JCTC Journal of Chemical Theory and Computation

Improving Docking Accuracy through Molecular Mechanics Generalized Born Optimization and Scoring

Matthew R. Lee* and Yaxiong Sun

Department of Molecular Structure, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1799

Received November 20, 2006

Abstract: Docking methods are typically used within the biopharmaceutical industry for the challenging purposes of suggesting putative binding modes of new chemotypes and for virtual screening. When attempting to satisfy the far more simplistic yet fundamentally important goal of reproducing and identifying the correct binding mode from a cocrystal, all docking methods fail at a rather significant rate, demonstrating room for further improvement in docking methodology. We report a hierarchical method that yields results comparable to the industry-leading docking packages GOLD, Glide, and Surflex. By first using a fast, simple, well-established method, UCSF DOCK 4.0, to rigidly dock conformational ensembles, we successfully generate the correct binding mode in all but 4 of a standard, publicly available set of 79 cocrystals from the PDB. Among these 4 failures (1glq, 1tmn, 1rds, and 8gch), all are highly flexible, highly charged, and not druglike. Subsequently, all resultant docking poses were optimized and scored in the protein with molecular mechanics, using a standard MMGB energy function. In total, this hierarchical method identified the correct binding in 71 of 79 cases (90%), an unprecedented level of accuracy on this highly benchmarked test set. Furthermore, the publicly available energy functions employ only physically based force fields without parameter fitting from this or any other docking test sets.

Introduction

A fundamentally important challenge facing any docking algorithm is the ability to generate and identify the native binding mode of a ligand in its protein target, since failure to reliably accomplish this objective necessarily implies greater difficulty in identifying the correct binding mode for other small molecules, as these programs have been designed to do. However, three known issues have precluded levels of redocking accuracy much greater than 70−80%: (i) insufficient sampling of a complex energy landscape, (ii) inaccuracies in the scoring functions, and (iii) uncertainties in the experimental structures themselves. While the third issue cannot be addressed, other than to identify and discard such problematic cases from consideration, the docking community has long understood the first two issues as distinct from one another and hence worked on ways to address them independently.

In the vast majority of cases, particularly for reasonably sized small molecules, the orientational space for a ligand in a protein is not an intractable problem; only a limited number of low-energy ligand conformers exist at room temperature, and a limited number of poses for each conformer will fit in the protein. In most instances, the total number of sterically permissible poses should be well within the accessible computational time frame of any docking package. Failure to generate the native pose, in most instances, likely stems from an arbitrary limitation on the docking engine resulting from overly restrictive default parameters. In this work, we assess this issue of docking completeness with the earliest, most longstanding of docking methods, UCSF DOCK 4.0,[1] which also is one of the simpler approaches.

In order to address the second issue of scoring accuracy, the docking community has recently drifted away from

* Corresponding author e-mail: mrlee@amgen.com.

10.1021/ct6003406 CCC: $37.00
Published on Web 03/09/2007

energy functions that are entirely physically based, toward empirical scoring functions that involve parameter fitting to reproduce experimental binding energies and/or statistical distributions of atomic interactions in protein−ligand complexes. Like all methods that involve parameter fitting, however, the quality of the results depends on the robustness of the training set. Validations on test sets that do not exhibit significant overlap with the training sets fare worse than validations on test sets possessing greater resemblance to the training sets. One of the potential problems often overlooked is the druglikeness characterization of the training and test sets. In this regard, accurate physically based molecular mechanics energy functions can be applied with a greater level of confidence onto new systems not represented by protein−ligand databases, which have been used for training empirical functions. Historically, one of the major obstacles in using molecular mechanics force fields has been how to account for the solvation energy. Over the past decade, however, advances in the field of continuum solvent models have been made and combined successfully with molecular mechanics energies for accurately reproducing relative binding energies. Perhaps the most popular continuum solvent model has been the Generalized Born approximation which, together with a molecular mechanics energy, gives rise to Molecular Mechanics Generalized Born (MMGB) binding energies. This was first summarized in a review article by Kollman and co-workers[2] and has been applied successfully over the past couple of years on an increasingly diverse set of molecular recognition cases for accurate binding energy predictions,[3−11] for virtual screening,[12] and compared alongside other scoring functions for discriminating native from binding site decoy conformations.[13] However, to date these methods have not been applied explicitly to docking validation studies, in large part due to a perception that reliable implementation of such methods would be computationally prohibitive. In this work, we present a fast, accurate, general hierarchical approach that incorporates an MMGB energy into optimization and scoring procedures as the final stages of a docking process, the culmination of many different attempts on how best to fold these latter methods into the overall docking protocol.

Methods

Test Sets. Over the past decade, a number of papers have assessed the accuracy of different docking methods in correctly predicting the experimental binding mode, with many of these papers using all or part of the publicly available "GOLD benchmark" data set, a total of 134 complexes from the PDB, most of which were curated and reported in 1997 in order to validate GOLD's effectiveness.[11] In 2003, a reduced "81 complex set" was published[12] that removed complexes with more than 15 rotatable bonds, covalent attachments, and obvious errors in structure and that has published data for many other docking methods. With an interest in comparatively assessing the binding mode accuracy of our docking approach with that of others, we also selected this highly benchmarked test set for validation purposes. However, we removed two additional entries: (i) 1lpm, because the ligand menthyl hexyl phosphonate is covalently bound to the gamma oxygen of Ser269 and (ii) 6rsa, because the ligand uridine vanadate contains an unusual vanadium atom, for which the methods in this work lack parameters. We refer to the remaining complexes as the "GOLD 79 set"; note this includes the newer PDB entries of 2ack instead of 1ack and 4aah instead of 3aah.

While the GOLD 79 set spans a diverse range of protein classes and its ligands have a reasonable average of 5 rotatable bonds, the vast majority of these ligands show little resemblance to compounds that would result from a medicinal chemistry effort. Many of them look to be closely related to the endogenous ligands for the protein targets, which are generally of high concentration in vivo and hence low affinity. Among oral drugs, the normal distributions for a number of descriptors that are widely recognized to be important for oral bioavailability have been characterized.[13] The normal range, defined as the limits spanning two standard deviations from the average, the 5th to 95th percentile, for molecular weight is 164 to 589, and for cLogP −1.9 to 6.3. In contrast, this test set consists of compounds for which 67% are in the normal range of molecular weight range of oral drugs and 70% are in the normal range of cLogP, as summarized in Table 1. In addition, and perhaps more problematic, is that these ligands have an average of 1.5 charged groups, with 36 (46%) having two or more charges and 16 (20%) having three or more charges.

In the interest of assessing accuracy of our docking method on a more pharmaceutically relevant test set, we culled through the PDB database and identified 14 targets with the greatest number of cocrystal entries, all of which are well validated targets of historical interest in the biopharmaceutical industry (Table 5). For these targets, we selected one representative complex from each and created what we refer to as the "PDB recurrent 14 set". Not surprisingly, 13 of 14 fall within two standard deviations from the average cLogP of oral drugs, and all 14 are within the statistically normal distribution of oral drugs for molecular weight, number of rotatable bonds, and numbers of hydrogen bond donors and acceptors. In addition, these compounds have an average of 0.8 total charges groups, with two (14%) having two or more charges and one (7%) having three or more charges.

Ligand Conformational Ensembles and Ligand Preparation. While the redocking of a ligand's bound conformation from the native complex has been used to compare accuracies in satisfying a fundamental docking objective of correctly predicting the native binding mode,[14] a more challenging and practical comparison is that of redocking a ligand's conformational ensemble. Without any a priori knowledge of which small molecule conformer a protein prefers, a truly objective method must consider all low-energy conformers. Using the bound conformer from the native complex does not provide a useful indication for the power of a docking run, because it greatly simplifies the problem by not having to deal with the following known difficulties: (i) comparing the native conformer's binding energy with that from other ligand conformations that can fit in the active site pocket in dissimilar ways and (ii) characterizing the energy landscape of a ligand and incorporating each conformer's energy of stability penalty into the scoring function.

Table 1. Properties of the GOLD 79 Set[a]

PDB code	MW	cLogP	# rotatable bonds	# H-bond donors	# H-bond acceptors	# + charges	# - charges
1abe	150	-2.2	0	4	5	0	0
1acj	199	3.5	0	2	1	0	1
1ack	166	-2.4	2	1	1	0	1
1acm	255	-3.2	6	1	8	4	0
1aco	174	-1.6	4	0	6	3	0
1aha	135	-0.1	0	2	4	0	0
1atl	339	2.7	9	2	5	1	0
1baf	397	1.9	7	4	8	0	1
1bbp	583	2.7	11	3	8	2	0
1bma	522	7.0	12	3	3	0	1
1cbs	300	6.4	5	0	2	1	0
1cbx	208	1.0	5	0	4	2	0
1com	226	-1.8	4	1	6	2	0
1coy	288	3.1	0	1	2	0	0
1dbb	314	3.8	1	0	2	0	0
1dbj	290	3.6	0	1	2	0	0
1dr1	237	-3.2	2	4	7	0	0
1dwd	523	1.5	9	4	6	0	1
1eap	343	1.5	10	1	6	3	0
1epb	300	6.4	5	0	2	1	0
1etr	509	1.3	10	5	6	1	1
1fen	270	8.2	4	0	0	0	0
1fkg	450	6.3	10	0	4	0	0
1fki	438	5.1	0	0	6	0	0
1frp	336	-3.5	6	3	12	4	0
1glq	443	-1.3	13	3	9	2	1
1hdc	571	6.6	6	0	7	2	0
1hdy	194	1.6	0	1	1	0	0
1hri	309	4.1	9	0	3	0	0
1hsl	156	-2.1	3	2	3	1	1
1hyt	208	1.0	5	0	4	2	0
1lah	134	-1.9	4	2	2	1	2
1lcp	168	-0.5	3	1	3	2	1
1ldm	89	-1.7	1	1	3	1	0
1lic	307	6.2	15	0	3	0	0
1lna	247	-0.7	8	3	3	1	2
1lst	148	-1.3	5	2	2	1	2
1mdr	166	0.8	2	1	3	1	0
1mrg	135	-0.1	0	2	4	0	0
1mrk	267	-1.7	2	5	8	0	0
1mrk	267	-1.7	2	5	8	0	0
1nco	661	1.9	8	4	12	0	1
1phg	226	1.5	3	0	3	0	0
1rds	589	-10.5	8	6	16	2	0
1rob	323	-2.5	4	3	10	2	0
1snc	402	-2.9	6	1	11	4	0
1srj	292	4.0	3	1	5	1	0
1stp	244	-0.3	5	2	4	1	0
1tka	424	-0.8	8	5	10	4	1
1tmn	481	3.8	13	3	5	2	1
1tng	114	2.0	1	1	0	0	1
1tni	150	2.4	4	1	0	0	1
1tnl	134	1.5	1	1	0	0	1
1trk	425	-1.4	8	2	11	4	1
1ukz	427	-3.9	6	3	14	4	0
1ulb	151	-1.2	0	3	4	0	0
1wap	205	0.3	3	2	2	0	1
2ada	270	-1.7	2	4	9	0	0
2ak3	347	-3.2	4	3	11	2	0
2cgr	384	4.0	8	2	6	1	0
2cht	228	-1.8	2	1	6	2	0
2cmd	192	-2.4	5	1	7	3	0
2ctc	166	0.5	3	1	3	1	0
2dbl	419	5.0	6	0	5	1	0
2gbp	180	-2.2	1	5	6	0	0
2lgs	148	-2.2	4	1	4	2	1
2phh	138	0.3	1	1	3	1	0
2r07	347	4.1	8	0	3	0	0
2sim	291	-3.1	5	5	8	1	0
3aah	238	-0.9	3	1	9	3	0
3cpa	266	2.7	5	3	5	1	1
3hvt	120	0.8	1	1	4	0	0
3ptb	231	1.1	1	2	0	0	1
3tpi	330	-2.1	6	2	3	1	1
4cts	132	-1.6	3	0	5	2	0
4dfr	454	-1.2	9	3	12	2	0
6abp	150	-2.2	0	4	5	0	0
6rnt	345	-2.4	4	3	11	2	0
7tim	171	-2.6	3	2	6	2	0
8gch	333	-0.5	7	4	4	1	1
AVG	286	0.5	4.7	1.9	5.3	1.1	0.4
Median	267	-0.1	4	2	5	1	0

[a] The four DOCK 4.0 failures are highlighted in red (see Table 2).

For the sake of completeness, we sought to create exhaustive conformational ensembles. To this end, we used Amgen's internally developed program FLAME,[15] which uses the OpenEye toolkits[16] OEChem and CASE and incorporates a genetic algorithm to search the torsional space resulting from using discrete 10° dihedral angles for each rotatable bond (36 total angles per torsion). With the exception of hydroxyl groups, dihedral angles that included hydrogen atoms were not considered rotatable during the conformational search. Conformers were discarded if they were either less than 0.2 Å away of an existing member of the ensemble, in terms of root-mean-square deviation of all heavy atom Cartesian coordinates, or 10 kcal/mol above the global energy minimum conformer, according to the MMFF94[17] force field, including the partial atomic charges, with a Sheffield solvation term.[16] In addition, the maximum number of conformers per ligand was set to 200; in cases where less than 200 conformers were generated, the conformational ensemble is considered exhaustive.

Ligands were extracted from the PDB and converted to mol2 format with explicit hydrogens, using the OEChem toolkit,[16] which has the advantages of assigning bond orders and of ionizing basic amines and acids. In each case, the resulting mol2 file was visually inspected to ensure that the bond order was in agreement with its 2D representation on the Brookhaven PDB Web page, with a few rare instances leading to incorrect bond orders; in every instance, the ionization states were in agreement. In a few cases, the ionization state reported in the complex's original reference did not match with this canonical ionization state. Tacrine, the ligand in 1acj, contains a 4-aminoisoquinoline that is considered neutral on both the Brookhaven Web page and by the OEChem toolkit but has experimentally been determined to contain a highly basic nitrogen in the ring,[18] which in turn forms a hydrogen bond with a backbone carbonyl oxygen in the complex. Unexpectedly, L-leucine phosphonic acid, the ligand in 1lcp, reportedly binds with a neutralized phosphate group.[19] Last, the ligand from 4aah, pyrroloquino-

line quinoline, reportedly employs a neutralized carboxylate oxygen (O2A) in order to hydrogen bond with the carboxylate group of Glu55.[20]

Protein Preparation. To prepare the proteins for redocking, all water molecules, cofactors, counterions, and ligands were stripped from each complex. (As a lone exception, two crystallographic water molecules were kept in 1lna, which are tightly bound to the counterion.) Hydrogen atoms were added and charges were added according to the Cornell et al.[21] all-atom force field. All histidine residues in direct or indirect contact with the ligands, cofactors, or counterions were visually inspected and assigned the appropriate tautomeric or charged state, based on what appeared to be most energetically favorable in the absence of ligands and cofactors. By default, histidine residues were assigned the tautomeric state in which the delta nitrogen is protonated (HID), modified to the alternative tautomer if the epsilon nitrogen was better served as the lone protonated nitrogen (HIE) or in rare instances assigned the charged state when both nitrogen atoms needed to serve as hydrogen bond donors (HIP) in the absence of ligand and cofactors. Similarly, tyrosine, serine, and threonine residues in close proximity to the active site small molecules were visually inspected to make sure that the dihedral of the hydroxyl group placed the proton in the more appropriate environment. Finally, fully charged residues, such as aspartic acids under physiological conditions, were also carefully inspected to see if two residues with like charges directed the termini of their side chains toward each other. In such cases, we neutralize one of the two residues, as in the case of 1trk, where we used the neutralized form of Glu162. Prior to the docking and optimization, cofactors and counterions in the active site were added back in.

DOCK 4.0. Regardless of which software is used, generating the native pose upon docking an exhaustive conformational ensemble of the native ligand into its cognate protein structure should be a solvable task. If the docking run fails to produce the native pose, this likely reflects a sampling issue, either resulting from an insufficient ligand conformational ensemble or from sampling issues within the docking engine itself, as is the case with the default parameters in DOCK 4.0, where we made changes to a number of the input parameters and added an additional parameter. Perhaps most importantly, we affect the sampling by turning on the bump filter and minimization and have DOCK 4.0 search along a largely shape-based energy function; the intermolecular energy score is turned on, while leaving electrostatics off by setting the electrostatic scale to 0. Ligand orientation is also turned on, and, in order to reduce the probability of insufficiently exploring the entire orientational space, we increase the default of 100 maximum orientations to 50 000 while saving the top 15 minimized orientations in terms of intermolecular energy. The chief sampling parameter for matching, the distance tolerance is also modified from 0.25 to 0.6 to allow for greater variance from the site points. Because there are several very small ligands with less than 10 heavy atoms, the distance minimum parameter is set to 0 in order to allow all pairs of atoms to match. Finally, in order to effect a soft-shell steric overlap potential to allow for resultant poses that might be more amenable to subsequent optimization into the native state, we leave the vdW scaling factor at the default value of 1 but have modified the source code to allow for a user-specified maximum energy for any atom, which we set at 3 kcal/mol in this work.

With DOCK 4.0, one must generate site points that create an inverse image of the binding site on which each conformer is docked. In this study, the site points were obtained by randomly perturbing the Cartesian coordinates for each heavy atom of the bound conformer, by up to 1 Å in each dimension for up to a total of 1.73 Å, and saving one-third of these resulting coordinates as site points, albeit with a minimum of five points. In our experience, the exact site point positions do not affect docking results significantly for the protocols used in this study. For comparison, we verified this on the PDB Recurrent 14 Set, by using the sphgen accesory that comes with DOCK 4.0 to generate site points, and this had no impact on the quality of docked results in terms of the ability to generate the native binding mode.

Optimization. An approach with growing popularity is to rely on a docking program to generate final poses that are subsequently used for scoring by an energy function which is independent from that which was used to dock the compounds. One of the known limitations of DOCK 4.0 and many other algorithms used for rigid-body docking is the inability to effect the requisite gently relaxation of torsions, angles, and bond lengths in response to the local environment of the protein binding site that distinguishes the bound conformation from any member of its conformational ensemble. (These discrepancies can arise because, while any method for generating conformational ensembles can sample exhaustively over the conformational space resulting from a given number of discrete torsions over each rotatable bond, it cannot sample exhaustively over a continuum of torsions.) Another widespread problem is not accounting for the plasticity of a protein in response to the ligand, although this type of simplification does not present a considerable obstacle on redocking exercises, except in cases where steric clashes and other local inaccuracies in the experimental structures exist. Subsequent to rigid-body docking of a conformational ensemble in the rigid protein and prior to scoring, in order to help address the limited sampling issues discussed above, we subjected each pose that emerged from DOCK 4.0 to a molecular mechanics optimization routine, comparing the effectiveness of several different protocols, varying ligand and protein flexibility.

In the Glide validation paper,[22] the authors discuss the importance of relaxing a protein in the presence of its cognate ligand, prior to the redocking exercise, in order to anneal away "untenable steric clashes often found in crystallographically determined protein sites", to which hard 12−6 Lennard-Jones vdW potentials are extremely sensitive. Like Glide, the methods described in this work also use the hard repulsive potentials in both the optimization and in the scoring. We evaluate our method, by performing the ligand optimization both in the presence of a rigid relaxed protein, analogous to the Glide validation study, and in the presence of the less biased original PDB heavy atom positions that

are flexible during the optimization. All optimizations in this work were carried out with Amber 7,[23] using the Cornell et al.[21] force field for proteins, the GAFF[24] force field for small molecules, and the default continuum solvent implementation, which is described in greater detail below in the following section. In our early implementation of the MMGB optimization method, we observed a number of resultant small molecules with conformations that we deemed unacceptable, particularly for the following functional groups: sulfonamides, phosphates, anilinic nitrogens, amides, nitriles, and alkynes. To correct for these problems, we commented out some of the GAFF parameters and modified a number of others by increasing the force constant to exaggerate preferences for desired small molecule free-state conformations (available in the Supporting Information).

In order to generate a relaxed representation on the protein active site that is subsequently used for ligand optimization in a rigid protein, we first defined a flexible residue list to include amino acids in the original PDB complex that show steric overlap with the native ligand pose and those fully charged amino acids that are in direct contact with the inhibitor. Because both proteins and ligands include explicit hydrogen atoms, we define amino acids in overlap with the ligand as those containing any atom less than 2 Å from any atom of the ligand. Starting with the original PDB complex, these flexible residues were crudely minimized in the presence of the entire complex until a loose convergence criterion of 5.0 kcal/mol·Å for the root-mean-square of the Cartesian elements of the energy gradient (DRMS) was reached. The flexible residues in the complex were then subjected to a 2 ps molecular dynamics simulated annealing schedule from 100 to 400 K over the first picosecond and from 400 K down to 5 K over the second picosecond, followed by a final minimization with a tighter convergence criterion, a DRMS of 0.6 kcal/mol·Å. Ligands were removed, and the resulting protein structure was used for rigid protein optimization of every resulting pose produced by Dock. In the presence of this relaxed protein, each docked pose is crudely minimized restraint-free, while the protein is held frozen, until the DRMS reaches 5.0 kcal/mol·Å. At this point, final optimized poses were created by either (i) restraint-free minimization alone until the DRMS reaches 0.6 kcal/mol·Å or (ii) 2 ps of restraint-free simulated annealing molecular dynamics followed by restraint-free minimization. In total, we ran each docked conformer through two separate rigid protein optimizations.

As an alternative approach to optimizing the docked poses in a rigid relaxed protein, we also optimized them in the original PDB protein conformation, while allowing for plasticity of the protein in response to the ligand, a protocol which we refer to as flexible protein optimization. After placing each pose in the protein, a flexible residue list was determined based solely on the 2 Å distance cutoff designation for steric clashes. These flexible residues then conform to each individual docked pose through crude minimization (DRMS of 5.0 kcal/mol·Å) and 2 ps of simulated annealing molecular dynamics, while keeping the ligand and all other residues frozen. We then add ligand flexibility back into the system by allowing both ligand and flexible protein residues to respond to each other and move simultaneously through a crude minimization (DRMS of 5.0 kcal/mol·Å), followed by either minimization alone or by simulated annealing molecular dynamics and minimization. During the latter part of the flexible protein optimization, ligand flexibility was run either restraint-free or with a positional restraint of 0.5 kcal/ mol·Å on all heavy atoms. In total, we ran each docked conformer through four separate flexible protein optimizations: (1) minimization alone with ligand restraint, (2) minimization alone without ligand restraint, (3) dynamics and minimization with ligand restraint, and (4) dynamics and minimization without ligand restraint. All molecular dynamics simulations were carried out in the presence of an implicit continuum solvent model, which we describe in the next section.

During the validation phase of this approach, we experimented with many alternative optimization protocols, varying the equilibration and relaxation, the simulated annealing schedule, the length of the molecular dynamics production run, and the extent of minimization. We arrived at the protocol described in this work as that which involved the minimum threshold number of CPU cycles necessary to lead to robust optimization of bound poses. Given that the production dynamics run entails only 2 ps of simulation time, the optional molecular dynamics addition adds minimal overhead to this approach (discussed in greater detail below).

MMGB Energy for Scoring. Over the past several years, molecular mechanics, together with continuum solvent representations, have been used to predict binding energies,[2] according to the following series of equations:

$$\Delta G_{\text{bind}} = (\Delta H_{\text{bind}} + \Delta\Delta G_{\text{solv}}) - T\Delta S \tag{1}$$

$$\Delta H_{\text{bind}} = H_{\text{complex}} - (H_{\text{protein}} + H_{\text{ligand}}) \tag{2}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{solv,np}} + \Delta G_{\text{solv,polar}} \tag{3}$$

The enthalpy terms of eq 2 are taken from the molecular mechanics energy. For estimating the polar contribution to the free energy of solvation ($\Delta G_{\text{solv,polar}}$), bulk medium can be treated as a continuum solvent, and the generalized Born equation used[25] is

$$\Delta G_{\text{elec}} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{\epsilon r_{ij}} - \frac{1}{2}\left(1 - \frac{1}{\epsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{4}$$

where r_{ij} represents the interatomic distance between atoms i and j, and a_i is the effective Born radius for atom i. In its true form, the effective Born radius for a given atom in a system is that which would lead the generalized Born equation to return the correct electrostatic energy of the system when all other atoms are discharged and serve only as an effective dielectric medium between the atom and the solvent.

Still and co-workers subsequently incorporated an approximation to the generalized Born equation into a molecular mechanics equation,[26] in which the two terms are

combined into a single expression that is a function of both r_{ij} and a_i:

$$\Delta G_{\text{solv,polar}} = -\frac{1}{2}\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{f(r_{ij},a_{ij})} \quad (5a)$$

$$f(r_{ij},a_{ij}) = \sqrt{(r_{ij}^{\,2} + a_{ij}^{\,2}e^{-D})} \quad (5b)$$

$$a_{ij} = \sqrt{a_i a_j} \quad (5c)$$

$$D = \frac{r_{ij}^{\,2}}{(2a_{ij})^2} \quad (5d)$$

For all of our calculations, we used the default generalized Born approximation model[27] that comes with AMBER 7.

The nonpolar contribution to solvation includes the cost of creating a solute-sized cavity in solvent and the free energy of inserting the discharged solute into that cavity, which has been found experimentally in hydrocarbons to be linearly related to the solvent accessible surface area (SASA):

$$\Delta G_{\text{solv,nonpol}} = \gamma\bullet\text{SASA} + b \quad (6)$$

The γ coefficient is set to 5.42 cal/mol·Å^2, and b is set to 920 cal/mol. While we included this nonpolar solvation term in our binding energy calculations, the differences in surface areas between various bound poses of a single ligand are minor and do not have a significant effect on the relative binding energies of different poses. Even when comparing different ligands, this effect is generally insignificant.

In this work, we ignore the ΔS term of eq 1, which is a valid approximation when comparing predicted binding energies among different conformers of the same ligand. This omission, however, can bias toward larger, more flexible ligands when comparing different compounds with largely varying degrees of freedom.

Because our final binding energy score is predominantly a measure of the molecular mechanics enthalpic energy and the generalized Born polar solvation terms, we refer to it as the MMGB energy, which is of the following form

$$\begin{aligned}\Delta G_{\text{bind}} \cong \Delta\text{MMGB}_{\text{bind}} = \Delta H_{\text{bind}} + \Delta\Delta G_{\text{solv}} = \\ (E_{\text{MM}} + \Delta G_{\text{solv,GBA}} + \gamma\bullet\text{SASA} + b)_{\text{complex}} - (E_{\text{MM}} + \\ \Delta G_{\text{solv,GBA}} + \gamma\bullet\text{SASA} + b)_{\text{ligand}} - (E_{\text{MM}} + \Delta G_{\text{solv,GBA}} + \\ \gamma\bullet\text{SASA} + b)_{\text{protein}} \quad (7)\end{aligned}$$

where E_{MM} is the AMBER molecular mechanics energy resulting from the Cornell et al. force field[21] for proteins and the GAFF force field[24] for small molecules, and $\Delta G_{\text{solv,GBA}}$ is the AMBER implementation of the generalized Born approximation equation.[27]

Four Different MMGB Optimization and Scoring Protocols. For each system in the test sets, we performed one docking run and passed each of the resultant docked poses into rigid protein optimizations and flexible protein optimizations. Rigid protein optimizations were run restraint free with minimization alone or with molecular dynamics followed by minimization. Because molecular dynamics involves thermal energy, thereby sampling along a free energy landscape rather than along the enthalpic energy landscape, it populates local minima differently than a minimization alone optimization. As a result, we separate the scoring protocols in accordance to the extent of optimization carried out, with each rigid protein optimization resulting in its own independent score, "MMGB Rigid Min" or "MMGB Rigid MD/Min" in Tables 2−5.

Flexible protein optimizations were also run with minimization alone or with molecular dynamics followed by minimization. However, distinct from the rigid protein optimizations, flexible protein optimizations were run both with and without a positional restraint on the ligand and flexible protein residues, for a total of 4 flexible protein optimizations. One of the potential problems during the dynamics optimization stage is that slightly unfavorable contacts with the protein can drive the ligand out of the binding site, despite its geometric proximity to the native state. For the rigid protein scheme, the relaxation of untenable interactions precluded such problematic contacts from arising during the optimization, but for the flexible protein scheme, ligands were run both with and without the aforementioned 0.5 kcal/mol·Å positional restraints on ligand heavy atoms. This assures that each docked pose can respond to the protein environment resulting from soft-shell docking both restraint free and also with a gentle tether in place. But while the flexible protein optimization with and without positional restraints on the ligand leads to different final optimized poses, they are combined into a single set with MMGB energies compared alongside one another. Thus, the two minimization alone flexible protein optimizations for each docked pose (with and without restraint) are considered together as the "MMGB Flex Min" protocol, and the two molecular dynamics and minimization flexible protein optimizations are considered together as the "MMGB Flex MD/Min" protocol. For all 4 protocols, the final MMGB binding energy is calculated on the final minimized structure in the absence of restraints. The entire process flow of our 4 protocols is summarized in Figure 1.

In addition to the four protocols compared in this work, our implementation of the MMGB optimization and scoring stages allows for even greater user control over how to run the optimization, all from the command line. The user can select specific residues in addition to or in place of those within a specified distance of the ligand's input pose. Water molecules can be explicitly added and treated as part of the "protein"; when treated as rigid, the hydrogen atoms remain flexible to optimize their positions that are generally not included in the PDB entry. Instead of running molecular dynamics for 2 ps, when greater conformational motion is desired, the user can increase the time frame to the desired level. Side-chain-only motion for flexible residues can be specified. The magnitude of the positional restraint on the inhibitor is also modifiable. Depending on the nature of the problem, we routinely generate models using various permutations of these variables.

Results

Completeness of Conformational Ensembles and Rigid Docking. By imposing an arbitrary limit of 200 maximum conformers per ensemble, we knowingly precluded en-

Table 3. rmsds of Best Scoring Poses[a]

PDB code	FlexX	GOLD	Surflex	Glide	UCSF Dock 4.0	MMGB Rigid		MMGB Flex	
						Min	MD/Min	Min	MD/Min
1abe	1.16	0.86	0.27	0.17	3.34	0.46	0.30	0.54	2.69
1acj	0.49	4.00	3.89	0.28	5.31	0.35	0.25	4.75	4.75
1ack	2.21	4.99	1.18	0.97	4.74	1.82	1.42	1.85	0.89
1acm	1.39	0.81	1.43	0.29	5.50	1.03	1.08	1.51	1.04
1aco	0.96	0.86	3.39	1.02	2.43	0.35	0.36	0.68	0.43
1aha	0.56	0.51	0.37	0.11	2.98	0.36	0.41	0.37	0.48
1atl	2.06	1.37	7.01	0.94	3.51	2.55	0.47	0.60	3.51
1baf	8.27	6.12	6.52	0.76	2.10	6.14	6.47	5.72	5.72
1bbp	3.75	8.22	1.07	4.96	8.36	1.55	1.62	4.45	1.58
1bma	13.41	1.03	1.00	9.31	5.05	0.85	0.63	1.73	0.97
1cbs	1.68	2.05	1.77	1.96	1.93	0.42	0.67	2.16	2.82
1cbx	1.35	0.54	0.70	0.36	5.83	1.05	1.04	0.75	7.89
1com	1.62	1.28	0.86	3.64	1.03	1.46	1.78	0.73	3.97
1coy	1.06	0.86	0.54	0.28	7.91	1.27	0.49	0.65	1.14
1dbb	0.81	1.17	0.54	0.41	0.99	0.59	0.85	0.71	1.30
1dbj	1.22	0.72	0.88	0.20	5.99	0.85	0.46	0.78	0.57
1dr1	5.64	1.41	1.25	1.47	5.66	1.53	0.47	1.58	0.71
1dwd	1.66	1.71	1.68	1.32	4.64	0.62	1.34	0.63	0.60
1eap	3.72	3.00	4.89	2.32	5.03	0.88	0.82	1.14	2.17
1epb	2.77	2.08	2.87	1.78	3.17	2.60	1.39	2.68	2.23
1etr	7.24	4.23	4.05	1.48	3.54	0.41	0.69	2.16	2.09
1fen	1.39	1.19	1.18	0.66	0.51	1.32	1.18	0.62	0.49
1fkg	7.59	1.81	1.81	1.25	6.23	1.42	1.99	1.61	1.61
1fki	0.59	0.71	0.70	1.92	0.73	0.51	1.06	0.47	0.41
1frp	1.89	0.66	0.75	0.27	3.85	0.55	0.68	2.49	1.85
1glq	6.43	1.35	5.68	0.29	7.37	1.91	0.52	4.92	2.65
1hdc	11.74	10.49	1.80	0.58	11.05	2.40	1.03	11.14	10.75
1hdy		0.94	0.66	1.74	0.38	1.61	0.52	1.74	1.75
1hri	10.23	14.01	1.98	1.59	2.43	1.65	2.41	1.70	2.33
1hsl	0.59	0.97	0.51	1.31	1.25	1.87	1.13	2.01	1.87
1hyt	1.62	1.10	0.55	0.28	5.45	1.18	1.61	1.17	0.62
1lah	0.28	0.27	0.30	0.13	1.03	0.24	0.23	0.33	0.45
1lcp	1.65	0.98	2.01	1.98	2.48	0.93	1.46	1.32	1.30
1ldm	0.74	1.00	0.44	0.30	5.52	1.35	1.35	1.53	4.70
1lic	5.07	10.78	3.46	4.87	5.03	1.23	1.25	1.33	2.35
1lna	5.40	3.44	0.88	0.95	5.51	1.04	2.56	1.18	5.82
1lst	0.71	0.87	0.33	0.14	1.24	0.67	0.58	0.42	0.68
1mdr	0.88	0.36	0.68	0.52	1.99	1.57	1.58	1.50	1.55
1mrg	0.81	0.57	0.70	0.30	2.96	0.50	0.31	3.03	0.36
1mrk	3.55	1.01	0.85	1.20	3.55	0.84	1.37	2.21	2.75
1nco	8.26	1.18	8.26	6.99	2.00	10.47	0.64	2.32	1.40
1phg	4.44	1.35	4.44	4.32	1.42	0.40	0.37	0.52	4.93
1rds	9.83	4.78	9.83	3.75	10.46	9.29	2.03	6.11	2.03
1rob	0.82	3.75	0.82	1.85	5.51	0.75	0.51	1.13	2.99
1snc	4.92	2.36	4.92	1.91	4.88	1.25	1.20	1.69	1.75
1srj	0.39	0.42	0.39	0.58	7.01	2.35	0.83	0.67	1.07
1stp	0.51	0.69	0.51	0.59	0.76	0.55	0.53	0.56	0.69
1tka	1.96	1.88	1.96	2.28	1.64	1.69	1.77	1.63	1.51
1tmn	1.30	1.68	1.30	2.80	9.65	3.94	3.48	4.06	2..35
1tng	0.22	0.95	0.22	0.19	3.18	0.65	0.32	0.66	0.26
1tni	2.97	1.98	2.97	2.18	3.68	1.65	2.00	2.54	2.98
1tnl	2.26	0.56	2.26	0.23	4.35	0.65	0.71	1.61	4.42
1trk	1.22	0.76	1.22	1.64	1.08	0.86	1.55	1.63	1.87
1ukz	0.77	2.87	0.77	0.37	1.04	0.46	0.59	1.90	1.12
1ulb	0.77	0.32	0.77	0.28	1.07	0.56	0.63	0.62	0.54
1wap	0.30	0.42	0.30	0.12	0.83	0.30	0.25	0.29	0.31
2ada	0.32	0.40	0.32	0.53	6.13	0.44	0.40	0.59	0.33
2ak3	0.60	5.08	0.60	0.71	3.75	0.91	0.72	0.90	0.68
2cgr	1.63	0.99	1.63	0.38	1.71	0.93	1.38	0.93	4.87
2cht	0.42	0.59	0.42	0.42	1.00	0.70	0.48	0.53	0.56
2cmd	1.60	1.68	1.60	0.65	3.78	0.65	0.48	1.43	3.17
2ctc	0.38	0.32	0.38	1.61	2.78	0.41	0.51	1.81	0.31
2dbl	0.81	1.31	0.81	0.69	1.40	0.97	0.92	1.21	1.48
2gbp	0.63	0.39	0.63	0.15	0.45	0.50	0.31	0.55	0.68
2lgs	1.22	4.41	1.22	7.55	2.62	1.48	1.67	4.53	1.71
2phh	0.44	0.72	0.44	0.38	0.89	0.60	0.59	0.59	0.61
2r07	1.35	8.23	1.35	0.48	1.14	0.99	0.99	1.08	2.91
2sim	1.10	0.92	1.10	0.92	4.34	1.01	1.28	1.01	1.35
3aah	5.93	0.42	0.68	0.30	0.32	0.32	0.33	0.29	0.35
3cpa	2.53	1.58	1.90	2.40	1.46	4.14	8.44	3.66	1.98
3hvt	10.26	1.12	1.64	0.77	0.80	0.82	0.62	0.94	0.98
3ptb	0.55	0.96	0.54	0.27	3.12	0.23	0.35	0.27	0.18
3tpi	1.07	0.80	0.52	0.49	2.09	0.27	0.23	0.46	0.49
4cts	1.53	1.57	2.20	0.19	2.11	0.30	0.30	3.67	3.69
4dfr	1.40	1.44	1.60	1.12	8.56	1.16	0.64	4.31	3.36
6abp	1.12	1.08	0.28	0.40	3.13	2.76	0.28	2.06	3.23
6rnt	4.79	1.20	7.03	2.22	7.15	2.79	1.01	1.78	1.77
7tim	1.49	0.78	1.20	0.14	4.60	0.60	0.47	1.04	0.73
8gch	8.91	0.86	4.51	0.30	4.18	3.57	4.77	3.59	2.96
AVG	2.76	2.06	1.84	1.36	3.59	1.42	1.14	1.80	2.03
Median	1.40	1.11	1.14	0.70	3.15	0.93	0.72	1.38	1.55
n	78	79	79	79	79	79	79	79	79
n (< 2.0)	51	60	60	65	26	67	71	57	49
% (< 2.0)	65%	76%	76%	82%	33%	85%	90%	72%	62%
n (< 2.5)	54	63	63	70	33	70	74	64	55
% (< 2.5)	69%	80%	80%	89%	42%	89%	94%	81%	70%

[a] Coloring is the same as in Table 2. In addition, for cases that MMGB scoring failed with the MMGB Rigid MD/Min protocol, their PDB codes are in magenta cells.

sembles of highly flexible molecules from being exhaustive, with the expectation that 200 would still be sufficient in the majority of cases for DOCK 4.0 to find the native binding mode. In the GOLD 79 set, ligands from 64 of the complexes (81%) were exhaustive, as is the case when the FLAME run completes prior to reaching the limit of 200 conformers. Among the 15 ligands where the ensemble was not complete, the 200 representative conformers still provided sufficient coverage of conformational space for DOCK 4.0 to generate the native binding pose in all but two exceptions, thermolysin (1tmn) and ribonuclease Ms (1rds), which contain exceptionally flexible ligands having 13 and 8 rotatable bonds, respectively. In addition, among the 64 exhaustive conformational ensembles, DOCK 4.0 failed to produce the binding pose on two other cases, gamma-chymotrypsin (8gch) and glutathione S-transferase (1glq), which also contain highly flexible ligands with 7 and 13 rotatable bonds, respectively. The structures of these four docking failures are shown in Figure 2. In total, the rigid docking by DOCK 4.0 of FLAME conformational ensembles described in this work succeeded at the 2 Å rmsd level in 75 of GOLD 79 set cases (95%). One of the key results reported in the validation paper of Surflex[12] is that its failure rate for returning a pose within 2.5 Å was 5/81 cases (6%) and half that of GOLD. By that measure of evaluation, our conventional use of DOCK 4.0, as described in the methods section, demonstrates a 0% failure rate, the two MMGB Min scoring protocols maintained the 0% failure rate, while the two MMGB MD/Min protocols introduced one failure (8gch) for a 1% failure rate (Table 2).

Among the ligands in the PDB recurrent 14 set, all conformational ensembles were exhaustive, and rigid docking succeeded at the 2 Å rmsd level in all cases in finding the native state.

Accuracy of MMGB Optimization and Scoring. While docking methods should be expected to generate the native

Table 2. Lowest rmsds Found in Surflex and Four MMGB Protocols[a]

PDB code	Surflex	UCSF Dock 4.0	MMGB Rigid		MMGB Flex	
			Min	MD/Min	Min	MD/Min
1abe	0.18	1.81	0.46	0.30	0.53	0.18
1acj	0.81	0.43	0.28	0.25	0.27	0.29
1ack	0.50	0.82	0.72	1.06	0.76	0.75
1acm	1.35	1.27	0.75	0.64	1.15	0.76
1aco	0.63	0.77	0.34	0.34	0.51	0.32
1aha	0.26	0.69	0.28	0.35	0.30	0.29
1atl	1.05	1.16	0.55	0.47	0.54	0.71
1baf	1.10	1.04	0.69	0.69	0.52	0.80
1bbp	1.06	1.96	0.59	1.09	1.59	1.30
1bma	0.86	1.02	0.83	0.63	0.95	0.77
1cbs	0.52	0.75	0.42	0.40	0.52	0.45
1cbx	0.70	0.75	1.03	1.04	0.75	0.81
1com	0.64	0.94	0.49	0.91	0.64	0.45
1coy	0.32	1.48	1.13	0.36	0.47	0.49
1dbb	0.28	0.89	0.57	0.58	0.71	0.38
1dbj	0.57	1.00	0.81	0.46	0.78	0.57
1dr1	0.29	0.96	0.53	0.36	0.76	0.32
1dwd	0.93	0.74	0.62	0.92	0.63	0.48
1eap	0.92	1.09	0.86	0.59	0.99	0.57
1epb	1.52	1.08	0.82	0.51	0.93	0.72
1etr	3.01	1.57	0.41	0.65	0.49	0.44
1fen	0.79	1.42	0.36	0.38	1.00	0.36
1fkg	1.52	1.35	1.26	0.61	1.26	0.68
1fki	0.64	0.76	0.51	1.06	0.43	0.41
1frp	0.50	0.69	0.52	0.51	0.58	0.58
1glq	3.49	2.07	1.91	0.52	2.04	0.95
1hdc	1.47	1.84	1.60	0.97	1.72	1.26
1hdy	0.41	0.64	0.41	0.44	0.83	0.70
1hri	1.96	1.22	0.92	0.64	1.04	0.91
1hsl	0.35	1.25	0.81	0.27	1.09	0.87
1hyt	0.53	0.79	0.58	0.51	0.54	0.49
1lah	0.25	0.68	0.20	0.21	0.33	0.21
1lcp	0.54	0.61	0.56	0.35	0.45	0.44
1ldm	0.38	1.21	0.86	0.36	1.21	0.92
1lic	2.19	1.00	0.53	0.96	0.62	0.57
1lna	0.60	0.98	0.70	1.60	0.56	1.00
1lst	0.23	0.55	0.29	0.18	0.19	0.18
1mdr	0.45	0.74	0.92	0.87	0.72	0.77
1mrg	0.60	0.76	0.49	0.31	0.65	0.36
1mrk	0.75	0.77	0.84	1.12	0.95	0.90
1nco	6.69	1.89	1.84	0.64	1.87	0.78
1phg	0.41	0.53	0.35	0.36	0.38	0.37
1rds	4.79	2.25	2.34	1.23	2.29	1.24
1rob	0.56	0.67	0.50	0.44	0.56	0.31
1snc	2.44	1.36	1.25	1.10	1.43	1.10
1srj	0.35	0.61	0.44	0.57	0.55	0.52
1stp	0.46	0.70	0.30	0.33	0.43	0.28
1tka	1.49	1.12	0.90	1.36	0.87	0.51
1tmn	0.65	2.24	2.14	1.21	2.17	0.98
1tng	0.20	0.78	0.44	0.23	0.66	0.26
1tni	1.33	0.83	0.70	0.96	0.70	0.76
1tnl	0.23	0.66	0.58	0.42	0.59	0.50
1trk	0.78	1.00	0.51	0.43	0.76	0.50
1ukz	0.25	1.01	0.42	0.54	0.78	0.52
1ulb	0.58	0.75	0.56	0.33	0.50	0.54
1wap	0.24	0.46	0.29	0.24	0.29	0.23
2ada	0.29	0.82	0.44	0.34	0.39	0.32
2ak3	0.58	0.65	0.37	0.58	0.50	0.47
2cgr	0.89	0.71	0.77	0.77	0.86	0.72
2cht	0.42	0.55	0.49	0.45	0.45	0.49
2cmd	1.49	0.74	0.65	0.47	0.70	1.02
2ctc	0.32	1.19	0.41	0.31	0.95	0.31
2dbl	0.66	0.93	0.89	0.67	0.90	0.61
2gbp	0.27	0.45	0.50	0.27	0.45	0.31
2lgs	0.79	1.69	1.30	1.36	1.24	1.19
2phh	0.41	0.48	0.45	0.47	0.47	0.43
2r07	1.09	0.65	0.57	0.57	0.61	0.64
2sim	0.35	0.46	0.36	0.42	0.38	0.24
3aah	0.33	0.32	0.27	0.31	0.26	0.26
3cpa	0.66	1.02	0.57	0.58	0.62	0.55
3hvt	1.61	0.38	0.31	0.28	0.36	0.33
3ptb	0.37	0.67	0.21	0.21	0.27	0.18
3tpi	0.37	0.36	0.27	0.21	0.37	0.49
4cts	0.53	1.37	0.30	0.29	1.16	0.38
4dfr	1.24	1.53	1.00	0.58	1.41	1.25
6abp	0.14	1.56	1.52	0.28	1.54	0.24
6rnt	4.68	1.72	1.61	0.80	1.66	1.13
7tim	0.34	0.90	0.50	0.38	0.48	0.60
8gch	0.96	2.35	2.45	2.53	2.38	2.85
AVG	0.95	1.01	0.72	0.61	0.81	0.62
Median	0.60	0.86	0.57	0.51	0.66	0.52
n	79	79	79	79	79	79
n (< 2.0)	72	75	76	78	75	78
% (< 2.0)	91%	95%	96%	99%	95%	99%
n (< 2.5)	74	79	79	78	79	78
% (< 2.5)	94%	100%	100%	99%	100%	99%

[a] The four cases that DOCK 4.0 failed on have their PDB codes highlighted in red. In the table body, rmsd values exceeding 3.0 Å are in red cells; those that are less than 2.0 Å are considered correct and are in cyan cells.

pose in the vast majority if not all cases, the more challenging problem of identifying the native pose as the most energetically favorable is considerably more formidable and can depend heavily on how one handles characterization of the protein. FlexX, GOLD, Surflex, and the MMGB Flex protocols make no modifications to the original PDB coordinates, while Glide and the MMGB Rigid protocols relax steric overlaps and other unfavorable geometries of the protein in the presence of the native ligand pose found in the PDB entry. Not surprisingly, protocols which reorganize the protein around the native ligand pose perform significantly better. Table 3 summarizes and compares the results of our 4 different MMGB scoring protocols with DOCK 4.0 itself, and the literature results that have been reported for FlexX, Gold, Surflex, and Glide.

Two key findings reported in the Glide validation paper are that (1) the top scoring poses by Glide exceeded the 2 Å rmsd at a significantly lower rate than that of FlexX, Gold, and Surflex and that (2) Glide results in a lower average rmsd than the other methods. While Glide prepares the protein sites by performing a series of restrained minimizations on the entire protein−ligand complex, with a 10 kcal/mol·Å^2 positional restraint on all heavy atoms, we relax only those residues within 2 Å of the inhibitor and other charged residues which are in direct contact with the inhibitor. On the GOLD 79 set, Glide reports failure at the 2 Å level on 14 cases[22] (18%), while the closely related MMGB Rigid Min protocol fails on 12 cases (15%) and the MMGB Rigid MD/Min protocol fails on only 8 cases (10%). In terms of rmsd for the top scoring poses, an average of 1.36 Å is accomplished by Glide, 1.42 Å for the MMGB Rigid Min protocol, and 1.15 Å for the MMGB Rigid MD/Min protocol (Table 3). As Glide only minimizes the ligand in the presence of the rigid relaxed protein, it is not surprising that its results are essentially equivalent to that of the MMGB Rigid Min protocol. In contrast, the greater sampling afforded by

Table 4. Effect of Including Poses within 3 kcal/mol of the MMGB Global Minimum[a]

	MMGB Rigid						MMGB Flex					
	Min			MD/Min			Min			MD/Min		
PDB code	lowest MMGB	lowest RMSD within 3 kcal/mol	# binding modes within 3 kcal/mol	lowest MMGB	lowest RMSD within 3 kcal/mol	# binding modes within 3 kcal/mol	lowest MMGB	lowest RMSD within 3 kcal/mol	# binding modes within 3 kcal/mol	lowest MMGB	lowest RMSD within 3 kcal/mol	# binding modes within 3 kcal/mol
AVG	1.42	1.16	1.2	1.15	0.91	1.3	1.80	1.53	1.4	2.03	1.49	1.7
Median	0.93	0.80	1	0.72	0.63	1	1.38	1.14	1	1.55	1.02	1
n	79	79	79	79	79	79	79	79	79	79	79	79
n (< 2.0)	67	71	68	71	75	60	57	63	58	49	63	51
% (< 2.0)	85%	90%	86%	90%	95%	76%	72%	80%	73%	62%	80%	65%
n (< 2.5)	70	73		73	76		64	67		55	68	
% (< 2.5)	89%	92%		92%	96%		81%	85%		70%	86%	

[a] Binding modes are considered new and distinct when they are more than 2 Å rmsd away from any other pose in an existing binding mode. As greater flexibility is incorporated during the optimization, more binding modes emerge within 3 kcal/mol of the global MMGB energy minimum.

molecular dynamics leads to a marked improvement in both the failure rate at the 2 Å rmsd level and in the average rmsd of the lowest energy poses.

While comparing the single pose of the lowest MMGB energy with the native pose is unequivocally objective, when using this method in support of lead optimization, we share not only the model of the absolute global MMGB minimum but also other binding modes residing within a reasonably small yet somewhat arbitrary range above the global minimum. Among a set of bound poses, one can define unique binding modes by first calculating the pairwise heavy atom rmsds and then requiring that no two members within a single binding mode (conformational family) be within a certain arbitrary rmsd distance apart. Using a 2 Å rmsd cutoff, we find that the majority of the test cases (ranging from 65−86% as shown in Table 4) result in a single binding mode for all poses within 3 kcal/mol of the global MMGB minimum, regardless of which MMGB optimization and scoring protocol we employ. Not surprisingly, as more flexibility is added into the system, the average number of discrete binding modes within 3 kcal/mol of the global minimum increases, and the number of cases with a single binding mode decreases accordingly, as seen in Table 4, reading from left to right. Accounting for protein flexibility as well as the use of molecular dynamics during optimization both lead to greater binding mode diversity within 3 kcal/mol of the global minimum. When taking into account additional binding modes within a 3 kcal/mol cutoff on the GOLD 79 set, we find that the MMGB Rigid Min failure rate improves from 15% to 10%, while that of the MMGB Rigid MD/Min failure rate improves from 10% to 5%.

When assessing closely related analogs of a single series, for which one member has been cocrystallized, the use of the MMGB Rigid protocols may be appropriate. However, optimization of the protein around the ligand introduces bias toward the correct answer in accordance with preferences of the force field that is used for the protein relaxation. Moreover, this treatment can lead to greater difficulty when attempting to dock more distantly related analogues. Our main goal at the onset of implementing this MMGB optimization/scoring stage was to allow for protein residues to respond on-the-fly to the presence of each ligand's docked poses. While the option to rigidify the protein leads to unparalleled results in docking accuracy, the freedom to allow for dynamic protein flexibility has proven more useful in our work, albeit less accurate, when attempting to dock ligands disparate from the cognate one (data not presented). Encouragingly, when comparing the MMGB flexible protein optimization protocols to that of the other popular docking methods which use the original PDB coordinates, we find comparable accuracy. While FlexX, GOLD, and Surflex report failure rates at the 2 Å level of 35%, 24%, and 24% for the best scoring poses, respectively, the MMGB Flex Min protocol fails at a rate of 28%, and the MMGB Flex MD/Min fails at a rate of 37% (Table 3). When also considering conformers within 3 kcal/mol of the global minimum, these rates drop to 20% for both MMGB protocols (Table 4). When comparing the average rmsd for the poses with the best scores, those from the MMGB flexible protein protocols are highly similar to those from GOLD and Surflex.

While the MMGB Flex approach performs favorably on the nondruglike GOLD 79 set, Table 5 illustrates that it performs just as well on the pharmaceutically relevant PDB recurrent 14 set. Interestingly, while the Min outperformed the MD/Min optimization on the GOLD 79 set, we find little difference between the two optimization schedules on this more well-behaved test set; both fail at the 2 Å level on 4 of the 14 cases (29%), with 1owj as the single failure in common, and the average rmsds for the lowest energy MMGB poses are 1.5 Å and 1.4 Å. One aspect, that the two optimization routines show some disparity in, is the quality of conformers within 3 kcal/mol of the global MMGB minimum, where the added molecular dynamics lead to a marked improvement. While taking these additional bound poses into consideration leads to better docking accuracy in both cases, decreasing the number of failures from 4 down to two (Min) and to one (MD/Min), the average rmsd of the best conformers drops considerably more in the MD/Min protocol from 1.52 Å to 0.86 Å (the mean improves from 1.43 Å to 0.49 Å).

While we ran the two MMGB Rigid protocols on the GOLD 79 set in order to run our approach in a manner comparable to that reported by Glide, we ran only the more unbiased MMGB Flex protocols on PDB recurrent 14 set. In the same way, we routinely generate models for medicinal chemistry compounds by using the MMGB protocols on original PDB coordinates, without relaxing them around the cognate ligand.

Limitations and Sources of Failure. In order to better understand the limitations of this approach, we attempt to identify the sources of failures. As alluded to above, DOCK 4.0 failed to generate a pose within 2 Å rmsd of the native pose in 4 cases. In one (1glq) of those four DOCK 4.0 failures, MMGB Rigid optimization successfully drew a non-native starting pose to the native state and scored it with the most favorable MMGB energy. In the other three cases, all of the docked poses were beyond the radius of convergence of our optimization routines. Among the remaining 75 in the GOLD 79 set, the most successful protocol, MMGB Rigid MD/Min, failed on 5 additional cases (7%) due to unambiguous scoring problems with MMGB: 1baf, 1hri,

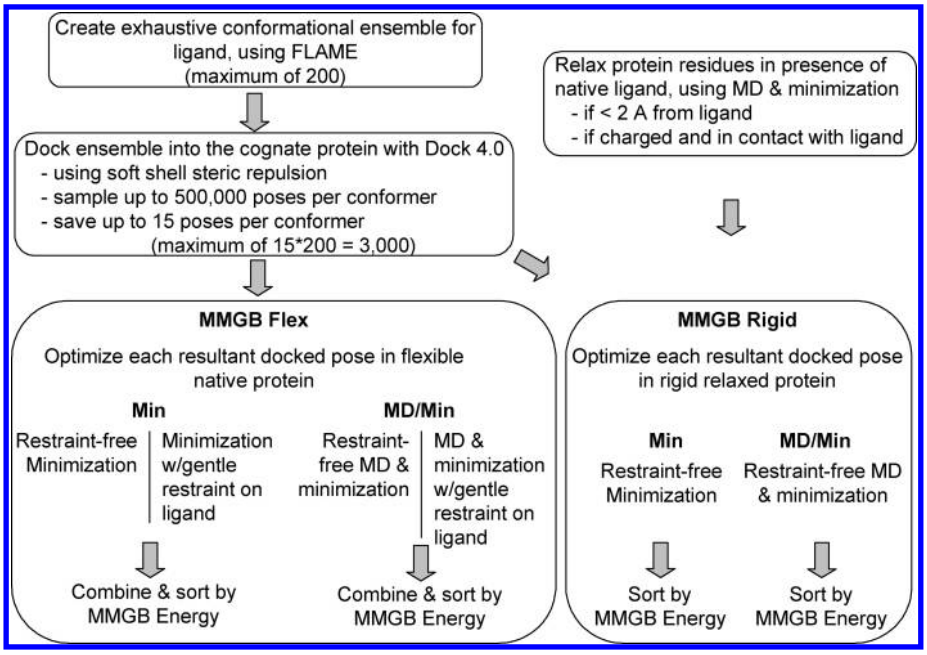

Figure 1. Process flow of the hierarchical dock/optimization/MMGB protocol. A total of 4 different MMGB optimization and scoring protocols were compared. When using a flexible protein, incoming poses for each of the two protocols were optimized both with and without restraints, and resultant poses were combined into a single protocol-specific pool and sorted according to their MMGB energies which are calulated in each of the 4 protocols on the final minimized snapshot in the absence of restraints.

Table 5. MMGB Flex Results on the PDB Recurrent 14 Set[a]

									UCSF Dock 4.0		MMGB Flex					
											Min			MD/Min		
PDB code	MW	cLogP	# rot. bonds	# H-bond don.	# H-bond acc.	# - charges	# + charges	Protein	best RMSD	best score	best RMSD	lowest MMGB	lowest RMSD within 3 kcal/mol	best RMSD	lowest MMGB	best RMSD within 3 kcal/mol of top
1c88	271	-0.8	3	2	5	2	1	Tyrosine Phosphatase 1B	0.36	0.43	0.20	0.38	0.38	0.21	0.32	0.32
1df8	244	-0.3	5	2	4	1	0	Streptavidin	0.46	2.46	0.35	0.37	0.35	0.21	0.38	0.38
1iqm	462	4.7	6	1	4	0	1	Coagulation Factor XA	0.88	9.11	0.82	1.84	1.22	0.58	2.00	1.21
1ke7	393	0.6	3	2	5	0	0	CDK2	1.71	4.91	1.70	5.24	5.24	1.54	1.98	1.54
1o2r	382	4.2	4	4	3	0	0	Beta-Trypsin	0.57	9.43	0.37	1.84	1.25	0.24	1.54	0.34
1oq5	381	4.4	4	1	3	0	0	Carbonic Anhydrase II	0.61	5.43	0.46	0.50	0.46	0.41	2.83	2.72
1owj	384	3.5	4	3	2	0	0	Urokinase-type Plasminogen Activator	1.39	1.55	1.29	2.96	1.56	0.69	3.63	1.96
1rt4	336	4.9	6	1	2	0	0	HIV-1 Rev Transcriptase	0.74	0.92	0.63	0.74	0.74	0.23	0.62	0.23
1sj0	465	6.3	6	3	5	0	1	Estrogen Receptor	1.01	2.57	0.87	2.54	0.91	0.42	1.32	0.85
1vcj	350	1.5	8	3	5	1	1	Neuraminidase	0.65	1.00	0.46	0.88	0.88	0.47	1.73	0.47
1x70	408	0.7	5	1	3	0	1	Dipeptidyl Peptidase IV	0.77	2.05	0.60	1.56	1.56	0.35	1.18	0.50
1xoq	403	3.0	7	1	4	0	0	cAMP-specific 3',5' cyclic PDE4d	0.48	1.27	0.40	0.45	0.40	0.35	0.61	0.38
1yc1	368	1.9	4	3	7	1	0	HSP90-alpha	1.38	2.56	1.16	2.45	2.45	0.67	0.71	0.71
7upj	480	6.4	6	2	5	0	0	HIV-1 Protease	1.34	3.61	1.21	1.49	1.49	0.48	2.42	0.48
	381	2.9	5.1	2.1	4.1	0.4	0.4	**AVG**	0.88	3.38	0.75	1.66	1.35	0.49	1.52	0.86
	383	3.2	5	2	4	0	0	**Median**	0.76	2.51	0.62	1.53	1.07	0.42	1.43	0.49
	14	14	14	14	14	14	14	**n**	14	14	14	14	14	14	14	14
								n (< 2.0)	14	5	14	10	12	14	10	13
								% (< 2.0)	100%	36%	100%	71%	86%	100%	71%	93%

[a] Among this smaller, pharmaceutically relevant test set, MMGB Flex protocols share only one failure in common when considering the lowest MMGB energy, 1owj. The correct binding mode is present within 3 kcal/mol of the global minimum for both protocols.

1lna, 1tni, and 3cpa, which are illustrated in Figure 3. Encouragingly, among four of these five MMGB scoring failures, at least one of the other MMGB protocols succeeded in correctly identifying the binding mode.

In 1baf, the native pose of the hapten ligand for the murine monoclonal antibody AN02 fails to rank as best with any of the MMGB optimization/scoring protocols. While the crystallographic structure stacks the hydrophobic center of the ligand's piperidin-1-ol functional group against the protein's negatively charged carboxylate oxygen atom from Asp49, the MMGB energy prefers a bound orientation in which the ligand maintains the same general shape by flipping the

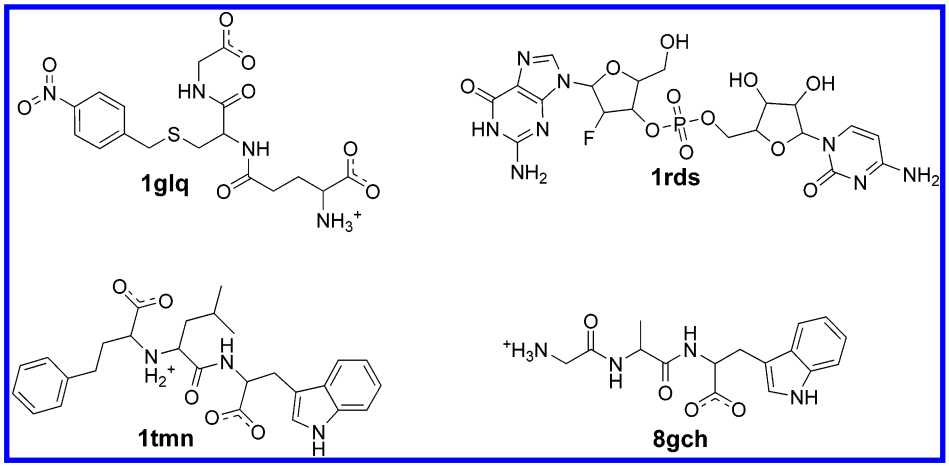

Figure 2. Structures of 4 docking failures from the GOLD 79 set. The four cases from the GOLD 79 set that DOCK 4.0 failed to generate the native pose for are all highly flexible. Conformational ensembles for 1rds and 1tmn did not run exhaustively within the arbitrary limit of 200 maximum conformers per small molecule.

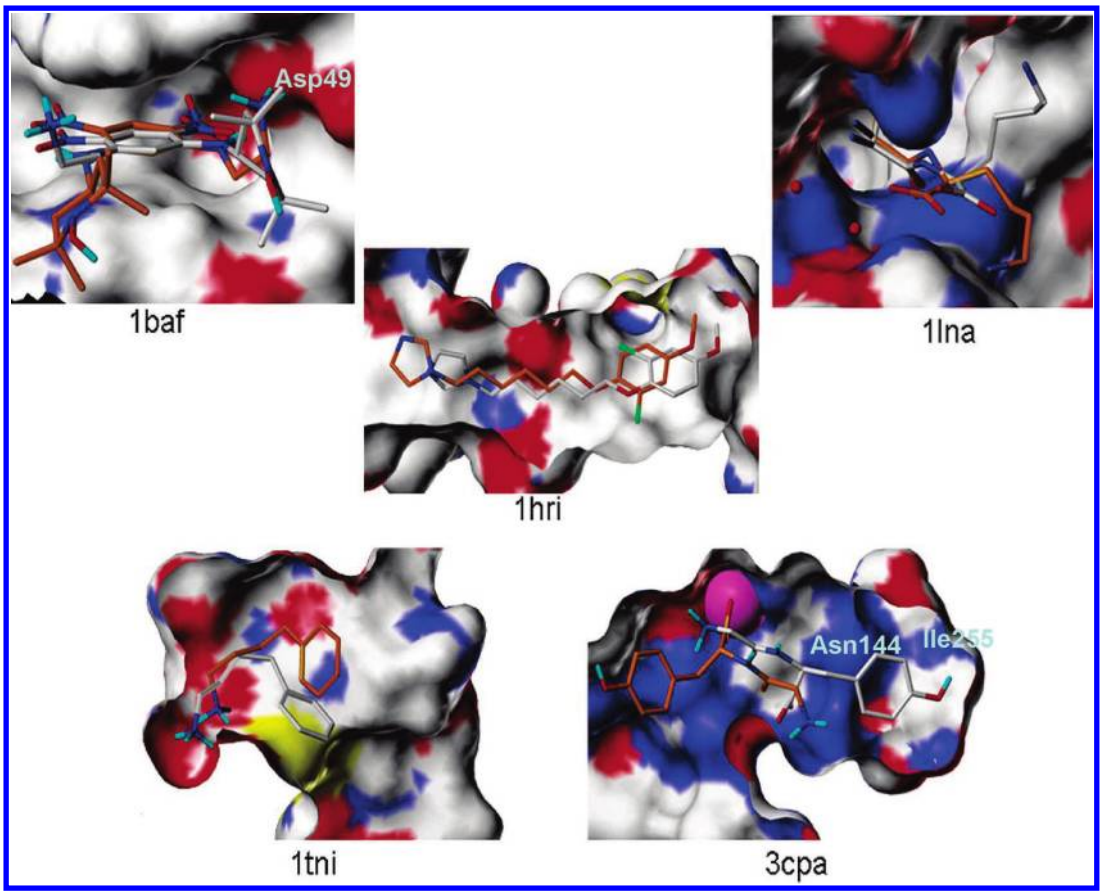

Figure 3. Five MMGB scoring failures. Solvent accessible surfaces are shown for the five MMGB scoring failures from the MMGB Rigid MD/Min protocol with both the native crystallographic binding mode of the ligand, colored in white, and for the pose at the global MMGB binding energy minimum, colored in orange. 1baf and 1lna have shallow binding pockets and open to bulk solvent, whereas with 1hri, 1tni, and 3cpa, surface cutouts are rendered in order to better illustrate the binding modes. The magenta sphere in 3cpa is the tightly bound zinc counterion.

dinitrobenzene 180° and satisfying Asp49 with an ionic interaction involving its positively charged basic ethyl amine. By analyzing the various components of the MMGB energy, we find that the crystallographic binding mode exhibits a largely positive (repulsive) electrostatic interaction energy, which agrees with the intuitive response one would have upon visual inspection of this binding mode. Important to note, however, is that the docking approach does find the crystallographic binding mode for the hapten and assigns a net negative (favorable) binding energy, despite the positive electrostatics term. When given a choice between satisfying a protein side-chain carboxylate with a charged group of opposite sign or an aromatic ring, MMGB and all physically based energy functions as well as most knowledge-based scoring functions will prefer the former, although knowledge-based methods may be able to dampen this effect

out more. Thus, it is not surprising that most docking methods had problems with this target. As we occasionally find in cases such as 1baf, some crystallographic structures seem to go against the grain of a conventional understanding. In these rare instances, physically based energy functions will invariably fail to properly rank order the native state as best unless artificial tricks are implemented.

In two of the MMGB scoring failures, 1lna and 1tni, MMGB favors poses that differ from the crystallographic structures only in the positioning of highly flexible groups directed out toward the solvent front. It is not surprising that MMGB scoring fails on cases like these, where the native ligand pose only partially occupies a wide pocket and there are few stabilizing interactions from the protein on the highly flexible ligand side chain. In a somewhat related scoring failure, MMGB prefers a pose for 1hri that is a rigid body translation of the native pose in a loose cylindrically shaped binding pocket, with a long aliphatic chain spanning the bulk of the long axis. In terms of lead optimization, the MMGB-preferred poses for these three cases are roughly correct, showing accurate geometries for the anchoring regions of the ligands in the binding sites, with binding modes that would be sufficient to generate new ideas for improving potency.

The fifth MMGB scoring failure comes from 3cpa, a highly charged system in which the native pose places a neutralized basic amine in contact with a zinc counterion, while also desolvating the amide NH2 of the Asn144 side chain and the backbone NH of Ile255 with a hydrophobic edge of the ligand's ethylphenol. MMGB prefers a pose in which the ligand's carboxylate group satisfies the positively charged zinc counterion and the polar pocket, lined by Asn144 and Ile255, is left unfilled by the ligand and thereby blocked from bulk solvent.

Computational Cost of Optimization and Scoring. With our arbitrary limits of 200 maximum conformers per ensemble and 15 poses saved per conformer, any given complex will pass on a maximum of 3000 poses to the optimization and scoring phase (Figure 1), which for a typical system consumes roughly 3−10 min of CPU-time per pose on an Intel Pentium III 1 GHz processor. Thus, on a standard MMGB Flex run, where we employ 200 CPU's from our Pentium III Linux cluster, each optimization protocol takes at most 3 h of wall-clock time per system (for both Flex Min and Flex MD/Min); simpler runs such as the MMGB Rigid ones take at most 1 h of wall-clock timer per system. While a detailed description is not within the scope of this work, we also use the flexible protein MMGB optimization and scoring protocol as a follow-up for database virtual screens, typically evaluating 50 000 to 100 000 docked poses in an overnight run on our cluster.

Discussion

Performance on GOLD 79 Set Does Not Necessarily Carry Over to a More Druglike Test Set. While the compounds in the PDB recurrent 14 set have properties consistent with those that we model on a regular basis in support of drug discovery projects, the GOLD 79 set shows little similarity. Because the energy functions in our general approach were never trained on protein−ligand complexes, they are insensitive to the disparity between the druglike PDB recurrent 14 set and the nondruglike GOLD 79 set. This can be seen in the comparable docking accuracy performance of the unbiased flexible protein among the two sets. In contrast, virtually all the scoring functions that are currently implemented in docking suites have been largely trained to perform well on test sets that show high similarity to and overlap with the GOLD 79 set. Unfortunately, there is not a pharmaceutically relevant test set that has been well benchmarked, and the transferability of high accuracy for empirical and knowledge-based scoring functions to test sets consisting of only druglike compounds remains to be seen.

Increasing Protein Flexibility during Optimization Extends Radius of Convergence, While Decreasing Accuracy of Scoring. With our implementation of the MMGB optimization, the degree of protein flexibility can be fine-tuned to appropriately take into account the nature of problem. In some instances, where medicinal chemists seek to probe the space of a deep, well-defined pocket, Rigid MMGB Min has functioned best and provided excellent agreement with experimental structure−activity relationships (data not presented). While Tables 3−5 indicate that added protein flexibility tends, in general, to decrease overall docking accuracy for redocking exercises, in general by introducing more false positives, certain instances clearly benefit from the added sampling afforded by the use of molecular dynamics during the optimization. On the druglike targets in Table 5, three of the four minimization alone failures were successful with the Flex MD/Min protocol. In many internal drug discovery projects, we have applied the MMGB Flex MD/Min protocol to a chemical series very different from one for which the cocrystal has been solved. In some instances, we predicted new binding modes, which were verified as correct by subsequent crystallographic structures and were made possible only after substantial movement of the protein, which the MMGB optimization captured. For lead optimization, we often use FLAME[15] to align new compounds against the native pose from a lead compound's cocrystal and then rely on the MMGB Flex MD/Min to tell us which compounds align well and fit without any protein movement of the existing protein crystal structure, which compounds align with poses having steric clashes with the protein that can be annealed away on-the-fly, and which compounds lead only to poses with insurmountable overlap with the protein or otherwise unfavorable MMGB binding energies. Increasing protein flexibility during the optimization phase increases the power at the expense of some accuracy.

Sources of Greater Uncertainty in the MMGB Scoring: Highly Charged Systems and Trapped Solvent Pockets. MMGB has difficulty with highly charged systems, particularly those in which the noncharged regions show loose fitting. The correct balance between desolvation of protein and ligand with that of the interaction energy is hard to achieve. Also, we find that MMGB scoring does not sufficiently penalize buried channels or pockets that are sufficiently large enough for water molecules, as is the case in 3cpa. While our discussion above of MMGB scoring failures focused only on those from the MMGB Rigid MD/

Min protocol, Tables 2 and 3 indicate that additional MMGB scoring failures arose during the MMGB Flex protocols where some alternate low-energy binding modes emerge as more favorable, as was often the case on highly charged systems and other instances, in which the lowest energy binding mode buried a water pocket. A recent advance in the Generalized Born solvation model that corrects the molecular volume[31] by using the molecular surface as opposed to the van der Waals surface may help to alleviate some of this error.

An important consequence of protein relaxation is the repositioning of amino acids in accordance with the preferences of the force field, a change that can have a profound effect on the local position of fully charged residues in a highly charged complex, which can be significant despite a small rmsd between all of the protein's relaxed and original non-hydrogen atoms of 0.3 Å or less. While the elimination of troublesome steric clashes can be required for favorable dispersion contact energies of properly docked conformers, any relaxation of the protein active site can introduce bias toward the shape of the native ligand state in accordance with the energy function used to optimize the protein, particularly for highly charged complexes.

Important Points To Consider When Incorporating MMGB into the Docking Process. While the methods described are general and should work, in principle, with any robust physically based energy function and implicit solvent model, the way that these physically based energy functions are implemented is important. The protocols we report in this paper are the result of significant trial and error. In particular, the most important consideration is keeping the vast majority of the protein fixed during the optimization. This serves to reduce the noise, particularly those regions not comprising the binding site. Otherwise, the thermal fluctuation of an entire protein introduces error bars with magnitudes that dwarf the magnitude of the binding energies, thereby making it too difficult to detect the signal among the noise. This approximation of keeping the protein largely rigid naturally does not account for differing degrees of protein distortion when comparing MMGB energies from different binding modes, thereby leading to larger error bars when comparing ligands that bind very differently to the same protein and when comparing MMGB binding energies of different proteins. Alternatively stated, the MMGB energies in this work are not useful for comparing relative binding energies between different proteins or between very different conformations of a single protein. Another important consideration is a final minimization following any molecular dynamics run to a given gradient, so that local minima are compared, instead of relying on the approximation that all systems will reach the bottom of their energy wells in a fixed number of minimization steps or relying on the more egregious approximation that minimization need not be run at all. Given that the depth of an energy well on a molecular dynamics free energy landscape is exceedingly greater than relative differences in binding energies between different poses, one would have to adequately sample a given energy well and capture a sufficient number of snapshots to make a Boltzmann-weighted average well represented. This would presumably entail molecular dynamics simulations significantly longer than the 2 ps used in this work. We instead make the approximation that most energy wells are deep and hence have Boltzmann-weighted averages that are dominated by energies at the bottom of the well, which we capture through simple minimizations. The advantage of our approximation is higher throughput and applicability to virtual screening in a reasonable time frame.

The approach in this work is not tied to DOCK for the generation of poses. Most of the popular docking packages should be capable of exhaustive runs that consistently generate the native pose in a maximum of a few CPU hours, as DOCK 4 was able to in this work, despite its reported inability to do effectively while using the default parameters[32] for nonexhaustive runs. The optimization aspect of this approach with user control over protein flexibility at a contact distance and/or residue level does add more computer time but provides features not available with other methods. While the optimization and scoring may require orders of magnitude more total CPU time, given the ubiquity and cost-effectiveness of Linux-based clusters, the added value of MMGB optimization and scoring requires less than a few hours of wall-clock time. When the goal is to predict the binding mode as accurately as possible, sacrificing a few hours on a Linux cluster as opposed to a few hours on a single machine is a small sacrifice, particularly when this allows for protein flexibility and use of a reliable physically-based scoring function.

Summary

With recent advances in the field of continuum solvent models, we explored the idea of incorporating a purely physically based MMGB energy into the docking process not only that can lead to unparalleled levels of success in redocking accuracy but also that can be run with varying degrees of protein and/or ligand flexibility during an optimization stage. This, in turn, makes for a robust hierarchical docking approach which satisfies a gamut of docking problems faced during the lead identification to lead optimization stages of drug discovery.

Supporting Information Available: Changes made to the GAFF parameters to exaggerate preferences for more suitable small molecule conformations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15* (5), 411−428.

(2) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; 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.

(3) Adekoya, O. A.; Willassen, N. P.; Sylte, I. Molecular insight into pseudolysin inhibition using the MM-PBSA and LIE methods. *J. Struct. Biol.* **2006**, *153* (2), 129−144.

(4) Coi, A.; Tonelli, M.; Ganadu, M. L.; Bianucci, A. M. Binding free energy calculations of adenosine deaminase inhibitors. *Bioorg. Med. Chem.* **2006**, *14* (8), 2636−2641.

(5) Luo, C.; Xu, L. F.; Zheng, S. X.; Luo, Z.; Jiang, X. M.; Shen, J. H.; Jiang, H. L.; Liu, X. F.; Zhou, M. D. Computational analysis of molecular basis of 1 : 1 interactions of NRG-1 beta wild-type and variants with ErbB3 and ErbB4. *Proteins, Struct., Funct., Bioinformatics* **2005**, *59* (4), 742−756.

(6) Mamolo, M. G.; Zampieri, D.; Vio, L.; Fermeglia, M.; Ferrone, M.; Pricl, S.; Scialino, G.; Banfi, E. Antimycobacterial activity of new 3-substituted 5-(pyridin-4-yl)-3H-1,3,4-oxadiazol-2-one and 2-thione derivatives. Preliminary molecular modeling investigations. *Bioorg. Med. Chem.* **2005**, *13* (11), 3797−3809.

(7) Murray, J. B.; Meroueh, S. O.; Russell, R. J. M.; Lentzen, G.; Haddad, J.; Mobashery, S. Interactions of designer antibiotics and the bacterial ribosomal aminoacyl-tRNA site. *Chem. Biol.* **2006**, *13* (2), 129−138.

(8) Nunez-Aguero, C. J.; Escobar-Llanos, C. M.; Diaz, D.; Jaime, C.; Garduno-Juarez, R. Chiral discrimination of ibuprofen isomers in beta-cyclodextrin inclusion complexes: experimental (NMR) and theoretical (MD, MM/GBSA) studies. *Tetrahedron* **2006**, *62* (17), 4162−4172.

(9) Villacanas, O.; Rubio-Martinez, J. Reducing CDK4/6−p16.-(INK4a) interface: Computational alanine scanning of a peptide bound to CDK6 protein. *Proteins, Struct., Funct., Bioinformatics* **2006**, *63* (4), 797−810.

(10) Xu, Y.; Wang, R. X. A computational analysis of the binding affinities of FKBP12 inhibitors using the MM-PB/SA method. *Proteins, Struct., Funct., Bioinformatics* **2006**, *64* (4), 1058−1068.

(11) 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* (16), 4805−4808.

(12) Huang, N.; Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. Molecular mechanics methods for predicting protein-ligand binding. *Phys. Chem. Chem. Phys* **2006**, *8* (44), 5166−5177.

(13) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. Assessing Scoring Functions for Protein-Ligand Interactions. *J. Med. Chem.* **2004**, *47* (12), 3032−3047.

(14) 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* (3), 727−748.

(15) Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46* (4), 499−511.

(16) Vieth, M.; Siegel, M. G.; Higgs, R. E.; Watson, I. A.; Robertson, D. H.; Savin, K. A.; Durst, G. L.; Hipskind, P. A. Characteristic physical properties and structural fragments of marketed oral drugs. *J. Med. Chem.* **2004**, *47* (1), 224−232.

(17) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57* (2), 225−242.

(18) Cho, S. J.; Sun, Y. X. FLAME: A program to flexibly align molecules. *J. Chem Inf. Model* **2006**, *46* (1), 298−306.

(19) *OEChem Toolkits*; OpenEye Scientific Software: Santa Fe, NM.

(20) Halgren, T. A. Merck Molecular Force Field .1. Basis, Form, Scope, Parameterization, and Performance of Mmff94. *J. Comput. Chem.* **1996**, *17* (5−6), 490−519.

(21) Albert, A.; Goldace, R. J. The strength of heterocyclic bases. *J. Chem. Soc.* **1948**, 2240−2249.

(22) Strater, N.; Lipscomb, W. N. Transition State Analogue L-Leucinephosphonic Acid Bound to Bovine Lens Leucine Aminopeptidase - X-Ray Structure at 1.65 Angstrom Resolution in a New Crystal Form. *Biochemistry* **1995**, *34* (28), 9200−9210.

(23) Xia, Z. X.; Dai, W. W.; Zhang, Y. F.; White, S. A.; Boyd, G. D.; Mathews, F. S. Determination of the Gene Sequence and the Three-Dimensional Structure at 2.4 Angstrom Resolution of Methanol Dehydrogenase from Methylophilus W3a1. *J. Mol. Biol.* **1996**, *259* (3), 480−501.

(24) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.

(25) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739−49.

(26) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Wang, J. M.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002.

(27) 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* (9), 1157−1174.

(28) Constanciel, R.; Contreras, R. Self-Consistent Field-Theory of Solvent Effects Representation by Continuum Models - Introduction of Desolvation Contribution. *Theor. Chim. Acta* **1984**, *65*, 1−11.

(29) 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.

(30) Tsui, V.; Case, D. A. Calculations of the absolute free energies of binding between RNA and metal ions using molecular dynamics simulations and continuum electrostatics. *J. Phys. Chem. B* **2001**, *105* (45), 11314−11325.

(31) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. Generalized Born Model with a Simple, Robust Molecular Volume Correction. *J. Chem. Theory Comput.* **2007**, *3* (1), 156−169.

(32) 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.; Wolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49* (20), 5912−5931.

CT6003406