Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites

William Welch, Jim Ruppert and Ajay N Jain

Background: Molecular docking seeks to predict the geometry and affinity of the binding of a small molecule to a given protein of known structure. Rigid docking has long been used to screen databases of small molecules, because docking techniques that account for ligand flexibility have either been too slow or have required significant human intervention. Here we describe a docking algorithm, Hammerhead, which is a fast, automated tool to screen for the binding of flexible molecules to protein binding sites.

Results: We used Hammerhead to successfully dock a variety of positive control ligands into their cognate proteins. The empirically tuned scoring function of the algorithm predicted binding affinities within 1.3 log units of the known affinities for these ligands. Conformations and alignments close to those determined crystallographically received the highest scores. We screened 80 000 compounds for binding to streptavidin, and biotin was predicted as the top-scoring ligand, with other known ligands included among the highest-scoring dockings. The screen ran in a few days on commonly available hardware.

Conclusions: Hammerhead is suitable for screening large databases of flexible molecules for binding to a protein of known structure. It correctly docks a variety of known flexible ligands, and it spends an average of only a few seconds on each compound during a screen. The approach is completely automated, from the elucidation of protein binding sites, through the docking of molecules, to the final selection of compounds for assay.

Address: Arris Pharmaceutical Corporation, 385 Oyster Point Boulevard, South San Francisco, CA 94080, USA.

Correspondence: Ajay N Jain
e-mail: jain@arris.com

Key words: computational screening, conformational analysis, flexible docking, ligand binding, molecular docking

Received: **27 Dec 1995**
Revisions requested: **22 Jan 1996**
Revisions received: **13 May 1996**
Accepted: **15 May 1996**

Chemistry & Biology June 1996, **3**:449–462

 ISSN 1074-5521

Introduction

As the number of proteins with known structure has increased, molecular docking techniques have become prominent in lead compound discovery. The molecular docking problem begins with a three-dimensional structure for an enzyme or receptor whose biological function can be modulated through the binding of small molecules. The goal of a docking algorithm is to generate the conformation and alignment (pose) of a small molecule that maximizes the value of a scoring function intended to predict the relative or absolute binding affinity.

Here, we focus on using docking as a tool to screen large databases of small molecules against a given protein target, with the expectation of producing a set of potential lead compounds for assay and follow-up development. There have been a wide variety of docking algorithms proposed in the literature [1–8], but few are practical for large database screening due to long execution times and considerable human effort per ligand. A database screen typically produces a large number of successfully docked ligands, many of which are geometrically possible yet chemically unfavorable, and many more than it is generally practical to assay. The docking algorithm must rank the molecules, and only the most favorable will be assayed for activity. To be an effective screening tool, it is not enough for a docking algorithm to reproduce the known geometry of a flexible ligand in a protein binding site; it must also be able to recognize this geometry as the optimal pose and generate a reasonably accurate prediction of the binding affinity. If a non-crystallographic pose of a ligand scores significantly better than the crystallographic one, the docking algorithm's ability to map contact geometries into binding affinities is suspect. If artificially high scores are assigned to non-binding ligand geometries, a high false-positive rate may result from the screen.

A variety of scoring functions have been used to rank docked poses, including those using as their primary criterion steric contact [1,9], molecular mechanics [10], electrostatic complementarity [11] (the comparative performance of these is discussed in [12]), *ad hoc* measures of steric and polar complementarity [2,4], and an empirically derived estimator of binding affinities [13]. Unfortunately, none of these scoring functions reliably ranks the crystallographically determined poses as being the best possible poses. A related problem is that the functions are unreliable at ranking the binding affinities of different ligands in the same site, and this can lead to a

large number of false positives in a screen. If the favored pose for a ligand diverges significantly from the experimentally determined one, the scoring function's ability to rank candidate compounds for assay is questionable. Because of the general unreliability of scoring functions at ranking ligand binding, it is customary to use screening tools in a coarse-grained way, by selecting a large number of the top-scoring ligands and then relying on visual inspection and modeling techniques to manually select a small subset for actual assay. When using the DOCK program, for example, a great deal of human effort goes into this selection process [14].

Even given a perfect scoring function where manual selection would not be necessary, docking is a difficult search problem. The number of possible poses (the pose space) that must be considered is huge. Disregarding protein flexibility, a pose includes the torsion angles of the ligand's rotatable bonds, and the six rigid-body alignment parameters that position the ligand relative to the protein. It would take a prohibitive amount of time to score every possible pose. Suppose, for example, a ligand with four rotatable bonds is to be docked into a protein's active site measuring roughly 10 Å on a side. Considering angles in 10-degree increments and translational parameters on a 0.5 Å grid, there are over 6×10^{14} poses to be tested, a computation requiring ~2 000 000 years at the rate of 10 poses per second. Thus, only a very small proportion of pose space can be explored, and there must be a trade-off between search time and search thoroughness.

Here we describe Hammerhead, a completely automatic, fast docking procedure for flexible ligands that uses an empirically tuned scoring function and an automatic method for identifying and characterizing the binding site on a protein. In a screen for compounds that bind to proteins with known ligands, Hammerhead predicts the natural ligand as the top-scoring compound in several cases, ranging from relatively weak interactions with limited ligand flexibility (e.g., trypsin–benzamidine) to strong interactions with significant ligand flexibility (e.g., dihydrofolate reductase (DHFR)–methotrexate). The algorithm is fast enough to allow screening of a library of roughly 10 000–100 000 small organic compounds in a few days, yet thorough and accurate enough that it correctly predicts crystallographic poses and binding affinities for a variety of known ligands.

Results and discussion

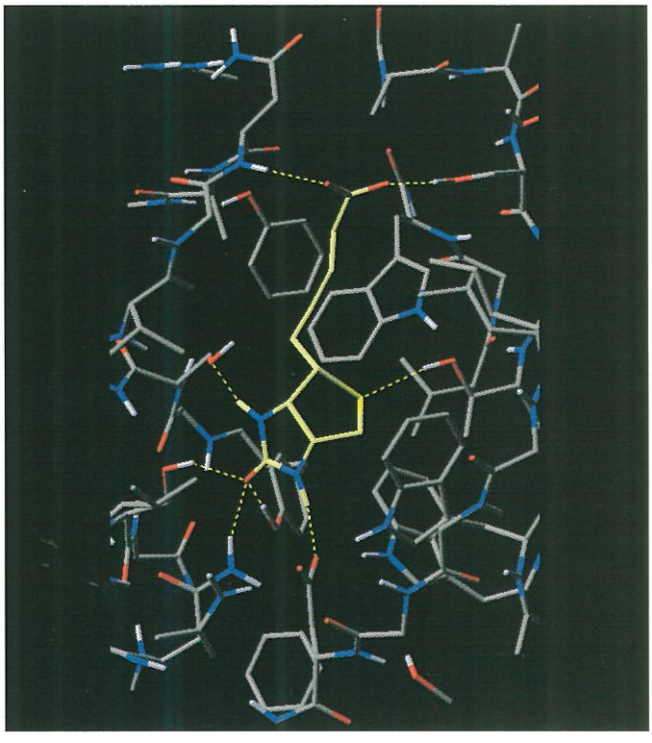
Hammerhead is composed of three principal computational elements: an empirical scoring function that accurately predicts the affinities of a wide variety of known ligands given their crystallographic poses, a protein 'pocket finder' that automatically identifies binding regions and ideal contact geometries on target proteins, and a fragment-based alignment and conformational search procedure for small molecules.

The scoring function and the automatic pocket finder

The scoring function is presented in detail elsewhere [15]. Briefly, the scoring function is a smooth nonlinear function that, given a ligand pose and a protein structure, estimates a binding affinity for the complex. The parameters of the function were tuned using a wide range of co-crystal complexes with known affinities (e.g., Fig. 1). The scoring function estimates binding affinities in units of $-\log K_d$, with an expected mean error of 1.0 log unit based on cross-validation. The function is differentiable with respect to ligand pose, so it is possible to refine a pose by performing gradient descent on the score, from a starting point some distance from the optimal pose.

The automatic pocket finder uses the scoring function to determine where a ligand might make energetically favorable interactions with the protein. This is done by coating the protein surface with steric and polar probes (H, C=O, and N–H fragments), positioned so that each makes an optimal interaction with the protein, then searching for clusters of probes having high predicted affinities. From these clusters, a pocket (a focused subset of the probes) is automatically constructed, with the requirement that all portions of the pocket must be well

Figure 1

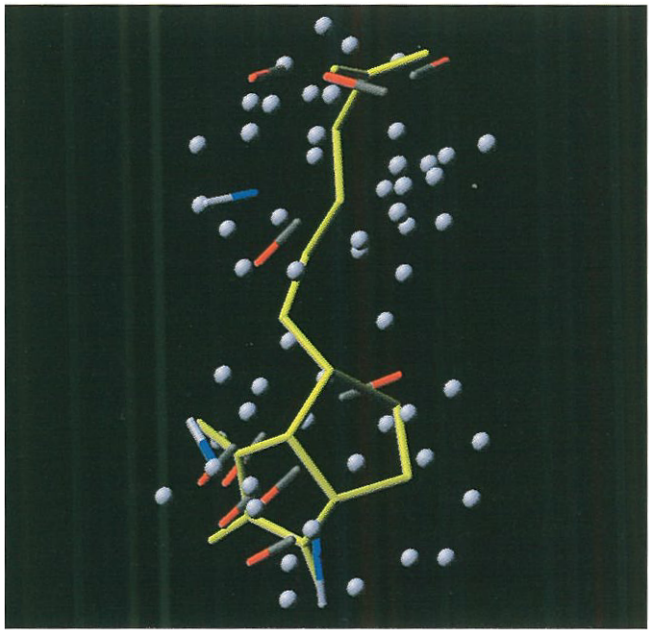


Co-crystal structure of biotin (yellow backbone) bound to streptavidin. Red, oxygen; blue, nitrogen; white, hydrogen.

connected, and that the overall pocket geometry must be amenable to small molecule ligands. The pocket finder designates particularly high-scoring probes, such as those making multiple hydrogen bonds or favorable hydrophobic ring face contacts, as anchor probes. A docking algorithm may use these anchor probes to focus the ligand alignment process by requiring that at least one of these energetically important interactions be made. The pocket produced for streptavidin is illustrated in Figure 2. The pocket finder has been applied to a wide range of crystal structures, and has successfully identified the known ligand-binding regions as the best-scoring pockets in every case. It is also a principled way of finding small-molecule binding sites when no sites are known, as is often the case with protein–protein interactions for which small-molecule antagonists are sought.

Here, we focus on ligand alignment and conformational search procedures and how they combine with the scoring function and pocket finder to yield a fast, automated flexible docking algorithm. The approach for conformational searching breaks ligands into fragments and constructs poses one fragment at a time. The scoring function is used throughout the reassembly process to optimize partial poses and to score the current pool of poses so that only the best-scoring candidates are followed up. This search strategy reins in the combinatorial explosion in the number of poses that would accompany exhaustive follow-up.

Figure 2



The pocket finder locates a binding pocket on the protein by coating the protein surface with steric and polar probes and identifying clusters of probes with high predicted affinities. The streptavidin pocket probe set is shown, with the conformation of biotin from the co-crystal structure superimposed in yellow. N–H groups indicate ideal positions for hydrogen-bond donors; C=O groups indicate ideal positions for hydrogen-bond acceptors; and spheres indicate favorable hydrophobic contact points.

Docking algorithm overview

Given a three-dimensional structure for a target protein, and an arbitrary conformation of a putative ligand, the goal is to compute a complementary pose for the ligand that maximizes the scoring function. To begin constructing a pose, it is assumed that, for a high-affinity small molecule, some portion will make a specific binding interaction with the protein above some threshold. This portion of the ligand we refer to as the head, because such functionality is often found at one of the ligand's extremities. Head fragments are determined automatically by breaking ligands into fragments, aligning each of these onto the protein by matching ligand atoms with probe atoms, and retaining the best-scoring ligand alignments. For each possible head, the remaining molecule fragments (the tail) are aligned one fragment at a time into the probe neighborhood about the end of the current head, and then merged with the head to yield correct bond angles (a process that we term 'chaining'). At each stage of the fragment alignment computation, gradient-descent pose optimization improves the conformation and alignment of the growing ligand, relaxing van der Waals surface interpenetrations of the ligand with itself and the protein, and improving hydrogen bond and hydrophobic surface contact geometries.

For each partially constructed pose, there will typically be many possible alignments of the next tail fragment. This can lead to a combinatorial explosion in the number of alignments as a fragment chain is being assembled, just as can happen in a pure conformational search. But, because the partial conformations are situated in the protein, the scoring function can be used to rank them. Only the top-scoring partial poses are followed up, based on the principle that a high-scoring ligand will generally have good partial scores as well. This yields a search that grows roughly linearly with the number of fragments, although this is at the risk of losing some poses that do not distinguish themselves early on as winners.

As an example of the potential savings of such a fragment-based approach over an exhaustive conformational search followed by rigid docking, suppose a ligand with 8 rotatable bonds is to be conformationally sampled at 120° increments at each bond. This is extremely coarse sampling, but nevertheless yields 3^8 conformations (roughly 6500) to be aligned to the protein and scored. If instead the ligand is split into three fragments having two torsions each (two torsions are eliminated by the fragmentation), and bonds are again sampled at 120° angles, there will be 3(fragments)•3^2(conformations) = 27 total fragment conformations. Using a piecewise alignment

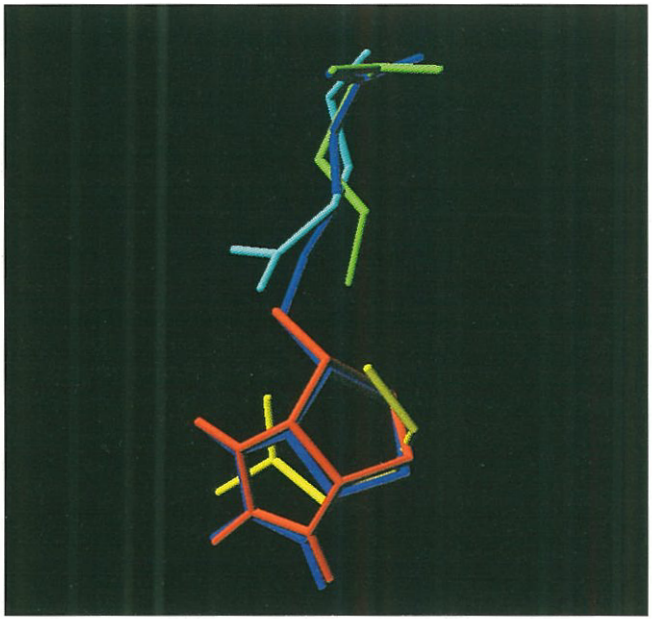
approach, and using the scoring function to rank and follow-up only the five best-scoring partial alignments at each stage, the entire process requires only 27(initial fragment alignments) + 5(follow-ups)•3^2(second fragment conformations) + 5(follow-ups)•3^2(third fragment conformations) = 117 total alignment calculations.

Various stages of this fragment-based docking approach are illustrated in Figures 3–5. A detailed description of the docking algorithm is given in Materials and methods.

Tuning Hammerhead's search parameters

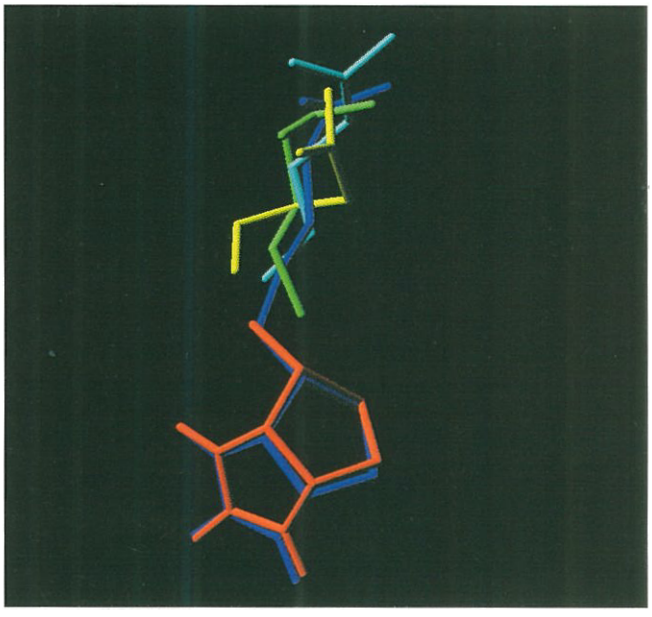
Pose space is much too large to search exhaustively when docking a ligand, so it is crucial to have a principled way of both focusing the search and of terminating it early for non-binding ligands, while ensuring the recognition of ligands that do bind. Hammerhead uses a variety of numerical search cutoff parameters to discard unpromising poses at various points in the docking process. These include score cutoffs for head fragments, root-mean-square deviation (RMSD) filtering thresholds for collections of aligned fragments, and the number of top-scoring fragments that will be followed up whenever a larger number is produced. These are described in detail in Materials and methods.

Figure 3



Hammerhead determines the optimal pose of a ligand by breaking it into fragments and aligning them with the pocket probes. The fragments that make specific binding interactions above a certain threshold are referred to as head fragments. The alignment of biotin head fragments to the streptavidin probes is shown. The probe atoms are not shown; the pose of biotin from the co-crystal structure is shown in blue for reference. The top-scoring head is the hydantoin fragment (red). Next best is the flexible methylene chain in green. The two 'incorrect' alignments of this fragment (yellow, cyan) score significantly lower than the other alignments.

Figure 4

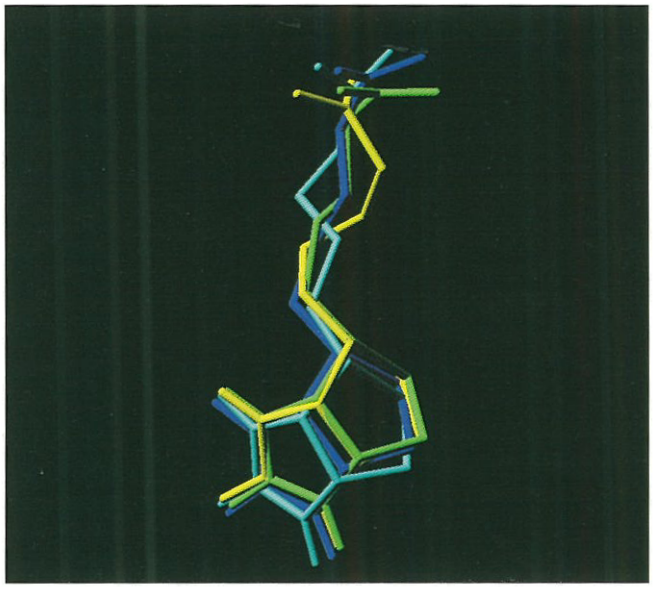


After aligning the head fragments, other fragments of the ligand (tail fragments) are aligned to the head fragment and the pocket probes. These fragments are later merged with the head (see Fig. 5). Alignments of biotin tail fragments onto the aligned hydantoin head and streptavidin pocket probes are shown. The probe atoms are not shown; the pose of biotin in the co-crystal structure is shown in blue for reference.

Choosing values for these parameters is an empirical issue, and it is not particularly difficult to tune Hammerhead on a single complex to generate dockings with 1 or 2 Å RMSD from the crystal structure with a search time of a few seconds, for that lone example. But such focused tuning is of little value for database screening as, in this situation, the algorithm must be able to find many different kinds of binding motifs and search many different chemical functionalities. By seeking parameter settings that allow high-scoring poses to be discovered over a broad range of controls (at the expense of longer search times), the false negative rate of the screening tool on novel complexes should be decreased.

We used four protein–ligand crystal structures as positive computational controls for tuning Hammerhead's search parameters: DHFR–methotrexate, trypsin–benzamidine, streptavidin–biotin, and thrombin–NAPAP (*N*-alpha-(2-naphthyl-sulphonyl-glycyl)-DL-*p*-amidinophenylalanyl-piperidine) [15]. The ligands have from 1 to 11 freely rotatable bonds, and binding affinities ranging from $10^{-4.7}$ to $10^{-13.4}$, and form a variety of hydrogen bonds, salt bridges, and hydrophobic interactions with their respective proteins. Geometrically, this is a diverse set, with each of the ligands exhibiting a different kind of flexibility and binding mode: benzamidine is small and mostly rigid, methotrexate has two rigid fragments and a flexible tail bound in a large pocket, biotin has a rigid head and a

Figure 5



The tail alignments from Figure 4 are merged with the head, yielding completed poses for biotin. The pose of biotin in the co-crystal structure is shown in blue for reference.

flexible tail bound in a very constricted pocket, and NAPAP contains two hydrophobic ring systems (one aromatic, one aliphatic) that must be inserted into a large pocket in a highly constrained manner. Pockets (probe sets) were generated for each of these proteins using the automatic pocket finder; some of the probe statistics are summarized in Table 1. Because Hammerhead's fragment breaking/chaining and conformational search procedures take their bond lengths and bond angles from the input conformation of the ligand, the ligand must have an energetically reasonable geometry. Therefore, the crystallographic ligand conformations were energy minimized *in vacuo* prior to docking.

The parameter settings used to successfully dock all of the test cases in a minimum amount of time are listed in Table 2. These were determined by first tuning those parameters related to rigid alignment, so that a favorable head alignment for each of the controls was produced in a minimum amount of time. Geometric chaining parameters were then established that allowed the tails to be

Table 1

Breakdown of pocket probes.

Protein	Donors	Acceptors	Sterics	Total	Anchors
DHFR	34	24	205	263	42
Trypsin	35	17	69	121	43
Streptavidin	19	16	146	181	51
Thrombin	51	15	233	299	50

Table 2

Hammerhead search parameter settings.[a]

Min break torsions	3	Nhood points	20
Min break atoms	20	Match tries	100
Sample angle	120°	Match error	1.5 Å
Sample min RMS	1.0 Å	Match RMS	1.0 Å
Sample min linker	30°	Head min	3.0[b]
Match min	4	Optimize heads	20
Match max	8	Keep heads	3
Ligand points	15	Keep tails	20
Merge distance	2.6 Å	Merge angle	57°

[a]For definitions, see Materials and methods.
[b]Reported as $-\log K_d$.

successfully aligned onto the head fragments (whose poses typically deviated slightly from the poses in the crystal structures). Finally, score-related search cutoffs were chosen to increase overall search speed while retaining the successful poses.

Positive controls: performance on known complexes

The search parameters listed in Table 2 were set by tuning the algorithm using a single minimized conformation of each control ligand. As a more thorough positive control, we explored Hammerhead's sensitivity to the conformation of the input ligand. The conformations of biotin, methotrexate and NAPAP were stochastically sampled, and 20 conformations were generated for each ligand; the conformations were then energy minimized *in vacuo* using MM2 or AMBER force-fields (the fourth control, benzamidine, had only one conformation after minimization). In an actual screen of a chemical database, of course, each compound would appear in only one, arbitrary, energy-minimized conformation.

Hammerhead was applied to these sampled conformations using the probe sets generated by the pocket-finder, with and without the anchor-probe restriction, which stipulates that at least one head fragment atom be aligned with an anchor probe. The dockings produced were generally identical, with or without the anchor restriction, although the search was faster with the restriction. A variety of torsion angle sampling densities were tried, but the original sparse sampling at 120° increments (followed by van der Waals relaxation) proved to be the most time-effective. The results of the dockings of the multiple conformations of the four controls are summarized in Table 3.

There are two key considerations in assessing the performance of the docking algorithm as a screening tool, based on these positive controls. The most critical issue is whether the docking of each of the ligands results in a computed binding affinity that is above a threshold to be assayed, implying that we would actually test the compound as a result. In a screening run, ligands that

Table 3

Performance with positive controls.

Target	$-\log K_d$ actual	$-\log K_d$ max, found	RMSD from crystal structure[a]	1.0 log average RMSD between high-scoring poses	$-\text{Log}K_d$ min	% of hits close to crystal pose	Find compound in screen?
Trypsin/benzamidine	4.7	6.0	0.6	0.6	6.0	100 %	100 %
Streptavidin/biotin	13.4	12.5	0.6	0.5	10.2	100 %	100 %
DHFR/methotrexate	9.7	10.1	1.7	1.5	6.4	60 %	100 %
Thrombin/NAPAP	8.5	8.8	1.3	2.9	4.9	25 %	100 %

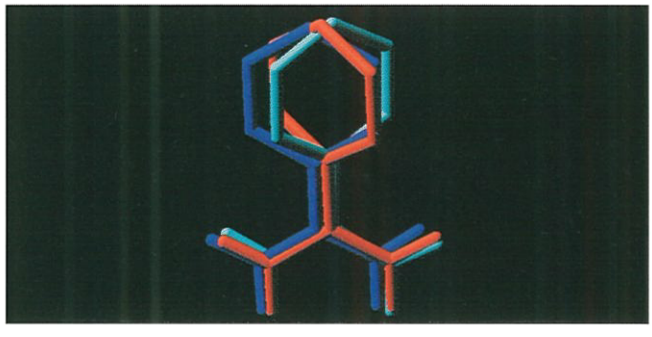
[a]RMSD, root mean square deviation. A heavy-atom RMS measures the RMS distance between the nearest atoms of the same type on two posed ligands. It is used to avoid atom-numbering artifacts when symmetric sub-structures are mirror-aligned.

dock with scores above some threshold are considered 'hits' and are investigated further with assays. Given that the best expected error of the scoring function is 1.0 log units [15], the threshold on scores is set at 4.0 (score units are $-\log K_d$) to reliably detect ligands that bind with affinities of 10 μM or better. Thus, it is not absolutely necessary that the pose with the maximum possible score be generated, so long as a pose is found that demonstrates the ligand can score above the screening threshold. In all four positive-control screens, beginning with all randomly selected conformations, Hammerhead generated poses for the ligands that exceeded the screening threshold.

To determine the validity of the scoring function, it is also important to determine that the best-scoring poses of the ligands correspond to the crystallographically determined poses. If poses divergent from this correct pose yield significantly better scores than those near the correct pose, it is very likely that the scoring function is not performing adequately and will generate false positives. In all four cases, across all dockings, the best-scoring poses were within 1.8 Å of those determined crystallographically and the scores were within 1.3 log units of the correct affinities. The mean RMSD of the highest-scoring poses was 1.0 Å, and the mean error of computed affinity was 0.7 log units. Thus, Hammerhead successfully reconstructed the optimal binding modes for the controls and recognized them as such. No alternative poses were judged to have higher affinities. Additionally, for benzamidine, biotin and methotrexate, any pose that deviated significantly from the correct pose did not score nearly as well, as shown by the low average RMSD for poses that scored within 1.0 log unit of the top-scoring pose. The scoring function maxima are narrow enough that the correct poses were clearly distinguishable from any others solely on the basis of their scores. Examples of these maximal and near-maximal docked poses appear in Figures 6–9.

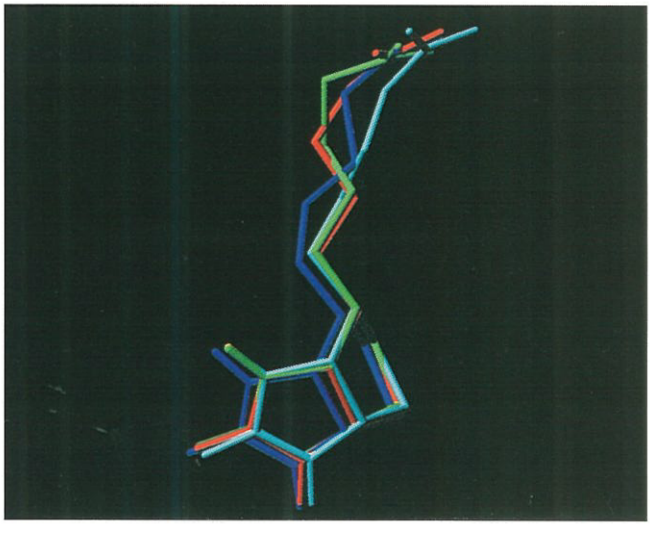
Considering all poses that scored within 1.0 log units of the maximal pose in affinity, the mean RMSD is extremely

Figure 6



Hammerhead accurately reconstructs the crystallographically determined pose for benzamidine bound to trypsin. The top-scoring poses are shown. The pose of benzamidine in the co-crystal structure is shown in blue for reference; the highest-scoring docked pose is in red. The phenyl group in the docked poses is rotated out of the plane of the amines to avoid a van der Waals clash. The planar conformation in the crystal structure is not energetically favorable, and the underlying density map may be an average of the docked conformations.

Figure 7



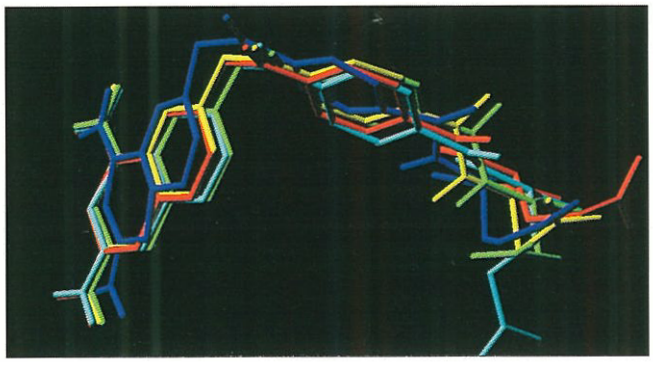
Hammerhead accurately reconstructs the crystallographically determined pose for biotin bound to streptavidin. The top-scoring poses are shown. The pose of biotin in the co-crystal structure is shown in blue for reference; the highest-scoring docked pose is shown in red.

low for all but the thrombin–NAPAP complex. For this control, Hammerhead always placed the benzamidine and piperidine portions of NAPAP in the correct binding positions, but often left the naphthalene moiety rotated out into solvent. This explains the low percentage of docked poses that were close to the crystallographically correct poses (Table 3). An alternative binding mode was also discovered that placed the naphthalene group on the outside of Trp600, effectively sandwiching the tryptophan between NAPAP's hydrophobic ring systems. The predicted affinity for this pose is very nearly that of the top-scoring pose, explaining the larger average RMSD for high-scoring poses. The methotrexate example showed some similar effects. All methotrexate dockings correctly positioned the pteridine fragment, but several dockings left part of the acidic tail rotated out into solvent. It is important to note that these alternative poses for NAPAP and methotrexate scored well enough to exceed the screening threshold, so the compounds would have been assayed as a result of docking even in cases where their predicted poses were only partially correct.

Run-times on compound libraries

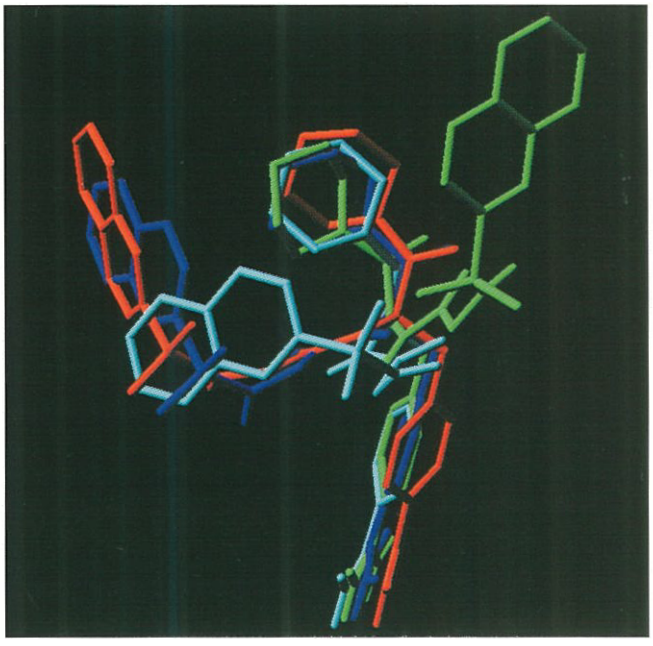
In a screen of a large chemical database, Hammerhead's run-time is an important consideration. The examples discussed in the previous sections range from 20s to 400s docking time per ligand on an SGI R4400 150 MHz processor. In a screen of a large database, docking times of hundreds of seconds per compound would be prohibitive. Fortunately, Hammerhead's search algorithm typically takes much less time to decide that something is a docking miss than it does to produce an optimal pose for a hit; and the vast majority of compounds in a database screen will be misses. The average run-times for misses from a collection of 1000 compounds, randomly selected from the Available Chemicals Directory (ACD), are shown in Table 4. The compounds all have six or fewer rotatable bonds, and the distribution of flexibilities matches that of the complete ACD. The fragment breaking/chaining procedure is used for ligands with more than three rotatable bonds, and straightforward conformational sampling followed by rigid docking is

Figure 8



Hammerhead accurately reconstructs the crystallographically determined pose for methotrexate bound to DHFR. The top-scoring poses are shown. The pose of methotrexate in the co-crystal structure is shown in blue for reference, the highest-scoring docked pose is shown in red. Several alignments of the pteridine fragment emerged as the best-scoring heads, and the phenyl and carboxy tail fragments were chained from these.

Figure 9



Hammerhead detects an alternative binding mode for NAPAP in thrombin. The top-scoring poses are shown. The pose of NAPAP in the co-crystal structure is shown in blue for reference; the highest-scoring docked pose is shown in red. Several alignments of the benzamidine fragment emerged as the best-scoring heads, and the piperidine and naphthalene groups were chained from these. The poses that rotate the naphthalene group out into solution (green and cyan) rather than tucking it into the pocket score significantly lower than the pose with the correct binding mode (red), but high enough that NAPAP bound in these poses would still be selected in a screen of thrombin inhibitors.

Table 4

Average docking time versus ligand torsions for computational 'misses' from 1000 random compounds against streptavidin.

#Torsions	#Ligands	Time (s)[a]	Time (s)[b]
0	135	13	7
1	192	19	9
2	205	47	22
3	187	89	36
4	134	66	39
5	63	105	61
6	10	115	67

[a]Without anchor-probe restriction.
[b]With anchor-probe restriction.

5

ge docking times for computational 'misses' for the ɔl pockets.

	Time (s)[a]	Time (s)[b]
avidin	52	25
ı	56	9
	71	12
bin	89	9

ɔut anchor-probe restriction.
anchor-probe restriction.

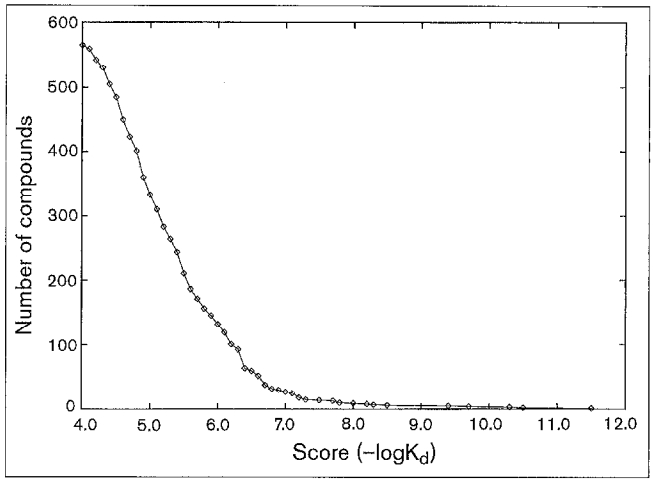
elsewhere. The average time to detect a docking within this set of 1000 ligands is shown in Table 5 .ch of the control proteins. These dockings were run with and without the anchor-probe restriction. Using ɔr probes makes a dramatic improvement in the ge run-times for these misses because the ·utational time for a miss consists mainly in vering that there are no suitable head fragment ments, a search that is expedited by restricting ments to using at least one anchor probe. It is this ge time performance that makes Hammerhead cal for screening large libraries of compounds.

tive controls: a computational screen of the ACD

nore ambitious computational test of Hammerhead, n a screen of the ACD against streptavidin. As the tavidin–biotin interaction is so strong, this experi- provides a nice set of computational negative ols for Hammerhead. It is extremely unlikely that .CD contains molecules that are better ligands for tavidin than biotin and its close analogs. The ACD ins over 100 000 compounds, including biotin and a ɔer of other known ligands to streptavidin. A set of 0 compounds was generated by protonating and nizing the three-dimensional structures generated by rogram CONCORD for molecules containing seven wer torsion bonds. These were docked into the tavidin pocket. The screen ran in under five days, a variety of SGI Indigo and Challenger processors, about four processors running at any given time nines drew ligands from a shared pool). The screen ıced 565 compounds with scores predicted better 4.0 (100 μM). A histogram of the scores of the ɔd compounds is shown in Figure 10.

ıl binding assays for compounds discovered in this n are ongoing and will be presented elsewhere. ugh it is not expected that a majority of these 565 ounds will actually bind to streptavidin, preliminary indicate the presence of novel specific ligands. nerhead's ranking of known streptavidin ligands in ɔreen was an important computational control. The of biotin, iminobiotin, desthiobiotin, and HABA ·droxyazobenzene-2-carboxylic acid) amongst the

Figure 10



Hammerhead selects molecules from the ACD with high affinity for streptavidin. The number of molecules scoring greater than or equal to a given score (out of 80 000 molecules screened) is plotted. Biotin and its analogs scored higher than any other compounds in the screen (see also Table 6).

screening hits is shown in Table 6. (Thiobiotin, another high-affinity streptavidin ligand, is not present in the ACD, although Hammerhead successfully docks it with a score of 7.8). Biotin was predicted to bind with the highest affinity, scoring 2 log units better than any non-biotin derivative. Iminobiotin docked within 0.7 Å of biotin's alignment. Desthiobiotin did not successfully dock at first, but examination of the ACD-provided structure revealed the wrong stereoisomer was present (optical isomers are not enumerated during a screen). A conformation of desthiobiotin with chirality similar to biotin docked within 0.7 Å of the biotin structure and scored 10.5. Thus, the docking procedure is specific enough to distinguish the handedness of the streptavidin pocket. HABA scored close to its assayed value, but the ligand clashed with the protein badly enough that it was not retained in the screen. This could be due to the treatment of the protein as rigid.

Table 6

Affinities and ranks of biotin and derivatives for binding to streptavidin in a screen of the ACD.

Ligand	–logK$_d$ actual	–logK$_d$ predicted	Rank/565
Biotin	13.4	11.5	1
Desthiobiotin[a]	[b]	10.5	2
Iminobiotin	10.5[c]	10.3	3
HABA	4.0	4.4	[d]

[a]Stereoisomer of ACD CONCORD structure.
[b]Affinity unknown.
[c]Affinity of binding to avidin.
[d]Excessive protein interpenetration; pose was not retained in screen.

Other considerations for large-scale screening

One consideration that is not addressed in the positive controls, but which becomes important in large-scale screening, is protein flexibility. Small conformational changes are expected to occur in the protein upon ligand binding, which relieve protein–ligand clashes and optimize the particular ligand's binding. Hammerhead treats the protein as rigid during the docking process; however it does contain a separate scoring function, which provides a measure of protein–ligand interpenetration, computed from the penetration of the ligand's van der Waals surface into the protein surface [15]. This interpenetration measure is not used when computing the binding affinity. Unlike other scoring function parameters, such as hydrogen bond strengths and distances, the range and relative importance of this interpenetration score cannot be deduced using the positive controls, since none of them exhibits significant interpenetration. Thus, the interpenetration score is treated as an independent quality measure, and is used during docking to discard 'hopelessly' interpenetrating poses. Hammerhead will ultimately produce poses with a range of interpenetration scores under this discard threshold. A high score implies that the ligand does not actually fit into the pocket, and its affinity prediction is therefore questionable; a ligand with a small value, on the other hand, might be accommodated by the protein. Although the accuracy of the binding affinity predictions will probably suffer by not considering protein flexibility, the estimates should still be useful for selecting a set of compounds for assay. The discard threshold for the interpenetration score can only be determined empirically, and will probably be a function of the target protein's binding-site geometry. To gain an understanding of this parameter's effect, Hammerhead is being run on a variety of test systems, and computational hits having a range of interpenetration scores are being assayed.

Another consideration that is not resolved by examination of crystal structures is the cost of ligand and protein desolvation. None of the 34 crystal structures used to tune the scoring function shows significant contact between polar atoms on the ligand or protein and an opposing hydrophobic surface, so it was not possible to deduce the cost of displacing solvent molecules from unmatched polar sites during ligand binding. Many of the compounds docked in the streptavidin screening run have unmatched polar features, and the expected reduction in binding due to solvation effects is not well modeled by the scoring function. As with the interpenetration discard threshold, it may be possible to empirically calibrate the solvation threshold, if the screens and assays of test systems produce a variety of ligands having buried polar surfaces. As solvation effects do not vary much as a function of pose (for a fixed binding mode), we expect that the solvation threshold could be applied as a final screening filter, rather than having to be integrated into the docking and scoring processes.

Relation to previous flexible docking approaches

Several docking approaches have been discussed in the literature [1–8]. Our approach has elements in common with each of these. One of the earliest published approaches, DOCK [1], is commonly used as a screening tool. The approach begins with the construction of a negative image of the target site — a set of spheres positioned outside the target against its solvent-accessible surface. Typically, a large set of spheres (300–400) is generated automatically, and these are edited by hand down to a manageable set (40–60), focusing the search towards particular regions of the protein surface. To dock a ligand, internal distances between ligand atoms are matched against distances between subsets of the sphere centers, and, when these agree within some tolerance level, the ligand atoms are aligned onto the matching spheres. Finally, the aligned ligand is tested for protein interpenetration, and the alignment is discarded if it interpenetrates excessively. A more recent screening tool, FLOG [2], uses a similar matching procedure to align ligands, and refines the alignments online, allowing some poor alignments to be salvaged, as well as improving the complementary positioning of the hydrophobic and polar features of the ligand.

Both of these approaches focus on rigid alignment of a given ligand conformation. The docking algorithm FLOG accounts for some ligand flexibility within this rigid alignment framework by storing a number of different conformations (10–20) of a ligand and docking each one separately. The problem with this approach to ligand flexibility is that the number of conformations required to span the conformation space scales up exponentially with increasing numbers of rotatable bonds, and thus there is good chance that the correct conformation will have been missed, even for ligands of limited flexibility.

A number of docking approaches have attempted to incorporate conformational search directly into their ligand placement procedure by using, for example, stochastic optimization techniques such as simulated annealing [3] or genetic algorithms [4]. Although each of these approaches has been shown to be viable for flexible docking, they are both extremely slow. Fragment-based approaches, similar to Hammerhead's, have also been proposed. The algorithm developed by DesJarlais *et al.* [5] breaks a ligand into rigid fragments, aligns all of the fragments onto the protein, and merges those fragments whose linkers are close enough to one another. Gradient descent on an energy field is used to adjust the positions of merged fragments. The dockings performed by DesJarlais *et al.* [5] use ligands with only two rigid fragments, and this approach will probably suffer from a

combinatorial explosion of the size of the pose space for very flexible ligands. Our approach narrows its search by selecting a few promising head-fragment alignments, then focusing successive fragment alignments onto the linkers to which they must merge. Because the scoring function is reliable at ranking (and improving) the poses of intermediate fragments, this search strategy can succeed without examining an exponentially large number of poses, and many more flexible ligands can be docked successfully.

The docking algorithms DOCK [6], FLEXX [7], and ADAM [8] all reconstruct a ligand by chaining from of an aligned head fragment. They attach fragments to the rooted ligand by re-joining the linkers at a discrete set of torsion angles. This approach may miss the exact angle needed to thread a flexible ligand into a very specific pocket. Our method of aligning a fragment onto a linker neighborhood and then merging it with the aligned head allows the protein itself to suggest the best angle for the joint. Further, the automatic fragment decomposition and head alignment procedure obviates the need for a human to select the head fragment, a limitation of each of these earlier fragment-based approaches, which precludes their use as screening tools.

FLEXX and ADAM both make use of a scoring function to rank intermediate poses and follow only the best-scoring candidates, as a way of controlling combinatorial expansion of the search. ADAM is similar to our approach in the way it uses a scoring function to refine intermediate poses. However, as in the work of DesJarlais *et al.* [5], ADAM's docking of methotrexate in DHFR disregards its flexible acid tail, so it is not clear how well the approach works for truly flexible ligands.

Finally, the previously described docking approaches select protein regions for docking either as neighborhoods around the structure of the bound ligand or by *ad hoc* modeling techniques. Hammerhead's use of automatically generated pocket probes eliminates human analysis and bias in protein docking site selection, and makes it possible to determine potential binding sites in a principled way.

Significance

We have presented a fully automated, flexible, molecular-docking procedure, Hammerhead, that is suitable for screening databases of tens of thousands of compounds for binding to a protein. Potential binding sites on a target protein are automatically determined and collected into geometrically plausible pockets. Flexible ligands are automatically broken into fragments, docked to the protein, and reassembled in place to find optimal poses. All of this is overseen by an empirical scoring function that reliably predicts binding affinities for a wide range of known small-molecule ligands. We expect that high-scoring ligands from a screen of a large chemical database could be ordered and assayed without additional modeling or selection effort by a human.

Hammerhead performs well on predicting the pose of several known flexible ligands bound to various proteins: benzamidine to trypsin, biotin to streptavidin, methotrexate to DHFR, and NAPAP to thrombin. The top-scoring poses are very close to the experimentally determined ones and the predicted binding affinities of these poses are within 1.5 log units of experimentally measured affinities. With the exception of a reasonable alternative binding mode for NAPAP, poses that diverge from that in the co-crystal structure score substantially less than those that are similar. This ability to rank poses is critical for controlling the number of false positives in a database screen.

Hammerhead also performs well with a variety of negative controls. A screen of the 80 000 commercially available flexible molecules against streptavidin was completed in a few days. The top-scoring ligand was biotin, in its experimentally determined conformation, closely followed by biotin derivatives. We are currently using Hammerhead to screen various databases for novel ligands of proteases and cytokine receptors.

Materials and methods

Detailed description of the docking algorithm

The fragment-based alignment algorithm is outlined below, followed by a more detailed description of the steps and parameters involved.

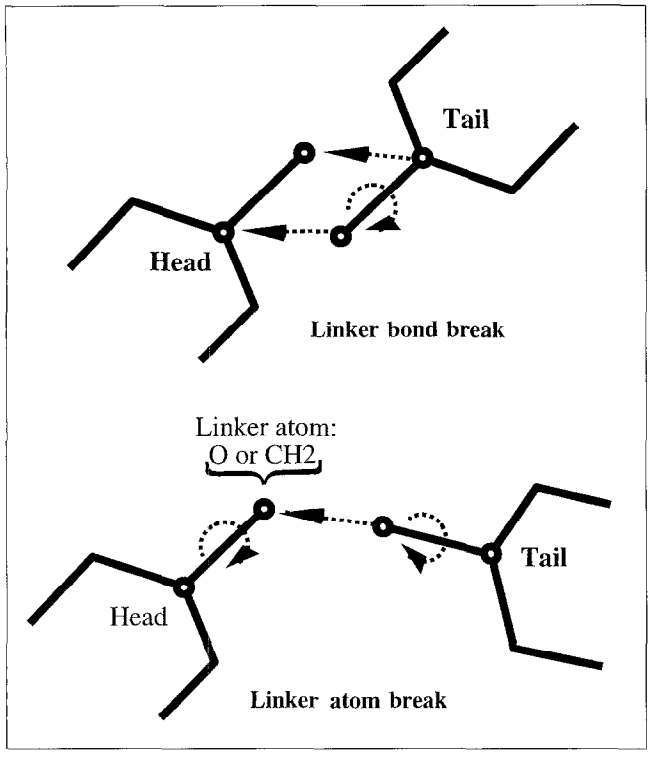
Input: a protein structure, a set of pocket probes (H, C=0, and N–H fragments) complementary to the active site, and a ligand in an arbitrary, *in vacuo* minimized, conformation.

Output: A set of ligand poses and a predicted binding affinity for each.

1. Generate ligand fragments:
   a. Break the ligand into fragments, each containing a limited number of rotatable bonds.
   b. Systematically search each fragment's conformations and retain a diverse set.
2. Compute a set of head fragments and poses:
   a. Rigidly align each of the ligand fragment conformations onto the pocket probes.
   b. Refine each pose, and eliminate each alignment that excessively interpenetrates the protein.
   c. Retain the top-scoring aligned fragments as heads.
3. For each of the heads align each successive ligand fragment
   a. Rigidly align the fragment's sampled conformations onto the probes at the tail of the partial pose.
   b. Merge each aligned fragment with the partial pose so their shared bond geometries agree.
   c. Refine the new partial pose, then eliminate those that excessively interpenetrate protein.
   d. Retain the best-scoring partial poses for addition of the next fragment.

There are a number of parameters that control this procedure, trading off search thoroughness for speed. Nominal values for each parameter

Figure 11



Breaking a molecule into fragments. For a break at a single rotatable bond (a linker bond break), each fragment retains a copy of the linker bond, and these will be overlaid when the fragments are eventually merged. For a break at a methylene carbon or ether oxygen (a linker atom break), each fragment retains a copy of the shared linker atom, and the methyl hydrogens are temporarily discarded. When the fragments are later merged (see Figs 14,15), the linker atoms are overlaid and the methyl hydrogens are then replaced.

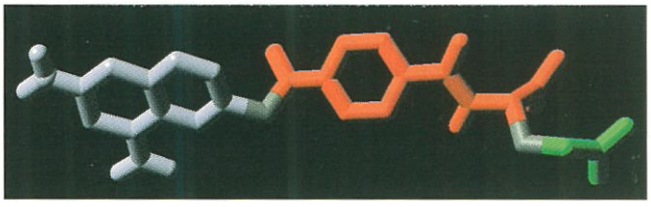
name in the discussion below (obtained by tuning the search procedure on a diverse set of co-crystal structures) appear next to the parameter name and also in Table 2.

Generating conformationally sampled ligand fragments

To break a ligand into fragments, first the set of freely rotatable bonds is determined. These are single bonds between non-terminal atoms that separate the ligand into two disjoint pieces (this excludes bonds that are part of ring systems, which will be addressed in future work). Bonds to symmetric end-groups like methyl groups are excluded, since it is not necessary to enumerate these rotamers. Any non-terminal methylene carbons or ether oxygens are identified. These are referred to as linker atoms and are always situated between two rotatable bonds. Ligands are preferentially broken at linker atoms (a linker atom break), but are broken at single rotatable bonds (a linker bond break) when no suitable linker atom is present (Fig. 11). Linker atom breaks are preferable to bond breaks because two unspecified torsion angles are solved simultaneously, and the alignments of the two fragments are less likely to be disturbed than with a bond break, which has a single free torsion angle.

Once linker atoms and breakable bonds have been identified, the ligand is broken into halves, using the linker break that yields the most balanced number of atoms in each fragment. If no balanced linker break can be found (nominally, each new fragment must be greater than 1/4 the size of the whole), the most balanced bond break is used

Figure 12



The fragment breaking algorithm applied to methotrexate yields three fragments joined by two linker carbons. The pteridine fragment is on the left, the acidic tail on the right.

instead. This breaking continues recursively until either the fragment has become too small (min break atoms = 20) or the fragment has few enough rotatable bonds that it can be conformationally searched as a whole (min break torsions = 3). Then, conformations for each fragment are systematically generated by rotating bonds in fixed angular increments (sample angle = 120°) and relaxing away from internal van der Waals clashes. Figures 12 and 13 illustrate the fragment decompositions of methotrexate and NAPAP.

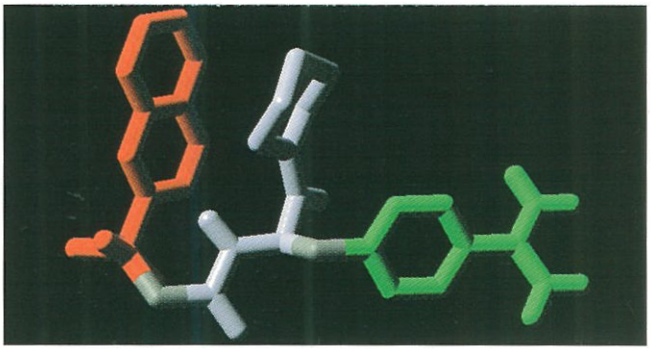
When internal van der Waals clashes restrict the range of fragment conformations, the systematic conformational search can yield clusters of similar conformations (after clashes have been relaxed). To eliminate this redundancy, the sampled conformations are filtered based on their atom center RMSD, measured after performing a least-squares alignment of the conformations onto one another.

During this fragment filtering process, it is important to protect diverse linker orientations, since these determine the range of directions in which successive fragments can be joined to the current one. To maintain a diverse linker population, the linker portion of the fragments is not used in computing the least-squares alignments of the fragments onto one another, or in computing the RMSD. Instead linker angular differences are measured separately (after alignment). Only fragments whose RMSD and linker direction differences are small enough (sample min RMS = 1.0 Å, sample min linker = 30°) are eliminated.

Rigid alignment onto pocket probes

One of Hammerhead's fundamental operations is rigid fragment alignment. Fragment conformations must be aligned to the protein to search for a head fragment and to generate alignments of the tails of

Figure 13



The fragment breaking algorithm applied to NAPAP yields three fragments joined by two linker carbons. The naphthalene group is on the left, the piperidine in the center, and the benzamidine group on the right.

partial poses to reconstruct full poses. In this rigid alignment procedure, the pocket probes serve as protein 'handles' to which fragments are aligned (Figs 1,2 show the crystal structure of biotin bound to streptavidin and the automatically generated probe set for streptavidin, respectively). Once an initial alignment has been computed that overlays some subset of the probes, the actual protein structure is used to refine the pose through gradient descent on the scoring function, simultaneously adjusting the fragment's bond angles and alignment.

The rigid alignment approach is related to that used in earlier docking work [1,2,9]. The idea is to match points on the fragment with pocket probes having compatible internal distances, then compute a rigid-body transformation that overlays the corresponding points as closely as possible. The rigid-body calculation is straightforward, given a set of pairs of ligand–probe points (a match). It is the search for matches that requires most of the work.

Only three pairs of points are needed to uniquely determine a rigid alignment, so some rigid alignment approaches simply enumerate all such matches. This leads to a large number of alignments, many of them duplicates (computed from disjointed subsets of paired points) and many of them nonsensical (having extreme ligand–protein interpenetration). Other approaches search for much larger matches to align multiple ligand features onto protein features, and thus these approaches produce fewer high quality alignments. But the number of different pairings that must be considered grows factorially with the size of the match desired. In our experience, the quality of the alignments does not correlate well with the size of the match beyond some fragment-specific size.

Hammerhead searches for relatively small matches (having match min = 4 to match max = 8 pairs), but enumerates them in a geometrically diverse order. Below, an outline of the alignment computation is provided, followed by a discussion of each step involved. The enumeration can be terminated early (typically with 50–200 matches) rather than having to blindly generate all possible matches (typically numbering in the tens of thousands), and this means that fewer alignments must be generated and tested later, in the fragment chaining process.

The rigid alignment algorithm proceeds as follows:

1. Sample the fragment at hydrogens, ring faces, and heteroatoms.
2. For each pocket probe, root a match search:
    a. Collect the probes near the root probe and sample them as was done with the ligand.
    b. For each ligand point of the same type as the root probe (steric, donor, or acceptor):
        i. Pair the ligand point and root probe.
        ii. Enumerate additional point–probe pairings having compatible internal distances in decreasing order of their distance from the points already in the match, and retain any matches of acceptable length.
        iii. Terminate the search after a predetermined number of pairings have been considered.
3. Compute rigid body alignments from the matches, and cluster the alignments.

Fragment and probe sampling

For rigid alignment, fragment shapes are represented as a collection of discrete sample points. Samples are placed at the center of each hydrogen atom (a steric point), and also in the centers of aromatic ring faces, offset one hydrogen radius from the surface for compatibility with the hydrogen-centered steric points. This procedure can yield many sample points, but it is the points at ligand extremities that are most important for constructing productive alignments. Therefore, to speed the match search, a subset of these points (ligand points = 15) that are maximally distant from one another is collected, starting with the point most distant from all others. Finally, samples are added at the center of each atom capable of donating or receiving a hydrogen bond or of being involved in salt-bridging (polar points).

Each match search is rooted at a particular probe. Probes more distant from the root probe than the ligand's maximum width cannot be paired with any ligand point and are thus ignored, leaving a smaller probe neighborhood to match against. This neighborhood is sub-sampled as was done for the ligand (nhood points = 20). For efficiency, this is pre-computed and cached for neighborhood samplings for a small set of fixed radii. The neighborhood sampling having the smallest inclusive radius is used when matching ligand points.

Match enumeration

A match is iteratively lengthened by searching for a pairing of ligand–probe points whose internal distances to points already in the match are equal, within a given error tolerance (match error = 1.5 Å). The search is depth first; pairs are added to a given match until its size reaches the match max threshold, or there are no more compatible pairs, at which point the match is recorded. When there are no more compatible pairs to lengthen a given match, the search backtracks and resumes with earlier, shorter sets of pairings.

Rather than exhaustively enumerating matches, Hammerhead terminates branches of this search early. For each possible initial match of length 2, a fixed maximum number of attempts are made to extend the match length by considering new pairs (match tries = 100). To bias this prematurely terminated search towards matches that more thoroughly cover the ligand, new pairs are tested in decreasing order of their distance from the points already in the match. This leads to higher quality alignments than matches that focus on localized subsets of the ligand points. For this reason it performs better than an enumeration that adds pairs in an order that minimizes accumulated match error, since the set of match points that gives the least total error may be clustered about only a portion of the molecule.

Computing and clustering alignments

Given a list of corresponding fragment–probe points, it is a straightforward calculation to compute a rigid-body transformation that aligns the fragment with the probes while minimizing the RMSD of corresponding points. Although match enumeration has been structured to generate a diverse collection of alignments, there will still be some redundant alignments generated (from mutually consistent but disjointed sets of points); thus additional filtering is warranted. It is not safe to filter alignments by computing the RMSD of their rotational and translational parameters. Instead, the computed alignments are applied to the fragment and a final RMSD filtering is performed based on the transformed fragment atom centers (match RMS = 1.0 Å).

Computing head fragment poses

We assume that there is at least one fragment of the ligand that makes a high-scoring interaction with the protein, specific enough that the fragment's optimal positioning is highly constrained. Rather than attempt to decide *a priori* which (if any) fragment of the ligand functions in this way, such head fragments are discovered automatically by docking all conformations of all fragments into the pocket and retaining the highest-scoring fragment poses. The steps in this process are outlined below:

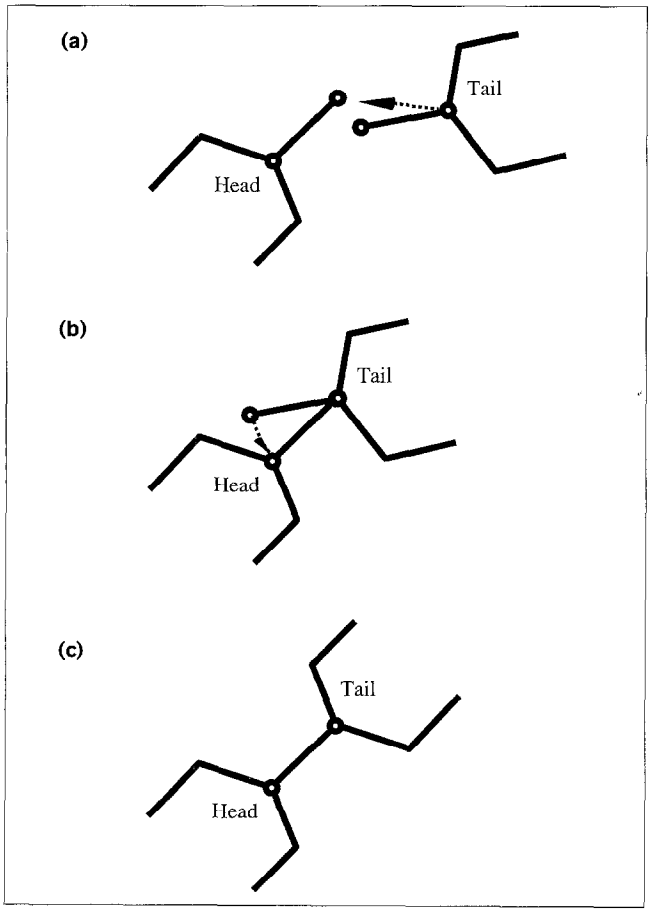
1. Align each fragment conformation onto the pocket probes, and eliminate those that excessively interpenetrate the protein.
2. Sort them by score, and RMS-filter similar poses (match RMS), preferentially retaining the higher-scoring poses.
3. Briefly refine the top scoring poses. (optimized heads = 20).
4. Sort the aligned conformations by score, and retain the top scorers as aligned head fragments (keep heads = 3). Discard any head fragments that score too low (head min = 3.0) .

The filtering step is important, because it leaves only the best-scoring exemplar of each fragment binding mode, reducing the number of candidates to be considered as possible heads. There is no guarantee that the chosen exemplar is the 'correct' one, but the expectation is that the retained one is close enough that subsequent scoring-function refinement during the reconstruction of the ligand (which adjusts torsion angles and the fragment alignment) will yield a suitable conformation.

The overall docking algorithm depends on one of the high-scoring fragment poses leading to the highest-scoring pose for the whole ligand; nevertheless, the top-scoring fragment need not be expected to do so. In addition to its correct pose, a fragment may have unrelated, better-scoring poses, and therefore a number of high-scoring head fragment poses (keep heads = 3) are retained.

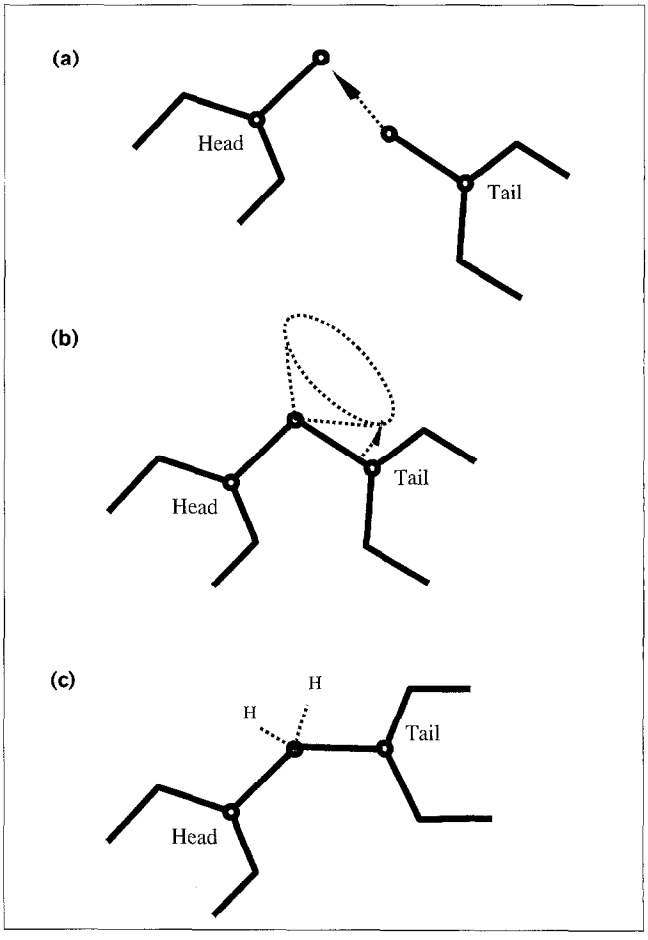
As a final shortcut in computing head alignments, a restricted set of anchor probes is used instead of the entire pocket (see Table 2). These are probes that have been designated by the automatic pocket finder as making a particularly favorable set of interactions with the protein. Matches are rooted only at these anchor probes, although the entire set of probes is still used for constructing the remainder of the match. The assumption is that a high-affinity ligand will take advantage of at least one of the pocket's most profitable interactions. One may choose to designate a root probe only at polar anchor probes in pockets where polar specificity plays a role in ligand recognition. Several head alignments for biotin in the streptavidin pocket are shown in Figure 3.

Figure 14



Steps in a linker-bond merge. **(a)** Translate the tail to align shared bond atoms. **(b)** Rotate the tail about the vector orthogonal to the shared bond vectors to orient the tail bond. **(c)** Completed bond merge.

Figure 15



Steps in a linker-atom merge. **(a)** Translate the tail to align linker atoms in head and tail fragments. **(b)** Rotate the tail about the vector orthogonal to the linker-bond vectors to restore the original bond angle (measured before the break). **(c)** Restore hydrogens for methyl carbons.

Chaining

Once a set of head fragments has been computed, the remainder of the ligand is docked into place, one fragment at a time, using the following procedure:

1. Align the fragment to the site, matching its linker with the linker of the existing partial pose.
2. For linkers close enough (merge distance = 2.6 Å), with similar enough orientations (merge angle = 57°), merge the newly aligned tail fragment to the existing partial pose so that bond angles are restored (Figs 14,15).
3. Sort the partial poses by score and refine the top-scoring merged fragments, then discard those that excessively interpenetrate the protein.
4. At each merge point, propagate only the top scores to continue to the next merge point (keep tails = 20).

Figures 4 and 5 illustrate tail alignment and merging. Pose refinement alleviates intra-ligand van der Waals clashes resulting from the merge, as well as clashes with the protein. The thoroughness of the flexible

search is controllable by the number of partial solutions that are propagated forward at each level. The intent is that sparse, fast searches will still pick out high-affinity ligands, on the assumption that they tend to have good partial scores even before the ligand has been completely docked.

Acknowledgements

We thank Teri Klein for collaborative efforts and for comments on the manuscript, and Mike Ross and Mike Venuti for support.

References

1. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. & Ferrin, T.E. (1982). A geometric approach to macromolecule–ligand interaction. *J. Mol. Biol.* **161**, 269–288.
2. Miller, M.D., Sheridan, R.P., Kearsley, S.K. & Underwood, D.J., (1994). 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.
3. Goodsell, D.S. & Olson, A.J. (1990). Automated docking of substrates to proteins by simulated annealing. *Proteins* **8**, 195–202.
4. Payne, A.W.R. & Glen, R.C. (1983). Molecular recognition using a binary genetic search algorithm. *J. Mol. Graphics* **11**, 74–123.
5. DesJarlais, R.L., Sheridan, R.P., Dixon, J.S., Kuntz, I.D. & Venkataraghavan, R. (1986). Docking flexible ligands to macromolecular receptors by molecular shape. *J. Med. Chem.* **29**, 2149–2153.
6. Leach, A.R. & Kuntz, I.D. (1992). Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comp. Chem.* **13**, 730–748.
7. Rarey, M., Kramer, B. & Lengauer, T. (1995). Time-efficient docking of flexible ligands into active sites of proteins. In *ISMB-95 Proceedings. Third International Conference on Intelligent Systems for Molecular Biology* (Rawlings, C., Clark, D., Altman, R., Hunter, L., Lengauer, T. & Wodak, S., eds), pp.300–308, AAAI Press, Menlo Park, CA.
8. Mizutani, M.Y., Tomioka, N. & Itai, A. (1994). Rational automatic search method for stable docking models of protein and ligand. *J. Mol. Biol.* **243**, 310–326.
9. Shoichet, B.K., Bodian, D.L. & Kuntz, I.D. (1992). Molecular docking using shape descriptors. *J. Comp. Chem.* **13**, 380–397.
10. Weiner, S.J., *et. al.*, & Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **107**, 765–784.
11. Gilson, M. & Honig, B. (1988). Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* **4**, 7–18.
12. Shoichet, B.K. & Kuntz, I.D. (1991). Protein docking and complementarity. *J. Mol. Biol.* **221**, 327–346.
13. Bohm, H.J. (1994). 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.
14. Ring, C.S., *et. al.*, & Cohen, F.E. (1993). Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc. Natl. Acad. Sci. USA* **90**, 3583–3587.
15. Jain, A. (1996). Scoring non-covalent ligand–protein interactions: a continuous differentiable function tuned to compute binding affinities. *J. Comput. Aided Mol. Des.*, in press.