Large-Scale Application of High-Throughput Molecular Mechanics with Poisson−Boltzmann Surface Area for Routine Physics-Based Scoring of Protein−Ligand Complexes

Scott P. Brown and Steven W. Muchmore*

Structural Biology, Abbott Laboratories, 100 Abbott Park Road, Abbott Park, Illinois 60064-3500

Received November 16, 2008

We apply a high-throughput formulation of the molecular mechanics with Poisson−Boltzmann surface area (htMM-PBSA) to estimate relative binding potencies on a set of 308 small-molecule ligands in complex with the proteins urokinase, PTP-1B, and Chk-1. We observe statistically significant correlation to experimentally measured potencies and report correlation coefficients for the three proteins in the range 0.72−0.83. The htMM-PBSA calculations illustrate the feasibility of procedural automation of physics-based scoring calculations to produce rank-ordered binding-potency estimates for protein−ligand complexes, with sufficient throughput for realization of practical implementation into scientist workflows in an industrial drug discovery research setting.

Introduction

The ability of numerical computation to impact decision-making processes in drug-discovery research is dependent on both the quality and timeliness of the delivered results. A robust algorithm capable of returning reliable answers to research hypotheses will have difficulty realizing impact if its time frame for delivery is significantly longer than the time scale of the decision-making that drives future directions of research. On the other hand, a computational algorithm that delivers rapid low-quality answers will have a correspondingly low likelihood of incorporation into any downstream decisions. Therein lies one of the substantial challenges in using computational analyses in industrial research, that is, the need to produce reliable results under the constraint of timely delivery.

In the arena of computationally inexpensive approaches, methodologies for docking and subsequent scoring of protein−ligand complexes have seen expansive development, with many reported improvements and successes appearing in the recent literature.[1−10] In regard to the general applicability of these methods, several critical studies have questioned the reliability and, in particular, the utility of computationally inexpensive scoring methodologies as tools to engage in *prospective* scientific inquiry.[11−13] The implicit philosophy behind empirical scoring methods is the rapid numerical evaluation of single (prepared) conformations of protein−ligand complexes, which produces a score that is simply a number obtained from a parametrized function. The quality and quantity of available experimental data are crucial to the parametrization, or training, of the empirical scoring function, which ultimately dictates the viability of the particular method. To achieve reliable prospective predictions, it must be the case that the training data thoroughly represent the particular system in question. Yet, typically, targets of interest are often novel systems that have yet to be fully characterized. Additionally, as with any numerical fit to a model, there exist certain situations in which it can be challenging to discriminate between satisfactory sets of parameters and simple chance correlation.[14]

This is not to say that empirical scoring has no suitable context (e.g., it can have utility as a tool for large-scale statistical enrichment[15]), but a lack of foundation on first principles limits the potential for further growth toward becoming predictive in a *truly prospective* sense. Underlying all empirical scoring methods is a fundamental limitation due to the lack of rigor in method construction; e.g., the very concept of free energy is inherently an *ensemble* property,[16] and thus, the general philosophy of empirical scoring is seriously flawed. A single conformation (i.e., one microstate) simply does not encode sufficient information about the distribution of thermally accessible conformations in systems whose states are distributed across high-dimensional energy landscapes (as is the case with solvated protein−ligand complexes).

From the point of view of prospective analysis, it is far more appealing to invest future development efforts into more rigorous approaches to free-energy estimations. Simply put, realistic estimates of binding potencies must involve (in some way) attempts at sampling physically reasonable representations of accessible configuration space of the protein−ligand complex, and methods built upon this concept hold greater potential for attaining generalizable predictive power. Unfortunately, in practice there exist many obstacles to realizing such approaches in typical research workflows of industrial drug-design scientists. For instance, while computationally inexpensive methodologies have a substantial tolerance for errors in input data, it is quite the opposite case as one tends toward more rigorous methods, which can be highly susceptible to minute inconsistencies in input data.

As a step toward this goal, we demonstrate here that, provided one does not require an *absolute* free energy, it is feasible (and computationally tractable) to perform automated physics-based calculations to estimate binding potencies and obtain meaningful *relative* differences in potency across a set of ligands for a particular protein target. We employ a version of the method developed by David Case and the late Peter Kollman termed molecular mechanics with Poisson−Boltzmann surface area (MM-PBSA[a]),[17] which we have modified to produce what we call a "high-throughput version" (htMM-PBSA)[18] tailored to

* To whom correspondence should be addressed. Phone: 847-938-4421. Fax: 847-938-2478. E-mail: steve.w.muchmore@abbott.com.

[a] Abbreviations: ECFP-6, extended connectivity fingerprint; MM-PBSA, molecular mechanics with Poisson−Boltzmann surface area; htMM-PBSA, high-throughput molecular mechanics with Poisson−Boltzmann surface area; GB, generalized Born; MD, molecular dynamics; MAC, mean anticipated correlation.

10.1021/jm801444x CCC: $40.75
Published on Web 04/22/2009

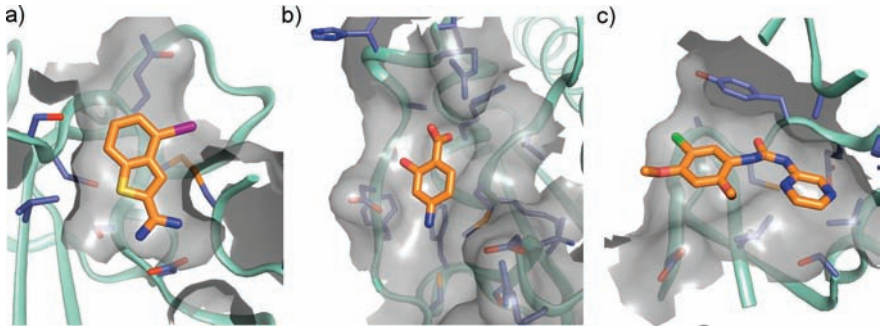

Figure 1. Depictions of the three-dimensional structure of bound inhibitors determined from in-house X-ray crystallography: (a) urokinase; (b) PTP-1B; (c) Chk-1.

rapid deployment onto a distributed-computing architecture. In this study, we describe the application of htMM-PBSA to a data set of 308 compounds across 3 protein targets and demonstrate that it is capable of producing rank-ordered binding-potency estimates with statistically significant correlation to experimental data such that it holds the potential to be used as a practical tool in industrial drug-discovery research.

Methods

Description of Data Set. We perform calculations on 308 small-molecule ligands in complex with three diverse protein targets: urokinase (75 ligands), PTP-1B (110 ligands), and Chk-1 (123 ligands). Urokinase, which is not a kinase but a serine protease, is implicated in a number of tumor-related activities.[19,20] Protein tyrosine phosphatase 1B (PTP-1B) is a regulator involved in insulin signaling[21] and has been implicated as a potential therapeutic target for treatment of type II diabetes.[22] Checkpoint kinase 1 (Chk-1) is a G2 cell-cycle checkpoint regulator,[23] thought to be a relevant target for cancer treatment through its potentiation of the effects of DNA damaging agents.[24]

Shown in Figure 1 are graphical depictions of examples of bound inhibitors for each protein target. The binding pocket for urokinase forms a narrow channel (Figure 1a), which serves to provide similar orientations for all of the urokinase compounds. In the case of PTP-1B (Figure 1b), the binding pocket forms a shallow cleft that contains the positively charged catalytic phosphate-binding site. Finally, Figure 1c shows an ATP-competitive inhibitor occupying the Chk-1 ATP binding site.

The particular targets, Uurokinase, PTP-1B, and Chk-1, were chosen because internal data exist for both the three-dimensional structure (obtained from in-house X-ray crystallography[25−27]) and assay measurement of the corresponding ligand potency (obtained from in-house enzyme-based assays[28−30]). The inhibitors used in this investigation were developed internally at Abbott (coordinates for which are provided in Supporting Information). In order to qualify for this analysis, each ligand was required to have data at acceptable levels of quality and consistency for each target, e.g., potencies obtained in a reproducible manner using consistent assay conditions. The potencies were initially obtained as IC_{50} values and then converted into pK_I units. Figure 2 shows the distribution of potencies for the small-molecule inhibitors against each target. A listing of the specific pK_I values for each compound can be found in Supporting Information.

Figure 3 shows the molecular-weight (MW) distributions of the compounds, organized by protein target. For each target, the MW and pK_I distributions span the range from early molecular-fragment hits all the way to mature lead-series compounds.

The molecules in each series span multiple chemotypes, the diversity of which we can characterize by looking at the pairwise similarities of molecules in each set. Shown in Figure 4 are the distributions of similarity values, determined by calculating an extended connectivity fingerprint (ECFP-6)[31] Tanimoto for all pairs within each target. The molecules for urokinase and PTP-1B have roughly similar distributions of molecular diversity, whereas for

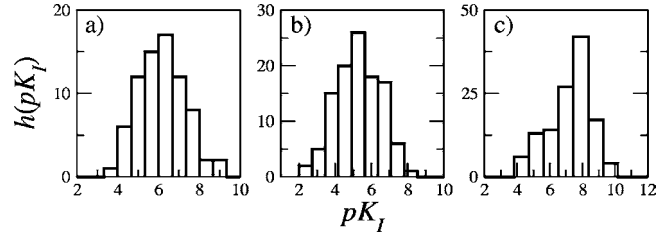

Figure 2. Histograms showing the distribution of experimentally measured potencies in pK_I units for the small-molecule ligands used in this analysis, organized by target: (a) urokinase; (b) PTP-1B; (c) Chk-1.

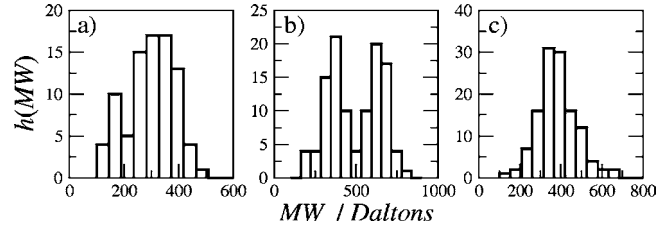

Figure 3. Histograms showing the distribution of molecular weights MW for the small-molecule ligands used in this analysis, broken down by target: (a) urokinase; (b) PTP-1B; (c) Chk-1.

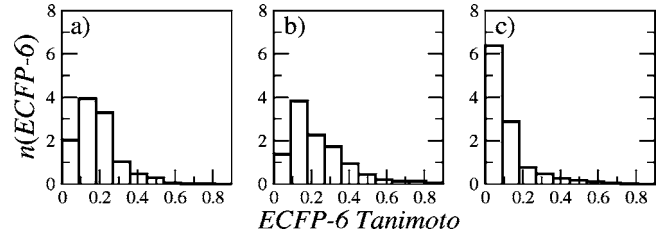

Figure 4. Normalized histograms showing the distribution of pairwise ECFP-6 Tanimoto similarities for the three ligand sets: (a) urokinase; (b) PTP-1B; (c) Chk-1. The mean values for each set are ⟨ECFP-6⟩ = 0.19 for urokinase, ⟨ECFP-6⟩ = 0.23 for PTP-1B, and ⟨ECFP-6⟩ = 0.12 for Chk-1.

Chk-1 the distribution is substantially more diverse. The mean ECFP-6 Tanimoto values are ⟨ECFP-6⟩ = 0.19 for urokinase, ⟨ECFP-6⟩ = 0.23 for PTP-1B, and ⟨ECFP-6⟩ = 0.12 for Chk-1. For comparison, the average pairwise ECFP-6 Tanimoto for lead-optimization programs is 0.36, which corresponds roughly to a 20% probability of being isoactive.[32]

Finally, in Figure 5 we characterize the distribution of molecular charge for the ligands. Each target has a different distribution in the net charge of its inhibitors, reflecting the distinct modes of interaction in the mechanisms of inhibition in each protein active site (details on the assignment of protonation states for the ligands are presented below). The urokinase inhibitors (Figure 5a) exhibit a skew in their distribution biased toward more positively charged

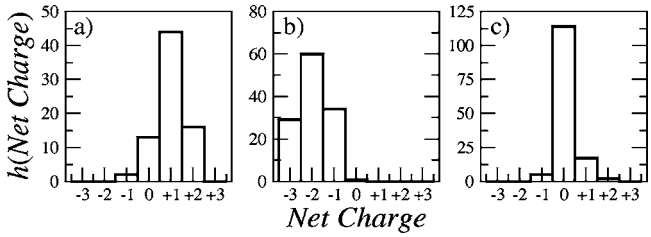

Figure 5. Histograms showing the distribution of net molecular charge for the small-molecule sets used in this analysis: (a) urokinase; (b) PTP-1B; (c) Chk-1.

ligands, while for PTP-1B (Figure 5b) an opposite skew is present, indicating a bias toward more negatively charged ligands. Finally, for Chk-1 (Figure 5c) the vast majority of inhibitors have no net charge.

Protein and Ligand Preparation. To prepare the proteins, we begin by taking the crystal-structure-coordinate file of each protein−ligand complex (acquired from in-house crystallography) and remove the ligand, water molecules and all non-peptide atoms, after which point we add hydrogens to the protein and save the resulting protein coordinates. The ligands are then processed to assign atom-centered partial charges using the AM1-BCC method.[33] AM1-BCC calculates very efficient and relatively high-quality partial charges and, as such, is ideal for automated handling of large sets of molecules. To explore reasonable protonation states for the ligands we used the pkatyper program from OpenEye Scientific Software,[34] for which a pH of 7 is assumed.

htMM-PBSA Procedure. A thorough, detailed account of the method formulation used here has been presented elsewhere.[35] Therefore, what follows will be a concise description. The modifications we make to Case and Kollman's original formulation of MM-PBSA are procedural in nature and do not alter the conceptual framework of the method. Briefly, one first performs some type of sampling of the degrees of freedom in the protein−ligand complex to generate a set of structures. This is followed by an energetic analysis on each structure in the set, from which ensemble averages are derived and subsequently used to obtain the final estimated binding potency. The two departures of our formulation from what is traditionally used in MM-PBSA are as follows. First, we use generalized Born (GB) implicit solvation (see below) during the compute-intensive ensemble-generating molecular dynamics (MD) runs, which allows the total number of MD steps required to obtain convergence to be shortened by approximately an order of magnitude compared to what is typically chosen (i.e., in the presence of explicit solvent molecules). Second, in the postproduction analysis, we use a Poisson−Boltzmann (PB) solver that employs a diffuse representation of the dielectric boundary (instead of the relatively more common discrete transition between solute and solvent). The use of this type of PB solver is significant here, as it is algorithmically more robust to the surface flexibility of the GB model used to sample conformational states. Details covering each of these methods are presented below.

Decomposition of the Free Energy. The net binding free energy accompanying the state change upon going from a solvated system with separate (unbound) protein and ligand to a solvated system containing the associated (bound) protein−ligand complex is decomposed into component contributions given by

$$\langle \Delta G_{\text{MM-PBSA}} \rangle = \langle \Delta E_{\text{MM}} \rangle + \langle \Delta G_{\text{PBSA}} \rangle - T\langle \Delta S_{\text{solute}} \rangle \quad (1)$$

where the angular brackets indicate the ensemble average of the contained quantity, ΔE_{MM} is the change in molecular mechanics (MM) energy, ΔG_{PBSA} is the net change in solvation free energy upon binding, T is the absolute temperature of the system, and ΔS_{solute} is the internal entropy change of the solutes upon binding.

In eq 1, E_{MM} is obtained from the molecular mechanics force field used in the MD simulation. It is evaluated after the completion of the MD runs, in which the MM energies are calculated in the gas phase with no distance cut-off on nonbonded interactions.

The next term on the right-hand side in eq 1, ΔG_{PBSA}, is composed of two separate parts, given by

$$\Delta G_{\text{PBSA}} = \Delta E_{\text{PB}} + \Delta G_{\text{SA}} \quad (2)$$

The ΔE_{PB} term in eq 2 is obtained by solution of the PB equation, which evaluates the energy required to charge a solute embedded within a bulk dielectric medium. Solutions are typically obtained on a grid that encompasses the solute. The PB calculations are performed in separate stages for the ligand, protein, and bound protein−ligand complex.[36] The PB solutions are obtained with the ZAP module from OpenEye Scientific Software.[34] The charges used in the PB calculations are those assigned during the ligand-preparation portion of the prodecure (see "Protein and Ligand Preparation" above). The atomic radii used are the default set implemented within ZAP (which appear to have been selected from the sets of Bondi[37] and PARSE[38] radii). An internal (solute) dielectric of 1.0 is used, with an external (bulk) dielectric value of 80.0. We use a focused grid centered on the ligand with grid spacing of 0.5 Å and a boundary spacing of 3.0 Å. This construction results in very rapid solutions of the PB equation, with timings of under 1 s per protein−ligand complex and typical errors of only a few percent.

The last term on the right-hand side of eq 2, ΔG_{SA}, is the net change in accessible, nonpolar surface area upon binding, and is calculated from

$$\Delta G_{\text{SA}} = \gamma \Delta A \quad (3)$$

where ΔA is $A^{\text{complex}} - A^{\text{protein}} - A^{\text{ligand}}$ and $\gamma = 0.005$ kcal/(mol·Å^2) is the free energy for surface formation of a hydrophobic cavity in water.[39,40] Surface-area values for the ligand, protein, and protein−ligand complex are obtained from the ZAP module for each snapshot saved from MD. The difference is then calculated to give the set of ΔA values for eq 3.

Finally, the last term on the right-hand side of eq 1, ΔS_{solute}, is the change in internal entropy due to formation of the protein−ligand complex. The solute entropy must be calculated in an additional step from the above analysis, which is typically highly computationally expensive to perform with unreliable results, e.g., using normal-mode analysis.[41] In this work we explicitly neglect changes in solute entropy.

The final potency estimate is then obtained by putting all of the pieces together in eq 1 to give a value for $\Delta G_{\text{MM-PBSA}}$. For comparison to experiment this estimated potency is then converted into units of pK_{I} using the relationship $\text{p}K_{\text{I}} = -\log_{10}[\exp(\Delta G_{\text{MM-PBSA}}/(RT))]$, where R is the universal gas constant and T is the absolute temperature (300 K).

Molecular Dynamics Procedure. To generate an ensemble of conformations, we perform MD runs using SANDER from version 8.0 of the AMBER[42] simulation package. For the protein, we use force field parameters and partial charges from the ff99 force field.[43] The force field parametrization for the small molecules comes from the generalized AMBER force field (GAFF) of Wang et al.[44] We perform the GAFF parameter assignments using the ANTECHAMBER[45] module (version 1.27). The above steps are scripted and form part of an automated procedure.

All hydrogen bonds during the MD runs are constrained using SHAKE,[46] with a time step of 2 fs and a distance cutoff of 12.0 Å for the nonbonded interactions. To account for the nontrivial effects of aqueous solvation during the MD runs, we use the GB implicit solvent model of Onufriev et al.,[47] with the internal (solute) dielectric set to 1.0 and the external (bulk) dielectric set to 80.0.

The MD procedure begins by performing a steepest-decent minimization on the prepared protein−ligand complex for 1000 steps. The final structure from minimization is passed on to an equilibration phase, in which we heat the system from 0 to 300 K over 6 ps using constant-temperature Langevin dynamics. The final structure from equilibration is then input as the initial structure into the MD production runs. We perform production MD runs for 13

ps and discard data from the first 3 ps. Over the course of the final 10 ps we save a series of 10 coordinate "snapshots" of the protein−ligand complex, taken at evenly spaced intervals in time along the trajectory. We do not perform separate MD simulations for the isolated ligand and unbound protein and hence use the MD trajectory of the protein−ligand complex as the source of solution conformations for the isolated ligand and apo protein.

By use of this construction, the total CPU time required to process a single PTP-1B structure is around 200 min. While this figure is certainly dependent on the particular system being analyzed, it is sufficiently low such that reasonable throughput can be realized, especially within a grid-computing architecture. htMM-PBSA calculations are well suited to deployment onto a coarse-grained parallel computing architecture. All results reported herein were obtained from calculations on an enterprise grid, in which idle CPU cycles are harvested from employee desktop personal computers using the freely distributed package CONDOR[48] from the Department of Computer Science at the University of Wisconsin.

X-Score Procedure. For purposes of comparing our htMM-PBSA results to a typical empirical scoring function, we have chosen the program X-Score[49] (version 1.2). The scoring function in X-Score is a linear equation with 12 adjustable parameters determined by fitting to approximately 800 protein−ligand structures with potency measurements obtained from literature reports. Preparation of the ligands for input into X-Score requires that SYBYL atom types are assigned according to the Mol2 molecular file format from Tripos, Inc.[50] X-Score was used to perform this task with a utility provided by the program. For the proteins no additional preparation steps were required (beyond what is described in "Protein and Ligand Preparation" above). The ligand Mol2 files with corresponding protein coordinate files are input into the X-Score program and numerically evaluated. The resulting output from X-Score is a direct prediction of the ligand pK_I for the particular protein input into the calculation. Many protein−ligand pairs can be evaluated this way in a short period of time, with typical timings on a 3 GHz CPU of approximately 5 complexes per second.

Results and Discussion

We measure the quality of prediction by calculating the Pearson product-moment correlation coefficient,[51] R, which is a reasonable metric in this case given the distributions of potencies. Values of R span the interval $[-1, 1]$, with those greater than $|R| \approx 0.8$ generally thought to indicate a significant degree of correlation. This is a somewhat subjective assessment; however, for the typical data set sizes processed in physics-based models (tens of data points), it does represent a statistically significant predictive probability.[52]

In Figure 6 we show the main result of this work, in which we plot the potency estimates from htMM-PBSA versus the experimentally measured potencies for the 308 protein−ligand complexes. For all protein targets investigated we obtain potency estimates with significant correlation to experiment: for urokinase we obtain a correlation coefficient of 0.78; for PTP-1B we find a correlation coefficient of 0.83; for Chk-1 we obtain a correlation of 0.72. All of the results displayed in Figure 6 were produced by calculations from a single overnight submission to our enterprise computing grid.

An inspection of the slopes of the lines fit to the data in Figure 6 reveals significant deviation from an ideal value of 1. There are a number of potential contributors to the observed slope (e.g., lack of solute-entropy changes), but for our purposes here the value of the slope is not particularly important, as we only require an ability to rank-order the list of compounds based on estimated relative binding potencies.

For the data in Figure 6, as with any set of inhibitor potencies used in method validation, it is important to check for the existence of the trivial correlation with MW. It is a common occurrence that ligands exhibit an increase in potency for a protein target with increasing MW, and any potency-ranking procedure should out-perform the MW dependence in the set. Table 1 shows the absolute value of the correlation coefficients between ligand potency and MW. Each set of compounds has a clear inverse correlation between molecular weight and potency; however, the correlations produced by htMM-PBSA are all greater than those observed with simple MW scaling. Also shown in Table 1 are the correlations produced by the empirical scoring function, X-Score, which has a performance on par with MW. The computational cost of the X-Score calculation is on the order of 1000 times less expensive than htMM-PBSA, yet it appears that its signal provides no additional information beyond what is already encoded in the simple MW dependence in the data.

A recent analysis[53] recommends inspection of the expected correlation based on the quality of the underlying data. Shown in Table 1 is a calculated "mean anticipated correlation", R_{MAC}, the full estimation of which is based on the pK_I span and total number of points in the data, as well as a knowledge of the experimental and prediction error. We report the error in the predictions next but will note here that the values for all htMM-PBSA correlation coefficients are consistent with the quality of the underlying data. This is seen by noting that the

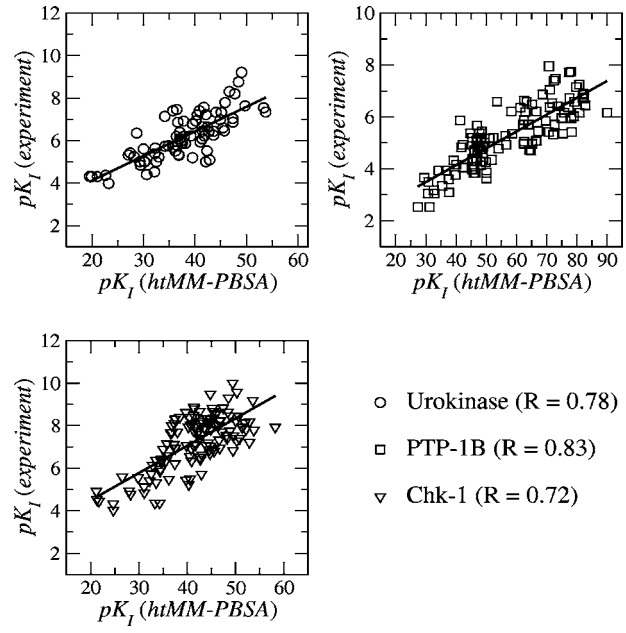

Figure 6. Estimated potencies versus experimentally measured potencies for small-molecule ligands grouped by target: urokinase, circles; PTP-1B, squares; Chk-1, triangles. The lines show the linear fits to the data, with the corresponding correlation coefficient indicated in the legend.

Table 1. Comparison of Pearson Correlation Coefficients, R, for Ligand-Potency Prediction versus Experimental Data for Each Target[a]

correlation	urokinase	PTP-1B	Chk-1
$R_{htMM-PBSA}$	0.78	0.83	0.72
$\|R_{MW}\|$	0.61	0.69	0.59
$R_{X-Score}$	0.62	0.71	0.56
R_{MAC}	0.88	0.90	0.87

[a] Shown is the correlation coefficient from htMM-PBSA calculations, $R_{htMM-PBSA}$, the absolute value of ligand molecular-weight correlation coefficient, $|R_{MW}|$, and the correlation coefficient from the empirical scoring method X-Score, $R_{X-Score}$. Also shown is the mean anticipated correlation coefficient, R_{MAC}, which indicates the value of the correlation coefficient that is anticipated based upon the quality of the underlying data.[53]

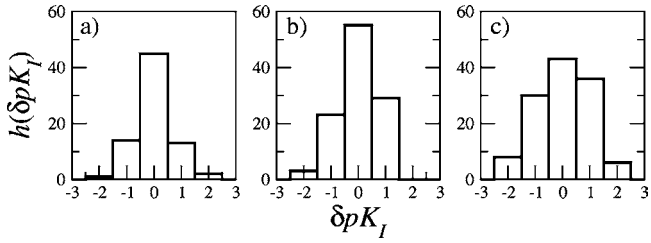

Figure 7. Histograms showing the distribution of estimated errors in the calculated binding potencies for htMM-PBSA: (a) urokinase; (b) PTP-1B; (c) Chk-1. $\delta pK_I = pK_I$ (experiment) $- pK_I$* (htMM-PBSA), where pK_I* is the linearly rescaled value. The rescaling is performed in order to bring the htMM-PBSA pK_I values into a similar range as the experimental pK_I values (note that this does not alter the correlation in the data). The root-mean-square error for each target is $\langle \delta^2 pK_I \rangle^{1/2} = 0.69$ for urokinase, $\langle \delta^2 pK_I \rangle^{1/2} = 0.66$ for PTP-1B, and $\langle \delta^2 pK_I \rangle^{1/2} = 0.89$ for Chk-1.

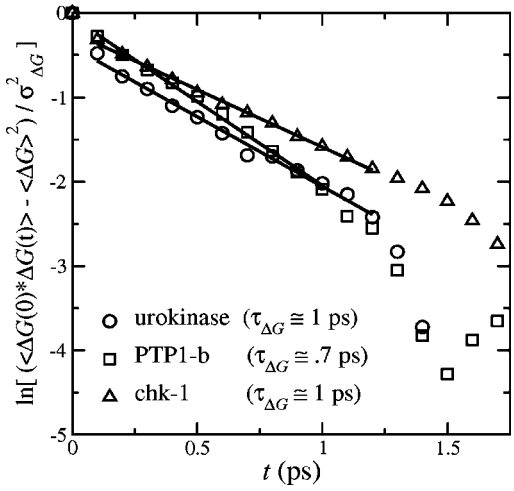

Figure 8. Decay in correlation of htMM-PBSA free-energy estimates $\ln[(\langle \Delta G(0)\, \Delta G(t) \rangle - \langle \Delta G \rangle^2)/\sigma_{\Delta G}{}^2]$ as a function of MD simulation time t, where $\sigma_{\Delta G}{}^2$ is a scaling factor given by $\sigma_{\Delta G}{}^2 = \langle \Delta G^2 \rangle - \langle \Delta G \rangle^2$. The time constants for decay are obtained by linear fit to the data in the plot (indicated by the lines), with the corresponding decay constants shown in the legend.

$R_{\text{htMM-PBSA}}$ values are all close to, but do not exceed, the R_{MAC} values for each set.

To inspect an estimate of the error in the htMM-PBSA calculations as a function of target identity, we need to rescale the calculated values to bring them into a common scale with the experimental values (note the range of the predicted values in Figure 6). This rescaling does not affect any correlation present in the data but merely serves to change the magnitude of the values. We then estimate the error in the predictions by simply taking the relative difference between the calculated and measured values. The results are shown in Figure 7. It can be seen that the distributions across all the targets are roughly centered about the origin, with the Chk-1 numbers showing a larger spread in the error. It is interesting to note that the set with the largest error is the one that has the least net charge. One might have expected to find larger error in the ligand sets with the greater numbers of charged species (see Figure 5), as mistakes in charge assignments are squared in calculating electrostatic energies. However, there are many likely culprits for significant sources of error in these calculations, such as lack of solute entropy, force-field inadequacies, and GB sampling artifacts. We do note that the Chk-1 set has the greatest chemical diversity of any of the sets (as seen in Figure 4c). This is likely a significant contributor to the greater spread in error; however, it is worth also pointing out here that it can be challenging to unambiguously assign sources to the errors in these calculations.

We do find that the error distributions do not significantly change with substantial increases in run times for the MD trajectories. For 100-fold increases in run time we observe qualitatively similar spreads in error estimates. Thus, one potentially major contributor to the error is likely not due to sampling inadequacies in that at least the local energy-basin appears to have been representatively sampled. In fact, for local regions the degree of sampling achieved can be quantitatively inspected by looking at a measure of convergence based on the relaxation of characteristic fluctuations in the system (Figure 8). We find fairly rapid decorrelation times for all targets, indicating that the duration of the MD trajectories (10 ps) spans roughly 10 characteristic relaxation times ($10\tau_{\Delta G}$). Also, if only a single snapshot is used to estimate the potency for each of the compounds, we observe significant degradation in the ability to produce rank-ordered lists by relative binding potency against each target.

Conclusions

Using a procedure for performing high-throughput MM-PBSA (htMM-PBSA) in the context of a distributed-computing paradigm, we show that algorithmically one can produce binding-potency estimates having statistically significant correlation to experimental data, with correlation coefficients in the range of 0.72−0.83 for the targets urokinase, PTP-1B, and Chk-1. In all cases we demonstrate improved signal relative to a typical empirical scoring function X-Score, which performs at the same level as the simple correlation present with molecular weight.

The larger computational cost of the htMM-PBSA calculations is justified by the improved signal, yet it is still low enough to potentially realize throughputs on the order of thousands of structures per day. In fact, the calculations for the 308 structures were performed overnight with a single submission to our enterprise grid, which comprises approximately 400 employee desktop computers. It is worth emphasizing here that the method we employ uses a fully automated procedure for parametrizing the protein−ligand complexes, in addition to an automated deployment and reconciliation routines. This is significant in that we wish to apply this method systematically to sizable compound libraries for which parametrization by hand would be a very tedious, if not untenable, process. In combination with the potential for physics-based generalizability, reflected in the fairly consistent correlations across the three targets, we find this to be a promising method for routine scoring evaluation of candidate ligands for structure-based drug design.

Every calculation involves some degree of error in the predicted estimates. An understanding of this error is vital to the successful deployment of any method. Following a recent analysis of expected correlation,[53] we find that the model performance of htMM-PBSA is consistent with what is expected on the basis of the quality of the underlying data used in the calculation. Additionally, on the basis of the root mean-square error we observe (see Figure 7), there appears to be sufficient signal for discrimination of compounds pulled from lead-optimization-type potency distributions.[53]

In continuance of this work, we plan to perform ht-MMPBSA analysis on a majority of Abbott's internal data, which represent around ∼1800 structures over roughly 3 dozen targets with accompanying binding potency measurements for each of the complexes. This will allow us to delineate a clearer picture of

the domain of applicability for the method, which will be invaluable in constructing confidence measures for guiding future *prospective* analyses.

Note Added after Print Publication. For this manuscript posted on the Web on April 22, 2009, and published in *J. Med. Chem.* **2009**, *52*, 3159−3165, the updated Supporting Information files containing added target identification and MMPBSA scores were posted on the Web on June 5, 2009.

Supporting Information Available: Coordinates and potency data for the compounds used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Thomsen, R.; Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **2006**, *49* (11), 3315–3321.

(2) Yang, C.-Y.; Wang, R.; Wang, S. M-Score: a knowledge-based potential scoring function accounting for protein atom mobility. *J. Med. Chem.* **2006**, *49* (20), 5903–5911.

(3) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196.

(4) Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, B. S.; Johnson, P. eHiTS: an innovative approach to the docking and scoring function problems. *Curr. Protein Pept. Sci.* **2006**, *7* (5), 421–435.

(5) Tripathi, A.; Fornabaio, M.; Kellogg, G. E.; Gupton, J. T.; Gewirtz, D. A.; Yeudall, W. A.; Vega, N. E.; Mooberry, S. L. Docking and hydropathic scoring of polysubstituted pyrrole compounds with anti-tubulin activity. *Bioorg. Med. Chem.* **2007**, doi:10.1016/j.bmc.2007.11.076.

(6) Taylor, P.; Blackburn, E.; Sheng, Y. G.; Harding, S.; Hsin, K.-Y.; Kan, D.; Shave, S.; Walkinshaw, M. D. Ligand discovery and virtual screening using the program LIDAEUS. *Br. J. Pharmacol.* **2007**, doi: 10.1038/sj.bjp.0707532.

(7) Catana, C.; Stouten, P. F. W. Novel, customizable scoring functions, parameterized using N-PLS, for structure-based drug discovery. *J. Chem. Inf. Model.* **2007**, *47* (1), 85–91.

(8) Corbeil, C. R.; Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 1. Development and validation of FITTED 1.0. *J. Chem. Inf. Model.* **2007**, *47* (2), 435–449.

(9) Chen, H.-M.; Liu, B.-F.; Huang, H.-L.; Hwang, S.-F.; Ho, S.-Y. SODOCK: swarm optimization for highly flexible protein-ligand docking. *J. Comput. Chem.* **2007**, *28* (2), 612–623.

(10) Jain, A. N. Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J. Comput.-Aided Mol. Des.* **2007**, *21* (5), 281–306.

(11) Perola, E.; Walters, W.; Charifson, P. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56* (2), 235–249.

(12) 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.

(13) Enyedy, I.; Egan, W. Can we use docking and scoring for hit-to-lead optimization? *J. Comput.-Aided Mol. Des.* **2008**, doi:10.1007/s10822-007-9165-4.

(14) Hawkins, D. M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (1), 1–12.

(15) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44* (7), 1035–1042.

(16) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.

(17) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897.

(18) Brown, S. P.; Muchmore, S. W. High-throughput calculation of protein−ligand binding affinities: modification and adaptation of the MM-PBSA protocol to enterprise grid computing. *J. Chem. Inf. Model.* **2006**, *46* (3), 999–1005.

(19) Andreasen, P. A.; Kjoller, L.; Christensen, L.; Duffy, M. J. The urokinase-type plasminogen activator system in cancer metastasis: a review. *Int. J. Cancer* **1997**, *72* (1), 1–22.

(20) Duffy, M. J.; Maguire, T. M.; McDermott, E. W.; O'Higgins, N. Urokinase plasminogen activator: a prognostic marker in multiple types of cancer. *J. Surg. Oncol.* **1999**, *71* (2), 130–135.

(21) Elchebly, M.; Payette, P.; Michaliszyn, E.; Cromlish, W.; Collins, S.; Loy, A. L.; Normandin, D.; Cheng, A.; Himms-Hagen, J.; Chan, C.-C.; Ramachandran, C.; Gresser, M. J.; Tremblay, M. L.; Kennedy, B. P. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* **1999**, *283* (5407), 1544.

(22) Zinker, B. A.; Rondinone, C. M.; Trevillyan, J. M.; Gum, R. J.; Clampit, J. E.; Waring, J. F.; Xie, N.; Wilcox, D.; Jacobson, P.; Frost, L.; Kroeger, P. E.; Reilly, R. M.; Koterski, S.; Opgenorth, T. J.; Ulrich, R. G.; Crosby, S.; Butler, M.; Murray, S. F.; McKay, R. A.; Bhanot, S.; Monia, B. P.; Jirousek, M. R. PTP1B antisense oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves insulin sensitivity in diabetic mice. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (17), 11357–11362.

(23) Sanchez, Y.; Wong, C.; Thoma, R. S.; Richman, R.; Wu, Z.; Piwnica-Worms, H.; Elledge, S. J. Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. *Science* **1997**, *277* (5331), 1497–1501.

(24) Tenzer, A.; Pruschy, M. Potentiation of DNA-damage-induced cytotoxicity by G2 checkpoint abrogators. *Curr. Med. Chem.: Anti-Cancer Agents* **2003**, *3* (1), 35–46.

(25) Nienaber, V.; Wang, J.; Davidson, D.; Henkin, J. Re-engineering of human urokinase provides a system for structure-based drug design at high resolution and reveals a novel structural subsite. *J. Biol. Chem.* **2000**, *275* (10), 7239–7248.

(26) Liu, G.; Xin, Z.; Pei, Z.; Hajduk, P. J.; Abad-Zapatero, C.; Hutchins, C. W.; Zhao, H.; Lubben, T. H.; Ballaron, S. J.; Haasch, D. L.; Kaszubska, W.; Rondinone, C. M.; Trevillyan, J. M.; Jirousek, M. R. Fragment screening and assembly: a highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor. *J. Med. Chem.* **2003**, *46* (20), 4232–4235.

(27) Li, G.; Hasvold, L. A.; Tao, Z. F.; Wang, G. T.; Gwaltney, S. L., 2nd; Patel, J.; Kovar, P.; Credo, R. B.; Chen, Z.; Zhang, H.; Park, C.; Sham, H. L.; Sowin, T.; Rosenberg, S. H.; Lin, N. H. Synthesis and biological evaluation of 1-(2,4,5-trisubstituted phenyl)-3-(5-cyanopyrazin-2-yl)ureas as potent Chk1 kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16* (8), 2293–2298.

(28) Wendt, M. D.; Rockway, T. W.; Geyer, A.; McClellan, W.; Weitzberg, M.; Zhao, X.; Mantei, R.; Nienaber, V. L.; Stewart, K.; Klinghofer, V.; Giranda, V. L. Identification of novel binding interactions in the development of potent, selective 2-naphthamidine inhibitors of urokinase. Synthesis, structural analysis, and SAR of *N*-phenyl amide 6-substitution. *J. Med. Chem.* **2004**, *47* (2), 303–324.

(29) Szczepankiewicz, B. G.; Liu, G.; Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z.; Xin, Z.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J.; Liang, H.; Huang, F.; Hutchins, C. W.; Fesik, S. W.; Jirousek, M. R. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J. Am. Chem. Soc.* **2003**, *125* (14), 4087–4096.

(30) Chen, Z.; Xiao, Z.; Gu, W. Z.; Xue, J.; Bui, M. H.; Kovar, P.; Li, G.; Wang, G.; Tao, Z. F.; Tong, Y.; Lin, N. H.; Sham, H. L.; Wang, J. Y.; Sowin, T. J.; Rosenberg, S. H.; Zhang, H. Selective Chk1 inhibitors differentially sensitize p53-deficient cancer cells to cancer therapeutics. *Int. J. Cancer* **2006**, *119* (12), 2784–2794.

(31) Hassan, M.; Brown, R. D.; Varma-O'brien, S.; Rogers, D. Cheminformatics analysis and learning in a data pipelining environment. *Mol. Diversity* **2006**, *10* (3), 283–299.

(32) Muchmore, S. W.; Debe, D. A.; Metz, J. T.; Brown, S. P.; Martin, Y. C.; Hajduk, P. J. Application of belief theory to similarity data fusion for use in analog searching and lead hopping. *J. Chem. Inf. Model.* **2008**, *48* (5), 941–948.

(33) 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* (2), 132–146.

(34) OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com/ (accessed October 31, 2008).

(35) Brown, S. P.; Muchmore, S. W. Rapid estimation of relative protein−ligand binding affinities using a high-throughput version of MM-PBSA. *J. Chem. Inf. Model.* **2007**, *47* (4), 1493–1503.

(36) Gilson, M. K.; Honig, B. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* **1988**, *4* (1), 7–18.

(37) Bondi, A. van der Waals volumes and radii. *J. Phys. Chem.* **1964**, *68* (3), 441–451.

(38) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.

(39) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71* (8), 2925–2927.

(40) Sitkoff, D.; Sharp, K. A.; Honig, B. Correlating solvation free energies and surface tensions of hydrocarbon solutes. *Biophys. Chem.* **1994**, *51* (2−3), 397–409.
(41) Case, D. A. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* **1994**, *4*, 285–290.
(42) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structureal and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.
(43) Wang, J.; 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* (12), 1049–1074.
(44) 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* (9), 1157–1174.
(45) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247–260.
(46) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
(47) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct., Funct., Bioinf.* **2004**, *55* (2), 383–394.
(48) Thain, D.; Tannenbaum, T.; Livny, M. *Condor and the Grid*; John Wiley & Sons Inc.: Hoboken, NJ, 2003; p 1080.
(49) Wang, R.; Lu, Y.; Wang, S. Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* **2003**, *46* (12), 2287–2303.
(50) Tripos Inc., St. Louis, MO. http://www.tripos.com/ (accessed October 31, 2008).
(51) Falk, R.; Well, A. D. Many faces of the correlation coefficient. *J. Stat. Educ.* **1997**, *5* (3), 1–12.
(52) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; L. Erlbaum Associates: Hillsdale, NJ, 1988.
(53) Brown, S. P.; Muchmore, S. W.; Hajduk, P. J. Healthy skepticism: assessing realistic model performance. *Drug Discovery Today*, in press.

JM801444X