Maximum Common Binding Modes (MCBM): Consensus Docking Scoring Using Multiple Ligand Information and Interaction Fingerprints

Steffen Renner,*[,†,||] Swetlana Derksen,[†,‡] Sebastian Radestock,[†,⊥] and Fabian Mörchen[§]

Chemical R&D, Merz Pharmaceuticals GmbH, Eckenheimer Landstrasse 100, D-60318 Frankfurt am Main, Germany, Institut für Organische Chemie und Chemische Biologie, Johann Wolfgang Goethe-Universität, Siesmayerstrasse 70, D-60323 Frankfurt am Main, Germany, and Siemens Corporate Research, 755 College Road East, Princeton, New Jersey 08540

Received October 7, 2007

Improving the scoring functions for small molecule-protein docking is a highly challenging task in current computational drug design. Here we present a novel consensus scoring concept for the prediction of binding modes for multiple known active ligands. Similar ligands are generally believed to bind to their receptor in a similar fashion. The presumption of our approach was that the true binding modes of similar ligands should be more similar to each other compared to false positive binding modes. The number of conserved (consensus) interactions between similar ligands was used as a docking score. Patterns of interactions were modeled using ligand receptor interaction fingerprints. Our approach was evaluated for four different data sets of known cocrystal structures (CDK-2, dihydrofolate reductase, HIV-1 protease, and thrombin). Docking poses were generated with FlexX and rescored by our approach. For comparison the CScore scoring functions from Sybyl were used, and consensus scores were calculated thereof. Our approach performed better than individual scoring functions and was comparable to consensus scoring. Analysis of the distribution of docking poses by self-organizing maps (SOM) and interaction fingerprints confirmed that clusters of docking poses composed of multiple ligands were preferentially observed near the native binding mode. Being conceptually unrelated to commonly used docking scoring functions our approach provides a powerful method to complement and improve computational docking experiments.

INTRODUCTION

Docking small molecule ligands into protein receptors consists of two different tasks: the generation of reasonable docking poses and scoring these poses. While docking programs are often found to solve the former task to an acceptable degree, there is frequently the problem that current scoring functions are not able to identify the best results among the set of generated docking poses.[1,2] Therefore improving scoring functions is a major challenge to increase the usability of computational docking programs. Commonly used scoring functions include force-field-based, empirical, and knowledge-based functions.[3,4] All these approaches suffer from simplifications for the sake of computational speed or insufficient training data, which results in an unpredictable performance of the scoring functions for a particular target. Consensus scoring methods that integrate predictions of several independent scoring functions have gained a strong interest.[5−7] The idea of such approaches is that individual weaknesses of each method are averaged out by the use of multiple scoring functions that provide different views on the problem. However if none of the applied underlying scoring functions performs well for a particular target, the consensus will also perform poorly.[8] Consequently there is still a need for new docking scoring functions that provide more accurate or conceptually new views on the quality of a docked pose.

Consensus methods were also applied successfully in ligand-based virtual screening, e.g., similarity searching using different descriptors,[9,10] different similarity metrics,[11,12] or with ensembles of machine learning models.[13−15] In ligand-based approaches it is common practice to use not only a consensus of scoring functions (here, e.g., similarity metrics) but also a consensus of multiple known active reference ligands, either based on molecular descriptors,[16,17] machine learning methods,[18−20] or on 3D pharmacophore models.[21−23] Here features that are infrequently found in the reference molecules ("weaknesses") might not be important for a particular biological activity and are averaged out by the consensus, and the method focuses on "important" features present in the majority of the ligands. In the present study we aimed to transfer this concept to the selection of docking poses: consensus binding modes extracted from docking poses of multiple known active ligands (i.e., clusters of multiple ligands with similar binding modes) were assumed to be likely near to the correct binding mode. Binding modes that were only found for single ligands or that were not very similar between multiple molecules were considered to be false positives ("weaknesses") of the docking procedure. Actually a recent study investigated whether similar ligands

* Corresponding author phone: +49-231-133-2432; e-mail: steffen.renner@mpi-dortmund.mpg.de.
† Merz Pharmaceuticals GmbH.
‡ Johann Wolfgang Goethe-Universität.
§ Siemens Corporate Research.
|| Current address: Department IV - Chemical Biology, Max-Planck-Institut für molekulare Physiologie, Otto-Hahn-Strasse 11, D-44227 Dortmund, Germany.
⊥ Current address: Fachbereich Biowissenschaften, Molekulare Bioinformatik, Johann Wolfgang Goethe-Universität, Max-von-Laue-Strasse 9, D-60438 Frankfurt am Main, Germany.

10.1021/ci7003626 CCC: $40.75
Published on Web 01/23/2008

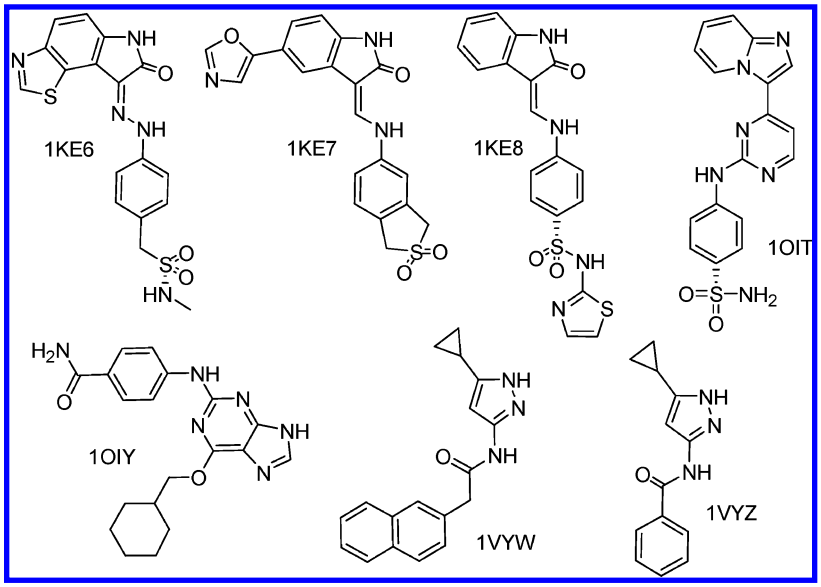

Figure 1. Ligands from the cyclin-dependent kinase 2 (CDK-2) data set with PDB IDs.

bind in a similar fashion.[24] While slight changes in the interactions between structurally similar ligands and their receptor were found in the majority of cases a general high conservation of the binding mode was observed.

It is common strategy to cluster the docking poses generated for a single ligand to isolate more populated clusters from smaller ones which are assumed to be less likely the bioactive binding mode.[25] The similarity between docking poses is usually calculated by the rmsd (root-mean-square distance) value. However this strategy is not possible using nonidentical ligands. The first promising consensus approach overcoming this limitation was reported recently.[26] It used docking poses of a series of different ligands containing a common scaffold. Here only the rmsd values of the common scaffold were used to obtain relevant clusters of docking poses. However this strategy was still not amendable for ligands that did not contain a common scaffold, which is an often found situation in computational chemistry.

A recently published novel docking scoring strategy applied structure interaction fingerprints (SIFt)[27] representing the interactions found between a ligand and a receptor in the form of a fingerprint. The idea of this approach was to use a ligand from a cocrystal structure as a reference to score docking poses of other ligands by fingerprint similarity. The SIFt approach was shown to outperform traditional scoring functions in virtual screening and was applied successfully to cluster docking poses of individual ligands.[27] Several extensions of these algorithms were reported, extending the accuracy of the definition of the interactions used for the fingerprint,[28,29] evaluating the approach for virtual screening,[30] scaffold docking,[31] or the use of profile SIFt for virtual screening.[32] The largest obstacle of this promising approach is that a reference ligand is mandatory. The same is true for a set of other methods that are able to incorporate available information on binding modes for the scoring of docking poses, like using the 3D similarity of docking poses to a bound ligand,[33,34] pharmacophore constraints,[35] or target-biased scoring functions.[36] Having a more reliable scoring function to predict binding modes of known ligands would also increase the applicability of these tools.

In our approach we used structure interaction fingerprints to describe the binding mode of docking poses. Using these fingerprints docking poses of ligands were scored according to the number of conserved interactions they have in common with docking poses of other ligands. Poses with many conserved interactions were considered to be more likely near-native compared to poses with no or few conserved interactions.

Coding docking poses using fingerprint representations offers also the possibility to visualize the distribution of the poses in high-dimensional descriptor space using projection methods to lower dimensions. One popular method is the self-organizing map (SOM)[37] that provides a nonlinear topology preserving projection into 2-dimensional space. Based on small molecule descriptor vectors SOMs have already been applied successfully, e.g., for the analysis of small molecule activity,[38] selectivity,[39] or for the selection of diverse subsets of ligands.[15]

Our algorithm was evaluated using several cocrystal structures from four different targets: cyclin dependent kinase 2 (CDK2), dihydrofolate reductase (DHFR), HIV-1 protease (HIVP), and thrombin (THR). All ligands of a target were docked into one receptor structure, and our approach was used to identify retrospectively the docking poses most similar to the crystal structure references.

MATERIALS AND METHODS

Data Sets. Protein structures were downloaded from the Brookhaven PDB server. We used data sets of protein ligands complexes that were previously reported for the evaluation of pharmacophore elucidation algorithms[40,41] or for the development of target-specific scoring functions.[36] We used the following data sets specified by PDB identifiers: cyclin-dependent kinase 2 (CDK2): 1KE6,[42] 1KE7,[42] 1KE8,[42] 1OIT,[43] 1OIY,[44] 1VYW,[45] 1VYZ[45] (see Figure 1); dihydrofolate reductase (DHFR): 1BOZ,[46] 1HFP,[47] 1DRF,[48] 1OHK,[49] 2DHF[50] (Figure 2); HIV-1 protease (HIVP): 1AJV,[51] 1AJX,[51] 1DMP,[52] 1HVR,[53] 1HWR,[54] 1PRO,[55] 1QBS[56] (Figure 3); thrombin (THR): 1C4V,[57] 1D6W,[57] 1D9I,[57] 1D4P,[58] 1DWD,[59] 1FPC,[60] 1TOM[61] (Figure 4). Pteridin nitrogens on

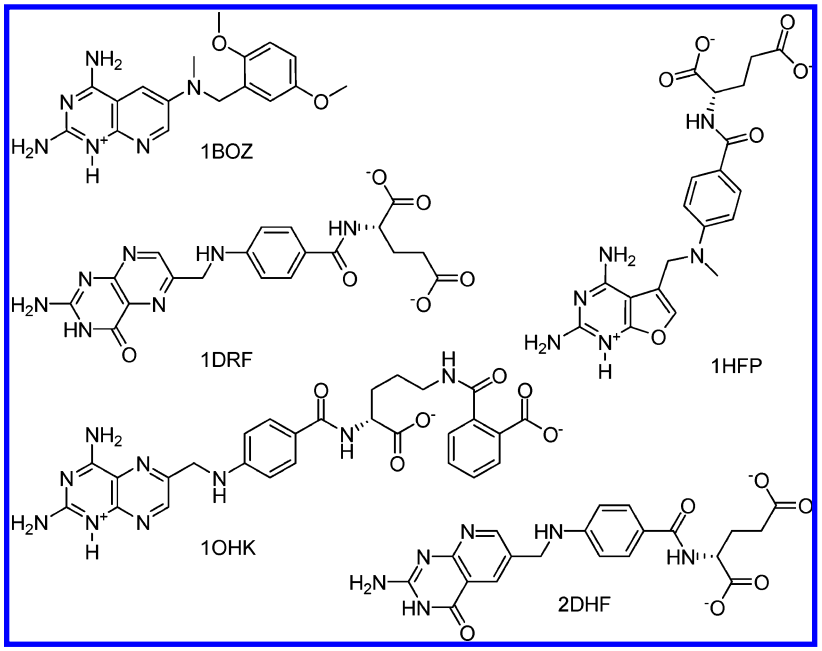

Figure 2. Ligands from the dihydrofolate reductase (DHFR) data set.

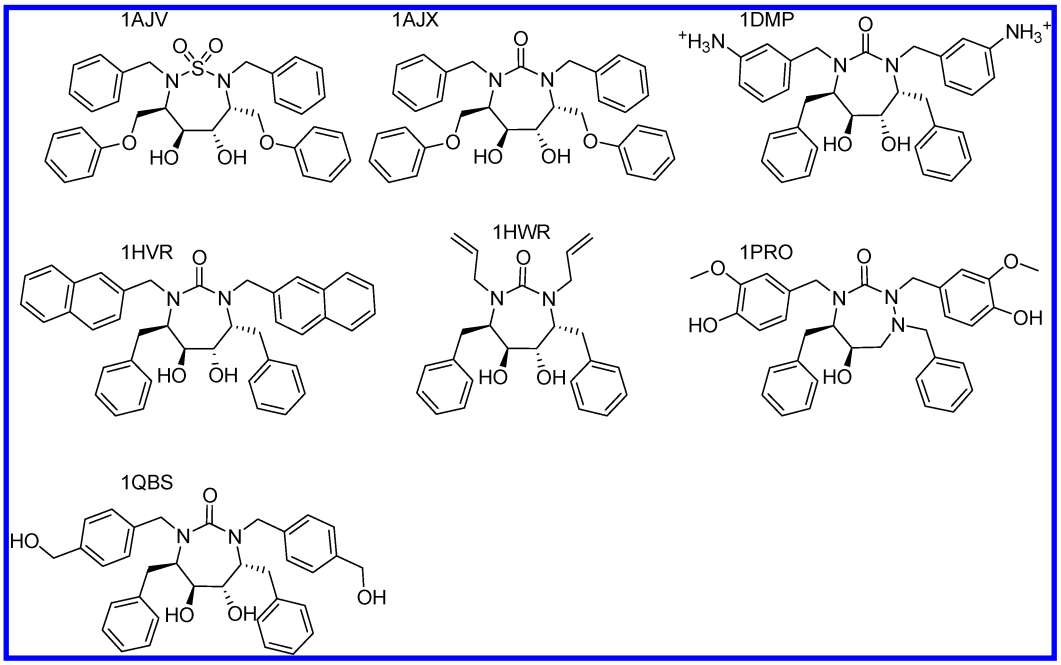

Figure 3. Ligands from the HIV-1 protease (HIVP) data set.

the DHFR ligands were protonated according to Richmond et al.[41] For docking all water molecules, ions and cofactors were removed from the structures.

Cross-Docking and Ranking with Traditional Scoring Functions. For docking we used FlexX[62,63] (version 1.20) in Sybyl 7.1.[64] Docking was performed using the default settings in FlexX. 100 poses were generated in each run to provide an extensive sampling of the binding site. To define the binding site we used a sphere of 15 Å around the ligand of the crystal structure. We used this rather large radius to prevent the possibility that a strict selection of the binding site already biased the generated poses toward the correct poses without using any scoring function. For each of the target classes, one protein structure was selected to serve as receptor in cross docking: CDK2, 1OIT; DHFR, 1DRF and 1OHK; HIVP, 1AJV; and THR, 1TOM. For the calculation of RMSD values all protein structures of a target class including their complexed ligands were aligned to the reference structure based on the backbone atoms of the protein, using Sybyl. The aligned ligands were subsequently used as references for the calculation of RMSD values of docked ligands. The HIV-1 protease inhibitors in our set (except for 1PRO) were found to be fully symmetric, i.e., the true binding mode of these inhibitors could be identified by two symmetric orientations of the docked ligands.[65] To take into account this symmetry, both correct orientations of the ligands were used as reference for the RMSD calculation. Therefore the representative protein active site residues of HIV-1 protease chain A were superimposed onto the corresponding residues of chain B, and the resulting transformation was applied to the ligands as well.

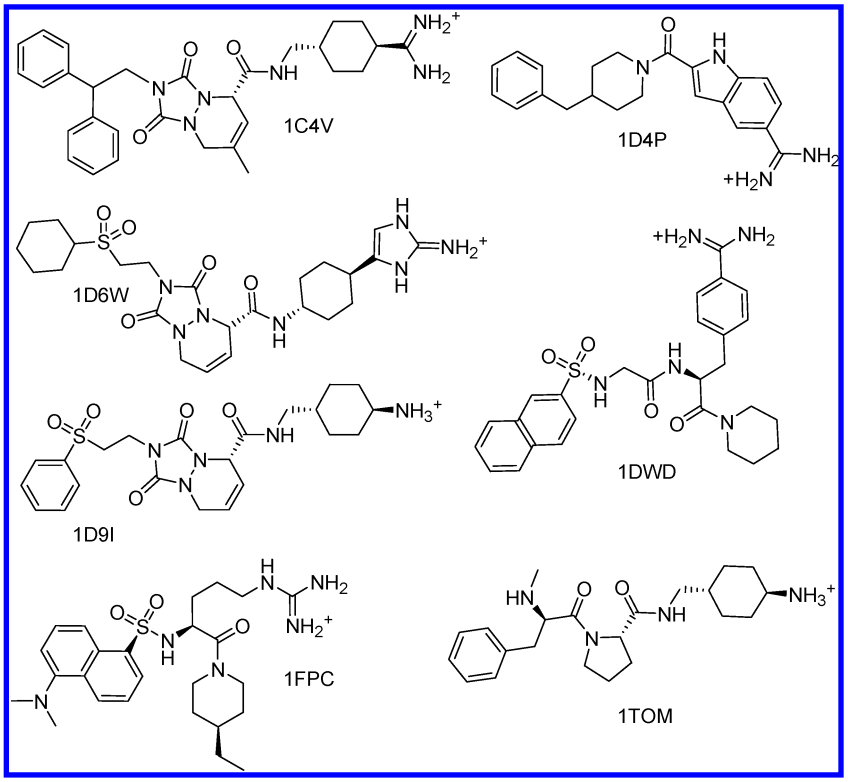

Figure 4. Ligands from the thrombin (THR) data set.

All docking poses were rescored using the scoring functions provided with the CScore module of Sybyl 7.1. All scoring functions (Chemscore,[66] Dock Score,[67] Gold Score,[68] and PMF Score[69]) were applied individually and incorporated into consensus scores including the FlexX scoring function. Two consensus scores were applied. The first one was a "rank sum" strategy that has found a widespread application in docking applications.[6,7] The idea of the "rank sum" strategy is to rank all molecules according to each individual scoring function and to use the sum of the rank positions of the poses as a score. This kind of consensus scoring function is independent of the scaling of different scoring functions and can therefore be applied for conceptually different scoring functions. The second consensus score was a voting strategy identical to the consensus score implemented in the CScore module in Sybyl,[64] further referred to as "vote rank". The idea of the "vote rank" strategy is that each scoring function votes for a pose to be a hit if the pose obtained a score in the top 50% of the score value range for all poses of a molecule. The number of votes for each pose finally serves as the consensus score. It often appears that more than a single pose obtained the maximum number of votes. In these situations an additional scoring was used to rank the molecules with identical "vote rank" scores. Two scoring functions were tested for that purpose: the FlexX score that already served for in the placement of the poses and the score obtained from the "sum rank" strategy.

Interaction Fingerprints. In the original publication of Deng et al.[27] seven binary flags were used to describe the interaction between the ligand and each residue of the receptor. The first three bits state whether a residue participates in an interaction and whether main chain atoms and/or side chain atoms are involved. The next two bits indicate whether hydrophobic and/or polar interactions were found, and the last two bits state the presence and absence of hydrogen bond donors and acceptors on the residues participating in hydrogen bonding to the ligand. However the authors stated in their article that only a single bit for the presence of interactions worked already comparably to the more complex fingerprints.[27] Therefore we decided to use a less complex fingerprint containing only 4-bit positions, stating whether there is an interaction, whether there is a hydrophobic interaction, whether there is a hydrogen bonding acceptor, and whether there is a hydrogen-bonding donor on the residue involved in a hydrogen bonding interaction.

Ligand receptor interactions were extracted from the matching files generated by FlexX (files ending with *_mat.log). These files contain all typed interactions that were found between the receptor and the docking pose and that were used for the FlexX scoring function. Receptor ligand interactions that were typed by FlexX as "h_don" − "h_acc" and "metal" − "metal_acc" were considered as hydrogen bonding interactions. Hydrophobic interactions were assigned for FlexX type interactions between "phenyl-_centers" and one of the following atom types "phenyl_ring", "ch3_phe", "amide", "ch", "ch2", "ch3", or "sulfur aro", in accordance to the definition of hydrophobic interactions in FlexX.

Maximum Common Binding Mode Scoring. The principle idea of the "maximum common binding mode" (MCBM) approach was that similar ligands with activity on the same receptor should have a similar set of interactions between ligand and receptor. The set of identical interactions should likely be large if the ligands were docked in near-native poses compared to random false positive poses. An overview over the algorithm for scoring is given in Figure 5.

The algorithm was based on the following steps: a) Docking *n* ligands, generating *m* poses per ligand. b) Interaction fingerprints were calculated for all docking poses.

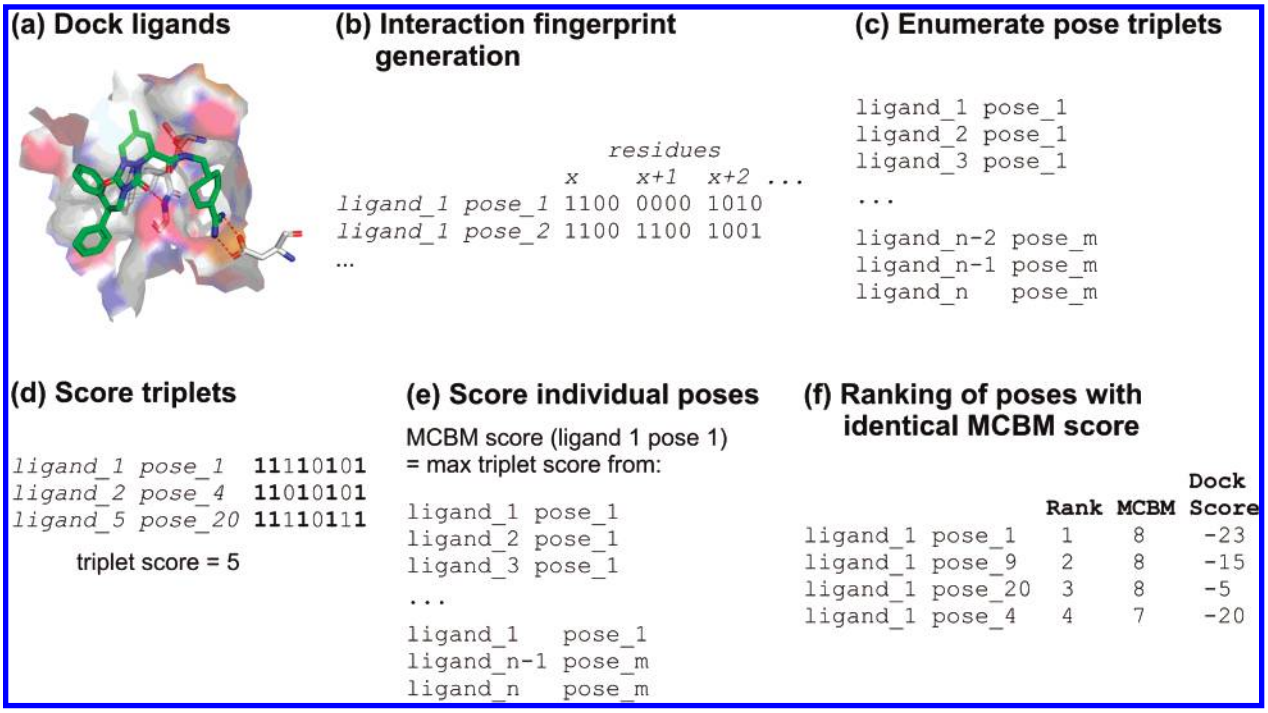

Figure 5. Overview of the calculation of the "maximum common binding mode" (MCBM) docking score: (a) Docking of *n* ligands, generating *m* poses per ligand. (b) Generation of interaction fingerprints, recording interactions of a ligand with each residue of the protein. (c) Enumeration of all triplets of poses composed of three different ligands. (d) Scoring of triplets by the number of fully conserved interactions. (e) Scoring individual poses (MCBM score) by the best triplet score of the respective pose. (f) Poses with identical MCBM score were further ranked by conventional docking scores.

c) All permutations of triplets of docking poses of three different ligands were generated. d) The triplets were ranked by the number of interactions (i.e., bits in the fingerprint) that were fully conserved over all three ligands. e) Each individual docking pose of each ligand was finally ranked by the highest number of interactions resulting from any of the triplets containing the particular docking pose. f) Poses with identical MCBM score were further ranked by conventional scoring functions.

Using triplets instead of any other fixed number of multiplets or a variable number of different ligands used for the docking experiment might seem at first a bit arbitrary. Several aspects can be considered for the selection of the set size. Based on our presumption that overlapping binding modes should indicate the correct binding mode, the reliability of the solution should increase with larger amounts of different ligands. On the other hand, it cannot be guaranteed for any docking algorithm to generate near-native docking solutions at all for a particular ligand and receptor pair.[2] In such a situation a multiplet that must contain poses from all docked ligands could not retrieve the correct poses of the other ligands even if these were docked correctly. The risk for such a scenario is reduced using only triplets. It is also a common situation that not all ligands share exactly the same set of interactions. This might lead to overpessimistic numbers of conserved interactions if conservation is counted over many different ligands. For the evaluation of the approach doublets and quartets of ligands were also considered.

One straightforward approach would have been to search for a maximum similar binding mode using, e.g., the Tanimoto coefficient, like it is done with the virtual screening applications of structure interaction fingerprints.[27,30,31] However in our setup the most likely way to obtain a perfectly similar interaction pattern (Tanimoto coefficient = 1) is having a single fully conserved interaction, which might occur for poses that hardly interact with the receptor at all. This effect is circumvented in virtual screening with interaction fingerprint scoring since reference ligands from crystal structures are used that predefine the "ideal" number of interactions. Therefore the use of maximum common interactions is the more reasonable measure in our approach.

Like with the "vote rank" strategy we found that in many cases a small set of poses was identified having identical MCBM scores, rather than a single best scoring solution. Therefore the FlexX score and the "sum rank" score were tested for further ranking the solution with identical MCBM scores.

Self-Organizing Maps (SOM). The SOM is a nonlinear topology preserving projection method that allows for the visualization of the distributions of objects in an original high dimensional space in a low (here 2D) dimensional space.[37] The original data space is represented by a set of prototype vectors (neurons), and each original data point is assigned to its nearest neuron. The most common usage of SOM is to use very few neurons resulting in a procedure similar to k-Means clustering, where each neuron corresponds to a cluster of data points. In contrast Emergent SOM[70] use a large number of neurons and provide a more detailed view into the similarity structure of the high dimensional space. In addition to a low dimensional projection preserving the topology of the input space, the original high dimensional distances can be visualized with the U-Matrix[70] display.

For the calculation of SOMs the ESOM Tools[71] software was used. Toroidal two-dimensional SOMs were calculated with 50 * 82 neurons and default settings. For the training of the map the distance or similarity of a training pattern to the neurons of the map had to be computed. To be consistent

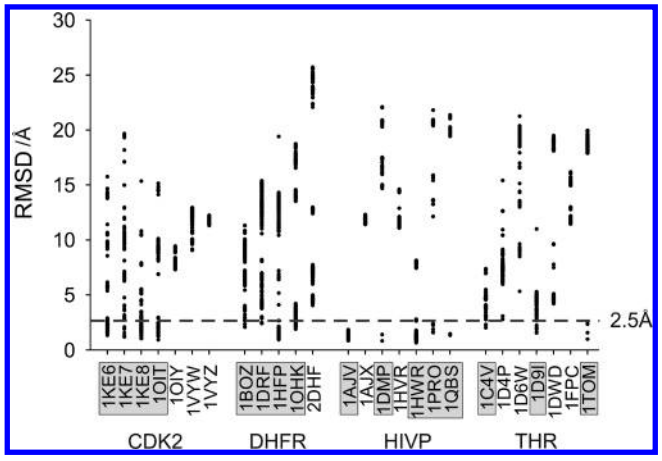

Figure 6. Distribution of RMSDs of ligands in the cross docking experiments. For ligands from crystal structures with names highlighted in gray docking poses with RMSD better than 2.5 Å were generated.

with the MCBM scoring function we used a similar similarity function for the calculation of the map. Since data points are compared with SOM neurons in a pairwise manner, the number of common bits set between a docking pose and a SOM neuron was used as similarity criterion.

The background of the SOM was colored according to the U-matrix,[70] that gives the local distance structure of the data in the SOM. The U-matrix is calculated from the sum of the distances to all immediate neighbors normalized by the largest occurring distance. The U-matrix is then displayed as a map of a landscape. Clearly defined borders between clusters, where large distances in data space are present, are visualized in the form of high mountains. Smaller intracluster distances or borders of overlapping clusters form smaller hills. Homogeneous regions of data space are placed in flat valleys. In this way clusters of data can be identified using the U-matrix.

The final maps were used to visualize two aspects of the docking poses: the distribution of the docking results for different molecules and the distribution of the rmsd of the poses. SOM neurons with at least one data point assigned were represented as colored circles. If poses of different molecules were assigned to the same neuron, then the neuron was colored by the molecule with the most poses assigned to the particular neuron. For the rmsd projection the lowest value was used that was found for the poses assigned to a neuron.

RESULTS AND DISCUSSION

Cross-Docking. All ligands of each target class were docked into one representative protein structure for the respective target. An overview over the results and the resulting RMSD values of the docked poses is given in Figure 6. Docking solutions were considered successful if the pose was found within an RMSD < 2.5 Å to the reference pose from the crystal structure. For all four data sets we found at least one ligand that was not docked with a RMSD < 2.5 Å. The number of successful dockings ranged from 5 of 7 ligands for HIV-1 protease over 4 of 7 for CDK2 to 3 of 7 for thrombin. For DHFR we were not able to find more than two successfully docked ligands for a single structure. Using triplets of ligands for the calculation of the MCBM score, correct poses would not be able to be found, in these cases. Therefore we merged the results of two reference protein structures. In this way the double amount of docking solutions was used compared to the other proteins. Using this setup 4 of 5 DHFR ligands were docked successfully. The variance of RMSDs for the ligands was large for most of the ligands. Only for the HIV-1 protease inhibitor from the structure 1AJV (the reference for HIV-1 protease) no pose with RMSD > 2.5 Å was generated. For all other ligands worst solutions with at least an RMSD of 7.4 Å (1C4V) were generated, indicating a large variance in the obtained docking poses that might render the selection of near-native docking poses a complicated task. It has to be stressed that ligands for which no poses were generated with RMSD < 2.5 Å were not removed from the data set. Therefore these ligands were also considered for the triplets, and the MCBM score had to discriminate these ligands from ligands that had poses with RMSD < 2.5 Å.

A comparison of the cocrystal structure binding modes compared to the binding modes resulting from the docking calculations is given in Figure 7 for the thrombin inhibitors. Figure 7a,c shows an overlay of the cocrystal structures of the thrombin inhibitors, compared to the docking solutions generated for the same ligands (Figure 7b,d). While the binding site of the cocrystallized ligands was well defined in the way that all ligands shared the same binding site (Figure 7a) the docking solutions were spread all over the surface within the restricted area of 15 Å around the native binding site (Figure 7b). Many of the clusters of the binding modes on the surface only consisted of docking solutions of one ligand. Figure 7c,d shows the ligands from the viewpoint of the receptor. The set of spatially aligned conserved functional groups (e.g., the central carbonyl) within the set of ligands from the cocrystal structures confirms the presence of similar binding modes for the thrombin inhibitors (Figure 7c). Within the docking solutions the same functional groups were found to be conserved between ligands with correctly predicted binding mode (Figure 7d); however, it appeared that this was not necessarily true for the false positives in the same figure. These findings support the hypothesis that the similarity of the binding modes of similar ligands might be useful to discriminate true from false positive binding modes, obtained from small molecule docking. Furthermore searching for conserved interactions or potential interaction points within different ligands closely resembles the approach commonly used for pharmacophore elucidation.[21−23]

Maximum Common Binding Mode (MCBM) Scoring − Ranking of Docking Poses. All generated docking poses from the entire range of RMSDs were ranked by the MCBM score using single- and four-bit fingerprints. In most cases more than a single pose per ligand obtained the maximum (best) score, i.e., had the maximum number of interaction identical to docking poses of at least two other ligands. To end up with a fully ranked list of poses, all poses having the same MCBM score were further ranked using conventional docking scoring functions. However such an approach has the risk to result in a scenario where the majority of poses has the same MCBM score, and the final ranking of poses would be dominated by the conventional scoring function. In this case the evaluation would not be valid. To ensure that this scenario was not true for the targets and ligands in the current study, the RMSD values of all poses that obtained

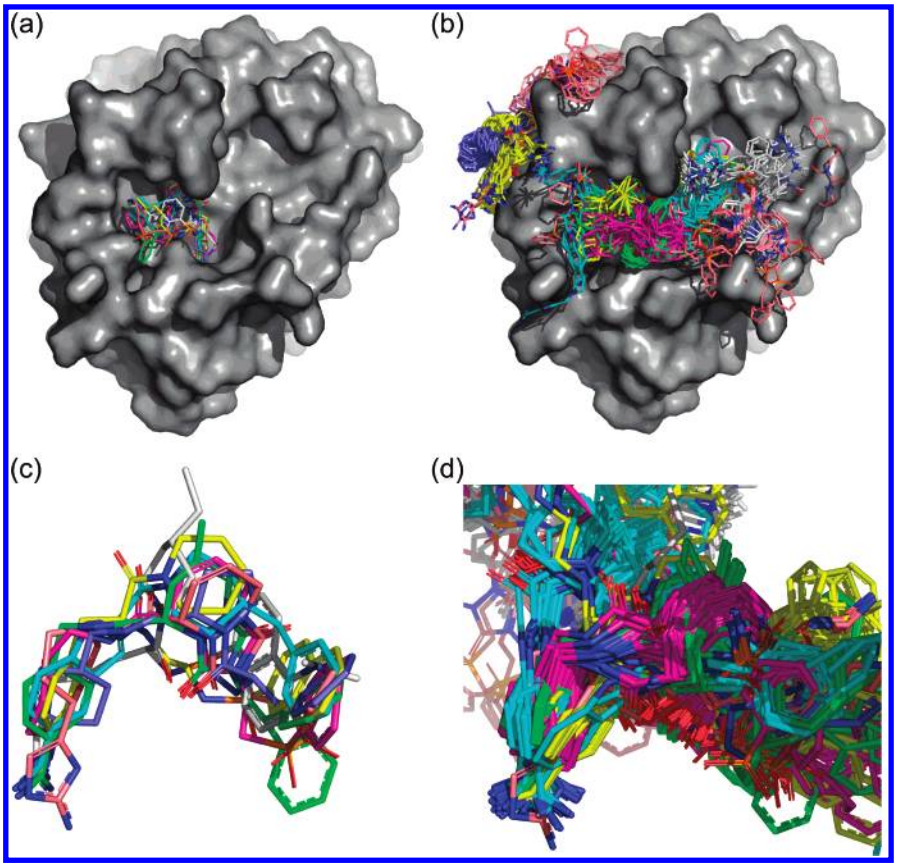

Figure 7. Comparison of overlaid crystal structure conformations (a,c) and docking solutions (b,d) of thrombin inhibitors on the surface of the protein (a,b) and shown from the perspective of within the binding site (c,d). (Atoms colors: blue = nitrogen, red = oxygen, other colors (= carbon) were used to discriminate different inhibitors: green = ligand from crystal structure with pdb code 1C4V; cyan = 1D4P; salmon = 1D6W; magenta = 1D9I; yellow = 1DWD; gray = 1FPC; blue = 1TOM).

the maximum MCBM score were investigated irrespective of any further ranking using conventional scoring (Figure 8). The number of poses having the maximum score varied from 1 for 1DRF and 1HFP to 64 for 1HWR (one-bit fingerprints) or from 1 for 1BOZ to 52 for 1KE8 (four-bit fingerprint). Only ligands that had poses with an RMSD < 2.5 Å were considered here. Using one-bit fingerprints, for all ligands except for 1DRF and 1HFP best scoring poses were found that had an RMSD < 2.5 Å. Furthermore in most cases the majority of the best scoring poses were found with RMSD < 2.5 Å, but still in 5 cases (1BOZ, 1DRF, 1HFP, 1HWR, 1C4V) best scoring solutions were also present in a clear non-native binding mode (RMSD > 5 Å). Using four-bit fingerprints, similar results were observed. However not a single correct result was retrieved for DHFR anymore, but strongly false poses with (with RMSD > 5 Å) were not present anymore for all other classes. These results indicated that the MCBM scoring function already has a discriminative power to retrieve near-native docking poses, and the obtained results were not dominated by the docking scoring function used for final ranking of the poses.

As a benchmark for the MCBM docking scoring function we applied all docking scoring functions available in the CScore module of Sybyl.[43] Furthermore two consensus scoring strategies were used: "vote rank" and "sum rank". Retrieval of correct poses with RMSD < 2.5 Å was compared for the best scoring poses as well as the best pose among the best five scoring poses (Figure 9).

Considering only the RMSD of the top scoring poses (Figure 9a), the individual CScore scoring functions performed worse in the retrieval of correct poses over all four target classes compared to the "sum rank" and "vote rank" consensus scoring strategies. This observation is in agreement with other studies[5−7] that already found that consensus scoring performs superior to individual scoring functions. The "sum rank" strategy performed best in the identification of near-native poses: for 12 out of 16 ligands correct poses were found. The "vote rank" strategy and the Gold scoring function each retrieved 10 correct solutions. PMF performed worst with only 5 correct solutions. In comparison the retrieval rates for the different tastes of the MCBM scoring function ranged from 12 correct hits for the MCBM with one-bit fingerprints and "sum rank" (MCBM 1BSR) to 9 correct poses for the version with a four-bit fingerprint and the FlexX score (MCBM 4BF). The better performance of the "sum rank" rescoring compared to FlexX rescoring is not surprising due to the general better performance of the former scoring function. However it must be noted that using the FlexX score in combination with MCBM already resulted in 9 and 10 hits (for one- and four-bit fingerprints, respectively) and thus performed better than all of the CScore scoring functions except for Gold score. Also the one-bit fingerprint at least performed equally well compared to the four-bit version. A similar effect was already observed for structure interaction fingerprint-based virtual screening.[27] If two ligands are able to form a set of well defined interactions with the same set of residues of a receptor, then it might be already an indication that the binding mode is reasonable, even if the interactions are not identical in nature.

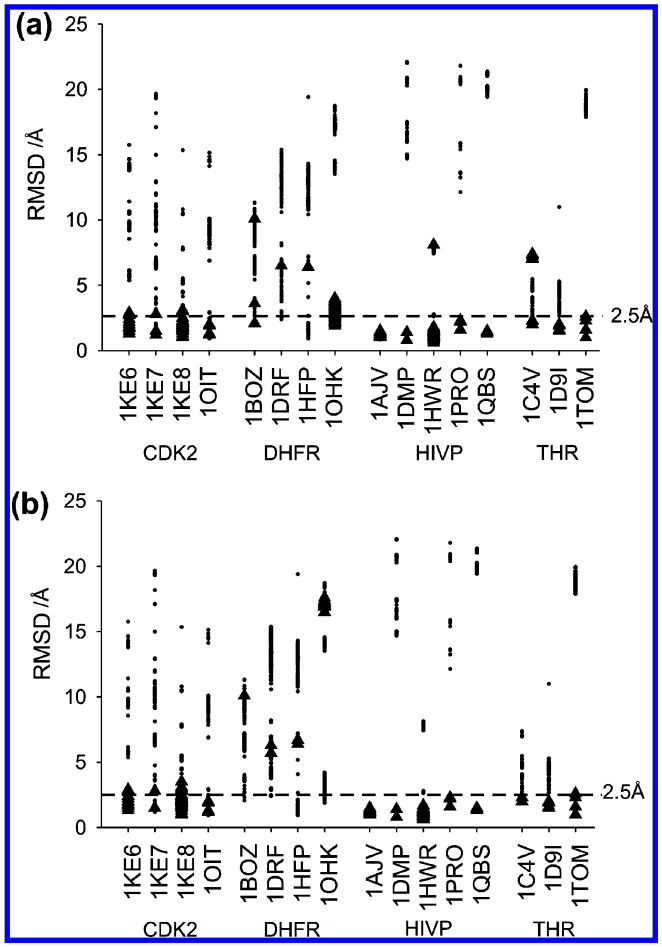

Figure 8. Distribution of RMSDs of docking poses having the maximum MCBM score with (a) 1-bit fingerprints or (b) 4-bit fingerprints (black triangles) relative to all docking poses for the respective ligand (black dots).

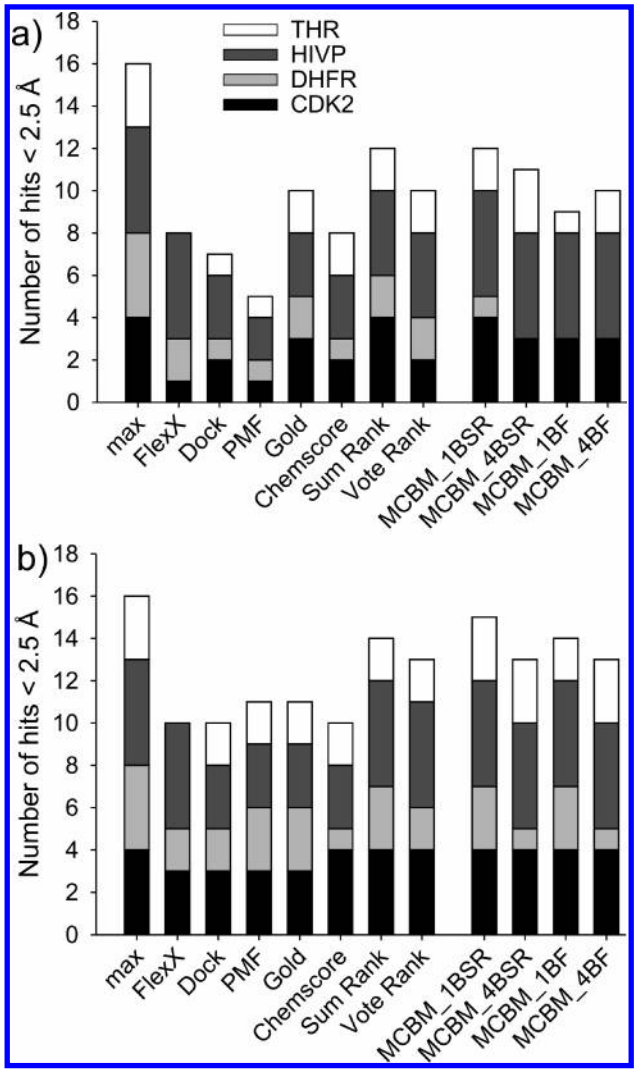

Figure 9. Comparison of MCBM with established docking scoring functions. Retrieval of poses with RMSD < 2.5 Å using the single best scoring pose (a) or the best RMSD among the five best scoring poses (b). MCBM was calculated using one bit (MCBM 1BSR and 1BF) and four bits (MCBM 4BSR and 4BF). Hits with identical MCBM score were further ranked by FlexX score (MCBM 1BF and 4BF) or by the "sum rank" score (MCBM 1BSR and 4BSR). "Max" gives the number of ligands that had poses with RMSD < 2.5 Å.

Considering the best RMSD among the five top scoring docking poses per ligand (Figure 9b) did not change the relative performance between individual scoring functions (10−11 correct poses), consensus scores (13−14), and the MCBM scores (13−15). The most prominent effect was that the large differences in the performance of the individual scoring functions decreased to a similar level of 10−11 hits. The better performance of the individual scoring functions resulted also in a better performance of FlexX rescoring compared to "sum rank" rescoring, though the latter was still found to be slightly better. The one-bit fingerprint also performed slightly better compared to the four-bit version.

A detailed overview over the obtained RMSD values of the best scoring poses for the individual ligands is given in Figure 10. These results demonstrate that MCBM not only performs comparable to the compared scoring functions but also complements these. For CDK2, HIVP, and THR the MCBM were found among the best scoring functions and obtained correct poses for ligands that were not predicted correctly by the majority of other scoring functions (1DMP, 1C4V). Considering the five best poses per ligand, this effect was less pronounced, because more ligands were retrieved by all scoring functions (see the Supporting Information). For DHFR only a single ligand was identified using MCBM_1BSR, while the other MCBI versions did not identify a single well docked pose. Investigating the crystal structures with cocrystallized ligands of DHFR we found that the inhibitors were interacting directly with the cofactor NADPH within the binding site of DHFR. Since all cofactors were removed prior to the docking experiments, a large additional binding site was provided in these cases, while at the same time interactions with the cofactor were not able to be formed anymore. Still correct poses were generated for all ligands expect 2DHF, and inspection of the best 5 poses lead to the identification of correct poses.

These results show that the MCBM scores performed at least equally well compared to the best consensus scoring function ("sum rank") composed of individual scoring functions. This is especially remarkable since the MCBM score quantifies rather the consensus of conserved interaction patterns than to provide a quantification of the binding constant, like it is done in traditional docking scoring functions.

To investigate the influence of the selection of triplets instead of other potential multimers of ligand docking poses an analysis using doublets and quartets of ligands was also

	FlexX	Dock	PMF	Gold	Chem-score	Sum Rank	Vote Rank	MCBM 1B Sum	MCBM 4B Sum	MCBM 1B FlexX	MCBM 4B FlexX
1KE6	2.8	2.7	14.1	1.7	2.6	2.4	2.4	2.4	2.4	2.8	2.8
1KE7	3.1	1.3	13.0	1.2	1.3	1.2	1.2	1.2	2.8	1.2	1.5
1KE8	1.1	1.1	7.9	1.9	1.5	1.1	1.1	1.1	1.1	1.1	1.1
1OIT	8.3	2.5	1.2	2.9	8.9	1.8	1.9	1.9	1.9	1.2	1.2
1BOZ	10.1	5.4	3.6	2.1	3.8	2.1	2.1	2.1	10.1	10.1	10.1
1DRF	2.4	3.0	2.5	3.0	3.0	3.0	3.0	6.5	6.3	6.5	6.3
1HFP	1.2	1.2	1.0	1.7	1.3	1.2	1.2	6.4	6.7	6.4	6.7
1OHK	3.7	16.7	3.5	3.6	18.5	3.1	3.1	3.1	16.9	3.1	16.9
1AJV	1.1	1.8	1.0	1.8	1.6	1.3	1.7	1.3	1.3	1.1	1.1
1DMP	0.8	20.3	17.5	20.7	20.3	20.8	20.7	0.8	0.8	0.8	0.8
1HWR	0.6	1.0	1.2	1.4	1.1	1.1	1.1	1.1	1.1	0.7	0.7
1PRO	1.6	1.6	13.2	1.6	1.9	1.6	1.6	1.6	1.6	1.6	1.6
1QBS	1.3	20.3	21.2	20.1	21.0	1.3	1.3	1.3	1.3	1.3	1.3
1C4V	4.7	3.8	4.0	3.8	7.0	7.0	7.0	7.0	2.3	7.0	2.0
1D9I	4.3	2.0	2.0	2.2	1.5	2.2	2.2	1.5	1.5	1.6	1.5
1TOM	18.2	2.6	19.9	2.3	2.3	2.3	2.3	2.3	2.3	2.6	2.6

Figure 10. RMSD values of the best docking poses according to the different docking scoring functions compared. Poses with an RMSD below 2.5 Å were considered to be correct (highlighted in gray).

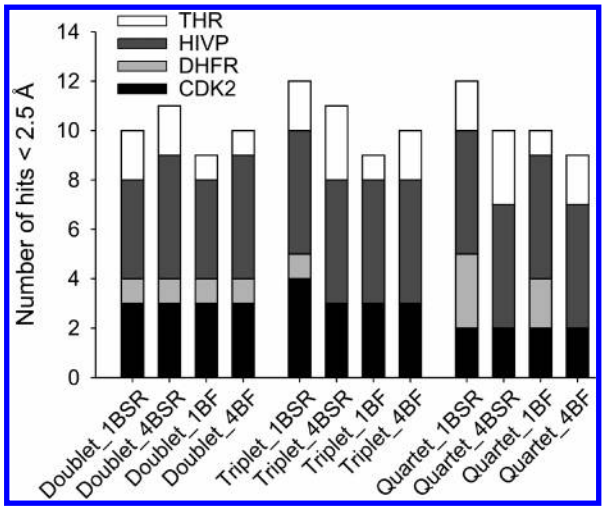

Figure 11. Comparison of doublets, triplets, and quartets of different ligands for the determination of commom binding mode for docking scoring. The number of best scoring hits were considered.

performed for comparison (Figure 11). No large differences were observed between the different setups with respect to the observed numbers of ligands with correctly predicted docking poses. The best result with a number of 12 retrieved hits was obtained using triplets or quartets in combination with one-bit fingerprints and "sum rank" scoring. Quartets solely identified three correct poses for DHFR, however, at the tradeoff of getting fewer correct poses for the remaining classes. Therefore triplets performed best for three out of the four target classes and were further considered as the generally best performing setup for the algorithm. For targets where triplets performed badly a consideration of quartets might be worth.

Maximum Common Binding Mode (MCBM) Scoring − Ranking of Ligands. Though we have demonstrated that the MCBM score is able to discriminate near-native poses from false positive poses of individual ligands, it is not clear whether the score can be applied to rank different ligands with respect to the quality of their predicted binding mode. A possible scenario is that more than one cluster of high scoring ligands was predicted, and these clusters were composed of nonidentical sets of ligands. This might occur for example in situations where only incorrect binding poses were generated for a set of ligands that cluster together in a false positive cluster. In such a case the ligands of the correct cluster should have higher scores compared to the ligands of the incorrect cluster. In other words the question was whether ligands with a high MCBM score tend to have better RMSD values compared to ligands with lower RMSD values. Figure 12 shows the relation of the best MCBM 1B score (of the best "sum rank" reranked pose) for each individual ligand docked in this study and the respective RMSD value of this pose. For the analysis, the resulting plots were further subdivided into four quadrants. These were defined by the RMSD threshold of 2 Å that was used to discriminate correct from incorrect poses and by a line separating the best MCBM scoring ligands from ligands with lower score. An ideal scoring function would calculate high scores for correctly docked ligands and low scores for incorrectly docked ligands. This should be reflected by high fractions of correctly predicted ligands in the lower right quadrant and by high fractions of falsely predicted ligands in the upper left quadrant.

For CDK2 an ideal distribution was observed. All correctly docked ligands had an identical top score of 12, placing these ligands in the lower right quadrant. All false positives obtained lower scores between 3 and 9, placing these ligands in the upper left quadrant. For HIVP an almost ideal performance was found. A single ligand (1HWR) was found in the lower left quadrant with a nonoptimal score but a good RMSD. Since this ligand would be placed in the binding site similar to the top scoring ligands, 1HWR would not be missed in this example. THR performed also almost ideally, with two correctly predicted ligands 1D9I and 1TOM placed

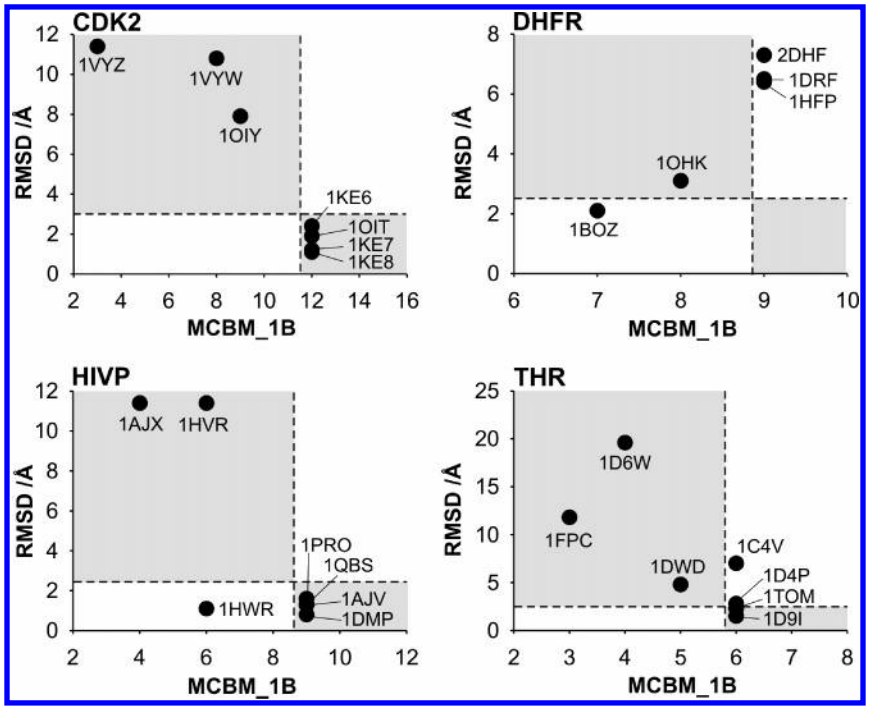

Figure 12. Ranking ligands using MCBM. Relationship between MCBM scores (the number of conserved interactions) versus the RMSD of the individual ligands of one target. The RMSD is given for the best scoring pose according to MCBM_1BSR. Each graph is divided into four quadrants: first by the RMSD threshold (2.5 Å) used to discriminate between correct and incorrect poses and second by a line separating the top MCBM_1B score from lower scores. Ideally all ligands would be placed into the quadrants highlighted in gray.

into the lower right quadrant. However two ligands (1D4P and 1C4V) were also found in the upper right quadrant, which means that they also obtained the maximum MCBM score but were not considered to be correct by RMSD. While the predicted pose of 1C4V had a real false binding mode (still within the correct binding site), the predicted pose of 1D4P (with an RMSD of 2.8 Å only slightly above the threshold of 2.5 Å) was seemingly still similar from the point of interaction patterns. Only in the case of DHFR the ranking did not work at all. The ligands with the worst RMSD (2DHF, 1DRF, 1HFP) turned out to have the best MCBM score, and the best ligand (1BOZ) had the worst docking score. This observation might be explained with the absence of the cofactor in the crystal structure. In general, however, the MCBM score seems to be well suited to rank not only poses of individual ligands but also the best poses of different ligands. This ability is crucial when different clusters of high scoring ligands have to be ranked. Comparable results were obtained using four-bit fingerprints (see the Supporting Information).

Ligand Similarity Analysis. The basic assumption of the MCBM approach, that similar molecules have similar binding modes, also applies a restriction to the molecules that could be scored using the methods. The method might not work for molecules that are very dissimilar but occupy the same part of the binding pocket and should not work for molecules that have different binding modes in the sense that molecules bind in parts of the binding pockets that only partly or hardly overlap or in separate binding sites.

To estimate to the ability of the MCBM approach to score diverse molecules, average similarities between the ligands of each of the targets used in this study were calculated. MACCS keys[72] as implemented in PipelinePilot[73] in combination with the Tanimoto coefficient[74] were used for this purpose. The averages over all ligands of each target ranged from 0.55 for CDK2 to 0.79 for HIVP (Table 1). Restricting the analysis only to ligands that were docked successfully narrowed the range of similarities down to values from 0.67 for CDK2 to 0.74 for HIVP. The highest observed similarity of the HIVP ligands correlates well with the best performance in MCBM scoring of this target. However the bad performance of DHFR seemed not to be attributed to a lower similarity of the ligands compared to the other classes, supporting the interpretation that the missing cofactor was the major problem in the scoring of DHFR ligands.

Table 1. Average Similarity of the Ligands for Each Target Class

target class	average MACCS similarity
All Ligands	
CDK2	0.55 ± 0.13
DHFR	0.68 ± 0.12
HIVP	0.79 ± 0.12
THR	0.63 ± 0.11
Correctly Predicted Ligands	
CDK2 (1KE6, 1KE7, 1KE8, 1OIT)	0.67 ± 0.08
HIVP (1AJV, 1DMP, 1HWR, 1PRO, 1QBS)	0.74 ± 0.12
THR (1C4V, 1D9I, 1TOM)	0.64 ± 0.07

Defining fingerprint similarity thresholds for similar biological activity of molecules is challenging and depends both on the applied fingerprint and the target under consideration.[75] For MACCS similarity a threshold of ca. 0.8 was proposed for optimal retrieval of molecules with similar activity in ligand-based virtual screening.[76] This threshold seemed to be well suited to retrieve a series of close analogues in databases. A threshold of ca. 0.7 was proposed for the retrieval of structurally diverse molecules with similar activity.[77] According to these thresholds only the average similarity of the HIVP ligands indicates a series of close analogues. The average similarities of the well docked ligands from CDK2 (0.67) and THR (0.64) were in the range

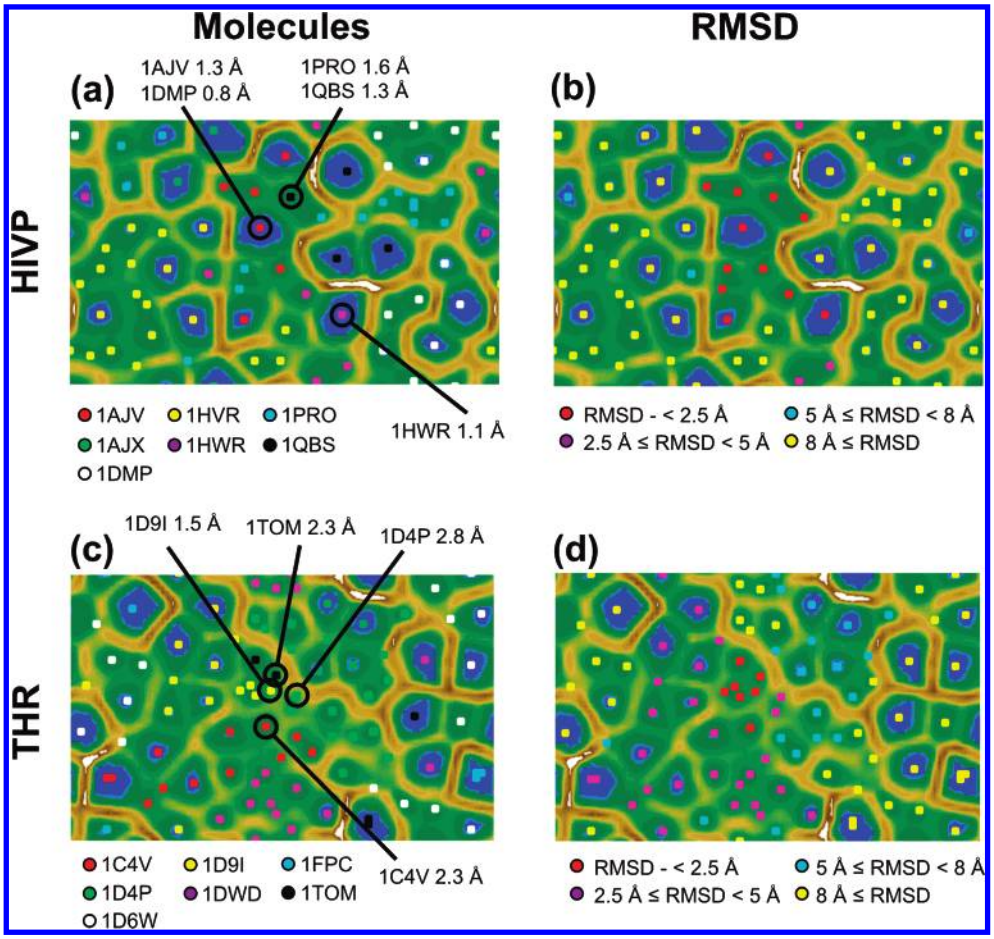

Figure 13. SOM projection of the HIVP (a,b) and THR (c,d) ligand docking results. Docking poses were coded using one-bit interaction fingerprints. Molecule color scheme (a,c): neurons with at least one docking pose assigned are colored according to the ligand molecule with the majority of poses assigned. RMSD color scheme (b,d): neurons with at least one docking pose assigned are colored according to the lowest found RMSD values range. The background is colored according to the U-matrix. Valleys denote clusters of similar docking solutions.

of molecules that have only moderately similar chemical structures. Therefore the similarity analysis indicates that it might be advantageous to use ligands with a higher similarity (in the range of 0.8 like in HIVP); however, it was also demonstrated that MCBM scoring also worked well with less similar ligands with similarities between 0.6 and 0.7 (like in CDK2 and THR).

SOM Analysis of the Distribution of Docking Poses. Coding docking poses using fingerprint representations offers also the possibility to visualize the distribution of the poses in high-dimensional descriptor space in a two-dimensional space using self-organizing maps (SOM). SOM projections are shown for THR and HIVP (Figure 13). The background of the maps is colored according to the U-matrix that gives the local distance structure of the maps. The colors of the background resemble the coloring of geographical maps. Accordingly blue and green regions (valleys) that are surrounded by brown and white regions (hills and mountains) denote clusters of similar objects. The neurons of the map that contained as least one ligand pose were colored either according to the ligand (Figure 13a) or the best RMSD that was achieved by the ligands in the neuron (Figure 13b). Comparing the ligand-composition of the SOM-generated clusters (Figure 13a,c), it is apparent that most clusters contained only poses of one particular ligand, and only a few clusters contained poses of more than one ligand. For both HIVP and THR the poses with the best MCBM_1BSR score of the ligands for which poses were generated with RMSD < 2.5 Å were highlighted in the map (Figure 13a,c). All highlighted poses of HIVP were found in the same cluster, and this cluster was found to be the largest clusters in the maps composed of multiple ligands. Similar results were found for THR. The best scoring poses (according to MCBM_1BSR) for the correctly predicted ligands 1D9I and 1TOM were highlighted in the map (Figure 13c). Also shown is the best predicted pose for 1D4P that failed the RMSD threshold only slightly. Not only this pose is in the same cluster with the two correct ligands but also the three poses form the interface of three sets of poses of the three ligands. The first correctly predicted pose for 1C4V (on rank 5) is also highlighted and is also placed near the interface to the other ligands. Indeed this area was the only area on the THR map where similar poses were found for four different ligands. Coloring the neurons according to the best found RMSD of the ligands in the respective neurons (Figure 13b,d) verified that the clusters that were composed of multiple ligands (containing the highlighted poses) had RMSD values mainly below 2.5 Å in case of HIVP. For THR RMSD values below 2.5 Å were exclusively found in the region where poses of four different ligands were found to be very similar to each other. The remainder of the SOM cluster also contained many poses with RMSD values between 2.5 Å and 5 Å. This indicates that for THR there was a continuous distribution of docking poses from good to bad solutions.

For HIVP the difference between good and bad poses seemed to be pronounced. For both HIVP and THR all clusters except for the ones with the highlighted structures had poses with RMSD values larger than 5 Å. Comparable results were obtained using four-bit fingerprints (see the Supporting Information). The results of the SOM analysis support the concept that similar binding patterns between multiple ligands were indeed only formed in correct or near-correct poses, whereas false poses seemed to form clusters preferentially with other poses of the same ligand.

Outlook. The proposed docking scoring function might serve as an additional scoring function complementing traditional scoring functions. The fact that it is not conceptually related with conventionally used force-field-based, empirical, or knowledge-based scoring functions provides an orthogonal view on generated docking poses. Having identified a set of reliable docking poses, these poses might then be used as references for scoring methods that are able to incorporate information from reference ligands like structure interaction fingerprint scoring, 3D similarity to known ligands, as the basis for pharmacophore constraints or as a training set for target-biased scoring functions. Though we have evaluated our approach only for known active ligands the ability to rank ligands with respect to the RMSD of their best binding mode indicates the applicability of the approach to cluster docking poses of virtual screening results to identify prominent binding modes that were not present among known active ligands. Analysis of docking poses with self-organizing maps reveals the underlying distribution of the binding modes and might also serve as a means for visual selection of promising clusters of ligands with a high similarity of binding modes.

CONCLUSION

We have proposed a novel scoring function for small molecule docking that was designed for the prediction of binding modes, integrating information from multiple known active molecules. The scoring function is based on the assumption that binding modes should be more conserved between multiple ligands of a receptor in near-native docking poses of the ligands compared to false positive poses. Therefore we used the number of identical interactions found between docking poses of multiple ligands as scoring function. Though this approach differs substantially from commonly used force-field-based, empirical, or knowledge-based scoring functions, our algorithm performed better in predicting correct docking poses than the CScore scoring functions from Sybyl and was comparable to consensus scoring schemes thereof. Furthermore we have shown that our scoring function not only was able to identify the best pose for individual ligands but also was able to separate well docked ligands from ligands with incorrect poses. Analysis of the similarity of the ligands that were docked successfully indicated that the MCBM approach works not only best for highly similar ligands but also well for ligands that are less similar to each other. Analysis of the similarity of the docking poses using self-organizing maps verified that clusters composed of multiple ligands were likely to contain near-native poses and clusters with only a single ligand were likely to contain incorrect poses.

The use of binding mode similarity as docking scoring function induces a strong bias toward binding modes that are compatible with the majority of ligands applied in a study. This might prevent the identification of alternate binding modes for similar ligands, which might be identified using traditional scoring functions. In cases where alternate binding modes are not present the MCBM approach enables the integration of additional information into the scoring calculation that is contained in sets of similar active ligands. This is currently not possible in the majority of docking scoring functions.

ACKNOWLEDGMENT

We would like to thank Dr. Tanja Weil for many stimulating discussions and valuable comments on the manuscript.

Supporting Information Available: Table of the best RMSD values among the five best scoring poses equivalent to Figure 10 with RMSD values for the best scoring poses, figure like Figure 12 using four-bit fingerprints for the ranking of molecules instead of one-bit fingerprints, and results of a SOM analysis like in Figure 13 using four-bit fingerprints instead of one-bit fingerprints. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) 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.
(2) Warren, G. L., et al. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912−5931.
(3) 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.
(4) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2644−2676.
(5) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100−5109.
(6) 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.
(7) Wang, R.; Wang, S. How does consensus scoring work for virtual library screening? An idealized computer experiment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422−1426.
(8) Yang, J. M.; Chen, Y. F.; Shen, T. W.; Kristal, B. S.; Hsu, D. F. Consensus scoring criteria for improving enrichment in virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 1134−1146.
(9) Baber, J. C.; Shirley, W. A.; Gao, Y.; Feher, M. The use of consensus scoring in ligand-based virtual screening. *J. Chem. Inf. Model.* **2006**, *46*, 277−288.
(10) Zhang, Q.; Muegge, I. Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring. *J. Med. Chem.* **2006**, *49*, 1536−1548.
(11) Salim, N.; Holliday, J.; Willett, P. Combination of fingerprint-based similarity coefficients using data fusion. *J Chem. Inf. Comput. Sci.* **2003**, *43*, 435−342.
(12) Fechner, U.; Schneider, G. Evaluation of distance metrics for ligand-based similarity searching. *ChemBioChem* **2004**, *5*, 538−540.
(13) So, S.-S.; Karplus, M. Evolutionary optimization in quantitative structure-activity relationship: an application of genetic neural networks. *J. Med. Chem.* **1996**, *39*, 1521−1530.
(14) Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5−32.
(15) Renner, S.; Hechenberger, M.; Noeske, T.; Böcker, A.; Jatzke, C.; Schmuker, M.; Parsons, C. G.; Weil, T.; Schneider, G. Searching for Drug Scaffolds with 3D Pharmacophores and Neural Network Ensembles. *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 5336−5339.
(16) Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A. Comparison of fingerprint-based methods for virtual screening using multiple bioactive reference structures. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1177−1185.
(17) Whittle, M.; Gillet, V. J.; Willett, P.; Alex, A.; Loesel, J. Enhancing the effectiveness of virtual screening by fusing nearest neighbor lists:

a comparison of similarity coefficients. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1840−1848.
(18) Bender, A.; Mussa, H. Y.; Glen, R. C.; Reiling, S. Molecular similarity searching using atom environments, information-based feature selection, and a naive Bayesian classifier. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 170−178.
(19) Franke, L.; Byvatov, E.; Werz, O.; Steinhilber, D.; Schneider, P.; Schneider, G. Extraction and visualization of potential pharmacophore points using support vector machines: application to ligand-based virtual screening for COX-2 inhibitors. *J. Med. Chem.* **2005**, *48*, 6997−7004.
(20) Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A. New methods for ligand-based virtual screening: use of data fusion and machine learning to enhance the effectiveness of similarity searching. *J. Chem. Inf. Model.* **2006**, *46*, 462−470.
(21) *Pharmacophore Perception, Development and Use in Drug Design*; Guner, O. Ed.; International University Line: La Jolla, CA, 2000.
(22) *Pharmacophores and Pharmacophore Searches*; Langer, T., Hoffmann, R. D., Eds.; Wiley-VCH: Weinheim, 2006.
(23) Renner, S.; Schneider, G. Fuzzy pharmacophore models from molecular alignments for correlation-vector-based virtual screening. *J. Med. Chem.* **2004**, *47*, 4653−4664.
(24) Boström, J.; Hogner, A.; Schmitt, S. Do structurally similar ligands bind in a similar fashion. *J. Med. Chem.* **2006**, *49*, 6716−6725.
(25) Kozakov, A.; Clodfelter, K. H.; Vajda, S.; Camacho, C. J. Optimal clustering for detecting near-native conformations in protein docking. *Biophys. J.* **2005**, *89*, 867−875.
(26) Chema, D.; Eren, D.; Yayon, A.; Goldblum, A.; Zaliani, A. Identifying the binding mode of a molecular scaffold. *J. Comput.-Aided. Mol. Des.* **2004**, *18*, 23−40.
(27) Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **2004**, *47*, 337−344.
(28) Kelly, M. D.; Mancera, R. L. Expanded interaction fingerprint method for analyzing ligand binding modes in docking and structure-based drug design. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1942−1951.
(29) Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; Iwasawa, Y.; Nakata, K.; Chuman, H.; Nakano, T. VISCANA: visualized cluster analysis of protein-ligand interaction based on the ab initio fragment molecular orbital method for virtual ligand screening. *J. Chem. Inf. Model.* **2006**, *46*, 221−230.
(30) Mpamhanga, C. P.; Chen, B.; McLay, I. M.; Willett, P. Knowledge-based interaction fingerprint scoring: a simple method for improving the effectiveness of fast scoring functions. *J. Chem. Inf. Model.* **2006**, *46*, 686−698.
(31) Marcou, G.; Rognan, D. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47*, 195−207.
(32) Chuaqui, C.; Deng, Z.; Singh, J. Interaction profiles of protein kinase-inhibitor complexes and their application to virtual screening. *J. Med. Chem.* **2005**, *48*, 121−33.
(33) Fradera, X.; Knetgel, R. M.; Mestres, J. Similarity driven flexible ligand docking. *Proteins* **2000**, *40*, 623−636.
(34) Wu, G.; Vieth, M. SDOCKER: a method utilizing existing X-ray structures to improve docking accuracy. *J. Med. Chem.* **2004**, *47*, 3142−3148.
(35) Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129−149.
(36) Radestock, S.; Bohm, M.; Gohlke, H. Improving binding mode predictions by docking into protein-specifically adapted potential fields. *J. Med. Chem.* **2005**, *48* 5466−5479.
(37) Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **1982**, *43*, 59−69.
(38) Anzali, S.; Barnickel, G.; Krug, M.; Sadowski, J.; Wagener, M.; Gasteiger, J.; Polanski, J. The comparison of geometric and electronic properties of molecular surfaces by neural networks: application to the analysis of corticosteroid-binding globulin activity of steroids. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 521−534.
(39) Noske, T.; Sasse, B. C.; Stark, H.; Parsons, C. G.; Weil, T.; Schneider, G. Predicting compound selectivity by self-organizing maps: cross-activities of metabotropic glutamate receptor antagonists. *Chem. Med. Chem.* **2006**, *1*, 1066−1068.
(40) Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 653−681.
(41) Richmond, N. J.; Abrams, C. A.; Wolohan, P. R.; Abrahamian, E.; Willett, P.; Clark, R. D. GALAHAD: 1. pharmacophore identification by hypermolecular alignment of ligands in 3D. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 567−587.
(42) Bramson, H. N.; Corona, J.; Davis, S. T.; Dickerson, S. H.; Edelstein, M.; Frye, S. V.; Gampe, R. T., Jr; Harris, P. A.; Hassell, A.; Holmes, W. D.; Hunter, R. N.; Lackey, K. E; Lovejoy, B.; Luzzio, M. J.; Montana, V.; Rocque, W. J.; Rusnak, D.; Shewchuk, L.; Veal, J. M.; Walker, D. H.; Kuyper, L. F. Oxindole-based inhibitors of cyclin-dependent kinase 2 (CDK2): design, synthesis, enzymatic activities, and X-ray crystallographic analysis. *J. Med. Chem.* **2001**, *44*, 4339−4358.
(43) Anderson, M.; Beattie, J. F.; Breault, G. A.; Breed, J.; Byth, K. F.; Culshaw, J. D.; Ellston, R. P.; Green, S.; Minshull, C. A.; Norman, R. A.; Pauptit, R. A.; Stanway, J.; Thomas, A. P.; Jewsbury, P. J. Imidazo[1,2-a]pyridines: a potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hybridisation. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3021−3026.
(44) Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. N2-substituted O6-cyclohexylmethylguanine derivatives: potent inhibitors of cyclin-dependent kinases 1 and 2. *J. Med. Chem.* **2004**, *47*, 3710−3722.
(45) Pevarello, P.; Brasca, M. G.; Amici, R.; Orsini, P.; Traquandi, G.; Corti, L.; Piutti, C.; Sansonna, P.; Villa, M.; Pierce, B. S.; Pulici, M.; Giordano, P.; Martina, K.; Fritzen, E. L.; Nugent, R. A.; Casale, E.; Cameron, A.; Ciomei, M.; Roletto, F.; Isacchi, A.; Fogliatto, G.; Pesenti, E.; Pastori, W.; Marsiglio, A.; Leach, K. L.; Clare, P. M.; Fiorentini, F.; Varasi, M.; Vulpetti, A.; Warpehoski, M. A. 3-Aminopyrazole inhibitors of CDK2/cyclin A as antitumor agents. 1. Lead finding. *J. Med. Chem.* **2004**, *47*, 3367−3380.
(46) Gangjee, A.; Vidwans, A. P.; Vasudevan, A.; Queener, S. F.; Kisliuk, R. L.; Cody, V.; Li, R.; Galitsky, N.; Luft, J. R.; Pangborn, W. Structure-based design and synthesis of lipophilic 2,4-diamino-6-substituted quinazolines and their evaluation as inhibitors of dihydrofolate reductases and potential antitumor agents. *J. Med. Chem.* **1998**, *41*, 3426−3434.
(47) Cody, V.; Galitsky, N.; Luft, J. R.; Pangborn, W.; Blakley, R. L.; Gangjee, A. Comparison of ternary crystal complexes of F31 variants of human dihydrofolate reductase with NADPH and a classical antitumor furopyrimidine. *Anticancer Drug. Des.* **1998**, *13*, 307−315.
(48) Oefner, C.; D'Arcy, A.; Winkler, F. K. Crystal structure of human dihydrofolate reductase complexed with folate. *Eur. J. Biochem.* **1988**, *174*, 377−385.
(49) Cody, V.; Galitsky, N.; Luft, J. R.; Pangborn, W.; Rosowsky, A.; Blakley, R. L. Comparison of two independent crystal structures of human dihydrofolate reductase ternary complexes reduced with nicotinamide adenine dinucleotide phosphate and the very tight-binding inhibitor PT523. *Biochemistry* **1997**, *36*, 13897−13903.
(50) Davies, J. F., II; Delcamp, T. J.; Prendergast, N. J.; Ashford, V. A.; Freisheim, J. H.; Kraut, J. Crystal structures of recombinant human dihydrofolate reductase complexed with folate and 5-deazafolate. *Biochemistry* **1990**, *29*, 9467−9479.
(51) Bäckbro, K.; Löwgren, S.; Osterlund, K.; Atepo, J.; Unge, T.; Hultén, J.; Bonham, N. M.; Schaal, W.; Karlén, A.; Hallberg, A. Unexpected binding mode of a cyclic sulfamide HIV-1 protease inhibitor. *J. Med. Chem.* **1997**, *40*, 898−902.
(52) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance profile, human pharmacokinetics and X-ray crystal structure of DMP 450. *Chem. Biol.* **1996**, *3*, 301−314.
(53) Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N. et al. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **1994**, *263*, 380−384.
(54) Ala, P. J.; DeLoskey, R. J.; Huston, E. E.; Jadhav, P. K.; Lam, P. Y.; Eyermann, C. J.; Hodge, C. N.; Schadt, M. C.; Lewandowski, F. A.; Weber, P. C.; McCabe, D. D.; Duke, J. L.; Chang, C. H. Molecular recognition of cyclic urea HIV-1 protease inhibitors. *J. Biol. Chem.* **1998**, *273*, 12325−12331.
(55) Sham, H. L.; Zhao, C.; Stewart, K. D.; Betebenner, D. A.; Lin, S.; Park, C. H.; Kong, X. P.; Rosenbrook, W., Jr.; Herrin, T.; Madigan, D.; Vasavanonda, S.; Lyons, N.; Molla, A.; Saldivar, A.; Marsh, K. C.; McDonald, E.; Wideburg, N. E.; Denissen, J. F.; Robins, T.; Kempf, D. J.; Plattner, J. J.; Norbeck, D. W. A novel, picomolar inhibitor of human immunodeficiency virus type 1 protease. *J. Med. Chem.* **1996**, *39*, 392−397.
(56) Lam, P. Y.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C. H.; Emmett, G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Hodge, C. N. et al. Cyclic HIV protease inhibitors: synthesis,

conformational analysis, P2/P2′ structure-activity relationship, and molecular recognition of cyclic ureas. *J. Med. Chem.* **1996**, *39*, 3514−3525.
(57) Krishnan, R.; Mochalkin, I.; Arni, R.; Tulinsky, A. Structure of thrombin complexed with selective non-electrophilic inhibitors having cyclohexyl moieties at P1. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 294−303.
(58) Chirgadze, N. Y.; Sall, D. J.; Klimkowski, V. J.; Clawson, D. K.; Briggs, S. L.; Hermann, R.; Smith, G. F.; Gifford-Moore, D. S.; Wery, J. P. The crystal structure of human alpha-thrombin complexed with LY178550, a nonpeptidyl, active site-directed inhibitor. *Protein Sci.* **1997**, *6*, 1412−1417.
(59) Banner, D. W.; Hadváry, P. Crystallographic analysis at 3.0-A resolution of the binding to human thrombin of four active site-directed inhibitors. *J. Biol. Chem.* **1991**, *266*, 20085−20093.
(60) Mathews, II.; Tulinsky, A. Active-site mimetic inhibition of thrombin. *Acta Crystallogr.,Sect. D: Biol. Crystallogr.* **1995**, *51*, 550−559.
(61) Lyle, T. A.; Chen, Z. G.; Appleby, S. D.; Freidinger, R. M.; Gardell, S. J.; Lewis, S. D.; Li, Y.; Lyle, E. A.; Lynch, J. J.; Mulichak, A. M.; Ng, A. S.; NaylorOlsen, A. M.; Sanders, W. M. Synthesis, evaluation, and crystallographic analysis of L-371,912: A potent and selective active-site thrombin inhibitor. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 67−72.
(62) 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.
(63) *FlexX, version 1.2*; BioSolveIT: Sankt Augustin, Germany.
(64) *Sybyl, version 7.1*; Tripos, Inc.: St. Louis, MO, 2005.
(65) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L. et al. Structure at 2.5-a Resolution of Chemically Synthesized Human-Immunodeficiency-Virus Type-1 Protease Complexed with a Hydroxyethylene-Based Inhibitor. *Biochemistry* **1991**, *30*, 1600−1609.
(66) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.
(67) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **1992**, *13*, 505−524.
(68) Jones, G.; Willett, P.; Glen, G. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43−53.
(69) 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.
(70) Ultsch, A., Self-organizing neural networks for visualization and classification. In *Information and Classification*; Opitz, O., Lausen, B., Klar, R., Eds.; Springer: Berlin, 1993; pp 307−313.
(71) Ultsch, A.; Moerchen, F. ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM. In *Technical Report Dept. of Mathematics and Computer Science*; University of Marburg: Germany, 2005; pp 46−53.
(72) *MACCS keys*; MDL Informations Systems Inc.: San Leandro, CA.
(73) *Pipeline Pilot, version 6.0*; Accelrys Inc.: San Diego, CA.
(74) Willett, P.; Barnard, J. M.; Downs, G. M. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983−996.
(75) Kogej, T.; Engkvist, O.; Blomberg, N.; Muresan, S. Multifingerprint based similarity searches for targeted class compound selection. *J. Chem. Inf. Model.* **2006**, *46*, 1201−1213.
(76) Xue, L.; Godden, J. W.; Stahura, F. L.; Bajorath, J. Design and Evaluation of a Molecular Fingerprint Involving the Transformation of Property Descriptor Values into a Binary Classification Scheme. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1151−1157.
(77) Godden, J. W.; Stahura, F. L.; Bajorath, J. Anatomy of fingerprint search calculations on structurally diverse sets of active compounds. *J. Chem. Inf. Model.* **2005**, *45*, 1812−1819.

CI7003626