Protein-Ligand Binding Affinity Predictions by Implicit Solvent Simulations: A Tool for Lead Optimization?

Julien Michel,[†] Marcel L. Verdonk,[‡] and Jonathan W. Essex*[,†]

School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom, and Astex Therapeutics, Ltd., 436 Cambridge Science Park, Cambridge, CB4 0QA, United Kingdom

Received August 24, 2006

Continuum electrostatics is combined with rigorous free-energy calculations in an effort to deliver a reliable and efficient method for in silico lead optimization. The methodology is tested by calculation of the relative binding free energies of a set of inhibitors of neuraminidase, cyclooxygenase2, and cyclin-dependent kinase 2. The calculated free energies are compared to the results obtained with explicit solvent simulations and empirical scoring functions. For cyclooxygenase2, deficiencies in the continuum electrostatics theory are identified and corrected with a modified simulation protocol. For neuraminidase, it is shown that a continuum representation of the solvent leads to markedly different protein−ligand interactions compared to the explicit solvent simulations, and a reconciliation of the two protocols is problematic. Cyclin-dependent kinase 2 proves more challenging, and none of the methods employed in this study yield high quality predictions. Despite the differences observed, for these systems, the use of an implicit solvent framework to predict the ranking of congeneric inhibitors to a protein is shown to be faster, as accurate or more accurate than the explicit solvent protocol, and superior to empirical scoring schemes.

Introduction

Virtual screening of large databases of compounds against a target protein has become an essential component of drug discovery.[1,2] Docking algorithms can suggest binding modes for thousands of compounds per day.[3] Once a docked pose is obtained, the affinity of the ligand for this target is estimated by some form of equation that attempts to relate the nature of predicted protein−ligand interactions in the complex to the experimental binding affinity of the ligand.[4] In principle, the highest scoring compounds are flagged for bench synthesis, and biological assay confirms their high potency, saving months of work and large quantities of money. In reality, while state of the art protein−ligand docking programs predict correctly the binding mode of a compound 70−80% of the time, few modelers are lucky enough to observe a modest correlation between predicted and experimental binding affinities. Often compounds that do not bind at all are discriminated from those that do bind, but binding affinities are not predicted with an accuracy that would allow in silico lead optimization. Yet, a theoretical method that is able to *reliably* propose appropriate substituent replacements would be extremely useful.

The theory of ligand binding can be understood by the laws of statistical mechanics and is a complex process that requires proper attention to solvation and entropic effects, precisely the factors that are often lacking in empirical scoring functions. In principle, these statistical mechanics equations allow us to calculate exactly the binding free energy of a ligand, and examples of such applications have been reported for two decades.[5−13] However, for all but the simplest systems, assumptions are made so that the calculations are feasible on a practical time scale. First, the force field that describes the molecular interactions may not adequately reflect reality, making the predictions inaccurate. Second, the sampling performed by the molecular dynamics (MD) or Monte Carlo (MC) algorithms may not be sufficient to observe a representative sample of protein−ligand interactions that the system can adopt. In this case, the calculated free energies will be imprecise. The latter point is particularly important: converged predictions have often proven difficult to obtain.[14−16] These considerations mean that in practice a free energy calculation study can fail to give accurate and reliable answers.

Because of the difficulties associated with the practical application of rigorous free-energy calculations, faster, more generally applicable methods have been developed. The MM/PBSA methodology predicts the absolute binding free energy ΔG_{bind} of a ligand by combining molecular mechanics energy, solvation free energies with Poisson Boltzmann or generalized Born calculations, and entropy estimates from normal mode calculations.[17] MM/PBSA simulations of several protein−ligand systems have been reported in the literature. The results were often encouraging[18−22] but sometimes unsatisfactory,[23−25] and the method has drawn some criticism due to its lack of clear theoretical foundation.[24,25] In another popular approach, the linear interaction energy (LIE) method, the absolute binding free energy of a ligand to a protein is obtained by running two independent simulations; one is the ligand free in solution and the other is the solvated protein−ligand complex.[26] An important drawback of the LIE method is that energy descriptors have to be extracted from the simulations and correlated empirically to the activity of a training set of compounds. This limits the application of LIE to systems for which sufficient experimental data is available.[27] Continuum solvent models have been introduced in the LIE methodology to increase efficiency and test their accuracy.[28−30]

The great appeal of implicit solvent techniques is the simplification that arises when thousands of solvent molecules surrounding the protein−ligand complex are abstracted into a continuum. This simplification translates into a faster potential energy evaluation and rapid convergence of thermodynamic properties. Despite the potential benefits, very few studies have considered the combination of implicit solvent theories with

* To whom correspondence should be addressed. Phone: +44 238059 2794. Fax: +44 238059 3781. E-mail: j.w.essex@soton.ac.uk.
[†] School of Chemistry, University of Southampton.
[‡] Astex Therapeutics, Ltd.

10.1021/jm061021s CCC: $33.50
Published on Web 11/21/2006

rigorous free-energy simulations, and none have attempted to calculate complete binding free energies for protein−ligand complexes.[31−34] Such a method would have the benefit of relying on a sound statistical thermodynamics basis and may prove advantageous over more complex explicit solvent simulations where convergence of the free energies takes longer because of the need to simulate the degrees of freedom of thousands of solvent molecules.

We report here in detail Monte Carlo (MC) free-energy simulations of protein−ligand complexes in an implicit solvent model. In addition, the calculated binding free energies are compared to those obtained by explicit solvent simulations and empirical scoring schemes. The systems selected are 10 inhibitors of cyclooxygenase2 (COX2),[35] 10 inhibitors of neuraminidase,[36] and 18 inhibitors of cyclin-dependent kinase2 (CDK2).[37] Potent inhibitors of these three proteins could provide treatment against pain, influenza, or cancer, respectively, and unsurprisingly, numerous drug design programs target them. From a methodological perspective, each system offers different challenges. The binding site of COX2 is buried and hydrophobic, which makes it an interesting test case to assess the ability of an implicit solvent methodology to treat ligand desolvation. In neuraminidase, the binding site is very polar, solvent exposed, and, depending on the ligand substituent, one or two crystallographic waters bridge interactions between the ligands and the protein. It is expected that such a binding site will cause difficulties for an implicit solvent methodology. The CDK2 system should prove particularly challenging because the structure−activity relationships are complex, and several compounds have a similar binding affinity. Predictions of high accuracy and precision are therefore necessary to rank properly the inhibitors of this series.

Methods

Free Energy Calculations. Relative binding free energies can be calculated by constructing a thermodynamic cycle. The principles behind this methodology have been reviewed elsewhere.[38] The free energy change for the mutation of ligand A into B in one medium is obtained by application of the thermodynamic integration method[39]

$$\Delta G_{\text{medium,A}\rightarrow\text{B}} = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda} \mathrm{d}\lambda = \int_0^1 \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \qquad (1)$$

where λ is a coupling parameter that allows the smooth transformation of the potential energy function $U(\lambda = 0)$ appropriate for ligand A, into a potential energy function appropriate for ligand B, $U(\lambda = 1)$. The brackets denote an ensemble average corresponding to the derivative of the potential energy function $U(\lambda)$ with respect to λ (free energy gradients). In practice this quantity is calculated by averaging the value of the free energy gradients calculated over several snapshots of the protein ligand complex generated from a MC or MD simulation. The free energy gradients can be calculated directly if the functional form of the derivatives of the potential energy function with respect to the coupling parameter λ have been implemented in the simulation program. They can also be calculated by a finite difference scheme, for example, $(\partial G(\lambda)/\partial\lambda) = [\Delta G(\lambda + \Delta\lambda) - \Delta G(\lambda - \Delta\lambda)]/[2\Delta\lambda]$, provided $\Delta\lambda$ is small enough. This approach was adopted here where the Zwanzig equation[40] was used to calculate the free energies $\Delta G(\lambda + \Delta\lambda)$ and $\Delta G(\lambda - \Delta\lambda)$. These free energy gradients are calculated at several values of the parameter λ, and the integral is then estimated by trapezoidal numerical integration.

The relative binding free energy, $\Delta\Delta G_{\text{bind,A}\rightarrow\text{B}}$, is the difference between the free energy change in the protein environment, $\Delta G_{\text{protein,A}\rightarrow\text{B}}$, and the free energy change in the aqueous environment, $\Delta G_{\text{aqueous,A}\rightarrow\text{B}}$. The relative hydration free energy, $\Delta\Delta G_{\text{hydr,A}\rightarrow\text{B}}$, is the difference between the free energy change in the aqueous environment, $\Delta G_{\text{aqueous,A}\rightarrow\text{B}}$, and the free energy change in vacuum, $\Delta G_{\text{vacuum,A}\rightarrow\text{B}}$.

In this study, the replica exchange thermodynamic integration (RETI) method[41,42] was used to construct the free energy profiles, and the necessary ensemble of states were formed using Metropolis Monte Carlo sampling.[43] In the RETI protocol, standard thermodynamic integration is performed at each value of the coupling parameter λ. In addition, moves that exchange system coordinates between replica i at $\lambda = A$ of energy $E_A(i)$ and replica j at $\lambda = B$ of energy $E_B(j)$ are occasionally attempted, subject to the following acceptance test.

$$\exp[\beta([E_B(j) - E_B(i)] - [E_A(j) - E_A(i)])] \geq \text{rand}(0,1) \qquad (2)$$

The occasional exchange of coordinates between the different simulations enhances configurational sampling and, hence, convergence of the calculated properties, while the acceptance test ensures that each replica converges the simulation to the correct distribution of states.[41,42]

Implicit Solvent Model. In an implicit solvent simulation, a solvation free energy term, ΔG_{solv}, is added to the potential energy function U describing the protein−ligand complex in vacuum. In this study, the generalized Born surface area theory[44,45] (GBSA) was adopted because of its efficiency. While not the most rigorous method, the GBSA treatment of electrostatics includes the screening of charge−charge interactions by the solvent and an atomic solvation energy term. With this theory, the solvation free energy, ΔG_{solv}, is given by eq 3

$$\Delta G_{\text{solv}} = -\frac{1}{2}\left(\frac{1}{\epsilon_{\text{vac}}} - \frac{1}{\epsilon_{\text{solv}}}\right)\sum_i\sum_j \frac{q_i q_j}{\sqrt{r_{ij}^2 + B_i B_j e^{-r_{ij}^2/4B_iB_j}}} + \sum_i \delta_i SASA_i \qquad (3)$$

where δ_i is an atom-type dependent empirical term, $SASA_i$ is the solvent-accessible surface area of that atom, ϵ_{vac} and ϵ_{solv} are the dielectric constants of the vacuum and the solvent, respectively, q_i is the atomic partial charge of atom i, r_{ij} is the distance between a pair of atoms ij, and B_i is the effective Born radius of atom i. The effective Born radius can be thought of as the spherically averaged distance of a solute atom to the solvent. Several algorithms exist to calculate this quantity, and we use the pairwise descreening approximation (PDA) proposed by Hawkins et al.[46] The parameters for the GBSA model were taken from a previous study.[47] Molecular dynamics simulation of biomolecules in a GBSA solvent are very efficient and often only 4−5 times slower than vacuum simulations.[48] Owing to the nonlocal nature of the GB energy, such efficiency is lost when the GB algorithms are combined with a Monte Carlo sampling approach. This is particularly notable as the system size increases. We have, however, developed a fast GBSA scheme that can be used in MC simulations of large proteins with little or no loss of accuracy and results in simulations only 4−5 times slower than vacuum condition.[49] The method relies on the introduction of a cutoff in the GB energy calculation based on the change of Born radius of one atom after a Monte Carlo move. Further efficiency is obtained by adopting a simplified sampling potential approach, where the sampling is conducted with a cheap GBSA potential E_ξ that uses reduced cutoffs and the resulting configurations are then accepted periodically into an ensemble corresponding to the rigorous GBSA potential E_π.[50,51] This method makes use of a special acceptance test for the Monte Carlo moves

$$\exp[\beta([E_\pi(j) - E_\pi(i)] - [E_\xi(j) - E_\xi(i)])] \geq \text{rand}(0,1) \qquad (4)$$

where E_π and E_ξ denote a rigorous and approximate GBSA potential, respectively, and i,j are the coordinates of two trial configurations.

The role of the acceptance test is to remove any bias in the distribution of states that would be introduced by the cheap potential

E_ζ and ensures the converged properties are those that would be obtained with the correct GBSA potential E_π. Thus, only states that pass this acceptance test contribute to the ensemble averages. It is interesting to note the similarities between eqs 4 and 2.

Ranking of the Compounds. The method of the predictive index (PI), defined by eqs 5−7 has been proposed by Pearlman and Charifson[11] to measure the ability of a predictive method to rank a series of inhibitors according to their order of binding affinity

$$\mathrm{PI} = \frac{\sum_{j>i}\sum_{i} w_{ij}C_{ij}}{\sum_{j>i}\sum_{i} w_{ij}} \tag{5}$$

with

$$w_{ij} = |E(j) - E(i)| \tag{6}$$

and

$$\begin{aligned} C_{ij} &= -1 \qquad \text{if } \frac{E(j)-E(i)}{P(j)-P(i)} < 0 \\ &= +1 \qquad \text{if } \frac{E(j)-E(i)}{P(j)-P(i)} > 0 \\ &= 0 \qquad \text{if } P(j)-P(i) = 0 \end{aligned} \tag{7}$$

where $E(i)$ and $P(i)$ are the experimental and predicted binding free energies of compound i. This index ranges from -1 to $+1$, depending on how well the predicted ranking matches the experimental ordering. A value of $+1$ indicates perfect predictions, a value of -1 indicates predictions are always wrong, and a value of 0 arises from predictions that are completely random. The predictive index method essentially considers each pair of compounds i and j in turn. Large differences in binding free energies will have a large weight w_{ij} and successfully predicting which of the two compounds is the more potent will provide a large positive contribution to the final PI. If i and j have a small difference in binding free energy, an incorrect prediction of the most potent binder will have a minor impact on the predictive index.

System Preparation. The PDB structure of murine COX2, a structure of N9 neuraminidase, and a structure of human CDK2 extracted from a CDK2/cyclin A complex were selected as starting points for this study (PDB code 1CX2,[52] 1BJI,[53] and 2C5P[54]). For the model of COX2, hydrogen atoms had already been assigned by the crystallographers[52] and were added to the other protein models with the program reduce.[55] Previous theoretical studies and crystallographic evidence have pointed out that the conformation of the sulfonamide moiety in SC-558 in this crystal structure of COX2 is incorrect.[10] As done in a previous study,[10] the N−S−C−C torsion around this functional group was rotated to position it to interact favorably with neighboring residues, and a nearby heme was also removed as it is not involved in any direct interactions with the binding site. The protonation state of histidines was decided by visual inspection. The proteins were setup with the AMBER99 force field,[56] inhibitors were setup with the GAFF[57] force field, and the atomic partial charges were derived using the AM1/BCC method,[58] as implemented in the package AMBER8.[48] To avoid steric clashes, each protein complexed to a representative ligand (compounds **2** for COX2, **20** for neuraminidase, and **32** for CDK2) was energy minimized using the Sander module of AMBER8 and a generalized Born force field (the igb keyword was set to 1).[48] The backbone of the energy minimized protein was kept rigid for subsequent Monte Carlo simulations, which were conducted with a modified version of the ProtoMS2.1 package.[59] To reduce the computational cost, only the protein residues that have one heavy atom within 15 Å of any heavy atom of a representative ligand were retained. The resulting protein scoops consisted of 155 residues for COX2, 145 residues for neuraminidase, and 115 residues for CDK2. The ligands were modeled in their binding site on the basis of the binding mode of the energy minimized representative ligands. For the explicit solvent simulations, crystallographic waters were retained, and the complex was hydrated by a sphere of TIP4P water molecules[60] of 22 Å radius and centered on the geometric center of the ligand. To prevent evaporation, a half-harmonic potential with a 1.5 kcal·mol^{-1} force constant was applied to water molecules whose oxygen atom distance to the ligand center of geometry was greater than 22 Å. A similar sphere of water was employed to solvate the ligands in the unbound state. For the implicit solvent simulations, all the crystallographic waters, including those bridging interactions between the ligands and the protein in neuraminidase, were removed.

The reported IC_{50} value for the same ligand can vary by several orders of magnitude, depending on the assay conditions (see for instance refs 61 and 62). Thus, it is difficult to relate the IC_{50} values of inhibitors reported in different studies, and it may be better to avoid converting those to an absolute binding free energy scale. To avoid these issues, data from a single assay was used, and the experimental IC_{50} values were converted to binding free energies relative to a reference compound in each series.[63] Sets of perturbations were then selected so as to obtain calculated binding free energies with respect to one reference compound in each series. The statistical errors were determined by the batch average method for each value of λ and were then propagated across the entire λ-coordinate to yield the maximum error. This error analysis is known to overestimate the statistical error.[41] Additional simulations were also run so as to close a number of thermodynamic cycles. The thermodynamic cycles closed to within 1 kcal·mol^{-1} or less in most cases, with deviations up to 2 kcal·mol^{-1} for some cycles involving the neuraminidase inhibitors.

Monte Carlo Simulation Protocol. The bond angles and torsions for the side chains of residues within 10 Å of any heavy atom of the ligand and all the bond angles and torsions of the ligand were sampled during the simulation, with the exception of rings. The bond lengths of the protein and ligand were constrained. The total charge of the system was brought to zero by neutralizing lysine residues lying in the outer (frozen) part of the scoop (511 and 532 for COX2, 273 and 432 for neuraminidase, and 6, 34, and 56 for CDK2). A 10 Å residue based cutoff was employed in all simulations. In the generalized Born simulations, a cutoff of 20 Å for the calculation of the Born radii was applied. To make the implicit solvent simulations more efficient, the generalized Born scheme described previously was adopted.[49]

For the explicit solvent simulations in the bound state, solvent moves were attempted with a probability of 85.7%, protein side-chain move with a probability of 12.8% and solute move with a probability of 1.4%. In the unbound state, solvent moves were attempted 98.4% of the time. Replica exchange moves were attempted every 200 thousand (K) moves. The solvent was equilibrated for 20 million (M) configurations to remove any repulsive contact with the solute(s). The system was then equilibrated in one end state (typically corresponding to that of the largest ligand) for 20M further moves where solute, protein, and solvent moves were attempted. The resulting configuration was distributed over 12 values of the coupling parameter λ (0.00, 0.10, ..., 0.90, 0.95, 1.00) and equilibrated for 10M moves before collecting statistics for 30M moves.

In the implicit solvent simulations, solute moves were attempted 10% of the time, with the remainder being protein side chain moves. In the unbound state, 2K moves of equilibration were performed before 200K moves of data collection. Replica exchange moves were attempted every 6K moves. In the bound state, the system was pre-equilibrated at one value of λ for 600K moves. The resulting configuration was distributed over the 12 values of λ, and further equilibration was performed for 100K moves. Data was collected over the remaining 900K moves.

Vacuum simulations were also conducted to obtain relative hydration free energies. In this instance, each simulation performed at a value of λ was equilibrated for 2K moves and data was collected for 200K moves.

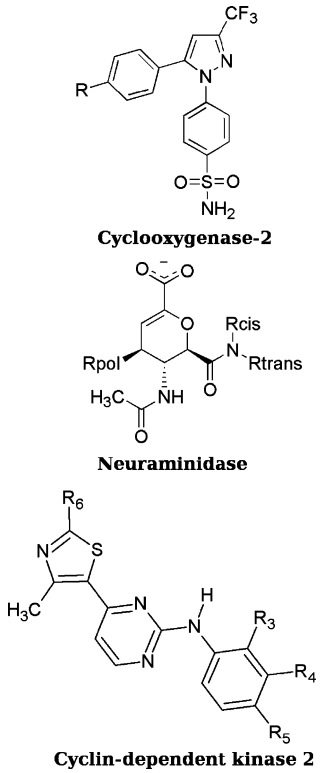

Figure 1. Structure of the inhibitors of the three proteins considered in this study.

Empirical Scoring. The modeled binding mode of each ligand was scored using the Goldscore function and the Chemscore function[64,65] in the protein−ligand docking program GOLD.[66] Only the Chemscore function was parametrized specifically to reproduce binding affinities, although Goldscore has been shown in another study to yield scores that correlate as well as Chemscore with experimental binding affinities.[67] To avoid artifacts due to the use of a different force field, the ligand scoring was done using the "local scoring" protocol described previously,[67] which only optimizes terminal groups on the protein and ligand during the searching part of the docking algorithm and then allows the ligand position and torsional degrees of freedom to relax during the SIMPLEX optimization. The terms in the scoring functions that relate to the internal energy of the ligand were not included in the score, as these terms have arbitrary reference points.

Results and Discussion

The structures of the inhibitors of the three proteins are shown in Figure 1. The perturbations selected in this work are shown in Figure 2. They are typical of those carried out in a binding free energy study, and the largest perturbation attempted involves the growth of a phenyl ring (**12** to **19**, **14** to **20**, and **17** to **18**). Relative hydration and binding free energies for each individual perturbation are listed in the Supporting Information. The simulation results will be discussed for each system independently before considering the broader lessons gained from this study.

Cyclooxygenase2. The calculated relative binding free energies with the explicit solvent protocol for the series of celecoxib derivative are shown in Figure 3a. The coefficient of determination r^2 has a value of 0.85, which suggests a respectable correlation between experiment and theory. The mean unsigned error (MUE) is 0.76 kcal·mol^{-1}, well within the so-called "chemical accuracy". The calculated PI stands at 0.96, denoting an excellent ability for the free energy simulations to rank the inhibitors according to their potency. Price and Jorgensen studied this system with a similar protocol and reported results in

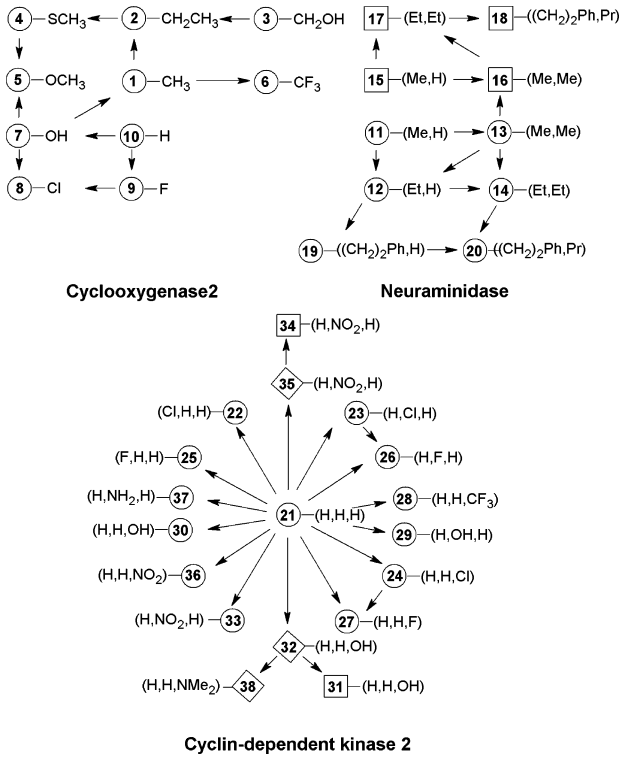

Figure 2. Series of relative binding and solvation free energy calculations performed in this study. For COX2, the substituent R is shown for each compound. For neuraminidase, the circles denote ammonium-substituted compounds, the squares denote guanadinium-substituted compounds, and substituents R_{trans},R_{cis} are listed in order for each compound. For CDK2, the circle, square, and diamond symbols denote the nature of the substituent on position R_6: methyl, monomethylated amino, or amino group, respectively. The three substituents R_3, R_4, and R_5 are listed in order for each compound.

somewhat better agreement with experimental data, with a mean unsigned error of 0.4 kcal·mol^{-1}, a coefficient of determination of 0.96, and a PI of 0.96.[10] Their more accurate simulation results may be due to the different force field that was employed (OPLS/AA[68] with CM1A[69] atomic partial charges against AMBER99/GAFF[56,57] and AM1/BCC[58] atomic partial charges for the ligands in this study). In addition, in our simulations, no water molecules were present in the binding site of COX2, while depending on the perturbation studied, one or two water molecules were present in the simulations of Price et al. There is no structural evidence supporting the presence of water molecules in this buried, hydrophobic binding site, and Price et al. could not rule out the possibility that the water molecule was an artifact of the procedure used to build the water cap in their simulations. Despite these differences, the overall ordering of the inhibitors is of similar quality.

The calculated relative binding free energies with the implicit solvent protocol for the series of celecoxib derivative are shown in Figure 3b. The MUE at 1.08 kcal·mol^{-1} is higher than that obtained for the explicit solvent simulations, and accordingly, the coefficient of determination has dropped to 0.70. However, the calculated predictive index stands at 0.96 and is identical to that obtained with the other protocols. Thus, while the predicted binding free energies deviate more from their experimental figure, the ordering of the compounds is as good as with the previous explicit water protocol.

There is a strong correlation between the hydration free energies predicted by the implicit and explicit solvent protocol ($r^2 = 0.97$). It is well-known that hydration free energies obtained by a generalized Born approach correlate very well

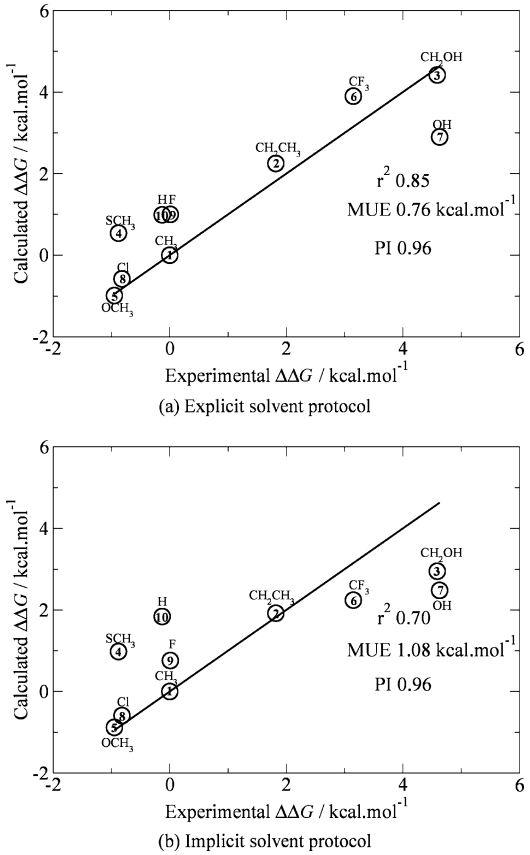

Figure 3. Simulation results for COX2. The substituent R is shown for each compound. The binding free energies are relative to compound **1**.

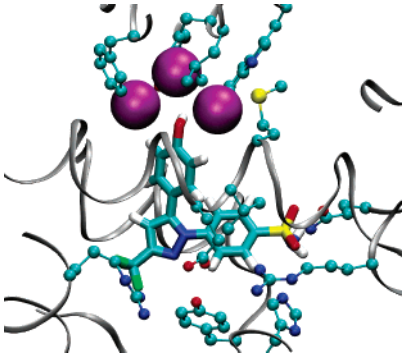

Figure 4. Model of compound **7** in the binding site of COX2, with the addition of three 2 Å radii spheres that cover approximately the small regions of the void left between the ligand and the pocket where the 5-aryl ring extends. The ligand atoms are in licorice representation, and the protein side chain atoms are in ball and stick representation. Hydrogen atoms on the amino acid side chains are not shown for clarity. Figure created with the program VMD.[79]

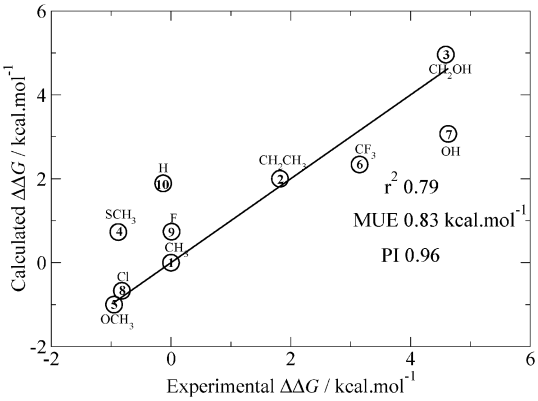

Figure 5. Modified implicit solvent protocol results for COX2. The binding free energies are relative to compound **1**.

with the hydration free energies of small molecules calculated by explicit solvent simulations.[70] Here this relationship still holds true in the case of more complex, flexible molecules. However, the correlation between the binding free energies is lower (r^2 = 0.92). This suggests that some aspects of solvation in the protein complex are not captured similarly by the two simulation methods.

Inspection of Figure 3a,b shows that compounds with polar substituents (**3** and **7**) are more stable in the implicit solvent simulations. This behavior is observed due to the treatment of desolvation by the algorithms employed to calculate the Born radii. In the binding pocket, small regions of void exist between the hydroxy group of **7** and the protein side chains. These regions of space are not occupied by water in the explicit solvent simulations. In the generalized Born protocol, however, these small regions are treated as regions of high dielectric ($\epsilon = 78.5$). As a result, the hydroxy group is still partially solvated even in the binding site. This leads to the relative stabilization of the polar hydroxy group with respect to the other, less-polar groups. Artifacts in solvation due to the presence of small pockets of high dielectric in the interior of proteins have been noted by other workers.[71−73] Here, we investigate a simple method that attempts to compensate for the improper treatment of desolvation by the generalized Born approach. By visualizing the binding site of COX2, we locate three small pockets of void that surround the 5-aryl group of the ligand, and we position a sphere of radius 2 Å in each pocket (see Figure 4 for clarity). The spheres are assigned generalized Born parameters suitable for a carbon atom. Other force field parameters are set to 0. As a result, the only impact of these spheres on the simulation is that they displace a volume of dielectric. Because these spheres make close contact with the parts of the ligands that are subject to a perturbation, they affect their Born radii, which in turn changes the generalized Born energy of the ligands. This protocol bears some resemblance to the method proposed by Liu et al.[74] to take into account the presence of small voids between the ligand and the receptor atoms.

Under these conditions, a mean unsigned error of 0.83 kcal·mol^{-1} is obtained. The predictive index is 0.96, and the coefficient of determination is 0.79. The correlation of the binding free energies between the explicit and the implicit protocols has also increased to a value practically identical (r^2 = 0.96) to the correlation observed between the hydration free energies. The better results shown in Figure 5 are obtained because the more polar compounds **3** and **7** are destabilized by 2.0 and 0.6 kcal·mol^{-1}, while the binding affinity of the other compounds is essentially unchanged. A better treatment of desolvation has therefore increased the quantitative accuracy of the implicit solvent calculations, even though they remain, overall, slightly inferior to the explicit solvent calculations. The modified generalized Born protocol depicted in Figure 4 is very simple as it only involves filling pockets of void with spheres. In principle, there are several pockets in the protein that would need to be filled that could render the protocol cumbersome. Because it amounts to a better calculation of the Born radii, similar results could be achieved more generally with a properly parametrized empirical function that rescales the Born radii obtained by the PDA algorithm.[72]

Neuraminidase. The calculated relative binding free energies with the explicit solvent protocol for the series of DANA derivatives are shown in Figure 6a. At 3.31 kcal·mol^{-1}, the

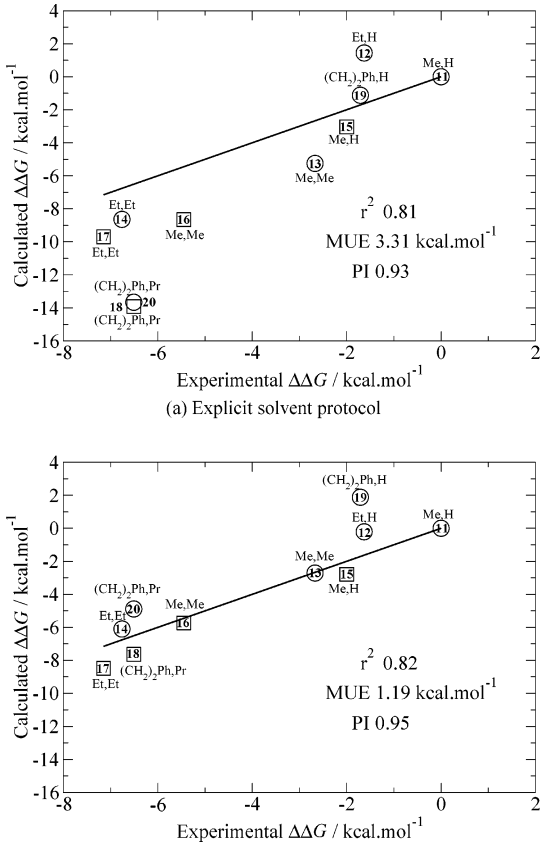

Figure 6. Simulation results for neuraminidase. The substituent R_{pol} is either an ammonium (circle) or a guanadinium group (square) and the R_{trans} and R_{cis} substituents listed for each compound. The binding free energies are relative to compound **11**.

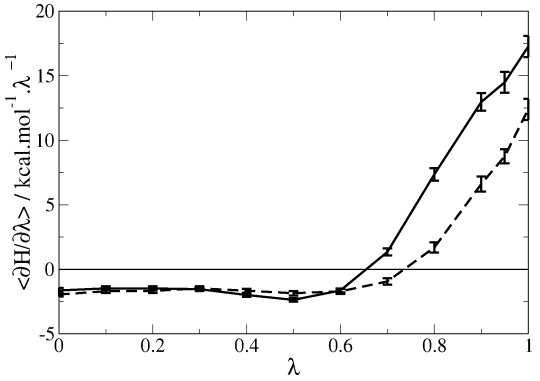

Figure 7. Free energy gradients recorded in the perturbation of **11** into **12** bound to neuraminidase. The solid line corresponds to the perturbation carried out in explicit solvent. The dashed line is the perturbation in implicit solvent. The error bars represent the standard error.

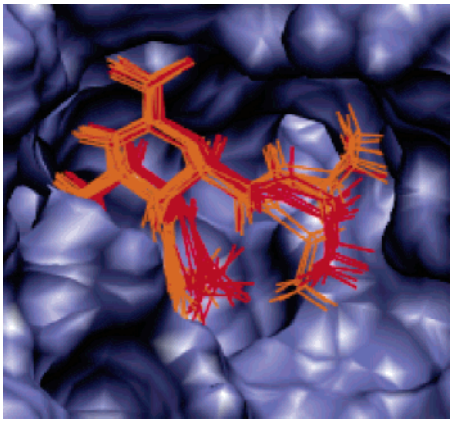

Figure 8. Overlay of 10 ligand snapshots evenly sampled from a trajectory recorded at a value of λ set to 1.00 for the perturbation of **11** into **12** bound to neuraminidase. In orange are the snapshots from the implicit solvent simulation; in red are the snapshots from the explicit solvent simulation. The solvent accessible surface area of the binding site is represented to indicate the position of the two pockets that can be occupied by the substituents R_{cis} and R_{trans}. Figure created with the program VMD.[79]

MUE is relatively high. This is essentially because the binding energies of the two potent binders **20** and **18** are overestimated. If these two compounds are excluded, the MUE drops to 2.16 kcal·mol^{-1}. The predictions for the whole set follow nonetheless closely the experimental trends and the coefficient of determination is 0.81 and the predictive index 0.93. A difficulty arises in the perturbation of compound **13** into compound **16**. Crystallographic evidence suggests that the bulkier guanadinium group of **16** must expel a crystallographic water that is present when **13** is bound.[53] This would require the annihilation of the crystallographic water prior to the perturbation of compound **13** into **16**. Such free energy calculations would require a more elaborate treatment that is beyond the scope of the present simulations. This system was however studied by Barillari with an identical force field and a much more refined simulation protocol, including a 30 Å nonbonded cutoff, protein backbone motion, and periodic boundary conditions.[75] The authors reported a binding free energy of -3.4 ± 1.1 kcal·mol^{-1}, which includes the effect of water displacement. This value was adopted for this study.

The calculated relative binding free energies with the implicit solvent protocol for the same series of perturbations are presented in Figure 6b. The results match closely the experimental trend and, surprisingly, are in much better quantitative agreement with experiment than was the case for the explicit solvent simulations. The MUE at 1.19 kcal·mol^{-1} is much lower than that obtained for the explicit solvent simulations. The coefficient of determination is 0.82 and is not significantly different from the explicit solvent results. The calculated predictive index stands at 0.95 and is nearly identical to that obtained with the explicit solvent protocol. Qualitatively, the explicit and implicit solvent protocols perform similarly. Quantitatively, the implicit solvent protocol performs significantly better. Such behavior is unexpected as the high degree of solvent exposure, and the presence of protein−ligand interactions, bridged by specific water molecules, was expected to be challenging for the implicit solvent methodology.

The relative hydration free energies obtained from the individual perturbations with the two different methodologies are strongly correlated ($r^2 = 0.95$), but the correlation is poorer for the binding free energies ($r^2 = 0.61$). The two solvent models reproduce solvation similarly only in bulk water. Such different behavior is best illustrated by a striking example. In Figure 7, the free energy gradients for the perturbation of **11** to **12** in the bound state with an implicit and explicit solvent model is reported. Initially very similar, the free energy gradients increase less in the implicit solvent simulation toward the end of the simulation, and this results in a more negative relative binding free energy. The difference in the free energy gradients is understood by observing simulation snapshots recorded at the end of the perturbation and shown in Figure 8. In the generalized Born simulations, the ethyl substituent on the amide group on the ligand can sample equally two pockets during the simulation, while it samples almost exclusively one pocket in the explicit

solvent simulation. That pocket is larger and more shallow than the previous, and the ethyl group there experiences weaker interactions with the protein than it would in the other pocket. This observation is supported by experimental binding free energy evidence.[36] Also, it can be seen that the position of the central ring of the ligand differs between the solvation models. In the explicit solvent simulations, the ring is slightly tilted compared to the ring in the implicit solvent simulations, and the amide group that bears the ethyl substituent is projected closer to the edge connecting the two pockets by about 0.8 Å.

The preceding observations suggest that the origin of the differences in the free energy gradients between the implicit and the explicit solvent approaches can be due to two factors. First, the solvation model can have such an influence on the potential energy surface that the two ligands adopt different configurations in the binding site. As a result, when the extra methyl group is grown, it experiences a different environment. The discrepancy between the two simulations would therefore be caused by force field effects. Second, in the explicit solvent simulations, conversion of the ligand between the two configurations can be hindered by the presence of several water molecules around the binding site. With simple Monte Carlo moves that randomly displace/rotate one water molecule at a time, it will be difficult for the solvent to let the ethyl group rotate freely. The differences in the free energy gradients would then be caused by an incomplete sampling of the thermally accessible states for the ligand in the binding site. A combination of these two factors is also possible.

The perturbation of **13** into **16** does not pose additional difficulties in a generalized Born force field, as a crystallographic water does not have to be annihilated. Intuitively, one would expect the implicit solvent simulation to yield results in disagreement with the observed change in binding free energy. This is because crystallographic water molecules bridging interactions between the ligand and the protein should exhibit a behavior very different from bulk water. The experimental change in binding free energy is -2.8 kcal·mol^{-1}. The generalized Born simulation yields a result of -3.03 ± 1.27 kcal·mol^{-1}, which is in very good agreement. With concern that this result might be fortuitous, the simulation length was doubled for each window. The final results, -2.86 ± 1.04 kcal·mol^{-1} is not different. It is tempting to argue that in the process of growing the guanadinium group, a volume of high dielectric space has been replaced by a low dielectric space. Thus, to some extent, the desolvation of the pocket is taken into account by the generalized Born theory. The perturbation of **13** into **16** was then carried out with the explicit solvent protocol in the absence of the crystallographic water and, hence, does not include contributions to the binding free energy due to the displacement of this water. A binding free energy of -5.14 ± 1.30 kcal·mol^{-1} was obtained. This figure is lower than those obtained with the generalized Born protocol by over 2 kcal·mol^{-1}, presumably because the replacement of the ammonium group by the bulkier guanadinium group is no longer penalized by desolvation of the pocket with this simulation protocol. It is surprising, however, that such a simple treatment of water expulsion by the implicit solvent methodology would lead to a good agreement with experiment, and in the absence of other systems to test the methodology, one must keep in mind that such agreement between the observed and the calculated binding free energy change may be fortuitous.

Cyclin-Dependent Kinase 2. The calculated relative binding free energies of the CDK2 inhibitors are plotted in Figure 9. The predictions are clearly in poor agreement with experiment.

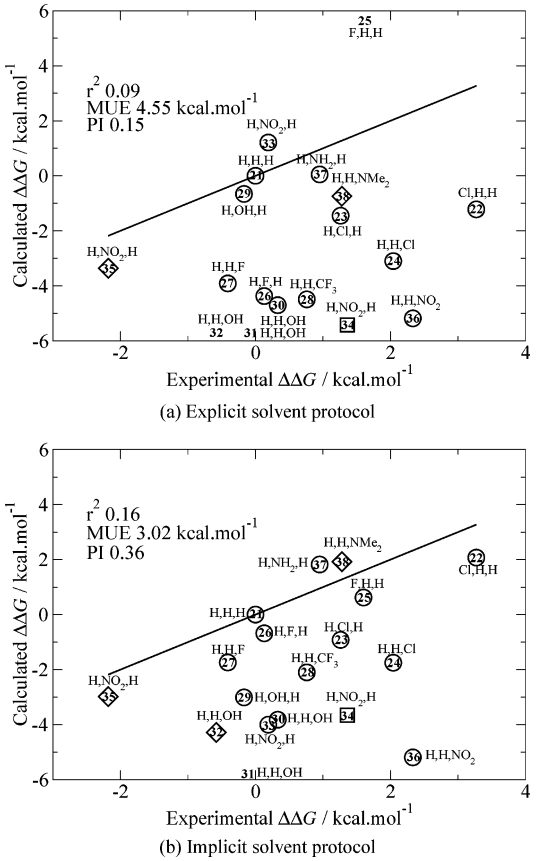

Figure 9. Simulation results for cyclin-dependent kinase 2. The substituent R_6 is a methyl (circle) or monomethylated amino (square) or amino (diamond) group, and the three substituents R_3, R_4, and R_5 are listed for each compound. The binding free energies are relative to compound **21**. Some compounds are off the scale.

For the explicit solvent simulations, the MUE is 4.55 kcal·mol^{-1}, the r^2 is 0.09, and the PI is 0.15. For the implicit solvent simulations, the MUE is 3.02 kcal·mol^{-1}, the r^2 is 0.16, and the PI is 0.36. It is interesting that the implicit solvent protocol fares better than the explicit solvent protocol. The overall performance is, however, too low to consider the predictions successful. The correlation of the hydration free energies calculated by each protocol is high ($r^2 = 0.85$), but much lower for the binding free energies ($r^2 = 0.53$), a trend already observed in neuraminidase, where the binding site is also solvent exposed.

Several factors make this series of CDK2 inhibitors a challenging test. First, there are twice as many compounds to rank as in the previous systems. Second, the span of experimental binding affinities is smaller (about 5 kcal·mol^{-1}), and half of the compounds in the series have a relative binding energy within 1 kcal·mol^{-1} of compound **21**. An accurate force field is therefore needed to obtain high quality ranking. It is also likely that other factors affect the quality of the results. The crystallographic structures of an analogue complexed to CDK2 (PDB code 2C5N) shows that when the phenyl ring is substituted by bulky groups on position R_5, the side chain of lysine 89 can adopt an alternative conformer to form a salt bridge with nearby aspartic acid 86. In the PDB structure adopted in this study, Lys89 extends over the ligand and does not interact directly with Asp86. In the explicit solvent simulations, when substitutions are made on position R_5, the side chain of Lys89 is pushed back into the solvent, but does not form a salt bridge with Asp86. In the implicit solvent simulations, Lys89 is considerably more flexible and it often adopts an

extended conformation that increases its solvation. A salt bridge between Lys89 and Asp86 is also observed, albeit infrequently. This observation suggests that a wider range of configurations of Lys89 are sampled during the implicit solvent simulations, but with the present force field, solvation of the lysine side chain is preferred over interactions with Asp86. In addition, crystallographic evidence does not clarify the precise binding mode of some inhibitors. The phenyl ring that bears substituents R_3-R_5 can be flipped by 180 degrees in the binding site, projecting the substituents into different environments. Additional calculations were run for selected perturbations with different orientations of the phenyl ring, but no significant improvement in ranking was observed.

In addition to the sampling and force field difficulties observed with this system, several assumptions have been made to create a model of the protein ligand complex (initial coordinates of the protein−ligand complex, rigidity of the protein backbone, and protonation states of histidines). It cannot be ruled out that these are directly responsible for the poor quality of the predictions and that a more careful treatment of this system would provide more accurate answers. The fact remains that the protocol employed to calculate binding free energies in this study was successfully applied to cyclooxygenase2 and neuraminidase, but not cyclin-dependent kinase 2.

Influence of Protein Flexibility. A significant difficulty in the calculation of protein−ligand relative binding free energies arises from the sampling of the many protein and solvent degrees of freedom, in addition to the ligand degrees of freedom. An implicit solvent framework reduces such complexity, but the degrees of freedom of protein side chains must still be sampled. Here we consider the impact of such protein side chain flexibility on the calculated binding free energies. The implicit solvent simulations were repeated with a completely rigid protein model, and the results were compared to those obtained with the implicit solvent simulations of a flexible protein. Because there were fewer degrees of freedom to average over, the simulations were run for only 300K moves for each window.

For COX2, the celecoxib analogues with the larger substituent **2** and **3** are more stable by about 1 kcal·mol^{-1}, while the smaller substituents are destabilized by 0.5−1 kcal·mol^{-1} (**8**, **9**, and **10**). The mean unsigned error is 1.03 kcal·mol^{-1}, the predictive index is 0.93, and the coefficient of determination is 0.67. For neuraminidase, the same trends are observed. The larger compounds are seen to bind even more favorably if no protein flexibility is considered. The MUE is 1.69 kcal·mol^{-1} because the affinity of compounds **18** and **20** are now overestimated, but the coefficient of determination is 0.82 and the predictive index still stands at a very high value of 0.96.

Before conducting the Monte Carlo simulations, the protein binding site was energy minimized in the presence of one compound in each series (**2** for COX2 and **20** for neuraminidase). Compounds **2** and **20** are the largest molecules in their series, and it is thus possible that the protein binding site has been optimized to interact with the larger compounds of the set. This could explain the observed trends.

Interestingly, simulations of a rigid model of CDK2 yield a much higher predictive index of 0.66. Inspection of the predictivity plots shows indeed a better agreement of the calculated binding affinities with experiment. The mean unsigned error is 3.45 kcal·mol^{-1} and the coefficient of determination is 0.18, but this is mainly because the binding affinity of compound **40** is largely overestimated (+23 kcal·mol^{-1}). If this compound is ignored, the MUE drops to 2.32 kcal·mol^{-1} and r^2 increases to 0.36. In general, R_5-substituted compounds are now much less favorable than when protein flexibility was enabled. This is because in the CDK2 model constructed from structure 2C5P, substituents on position R_5 form bad contacts with the side chain of Lys89. In the previous simulations, Lys89 was able to move away to accommodate the substituents but, with the present protocol, is unable to do so, and the bad contacts cannot be alleviated.

To test the influence of the initial protein model on the predictions, the calculations were repeated with the PDB structure 2C5N. No correlation or predictive power was observed (MUE, 2.32 kcal·mol^{-1}; PI, 0.04; r^2, 0.00), and R_5-substituted compounds are now predicted to bind more favorably. This is because in the CDK2 model constructed from PDB structure 2C5N, Lys89 has adopted a retracted conformation to interact with Asp86. As a result, substituents can be added to position R_5 without forming bad contacts with Lys89.

It is noteworthy that, for COX2 and neuraminidase, quantitative agreement has worsened in both cases but qualitative predictions, as judged by the coefficient of determination and the predictive index, are of similar quality. However, in the case of CDK2, a rigid protein protocol could not be applied reliably as the results depend markedly on the initial selection of a protein structure.

Convergence of the Free Energies. No specific rule dictated the choice of the number of Monte Carlo moves employed to calculate the free energy changes reported in the previous sections. It is interesting to evaluate, a posteriori, the quality of the predictions as a function of the amount of computational resources invested. This would also provide a fair comparison of the implicit and explicit solvent protocols.

In Figure 10, the MUE and the PI are plotted as a function of the time taken to complete the simulations for each protein. It assumes enough CPUs are available to run all the perturbations simultaneously. In addition, while simulations in the unbound state are very fast in the implicit solvent simulations (about 20 min), they do take longer in the explicit solvent simulations and their cost has to be considered.

The mean unsigned error converges very quickly for COX2. For the explicit solvent simulations, stable results require about 12 h of simulation. The MUE of the implicit and modified implicit simulation protocols does not evolve much after 5−7 h of simulation. For the simulations conducted with a rigid protein, the MUE is stable after 2−3 h. All methods yield a PI greater than 0.90 after only 3 h of simulation. The PI for the explicit solvent simulations is stable after about 5 h of simulation. This varies between 1 and 4 h for the implicit and modified implicit solvent simulations. The PI for the simulations conducted with a rigid protein is essentially stable after 2 h.

In neuraminidase, for the explicit solvent simulations, after 10 h of simulation, the mean unsigned error stabilizes around 2.6 kcal·mol^{-1}. However, it steadily increases after 20 h. This suggest that all the calculated individual free energy differences may not be fully converged. The opposite behavior is observed with the implicit solvent simulations and the mean unsigned error peaks at 2.3 kcal·mol^{-1} after about 3 h and then steadily decreases to about 1.0 kcal·mol^{-1}. If no protein flexibility is allowed, the mean unsigned error is seen to rapidly oscillate around 1.6 kcal·mol^{-1} after about 3 h. The PI obtained with the explicit solvent simulations is stable after about 8 h of simulation. Both implicit solvent protocols yield relatively stable PIs quickly, in about 2 h.

For CDK2, the MUE of the implicit solvent protocols is stable after a few hours and approximately 14 h for the explicit solvent protocol. The plot of the MUE for the rigid protein simulations

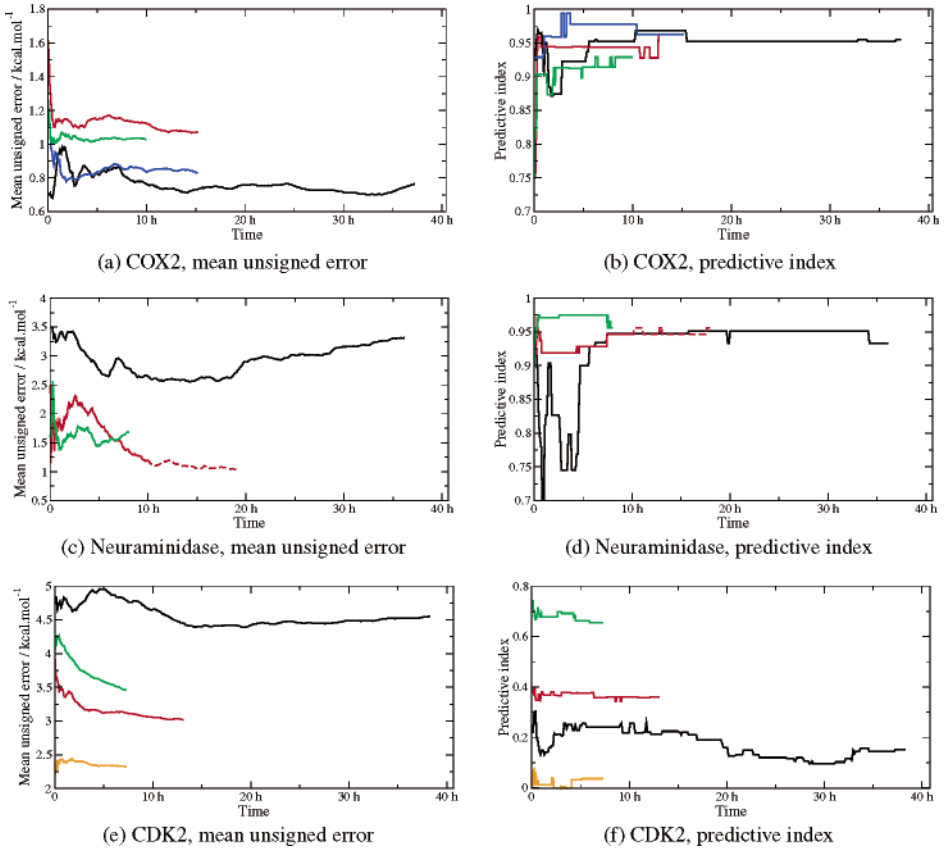

Figure 10. Convergence of the mean unsigned error and the predictive index as a function of the time taken to complete a single simulation at one λ value. In black are the explicit solvent simulations; in red are the implicit solvent simulations; in green are the implicit solvent simulations with a rigid protein; in blue are the modified implicit solvent simulations (COX2 only); and in orange are the implicit solvent simulations in an alternative rigid model of a protein (CDK2 only). The timings assume that sufficient CPUs are available to run all the perturbations simultaneously (132 CPUs for COX2, 156 CPUs for neuraminidase, 228 CPUs for cyclin-dependent kinase 2). All the timings were recorded on 2.2 GHz AMD Opteron processors. For neuraminidase, it was initially unclear if the mean unsigned error had converged at the end of the implicit solvent simulations, and these were run for twice their original duration, showing no significant deviation from the initial results (red-dashed line).

of 2C5P (green curve) is slightly misleading because it includes compound **40**, whose binding affinity is largely overestimated (+23 kcal·mol^{-1}). This is because the large group NMe_2 of **40** clashes with Lys89, and the calculated free energy change never stabilizes. If **40** is removed from the set, a much more stable curve, similar to the rigid protein simulation of 2C5N (orange curve) is observed. Finally, the PIs for all the methods are relatively stable, with the explicit solvent protocol showing the most variations.

The preceding results suggest that the implicit solvent simulations provide stable predictions more quickly than the explicit solvent simulations. It is also clear that predictions of similar quality could have been obtained at a fraction of the computational expense associated with the simulation protocols adopted for this study. Interestingly, analysis of the convergence of the free energy change for individual perturbations shows that some require longer simulation time to be well converged. Thus, the more rapid convergence of the mean unsigned error and predictive index must reflect a cancellation of errors between individual perturbations.

Importance of Configurational Averaging. Some workers have suggested incorporating solvation effects into empirical scoring functions by calculating the total electrostatic energy (Coulombic and generalized Born energy) of a single snapshot of a ligand, protein, and ligand−protein complex.[74,76] In most empirical scoring functions, a single configuration of the ligand bound in the protein binding site is usually considered. In Figure 11, the average of the total electrostatic energy of compound **11** in the unbound state and bound to neuraminidase is plotted as a function of the number of Monte Carlo moves.

It is important to remember that the configurations generated during the simulation are those that are thermally accessible to the system at a biological temperature. Because the ligand is relatively stable in the binding site and was manually docked such that it reproduces the binding mode of an analogue, the vast majority of the ligand configurations would be equivalent to acceptable docking results and could have been used to obtain a score based on that single configuration. From the plots above, it is seen that the electrostatic energies fluctuate significantly. Even in the unbound state, which consists of the ligand isolated in solution, the electrostatic energy can fluctuate by 1−2 kcal·mol^{-1} between two different blocks of simulations. As each block is the average of the electrostatic energy over 1K MC moves, the fluctuations between different snapshots would be even larger. All the points along these trajectories would be suitable candidates for scoring and yet the fluctuations of the total electrostatic energy are on the scale of a typical binding free energy. For this system, any binding score obtained from a single snapshot analysis would arguably be unreliable. A more reliable estimate can be obtained by calculating the cumulative average of the electrostatic energy, which requires averaging over several uncorrelated snapshots to converge to within a sufficient precision.[77]

Empirical Scoring Models. Could predictions of similar quality have been obtained with simple empirical scoring functions instead of the more expensive free energy protocols?

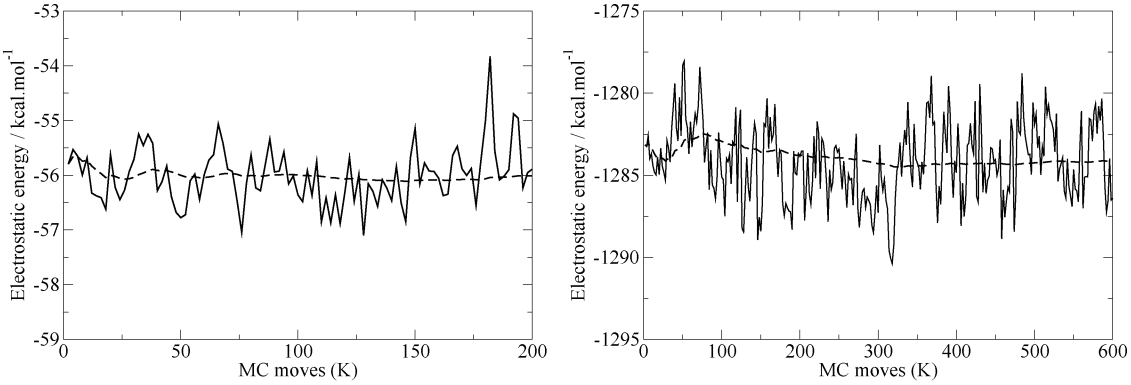

Figure 11. Fluctuations in the total electrostatic energy of the system during the simulation of compound **11**. The dashed line is the cumulative average. Each point is the average of 1K MC moves.

Because it is not always clear how to relate an empirical score to a binding free energy, only predictive indexes were measured with these methods. With the Chemscore scoring function,[64,65] PIs of 0.58 and 0.00 were obtained for COX2 and neuraminidase, respectively. With the Goldscore scoring function,[66] PIs of −0.26 and 0.75 were obtained for COX2 and neuraminidase, respectively. Because it was found for CDK2 that the free energy simulations with a rigid protein model were markedly sensitive to the initial protein coordinates, the scoring functions were applied to the two CDK2 structures used previously. Chemscore and Goldscore gives PI values of 0.49 and 0.21 for the series of compounds bound to PDB structure 2C5P. These PIs drop to 0.00 and −0.28, respectively, when the compounds are scored against PDB structure 2C5N.

In general, the PIs are much lower than those obtained with the free energy protocols. In addition, the performance of each scoring function appears to depend on the nature of the binding site. Our results agree with a study from Verdonk et al., which found in virtual screening applications that Chemscore gave better enrichments than Goldscore for lipophilic binding sites (COX2), while the opposite was found for polar sites (neuraminidase).[78] The behavior of both scoring functions on CDK2 is similar to the one observed for the free energy simulations with a rigid protein model: reasonable PIs can be obtained if the compounds are scored against a particular CDK2 structure. This raises the question of how such protein structure could be singled out from other CDK2 structures bound to ligands from the same series.

Conclusions

Protein−ligand binding free energies have been calculated by computer simulations for three different protein−ligand systems within the rigorous framework of statistical thermodynamics. The influence of water was represented by explicit and implicit solvent theories. Both methods give predictions in excellent qualitative agreement for COX2 and neuraminidase, but not CDK2. For COX2, the implicit solvent simulations are in slightly less quantitative agreement with experiment than the explicit solvent simulations. For neuraminidase and CDK2, the implicit solvent simulations are more accurate than the explicit solvent simulations. In addition, converged predictions are obtained more rapidly with the implicit solvent simulations. Compared with other binding free energy calculation methods that make use of implicit solvation,[17,28,30,74] the present methodology has the advantage of relying on a clear theoretical framework to incorporate correctly entropic effects into the computed binding free energies. It is intriguing that the implicit solvent simulations fare better in two of the three systems studied. There is no reason to expect that an implicit treatment of solvation would be in general more accurate than an explicit treatment of solvation, and the present results may be fortuitous. It may also be that the other approximations in this study (other terms in the force field equations, protein model, extent of sampling) are more important for a good prediction of binding energies than the differences between implicit and explicit theories of solvation. On the systems tested, the scoring functions were found to yield in general lower quality predictions. This emphasizes only that they cannot be used reliably for lead optimization.

When high quality predictions are obtained for a given dataset, one should wonder if the system investigated was not, in fact, too simple. For COX2, it is rather obvious that the hydroxyl analogues **3** and **7** should not be potent inhibitors, as they cannot form a hydrogen bond in the hydrophobic pocket in which they are located. If the predictions are reanalysed in the absence of these two compounds, PI values of 0.95, 0.91, and 0.95 are obtained for the explicit solvent, implicit solvent, and modified implicit solvent protocols. At 0.15, the PI of Goldscore for this modified dataset is still extremely low and for Chemscore the PI drops from 0.58 to −0.07. Thus, in this system, if the "obvious" compounds are not considered, the performance of the free energy methods does not degrade, unlike the empirical scoring functions tested here.

A number of useful methodological conclusions can also be drawn from this study. For instance, the relative hydration free energies of the ligands predicted by the generalized Born model were found to be in very good agreement with those predicted by the explicit solvent protocol, yet the binding free energies were in poorer agreement. Thus, future implicit solvent parametrizations aimed at ligand binding free energy calculations cannot rely solely on a good prediction of hydration free energies. In the case of COX2, the agreement could be improved by modifying the simulation conditions such that more accurate Born radii are calculated. However, in the case of neuraminidase, the absence of explicitly modeled waters leads to markedly different protein−ligand interactions, and it is unclear if both methodologies can be reconciled. Another important observation is that when simulations are carried out with a rigid protein model the predictions are found to converge quickly, but averaging over the ligand degrees of freedom is still necessary to obtain precise results in general. The predictive indices obtained by this approach were consistently superior to those obtained by either Chemscore or Goldscore. However, if protein flexibility is neglected, care must be taken to establish that the

predictions are not markedly affected by the choice of the protein structure and the conditions of preparation of the protein−ligand model. This remark also applies to empirical scoring functions based on a single snapshot analysis of protein−ligand complexes.

Importantly, with modern computing facilities, stable predictive indices, and mean unsigned errors can be obtained in just a few hours before all the individual perturbations are fully converged. With the implicit solvent protocol and about 150 CPUs, available to many organizations or through distributed computing projects, it seems feasible to study two dozen substituent replacements a day. This appears sufficiently fast to provide valuable insights to help optimize a promising hit into a potent lead. Thus, free energy calculations should not be considered too time-consuming to be of practical use for drug design studies.

Other limitations make free energy calculations less applicable than LIE or MM/PBSA. The present study is limited to congeneric inhibitors due to the difficulty of perturbating one ligand into another unrelated ligand. This is partly an issue of convergence of the free energies, which can be solved by running longer simulations, and complex system setup, which prevents automation. It is, therefore, important to concentrate methodological efforts on the development of rigorous free energy methods that allow fast, reliable calculations of free energy differences between structurally dissimilar ligands.

Do these results mean free energy calculations should be used more widely? We believe the answer is yes *if* crystallographic evidence or reliable docking predictions are available, the class of ligands studied do not cause significant conformational changes of the binding site, and the goal is to optimize substituent placement on a scaffold. Because of limitations in sampling algorithms, force fields, and protein models, it is unrealistic to expect high accuracy predictions to be obtained all the time, but the data presented here suggest that the calculations work better than the empirical scoring functions typically available to a modeler and at a computational cost that has become affordable.

Acknowledgment. We thank the University of Southampton and Astex Therapeutics for funding this work and EPSRC (GR/R06137/01) for providing computational resources.

Supporting Information Available: Tables of relative hydration and binding free energies for individual perturbations with associated standard errors. Closure of several thermodynamic cycles to assess the convergence of the simulations. Correlation plots of the predicted relative hydration and binding free energies between the explicit and the implicit solvent protocol. Correlation plots for free energy simulations conducted with a rigid protein and the empirical scoring functions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gane, P. J.; Dean, P. M. Recent advances in structure-based rational drug design. *Curr. Opin. Struct. Biol.* **2000**, *10*, 401−404.
(2) von Dongen, M.; Weigelt, J.; Uppenberg, J.; Schultz, J.; Wikstrom, M. Structure-based screening and design in drug discovery. *Drug Discovery Today* **2002**, *7*, 471−478.
(3) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein−small molecule docking methods *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(4) Krovat, E. M.; Steindl, T.; Langer, T. Recent advances in docking and scoring. *Curr. Comput.-Aided Drug Des.* **2005**, *1*, 93−102.
(5) Rao, S. A.; Singh, U. C.; Bash, P. A.; Kollman, P. A. Free-energy perturbation calculations on binding and catalysis after mutating Asn-155 in subtilisin. *Nature* **1987**, *328*, 551−554.
(6) Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. Monte Carlo simulations for proteins: Binding affinities for trypsin−benzamidine complexes via free-energy perturbations. *J. Phys. Chem. B* **1997**, *101*, 9663−9669.
(7) Fox, T.; Scanlan, T. S.; Kollman, P. A. Ligand binding in the catalytic antibody 17E8. A free energy perturbation calculation study. *J. Am. Chem. Soc.* **1997**, *119*, 11571−11577.
(8) Lamb, M. L.; Jorgensen, W. L. Investigations of neurotrophic inhibitors of FK506 binding protein via Monte Carlo simulations. *J. Med. Chem.* **1998**, *41*, 3928−3939.
(9) McCarrick, M. A.; Kollman, P. A. Predicting relative binding affinities of non-peptide HIV protease inhibitors with free energy perturbation calculations. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 109−121.
(10) Price, M. L. P.; Jorgensen, W. L. Analysis of binding affinities for celecoxib analogues with COX-1 and COX-2 from combined docking and Monte Carlo simulations and insight into the COX-2/COX-1 selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 9455−9466.
(11) Pearlman, D. A.; Charifson, P. S. Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(12) Udier-Blagovic, M.; Tirado-Rives, J.; Jorgensen, W. L. Structural and energetic analyses of the effects of the K103N mutation of HIV-1 reverse transcriptase on efavirenz analogues. *J. Med. Chem.* **2004**, *47*, 2389−2392.
(13) Guimaraes, C. R. W.; Boger, D. L.; Jorgensen, W. L. Elucidation of fatty acid amide hydrolase inhibition by potent α-ketoheterocycle derivatives from Monte Carlo simulations. *J. Am. Chem. Soc.* **2005**, *127*, 17377−17384.
(14) Mitchell, M. J.; McCammon, J. A. Free-energy difference calculations by thermodynamic integration−difficulties in obtaining a precise value. *J. Comput. Chem.* **1991**, *12*, 271−275.
(15) Hodel, A.; Simonson, T.; Fox, R. O.; Brunger, A. T. Conformational substates and uncertainty in macromolecular free energy calculations *J. Phys. Chem.* **1993**, *97*, 3409−3417.
(16) Pearlman, D. A. Free-energy derivatives−A new method for probing the convergence problem in free-energy calculations *J. Comput. Chem.* **1994**, *15*, 105−123.
(17) Massova, I.; Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.
(18) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330−14335.
(19) Reyes, C. M.; Kollman, P. A. Investigating the binding specificity of U1A-RNA by computational mutagenesis. *J. Mol. Biol.* **2000**, *295*, 1−6.
(20) Gouda, H.; Kuntz, I. D.; Case, D. A.; Kollman, P. A. Free energy calculations for theophylline binding to an RNA aptamer: MM-PBSA and comparison of thermodynamic integration methods *Biopolymers* **2003**, *68*, 16−34.
(21) Wang, W.; Lim, W. A.; Jakalian, A.; Wang, J.; Wang, J. M.; Luo, R.; Bayly, C. T.; Kollman, P. A. An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. *J. Am. Chem. Soc.* **2001**, *123*, 3986−3994.
(22) Fogolari, F.; Moroni, E.; Wojciechowski, M.; Baginski, M.; Ragona, L.; Molinari, H. MM/PBSA analysis of molecular dynamics simulations of bovine β-lactoglobulin: Free energy gradients in conformational transitions? *Proteins* **2005**, *59*, 91−103.
(23) Kuhn, B.; Gerber, P.; Schulz-gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **2005**, *48*, 4040−4048.
(24) Pearlman, D. A. Evaluating the molecular mechanics Poisson−Boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase. *J. Med. Chem.* **2005**, *48*, 7796−7807.
(25) Woo, H. J.; Roux, B. Calculation of absolute protein−ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825−6830.
(26) Åqvist, J.; Medina, C.; Samuelsson, J. E. New method for predicting binding-affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385−391.
(27) Wall, I. D.; Leach, A. R.; Salt, D. W.; Ford, M. G.; Essex, J. W. Binding constants of neuraminidase inhibitors: An investigation of the linear interaction energy method. *J. Med. Chem.* **1999**, *42*, 5142−5152.
(28) Zhou, R. H.; Friesner, R. A.; Ghosh, A.; Rizzo, R. C.; Jorgensen, W. L.; Levy, R. M. New linear interaction method for binding affinity calculations using a continuum solvent model. *J. Phys. Chem. B* **2001**, *105*, 10388−10397.

(29) Huang, D.; Caflisch, A. Efficient evaluation of binding free energy using continuum electrostatics solvation. *J. Med. Chem.* **2004**, *47*, 5791−5797.

(30) Carlsson, J.; Ander, M.; Nervall, M.; Åqvist, J. Continuum solvation models in the linear interaction energy method. *J. Phys. Chem. B* **2006**, *110*, 12034−12041.

(31) Simonson, T.; Carlsson, J.; Case, D. A. Proton binding to proteins: pK_a calculations with explicit and implicit solvent models. *J. Am. Chem. Soc.* **2004**, *126*, 4167−4180.

(32) Henchman, R. H.; Kilburn, J. A.; Turner, D. L.; Essex, J. W. Conformational and enantioselectivity in host−guest chemistry: The selective binding of cis amides examined by free energy calculations. *J. Phys. Chem. B* **2004**, *108*, 17571−17582.

(33) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *J. Comput. Chem.* **2002**, *23*, 517−529.

(34) Zhang, L. Y.; Gallicchio, E.; Friesner, R. A.; Levy, R. M. Solvent models for protein−ligand binding: Comparison of implicit solvent Poisson and surface generalized Born models with explicit solvent simulations. *J. Comput. Chem.* **2001**, *22*, 591−607.

(35) Penning, T. D.; Talley, J. J.; Bertenshaw, S. R.; Carter, J. S.; Collins, P. W.; Docter, S.; Graneto, M. J.; Lee, L. F.; Malecha, J. W.; Miyashiro, J. M.; Rogers, R. S.; Rogier, D. J.; Yu, S. S.; Anderson, G. D.; Burton, E. G.; Cogburn, J. N.; Gregory, S. A.; Koboldt, C. M.; Perkins, W. E.; Seibert, K.; Veenhuizen, A. W.; Zhang, Y. Y.; Isakson, P. C. Synthesis and biological evaluation of the 1,5-diarylpyrazole class of cyclooxygenase-2 inhibitors: Identification of 4-[5-(4-methylphenyl)-3-(trifluoromethyl)-1*H*-pyrazol-1-yl]benzenesulfonamide (SC-58635, Celecoxib). *J. Med. Chem.* **1997**, *40*, 1347−1365.

(36) Smith, P. W.; Sollis, S. L.; Howes, P. D.; Cherry, P. C.; Starkey, I. D; Cobley, K. N.; Weston, H.; Scicinski, J.; Merritt, A.; Whittington, A.; Wyatt, P.; Taylor, N.; Green, D.; Bethell, R.; Madar, S.; Fenton, R. J.; Morley, P. J.; Pateman, T.; Beresford, A. Dihydropyrancarboxamides related to zanamivir: A new series of inhibitors of influenza virus sialidases. 1. Discovery, synthesis, biological activity, and structure−activity relationships of 4-guanidino- and 4-amino-4*H*-pyran-6-carboxamides. *J. Med. Chem.* **1998**, *41*, 787−797.

(37) Wang, S. D.; Meades, C.; Wood, G.; Osnowski, A.; Anderson, S.; Yuill, R.; Thomas, M.; Mezna, M.; Jackson, W.; Midgley, C.; Griffiths, G.; Fleming, I.; Green, S.; Mcnae, I.; Wu, S. Y.; Mcinnes, C.; Zheleva, D.; Walkinshaw, M. D.; Fischer, P. M. 2-Anilino-4-(thiazol-5-yl)pyrimidine CDK inhibitors: Synthesis, SAR analysis, X-ray crystallography, and biological activity. *J. Med. Chem.* **2004**, *47*, 1662−1675.

(38) Kollman, P. Free-energy calculations−applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.

(39) Leach, A. R. *Molecular Modelling, Principles and Applications*; Longman: Harlow, U.K., 1996.

(40) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(41) Woods, C. J.; Essex, J. W.; King, M. A. The development of replica-exchange-based free-energy methods. *J. Phys. Chem. B* **2003**, *107*, 13703−13710.

(42) Woods, C. J.; Essex, J. W.; King, M. A. Enhanced configurational sampling in binding free-energy calculations. *J. Phys. Chem. B* **2003**, *107*, 13711−13718.

(43) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(44) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(45) Bashford, D.; Case, D. A. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129−152.

(46) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.* **1995**, *246*, 122−129.

(47) Michel, J.; Taylor, R. D.; Essex, J. W. The parameterization and validation of generalized Born models using the pairwise descreening approximation. *J. Comput. Chem.* **2004**, *25*, 1760−1770.

(48) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004.

(49) Michel, J.; Taylor, R. D.; Essex, J. W. Efficient generalized Born models for Monte Carlo simulations. *J. Chem. Theory Comput.* **2006**, *2*, 732−739.

(50) Gelb, L. D. Monte Carlo simulations using sampling from an approximate potential. *J. Chem. Phys.* **2003**, *118*, 7747−7750.

(51) Iftimie, R.; Salahub, D.; Wei, D. Q.; Schofield, J. Using a classical potential as an efficient importance function for sampling from an ab initio potential. *J. Chem. Phys.* **2000**, *113*, 4852−4862.

(52) Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; Mcdonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Miyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C. *Nature* **1996**, *384*, 644−648.

(53) Taylor, N. R.; Cleasby, A.; Singh, O.; Skarzynski, T.; Wonacott, A. J.; Smith, P. W.; Sollis, S. L.; Howes, P. D.; Cherry, P. C.; Bethell, R.; Colman, P.; Varghese, J. Dihydropyrancarboxamides related to zanamivir: A new series of inhibitors of influenza virus sialidases. 2. Crystallographic and molecular modeling study of complexes of 4-amino-4*H*-pyran-6-carboxamides and sialidase from influenza virus types A and B. *J. Med. Chem.* **1998**, *41*, 798−807.

(54) Kontopidis, G.; McInnes, C.; Pandalaneni, S. R.; McNae, I.; Gibson, D.; Mezna, M.; Thomas, M.; Wood, G.; Wang, S.; Walkinshaw, M. D.; Fischer, P. M. Differential binding of inhibitors to active and inactive CDK2 provides insights for drug design. *Chem. Biol.* **2006**, *13*, 201−211.

(55) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **1999**, *285*, 1735−1747.

(56) 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. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(57) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(58) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.

(59) Woods, C. J.; Michel, J. *ProtoMS2.1*, A Fortran program for Monte Carlo simulations of chemical systems; 2005.

(60) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(61) Laneuville, O.; Breuer, D. K.; Dewitt, D. L.; Hla, T.; Funk, C. D.; Smith, W. L. Differential inhibition of human prostaglandin endoperoxide H synthase-1 and -2 by nonsteroidal anti-inflammatory drugs. *J. Pharmacol. Exp. Ther.* **1994**, *271*, 927−934.

(62) Futaki, N.; Takahashi, S.; Yokoyama, M.; Arai, I.; Higuchi, S.; Otomo, S. NS-398, a new anti-inflammatory agent, selectively inhibits prostaglandin-G/H synthase cyclooxygenase (COX-2) activity in-vitro. *Prostaglandins* **1994**, *47*, 55−59.

(63) Cheng, Y.; Prusoff, W. H. Relationship between inhibition constant (K_i) and concentration of inhibitor which causes 50 percent inhibition (I_{50}) of an enzymatic-reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099−3108.

(64) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions. 1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.

(65) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand-receptor binding affinities and the use of Bayesian regression to improve the quality of the model. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 503−519.

(66) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.

(67) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein−ligand docking using GOLD. *Proteins* **2003**, *52*, 609−623.

(68) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(69) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. Class-IV charge models−a new semiempirical approach in quantum chemistry. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87−110.

(70) Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D. Estimation of absolute free energies of hydration using continuum methods: Accuracy of partial, charge models and optimization of nonpolar contributions. *J. Chem. Theory Comput.* **2006**, *2*, 128−139.

(71) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(72) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins* **2004**, *55*, 383−394.
(73) Swanson, J. M. J.; Mongan, J.; McCammon, J. A. Limitations of atom-centered dielectric functions in implicit solvent models. *J. Phys. Chem. B* **2005**, *109*, 14769−14772.
(74) Liu, H. Y.; Kuntz, I. D.; Zou, X. Q. Pairwise GB/SA scoring function for structure-based drug design. *J. Phys. Chem. B* **2004**, *108*, 5453−5462.
(75) Barillari, C. *The Role of Water in Protein-Ligand Interactions: Implications for Rational Drug Design*. Ph.D. Thesis, University of Southampton, Southampton, United Kindgom, 2006.
(76) Zou, X. Q.; Sun, Y. X.; Kuntz, I. D. Inclusion of solvation in ligand binding free energy calculations using the generalized-Born model. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.
(77) Woods, C. J.; King, M. A.; Essex, J. W. The configurational dependence of binding free energies: A Poisson-Boltzmann study of neuraminidase inhibitors. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 129−144.
(78) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein−ligand docking: Avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−806.
(79) Humphrey, W.; Dalke, A.; Schulten, K. VMD−Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

JM061021S