PROTEINS: Structure, Function, and Genetics 37:228–241 (1999)

Evaluation of the FLEXX Incremental Construction Algorithm for Protein–Ligand Docking

Bernd Kramer,* Matthias Rarey, and Thomas Lengauer
German National Research Center for Information Technology (GMD), Institute for Algorithms and Scientific Computing (SCAI), Sankt Augustin, Germany

ABSTRACT **We report on a test of FLEXX, a fully automatic docking tool for flexible ligands, on a highly diverse data set of 200 protein–ligand complexes from the Protein Data Bank. In total 46.5% of the complexes of the data set can be reproduced by a FLEXX docking solution at rank 1 with an rms deviation (RMSD) from the observed structure of less than 2 Å. This rate rises to 70% if one looks at the entire generated solution set. FLEXX produces reliable results for ligands with up to 15 components which can be docked in 80% of the cases with acceptable accuracy. Ligands with more than 15 components tend to generate wrong solutions more often. The average runtime of FLEXX on this test set is 93 seconds per complex on a SUN Ultra-30 workstation. In addition, we report on "cross-docking" experiments, in which several receptor structures of complexes with identical proteins have been used for docking all cocrystallized ligands of these complexes. In most cases, these experiments show that FLEXX can acceptably dock a ligand into a foreign receptor structure. Finally we report on screening runs of ligands out of a library with 556 entries against ten different proteins. In eight cases FLEXX is able to find the original inhibitor within the top 7% of the total library. Proteins 1999;37:228–241.**

Key words: molecular docking; flexible docking; protein–ligand interaction; molecular flexibility; validation; drug design

INTRODUCTION

In one scenario in rational drug design, the three-dimensional (3D) structure of a pharmacological target protein has been solved, and computer methods are applied to find lead compounds. This is done by solving the docking problem for a number of potential drug molecules accumulated in a database. Nowadays there are a wide variety of docking methods, not all suitable or even conceived for this kind of ligand screening (see Blaney and Dixon,[1] Colman,[2] Kuntz,[3] Lengauer and Rarey,[4] and Lybrand[5] for overviews).

Most of those docking tools have been examined with predominantly smaller test sets of structurally known protein–ligand complexes, e.g., we have evaluated our own docking tool FLEXX with 19 complexes in Rarey et al.[6–8]

From our point of view a first objective test for docking methods has been the CASP2 experiment[9] in 1996, which included a docking section with six non-covalently bound protein–ligand complexes. The task was to submit blind predictions of the ligand structures to the organizers who made the experimental structures available for an evaluation later on. Aside from us, ten other groups entered the contest with their docking methods. Detailed information on the results can be found in a special CASP2 issue.[10,11] FLEXX turned out to be the fastest tool while exhibiting a prediction quality that was well placed among the leading groups. In comparison to our test set with 19 entries the CASP2 test led to a few surprisingly low-quality predictions. Such information is very important in the development process of a docking tool. Of course a disadvantage is the small diversity of the CASP2 example set.

An impressive validation of the docking tool GOLD has been reported by Jones et al.[12] They tested their method which treats the ligands flexibly and which is based on a genetic algorithm with a data set of 100 complexes extracted from the Brookhaven Protein Data Bank.[13] GOLD achieved a 71% success rate in identifying the experimental binding mode. The diversity of this data set seems to be satisfying.

By merging the examples of the GOLD data set and of our own test set (which had been enlarged significantly in the meantime) we now have a data set of 200 protein–ligand complexes. Here we report the results of FLEXX in reproducing these complexes. Besides the illustration of several examples of successful docking we will analyze the limitations of the FLEXX docking-algorithm. In addition we performed "cross-docking" experiments involving protein–ligand complexes with different ligands but the same protein. In a cross-docking experiment all involved ligands are docked into all relevant receptor structures. The receptor structures (of the same protein) differ by small margins because of induced fit. In our opinion, this kind of experiment is preferable to docking into uncomplexed receptor structures. In uncomplexed protein structures there is a higher risk of unduly large structural changes owing to the unbound state. In addition, the situation that only crystal structures of apoproteins are available is rare in practice. Within the data set there are 21 groups of complexes with identical proteins. Each group has be-

*Correspondence to: Bernd Kramer, 4SC GmbH, Am Klopferspitz 19, D-82152 Planegg-Martinsried, Germany. E-mail: kramer@4sc.com
Received 3 September 1998; Accepted 1 June 1999

tween two and (in the case of trypsin) nine members. In this report we present the results for seven of these groups. And last we will present screening experiments with ten proteins and a small ligand library with 556 entries containing the ligands of the complex data set and additional drug-like molecules.

MATERIALS AND METHODS

Below we give a brief description of the models and the algorithms underlying the FLEXX docking tool. All docking experiments have been performed with version 1.6.5 on a SUN Ultra-30 with a single 296-MHz processor and 128 MB main memory.

Modeling Ligand Flexibility

The conformational flexibility of the ligand is modeled by a discrete set of preferred torsion angles at acyclic single bonds, and multiple conformations for ring systems. Torsion angles at multiple bonds, bond lengths, and bond angles are used as given in the input structure. Therefore, reasonably minimized geometries should be used. The torsional angles are taken from a database containing about 900 molecular fragments with a central single bond which has been derived from the Cambridge Structure Database (CSD)[14] by Klebe and Mietzner.[15] By this method up to 12 low-energy torsion angles have been assigned to each single bond.

Multiple conformations for rings are computed with the program CORINA.[16] The number of ring atoms is limited to seven. Larger rings are considered rigid, and the input structure is used. There are three examples (1fki, 1mmq, and 1nco) in the data set with large rings that must be treated in this dissatisfying way.

The described model of ligand flexibility causes RMS deviations only in terms of conformational differences of less than 1 Å (e.g., methotrexate: 0.4 Å).

Modeling Receptor–Ligand Interactions

The model of molecular interactions used in FLEXX has been adopted from Böhm[17,18] and Klebe.[19] For each group that is able to form an interaction a special interaction geometry is defined: the position of a center and the shape of an interaction surface which is usually part of a sphere. Two groups have an interaction if the interaction center of each group is lying (approximately) on the interaction surface of the counter group (Fig. 3a). The interaction surface of the receptor group is approximated by a finite set of points, for algorithmic purposes. The modeling of short-range interactions has been extended in order to include hydrophobic interactions. Different types of interactions are arranged on three levels, from level 3 for highly directional bond such as H-bonds down to level 1 for directionally unspecific bonds such as hydrophobic interactions 1 (Table I). In placing the base fragment, FLEXX first tries to direct itself by interactions of high-level types. If there are not enough such interactions, the algorithm descends to lower-level interaction types.[20]

TABLE I. Interaction Types of FLEXX

Compatible interaction types	Interaction distance	$\Delta G_{neutral}$	ΔG_{ionic}	Level
H-acceptor/ H-donor	1.9 Å	−4.7 kJ/mol	−8.3 kJ/mol	3
Metal acceptor/ metal	2.0 Å	−4.7 kJ/mol	−8.3 kJ/mol	3
Aromatic-ring-atom, methyl, amide/aromatic-ring-center	4.5 Å	−0.7 kJ/mol	—	2
Aliphatic and aromatic carbon atoms, sulfur	4.5 Å	—	—	1

Selection of Base Fragments

The ligand is fragmented into components by severing at all acyclic single bonds. Then FLEXX automatically forms a set of alternative base fragments by selecting single components or combinations of them. The maximum number of different base fragments used in this test is four. This part of the algorithm has been described in more detail elsewhere.[8]

Placing Base Fragments

The base fragments are the first parts of the ligand that are placed into the active site. Actually two algorithms[7,8,21] are in use. The first one superposes triples of interaction centers of a base fragment with triples of compatible interaction points in the active site. If a base fragment has fewer than three interaction centers or if the number of placements is too low, the second algorithm, called line matching, is started. This one matches pairs of interaction centers with pairs of interaction points. Because of geometric ambiguity, multiple placements are generated by rotation around the axis defined by the interaction points and centers. Both base placement algorithms typically generate a large number of solutions. A reduction by clash tests and clustering follows. All important parameters of the base placement algorithm used in this report are listed in Table II.

Incremental Construction

Starting with the different base placements the complete ligand is constructed by linking the remaining components in compliance with the torsional database step by step. After adding one component new interactions are searched and the scoring function is used to select the best partial solutions which are used for the next extension step. The maximum number of solutions taken into account in the next iteration is $400 + 100n_f$, with n_f counting the different base fragments. The parameters which are used in the construction phase are given in Table II.

TABLE II. Parameters Used in the Docking Experiments

General	
Maximum allowed overlap volume between particular ligand and receptor atoms	2.5 Å^3
Maximum allowed average overlap volume between all ligand and receptor atoms	1.0 Å^3
Maximum number of different base fragments determined automatically	4
Placement of base fragments	
Maximum RMSD used for clustering placements of triangles of ligand interaction centers	1.1 Å
Maximum RMSD used for clustering placements of pairs of ligand interaction centers	0.4 Å
Incremental construction	
Number of solutions used in the next iteration step	400
Additional number of solutions for each base fragment used in the next iteration step	100
Maximum RMSD used for clustering placements	0.7 Å

Docking Covalently Bound Ligands

Covalent docking by FLEXX can only be managed if the covalent bond between the ligand and the receptor is manually specified. The ligand structure input file is extended by including the two atoms closest to the covalent bond on the receptor side. The first placement is performed by superimposing these atoms with their positions in the receptor structure. The normal incremental construction algorithm leads to the final placement. There are seven covalently bound complexes in the data set: 1aec, 1ase, 1blh, 1lpm, 1tpp, 3gch, and 4est.

Scoring Function

Ranking of the docking results is done with a modification of the scoring function developed by Böhm.[22]

$$\Delta G = \Delta G_0 + \Delta G_{rot} \times N_{rot} \quad (1)$$

$$+ \Delta G_{hb} \sum_{\text{neutral H-bonds}} f(\Delta R, \Delta \alpha) \quad (2)$$

$$+ \Delta G_{io} \sum_{\text{ionic int.}} f(\Delta R, \Delta \alpha) \quad (3)$$

$$+ \Delta G_{aro} \sum_{\text{aro int.}} f(\Delta R, \Delta \alpha) \quad (4)$$

$$+ \Delta G_{lipo} \sum_{\text{lipo. cont.}} f^*(\Delta R) \quad (5)$$

The first two terms (1) of the function are a fixed ground term (ΔG_0 = 5.4 kJ/mol) and a term taking into account the loss of entropy during ligand binding due to the hindrance of rotatable bonds (ΔG_{rot} = 1.4 kJ/mol). The following terms (2–4) are sums over all pairwise interactions (see Table I). The last part (5) of the scoring function rates the atom–atom contacts between protein and ligand, which are hydrophobic contacts and forbiddingly close contacts (clashes). The functions f, f* are heuristic distance and angledependent penalties (see refs. 7 and 22).

RESULTS

Preparation of Input Data

Since the publication of the complete docking algorithm[8] the preparation of input data has been changed and adapted to the further developments of FLEXX. Therefore, we summarize the whole process again. The ligand input files have been generated with SYBYL.[23] First, the ligand structure containing only non-hydrogen atoms has been extracted from the PDB file. This is the reference structure that is used for the calculation of RMSD values later on. By defining correct atom types including hybridization states, as well as correct bond types, adding hydrogen atoms, assigning formal charges to each atom, and finally energy-minimizing the structure we obtain the ligand input files. The energy minimization guarantees a low-energy conformation with suitable bond distances and angles. In principle it guarantees also that there is a completely new geometry with no "docking information" of the pdb structure. Generally, all carboxylic-acid and phosphoric-acid groups have been ionized, and all amino, amidino, and guanidino groups, but no amide groups, have been protonated.

The geometric input data of the receptor atoms which are taken from the pdb files are treated rigid during the docking procedure. Therefore, currently a study of an induced-fit docking is not possible with FLEXX. The preparation of the receptor input data requires the definition of the receptor atoms (via chain identifiers and hetero groups), the resolution of ambiguities in the PDB file (alternate location indicators, etc.), the determination of the positions of the essential hydrogen atoms, and the definition of the active site atoms. The assignments of the hydrogen positions are made on the basis of default rules except for the definition of the torsion angles at the hydroxyl groups of the amino acids serine, threonine, tyrosine, and the hydrogen position inside the histidine side chain. Here, suitable torsion angles and the optimal tautomeric histidine state, respectively, have been selected by visual inspection of the protein. The residues of lysine and arginine are protonated and the acid groups of aspartic and glutamic acid are ionized.

In order to define the active site of the protein all atoms are selected that are located no farther than 6.5 Å apart from an ligand atom at its crystalline position. In the case of the docking experiments with foreign ligands the active sites have been enlarged by using a distance criterion of 12 Å. This enlargement had only minor influence on the runtimes. The number of base placements certainly may become larger, but also the fraction of wrong construction paths leading to an early break-off is larger. Both effects compensate each other.

Generally, all water molecules of the PDB file have been removed, apart from the following exceptions: 1aaq and 4phv, both HIV proteases, here the water molecule HOH1 remains which is known to play a critical role in ligand binding[24]; 1lna and 1xie, here water molecules are bound to metal ions.

TABLE III. Statistics of the Docking Experiments in the Data Set

	Average	Maximum	Minimum
Run time	92.5 s	492.21 s	0.11 s
No. of solutions	245.84	624	2
RMSD of solution with rank 1	3.97 Å	15.10 Å	0.31 Å
Best RMSD of the ten highest-ranking solutions	2.98 Å	15.10 Å	0.22 Å
RMSD of solution with smallest RMSD	2.16 Å	14.76 Å	0.22 Å
Ligand size:			
No. of atoms	22.83	65	5
No. of components	9.35	37	1

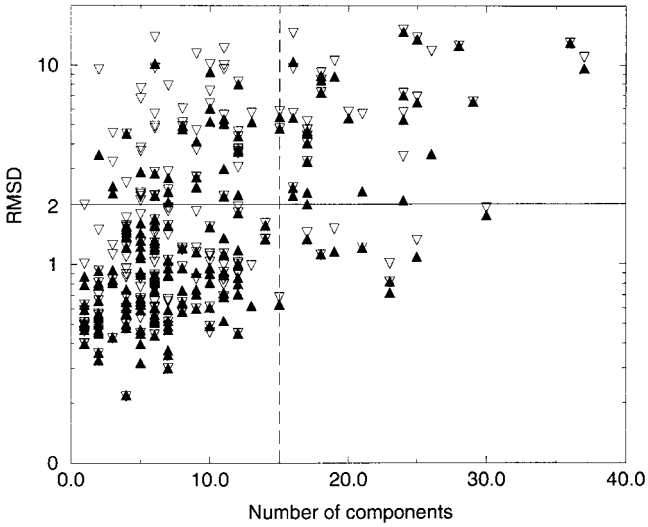

Fig. 1. Scatter plot of the RMSD values of all 198 examples that could be docked versus the number of components of the ligands. Solutions with best RMSD of the tenth highest-ranking solutions are represented by open triangles, and solutions with the best RMSD of the entire solution set are represented by solid triangles.

Results on the 200 Complexes Data Set

Table III gives initial statistics on the docking experiments. The parameter mainly determining the accuracy and run time of a docking experiment is the number of components of the ligand. This number varies between 1 to 37 in the data set. The average number of components is about 9. The docking experiments with FLEXX showed an average run time of 93 seconds on the SUN Ultra-30 workstation. The maximum run time of 8.2 minutes was used for the docking run of 1ack. Although the ligand of this complex (ethyl(3-hydroxyphenyl)dimethylammonium ion) has only four components, the run time is large because the more time-consuming second variant of the base placement algorithm has to be used on a base fragment with many hydrophobic interactions. On the other hand many complexes of the data set can be docked in less than 30 seconds.

Our analysis of the accuracy of the results is based on the RMS-deviation of the locations of all non-hydrogen ligand atoms to the crystal (RMSD) structure, in contrast to the subjective categories (good, close, errors, wrong) introduced by Jones et al.[12] A visual assessment of the FLEXX docking results was not feasible because of the large number of docking solutions (246 on the average per complex, see Table III). We consider results up to 2.0 Å RMSD acceptable docking results. This is in accordance with the classification of Jones et al.[12] who found only good or close solutions if the RMSD is below 2.0 Å. These acceptable solutions are in the range in which refinement to experimental accuracy with classical force-field methods should be possible.

The average RMSD of the highest-ranking solution of all docked examples of the data set is 3.97 Å. Owing to the FLEXX algorithm a docking result consists of a set of different solutions. Among these, very often solutions with lower RMSDs can be found at lower ranks. It is not a good idea to reject these lower-ranking solutions, because often it is feasible to filter out bad solutions among these by visual inspection or by extended computer methods. In any case, an analysis of lower ranks even with a rough scoring is worthwhile because of the great number of differing binding modi generated by FLEXX. The average RMSD of the solutions with best RMSD—disregarding rank—is 2.16 Å. A perhaps more practical restriction to the ten highest-ranking solutions leads to the still acceptable average RMSD of 2.98 Å.

The scatter plots of the best RMSD values of each solution set and the best RMSD values of the ten highest-ranking solutions of all examples versus the number of components of the ligands (Fig. 1) give an impression of the confidence that can be placed in a docking prediction generated by FLEXX. Most of the ligands with up to five components (91%) can be docked with acceptable accuracy (less than 2 Å), whereby 77% are among the ten highest-ranking solutions. In the range of six to 15 components 73% of the examples can be reproduced with this accuracy. Ligands with more than 15 components tend to generate wrong solutions (about 75% of the examples).

In Figure 2 we show an accumulated plot of the rate of those examples placed with an RMSD of less than or equal to a certain constant value. In total, 51% of the examples can be docked with an RMSD of less than 1.0 Å and 70% with an RMSD of less than 2.0 Å. The rate of acceptable docked (RMSD $\leq$ 2.0 Å) examples rises to about 80% if one only looks at those with ligands with fewer than 16 components.

Table IV gives an overview of the number of complexes that can be docked within a certain RMSD range. Summarizing the numbers up to an RMSD of 2.0 Å leads to a rate of successful dockings of 46.5% if one considers deviations of the rank 1 solutions. This rate rises to 70% if the solutions with best RMSD are taken into account. In addition, an average rank of the solution with best RMSD is given. The average rank of the examples with an RMSD of less the 2.0 Å is about 40. In the following, we will augment these statistics with discussions of selected ex-

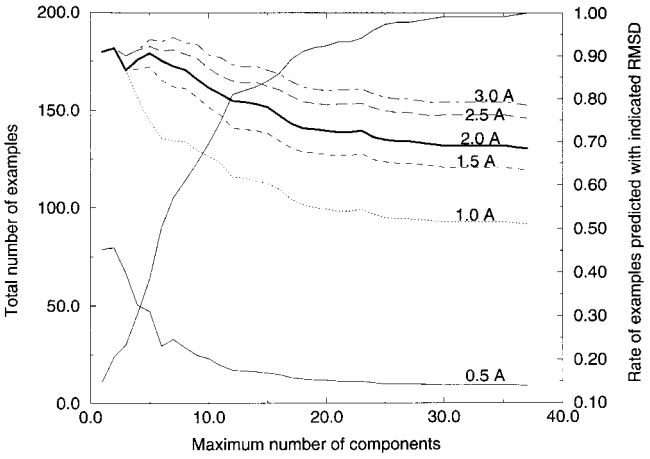

Fig. 2. Accumulated plots of the rate of examples that have solutions with RMSD values below 0.5 Å, 1.0 Å, 1.5 Å, 2.0 Å, 2.5 Å, and 3.0 Å versus the number of components of the ligands. Additionally the total number of examples is shown.

amples. We will analyze acceptable solutions as well as give reasons for the generation of unacceptable predictions and cases in which the FLEXX docking algorithm was not able to find any placements.

Examples in Which FLEXX Finds Acceptable Solutions

1aaq

FLEXX divides the peptide-like ligand of this HIV protease complex into 23 components. This size is above the aforementioned limit for reliable predictions of the algorithm. This is supported by the observation that a small change of the geometry, namely the inclusion of HOH1 to the receptor structure, makes the difference between a totally wrong placement and a nearly perfect one. Figure 3b shows the crystal structure in comparison with the docking solutions on rank 1 and 2, and the included HOH1. The RMSD values are 1.71 Å and 1.03 Å, respectively. Without the presence of the water molecule the algorithm is not able to complete the correct construction path after two-thirds of the ligand fragments have been placed. At this point the subsequent fragments are leading to clashes, and only solutions with very large RMSD survive.

The inclusion of HOH1 seems to support some correct interim solutions leading ultimately to correct placements, although the energy contribution of the H-bond between ligand and water molecule is not very important.

1acj

The complex of acetylcholinesterase with tacrine has been reproduced with an RMSD of 0.64 Å by the solution at rank 2 (see Fig. 3c). At rank 1 FLEXX generates a placement that might be an alternative binding mode. In the crystal structure and in solution 2 the NH-group within the ring system of tacrine binds to HIS 440, whereas in solution 1 the ligand is flipped by 180° such that the amino group at the ring system is able to form the hydrogen bond to HIS 440.

1tdb

The first acceptable solution of the thymidylate synthase complex 1tdb has rank 3 and an RMSD of 1.93 Å (Fig. 3d). The binding mode of the ligand 5-fluoro 2′-deoxyuridine 5′-monophosphate differs mainly in the ribose ring system. Both better-ranking solutions have different binding modes (for solution 1, see Fig. 3d) and very large RMSD values (10.3 Å and 10.2 Å). They do not occupy the pocket of the uridine moiety but form one additional hydrogen bond. The lack of a penalty term in the scoring function for unfilled holes in the receptor is the major reason of the wrong ranking.

1abe

In the case of small sugar molecules like l-arabinose the FLEXX docking algorithm usually finds the correct pocket but gets into difficulties in estimating the exact binding mode because of the flexibility of both the ring system and its hydroxyl groups. In the docking experiment with the l-arabinose–binding protein complex with l-arabinose (1abe) all placements of the solution set have nearly the same center of mass and differ only in the ring conformation and the torsion angles of the hydroxyl groups. The highest-ranking solution has an acceptable RMSD of 1.80 Å. A better placement with a correct binding mode can be found at rank 3 (0.90 Å RMSD). The solution at rank 2 shows a large RMSD of 3.03 Å. All three geometries can be compared with the crystal structure in Figure 3e. Here the values of the scoring function lie in a narrow interval of 0.6 kJ/mol.

1glp

The complex glutathione-S-transferase with glutathione sulphonic acid is a very nice example of a successful docking. With 12 components the ligand size is in the range for which the algorithm typically generates reliable predictions. The best-ranking solution (see Fig. 3f) is also the solution with the lowest RMSD (0.45 Å) of the solution set. The binding mode is completely identical with the crystal structure.

1cbs

The docking of the complex 1cbs was quite difficult before the concept of multi-level interactions was introduced into the FLEXX base placement algorithm. The ligand, retinoic acid, is mainly hydrophobic and has only two polar interacting atoms, the oxygen atoms of the carboxylate group. One of them binds to Tyr 134 and Arg 132. Using the hydrophobic aromatic ring at the opposite side of the molecule as the decisive base fragment, the algorithm finds a good placement (Fig. 6a) with an RMSD of 1.4 Å. This solution includes an exact reproduction of the binding mode of the carboxylate group.

TABLE IV. Results of the Docking Experiments in the Data Set

RMSD	No. of examples with rank 1	No. of examples with any rank	Average rank of best RMSD sol.	Examples with any rank
≤0.5	6	25	17.4	1abe 1abf 1acj 1ack 1aco 1dbj 1dwb 1glp 1lah 1lst 1mbi 1mld 1mmq 1mrg 1pbd 1phd 1tng 1tni 1wap 2ada 2cpp 2gbp 2phh 5abp 6abp
>0.5, ≤1.0	27	77	53.4	121p 1aaq 1acm 1aha 1ase 1atl 1avd 1azm 1cbx 1com 1coy 1cps 1dbb 1dbk 1dr1 1dwc 1fki 1frp 1hfc 1hsl 1hti 1hyt 1imb 1ivb 1ivd 1ivf 1lcp 1ldm 1lna 1mdr 1nis 1nsc 1phf 1ppc 1rbp 1rnt 1stp 1tmn 1tnh 1tnj 1tnk 1tnl 1tph 1trk 1ukz 1ulb 1xid 1xie 2ak3 2cgr 2cmd 2ctc 2lgs 2mth 2pk4 2sim 2tmn 2xis 2yhx 2ypi 3aah 3cpa 3ptb 3tpi 4cts 4dfr 4fxn 4hmg 4tim 4tln 4ts1 5cpp 5p2p 5tim 6rsa 7tim 8atc
>1.0, ≤1.5	31	25	17.4	1ake 1blh 1byb 1cbs 1cde 1cil 1dbm 1dwd 1epb 1fen 1ghb 1ivc 1ive 1ppk 1rds 1slt 1tka 1tpp 1tyl 2dbl 2mcp 2r04 4est 4phv 6tim
>1.5, ≤2.0	29	11	85.8	1ctr 1ela 1hgi 1lpm 1pso 1rob 1tdb 1thy 2cht 3gch 4fbp
>2.0, ≤2.5	6	15	92.4	1die 1eta 1hgh 1hgj 1icn 1mup 1poc 1pph 1ppl 1tlp 2r07 4fab 4tmn 6rnt 6tmn
>2.5, ≤3.0	10	4	83.2	1did 1mrk 1snc 1srj
>3.0	89	41	85.0	1aec 1apt 1baf 1bbp 1bma 1cdg 1eap 1eed 1elb 1elc 1eld 1ele 1etr 1fkg 1glq 1hdc 1hef 1hgg 1hri 1hvr 1ida 1igj 1lic 1lmo 1mcr 1nco 1phg 1ppi 1ppm 1rne 2er6 2plv 3cla 3hvt 4hvp 5cts 5tmn 6cpa 7cpa 8gch 9hvp

Examples in Which FLEXX Finds no Solutions

There are two complexes for which the FLEXX docking algorithm finds no solutions: 1hdy and 1pha. Furthermore, the cross-docking of 7cpa fails. Owing to the larger active site used for cross-docking, there is a larger number of configurations to be considered, which, because of the greedy nature of the algorithm, eventually leads to docking failure. The case of 7cpa will be discussed in the next section.

In principle, there are several reasons why the FLEXX docking algorithm may not find solutions. The first used to be a lack of interacting groups of the ligand. However, the concept of multilevel interactions now prevents these problems in all test cases.

In some cases the restrictions of the model concerning the geometries of H-bonds or the internal torsion angles lead to placements that overlap with the receptor, although the structure found by the algorithm is close to the crystal. For instance, in the complexes 1hdy and 1pha FLEXX generates only clashing placements of the base fragments.

A further reason why docking of 1hdy fails is a forbiddingly close overlap of the ligand and the cofactor NAD^+ even in the crystal structure.

Examples in Which FLEXX Finds Only Unacceptable Solutions

1hvr

This is another HIV protease complex. Its ligand, a 1,3-diazepin-2-one derivate, consists of a central seven-membered ring which leads to four-ring conformations in the geometry generation with CORINA. The FLEXX prediction with the best score has an RMSD of more than 10 Å. Its main failure is that both naphthyl groups are sticking out of the protein pocket.

Solution 14 which has an energy lying 2 kJ/mol above the highest-ranking solution is the first placement with a much smaller RMSD (3.65 Å). This solution is an alternative binding mode; a naphthyl and a phenyl group have changed their hydrophobic pockets (see Fig. 6b). Although the scoring of this solution has a large lipophilic contact energy contribution, the overall value of the solution with rank 1 is smaller because there are additional h-bonds and there is no term in the actual scoring function that punishes for large hydrophobic groups reaching out into the solvent.

An interesting observation is worth mentioning: using the ring conformation of the crystal structure leads to an excellent placement with an RMSD of only 0.82 Å at rank 1. Although one of the four Corina-generated conformations is close to the crystal structure, the additional degree of freedom during the base placement seems to cause the wrong solution.

7cpa

The ligand of the carboxypeptidase A complex 7cpa is peptide-like and has one phosphate group. Its size (18 components) is at the margin of the range for reliable docking predictions. The algorithm starts with four different base fragments and finds excellent positions for them with RMSD of less than 1 Å. After the addition of eight fragments in the ligand construction phase, the generated partial solutions close to the crystal structure are filtered out because of bad scoring or forbiddingly large clashes with the receptor. The solution that is finally generated with rank 1 has an RMSD of 9.3 Å. Figure 6c shows that the hydrophobic pockets are occupied by the phenyl groups of the ligand, but in a modified order.

The reason for misplacing this ligand is a combination of bad ranking and the greedy character of the construction

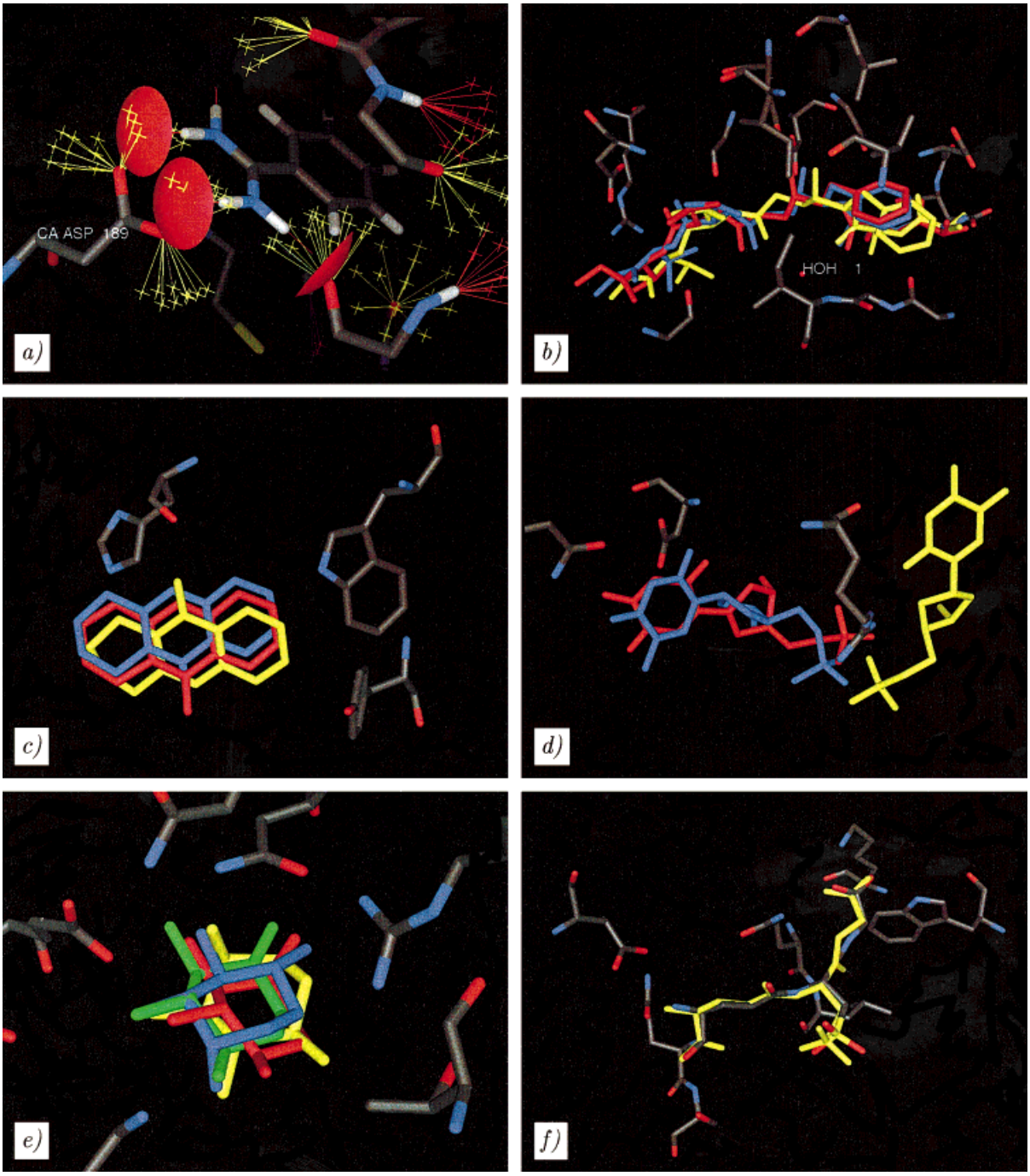

Fig. 3. **a:** Interaction of the amidino group of benzamidine with the carboxylate group of Asp 189 in tryspsin (3ptb). The oxygen atoms of Asp 189 are lying on the red interaction surfaces of the hydrogen atoms of N1 and N2. Vice versa the hydrogen atoms reside among the yellow interaction point clouds of the oxygen atoms. **b:** Docking of 1aaq; crystal structure in red, solution with rank 1 in yellow, solution with rank 2 in blue, included HOH1 in red, active site amino acids colored by atom type. **c:** Docking of 1acj; crystal structure in red, solution with rank 1 in yellow, solution with rank 2 in blue, active site amino acids colored by atom type. **d:** Docking of 1tdb; crystal structure in red, solution with rank 1 in yellow, solution with rank 3 in blue, active site amino acids colored by atom type. **e:** Docking of 1abe; crystal structure in red, solution with rank 1 in yellow, solution with rank 2 in blue, solution with rank 3 in green, active site amino acids colored by atom type. **f:** Docking of 1glp; ligand crystal structure and active site amino acids colored by atom type, solution with rank 1 in yellow.

algorithm. The complex 7cpa is also a part of the cross-docking experiments. But here we use larger active sites (all receptor atom in 12 Å surroundings of the ligand) associated with a larger complexity of the docking problem. FLEXX does not generate any docking solution. The ligand construction algorithm has to handle a great num-

ber of alternative partial solutions which, at least, all lead to clashes with the receptor.

1srj

The ligand 2-((4′-hydroxynaphthyl)-azo)benzoate has been docked into streptavidin with an RMSD of 8.03 Å for the solution at rank 1. The solution at rank 154 (see Fig. 6d) has an RMSD of 2.9 Å. This is worse than the limit we had set for unacceptable solutions, but, to some extent, this placement is close to the experimental structure. Nevertheless, there are some structural features in the experimental structure of the complex which are unusual and which cannot be generated by FLEXX (Fig. 6d). There are clashes between the oxygen atom at the naphthyl group of the ligand and the backbone atoms of Asn 49 (e.g., there is an atom–atom distance of 2.18 Å to the carbon atom).

Here, structural problems in the experimental structure may be the reasons for the misplacement. More examples of poorly determined geometries of ligands in proteins can be found in references 7, 11, and 12.

1fkg

The ligand of this complex of the FK506-binding protein consists of three medium-sized hydrophobic groups bonded to a hydrophilic core. In this core there are four hydrogen acceptor atoms (all oxygen atoms) that bind to the receptor (Tyr 82 and Ile 56). The docking solution at rank 1 shows a similar but different pattern of hydrogen bonds of these atoms. Therefore, the hydrophobic groups are placed in the wrong hydrophobic pockets, as well, which can be seen by comparison of the experimental and calculated ligand geometry shown in Figure 6e. The resulting RMSD is 5.9 Å. There are no obviously better placements in the generated solution set. A reason seems to be the large variety of different placements of the hydrophilic core which enlarges the complexity of this docking problem.

Cross-Docking Experiments

The following seven proteins have been selected for this experiment: trypsin, α-thrombin, carboxypeptidase A, cytochrome P-450, a FAB fragment, L-arabinose-binding protein, and triosephosphate isomerase. The pdb codes of the complexes containing theses proteins are listed in Table V. There are a few more groups of complexes with identical proteins in the data set, but they have been disregarded here, because they either had only two members or most of the examples could only be docked with large errors.

All ligands of one group have been docked into all protein structures of the same group. For this purpose all receptor structures used for the cross-docking experiment have been superposed with the first structure in the list. Table V shows the RMSDs of the highest-scoring docking solutions of a ligand in its own complex structure and the best result in a foreign structure. The differences of these values, given in the last column, indicate (by positive values) whether the docking gets better by use of a foreign structure. Only in the case of the ligand of 1dbj are the docking solutions with rank 1 of all foreign structures much worse than the original docking result. Three-fourths of the examples can be docked into at least one foreign structure equally well or better. The results improve significantly in the case of 1cbx, 1cps, 1phf, 2cpp, and 4tim. Here we obtain acceptable placements with foreign structures, whereas the docking solution in the own structures are wrong. On the one hand, this is a scoring problem. As Table IV shows, 1cbx, 1cps, 1phf, 2cpp,

TABLE V. Cross-Docking Results

Original complex		Best docking receptor structure		
pdb-code	RMSD [Å]	RMSD [Å]	pdb-code	RMSD difference [Å]
Trypsin				
1ppc	2.800	2.830	(1tnk)	−0.030
1pph	4.550	3.270	(1tnl)	1.280
1tng	0.320	0.490	(1tni)	−0.170
1tnh	0.690	0.460	(1tni)	0.230
1tni	2.770	2.360	(1tnl)	0.410
1tnj	1.280	0.620	(1tni)	0.660
1tnk	1.840	1.510	(1tnl)	0.330
1tnl	0.620	0.480	(1tnj)	0.140
3ptb	0.430	0.340	(1ppc)	0.090
α-Thrombin				
1dwb	0.470	0.490	(1dwc)	−0.020
1dwc	1.370	1.880	(1dwd)	−0.510
1dwd	1.010	1.780	(1dwc)	−0.770
Carboxypeptidase A				
1cbx	5.220	1.140	(2ctc)	4.080
1cps	4.850	0.780	(1cbx)	4.070
2ctc	1.980	1.880	(1cbx)	0.100
3cpa	2.540	1.840	(2ctc)	0.700
6cpa	6.030	4.290	(2ctc)	1.740
7cpa	—	—	—	—
Cytochrome P-450				
1pha	17.950	17.300	(1phg)	0.650
1phd	0.770	0.650	(1pha)	0.120
1phf	4.230	1.900	(1pha)	2.330
1phg	17.580	5.080	(1pha)	12.500
2cpp	2.920	0.600	(1phg)	2.320
5cpp	1.600	0.720	(1phd)	0.880
FAB fragment				
1dbb	0.810	0.910	(1dbk)	−0.100
1dbj	1.410	4.340	(1dbk)	−2.930
1dbk	1.140	0.390	(1dbj)	0.750
1dbm	2.070	2.360	(1dbj)	−0.290
2dbl	1.640	1.150	(1dbm)	0.490
L-Arabinose–binding protein				
1abe	1.800	0.600	(5abp)	1.200
1abf	0.680	0.600	(1abe)	0.080
5abp	1.320	1.310	(1abf)	0.010
Triosephosphate isomerase				
4tim	3.990	0.770	(6tim)	3.220
5tim	1.990	1.200	(4tim)	0.790
6tim	1.620	1.420	(5tim)	0.200

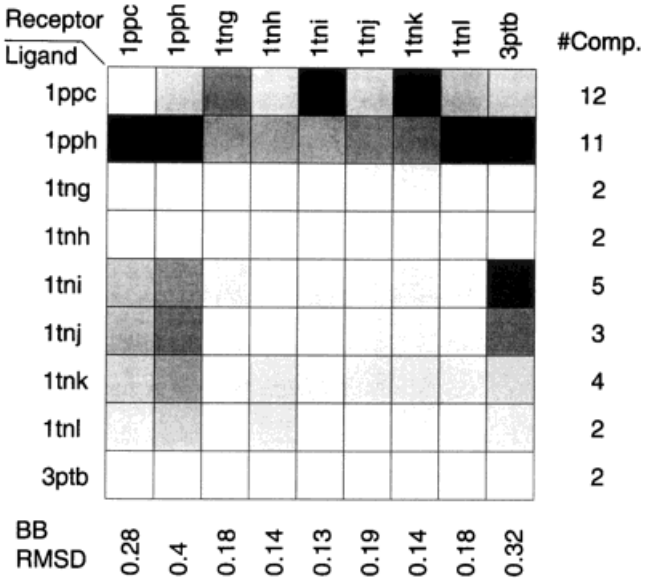

Fig. 4. All-pairs docking experiment for nine trypsin complexes. For gray-scaling of RMSD values, see Figure 5.

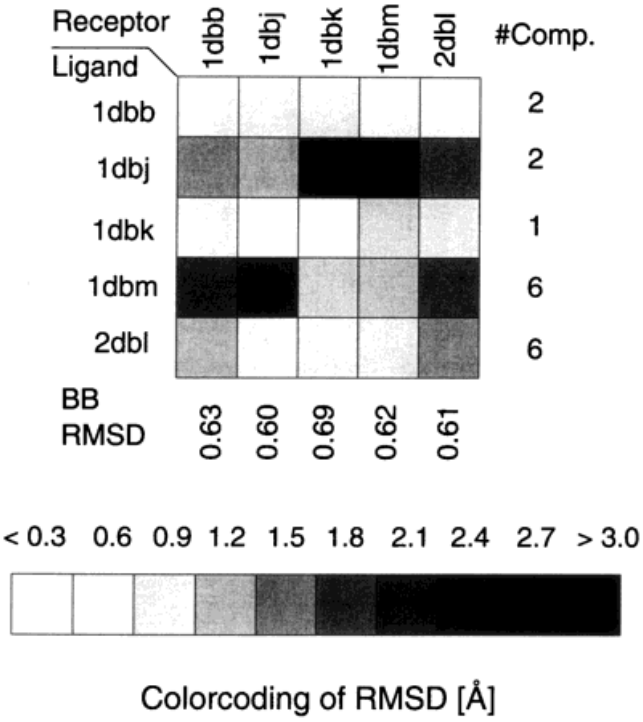

Fig. 5. All-pairs docking experiment for five FAB complexes and gray-scaling of RMSD values.

and 4tim can be docked with RMSDs less than 1 Å at a lower rank. On the other hand, the results demonstrate that it is worthwhile to use many available protein structures in a screening experiment.

Figures 4 and 5 show the RMSD values of the solutions with lowest RMSD of the cross-docking experiments of all complexes with trypsin and with the FAB fragment in color-coded representations. In addition, at the bottom of each receptor column the average backbone RMSD to the other receptor structures are given. These values are about 0.5 Å and are never above 1 Å. Again the diagrams point out the type of problem arising in docking experiments with FLEXX. Special ligands (represented by rows in the matrices) tend to cause misplacements (represented by dark colors) with many receptor structures. For example, the ligands of 1ppc and 1pph cannot be placed into approximately 30% of the foreign receptor structures (Fig. 4). Their sizes (12 and 11 components, respectively) are not exceedingly large. The misplacements with the foreign structures probably stem from the greedy construction algorithm and the scoring problem. In the case of the complexes with the FAB fragment there is the same tendency to obtain somewhat worse placements for the ligands with six components (1dbm and 2dbl) but also for the ligand of 1dbj with only two components.

Nevertheless, most of the placements are acceptable, with one exception: 1pph in its own receptor structure. The placements with best RMSD of the ligand of 1pph 3-TAPAP in the receptor structure of 1pph (rank 174, 2.44 Å RMSD) and in the receptor structure of 1tni (rank 8, 1.17 Å RMSD) are shown in Figure 6f. In both solutions the benzamidino group binds in very good agreement with the crystal structure to Asp 189. The lack of the solution generated in the 1pph structure is a misplacement of the hydrophobic groups of the ligand. Only in the 1tni-structure the methyl-phenyl group and the piperidino group are lying at the correct position. A reason might be a change of the side chain torsion angle at Leu 99 which is close to the the hydrophobic groups of the ligand. In the remaining part of the pocket there are only marginal differences between both protein structures.

Screening Experiments

The proteins of ten complexes with experimentally known strong binding activities have been chosen to perform screening experiments on a ligand library of 556 molecules. All of these ten complexes taken out of our data set have nanomolar binding constants (which means ΔG values of about -50 kJ/mol). The examples have been chosen such that there are no similar or identical proteins. The ligand library consists of 200 ligand structure files out of the complex data set and of a set of 356 ligand files taken from the CSD library by searching for the keyword "drug" (G. Klebe, personal communication). The average run time over all 5,560 docking runs has been 190 seconds per protein–ligand pair. In Table VI the results of the screening experiments are summarized.

In this experiment we wanted to investigate whether FLEXX is able to find the original ligand of the complex as the strongest or at least as one of the strongest binding compounds of our library. There were five complexes where our expectations have been fulfilled. The original ligands of 1ake, 1dwd, and 4dfr are the best binding ones in the

TABLE VI. Results of the Screening Experiments[†]

			Docking of the original ligand		Ligand with best docking score	
Pdb code	$\Delta G_{exp.}$	$\Delta G^{FlexX}_{exp.struct.}$	ΔG	Rank (%)[a]	Pdb code/ Lig. no.	ΔG
1ake	−49.07	−118.94	−103.31	1 (0.2)	—	—
1dwd	−48.60	−30.95	−47.62	1 (0.2)	—	—
1hvr	−54.25	−38.85	−20.29	284 (51.1)	c291	−46.85
1hvr[b]	−54.25	−38.85	−46.85	45 (8.1)	c291	−46.85
1pso	−58.99	−31.81	−40.04	5 (0.9)	1glq	−45.78
1rne	−53.63	−24.41	−22.92	153 (27.5)	c383	−41.01
1stp	−76.45	−33.38	−29.95	18 (3.2)	c010	−43.49
2ada	−73.00	−40.76	−33.38	3 (0.5)	c209	−37.63
4dfr	−55.34	−54.70	−66.36	1 (0.2)	—	—
4tmn	−58.13	−35.65	−37.91	35 (6.3)	c383	−54.20
6cpa	−65.72	−44.21	−42.25	34 (6.1)	1glp	−56.24

[†]All ΔG values in kJ/mol.
[a]In parentheses: fraction of high scoring molecules on the total data set which includes the original ligand.
[b] In this calculation the experimental ring structure of the 1hvr ligand has been used.

corresponding proteins. In 1pso and 2ada the original ligands of the complexes have rank 5 and rank 3, respectively. This means that, by a screening run with FLEXX, the fraction of compounds of the ligand library that should be analyzed experimentally could be reduced to less than 1% of the whole ligand database. For the complexes 1stp, 4tmn, and 6cpa this amount is less than 7%. Even these examples show a very good enrichment, and screening with FLEXX thus proves its worth in a drug discovery process.

Another aspect which can be analyzed with the data of Table VI is the energy. In columns 2, 3, and 4 the experimental binding energy ΔG_{exp}, and the FLEXX scores of the crystallographic ligand structure $\Delta G^{FlexX}_{exp.struc.}$ and the first docking solution ΔG are shown. Although the scoring function was useful in distinguishing roughly between good and bad inhibitors, the absolute values often show large differences (about 20 kJ/mol) from the experimental values in both directions. There are also differences in both directions between the scoring values of the crystallographic structure and the first docking solution. A worse docking energy indicates that the algorithm was not able to find a good docking solution, e.g., in the case of 1hvr. A worse score of the crystallographic structure can be an effect of the inaccuracy of the FLEXX scoring function as well as of an error in the experimental structure. In the following we will discuss some details of the screening experiments which cannot be derived from Table VI alone.

1ake

Although the rank of the original ligand is best there is a very large difference (50 kJ/mol) between the experimental binding energy and that estimated by the scoring function. We see two possible reasons. First the compound (see Fig. 7) consists of five phosphate units, all of which are deprotonated. Although this seems to be unrealistic we did so, because currently FLEXX is not able to handle tautomeric structures of a partially protonated molecule. Therefore, the energy is overestimated by several ionic hydrogen bonds. The second reason may be a movement of the residue of Arg 167 during complexation. This is indicated by two alternative locations of this amino acid in the pdb data. The movement builds up strong interactions to the ligand but simultaneously breakes some salt bridges within the receptor which therefore are not included into the energy calculation.

1hvr

The original ligand (Fig. 7) of the HIV-1 protease complex has rank 284 for a badly placed docking solution. As mentioned above we get a much better docking result when using only the experimental conformation of the central 1,3-diazepin-2-one ring (rank 45). The best binding ligand of the library is the tetra peptide c291 (Fig. 7). A further HIV-1 protease inhibitor can be found at rank 44: the ligand of 4hvp.

1pso

Pepstatin (Fig. 7), the ligand of the pepsin complex 1pso, has rank 5. The energy of the best-binding ligand (Fig. 7) only differs by 5 kJ/mol. This ligand is a glutathione derivate which inhibits the glutathione-S-transferase complex 1glq.

1rne

The original ligand of the renin complex 1rne is a transition state analogon (Fig. 7). Its rank in the screening experiment is only 153 because the docking algorithm does not find the correct placement. The best-binding ligand of the library is the pentapeptide c383 (Fig. 8) whose score is better by 18 kJ/mol.

1stp

A bad energy estimate combined with a very accurate placement by docking is a well-known discrepancy[7] in the

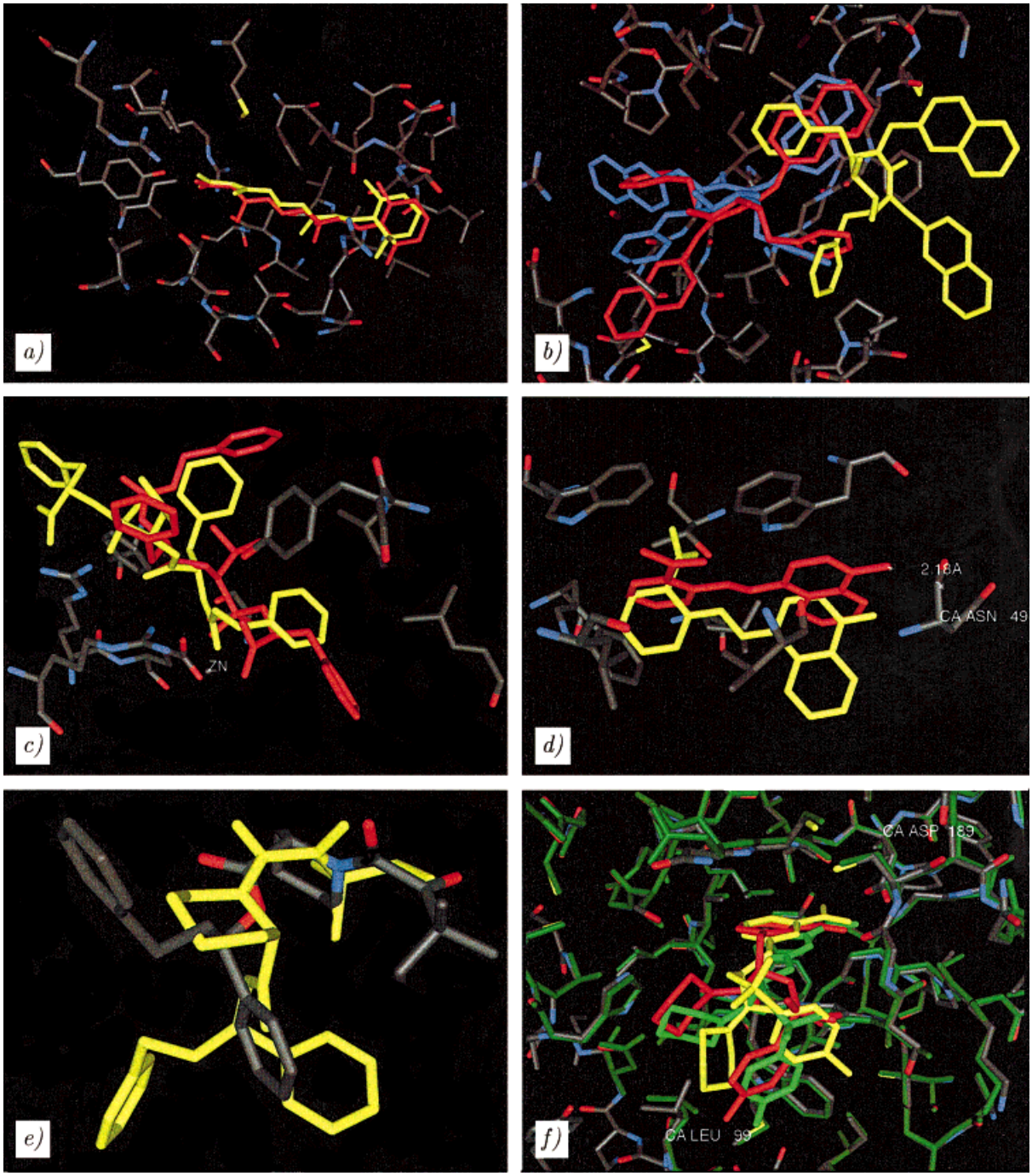

Fig. 6. **a:** Docking of 1cbs; crystal structure in red, solution with rank 1 in yellow, active site amino acids colored by atom type. **b:** Docking of 1hvr; crystal structure in red, solution with rank 1 in yellow, solution with rank 14 in blue, active site amino acids colored by atom type. **c:** Docking of 7cpa; crystal structure in red, solution with rank 1 in yellow, active site amino acids colored by atom type. **d:** Docking of 1srj; ligand crystal structure in red, active site amino acids colored by atom type, solution with rank 1 in yellow. **e:** Docking of 1fkg; ligand crystal structure colored by atom type, solution with rank 1 in yellow. **f:** Docking of the ligand of 1pph into its own receptor structure (site amino acids colored by atom type) and the receptor structure of 1tni (site amino acids in dark green); ligand crystal structure in red, solution (docked into 1pph-structure) with rank 174 in yellow, solution (docked into 1tni-structure) with rank 8 in green.

case of the streptavidin–biotin complex 1stp. But rank 18 is still a good result in this screening experiment. Interestingly the best-binding ligand of the library c010 (Fig. 7) shows the same ureido moiety as biotin (Fig. 7). And as one expects, the placements show a great similarity, with both ureido groups lying exactly at the same place.

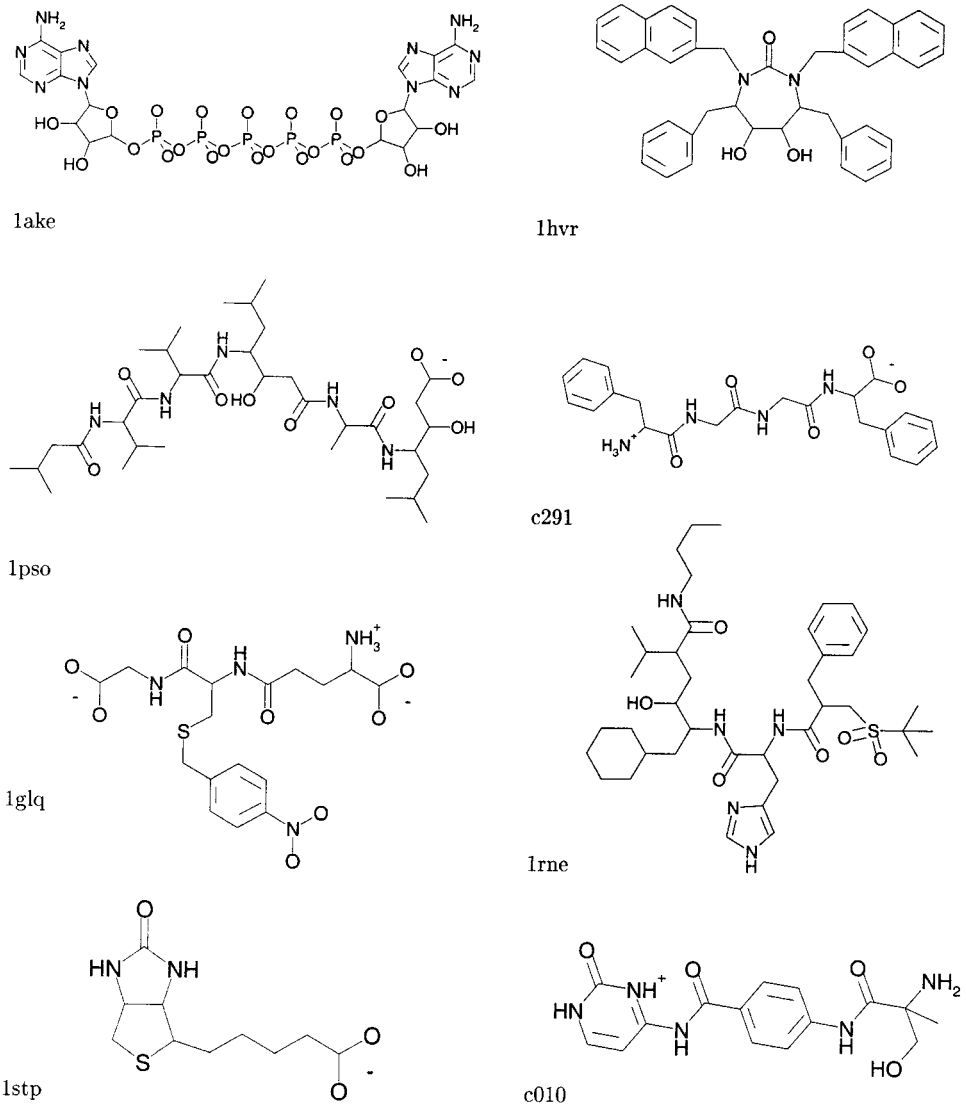

Fig. 7. Structural formulas of some ligands of the dataset.

2ada

The original ligand of this adenosin-deaminase is a hydroxyl purine derivate (Fig. 8). It could be found at rank 3 in this screening experiment. The best-binding ligand with a 4 kJ/mol better binding energy is c209 (Fig. 8). This pterin derivate can be placed by FLEXX in a way that allows the build-up of a similar number of hydrogen bonds as with the purine group.

4tmn

In the complex 4tmn thermolysin is inhibited by a ligand with a central phosphonate group (Fig. 8) which is bonded to the Zn ion of the active site. Despite this strong binding moiety of the ligand the docking algorithm does not find the correct receptor-ligand interactions. A much better scoring ligand is the penta peptide c383 (Fig. 8), which as in the renin case is no candidate for an inhibitor of course.

6cpa

6cpa is a very strong complex of carboxypeptidase A and a peptidic ligand with a central phosphonate group (Fig. 8). Because of a bad docking result with a RMSD of 7 Å FLEXX obtains a bad score and a rank of 34. A much better binding energy is achieved by glutathionic acid (Fig. 8), which is the ligand of the glutathione-S transferase complex 1glp.

CONCLUSIONS AND FURTHER DEVELOPMENTS

In total, 46.5% of the complexes of the data set can be reproduced with an acceptable accuracy by a FLEXX docking solution at rank 1. This fraction rises to 70% if we consider the best placement of each generated solution set. This shows that there is a problem with the correct ranking of the solutions. The scoring function has been chosen because it is relatively accurate in estimating

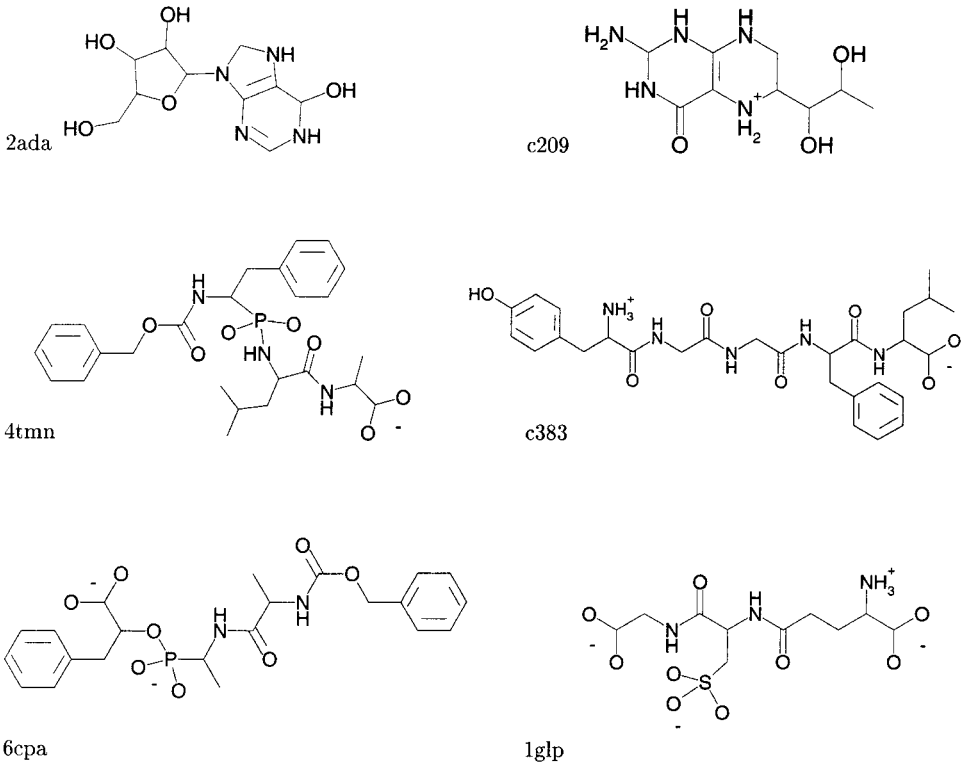

Fig. 8. Structural formulas of some ligands of the dataset.

binding energies of protein–ligand complexes and it is fast and easy to compute. The latter feature and the high efficiency of the algorithm are the reason for the extremely short runtimes of the docking experiments: 93 seconds on average.

Nevertheless, changes or extensions of the scoring functions are necessary. One concept, which has been discussed,[25] is filter functions. These functions are used to extract placements with "undesirable properties," which are not shared by experimentally observed structures. Such undesirable properties are, e.g., insufficient steric complementarity between ligand and protein and large lipophilic regions of the ligand being exposed to the solvent. After filtering the FLEXX solution set, the remaining low-energy placements mostly have small RMSD values. By this method one can only improve the highest-ranking solution or the rank of the best placements, respectively. New good solutions can only be generated by implementing suitable filters into the construction algorithm which is, in fact, a part of our actual development on FLEXX. After completing this development a new test will lead to a decision as to whether the greedy strategy has to be reconsidered or not.

A further aspect which leads to problems in the docking experiments is the fact that water molecules can mediate binding between ligand and receptor atoms. In this study we have added explicit water molecules to the receptor structure in some cases. A more elegant procedure which is able to place additional water molecules if they are necessary during the docking algorithm has already been implemented and tested.[26] Furthermore we are developing alternative docking algorithms for base fragments.

The cross-docking experiments indicate that, in most cases, FLEXX is able to generate similar docking results by use of the foreign receptor structure. Sometimes the docking results in foreign receptor structures are even closer to the experimental geometry than in the original structure. As long as the receptor is treated as rigid it seems to be a good idea to use more than a single receptor structure in screening experiments. Despite many shortcomings of the actual scoring function within the screening experiments FLEXX showed a reasonable enrichment behavior when tested with a medium-sized ligand library.

ACKNOWLEDGMENTS

The authors thank Manfred Hendlich and Gerd Klebe (both University of Marburg) for making a large set of ligand structure files available and Daniel Hoffmann (GMD SCAI) and Martin Stahl (Hoffmann-La Roche) for helpful comments on this work. This work was performed as part of the RELIMO-Project, which is funded in part by the German Federal Ministery for Education, Science, Research and Technology (BMBF) under grant no. 0311620 and the participating industrial partners Merck Preclinical Research (Darmstadt, Germany) and Boehringer Ingelheim Pharmaceutical Research (Ingelheim, Germany).

ADDENDUM

The FLEXX software package is available for SUN, SGI, and PCs running the Linux operation system. Interested readers should visit our WWW page http://cartan.gmd.de/FlexX or contact the corresponding author. The input files of the presented data set will be made available soon on the FLEXX web-pages.

REFERENCES

1. Blaney JM, Dixon JS. A good ligand is hard to find: automated docking methods. Perspect Drug Discovery Design 1993;1:301–319.
2. Colman PM. Structure-based drug design. Curr Opin Struct Biol 1994;4:868–874.
3. Kuntz ID. Structure-based strategies for drug design and discovery. Science 1992;257:1078–1082.
4. Lengauer T, Rarey M. Computational methods for biomolecular docking. Curr Opin Struct Biol 1996;6:402–406.
5. Lybrand TP. Ligand-protein docking and rational drug design. Curr Opin Struct Biol 1995;5:224–228.
6. Rarey M, Kramer B, Lengauer T. Time-efficient docking of flexible ligands into active sites of proteins. In: Rawlings C, et al., editors. Proceedings of the Third International Conference on Intelligent Systems in Molecular Biology. Menlo Park, CA: AAAI Press; 1995. p 300–308.
7. 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.
8. Rarey M, Kramer B, Lengauer T. Multiple automatic base selection: protein-ligand docking based on incremental construction without manual intervention. J Comput Aided Mol Design 1997;11: 369–384.
9. Moult J, Hubbard T, Bryant SH, Fidelis K, Pedersen JT. Critical assessment of methods of protein structure prediction (casp): Round ii. Proteins Supplement 1 1997;1(1):2–6.
10. Dixon JS. Evaluation of the casp2 docking section. Proteins Supplement 1; 1997;1(1):198–204.
11. Kramer B, Rarey M, Lengauer T. Casp-2 experiences with docking flexible ligands using flexx. PROTEINS Supplement 1, 1997;1(1): 221–225.
12. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267:727–748.
13. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The protein data bank: a computer based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
14. Allen FH, Bellard S, Brice MD, Cartwright BA, Doubleday A, Higgs H, Hummelink-Peters T, Kennard O, Motherwell WDS, Rodgers JR, Watson DG. The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information. Acta Crystallograph 1979;B35:2331–2339.
15. Klebe G, Mietzner T. A fast and efficient method to generate biologically relevant conformations. J Comput Aided Mol Design 1994;8:583–606.
16. Sadowski J, Gasteiger J. From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. Chem Rev 1993;93:2567–2581.
17. Böhm H-J. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. J Comput Aided Mol Design 1992;6:61–78.
18. Böhm H-J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. J Comput Aided Mol Design 1992;6:593–606.
19. Klebe G. The use of composite crystal-field environments in molecular recognition and the de-novo design of protein ligands. J Mol Biol 1994;237:221–235.
20. Rarey M, Kramer B, Lengauer T. Docking of hydrophobic ligands with interaction-based matching algorithms. Comp Appl Biol Sci 1999; in press.
21. Rarey M, Wefing S, Lengauer T. Placement of medium-sized molecular fragments into active sites of proteins. J Comput Aided Mol Design 1996;10:41–54.
22. Böhm H-J. The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. J Comput Aided Mol Design 1994;8:243–256.
23. SYBYL Molecular Modeling Software, Version 6.x. St. Louis, MO: Tripos Associates, Inc.; 1994.
24. Wlodawer A. Rational drug design: the proteinase inhibitors. Pharmacotherapy 1994;14(6):9S–20S.
25. Stahl M, Böhm HJ. Development of filter functions for protein-ligand docking. J Mol Graph Modelling 1998; in press.
26. Rarey M, Kramer B, Lengauer T. The particle concept: Placing discrete water molecules during protein-ligand docking predictions. Proteins 1998;34:17–28.