) of the inhibitor. Protein concentrations were determined using a NanoDrop spectrophotometer (Thermo Scientific, Model ND-1000) with molecular weight and extinction coefficient set to 43.8 kDa and 53000 M^{-1} cm^{-1} respectively, both of which were estimated using the ProtParam web server (http://ca.expasy.org/tools/protparam.html). The concentration of SB203580 (LC Laboratories) was set by weighing out appropriate amounts into an appropriate volume of buffer. The emission signal at 394 nM, following excitation at 320 nm, was measured for each of the 12 titrations using a FluoroLog-3 spectrofluorometer (Horiba Scientific). All experiments were performed at 25 °C in the following buffer: 5 mM HEPES, 150 mM NaCl, 20 mM $MgCl_2 \cdot 6H_2O$, 2% dimethyl sulfoxide (DMSO), and 2 mM TCEP at pH 7.4.

For each mutant, the raw fluorescence intensities at 394 nm were first normalized by the highest fluorescence intensity obtained in the set of 12 titrations. The normalized fluorescence intensities were then fit to a quadratic binding function in SigmaPlot 10.0[23] in order to obtain the dissociation constant (K_d):

$$f = a\left\{\left[([P]_T + [L]_T + K_d) - \sqrt{([P]_T + [L]_T + K_d)^2 - 4[P]_T[L]_T}\right]/\left(2[L]_T\right)\right\} + b$$

Here, $[P]_T$ is the total protein concentration, $[L]_T$ the total inhibitor concentration, K_d the dissociation constant, and f the fractional saturation. The two additional fitting parameters a and b allow, respectively, the relative fluorescence intensity at saturation and the relative fluorescence in the absence of protein to both float during the fit; the former parameter was found to be important in helping to determine the K_d value of the weaker-binding mutants. In all cases, fits to the above equation were excellent ($r^2 > 0.98$). Error bars for all measurements were obtained from the standard deviation of three independent repeats of the titration experiments. Finally, experimental binding free energies (ΔG) were calculated from the K_d value using the standard relation

$$\Delta G = -RT \ln\left(\frac{1}{K_d}\right)$$

The relative binding free energies (denoted as $\Delta\Delta G$) of the mutants to wild-type p38α were then obtained using the equation

$$\Delta\Delta G = \Delta G(\text{mutant}) - \Delta G(\text{wild-type})$$

Simulation System Preparation. Residues Val38 and Ala51 in p38α were selected for mutation, because of their proximity to the inhibitor in the crystal structure of the p38α–SB203580 complex (see Figure 1, as well as the work of Wang et al.[18]). For each of the five mutations studied experimentally—i.e., A51V, A51G, V38A, V38I, and A51V/V38I—separate free-energy calculations were performed for both inhibitor-*bound* and inhibitor-*unbound* states in order to obtain the relative binding free energy, $\Delta\Delta G$. For simulations of the bound state, the p38α crystal structure with SB203580 bound (Protein Databank (PDB) code: 1A9U) was used as the initial structure.[18] Since this structure is of the human form of p38α, two minor modifications (L48H and T263A) were made to the structure using SCWRL 4.0[24] to ensure that its sequence was identical to that of the mouse form of p38α (the form used in our fluorescence experiments; see above). Both of these residues are distant from the inhibitor's binding site, so we do not expect them to make a significant contribution to the binding free energy. For simulations of the *unbound* state, the apo p38α crystal structure (PDB code: 1P38) was used as the

initial structure;[25] since this structure is of the mouse form, no structural modification was required.

Molecular Dynamics (MD) Simulations. Explicit-solvent molecular dynamics (MD) simulations were performed using GROMACS 4.5.1[26,27] on Kraken, which is a Cray XT5 supercomputer hosted by the National Institute for Computational Sciences at the Oak Ridge National Laboratory (Oak Ridge, TN, USA). The simulation system was set up in a cubic box (each side 85 Å) solvated with ∼13 000 TIP3P water molecules;[28] all simulations contained a total of ∼43 000 atoms. Separate sets of simulations were performed with the OPLS-AA/L[20] and Amber ff99SB-ILDN[22] force fields. The former force field is a modification of the original OPLS all-atom force field[29] refitted to high-level quantum-mechanical (QM) data.[20] The latter force field is based on the Amber ff99SB force field,[21] which, in turn, was based on a modified Amber ff94 force field;[30] again, improvement was obtained by matching several backbone and side chain dihedrals to high-level QM calculations.[22]

Prior to deriving force field parameters for the SB203580, it was necessary to decide upon the likely protonation states of the (three) ionizable nitrogen atoms in its imidazole and pyridine rings. To this end, pK_a values were predicted using the SPARC web server (http://sparc.chem.uga.edu/sparc/; see the 1995 work of Hilal et al.[31]). The predicted pK_a value of the pyridine nitrogen was 5.77, suggesting that it should be unprotonated at neutral pH; this assignment is consistent with the fact that it accepts a hydrogen bond from the backbone NH of Met109 in the crystal structure of the complex with p38α.[18] The two imidazole nitrogen atoms were assigned pK_a values of 2.62 and 10.75, respectively (when viewed from top to bottom in Figure S1 in the Supporting Information); these assignments were again found to be consistent with the arrangement seen in the crystal structure of the p38α–SB203580 complex: the predicted unprotonated nitrogen accepts a (charged) hydrogen bond from the side chain of Lys53.

Having decided on protonation states, it was possible to derive partial charge sets for the inhibitor for simulations with the two force fields. For simulations using the OPLS-AA/L force field, the partial charges for the inhibitor were initially obtained using the CM1A method[32] (see Table S1 in the Supporting Information), as has been previously used successfully by others.[33] However, in preliminary simulations performed with this charge set, the key hydrogen bond between the pyridine ring of the inhibitor and the backbone NH group of Met109 was rapidly lost, suggesting that the charge set significantly underestimates the strength of this interaction (see Figure S2 in the Supporting Information). Therefore, using an approach that we described previously,[34] we instead opted to derive partial charges for the inhibitor on the basis of analogy with functional groups already parametrized within the OPLS-AA/L force field. The inhibitor's structure can be usefully decomposed into five principal moieties: (1) imidazole, (2) fluorobenzene, (3) pyridine, (4) benzene, and (5) sulfoxide (see Figure S1 in the Supporting Information). Since each of these groups has already been parametrized in the OPLS-AA/L force field,[20] it is straightforward to merge their partial charge sets to obtain a complete set of charges for the inhibitor. For simulations using the Amber ff99SB-ILDN force field, the partial charges of the inhibitor were obtained using the standard protocol[35] of performing restrained electrostatic potential (RESP) fits; these calculations were performed using the RED web server[36] (http://q4md-forcefieldtools.org/REDS/).

For simulations using the OPLS-AA/L force field, the bonded and nonbonded parameters for SB203580 were also derived by analogy with functional groups already present in the OPLS-AA/L force field. For simulations using the Amber ff99SB-ILDN force field, the bonded and nonbonded parameters for SB203580 were parametrized using the generalized AMBER force field (GAFF) approach.[37] In order to maintain the correct "propeller-like" arrangement of the inhibitor's aromatic groups seen in its crystal structure with p38α, additional dihedral restraints were applied to the ring atoms in all simulations (see Table S2 in the Supporting Information).

All MD simulations were performed in the NVT ensemble. The temperature was maintained at 298 K using the Langevin thermostat with a friction coefficient of 1 ps^{-1}. For van der Waals interactions and short-range electrostatic interactions, a 10 Å cutoff was used; for long-range electrostatic interactions, the smooth Particle Mesh Ewald (PME) method was used.[38] LINCS[39] was used to constrain all bonds, enabling a 2 fs time step to be used. For each simulation, energy minimization using the steepest descent algorithm was first carried out for 1000 steps, followed by 600 ps of "equilibration" that involved a stepwise heating of the system from 50 K to 298 K (through 100 K, 150 K, 200 K, and 250 K). Following this period of equilibration, "production" simulations were run for 12 ns with snapshots saved every 1 ps and energies saved every 0.1 ps for further analysis.

Free-Energy Calculations. The use of thermodynamic integration (TI) methods allows us to calculate the free-energy changes associated with mutations in both inhibitor-bound and inhibitor-unbound states of the protein (i.e., ΔG_{bound} and $\Delta G_{unbound}$, respectively). These free-energy changes were calculated by performing several independent MD simulations at different values of the coupling parameter λ that interpolates between the wild-type and mutant forms of the kinase. In order to practically implement this, many new hybrid residue types were defined in GROMACS. As shown in Figure S3 in the Supporting Information, for example, for the A51V mutation, a hybrid residue type was constructed by combining the atoms of alanine and valine into a single topology file; atoms not present in either of the two residue types are represented by noninteracting "dummy" atoms in such a way that the hybrid residue is effectively an alanine at $\lambda = 0$ (wild-type) and a valine at $\lambda = 1$. The free-energy changes (ΔG_{bound} and $\Delta G_{unbound}$) were each found by integrating the ensemble average of $dH/d\lambda$ accumulated in independent MD simulations performed at each of the following 11 values of λ: 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0. All integrations were performed using Simpson's rule, since this has been shown by the Boresch group to be a considerably more robust method of integration than the more commonly used trapezoid method.[40,41] To avoid numerical instabilities at the end points (i.e., near $\lambda = 0$ and $\lambda = 1$), soft-core potentials[42] were applied to both van der Waals and Coulombic interactions with the coefficients α, p, and σ set to 0.5, 1.0, and 0.3, respectively. (For a good tutorial on the selection of parameters, see http://www.dillgroup.ucsf.edu/group/wiki/index.php?title=Free_Energy:_Tutorial.)

To estimate statistical uncertainties, three independent 12-ns MD simulations were performed at each λ-value for each mutant, where each case starts with a different set of initial velocities assigned to the atoms in the system. If we assume that the three estimates of $dH/d\lambda$ obtained at each λ are truly independent, it is possible to obtain a total number of 3^{11}

Table 1. Comparison of Experimental and Computational ΔG and $\Delta\Delta G$ Values

protein	ΔG_{expt} [a] (kcal/mol)	$\Delta\Delta G_{expt}$ [b] (kcal/mol)	$\Delta\Delta G_{OPLS}$ [c] (kcal/mol)	$\Delta\Delta G_{AMBER}$ [d] (kcal/mol)	$\Delta\Delta G_{OPLS}$ [e] (kcal/mol)	$\Delta\Delta G_{AMBER(R)}$ [f] (kcal/mol)
WT	−11.0 ± 0.1	N/A	N/A	N/A	N/A	N/A
A51V	−8.9 ± 0.1	2.0 ± 0.1	−0.3 ± 0.4	3.1 ± 0.4	2.3 ± 0.2	3.2 ± 0.2
A51G	−10.1 ± 0.1	0.8 ± 0.1	0.6 ± 0.2	0.6 ± 0.2	0.8 ± 0.1	0.5 ± 0.1
V38A	−8.5 ± 0.1	2.5 ± 0.1	1.4 ± 0.3	1.4 ± 0.3	1.9 ± 0.1	2.1 ± 0.2
V38I	−11.6 ± 0.2	−0.7 ± 0.3	−0.4 ± 0.2	−0.2 ± 0.4	−0.4 ± 0.1	−0.3 ± 0.2
A51V/V38I	−10.0 ± 0.1	1.0 ± 0.2	−0.5 ± 0.5	1.6 ± 0.6	1.3 ± 0.2	2.3 ± 0.3

[a]Experimental absolute binding free energy. [b]Experimental relative binding free energy. [c]Computational relative binding free energy obtained using OPLS-AA/L force field without position restraints. [d]Computational relative binding free energy using Amber ff99SB-ILDN force field without position restraints. [e]Computational relative binding free energy obtained using OPLS-AA/L force field with position restraints (0.239 kcal/mol/Å^2). [f]Computational relative binding free energy using Amber ff99SB-ILDN force field with position restraints (0.239 kcal/mol/Å^2).

estimates of both ΔG_{bound} and $\Delta G_{unbound}$; therefore, we chose to estimate the statistical error of ΔG_{bound} and $\Delta G_{unbound}$ from the standard deviation of these 3^{11} integrations of $dH/d\lambda$. Finally, following standard statistics,[43] the standard deviation of $\Delta\Delta G$ was obtained as the square root of the sum of the squared standard deviations of ΔG_{bound} and $\Delta G_{unbound}$.

The total amount of simulation time accrued during this project can be computed as follows. Each $\Delta\Delta G$ estimate required 22 × 12 ns = 264 ns of MD simulation (11 λ-values for the bound state, 11 λ-values for the unbound state). For the OPLS-AA/L force field, we obtained 40 $\Delta\Delta G$ estimates: 3 replicates of 5 mutants for unrestrained simulations, 3 replicates of 5 mutants for restrained simulations using a force constant of 0.239 kcal/mol/Å^2 (i.e., 100 kJ/mol/nm^2; see the Results section), 1 replicate of 5 mutants using a force constant of 0.0239 kcal/mol/Å^2, and 1 replicate of 5 mutants using a force constant of 2.39 kcal/mol/Å^2. For the Amber ff99SB-ILDN force field, we obtained 30 $\Delta\Delta G$ estimates, since we did not perform position-restrained calculations using force constants of 0.0239 and 2.39 kcal/mol/Å^2. Therefore, the total simulation time used in this study amounts to 70 × 264 ns = 18.48 μs.

Calculations Using the MM/GBSA Method. An alternative to computing the relative free energies of binding by performing MD simulations at several independent, intermediate λ-values is to use a so-called "end point" method that involves calculations only at the extreme λ-values.[44] One such method is the molecular mechanics/generalized Born surface area (MM/GBSA) method;[45−47] a large number of studies have used this approach in ligand−receptor calculations.[48−55] The MM/GBSA method estimates the binding free energy (ΔG) as the difference in free energies of the ligand−receptor complex and its separated components, with each free energy being modeled as a sum of the following terms:

$$G = E_{int} + E_{elec} + E_{vdW} + G_{sol,p} + G_{sol,np} + TS$$

E_{int} is the internal energy due to the traditional bonded terms of the force field (i.e., bond stretches, angles, and dihedrals), and E_{elec} and E_{vdW} are the electrostatic and van der Waals nonbonded energies, respectively. $G_{sol,p}$ is the polar component of the solvation free energy, which is calculated here using the OBC generalized Born solvation model,[56] with parameters α = 1, β = 0.8, and γ = 4.85, as these parameters have been shown to work well in similar systems.[53] $G_{sol,np}$ is the nonpolar component of the solvation free energy, calculated using $G_{sol,np} = \gamma \times SASA$, where γ = 0.0054 kcal/mol (from ref 57) and SASA is the solvent-accessible surface area. As is common in applications of the MM/GBSA method, the entropy term (TS) was omitted from the calculations reported here.

For each mutant, free energies were computed at the λ = 1 end states using snapshots sampled at 10-ps intervals from 12-ns MD trajectories; since each trajectory was performed in triplicate, a total of 360 snapshots were used to obtain the binding free energy of each mutant. For the wild-type, free energies were computed at the λ = 0 end state using snapshots sampled from all independent simulations; since five mutants were studied, and since each was simulated in triplicate, a total of 5 × 3 × 120 = 1800 snapshots were available for calculating the binding free energy of the wild-type protein. The relative binding free energies, $\Delta\Delta G$, of all mutants were obtained by subtracting the binding free energy of the wild-type kinase. All calculations were performed using the "rerun" option of the main MD engine, "mdrun", in GROMACS.[26,27]

RESULTS

Experimental Measurements of $\Delta\Delta G$ of p38α Mutants. In order to determine the binding affinities of p38α mutants with the inhibitor SB203580 experimentally, we have developed a fluorescence-based binding assay; this assay exploits the fact that the fluorescence intensity of the inhibitor changes significantly when it binds to the protein (Figure 2a). To validate the assay, it was first applied to the wild-type p38α, so that comparison with previously reported binding affinities for SB203580 could be made. The dissociation constant (K_d) measured here for SB203580 with wild-type p38α is 9.0 ± 1.7 nM (Figure 2b). This agrees well with the previously reported values of 10 nM,[58] 17 nM,[59] 12 nM,[60] and 4 nM;[61] the largest difference in free energy between the present value and any of the other reported values, therefore, is 0.48 kcal/mol (i.e., a 4% difference). Having validated the assay for the wild-type protein, binding affinities of SB203580 for the five mutants examined here were determined using the same assay; the measured ΔG values are listed in Table 1 (see Figure S4 in the Supporting Information for fits of the experimental data for all mutants).

The two mutations that decrease the size of the side chains contacting the inhibitor, V38A and A51G, decrease the binding affinity of the inhibitor by 2.5 and 0.8 kcal/mol, respectively (see Table 1). The former result is qualitatively consistent with a previous study showing that this mutation reduces cross-linking of the kinase to the similar ligand ^{125}I-SB206718 (identical to SB203580, except that the latter's fluorine is replaced by iodine[62]), and both results are consistent with our naive expectations, given that they entail the loss of hydrophobic contacts with the inhibitor. It is not immediately obvious why the V38A mutation should be so much more deleterious to binding than the A51G mutation, since, according to the Ligand-Protein Contacts web server,[63] the

two wild-type side chains bury essentially the same amounts of surface area in binding to the inhibitor (22 and 20 Å^2 for V38 and A51, respectively). However, this absence of an obvious explanation suggests that correctly reproducing the relative magnitudes of the effects of these two mutations should be a challenge for the computer simulations.

In contrast to what was seen with the above mutations, the two mutations that increase the size of the side chains contacting the inhibitor, A51V and V38I, have qualitatively quite different effects on the binding affinity of the inhibitor: the A51V mutation decreases the binding affinity of the inhibitor by 2.0 kcal/mol, while the V38I mutation increases the binding affinity by 0.7 kcal/mol (see Table 1). These two results can be at least qualitatively rationalized on the basis of simple molecular modeling performed on the crystal structure of the p38α–SB203580 complex (PDB code: 1A9U). When we use the side chain modeling program SCWRL4[24] to model the A51V mutation (Figure 3), pronounced steric clashes are

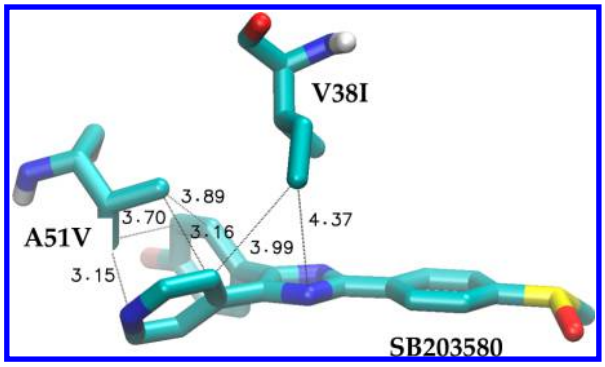


Figure 3. Putative close contacts between the mutation sites of p38α kinase and the inhibitor SB203580. All distances are measured in Ångstroms.

apparent between the valine side chain and the fluorophenyl and pyridine rings of the inhibitor: the closest distance between heavy atoms is 3.15 Å. In contrast, when we carry out the same procedure with the V38I mutation (Figure 3), no clashes are apparent between the isoleucine side chain and the inhibitor, and the side chain appears much freer to rotate. The fact that the V38I mutation introduces an additional CH_2 group to contact the inhibitor without introducing any new steric clashes therefore provides a reasonable explanation for the finding that the mutation increases the binding affinity of the inhibitor. Interestingly, the idea that increasing the size of side chains contacting the inhibitor can cause an increase in the binding affinity of the p38α–SB203580 complex is not without precedent: it has been shown previously that the A157V mutation (also in the active site of p38α) decreases the IC_{50} for the inhibitor from 30 nM to 6 nM.[64] Evidently, therefore, despite the inhibitor's high affinity for p38α, there is room to improve its affinity by improving the packing of the protein's side chains around it.

Since the two size-increasing single mutations, A51V and V38I, have qualitatively different thermodynamic effects on SB203580s binding affinity, it was of interest to establish whether the two effects were additive, especially since the two sites are physically close to each other: the distance between the Cα atom of Val38 and Ala51 is 6.4 Å in the p38α–SB203580 complex crystal structure. Therefore, we measured the binding affinity of the double mutant, A51V/V38I, and found that the overall effect is to decrease the binding affinity by 1.0 kcal/mol; this is to be compared with the predicted combined effects of the two single mutations ($\Delta\Delta G$ = 2.0 − 0.7 kcal/mol = 1.3 kcal/mol). The result is therefore somewhat equivocal: a very modest degree of cooperativity (~0.3 kcal/mol) may be present but it is not sufficiently large in magnitude to be clearly outside of the error in the experimental measurements.

Free-Energy Calculations of $\Delta\Delta G$ of p38α Mutants. Having experimentally measured the binding free energy of SB203580 to several site-directed mutants of the active site of p38α, we next attempted to determine whether these data could be reproduced by explicit-solvent TI/MD simulations. Our initial round of free-energy simulations used the OPLS-AA/L force field with parameters for the inhibitor devised to be as consistent as possible with the rest of the force field (see the Experimental Section). TI calculations with this force field produced very reasonable results for three of the mutants—A51G, V38A, and V38I (cyan, green, and yellow symbols in Figure 4a)—and, in particular, were able to capture the finding that the V38A mutation is significantly more damaging to binding than the A51G mutation is (see above). However, the

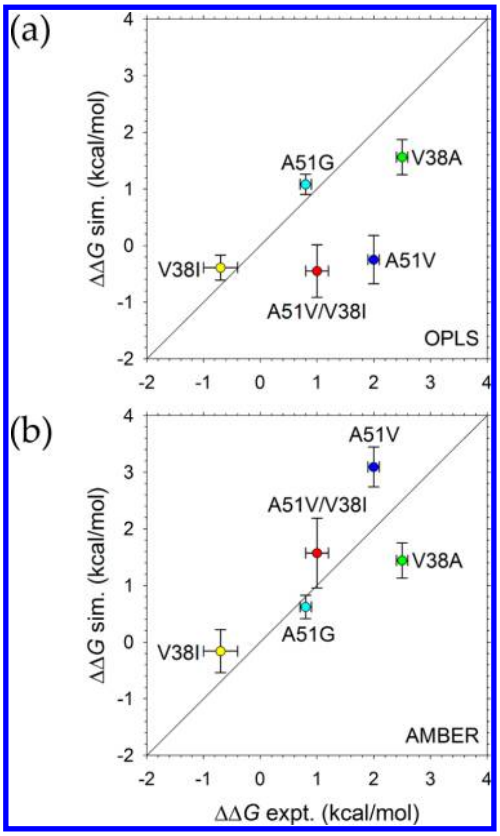


Figure 4. Correlation of the $\Delta\Delta G$ values calculated from unrestrained simulations with the experimentally measured $\Delta\Delta G$ values for p38α mutants: A51V, A51G, V38A, V38I, and A51V/V38I. The results of the simulations using the following force fields are shown: (a) OPLS-AA/L and (b) Amber ff99SB-ILDN. Each vertical error bar represents the standard deviation of 3[11] independent $\Delta\Delta G$ estimates from three 12-ns simulations with different initial random velocities. Each horizontal error bar represents the standard deviation of $\Delta\Delta G$ from three independent fluorescence binding experiments. The diagonal line representing perfect correlation is also shown. The linear regressions obtained for OPLS and AMBER force fields are y = 0.38x − 0.12 (R^2 = 0.26) and y = 0.75x + 0.47 (R^2 = 0.59), respectively.

TI results for the remaining two mutants, both of which involve the A51V substitution, were qualitatively incorrect (blue and red symbols in Figure 4a; Table 1): the simulations predict that both mutants should have an increased binding affinity for the inhibitor, when, in fact, their affinities are substantially decreased. We consider possible causes of these discrepancies in the Discussion section. To explore whether similar results might be obtained with an alternative force field, we repeated the entire set of free-energy calculations using the Amber ff99SB-ILDN force field (see the Experimental Section). In this case, the calculations performed considerably better (Figure 4b): qualitatively correct results were obtained for all five mutants, and, although the calculations now overestimated the extent to which the A51V mutations disfavor inhibitor binding, the correspondence with the experimental data was markedly improved. Overall, the correlation coefficient increased from 0.49 (with OPLS-AA/L) to 0.76 (with Amber ff99SB-ILDN), and the average unsigned error in the computed $\Delta\Delta G$ values decreased from 1.1 kcal/mol to 0.7 kcal/mol.

Imposing Harmonic Position Restraints Can Significantly Improve Calculated $\Delta\Delta G$ Values. While there are several likely causes of the poor results obtained when the OPLS-AA/L force field is applied to the A51V-containing mutants (see below), one possible interpretation is that the force field may be allowing the protein−inhibitor complex to adopt conformations that, in reality, would not be sampled. If so, and assuming that the crystal structure of the wild-type kinase in complex with the inhibitor provides a good guide to what the A51V mutant *should* look like in complex with the inhibitor, then it might be expected that adding modest position restraints to limit the conformational freedom of the complex might lead to somewhat improved binding affinity predictions. To test this idea, we repeated the TI calculations with harmonic restraints applied to the backbone atoms of the protein and to the heavy atoms of the inhibitor; no harmonic restraints were imposed on the protein side chains, because it was believed important to allow them to sample alternative conformations freely (see Figure S5 in the Supporting Information for plots comparing the relative flexibilities of these groups).

Using an arbitrarily selected force constant of 0.239 kcal/mol/Å^2 (i.e., 100 kJ/mol/nm^2), the relative binding free energies of all five p38α mutants were calculated in exactly the same way as for the unrestrained simulations. As expected, the primary effect of the position restraints was to affect the calculation of the free-energy changes in the inhibitor-bound state (ΔG_{bound}) while making little difference to the calculation of the free-energy changes in the inhibitor-free state ($\Delta G_{unbound}$) (see Figure S6 in the Supporting Information). More importantly, however, the imposition of harmonic restraints hugely improved the calculated $\Delta\Delta G$ values for the A51V and A51V/V38I mutants without adversely affecting the calculated $\Delta\Delta G$ values for the other mutants (see Figure 5a). In fact, the correlation coefficient for the OPLS-AA/L results improved from 0.49 without restraints to 0.95 with restraints; similarly, the average unsigned error improved from 1.1 kcal/mol to only 0.3 kcal/mol. Interestingly, we found that the same restraints could also be imposed in simulations with the Amber ff99SB-ILDN force field without causing a significant adverse change in the force field's ability to reproduce the experimental data (see Figure 5b): the correlation coefficient and average unsigned errors changed to 0.83 and 0.7 kcal/mol, respectively, when restraints were included.

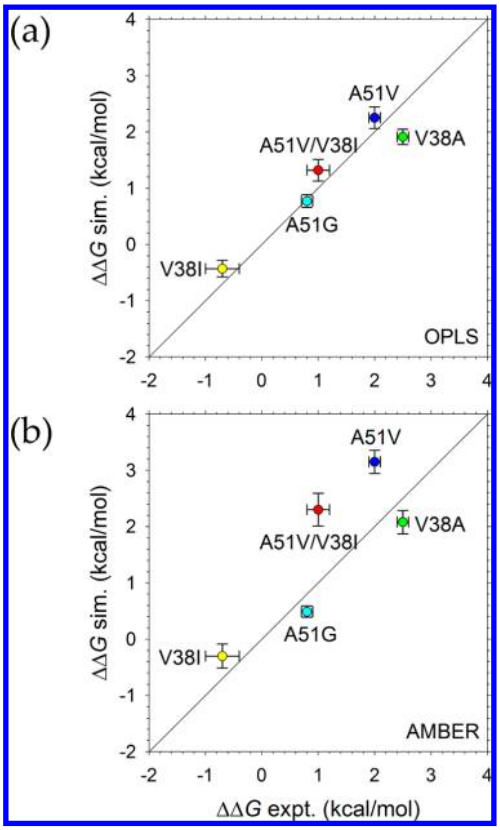


Figure 5. Correlation of the $\Delta\Delta G$ values calculated from restrained simulations with the experimentally measured $\Delta\Delta G$ values for p38α mutants: A51V, A51G, V38A, V38I, and A51V/V38I. The results of the simulations using the following force fields are shown: (a) OPLS-AA/L and (b) Amber ff99SB-ILDN. Position restraints with a force constant of 0.239 kcal/mol/Å^2 were applied to the backbone of the protein and heavy atoms of the drug. Each vertical error bar represents the standard deviation of 3[11] independent $\Delta\Delta G$ estimates from three 12-ns simulations with different initial random velocities. Each horizontal error bar represents the standard deviation of $\Delta\Delta G$ from three independent fluorescence binding experiments. The diagonal line representing perfect correlation is also shown. The linear regressions obtained for OPLS and AMBER force fields are $y = 0.82x + 0.25$ ($R^2 = 0.92$) and $y = 0.94x + 0.49$ ($R^2 = 0.68$), respectively.

Although it is encouraging that the mixed performance of the OPLS-AA/L force field can be improved by adding position restraints, it is important from a practical perspective to ask how strong these restraints must be in order to have a "correcting" effect. To address this question, we again repeated our TI calculations with the OPLS-AA/L force field, but using instead force constants that were either an order of magnitude weaker or stronger (i.e., 0.0239 and 2.39 kcal/mol/Å^2 respectively); for reference, a plot showing the effects of increasing force constant on the flexibilities of backbone, side chain and inhibitor atoms is shown in Figure S7. Interestingly, the computed $\Delta\Delta G$ values of the two A51V-containing mutants show a very significant dependence on the magnitude of the force constant, increasing linearly with the logarithm of the force constant (Figure 6). In contrast, the three non-A51V-containing mutants, A51G, V38A, and V38I, show much less dependence on the force constant, although the thermodynamically favorable V38I mutation is somewhat more poorly reproduced as the strength of the position restraints increases. The latter result suggests that correctly reproducing the favorable effect of this particular mutation requires a degree

dx.doi.org/10.1021/ct400104x | J. Chem. Theory Comput. 2013, 9, 3151−3164

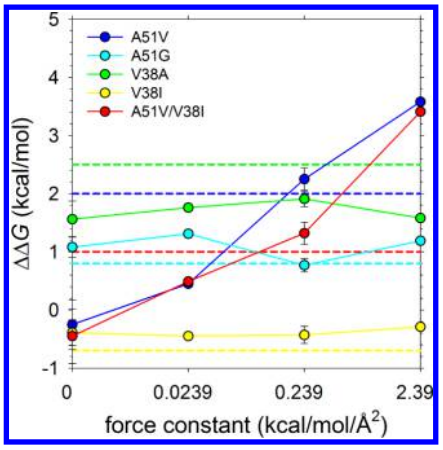


Figure 6. Force constant dependence of $\Delta\Delta G$ for p38α mutants obtained from simulations using OPLS-AA/L force field. Solid circles indicate the calculated $\Delta\Delta G$ values and dashed lines indicate experimentally measured $\Delta\Delta G$ values. The calculated $\Delta\Delta G$s were obtained from three 12-ns simulations using the following four different force constants: (1) 0 kcal/mol/Å^2 (unrestrained), (2) 0.0239 kcal/mol/Å^2, (3) 0.239 kcal/mol/Å^2, and (4) 2.39 kcal/mol/Å^2. The error bars for (1) and (3) were obtained from the standard deviation of three simulation replicates; the error bars for (2) and (4) were not calculated, because only one simulation was performed.

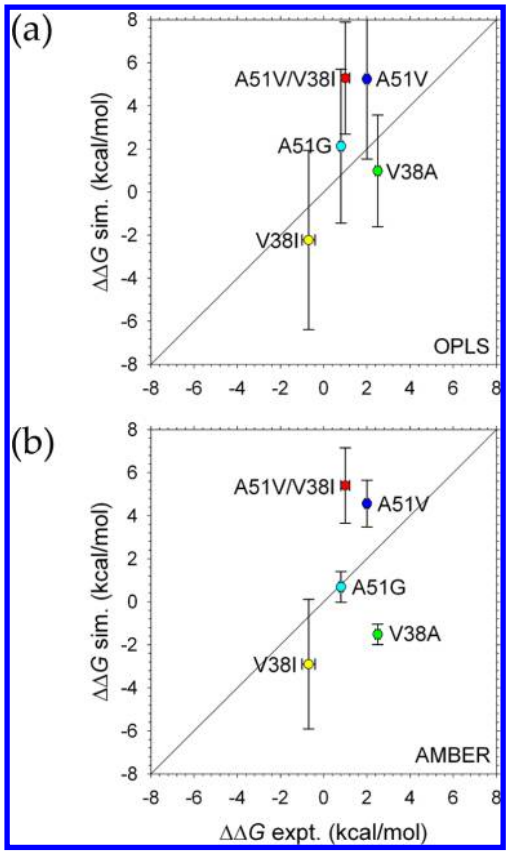


Figure 7. Correlation of the $\Delta\Delta G$ values obtained from MM/GBSA calculations for unrestrained simulations with the experimentally measured $\Delta\Delta G$ values for p38α mutants: A51V, A51G, V38A, V38I, and A51V/V38I. The results of the simulations using the following force fields are shown: (a) OPLS-AA/L and (b) Amber ff99SB-ILDN. Each vertical error bar is obtained using the procedure as described in the Experimental Section. Each horizontal error bar represents the standard deviation of $\Delta\Delta G$ from three independent fluorescence binding experiments. The diagonal line representing perfect correlation is also shown. The linear regressions obtained for OPLS and AMBER force fields are $y = 1.43x + 0.68$ ($R^2 = 0.31$) and $y = 1.04x + 0.09$ ($R^2 = 0.12$), respectively.

of conformational adjustment on the part of either (or both) the protein and the inhibitor. Overall, however, the results indicate that while the imposition of position restraints can overcome apparent limitations in a simulation force field, the strength of the restraints must be carefully chosen.

One final question that we have addressed is whether similar results could be obtained using a simpler computational approach that avoids the need to simulate intermediate λ values in the calculation of relative binding free energies. To this end, we have used the MM/GBSA method,[44−47] in which the energies of MD snapshots sampled at $\lambda = 0$ and $\lambda = 1$ are calculated as a sum of a molecular mechanics component and an implicit solvent component that combines continuum electrostatic and surface area-dependent terms (see the Experimental Section). The results obtained when the MM/GBSA method is used to compute energies for snapshots sampled during the unrestrained simulations with both force fields are shown in Figure 7; corresponding results obtained using snapshots sampled during the simulations restrained with a force constant of 0.239 kcal/mol/Å^2 are shown in Figure 8. Interestingly, in the absence of position restraints, the correlation between the MM/GBSA calculations and experiment is quite poor with both force fields: correlation coefficients of 0.56 and 0.35 are obtained for the OPLS-AA/L and Amber ff99SB-ILDN force fields, respectively. When position restraints are included, however, both force fields produce much better agreement with experiment: the correlation coefficients become 0.79 and 0.75 for the OPLS-AA/L and Amber ff99SB-ILDN force fields, respectively, which are comparable to but somewhat lower than those obtained using the much more expensive TI calculations.

DISCUSSION

The potential utility of free-energy calculation methods such as FEP and TI for applications such as computer-aided drug design was first suggested and demonstrated in the mid-1980s,[65−70] and such methods continue to be widely used as tools for studying the thermodynamics of biomolecular interactions. In recent years, considerable interest has been focused on understanding the *absolute* thermodynamics of ligand binding to a variety of model protein systems such as T4 lysozyme,[71−78] the FK-506 binding protein (FKBP),[79−81] and cytochrome P450.[82] However, there also remains significant interest in using such methods to compute *relative* binding free energies, for example, of closely related drug analogues binding to the same protein.[83−89]

Conceptually, the present study belongs with other computational studies exploring the effects of mutations on a protein binding to a common ligand. Previous studies along these lines have included cases where the ligand is another protein,[90,91] a sugar,[92] or an enzyme substrate.[93−96] However, of most relevance to the current work are those computational studies that have focused specifically on reproducing mutational effects on the binding thermodynamics of small molecules to proteins.[97−102] In comparison with all but the most recent of these studies, the simulations reported here involve much longer simulation times; nevertheless, the overall accuracy of the results that we have obtained is broadly consistent with that obtained in these previous studies. For example, a study of the

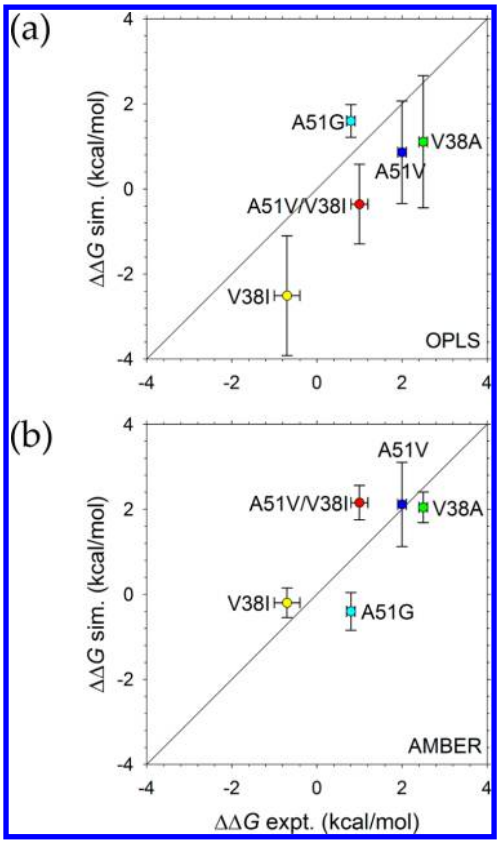


Figure 8. Correlation of the $\Delta\Delta G$ values obtained from MM/GBSA calculations for restrained simulations with the experimentally measured $\Delta\Delta G$ values for p38α mutants: A51V, A51G, V38A, V38I, and A51V/V38I. The results of the simulations using the following force fields are shown: (a) OPLS-AA/L and (b) Amber ff99SB-ILDN. Position restraints with a force constant of 0.239 kcal/mol/Å^{-2} were applied to the backbone of the protein and heavy atoms of the drug. Each vertical error bar is obtained using the procedure as described in the Experimental Section. Each horizontal error bar represents the standard deviation of $\Delta\Delta G$ from three independent fluorescence binding experiments. The diagonal line representing perfect correlation is also shown. The linear regressions obtained for OPLS and AMBER force fields are $y = 1.05x - 1.03$ ($R^2 = 0.62$) and $y = 0.80x + 0.24$ ($R^2 = 0.57$), respectively.

effects of five mutations on the binding of the enzyme fructose 1,6-biphosphatase to adenosine monophosphate obtained excellent agreement with experiment: despite the use of comparatively short production periods of 306 ps, the error in the computed $\Delta\Delta G$ values in that study was in the range of only 0.2−0.4 kcal/mol.[101] As a second example, in a study examining the interaction between protease inhibitors and three resistance-causing mutants of the Hepatitis C virus protease, errors in the range of 0.3−0.9 kcal/mol were obtained; interestingly, that study also used position restraints similar to those applied in the present study, although, in that case, the CHARMM force field was used.[100]

In the present work, we have used both unrestrained and restrained simulations in order to calculate $\Delta\Delta G$ values. We discuss the latter in detail below, but first we consider the results obtained from unrestrained simulations. One important point to note concerns the magnitudes of the discrepancies between the computed $\Delta\Delta G$ values and our experimentally measured values. As shown in Table 1, in the unrestrained calculations, these discrepancies appear to scale with the size of the mutation—larger changes to side chains appear more difficult for the calculations to reproduce quantitatively—and this trend holds true regardless of which force field is used in the simulations. Specifically, the mutations that involve the addition or loss of only a single carbon (A51G and V38I) are both reproduced very accurately: the unsigned error for both mutations with both force fields is only 0.3 ± 0.1 kcal/mol. In contrast, the mutants that involve the addition or loss of two carbons (A51V and V38A) are reproduced much less accurately: the unsigned error for these mutations with the two force fields is 1.3 ± 0.6 kcal/mol. We have already noted that the results for A51V with OPLS-AA/L are poor (Table 1), but at a quantitative level, this mutation also appears to be difficult for the Amber ff99SB-ILDN force field, since it overestimates the change in binding affinity by 1.1 kcal/mol. In addition, while both force fields also qualitatively reproduce the lower binding affinity of V38A, the magnitude of the effect is underestimated by 0.9 kcal/mol with OPLS-AA/L and 1.1 kcal/mol Amber ff99SB-ILDN. The fact that the thermodynamic effects of both of these mutations are difficult to reproduce indicates that mutations that introduce steric insults to inhibitor binding (A51V) may not be any more challenging than mutations that involve the removal of interactions (V38A), at least from the point of view of producing quantitative agreement with the experiment. This finding may have implications for attempts to model the effects of mutations on protein stability where, because of difficulties in modeling the side chains, attention has previously been focused on those mutations that involve the deletion of atoms.[103,104]

A second result of the unrestrained simulations is that we find that increasing the "production" period of the simulations (i.e., the period of time over which data are recorded) has equivocal effects on the calculated $\Delta\Delta G$ values. Most of the computed $\Delta\Delta G$ values are, perhaps surprisingly, relatively insensitive to changes in the duration of the production period (see Figures S8 and S9 in the Supporting Information). In the case of unrestrained simulations using the OPLS-AA/L force field (Figure S8a in the Supporting Information), for example, the most noticeable effect is on the computed $\Delta\Delta G$ value for the A51V mutant, which decreases from 1.0 kcal/mol to −0.3 kcal/mol as the production period increases from 1 ns to 12 ns. Given that our experimentally determined value is 2.0 kcal/mol (Table 1), the calculated result clearly deviates further from the experimental value as the simulation production period increases. In the case of unrestrained simulations using the Amber ff99SB-ILDN force field, on the other hand (Figure S9a in the Supporting Information), the most noticeable changes are to the $\Delta\Delta G$ values of the V38I and A51V/V38I mutants, both of which become closer to the experimental value over time.

While the dependence of the results on the duration of the production period is generally more muted in the simulations that include position restraints on the backbone atoms (Figures S8b and S9b in the Supporting Information), in certain cases (e.g., for the A51V/V38I mutant with the Amber ff99SB-ILDN force field), a clear time dependence remains. One likely cause of this remaining time dependence, and a significant possible source of discrepancies with experiment, is the relatively infrequent sampling of alternative rotameric states of side chains that contact the inhibitor. In fact, infrequent sampling of side chain rotamers appears to be an issue in all of our simulations, regardless of the force field used and regardless of whether or not position restraints are applied to the inhibitor

and the protein backbone. As shown in Figures S10 and S11 in the Supporting Information, for example, at position 38, we see very few transitions in the χ1 dihedral over the course of the 12-ns production period in simulations of A51V-containing mutants; sampling is slightly better at position 51 in the same mutants (see Figures S12 and S13 in the Supporting Information), but it remains infrequent. Such slow sampling of alternative side chain conformations is not necessarily unrealistic, since (i) significant energy barriers to side chain rotations are common[105] and (ii) in seven different crystal structures of SB203580 in complex with p38 kinases (pdb codes 1A9U, 1PME, 2EWA, 3GCP, 3OBG, 3MPA, and 3ZS5), the side chain of V38 is always in the *trans* conformation. Nevertheless, it makes it difficult to be confident in the convergence of the computed free-energy values.

This is especially a concern given that others have shown that proper consideration of alternative side chain rotamers can be an important factor in obtaining accurate absolute binding free energies in ligand−receptor systems. For example, many studies have shown that inadequate sampling of the χ_1 angle at Val111—for which conformational rearrangements are slow—can have a critical effect on the computed binding free energies of aromatic molecules to T4 lysozyme.[73,75,78,82,106] A variety of different ways to correct for this problem have been proposed and have been explicitly demonstrated in the same system. One such method is the "confine-and-release" protocol developed by Mobley, Chodera, and Dill,[76] in which the free-energy contribution from a side chain that is kinetically trapped within a single rotameric state is explicitly computed. An alternative comes the Roux group,[78] showing that it is possible to accelerate convergence in absolute binding free-energy calculations by using Hamiltonian exchange methods in combination with biasing potentials that are explicitly devised to improve sampling of the χ_1 dihedral angles. Finally, the Berne and Friesner groups have shown that free energies that are independent of the initial rotameric states of side chains can be achieved in simulations of only ∼2 ns duration, using a replica exchange solute-tempering method[106] that, in effect, simulates the ligand and its immediate binding site at a higher temperature; this same method has already been applied to a kinase system.[55] Any or all of these methods could be especially helpful for the present system.

It is perhaps because of the infrequent sampling of side chain conformations, therefore, that we do not see any clear relationship between the length of the production period in the simulations and the accuracy with which the experimental results are reproduced. Because of this, therefore, we cannot conclusively rule out the possibility that increasing the production period of the simulations to a much longer time scale (e.g., microseconds) might change the results, either for better or for worse. The ability to perform much longer simulations, thanks to the advent of superfast MD engines such as the purpose-built supercomputer Anton created by the Shaw group,[107,108] should make it possible to eliminate incomplete sampling as a major concern in the near future. In the meantime, however, and for the general user of simulations who does not have access to such machines, it appears worthwhile to consider methods such as those outlined above, as well as others that improve sampling in the computation of binding free energies.[76,78,106,109]

While the above discussion indicates that we should be cautious about the convergence of our computed $\Delta\Delta G$ values, it remains reasonably clear that the results we obtain for the A51V-containing mutants from unrestrained simulations with the OPLS-AA/L force field are qualitatively incorrect while those obtained with the Amber ff99SB-ILDN force field are qualitatively correct (see Table 1). Given the highly interconnected nature of the parameters in current energy functions, it can be extremely difficult to cleanly identify problem areas in a given force field. One potential contributor that we can probably rule out is differences in the conformational preferences of the inhibitor, since, with both force fields, we have imposed the same set of dihedral restraints in order to ensure that the relative orientations of the fluorophenyl, pyridine, and imidazole rings are preserved; as a result, the sampled distributions of these dihedral angles are essentially identical with the two force fields (see Figure S14 in the Supporting Information). Another possibility is the description of the kinase's conformational preferences, and, in particular, those of the side chains that contact the inhibitor: the OPLS and AMBER force fields have been shown by others, for example, to produce somewhat different χ_1 rotamer distributions for residue types such as valine.[110] The infrequent sampling of these rotamers noted above—which is only one aspect of the more general problem of sampling that has been emphasized recently[111]—makes it difficult to determine if this is a major issue but, in any case, it should be noted that we can get excellent agreement with experiment from simulations that include backbone position restraints, despite the fact that any errors in rotamer preferences would still be present in such simulations (restraints were not applied to side chain atoms).

Still another possible contribution comes from the partial charges assigned to the inhibitor. In the Experimental Section, we note that in preliminary MD simulations with the OPLS-AA/L force field, the use of CM1A charges for the inhibitor led to a rapid loss of the key hydrogen bonding interaction between the pyridine ring of the inhibitor and the amide hydrogen of the backbone of Met109. This led us to instead derive partial charges for the inhibitor by analogy with functional groups already parametrized in the OPLS-AA/L force field, and while these analogy-derived charges improved the stability of the hydrogen bond significantly, its stability in 12-ns MD simulations of the wild-type kinase is still clearly somewhat lower than that obtained with the Amber ff99SB-ILDN force field (see Figure S2 in the Supporting Information). Notably, the A51V mutation is the only one considered here that introduces a clear steric clash into the system and, given the proximity to the pyridine ring, it is quite possible that adjustments to the hydrogen bond might be required in order to accommodate it (recall Figure 3). If the simulation force field's description of the hydrogen bond interaction is unrealistically weak, it might not correctly account for the fact that there should be a significant energetic cost to altering the geometry of the hydrogen bond and might, therefore, underestimate how unfavorable the mutation is for binding of the inhibitor. Arguing against this possible explanation, however, is the fact that, in additional simulations that we have performed using the OPLS-AA/L force field but with the Amber charges assigned to the inhibitor, we obtain a $\Delta\Delta G$ value of 0.3 ± 0.4 kcal/mol for the A51V mutant, which is only marginally better than the result obtained when the analogy-derived OPLS-AA/L charges are used ($\Delta\Delta G$ = −0.3 ± 0.4 kcal/mol; see Table 1).

Although the exact cause of the OPLS-AA/L force field's problems with the A51V-containing mutations remains difficult to establish unambiguously, it is important to note that these

problems can be made to disappear when position restraints are applied to the protein backbone and the inhibitor: in fact, when set appropriately, restraints allow average unsigned errors as low as 0.3 kcal/mol to be achieved. The addition of position restraints to simulations with the Amber ff99SB-ILDN force field made little difference to what was already a rather good agreement with the experiment, although there was a slight increase in the correlation. There was, however, one important respect in which the Amber ff99SB-ILDN results were also markedly improved by the addition of restraints: this was in the reproduction of the cooperativity of the A51V/V38I double mutant. Without restraints, the "coupling energy" between the two mutation sites, defined as $\Delta\Delta G_{A51V/V38I} - \Delta\Delta G_{V38I} - \Delta\Delta G_{A51V}$, was computed to be −1.4 kcal/mol; with restraints, the value reduces to −0.6 kcal/mol, which is much closer to the experimental value of −0.3 kcal/mol. Similarly, with OPLS-AA/L, the computed coupling energy is +0.2 kcal/mol (i.e., qualitatively incorrect) in the absence of restraints, and −0.6 kcal/mol in the presence of restraints. Interestingly, the imposition of position restraints also markedly improved the results obtained with the faster but more approximate MM/GBSA method for *both* force fields (compare Figures 7 and 8).

The fact that imposing position restraints can lead to better calculations of mutational effects echoes a previous result from our laboratory. In work aimed at developing rapid computational methods for screening the entire complement of human protein kinases, we showed that it was possible to use homology models of protein kinases, together with a simple physical energy function, to qualitatively predict which would be strongly bound by a chosen small-molecule inhibitor.[34] Importantly, there were two key features of that approach: (a) to assume that all kinases use the exact same backbone conformation to bind the putative inhibitor, and (b) to assume that the inhibitor binds in an identical orientation to all kinases. Notably, when either of these apparently drastic and unrealistic restrictions was removed, significantly poorer results were obtained. In comparison with the explicit-solvent TI/MD simulations used here, the previous approach was obviously highly simplified. Despite that, the key result of that study was essentially the same as that obtained here: in both cases, better results can be obtained when a simulation force field is denied the opportunity to lead the modeled protein−inhibitor complex into different, and possibly wrong, conformations. Obviously, in an ideal situation, position restraints would not be necessary at all, and as force fields continue to improve, the advantages of including them will likely disappear; as a temporary stop-gap measure, however, they appear to provide a reasonable solution when the force field has clearly identified limitations. Importantly, the conclusion that position restraints can improve sampling results in MD also emerges from a recent study by the Shaw group reporting very-long-time-scale MD simulations of homology-modeled protein structures.[112] It is also important to note, however, that there are clear cases where the addition of position restraints will likely worsen rather than improve results. Here again, the T4 lysozyme system offers a good example, because ligands binding to the T4 lysozyme system can adopt several different orientations and correctly accounting for them can be important for correctly reproducing experimental binding free energies.[113] An additional point to note is that, as shown here, the best choice of force constant to use is likely to be dependent on the specific system and force field under study.

Finally, it is worth noting that another factor that might affect the ability to produce good agreement with experiment is the presence of multiple conformational equilibria within the kinase. It has been known for several years, for example, that a flipping of the D168-F169-G170 motif in p38α from a so-called "DFG-in" to a "DFG-out" conformation, in which a cryptic hydrophobic binding site is exposed, can be exploited by so-called Class II inhibitors.[114] Kinetic data[115] have shown that the binding of such inhibitors is much slower than that of the more conventional Class I (i.e., ATP-binding site) inhibitors such as SB203580, suggesting that the DFG-in and DFG-out conformations interconvert slowly; NMR data indicate that, in the ligand-unbound state, this exchange occurs on a millisecond time scale.[116] Interestingly, NMR studies have indicated that the binding of SB203580 does not affect the equilibrium between the DFG-in and DFG-out conformations,[116] and crystallographic studies have shown that SB203580 can in fact bind to both conformations.[116,117] Evidence for additional conformational dynamics within p38α comes from another NMR study,[118] showing that significant line broadening of a number of resonances occurs when SB203580 binds; notably, these include many amino acids that are close to the inhibitor (e.g., G33-V38; I84-L86). Interestingly, residual dipolar coupling (RDC) data reported in the same study also indicate that the backbone conformation of unbound p38α in solution might differ from that observed in the crystal structure at residues G110-A111. Still another feature to potentially consider is that, in its physiological state, p38α is activated by phosphorylation at T180 and Y182;[119] NMR and H/D exchange studies have shown that this double phosphorylation has effects distributed throughout the protein and causes the activation loop to switch from the intermediate to the fast exchange regime.[120] Note that phosphorylation is not an issue in the present study, since both the experiments and simulations have been performed in conditions in which p38α is unphosphorylated; in any case, it has been shown that SB203580 binds with similar affinities to both forms of the enzyme.[121]

Many computational studies have already considered some of the above issues. The transition between the DFG-in and DFG-out conformations of p38α, for example, has been directly observed in artificially accelerated MD simulations performed by our group[122] and by the Menziani group.[123] Very long, unforced MD simulations of the same transition in other kinases have also been reported by the Shaw group,[124,125] and several attempts to model the DFG-out conformations of other kinases for use in virtual screening applications have also been reported.[126−128] The finding that SB203580 binds with equal affinities to the DFG-in and DFG-out conformations of p38α has been reproduced by the Chang group in recent computer simulations, using the "mining minima" method,[129] in conjunction with a continuum solvent model.[130] Finally, many other computational studies focusing on p38α as a model system for virtual screening applications have also been reported.[128,131−134]

Given the significant conformational transitions that occur within p38α, and the time scale over which they occur, it is perhaps surprising that comparatively short simulations coupled with position restraints can yield relative binding free energies that are in good agreement with the experiment. Presumably, this is because the mutations studied here do not significantly affect either the conformational dynamics of p38α or its conformational free-energy landscape; the effects of mutations

that are closer to the site of action (e.g., mutations of the DFG motif itself[135]) might well be much more difficult to reproduce. For the future, therefore, experimental measurements of the binding thermodynamics of SB203580 to a wider range of mutants might enable the p38α−SB203580 system to serve as a very challenging model for testing both enhanced conformational sampling methods and simulation force fields.

In summary, we have reported here a combined experimental and computational approach to study the effects of active-site mutations on the binding free energy of p38α MAP kinase to the inhibitor SB203580. We have shown that, depending on the site of the mutation, increasing the size of a side chain contacting the inhibitor can either increase or decrease the binding affinity of a small-molecule inhibitor. We have also shown that, under the right circumstances, these changes in binding affinity can be quite accurately predicted by TI/MD simulations. Given that the conservative mutations studied here are quite common in cases of drug resistance, the results obtained here suggest that TI/MD simulations could eventually find use in predicting potential resistance-causing mutations prior to their occurrence in the clinic; the results reported by others in such applications[100,102] support this notion. However, to do this in a comprehensive fashion, at least for protein kinases, it will be important also to determine the effects of the mutations on the binding of the kinase's substrate ATP. This is especially true given that resistance-causing mutations have already been identified that exert their effects by increasing the affinity for ATP without affecting the affinity for the inhibitor.[136]

■ ASSOCIATED CONTENT

Supporting Information

Principal moieties of the inhibitor SB203580; probability distribution function of the key hydrogen bond observed in simulations of the wild-type p38α−SB203580 complex; scheme for performing alchemical transformations of p38α mutants; quadratic fits of the normalized fluorescence data obtained from titrations of p38α protein with SB203580; time dependence of root-mean-square deviations (RMSDs) for selected groups of atoms in harmonically restrained simulations of wild-type p38α in complex with SB203580; plots of $dH/d\lambda$ vs λ for unrestrained and restrained simulations of p38α mutants using the OPLS-AA/L force field; RMSD versus force constant for 12-ns restrained simulations of wild-type p38α kinase in complex with SB203580 using the OPLS-AA/L force field; time dependence of $\Delta\Delta G$ values calculated from simulations using the OPLS-AA/L force field; time dependence of $\Delta\Delta G$ values calculated from simulations using the Amber ff99SB-ILDN force field; time dependence of χ_1 angles at residue 38 in simulations of the A51V mutant; time dependence of χ_1 angles at residue 38 in simulations of the A51V/V38I mutant; time dependence of χ_1 angles at residue 51 in simulations of the A51V mutant; time dependence of χ_1 angles at residue 51 in simulations of the A51V/V38I mutant; dihedral distributions for inter-ring angles in the inhibitor sampled during unrestrained simulations using the OPLS-AA/L and Amber ff99SB-ILDN force fields; three partial charge sets for the inhibitor SB203580; additional dihedral restraints applied to SB203580 to maintain its "propeller-like" shape. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: adrian-elcock@uiowa.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was made possible by the support of NIH R01 GM068648 to A.H.E. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation (under Grant No. OCI-1053575); generous allocations of time on the supercomputers Kraken and Ranger were awarded to A.H.E. under Project No. TG-MCB090131.

■ REFERENCES

(1) McLean, G. W.; Carragher, N. O.; Avizienyte, E.; Evans, J.; Brunton, V. G.; Frame, M. C. *Nat. Rev. Cancer* **2005**, *5*, 505−515.
(2) Shapiro, G. I. *J. Clin. Oncol.* **2006**, *24*, 1770−1783.
(3) Schindler, J. F.; Monahan, J. B.; Smith, W. G. *J. Dent. Res.* **2007**, *86*, 800−811.
(4) Deininger, M.; Buchdunger, E.; Druker, B. J. *Blood* **2005**, *105*, 2640−2653.
(5) Daub, H.; Specht, K.; Ullrich, A. *Nat. Rev. Drug Discovery* **2004**, *3*, 1001−1010.
(6) Nardi, V.; Azam, M.; Daley, G. Q. *Curr. Opin. Hematol.* **2004**, *11*, 35−43.
(7) Jänne, P. A.; Gray, N.; Settleman, J. *Nat. Rev. Drug Discovery* **2009**, *8*, 709−723.
(8) Barouch-Bentov, R.; Sauer, K. *Expert Opin. Invest. Drugs* **2011**, *20*, 153−208.
(9) Bikker, J. A.; Brooijmans, N.; Wissner, A.; Mansour, T. S. *J. Med. Chem.* **2009**, *52*, 1493−1509.
(10) Krishnamurty, R.; Maly, D. J. *ACS Chem. Biol.* **2010**, *5*, 121−138.
(11) Girdler, F.; Sessa, F.; Patercoli, S.; Villa, F.; Musacchio, A.; Taylor, S. *Chem. Biol.* **2008**, *15*, 552−562.
(12) Aleksandrov, A.; Thompson, D.; Simonson, T. *J. Mol. Recognit.* **2010**, *23*, 117−127.
(13) Steinbrecher, T.; Labahn, A. *Curr. Med. Chem.* **2010**, *17*, 767−785.
(14) Michel, J.; Essex, J. W. *J. Comput. Aid. Mol. Des.* **2010**, *24*, 639−658.
(15) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.
(16) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(17) Tong, L.; Pav, S.; White, D. M.; Rogers, S.; Crane, K. M.; Cywin, C. L.; Brown, M. L.; Pargellis, C. A. *Nat. Struct. Biol.* **1997**, *4*, 311−316.
(18) Wang, Z. L.; Canagarajah, B. J.; Boehm, J. C.; Kassisà, S.; Cobb, M. H.; Young, P. R.; Abdel-Meguid, S.; Adams, J. L.; Goldsmith, E. J. *Structure* **1998**, *6*, 1117−1128.
(19) Stephens, P.; Hunter, C.; Bignell, G.; Edkins, S.; Davies, H.; Teague, J.; Stevens, C.; O'Meara, S.; Smith, R.; Parker, A.; Barthorpe, A.; Blow, M.; Brackenbury, L.; Butler, A.; Clarke, O.; Cole, J.; Dicks, E.; Dike, A.; Drozd, A.; Edwards, K.; Forbes, S.; Foster, R.; Gray, K.; Greenman, C.; Halliday, K.; Hills, K.; Kosmidou, V.; Lugg, R.; Menzies, A.; Perry, J.; Petty, R.; Raine, K.; Ratford, L.; Shepherd, R.; Small, A.; Stephens, Y.; Tofts, C.; Varian, J.; West, S.; Widaa, S.; Yates, A.; Brasseur, F.; Cooper, C. S.; Flanagan, A. M.; Knowles, M.; Leung, S. Y.; Louis, D. N.; Looijenga, L. H. J.; Malkowicz, B.; Pierotti, M. A.; Teh, B.; Chenevix-Trench, G.; Weber, B. L.; Yuen, S. T.; Harris, G.; Goldstraw, P.; Nicholson, A. G.; Futreal, P. A.; Wooster, R.; Stratton, M. R. *Nature* **2004**, *431*, 525−526.
(20) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(21) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins—Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(22) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins—Struct., Funct., Bioinf.* **2010**, *78*, 1950−1958.
(23) *SigmaPlot V.10*; Systat Software, Inc.: Richmond, CA.
(24) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L. *Proteins—Struct., Funct., Bioinf.* **2009**, *77*, 778−795.
(25) Wang, Z. L.; Harkins, P. C.; Ulevitch, R. J.; Han, J. H.; Cobb, M. H.; Goldsmith, E. J. *Proc. Natl. Acad. Sci., U.S.A.* **1997**, *94*, 2327−2332.
(26) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.
(27) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(30) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(31) Hilal, S. H.; Karickhoff, S. W.; Carreira, L. A. *Quant. Struct. Act. Relat.* **1995**, *14*, 348.
(32) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1995**, *9*, 87.
(33) Tominaga, Y.; Jorgensen, W. L. *J. Med. Chem.* **2004**, *47*, 2534.
(34) Rockey, W. M.; Elcock, A. H. *J. Med. Chem.* **2005**, *48*, 4138.
(35) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
(36) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F. Y. *Nucleic Acids Res.* **2011**, *39*, W511.
(37) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(38) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(39) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463.
(40) Bruckner, S.; Boresch, S. *J. Comput. Chem.* **2011**, *32*, 1320−1333.
(41) Bruckner, S.; Boresch, S. *J. Comput. Chem.* **2011**, *32*, 1303−1319.
(42) Beutler, T. C.; Mark, A. E.; Vanschaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529.
(43) Barlow, R. J. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*; Wiley: Chichester, U.K., 1989.
(44) Hayes, J. M.; Archontis, G. MM-GB(PB)SA calculations of protein−ligand binding free energies. In *Molecular Dynamics—Studies of Synthetic and Biological Macromolecules*, Wang, L., Ed.; InTech, Rijeka, Croatia, 2012. (ISBN: 978-953-51-0444-5; DOI:10.5772/37107.)
(45) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(46) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(47) Massova, I.; Kollman, P. A. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.
(48) Gohlke, H.; Kiel, C.; Case, D. A. *J. Mol. Biol.* **2003**, *18*, 891−913.
(49) Gohlke, H.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 238−250.
(50) Chachra, R.; Rizzo, R. C. *J. Chem. Theory Comput.* **2008**, *4*, 1526−1540.
(51) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.
(52) Ng, C. A.; Oehme, D. P.; Kato, Y.; Tanokura, M.; Brownlee, R. T. C. *J. Chem. Theory Comput.* **2009**, *5*, 2886−2897.
(53) Balius, T. E.; Rizzo, R. C. *Biochemistry* **2009**, *48*, 8435−8448.
(54) Newhouse, E. I.; Xu, D.; Markwick, P. R. L.; Amaro, R. E.; Pao, H. C.; Wu, K. J.; Alam, M.; McCammon, J. A.; Li, W. W. *J. Am. Chem. Soc.* **2009**, *131*, 17430−17442.
(55) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. *J. Chem. Theory Comput.* **2013**, *9*, 1282−1293.
(56) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Struct. Funct. Bioinf.* **2004**, *55*, 383−394.
(57) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(58) Regan, J.; Breitfelder, S.; Cirillo, P.; Gilmore, T.; Graham, A. G.; Hickey, E.; Klaus, B.; Madwed, J.; Moriak, M.; Moss, N.; Pargellis, C.; Pav, S.; Proto, A.; Swinamer, A.; Tong, L.; Torcellini, C. *J. Med. Chem.* **2002**, *45*, 2994.
(59) Fabian, M. A.; Biggs, W. H.; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. *Nat. Biotechnol.* **2005**, *23*, 329.
(60) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. *Nat. Biotechnol.* **2008**, *26*, 127.
(61) Swann, S. L.; Merta, P. J.; Kifle, L.; Groebe, D.; Sarris, K.; Hajduk, P. J.; Sun, C. H. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5787.
(62) Gum, R. J.; McLaughlin, M. M.; Kumar, S.; Wang, Z. L.; Bower, M. J.; Lee, J. C.; Adams, J. L.; Livi, G. P.; Goldsmith, E. J.; Young, P. R. *J. Biol. Chem.* **1998**, *273*, 15605.
(63) Sobolev, V.; Sorokine, A.; Prilusky, J.; Abola, E. E.; Edelman, M. *Bioinformatics* **1999**, *15*, 327.
(64) Lisnock, J. M.; Tebben, A.; Frantz, B.; O'Neill, E. A.; Croft, G.; O'Keefe, S. J.; Li, B.; Hacker, C.; de Laszlo, S.; Smith, A.; Libby, B.; Liverton, N.; Hermes, J.; LoGrasso, P. *Biochemistry* **1998**, *37*, 16573.
(65) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281.
(66) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.
(67) Lybrand, T. P.; McCammon, J. A.; Wipff, G. *Proc. Natl. Acad. Sci., U.S.A.* **1986**, *83*, 833.
(68) Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830.
(69) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574.
(70) McCammon, J. A. *Science* **1987**, *238*, 486.
(71) Mann, G.; Hermans, J. *J. Mol. Biol.* **2000**, *302*, 979.
(72) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535.
(73) Deng, Y. Q.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255.
(74) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902.
(75) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118−1134.
(76) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 1231.
(77) Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A.; Shoichet, B. K. *J. Mol. Biol.* **2009**, *394*, 747.
(78) Jiang, W.; Roux, B. *J. Chem. Theory Comput.* **2010**, *6*, 2559−2565.
(79) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 5.
(80) Wang, J. Y.; Deng, Y. Q.; Roux, B. *Biophys. J.* **2006**, *91*, 2798.
(81) Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. *J. Chem. Phys.* **2006**, *125*, 12.
(82) Deng, Y. Q.; Roux, B. *J. Chem. Phys.* **2008**, *128*, 8.
(83) Khavrutskii, I. V.; Wallqvist, A. *J. Chem. Theory Comput.* **2011**, *7*, 3001.
(84) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2001**, *123*, 6246.

(85) Erion, M. D.; Dang, Q.; Reddy, M. R.; Kasibhatla, S. R.; Huang, J.; Lipscomb, W. N.; van Poelje, P. D. *J. Am. Chem. Soc.* **2007**, *129*, 15480.
(86) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296.
(87) Zeevaart, J. G.; Wang, L. G.; Thakur, V. V.; Leung, C. S.; Tirado-Rives, J.; Bailey, C. M.; Domaoal, R. A.; Anderson, K. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2008**, *130*, 9492.
(88) Jorgensen, W. L.; Bollini, M.; Thakur, V. V.; Domaoal, R. A.; Spasov, K. A.; Anderson, K. S. *J. Am. Chem. Soc.* **2011**, *133*, 15686.
(89) Genheden, S.; Nilsson, I.; Ryde, U. *J. Chem. Inf. Model.* **2011**, *51*, 947.
(90) Zhou, R. H.; Das, P.; Royyuru, A. K. *J. Phys. Chem. B* **2008**, *112*, 15813.
(91) Park, H.; Jeon, Y. H. *J. Mol. Graphics Model.* **2011**, *29*, 643.
(92) Das, P.; Li, J. Y.; Royyuru, A. K.; Zhou, R. H. *J. Comput. Chem.* **2009**, *30*, 1654.
(93) Zheng, Z. L.; Ye, M. Q.; Zuo, Z. Y.; Liu, Z. G.; Tai, K. C.; Zu, G. L. *Biochem. J.* **2006**, *395*, 509.
(94) Pan, Y. M.; Gao, D. Q.; Yang, W. C.; Cho, H.; Zhan, C. G. *J. Am. Chem. Soc.* **2007**, *129*, 13537.
(95) Yang, W. C.; Pan, Y. M.; Zheng, F.; Cho, H.; Tai, H. H.; Zhan, C. G. *Biophys. J.* **2009**, *96*, 1931.
(96) Yang, W. C.; Pan, Y. M.; Fang, L.; Gao, D. Q.; Zheng, F.; Zhan, C. G. *J. Phys. Chem. B* **2010**, *114*, 10889.
(97) Rizzo, R. C.; Wang, D. P.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 12898.
(98) Wang, D. P.; Rizzo, R. C.; Tirado-Rives, J.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2799.
(99) Park, H.; Lee, S. *J. Comput. Aid. Mol. Des.* **2005**, *19*, 17.
(100) Guo, Z. Y.; Prongay, A.; Tong, X.; Fischmann, T.; Bogen, S.; Velazquez, F.; Venkatraman, S.; Njoroge, F. G.; Madison, V. *J. Chem. Theory Comput.* **2006**, *2*, 1657.
(101) Mutyala, R.; Reddy, R. N.; Sumakanth, M.; Reddanna, P.; Reddy, M. R. *J. Comput. Chem.* **2007**, *28*, 932.
(102) Ripoll, D. R.; Khavrutskii, I. V.; Chaudhury, S.; Liu, J.; Kuschner, R. A.; Wallqvist, A.; Reifman, J. *PLoS Comput. Biol.* **2012**, *8*, e1002665.
(103) Guerois, R.; Nielsen, J. E.; Serrano, L. *J. Mol. Biol.* **2002**, *320*, 369−387.
(104) Wickstrom, L.; Gallichio, E.; Levy, R. M. *Proteins: Struct., Funct., Bioinf.* **2011**, *80*, 111−125.
(105) Petrella, R. J.; Karplus, M. *J. Mol. Biol.* **2001**, *312*, 1161−1175.
(106) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci., U.S.A.* **2012**, *109*, 1937−1942.
(107) Shaw, D. E.; Dror, R. O.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B.; Bowers, K. J.; Chow, E.; Eastwood, M. P.; Ierardi, D. J.; Klepeis, J. L.; Kuskin, J. S.; Larson, R. H.; Lindorff-Larsen, K.; Maragakis, P.; Moraes, M. A.; Piana, S.; Shan, Y.; Towles, B. In *Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis (SC09)*; Association for Computing Machinery (ACM): New York, 2009.
(108) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W. *Science* **2010**, *330*, 341.
(109) Zheng, L.; Yang, W. *J. Chem. Theory Comput.* **2012**, *8*, 810.
(110) Jiang, F.; Han, W.; Wu, Y. D. *J. Phys. Chem. B* **2010**, *114*, 5840.
(111) Mobley, D. L. *J. Comput. Aided Mol. Des.* **2012**, *26*, 93−95.
(112) Raval, A.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *Proteins—Struct., Funct., Bioinf.* **2012**, *80*, 2071.
(113) Gallicchio, E.; Lapelosa, M.; Levy, R. M. *J. Chem. Theory Comput.* **2010**, *6*, 2961−2977.
(114) Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan. *Nat. Struct. Biol.* **2002**, *9*, 268−272.
(115) Regan, J.; Pargellis, C.; Cirillo, P. F.; Gilmore, T.; Hickey, E. R.; Peet, G. W.; Proto, A.; Swinamer, A.; Moss, A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3101−3104.
(116) Vogtherr, M.; Saxena, K.; Hoelder, S.; Grimme, S.; Betz, M.; Schieborr, U.; Pescatore, B.; Robin, M.; Delarbre, L.; Langer, T.; Wendt, K. U.; Schwalbe, H. *Angew. Chem., Int. Ed.* **2006**, *45*, 993−997.
(117) Simard, J. R.; Getlik, M.; Grtter, C.; Pawar, V.; Wulfert, S.; Rabiller, M.; Rauh, D. *J. Am. Chem. Soc.* **2009**, *131*, 13286−13296.
(118) Honndorf, V. S.; Coudevylle, N.; Laufer, S.; Becker, S.; Griesinger, C. *Angew. Chem., Int. Ed.* **2008**, *47*, 3548−3551.
(119) Raingeaud, J.; Gupta, S.; Rogers, J. S.; Dickens, M.; Han, J.; Ulevitch, R. J.; Davis, R. J. *J. Biol. Chem.* **1995**, *270*, 7420−7426.
(120) Nielsen, G.; Schwalbe, H. *ChemBioChem* **2011**, *12*, 2599−2607.
(121) Young, P. R.; McLaughlin, M. M.; Kumar, S.; Kassis, S.; Doyle, M. L.; McNulty, D.; Gallagher, T. F.; Fisher, S.; McDonnell, P. C.; Carr, S. A.; Huddleston, M. J.; Seibel, G.; Porter, T. G.; Livi, G. P.; Adams, J. L.; Lee, J. C. *J. Biol. Chem.* **1997**, *272*, 12116−12121.
(122) Frembgen-Kesner, T.; Elcock, A. H. *J. Mol. Biol.* **2006**, *359*, 202−214.
(123) Filomia, F.; De Rienzo, F.; Menziani, M. C. *Bioorg. Med. Chem.* **2010**, *18*, 6805−6812.
(124) Shan, A.; Seeliger, M.; Eastwood, M. P.; Frank, F.; Xu, H.; Jensen, M. O.; Dror, R. O.; Kuriyan, J.; Shaw, D. E. *Proc. Natl. Acad. Sci., U.S.A.* **2009**, *106*, 139−144.
(125) Shan, Y.; Arkhipov, A.; Kim, E. T.; Pan, A. C.; Shaw, D. E. *Proc. Natl. Acad. Sci., U.S.A.* **2013**, *110*, 7270−7275.
(126) Kufareva, I.; Abagyan, R. *J. Med. Chem.* **2008**, *51*, 7921−7932.
(127) Xu, M.; Yu, L.; Wan, B.; Yu, L.; Huang, Q. *PLoS One* **2011**, *6*, e22644.
(128) Flick, J.; Tristram, F.; Wenzel, W. *J. Comput. Chem.* **2012**, *33*, 2504−2515.
(129) Chen, W.; Chang, C. E.; Gilson, M. K. *Biophys. J.* **2004**, *87*, 3035−3049.
(130) Huang, Y. H.; Chen, W.; Potter, M. J.; Chang, C. A. *Biophys. J.* **2012**, *103*, 342−351.
(131) Rao, S.; Sanschagrin, P. C.; Greenwood, J. R.; Repasky, M. P.; Sherman, W.; Farid, R. *J. Comput. Aided Mol. Des.* **2008**, *22*, 621−627.
(132) Cheeseright, T. J.; Holm, M.; Lehmann, F.; Luik, S.; Göttert, M.; Melville, J. L.; Laufer, S. *J. Med. Chem.* **2009**, *52*, 4200−4209.
(133) Armen, R. S.; Chen, J.; Brooks, C. L., III. *J. Chem. Theory Comput.* **2009**, *13*, 2909−2923.
(134) Vinh, N. B.; Simpson, J. S.; Scammells, P. J.; Chalmers, D. K. *J. Comput. Aided Mol. Des.* **2012**, *26*, 409−423.
(135) Bukhtiyarova, M.; Karpusas, M.; Northrop, K.; Namboodiri, H. V. M; Springman, E. B. *Biochemistry* **2007**, *46*, 5687−5696.
(136) Yun, C. H.; Mengwasser, K. E.; Toms, A. V.; Woo, M. S.; Greulich, H.; Wong, K. K.; Meyerson, M.; Eck, M. J. *Proc. Natl. Acad. Sci., U.S.A.* **2008**, *105*, 2070.