Similarity Based Docking

J. Marialke, S. Tietze, and Joannis Apostolakis*

Institute for Informatics, Research and Educational Unit for Bioinformatics and Practical Informatics, Ludwig-Maximilians-University, Amalienstrasse 17, D-80333 Munich, Germany

Received April 5, 2007

We have recently introduced GMA, a highly efficient method for flexible molecular alignment. Here we show how this approach can be used to improve docking accuracy and efficiency, in cases where a complex structure of a ligand with the target protein is known. In cases where a known ligand exists, yet the complex structure is unknown it is possible to make use of the advantages offered by this approach, by combining it with standard ligand docking.

INTRODUCTION

The prediction of the conformation of a particular ligand in the binding site of a protein is important for two reasons: It provides information on the so-called bioactive conformation of the ligand and on the essential interactions between the ligand and the protein. The bioactive conformation can be used for deriving quantitative structure−activity relationships (QSAR) and for similarity based screening, while knowledge of the interactions between a protein and a ligand can guide protein and ligand design, provide insight into specificity of molecular recognition, and shed light into the function of a particular protein as well as the mechanism behind it. Typically, prediction of the active conformation of a ligand is addressed by searching for a low energy (high score) conformation of the ligand in the field of the protein. In most cases where a target structure is available, there also exist known ligands and often even the structure of the complex of the protein with a known ligand. It is generally accepted that the difference in the structure of the apoprotein (the protein with an empty binding site) and the structure of the protein with a bound ligand is larger and more significant than the difference in the structure of a protein complexed with two different ligands. Therefore, whenever possible, the structure of the binding site is taken from a complex with a known ligand.

The information on the known ligand is generally discarded, while it could in principle be used on its own to obtain a reasonable prediction of the bioactive conformation of new ligands. It has been shown that the quality of structure prediction based on structural superposition[1] is comparable to that of docking, at least when evaluated in terms of structural deviation from the true solution, for example by root-mean-square deviation of the coordinates of the ligand heavy atoms. The disadvantage of similarity based structure prediction of the bioactive conformation of ligands is that it completely neglects the interaction with the protein field.

Traditionally the methods for structural superposition and docking are very similar to each other and differ mainly in the objective function being optimized, which in the first case is some type of superposition integral, while in the latter it is generally an interaction energy or score. The sampling methods are generally very similar, which is partly due to the parallel development of algorithms for docking and superposition: FlexS[1] and FlexX[2] from the Lengauer lab, the genetic algorithm for molecular superposition[3] and Gold[4] by Jones et al., and morphological similarity[5] and Surflex[6] by Jain are some of the best known examples. The similarity of the sampling algorithms may also be responsible for the comparable efficiency of the corresponding "sister" tools.

The parallel structure of the problems of superposition and docking as well as the current tools for solving them has inspired a number of studies, which have recently investigated the possibility of using a combination of structural and similarity based techniques to improve structure prediction accuracy and efficiency.

FlexX Pharm[7] has been introduced as an extension of the FlexX[2] docking method which takes pharmacophoric constraints into account during the buildup of the ligands. The constraints are coded either as selected interactions that need to be fulfilled or by inclusion volumes in the receptor binding site. It has been validated on a number of complexes, where FlexX has shown poor performance, and on screening applications. While the introduction of the constraints does make the search more focused, the gain in efficiency is relatively small compared to normal FlexX (unto a factor of 3, though on average significantly less).

Cross has recently combined FlexX with FlexS for base fragment placement, thus using information from a known ligand to decide the initial placement of an anchor fragment for the buildup in the docking algorithm.[8] The method was shown to reproduce the FlexX200 redocking benchmark[9] with high accuracy, which is not surprising given the fact that information of the complex itself is used to seed the search. More interestingly cross docking accuracy (i.e., docking a ligand into the protein conformation from the complex of its protein receptor with a similar ligand) was evaluated. This benchmark will also be used in this work.

In a very recent paper[10] Marcou and Rognan have suggested the use of interaction fingerprints for postprocessing docking solutions comparing them to the interaction fingerprint derived from a pharmacophore or a given ligand.

* Corresponding author e-mail: Joannis.Apostolakis@bio.ifi.lmu.de; apostola@bio.ifi.lmu.de.

10.1021/ci700124r CCC: $40.75
Published on Web 11/29/2007

We have recently introduced a highly efficient[11] method for 3d flexible superposition, based on 2d graph matching. In our evaluation of the method on benchmarks taken from the literature it was shown that the method leads to comparable results to recent methods,[1,12−14] at significantly higher efficiency. It allows the computation of superpositions in 0.1 s on average, which is 2 orders of magnitude faster than most published methods for structural superposition and all known flexible docking tools. In this work we investigate the use of this approach in the context of molecular docking, under the assumption that both the target structure and at least one ligand are known. We call the approach homology docking, due to its conceptual similarity to homology modeling of protein structures. The aim of the approach is to increase the efficiency and accuracy of structure based screening, while still providing structural detail. We will further investigate how the method can be used in cases where a known ligand exists, whose structure in complex with the protein is unknown. The method is conceptually simpler than normal pharmacophore based approaches, since it does not require the identification of the pharmacophore itself and instead uses the available ligands to direct the placement of new candidates. On the other hand, the extension of the method to work for pharmacophores instead of ligands is straightforward and will be addressed in forthcoming work. We expect that the main application of the method will be at the rational optimization stage of ligand design, where, starting from a known candidate, molecules with more favorable properties are sought.

METHODS

We compare four different methods for structure prediction of the ligand in the binding site, utilizing similarity information to different extents. At one end are methods that use only ligand information, while at the other end are methods that use only receptor information. The first approach is the recently published graph based molecular alignment (GMA),[11] which flexibly aligns molecules on a rigid template of a second ligand. In the second approach GMA is performed initially, followed by relaxation of the ligand in the field of the protein. In the third approach (PoseDock) only knowledge of the identity of a true ligand is used, not of its conformation in the complex. Finally, as a comparison we employ standard protein−ligand docking (GlamDock) which uses no explicit information on existing ligands. In the following we describe the four methods.

GMA. We only give a very short description of GMA, as it has recently been presented elsewhere.[11] GMA or graph based molecular alignment flexibly superposes a query molecule on a rigid template. It first performs a maximum common subgraph (MCS) search of the corresponding chemical graphs, with the help of a simplified version of the RASCAL algorithm.[15] Similar to the original approach by Raymond the MCS search is performed on line graphs, which are derived from the chemical graphs. The line graph of a chemical graph contains a vertex for every edge (bond) in the chemical graph. Two line graph vertices are connected by an edge if the corresponding bonds in the chemical graph are incident to the same atom. The MCS defines a mapping between the two line graphs which is then transformed into an atomic mapping between the corresponding molecules. The atomic mapping is then used to obtain the structural superposition. To this end, the root-mean-square deviation (RMSD) of the mapped atoms is used as an objective function. First a crude approach called stamping is used to approximately adapt the conformation of the query molecule to that of the template: Torsional angles in the query molecule are mapped to corresponding angles in the template and set to the value in the template. A rigid superposition is performed,[16] and the resulting conformation of the query molecule is used as a starting conformation for a gradient minimization of the RMSD of all aligned atoms. The minimization takes place in torsion space with a combination of the RMSD and an internal steric overlap term as energy. This optimization is performed for 200 steps.

HomDock. HomDock is a combination of GMA with optimization methods used in docking. The starting point for a HomDock application is the crystal structure of the target protein and the structure of the template ligand in the binding site. The main idea of the approach is to use GMA to place candidate ligands onto the template, then optimize their placement in the field of the protein, and, finally, rank the ligands according to their interaction with the protein and/or their similarity with the ligand.

To this end, the GMA superposition is performed as above. Again the gradient optimization is set to 200 steps. After the superposition, the ligands generally show poor interactions with the protein. Hydrogen bond donors or acceptors in the ligand are not necessarily optimally oriented with respect to the corresponding groups in the protein, and clashes between protein and ligand may exist in parts where the ligand could not be optimally superposed to the template. To alleviate these problems a second optimization in the field of the protein is performed. We have recently[17] developed and implemented a continuous derivative empirical scoring function similar to ChemScore,[18] which we denote as ChillScore to distinguish it from the original. This scoring function has a continuous gradient and can also be minimized with the torsion space optimization scheme used in GMA. The second optimization using ChillScore as an objective function is performed for 200 steps with the previously described minimization algorithm.[11,17] Since in HomDock efficiency is an important issue, the length of the minimization has been kept relatively short. This may lead to structures with partial overlap between ligand and protein. It was found empirically that scaling the clash term by a factor of 0.9 is a good compromise between accuracy and efficiency of optimization.

Pose Dock. PoseDock requires information about a known ligand; however, its binding conformation is assumed to be unknown. The known ligand is docked with our in-house docking tool GlamDock,[17] and a number of high ranking conformations of sufficient structural variability are kept and subsequently used as templates in a HomDock calculation. Here we use the standard pose clustering of GlamDock: Conformations obtained in the docking are sorted in order of increasing energy. The best conformation is placed into the set of cluster representatives. Starting from the second best conformation, all conformations with an RMSD less than 1.5 Å to any cluster representative are discarded, and the rest become cluster representatives. We tested two settings for Pose Dock: PD1 uses only the best docked conformation, as a template, and is in every other way identical to

HomDock. PD10, the second variant uses the best ten pose cluster representatives for placing the template. In PD10, each query ligand is superposed onto the alternative conformations of the template ligand and minimized as in HomDock. The corresponding conformations for the query ligand are then ranked according to interaction energy, and the best is kept for further analysis. PD10 thus allows a restricted form of conformational sampling, as the query ligand is sequentially superposed on 10 different conformations of the template.

GlamDock. As a comparison to the similarity based docking procedures we have used our recently developed docking program, which has shown high accuracy in standard benchmarks.[17] GlamDock is a Monte Carlo with Minimization[19] based docking algorithm, which combines a number of previously suggested methods to obtain high accuracy and high efficiency predictions. The Monte Carlo search takes place in a nonorthogonal, generally redundant space of ligand orientations: The space consists of continuous coordinates on the one hand (the torsional degrees of freedom of the ligand), and integer degrees of freedom coding for direct mappings between interaction sites on the ligand and preplaced probes of corresponding chemistry to the ligand interaction sites in the protein binding site. The preplaced probes are obtained in a procedure similar to the Protomol[6] generation by Surflex: Around each possible interaction site in the protein binding site, complementary interaction probes are placed in a uniform manner and optimal interaction geometry. The placed probes are scored according to the normal interaction score and pose clustered keeping only the best scoring representatives of each cluster. The retained probes are indexed. For each interaction site on the ligand an integer degree of freedom is generated which can take on the value of any of the indices of compatible probes, indicating that the ligand is to be placed in a way that the corresponding interaction site overlaps with the probe corresponding to that index. It is also possible that the mapping takes the value -1 indicating no mapping for the particular interaction site.

A given point in the composite search space is translated into a ligand conformation, first by setting the torsional degrees of freedom to the corresponding coordinates of the composite space and then by rigidly placing the ligand in a way that the mappings coded in the composite space point are optimally fulfilled. This is achieved with an iteration of simple Kabsch rotations. Kabsch rotations minimize the distance between mapped points on two different objects (in this case interaction sites in the ligand and probes in the binding site). Unfulfilled mappings are removed after each rotation, and the procedure is reiterated. A subsequent torsion space minimization with the full ChillScore function is then performed for 15 steps, to relax the conformation in the field of the protein. The remaining (fulfilled) mappings are re-encoded back into the composite space point that lead to this placement in the first place. Starting from a given point in the composite search space, the Monte Carlo move consists of a random change in the coordinates of the configuration, minimization, and the re-encoding described above.

A docking simulation consists of 5 runs of 300 Monte Carlo steps each, where local minima are kept during all simulations, pose clustered, and postminimized with 80 steps of torsion space minimization. The best ranking conformation is predicted as the structure of the complex. More details on the algorithm and a wide ranging validation are given elsewhere.[17]

Benchmark. For the comparison between the different docking protocols we have taken the cross docking data set recently used by Simon Cross for demonstrating the improved FlexS/FlexX docking approach.[8] This benchmark consists of 7 different protein sets with a total of 35 different ligands. In each protein set each ligand is docked against each protein leading to 205 different dockings, 35 of which are redockings and 170 cross dockings. The structures for the protein ligand complexes were taken from previously prepared data sets to avoid effects of preparation customization. In his study Cross protonated and optimized the protein binding sites with Sybyl, and these structures were not publically available. Instead we used the structures as they had been prepared in the Astex data set.[20] The only exceptions were structures which were not included in that data set, namely the triose isomerase data set and 1dbk, which were taken from the FlexX-200 data set[9] and 1tnj and 1tnk which were taken from the GlamDock paper.[17] Visual inspection of structures from different sources did not identify any differences in preparation. Ligands and protein structures were not further modified apart from superposing the structures within the single protein sets to allow comparison of cross docking results.

For the purposes of our comparison mainly cross docking results are relevant since the methods we are comparing require the existence of a complex structure, on which the others are modeled, or, in the case of PoseDock, practically reduce to normal docking in a redocking benchmark.

Cross docking benchmarks consist of groups of PDB complexes of different ligands with the same protein. The complexes are superposed with respect to the binding site residues. The conformation of the binding site may differ in the different complexes, due to induced fit, crystallographic conditions, or dynamic effects of the structure. Nevertheless in general a reasonable alignment should be possible. This alignment allows the comparison of the different ligand poses with respect to each other.

In each protein group, each ligand is docked against all protein structures taken from complexes in which that ligand is not present. The quality of each docking can be evaluated by comparing the RMSD of the docked ligand to the ligand pose obtained by the alignment of the native complex (i.e., the complex with that particular ligand) to the protein conformation in which the docking had been performed. Cross docking is generally significantly more difficult than redocking, at the same time however also more relevant. In this study the difference between the two scenarios is even more pronounced, as (some of) the studied methods explicitly make use of the ligand structure in the complex in which a new ligand is docked.

Cross[8] has evaluated the results of cross docking by reporting the best RMSD of the ligand docked in any of the non-native protein structures from the set of complexes belonging to the same protein. This evaluation is possibly more relevant than pure redocking studies, as it shows that the docking approach finds at least one reasonable solution in a number of different candidate protein structures. However, in practice it does not provide any means to identify the best docking structure in the absence of the native ligand conformation. We will use this evaluation here in

order to compare with the results reported for FlexX and FlexS/FlexX. In the next section we will also report the results for overall cross docking accuracy and compare the variants introduced in this paper.

RESULTS

In each benchmark scenario, the results from FlexX and the combined FlexS/FlexX approach are described first as they serve for comparison purposes. Our own methods will be discussed in order of increasing usage of protein field information and decreasing ligand structure information: GMA uses only the information on the conformation of the ligand in the complex and no information on the interaction with the protein. HomDock uses the field of the protein to postoptimize the ligand conformation obtained from the alignment by GMA. PoseDock uses no information on the conformation of the ligand from the crystal structure. Instead it takes the docked structures obtained from docking a reference ligand, and finally GlamDock uses only the structure of the protein and the chemical structure of the ligand.

Redocking. The redocking results (i.e., the prediction of the conformation based on the structure from the complex corresponding to the ligand that is to be predicted) are at first sight not particularly important in terms of benchmarking the different methods, since they have no relevant application in a real world scenario: They explicitly (GMA, HomDock, FlexS/FlexX) or implicitly (PoseDock, GlamDock) require knowledge of the complex structure they are supposed to predict. In the latter case, the structure of the protein in the complex with the ligand is used, which biases to a certain extent the solution toward the correct ligand conformation. On the other hand, in docking benchmarks redocking is still the standard approach.[2,4,6,9,21] In the redocking experiment the reference ligand and the ligand to be docked are identical. The average RMSDs obtained for the redocking along with the number of dockings with an RMSD below 2.0 Å to the crystal structure are shown for the different methods in Table 1.

Overall the results separate the different methods into the two classes already mentioned. In the first, GMA, HomDock, and FlexS/FlexX show almost perfect reconstruction of the ligand in its complex conformation, with at least 31 (89%, FlexS/FlexX) complexes reproduced at less than 2.0 Å deviation. Clearly the results obtained here are a trivial validation of the fact that the ligand conformation can be reproduced, when its conformation is known. For FlexS/FlexX, this point needs to be weakened, since it only uses part of the ligand. However, the results demonstrate that when the anchor fragment is placed correctly, in most cases the overall construction of the ligand is successful.

The difference in the average RMSD between GMA (0.24) and HomDock (0.55) shows the average structural change upon minimization of an essentially native structure (as obtained by superposing the ligand onto itself with GMA) in the field of the protein. The fact that on average the minimum of the scoring function lies close to the native structure indicates that the energy function is suitable for these complexes and that the structures themselves do not contain significant clashes between ligand and protein.

Table 1. Redocking Experiment: Comparison over 7 Different Methods[a]

	FlexS/X	FlexX	GMA	HD	PD1	PD10	Glam
Trypsin							
1ppc	0,68	3,12	0,02	0,73	1,54	1,54	1,74
1pph	4,04	4,52	0,14	0,92	1,62	1,62	1,73
1tng	1,94	1,89	0,50	0,69	0,53	0,61	1,01
1tnh	0,20	0,55	0,05	0,47	0,45	0,67	0,48
1tni	1,87	2,67	0,20	0,30	2,44	2,44	2,47
1tnj	0,93	1,30	0,04	0,53	1,11	1,11	1,08
1tnk	1,25	1,85	0,62	0,71	0,75	0,75	0,78
1tnl	0,66	0,62	0,11	0,21	0,25	4,26	0,20
3ptb	3,61	0,70	0,03	0,20	0,33	0,33	1,12
Thrombin							
1dwb	0,35	0,55	0,02	0,52	7,10	8,57	7,12
1dwc	0,88	6,68	0,12	0,36	3,15	2,97	2,96
1dwd	0,88	4,81	0,07	0,43	7,76	1,86	7,77
Carboxyptd-A							
1cbx	1,06	6,31	0,12	0,30	0,96	0,96	1,01
1cps	0,45	0,96	1,27	1,39	1,66	0,58	1,79
2ctc	1,47	1,97	0,14	0,57	0,65	0,65	0,67
3cpa	1,35	2,52	0,19	0,48	1,08	2,12	1,15
6cpa	1,36	5,92	0,94	0,82	3,61	3,61	3,63
7cpa	4,47		0,54	0,99	3,90	4,18	3,93
Cyt P450							
1pha			0,34	0,74	4,99	5,78	4,99
1phd	0,47		0,56	0,65	4,54	4,48	4,48
1phf	0,46	4,54	0,26	0,50	4,54	4,55	4,55
1phg	1,25		1,33	0,49	0,54	0,54	0,57
2cpp	0,38	2,92	0,01	0,56	0,95	3,05	2,43
5cpp	0,41	1,34	0,01	0,96	1,42	1,42	1,42
FAB Fragment							
1dbb	0,53	6,64	0,03	0,55	0,58	0,58	0,61
1dbj	0,21	8,32	0,01	0,22	1,05	1,05	1,06
1dbk	0,78	6,31	0,01	0,87	1,01	0,87	1,01
1dbm	1,97	6,45	0,06	0,38	1,89	1,89	1,89
2dbl	1,56	5,61	0,31	0,74	1,12	1,12	1,06
Arabinose							
1abe	0,66	1,20	0,01	0,08	0,29	0,29	0,41
1abf	0,81	0,92	0,01	0,11	0,14	0,14	0,15
5abp	0,48	0,55	0,01	0,34	0,32	0,32	0,42
Triose phopsphate isomerase							
4tim	1,06	4,09	0,07	0,38	2,90	4,29	2,94
5tim	0,34	1,99	0,04	0,72	1,15	1,15	1,15
6tim	0,72	1,62	0,10	0,45	0,78	1,24	0,86
RMSD	1,16	3,21	0,24	0,55	1,92	2,05	2,02
# success	31,00	15,00	35,00	35,00	25,00	23,00	24,00
ratio	0,89	0,43	1,00	1,00	0,71	0,66	0,69

[a] Left to right: combined FlexS/FlexX, FlexX, GMA, HomDock, PoseDock (only best ranked GlamDock solution), PoseDock (ten highest ranked solutions), GlamDock. RMSD below 2.0 Å encoded green, >=2.0 red. Bottom: average RMSD over all redockings. # success : count of all elements that show an RMSD below 2.0 Å. Ratio : the success ratio over the whole data set (35 elements).

In the second class the PoseDock variants, GlamDock and FlexX all show significantly lower accuracy, ranging from 43% (FlexX[2]) to 71% (PoseDock1).

As GlamDock is a stochastic algorithm, it is important to average the obtained results. To this end we performed 5 repeats of each docking, and in Table 1 we report the average results for the success rate. The RMSDs reported for each ligand is not an average, rather it shows the results of the repeat which lead to the success ratio most similar to the average success ratio. PoseDock is in principle deterministic; however, it uses the poses obtained from GlamDock. Therefore the same averaging over the 5 repeats was performed. Statistical deviations of the results are given in Table 3, in the overall comparison of the different benchmark variants. In the redocking, PoseDock aligns the ligand on template poses obtained from docking the ligand itself and then applies a further minimization. In principle this is redundant, as the poses are already minimized by GlamDock; this redundancy can serve, however, as a stability check on the method. Interestingly, PoseDock10 actually leads to

Table 2. Cross Docking Experiment: Comparison over 7 Different Methods[a]

	FlexS/X	FlexX	GMA	HD	PD1	PD10	Glam
Trypsin							
1ppc	2,24	2,24	2,01	1,78	2,51	3,04	2,61
1pph	1,81	1,43	1,56	1,10	3,87	1,80	1,53
1tng	0,73	0,49	0,33	0,30	0,21	0,21	0,34
1tnh	0,58	0,47	0,29	0,50	0,40	0,40	0,51
1tni	1,16	2,53	1,95	2,44	2,29	2,20	0,88
1tnj	0,96	1,09	0,73	0,59	0,94	0,94	0,94
1tnk	1,15	1,14	0,95	0,77	0,77	0,78	0,69
1tnl	0,75	0,63	0,49	0,19	0,29	1,59	0,27
3ptb	0,47	0,33	0,20	0,27	0,47	0,43	0,54
Thrombin							
1dwb	0,45	0,50	0,24	0,56	0,74	0,72	0,74
1dwc	4,52	1,31	2,28	1,82	3,02	4,50	3,36
1dwd	8,52	8,19	2,65	2,69	5,45	1,36	0,74
Carboxyptd-A							
1cbx	1,10	1,30	0,57	0,76	1,20	1,01	0,79
1cps	0,73	1,06	1,10	1,09	1,13	0,86	0,69
2ctc	0,85	1,87	0,43	0,92	0,84	1,11	1,13
3cpa	1,67	1,72	1,06	1,56	1,87	1,94	1,00
6cpa	2,97	2,77	1,12	0,94	3,17	3,17	3,36
7cpa	3,23		2,13	3,53	3,84	3,84	1,57
Cyt P450							
1pha	7,05	7,39	4,08	3,01	4,64	4,23	5,12
1phd	1,61	1,73	1,34	1,54	3,24	3,10	4,44
1phf	1,24	4,38	1,71	2,08	3,29	2,12	4,65
1phg	5,22	5,21	4,51	3,77	3,49	2,16	0,39
2cpp	2,88	0,47	2,71	2,58	1,63	2,54	2,30
5cpp	0,65	0,75	1,98	2,19	2,28	1,43	0,95
FAB Fragment							
1dbb	0,48	1,18	0,45	0,41	0,57	0,62	0,59
1dbj	0,61	4,57	0,28	0,59	0,65	0,77	0,52
1dbk	0,53	6,36	0,29	0,55	0,74	1,07	0,88
1dbm	2,14	2,45	1,24	1,30	1,21	0,96	2,17
2dbl	1,19	1,12	1,44	1,32	1,27	1,09	0,77
Arabinose							
1abe	0,38	1,18	0,36	0,18	0,28	0,28	0,35
1abf	0,30	0,58	0,23	0,32	0,33	0,33	0,41
5abp	0,49	1,19	0,23	0,20	0,21	0,21	0,81
Triose phopsphate isomerase							
4tim	1,06	0,96	0,79	0,52	1,54	0,86	0,88
5tim	0,34	1,20	1,29	1,54	1,55	1,48	1,61
6tim	0,72	0,75	0,50	0,53	1,26	1,26	0,85
RMSD	1,74	2,07	1,24	1,27	1,75	1,55	1,41
# success	26,00	24,00	28,00	27,00	23,00	25,00	27,00
ratio	0,74	0,69	0,80	0,77	0,66	0,71	0,77

[a] Left to right: combined FlexS/FlexX approach, FlexX, GMA, HomDock, PoseDock_1, PoseDock10, GlamDock. RMSD below 2.0 Å encoded green, >=2.0 red. Bottom: RMSD average RMSD over all elements. # success : count of all elements that show an RMSD below 2.0 Å. Ratio: the success rate over the whole data set (35 elements).

somewhat poorer results, suggesting a low difference between the native conformation of the ligand and the next best conformations, for at least some of the complexes (e.g., 1tnl, 3cpa in Table 1), while PoseDock1 shows a slightly better success ratio (0.71) than GlamDock (0.69). This is due to the fact that upon minimization the GlamDock pose of 2cpp improves from 2.43 to 0.95 Å RMSD. However, a comparison of average RMSDs for GlamDock and PoseDock shows only a minor effect of the additional minimization in PoseDock1.

The redocking results explicitly demonstrate the dependence of the different methods on the quality of the complex used for docking. In the particular case of GlamDock they are particularly relevant as the generated poses are used for PoseDock. The accuracy of GlamDock redocking (69%) suggests that overall this is a typical set of complexes for redocking, slightly on the easy side.

Cross Docking (Best RMSD Evaluation). In Table 2 the cross docking results as evaluated by Simon Cross are reported. The results for FlexX and FlexS/FlexX are taken from reference 8. Each ligand from a cross docking set is docked into the protein structures of the different complexes. For the two methods that produce more than one structure per ligand/protein combination (PoseDock10 and GlamDock) only the best scoring structure is used.

The RMSDs shown correspond to the best result (with respect to RMSD) of a single ligand over all protein conformations in the protein group, except the native (i.e., the one in which the ligand is contained). In general, cross docking is more difficult than redocking; however, in this particular benchmark the results appear better, as the best results over N dockings are reported compared to a single docking in the redocking case.

The comparison shows that for the different ligand sets GMA and HomDock lead to quite similar results. In 28 (80%, GMA) and 27 (77%, HomDock) of 35 cases they find a structure with a RMSD below 2.0 Å to the crystal structure. This is somewhat better than the results reported for the FlexS/FlexX combination (26 of 35, 74%). If the success criterion is relaxed to a threshold of 2.5 Å the difference between GMA/HomDock and the Cross results is more significant with 31 (89%) and 30 (86%) of the complexes having been docked within the threshold, respectively, compared to 28 (80%) for FlexS/FlexX. This improvement is evident also at the average RMSD level: the average RMSD for the GMA/HomDock approach lies 0.5 Å lower than for FlexS/FlexX, which is not only due to the very poor performance of the FlexS/FlexX method in thrombin. In all other data sets the average RMSD of GMA and HomDock lies at least 0.1 Å lower than the FlexS/FlexX combination.

Somewhat disappointingly there is little difference between the success rates of GMA and HomDock, which is surprising considering that in HomDock the ligand can adapt to the structure of the protein into which it is docked. From a structural point of view GMA appears superior to HomDock, as it leads to slightly better or at least comparable results, in a fifth of the time.

PoseDock1 uses as a template the best scoring structure obtained from GlamDock and performs a superposition and relaxation just like HomDock. Its performance suffers from the combination of errors: For PoseDock to be successful it is necessary that both GlamDock and HomDock work well. While it is not possible to directly compare the results shown in Table 3, because in different methods the results shown are based on different templates, the tendency for PoseDock1 to show high RMSD deviation when either GlamDock or HomDock do not find a good structure is evident. Nevertheless, as a fast approximate method for finding a reasonably good structure it fares reasonably well with a success ratio of 66%.

PoseDock10 appears to be a good compromise between efficiency and accuracy leading to a slightly higher success ratio (25 of 35, 71%).

In single cases the advantage of using alternative template conformations can be seen in this benchmark. In Figure 1 the difference between PoseDock1 and PoseDock10 is exemplified in the case of docking the 1dwd ligand into thrombin. Both PoseDock1 and PoseDock10 use the 1dwc ligand as a template. PoseDock1 suffers from the fact that the docking of 1dwc in its own structure leads to a poor structure at rank 1. Among the ten best ranks there is at least

Table 3. Comparison of the Different Docking Protocols[a]

	FlexS/X	FlexX	GMA	HDock	PDock 1	PDock 10	GlamDock
redocking	0.89	0.43	1.00	1.00	0.72 (0.02)	0.65 (0.02)	0.69 (0.03)
cross best RMSD	0.74	0.69	0.80	0.77	0.66 (0.01)	0.72 (0.03)	0.74 (0.03)
cross all			0.49	0.49	0.43(0.03)	0.44 (0.01)	0.51 (0.01)
highest sim.			0.68	0.71	0.55	0.54	
rb bestRMSD	0.08	−0.20	0.10	0.11	0.04	−0.04	−0.03
rb all			0.31	0.31	0.22	0.17	0.13
time [s]			0.10	0.65	0.65	1.40	40.00

[a] Success rates for the different evaluation methods are shown. *rb* is the calculated redocking bias. Numbers in parentheses are the standard deviations over five runs, for those scenarios that depend on GlamDock.

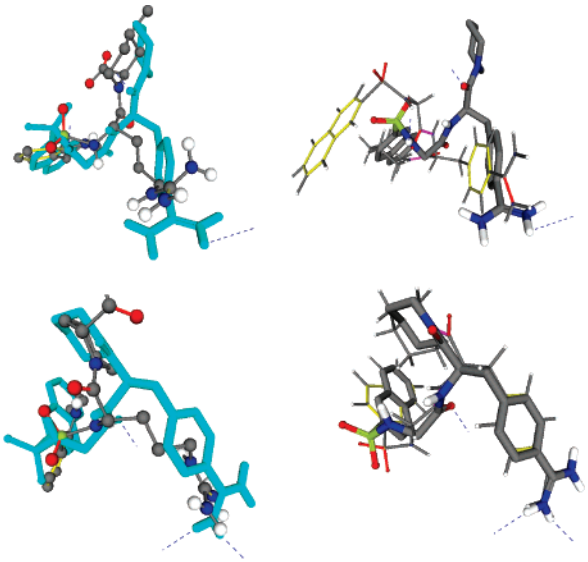

Figure 1. Top left: (a) PoseDock1 solution for ligand 1dwd. The ligand (light blue thick stick) is aligned on the best ranked pose for the template 1dwc ligand. Top right: (b) Comparison between aligned pose (thick stick) and crystal structure for ligand 1dwd (sticks) RMSD 5.45 Å. Bottom left: (c) PoseDock10 solution 1dwd. The ligand (light blue thick stick) is aligned on the sixth ranked solution of the 1dwc ligand (thin sticks). Bottom right: (d) Comparison between aligned pose (thick stick) and crystal structure for ligand 1dwd (sticks) RMSD 1.36 Å.

one structure, which is close enough to the native structure. PoseDock10 uses that structure for alignment and identifies the corresponding placement for 1dwd as the one interacting best with the protein. The use of alternative template conformations corresponds to a conformational search for the query ligand. In general the improvement in structure prediction accuracy is however relatively low. We have also tested other alternatives for conformational sampling, such as Monte Carlo sampling around the conformation obtained from HomDock; however, also this approach did not lead to significant improvement in overall results on the current benchmark, in spite of significant additional computational overhead (results not shown).

Overall Cross Docking Evaluation. The overall cross docking accuracy data for FlexX and FlexS/FlexX are not available; we therefore can only compare the performance of our own methods in this respect. For each ligand-complex pair in each protein group the ligand is docked into the complex, and the RMSD to the native structure of the ligand is evaluated. The overall success rates are given in Table 3. The accuracy ranges from 0.43 (PoseDock) to 0.51 for GlamDock. The first interesting observation is that, when a structure of a protein complex exists, molecular superposition

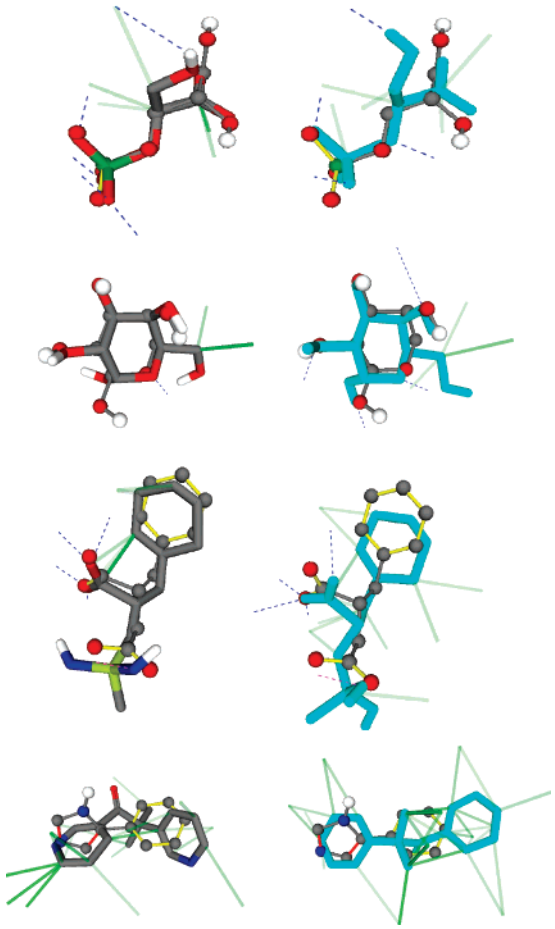

Figure 2. Examples of structure predictions with HomDock. On the left side panels with the corresponding crystallographic alignment are shown. On the right side panels with the alignment of the query (thick sticks) on the template crystal structure (ball and stick) are shown. Green transparent bars denote clashes, while blue lines denote H-bonds to the protein. From top to bottom the pairs are 4tim/6tim, 5abp/1abe, 1cps/1cbx, and 1phg/1phf.

with GMA (success rate 0.49) and homology docking (0.49) are practically as accurate as docking with GlamDock (0.51).

In Figure 2 different examples of structure prediction with HomDock are shown. In general when the similarity between template and ligand is high enough the prediction is highly accurate. This is exemplified in the first three cases all of which show low RMSDs (0.52, 0.68, and 1.32 Å, respectively). The interactions formed with the binding site are in general comparable at a qualitative level showing more or less similar clashes and H-bonds formed as in the crystal structure. Deviations from the crystallographic structure can generally be attributed to small differences in the structure of the binding site (as these are cross docking results) and

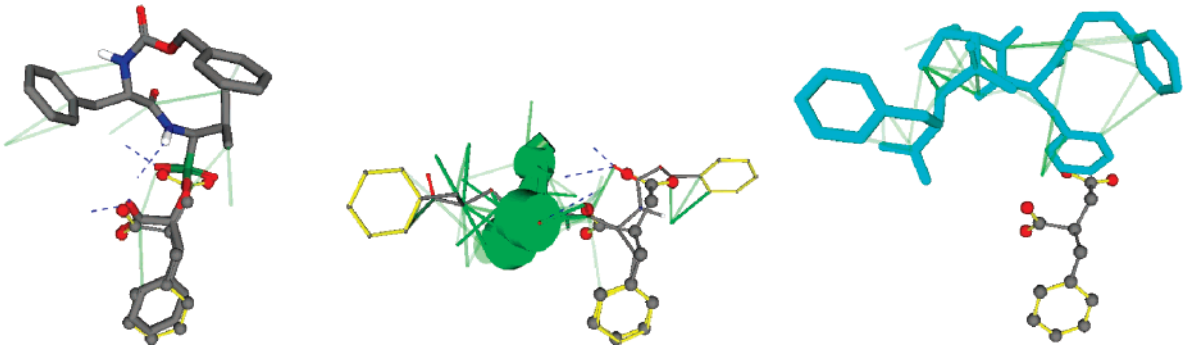

Figure 3. An example where the query (7cpa) is significantly larger than the template (1cbx) is shown. The leftmost and rightmost panel show the crystal structure and the aligned structure of 7cpa in the 1cbx binding site, with the same definitions as in Figure 2. In the middle panel the structure of the ligand 7cpa aligned by GMA on ligand 1cbx is shown.

the effect of the minimization of the query in the binding site. In the example of the arabinose binding protein (second row in Figure 2) the minimization leads to the identification of additional H-bonds not readily observed in the crystal structure. However this difference is artificial as these bonds also appear in the crystal structure upon minimization. We have kept the unminimized structures as reference to avoid introducing force field bias into the reference states. In the last case (bottom row, the prediction of 1phg in the 1phf conformation) the query is placed in an inverted pose compared to the known conformation due to the pseudo-symmetry in the query molecule. This leads to a poor RMSD of 5.5 Å.

The example in Figure 3 shows the problems of the approach for query ligands that are significantly larger than the template. GMA places ligand 7cpa onto the template 1cpx leading to significant clashes with the binding site, which in the minimization leads to the ligand being thrown out of the binding site. The problem could be solved by a substantial conformational search of the unmatched parts of the query ligand in the binding site; however, this leads to significant loss of efficiency. The protocols studied here for the ligand based methods aim at high efficiency at the cost of accuracy in a few cases. It is important to note in this context that this type of situation is readily identifiable by the low similarity between query and template.

PoseDock performance suffers from the additional inaccuracies of docking with GlamDock and therefore shows lower accuracy (0.43 and 0.44 for PoseDock1 and PoseDock10, respectively). Nevertheless, single examples exist where PoseDock10 shows that alternative binding modes found in nature for similar ligands are actually mirrored in the candidate conformations for a single ligand.

This is demonstrated in the example of the placement of 1tni on the 1tnk template as shown in Figure 4. 1tnk is redocked well by GlamDock (RMSD $<$ 1 Å). Furthermore 1tnk and 1tni ligands are chemically very similar (the linker between the amino group and the phenol ring is longer by 1 in 1tnk). However, with 1tnk as template, both HomDock and PoseDock1 predict for 1tni a structure deviating from the crystal structure as shown in Figure 4c with an RMSD $>$ 2 Å. The difference lies in the aromatic ring placement. PoseDock10 finds the best placement for 1tni based on the second GlamDock pose for 1tnk.

A more interesting example is that of ligands 1dbm and 1dbk in Figure 5. The steroid binding antibody is known to bind different steroids in different binding modes.[22] In fact the ligands in 1dbm and 1dbk show a difference in the rotation of the steroid plane of more than 120° around the longest axis of the steroid. Interestingly, GlamDock can reproduce the correct orientations of the different steroids in the binding site (RMSD $<$ 1.85 Å for all complexes). However, using PoseDock1 with 1dbk as a template for 1dbm leads to a very poor RMSD, simply because the template binds in the other binding mode. The same is true for HomDock, demonstrating that in this case the problem does not lie in the structure prediction of the template. For the ligand 1dbk GlamDock finds the alternative binding mode at the second rank, and PoseDock10 uses it to make an almost perfect prediction for 1dbm (RMSD $=$ 0.97 Å).

These two examples support the notion that alternative binding modes found in complex structures for similar ligands exist for each of the ligands. They are simply differentially stabilized by the different substitutions in the ligands.

Predictive Choice of Best Template. In the real life scenario the choice of the template ligand cannot be made based on the best RMSD obtained. For GMA and HomDock the choice is limited by the existing crystal structures, while in PoseDock we are free to choose the template ligand. One possible criterion for this choice is the interaction energy. One can use the template that leads to the best scoring placements for a particular test ligand. We obtain better results, however, by choosing the template according to the similarity between the test ligand and the template. For the similarity we simply use the size of the common subgraph, obtained in the MCS detection step of GMA.[11] With this choice the accuracy for PoseDock1 and PoseDock10 rises to 55% and 54%, respectively. A similar behavior is observed for GMA and HomDock, where the corresponding success ratio rises from 49% to 68% and 71%, respectively. The choice of the most similar template thus leads to an accuracy which lies between the overall accuracies and the ideal accuracy obtained in the best RMSD scenario.

Comparison of Different Evaluations and Redocking Bias. The comparison between redocking and overall cross docking results demonstrates the higher complexity of the cross docking problem: All methods (including normal docking) show significantly lower accuracy. This is seen in Table 3, which shows the overall success rate for cross docking with the different methods.

In comparison to the redocking experiments the success rate of protein interaction based methods (GlamDock) is significantly more stable than for ligand based methods such

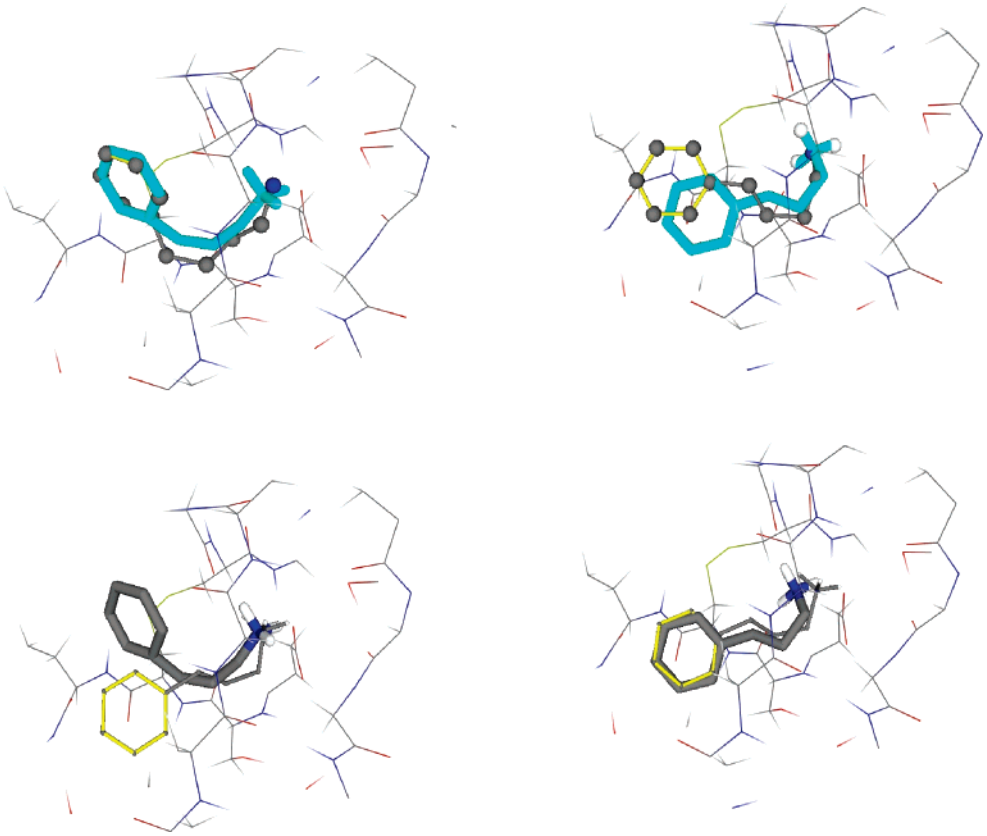

Figure 4. Top left: (a) 1tnk ligand (light blue thick sticks) aligned on template 1tni (ball and sticks). The ligand 1tnk is almost perfectly aligned on the template molecule. Top right: (b) 1tnk ligand (light blue thick stick) aligned on the second ranking pose of 1tni ligand in PoseDock10 (ball and sticks). The ligand is shifted away from the template after gradient optimization. Bottom left: (c) HomDock aligned 1tnk (thick sticks) compared to its crystal structure (thin sticks) RMSD = 2.89 Å. Bottom right: (d) PoseDock10 1tnk (thick sticks) compared to its crystal structure (thin sticks) RMSD = 0.82 Å.

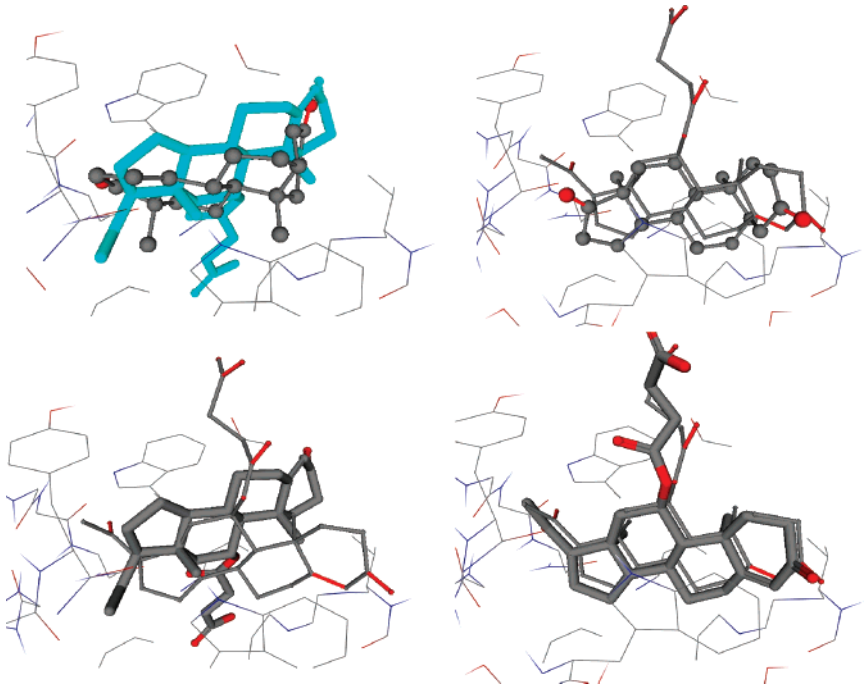

Figure 5. Top left: (a) 1dbm ligand (light blue thick stick) aligned by PoseDock1 on a template 1dbk ligand (ball and sticks). Top right: (b) 1dbm ligand (light blue thick stick) aligned by PoseDock10 on the second pose of 1dbk (ball and sticks). Bottom left: (c) Aligned HomDock 1dbm structure compared to observed crystal structure RMSD = 6.57 Å. Bottom right: (d) Aligned PoseDock10 1dbm structure compared to observed crystal structure RMSD = 0.97 Å.

as GMA. It is interesting to note that the FlexS/FlexX method also shows significant loss of accuracy in the comparison to the redocking results, even though it is not as large as for GMA.

The decrease in the accuracy from redocking to cross docking indicates the dependence of the docking methods on a rather exact structure of the receptor and their intolerance with respect to structural changes in the binding site. Additionally for the methods that use a template ligand the structural difference between the test ligand and the template also has a detrimental effect on docking accuracy. We quantify this dependence of the different methods, by introducing a novel measure called the *redocking bias rb*. For a given cross docking benchmark, the *redocking bias* is given simply by the quotient of the success ratios for redocking and cross docking. In general, we will report the logarithm of the *redocking bias*. The reason for this is simply that a *redocking bias* of 0 suggests no bias as it means that the success ratios are identical in the redocking and cross docking experiment. The *redocking bias* characterizes both the benchmark and the docking algorithms: If the protein structures in every group of the cross docking benchmark practically show no structural variation, the *redocking bias* should be 0. If they do show structural variation but the docking algorithm is ideal, in the sense that it can take protein flexibility correctly into account, then again the *redocking bias* is 0. Thus this measure can be used both for comparing different benchmark sets based on the results of (preferably a few) docking algorithms and for comparing the robustness of different algorithms on the same benchmark set. In our case the *redocking bias* also measures the general dependence on the information about the complex used for docking: Apart from small changes in the structure of the protein, the dependence on the template ligand is also mirrored in the *redocking bias*. In Table 3 we report the *redocking bias* of the two different benchmark evaluations. In the best RMSD evaluation the *redocking bias* often appears negative, which is due to the fact that the cross docking experiment has the unfair advantage of choosing the complex that leads to the best RMSD for the docking. Most notably for FlexX the *redocking bias* of −0.2 appears to be due to the rather poor redocking accuracy in FlexX. Three complexes could not be redocked in the study by Cross, decreasing redocking performance. Assuming that these complexes could have been docked correctly would improve the redocking accuracy of FlexX enough to explain the negative *redocking bias*. The *redocking bias* is largest for GMA and HomDock (0.10 and 0.11) followed by FlexS/FlexX (0.08), PoseDock1 (0.04), and PoseDock10 and GlamDock (−0.04 and −0.03). The relevance of this analysis is that it allows us to rank FlexS/FlexX within the spectrum of our variants. It is somewhat surprising that FlexS/FlexX shows a relatively strong *redocking bias*, as it is a normal docking approach with an initial constraint on the choice of the base fragment placement. It is even more surprising given the apparent low *redocking bias* for FlexX alone. One possibility is that this discrepancy is an artifact of the base fragment placement procedure in FlexX. The explanation would be that the low *redocking bias* for FlexX is mainly due to poor base placement in redocking, whose effect can be removed either by using the FlexS/FlexX approach or by the bestRMSD evaluation of the cross docking benchmark.

The true *redocking bias* is seen in the comparison between redocking accuracy and average cross docking accuracy over all protein ligand combinations. This analysis can however not be performed for FlexX and FlexS/FlexX as we do not have the corresponding data. GMA and HomDock show the highest bias (0.31), with PoseDock1 and PoseDock10 at 0.22 and 0.17, respectively, followed by GlamDock with 0.13. The order of the *redocking bias* corresponds to that obtained in the best RMSD evaluation and suggests that also under this evaluation the general trends would not change also for FlexS/FlexX.

Table 4. Correlation of Energies for 5 Different Methods in the Best RMSD Analysis

	GMA	HDock	PD 1	PD 10	GlamDock
GMA	1	0.6689	0.70258	0.54626	−0.03498
Hdock	0.6689	1	0.8473	0.82006	0.4892
PD 1	0.70258	0.8473	1	0.80214	0.43754
PD 10	0.54626	0.82006	0.80214	1	0.69854
GlamDock	−0.03498	0.4892	0.43754	0.69854	1

Energetic Comparison. While a good structure prediction is important in itself, the identification of key interactions between the ligand and the protein is an additional point relevant in modeling studies. Structures within 2 Å from each other can differ significantly with respect to the interactions they form with the binding site. Directed polar interactions such as hydrogen bonds often depend on relatively small changes in the structure of the complex. This is not only true for binding interactions. Steric overlap between protein and ligand can often be resolved by small changes in the coordinates in the order of a few tenths of an Ångstrom. On the other hand, in some cases the binding site may be so tight around the ligand that there is no way for the ligand to be accommodated, even though the overlap is relatively small.

In order to test the quality of the interaction prediction by the different docking variants discussed in this work we compared the score of the predicted structures. The score represents a summary of the quality of the interactions formed between the ligand and the protein. Furthermore, in virtual screening and related studies the interaction score is used for deciding whether a particular molecule could be a ligand of the protein. As a standard of truth we used the interaction score from the GlamDock docking run, as the most independent (and CPU intensive) of the tested methods. In Table 4 the correlation between the scores obtained from the different methods is shown. The last column shows the correlation to docking scores, which is the score commonly used in structure based screening. Correlations suffer from outliers, usually caused by clashing structures. This is shown explicitly in Figure 6, where the energies for the bestRMSD structures are plotted, sorted by (Max energy − Min energy). In all but 4 cases the maximum energy is given by GMA, the minimum by GlamDock. Both with respect to correlation and in the plot PoseDock10 yields the scores most similar to GlamDock. Thus while the additional effort invested in minimizing up to 10 different conformations in PoseDock does not show clear advantages for structure prediction, for ligand scoring it appears significantly better than the other methods.

Times. The CPU time requirement of the methods increases with decreasing usage of ligand information. GMA takes approximately 0.1 s per structural alignment, HomDock, and PoseDock1 approximately 0.65 s for the alignment and the subsequent minimization in the field of the protein. PoseDock10 takes less than 2 s (∼1.4 s) for the alignment

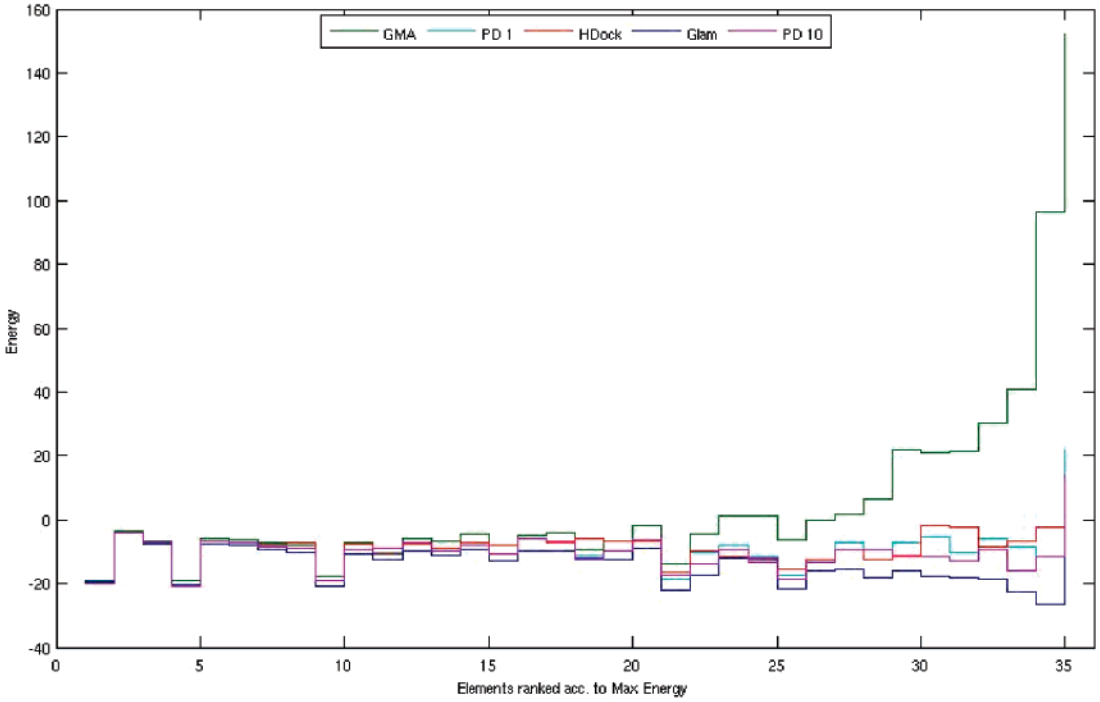

Figure 6. Energy comparison of ChillScore energies for 35 compounds docked with 5 different methods. The elements are sorted according to the maximum energy difference between GlamDock docking and the homology based methods (including GMA).

and minimization of the ligand in all 10 poses of the template ligand. The reason for the apparent sublinear dependence on the number of poses used in PoseDock1 and PoseDock10 lies with the fact that a number of time intensive steps need to be performed only once in the PoseDock10 run: The MCS procedure for finding a mapping between the test ligand and the template, the preprocessing of the protein binding site, and the initialization of the energy function. The only parts of the calculation that need to be performed for each template pose are the minimizations in torsion space, first with respect to the RMSD to the template pose and second with the interaction with the protein as an objective function.

GlamDock takes on average 40 s for a single docking. The time for the docking of the template ligand is not considered for PoseDock, as it is assumed that the number of test ligands is significantly higher than the template ligands used. The times are reported for a 2.8 GHz Xeon CPU.

CONCLUSION

We have evaluated different structure prediction protocols ranging from similarity based to receptor based approaches, whereby we have focused on highly efficient methods that combine information on both the ligand and the receptor. The main results of this study are the introduction of a novel measure for comparing data sets and docking methods, the introduction of simple methods for taking prior knowledge on known ligands into account, and their comparison both among each other as well as against existing methods from the literature. The latter comparison shows the methods to be of at least comparable accuracy to the existing methods on standard benchmark sets at significantly higher efficiency.

For the analysis of the methods we have introduced the *redocking bias*, a simple measure of the dependency of docking accuracy on the knowledge of structure of the complex to be predicted. We believe this to be a generally relevant quantity for the comparison of docking methods as the development focus moves from improving redocking performance to achieving higher prediction accuracy in cross docking experiments. Within a benchmark the *redocking bias* can be used to compare different methods on their robustness toward protein flexibility, and between benchmarks it can be used for comparing their relative complexities (the amount and relevance of protein flexibility within a group of complexes) based on the performance of the same docking methods on them.

The approaches suggested here allow accurate structure prediction taking 1 s or less per prediction. When reasonably similar templates are available, GMA and HomDock provide significantly higher prediction accuracy than FlexS/FlexX, FlexX, and GlamDock. They further provide an approximation of the interaction score between protein and the ligand. In the type of evaluation found in the literature,[8] the methods suggested here lead to very good results in a fraction of the time of previous suggested approaches.

Unlike the methods FlexS/FlexX, GMA, and HomDock, the PoseDock approach does not require an existing crystal structure for the template ligand. It takes the template poses from docking simulations with GlamDock. While less accurate than the other methods from a structural point of view, PoseDock can be used in high throughput scenarios to dock large libraries of similar ligands in a focused virtual screening scenario, e.g., in the rational optimization stage of lead development or in docking preclustered libraries of compounds. In the former case the known leads can be used as templates to direct the placement of the other compounds in the library. In the latter case the cluster representatives can be used as templates.

We are currently assessing the performance of those methods in screening, as they automatically provide similarity and structure based criteria for identifying active molecules.

ACKNOWLEDGMENT

The authors acknowledge partial funding by the DFG project AP 101/1-2.

REFERENCES AND NOTES

(1) Lemmen, C.; Lengauer, T.; Klebe, C. FLEXS: a method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502−20.
(2) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261* (3), 470−489.
(3) Jones, G.; Willett, P.; Glen, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9* (6), 532−49.
(4) Jones, G.; Willett, P.; Glen. R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245* (1), 43−53.
(5) Jain, A. N. Morphological similarity: a 3D molecular similarity method correlated with protein-ligand recognition. *J. Comput.-Aided Mol. Des.* **2000**, *14* (2), 199−213.
(6) Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46* (4), 499−511.
(7) Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* **2002**, *16* (2), 129−149.
(8) Cross, S. S. J. Improved Flexx docking using Flexs-determined base fragment placement. *J. Chem. Inf. Model.* **2002**, *45* (4), 993−1001.
(9) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the Flexx incremental construction algorithm for protein-ligand docking. *Proteins* **1999**, *37* (2), 228−241. Marcou, G.; Rognan, D. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195−207.
(10) Marcou, G.; Rognan, D. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195−207.
(11) Marialke, J.; Koerner, R.; Tietze, S.; Apostolakis, J. Graph-based molecular alignment (GMA). *J. Chem. Inf. Model.* **2007**, *47* (2), 591−601.
(12) Kramer, A.; Horn, H. W.; Rice, J. E. Fast 3D molecular superposition and similarity search in databases of flexible molecules. *J. Comput.-Aided Mol. Des.* **2003**, *17* (1), 13−38.
(13) Labute, P.; Williams, C.; Feher, M.; Sourial, E.; Schmidt, J. M. Flexible alignment of small molecules. *J. Med. Chem.* **2001**, *44* (10), 1483−1490.
(14) Cho, S. J.; Sun, Y. Flame: A program to flexibly align molecules. *J. Chem. Inf. Model.* **2006**, *46* (1), 298−306.
(15) Raymond, J.; Gardiner, E.; Willett, P. RASCAL: Calculation of Graph Similarity Using Maximum Common Edge Subgraphs. *Comput. J.* **2002**, *45*, 631−644.
(16) Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr.* **1976**, *32*, 922−923.
(17) Tietze, S.; and Apostolakis, J. Glamdock: Development and validation of a new docking tool on several thousand protein-ligand complexes. *J. Chem. Inf. Model.* **2007**, *47* (4), 1657−72.
(18) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolinim, 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* (5), 425−445.
(19) Apostolakis, J.; Plückthun, A.; Caflisch, A. Docking small ligands in flexible binding sites. *J. Comput. Chem.* **1998**, *19* (1), 21−37
(20) Nissink, J. W.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A new test set for validating predictions of protein-ligand interaction. *Proteins* **2002**, *49* (4), 457−71.
(21) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57* (2), 225−242.
(22) Arevalo, J. H.; Hassig, C. A.; Stura, E. A.; Sims, M. J.; Taussig, M. J.; Wilson, I. A. Structural analysis of antibody specificity. Detailed comparison of five Fab′-steroid complexes. *J. Mol. Biol.* **1994**, *241* (5), 663−90.

CI700124R