Validation Studies of the Site-Directed Docking Program LibDock

Shashidhar N. Rao,[‡,||] Martha S. Head,[†] Amit Kulkarni,[‡] and Judith M. LaLonde*[,§]

GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, Pennsylvania 19426, Accelrys Inc., 9685 Scranton Road, San Diego, California 92121, and Chemistry Department, Bryn Mawr College, 101 North Merion Avenue, Bryn Mawr, Pennsylvania 19010

Received October 4, 2006

The performance of the site-features docking algorithm LibDock has been evaluated across eight GlaxoSmithKline targets as a follow-up to a broad validation study of docking and scoring software (Warren, G. L.; Andrews, W. C.; Capelli, A.; Clarke, B.; Lalonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Walls, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912−5931). Docking experiments were performed to assess both the accuracy in reproducing the binding mode of the ligand and the retrieval of active compounds in a virtual screening protocol using both the DJD (Diller, D. J.; Merz, K. M., Jr. *Proteins* **2001**, *43*, 113−124) and LigScore2 (Krammer, A. K.; Kirchoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. *J. Mol. Graphics Modell.* **2005**, *23*, 395−407) scoring functions. This study was conducted using DJD scoring, and poses were rescored using all available scoring functions in the Accelrys LigandFit module, including LigScore2. For six out of eight targets at least 30% of the ligands were docked within a root-mean-square difference (RMSD) of 2.0 Å for the crystallographic poses when the LigScore2 scoring function was used. LibDock retrieved at least 20% of active compounds in the top 10% of screened ligands for four of the eight targets in the virtual screening protocol. In both studies the LigScore2 scoring function enhanced the retrieval of crystallographic poses or active compounds in comparison with the results obtained using the DJD scoring function. The results for LibDock accuracy and ligand retrieval in virtual screening are compared to 10 other docking and scoring programs. These studies demonstrate the utility of the LigScore2 scoring function and that LibDock as a feature directed docking method performs as well as docking programs that use genetic/growing and Monte Carlo driven algorithms.

INTRODUCTION

Docking and scoring software is used widely to enhance the drug design and development processes in the pharmaceutical industry.[5−8] These programs can be divided into three major subtypes: *genetic/growing* (Flexx,[9] Gold,[10,11] MVP[12]), *Monte Carlo driven* (Autodock,[12] Flo,[14] Glide,[15,16] LigandFit[17]), and *geometric and shape complimentarity* (Dock4,[14] DockIt,[15] Fred,[16] ICM[17]) based algorithms.[22−24] One major use of these programs is in virtual screening exercises to increase the number of hits versus the number of compounds tested in high throughput biological assays. A second use is predicting the binding modes of small molecules in protein−ligand complexes in structure based drug design efforts. A number of different metrics are used to measure the success of these tools. The most commonly used measures are the compatibility of docked poses with the corresponding poses in the X-ray structures, retrieval of active compounds over inactive ones, and the correlation between experimental binding constants and computed scores.

A number of validation studies of the various docking and scoring functions have been reported in the literature.[25−34] These have generally included a variety of publicly available receptor−ligand complexes spanning a large range of binding affinities. A recent report compares several docking and scoring methods on 50 protein−ligand complexes from three systems of pharmaceutical relevance[27] as well as including 100 receptor ligand complexes extracted from the Protein Data Bank (PDB).[35] A related approach has been used in a large docking−scoring evaluation of eight proprietary GlaxoSmithKline (GSK) targets using 11 commercially available software programs.[1] As described previously, these protein targets were chosen for their pharmaceutical relevance, diversity in binding sites and functions, and the number of protein−ligand crystal structures available. The ligand set is distinctive because it contains for each protein target up to five compound chemical classes and at least two congeneric series with binding affinities spanning 3 orders of magnitude. Evaluations using this data set reported thus far have yielded useful information on the applicability and conditions required for using docking−scoring tools successfully.

As a follow-up to the original study, the site-feature docking algorithm LibDock[2] has been applied to the GSK validation data set. LibDock is based on the algorithm developed by Diller and Merz[2] and is one of the few commercially available docking programs[35] that uses protein binding site features to guide docking.[2,37] The LibDock methodology was originally developed to handle the rapid

* Corresponding author e-mail: jlalonde@byrnmawr.edu.
[†] GlaxoSmithKline.
[‡] Accelrys Inc.
[§] Bryn Mawr College.
[||] Current address: Schrodinger Inc., 120 W. 45th St., New York, NY 10036.

10.1021/ci6004299 CCC: $37.00
Published on Web 11/07/2007

Table 1. Enzyme Targets Used in the LibDock Docking of the 1303 Compound Database

protein target (class) (abbrev)	no. of X-ray complexes	no. of structural classes	minimum affinity[a]	maximum affinity[a]	LibDock box ($X-Y-Z$ Å)
Chk1 kinase (CHK1)	15	3	5.00	8.16	$36 \times 26 \times 30$
HCV polymerase (HCVP)	13	2	4.22	8.26	$40 \times 40 \times 32$
Met tRNA synthetase (MRS)	31	4	4.91	9.00	$40 \times 40 \times 40$
PPARδ nuclear hormone receptor (PPARδ)	53	9	4.30	9.48	$32 \times 28 \times 32$
Factor XA serine protease (FACTOR Xa)	10	8	5.32	9.96	$40 \times 40 \times 30$
Gyrase B isomerase (GYRASE B)	7	2	4.37	8.40	$38 \times 38 \times 38$
E. coli PDF metalloprotease (*e*PDF)	6	3	4.00	9.00	$24 \times 28 \times 24$
Strep PDF metalloprotease (*s*PDF)	2	2	4.00	9.00	$28 \times 28 \times 28$

[a] Expressed as log(binding constant in M).

docking of combinatorial libraries of compounds with the goal of prioritizing the selection of libraries rather than rank ordering the compounds themselves.[2] The algorithm has four functional aspects: conformation generation of the ligands, creating a binding site image (hot spot identification), matching the binding site image and the ligand, and a final optimization stage and scoring. The binding site image consists of lists of polar and apolar hot spots. These are generated by laying a grid in the binding site volume and then scoring an apolar and polar probe at each grid point. The apolar and polar hot spots are clustered to generate their respective hot spot maps, which are then used to match the ligand apolar and polar features to the binding site. The optimization step includes pruning the list of matches for steric clashes, ranking using an atom pairwise score, and clustering. In the final stage of refinement hydrogen bonding and steric potentials are used with the BFGS optimization algorithm.[38] The scoring algorithm originally developed by Diller et al.[2] is named "DJD score" in the LibDock implementation. The scoring function resembles a piecewise linear potential (PLP)[45] summed over interacting atoms in the protein−ligand complex. The atoms are divided into four atom types—apolar, acceptor, donor, and donor/acceptor—and the score between interacting atoms is scored using either a hydrogen binding potential or a steric potential.[2]

Diller et al.[2] report that LibDock reproduced 75% of the docked protein−ligand complexes within a 2.0 Å root-mean-square difference (RMSD) of the known crystallographically determined positions for any docked pose within the ensemble of docked ligand conformations.[2] However, when considering only the top ranked poses, just 46% fall within 2.0 Å RMSD of the known crystal structures. In analyzing LibDock's performance in a combinatorial library docking exercise of 13 020 compounds approximately one-third of the top ranked poses docked within 1.5 Å of the crystallographically determined compound. The enrichment factor was 3.3 for retrieving active compounds. This level of success is noteworthy because retrieving actives from a combinatorial library of similar compounds would appear to be more challenging than the generalized virtual screening study of a diverse compound collection.

The recently reported GSK docking−scoring study[1] included two programs, MVP[12] and Fred,[20] which used binding site information in their docking−scoring algorithms. The first program, MVP, used atom-type target points to superimpose ligands in the binding site and performed well in both docking accuracy and retrieval of actives in a virtual screening. Similarly, including a pharmacophore constraint in the Fred docking algorithm for the metal binding site of *E. coli* polypeptide deformylase (PDF) and *Streptococcus pneumoniae* (pneumococcus) PDF improved docking performance. LibDock uses the physicochemical properties of the ligands to guide docking to corresponding features in the protein binding sites by matching a triplet of ligand atoms to a triplet of protein hot spots. Unlike most docking and scoring algorithms, the initial docking steps in LibDock are not driven by a molecular mechanics force field score. Instead, LibDock is based on matching polar and apolar binding site features of the protein−ligand complex. Hence, the performance of an algorithm such as LibDock that is solely driven by matching features would be of interest. Given the uniqueness of LibDock's underlying algorithm and its reported success, a full validation study was warranted to gauge its performance against other classes of docking−scoring algorithms. Here, we report LibDock's performance across the GSK validation data set and compare it to the performance of nine other docking and scoring algorithms.

MATERIALS AND METHODS

Protein and Ligand Preparation. Docking studies on a collection of 1303 ligands[1] were carried out against the eight targets listed in Table 1. The choice of these targets has been described previously.[1] The target proteins are abbreviated as follows: Chk1 kinase, CHK1; HCV polymerase, HCVP; methionyl tRNA synthetase, MRS; PPARδ nuclear hormone receptor, PPARδ; serine protease factor Xa, FACTOR Xa; gyrase B isomerase, GYRASE B; *E. coli* and *Strep* PDF metalloproteases, *e*PDF and *s*PDF, respectively. The proteins were prepared in the following steps: All hydrogens were added using the program "REDUCE"[39] consistent with the physiological pH (7.0). Thus, arginine and lysine side chains were protonated, while the glutamate and asparate side chains were deprotonated. The histidine side chains were selectively protonated depending on their environment in the targets. Flipping of Asn, Gln, and His side chains were allowed to promote appropriate intramolecular interactions in the protein environment. The protein atoms were then typed using the CFF force field.[40] In the case of *e*PDF and *s*PDF, the binding site nickel ion was not bonded to the amino acid side chains and was typed as a Zn^{2+}. The hydrogen positions were energy minimized using either the CFF or the Dreiding force field[41] in the absence of any ligands, while holding the heavy atoms of the protein backbone and side chains fixed.

The database of 1303 ligands has been described previously.[1] Briefly, the compounds in the data sets were synthesized in support of active GSK targets and show

biological activity in in vitro assays. No compounds were added as decoys from publicly available compound databases. For each protein target/compound set, a number of crystallographically determined protein/ligand structures, ranging from 6 to 54, were chosen. For each target set there were 150−200 compounds chosen ranging over 4 orders of magnitude in ligand affinity and containing at least one representative crystal structure for each representative class.

The starting SD file used to build the CATALYST[42] database of ligand conformations was constructed from SMILES strings of the ligands Therefore, the ligand database did not include conformations found in the X-ray crystallographic complexes of any of the molecules. Thus, any agreement between the docked conformation and the X-ray data is due to the conditions and constraints of docking. The ligand database contained cognate ligands for all eight targets and was prepared by adding and deleting hydrogen atoms, consistent with the ionization states of nitrogen and oxygen atoms at physiological pH. Ligand conformations were randomly generated starting from their SMILES strings and energy minimized using the CFF force field[40] and stored in an SD file. The latter was used to generate a database using the CATALYST[42] BEST conformational generation paradigm. The CATALYST BEST method uses a poling algorithm[49] to generate a diverse set of low-energy conformations. We set the maximum number of conformations generated to 300 (MaxConfs = 300) with a 20 kcal/mol energy window to generate approximately 50 lowest energy conformers per ligand. The maximum number of conformations was chosen to ensure adequate coverage of conformational space.

Docking Procedure. The site-features directed docking (LibDock) studies were carried out on the eight targets using the same procedure. Binding site models that were used in the LigandFit portion of the GSK docking−scoring study[1] were used as starting points for LibDock docking studies. Using LigandFit, the initial protein-shape-based binding sites were identified using the "site finding" module with a value of 7 Å for the site opening. These sites were edited manually with selective additions and deletions in the Cerius2[36] interface to remove artifacts and to ensure that site points were present around key active site residues. The centroid of the site model was used to define an active site box required by LibDock. Typically, the box dimensions were between 24 and 40 Å in the *X*, *Y*, and *Z* directions (Table 1). Unless stated otherwise, hydrophilic and lipophilic hot spots were generated with a density of 250.

Analysis of Docking Poses. All the docked poses were ranked by the DJD score,[2] and the top 25 were retained in the output SD file. The output SD file from a LibDock run does not contain any hydrogens, including those on heteroatoms. Therefore, docked poses were processed for further minimization through addition of hydrogen atoms, taking into account the ionizable atoms in the starting structures. The hydrogen-filled docked poses were minimized in the protein targets using the CFF force field.[40] In the case of enzymes *e*PDF and *s*PDF, the Dreiding force field[41] was employed as the CFF lacks suitable parameters for highly coordinated metal ions.

In the cases of ligands that have X-ray data, the docked and minimized poses were compared with the X-ray poses using an SVL script in the MOE interface.[43] This script computes a symmetry-corrected root-mean-square distance (RMSD) for heavy atoms in the docked and crystal conformations. For each compound involved in an X-ray complex, the following parameters were computed: (a) the lowest value of RMSD in the ensemble of 25 saved poses abbreviated as the "RMSD for any pose"; (b) the RMSD of the first DJD pose (RMSD DJD); (c) the RMSD of the first LigScore2 pose (RMSD LigScore2). In some cases the mean of these metrics across a protein target has also been calculated. The Tanimoto coefficient[44] for volume overlap was calculated between the lowest RMSD pose in the ensemble or with the first DJD scored pose with the ligand's crystal structure. For each of the enzyme targets, the percentages of compounds docked with RMSD values of less than 0.5, 1.0, 2.0, 3.0, and 4.0 Å have been listed under varying pose scoring constraints (none, first poses based on DJD,[2] LigScore2,[3] LigScore1,[3] PLP,[45] PMF,[46] and JAIN[47] scores).

Analysis of Retrieval of Active Ligands. In addition to the investigations on the consistency of docked poses with X-ray structures, we have analyzed the retrieval of active compounds on the basis of DJD score and LigScore2 ranks. For these analyses, the compounds were divided into two classes of activity, active and inactive. The latter class corresponds to binding constants of $>1.0\ \mu$M and the former corresponds to $\leq 1.0\ \mu$M.

The docked and minimized compounds were ranked separately by their DJD and LigScore2 scores. The enrichment factors corresponding to the entire database and for each class-specific database were calculated as defined by Pearlman and Charifson.[48] The percent actives retrieved for the percentage of compounds screened were computed. The percentages of active compounds retrieved in the top 10% of the compound database and the percent of the database that was examined to retrieve 90% of the compounds were calculated as indicated by

percent yield (%Y); percentage of the known actives in the hit list:

$$\%\mathrm{Y} = \frac{H_\mathrm{a}}{H_\mathrm{t}} \times 100$$

percent of actives (%A); percentage of known active compounds retrieved from the database:

$$\%\mathrm{A} = \frac{H_\mathrm{a}}{A} \times 100$$

RESULTS AND DISCUSSION

Information pertaining to the protein−ligand complexes is presented in Table 1. Summarized within the table are (a) the number of X-ray structures of the protein−ligand complexes, (b) the number of structural classes of ligands, (c) the range of binding affinities expressed as log(binding constant in M), and (d) the box dimensions in the *X*, *Y*, and *Z* directions used in the LibDock binding site definition. The performance of LibDock has been assessed for the retrieval of poses resembling those observed in the crystal structures as well as the retrieval of active compounds in a virtual high throughput screen. These topics are discussed separately in the sections below.

Table 2. Mean RMSDs and Tanimoto Coefficients for the Eight Protein Targets[a]

protein target	no. of X-ray complexes	mean RMSD for any pose	mean RMSD LigScore2	mean RMSD DJD	mean Tanimoto for first DJD pose
CHK1	15	1.48	3.45	5.79	0.366
HCVP	13	7.59	13.2	14.4	N.D.
MRS	31	1.81	4.46	8.19	0.37
PPARδ	53	4.19	6.89	6.43	0.43
FACTOR Xa	10	1.94	3.1	5.76	0.38
GYRASE B	7	2.76	4.58	6.84	0.43
*e*PDF/*s*PDF[b]	8	2.34	3.32	4.03	0.45

[a] Abbreviations: RMSD for any pose, the lowest value of RMSD from any of the poses in the ensemble of 25 saved docked conformations; RMSD DJD, the RMSDof the first DJD ranked pose; N.D., not determined. The Tanimoto coefficient[39] for volume overlap was calculated between the first rank DJD score pose in the ensemble and the ligand's crystal structure. [b] The values for the ligands of *e*PDF and *s*PDF are reported together because of the small number of structures available and the similarity of the bound ligands.

Table 3. LibDock Poses and Scoring Functions: Percent Retrieved

target	RMSD (Å)	none[a] (%)	first DJD (%)	first LigScore1 (%)	first LigScore2 (%)	first JAIN (%)	first PMF (%)	first PLP1 (%)
CHK1	≤2.0	73	7	33	60	33	20	13
HCVP[b]	≤2.0	9	0	N.D.	9	N.D.	N.D.	N.D.
MRS	≤2.0	65	19	29	42	19	6	23
PPARδ[c]	≤2.0	55	36	38	40	30	49	36
FACTOR Xa	≤2.0	50	0	10	30	10	0	0
GYRASE B	≤2.0	43	14	0	0	0	0	0
*e*PDF/*s*PDF	≤2.0	50	13	13	13	13	13	0

[a] No constraints of any score, meaning the percent of ligands that had a pose in the docked ensemble with an RMSD to the crystal structure of ≤2.0 Å. [b] LibDock poses for the HCV polymerase were not rescored because an accurate pose was not obtained for any of its ligands. [c] For PPARδ the set of compounds (26 of 53) with RMSD > 8.0 Å were not included in the recompiling of rescoring statistics.

Accuracy in Reproducing Crystal Structure Conformations. The calculated RMSD between the docked pose and its corresponding crystal structure assesses the accuracy of LibDock in reproducing the small molecule conformation observed in the protein−ligand complex. The RMSD gives a good measure of the match between their atomic positions on an atom by atom basis. A docking algorithm reproduces a crystallographic conformation if its calculated RMSD is below 2.0 Å. A second measure used to assess the positions of the ligands is the Tanimoto volume overlap between the docked conformation and crystallographically observed ligand positions. The RMSD indicates whether the algorithm was able to match atom position and by inference the atomic properties of the ligand deemed important in protein−ligand recognition, while the Tanimoto overlap reveals how close the algorithm reproduced the general fit in the active site with regard to shape and location of the docked pose. The RMSDs and Tanimoto coefficients were calculated a number of different ways corresponding to a particular pose rank and the scoring function used for the ligand. The RMSDs used in this analysis are RMSD for any pose, RMSD DJD, and RMSD LigScore2 as described under Materials and Methods. The mean RMSD and Tanimoto values for the eight protein−ligand systems are reported in Table 2. The values for individual compounds of a target set are reported in separate tables in the Supporting Information. The overall performance of LibDock as indicated from the mean RMSD for any pose is quite reasonable with six out of eight targets yielding mean RMSDs below 3.0 Å. Based both on Tanimoto value and the mean RMSD for any pose, LibDock produced the most accurately docked poses for the CHK1 target. The poor mean RMSD for any pose (7.59 Å) of HCVP suggests that on the whole the LibDock algorithm failed for this target (see below).

Docking Accuracy and Rank Score. A compound is considered to be docked accurately if the RMSD with its crystal structure is ≤2.0 Å. The actual percentages of accurate poses retrieved for the protein−ligand systems at various RMSD cutoffs are plotted in Figure 1. The data were compiled using the RMSD for any pose (Figure 1A), RMSD DJD (Figure 1B), and RMSD LigScore2 (Figure 1C). For all protein targets, with the exception of HCVP, 40% of the crystallographic ligand conformations are reproduced within the ensemble of 25 docked conformations with almost 80% of those recovered for the CHK1, Figure 1A. The utility of a docking−scoring algorithm, however, is in the ability to accucarately produce binding modes among the highest-scoring conformations produced. The accuracies of the first ranked pose when using either DJD or LigScore2 functions are shown in Figure 1B and Figure 1C, respectively. Recovery of crystallographic poses from the highest DJD scored ranked poses was unsuccessful (Figure 1B). Less than 20% of the compounds are docked with a RMSD below 2.0 Å for all but PPARδ. In the absence of a crystal structure defining the binding mode, the performance of the scoring function is critical.

The deficiencies of scoring functions are a generalized phenomenon, with some scoring functions producing better results for some targets than for other targets. In the case of LibDock, for most of the targets a true binding mode can be identified within the group of 25 docked poses. The question remains as to whether any other scoring function would have been more successful in locating the true binding mode as a top ranked pose. Within the Cerius2 module there are several

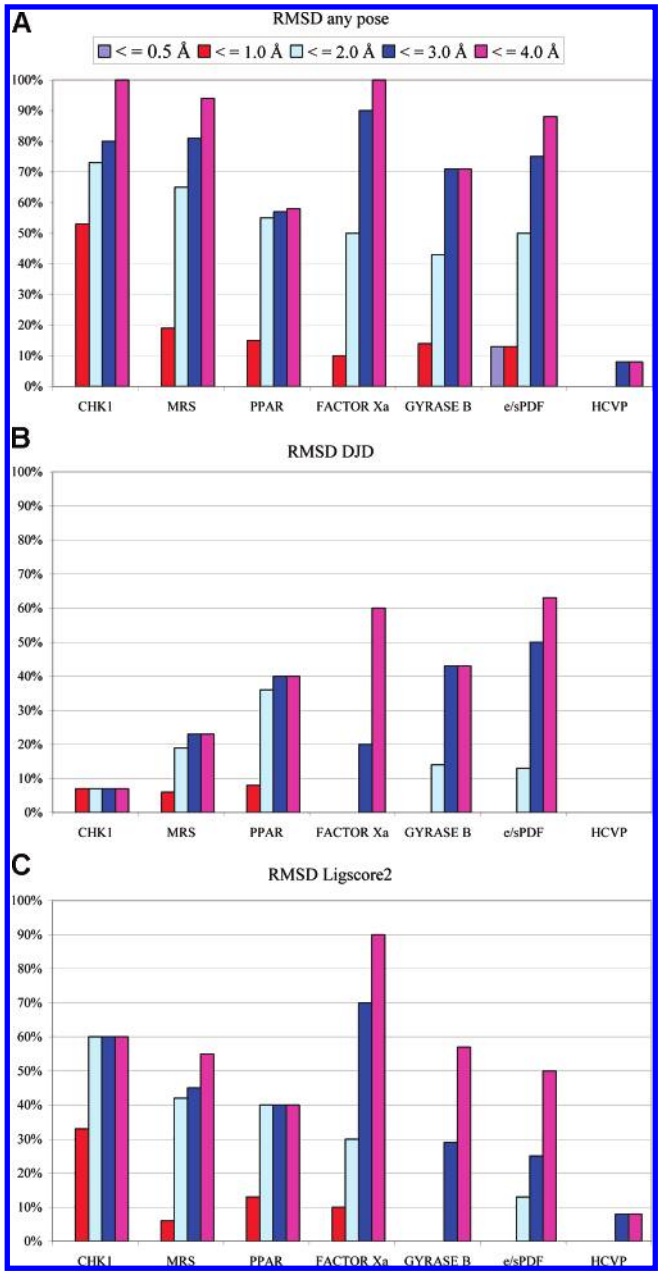

Figure 1. (A) Recovery of X-ray poses from all the LibDock docked poses where the RMSDs are tabulated. (B) Recovery of X-ray poses from the highest DJD score ranked poses. (C) Recovery of X-ray poses from the top-LigScore2 ranked LibDock docked poses.

scoring functions for which the DJD docked poses can be rescored. These are LigScore2,[3] LigScore1,[3] PLP,[45] PMF,[46] and JAIN.[47] In Table 3, the percent of ligands retrieved as a top rank score with a RMSD below 2.0 Å are shown for each of the eight targets and the various scoring functions. Except for GYRASE B (Table 3), LigScore2 is more successful in retrieving an accurate pose as the top scoring pose in the LibDock docking ensemble than using the top ranked DJD pose (Figure 1C vs Figure 1B) or any other score. A vast improvement is observed with LigScore2 scoring (Figure 1C) for CHK1, MRS, PPARδ, and FACTOR Xa targets. The biggest gain in using a top ranked LigScore2 over DJD pose is for CHK1, where retrieval increased from a meager 7% for top ranked DJD scored poses to 60% for top LigScore2 poses. This trend of improved retrieval for LigScore2 is also observed for MRS, PPARδ, and to a lesser extent FACTOR Xa. The Jain and PLP1 scoring functions do not enhance the retrieval rate over LigScore2 for any of the protein systems. Interestingly, the PMF score retrieves 49% of the PPARδ ligands in comparison to 40% for LigScore2 and 36% for the top DJD score. A plausible explanation for improved retrieval is that the PMF score may more accurately capture the protein−ligand interactions of polar functional ligand groups within the mostly hydrophobic nuclear hormone receptor site. The accurate retrieval of metalloprotease (*e*PDF/*s*PDF) ligands is poor for the top ranked poses of all scoring functions.

Results from Individual Targets. The protein systems of the GSK validation study were chosen to represent a wide variety of relevant pharmaceutical targets for which there are active drug design programs. Published validation studies have generally reported good results for many publicly available protein−ligand complexes. Some programs have been reported to be more suitable for some protein systems than other proteins.[1] Understanding the performance of LibDock for a particular protein system may provide information on how best to use the algorithm in the future.

The retrieval of metalloprotease (*e*PDF/*s*PDF) ligands is poor: only 13% are accurately docked in the top ranked poses of both the DJD and LigScore2 scoring functions. The performance tables of the individual crystallographic ligands *e*PDF and *s*PDF are provided in the Supporting Information. No difference in performance over the compound classes for *e*PDF/*s*PDF ligands was observed. However, 40% were docked accurately irrespective of scoring (RMSD for any pose). The failure of both scoring functions can be attributed to the presence of the metal−ligand interactions in the active site. The Dreiding force field[41] was used for the PDF targets because the CFF[40] did not contain appropriate metal parameters. The poor results obtained here suggest that even the Drieding force field may not adequately handle the metal−ligand interactions.

The poorest performing target is HCVP, for which only two of the 13 compounds have docked poses with RMSD for any pose values of <4 Å relative to the corresponding X-ray data. The remaining 11 compounds have RMSD for any pose values of >4.5 Å. Docking experiments conducted with a hot spot density of 500 did not yield significantly different results with respect to the root-mean-square deviations from the X-ray poses. Inspection of the lowest RMSD pose in the ensemble indicates that the majority of the compounds are actually docked entirely within the large polymerase-binding site. The HCVP protein−ligand crystal structure complexes indicate that the ligands protrude from the binding site and are partially solvent exposed. Furthermore, the X-ray structures indicate that important hydrogen bonds exist between the one class of compounds and a water molecule. For the second class of compounds, two portions of the molecule are positioned in the correct binding pocket; however, there is an inversion in comparison to the crystal structure of a polar group from facing solvent to being buried in the best docked posed.

The failure in reproducing the crystallographic binding modes for the HCVP target cannot be explained by the size of the active site. If this were the case, then docked conformations would show a more varied distribution of positions rather than having only some portions of the ligands

incorrectly docked. A possible explanation may be that a number of crystallographic water molecules found in the binding site were not included in the target protein during docking. A second explanation indicated by the crystal structures is that the compounds are actually solvent exposed. Binding modes in a solvent-exposed region of a binding may not be adequately recognized in the LibDock algorithm because it maps apolar and polar interaction hot spots at the protein−ligand interface. Furthermore, LibDock docks the ligands to the active site via these hot spots. Therefore, the resulting docked poses would select orientations and conformations near amino acid residues rather than interacting with solvent. This explains why many of the incorrectly docked poses for this target have portions of the compound buried within the site instead of having solvent interactions as revealed in the crystal structures.

In the case of GYRASE B, no accurate poses were identified by LigScore2 ranking while a slight improvement with DJD scoring identified one of seven compounds with RMSD ≤ 2.0 Å. The slight improvement with DJD scoring over LigScore2 for GYRASE B was the exception among the eight targets. For FACTOR Xa LibDock recovers approximately 30% of the X-ray conformations within 2.0 Å of the crystal structure. No explanation of docking failures is apparent among the compound classes for this serine protease.

For MRS and PPARδ a moderate improvement is observed using LigScore2. For MRS, an accurate pose is obtained for the crystallographic ligands in the ensemble of 25 docked poses for 21 of 30 compounds. However, only 13 of the 31 ligands are identified by subsequent LigScore2 rescoring. Inspection of the docked poses for the individual compounds indicates that the top scoring pose is not simply an inversion of the compound in its binding site but rather is a docked conformation that is outside the binding site and solvent exposed. In contrast to this top scoring pose docked outside the binding site, lower ranking poses are simply inverted in the binding site. For this target and some of its ligands, a docked conformation that is outside the binding site is scored higher than a docked conformation that occupies the binding site and is simply inverted from the true binding mode. This can be interpreted as a scoring problem rather than a docking problem which even the better performing LigScore2 scoring function does not improve.

The PPARδ protein target has a set of compounds (26 of 53) that are docked inaccurately. The correct binding mode for these compounds is not even predicted within the 25 saved poses (mean RMSD for any pose of 8.0 Å). This outlier set consists of larger compounds, all of which are from the same structural class. This is also reflected in the Tanimoto coefficients for the lowest RMSD pose, which fall within the range 0.75−0.4 Å (Table 4). LibDock is able to position some of these compounds correctly in the active site, (compound numbers 27−30, 34, 35, 38, 39, and 43). However, there is an inversion of two similar aromatic rings about the center of the compound. With these compounds we observe a docking failure for large compounds within the enclosed hydrophobic binding site. An interpretation of these failures is that the scoring function cannot distinguish these slightly differing, hydrophobic groups in the context of a hydrophobic binding site. For the remaining outlier compounds with a Tanimoto below 0.4, the algorithm is not able to produce the correct pose within the closed active site and positions the compound partially exposed to solvent. For the PPARδ ligands, therefore, LibDock success is limited to compound class.

Table 4. PPARδ: Nuclear Hormone Receptor, RMSD, and Tanimoto Values

compd no.	RMSD any pose	RMSD LigScore2	RMSD DJD	Tanimoto for the pose with the lowest RMSD
1	0.67	2.50	0.67	0.88
2	0.68	13.51	0.98	0.89
3	0.70	9.36	0.70	0.88
4	0.75	3.43	10.12	0.79
5	0.86	3.20	1.18	0.78
6	0.96	3.57	1.83	0.78
7	0.99	2.50	0.99	0.73
8	0.99	8.00	1.54	0.83
9	1.05	2.09	1.61	0.74
10	1.09	2.01	1.11	0.77
11	1.10	6.50	8.85	0.77
12	1.11	3.33	1.91	0.80
13	1.13	7.13	1.13	0.75
14	1.13	7.41	1.79	0.77
15	1.13	7.99	1.20	0.70
16	1.17	7.14	1.17	0.72
17	1.17	3.38	1.73	0.73
18	1.25	7.03	1.25	0.67
19	1.29	3.52	1.33	0.68
20	1.39	7.65	10.15	0.68
21	1.42	12.91	1.75	0.69
22	1.47	11.06	11.29	0.63
23	1.54	14.82	11.62	0.64
24	1.66	8.07	9.72	0.72
25	1.69	8.37	1.76	0.75
26	1.72	9.52	2.24	0.52
27	1.79	10.20	9.39	0.56
28	1.85	8.67	10.20	0.66
29	2.07	10.52	2.01	0.56
30	2.69	2.69	10.78	0.45
31	3.86	14.56	8.51	0.37
32	4.43	13.28	4.43	0.28
33	6.48	6.84	8.09	0.18
34	6.57	9.46	7.34	0.74
35	6.66	13.50	6.66	0.68
36	6.79	14.26	6.79	0.24
37	7.33	12.87	9.36	0.21
38	7.45	7.75	7.68	0.52
39	7.68	10.55	9.87	0.53
40	7.74	10.56	9.81	0.22
41	7.92	9.86	10.09	0.47
42	8.02	8.45	8.02	0.45
43	8.17	10.72	9.82	0.54
44	8.21	12.10	10.29	0.24
45	8.32	10.32	8.66	0.56
46	8.34	9.07	8.74	0.06
47	8.57	9.34	9.25	0.28
48	8.79	13.09	8.79	0.25
49	9.39	9.39	14.81	0.07
50	9.86	13.82	10.77	0.18
51	10.47	10.47	11.12	0.13
52	10.78	13.19	14.10	0.13
53	11.52	12.42	14.36	0.05

The best results were produced with CHK1, where 67% of the compounds were recovered with positions within 2.0 Å of the crystal structures. The individual compound RMSDs for any pose are shown in Table 5. Nine of the 15 compounds have an RMSD LigScore2 of 1.5 Å or less. A second set of molecules (10−15) displays much poorer RMSD LigScore2 values (between 4.4 and 9.5 Å). The poorly docked compounds are from several chemical classes. The docked ensemble does contain a pose within the vicinity of the

Table 5. RMSD LigScore2 values and Tanimoto Coefficients[a]

compd no.	pAffinity	LigScore2 pK_i	RMSD LigScore2 (Å)	Tanimoto coeff	rank of pose with the lowest RMSD
1	7.08	6.49	0.54	0.92	1
2	7.96	6.73	0.56	0.92	1
3	7.74	6.82	0.59	0.92	1
4	8.15	6.89	0.84	0.84	1
5	5.71	5.75	0.92	0.81	8
6	7.54	7.82	1.14	0.82	4
7	6.79	6.12	1.22	0.78	6
8	7.72	6.81	1.32	0.90	10
9	7.92	6.95	1.33	0.69	1
10	7.05	7.01	4.43	0.62	5
11	5.55	5.53	6.48	0.57	17
12	6.4	6.33	6.71	0.19	20
13	6.71	5.5	7.82	0.12	16
14	6.79	7.04	8.88	0.80	8
15	5.6	5.96	9.51	0.19	13

[a] The RMSD LigScore2 values and Tanimoto coefficients for 15 ligands of CHK1 (columns 4 and 5) together with their affinities and score (columns 2 and 3). Also listed is the rank of the pose with the lowest RMSD among the ensemble of 25 docked conformations (column 6).

crystal structure (less than 4.0 Å). However, neither the DJD nor LigScore2 scoring can identify this pose within the top rank of scored poses. Four of these six poorly docked compounds (11−13 and 15, RMSD LigScore2 > 6.5 Å) are from the same thiophene class of compounds. When considering compound properties for this class of compounds, the poorest docked compound also has the lowest ratio of hydrophobic accessible surface area to polar accessible surface area (data not shown). It appears that the more hydrophobic compounds are docked more accurately for the CHK1 target.

Aside from the improvement for the nuclear hormone receptor when using PMF scoring, LigScore2 enhances the overall performance of LibDock. LigScore2 is an empirically trained scoring function used to predict binding affinities.[3] As described by Krammer et al.,[3] it is based on a QSAR equation derived from approximately 120 protein−ligand complexes using intermolecular descriptors. The statistical relevant QSAR model is represented by

$$pK_i = (1/4)C - \beta_1 \langle Evdw \rangle + \beta_2 Cpos\_tot - \beta_3 (SolvPlty\_lig^2 + SolvPlty\_prot^2)$$

The LigScore2 scoring function contains three terms, a softened Lennard-Jones potential (<Evdw>), a count of the buried polar surface area between the complex involving attractive protein−ligand interactions (β_2Cpos_tot), and the square of the protein−ligand surface descriptors (SolvPlty_lig^2 + SolvPlty_prot^2). It is likely that the two terms representing buried polar interactions improve the rank ordering of accurately docked poses. In the absence of a known crystal structure, we recommend that LibDock docked poses be rescored using LigScore2 and that for selected targets with very hydrophobic binding sites PMF score be used as well.

Comparison of LibDock Accuracy. An improvement of docking accuracy was obtained using the LigScore2 scoring functions for most of the protein targets. As noted by Warren et al.,[1] docking algorithms adequately sample conformational space and are able to reproduce ligand binding modes observed in crystal structures of protein−ligand complexes. However, the scoring functions were less able to distinguish the crystallographic binding mode among the ranks of the docked poses. For the current user of docking and scoring technology, the question remains on how to best use the various docking programs and scoring functions. Practically speaking, if a crystal structure exists of a similar ligand bound to the protein target of interest, then one can judge docked poses by how well the docked pose reiterates what is seen in the crystal structure. However, docking and scoring experiments are often used prior to crystallographic knowledge to generate a hypothesis for binding mode and protein−ligand interactions in order to guide synthesis. Knowing which docking and scoring protocol has the highest chance of producing accurately docked poses for a particular protein−ligand system would be beneficial to most users of docking and scoring algorithms. The percentage of accurately docked poses (RMSD < 2.0 Å) for the first pose obtained in LigScore2 is plotted relative to 10 other docking and scoring algorithms (Dock4,[18] Dockit,[19] Fred,[16,17] Flexx,[9] Flo,[14] Glide,[15,16] Gold,[11] LigandFit,[17] MOE,[43] and MVP[12]) for each of the proteins; see Figure 2. LibDock's performance is on par with the better performing docking and scoring programs for CHK1, FACTOR Xa, HCVP, MRS, and PPARδ. In comparison, for GYRASE B LibDock's performance is mediocre with compounds only docked within 4.0 Å RMSD accuracy. LibDock performs much worse for PDF in comparison to the best performing docking and scoring programs, and we conclude that LibDock is not suitable for ligands bound in metal-containing active sites. As previously stated, we suspect scoring failure of the metal−ligand interactions in the active site. However, for the five other protein targets, when scored post-docking with LigScore2, results from LibDock are similar to the other well performing docking programs.

Recovery of Active Ligands in a Virtual Screening Protocol. In addition to examining the consistency of LibDock docked poses with the ligand X-ray structures, we have also analyzed the recovery of ligands from a database of both active and inactive compounds. The aim of this study is to ascertain how well LibDock distinguishes active from inactive compounds, hence mimicking a virtual high throughput screening. The compound database contains cognate ligands belonging to each target protein, while the entire ligand database, numbering 1303 ligands, consists of the collection of cognate ligands for all the protein systems. Table 6 shows the number of cognate ligands for each target protein. This study assumes that cognate ligands are active only against their own targets.

The database of compounds was docked to each target and was then scored with both the DJD and LigScore2 scoring functions. For the purposes of this analysis, within a protein cognate ligand set, we have defined the active class of compounds as those characterized with IC_{50} values ≤ 1.0 μM. The remaining compounds are considered inactive.

In Table 7 the percentage of active compounds retrieved in the top 10% of the ranked compounds for each target is reported. This serves as a simplified numerical value to quantify the improvements observed using LigScore2 verses DJD scoring. We reiterate that the rescoring of the database with LigScore2 occurs after docking and optimizing with

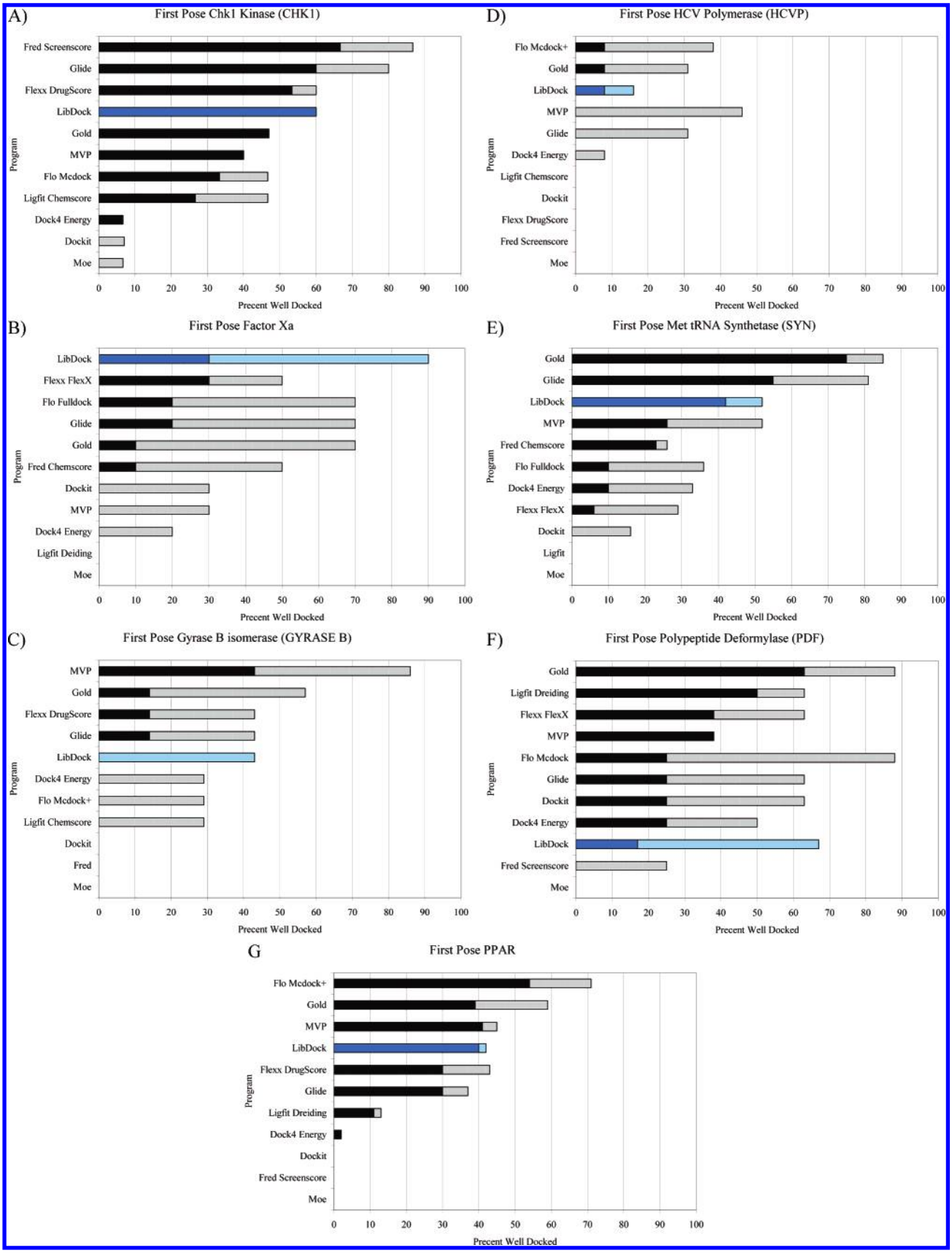

Figure 2. Comparison of RMSD results obtained from LibDock with LigScore2 for each protein target compared to 10 other docking programs reported by Warren et al.[1] Black/dark blue bars indicate the percentage of compounds for which a docked pose was found within 2 Å of the crystal structure, while gray/light blue bars indicate the percentage of compounds for which a docked pose was found within 4 Å of the crystal structure. The graph for each target shows the RMSD for the first pose returned by a particular docking program, For five of the seven protein targets, LibDock with LigScore2 ranks within the top 4 of the 11 docking programs for reproducing a crystallographic pose for the first ranked docking posed. Programs used by Warren et al.: Dock4,[18] Dockit,[19] Fred,[20,21] Flexx,[9] Flo,[12] Glide,[15,16] Gold,[10,11] LigandFit,[17] MOE,[43] MVP.[12]

Table 6. Distribution of Active and Inactive Classes of Ligands Docked against All Targets

enzyme target	no. of cognate compds	min affinity (nM)	max affinity (nM)	no. of active (1.0 μM)	active % (class)	active % (total DB)
CHK1	193	>10 000	7	79	40.9	6.1
HCVP	205	>10 000	5.6	137	66.8	10.5
MRS	144	>10 000	1	115	79.9	8.8
PPARδ	206	>10 000	0.3	151	73.3	11.6
FACTOR Xa	218	5 000	<1	133	61.0	10.2
GYRASE B	138	>10 000	4	95	68.8	7.3
*e*PDF and *s*PDF	199	>10 000	1 and <2	171	85.9	13.1

Table 7. Percent Active Compounds Retrieved[a]

target	DJD	LigScore2	% docked accurately LigScore2 poses
CHK1	5	18	60
HCVP	1	56	9
MRS	32	22	42
PPARδ	13	33	40
FACTOR Xa	26	14	30
GYRASE B	53	75	0
*e*PDF	0	9	13
*s*PDF	1	1	13

[a] The percent of actives retrieved in the top 10% of ranked poses in the database using either DJD or LigScore2. The % docked accurately reiterates the data from Table 5 for the number of top rank LigScore2 poses with RMSD $\leq$ 2.0 Å.

the DJD scoring function. The most dramatic improvement when using the LigScore2 instead of DJD scores is observed for HCVP. Using LigScore2, 56% of the actives are found in the top 10% of the database while only 1% is retrieved using the DJD scoring. This trend of improvement with LigScore2 over DJD scoring is also observed in the other targets, but to a lesser extent. A modest improvement in retrieval for of CHK1 cognate ligands was observed with LigScore2. The isomerase (GYRASE B) and the nuclear hormone receptor (PPARδ) show some enhancement with LigScore2. MRS performs approximately the same for either DJD or LigScore2 scoring. The two metalloproteases (*e*PDF and *s*PDF) have below random retrieval rates for both DJD and LigScore2 scoring. FACTOR Xa is the only target that shows a decrease in retrieval rate with LigScore2 compared to using DJD scores. For HCVP a vast improvement was observed using LigScore2 over DJD scoring with 56% of the active ligands retrieved. Likewise, LigScore2 performed very well for the GYRASE B target, identifying 75% of the actives in the highest 10% scored.

Compared to the other docking−scoring algorithm used by Warren et al.[1] (Figure 3), LibDock performs rather poorly with less than 30% capture of active compounds for CHK1, FACTOR Xa, PDF, and PPARδ. However, there are two exceptions. LibDock is one of the best performing algorithms for both GYRASE B and HCVP (Figure 3D,E). For HCVP approximately the same number of active compounds is retrieved with LibDock as with Flexx Gscore and MVP. For GYRASE B LibDock outperforms even MVP, which consistently performed well in virtual screening across the eight target proteins.

One would argue that the ability of docking−scoring algorithms to distinguish cognate ligands from decoy ligands would be dependent upon its ability to reproduce true binding modes observed in crystal structure complexes. The argument follows that if LibDock reproduced the crystal structure conformation, then components deemed important in binding would be accurately modeled as well. Hence, if a docking−scoring algorithm reproduces a crystal structure, then it should be capable of distinguishing an active from an inactive compound. As discussed in the previous section, a marked improvement was observed in predicating crystal structure conformations when using top ranked LigScore2 instead of top ranked DJD poses. In Table 7 (column 3) the docking accuracy, as judged as the percent top scoring LigScore2 poses with RMSD $\leq$ 2.0 Å, is tabulated along with the retrieval rates in the top 10% of the database. The CHK1 target with the greatest number of accurately docked LigScore2 poses has a poor retrieval rate of active over decoy compounds. This was not observed for CHK1 by Warren et al.[1] in a larger docking−scoring study with various algorithms. In that study, the docking−scoring algorithms that produced the most accurate conformations for CHK1 also produced the best retrieval of actives in the virtual screening protocol.

In contrast to the docking accuracy observed for CHK1, GYRASE B did not have accurately docked poses, but it has the best retrieval rates for virtual screening. In these instances the overall docking accuracy for a target is not predictive of the success in virtual screening when using LibDock with either LigScore2 or even DJD scoring. In the case of GYRASE B with good retrieval of actives in the LibDock virtual screen, none of the crystal structure ligands are docked accurately (RMSD $\leq$ 2.0 Å). However, using a lower RMSD criterion (RMSD $\leq$ 4.0 Å), 57% of the GYRASE B crystal structure ligands are docked with a RMSD $\leq$ 4.0 Å. The Tanimoto coefficients are in the range 0.3−0.5 for these ligands, indicating that portions of the docked ligand do correspond to the binding mode of the crystal structure. Examination of the individual actives retrieved in the top 10% of the database screened indicates that they are of the same structural class. Furthermore, this group of GYRASE B compounds had five representative crystal structures with moderate RMSDs ranging from 2.5 to 4.1 Å. The Tanimoto coefficients and visual inspection indicate that portions of the compounds are simply inverted in the binding site. It appears that at least for the GYRASE B target moderate docking accuracy (RMSDs of 2.0−4.0 Å) is sufficient to distinguish active compounds from decoy compounds using the LigScore2 scoring function. However, this is not applicable to the use of LibDock with the other targets, where despite inaccurate docking the matching and scoring algorithm cannot distinguish active cognate ligands from noncognate ligands in virtual screening.

A direct comparison of the results presented here with the previous validation study reported by Diller and Merz[2,38] is

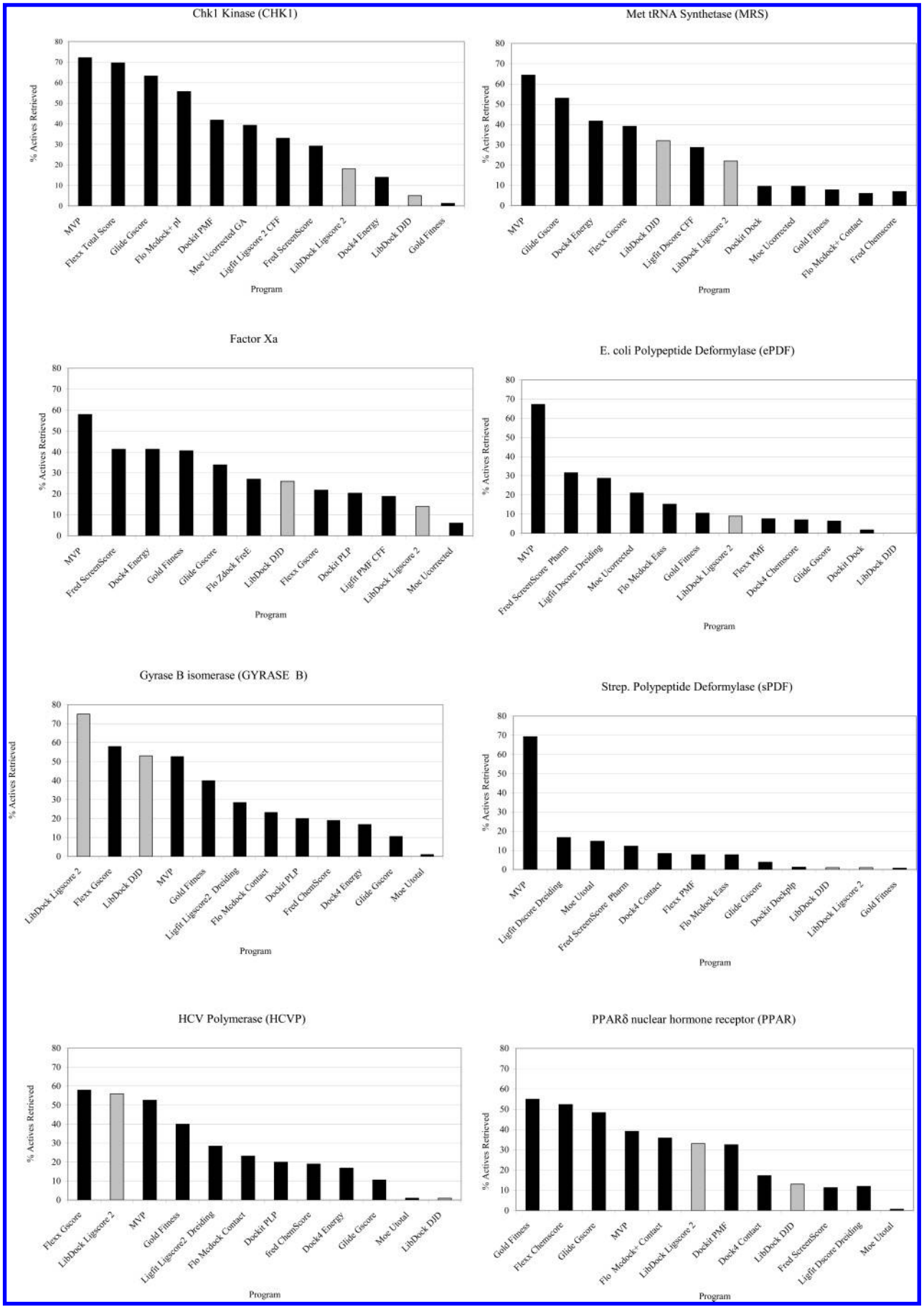

Figure 3. Comparison in the percent of active cognate ligands retrieved in the first 10% of the scored compounds in the virtual screening exercise among the ten docking−scoring algorithms. The most successful scoring protocol for each program is shown for each target.

difficult because the details of the virtual screening protocols were quite different. Diller and Merz used a combinatorial library of compounds to a single target, plasmepsin II. A moderate docking accuracy was obtained for top ranked DJD poses for the core of the compound in the combinatorial library and its reference crystal structure with 33% of docked compounds having a RMSD of 1.5 Å or less. The PLP scoring function was used to rescore these poses, giving an enrichment of 3.3. Under these conditions LibDock had moderate performance for this one target and a combinatorial library of compounds.

The performance in virtual screening as described here is dependent upon both the accuracy of the docked pose and the scoring function used. Conformations docked with an RMSD of 4.0 Å or less appear to be modeled adequately enough to distinguish active ligands from decoy compounds in a virtual database for targets with larger binding sites. However, as the CHK1 example indicates, poses with RMSDs of greater than 4.0 Å are not modeled well enough to be distinguished from inactives.

CONCLUSIONS

The key finding of this evaluation of LibDock for eight pharmaceutically pertinent protein−ligand targets is that the LigScore2 scoring function enhances both the recovery of accurately docked conformations and the retrieval of active ligands in a virtual screening protocol. LibDock starts with a precalculated set of conformers for a compound, matches the polar and apolar features of the protein−ligand interface, and then optimizes the putative complex with the DJD scoring function. This simple and efficient algorithm predicts accurate binding modes in the ensemble of 25 saved poses for most of the protein−ligand complexes examined. The accuracy obtained with LibDock is on par with that achieved by Warren et al.[1] for the same targets for several other popular docking and scoring algorithms. Subsequent rescoring of the ensemble of docked conformations with LigScore2 enhances the performance of the algorithm by increasing the probability that the most accurately docked pose will be ranked as the best scoring conformation.

The LibDock program can be used in two different applications. The first is when a protein−ligand crystal structure is known and prediction of binding modes of compounds prior to synthesis is desired. Calculation of the RMSD or a Tanimoto coefficient with the core features of the crystal structure ligand would give a clear indication of the docked pose closest to the true binding mode. In a second scenario where the crystal structure of a protein−ligand complex is not known, choosing the algorithm with the highest probability of returning accurate binding modes would yield the best outcome. The plots in Figure 2 are meant to serve as a guide to the user in choosing the best program for a specific target. In seven of the enzyme targets investigated, LibDock docking experiments are found to be successful in retrieving at least 50% of the compounds whose docked poses have RMSDs less than 3 Å to the respective X-ray structure poses. LibDock followed by LigScore2 rescoring would be a suitable tool for CHK1, FACTOR Xa, MRS, and PPARδ type protein−ligand complexes. LibDock is not suitable for metalloprotease targets such as PDF and is of limited utility for targets with large binding sites such as GYRASE B and HCVP.

As suggested by this study, using LibDock for virtual screening is problematic. There is no advantage to using LibDock over other popular docking−scoring algorithms for targets such as CHK1, FACTOR Xa, MRS, PDF, and PPARδ. However, for large targets such as HCVP and GYRASE B LibDock performs quite well. At the outset of this study, LibDock was evaluated to determine if using information about binding site features in the docking protocol would enhance results. However, this only appears to hold true for large protein binding sites in the case of virtual screening.

The results for LibDock and the various targets have illustrated both the strengths and weaknesses of the algorithm. For the targets examined in this docking−scoring study a range of successes has been reported for both docking accuracy and virtual screening. The most accurate docking was achieved for protein targets with smaller binding sites, while the best retrieval of active compounds was demonstrated for the targets with larger binding sites. Accurate docking for smaller binding sites makes intuitive sense. However, it is counterintuitive that targets with larger binding sites (GYRASE B and HCVP) demonstrate poorer docking accuracy but better retrieval of active compounds in virtual screening. The analysis of the retrieved compounds for these two targets indicates that only moderate accuracy was achieved (RMSD 2.0−4.0 Å) and that this is sufficient in distinguishing cognate from noncognate ligands. For GYRASE B and HCVP the DJD score also outranked moderately docked cognate compounds over inactive compounds. This suggests that the matching algorithm correctly captures the particular hydrophobic and polar interactions of the cognate compounds in the active site. Therefore, highly accurate docked conformations are not a prerequisite for the retrieval of active compounds in a virtual screen for these two targets.

In this study we found that the kinase performs the best in LibDock with crystallographically observed binding modes being reproduced. Docking and scoring algorithms often perform well in docking accuracy. A correlation plot of LigScore2 pK_i against measured pK_i for the 9 of 15 accurately docked CHK1 compounds (data compiled in Table 5) yields $R = 0.72$ and $R^2 = 0.52$. Correlations between measured pK_i and LigScore2 pK_i, however, were not observed among the other protein targets. Two hypotheses exist as to why kinases perform well in docking and scoring algorithms. The first is that kinases are well represented in the PDB and if the docking−scoring algorithm is trained on a large number of this type of molecules we can expect that a kinase would do well. In training Ligscore2, 112 protein−ligand complexes from the PDB[35] were used with a large number of these consisting of proteases;[3] kinases, however, were not overrepresented in the training set. The second hypothesis is that kinases inherently do well in docking−scoring algorithms because the features of the protein−ligand complex are well modeled by the force field scoring function. This suggests that LibDock is properly representing the kinase−ligand interactions in a simple apolar/polar field.

As noted by Warren et al.,[1] docking multiple compound classes to a single protein crystal structure did not skew the interpretation of docking accuracy as judged by RMSD values. Although one would expect some differences in protein conformation when binding various ligand classes,

these changes were modest for the proteins systems used here. Most of the protein crystal structures from a target series had calculated backbone RMSDs of less than 1.0 Å. For instance, in the case of the 15 protein crystal structures of CHK1, the RMSD of active site residues was 0.8 Å, with the largest difference manifested by tyrosine side-chain rotamers which did not impinge on ligand binding. Furthermore, as noted in the original study,[1] protein structures that could accommodate all of the compound classes under consideration were purposely selected for use in all docking calculations. Evaluating LibDock under the same conditions as the original study was therefore not expected to alter the general assessment of the algorithm's performance and also provides a baseline comparison to other docking and scoring algorithms under conditions similar to those currently used in the pharmaceutical industry.

As reported by Warren et al.,[1] there is a wide range in performance for various docking−scoring algorithms in reproducing the crystallographic binding modes of protein−ligand complexes. LibDock when used in conjunction with LigScore2 reproduces crystallographic poses as well as other docking−scoring algorithms for four of the seven protein target types, CHK1, MRS, PPARδ, and FACTOR Xa, reported in this study.

ACKNOWLEDGMENT

We thank our colleagues at Accelrys for comments and support and David J. Diller for comments and suggestions.

Supporting Information Available: Excel spreadsheets for LibDock data for each of the seven targets. Each spreadsheet lists the compound number, scores, and RMSD for compounds with a corresponding crystal structure. This information is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Warren, G. L.; Andrews, W. C.; Capelli, A.; Clarke, B.; Lalonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Walls, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2006**, *49*, 5912−5931.

(2) Diller, D. J.; Merz, K. M., Jr. High Throughput Docking for Library Design and Library Prioritization. *Proteins* **2001**, *43*, 113−124.

(3) Krammer, A. K.; Kirchoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. LigScore: A Novel Scoring Function For Predicting Binding Affinities. *J. Mol. Graphics Modell.* **2005**, *23*, 395−407.

(4) Fradera, X.; Mestres, J. Guided Docking Approaches to Structure-Based Design and Screening. *Curr. Top. Med. Chem.* **2004**, *4*, 687−700.

(5) Coupez, B.; Lewis, R. A. Docking and Scoring−Theoretically Easy, Practically Impossible. *Curr. Med. Chem.* **2006**, *13*, 2995−3003.

(6) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. Prediction of Protein-Ligand Interactions. Docking and Scoring: Successes and Gaps. *J. Med. Chem.* **2006**, *49*, 5851−5855.

(7) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein-Ligand Docking: Current Status and Future Challenges. *Proteins* **2006**, *65*, 15−26.

(8) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking And Scoring In Virtual Screening For Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discovery* **2004**, *3* (11), 935−949.

(9) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(10) Jones, G.; Willet, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development And Validation Of A Genetic Algorithm For Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.

(11) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins* **2003**, *52*, 609−623.

(12) Lambert, M. H. Docking Conformationally Flexible Molecules into Protein Binding Sites. In *Practical Application of Computer-Aided Drug Design*; Charifson, P. S., Ed.; Dekker: New York, 1997; pp 243−303.

(13) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(14) McMartin, C.; Bohacek, R. S. QXP: Powerful, Rapid Computer Algorithms for Structure-Based Drug Design. *J. Comput.-Aided Mol. Des.* **1997**, 333−344.

(15) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach For Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.

(16) 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*, 1739−1749.

(17) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. Ligandfit: A Novel Method for the Shape-Directed Rapid Docking of Ligands to Protein Active Sites. *J. Mol. Graphics Modell.* **2003**, *4*, 289−307.

(18) Ewing, T. J.; 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*, 411−428.

(19) Blaney, J. M; Dixon, J. S. *DockIt*, version 1.0, Metaphorics, LLC: Misson Viejo, CA; www.metaphorics.com/products/dockit.html.

(20) *Fred*, 1.2.1; OpenEye Scientific Software: Santa Fe, NM.; www.eyesopen.com/applications/fred.html. McGann, M.; Almond, H.; Nicholls, A.; Grant, J. A.; Brown, F. Gaussian Docking Functions. *Biopolymers* **2003**, *68*, 76−90.

(21) Totrov, M; Abagyan, R.; Flexible Protein-Ligand Docking by Global Energy Optimization in Internal Coordinates. *Proteins* **1997**, (Suppl 1), 215−220.

(22) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of Docking: An Overview of Search Algorithms and a Guide To Scoring Functions. *Proteins* **2002**, *47*, 409−443.

(23) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A Review of Protein-Small Molecule Docking Methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.

(24) Booeijmans, N.; Kuntz, I. D. Molecular Recognition and Docking Algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335−373.

(25) 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*, 3032−3047.

(26) Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *J. Med. Chem.* **2004**, *47*, 558−565.

(27) Perola, E.; Walters, W. P.; Charifson, P. S. A Detailed Comparison of Current Docking and Scoring Methods on Systems Of Pharmaceutical Relevance. *Proteins: Struct. Funct. Bioinf.* **2004**, *56*, 235−249.

(28) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M.; Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 45−55.

(29) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46*, 2287−2303.

(30) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative Evaluation of Eight Docking Tools for Docking and Virtual Screening Accuracy. *Proteins: Struct. Funct. Bioinf.* **2004**, *57*, 225−242.

(31) Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening Of Chemical Databases: 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.

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

(33) Ruvinsky, A. M. Role of Binding Entropy In The Refinement Of Protein-Ligand Docking Predictions: Analysis Based On The Use Of 11 Scoring Functions. *J. Comput. Chem.* **2007**, *8*, 1364−1372.

(34) Teramoto, R.; Fukunishi, H. Supervised Consensus Scoring For Docking and Virtual Screening. *J. Chem. Inf. Model.* **2007**, *47*, 526−34.

(35) Holloway, K. M.; McGaughey, G. B.; Coburn, C. A.; Stachel, S. J.; Jones, K. G.; Stanton, E. L.; Gregro, A. R.; Lai, M. T.; Crouthamel, M. C.; Pietrak, B. L.; Munshi, S. K. Evaluating Scoring Functions For Docking and Designing Beta-Secretase Inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 823−827.

(36) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C. The Protein Data Bank. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 899−907.
(37) *Cerius2*, V. c.; Accelrys, Inc.: San Diego, CA, 2003.
(38) Diller, D. J.; Li, R.; Kinases Homology Models, and High Throughput Docking. *J. Med. Chem.* **2003**, *43*, 4638−4647.
(39) Haug, E. J.; McCullough, M. K. A Variational-Vector Calculus Approach to Machine Dynamics. *J. Mech. Trans. Auto. Des.* **1986**, *108*, 25−30.
(40) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. Asparagine And Glutamine: Using Hydrogen Atom Contacts In The Choice Of Side-Chain Amide Orientation. *J. Mol. Biol.* **1999**, *285*, 1735−1747.
(41) Peng, Z.; Ewing, C. S.; Hwang, M.-J.; Waldman, M.; Hagler, A. T. Derivation of Class II Force Fields. Part 4. van der Waals Parameters of Alkali Metal Cations and Halide Anions. *J. Phys. Chem. A* **1997**, *101*, 7243−7252.
(42) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A Generic Force Field for Molecular Simulations. *J. Phys Chem.* **1990**, *94*, 8897−8909.
(43) *Catalyst*, version V; Accelrys Inc.: San Diego, CA, 2004.
(44) *MOE*, version 2002.03; Chemical Computing Group: Montreal, Canada, 2002.
(45) Willett, P.; Barnard, J. M.; Downs, J. M. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983−996.
(46) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular Recognition of The Inhibitor AG-1343 By HIV-1 Protease: Conformationally Flexible Docking by Evolutionary Programming. *Chem. Biol.* **1995**, *2*, 317−324.
(47) Muegge, I.; Martin, Y. C. A General And Fast Scoring Function For Protein-Ligand Interactions: A Simplified Potential Approach. *J. Med. Chem.* **1999**, *42*, 791−804.
(48) Jain, A. Scoring Noncovalent Protein-Ligand Interactions: A Continuous Differentiable Function Tuned To Compute Binding Affinities. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 427−440.
(49) Pearlman, D. A.; Charifson, P. S. Improved Scoring of Ligand-Protein Interactions Using OWFEG Free Energy Grids. *J. Med. Chem.* **2001**, *44*, 502−511.
(50) Smellie, A.; Teig, S. S.; Towbin, P. Poling: Promoting Conformational Variation. *J. Comput. Chem.* **1995**, *16*, 171−187.

CI6004299