PROTEINS: Structure, Function, and Genetics 33:367–382 (1998)

Flexible Docking Using Tabu Search and an Empirical Estimate of Binding Affinity

Carol A. Baxter, Christopher W. Murray,* David E. Clark, David R. Westhead, and Matthew D. Eldridge
Proteus Molecular Design Ltd., Macclesfield, Cheshire, United Kingdom

ABSTRACT **This article describes the implementation of a new docking approach. The method uses a Tabu search methodology to dock flexibly ligand molecules into rigid receptor structures. It uses an empirical objective function with a small number of physically based terms derived from fitting experimental binding affinities for crystallographic complexes. This means that docking energies produced by the searching algorithm provide direct estimates of the binding affinities of the ligands. The method has been tested on 50 ligand-receptor complexes for which the experimental binding affinity and binding geometry are known. All water molecules are removed from the structures and ligand molecules are minimized in vacuo before docking. The lowest energy geometry produced by the docking protocol is within 1.5 Å root-mean square of the experimental binding mode for 86% of the complexes. The lowest energies produced by the docking are in fair agreement with the known free energies of binding for the ligands. Proteins 33:367–382, 1998.**

Key words: ligand-protein docking; molecular recognition; tabu search; empirical scoring function; binding affinity prediction

INTRODUCTION

The safe and effective action of a pharmaceutical agent within the human body depends on the selective recognition of the drug molecule by the appropriate target protein. This molecular recognition is applied in structure-based drug design in which the structure of a target protein is used as a basis for designing new ligands. It would therefore be extremely useful if one could reliably estimate the binding mode of the ligand-receptor complex. In computer-aided molecular design (CAMD), the "docking problem" is a term given to the prediction of the binding mode of a ligand-receptor complex when the structure of the receptor is known or can be estimated. With X-ray crystallography revealing the structures of more protein binding sites, molecular docking is becoming an increasingly important technique during the early stages of drug design, and the progress made has been reviewed in a number of recent articles.[1–6]

Early attempts to solve the docking problem considered only the translational and orientational degrees of freedom of the ligand with respect to the receptor[7]; however, as ligand geometry is known to change often upon binding, this is an oversimplification. With recent advances in computer facilities most approaches to the docking problem now take into account conformational flexibility of the ligand, and in some methods[8–10] limited flexibility of the receptor is also permitted. Clearly, as the number of degrees of freedom increases, the size of search space rapidly becomes enormous. Simulated annealing (SA),[11,12] evolutionary programming (EP),[13,14] and genetic algorithms (GA)[8,9,15–17] have all been applied to the docking problem. Although these methods can produce accurate predictions of binding modes, they are computationally expensive and may therefore be unsuitable for application to large numbers of compounds. More efficient approaches use incremental construction methods,[18] such as FLEXX[19] and Hammerhead[20]; highly specialized sampling techniques such as QXP[21]; or fast shape matching algorithms similar to DOCK with multiple conformations for the screened molecules.[22] DOCK[7] still remains the most commonly used program for high throughput docking applications.

An important feature of any docking method is an energy function that is capable of predicting binding modes. The minimum values of the function should correspond to the preferred binding mode(s) of the ligand. In many applications, the docking energy function is not suitable as an estimate of the binding affinity, and a separate methodology is applied.[23] A

Abbreviations: CAMD, computer-aided molecular design; GA, genetic algorithm; EP, evolutionary programming; SA, simulated annealing; TS, tabu search; RMS, root-mean square; PDB, Protein Data Bank; NAPAP, N-alpha-(2-naphthylsulphonyl-glycyl)-DL-p-amidinophenylalanyl-piperidine.

D.E. Clark's present address is Dagenham Research Centre, Rhône-Poulenc Rorer Ltd., Rainham Road South, Dagenham, Essex RM10 7XS, United Kingdom.
D.R. Westhead's and M.D. Eldridge's present address is E.M.B.L. Outstation, European Bioinformatics Institute, Hinxton Hall, Hinxton, Cambridge CB10 1RQ, United Kingdom.
*Correspondence to: Christopher W. Murray, Proteus Molecular Design Ltd., Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire, SK11 0JL, United Kingdom.
Received 6 February 1998; Accepted 25 June 1998

more consistent procedure is to use the estimate of binding affinity as the objective function minimized in the docking procedure. This method has been used in FLEXX[19] and Hammerhead,[20] although there has been no systematic analysis of the performance of the functions (versus experimental affinity) for a substantial number of docked complexes.

Here we present a fast, automated flexible docking program, Ligand Evaluation by Automatic Docking Studies (PRO_LEADS). Our aims are to predict accurately the binding affinities and binding modes of flexible ligands, typical of the type encountered in our structure-based drug design projects. This means that the ligands are moderately flexible, small organic molecules typically no larger than four or five residue peptides. We would like the method to be reasonably fast so that ligands can be docked (including any necessary repetitions of the docking procedure) in a few minutes on a typical workstation, although more time could be taken when more accurate results are required. It is also important that the limitations of the method be properly investigated using a set of test cases as typical as possible of our target application. Initially, this means testing on a large dataset of protein-ligand complexes where the binding affinity is known; in this article, we examine the performance of the method on a test set of 50 such complexes. The test set is reasonably diverse and consists of a variety of protein classes and types as well as different ligands of varying size and flexibility. To make the test as meaningful as possible, all water molecules are removed from the complexes because, in general, it is difficult for the drug designer to predict which water molecules should be retained for any given molecular design.[24] Ligands are minimized in vacuo, and the ligand's rotatable bonds are randomized before the docking procedure.

PRO_LEADS uses tabu search (TS) to sample the conformational space. During a previous study,[25] the performances of SA, GA, EP, and TS were compared in molecular docking applications. It was found that TS performed at least as well as the other methods; of course, it is impossible to reach firm conclusions because the performance of the algorithms is highly implementation-dependent. TS has been applied to many problems, but to our knowledge this was its first application to the docking problem. The modern form of TS[26] operates by imposing restrictions to enable a search process to negotiate otherwise difficult regions. These restrictions take the form of a tabu list that stores a number of previously visited solutions. By preventing the search from revisiting these regions, the exploration of new search space is encouraged. The TS implemented in PRO_LEADS is described in Materials and Methods. It includes some new features that have been added to improve the local searching performance of the algorithm.

A simple, empirically derived scoring function[27] is used to estimate the binding affinities and binding modes. Functions with a similar motivation[28,29] have been used before as a basis for docking potentials.[19,20] The function described in Materials and Methods comprises simple contact terms to estimate lipophilic and metal-ligand binding contributions, a simple explicit form for hydrogen bonds and a term which penalizes flexibility. The function was optimized against a large training set of 82 protein-ligand complexes taken from the Protein Data Bank (PDB)[30] for which the binding affinities were known. The function reproduced the binding affinities of the complexes with a cross-validation error of 8.68 kJ/mol and performed well on a separate test set of a further 30 complexes. The optimized scoring function was augmented for application to PRO_LEADS by the addition of three terms deemed to be important to molecular docking. The terms calculate the ligand internal energy, penalize conformations that clash with the receptor, and penalize solutions that fall outside a user-defined active site. The docking is performed using a grid-based methodology with two grids containing the values for the lipophilic term and the clash term. The hydrogen bond term is short range so it can be efficiently calculated using a neighbor grid formalism that requires another two (smaller) grids to be stored. The hydrogen bond term has also been modified to inhibit the generation of some unrealistic contacts. Implementation details are given in Materials and Methods.

MATERIALS AND METHODS

Preparation of the Test Set

A database of 50 protein-ligand complexes for which the binding affinity is known and the binding geometry is reliable has been constructed. It contains only small noncovalently bound ligands of the type which de novo design or database searching might be expected to produce. We have excluded ligands with large cycles, large peptide ligands, and ligands containing several sugar monomers. It was also decided to use only PDB structures so the database contains no docked structures or structures obtained from proprietary or privileged sources. The main problem in constructing a test set is that binding affinities and binding geometries are needed. To build up a substantial test set, three additional structures (TMT1, TSC2, and DFR4) were obtained by removing part of the ligand in a PDB structure (1TMT, 2TSC, and 4DFR, respectively) in a manner described previously.[27]

The complexes in the test set can be split into the following five classes:

4 aspartic proteases,
9 serine proteases,
7 metalloproteases,
6 sugar binding proteins, and
24 others.

TABLE I. The 50 Complexes in the Dataset and the Success Rate Obtained With the Complexes

Number	Pcode	No. rot. bonds	No. heavy atoms	Receptor name	Success rate
1	1MBI	0	5	Myoglobin	100
2	3PTB	0	9	Trypsin	100
3	1TNG	2	8	Trypsin	100
4	1RBP	2	21	Retinol binding protein	100
5	1TNH	2	9	Trypsin	100
6	1STP	5	16	Streptavidin	100
7	2PHH	1	10	PHBH	100
8	1DWB	0	9	Thrombin	99
9	2YPI	3	9	TP isomerase	96
10	2CGR	5	29	Immunoglobulin	95
11	5CPP	0	11	Cytochrome P450cam	95
12	1ABF	4	11	ABP	95
13	2CPP	6	11	Cytochrome p450cam	94
14	1CBX	5	15	CPA	94
15	1DBB	1	23	DB3	93
16	2TMN	5	13	Thermolysin	93
17	5ABP	6	12	ABP	91
18	1ABE	4	10	ABP	90
19	2GBP	9	12	GBP	82
20	DFR4	0	12	DHFR	76
21	1ULB	0	11	PNP	73
22	1HSL	4	11	Histidine-binding protein	70
23	1DBJ	1	21	DB3	68
24	1HVR	10	46	HIV-1 protease	61
25	5TLN	9	23	Thermolysin	55
26	1ETR	7	35	Thrombin	52
27	2R04	9	25	Virus coat protein	52
28	1MNC	10	25	Neutrophil collagenase	38
29	1PPC	6	37	Trypsin	35
30	1ETS	6	37	Thrombin	31
31	TMT1	10	27	Alpha-thrombin	30
32	1PPH	4	30	Trypsin	29
33	1NSD	8	20	Neuraminidase	28
34	1PHG	2	17	Cytochrome p450cam	27
35	1PHF	0	11	Cytochrome P450cam	20
36	TSC2	3	23	Thymidylate synthase	20
37	2TSC	8	35	Thymidylate synthase	19
38	1HPV	11	35	HIV-1 protease	16
39	1TMN	13	35	Thermolysin	15
40	1EBG	2	9	Enolase	13
41	1DOG	5	11	Glucoamylase II 471	13
42	1ETT	4	30	Thrombin	11
43	2IFB	14	18	FABP	11
44	6CPA	11	33	CPA	11
45	7CPA	15	41	CPA	10
46	1EPO	17	46	Endothiapepsin	7
47	4DFR	7	33	DHFR	4
48	1APU	17	34	Penicillopepsin	3
49	1PGP	11	17	6-PGDH	3
50	1HTF	13	41	HIV-1 protease	1

Tables I and II give the important information on the complexes in the database. The information given for each protein-ligand complex is the PDB entry, the protein class, the size of the ligand, and the experimental binding affinity (in kJ/mol).

The preparation of the complexes in the database has been described in detail elsewhere.[27] Because there is no explicit treatment of receptor flexibility and the scoring function requires the positions of hydrogen atoms on the receptor, it is necessary to obtain good hydrogen atom positions for the receptor. Briefly, the crystal structures were extracted from the Brookhaven Databank,[30] hydrogen atoms were added using InsightII,[31] and the molecules had atom

TABLE II. The Top Two Ranking Clusters Obtained in the Docking Runs†

	Rank 1 cluster			Rank 2 cluster			
Pcode	Best energy (KJ/mol)	RMS (Å)	Size	Best energy (KJ/mol)	RMS (Å)	Size	Experimental binding affinity (KJ/mol)
1MBI	−21.648	0.52	100	0.000	0.00	0	−10.76
3PTB	−25.812	0.40	100	0.000	0.00	0	−27.04
1TNG	−22.278	1.03	100	0.000	0.00	0	−16.73
1RBP	−43.382	1.07	100	0.000	0.00	0	−38.35
1TNH	−21.525	1.14	100	0.000	0.00	0	−19.21
1STP	−40.644	0.69	100	0.000	0.00	0	−76.45
2PHH	−32.143	0.64	100	0.000	0.00	0	−26.69
1DWB	−22.320	0.52	99	−19.155	3.05	1	−16.66
2YPI	−22.639	0.69	96	−19.879	4.01	4	−27.52
2CGR	−36.060	0.72	95	−26.640	6.49	3	−41.51
5CPP	−32.212	1.22	100	0.000	0.00	0	−33.58
1ABF	−32.765	0.21	95	−27.674	2.85	5	−30.92
2CPP	−32.178	1.26	94	−30.953	2.70	5	−34.63
1CBX	−42.615	1.13	94	−31.005	1.97	2	−36.21
1DBB	−40.961	0.74	93	−37.646	6.90	2	−51.35
2TMN	−35.979	0.92	93	−30.958	2.80	6	−33.58
5ABP	−32.622	0.20	91	−27.311	2.73	8	−37.87
1ABE	−30.555	0.33	90	−26.111	2.76	10	−40.04
2GBP	−37.284	0.35	82	−27.578	3.69	2	−43.36
DFR4	−22.977	0.74	76	−19.904	3.34	15	−34.23
1ULB	−21.955	0.45	73	−20.559	3.34	25	−30.24
1HSL	−31.241	0.50	64	−27.867	1.55	36	−41.65
1DBJ	−41.882	0.69	68	−40.653	5.99	32	−43.80
1HVR	−63.699	1.16	60	−53.687	1.69	4	−54.25
5TLN	−40.172	1.02	57	−39.504	1.93	5	−36.32
1ETR	−41.505	0.85	52	−32.072	3.95	12	−42.21
2R04	−43.253	1.46	13	−42.816	1.30	5	−35.49
1MNC	−39.548	0.70	36	−36.828	4.26	11	−51.35
1PPC	−39.467	0.48	36	−39.097	1.68	1	−36.83
1ETS	−47.072	0.65	31	−38.758	2.94	14	−48.62
TMT1	−39.452	0.86	29	−34.920	8.44	2	−39.70
1PPH	−37.254	0.49	29	−30.366	4.51	14	−35.50
1NSD	−29.380	1.24	30	−24.703	4.43	5	−30.24
1PHG	−38.110	1.25	7	−37.728	5.24	42	−49.39
1PHF	−29.256	0.67	17	−27.104	3.08	13	−25.09
TSC2	−34.035	0.58	20	−32.026	6.44	4	−30.80
2TSC	−34.532	1.14	15	−33.111	3.09	14	−48.62
1HPV	−40.790	1.38	17	−37.786	9.18	15	−52.61
1TMN	−52.273	1.05	17	−49.212	1.78	18	−41.65
1EBG	−45.085	1.56	26	−37.334	0.43	71	−61.75
1DOG	−30.193	3.32	22	−29.036	3.80	61	−22.92
1ETT	−34.762	3.87	51	−34.540	0.57	11	−35.30
2IFB	−43.570	2.83	24	−43.218	1.97	47	−30.99
6CPA	−48.685	1.16	13	−48.431	1.61	7	−65.74
7CPA	−55.571	1.15	12	−54.462	5.09	8	−79.64
1EPO	−47.370	0.93	7	−45.001	1.74	4	−45.40
4DFR	−39.282	1.04	5	−32.728	3.67	4	−55.33
1APU	−36.931	1.74	7	−36.090	4.13	3	−43.92
1PGP	−23.238	1.72	18	−23.232	5.95	16	−32.51
1HTF	−49.138	10.44	17	−46.798	9.02	1	−46.18

†Each docking run used 100 repetitions. The best energy, the RMS to the crystal structure of the best energy solution, and the number of solutions are given for each cluster.

types and partial charges assigned using the CVFF forcefield. The complex was minimized with Discover[31] using tethers with large force constants to keep heavy atom positions very close to the crystallographically determined positions. All waters were then removed from the structures. The ligand was separated from the receptor and minimized in vacuo using the CVFF forcefield. Rotatable bonds were assigned using an automated procedure based on recognition of chemistries using SMILES[32] strings.

The rotatable bond assignment was checked and if necessary, changed manually. In general, all acyclic, nonterminal sp^3—sp^3, sp^3—sp^2 bonds were rotatable. Single bonds between two sp^2 atoms were defined as not rotatable (except for the N-benzene bond in 4DFR, 2TSC, and TSC2). Ring systems were treated as rigid entities. The fixed parts of the valence geometry were determined by the prior minimization of the isolated ligand. All docking variables (i.e., rigid body variables and flexible ligand torsions) were randomized before every docking run.

Degrees of Freedom

For solution of the docking problem to be feasible using currently available methods and computational resources, it is necessary to make certain simplifications. First, neither receptor nor ligand can be considered to be fully flexible. Second, an active site for the receptor must be defined to restrict the region of space in which solutions are sought. Only with these simplifications is the search space of the problem sufficiently small that a good heuristic algorithm could be expected to find optimal solutions quickly with a reasonable success rate. The docking methods implemented in PRO_LEADS allow the following degrees of freedom:

Ligand translation—the ligand is free to move within a user-specified box defining the active site; if the centroid of the ligand moves outside the box, a user-specified penalty is added to the energy value.
Ligand orientation—the ligand has full orientational freedom.
Ligand flexibility—the ligand is considered flexible through a list of rotatable bonds.

A restriction on the ligand variables is defined from the centroid of the ligand in its crystallographic conformation. The centroid of the ligand is calculated at every docking step, and if the centroid of the ligand goes outside a 64 Å^3 volume, a large penalty term is applied. This gives the ligand complete orientational freedom but constrains the translation of the ligand. The receptor is kept rigid during docking.

Docking Variable

The internal coordinate modeling tree[33] provides a complete internal coordinate description of an assembly of molecules, their internal conformations, and relative positions. This makes it a good basis for the choice for the variables for use in docking; thus, a very similar scheme has been implemented in this work. In PRO_LEADS, the docking variables representing the relative position of ligand and receptor and their internal conformations are a subset of the variables from the internal coordinate tree (see Fig. 9). The receptor is considered fixed in space, and the relative position of the receptor and ligand is governed by the rigid body variables attached to the ligand. In our notation these are $\{B1, V1, T1, V2, T2, T3\}$ (Fig. 1) and can be interpreted as bond lengths, valence angles and torsion angles with respect to the fixed triplet of virtual atoms at the root of the tree. The software automatically chooses an atom close to the centroid of the ligand as the first real atom in the tree. Variables representing flexibility of ligand and receptor are torsion angles taken from the internal coordinate tree.

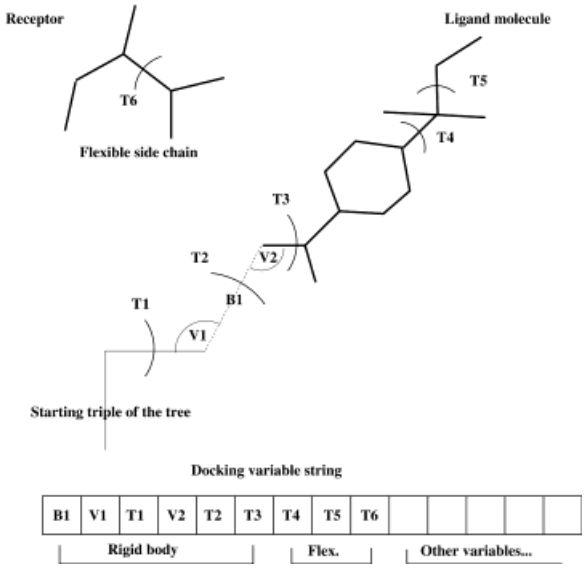

Fig. 1. Derivation of the docking variables string from variables present in the molecular internal coordinate tree. B denotes a bond length, V a valence angle, and T a torsion angle. The variable string is composed of real numbers, and only derivation of conformational variables is illustrated.

The docking variables, as manipulated by TS, are stored as a string of real numbers. The first six are the rigid body variables, and after these follow the variables for ligand flexibility. The energy function used in PRO_LEADS requires Cartesian coordinates for the ligand and receptor molecules. It is therefore necessary to carry out interconversions between the docking variable string and Cartesian coordinates. This is accomplished by a single software module of PRO_LEADS.[25]

Tabu Search

The tabu search method implemented in PRO_LEADS is shown in Figure 2. Tabu search maintains only one current solution during the course of a search. At the start of each run the current solution was initialized by randomizing the position and orientation of the ligand within the bounding box.

From the current solution, a user-defined number of "moves" is generated by a mutation-like procedure in which random numbers are drawn from an appro-

1. **Create initial solution** at random. Make this the current solution

2. **Evaluate current solution**. **If** the current solution is the **best** so far, **record** it

3. **Update tabu list**
   (a) If tabu list is not full (<25 members), add current solution to list
   (b) Else, replace oldest member of list with current solution.

4. **Generate** and evaluate **100 possible moves** from the current solution

5. **Rank 100 possible moves** in ascending order of energy

6. **Examine the moves in rank order**
   (a) If move has lower energy than **best so far**, **accept** it and go to 7
   (b) If move is **not tabu** (> 0.75 RMS), **accept** it and go to 7
   (c) If **no acceptable moves** are located, **terminate** algorithm

7. If the **iteration limit** of 1000 has been reached, **exit** with the best solution found. If the best solution so far has not changed for a given number of iterations (100), restart the whole procedure (go to 1). Otherwise, go to 2

Fig. 2. Schematic of the tabu search algorithm. The reader is refered to the text for more details.

priate distribution and are added to each of the docking variables in the current solution. In this work, a Cauchy distribution with a specified semi-interquartile range is used and more details are given elsewhere.[25] Each of the generated moves is scored using the energy function and they are then ranked in order, with the best move at the head of the list. The moves are examined in rank order. The TS maintains a tabu list that stores a number of previously visited solutions, and a move is considered "tabu" if it generates a solution that is not sufficiently different from the stored solutions. The threshold measure used in this work to determine the tabu status or otherwise of potential moves is a root-mean square (RMS) (measured over selected heavy atoms) of 0.75 Å or less between the two solutions being compared. The highest ranking move (tabu or not) is always accepted if its energy is lower than the lowest energy so far. Otherwise the algorithm chooses the best non-tabu move. If neither of these criteria can be met, the algorithm terminates.

If a new current solution can be found, it is added to the tabu list. In the current work, the length of the tabu list was 25. Solutions are simply added to the end of the list until it is full. Thereafter, the current solution must replace an existing solution stored in the tabu list. The tabu list is managed in a "first-in, first-out" manner, with the current solution replacing the tabu solution having the longest residence in the list.

Once the new current solution has been identified and stored, a new set of moves is generated from it and the search procedure continues with the next iteration. At each iteration, 100 moves were generated using Cauchy mutation with a σ value of 0.075.

A further mechanism that helps search exploration has also been implemented: if, after 100 iterations of the above procedure, it is observed that the best solution has not changed, then 100 iterations of fine searching are performed about the local minimum. The σ value of the Cauchy mutation and the size of the threshold measure are reduced by a factor of 4 in an attempt to find the lowest energy in the region before the search is randomly restarted at a new position in space.

The tabu search continues for a user-defined number of iterations. At the end of this time, it terminates and returns the best solution found during the search.

Local Minimization

A local minimizer that uses the Powell algorithm[34] has been implemented in PRO_LEADS. This is a nonderivative algorithm designed to move the solution to the nearest local minimum. It is used as a final stage minimization of the lowest energy conformation found after operation of the tabu search.

Energy Evaluation

Atom types

The scoring function assigns general atom types to all ligand atoms and receptor atoms in contact with the ligand. The atom types are as follows:

lipophilic chlorine, bromine, and iodine atoms that are not ions; sulphurs that are not acceptor or polar types; carbons that are not polar type.
H-bond donor hydrogens attached to oxygen or nitrogen.
H-bond donor/acceptor oxygens attached to hydrogen atoms; special case of imine nitrogen (i.e., C = NH nitrogen).
H-bond acceptor oxygens not attached to hydrogen; nitrogens with no hydrogens attached and one or two connections; halogens that are ions; sulphurs with only one connection (e.g., thioureas).
polar (*non–H-bonding*) nitrogens with no hydrogens attached and more than two connections; phosphorus; sulphurs attached to one or more polar atoms (including H-bonding atoms and not including polar carbon atoms or fluorine atoms); carbons attached to two or more polar atoms (including H-bonding atoms and not including polar carbon atoms or fluorine atoms); carbons in nitriles or carbonyls; N atoms with no hydrogens and four connections; fluorine atoms.
metal metal atoms.

It is hoped that these atom types are reasonable in most applications of the function, and they are thought to be reasonable compromises across the molecules in the ligand-receptor database.

ChemScore regression equation

In an earlier work, we developed an empirical scoring function (the ChemScore function) that, given a geometry for a ligand-receptor complex, calculates an estimate of the free energy of binding.[27]

There are several methods that follow a similar line.[29,35] Our empirical scoring function can be written in the form:

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{hbond}\Sigma_{iI} g_1(\Delta r) g_2(\Delta\alpha) + \Delta G_{metal}\Sigma_{aM} f(r_{aM}) + \Delta G_{lipo}\Sigma_{lL} f(r_{lL}) + \Delta G_{rot} H_{rot} \quad (1)$$

The form of the separate terms has been explained in detail elsewhere.[27] In the present docking work, the rotatable bond term is only evaluated at the end of the docking run and so is not treated as a variable during the search.

The ΔG coefficients were determined previously using Multiple Linear Regression on a training set of 82 complexes for which the binding affinity is known experimentally. The robustness of the equation was tested against test sets comprising a further 30 ligand-receptor complexes and by using extensive cross-validation on the original training set. The coefficients used in equation 1 are −5.48, −3.34, −6.03, −0.117, and 2.56 kJ/mol, respectively. The function has a (leave one out) cross-validated q^2 of 0.658 and s_{press} value of 8.68 kJ/mol.

The ChemScore function has been changed to facilitate its use in docking applications.

An extended hydrogen bond term was used at the beginning of the TS. The hydrogen bond term,[28] $\Sigma_{iI} g_1 g_2$, is calculated for all complementary possibilities of hydrogen bonds between ligand atoms, i, and receptor atoms.

$$g_1(\Delta r) = \begin{cases} 1 & \Delta r \le 0.25\text{Å} \\ 1 - (\Delta r - 0.25)/ & \\ (upper\_dist - 0.25) & 0.25\text{Å} < \Delta r \le \text{upper\_dist} \\ 0 & \Delta r > \text{upper\_dist} \end{cases}$$

and

$$g_2(\Delta\alpha) = \begin{cases} 1 & \Delta\alpha \le 30° \\ 1 - (\Delta\alpha - 30)/ & \\ (upper\_ang - 30) & 30° < \Delta\alpha \le \text{upper\_ang} \\ 0 & \Delta\alpha > \text{upper\_ang} \end{cases}$$

Δr is the deviation of the H . . . O/N hydrogen bond length from 1.85 Å, and $\Delta\alpha$ is the deviation of the hydrogen bond angle N/O-H . . . O/N from its ideal value of 180°. When using this term in the TS, upper_dist and upper_ang were 1.15 Å and 110°, respectively, at the beginning of the TS and changed to 0.65 Å and 80° after the application of 75% of the tabu iterations. It was found that docking performance was improved when an extra angle term was included (as a product) in the hydrogen bond term. The angle term is

$$g_3(\Delta\beta) = \begin{cases} 1 & \Delta\beta \le 30° \\ 1 - (\Delta\beta - 30)/10 & 30° < \Delta\beta \le 40° \\ 0 & \Delta\beta > 40° \end{cases}$$

where $\Delta\beta$ is the deviation of the hydrogen bond angle H . . . O/N-X from the value 140°. Additionally, a constraint was placed on hydrogen atoms so that they could only form one hydrogen bond. These terms reduce the number of nonnative contacts suggested by the docking.

The following terms that have been deemed to be important to the problem of molecular docking have also been added to the ChemScore regression equation.

The form of the clash term used to penalize close contacts between the ligand and receptor heavy atoms has been partially optimized on a small subset of the complexes in the dataset and depends on the atom types of the two atoms in clash. For two atoms that could be involved in a hydrogen bond contact, the term is zero when $r > 1.60$, and for $r \le 1.60$ is:

$$E = \frac{20.0}{\Delta G_{hbond}}\left(\frac{1.60 - r}{1.60}\right) \quad (2)$$

For a metal contact, the term is zero when $r > 1.38$, and for $r \le 1.38$ is:

$$E = \frac{20.0}{\Delta G_{metal}}\left(\frac{1.38 - r}{1.38}\right) \quad (3)$$

where r is the distance in Å between the ligand atom and the receptor atom. For all other heavy atoms the term is zero when $r > r_{clash}$, and for $r \le r_{clash}$ takes the form:

$$E = 1.0 + 4.0\left(\frac{r - r_{clash}}{-r_{clash}}\right) \quad (4)$$

r_{clash} = 3.35 Å if the receptor atom is a sulphur, otherwise r_{clash} = 3.10 Å.

The internal energy of the ligand is a sum of torsion and internal clash terms. The latter term takes of the same forms as the ligand-receptor clash terms but is only evaluated for ligand atoms which are more than 3 bonds apart (i.e., 1–2s, 1–3s, and 1–4s are excluded). The torsion term is

$$E = A[1 - cos\,(n\phi - \phi_0)] \quad (5)$$

where ϕ is the torsion angle, for $sp^3 - sp^3$ bonds $A = 0.1875$, $n = 3.0$, $\phi_0 = \pi$, for $sp^3 - sp^2$ bonds $A = 0.09375$, $n = 6.0$, $\phi_0 = 0.0$ and for $sp^2 - sp^2$ bonds $A = 0.1875$, $n = 2.0$, $\phi_0 = 0$ (although in this work,

single $sp^2 - sp^2$ bonds were mainly treated as nonrotatable).

The final term penalizes a solution if its centroid of the ligand falls outside the user-defined bounding box for the complex.

The addition of new terms did not significantly affect the robustness of the ChemScore regression equation, and recalibration was not necessary.

Grid representation

For computational speed, lipophilic interactions and ligand-receptor clashes are precalculated and stored on two grids covering the active site. The value at a point in the clash grid indicates the value of the clash term summed over all receptor atoms when an atom is placed at that grid point. Similarly, a point in the lipophilic grid indicates the value of the lipophilic term when a lipophilic ligand atom is placed at that point. When the grid is used, the values of the clash term and the lipophilic term for a given ligand heavy atom are taken directly from the values at the nearest grid point.

Because the hydrogen bond interaction term has an angle dependence, it cannot be treated in the same way as the above terms. Instead, "neighbor lists" are used. In the neighbor list, each grid point stores not an energy value, but a list of indices of receptor atoms that are capable of forming a hydrogen bond with a complementary ligand atom if it were to be placed on the grid point. The hydrogen bond term is then determined by performing an explicit calculation for each receptor atom stored in the list. The clash term for hydrogen bonding atoms is also evaluated explicitly. Two neighbor lists are required—one for hydrogen bond donors and one for acceptors. The lists have the same resolution as the lipophilic and clash grids but take up less space due to the high level of redundancy in the list. The neighbor list approach is exact and is computationally efficient because the hydrogen bond terms are short range.

In all the experiments reported in this article, a grid resolution of 0.2 Å was used. The grid dimensions ranged from 77.8 $Å^3$ to 7,598.9 $Å^3$, and the average time taken to generate the grids was 43 minutes on the Convex Exemplar.

Clustering

The analyses presented here use a specialized clustering and RMS calculation. In the RMS calculation it is important that the permutational symmetry of labeling of molecules be taken into account. For example, any nonsymmetric molecule that contains a rotatable bond to a carboxylate ion can rotate by 180° around that bond, to produce a molecule that will be identical in its energetics but in which the indexes of the two oxygens have swapped. In PRO_LEADS, a subgraph isomorphism of the heavy atom portion of the molecule back onto itself is performed, and all equivalent labeling schemes are obtained. When the RMS is calculated, each labeling scheme is looped over and the lowest RMS is reported. An upper diagonal matrix of heavy atom RMS between all docking solutions is produced and used as the basis for clustering. The clustering is performed with a tolerance 1.5 Å RMS using the average linkage method. This means that two neighboring clusters are joined together if the average RMS difference between the two clusters is less than 1.5 Å.

Selection Criterian

In judging the performance of the method we have concentrated mainly on the following five criteria:

1. The lowest energy solution produced by program is referred to as the predicted conformation. This would be the conformation that a drug designer would concentrate on in prioritizing the results from docking. Therefore, we believe that the result for a complex should only be judged acceptable if the lowest energy solution is close to the crystallographic minimum. (In our tests we use a heavy atom RMS of 1.5 Å as the definition of acceptable.)
2. It is also relevant to know how often any particular docking run produces the correct answer, i.e., the percentage of docking runs within 1.5 Å RMS of the crystallographic binding mode. If the right answer is only found a small percentage of the time, then a large number of docking runs will be required to find it.
3. There is also a comparison of the lowest energies produced by the docking runs with the experimental binding affinity. The relevant quantities are the R^2 value, which reflects how well the trends in activity are predicted, and s, the standard deviation, which reflects how accurately any one value is predicted.
4. The time taken for performing the docking is also considered because this determines the range of applicability of the program.
5. There is a consideration of the failures of the method because this points the way to understanding the limitations of the method and identifying improvements.

We believe the test set and the criteria form a good overall assessment of the performance of the program, although it is intended that other tests be reported in the future.

RESULTS

The 50 protein-ligand complexes in the dataset were selected from the PDB[30] to include a variety of protein classes with noncovalently bound ligands ranging in size from 0 to 17 rotatable bonds and from 5 to 46 heavy atoms. All the complexes have known

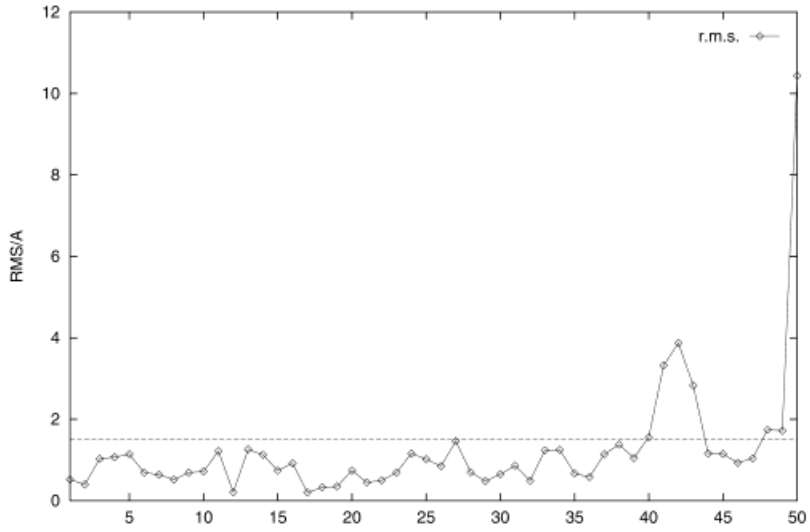

Fig. 3. Plot of heavy atom RMS deviations between the predicted conformations and the crystallographic ligand geometries for the 50 complexes in the dataset. (The predicted conformation is the one corresponding to the lowest energy solution produced by PRO_LEADS in 100 trials.)

dissociation constants. Preparation of the dataset is described in Materials and Methods and details of its contents are given in Table I, with the second column listing the PDB codes of the complexes.

Before performing the docking tests, all water molecules were removed and the ligand and receptor structures were separated. The ligand was minimized using Discover with the CVFF forcefield.[31] The ligand was docked into the receptor using PRO_LEADS. Each docking was repeated a specified number of times starting from random ligand conformations.

Prediction of the Binding Mode

In the first experiment, each complex in the dataset was docked back into the receptor 100 times and each docking trial was limited to 200,000 function evaluations. The solutions generated by the 100 trials were clustered with respect to their binding mode, and the clusters were ranked according to the lowest energy structure in each cluster. The solution in the top ranking cluster with the lowest docking energy was the predicted conformation. The heavy atom RMS has been calculated between the predicted conformation and the crystallographic geometry. Any solution with an RMS of less than 1.5 Å was judged to be an acceptable prediction of the binding mode. Table II reports the energy and RMS of the solutions with the lowest docking energy in the top two ranking clusters, together with the number of solutions in the clusters and the experimental binding affinities for each complex. The RMS of the predicted conformations (lowest energy solutions) for the 50 complexes have been plotted (Fig. 3), and a dotted line has been drawn at 1.5 Å. Only seven predicted conformations have an RMS greater than 1.5 Å; therefore, using this criterion to judge success, PRO_LEADS has achieved a 86% prediction rate in this test.

The success rate plotted (Fig. 4) is the number of trials in which the docking solution was within 1.5 Å RMS of the crystallographic configuration. Success rates of greater than 50% were achieved for 27 complexes. Although a large majority of the examples with good predicted conformations also had high success rates, there were a few exceptions. For example, of the 100 trials on the complex, 1ETT, 11 found acceptable solutions, but the lowest energy conformation had an RMS of 3.87 Å. In contrast, the predicted conformation of the docked ligand in the complex, 1EPO, was acceptable but the TS found this solution in only seven experiments. It would therefore be incorrect to assess the ability of PRO_LEADS to predict binding modes after the examination of just one set of data. The success rates were found to depend on the number of degrees of freedom of the problem. 1EPO was the largest ligand in the dataset with 17 rotatable bonds. The complexes in Table I are listed in order of decreasing success rate. The 9 complexes that contain ligands with more than 10 rotatable bonds all reported less than 20 acceptable solutions from the 100 trials.

The time taken to perform 100 docks of each complex was found to depend on the size and flexibility of the ligand and ranged from 139 to 1115 minutes on one processor of a Convex Exemplar (Convex Computer Corporation, Richardson, TX). The average time for 100 docks was 342 minutes, with each docking repetition taking 3 minutes 26 seconds on average. (The Convex Exemplar is about 2.4 times *slower* than a Silicon Graphics R10000 (Silicon Graphics, Inc., Mountain View, CA) workstation for these applications.) The two energy grids and the two neighbor lists took an average of 43 minutes to generate. As the grids and neighbor lists are only generated once for each protein, these times have not been included in the results presented.

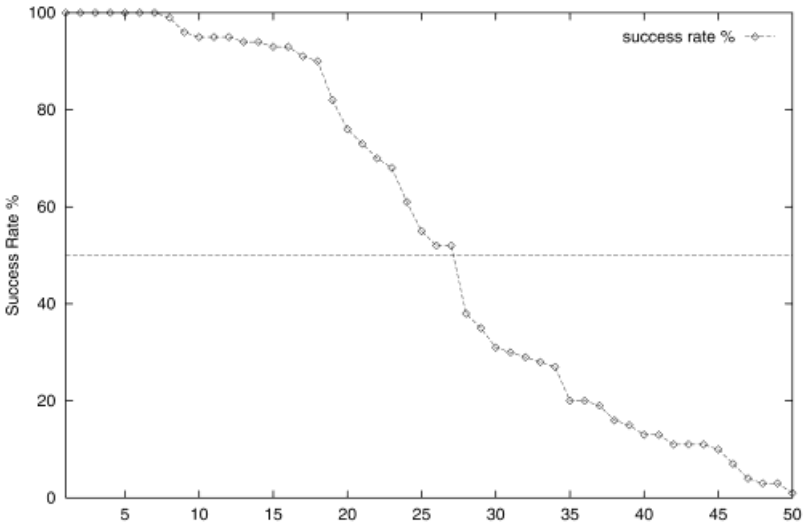

Fig. 4. Plot of success rates for the 50 complexes in the dataset. Success is defined as PRO_LEADS producing a geometry within 1.5 Å of the crystallographic conformation.

The above test is wasteful because so many repetitions would not be necessary in most applications. There are many situations in which one would wish to apply docking to a large number of ligands to the same complex, and a quicker protocol would be required. As a second test, the number of repetitions was set to 10 and the number of function evaluations was reduced to 100,000. The average total docking time for the 10 trials was 16 minutes per structure. The prediction rate fell from 86% to 72%.

Some of the results of the first docking experiment on a number of test cases are discussed in detail in the following sections. Only one example of successful docking is discussed because there is more to learn about the limitations of the method from failures. All the failures are discussed.

A successful docking: NAPAP-thrombin, 1ETS

The crystallographic geometry of bound NAPAP in thrombin is shown (Fig. 5) together with the geometry of the lowest energy solution produced by PRO_LEADS. The size of the problem is nontrivial, NAPAP has 37 heavy atoms and was described by 6 rotatable bonds. The complex is also difficult for some energy functions because NAPAP has a benzamidine and a naphthyl group that are similarly sized. Energy functions that are sterically dominated produce inverted modes in which the benzamidine is positioned in the D pocket and the naphthyl group in the polar S1 pocket.[25] The time taken to perform 100 docks was 414 minutes on the Convex Exemplar. The heavy atom RMS deviation between the predicted structure and the crystallographically determined ligand structure was 0.65 Å. Of the 100 docking trials, there were 31 solutions with RMS deviations of less than 1.5 Å. A good method of viewing docking results is to plot the scatter of energies against the RMS deviations for the 100 solutions. The scatter plot for 1ETS (Fig. 6) shows a large cluster of acceptable solutions with low energy that is well separated in energy from the other clusters.

Incorrectly docked complexes

To give detailed descriptions of all the experiments on the 43 complexes correctly predicted by PRO_LEADS would be both tedious and pointless. An examination of the failures is more useful as a better understanding of the limitations of the method can be gained by studying what went wrong.

Palmitic acid docked into a fatty acid-binding protein, 2IFB, is shown in Figure 7. Palmitic acid is a long hydrocarbon ligand with a carboxylate group at one end which makes hydrogen bonds with the receptor. The hydrocarbon chain lies in an extended conformation along a long cavity in the protein and only lipophilic interactions are made. The PRO_LEADS energy function scored these interactions with the nondirectional terms described in Materials and Methods. PRO_LEADS docking runs positioned the oxygen atoms correctly, but the hydrocarbon chain found a variety of paths through the protein cavity. These different pathways through the cavity were all assigned similar energies giving rise to the large range of RMS values with similar energies in the scatter plot (Fig. 8). With lipophilic contacts of this nature, it is sometimes justified to be circumspect about the crystallographic coordinates, because it is quite likely that the lipophilic tail in palmitic acid does occupy a variety of different conformations in the lipophilic channel. The electron density in the crystal is likely to represent the average of several different possible conformations. The mobility of the lipophilic tail is reflected in the B values of the carbon atoms.[36] The final carbon atom (C16) in palmitic acid has a fairly high temperature factor in the crystallographic structure (51.0 Å 2), whereas the penultimate carbon (C15) has a fairly

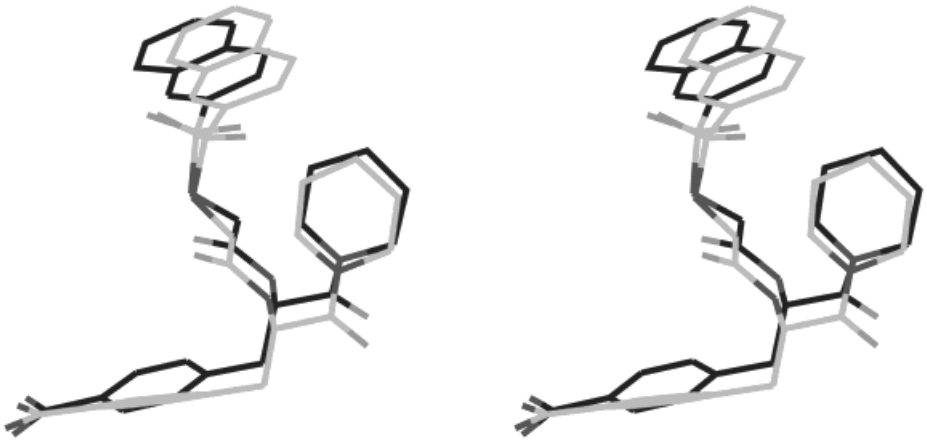

Fig. 5. Docking of NAPAP into thrombin (1ETS). The predicted conformation of the ligand is shown in black, and the crystallographic conformation is shown in grey.

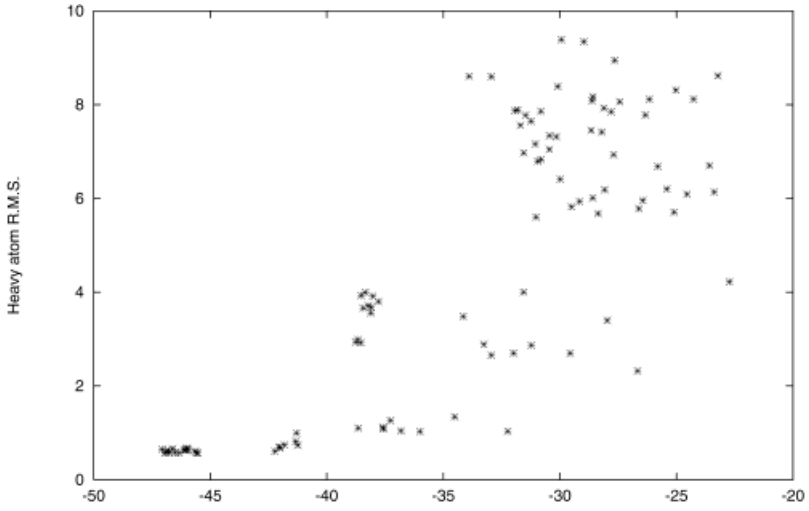

Fig. 6. Scatter plot of the predicted binding affinity against RMS for the complex 1ETS (NAPAP in thrombin). Each star corresponds to the RMS and affinity for 1 of the 100 docking runs performed by PRO_LEADS.

low one (24.7 Å 2). One interpretation of this is that the position of C16 in the crystallographic coordinates is only partially occupied, whereas the C15 position is usually occupied by either C15 *or* C16.

We also failed to predict the binding mode of the ligand GR126045 in the HIV-1 protease complex 1HTF. The solution with the lowest binding energy had an RMS of 10.44 Å. The fourth ranking cluster has four members, and the best energy in this cluster is within 2.0 Å RMS of the crystallographic solution. This latter solution differs from the crystallographic solution mainly due to differences in positioning of a portion of the ligand which is in solvent and makes no contact with the receptor. Crystallography has shown that the HIV-1 protease binding site contains a highly conserved water molecule that forms good hydrogen bonds to the two backbone N-Hs in the ILE-50 residues of each monomer. In the absence of any water molecules, the lowest energy solution obtained by PRO_LEADS identifies this water's position as an ideal place to put a hydrogen bond acceptor. The docking experiment was repeated with the inclusion of this highly conserved water molecule as part of the receptor. The results showed some improvement. The lowest ranking cluster was still not correct, but the second, third, and fourth clusters were all close to the crystallographic solution and only differed in the positioning of portions of the ligand which were in solvent. The difference in energy between the best solution of the first ranking cluster and that of the second ranking cluster was less than 1 kJ/mol.

PRO_LEADS also failed to predict another one of the larger ligands, 1APU (acetylpepstatin in penicillopepsin). In the preferred conformation (Fig. 9), the hydroxyl that contacts the aspartates of the aspartic protease is incorrectly orientated and makes only one hydrogen bond to one of the Asp groups. Additionally, the ester tail of the ligand is poorly positioned although it still retains its only polar contact. Apart from these failings, the solution produced by

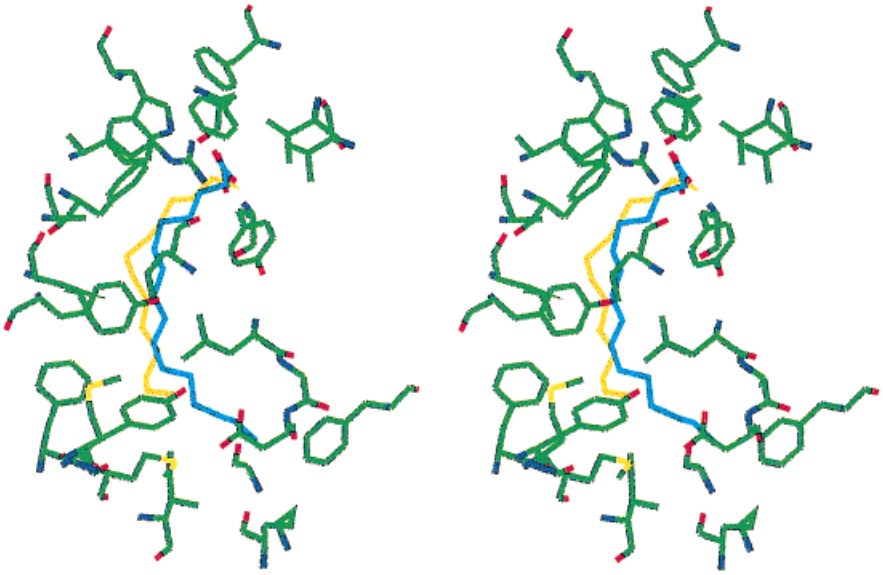

Fig. 7. Docking of fatty-acid binding protein into palmitic acid (2IFB). The predicted conformation of the ligand is shown in yellow, and the crystallographic conformation is shown in blue.

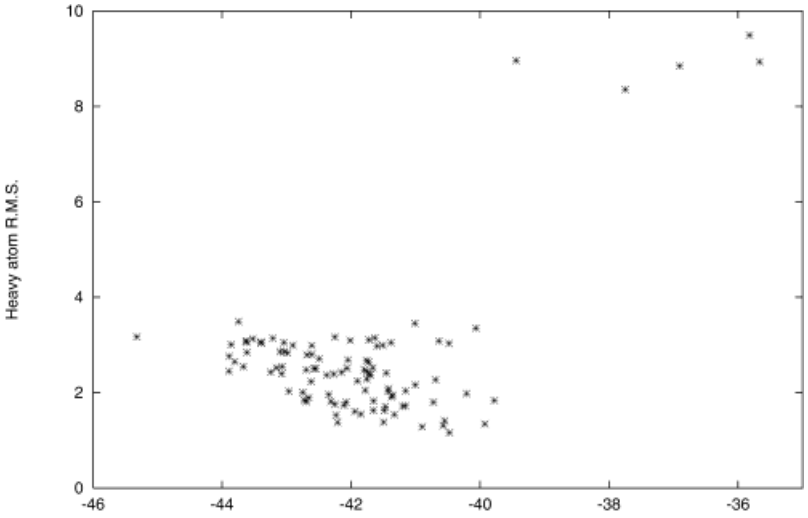

Fig. 8. Scatter plot of the predicted binding affinity against RMS for the complex 2IFB (palmitic acid in fatty-acid binding protein). Each star corresponds to the RMS and affinity for 1 of the 100 docking runs performed by PRO_LEADS.

PRO_LEADS is very close to the crystallographic conformation.

1ETT is the complex of 4-TAPAP with thrombin. The ligand contains a benzamidine group that enters the S1 subsite of thrombin at an unusual angle, forming an oblique contact with ASP-189 at the bottom of the S1 pocket. This is not expected to be the ideal orientation of contact. For example, the amidine forms a direct contact with the carboxylate of ASP-189 for the structures of thrombin or trypsin with benzamidine (1DWB and 3PTB). The preferred conformation from PRO_LEADS is the same as the crystal structure except that the benzamidine ignores the S1 pocket and enters the catalytic region. The energy gap between this wrong answer and the best answer from the second cluster (RMS to crystal structure of 0.57 Å) is less than 0.5 kJ/mol. In this second ranking solution, the benzamidine forms a direct contact with ASP-189 although in doing so it incurs some ligand-receptor clashes and is marked down accordingly. Clearly, the balance of clash terms versus the importance of the salt bridge is not quite right in this particular example. It is a difficult test case for PRO_LEADS.

1PGP (6-phosphogluconic acid in 6-phosphogluconate dehydrogenase) is not successfully docked by PRO_LEADS—the preferred conformation being 1.76 Å from the correct solution. The complex is characterized by a large number of polar interactions. PRO_LEADS correctly predicts the positioning of the phosphate group, but interactions at the tail of the molecule are switched around, leading to the

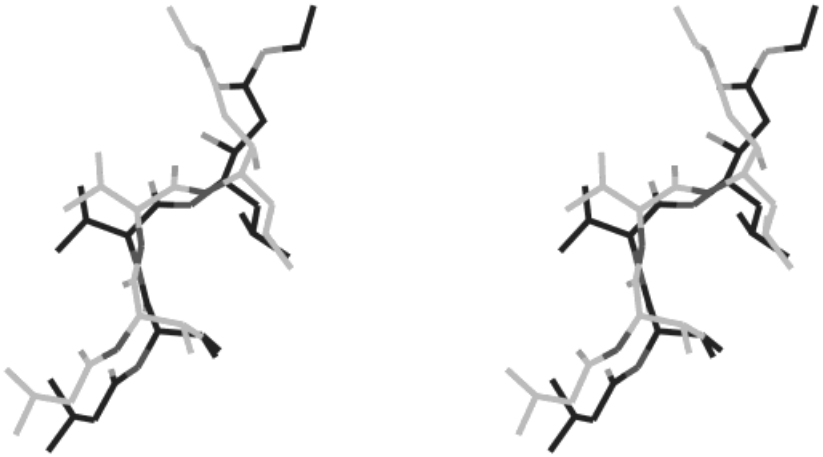

Fig. 9. Docking of acetylpepstatin into penicillopepsin (1APU). The predicted conformation of the ligand is shown in black, and the crystallographic conformation is shown in grey.

poor RMS. This failure indicates how PRO_LEADS can produce errors when choosing between a large number of alternative polar interactions.

A similar failure occurs with 1EBG (phosphonoacetohydroxamate in enolase). The preferred conformation is 1.56 Å RMS away from the experimental result and is characterized by correct positioning of the phosphate group but a 180° rotation of the C(=O)-N(OH)-H moiety. The second cluster contains the correct solution (0.43 Å RMS), although the energy gap between the two clusters is quite large (8 kJ/mol). The incorrect solution forms more hydrogen bonds with the receptor although in doing so incurs some clash penalties. Additionally, the correct solution contains a water-mediated hydrogen bond that cannot be scored in these "dry" runs.

1DOG (a sugar derivative in glucoamylase) is the final failure of the function. The predicted conformation occupies the same binding region as the crystallographic answer but is rotated with respect to it. The ligand does not attain many hydrogen bond contacts compared with other sugar-like complexes in the test set (e.g., 1ABE or 1NSD). These factors make it difficult for PRO_LEADS to choose the best orientation for the ligand.

Estimation of Binding Affinity

The PRO_LEADS energies of the predicted conformations have been plotted (Fig. 10) against the experimental binding affinities of the 50 complexes in the dataset. There is one severe outlier (point 6—streptavidin-biotin, 1STP). This complex is an outlier in the regressions and tests on the original scoring function,[27] and so it was expected to be an outlier in the current analysis. The reason for the extraordinarily large binding affinity of biotin to streptavidin is not well understood.[37] After omission of this outlier, the R^2 value is 0.577 and the s value is 8.6 kJ/mol. This can be compared with the performance of the ChemScore function on the crystallographic binding modes where R^2 is 0.532 and s is 10.3 kJ/mol. The small improvement is probably due to the regularization of the contacts observed in the docked structures or could simply be fortuitous. The fact that the lowest energies produced by the docking function are in fair agreement with experiment is a good validation of the docking function, because the optimization of the geometries has not resulted in large deviations of the energies.

It should be noted that 43 of the complexes used as the test set in the docking were also in the training set for establishing the preliminary form[27] of the scoring function. The other seven (complexes 7, 8, 15, 21, 23, 27, and 46) were in completely separate test sets. Inspection of Figure 10 reveals that the binding affinity of these seven complexes is well predicted, so one can conclude that the results have not been significantly biased by the original fitting process. Such bias would have been surprising for two reasons. First, the scoring function has changed quite a bit from that used in the original training through the introduction of penalty terms and the inclusion of a stricter hydrogen bond term. Second, the original work[27] used extensive cross-validation by checking the robustness of the regression when different classes of enzyme-complexes were omitted from the training set.

The predicted binding affinities have also been analyzed for the reduced central processing unit time protocol described above (i.e., 100,000 functions evaluations and 10 repetitions). As expected the agreement worsens, but the standard deviation is still a respectable 10.69 kJ/mol and there is still some correlation (R^2 of 0.39).

DISCUSSION

In this article, we have described the fast, flexible docking method, PRO_LEADS. PRO_LEADS uses a

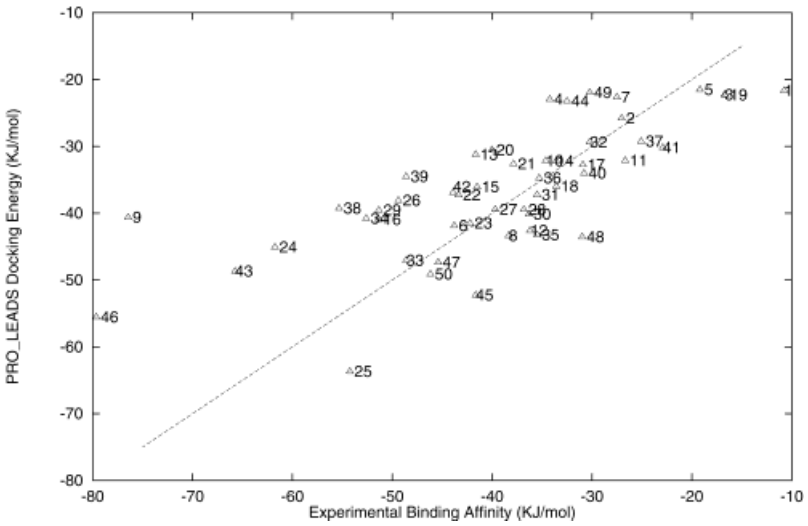

Fig. 10. A scatter plot showing the predicted binding affinities for the dataset versus the experimental binding affinities in kJ/mol. The predicted values were obtained using the PRO_LEADS scoring function and represent the best energies produced PRO_LEADS during the 100 repetitions of the docking for each complex. The labels on the plot correspond to the numbering of the complexes shown in Table I. Excluding the outlier, 1STP, the R^2 is 0.577 and the standard deviation is 8.6 kJ/mol.

novel TS algorithm to explore the extensive search space and an empirically derived scoring function that is capable of estimating binding affinities to rank the possible solutions. The method has been tested on a large dataset of 50 protein-ligand complexes. PRO_LEADS accurately predicted the binding mode of 86% of the complexes. The method also estimates the binding affinity to a fair accuracy with a standard deviation of less than 1.5 orders of magnitude. These results used long run times, but adaptation of the protocol reduces this time to a few minutes per docking on a good workstation. This adaptation causes only a modest drop in the prediction rate to 72% and still gives respectable agreement with the experimental binding affinities.

The complexes for which PRO_LEADS failed to predict the binding mode have been analyzed in an attempt to determine the limitations of the method. Several problems can be seen. The first is that the algorithm has difficulty in predicting the exact orientation of flexible lipophilic fragments. This can be seen in the failure to predict 2IFB correctly. It is not clear how this could be easily rectified, but from the drug design point of view, this is not a particularly serious deficiency. Another problem with the scoring function is that very flexible ligands that form a large number of polar contacts are difficult to position correctly. This can be seen in the failure to predict the ligand conformation for 1PGP and also to some extent for 1EBG. When faced with a large number of potential hydrogen bonds, PRO_LEADS has difficulty in choosing between them. This problem is likely to be shared by other objective functions used in molecular docking applications.

Another difficulty in docking applications is the treatment of structural or conserved water molecules. Currently, PRO_LEADS has no automatic way of treating water molecules, and the examples of 1HTF and 1EBG illustrate the problems that arise when one or more water molecules play an important role in determining the binding mode of a ligand-receptor complex. The treatment of important water molecules is a nontrivial problem[24] and is a focus for our future work. In a real design situation in which the binding mode is unknown, information about the location of water molecules in the active site would not generally be available. We therefore do not consider that including all crystallographic water molecules is an acceptable solution to the problem. The fact that good conformations are often produced in the absence of all water molecules is very encouraging, and indicates that the scoring function is addressing some of the important parts of the molecular recognition process.

In addition to the inclusion of crystallographic water molecules, there are several potential biases that can be introduced into a docking method to reduce the complexity of the problem. We have tried to develop an objective docking method, but a number of biases have been recognized. For example, PRO_LEADS treats the receptor as fixed and any induced fit that occurs between the receptor and ligand is factored out, because the receptor structure used in the docking is the one that is appropriate for the ligand. In other words, the program will not be expected to do as well if the a ligand is docked against a receptor structure obtained by cocrystalization with an alternative ligand. Additionally, the minimization protocol used in preparing the complexes in the test set chooses hydrogen positions for the receptor which are most appropriate for the correct ligand conformation. Any potential bias from this effect could be reduced by allowing limited receptor flexibility,[9] although the more general problem of induced fit would still remain. The effects of receptor flexibility on docking accuracy will be the subject of a future report.[38] Another potential source of bias in docking tests is restrictions on the transla-

tional freedom of the ligand. The centroid of the ligand is free to move within a 64 Å^3 box that defines the active site. This active site definition has been used by others[15] and in our view, is sufficient for many structure-based drug design applications where the receptor binding site is known or can be predicted.

A current limitation of the method is that there is no treatment of flexibility in ring systems. This may be addressed in future work, although it is a difficult problem. Allowing flexibility in ring system without a good treatment of the energetics could be dangerous. On the whole, most of the molecules in our target applications do not contain flexible ring systems and when they do, our current approach would be to perform a conformational analysis of the ring system and separately dock representative conformations. Only one complex (1HVR) in the test set contains a ring size greater than six.

The methodology used for the docking is grid-based. This is not ideal because in our implementation it places memory restrictions on the size of ligand which can be treated. With our current procedure, we would find it difficult to dock a ligand bigger than approximately seven residues in length, even after making what we would consider to be compromises on the grid resolution. It would therefore not be useful to compare PRO_LEADS to methods for docking sizable proteins, such as those applied in the recent blind docking studies of the β-lactamase inhibitory protein,[39] because PRO_LEADS is aimed at docking small drug-like molecules to receptors.

It is useful to compare the performance of the PRO_LEADS with other flexible docking codes. Here the main focus will be on methods that have been tested systematically on a large number of crystallographic complexes of small drug-like molecules. The need to test on a large number of complexes is important if comparative conclusions are to be drawn about performance. The FLEXX[19,40] incremental search strategy has been applied to 19 complexes. Here we consider the results obtained using their automated method of base fragment selection.[40] For 11 of these complexes (58%), the lowest scoring solution is within 1.5 Å heavy atom RMS of the crystallographic solution. The FLEXX method is fast, taking an average of 2 minutes on a workstation. Using the quick protocol, PRO_LEADS appears to take about 8 times longer than FLEXX but does give superior prediction rates (72%) and better estimates of the binding affinities. (Note that the PRO_LEADS timings exclude grid formation because in high throughput applications, the grid formation is a one-time cost that can be neglected.) It is certainly possible that more specialized protocols would speed up PRO_LEADS even more, because currently the same amount of searching is being performed for small rigid ligands as for large flexible ligands. The program, GOLD,[9] has been applied to a large dataset of 100 complexes, and the authors report a success rate of 71% using a subjective criterion for success. PRO_LEADS thus has a better prediction rate than GOLD. However, comparisons of this type are always dangerous. In particular, GOLD defines quite a large search space for each docking example, whereas ours is comparatively small. The GOLD tests are therefore intrinsically more objective than the ones given here. However, we are certainly encouraged by the higher success rate and tend to believe it may indicate a better objective function for docking. One clear advantage of PRO_LEADS over GOLD is that the energies produced can be correlated with experimental binding affinities. Additionally, the GOLD runs are considerably more expensive. The GOLD test protocol uses 20 repeats of runs, each taking an average of 12 minutes on an SGI 4400 (Silicon Graphics, Inc., Mountain View, CA) (which is about 1.3 times faster than the Convex Exemplar for PRO_LEADS runs). Twenty repeats of the standard PRO_LEADS run (ignoring the time for grid evaluation) would take about 69 minutes on average, making PRO_LEADS approximately 4.5 times quicker. Clearly, both GOLD and PRO_LEADS would use faster protocols in applications involving multiple ligands.

PRO_LEADS is a suitable tool for docking small ligands into proteins and assessing their binding affinity. We hope to extend it to consider more flexibility in the receptor and the ligand and to consider automatically the role of active site water molecules. We are also testing the performance of the method in docking different ligands into one model of a receptor.

REFERENCES

1. Blaney, J.M., Dixon, J.S. A good ligand is hard to find: Automatic docking methods. Perspect. Drug Discov. Res. 1:301–319, 1993.
2. Jones, G., Willett, P. Docking small-molecule ligands into active sites. Curr. Opin. Biotech. 6:652–656, 1995.
3. Lybrand, T.P. Ligand-protein docking and rational drug design. Curr. Opin. Struct. Biol. 5:224–228, 1995.
4. Rosenfeld, R., Vajda, S., Delisi, C. Flexible docking and design. Ann. Rev. Biophys. Biomol. Struct. 24:677–700, 1995.
5. Gschwend, D.A., Good, A.C., Kuntz, I.D. Molecular docking towards drug discovery. J. Mol. Recognit. 8:175–186, 1996.
6. Lengauer, T., Rarey, M. Computational methods for biomolecular docking. Curr. Opin. Struct. Biol. 6:402–406, 1996.
7. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R., Ferrin, T.E. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 161:269–288, 1982.
8. Jones, G., Willett, P., Glen, R.C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol. 245:43–53, 1995.
9. Jones, G., Willett, P., Glen, R.C., Leach, A.R., Taylor, R. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. 267:727–748, 1997.
10. Leach, A.R. Ligand docking to proteins with discrete side-chain flexibility. J. Mol. Biol. 235:345–356, 1994.
11. Goodsell, D.S., Olson, A.J. Automated docking of substrates to proteins by simulated annealing. Proteins 8:195–202, 1990.
12. Morris, G.M., Goodsell, D.S., Huey, R., Olson, A.J. Distrib-

uted automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. J. Comput. Aided Mol. Des. 10:293–304, 1996.
13. Gehlhaar, D.K., Verkhivker, G.M., Rejto, P.A., et al. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: Conformationally flexible docking by evolutionary programming. Chem. Biol. 2:317–324, 1995.
14. Gehlhaar, D.K., Verkhivker, G.M., Rejto, P.A., Fogel, D.B., Fogel, L.J., Freer, S.T. Docking conformationally flexible small molecules into a protein binding site through evolutionary programming. In: "Evolutionary Programming IV: Proceedings of the Fourth Annual Conference on Evolutionary Programming." McDonnell, J.R., Reynolds, R.G., Fogel, D.B. (eds.). Cambridge: MIT Press, 1995:615–627.
15. Judson, R.S., Jaeger, E.P., Treasurywala, A.M. A genetic algorithm method for molecular docking flexible molecules. J. Mol. Struct. 308:191–206, 1994.
16. Oshiro, C.M., Kuntz, I.D., Dixon, J.S. Flexible ligand docking using a genetic algorithm. J. Comput. Aided Mol. Des. 9:113–130, 1995.
17. Clark, K.P., Ajay. Flexible ligand docking without parameter adjustment across four ligand-receptor complexes. J. Comp. Chem. 16:1210–1226, 1995.
18. Leach, A.R., Kuntz, I.D. The conformational analysis of flexible ligands in macromolecular receptor sites. J. Comp. Chem. 13:730–748, 1992.
19. Rarey, M., Kramer, B., Lengauer, T., Klebe, G. Predicting receptor-ligand interactions by an incremental construction algorithm. J. Mol. Biol. 261:470–489, 1996.
20. Welch, W., Ruppert, J., Jain, A.N. Hammerhead: Fast, fully automated docking of flexible ligands to protein binding sites. Chem. Biol. 3:449–462, 1996.
21. McMartin, C., Bohacek, R.S. QXP: Powerful, rapid computer algorithms for structure-based drug design. J. Comput. Aided Mol. Des. 11:333–344, 1997.
22. Miller, M.D., Kearsley, S.K., Underwood, D.J., Sheridan, R.P. FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. J. Comput. Aided Mol. Des. 8:153–174, 1994.
23. Meng, E.C., Kuntz, I.D., Abraham, D.J., Kellogg, G.E. Evaluating docked complexes with the HINT exponential functions and empirical atomic hydrophobicities. J. Comput. Aided Mol. Des. 8:299–306, 1994.
24. Raymer, M.L., Sanschagrin, P.C., Punch, W.F., Venkataraman, S., Goodman, E.D., Kuhn, L.A. Predicting conserved water-mediated and polar ligand interactions in proteins using a k-nearest-neighbours genetic algorithm. J. Mol. Biol. 265:445–464, 1997.
25. Westhead, D.R., Clark, D.E., Murray, C.W. A comparison of heuristic search algorithms for molecular docking. J. Comput. Aided Mol. Des. 11:209–228, 1997.
26. Glover, F., Laguna, M. Tabu search. In: "Modern Heuristic Techniques for Combinatorial Problems." Reeves, C.R. (ed.). Oxford: Blackwell Scientific Publications, 1993:70–150.
27. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V., Mee, R.P. Empirical scoring functions. I: The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Comput. Aided Mol. Des. 11:425–445, 1997.
28. 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. Des. 8:243–256, 1994.
29. Jain, A.J. Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities. J. Comput. Aided Mol. Des. 10:427–440, 1996.
30. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., et al. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
31. InsightII/Discover, v. 95.0. San Diego, CA: Molecular Simulations Inc., 1996.
32. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J. Chem. Inf. Comput. Sci. 28:31–36, 1988.
33. Abagyan, R.A., Totrov, M., Kuznetsov D. ICM—a new method for protein modelling and design: Applications to docking and structure prediction from distorted native conformation. J. Comp. Chem. 15:488–506, 1994.
34. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. In: "Numerical Recipes in C: The Art of Scientific Computing." 2nd ed. Cambridge: Cambridge University Press, 1992:412–420.
35. Head, R.D., Smythe, M.L., Oprea, T.I., Waller, C.L., Green, S.M., Marshall, G.R. VALIDATE: A new method for the receptor-based prediction of binding affinities of novel ligands. J. Am. Chem. Soc. 118:3959–3969, 1996.
36. Sacchettini, J.C., Gordon, J.I., Banaszak, L.J. Crystal structure of rat intestinal fatty-acid-binding protein. J. Mol. Biol. 208:327–339, 1989.
37. Weber, P.C., Wendoloski, J.J., Pantoliano, M.W., Salemme, F.R. Crystallographic and thermodynamic comparison of natural and synthetic ligands bound to streptavidin. J. Am. Chem. Soc. 114:3197–3200, 1992.
38. Murray, C.W., Baxter, C.A. Application of molecular docking to thrombin and thermolysin. To be published.
39. Strynadka, N.C.J., Eisenstein, M., Katchalski-Katzir, E., et al. Molecular docking programs successfully predict the binding of a beta-lactamase inhibitory protein to TEM-1 beta-lactamase. Nat. Struct. Biol. 3:233–239, 1996.
40. Rarey, M., Kramer, B., Lengauer, T., Klebe, G. Multiple automatic base selection: Protein-ligand docking based on incremental construction without manual intervention. J. Comput. Aided Mol. Des. 11:369–384, 1997.