JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

A Molecular Mechanics Approach to Modeling Protein−Ligand Interactions: Relative Binding Affinities in Congeneric Series

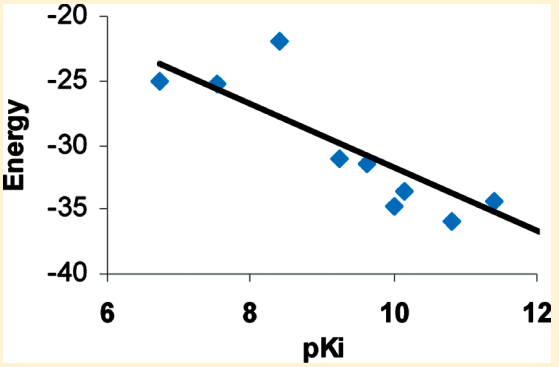
Chaya Rapp,[†] Chakrapani Kalyanaraman,[‡] Aviva Schiffmiller,[†] Esther Leah Schoenbrun,[†] and Matthew P. Jacobson*[,‡]

[†]Department of Chemistry, Stern College for Women, Yeshiva University, New York, New York, United States

[‡]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California, United States

Supporting Information

ABSTRACT: We introduce the "Prime-ligand" method for ranking ligands in congeneric series. The method employs a single scoring function, the OPLS-AA/GBSA molecular mechanics/implicit solvent model, for all stages of sampling and scoring. We evaluate the method using 12 test sets of congeneric series for which experimental binding data is available in the literature, as well as the structure of one member of the series bound to the protein. Ligands are "docked" by superimposing a common stem fragment among the compounds in the series using a crystal complex from the Protein Data Bank and sampling the conformational space of the variable region. Our results show good correlation between our predicted rankings and the experimental data for cases in which binding affinities differ by at least 1 order of magnitude. For 11 out of 12 cases, >90% of such ligand pairs could be correctly ranked, while for the remaining case, Factor Xa, 76% of such pairs were correctly ranked. A small number of compounds could not be docked using the current protocol because of the large size of functional groups that could not be accommodated by a rigid receptor. CPU requirements for the method, involving CPU minutes per ligand, are modest compared with more rigorous methods that use similar force fields, such as free energy perturbation. We also benchmark the scoring function using series of ligands bound to the same protein within the CSAR data set. We demonstrate that energy minimization of ligands in the crystal structures is critical to obtain any correlation with experimentally determined binding affinities.



INTRODUCTION

Recently, there has been a resurgence of interest in using molecular mechanics force fields, of the type used for molecular dynamics, for predicting protein−ligand interactions.[1−3] In the extreme case, these energy functions are used in conjunction with rigorous statistical mechanical methods for computing absolute or relative binding free energies.[4−8] These methods, however, remain very computationally expensive, and although some encouraging results have been reported, primarily on "model systems", much work remains to be done to establish their utility for drug discovery.[3,9,10] Another line of work by ourselves and others has examined whether molecular mechanics energy functions can be used in a much simpler and more approximate manner as a "scoring function" for protein−ligand docking.[10−17] Most of this work employs implicit solvent models, as opposed to the explicit treatment of water typically used in free energy simulations, for computational expediency. In the simplest case, the energy of the free protein and free ligand are simply subtracted from that of the protein−ligand complex, which means that entropic losses of the ligand and protein are neglected.[15]

Despite these limitations and inherent errors and approximations in the underlying force fields a number of promising results have been reported, suggesting that force fields can be successfully used as docking scoring functions. A number of studies have suggested that molecular mechanics force fields in conjunction with implicit solvent models can be used to identify protein−ligand binding poses.[18,19] Molecular mechanics scoring functions have also been used to rank ligands according to predicted relative binding affinity with respect to a particular protein binding site,[9,15,20−22] including some tests using ligand libraries with tens of thousands of members.[15,20−22] One of the most promising applications, which is also explored here, is using molecular mechanics scoring functions to rank compounds within congeneric series (or other chemically similar series of ligands) according to predicted binding affinity.

There are a number of potential advantages to using molecular mechanics scoring functions in the context of "docking". First, they capture many of the key physical forces driving binding, albeit approximately, including hydrogen bonding, desolvation of the protein and ligand, and internal strain.[13] In the longer term, these methods will be able to exploit advances in force fields and solvent models, such as the new generation of polarizable force fields.[23−25] One major limitation has simply

Special Issue: CSAR 2010 Scoring Exercise

Received: January 24, 2011
Published: July 22, 2011

Table 1. Test Sets and Summary of Results

target proteins					Prime-ligand				
name	PDB id	reference, table(s)	no.	experimental data (range)	r^2	slope	rs[d]	τ[e]	pairs[f]
P38 kinase	2bak	10, Table 1	13	pIC_{50} (5.7−8.0)	0.70	−14.8	1.00	0.72	28/28
Aurora A kinase	2c6e	10, Table 2	12	pIC_{50} (5.4−9.8)	0.63	−5.2	0.83	0.67	38/40
Cdk-2	1oiu	10, Table 3	11	pIC_{50} (4.9−8.3)	0.82	−7.3	0.75	0.64	26/27
Aurora A kinase	3d14	36, Tables 1 and 2	18	pIC_{50} (5.1−8.4)	0.49	−7.3	0.59	0.40	61/66
DAAO	3g3e	37, Table 1[a]	17	pIC_{50} (4.6−8.5)	0.50	−5.7	0.78	0.56	67/73
N-methyl-transferase	1hnn	38, Table 1[b]	9	pK_i (5.0−8.5)	0.45	−9.7	0.48	0.39	16/20
SYK	3emg	39, Table 1	8	pK_i (5.7−8.1)	0.51	−2.2	0.43	0.36	9/10
Pim-1 kinase	3bgq	40, Tables 1 and 2	16	pKi (5.7−8.0)	0.51	−3.4	0.70	0.47	30/31
Cdk-2	2clx	41, Table 1	12	pK_i (4.0−5.6)	0.25	−13.1	0.38	0.27	6/8
HCV NS5B polymerase	2hwi	42, Table 1[c]	13	pIC_{50} (4.2−5.5)	0.31	−5.7	0.58	0.38	5/5
HIV-1 protease	2i0d	51, Table 1 and 2	10	pK_i (6.7−12.1)	0.72	−2.5	0.82	0.64	34/42
Factor Xa	2boh	51, Tables 1−3, and 52, Tables 1 and 2	41	pK_i (6.0−10.1)	0.37	−0.1	0.58	0.39	907/1184

[a] Quinoline inhibitors. [b] Ligands 7−13 and 16−17. [c] Ligands 1 and 15−27. [d] Spearman's rank correlation coefficient. [e] Kendall τ rank correlation coefficient. [f] Fraction correctly ranked among pairs with binding energies differing by at least 1 order of magnitude.

been computational expense. Most docking scoring functions have much simpler functional forms;[26,27] the computational cost of implicit solvent models is particularly challenging. As a result, most of the results we cite above that have successfully used force fields with implicit solvent models did not use such energy functions *exclusively*. In most cases, a docking program employing a more common docking scoring function was used to position ligands in a protein binding site, and the molecular mechanics/implicit solvent scoring function was used only to "rescore" these poses.[15] Similarly, in a widely adopted approach to treating protein flexibility in ligand docking ("induced fit"), a docking program using an empirical scoring function and a protein modeling program using a molecular mechanics-generalized Born surface area (MM-GBSA) scoring function are used iteratively to treat ligand and protein degrees of freedom, respectively.[28] There is nothing inherently wrong with such an approach, but we posit that there may be advantages to generating, selecting, and ranking protein−ligand poses with the same energy function, which is the main novel aspect of this work.

We apply this approach here to ranking ligands in congeneric series. One simplification in this application is that the usual conformational search over three rotational and three translational degrees of freedom for the ligand in the protein binding site can be significantly constrained. We start with the conserved chemical substructure positioned in the binding site as determined by the X-ray crystallographic structure of one of the compounds in the series and sample the "ligand side chains" using torsion-angle based methods and energy minimization. This approach assumes that the position of the conserved substructure will be similar among the members of a series, i.e., that it will not "flip" to a completely different configuration. Although not guaranteed to be the case, this is a reasonable assumption in most cases and helps to maximize presumed "cancellation of error", by minimizing noise introduced by differences in the position of the core among members of the series.

We refer to our method as "Prime-ligand" because it is implemented in the code base of the Prime modeling package,[29] which was initially developed for protein modeling.[30,31] It takes advantage of some of the torsion angle-based sampling algorithms originally developed for protein side chain[30] and loop sampling,[32,33] applied here to the ligands (and, as desired, side chains in the protein binding site). Ultimately, we hope that this method will be useful for helping to predict chemical modifications that can improve binding affinity, possibly in conjunction with other methods,[34,35] but here our results are restricted to benchmarking against 12 test cases of congeneric series, mainly kinases for which appropriate binding data could be culled from the literature.

We also use the CSAR data set as an additional test, primarily of the scoring function, because crystal structures are available for all members. It is worth reiterating that the MM-GBSA scoring function as we use in the present calculation can compute only relative binding affinities of ligands. Therefore, we consider only a subset of the CSAR data, specifically those proteins with several bound ligands. The series of ligands are not necessarily congeneric, making the scoring more challenging. Scoring the protein−ligand complexes as originally provided yields very poor results. Applying receptor preparation protocols and energy minimization of ligands in the binding site is necessary to obtain any correlation with experimental binding affinities.

MATERIALS AND METHODS

Congeneric Series Test Sets. Our study includes 12 test sets listed in Table 1.[10,36−42] Test sets were identified by a search of recent literature for congeneric series in which a series of at least eight compounds binds to a target protein, and a crystal structure of one compound in the series in complex with the target protein is available in the Protein Data Bank.[43] Table 1 contains the reference for each series, the number of compounds in the series used in this study, the nature of the experimental data (pIC_{50} or pK_i values), and the range of the experimental data. For the DAAO[37] and HCV receptors,[42] some of the compounds in the referenced table do not share a common fragment with the compound in the crystal complex and are therefore excluded from this study. Compounds for which the experimental data is imprecise (e.g., >50 μM) are similarly excluded from our results. Compounds listed in two

tables are treated as a single series when they share a common fragment with the compound in the crystal complex.

Preparation of Inhibitor Compounds. All compounds were constructed in Maestro (windows version 9.0[44]) and energy minimized using Ligprep (Schrodinger software, version 21207). Protonation states of carboxylic acid groups were assumed to be charged, and piperidine and piperazine rings were assumed to be neutral. A utility script named hetgrp_ffgen available within the Schrodinger molecular modeling package was used to generate OPLS2005 force field parameters for each compound.

Docking and Scoring. Preparation of the target protein, docking, and calculation of all energies, were done using Protein Local Optimization Program (PLOP) (free academic version 8.1; commercial version marketed as Prime by Schrodinger LLC),[29,30,32] using the OPLS all-atom force field (OPLS-AA),[45] and the surface generalized Born implicit solvent model as described in earlier works.[46,47] The user defines the stem fragment for the series by identifying the fragment present in the cocrystallized compound that is a common fragment among all series compounds. The crystal structure of the target was prepared for docking by editing the crystal compound down to the stem fragment, removal of water molecules, and initial energy minimization in which all atoms not found in the ATOM records are optimized (primarily hydrogen atoms and missing side chains). Rotatable bonds for each compound were defined manually in this work.

Ligand conformations are generated by first superimposing the stem atoms of each compound on the stem atoms of the cocrystallized compound. Conformations of the ligand "side chains", i.e., the nonconserved portions, are sampled by varying the nonrigid dihedral angles. Conformations that result in conformational clashes are automatically eliminated using an "overlap factor" criterion that computes the distance between two atoms divided by the sum of their radii. The minimum acceptable overlap factor is user adjustable and was set to 0.65 in this work. The conformational search is performed in an exhaustive fashion subject to a sampling resolution. The sampling begins very coarse (only 0 and 180 degrees sampled for each dihedral angle) and then is gradually increased to a maximal resolution of 10 degrees (36 conformations per rotatable dihedral angle). Because conformations resulting in steric clashes are efficiently eliminated, all rotatable bonds are treated identically and evenly sampled. Sampling is terminated before reaching the maximal resolution if a user defined minimum number of conformations is generated (in this work, 1000 conformations).

Subsequently, the conformations are clustered, and one member of each cluster is subjected to energy minimization,[48] with a termination criterion of 0.001 kcal/mol/Å root-mean-squared force, which ensures that the minimization is well converged. Only the ligand is energy minimized in this work (i.e., not the protein). The clustering algorithm uses the K-means method, using Cartesian coordinates of the ligand atoms for clustering. In this work, 50 clusters were generated, and the representative member was chosen as the one closest to the geometric center of the cluster. The energy minimization is performed with an optimized version of the truncated Newton algorithm.[48] The lowest energy conformer is selected as the final pose, which is used for the energy-based ranking.

All atoms of the target protein are held rigid during docking. For each compound, "docking" and calculation of all energies requires between 100 and 500 CPU seconds. Binding energies

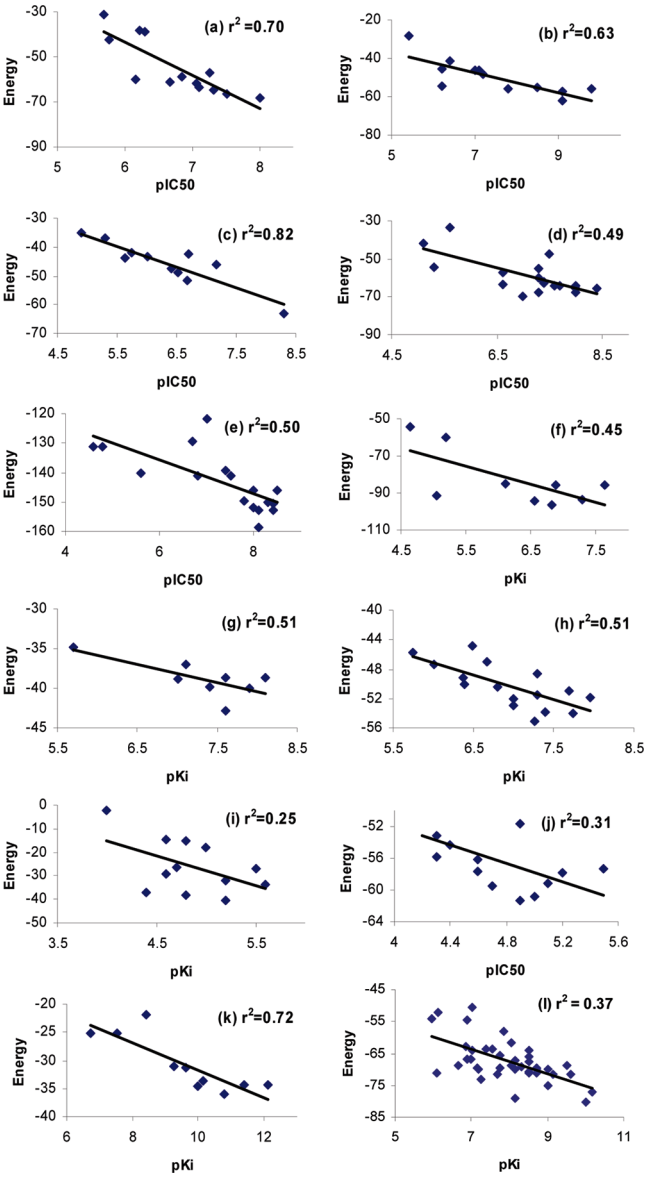


Figure 1. Plots of computed energy scores (kcal/mol) (*y*-axis) vs pIC_{50} or pK_i values (*x*-axis) for (a) 2bak, (b) 2c6e, (c) 1oiu, (d) 3d14, (e) 3g3e, (f) 1hnn, (g) 3emg, (h) 3bgq, (i) 2clx, (j) 2hwi, (k) 2i0d, and (l) 2boh.

are calculated as the difference between the energy of the bound complex and the energy of the unbound target and inhibitor compound, $E_{binding} = E_{complex} - E_{target} - E_{inhibitor}$. Specifically, after calculating the energy of the ligand−protein complex, the ligand and the protein are separated, and their energies are computed using OPLS-AA force-field with generalized Born implicit solvent model. After separating the protein and ligand, we have not subjected them to any additional energy minimizations. Note that the implicit solvent model estimates solvation free energies, and thus, this energy implicitly includes entropies associated with solvent. Binding energies are plotted against pIC_{50} or pK_i values for the series, and we report the degree of correlation between the two values using the correlation coefficient, *r*. We also report the Spearman's rank correlation coefficient, rs, which compares the position of each inhibitor compound when ranked by binding energy to its

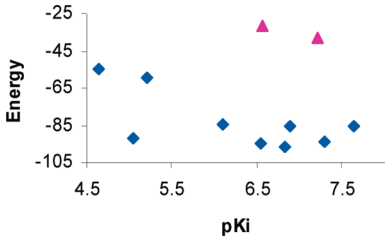


Figure 2. Plot of computed energy scores (kcal/mol) vs pIC_{50} values for 1hnn including compounds 14 and 15 (triangular data labels).

position when ranked by its pIC_{50} or pK_i value.[49] The Spearman's rank correlation coefficient is defined as

$$rs = 1 - \frac{6\sum di^2}{n(n^2-1)}$$

where d_i is difference in rank for the i^{th} compound under the two different criteria, i.e., binding energy and experimental binding constant, and n is the number of compounds in the series. In addition to the Spearman rank correlation coefficient (rs), we have also calculated the Kendall rank correlation coefficient (τ) using the formula,

$$\tau = \frac{2(C-D)}{n(n-1)}$$

where C and D are respectively the number of concordant and discordant pairs between experimental and calculated ranks, and n is the number of ligands in the data set. Large positive τ values indicate strong positive correlations between calculated and experimental ranking of ligands, and large negative τ values indicate strongly negative correlations between the two rankings. The value of τ ranges between -1 and 1.

RESULTS AND DISCUSSION

Plots of binding energy versus pIC_{50} or pK_i values for each series are shown in Figure 1. Correlation coefficients and slopes for the plots, as well as Spearman's rank correlation coefficients, are reported for each series in Table 1. In most cases, the results show reasonable correlation between relative binding energies and the experimental data, similar to that obtained by other methods.[10−12,14] As an additional metric of the ability to distinguish more potent versus less potent inhibitors, we identified all pairs of compounds with binding constants differing by at least 1 order of magnitude; the fraction of such pairs correctly ranked is reported in the last column in Table 1. Our data show that for 11 cases, a large majority, >90%, of such pairs were correctly ranked. For the remaining case, Factor Xa with 41 data points, 76% of such pairs were correctly ranked.

For the kinases p38 (PDB id 2bak), Aurora kinase (PDB id 2c6e), and CDK2 (PDB id 1oiu), we compare our results to those of Lyne et al.,[10] in which the compounds in each series were docked by Glide, and the three best poses were scored using the same MM-GBSA force field employed here. Inhibitors of four kinases were used as test sets in that study; our test set does not include Jnk-3 kinase because no appropriate crystal structure is available in the Protein Data Bank. For p38 kinase, we achieved the same result as Lyne (r^2 = 0.70). For Aurora A kinase and CDK2, our results of r^2 = 0.63 and r^2 = 0.82 are better than those reported by Lyne (r^2 = 0.56 and r^2 = 0.50 for Aurora A kinase and CDK2, respectively).[10] Additionally, for Aurora A kinase, we successfully docked all 12 compounds in the series, while Lyne et al.[10] were able to dock only eight out of 12.

For the N-methyltransferase,[38] PDB id 1hnn, the referenced table contains 11 sulfonamide compounds (numbered 7−17); our reported result (r^2 = 0.45) is for nine compounds (7−13 and 16−17). Compounds 14 and 15 generate binding energies that are notably higher than the other compounds in the series (Figure 2) and when included produce no overall correlation in the plot of binding energy vs pIC_{50} values (r^2 = 0.03). We understand this in light of the fact that an alternate crystal structure for the receptor (PDB id 2obf), which contains the receptor in complex with compound 15, shows "significant differences"[38] compared to 1hnn, which contains the same receptor in the complex with compound seven (and which was used to dock the ligands in the series). These differences include the shifting of a lysine side chain and the presence or absence of hydrogen bond with the backbone. Our use of a rigid receptor structure precludes the ability to model these induced fit effects. Although it is possible to treat protein side chains as flexible in our approach, our initial attempts to include receptor flexibility generally degraded the results of rank ordering the compounds by predicted binding affinity, presumably due in part to reduced "cancellation of error".

For HIV-1 protease, we used a set of 10 congeneric series compounds with binding affinities ranging from 0.0008 to 188.8 nM.[50] Recently, the same data set was also used in a study by Chen et al.[8] They used a more rigorous (and computationally more intensive) mining minima approach[8] to calculate relative binding energies of the ligands. We selected one of the ligands (21E K_i = 0.0008 nM) bound to the protein structure (PDB id 2i0d) as the starting structure. The computed relative binding energies correlated with r^2 = 0.72 with the experimental binding affinities as shown in Figure 1. The mining minima approach[8] also yielded a similar correlation coefficient.

The seriene protease, Factor Xa, is an important antithrombosis target. Nazare et al. have identified several low nanomolar compounds against this target.[51,52] In particular, using indole as a main fragment, they added various substituent groups and optimized interactions with the S1 and S4 pockets. They reported binding affinities of several ligands from their structure−activity relationship study while optimizing the S4 pocket.[51,52] Among these ligands, 41 of them shared a common structural scaffold, and we considered these in the present study (see Supporting Information for the compound identity and binding affinity). A structure of one of the ligands cocrystallized with the protein was available (PDB id 2boh), and we used it as the starting structure. The correlation coefficient (r^2) between the computed and experimental binding affinities is 0.37. There were a number of compounds in the data set with a very narrow range of binding affinities. For instance, in the 1−9 nM range, there were 17 compounds, in the 10−90 nM range, there were 11 compounds, and 100−900 nM range, there were 7 compounds. Although our procedure correctly distinguishes strong and weak binders, distinguishing compounds that narrowly differ in their binding affinity is challenging to the present scoring function. The Sperman and Kendal rank correlation coefficients are 0.58 and 0.39, respectively. There were 1184 pairs of ligands with binding affinities differing by an order of magnitude, and 907 of these pairs had been correctly ranked by our protocol.

Finally, the ability to rank inhibitor compounds in a series is limited to series in which there is a significant difference in binding affinities between the compounds. For the congeneric series with CDK2 (PDB id 2clx)[41] and HCV NS5B polymerase (PDB id 2hwi),[42] binding constants vary over less than 2 orders of magnitude. For these cases, we were able to achieve correlations

Table 2. Summary of Results Using Alternate Protocols

	Glide XP				Glide + MM-GBSA			Prime-ligand + Glide XP score		
PDB id	no. ligands	r^2	rs	τ	r^2	rs	τ	r^2	rs	τ
2bak	13	0.77	0.90	0.74	0.46	0.75	0.59	0.55	0.89	0.72
2c6e	10	0.16	0.36	0.24	0.17	0.32	0.24	0.37	0.69	0.52
1oiu	6	0.03	0.26	−0.20	0.04	0.26	0.20	0.49	0.66	0.49
3d14	18	0.08	0.40	0.22	0.56	0.42	0.24	—[a]	—	—
3g3e	15	0.09	0.23	0.05	0.22	0.65	0.47	0.00	—	—
1hnn	6	0.38	0.54	0.33	0.00	—	−0.07	0.40	0.45	0.33
3emg	8	0.34	0.19	0.14	0.44	0.24	0.14	0.01	0.17	0.14
3bgq	16	0.29	0.55	0.35	0.23	0.53	0.37	—	—	—
2clx	12	0.48	0.64	0.52	—	—	—	—	—	—
2hwi	13	0.01	0.15	0.08	—	—	—	0.33	0.49	0.41

[a] Slope of correlation has incorrect sign.

Table 3. Community Structure−Activity Resource (CSAR) Data Set Results

target proteins			MM-GBSA *as-is* scoring			MM-GBSA *ligand minimized* scoring		
name	PDB id	pKi range	r^2	rs	τ	r^2	rs	τ
HIV-1 protease wild type	2i0d, 2i0a, 2qi5, 2qi6, 2qi4, 2qi3, 2q54, 2qhy, 2qi1, 2qhz, 2psv	12.10−7.24	0.46	0.60	0.42	0.58	0.70	0.53
HIV-1 protease L63P mutant	1ec2, 1ebz, 1ec1, 1d4i, 1ec0, 1d4j, 2cen, 2cem, 1xl5	10−7.35	0.09	−0.33	−0.17	0.48	0.63	0.50
tyrosine phosphatase	2zn7, 2zmm, 2qbq, 2qbs, 2b07, 2qbr, 2hb1, 2azr	7.89−3.64	0.31	0.57	0.43	0.53	0.76	0.57
Factor Xa	2j4i, 2boh, 2uwl, 2cji, 2j34, 2j2u, 2uwp	9−6.81	0.08	−0.32	−0.24	0.62	0.93	0.81
t-RNA guanine transglycosylase	1s39, 1q4w, 1r5y, 1s38, 2bbf, 1enu	7.7−5.08	0.35	0.09	0.07	0.49	0.09	0.07

(r^2) of only 0.25 and 0.31, respectively, indicating that this dynamic range is the current limit of our protocol.

To compare our results with a widely used ligand docking method, we used Glide (xglide version 1.33.2.6) in extra precision (XP) mode[53] on our test set; results are reported in Table 2. Three different protocols involving Glide were used:

1. Docking in Glide using a core constraint option similar to the Prime-ligand protocol described here. For each congeneric series, the same common fragment used in Prime-ligand docking was used as the "core" in Glide, with the exception of CDK2 kinase[10] for which the ligands could only be docked by Glide when defining a smaller core fragment. All ligands successfully docked by Prime-ligand were subjected to GLIDE docking. The protein preparation module available in Maestro (Schrodinger LLC) was used to generate appropriate bond orders, protonation states of titratable residues, and to add hydrogens to the receptor[54] prior to docking.
2. Rescoring of the Glide docked complexes (generated in protocol 1) using energy minimization and recalculation of energies in Prime using the OPLS-AA force field and the GBSA continuum solvent model. Up to three poses per ligand were saved in Glide docking, and all poses were rescored by Prime. The lowest energy pose is included in our reported data.
3. Rescoring of the Prime-ligand docked complexes, generated using the protocol described in the Methods section, using refinement and calculation of energies in Glide ("opt and score" mode).

Table 2 reports correlation coefficients and Spearman's and Kendall τ rank correlation coefficients obtained from these three protocols. Overall, an improvement is seen is using MM-GBSA for all stages of the protocol, as in Table 1, and the improvement can be quite significant in some cases.

CPU time requirements using Prime-ligand docking are moderate; for example, calculations for the 12 compounds in the Aurora (PDB id 2c6e) congeneric series take less than 1 h, or an average of 5.3 min per ligand.

As has been observed previously, even when the MM-GBSA energy correlates reasonably well with the logarithm of the experimental binding affinities, the slope of the correlation is always large, i.e., the computed energy scores grossly overestimate the differences in binding affinity. This phenomenon is likely due in part to the neglect of ligand and protein entropy losses, which oppose binding. These neglected entropy losses must of course correlate significantly with the other components of the computed energies; otherwise, no correlation with experimental binding affinities would be possible. Entropy−enthalpy compensation of this type was reported by Gilson et al. on the basis of mining minimal calculations.[8] The neglect of electronic polarizability by the fixed charge force field may also be related to the observed slopes.

COMMUNITY STRUCTURE−ACTIVITY RESOURCE (CSAR) RESULTS

During the course of developing this method, the CSAR exercise was held.[55] Molecular mechanics-based scoring is highly

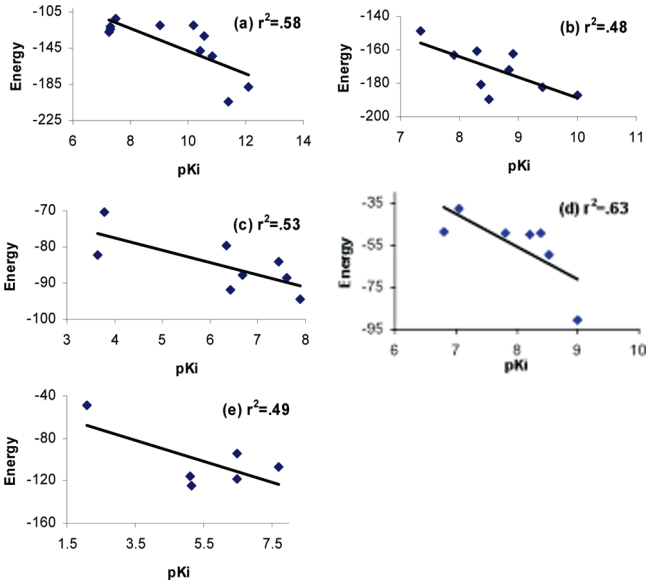


Figure 3. Plots of computed energy scores (kcal/mol) (*y*-axis) vs pK_i (*x*-axis) for (a) HIV-1 protease wild-type, (b) HIV-1 protease L63P mutant, (c) Tyrosine phosphatase, (d) Factor Xa, and (e) t-RNA guanine transglycosylase.

unlikely to correlate well with absolute binding affinities of diverse compounds binding to diverse receptors, especially because of the neglect of entropy losses. However, the CSAR data set contained several different crystal structures of certain proteins, with different ligands bound. These were not, in general, congeneric series in the sense of sharing a precisely conserved core scaffold but were more chemically homogeneous than the ligands in the data set as a whole. We tested the molecular mechanics-based scoring on this subset of the CSAR data. This test was thus different from the preceding results in two senses: (1) no sampling of the ligands was needed because crystal structures were available for all members of the series, and (2) the ligands were, in general, somewhat more chemically diverse within the series than in the congeneric series used in the preceding results.

The results are summarized in Table 3 and Figure 3. The most obvious conclusion is that energy minimization of the bound ligand is absolutely required to obtain any correlation at all with the experimental relative binding affinities. The results obtained with full energy minimization of the ligand (using the multiscale truncated Newton method) are qualitatively similar to those obtained for the congeneric series but slightly worse overall presumably because of the larger chemical diversity within each series. Again, the correlation between experimental and computed rankings are favorable when the data set contains ligands that differ in binding affinity by at least an order of magnitude.

CONCLUSION

These results represent proof-of-concept for a molecular mechanics-based docking and scoring method and a benchmark for further work aimed at improving the ability to rank relative binding affinities within congeneric series. Given the simplicity of the sampling and scoring scheme, and hence low computational expense (minutes per ligand) at least relative to rigorous free energy methods, the results are encouraging overall. However, the ability to rank-order compounds remains restricted to those differing by at least an order of magnitude in binding affinity. We expect that further improvements may require better treatment of water in the binding site cavities than can be achieved with an implicit solvent model, such as has been implemented in the Watermap approach.[34,35] Other obvious limitations of the scoring used here include the neglect of conformational entropy losses upon binding and neglect of electronic polarizability, and we suspect that these limitations likely are primarily responsible for the large slopes observed in the correlations between the computed and experimental binding free energies, as also discussed by Gilson and co-workers.[56]

Finally, we note that we have implemented the ability to simultaneously sample protein side chains concomitantly with the ligand side chains, and we will report results elsewhere. However, we note that allowing receptor flexibility in the test cases reported here reduced the correlations between computed and measured binding affinities, in some cases dramatically. Thus, at least in our hands, the ability to rank-order compounds within congeneric series requires use of a rigid receptor, presumably due in part to greater cancellation of error. The trade-off however is that compounds too large for the rigid binding site cannot be appropriately "docked", as happened here in a few cases.

ASSOCIATED CONTENT

Supporting Information. A table of Factor Xa ligands included in the present study along with their experimental binding affinities and references. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: Matt.Jacobson@ucsf.edu.

ACKNOWLEDGMENT

M.P.J. is a consultant to Schrödinger LLC. C.K. is grateful to Dr. Woody Sherman (Schrödinger LLC) for introducing the Factor Xa ligand data set and fruitful discussions. This work was supported by National Institutes of Health Grant P01-GM071790.

REFERENCES

(1) Jorgensen, W. L. Efficient drug lead discovery and optimization. *Acc. Chem. Res.* **2009**, *42*, 724–33.
(2) Steinbrecher, T.; Labahn, A. Towards accurate free energy calculations in ligand protein-binding studies. *Curr. Med. Chem.* **2010**, *17*, 767–85.
(3) Pohorille, A.; Jarzynski, C.; Chipot, C. Good practices in free-energy calculations. *J. Phys. Chem. B* **2010**, *114*, 10235–53.
(4) Singh, N.; Warshel, A. Absolute binding free energy calculations: On the accuracy of computational scoring of protein−ligand interactions. *Proteins* **2010**, *78*, 1705–23.
(5) Ge, X.; Roux, B. Absolute binding free energy calculations of sparsomycin analogs to the bacterial ribosome. *J. Phys. Chem. B* **2010**, *114*, 9525–39.
(6) Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A.; Shoichet, B. K. Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. *J. Mol. Biol.* **2009**, *394*, 747–63.
(7) Rodinger, T.; Howell, P. L.; Pomes, R. Calculation of absolute protein−ligand binding free energy using distributed replica sampling. *J. Chem. Phys.* **2008**, *129*, 155102.

(8) Chen, W.; Glison, M. K.; Webb, S. P.; Potter, M. J. Modeling protein−ligand binding my mining minima. *J. Chem. Theory Comput.* **2010**, *6*, 3540–3557.
(9) Michel, J.; Essex, J. W. Prediction of protein−ligand binding affinity by free energy simulations: Assumptions, pitfalls and expectations. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 639–58.
(10) Lyne, P. D.; Lamb, M. L.; Saeh, J. C. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **2006**, *49*, 4805–8.
(11) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J. Comput. Chem.* **2010**, *31*, 797–810.
(12) Guimaraes, C. R.; Cardozo, M. MM-GB/SA rescoring of docking poses in structure-based lead optimization. *J. Chem. Inf. Model.* **2008**, *48*, 958–70.
(13) Huang, N.; Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. Molecular mechanics methods for predicting protein−ligand binding. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5166–77.
(14) Ferrari, A. M.; Degliesposti, G.; Sgobba, M.; Rastelli, G. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorg. Med. Chem.* **2007**, *15*, 7865–77.
(15) Huang, N.; Kalyanaraman, C.; Irwin, J. J.; Jacobson, M. P. Physics-based scoring of protein−ligand complexes: Enrichment of known inhibitors in large-scale virtual screening. *J. Chem. Inf. Model.* **2006**, *46*, 243–53.
(16) Foloppe, N.; Hubbard, R. Towards predictive ligand design with free-energy based computational methods? *Curr. Med. Chem.* **2006**, *13*, 3583–608.
(17) Guimaraes, C. R.; Mathiowetz, A. M. Addressing limitations with the MM-GB/SA scoring procedure using the WaterMap method and free energy perturbation calculations. *J. Chem. Inf. Model.* **2010**, *50*, 547–59.
(18) Rapp, C. S.; Schonbrun, C.; Jacobson, M. P.; Kalyanaraman, C.; Huang, N. Automated site preparation in physics-based rescoring of receptor ligand complexes. *Proteins* **2009**, *77*, 52–61.
(19) Nervall, M.; Hanspers, P.; Carlsson, J.; Boukharta, L.; Aqvist, J. Predicting binding modes from free energy calculations. *J. Med. Chem.* **2008**, *51*, 2657–67.
(20) Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. Virtual screening against highly charged active sites: Identifying substrates of alpha-beta barrel enzymes. *Biochemistry* **2005**, *44*, 2059–71.
(21) Graves, A. P.; Shivakumar, D. M.; Boyce, S. E.; Jacobson, M. P.; Case, D. A.; Shoichet, B. K. Rescoring docking hit lists for model cavity sites: Predictions and experimental testing. *J. Mol. Biol.* **2008**, *377*, 914–34.
(22) Taufer, M.; Armen, R.; Chen, J.; Teller, P.; Brooks, C. Computational multiscale modeling in protein−ligand docking. *IEEE Eng. Med. Biol. Mag.* **2009**, *28*, 58–69.
(23) Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. Calculation of protein−ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. U S A* **2008**, *105*, 6290–5.
(24) Kaminski, G. A.; Ponomarev, S. Y.; Liu, A. B. Polarizable simulations with second order interaction model—force field and software for fast polarizable calculations: Parameters for small model systems and free energy calculations. *J. Chem. Theory Comput.* **2009**, *5*, 2935–2943.
(25) Friesner, R. A. Modeling polarization in proteins and protein−ligand complexes: Methods and preliminary results. *Adv. Protein Chem.* **2005**, *72*, 79–104.
(26) Ha, S.; Andreani, R.; Robbins, A.; Muegge, I. Evaluation of docking/scoring approaches: A comparative study based on MMP3 inhibitors. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 435–48.
(27) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L., III. Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032–47.
(28) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534–53.
(29) *Prime*, V2.2.108; Schrodinger LLC: New York, 2007.
(30) Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. Force field validation using protein side chain prediction. *J. Phys. Chem. B* **2002**, *106*, 11673–11680.
(31) Andrec, M.; Harano, Y.; Jacobson, M. P.; Friesner, R. A.; Levy, R. M. Complete protein structure determination using backbone residual dipolar couplings and sidechain rotamer prediction. *J. Struct. Funct. Genomics* **2002**, *2*, 103–11.
(32) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351–67.
(33) Felts, A. K.; Gallicchio, E.; Chekmarev, D.; Paris, K. A.; Friesner, R. A.; Levy, R. M. Prediction of protein loop conformations using the AGBNP implicit solvent model and torsion angle sampling. *J. Chem. Theory Comput.* **2008**, *4*, 855–868.
(34) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–31.
(35) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein−ligand binding. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 808–13.
(36) Oslob, J. D.; Romanowski, M. J.; Allen, D. A.; Baskaran, S.; Bui, M.; Elling, R. A.; Flanagan, W. M.; Fung, A. D.; Hanan, E. J.; Harris, S.; Heumann, S. A.; Hoch, U.; Jacobs, J. W.; Lam, J.; Lawrence, C. E.; McDowell, R. S.; Nannini, M. A.; Shen, W.; Silverman, J. A.; Sopko, M. M.; Tangonan, B. T.; Teague, J.; Yoburn, J. C.; Yu, C. H.; Zhong, M.; Zimmerman, K. M.; O'Brien, T.; Lew, W. Discovery of a potent and selective aurora kinase inhibitor. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4880–4.
(37) Duplantier, A. J.; Becker, S. L.; Bohanon, M. J.; Borzilleri, K. A.; Chrunyk, B. A.; Downs, J. T.; Hu, L. Y.; El-Kattan, A.; James, L. C.; Liu, S.; Lu, J.; Maklad, N.; Mansour, M. N.; Mente, S.; Piotrowski, M. A.; Sakya, S. M.; Sheehan, S.; Steyn, S. J.; Strick, C. A.; Williams, V. A.; Zhang, L. Discovery, SAR, and pharmacokinetics of a novel 3-hydroxyquinolin-2(1H)-one series of potent D-amino acid oxidase (DAAO) inhibitors. *J. Med. Chem.* **2009**, *52*, 3576–85.
(38) Grunewald, G. L.; Seim, M. R.; Regier, R. C.; Martin, J. L.; Gee, C. L.; Drinkwater, N.; Criscione, K. R. Comparison of the binding of 3-fluoromethyl-7-sulfonyl-1,2,3,4-tetrahydroisoquinolines with their isosteric sulfonamides to the active site of phenylethanolamine N-methyltransferase. *J. Med. Chem.* **2006**, *49*, 5424–33.
(39) Farmer, L. J.; Bemis, G.; Britt, S. D.; Cochran, J.; Connors, M.; Harrington, E. M.; Hoock, T.; Markland, W.; Nanthakumar, S.; Taslimi, P.; Ter Haar, E.; Wang, J.; Zhaveri, D.; Salituro, F. G. Discovery and SAR of novel 4-thiazolyl-2-phenylaminopyrimidines as potent inhibitors of spleen tyrosine kinase (SYK). *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6231–5.
(40) Grey, R.; Pierce, A. C.; Bemis, G. W.; Jacobs, M. D.; Moody, C. S.; Jajoo, R.; Mohal, N.; Green, J. Structure-based design of 3-aryl-6-amino-triazolo[4,3-b]pyridazine inhibitors of Pim-1 kinase. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3019–22.
(41) Krystof, V.; Cankar, P.; Frysova, I.; Slouka, J.; Kontopidis, G.; Dzubak, P.; Hajduch, M.; Srovnal, J.; de Azevedo, W. F., Jr.; Orsag, M.; Paprskarova, M.; Rolcik, J.; Latr, A.; Fischer, P. M.; Strnad, M. 4-Arylazo-3,5-diamino-1H-pyrazole CDK inhibitors: SAR study, crystal structure in complex with CDK2, selectivity, and cellular effects. *J. Med. Chem.* **2006**, *49*, 6500–9.
(42) Yan, S.; Larson, G.; Wu, J. Z.; Appleby, T.; Ding, Y.; Hamatake, R.; Hong, Z.; Yao, N. Novel thiazolones as HCV NS5B polymerase allosteric inhibitors: Further designs, SAR, and X-ray complex structure. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 63–7.
(43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
(44) *Maestro*; Schrodinger LLC: New York, 2007.
(45) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(46) Yu, Z.; Jacobson, M. P.; Friesner, R. A. What role do surfaces play in GB models? A new-generation of surface-generalized born model based on a novel gaussian surface for biomolecules. *J. Comput. Chem.* **2006**, *27*, 72–89.

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

(48) Zhu, K.; Shirts, M. R.; Friesner, R. A.; Jacobson, M. P. Multi-scale optimization of a truncated Newton minimization algorithm and application to proteins and protein−ligand complexes. *J. Chem. Theory Comput.* **2007**, *3*, 640–648.

(49) Wei, H. Y.; Tsai, K. C.; Lin, T. H. Modeling ligand-receptor interaction for some MHC class IIHLA-DR4 peptide mimetic inhibitors using several molecular docking and 3D QSAR techniques. *J. Chem. Inf. Model.* **2005**, *45*, 1343–1351.

(50) Ali, A.; Reddy, G. S.; Cao, H.; Anjum, S. G.; Nalam, M. N.; Schiffer, C. A.; Rana, T. M. Discovery of HIV-1 protease inhibitors with picomolar affinities incorporating N-aryl-oxazolidinone-5-carboxamides as novel P2 ligands. *J. Med. Chem.* **2006**, *49*, 7342–56.

(51) Nazare, M.; Essrich, M.; Will, D. W.; Matter, H.; Ritter, K.; Urmann, M.; Bauer, A.; Schreuder, H.; Czech, J.; Lorenz, M.; Laux, V.; Wehner, V. Novel factor Xa inhibitors based on a 2-carboxyindole scaffold: SAR of P4 substituents in combination with a neutral P1 ligand. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4197–201.

(52) Nazare, M.; Will, D. W.; Matter, H.; Schreuder, H.; Ritter, K.; Urmann, M.; Essrich, M.; Bauer, A.; Wagner, M.; Czech, J.; Lorenz, M.; Laux, V.; Wehner, V. Probing the subpockets of factor Xa reveals two binding modes for inhibitors based on a 2-carboxyindole scaffold: A study combining structure-activity relationship and X-ray crystallography. *J. Med. Chem.* **2005**, *48*, 4511–25.

(53) 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*, 6177–96.

(54) *Maestro*, V 9.1.107; Schrodinger LLC: New York, 2007.

(55) Community Structure-Activity Resource (CSAR). http://www.csardock.org/.

(56) Zhou, H. X.; Gilson, M. K. Theory of free energy and entropy in noncovalent binding. *Chem. Rev.* **2009**, *109*, 4092–107.