ARTICLES

Improving Virtual Screening Performance against Conformational Variations of Receptors by Shape Matching with Ligand Binding Pocket

Hui Sun Lee,[†] Cheol Soon Lee,[‡] Jeong Sook Kim,[¶] Dong Hou Kim,[§] and Han Choe*[,†]

Department of Physiology, Departments of Anatomy and Cell Biology and Research Institute for Biomacromolecules, University of Ulsan College of Medicine, Seoul 138-736, South Korea, Gachon Clinical Trials Center, Gachon University, Incheon 417-842, South Korea, and Nanomol Co., Ltd., Seoul 137-130, South Korea

Received July 5, 2009

In this report, we present a novel virtual high-throughput screening methodology to assist in computer-aided drug discovery. Our method, designated as SLIM, involves ligand-free shape and chemical feature matching. The procedure takes advantage of a negative image of a binding pocket in a target receptor. The negative image is a set of virtual atoms representing the inner shape and chemical features of the binding pocket. Using this image, SLIM implements a shape-based similarity search based on molecular volume superposition for the ensemble of conformers of each molecule. The superposed structures, prioritized by shape similarity, are subjected to comparison of chemical feature similarities. To validate the merits of the SLIM method, we compared its performance with those of three distinct widely used tools ROCS, GLIDE, and GOLD. ROCS was selected as a representative of the ligand-centric methods, and docking programs GLIDE and GOLD as representatives of the receptor-centric methods. Our data suggest that SLIM has overall hit ranking ability that is comparable to that of the docking method, retaining the high computational speed of the ligand-centric method. It is notable that the SLIM method offers consistently reliable screening quality against conformational variations of receptors, whereas the docking methods have limited screening performance.

INTRODUCTION

Virtual screening is an *in silico* approach for rapidly prioritizing molecules with relevant biological activities to a target receptor from a compound library for further experimental validation. Given that the very high cost of experimental high-throughput screening (HTS) necessitates the selective downsizing of libraries consisting of an enormous number of compounds into a small subset that may contain a specific target drug, virtual screening has been increasingly applied in drug discovery projects prior to experimental HTS. The well-directed application of diverse virtual screening tools may lead to a significant improvement in research efficiency in the drug discovery fields. Despite considerable efforts to develop virtual screening methodologies and analysis of their utility, there still remains an urgent need for computational tools with optimized sensitivity and specificity.[1−3]

Three-dimensional (3D) structure-based virtual screening approaches are mainly divided into two classes: those based on ligand coordinates and those based on receptor coordinates.[4] Ligand-based approaches facilitate the identification of library molecules analogous to experimentally determined ligands (actives) by comparing their 3D structural properties. These approaches are typically applied to cases where limited information on the target receptor of interest is available. Ligand-based methods essentially involve a comparative analysis of the structural properties, and thus acquisition of known actives is a prerequisite for their application.[5] One available program, ROCS (Rapid Overlay of Chemical Structures),[6] uses a superposition method for a large-scale shape-based comparison. ROCS compares a Gaussian-based overlap parametrized to reproduce hard-sphere volumes and obtains an optimal alignment with the highest volume overlap between two molecules. Because both the shape and chemical features of a molecule are critical for biological activity, overlap of functional groups is additionally employed, using a color force field. The conformational flexibility of a molecule can be taken into account by precomputing an ensemble of conformers and comparing each in turn. Several publications validate the strength of ROCS performance.[4,7]

In cases where the high-resolution coordinates of target receptors are available, a receptor-based method called molecular docking is a common method of choice for virtual screening.[8] The docking procedure computationally fits molecules from a compound library into the receptor using a suitable scoring function to predict binding affinity. Essentially, this method does not require information on

* To whom correspondence should be addressed. Tel.: +82 2 3010 4292. Fax: +82 2 3010 8029. E-mail: hchoe@amc.seoul.kr.
[†] Department of Physiology, University of Ulsan College of Medicine.
[‡] Gachon University.
[¶] Nanomol Co., Ltd.
[§] Department of Anatomy and Cell Biology, University of Ulsan College of Medicine.

10.1021/ci9002365 CCC: $40.75
Published on Web 10/23/2009

active compounds, although incorporating the binding characteristics of the known actives may enhance its performance.

Typically, docking calculations use an experimentally determined ligand-bound (holo) receptor structure. An important point to be considered in docking is the conformational flexibility of biological macromolecules. An identical target receptor may have different conformations at the ligand binding site, depending on the structures of the bound ligands. These conformational changes have been described using two different mechanisms: "induced fit"[9] and "pre-existing equilibrium dynamics".[10,11] The former postulated that the binding of a ligand induces the conformational change in the receptor, whereas in the latter, the ligand preferentially binds to a specific pre-existing conformation among the ensemble of conformational substates accessed by equilibrium dynamics. The conformational changes that can occur within the binding site of receptors upon ligand binding (by induced fit or/and pre-existing equilibrium dynamic mechanism) have increased the difficulty of docking.

Explicit incorporation of flexible residues at a ligand binding site is too computationally expensive to perform screening on a very large number of molecules in a compound database.[12] Thereby, in most cases, the receptor conformation is kept fixed according to the coordinates of a crystal structure. However, because even small conformational changes in main chains or/and side-chains may significantly influence the score of docked molecules, the single conformation docking approach often results in a reduced chance of finding potential actives.[13,14] Furthermore, when the holo-conformation of a receptor is unavailable for docking or when an interesting target has no experimental structure at all, decrease in the performance becomes more evident. McGovern and Shoichet reported that docking performance against ligand-unbound (apo) and homology-modeled receptors are typically inferior to that against holo-conformations.[15]

In this report, we present a ligand-free high-speed virtual screening methodology, for the purpose of developing a high-speed receptor-based virtual screening tool capable of providing consistently reliable performances against the conformational variation of the receptor. Our approach, designated SLIM (shape-based ligand matching with binding pocket), utilizes the coordinates of a ligand-binding pocket in a target receptor as a query for identifying potential actives. The procedure involves a 3D shape similarity comparison between the negative image of the binding pocket and molecules from a compound library, simultaneously considering their chemical similarities. We describe the computational procedures employed for SLIM and validate its potency in virtual screening by comparing its performance with those of other widely used tools.

MATERIALS AND METHODS

SLIM Methodology. The SLIM method is divided into two principal steps. The first involves generation of the negative image of a binding pocket in a receptor, and the second is a shape and chemical feature similarity comparison between the negative image and molecules from a compound library. Figure 1 is a schematic presentation of the procedures used to generate a negative image of the binding pocket. To construct a negative image, a box centered by a centroid

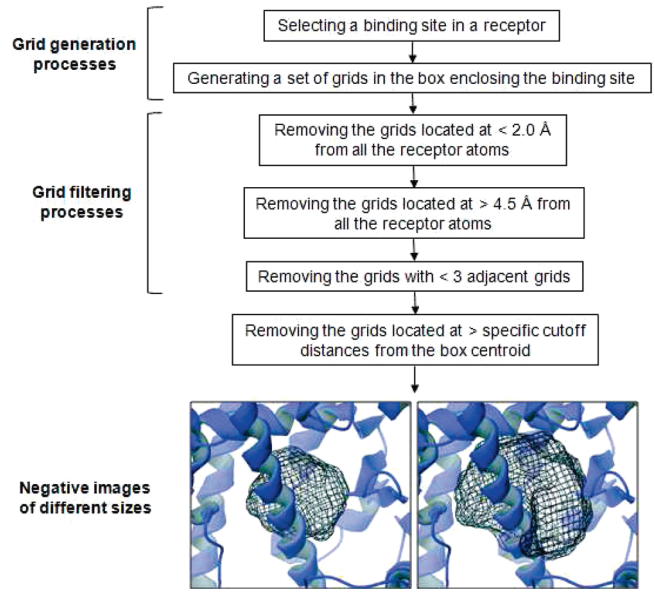

Figure 1. Schematic representation of the procedure used to generate the negative images of a binding pocket; the negative images displayed in this figure were obtained from a PPARγ holo-structure (Protein Data Bank (PDB) entry: 2PRG).

specified in the *X*, *Y*, and *Z* Cartesian coordinates was defined. One can freely determine the centroid at any position in the receptor structure. However, for straightforward comparison of the virtual screening performance, the centroid was determined by coordinates of the geometric center of the cognate ligand in the holo-structure. In the case of apo- and homology-modeled structures, they were superposed onto a holo-structure of the identical target, and the geometric center of the cognate ligand in the holo-structure was used as their centroid. The size of the boxes was set as 20 Å for *X*, *Y*, and *Z*. The defined box was divided into a set of grid points. The grid was cubic, with a grid spacing of 2.0 Å.

To specifically extract the shape of the binding pocket, the grid points in the box were successively discarded by grid filtering criteria in Figure 1. The cluster of remaining grid points corresponds to the negative image, representing the inner shape of the binding pocket. Finally, grid points located more than the specific cutoff distance from the box centroid were removed. In this study, we used the cutoff distance values of 4.0, 5.5, 7.0, and 8.5 Å to obtain four negative images of different sizes. For volume calculation of the negative image, the virtual atom type for each grid point was set to carbon.

To determine chemical complementarity between a binding pocket and library molecules, chemical features were incorporated onto the surface of the negative image. The chemical features were derived from the types of atoms consisting of the binding pocket, and this approach thus requires no cognate ligand in the holo-structure. All receptor atom types were determined according to CHARMM19 EEF1 topology parameters.[16] We then assigned four chemical features—positive, negative, ring, and hydrophobe—to specific atom types. The positive chemical features are represented by hydrogen-bond (H-bond) donor or positively ionizable atoms. The negative chemical features are represented by H-bond acceptor or negatively ionizable atoms. The chemical feature at each grid point constituting the negative image is complementarily assigned by that of the nearest receptor atom of

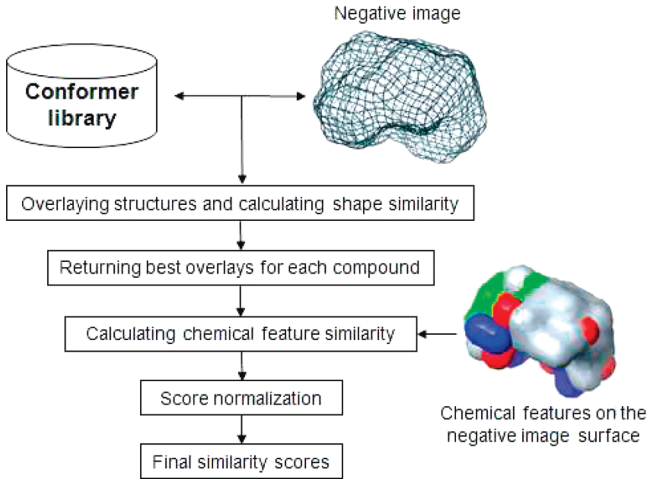

Figure 2. Schematic representation of SLIM method. In this study, we used four negative images of different sizes for shape and chemical feature similarity comparison. We only display a representative negative image, to simplify the figure. In the figure illustrating the chemical features on the negative image surface, positive is shown in blue, negative in red, ring in green, and hydrophobe in gray.

the grid point. The complementary chemical feature pairs between the receptor atom (R) and grid point (G) were defined as follows: positive (R)−negative (G), negative (R)−positive (G), ring (R)−ring (G), and hydrophobe (R)−hydrophobe (G). The positive/negative and ring/hydrophobe chemical features were only assigned to a grid point located within 2.0−4.0 Å and 2.5−4.5 Å from its nearest receptor atoms, respectively.

For shape and chemical feature comparison, in terms of the conformational flexibility of library molecules, multiple conformers of each molecule were generated using the OMEGA program (see the "Preparation of Conformer Library" section of the Materials and Methods for more detail). The shape and chemical feature similarities between the negative image and the library molecules were scored with another in-house program written using the OEChem toolkit (version 1.6).[17] The toolkit was used for structure file reading, 3D shape calculation, volume overlap calculation between two structures, and chemical feature assignment of a library molecule. The overall calculation of shape and chemical feature similarities are schematically presented in Figure 2. Initially, the program read the files for a set of the negative images and the multiple conformers for each molecule from the conformer library. The shape overlap between each negative image and each conformer was maximally optimized. All overlays for the set of conformers onto a negative image were sorted by their Tanimoto coefficient values (S_{shape}). After sorting, the best overlay was returned for each molecule. In this study, four negative images of different sizes were used for the overlays, and therefore the returned results were optimally overlaid four conformers of each library molecule on the four negative images. To assign the chemical features in each library molecule, we used the ImplicitMillsDean color force field, which is a simple pK_a model that defines the H-bond donor, H-bond acceptor, cation, anion, ring, and hydrophobe. The chemical feature similarity between the negative image and overlaid molecule (S_{CFF}) was defined as

$$S_{CFF} = \sum_{i,j} \frac{1}{\exp(r_{ij})} \quad (1)$$

where r_{ij} is the distance between the positions of the assigned chemical features i and j (in the negative image and overlaid compound, respectively). We calculated the chemical feature similarity only for pairs with distances of ≤3.0 Å, and only when their chemical features of the pairs were positive (i)-donor/cation (j), negative (i)-acceptor/anion (j), or ring/hydrophobe (i)-ring/hydrophobe (j). No weight parameter was applied for the specific chemical features.

To estimate the total similarity score (S_{total}), the means (μ) and standard deviations (σ) of all the scores of S_{shape} and S_{CFF} were calculated and Z-transformations were performed on each score, as follows:

$$S_{i,Z,shape} = \frac{S_{i,shape} - \mu_{shape}}{\sigma_{shape}} \quad (2)$$

$$S_{i,Z,CFF} = \frac{S_{i,CFF} - \mu_{CFF}}{\sigma_{CFF}} \quad (3)$$

where S_i is the similarity score for the ith overlaid structure. The S_{total} value of the ith overlaid structure was defined as the sum of the Z-transformed S_{shape} and S_{CFF} values:

$$S_{i,total} = S_{i,Z,shape} + S_{i,Z,CFF} \quad (4)$$

The final similarity score for each compound was determined from the best score among the S_{total} values calculated for the four returned overlays of the molecule.

Preparation of Validation Compound Sets. Distinct protein target systems were selected on the basis of crystal structure availability and previous screening studies by other research groups. All the X-ray structures, including holo- and apo-structures, were obtained from the Protein Data Bank (PDB).[18]

Compound sets consisting of actives and decoys (inactive compounds) for each target were obtained from the directory of useful decoys (DUD), which is a freely available compound database applied as a benchmarking set to evaluate the performance of virtual screening.[19] Targets and compound sets used in this study are summarized in Table 1SI in the Supporting Information.

Preparation of Homology-Modeled Structures. The query (target) sequences were searched to determine the related protein structures by the BLAST against the PDB database. Based on the BLAST search results, two X-ray crystallography structures were selected as the templates for generating two different homology-modeled structures for each target. The template structures used for the homology modeling are summarized in Table 3 (presented later in this work). We built the homology-modeled structures with the MODELLER 9v2 program,[20] using the automodel class. Alignments between target and template sequences were done using the auto_align method of the automodel class. All homology-modeled structures were generated using default parameters to avoid introducing user bias into the system.

Preparation of Conformer Library. For 3D shape comparison of molecules in terms of conformational flexibility, multiple conformers of each compound in the

validation sets were generated using the OMEGA program (version 2.2).[21] The maximum number of rotatable bonds was set sufficiently high to accept all compounds. A maximum of 100 conformers was allowed for each compound, based on a default root-mean-square deviation (rmsd) cutoff of 0.8 Å and an energy window of 10 kcal/mol. The prepared conformer libraries were used to estimate the performance of SLIM and ROCS.

ROCS Program Details. Pairwise 3D shape similarities between cognate ligand in a holo-structure and molecules in conformer libraries were quantitatively calculated using ROCS (version 2.3).[6] In all cases, the X-ray crystallographic conformation of the crystal ligand was used as a query for ROCS. The similarity between the two structures was calculated using the sum of the shape Tanimoto and scaled color values ranging from 0 to 2, where 2.0 represents an exact match of both the shape and functional groups between two molecules. To measure chemical complementarity, we used the ImplicitMillsDean color force field. All experiments were performed with default values for all other parameters.

GLIDE Program Details. The GLIDE program (version 4.0)[22] was used for docking studies. Prior to docking, all water molecules were removed and multimeric complexes simplified from the PDB structure. For GLIDE docking, receptor structures of all holo-structures were preprocessed using protein preparation and refinement components in the docking package. Hydrogen atoms were added using the all-atom force field. Sidechains that were not in close proximity to the ligand binding site and did not participate in salt bridges were neutralized. Restrained minimization using the OPLS-AA force field was performed for refinement of the complex structure. This minimization continued until an average rmsd of the non-hydrogen atoms reached the specified limit of 0.3 Å. Next, a set of grids for each receptor was generated. For apo- and homology-modeled structures, only the preparation procedure was performed prior to grid generation, and grid box sizes were set as 20 Å for *X*, *Y*, and *Z*. The geometric centers of the corresponding holo-structures were used as the centroid coordinates for grid box generation. All docking calculations were run in the "Standard Precision" (SP) mode,[23] with default values for all parameters.

GOLD Program Details. The GOLD program (version 4.0)[24] was also used for docking studies. Hydrogen atoms were added in the structure using the "Add Hydrogen" function. To define a binding site of a holo-structure, the geometric center of the cognate ligand was used as the centroid. As with GLIDE docking, the geometric center of the holo-structure was used to define the binding sites of the apo- and homology-modeled structures of the identical target. The parameters of the "genetic algorithm" were taken as medium default. No restraints were added to minimize the number of spurious docking solutions generated by the program. The "Goldscore" function was used to sort the docking results.

RESULTS

Overall Performance of SLIM for Diverse Targets. The performance of SLIM was compared to those of two widely applied virtual screening methods, ROCS and GLIDE, against specific holo-structures of various target systems. For quantitative comparison, we plotted receiver-operating-characteristic (ROC) curves from the prediction results and calculated the area-under-curve (AUC) values.[25] The AUC value is an objective measure of the overall performance of a given virtual screening tool used to discriminate real "active molecules" from real "decoy molecules". An AUC value of 1.0 signifies that the virtual screening tool perfectly prioritizes active compounds, whereas a value of 0.5 implies random prediction.

Table 1. Measured Performance of Three Different Virtual Screening Methods, SLIM, ROCS, and GLIDE for Various Targets[a]

target	Protein Databank (PDB) ID	Measured Virtual Screening Method Performance		
		SLIM	GLIDE	ROCS
CDK2	1ckp	**0.81**	0.56	0.74
PDE5	1xp0	**0.77**	0.70	0.65
ALR2	2pdg	**0.77**	0.65	0.62
DHFR	3dfr	0.75	0.63	**0.78**
ER	1l2i	0.73	0.72	**0.93**
p38 MAP	1kv2	**0.73**	0.57	0.44
PNP	1b8o	0.72	0.63	**0.93**
PARP	1efy	0.71	**0.96**	0.57
TK	1kim	0.68	0.72	**0.85**
PPARγ	1fm9	0.60	0.79	**0.96**
HIV-PR	1hpx	0.59	**0.64**	0.46
COX2	1cx2	0.50	0.84	**0.96**
Thrombin	1ba8	0.47	**0.81**	0.65
	average:	**0.68**	**0.71**	**0.73**

[a] The PDB ID numbers of the structures used for the program runs are also displayed in the table. The highest AUC in each row are represented in boldface font.

Table 2. Average Run Times of Three Different Virtual Screening Methods over 13 Target Systems

	SLIM	GLIDE	ROCS
average run time (s)	332	12108	87

The performances of each method against the diverse targets are illustrated in Table 1. SLIM showed the highest AUC values for 4 out of the 13 targets, while GLIDE showed the highest AUC values for 3 of the targets and ROCS showed the highest AUC values for 6 of the targets. For a straightforward comparison of the overall performance of each method, the average AUC for all the targets was calculated. On average, ROCS obtained the best performance. Although SLIM was the lowest average AUC among the three tested methods, the results indicate that this method offers relatively reliable performance, comparable to that of GLIDE.

Table 2 shows the average computation times of each method for one search running at 2.4 GHz over the 13 target systems. As expected, the virtual screening speed of the ligand-based method ROCS was the fastest among the tested methods. ROCS was ~3.8 times faster than SLIM. In contrast, the receptor-based method GLIDE showed the slowest calculation speed. The average running time of SLIM was ~36 times faster than that of GLIDE.

Application of SLIM to Diverse Conformations of a Target. One of the biggest limitations of structure-based virtual screening is that a receptor molecule can adopt a variety of conformational substates along a complex energy landscape, because of the flexible nature of

A

PDBid	SLIM	GLIDE	ROCS
2ath	0.75	0.38	0.70
2fvj	0.75	0.63	0.29
1i7i	0.67	0.59	0.98
1knu	0.66	0.57	0.90
2prg	0.64	0.63	0.84
2g0g	0.62	0.38	0.36
1wm0	0.62	0.27	0.46
1zeo	0.61	0.65	0.72
1fm9	0.60	0.79	0.96
2hfp	0.60	0.68	0.37
4prg	0.60	0.58	0.44

B

PDBid	SLIM	GLIDE	ROCS
1kv2	0.73	0.57	0.44
1bl7	0.67	0.62	0.86
1a9u	0.63	0.60	0.95
3ctq	0.61	0.56	0.24
1kv1	0.59	0.52	0.34
1di9	0.58	0.49	0.34
1m7q	0.56	0.52	0.65
1bl6	0.53	0.50	0.96

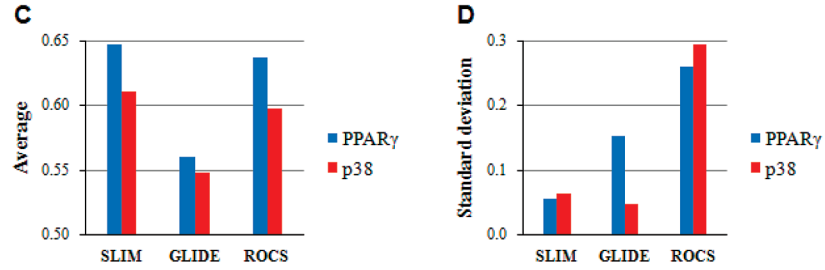

Figure 3. Overall performance of SLIM, GLIDE, and ROCS against various holo-PPARγ and holo-p38 MAP kinase structures, showing AUC raw data for (A) PPARγ and (B) p38 MAP kinase structures. (C) Average AUC values for all structures of each target. (D) Standard deviations for all AUCs of each target.

biological macromolecules.[26] In many cases, ligand binding to the receptor causes changes in the energy landscape, resulting in conformational alterations in the receptor structure. The X-ray crystallography structure of an interesting target represents a set of fixed coordinates and, therefore, only provides limited structural information on the target for virtual screening.[27] To estimate the performance of a virtual screening tool in detail, it is worth considering the structural diversity of ligands and the conformational changes in the receptor binding site. A simple approach is to use multiple sets of the holo-conformations of a target and their cognate ligands. For this purpose, we selected two target systems: peroxisome proliferator-activated receptor γ (PPARγ) and p38 mitogen-activated protein (p38 MAP) kinase.

PPARγ contains a deeply buried, rigid, and sterically constrained hydrophobic binding site. It has been typically known that nuclear receptors, such as PPARγ, are easy targets, particularly for docking.[28] Among the collected crystal holo-structures of PPARγ, 11 structures were finally selected based on the structural diversity of their cognate ligands. The cognate ligand structures in the selected holo-structures are illustrated in Figure 1SI in the Supporting Information. Performance comparison results of the three tools are shown in Figures 3A, 3C, and 3D. The calculated AUC values varied significantly, depending on the receptor conformations and the structures of cognate ligands. SLIM was better than random for all of the holo-conformations, although for some holo-conformations, such as 1fm9, 2hfp, and 4prg, the performance enhancement was relatively small. When averaged over all of the AUC values, SLIM displayed an overall best performance of ~0.65 AUC. ROCS displayed the most variable performance across cognate ligand structures used as a query: it performed excellently (i.e., AUC > 0.90) for cognate ligands in the 1i7i and 1fm9 holo-structure, while displaying exceptionally low virtual screening efficacy (i.e., AUC < 0.50) for 2fvj, 2g0g, 1wm0, 2hfp, and 4prg. The reason for this inferior performance is that the query structures used for the ROCS calculations have uncommon scaffolds for PPARγ ligands and, therefore, depicts the chemical space covered only by a small fraction of the actives.[29] The actives with significantly different shape from the query molecule are likely to be missed during screening. This bias in performance was evident in a plot of standard deviation values calculated over all of the holo-structures. In contrast, it is notable that SLIM showed the lowest fluctuations in performance.

The SLIM method was further examined against another target system: p38 MAP kinase. The p38 MAP kinase target system is a particularly challenging target, because of its structural flexibility, a high degree of solvent exposure of the binding site, and a relatively shallow hydrophobic cleft.[28] Eight p38 MAP kinase/ligand complex crystal structures were selected for the experiments. The structures of the cognate ligands in holo-structures are presented in Figure 2SI in the Supporting Information. Figures 3B, 3C, and 3D show performance comparison results for this target. As with PPARγ, performance was significantly affected by the receptor conformations and the structures of cognate ligands. On average, all three methods were relatively ineffective, compared to the case of PPARγ. The data corroborate once more that SLIM provides a consistent overall performance against diverse receptor conformations.

Good and Oprea investigated the analogue bias of the DUD active compound sets.[30] They systematically filtered the actives in the datasets using leadlike criteria and reduced graphs. This filtering significantly reduced the number of the actives, indicating that the DUD datasets have an analogue bias in the chemical structures of the actives. This bias in the validation sets may affect the performance of virtual screening tools. To examine whether large variations in the ROCS performance result

Table 3. Target Systems for Method Validation against Apo- and Homology-Modeled Structures

	Holo	Apo		Model	Template			
target	species	PDB	species	species	PDB	description	species	seq i.d. (%)
CDK2	*H. sapiens*	1pw2	*H. sapiens*	*H. sapiens*	2qkr	CDK	*C. parvum iowa II*	57.5
					1bi8	CDK6	*H. sapiens*	44.2
PDE5	*H. sapiens*	2h40	*H. sapiens*	*H. sapiens*	2o8h	PDE10	*R. norvegicus*	37.5
					2oun	PDE10	*H. sapiens*	34.8
ALR2	*H. sapiens*	1ads	*H. sapiens*	*H. sapiens*	1dla	aldose reductase	*S. scrofa*	85.8
					1q13	hydroxysteroid dehydrogenase	*O. cuniculus*	49.8
p38 MAP	*H. sapiens*	1wfc	*H. sapiens*	*H. sapiens*	3gp0	p38β	*H. sapiens*	72.8
					1pme	Erk2 MAP kinase	*H. sapiens*	44.1
PNP	*B. taurus*	1pbn	*B. taurus*	*B. taurus*	1pf7	PNP	*H. sapiens*	86.9
					2p4s	PNP	*A. gambiae*	41.6
PARP	*G. gallus*	2paw	*G. gallus*	*G. gallus*	1uk0	PARP	*H. sapiens*	86.9
					1gs0	PARP2	*M. musculus*	45.2
PPARγ	*H. sapiens*	1prg	*H. sapiens*	*H. sapiens*	1y0s	PPARδ	*H. sapiens*	62.5
					2v0v	Rev-erbβ	*H. sapiens*	27.5
Thrombin	*H. sapiens*	1hgt	*H. sapiens*	*H. sapiens*	1uvt	Thrombin	*B. taurus*	79.5
					1ezq	coagulation factor XA	*H. sapiens*	36.4

from the structural bias of the validation sets, we deliberately prepared diverse active subsets from the DUD datasets for PPARγ and p38 MAP kinase (see Figure 3SI in the Supporting Information). We then measured the overall hit ranking ability of ROCS using the newly prepared datasets. The average AUCs and standard deviations for all AUCs were comparable to those for the original DUD datasets (see Table 2SI in the Supporting Information). This clearly demonstrates that the highest fluctuations of the ROCS performance by query selection is not fundamentally attributed to an intrinsic bias in chemical structures of the actives in the DUD datasets.

Application of SLIM to Apo- and Homology-Modeled Structures. A major challenge in docking is to use receptor conformations that are not preorganized for ligand binding. When no structural information on biologically active molecule is available for an interesting target, ligand-based approaches and docking with a holo-receptor cannot be utilized for virtual screening. Under these circumstances, a ligand-unbound (apo) conformation often must be used to select lead compound candidates, if available. Furthermore, in cases where a target has no experimental structure at all, a structural model computationally generated by homology modeling techniques may be a choice for docking experiments.[31,32] However, these approaches may not be useful, because apo- and homology-modeled conformation analyses commonly yield lower enrichments of known actives in a compound database than holo-conformation. This is due to a decreased ability of these conformations to recognize ligands at the active site.[15]

Based on our data, which suggests that SLIM shows relatively reliable and homogeneous performance against diverse holo-conformations of a receptor, we expanded the application of SLIM into apo- and homology-modeled conformations. To evaluate the impact of the receptor conformation on the virtual screening performance systematically, eight target systems were screened in additionally prepared three different conformations each: an apo conformation and two homology-modeled conformations (see Table 3). The performance of SLIM was compared to those of two widely used docking tools:

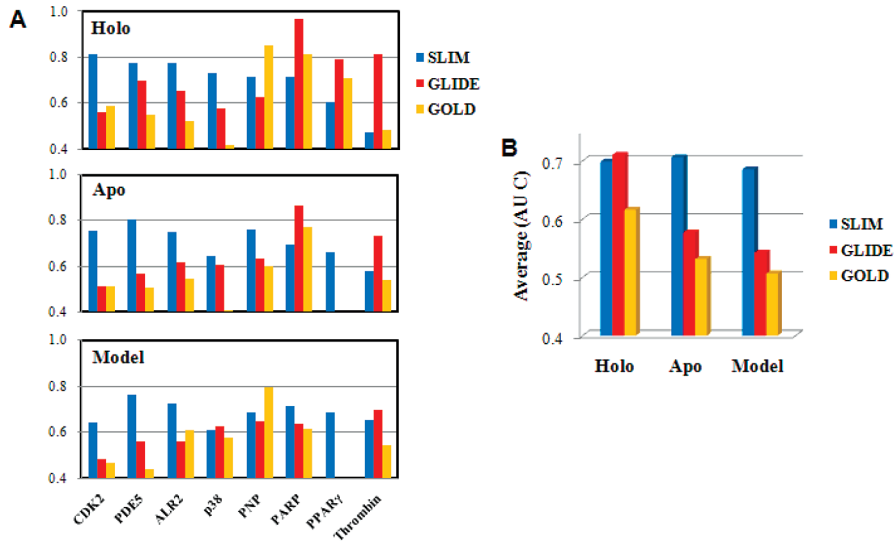

Figure 4. AUC values yielded by three receptor-centric virtual screening tools, SLIM, GLIDE, and GOLD: (A) the AUC values against holo-, apo- and homology-modeled conformations of each target receptor (in the case of the homology-modeled conformation, an average of the AUC values for two different models of each target receptor was used to produce this figure); (B) AUC average graph over all of the targets.

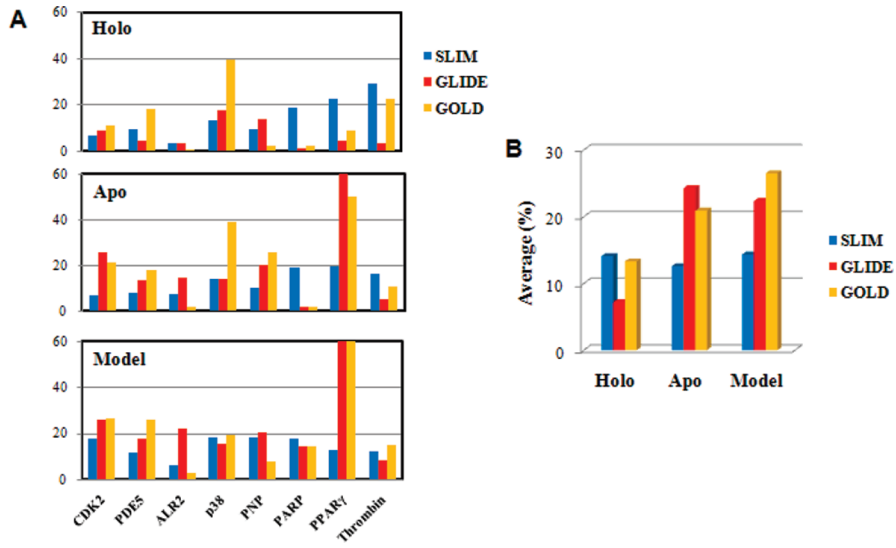

Figure 5. Percentage of the ranked compounds necessary to find 25% of the actives, yielded by SLIM, GLIDE, and GOLD: (A) the percentage of the database to find 25% of actives against holo-, apo- and homology-modeled conformations of each target receptor; (B) average percentage graph over all of the targets.

GLIDE and GOLD. All benchmark raw data are provided in Table 3SI in the Supporting Information.

SLIM obtained the highest AUC values for more targets than GLIDE and GOLD with apo- and homology-modeled conformations (see Figure 4A). The SLIM screening against the apo- and homology-modeled conformations yielded average AUC values of 0.71 and 0.69, respectively, which are similar to that against holo-conformations (AUC of 0.70) (see Figure 4B). This obviously shows that the overall hit ranking ability of SLIM is homogeneous over the holo-, apo-, and homology-modeled conformations. On the other hand, when we applied the docking programs, their overall hit ranking ability was significantly affected by the different types of conformations. In the case of GLIDE, its average AUC of 0.71 for the holo-conformations decreased to 0.54 for the homology-modeled conformations. The same trend was also observed in the GOLD docking results. Enrichment was additionally evaluated using another criterion: the percentage of the ranked database compounds required to find 25% of the actives (see Figure 5). On average, SLIM concentrated the top 25% of actives in a smaller fraction of the compound database than GLIDE and GOLD with the apo- and homology-modeled conformations. Together, the data indicate that, when SLIM is applied to less-defined receptor conformations (apo- and homology-modeled conformation), it can offer better enrichment than the docking method.

Detailed analysis of SLIM for various target structures illustrates some of the limitations of this method in regard to application to ligand discovery. SLIM showed a similar average AUC value to GLIDE for the holo-conformations (see Figure 4B). However, SLIM found 25% of actives in the top 14.09% of the database, on average, while GLIDE found 25% of actives in the top 7.18% (see Figure 5B). This indicates that the overall hit ranking ability of SLIM for the holo-conformations is comparable to that of GLIDE, but the ability of recognizing the actives at the beginning of a rank-ordered list is inferior. This trend can also be observed in the results of average maximum enrichment factor (EF) and average percentage of the database where maximum EF is observed (see Table 4). The EF value takes into account the improvement of the hit rate by a virtual screening tool, compared to a random selection, and is defined as

Table 4. Average Maximum Enrichment Factor (EF) and Average Percentage of the Database Where the Maximum EF is Observed for SLIM, GLIDE, and GOLD on the Eight Targets in Table 3

	Average Maximum EF Value			Average Percentage of Database Where Maximum EF Occurred		
	SLIM	GLIDE	GOLD	SLIM	GLIDE	GOLD
holo	4.96	15.35	15.63	14.12	1.42	0.70
apo	7.60	8.84	12.71	9.39	1.19	17.04
homology-modeled	5.38	9.30	15.14	8.24	14.90	17.79

$$\mathrm{EF} = \frac{\mathrm{Hits}_{\mathrm{sampled}}/N_{\mathrm{sampled}}}{\mathrm{Hits}_{\mathrm{total}}/N_{\mathrm{total}}} \quad (5)$$

where $\mathrm{Hits}_{\mathrm{sampled}}$ is the number of hits found, N_{sampled} the number of compounds screened at a given percentage of the database screened, $\mathrm{Hits}_{\mathrm{total}}$ the number of total actives in entire database, and N_{total} the number of compounds in the entire database. In holo-conformation, the average maximum EF value of SLIM is less than that of the docking programs. The docking programs GLIDE and GOLD showed average maximum EF values of 15.35 and 15.63 at the top 1.42% and 0.70%, respectively, whereas SLIM has an average maximum EF value of 4.96 at the top 14.12%. This means that the docking methods show better "early recognition" ability than SLIM, while their performances are significantly deteriorated with apo- and homology-modeled conformations.

Refinement of the SLIM Methodology. In this study, we defined the grid box (which is divided into a set of grid points with a grid spacing of 2.0 Å) to generate negative images of the target binding site. To examine whether a smaller grid spacing can improve the SLIM performance, we performed SLIM runs, using the negative images generated by a grid spacing of 1.0 Å and then measured the AUCs (see Table 4SI in the Supporting Information). The results showed that better performance is achieved for some target

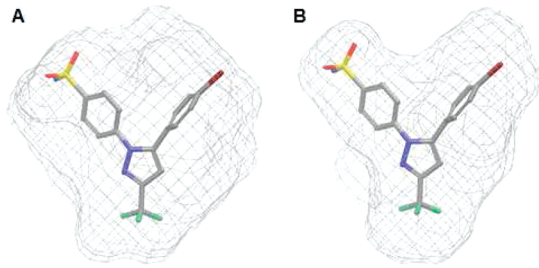

Figure 6. Comparison of the negative image shapes of COX2 binding site using different procedures. The negative images (mesh representation) displayed have been obtained from PDB entry 1CX2. The cognate ligand in the crystal structure is also shown in a stick representation. Panel (A) shows a grid spacing of 2.0 Å, and the cutoff for removal of grids close to the receptor atoms is 2.0 Å. Panel (B) shows a grid spacing of 1.0 Å, and the cutoff for removal of grids close to the receptor atoms is 3.0 Å.

systems such as PDE5, PARP, and thrombin. On average, however, a smaller grid spacing did not improve the performance, whereas average computational time, which is a strong point of the SLIM method, was almost two times longer.

As shown in Table 1, the AUCs for COX2 and thrombin were not better than random. We sought a way to improve the performance of SLIM for the two proteins. The ligand binding site of COX2 is burial and lipophilic. Major factors that determine the binding of the ligand are the overall shape fitting and hydrophobic interactions, rather than H-bond interactions.[33] To prepare negative images more similar to the shape of the cognate ligand, we generated a set of grids using a grid spacing of 1.0 Å and then removed the grids located <3.0 Å from all the receptor atoms (see Figure 6). We also applied a weight of 5 to interactions between rings and hydrophobes for the calculation of chemical feature similarity, instead of using the same weight. Indeed, this reparameterization improved the AUC from 0.50 to 0.65.

The binding site of thrombin is highly exposed to the solvent. Therefore, appropriate removal of grids located at the solvent-exposed regions may enhance the performance of SLIM. To determine the grids at the solvent-exposed regions, we calculated enclosure for each grid. The enclosure is defined by the fraction of radial rays that strike the receptor surface atoms among 146 evenly spaced radial rays from a grid point. If the enclosure is <0.5, the grid is removed.[34] Figure 7 shows the negative image before and after the grid filtering by the enclosure calculation. When this algorithm was added to the original procedure for the negative image generation, the AUC increased from 0.47 to 0.58.

We showed that modification of some parameters was able to contribute to the enhanced screening performance of SLIM. However, applying the reparameterization did not lead to universal improvement in the performance over all the target systems (data not shown). The reparameterization effects seem to be strongly dependent on the binding site characteristics of each target.

DISCUSSION AND CONCLUSIONS

In this study, we have designed a novel virtual screening methodology that is called SLIM. It takes advantage of a negative image of a binding pocket, which is a set of virtual atoms that represents the inner shape and chemical features of the binding pocket, to calculate the shape and chemical feature complementarity with library compounds. SLIM is a receptor coordinate-based method using the shape and chemical feature comparisons typically used in ligand coordinate-based methods. We have evaluated the SLIM method for different targets and diverse conformations of each target, and compared it to other widely used tools, with particular focus on the ability of SLIM to distinguish active compounds from a larger set of inactive compounds.

Docking performance often varies significantly with conformational changes in the ligand-binding site. The choice of receptor conformation for docking to enhance virtual screening efficiency is a particularly challenging issue. Our results indicate that docking has limited screening efficiency, depending on the receptor conformation. In contrast, SLIM has offered consistently reliable performance, despite the conformational variations of the target receptor. In particular, SLIM has showed improved performance with apo- and homology-modeled conformation, compared to the docking methods. We propose that the consistent virtual screening performance of SLIM is a result of simple parametrization. Relatively complicated parametrization in the docking program may retrieve only active compounds overspecialized for a receptor conformation and, thus, result in a reduction in true hits. We have showed that a simple approach based on a comparison of shape and chemical features (i) helps to overcome this problem of docking programs and (ii) can considerably reduce the computation time.

Hawkins et al. reported that the ROCS methodology consistently showed good performance against a diverse target system.[4] In our results, calculation of the average performance also revealed that ROCS performed consistently well. However, ROCS performance varied significantly, depending on the queries used for similarity searches. This

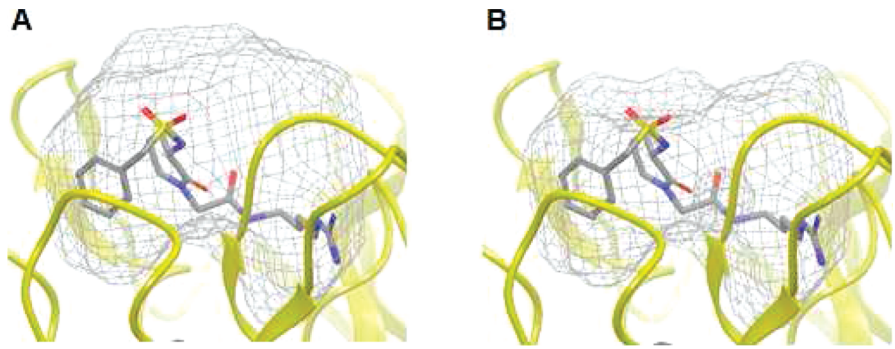

Figure 7. Comparison of the negative image shapes of thrombin binding site (A) before and (B) after the grid filtering by enclosure calculation. The negative images (mesh representation) were obtained from PDB entry 1BA8.

finding suggests that, in ligand three-dimensional (3D) shape-based virtual screening methods, the prudent selection of a query molecule is needed to extensively search the chemical space of diverse scaffold. In recent work by Kirchmair et al., they systematically evaluated the impact of multiple queries on the performance of ROCS, and they suggested that using the center molecule of the active compounds is a suitable way to optimize the screening performance.[29]

All of the virtual screening methods evaluated in this study have their own superior features, compared to another one. The screening setup for the best virtual screening performance relies largely on the focus of a screening campaign. The careful selection and combination of virtual screening approaches may facilitate research efficiency in drug discovery projects. To that purpose, we would like to propose potential applications of the SLIM method. SLIM is a receptor coordinate-based virtual screening method with the high computational speed of ligand coordinate-based methods. SLIM constitutes a substantial advantage for virtual screening as an alternative of docking in cases where there is no well-defined holo-conformation of an interesting target. SLIM can also be applicable to a target enzyme with high conformational flexibility, where improper selection of receptor conformation might severely impair docking performance, to downsize libraries that consist of an enormous number of compounds in a small subset at the early stage of drug discovery.

After completing this manuscript, we found that a conceptually similar study has been reported by Ebalunode et al.[35] Although they also used the OEChem toolkit for shape comparison between the negative image of a target binding site and library compounds, the processes for generating negative image, adding chemical features to the negative image, and prioritizing hit compounds are significantly different from our approaches. Moreover, a major goal of our research was the development of a high-speed virtual screening tool capable of providing consistently reliable performances against receptor conformational changes. Accordingly, our study was mainly focused on confirming the virtual screening efficiency against the diversity of receptor conformations. In addition, the performances were directly compared to those of representative docking tools, GLIDE and GOLD. The work by Ebalunode et al.,[35] on the other hand, only provided a comparison of the performances with a related tool, the FRED:ShapeGauss method, for the specific crystal holo-structures of five target systems.

Current version of SLIM utilizes simply classified chemical features; therefore, a more-sophisticated assignment of the chemical features may enhance performance. Furthermore, improved negative image generation, as well as better scoring systems for the calculation of shape and chemical feature similarities, are promising procedures to explore. Parameter optimization specialized for characteristics of each binding site can enhance the performance of SLIM, although it is a challenging issue.

ACKNOWLEDGMENT

This work was supported by the Korea Research Foundation (KRF-2007-412-J00304) and Brain Research Center of the 21st Century Frontier Research Program (2009K001276) funded by the Ministry of Education, Science and Technology, Republic of Korea. We thank Dr. Yoon for allowing us to use his GLIDE and Dr. Glickstein for careful English proofreading.

Supporting Information Available: The number of actives and decoys in each validation dataset, cognate ligands in PPARγ and p38 MAP kinase holo-structures used, diversity plots of actives in PPARγ and p38 MAP kinase validation datasets, ROCS performances on unbiased DUD datasets for PPARγ and p38 MAP kinase, SLIM performances on grid spacing 1.0 Å, and enrichment of actives by holo-, apo-, and homology-modeled receptors. This information is available free of charge via the Internet at http://pubs.acs.org/.

REFERENCES AND NOTES

(1) Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **2006**, *11*, 580–594.
(2) Taft, C. A.; Da Silva, V. B.; Da Silva, C. H. Current topics in computer-aided drug design. *J. Pharm. Sci.* **2008**, *97*, 1089–1098.
(3) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein-ligand docking: current status and future challenges. *Proteins* **2006**, *65*, 15–26.
(4) Hawkins, P. C.; Skillman, A. G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74–82.
(5) Stahura, F. L.; Bajorath, J. Virtual screening methods that complement HTS. *Comb. Chem. High Throughput Screen* **2004**, *7*, 259–269.
(6) *ROCS*, Version 2.3; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2008.
(7) Moffat, K.; Gillet, V. J.; Whittle, M.; Bravi, G.; Leach, A. R. A comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS. *J. Chem. Inf. Model* **2008**, *48*, 719–729.
(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*, 935–949.
(9) Koshland, D. E. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.
(10) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. Folding funnels and binding mechanisms. *Protein Eng.* **1999**, *12*, 713–20.
(11) Tsai, C. J.; Ma, B.; Sham, Y. Y.; Kumar, S.; Nussinov, R. Structured disorder and conformational selection. *Proteins* **2001**, *44*, 418–427.
(12) Abagyan, R.; Totrov, M. High-throughput docking for lead generation. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375–382.
(13) 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.
(14) Murray, C. W.; Baxter, C. A.; Frenkel, A. D. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 547–562.
(15) McGovern, S. L.; Shoichet, B. K. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J. Med. Chem.* **2003**, *46*, 2895–2907.
(16) Lazaridis, T.; Karplus, M. Effective energy function for proteins in solution. *Proteins* **1999**, *35*, 133–152.
(17) *OEChem Tool Kit*, Version 1.6; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2008.
(18) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
(19) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.
(20) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
(21) *OMEGA*, Version 2.2; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2008.
(22) *Glide*, Version 4.0; Schrödinger, LLC: Portland, OR, 2008.
(23) 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.
(24) 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.

(25) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J.-P.; Bertrand, H.-O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype. 4. *J. Med. Chem.* **2005**, *48*, 2534–2547.
(26) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **1991**, *254*, 1598–1603.
(27) B-Rao, C.; Subramanian, J.; Sharma, S. D. Managing protein flexibility in docking and its applications. *Drug Discovery Today* **2009**, *14*, 394–400.
(28) 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.
(29) Kirchmair, J.; Distinto, S.; Markt, P.; Schuster, D.; Spitzer, G. M.; Liedl, K. R.; Wolber, G. How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J. Chem. Inf. Model.* **2009**, *49*, 678–692.
(30) Good, A. C.; Oprea, T. I. Optimization of CAMD techniques. 3. Virtual screening enrichment studies: A help or hindrance in tool selection. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 169–178.
(31) Park, H.; Bahn, Y. J.; Jung, S. K.; Jeong, D. G.; Lee, S. H.; Seo, I.; Yoon, T. S.; Kim, S. J.; Ryu, S. E. Discovery of novel Cdc25 phosphatase inhibitors with micromolar activity based on the structure-based virtual screening. *J. Med. Chem.* **2008**, *51*, 5533–5541.
(32) Kenyon, V.; Chorny, I.; Carvajal, W. J.; Holman, T. R.; Jacobson, M. P. Novel human lipoxygenase inhibitors discovered using virtual screening with homology models. *J. Med. Chem.* **2006**, *49*, 1356–1363.
(33) Schulz-Gasch, T.; Stahl, M. Binding site characteristics in structure-based virtual screening: evaluation of current docking tools. *J. Mol. Model.* **2003**, *9*, 47–57.
(34) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.
(35) Ebalunode, J. O.; Ouyang, Z.; Liang, J.; Zheng, W. Novel approach to structure-based pharmacophore search using computational geometry and shape matching techniques. *J. Chem. Inf. Model.* **2008**, *48*, 889–901.

CI9002365