Articles

Hit Identification and Binding Mode Predictions by Rigorous Free Energy Simulations

Julien Michel[†] and Jonathan W. Essex*

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

Received May 7, 2008

The identification of lead molecules using computational modeling often relies on approximate, high-throughput approaches, of limited accuracy. We show here that, with a methodology we recently developed, it is possible to predict the relative binding free energies of structurally diverse ligands of the estrogen receptor-α using a rigorous statistical thermodynamics approach. Predictions obtained from the simulations with an explicit solvation model are in good qualitative agreement with experimental data, while simulations with implicit solvent models or rank ordering by empirical scoring functions yield predictions of lower quality. In addition, it is shown that free energy techniques can be used to select the most likely binding mode from a set of possible orientations generated by a docking program. It is suggested that the free energy techniques outlined in this study can be used to rank-order, by potency, structurally diverse compounds identified by virtual screening, de novo design or scaffold hopping programs.

Introduction

Computer programs that predict the orientation of a small molecule in a three-dimensional model of a protein structure have become an established tool for rational drug design, whether it be to analyze protein−ligand interactions of a putative drug in atomic details or to conduct a virtual screen of a large library of compounds in the search for novel small molecules that can inhibit proteins of biomedical relevance.[1−3] Numerous studies have attempted to compare the strengths and weaknesses of various docking programs and have suggested that redocking experiments with modern docking algorithms can predict correctly the binding mode of a small molecule about 80% of the time, which is sufficient to make docking a valuable tool in structure-based drug discovery.[4−7] However, accurate prediction of a ligand binding mode is not a sufficient condition to identify novel drug-like molecules. A strong binding affinity for its target is essential for a small molecule to become an effective drug and therefore successful computer-aided drug design methodologies must also be able to predict the potency of a small compound for a given biomolecule.

This issue, often referred to as the scoring problem, has traditionally been addressed with scoring functions that typically take the form of a series of equations that attempt to relate empirically the nature of the predicted protein−ligand interactions in the complex to the experimental binding affinity of the ligand.[6−8] Because scoring functions tend to be relatively efficient, they have been used extensively to process hundred of thousands of docking poses in virtual screening experiments. These studies have shown that properly parametrized scoring functions can discriminate to some degree compounds that bind a target from those that would not but their performance is often target dependent. In addition, scoring functions often have difficulties in rank ordering active compounds by their potency.[6,7] Because of these issues, considerable efforts have been devoted to improving the accuracy and reliability of binding affinity predictions for protein−ligand complexes.

The factors that govern binding affinity can be understood in the framework of statistical thermodynamics and the resulting equations allow, in principle, an exact computation of the binding free energy of a ligand.[9] In practice, these statistical mechanics equations are far too complex to be solved rigorously even with the aid of a computer and approximate solutions must be sought. For instance, molecular interactions are often described by molecular force fields rather than quantum theory.[10,11] As a result, if the force field does not reproduce well the intermolecular interactions, the accuracy of the predicted free energy will be lessened. Another longstanding issue in free energy simulations is that it can be difficult to observe over the course of a molecular dynamics (MD[a]) or Monte Carlo (MC) simulation a sufficiently large number of representative protein−ligand conformations, in which case entropic contributions to the binding free energy may be incorrectly calculated.[12,13] This problem is often seen when free energy simulations of the same protein−ligand complex, repeated independently, fail to give the same result. The solution to this predicament is simple, as precision can usually be increased by running longer simulations. However, this can often render a free energy simulation prohibitively expensive for routine use. Thus, while prediction of binding affinities by free energy computer simulations have been reported for two decades,[14] limitations in force field and sampling techniques have restricted widespread adoption of this methodology. We use the term rigorous to denote methodologies that are based on a sound statistical thermodynamic basis and only make approximations in their implementation. To overcome the

* To whom correspondence should be addressed. Phone: +44 238059 2794. Fax: +44 238059 3781. E-mail: J.W.Essex@soton.ac.uk.

[†] Current address: Department of Chemistry, Yale University, New Haven, Connecticut 06520.

[a] Abbreviations: MC, Monte Carlo; MD, molecular dynamics; ER-α-LBD, estrogen receptor-α ligand binding domain; LIE, linear interaction energy; MM/PBSA, molecular mechanics Poisson−Boltzmann surface area; GBSA, generalized Born surface area; RETI, replica exchange thermodynamic integration; rmsd, root-mean-square deviation; β, inverse temperature; M, million; K, thousand.

10.1021/jm800524s CCC: $40.75
Published on Web 10/04/2008

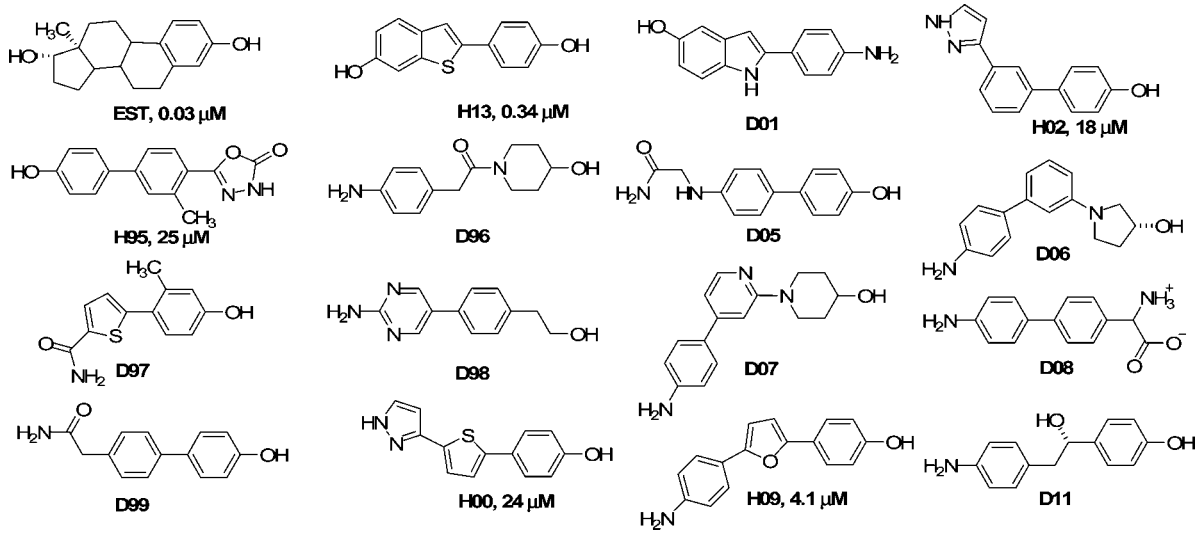

Figure 1. Structure of the compounds considered in this study. The natural agonist 17-β-estradiol is labeled EST. Compounds that were judged potent in the study of Firth-Clark et al.[33] are labeled with the letter H, while compounds that did not show a noticeable inhibition are labeled with the letter D. The IC_{50} of the active compounds reported by Firth-Clark et al. are given next to their label.

limitations of rigorous free energy simulations, approximate free energy calculation methodologies have been developed. The MM/PBSA method predicts absolute binding free energies by combining molecular mechanics interaction energies, implicit solvation free energies, and entropy estimates from normal-mode analysis.[15] The methodology appears more accurate than simpler scoring functions, but the performance can be highly variable across different protein−ligand complexes.[16,17] The linear interaction energy (LIE) method estimates absolute binding free energies from only two simulations of the solvated protein−ligand complex and the ligand in solution and is thus approximately 1 order of magnitude more efficient than rigorous free energy simulations.[18] However, LIE models need in general to be calibrated with a training set of inhibitors and thus require available experimental data to be used reliably.[19]

In recent years, the increased affordability of computing power, along with methodological advances, have renewed interest in the applicability of rigorous free energy simulations to drug design problems.[9,20,21] Different studies have demonstrated that with a cluster of commodity PCs it now appears possible to predict efficiently relative binding affinities for several dozens of structurally similar compounds over the course of a few days and with an accuracy superior to traditional scoring functions.[22−25] Key to the success of these studies is the structural similarity of the ligands investigated. It has been known for some time that free energy differences between similar compounds converge more rapidly than between dissimilar compounds. Series of similar compounds are often considered in lead optimization studies, and this is where free energy simulations usually have had an impact on drug design.[23,24]

There are, however, many instances where one would be interested in the binding affinity of sets of structurally diverse compounds. Practical free energy simulation techniques that readily allow binding affinity predictions for a range of structurally diverse inhibitors could prove particularly useful for rational drug design, for instance, when searching for novel scaffolds that could form a suitable starting point for a lead optimization program. In recent years, methodologies that allow the computation of absolute or relative binding free energies of a series of diverse compounds have been reported.[26−32] These studies have often been proofs of principle, and the methods were often validated on simple systems for which extensive experimental data were available. It is thus important to explore the applicability of these novel free energy simulation techniques to pharmaceutically relevant protein−ligand systems for which limited experimental data is available, as this is the situation typically faced by scientists engaged in rational drug design efforts.

In this paper, we demonstrate how modern free energy simulation techniques can be applied in a drug design context to identify promising scaffolds from a series of diverse compounds. We select for this a series of putative estrogen receptor α ligand binding domain (ER-α LBD) inhibitors reported by Firth-Clark et al.[33] The modulation of the activity of estrogen receptors forms the basis of many therapeutic interventions for the treatment of cancer or neurodegenerative diseases.[34]

This data set, shown in Figure 1, consists of 16 structurally diverse ligands. These compounds were generated and scored favorably by a de novo design program.[35] Subsequently, they were synthesized and tested for inhibition. Six compounds were found to be active, while the others did not show significant inhibition. Because the six actives comprise the native agonist and one structure that was already known to be active,[36] four novel actives were identified in this study. This result was encouraging, but it also showed the limitations of the scoring function employed in this study, as 10 ligands predicted to be active did not in fact show potent inhibition. We ask ourselves, could true actives have been identified in this data set using more sophisticated computational techniques? We seek answers to this question by using different free energy simulation methodologies and in doing so demonstrate how these methods can expand the range of problems molecular modelers can tackle with greater confidence.

Methods

Docking Protocol. The program GOLD was used to dock ligands into the estrogen receptor binding site.[37,38] The structure of 17-β estradiol complexed to ER-α LBD was used for this purpose (PDB 1GWR).[36] Fairly exhaustive search criteria were selected (population size of 100 individuals and a maximum of 100000 operations per docking). GOLD was configured to generate poses until the RMSDs of the top three scoring poses are within 1.5 Å. Otherwise,

a maximum of 10 poses were generated. The top scoring pose was retained as a starting point for subsequent free energy simulations. In some dockings, to bias the search toward particular binding modes, a constraint was applied to force the formation of a hydrogen bond between particular protein and ligand atoms. Previous literature evidence[39] has suggested that the Chemscore[40] scoring function produces more accurate binding modes over the Goldscore[37] scoring function for hydrophobic binding sites, and the former was therefore employed in this study.

Free Energy Calculations. Relative binding free energies can be calculated by constructing a thermodynamic cycle. The principles behind this methodology have been reviewed elsewhere.[41] The free energy change for the mutation of ligand A into B in one medium is obtained by application of the thermodynamic integration method.[42]

$$\Delta G_{\text{medium,A}\rightarrow\text{B}} = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda}\, \mathrm{d}\lambda \tag{1}$$

λ is a coupling parameter that allows the smooth transformation of the potential energy function appropriate for ligand A into a potential energy function appropriate for ligand B. To accurately estimate the free energy change, the free energy gradients are calculated at several values of the parameter λ (one simulation must therefore be conducted for each λ value) and the integral in eq 1 is then estimated by numerical integration using the trapezium rule. In this work, the free energy gradients were calculated by a finite difference scheme, i.e., $(\partial G/\partial\lambda)_\lambda = [\Delta G(\lambda \rightarrow \lambda + \Delta\lambda) - \Delta G(\lambda \rightarrow \lambda - \Delta\lambda)]/2\Delta\lambda$ provided $\Delta\lambda$ is small.[43] For each simulation run at a particular value of the coupling parameter λ, the Zwanzig equation (eq 2) was used to calculate the free energy changes $\Delta G(\lambda \rightarrow \lambda + \Delta\lambda)$ and $\Delta G(\lambda \rightarrow \lambda - \Delta\lambda)$, where $\Delta\lambda$ was set to 0.001.[44]

$$\Delta G(\lambda \rightarrow \lambda + \Delta\lambda) = -\beta^{-1} \ln\langle \exp(-\beta(U_{\lambda+\Delta\lambda} - U_\lambda))\rangle_\lambda \tag{2}$$

where the angular brackets denote an ensemble average and U_λ is the potential energy of the system simulated at λ. The potential energy function U_λ is a function of the intra- and intermolecular degrees of freedom of the simulated system and an estimate of the ensemble average in eq 2 requires the enumeration of the low energy configurations of the system. These configurations were generated using the replica exchange thermodynamic integration (RETI) method[45,46] and Metropolis Monte Carlo sampling.[47] In the RETI method, a standard Monte Carlo simulation is performed as usual at each value of the coupling parameter λ. In addition, moves that exchange system coordinates between replica i at $\lambda = A$ of energy $U_A(i)$ and replica j at $\lambda = B$ of energy $U_B(j)$ are occasionally attempted, subject to the acceptance test described by eq 3.

$$\exp(\beta([U_B(j) - U_B(i)] - [U_A(j) - U_A(i)])) \geq \text{rand}(0, 1) \tag{3}$$

The exchange of coordinates between the different simulations run in parallel has been shown to enhance 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.[45,46]

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 $\Delta G_{\text{aqueous,A}\rightarrow\text{B}}$. Relative free energies are often computed using a single topology approach whereby the molecular geometry and force field parameters of two ligands A and B are linearly interpolated to allow for the smooth alchemical transformation of ligand A into ligand B as the coupling parameter λ is increased.[48] While this approach allows the efficient computation of free energy changes between congeneric ligands (e.g., substituents on a ring), it is usually difficult to set up large perturbations that change significantly the topology of the solute (e.g., inserting a heterocycle). To overcome this limitation, a dual topology approach can be used to compute relative free energies.[49] With this technique, the two ligands of interest A and B are simulated together. No intermolecular interaction energies are computed between the two ligands and the other intermolecular energy terms are coupled to λ such that the total energy U of the system is given by eq 4

$$U(\lambda) = U_0 + \lambda U(\text{B}) + (1 - \lambda)U(\text{A}) \tag{4}$$

where U(B) is the intermolecular energy of the ligand B that is being turned on, and U(A) the intermolecular energy of the ligand A that is being turned off. U_0 accounts for all other energy terms.

We have recently shown that by using a dual topology approach in conjunction with a soft-core energy function[50] and a constraint that prevents the drift of one ligand away from the binding site as it becomes decoupled from its surrounding environment, it is possible to compute efficiently relative binding free energies between structurally diverse compounds.[30] This technique is especially advantageous over absolute binding free energy calculation schemes when there is a significant conformational or hydration change in a protein binding site upon ligand binding, as large structural rearrangement and water diffusion in and out of cavities are not sampled easily in typical MC or MD simulations. This dual topology approach was adopted here, not only to calculate the relative binding free energies of the ligands but also to predict the relative stabilities of the docked poses.

With some docking protocols, it was found that some compounds could bind in two binding modes. In this case, the relative binding affinities were computed from the most stable binding mode. In principle, a ligand able to bind in two binding modes of similar free energy would have an additional entropic contribution to its relative binding free energy of up to $-RT \ln 2$ kcal·mol^{-1}.[51] Such a contribution is however small and would not have affected any of the qualitative rankings obtained in this study.

Implicit Solvent Model. In biomolecular simulations, solvent effects are usually reproduced by explicitly considering thousands of solvent molecules in the simulation. This approach is generally considered accurate and faithful to statistical thermodynamics. It is, however, possible to model solvent effects implicitly by application of the laws of classical electrostatics.[42] This approach has often been shown to be an effective means to speed up significantly simulations with only a moderate loss of accuracy. Implicit solvent models typically decompose a hydration free energy into a sum of free energy terms arising from polar and non polar effects as shown in eq 5:

$$\Delta G_{\text{hyd}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \tag{5}$$

The latter term is often taken as proportional to the solvent accessible surface area of the system. The former term can be obtained from Poisson−Boltzmann or Generalized Born theories.[52−54] To obtain the ΔG_{polar} term by a GBSA theory requires the computation of atomic Born radii. In essence, a Born radius measures the spherically averaged distance of a solute atom to the solvent. This quantity is difficult to calculate accurately because it formally involves the solution of a complex integral over the position of all the atoms present in the system. Fast, effective approaches that estimate this quantity by a sum of pairwise terms have proven a popular alternative.[55] A drawback of these schemes is that they incorrectly attribute a bulk dielectric constant to numerous small voids in a protein that are not occupied by solvent. As a result, it is possible for a polar group to be still considered partially solvated, even in purely hydrophobic environments. This can affect predicted binding free energies by several kcal·mol^{-1}.[22] Given the hydrophobic nature of the ER-α LBD, this error was of concern. As the error on the Born radii tends to be systematic, Onufriev et al. have proposed an empirical formula that adjusts the values of the atomic Born radii to compensate for their underestimation in protein.[56] This approach was implemented and tested in this study.

System Preparation. The PDB structure of human ER-α LBD was selected as the starting point for this study (PDB code 1GWR).[36] Hydrogen atoms were added to the protein with

the program WHATIF, with which the protonation state of the histidines was also decided.[57] The protein was setup with the AMBER99 force field.[58] Models of the set of inhibitors reported by Firth-Clark et al.[33] were constructed using the program Molden.[59] Relative free energy calculations, which involve a change of net charge, are difficult to carry out reliably with the simple nonbonded cutoff employed in this study, and compounds 1-(4-amino-1-piperidinyl)-2-(4-hydroxyphenyl)-ethanone (DNP-13794) and 2-hydroxy-5-[3-(4-hydroxyphenyl)-1-oxopropyl]-benzoic acid (DNP-13810) from the work of Firth-Clark et al. were therefore not included because they would carry a net charge at physiological pH, unlike all the other inhibitors that are likely to be neutral at physiological pH.[33] Compounds that could adopt multiple tautomers were modeled in the tautomeric form used in the study of Firth-Clark et al. The ligands were setup with the GAFF[60] force field and the atomic partial charges were derived using the AM1/BCC method[61] as implemented in the package AMBER8.[62] To avoid steric clashes, the protein complexed to 17-β estradiol was energy minimized using the Sander module of AMBER8.[62] Subsequent Monte Carlo simulations were conducted with a modified version of the ProtoMS2.1 package.[63] To reduce the computational cost, only the protein residues that have one heavy atom within 15 Å of any heavy atom of the ligand 17-β-estradiol were retained. The protein scoop was inspected visually, and additional residues distant from the binding site were added to the scoop to neutralize its total charge and avoid excessive fragmentation of the protein backbone. The resulting protein scoop consisted of 123 residues. The ligands were initially modeled in the binding site in the orientations predicted by the docking program GOLD. For the explicit solvent simulations, crystallographic waters were retained and the complex was hydrated by a sphere of TIP4P water molecules[64] 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}·Å^{-2} 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 were removed. IC_{50}s have been reported for a subset of the inhibitors considered in this study. These were converted to a binding affinity relative to ligand H13 using eq 6, which is derived from the Cheng−Prusoff equation:[65]

$$\Delta\Delta G_{\mathrm{bind,A\rightarrow B}} = -RT \ln \frac{\mathrm{IC}_{50}(\mathrm{B})}{\mathrm{IC}_{50}(\mathrm{A})} \qquad (6)$$

Monte Carlo Simulation Protocol. The bond angles and torsions for the side chains of 49 residues within about 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. To increase flexibility, the main chain of the flexible protein side chains was also allowed to move by rigid body rotations and translations around the C_α atom.

A 10 Å residue-based cutoff, feathered over the last 0.5 Å, was employed in all simulations.[42] In the generalized Born simulations, a cutoff of 20 Å for the calculation of the Born radii was applied. To render the implicit solvent simulations more efficient, a generalized Born scheme described previously was adopted.[66] For the explicit solvent simulations in the bound state, solvent moves were attempted with a probability of 85.7%, protein residue moves with a probability of 12.8% and solute moves with a probability of 1.4%. In the unbound state, solvent moves were attempted 98.4% of the time. The temperature was set to 25 °C. Replica exchange moves were attempted every 200 thousand (K) moves. The solvent was equilibrated for 20 million (M) configurations to remove any repulsive contacts with the solute(s). The system was then simulated at 16 evenly spaced values of the coupling parameter λ for 70 M moves where solute, protein, and solvent moves were attempted and statistics were collected over the last 60 M moves. In the implicit solvent simulations, solute moves were attempted 10% of the time, with the remainder being protein residue moves. In the unbound state, 1 K moves of equilibration were performed before 200 K moves of data collection. Replica exchange moves were attempted every 6 K moves. In the bound state, 2 M moves were conducted at each of the 16 values of λ, and statistics were collected over the last 1.8 M moves. The parameters for the GBSA force field were taken from a previous study.[67]

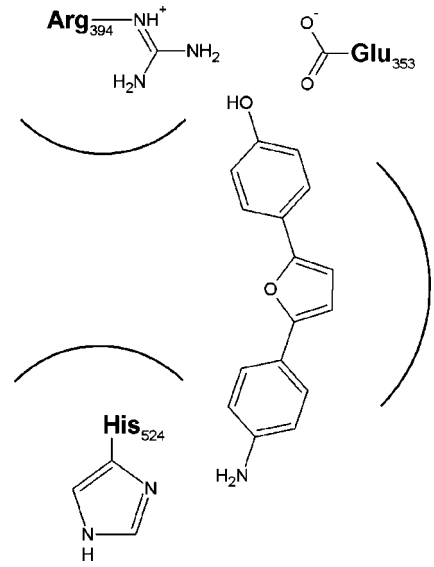

Figure 2. Representation of compound H09 bound to the ligand binding domain of the estrogen receptor α. The ligands in the data set shown in Figure 1 can donate or receive hydrogen bonds to Arg_{394}, Glu_{353}, and His_{524}. In addition, they form mainly hydrophobic interactions with the other nearby protein residues.

Unless otherwise noted, each simulation was repeated three times with a different random number seed and the free energy change and statistical error estimate was taken as the mean and the standard error of the free energies from these simulations. The free energy changes computed in this study, as well as Chemscore binding energies, are listed in the Supporting Information, and plots summarizing these data are presented in the Results section.

Results

Refinements of Binding Modes Predicted by Docking. The chemical structures of the ligands considered in this study are shown in Figure 1. It can be seen that the compounds present in this data set contain two to four rings and several different functional groups commonly encountered in drug-like molecules. All the compounds are neutral, but compound D08 is a zwitterion at physiological pH. The molecular weight of the compounds is not high, and the IC_{50}s of the actives range from low nanomolar to micromolar. The novel actives identified in this data set by Firth-Clark et al.[33] could be suitable candidates for subsequent lead optimization.

Figure 2 shows a sketch of the estrogen receptor α ligand binding domain binding site that depicts a possible binding mode of ligand H09. With the exception of 17-β-estradiol, the compounds in this data set were generated by a de novo design program,[35] and it is apparent that the program sought to satisfy hydrogen bonding with Glu_{353}, Arg_{394}, and His_{524} and identified a diverse range of hydrophobic and aromatic moieties that could link the hydrogen bond donating/receiving groups. This explains the relative diversity of scaffolds that are present in this data set but also why they tend to have roughly the same size. Even then, the structural diversity of the compounds shown in Figure 1 is large for a traditional relative free energy simulation study,

and it would have been difficult to compute binding affinities for this data set until recent methodological advances.[27,28,30]

Before binding affinities can be predicted by free energy simulations, a structure of each ligand complexed in the binding site of the ligand binding domain of the estrogen receptor α must be available. Because in this data set only the crystal structures of ligand H13 and 17-β-estradiol (EST) have been solved,[36] plausible binding modes for the 14 other compounds must be determined by a docking program.

In Figure 2, ligand H09 has its hydroxyl group oriented toward the Glu_{353}/Arg_{394} pair. Because of the relatively few possible hydrogen bond partners in the protein binding site and the nature of the compounds in this data set, it is actually possible to generate a plausible binding mode by flipping ligand H09 by 180° such that its hydroxyl group interacts with His_{524}. Apart from ligand H13 and EST, for whom a crystal structure has been solved, a similar situation arises for all other ligands in the data set (compound D08 does not have a hydroxyl group, but the same situation arises for the orientation of its amino acid group). This feature of the binding site of the estrogen receptor α is well documented. For instance, crystallographic data shows that the orientation of raloxifene, a selective estrogen receptor modulator that differs from H13 by the addition of a side chain to the benzothiophene ring,[68] is flipped by 180° compared with H13.[36,69]

To simplify the discussion, we arbitrarily decide that the pair Glu_{353}/Arg_{394} is at the "top" of the binding site and His_{524} at the "bottom". The first orientation where the hydroxyl group (or amino acid group for compound D08) is oriented toward Glu_{353}/Arg_{394} will be referred to as "top", and the second orientation where the hydroxyl group (or amino acid group for compound D08) is oriented toward His_{524} will be referred as "bottom".

To assess whether the free energy simulations can correctly identify the crystallographic binding mode for the two ligands EST and H13 for whom a structure is available, poses in top or bottom orientations of the phenol hydroxyl group were generated with the program GOLD. The difference in binding free energy between these two alternative binding geometries was then calculated using the dual-topology free energy methodology described earlier. Experimentally, both compounds are known to adopt a top orientation in the crystal structures.

EST is larger than any other ligand in the data set and does not fit well in a bottom orientation. Accordingly, the top pose is favored by Chemscore by 3.7 score units and by a large free energy difference of about 34 and 31 kcal·mol^{-1} with the implicit and explicit solvent protocols. This is in agreement with the experimentally observed binding mode.

Interesting results were obtained for ligand H13. GOLD predicts that the top pose is favorable by 1.4 score units. However the implicit solvent simulations predict that the bottom pose is favored by 0.8 ± 0.3 kcal·mol^{-1}, while the explicit solvent simulations predict the top pose to be favored by 0.6 ± 0.4 kcal·mol^{-1}. Therefore, given the limits in accuracy and precision of the present simulations, it is difficult to assign unambiguously a single binding mode to H13. Because crystallographic studies have suggested that, depending on the nature of an additional substituent, this scaffold can bind in a top or bottom orientation, it is perhaps not surprising that the free energy simulations do not indicate a strong preference for either orientation. If the bottom orientation of this scaffold was strongly disfavored, then analogues adopting this binding mode would not be expected to be potent ligands.[69]

Table 1. Orientation of Each Ligand in the Binding Site of the LBD of ER-α as Predicted by GOLD or by Free Energy Simulations[a]

compd	GOLD/Chemscore	Δ*G*, explicit	Δ*G*, implicit
H95	both	top	top
D96	top	bottom	bottom
D97	top	bottom	both
D98	top	bottom	bottom
D99	top	top	top
H00	top	top	top
D01	top	top	both
H02	top	bottom	bottom
D05	top	top	top
D06	top	bottom	bottom
D07	top	top	top
D08	top	top	top
H09	top	bottom	bottom
D11	top	top	top

[a] If the top and bottom orientations differed by less than 0.5 score units or 1 kcal·mol^{-1}, both orientations were deemed possible.

The remaining compounds were then docked using the program GOLD. The best scoring pose was assigned to be top or bottom depending on the orientation of the ligand hydroxyl group. Table 1 shows that GOLD predicted that all of the ligands, regardless of their chemical structure, would bind with their hydroxyl group oriented toward the top of the binding site, with the exception of H95 for which the docked energies of a top and bottom pose were within 0.5 score units, and hence the docking was judged to be indecisive.

To assess the predictions of the docking program, the relative free energy of the top and bottom binding modes for each ligand was computed in implicit and explicit solvent. The data in Table 1 shows that the predictions of the free energy simulations differ from Chemscore in a number of instances. Using implicit solvent, five compounds are predicted to bind with their hydroxyl group at the bottom of the binding site and the relative free energies of the top and bottom binding modes for two compounds is less than 1 kcal·mol^{-1}, suggesting they could bind in both orientations. The explicit solvent simulations mirror this trend well but with two exceptions, ligands D97 and D01. They were predicted to bind in both orientations by the implicit solvent simulations but are found to bind only in bottom and top orientations respectively, in explicit solvent.

Prediction of Relative Binding Affinities. For each ligand, the most likely binding mode, as suggested by the relative free energies of the top and bottom orientations of each compound, was retained for the prediction of binding affinities. Given the experimental evidence and the slight preference for a top orientation as predicted by the free energy simulations, ligand H13 in its orientation observed in a crystal structure[36] was selected as a reference compound and relative binding free energies between ligand H13 and every other ligand were computed in implicit and explicit solvent. Figure 3 plots the relative binding affinities predicted by each approach and compares them with experiment. Because IC_{50}s were measured only for the compounds judged to be active, quantitative comparisons and rank ordering are only possible in the set of six actives. If the 10 other compounds are to be confidently identified as inactives, then they should have a relative binding affinity higher than those of the actives.

With the explicit solvent simulations, EST is correctly identified as the most potent ligand in this data set; however, the magnitude of its predicted binding affinity is much too negative. The rank-ordering of the actives is not very accurate either. Ligand H95 is predicted to be more potent than ligand H13, and ligand H09 is predicted to be less potent than several decoys. However, the discrimination between actives and

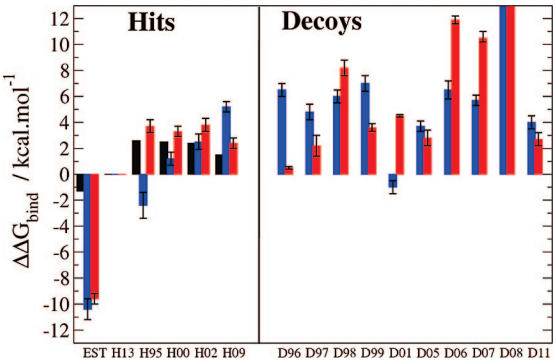

Figure 3. Binding affinities predicted by explicit and implicit solvent simulations. In black, the experimental IC_{50}s converted to relative binding affinities with eq 6, in blue, the binding affinities predicted by explicit solvent simulations, and in red, the binding affinities predicted by implicit solvent simulations. The binding affinities of D08 are off-scale and stand at 21.8 ± 0.8 and 18.5 ± 0.3 kcal·mol^{-1} for the explicit and implicit solvent simulations, respectively. All the figures are relative to ligand H13. The error bars give one standard error.

inactives is high, and five of the six actives (EST, H95, H13, H00, H02) are ranked among the top six compounds of the data set. This result is very encouraging because the IC_{50}s of several of the hits are in the micromolar range, and such compounds can be regarded as moderately potent inhibitors that may not be easily discriminated from other inactive compounds. Given that all the compounds in the data set were scored favorably by Screenscore,[33] it is clear that this rescoring by free energy simulations has markedly improved the discrimination between hits and inactives.

The results of the implicit solvent simulations differ for several ligands. EST is still predicted the most potent, and the relative binding affinities of the hits are now in better agreement with experiment than with the explicit solvent simulations. For instance, H95 is now correctly predicted to be less potent than H13, and the computed binding affinity for H09 is now in closer agreement with experiment. The six hits are now correctly rank-ordered with the implicit solvent simulations, which was not the case for the explicit solvent simulations.

On the other hand, the computed binding affinities of several decoys are either much lower (D96, D97) or much higher than with explicit solvent simulations (D98, D01, D06, D07). As a result, only three hits are among the top six compounds (EST, H13, H09) in this data set, and even then, the relative potencies of H09, D05, and D11 cannot be well discriminated because of overlapping error bars.

The binding affinity of compound D08 is very high with both simulation protocols. That such a compound would bind poorly was not unexpected, as it is the only zwitterionic compound in the data set. As a result, it is much better solvated by water and its transfer to a buried, hydrophobic binding site is therefore strongly disfavored. It is also possible that D08 would adopt a different protonation state in the binding site. Unfortunately, this possibility can not easily be accounted for with the classical potential energy function employed in this study. Firth-Clark et al. noted that Screenscore does not consider the acidity of functional groups, and this could explain why D08 scored well with Screenscore but very poorly in the present free energy simulations.[33]

In the previous section, free energy simulations suggested that a number of ligands would bind in a different orientation to the one predicted by docking. In particular, two of the hits, H02 and H09, were predicted to bind in a top orientation by Chemscore but in a bottom orientation by free energy simulations. Chemscore was not able to discriminate between top and bottom orientations for H95, while the explicit-solvent free energy simulations indicate that this compound would bind in a top orientation. We computed the binding affinity of these ligands relative to H13 in explicit solvent, but starting from the top orientation instead for H02 and H09 and bottom orientation for H95, i.e., we used the opposite geometries to those suggested by the free energy calculations. The computed binding affinities were 8.2 ± 0.3, 13.3 ± 0.6, and 13.2 ± 0.6 kcal·mol^{-1}, respectively. These binding affinities are 5.9, 8.1, and 15.6 kcal·mol^{-1} higher than those obtained when the ligands are modeled in the orientation predicted by the explicit solvent free energy simulations. With such poor relative affinities, these three hits would not have been discriminated from the decoys. In the absence of crystallographic data, these additional simulations provide further confidence in the binding modes predicted by the free energy simulations.

Table 2. Hydration Free Energies of Compounds Containing Amide Groups[a]

compound	ΔGsolv[b]	ΔGsolv$_{GBSA}$[c]	ΔGsolv$_{GBSA,\ rev}$[d]
acetamide	−9.7	-8.5 ± 0.1	-11.6 ± 0.1
N-methylacetamide	−10.1	-6.8 ± 0.1	-9.7 ± 0.1
N,*N*-dimethylacetamide	−8.5	-4.9 ± 0.1	-7.6 ± 0.1
propionamide	−9.4	-7.3 ± 0.1	-10.3 ± 0.1
D96	-6.1 ± 0.2	-1.8 ± 0.1	-5.0 ± 0.1
D97	-3.7 ± 0.2	-0.9 ± 0.1	-3.9 ± 0.1
D99	-4.7 ± 0.4	-1.4 ± 0.1	-4.5 ± 0.1
D05	-7.9 ± 0.2	-4.5 ± 0.1	-7.9 ± 0.1

[a] All the figures are in kcal·mol^{-1}. The free energies are absolute hydration free energies for the first four compounds, and relative to ligand H13 for the last four compounds. [b] From experiments (small molecules)[70] or from explicit solvent simulations (ligands). [c] Default GBSA parameter set. [d] Revised GBSA parameter set (see text).

Outliers in the Implicit Solvent Predictions. It is apparent that the implicit solvent protocol has not been as successful as the explicit solvent protocol to discriminate hits from decoys in this data set. Interestingly, as shown in Figure 3, several compounds containing an amide group score too favorably with the implicit solvent protocol but score badly with the explicit solvent protocol (D96, D97, D99, D05). Table 2 lists the absolute hydration free energies of simple amide groups, and the hydration free energy relative to H13 for some of the ligands containing an amide group, calculated using our GBSA protocol. These figures are compared to the absolute hydration free energy measured by experiment, or the hydration free energy relative to H13 predicted by explicit solvent simulations. With the GBSA parametrization employed here,[67] it is apparent that the hydration of amide groups is systematically underestimated. This would mean that the cost of desolvating an amide group with the current GBSA model would be less than in the explicit solvent simulations. This may explain then why decoys containing amide group are predicted to bind too favorably with the implicit solvent model. Because amide groups are ubiquitous in proteins, this observation was of concern and it was decided to reparameterize the GBSA model to improve predictions for the hydration free energy of amides. Table 2 shows that simple adjustments in the Born radius parameters for the oxygen atom of the amide group (the radius was reduced from 1.65 to 1.37 Å) allowed considerable improvement in the prediction of the hydration free energies. In addition, the relative hydration free energies of the ligands containing amide groups predicted with the revised GBSA parameter set were now found to be in much closer agreement with the explicit solvent results.

Another issue to be expected with the present implicit solvent model is its tendency to assign high dielectric constants to small

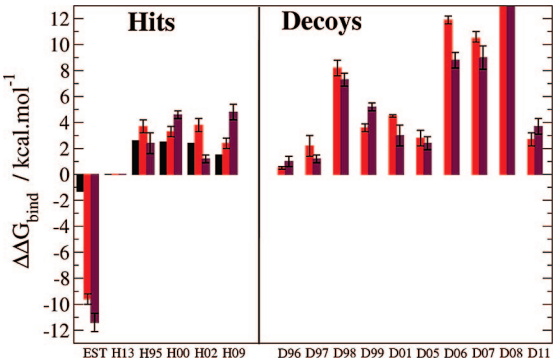

Figure 4. Binding affinities predicted by two different implicit solvent simulation protocols. In black, the experimental IC_{50}s converted to relative binding affinities with eq 6, in red, the binding affinities predicted by the default GBSA model, and in maroon, the binding affinities predicted by implicit solvent simulations with a revised GBSA model. The binding affinity of D08 is off-scale and stands at 18.5 ± 0.3 and 20.4 ± 0.1 kcal·mol^{-1} for the implicit and revised implicit solvent simulations, respectively. All the figures are relative to ligand H13. The error bars show one standard error.

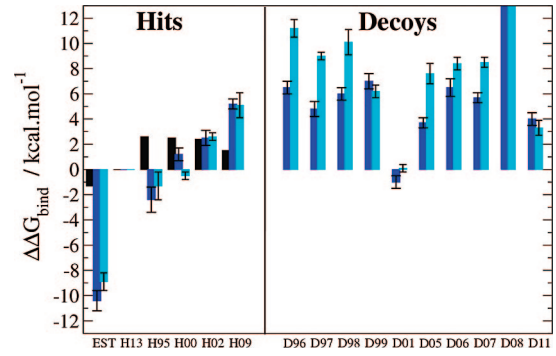

Figure 5. Binding affinities predicted by explicit solvent simulations conducted with two different protein models. In black, the experimental IC_{50}s converted to relative binding affinities with eq 6, in blue, the binding affinities predicted by the explicit solvent simulations conducted with a model constructed from 1GWR, and in cyan, the binding affinities predicted by explicit solvent simulations conducted with a model constructed from 1GWQ. The binding affinity of D08 is off-scale and stands at 21.8 ± 0.8 and 24.2 ± 0.6 kcal·mol^{-1} for the simulations in 1GWR and 1GWQ, respectively. All the figures are relative to ligand H13. The error bars show one standard error.

pockets of void that are left in the binding site but cannot in fact be occupied by water. While the modifications of Onufriev et al. tend to fix this error in an average way,[56] initial results we obtained with this protocol show that the methodology did not in fact improve the accuracy of the binding affinity predictions. Previously, it was found that by introducing in the simulation a set of spherical particles that do not have non-bonded terms but a Born radius and hence that solely displace the dielectric in a hydrophobic binding site, it was possible to improve the treatment of ligand desolvation by a Generalized Born technique and thereby improve the accuracy of the predicted binding affinities.[22] This approach was therefore adopted in this study and eight such particles were positioned in the binding site of the ER-α LBD, where pockets of void between protein residues and EST had been left.

Figure 4 plots the relative binding affinity obtained with this revised implicit solvent protocol that uses the revised GBSA parameter set and solvent displacing spheres and compares it to the results obtained with the previous implicit solvent simulations. In spite of the improved predictions of hydration free energies, the relative binding affinities are comparable to the previous simulation protocol. In fact, in light of the statistical uncertainties, the binding affinities of the decoys containing amides have not increased significantly even though they now have a more negative hydration free energy. Addition of solvent displacing spheres in the binding site was found previously to improve the prediction of binding affinities for a series of Celecoxib analogues bound to COX-2.[22] With the current data set, this protocol does not seem effective at improving the discrimination between hits and decoys. This indicates perhaps that this approach might not be systematically applicable and that, in fact, alternative algorithms that compute more accurate Born radii are necessary.

Even though the binding affinities of the decoys containing amide groups have not increased significantly, the revised implicit solvent protocol discriminate better hits from decoys in this data set. This is because H02 scores lower, and now four hits score among the top six compounds (EST, H13, H02, H95) although H95 cannot be well discriminated from D01 and D05 in light of the statistical uncertainty.

It is interesting that the modified implicit solvent protocol has improved the prediction of hydration free energies, but this improvement has not translated into better predictions of binding affinities or closer agreement with the explicit solvent results. In fact, we have noted previously that strong correlation between hydration free energies predicted by explicit or implicit solvent does not necessarily imply a strong correlation between binding free energies predicted by both techniques.[22] Similarly, we have shown that accurate prediction of hydration free energies by a GBSA model does not necessarily imply a correct treatment of the energetic of intermolecular interactions.[67] The current findings question which steps should to be taken to improve the accuracy and reliability of implicit solvent models.

Outliers in the Explicit Solvent Predictions. The explicit solvent simulation protocol was found to discriminate well actives from inactives in this data set. However, the binding affinity of EST is significantly overestimated and the binding affinity of compound D01 was too strong to discriminate it from the other actives.

We recall that the model of the protein binding site employed in these simulations was constructed from the crystallographic structure of ER-α complexed to EST. The protein residues in the binding site of this structure are prearranged to interact with EST. In addition, because the setup of the model of the protein binding site entails energy minimization of the structure complexed to EST, we expressed concern that such a protocol would create a bias for this particular ligand, which could explain why this compound scores very favorably.

To assess the impact of the protein setup on the results, free energy simulations in explicit solvent were repeated using a model of the protein binding site constructed from a crystallographic structure of ER-α complexed to ligand H13 (PDB code 1GWQ).[36] Figure 5 plots the results obtained with this protein structure and compares them to the results obtained previously.

It is apparent that EST still scores very well, regardless of the protein structure. This suggest that no significant bias toward this particular compound has been introduced with the present protein setup protocol and perhaps that its large computed binding affinity might be attributed to errors caused by the potential energy function.

However, several decoys score now much higher in binding affinity (D96, D97, D98, D05, D06, D07). The large, several kcal·mol^{-1}, differences in binding affinities obtained do indicate that the computed free energies depend on the selection of a

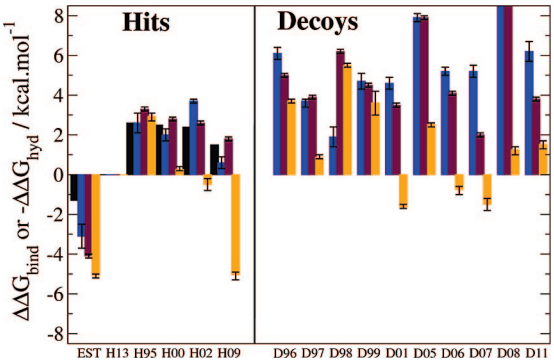

Figure 6. Discrimination of hits from decoys by plotting hydration free energies or Chemscore binding energies. In black, the experimental IC_{50}s converted to relative binding affinities with eq 6, in blue, the negative of the hydration free energies predicted by explicit solvent simulations, in maroon, the negative of the hydration free energies predicted by revised implicit solvent simulations, and in orange, the binding energies predicted by Chemscore. The relative hydration free energy of D08 is off-scale and stands at 38.4 ± 0.5 and 34.6 ± 0.1 kcal·mol^{-1} for the simulations in explicit and implicit solvent, respectively. All the figures are relative to ligand H13. The error bars show one standard error.

particular protein structure. It has been observed previously that the position of a single valine side chain in a mutant of T4-lyzozyme can affect absolute binding free energies by several kcal·mol^{-1} for a set of small molecules.[28,71] Here, overlay of the binding sites constructed from the two protein structures does not indicate a single obvious difference in side chain positions, which could explain the discrepancy in the simulation results. Rather, because the structures differ slightly in several locations (backbone position, side chain orientations) it is more likely that the differences in binding affinities arise from a combination of these small differences. In principle, the computed binding energies should not depend on the selection of a given three-dimensional model of the same protein if sufficient sampling of the protein and ligand degrees of freedom is achieved during the simulation. That such a dependency is in fact observed indicates that the sampling techniques used here are not able to generate completely identical protein−ligand configurations from both protein structures.

We note that, with the exception of compound H00, the binding affinities of the hits have not changed significantly and compound D01 still scores favorably in this new structure. As a result, while the rank-ordering of the actives is now slightly different, five still score among the top six compounds (EST, H95, H00, H13, H02). Thus, for this data set, selection of either protein structure leads to a similar discrimination of hits from decoys.

Correlation of Binding Affinities with Simpler Models. The free energy simulations in explicit solvent show good discrimination between actives and inactives in this data set, but their computational cost is significant. Because these free energy simulations generate a wealth of data that can be subsequently analyzed, it is interesting to determine if, in fact, correlations between binding affinity and other simple descriptors can be obtained.

Figure 6 shows in blue and magenta colors, the negative of the relative hydration free energies of the compounds in this data set predicted by explicit or implicit solvent techniques. With this simple property, it is apparent that most decoys score poorly while the hits tend to score well. In fact, rank-ordering according to this metric yields five hits among the top six compounds for the explicit solvent simulations (EST, H13, H09, H00, H95) and the revised implicit solvent simulations (EST, H13, H09, H02, H00). We note that the revised implicit solvent protocol fares as well as the explicit solvent protocol only with the revised GBSA parameter set. With the default GBSA parameters,[67] only three hits (EST, H13, H09) would be ranked among the top six compounds (data not shown, see Supporting Information).

It may appear at first sight surprising that the hydration free energies provide a means to discriminate actives from inactives that is as accurate as the more time-consuming computation of binding free energies. The answer to this observation lies probably in the nature of the data set. The compounds listed in Figure 1 were generated by a de novo design algorithm that seeks suitable chemical structures amenable to synthesis that satisfy hydrogen bonding and size constraints defined by the binding site. Therefore, all the optimum solutions of this algorithm yielded compounds of roughly similar size that satisfy all the important hydrogen bond constraints in the binding site. Because the binding site is fairly hydrophobic and not solvent exposed, it is likely that ligand desolvation is an important factor that controls the strength of binding. An accurate prediction of the relative hydration free energies can therefore provide a reasonable rank-ordering of the compounds in this data set. This is an interesting observation because the computation of relative hydration free energies is about twice as fast as the computation of the relative binding free energies with the explicit solvent protocol and hundreds of time faster with the implicit solvent protocol. We stress again however that only the revised implicit solvent protocol fares as well as the explicit solvent protocol and that the strong correlation between hydration and binding free energies is likely to be system dependent.

The free energy simulations were shown to discriminate actives from inactives in this data set better than the scoring function Screenscore, but this comparison might be unfair because this data set was selected precisely because Screenscore reported a majority of false positives.[33] To assess whether or not the free energy simulations do, in fact, show an improvement over more traditional scoring functions, we also plot in Figure 6 in orange the binding affinities computed with the scoring function Chemscore.

Chemscore appears to give better results than Screenscore; EST is identified as one of the most potent ligands, along with H09. However, three decoys (D01, D06, D07) score well and overall, only three actives appear among the top six compounds (EST, H09, H02).

Discussion

In drug design to date, rigorous free energy simulations have been mainly applied to lead optimization problems. A number of studies have suggested that several dozen substituents around an identical core can be screened by such techniques at a reasonable computational expense.[22−25] A difficult problem in drug design is often the selection of a chemical structure that is an adequate starting point for a lead optimization program. The aim of the present work was to explore if and how recent methodological advances in computation of binding affinities[30] would allow the screening and rank-ordering of structurally diverse scaffolds. The quality of the results reported here suggests that free energy simulations should be considered more often to assist rational drug design efforts. However, it is important to keep in mind the current limitations in order to decide where and when this technology can be applied effectively.

A source of concern for any molecular modeling study is the size and quality of the data set used to validate the different

computational techniques tested. The compounds reported by Firth-Clark et al.[33] were adopted in this work because they presented chemically diverse structures and had all been assayed in the same study. Only two compounds were excluded from our study (see the Methods section) for sound methodological reasons.

The computational expense of the present free energy simulation protocols was higher than those necessary for the prediction of relative binding affinities of structurally similar compounds. For a given ligand, a binding affinity prediction required 1300 CPU hours on a AMD Opteron 2.2 GHz processor (or 700 CPU hours for the implicit solvent protocol), but these could be spread on up to 32 processors. Thus, one data point could be acquired in ca. 1−2 days of computation. However, it was apparent that the binding affinities were not fully converged because independent repetitions of the same run gave different answers (see the Supporting Information for plots of the convergence of free energy against the number of Monte Carlo moves for selected systems). To obtain an estimate of the statistical error on the computed free energies, the binding affinity of each ligand for each protocol was computed three times and the mean binding affinity was reported. The average standard deviation was about 1.0 kcal·mol^{-1} for the explicit solvent simulations, and 0.7 kcal·mol^{-1} for the implicit solvent simulations. This standard deviation was higher than that estimated by block averaging techniques applied to a single run and is probably closer to the true statistical uncertainty of the calculated free energies. We note that, overall, the number of degrees of freedom of the ligands in this data set is not large. Previously, we have suggested that with the current protocol, the more degrees of freedom a ligand possesses, the more time-consuming it will be to converge its binding free energy to an acceptable level of precision.[30]

In a more pragmatic application, the simulation protocol could be modified to accelerate the speed at which results are obtained. For instance, if a compound scores very poorly in one run (like D08 in this study), additional runs are not necessary to flag it as inactive. The number of intermediate values of the coupling parameter λ (16 in this study) or Monte Carlo moves (70 M per λ value) could also be reduced. More refined protocols could be subsequently applied to rank order more precisely a subset of compounds whose predicted binding affinities lie in a narrow interval.

In this work, we found that free energy simulations could be used to predict the most likely orientation of a given ligand from a set of possible binding modes. Interestingly, some binding modes identified as most likely by the free energy simulations did not score favorably with Chemscore. Because the binding affinities subsequently computed from the incorrect binding mode (as identified by free energy simulations) led to poor discrimination of hits from decoys, it is important to ensure that the correct binding mode has been selected prior to conducting a free energy study. Because the binding modes of most of the compounds in this work are not known (a situation often faced in a real drug design context), we cannot determine which method is more accurate, and this work should be the focus of another more detailed study. We note that to use free energy simulations to score different binding modes, these must have been determined previously by a docking program or by the molecular modeler's intuition and/or knowledge of the binding site. The present technique cannot help locate a binding mode that has not been identified previously because the free energy simulations are too computationally intensive to drive the search for different poses by a docking program.

In many cases, we have tested the performance of a given technique by assessing its ability to score the active compounds among the top six compounds in the data set. Given the small size of the present data set, we must ask ourselves if this approach generated statistically meaningful observations. The probability of randomly picking six actives from a data set of six actives and 10 inactives (regardless of the order in which the actives are picked and picking one compound at a time) is 0.01%. The probabilities of randomly picking at least 5, 4, 3, 2, 1 actives among six compounds are 0.7%, 9.2%, 39.1%, 78.5%, and 97.4% respectively. With these considerations in mind, we can say that the discrimination between actives and inactives by explicit solvent simulations or by ordering according to the opposite of the computed hydration free energies is highly unlikely to be a chance effect but that the performance of Chemscore or the implicit solvent simulations is not guaranteed to be much more than chance.

In the present work, the explicit solvent simulation protocol outperforms the implicit solvent simulation protocol. In a previous study on another data set, both protocols were judged to give predictions of similar accuracy.[22] We note that in the absence of compounds containing an amide group, the quality of the implicit solvent predictions would have been comparable to the explicit solvent results. It is likely that a more definitive assessment of the strength of each approach will have to be deferred until the time that much larger data sets can be considered for a free energy simulation study. Alternative implicit solvent models that improved the prediction of hydration free energies for these compounds were not able to improve the prediction of binding affinities. This is an interesting observation that is of importance, as the quality of an implicit solvent model is often judged on its ability to reproduce hydration free energies.

The span of the binding affinities generated by implicit or explicit solvent simulations was larger than the range that would be expected on the basis of the ratio of the measured IC_{50}s for the actives. This trend was observed in other studies.[22−24] In addition, quantitative agreement between computed and measured binding affinities was not systematically achieved for the actives. A mean error of about 1.0 kcal·mol^{-1} is often regarded as successful in a typical relative binding energy study. Such studies were often concerned with free energy differences between structurally similar compounds that are usually easier to converge and that are less sensitive to the quality of the molecular force field, as errors in force field terms for the parts of the compounds that are similar can be expected to cancel. This situation will arise less frequently if one considers data sets of structurally diverse compounds, and with current biomolecular force fields, the average error may be expected to be higher than in studies concerned only with series of congeneric inhibitors. Also, the binding affinities were found to be sensitive to the protein structure selected, from which a model of the protein binding site was constructed. If sufficient sampling of the protein degrees of freedom is conducted, the binding affinities should be independent of the protein structure because it contains the same amino acid sequence. This problem points to the need for methodological advances that allow more effective sampling of the protein degrees of freedom.

All these observations suggest that it is challenging to systematically achieve quantitative agreement in the prediction of protein−ligand binding affinities with current simulation methodologies as applied to a set of diverse ligands. However, even if the computed binding free energies cannot be directly related to a binding affinity, in most instances, a suitably high

correlation between the two properties, or a statistically significant ability to discriminate actives from inactives, would be the most important goal. This is clearly achieved by the explicit solvent simulation protocol in this study. In addition, analysis of the simulations provide useful information that can assist ligand design; for instance, the contribution of desolvation to the binding affinities of the ligands in this data set and for this target. Given that the ligand data set is so diverse, this success is very encouraging and supports the use of free energy simulations for drug design studies.

Conclusions

The binding affinities for a set of 16 structurally diverse compounds to the ligand binding domain of the estrogen receptor α have been determined using statistical thermodynamics, classical force fields, and implicit or explicit solvent models. Free energy simulations conducted in explicit solvent were found to discriminate well actives from inactives and better than implicit solvent simulations or empirical scoring functions. The free energy simulations were also used to score possible binding modes for each compound, and it was found that the orientation favored by the free energy simulations was not always the orientation favored by the empirical scoring function Chemscore.

With computational facilities available to a typical research organization, the present free energy simulation technique can be used to screen a focused data set of chemically diverse, low molecular weight compounds. The current approach can be readily applied to rank by potency, compounds generated by virtual screening, de novo design, or scaffold hopping tools. It is thus expected that the free energy simulation methodology employed here will prove useful to identify lead molecules in rational drug design studies.

Acknowledgment. We thank the University of Southampton and BBSRC for funding this work.

Supporting Information Available: Tables of relative free energies for the different binding modes and ligands with the different solvent models. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935–949.
(2) Joseph-McCarthy, D.; Baber, J. C.; Feyfant, E.; Thompson, D. C.; Humblet, C. Lead optimization via high-throughput molecular docking. *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 264–274.
(3) Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **2006**, *11*, 580–594.
(4) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein−ligand docking: current status and future challenges. *Proteins* **2006**, *65*, 15–26.
(5) 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.
(6) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. Prediction of protein−ligand interactions. Docking and scoring: successes and gaps. *J. Med. Chem.* **2006**, *49*, 5851–5855.
(7) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
(8) Rajamani, R.; Good, A. C. Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development. *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 308–315.
(9) Gilson, M. K.; Zhou, H. X. Calculation of protein−ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
(10) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27–85.
(11) Mackerell, A. D. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
(12) 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.
(13) 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.
(14) Rao, S. N.; 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.
(15) 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.
(16) 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.
(17) 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.
(18) Aqvist, J.; Medina, C.; Samuelsson, J. E. New Method for Predicting Binding Affinity in Computer-Aided Drug Design. *Protein Eng.* **1994**, *7*, 385–391.
(19) 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.
(20) Huang, N.; Jacobson, M. P. Physics-based methods for studying protein−ligand interactions. *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 325–331.
(21) Foloppe, N.; Hubbard, R. Towards predictive ligand design with free-energy based computational methods. *Curr. Med. Chem.* **2006**, *13*, 3583–3608.
(22) Michel, J.; Verdonk, M. L.; Essex, J. W. Protein−ligand binding affinity predictions by implicit solvent simulations: a tool for lead optimization? *J. Med. Chem.* **2006**, *49*, 7427–7439.
(23) Jorgensen, W. L.; Ruiz-Caro, J.; Tirado-Rives, J.; Basavapathruni, A.; Anderson, K. S.; Hamilton, A. D. Computer-aided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 663–667.
(24) Kim, J. T.; Hamilton, A. D.; Bailey, C. M.; Domoal, R. A.; Wang, L. G.; Anderson, K. S.; Jorgensen, W. L. FEP-guided selection of bicyclic heterocycles in lead optimization for non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J. Am. Chem. Soc.* **2006**, *128*, 15372–15373.
(25) Oostenbrink, C.; van Gunsteren, W. F. Free energies of binding of polychlorinated biphenyls to the estrogen receptor from a single simulation. *Proteins* **2004**, *54*, 237–246.
(26) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical−thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* **1997**, *72*, 1047–1069.
(27) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. Direct calculation of the binding free energies of FKBP ligands. *J. Chem. Phys.* **2005**, *123*, 084108.
(28) Deng, Y. Q.; Roux, B. Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.
(29) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.
(30) Michel, J.; Verdonk, M. L.; Essex, J. W. Protein−ligand complexes: computation of the relative free energy of different scaffolds and binding modes. *J. Chem. Theory Comput.* **2007**, *3*, 1645–1655.
(31) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *J. Mol. Biol.* **2007**, *371*, 1118–1134.
(32) Oostenbrink, C.; van Gunsteren, W. F. Free energies of ligand binding for structurally diverse compounds. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750–6754.
(33) Firth-Clark, S.; Willems, H. M. G.; Williams, A.; Harris, W. Generation and selection of novel estrogen receptor ligands using the de novo structure-based design tool, SkelGen. *J. Chem. Inf. Model.* **2006**, *46*, 642–647.
(34) Deroo, B. J.; Korach, K. S. Estrogen receptors and human disease. *J. Clin. Invest.* **2006**, *116*, 561–570.
(35) Stahl, M.; Todorov, N. P.; James, T.; Mauser, H.; Boehm, H. J.; Dean, P. M. A validation study on the practical use of automated de novo design. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 459–478.
(36) Warnmark, A.; Treuter, E.; Gustafsson, J. A.; Hubbard, R. E.; Brzozowski, A. M.; Pike, A. C. W. Interaction of transcriptional intermediary factor 2 nuclear receptor box peptides with the coactivator binding site of estrogen receptor alpha. *J. Biol. Chem.* **2002**, *277*, 21862–21868.
(37) 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.

(38) 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.
(39) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; 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.
(40) 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.
(41) Kollman, P. Free-Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
(42) Leach, A. R. *Molecular Modelling, Principles and Applications*, 2nd ed.; Pearson/Prentice-Hall: Harlow, UK, 2001.
(43) Mezei, M. The Finite-Difference Thermodynamic Integration, Tested on Calculating the Hydration Free-Energy Difference between Acetone and Dimethylamine in Water. *J. Chem. Phys.* **1987**, *86*, 7084–7088.
(44) Zwanzig, R. W. High-temperature equation of state by a perturbation method. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(45) 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.
(46) 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.
(47) 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.
(48) Jorgensen, W. L.; Ravimohan, C. Monte Carlo Simulation of Differences in Free Energies of Hydration. *J. Chem. Phys.* **1985**, *83*, 3050–3054.
(49) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. Hidden Thermodynamics of Mutant Proteins: A Molecular-Dynamics Analysis. *Science* **1989**, *244*, 1069–1072.
(50) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-Shifted Scaling, a New Scaling Method for Lennard-Jones Interactions in Thermodynamic Integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
(51) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **2006**, 125.
(52) Roux, B.; Simonson, T. Implicit solvent models. *Biophys. Chem.* **1999**, *78*, 1–20.
(53) Bashford, D.; Case, D. A. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.
(54) 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.
(55) 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.
(56) 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.
(57) Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **1990**, *8*, 52–56.
(58) Wang, J. M.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(59) Schaftenaar, G.; Noordik, J. H. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123–134.
(60) Wang, J. M.; 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.
(61) 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.
(62) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(63) Woods, C. J.; Michel, J. *ProtoMS2.1*, 2006; in-house Monte Carlo software.
(64) 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.
(65) Cheng, Y.; Prusoff, W. H. Relationship between Inhibition Constant (K1) and Concentration of Inhibitor Which Causes 50 Per Cent Inhibition (I50) of an Enzymatic Reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
(66) Michel, J.; Taylor, R. D.; Essex, J. W. Efficient generalized Born models for Monte Carlo simulations. *J. Chem. Theory Comput.* **2006**, *2*, 732–739.
(67) 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.
(68) Meegan, M. J.; Lloyd, D. G. Advances in the science of estrogen receptor modulation. *Curr. Med. Chem.* **2003**, *10*, 181–210.
(69) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustafsson, J. A.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389*, 753–758.
(70) Zhu, T. H.; Li, J. B.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Density functional solvation model based on CM2 atomic charges. *J. Chem. Phys.* **1998**, *109*, 9117–9133.
(71) Mobley, D. L.; Chodera, J. D.; Dill, K. A. Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. *J. Chem. Theory Comput.* **2007**, *3*, 1231–1235.

JM800524S