PROTEINS: Structure, Function, and Genetics 43:113–124 (2001)

# High Throughput Docking for Library Design and Library Prioritization

**David J. Diller*** **and Kenneth M. Merz, Jr.**
*Center for Informatics and Drug Discovery, Pharmacopeia, Inc., CN5350, Princeton, New Jersey*

***ABSTRACT*** **The prioritization of the screening of combinatorial libraries is an extremely important task for the rapid identification of tight binding ligands and ultimately pharmaceutical compounds. When structural information for the target is available, molecular docking is an approach that can be used for prioritization. Here, we present the initial validation of a new rapid approach to molecular docking developed for prioritizing combinatorial libraries. The algorithm is tested on 103 individual cases from the protein data bank and in nearly 90% of these cases docks the ligand to within 2.0 Å of the observed binding mode. Because the mean CPU time is <5 s/mol, this approach can process hundreds of thousands of compounds per week. Furthermore, if a somewhat less thorough search is performed, the search time drops to 1 s/mol, thus allowing millions of compounds to be docked per week and tested for potential activity. Proteins 2001;43:113–124.** © 2001 Wiley-Liss, Inc.

**Key words: molecular docking; combinatorial chemistry; conformational analysis; flexible minimization**

## INTRODUCTION

With the advent of combinatorial chemistry and the resulting ability to synthesize large collections of compounds for a broad range of targets, it has become apparent that the capability to effectively prioritize screening efforts is crucial to the rapid identification of the appropriate region of chemical space for a given target. Because it has been generally observed that hits obtained against a given target are clustered in a finite region of chemical space, there is reason to believe that given the right computational tools it is possible to prioritize screening efforts so that only libraries containing active compounds are interrogated. Effective prioritization tools would allow scientists to both obtain leads in a cost-effective and efficient manner and to test virtual libraries against novel targets before actual synthesis and bioanalysis, thereby reducing synthesis costs. With the expected flood of new targets becoming available in the coming decade, it will be critical to focus screening efforts on target-appropriate regions of chemical space.

There are many challenges to overcome before being able to develop appropriate library prioritization tools. At one extreme are the screens for which there is no structural data for the target. In these cases, QSAR or other data mining tools are typically the method of choice for screening prioritization. At the opposite extreme are the structure-based approaches that rely on the availability of crystallographically determined structures of the target. Unfortunately, in most cases a crystal structure is not available. With the advent of proteomics and high-throughput protein crystallography, however, it is likely that for a given target a structure of a related protein will be available. In these cases, a homology model can be built starting from the structure of a related protein, and structure-based tools could be used in conjunction with QSAR or other data-mining tools.

When structural information for a target protein is available, molecular docking can be a useful tool for prioritizing screening efforts.[1–3] Operationally, this means that rather than assaying an entire collection of compounds, the compounds are first docked and ranked via some scoring function, and then only a subset of the compounds, usually the highest ranked, are assayed. This approach to prioritizing screening efforts usually gives an increase by a factor of 1–10 in the number of active compounds compared with a randomly selected subset of compounds.[4]

Our ultimate goal is to use molecular docking as a way to prioritize combinatorial library screening efforts, i.e., rather than ranking individual compounds, we want to rank combinatorial libraries of compounds. Compounds synthesized through combinatorial methods are often quite flexible compared with typical databases of compounds used for molecular docking studies. Thus, for a docking procedure to be useful for our purposes, it must be able to handle fairly flexible compounds (as many as 10–20 rotatable bonds), and it must be extremely fast (on the order of one million compounds a week). With these constraints in mind, we have developed and validated a new docking method.

The docking procedure discussed below is based on a simple conceptual picture of protein-ligand complex formation (see Fig. 1). Initially, the ligand adopts many conformations in solution. The protein then recognizes one or several of these conformations. On recognition, the ligand, protein, and solvent follow the local energy landscape to form the final complex.

*Correspondence to: David J. Diller, Pharmacopeia, Inc., CN 5350, Princeton, NJ 08543-5350. E-mail: ddiller@pharmacop.com

Received 13 July 2000; Accepted 30 November 2000

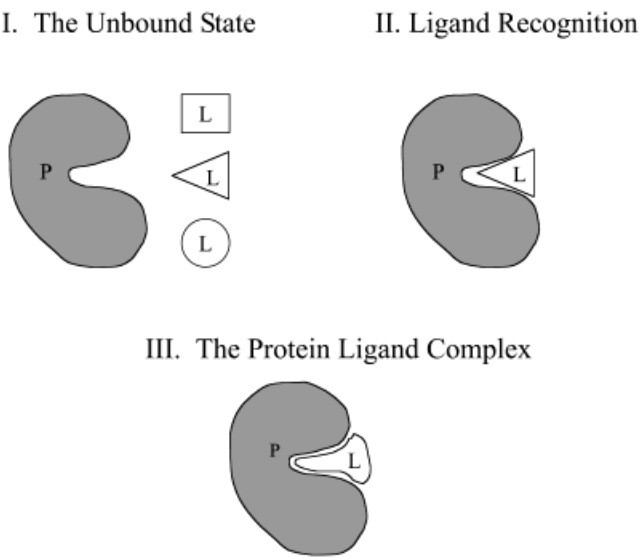


Fig. 1. A schematic of our model of protein-ligand complex formation. First, the protein and ligand are free in solution with the ligand adopting many conformations. Next, the protein recognizes one of the ligand conformations. Finally, the ligand conformation adapts to optimize the protein-ligand interactions.

This simple picture of protein/ligand complex formation was converted into an efficient computational model as follows. The initial solution conformations are generated by using a straightforward conformational search procedure. One might view the conformational search part of this algorithm as a part of the entire docking process, but because it involves only the ligand, it can be decoupled from the purely docking steps. This is justified because three-dimensional (3-D) databases of conformations for a collection of molecules can readily be generated and stored for use in numerous docking studies (e.g., using Catalyst). The recognition stage is modeled by matching atoms of the ligand to interaction "hot spots" in the binding site. The final complex formation is modeled by using a gradient-based optimization algorithm with a simple energy function. During this final stage, the translation, orientation, and rotatable bonds of the ligand are varied, whereas the protein and solvent are held fixed.

Most docking methods can be classified into one of two loosely defined categories: stochastic, for example AutoDock,[5,6] GOLD,[7] TABU,[8,9] LigandFit, and SAS,[10], or combinatorial, for example, DOCK,[11–13] FlexX,[14–16], HammerHead,[17] and Structure Based Focusing.[18] The stochastic methods, although often providing more accurate results, are typically too slow to search large databases. The method presented here falls into the combinatorial group. This approach is similar to FlexX and HammerHead in that it attempts to match interactions between the ligand and receptor. It differs from these and most other docking algorithms primarily in how it handles the flexibility of the ligand. Most current combinatorial docking algorithms handle flexibility by using an incremental construction approach, whereas our method uses an initial conformational search followed by a gradient-based minimization in the presence of the target protein.

The docking method presented here has several advantages. First, it is built from several independent pieces. This allows us to better take advantage of scientific breakthroughs. For example, when a better conformational search procedure (in our context this means more biologically relevant conformers) becomes available, it can be used to replace the current conformational search procedure by generating new 3-D databases. Second, we believe that this approach to ligand flexibility is better suited for the class of compounds synthesized through combinatorial methods. Compounds from combinatorial libraries frequently do not have a clear anchor fragment. Because finding and docking an anchor fragment from the ligand are key steps in the incremental construction algorithms, these algorithms may encounter difficulties with compounds commonly found in combinatorial libraries. Docking entire conformations overcomes this difficulty. In addition, including an efficient flexible optimization step removes a significant burden from the conformational search procedure. We can also take advantage of further improvements in energy minimization algorithms as they become available.

Our approach to ligand flexibility could be viewed as a liability because we rely on an initial conformational search. As indicated previously, to achieve maximum efficiency, the conformational search should be performed once for an entire library or collection and the resulting conformations stored for future use. For large collections, this would be a considerable investment in both computer time and disk space. Because a database will typically be used many times, the initial computer time for the conformational search can easily be justified. Moreover, with the availability of parallel computers and faster CPUs, the conformational search can be completed or occasionally redone in a reasonable amount of time. Because disk sizes are now approaching the tera-byte level, storing the conformations for millions of compounds presents no problems. Thus, we believe that the use of a conformational search is at worst a minor drawback.

## MATERIALS AND METHODS

There are four important aspects of our docking procedure: the conformational search procedure, the binding site image, i.e., the interaction hot spots, the matching procedure, and the final optimization and scoring procedure. Each of these aspects is described briefly below.

### Conformational Search Procedure

For this work, a straightforward but effective conformational search procedure was adopted. First, 1,000 uniformly distributed random conformations were generated (varying only the rotatable bonds). The internal energy, estimated by using van der Waals potentials and a dihedral angle term,[10] for each conformation was then minimized by using a BFGS optimization algorithm.[19] Any conformation with an internal energy of 15 kcal/mol above the conformation with the lowest internal energy was

eliminated. The conformations were then ranked via the score

$$\text{Score} = \text{Strain} - 0.1 \times \text{SASA} \quad (1)$$

where SASA is the solvent accessible surface area of a particular conformation measured in Å$^2$ and the strain is in units of kcal/mol. Finally, any conformation within a rms deviation of 1.0 Å of a higher ranked (i.e., better) conformation was removed. A maximum of 50 conformations was kept at the end of the conformational analysis step.

The process of a small molecule binding to a protein target is a balance between solvation by water versus "solvation" by the protein. With this in mind, the SASA term was chosen in analogy with simple aqueous solvation models,[20–22] The key difference in protein versus water "solvation" is that water competes for polar interactions only, whereas a protein effectively competes for both polar and hydrophobic interactions. Therefore, we chose to treat polar and apolar surface area identically. The choice of 0.1 as a weight for the surface area term is somewhat arbitrary but is comparable with the weights chosen for surface area-based solvation models.[20–22] Ultimately, conformations with more solvent-accessible surface area are going to be able to interact more extensively with a target protein and can, therefore, be of somewhat higher strain and still bind tightly. A more refined ranking system could be used, but this approach to ranking conformations supplies reasonable conformations.

### Binding Site Image—Locating the Hot Spots

The binding site image consists of a list of apolar hot spots, i.e., points in the binding site that are favorable for an apolar atom to bind, and polar hot spots, i.e., points in the binding site that are favorable for a hydrogen bond donor or acceptor to bind. The procedure used to create these two lists is as follows. First, to find the binding site, a grid was placed around the ligand. In all cases, the grid was at least 20 Å by 20 Å by 20 Å with at least 5.0 Å of extra space in each direction. A 0.2 Å spacing was used for the grid. Any grid point inside the protein was eliminated. Any point contained in a 6.0 Å or larger sphere not touching the protein was also eliminated. The remaining volume became the "hot spot search volume."

The hot spots were then determined by using a gridlike[5] search of the hot spot search volume, i.e., to find the apolar hot spots, an apolar probe was placed at each grid point in the hot spot search volume, and the probe score was calculated and stored. For each type of hot spot, the grid points were clustered with the top 30 being kept.

### Matching Procedure—Forming the Initial Complex

To initially position a given conformation of a ligand as a rigid body into the binding site, the atoms of the ligand are matched to the appropriate hot spots. More precisely, a triplet of atoms, $A_1$, $A_2$, $A_3$, is considered a match to a triplet of hot spots, $H_1$, $H_2$, $H_3$ if:

1. The type of $A_j$ matches the type of $H_j$ for each j = 1, 2, 3, i.e., apolar hot spots match apolar atoms and polar hot spots match polar atoms.
2. $D(A_j, A_k) = D(H_j, H_k) \pm \delta$ for all j, k = 1, 2, 3 where $D(A_j, A_k)$ and $D(H_j, H_k)$ are the distance from $A_j$ to $A_k$ and the distance from $H_j$ to $H_k$, respectively, and $\delta$ is some allowable amount of error usually between 0.25 Å and 0.5 Å.

A match then determines a unique rigid body transformation that minimizes

$$I(R, T) = \sum_{j=1}^{3} |H_j - RA_j - T|^2 \quad (2)$$

where R is a 3 × 3 rotation matrix and T is a translation vector. This rigid body transformation, which can be determined analytically, is then used to place the ligand conformation into the binding site. For this aspect of the calculation, several algorithms for finding all matches were tested. The geometric hashing algorithm developed for FlexX[23] proved to be the most efficient and was adopted.

### Final Optimization Stage

A single conformation can produce up to 10,000 matches. In the interest of efficiency, most of these matches cannot be optimized, so a pruning/scoring strategy is required. Initially, all matches for which >10% of the ligand atoms have a steric clash were eliminated. The remaining matches are ranked by using the atom pairwise score described below with an atom score cutoff of 1.0. The cutoff allows matches that fit reasonably well with a few steric clashes to survive to the final round. After being ranked, the matches are clustered, and only the top 25–100 move into the final stage.

Each remaining match is optimized by using a BFGS optimization algorithm with a simple atom pairwise score. The score is modeled after the Piecewise Linear Potential[24] with the key difference being that the score used here is differentiable. For this score, all hydrogens are ignored, and all non-hydrogen atoms are classified into one of four categories:

*Apolar*—anything that cannot form a hydrogen bond.
*Acceptor*—any atom that can act as a hydrogen bond acceptor but not as a donor.
*Donor*—any atom that can act as a hydrogen bond donor but not as an acceptor.
Donor/Acceptor—any atom that can act as both a hydrogen bond donor and acceptor.

The score between two atoms is calculated by using either a hydrogen-bonding potential or a steric potential. The two potentials, shown in Figure 2, have the mathematical form

$$F(r) = \varepsilon\left[\left(\frac{(1+\sigma)R_{\min}^2}{r^2 + \sigma R_{\min}^2}\right)^6 - 2\left(\frac{(1+\sigma)R_{\min}^2}{r^2 + \sigma R_{\min}^2}\right)^3\right]\Phi(r^2{:}r_1^2, r_0^2) \quad (3)$$

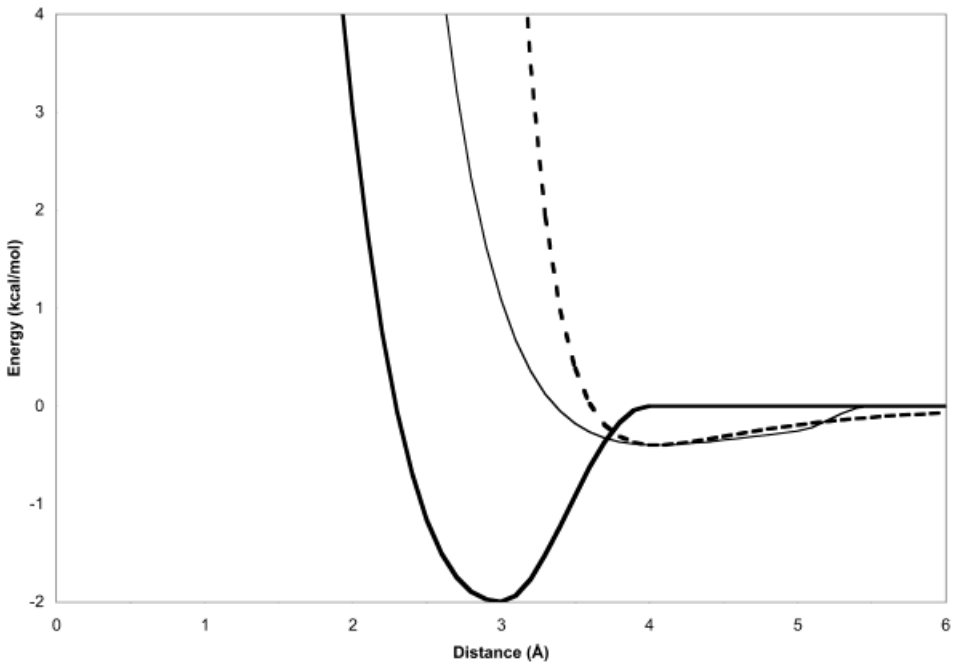


Fig. 2. The atom pairwise potentials used to score and optimize docked conformations. The dark solid line is the hydrogen-bonding potential. The light solid line is the steric potential. The dashed line is the 12-6 van der Waals potential with the same $\epsilon$ and $R_{min}$ as the steric potential. The 12-6 van der Waals potential is shown to demonstrate the effect of the softening parameter.

where $R_{min}$ is the position of the score minimum, $\epsilon$ is the depth of the minimum, $\sigma$ is a softening factor, and $\Phi(r; r_1,r_2)$ is a differentiable cutoff function of r having the properties that when $r < r_1$ $\Phi = 1$ and when $r > r_0$ $\Phi = 0$. Each potential, steric and hydrogen bonding, was assigned its own set of parameters. The exact choice of parameters are as follows. For the steric potential, $\epsilon = 0.5$, $\sigma = 1.0$, $R_{min} = 4.05$, $r_1 = 5.0$, $r_0 = 6.0$. For the hydrogen-bonding potential, $\epsilon = 2.0$, $\sigma = 1.5$, $R_{min} = 3.0$, $r_1 = 3.0$, $r_0 = 4.0$. The parameters for these potentials were chosen largely via intuition and some subsequent testing, but they are by no means fully optimized.

These potentials are very similar to the 12-6 van der Waals potentials used in many force fields with two key differences. First the softening factor, $\sigma$, makes the potentials significantly softer then the typical 12-6 van der Waals potentials (see Fig. 2), i.e., mild steric clashes common in docking runs are tolerated by this potential. In spirit, the softening factor implicitly models small induced fit effects of the protein that can be important,[25] and in practice, makes the potential much more error tolerant.

The second key difference is the cutoff function. Because the potentials are zero beyond a finite distance, typically between 5.0 Å and 6.0 Å, the direct calculation of the score can be significantly sped up in the following manner. First, a three-dimensional grid, with a grid spacing of 3.0 Å, is placed around the active site. Then, for each grid cube a linked list of the protein atoms within the cutoff distance of the cube is made. To calculate the score between a ligand atom and the protein, only the protein atoms in the cube containing the ligand atom need to be considered.

We tried calculating the scores both directly and through precalculated grids. The advantage of using the grids is the score can be calculated very rapidly: we found grids to be 5–10 times faster than the direct calculation. The advantage of the direct calculation is that effects such as protein flexibility and solvent mobility can be accommodated more easily. Because using the grids did not seem to cause any deterioration in the quality of the docking results and because we currently are not including protein flexibility or solvent mobility, for the results presented here, we calculated the scores through precalculated grids. For the purposes of the BFGS optimization algorithm, all derivatives were calculated analytically, including those with respect to the rotatable bonds.[26]

## RESULTS

To test our docking procedure, the GOLD test set was used.[7] Any covalently bound ligand or any ligand bound to a metal ion was removed because they cannot, at present, be modeled by our simple scoring function. In addition, any "surface sugars" were removed because they are not typical of the problems we encounter. This left a total of 103 cases (see Table I). No further individual processing of the test cases was performed.

### Conformational Search

The quality of the conformational search procedure is shown in Figure 3. As expected, the rms deviation between the bound conformation (X-ray) and the closest computationally generated conformation increases with the number of rotatable bonds. In all but five cases, at least one

**TABLE I. Test Cases Used for Docking Validation Studies**[†]

| PDB code | No. of rot bonds | Minimum RMSD | RMSD of top score | PDB code | No. of rot bonds | Minimum RMSD | RMSD of top score |
|---|---|---|---|---|---|---|---|
| 1aaq | 17 | 1.35 | 1.4 | 1lst | 5 | 0.58 | 1.43 |
| 1abe | 0 | 0.31 | 0.31 | 1mcr | 5 | 3.92 | 5.41 |
| 1acj | 0 | 0.59 | 0.71 | 1mdr | 2 | 0.41 | 0.78 |
| 1ack | 2 | 0.45 | 0.46 | 1mmq | 7 | 0.55 | 0.60 |
| 1acm | 6 | 0.31 | 0.31 | 1mrg | 0 | 0.45 | 3.42 |
| 1aha | 0 | 0.25 | 0.53 | 1mrk | 2 | 0.94 | 2.91 |
| 1apt | 18 | 1.10 | 1.63 | 1mup | 2 | 1.74 | 4.40 |
| 1atl | 9 | 1.05 | 4.24 | 1nco | 8 | 2.88 | 8.50 |
| 1azm | 1 | 1.40 | 2.33 | 1pbd | 1 | 0.29 | 0.38 |
| 1baf | 7 | 0.76 | 7.10 | 1poc | 23 | 2.81 | 8.62 |
| 1bbp | 11 | 1.45 | 1.55 | 1rne | 21 | 8.83 | 10.14 |
| 1cbs | 5 | 0.70 | 12.63 | 1rob | 4 | 0.83 | 1.17 |
| 1cbx | 5 | 0.53 | 2.30 | 1snc | 5 | 1.17 | 5.60 |
| 1cil | 3 | 1.07 | 5.94 | 1srj | 3 | 0.48 | 0.58 |
| 1com | 3 | 0.76 | 0.76 | 1stp | 5 | 0.33 | 0.48 |
| 1coy | 0 | 0.52 | 0.70 | 1tdb | 4 | 1.33 | 7.09 |
| 1cps | 5 | 0.85 | 0.97 | 1tka | 8 | 1.44 | 1.44 |
| 1dbb | 1 | 0.72 | 0.85 | 1tng | 1 | 0.35 | 0.42 |
| 1dbj | 0 | 0.64 | 5.90 | 1tnl | 1 | 0.45 | 4.25 |
| 1did | 2 | 2.76 | 3.65 | 1tph | 3 | 0.63 | 1.44 |
| 1die | 1 | 2.24 | 2.30 | 1ukz | 4 | 0.43 | 6.20 |
| 1dr1 | 2 | 1.02 | 1.61 | 1ulb | 0 | 1.22 | 4.19 |
| 1dwd | 9 | 0.75 | 7.98 | 1wap | 3 | 0.29 | 0.34 |
| 1eap | 10 | 0.79 | 3.95 | 1xid | 2 | 0.79 | 4.23 |
| 1eed | 19 | 3.41 | 3.41 | 1xie | 1 | 0.34 | 3.89 |
| 1epb | 5 | 0.75 | 2.86 | 2ada | 2 | 0.53 | 0.58 |
| 1eta | 5 | 5.48 | 7.29 | 2ak3 | 4 | 1.91 | 3.24 |
| 1etr | 9 | 2.70 | 7.06 | 2cgr | 7 | 0.61 | 3.46 |
| 1fen | 4 | 0.98 | 2.45 | 2cht | 2 | 0.18 | 0.40 |
| 1fkg | 10 | 1.68 | 1.72 | 2cmd | 5 | 0.50 | 2.36 |
| 1fki | 0 | 0.30 | 0.54 | 2ctc | 3 | 0.36 | 4.15 |
| 1frp | 6 | 0.67 | 1.13 | 2dbl | 6 | 0.40 | 0.96 |
| 1ghb | 4 | 0.90 | 0.94 | 2gbp | 1 | 0.17 | 0.17 |
| 1glp | 10 | 1.45 | 8.92 | 2lgs | 4 | 0.71 | 5.48 |
| 1glq | 13 | 1.91 | 9.96 | 2phh | 1 | 0.51 | 0.51 |
| 1hdc | 6 | 1.52 | 11.25 | 2plv | 5 | 1.98 | 7.40 |
| 1hef | 19 | 3.63 | 5.29 | 2r07 | 15 | 1.17 | 2.45 |
| 1hfc | 10 | 1.37 | 7.77 | 2sim | 8 | 0.92 | 1.37 |
| 1hri | 9 | 1.49 | 3.29 | 2yhx | 3 | 1.07 | 6.99 |
| 1hsl | 3 | 0.76 | 2.21 | 3aah | 3 | 0.48 | 0.68 |
| 1hyt | 5 | 0.79 | 1.56 | 3cpa | 5 | 0.92 | 1.40 |
| 1icn | 15 | 1.78 | 9.43 | 3hvt | 1 | 0.27 | 0.56 |
| 1ida | 15 | 1.32 | 1.38 | 3ptb | 0 | 0.22 | 0.28 |
| 1igj | 3 | 0.90 | 7.46 | 3tpi | 6 | 0.42 | 0.53 |
| 1imb | 2 | 1.64 | 4.48 | 4cts | 3 | 0.73 | 0.77 |
| 1ive | 2 | 2.55 | 6.63 | 4dfr | 9 | 2.05 | 8.72 |
| 1lah | 4 | 0.71 | 0.77 | 4fab | 2 | 2.52 | 4.45 |
| 1lcp | 3 | 0.53 | 4.65 | 4phv | 12 | 0.38 | 0.38 |
| 1ldm | 1 | 0.80 | 5.24 | 6abp | 0 | 0.34 | 0.34 |
| 1lic | 15 | 1.32 | 4.39 | 7tim | 3 | 0.40 | 0.98 |
| 1lmo | 6 | 5.00 | 8.40 | 8gch | 7 | 1.70 | 4.45 |
| 1lna | 6 | 1.35 | 1.46 | | | | |

[†]Given are the number of rotatable bonds of the small molecule, the minimum rms deviation, and the rms deviation of the minimum scoring docked conformation from the X-ray structure. The mean number of rotatable bonds is 5.6. All rms deviations are given in units of Å.

conformation was generated by the conformational search within 1.5 Å rms deviation of the bound conformation. The most interesting aspect of the conformational search results is that for some of the more rigid ligands the minimum rms deviation is quite large. For example, there are several ligands with fewer than five rotatable bonds but with a minimum rms deviation near 1.0 Å. This occurs for two reasons. First, a clustering radius of 1.0 Å in all cases was used. This prevented the conformational space of small ligands from being sufficiently sampled. For

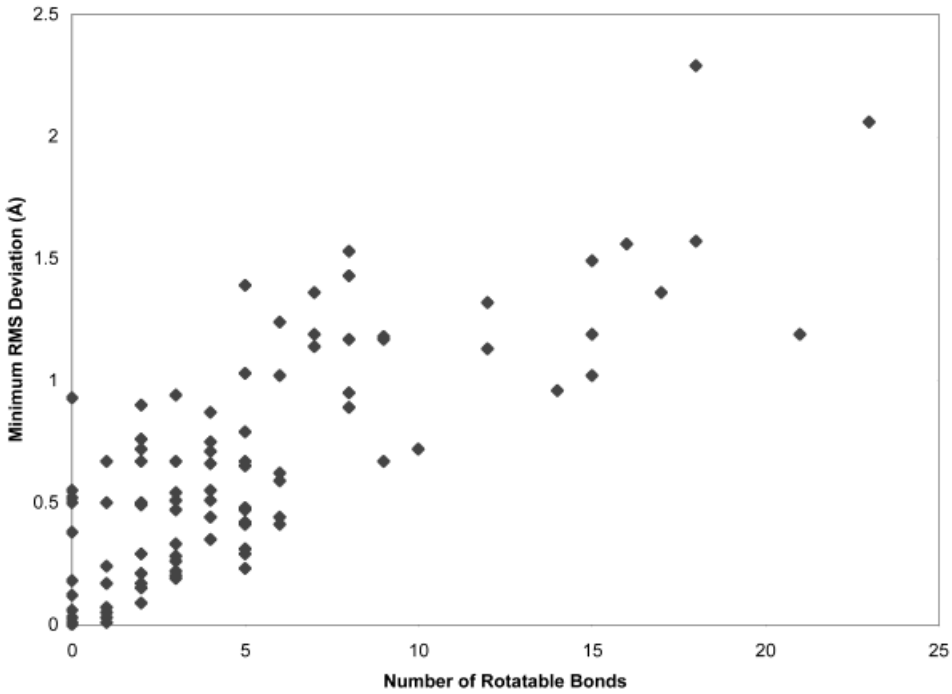


Fig. 3. The minimum rms deviations (in Å) between the bound conformation and the conformations generated by the conformational search versus the number of rotatable bonds of the ligand. Only the heavy atoms were used in counting the rotatable bonds. Rings were treated as rigid. Double bonds and bonds between 2 $Sp^2$ atoms were not counted as rotatable.

future work, a clustering radius dependent on the molecule size could be used to alleviate this particular problem. The second problem is that a bond between two Sp2 atoms was always treated as being conjugated. Thus, whenever this type of bond is encountered, it is strongly restrained to be planar. Although bonds between two Sp2 atoms are often conjugated, this is clearly an oversimplification and will necessarily be addressed in the future.

## Docking Results

For the docking runs, two different sets of parameters were tested to see their effects on the quality and speed of the docking runs: one for high-quality docking and one for rapid searches. The key differences between the two sets of parameters are the match tolerance and the number and length of the BFGS optimization runs. The match tolerance ranges from 0.5 Å for the high quality to 0.25 Å for the rapid searches. For the high-quality runs, a maximum of 100 matches per ligand were optimized for 100 steps compared to 25 matches per ligand for 20 steps for the rapid searches.

The first problem is to generate at least one docked position within a given rms deviation cutoff. Here we adopt the terminology that a ligand that is docked to within X Å of the crystallographically observed position of the ligand is referred to as an X Å hit. The cumulative histograms of the best rms deviations are shown for both sets of parameters in Figure 4 with the exact numbers for the high-quality runs in Table I. For the high-quality runs, 89 of the 103 cases produce at least one 2.0 Å hit. The numbers drops to 80 at 1.5 Å, 63 at 1.0 Å and 26 at 0.5 Å. For the rapid searches, 75 of the 103 cases produce a 2.0 Å hit, 65 produce a 1.5 Å hit, 42 produce a 1.0 Å hit, and 16 produce a 0.5 Å hit (see Fig. 4). In both cases, these numbers compare favorably with similar statistics from other dock-

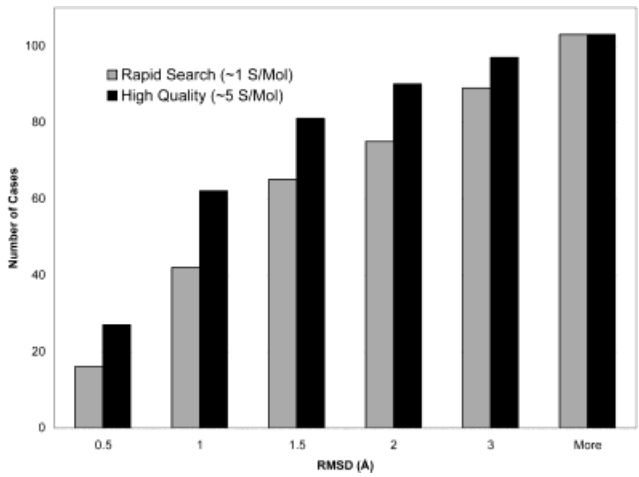


Fig. 4. The cumulative histogram of the minimum rms deviation (in Å) between the docked conformations and the crystallographically observed position of the ligand for the high-quality and rapid searches. This graph (and all subsequent cumulative histograms) should be interpreted as follows. The height, Y, of the bin at the point X is the number of cases in which a hit within X Å of the crystallographically observed position of the ligand was found.

ing packages that have been tested on the GOLD or similar test sets.[7,9,15,16,27]

The second problem is to correctly rank the docked compounds, i.e., is the top ranked conformation reasonably close to the crystallographically observed position for the ligand. This is a significantly more difficult problem than the first. The cumulative histograms of the rms deviation between the top scoring docked position and the observed position are shown for both sets of parameters in Figure 5 with the exact numbers for the high-quality runs given in Table I. In this case, there is little difference between the two sets of parameters. For the high-quality runs, 48 of the 103 cases produce a 2.0 Å hit as the top scoring docked position. This number drops to 41 at 1.5 Å, 34 at 1.0 Å, and 10 at 0.5 Å. For the rapid searches, 45 of the 103 cases produce a 2.0 Å hit as the top scoring docked position with 41 at 1.5 Å, 34 at 1.0 Å, and 10 at 0.5 Å.

The utility of the scoring function used in this study lies less as a tool to absolutely rank the docked conformations than as an initial filter to select only a few docked conformations. To show this, the cumulative histograms of the best rms deviations of all docked positions whose score was within 10% of the top scoring position are shown for both sets of runs in Figure 6. As is evident from comparing Figures 4 and 6, most of the well-docked positions, i.e., low rms deviations, survive this 10% cutoff. Most of the docked positions, however, do not. For the high-quality runs, on average, 74 positions are found, but after the 10% cutoff, on average, only 8 remain. For the rapid searches, on average, nearly 21 positions are found, but after the cutoff, on average, only 5 remain. At this point, the docked positions that survive the 10% score cutoff could be further optimized, visually screened, or passed to a more accurate but less efficient scoring function.

For the high-quality runs, the average CPU time (on an SGI R12000) per test case is approximately 4.5 s. At this rate, screening one million compounds with one CPU would take about 50 days. For the rapid searches, the average CPU time per test case drops to approximately 1.1 s per test case. At this rate, screening one million compounds with one CPU would take about 12 d. Because database docking is a highly parallel job, multiple CPUs could easily cut this to a reasonable amount of time (a day or so). In addition to the docking time, the hot spot search routines take between 1 and 3 min per case, and the grids take between 5 and 10 s to calculate. These final two aspects are required only once per case.

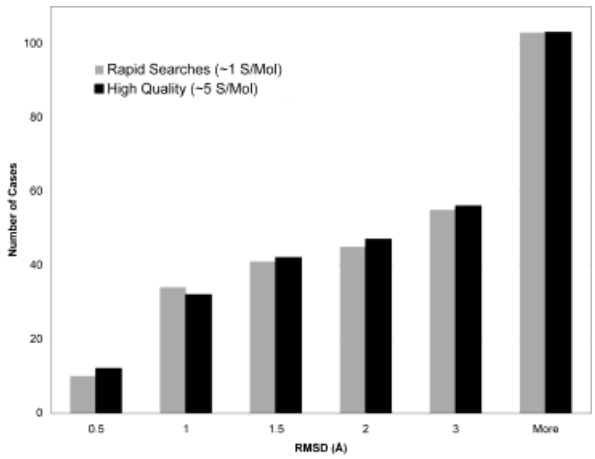


Fig. 5. The cumulative histogram of the rms deviations (in Å) between the top ranked docked conformation and the crystallographically observed position of the ligand for the high-quality and rapid searches.

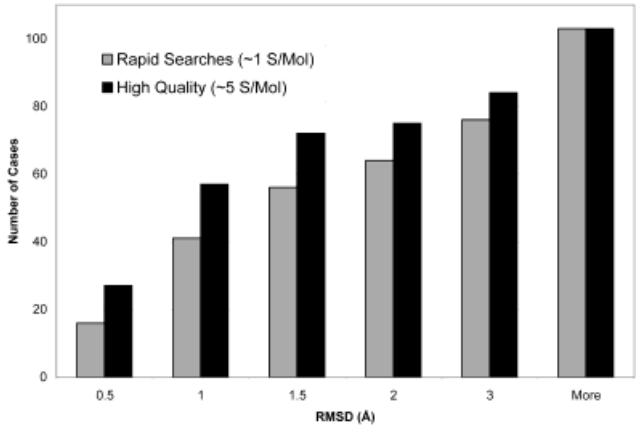


Fig. 6. The cumulative histograms of the best rms deviation between all docked positions whose scores are within 10% of the top ranked docked position and the crystallographically observed position for the ligand.

## Some Specific Successful Cases

In this section, a few of the successful cases show the strengths of our approach to docking small molecules. In all of these cases, the results are from the high-quality docking runs. The first case is the dipeptide ILE-VAL from the PDB entry 3tpi.[28] This case has no clear anchor fragment, and as a result, the incremental construction approach to docking might have difficulties with this ligand. Our conformational search procedure produced a conformation within 0.42 Å of the observed conformation. The rms deviation between the best scoring docked position and the observed position is 0.53 Å (see Fig. 7).

The second example, with a ligand having 15 rotatable bonds, is a much more difficult example. It is an HIV protease inhibitor from the PDB entry 1ida.[29] In this case, the conformational search procedure was able to generate a conformation with an rms deviation of 0.96 Å from the bound conformation. The best scoring docked conformation for this case is shown in Figure 8 with an overlay of the observed position for the ligand. The rms deviation for the top scoring docked position is 1.38 Å. In fact, the top 13 scoring docked positions are all within 2.0 Å of the observed position with the closest near 1.32 Å.

The final case is an HIV protease inhibitor from the PDB entry 4phv.[30] The ligand in this case has 12 rotatable bonds. This case clearly shows the value of including the final flexible gradient optimization of the ligand. The closest conformation produced from the conformational search procedure is 1.32 Å from the crystallographically

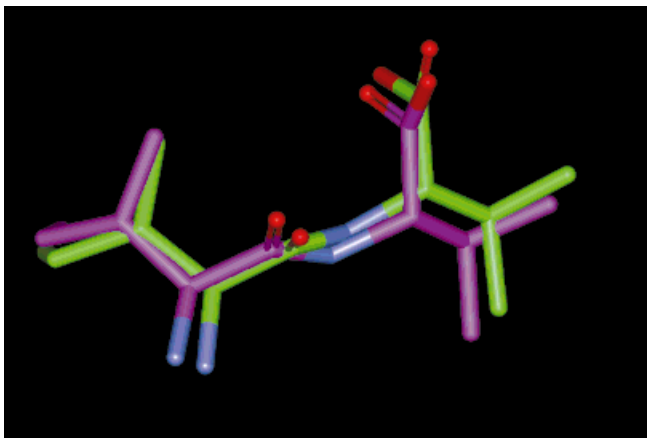

Fig. 7. PDB entry 3tpi.[28] The observed position for the ligand is shown with carbon atoms in green, and the top ranked docked position is shown with carbon atoms in magenta. Figures 7, 8, 9, and 13 were created with Web Lab Viewer Pro.

Fig. 8. PDB entry 1ida.[29] The top ranked docked position is shown with carbon atoms in magenta relative to the observed position for the ligand with carbon atoms in green.

observed conformation. An overlay of the top scoring docked position and the observed position is shown in Figure 9. With an rms deviation of 0.38 Å, the top scoring docked position is also the closest to the observed position. The smallest rms deviation that could have been obtained without the flexible optimization is that of the closest conformation generated by the conformational search procedure, i.e., 1.32 Å. Thus, in this case, the flexible optimization decreased the final rms deviation by nearly 1.0 Å.

### An Analysis of the Errors and Avenues for Improvement

It is often assumed that when a docking simulation fails, the score has failed, i.e., the global minimum of the scoring function did not correspond to the crystallographically determined position for the ligand. Because the docking problem involves many degrees of freedom, it is reasonable to believe that in many cases the failure can be attributed to insufficient search. It is the goal of this section to

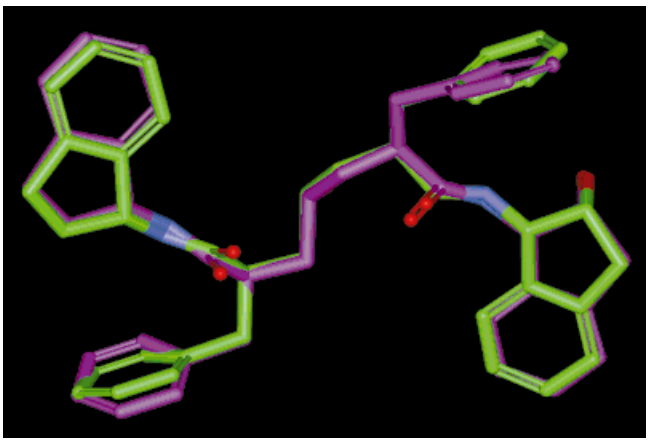

Fig. 9. PDB entry 4phv.[30] The top ranked docked position is shown with carbon atoms in magenta relative to the observed position for the ligand with carbon atoms in green.

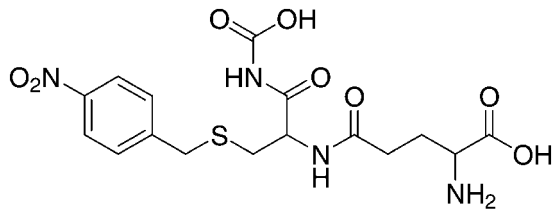


Fig. 10. The structure of the ligand in the PDB entry 1glq.

identify the cause of failure in the cases in which our procedure performed poorly.

To classify docking failures as either scoring failures or search failures, the ligand was taken as bound to the protein and a BFGS optimization was performed. If the resulting score was significantly less than the best score found from the docking runs, the failure is classified as a search failure. Every other failure is classified as a scoring failure.

The vast majority of the cases qualify as moderate scoring errors, i.e., the global minimum appears not to correspond to the crystallographically observed position of the ligand, but the percent difference between the global minimum and the best score near the crystallographic position of the ligand is $<10\%$. In these cases, it is difficult to decide which aspects of the score are failing, but it is reasonable to believe that many of these cases can be corrected simply by including some more detail in the scoring function, such as angular constraints on the hydrogen-bonding term or a solvation model. There are, however, a few cases with dramatic scoring errors. These cases provide some insight into the weakness of our score and the complexities of protein/ligand interactions.

The case 1glq[31] pointed out the main weakness of the score used in this study—hydrogen-bonding patterns. This is a polar ligand (see Fig. 10). The top ranked position for this ligand scores very well largely because of many "perceived" hydrogen bonds. In reality, these hydrogen bonds would be extremely weak because the angular

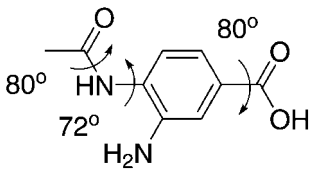


Fig. 11. The structure of the ligand in the PDB entry 1ive.

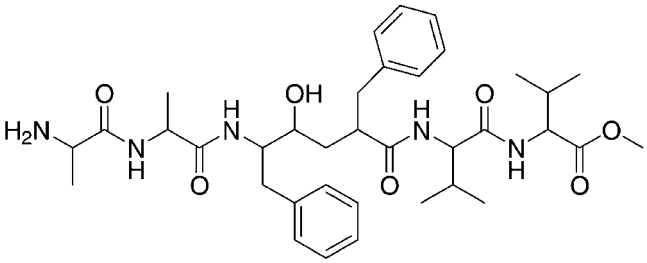


Fig. 12. The structure of the ligand in the PDB entry 1hef.

dependence of the interaction is poor. Moreover, the sulfur atom in the X-ray position is accepting a hydrogen bond from the OH of a tyrosine, and the carboxylic acid is involved in a salt bridge with a lysine. Neither of these interactions can be correctly treated by our scoring function.

In the case 1ive[32] (see Fig. 11), the correct position receives a relatively poor score largely because of the estimated strain of the observed conformation. Our program recognizes the three bonds marked in Figure 11 as being conjugated. Thus, a stiff penalty is applied when these bonds are not planar. In the observed conformation, the dihedral angles are all nearly 80° from planar. If these dihedral angles are forced to be near 0°, the conformation is no longer compatible with the observed interactions between the ligand and the protein. It would be difficult for any docking algorithm to predict these values for these dihedral angles.

The case 1hef,[33] an HIV protease inhibitor (see Fig. 12), is perhaps the most interesting of all of the dramatic scoring errors. The binding pocket is at the interface of a dimer with the protein monomers being related through a crystallographic symmetry operation. At the C-terminus of the ligand, a methyl group is within 2.0 Å of an $O_\delta$ of Asp 29 (see Fig. 13a). In addition, a carbonyl oxygen has an internal contact with a phenyl ring with a distance of 2.3 Å. These interactions would be extremely difficult to predict. Our program did come up with an interesting alternate conformation for the C-terminus of the ligand (see Fig. 13b). This conformation eliminates both the internal and external steric clashes and forms an additional hydrogen bond with the protein (see Fig. 13b).

There are two cases that can be classified as conformational search failures: 1hef and 1poc. In these cases, the best conformation produced is 2.1 Å and 2.3 Å, respectively. The ligand in the case 1poc has 23 rotatable bonds, and thus, it is very difficult to fully cover its conformational space with only 50 conformers. Although the ligand

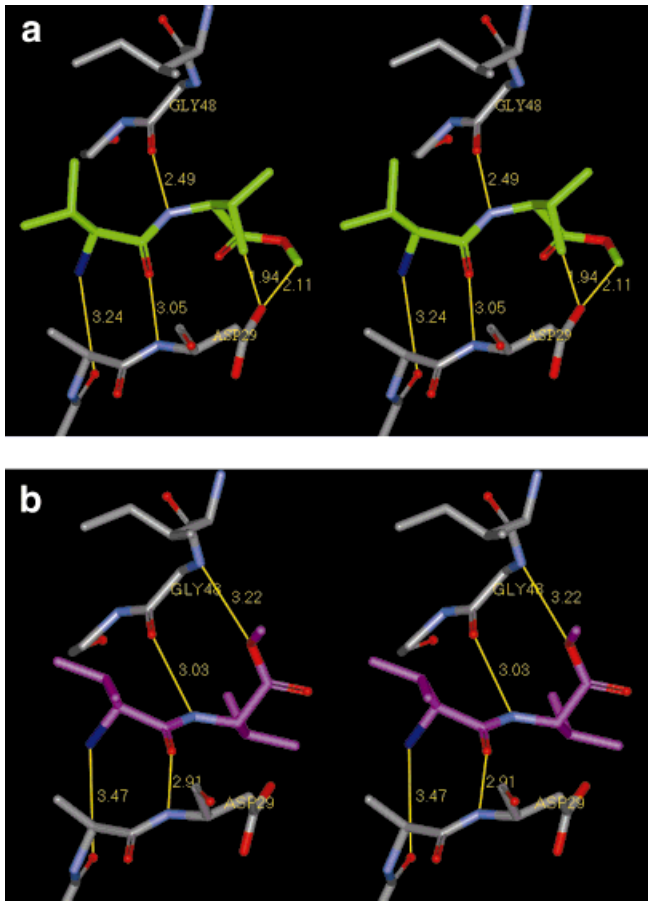


Fig. 13. A comparison of the crystallographically determined position for the ligand in PDB entry 1hef to the top ranked docked conformation. **a:** The crystallographically determined position for the ligand. The carbon atoms of the protein are colored gray. The carbon atoms of the ligand are colored green. The heteroatoms are given their usual color with the exception that the nitrogen at which the ligand was cut is colored a dark blue. **b:** The position of the top ranked docked conformation. The carbon atoms of the protein are colored gray. The carbon atoms of the ligand are colored magenta. The remaining atoms are colored as in Figure 13a.

in the case 1hef is also very flexible (18 rotatable bonds), the observed conformation, as described above, also has a serious steric clash. Thus, this case should be expected to be a very difficult challenge for any conformational search procedure.

Most of the remaining cases that are classified as search failures are polar molecules and, therefore, rely on matching some hydrogen bonding atoms to polar hot spots. Usually, the apolar hot spots are reasonable, but often the polar hot spots are poor. Many of these polar hot spots are in position to hydrogen bond to relatively exposed portions of the protein rather than in favorable pockets. Thus, in the future new methods for generating polar hot spots will be investigated.

### Library Docking and Analysis

As a simple application of this docking method to a combinatorial library, we dock the library PL792[34] into the binding site of plasmepsin II *of plasmodium falcipa-*

*rum* (pdb identifier 1sme[35]). For this case, a 20 Å × 32 Å × 22 Å box around the active site was used. The ECLIPS® combinatorial library PL792 consists of 13,020 compounds. These compounds were based on the core of pepstatin, and as a result, are relatively large: molecular weight = 530 ± 75 and number of rotatable bonds = 19 ± 3. The conformations were generated by using Catalyst version 4.5,[36,37] allowing a maximum of 100 conformations per molecule. The time to generate the conformations was 10.2 s per compound. The conformations were then docked by using the rapid search mode as described above. The time required to dock the compounds was 4.2 s per compound.

The library was designed around the core of pepstatin, and the crystal structure used in the study is a complex between plasmepsin and pepstatin. Thus, as a first measure of the quality of the docked compounds the rms deviation between the core of each compound and the core of the crystallographically observed position of pepstatin can be calculated. The results of this calculation are shown in Figure 14. For this, only the top ranked pose from each compound was used. One third of the compounds are docked to within 1.5 Å of the crystallographically observed binding mode of pepstatin. Thus, even with the high molecular weight and flexibility of the compounds, a large percentage of the compounds that fit into the active site are well docked.

From this library, high throughput screens discovered 134 active compounds. Thus, a second measure of the quality of results is the ability to preferentially pull active compounds from the library by using various scores. This is a particularly difficult challenge because all the compounds are rather similar to pepstatin, and many will likely receive good scores. For this we use the piecewise linear potential[24] and the potential of mean force.[38,39] Two values are reported as measures of the success of the two scores in pulling out actives. The first of these measurements is the enrichment factor,[40] which is given by

$$\text{Enrichment} = \frac{a/n}{A/N} \tag{4}$$

where $N$ is the number of compounds in the library (13,020 in this case), $A$ is the number of active compounds (134 in this case), and $a$ is the number of active compounds in the top $n$ compounds. The second of these measurements is the statistical significance of the enrichment, which is given by

$$\text{Significance} = \sum_{k=a}^{A} \frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}} \tag{5}$$

where $N,A,n,a$ are defined as for the enrichment. Loosely, the enrichment is the ratio of the number of actives found in the top $n$ compounds to the number of compounds that one would expect to find if n compounds were chosen at random, and the significance is the probability that a random selection of the same number of compounds would have done as well or better. The enrichment values and their significance are shown in Figure 15. For this case, the piecewise linear potential performs better than the potential of mean force. The piecewise linear potential achieves enrichments of greater than two with a maximum enrichment of 3.3, whereas the potential of mean force achieves a maximum enrichment of 1.7. For the piecewise linear potential, the statistical significance ranges from $10^{-3}$ to $10^{-5}$, meaning that it is very unlikely that the results are coincidental.

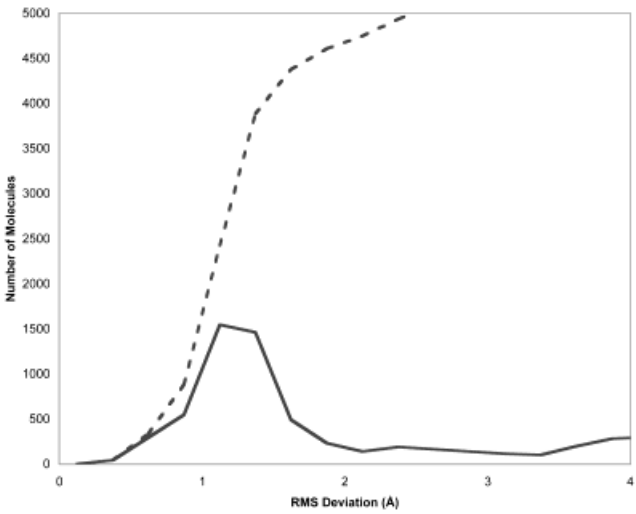


Fig. 14. The rms deviation distributions between the core of each compound from the library and the core of pepstatin as observed in the binding site of plasmepsin. The solid line is the histogram, whereas the dashed line is the cumulative histogram.

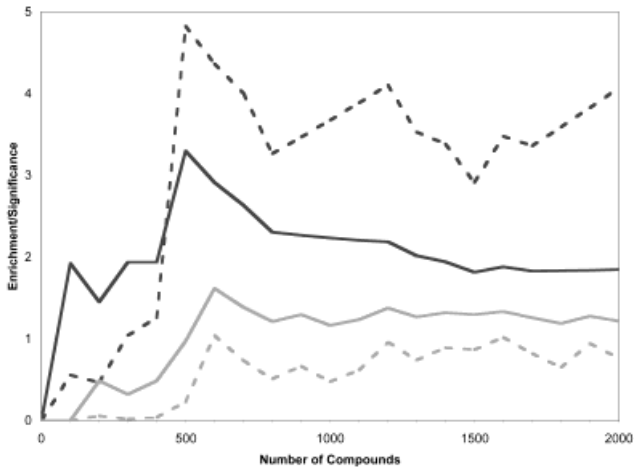


Fig. 15. The enrichment values (solid) and $-\log_{10}$(significance) (dashed) values for the piecewise linear potential (black) and the potential of mean force (gray).

## CONCLUSIONS

In this article, we have presented a new rapid method for docking flexible ligands into the binding sites of proteins. The method is based on a pregenerated set of conformations for the ligand and a final flexible gradient-

based optimization of the ligand in the binding site of the protein. Based on our results, this is a robust approach to handling ligand flexibility. With relatively few conformations (<50 per molecule), usually a conformation within 1.5 Å of the bound conformation can be generated. Applying the flexible optimization as the final step reduces the number of conformations required while maintaining high-quality final docked positions.

There are opportunities to improve nearly every step in the present docking method. The conformer generation, although reasonably successful, should treat small relatively rigid molecules and large flexible molecules differently. Because the conformational space of very large flexible molecules is too large to explore thoroughly, a Monte Carlo search algorithm is necessary. But for smaller molecules, a thorough search could be performed. In addition, the score used to rank the conformations is certainly too simplistic and can be improved. For example, variations of solvation models[20,41] would likely give better conformations. Finally, a better treatment of strain, particularly that for rotation about bonds between two Sp2 atoms, might lead to improved results.

The algorithm used to find the polar hot spots may need the greatest attention. In particular, the polar hot spots tend to find any hydrogen bond donor and acceptor rather than those buried in the binding site. Improving the hot spot search routine will not only increase the quality of our algorithm but will also decrease the number of hot spots needed and, thus, make the algorithm more efficient. Some available program, such as GRID[5] or the LUDI binding site description[42] or a documented method[43] would likely show some improvement. In addition, separating the polar hot spots into donor, acceptor, ionic, etc., hot spots might improve our results. Finally, in a practical application, most users would be willing to spend some time to enhance the image, i.e., eliminate by hand bad hot spots, and add hot spots where needed. In practice, this has significantly improved the docking runs.

As with all docking programs, the quality of the score is a significant problem. A good score should be efficient, error tolerant, and accurate. The score used here satisfies the first two qualities. These two qualities, however, are usually not compatible with the third. It appears that this score will still be useful as an initial screen after which a more accurate score can be applied. Some obvious areas for investigation include geometric constraints for the hydrogen-bonding term, recognition of ionic interactions and solvation effects, and terms for dealing with metals.

Even with the significant room for improvement, we believe that when a crystal structure is available, our approach to molecular docking will be useful in library screening prioritization. The real challenge from our perspective will be whether the method can be useful for such purposes when only a homology model is available. Because we use a scoring function that relies on "soft" potentials and no geometric constraints, we have reason to believe that even with lower quality structural information, such as a homology model, our method will still provide useful information.

## ACKNOWLEDGMENTS

We thank Marvin Waldman and C.M. Venkatachalam of MSI for helpful discussions about docking strategies.

## REFERENCES

1. Charifson PS. Practical application of computer-aided drug design. New York: Marcel Dekker; 1997. 551 p.
2. Knegtel RMA, Wagener M. Efficacy and selectivity in flexible database docking. Proteins Struct Funct Genet 1999;37:334–345.
3. Debnath AK, Radigan L, Jiang S. Structure-based identification of small molecule antiviral compounds targeted to the gp41 core structure of the human immunodeficiency virus type 1. J Med Chem 1999;42:3203–3209.
4. Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 1999;42:5100–5109.
5. Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 1985;28:849–857.
6. Goodsell DS, Olson AJ. Automated docking of substrates to proteins by simulated annealing. Proteins Struct Funct Genet 1990;8:195–202.
7. 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.
8. Westhead DR, Clark DE, Murray CW. A comparison of heuristic search algorithms for molecular docking. J Comput Aided Mol Des 1997;11:209–228.
9. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD. Flexible docking using Tabu search and an empirical estimate of binding affinity. Proteins Struct Funct Genet 1998;33:367–382.
10. Diller DJ, Verlinde CLMJ. A critical evaluation of several global optimization algorithms for the purpose of molecular docking. J Comp Chem 1999;20:1740–1751.
11. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecular-ligand interactions. J Mol Biol 1982;161:269–288.
12. Kuntz ID. Structure-based strategies for drug design and discovery. Science 1992;257:1078–1082.
13. Makino S, Kuntz ID. Automated flexible ligand docking method and its application for database search. J Comp Chem 1997;18:1812–1825.
14. Rarey M, Kramer D, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–489.
15. Rarey M, Kramer B, Lengauer T. The particle concept: placing discrete water molecules during protein-ligand docking predictions. Proteins Struct Funct Genet 1999;34:17–28.
16. Rarey M, Kramer B, Lengauer T. Docking of hydrophobic ligands with interaction-based matching algorithms. Bioinformatics 1999;15:243–250.
17. Welch W, Ruppert J, Jain AN. Hammerhead: fast fully automated docking of flexible ligands to protein binding sites. Chem Biol 1996;3:449–462.
18. Venkatachalam CM, Kirchhoff P, Waldman M. Receptor-based pharmacophore perception and modeling. In: Guner OF, editor. Pharmacophore perception, development and use in drug design. La Jolla, California: International University Line; 2000.
19. Press WH, Teulkolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C. 2nd ed. Cambridge: Cambridge University Press; 1997.
20. Eisenberg D, McLachlan AD. Solvation energy in protein folding and binding. Nature 1986;319:199–203.
21. Ooi T, Oobatake M, Nemethy G, Scheraga HA. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. Proc Natl Acad Sci 1987;84:3086–3090.
22. Vajda S, Weng Z, Rosenfeld R, DeLisi C. Effect of conformational flexibility and solvation on receptor-ligand binding free energies. Biochemistry 1994;33:13977–13988.
23. Rarey M, Wefing S, Lengauer T. Placement of medium-sized molecular fragments into active sites of proteins. J Comput Aided Mol Des 1996;10:41–54.

24. Gehlhaar DK, Verkhivker GM, Rejto PA, et al. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. Chem Biol 1995;2: 317–324.
25. Murray CW, Baxter CA, Frenkel D. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. J Comput Aided Mol Des 1999;13:547–562.
26. Haug EJ, McCullough MK. A variational-vector calculus approach to machine dynamics. J Mech Trans Auto Des 1986;108:25–30.
27. Kramer B, Rarey M, Lengauer T. Evaluation of the FlexX incremental construction algorithm for protein-ligand docking. Proteins Struct Funct Genet 1999;37:228–241.
28. Marquart M, Walter J, Deisenhofer J, Bode W, Huber R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr 1983;B39:480.
29. Tong L, Pav S, Mui S, et al. Crystal structures of HIV-2 protease in complex with inhibitors containing hydroxyethylamine dipeptide isostere. Structure 1995;3:33–40.
30. Bone R, Vacca JP, Anderson PS, Holloway MK. X-ray crystal structure of the HIV protease complex with L-700,417, an inhibitor with psuedo C2 symmetry. J Am Chem Soc 1991;113:9382–9384.
31. Garcia-Saez I, Parraga A, Phillips MF, Mantle TJ, Coll M. Molecular structure at 1.8 A of mouse liver class pi glutathione S-transferase complexed with S-(p-nitrobenzyl)glutathione and other inhibitors. J Mol Biol 1994;237:298–314.
32. Jedrzejas MJ, Singh S, Brouillette WJ, Laver WG, Air GM, Luo M. Structures of aromatic inhibitors of influenza virus neuraminidase. Biochemistry 1995;34:3144–3151.
33. Murthy KHM, Winborne EL, Minnich MD, Culp JS, Debouck C. The crystal strucutures ast 2.2-A resolution of hydroxyethyleene-based inhibitors bound to human immunodeficiency virus type 1 protease show that the inhibitors are present in two distinct orientations. J Biol Chem 1992;267:22770–22778.
34. Carroll CD, Patel H, Johnson TO, et al. Identification of potent inhibitors of plasmodium falciparum plasmepsin II from an encoded statine combinatorial library. Bioorg Med Chem Lett 1998;8:2315–2320.
35. Bernstein NK, Cherney MM, Loetscher H, Ridley RG, James MNG. Crystal structure of the novel aspartic proteinase zymogen proplasmepsin II from plasmodium falciparum. Nat Struct Biol 1999;6:32–37.
36. Smellie A, Kahn SD, Teig SL. Analysis of conformational coverage. 2. Applications of conformational models. J Chem Inf Comput Sci 1995;35:295–304.
37. Smellie A, Kahn SD, Teig SL. Analysis of conformational coverage. 1. Validation and estimation of coverage. J Chem Inf Comput Sci 1995;35:285–294.
38. Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. J Med Chem 1999;42:791–804.
39. Muegge I, Martin YC, Hajduk PJ, Fesik SW. Evaluation of PMF scoring in docking weak ligands to the FK506 binding protein. J Med Chem 1999;42:2498–2503.
40. Guner OF. Pharmacophore perception, development and use in drug design. La Jolla, CA: International University Line; 1999, 537 p.
41. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
42. Bohm H-J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. J Comput Aided Mol Des 1992;6: 593–606.
43. Mills JEJ, Perkins TDJ, Dean PM. An automated method for predicting the positions of hydrogen-bonding atoms in binding sites. J Comput Aided Mol Des 1997;11:229–242.