Enhanced Virtual Screening by Combined Use of Two Docking Methods: Getting the Most on a Limited Budget

Vladimir Maiorov* and Robert P. Sheridan

Molecular Systems, Merck Research Laboratories, Merck & Co., Inc., RY50SW-100, P.O. Box 2000, Rahway, New Jersey 07065

Received March 21, 2005

Flexible ligand docking is a routine part of a modern structure-based lead discovery process. As of today, there are quite a number of commercial docking programs that can be used to screen large databases (hundreds of thousands to millions of compounds). However, limiting factors such as the number of commercial software licenses needed to perform docking simultaneously on multiple processors ("software cost") and the relatively long time required per molecule to get good results ("quality-to-speed") should be taken into account when planning a large docking run. How can we optimize the efficiency of selecting lead candidates by docking, in respect to the quality of the results, search speed, and software cost? We present a combination of two methods, our "fast−free−approximate" in-house docking program (Miller et al. *J. Comp.-Aided Mol. Des.* **1994**, *8*, 153-174) and the "slow−costly−accurate" ICM-Dock (Totrov and Abagyan. *Proteins* **1997**, *1* (Suppl.), 215-220), as an example of one solution to the problem. Our proposed protocol is illustrated by a series of virtual screening experiments aimed at identifying active compounds in the MDL Drug Data Report database. In more than half of the 20 cases examined, at least several actives per protein target were identified in approximately 24 hours per target.

INTRODUCTION

The use of molecular docking for virtual ligand screening (VLS) over databases containing up to hundreds of thousands or millions of compounds is a routine part of modern structure-based lead discovery. It has been proven to be useful in identifying lead compounds.[1−8] The quality of the results strongly depends on the time required to do a thorough docking calculation. There are a number of docking programs, many commercial, that do a credible job of "flex-on-the-fly" docking, but typically these take a minute or more per molecule. Fortunately, VLS of databases is easily parallelized by distributing chunks of a whole database across a series of networked processors. If the docking engine is a commercial program, and there is a sufficient number of available processors (which is the case in many research institutions worldwide), then the limiting factor is the number of available software licenses. Unless there is a special arrangement with the vendor, a typical user may not be able to afford the number of licenses needed for parallelization on a scale that can complete the search in a reasonable time. We would like to present an example of a solution to this problem. Our proposed two-step protocol is to use a rapid, approximate, and free docking method to select a subset of a database and then to use a slower, more accurate, commercial docking program to perform a detailed docking on only that subset. The ultimate goal is to combine the speed of the first step with the quality of the second to extract the maximum number of leads in the minimal time with the least cost in terms of software licenses. The applicability of the protocol is shown by its ability to identify known actives of a given protein target within a large database.

* Corresponding author e-mail: vladimir_maiorov@merck.com; phone: 732-594-0908; fax: 732-594-4224.

METHODS

The two programs used in the work were FLOG[9] (flexible ligands oriented on grid) for the first step of the protocol and ICM-Dock for the second.[10−12]

FLOG. The method was originally developed in our laboratory. It is a hybrid of a pharmacophore search and simplistic docking algorithms. Details of the FLOG algorithm are described elsewhere;[9] therefore, only the most important features of the method are described below.

Similar to the well-known program DOCK,[13] FLOG uses a set of match centers to represent the volume of the binding cavity and a clique-finding algorithm to generate trial orientations in the binding site.[9] Match centers are classified into several conventional pharmacophoric types, namely, cation, anion, polar, hydrophobic, hydrogen-bond donor and acceptor, and none of the above. It is often desirable to filter out compounds without required pharmacophoric properties or to stress the significance of certain functional groups in the binding area. This is done by changing the status of certain match centers to make them so-called "essential points"[9] that must be paired with a ligand atom. There are three main types of the ligand-atom-to-essential-point matches available in FLOG: (a) nonspecific match (the ligand atom closest to the match center makes a match regardless of its type), (b) match by chemical element type, and (c) match by pharmacophoric type. Essential points are used to guide the docking process in such a way that functionally important areas in an active site are better sampled and are filled by a chemically appropriate ligand atom.

The receptor structure is represented as a series of grids, each grid corresponding to the pharmacophoric types listed above. Each docked pose of a conformer is assigned a FLOG score (S_{raw}), which is a sum of the grid energy contributions

10.1021/ci050089y CCC: $30.25
Published on Web 06/24/2005

for all ligand atoms according to their pharmacophoric types plus extra contributions for essential point matches if any occur.

The flexibility of a ligand in FLOG is taken into account by precalculating a representative sample of conformers, typically 20−100, for every candidate ligand. Conformations are stored in dedicated databases called "flexibases".[14] Generally, the conformational space of a candidate ligand cannot be fully covered by such a representation, but one inevitably sacrifices coverage to achieve the required search speed. To further accelerate the search, only the first 25 most dissimilar conformers of every compound are typically examined. Our experience shows that, in most cases, this is sufficient for a reasonable quality of results. Rigid-body docking of a conformer of a ligand is performed by a simplex algorithm that optimizes the interaction energy of the rigid conformer against the grids. The final score of a ligand is the score of the highest-scoring orientation of all conformers of that ligand.

Corporate database searches routinely done in our laboratory show that the ability of FLOG to select active compounds can be dramatically improved if functionally significant areas of the binding pocket of a protein are specified by properly chosen essential points. We therefore routinely use essential points, and we include them in the protocol for this work. One essential point of type c (pharmacophore type match, see above) was added to every target receptor to highlight one very important spot in the binding area. In most cases, the selected essential point corresponded to the known important functional group responsible for key interactions with known ligands (e.g., catalytic carboxylate in HIV-1 protease). If such selection could not be made unambiguously, a match center was chosen closest to the most favorable point in the energy grid.

In addition to the "raw" score S_{raw}, a "normalized" FLOG score S_{norm} is also calculated. This is done to compensate for the bias that larger compounds tend to have higher scores. The normalized score $S_{norm} = (S_{raw} - 1.4)/N^{1/3}$, where N is the number of non-hydrogen atoms. (Parameters of this relationship were derived by a linear regression of the raw score values to the number of atoms for a random sample of 10 000 molecules with N in the range 20−100. The idea to use the cube root is from reference 19.)

Cumulative recall curves are a conventional way to evaluate the results of a database search if active compounds are known beforehand. A cumulative recall curve shows the total number of active compounds found as a function of the fraction of the database searched when the database compounds are ordered by decreasing score (the convention in FLOG is higher scores are better). At one extreme, if a scoring function is perfectly discriminating, so that all actives get higher scores than any nonactive compound, then one would find all of the actives at the front of the list. The slope of the curve would be 1 until all actives are tested; thereafter, the slope would be 0 ("ideal recall"). At the other extreme, if the scoring function is not discriminating at all, the curve would be a straight line with the slope equal to the ratio of the number of actives to the total number of compounds in the database ("random recall"). In this paper, we used both the raw and normalized scores. From our previous experience with FLOG we learned that, in most cases, the normalized FLOG score shows better performance than the raw score

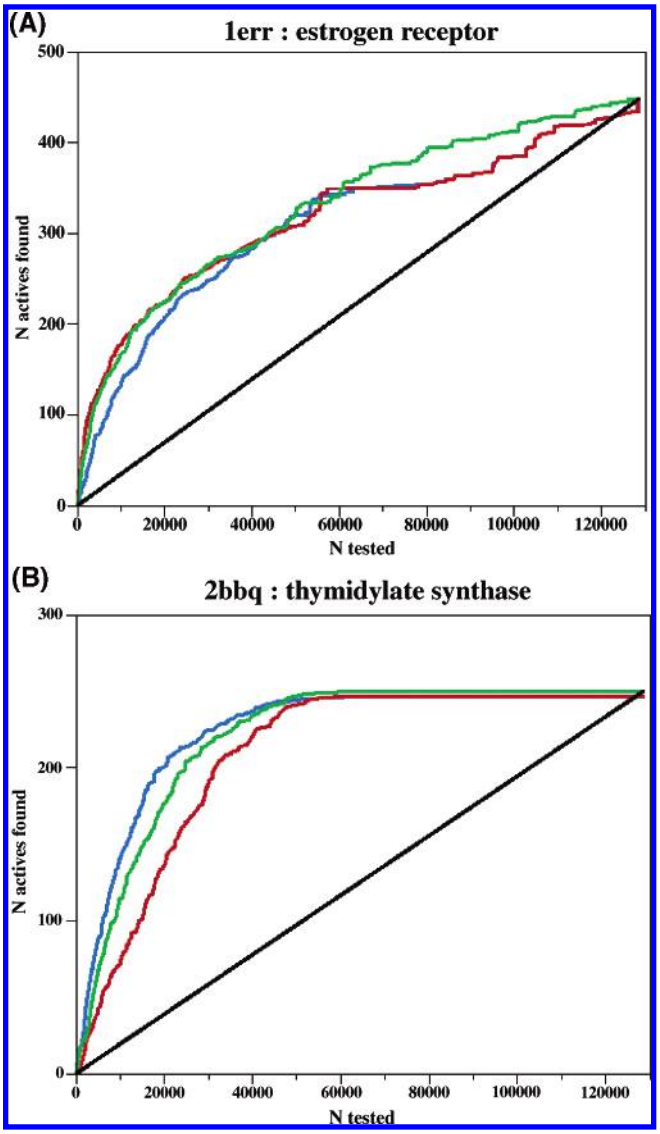

Figure 1. Cumulative recall curves: examples of different mutual positions of raw (blue) and normalized (red) FLOG scores (see text for details). The green line indicates the consensus score defined as the minimum rank of the ranks for the raw and normalized scores.[15] The enrichment level expected for "random recall" is shown as a black diagonal line; "ideal recall" is not shown.

(Figure 1A, red and blue lines, respectively). However, it is not always true, and there are cases where the raw score is better (Figure 1B). This is likely due to an interplay of the size of the actives relative to the size of an average compound in the database and to the volume and shape of the binding pocket. If actives are not known in advance, there is not a reliable way to predict which of the two scores should do a better job. Therefore, we combined both scores in a "consensus" score; that is, for each compound, we took the lowest rank of the two from the individually sorted lists.[15] Preliminary tests showed that this consensus score is a reasonable tradeoff (worse than the better of two scores, but better than the poorer score), resulting in a decent combined recall curve (Figure 1A and B, green lines).

In our computational environment, including IBM Power3/Power4 clusters, FLOG is very quick (dozens of compounds per second). Furthermore, parallelization allows us to run simultaneously up to 70 threads per user. Because FLOG is in-house software, there is no license issue. A search of the

Table 1. Test Set Used in the Work[a]

PDB code	protein	X-ray resolution Å	MDDR keyword	$N_{actives}$
1a7a	(*S*)-adenosylhomocysteine hydrolase	2.8	(*S*)-adenosyl-L-homocysteine hydrolase inhibitor	51
1acj	acetylcholinesterase	2.8	acetylcholinesterase inhibitor	748
1ah3	aldose reductase	2.3	aldose reductase inhibitor	929
1aq1	cyclin-dependent kinase 2	2.0	kinase C inhibitor	487
1azm	carbonic anhydrase I	2.0	carbonic anhydrase inhibitor	278
1b8o	purine nucleoside phosphorylase	1.5	purine nucleoside phosphorylase inhibitor	70
1c2t	GAR transformylase	2.1	glycinamide ribonucleotide formyltransferase inhibitor	50
1dwc	thrombin	3.0	thrombin inhibitor	1116
1efy	poly(ADP-ribose) polymerase	2.2	poly(ADP-ribose)synthethase inhibitor	158
1ep4	HIV-1 reverse transcriptase	2.5	reverse transcriptase inhibitor	634
1err	estrogen receptor	2.6	antiestrogen	443
1kr3	beta-lactamase type II	2.5	lactamase (beta) inhibitor	185
1pge	prostaglandin H2 synthase-1	3.5	cyclooxygenase-2 inhibitor	1473
1poc	phospholipase A2	2.0	phospholipase A2 inhibitor	853
1sln	fibroblast stromelysin-1	2.27	matrix metalloproteinase inhibitor	557
1tpp	trypsin	1.4	trypsin inhibitor	115
2bbq	thymidylate synthase	2.3	thymidylate synthetase inhibitor	247
3dfr	dihydrofolate reductase	1.7	dihydrofolate reductase inhibitor	172
4est	elastase	1.78	elastase inhibitor	634
9hvp	HIV-1 protease	2.8	HIV-1 protease inhibitor	890

[a] Protein structures were taken from the Protein Data Bank[18]. MDDR keywords were used to identify active database compounds by the associated corresponding activity type.

whole corporate database (more than a million compounds) typically takes several hours of elapsed time, or overnight in the worst cases.

ICM. We use the ICM-Dock module of the ICM-Pro 3.0 program.[10−12,16] During ICM docking, all rotatable bonds of a ligand and six positional degrees of freedom of a ligand are varied, a series of low energy conformations is generated, and a Monte Carlo optimization of a combined scoring function including the internal energy of the ligand and its interaction with the grid representation of the receptor is performed in torsion angle space. The best generated pose of each ligand is evaluated by a special ICM scoring function (which differs from the function evaluated during the docking process[12]). The scoring function includes a weighted sum of the internal energy of the ligand, grid-based van der Waals and hydrogen bond interactions, a hydrophobicity term calculated as solvent accessible surface buried upon binding, an electrostatic solvation term, and an entropy term proportional to the number of flexed torsion angles.

In preparation for ICM docking, ionizable groups in the protein structures were converted into the protonated states appropriate at neutral pH, and the ICM default partial atomic charges were set up. The active site for a protein was defined as being within 5 Å of a ligand in the PDB X-ray co-crystal structure. Energy grids representing the active site (van der Waals, hydrogen bonding, electrostatics, and hydrophobic interactions) were calculated with 0.5 Å grid spacing. The default "rule-of-five"[17] filter was disabled.

The docking of ligands with too many flexible torsion angles may be prohibitively expensive in ICM. Therefore, for the sake of speed, an empirically adjusted maximum of 14 rotatable torsion angles was used. Compounds with more than 14 torsions were not processed but were assigned an arbitrary poor ICM score and kept in the subsequent analysis. All other parameters were used at their default values. Compounds were ranked by the ICM docking score function.[12] More details regarding the ICM docking methodology can be found elsewhere.[10−12,16] All ICM docking calculations were performed on SGI 500 MHz Fuel machines with two ICM licenses.

Test Dataset. Table 1 lists 20 co-crystal protein structures taken from the Protein Data Bank.[18] These protein structures were used in earlier in-house docking experiments and were selected because receptor preparation had already been done. The proteins varied substantially in X-ray resolution (1.35−3.5 Å) and in the number of actives (51−1437) (Table 1). It should be emphasized that this set of proteins is not intended to be a representative selection of any particular protein structural classes, biological activities, or binding pocket types but should be considered as an ad hoc arbitrary sample of targets to illustrate the proposed methodology.

The MDL Drug Data Report (MDDR) database (release 2002.2, ca. 129 000 compounds) was used as a source of compounds to screen. Conformers of the compounds were available to us as an in-house precalculated database, so the only remaining step was to compile a list of active compounds for each protein target. This was done by analysis of the MDDR activity keywords and extracting those most appropriate for a given protein. Most often, a combination of the receptor name and the word "inhibitor" was enough. The number of such identified actives is shown in Table 1.

Protocol and Enrichment Ratios. The workflow of the proposed protocol is shown in Figure 2. At the first step ("search"), FLOG docking was performed against the whole MDDR database, the 1000 best FLOG hits (as defined by consensus score) were collected, and the first step enrichment ratio $P_{FLOG/random}$ (see below) was calculated. If the enrichment ratio was less than 2.0, then the target was discarded and nothing further was done with it. Preliminary studies showed that if the ratio was less than 2.0, the set of 1000 highest ranked compounds contained too few actives to make further processing meaningful (data not shown).

If a target survived the first step, then the second step ("refinement") was undertaken: the 1000 best FLOG hits were submitted for ICM docking. After the second step was complete, the 100 best ICM compounds (as defined by the

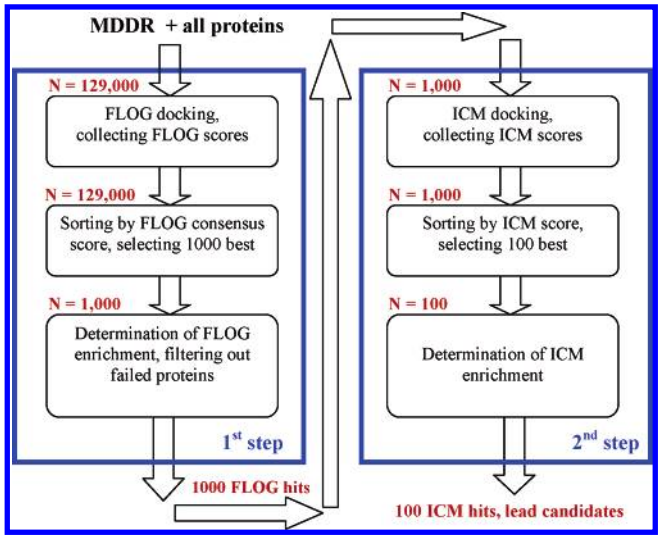

Figure 2. Schematic diagram of a combined protocol including FLOG docking of the whole MDDR (ca. 129 000 compounds) ("search step") followed by ICM postprocessing ("refinement step") of the 1000 best FLOG hits. Refinement is undertaken only if the FLOG-to-random enrichment ratio is not less than 2.

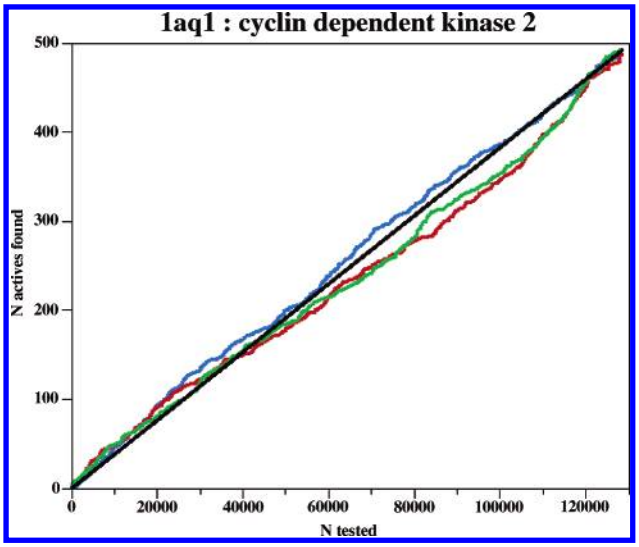

Figure 3. Example demonstrating poor enrichment during the FLOG search. Cumulative recall curves describing FLOG search results against the whole MDDR database for CDK2 (PDB code 1aq2) (compare to Figure 1).

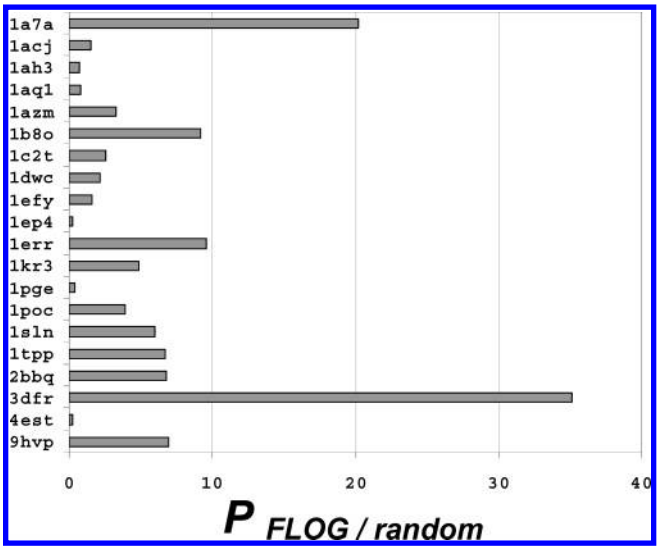

Figure 4. FLOG-to-random enrichment ratio, $P_{FLOG/random}$, defined for the 1000 best FLOG hits for 20 proteins from the test set.

ICM scoring function) were gathered and the second step enrichment ratio $P_{ICM/FLOG}$ (see below) was calculated. The choice of 1000 for the FLOG sample size and 100 for the ICM sample size were mainly due to practical considerations, namely, (a) how many compounds can be relatively rapidly processed by ICM (1000) and (b) how many compounds can typically be forwarded for experimental testing (100).

We introduce three measures of the distribution of active compounds among nonactives in our protocol. Each of them can be specified by the average concentration of the actives in a "sample": (1) C_{random}, total number of the actives in the database divided by the total number of MDDR entries; (2) C_{FLOG}, number of actives in the best 1000 FLOG hits divided by 1000; and (3) C_{ICM}, the number of actives in the best 100 ICM hits divided by 100. In turn, these concentrations can be related to each other to characterize the enrichment of the actives across the steps of the protocol, namely, FLOG enrichment, $P_{FLOG/random} = C_{FLOG}/C_{random}$, and ICM enrichment, $P_{ICM/FLOG} = C_{ICM}/C_{FLOG}$.

RESULTS AND DISCUSSION

The described protocol was applied to 20 targets in the test set. After the first step of the protocol, we found that there were many reasonable cumulative recall curves (e.g., 2bbq in Figure 1B) and some accumulation curves no better than random (e.g., 1aq1 in Figure 3). The average FLOG-to-random enrichment over all 20 processed proteins was found to be 6.4. Data on enrichment at this step are shown in Figure 4. Twelve proteins (60%) showed the enrichment to be 2 or higher and were forwarded to the "refinement" step. We could not find any obvious difference between the group of proteins that failed (1acj, 1ah3, 1aq1, 1c2t, 1efy, 1ep4, 1pge, 4est) and the group that succeeded. Both groups appear to be similar in respect to X-ray resolution, number of MDDR actives, and biological functions (Table 1). As of now, it seems impossible to predict what protein may fail, and further investigation of the issue is required.

ICM docking was performed on the 1000 best-scored FLOG compounds for all of the 12 proteins passing the first step. Figure 5A shows an example of an excellent ICM-to-FLOG enrichment manifested in the case of estrogen receptor (1err), the best among all of the cases processed during the refinement step.The first 100 ICM hits contain substantially more actives compared to the FLOG results. On the other hand, in several cases, we observed fairly poor enhancements of ICM docking over FLOG, such as in the case of HIV-1 protease (Figure 5B). Perhaps, this can at least partially be attributed to good work done by FLOG during the search step, so ICM refinement simply could not improve the enrichment any further. Averaged over 12 proteins, the ICM-to-FLOG enrichment ratio was found to be 4.7. Data on enrichment at the refinement step are shown in Figure 6.

Because of an upper limit imposed on the maximal number of rotatable bonds during ICM refinement, some actives for several proteins were assigned an arbitrarily poor score (see Methods) and, therefore, were effectively excluded from the final list of 100 best ICM hits. Naturally, this affects the distribution of the scores and, eventually, the evaluation of the performance of the protocol. We argue that this effect

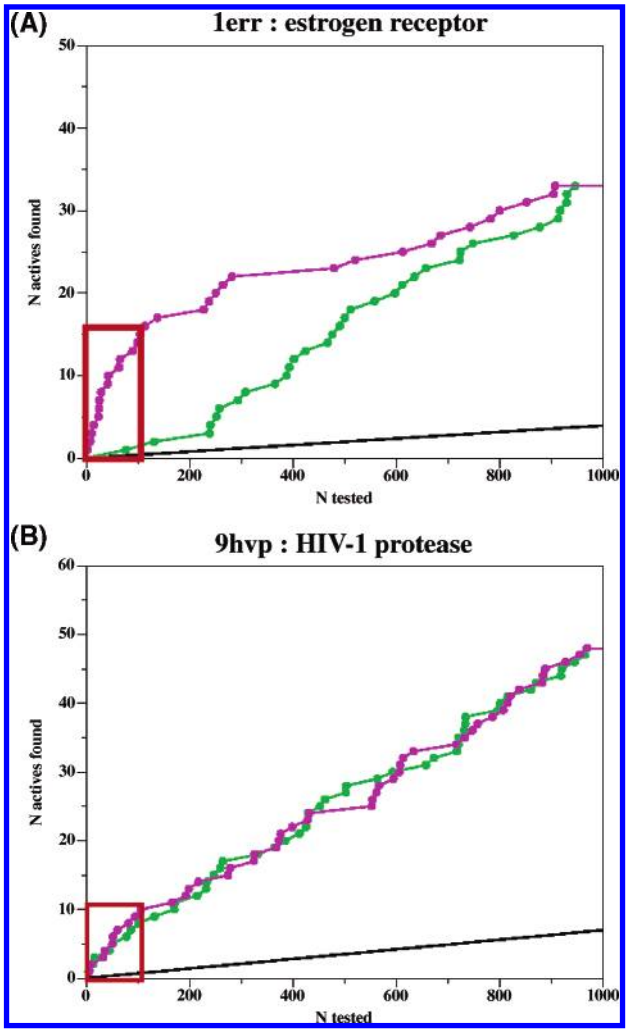

Figure 5. Examples of a comparison of cumulative recall curves of FLOG consensus (green) and ICM docking (magenta) scores. The 1000 best FLOG hits were submitted for ICM refinement. The focus of the comparison, highlighted by red rectangles, is on the 100 best ICM hits composing a hypothetical sample for the experimental testing of biological activity. (A) Successful ICM refinement (estrogen receptor antagonist, PDB structure 1err). (B) Failed ICM refinement (for HIV-1 protease inhibitor, 9hvp). The enrichment level expected for a random selection is indicated by a black line.

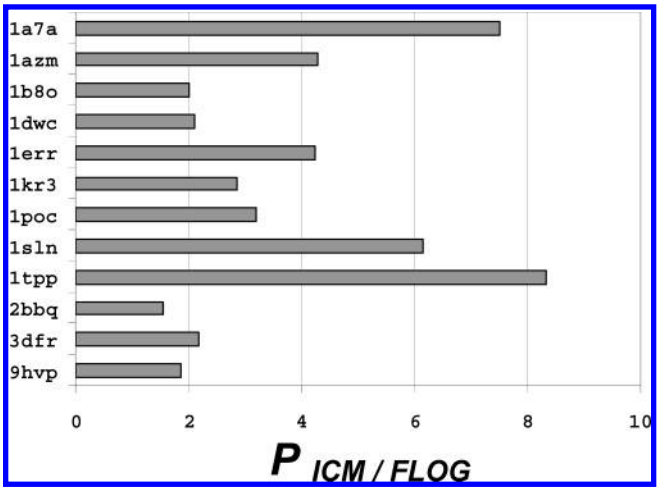

Figure 6. ICM-to-FLOG enrichment ratio, $P_{ICM/FLOG}$, for 12 proteins for which ICM docking ("refinement") was performed.

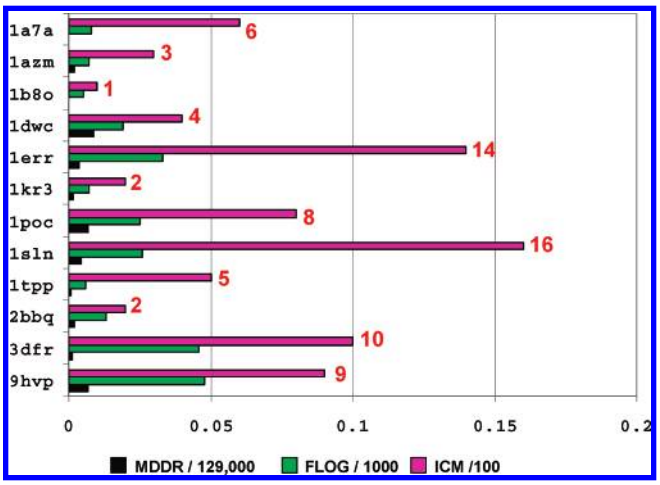

Figure 7. Comparison of the effective concentration of actives C_{random} (black bars), C_{FLOG} (green), and C_{ICM} (magenta) across the steps of the protocol. The numbers of actives found in the output 100-compound sample after the refinement step are indicated. The C_{random} value for (*S*)-adenosylhomocysteine hydrolase (PDB code 1a7a) is too small (~0.0004) to appear on the graph with the given scale.

seems not to be significant from a viewpoint of general assessment. First, such a change in the number of actives was relatively small: only in 4 out of 12 cases did we observe more than a 25% decrease in the effective number of actives, namely, 9hvp (50%, 24 actives filtered out of 48), 2bbq (46%, 6 of 13), 1dwc (37%, 7 of 19), and 1err (27%, 9 of 33). For all of the remaining cases, the difference was less than 12%. Second, ligands having many rotatable bonds have a large conformational space that is difficult to sample. Therefore, any docking program would have difficulty docking very flexible ligands. It seems reasonable to filter out large/flexible compounds (even sacrificing a certain fraction of flexible actives) and to focus on more promising, smaller molecules.

Figure 7 shows a combined plot of the concentration of actives in the whole MDDR database, 1000 best FLOG hits, and ICM best 100 hits. If the proposed protocol was applied in a real case, then with a probability of 60% we would expect to have at least several actives in the final 100 compound sample. (60% means that the same success-to-failure rate is assumed, viz., 12 out of 20 targets survived after the first step.)

An interesting question concerns how many actives one is missing simply because only a small fraction of the database is processed with a full-force docking program like ICM. However, in our setting (two ICM licenses), we could not address this question for exactly the same reason that we need to combine two different methods: a sufficient number of the commercial licenses is not available to search the entire database with ICM within reasonable time. Thus, we cannot claim that our two-step procedure gives results "almost as good as ICM", only that the results are much better than FLOG alone.

It is worthwhile to emphasize that both FLOG and ICM were used in an almost "out-of-the-box" mode. No special tuning was done in either case, except to add one essential point in the case of FLOG, switching off the "rule-of-five" filter and excluding compounds with more than 14 flexed torsion angles in the case of ICM. It looks very likely that the performance of the combined protocol can be improved if both approaches are tuned to their best level of performance for a specific target.

Our two-step approach of "docking" and "redocking" is somewhat new. However, there are many examples in the literature of "docking" followed by "rescoring and reranking the poses". The advantages of rescoring were demonstrated, for example, in a small-scale docking experiment on 10 targets with their native ligand where molecular-mechanics optimization of the complexes was used to rerank the original docked poses from FlexX.[20] We tried to evaluate if rescoring can give reasonable results compared to redocking. Poses of the 1000 best FLOG compounds for every one of the 12 protein targets considered at step 2 of the protocol were submitted "as is" to a rescoring protocol. Seven scoring functions were used: XSCORE,[21] LigScore1, LigScore2, LUDI, PLP1, PLP2, and PMF (the last six implemented in Cerius^2).[22] Unfortunately, the current version of ICM does not allow rescoring of existing poses. A rescoring quality index I_{SF} was defined as

$$I_{SF} = \left(\sum_{M} \frac{N_{act}(SF, 100)}{N_{act}(FLOG, 1000)} \right) / M$$

where N_{act}(SF, 100) is the number of actives in the best 100 hits obtained by rescoring with the scoring function SF the 1000 best FLOG hits, N_{act}(FLOG, 1000) is the total number of actives in the 1000 best FLOG hits, and M (equal to 12) is the total number of proteins in question. Although the quality index was 11% for the original FLOG score, it varied from 2% (XSCORE) to 24% (LigScore2 and LUDI) for rescoring. This is compared to 42% for ICM redocking. This is rather approximate, but we believe that it confirms the point that redocking is expected to do a better job than rescoring existing poses.

The described combination of FLOG and ICM is certainly not unique, and other similar docking methods can, in principle, be plugged into the protocol in a similar fashion. Instead of FLOG at the first step, one could use any docking/pharmacophore search program (giving reasonable initial enrichment) that is fast or cheap/free (allowing to run many jobs in parallel), for example, DOCK.[13] Likewise, any other good modern docking program can be tried for the refinement step.[1−3]

Wall clock timing of a typical processing of a protein target was as follows. Processing of the whole MDDR during the search step took several hours on an IBM "Power" cluster. At the same time, ICM processing of the best 1000 FLOG hits took about 20 to 30 h using two SGI 500 MHz "Fuel" machines. This second step was apparently the bottleneck of the whole workflow. Besides using CPUs with a higher clock speed, the timing can likely be improved by prefiltering database compounds, for example, by applying the "rule-of-five" filter[17] during the first step. Also, one could accelerate the ICM docking by lowering the maximal allowed number of flexible torsions. Even without the proposed tuning, the approximate turnaround time per protein was close to 24 h, and at least several actives were found for more than half of the examined proteins in a very small (100) set of candidates. We seem to have achieved a plausible level of selection of actives in a short amount of time.

CONCLUSION

The cost of commercial software licenses imposes serious limitations on the use of robust modern docking programs for VLS of large databases. A combination of two methods, "fast−free−approximate" FLOG (in-house software, a hybrid of a pharmacophoric search and simplistic docking algorithm) and "slow−costly−accurate" ICM-Dock (commercial docking program), is proposed as a solution to the problem. A two-step protocol selects the most likely lead candidates in the first ("search") step followed by a careful accurate docking of a limited number of compounds in the second ("refinement") step. In more than half of the examined 20 cases, at least several actives per protein target were obtained in approximately 24 h per protein turnaround time with only two commercial licenses. The results, an almost 5-fold enrichment relative to the yield from the first docking program, seem to validate the proposed protocol.

ACKNOWLEDGMENT

We thank Drs. W. Cornell, E. Gancia, and B. McKeever (Merck Research Laboratories) for reading the manuscript, our colleagues from Molecular Systems for numerous helpful discussions, and Drs. G. McGaughey (Merck Research Laboratories) and N. Nevins (GSK) for providing the opportunity to present the method in a session of the 228th ACS meeting's "Docking and Scoring" section. We are thankful to the reviewers for their helpful suggestions.

REFERENCES AND NOTES

(1) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(2) Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. Lead discovery using molecular docking. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439−446.
(3) Schneider, G.; Bohm, H. J. Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002**, *7*, 64−70.
(4) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(5) Lyne, P. D. Structure-based virtual screening: An overview. *Drug Discovery Today* **2002**, *7*, 1047−1055.
(6) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 234−249.
(7) Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L., III. Comparative study of several algorithms for flexible ligand docking. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 755−763.
(8) Abagyan, R.; Totrov, M. High-throughput docking for lead generation. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375−382.
(9) Miller, M. D.; Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P. FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 153−174.
(10) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM−a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, *15*, 488−506.
(11) Abagyan, R.; Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **1994**, *235*, 983−1002.
(12) Totrov, M.; Abagyan, R. Flexible protein−ligand docking by global energy optimization in internal coordinates. *Proteins* **1997**, *1* (Suppl.), 215−220.
(13) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269−288. Shoichet, B. K.; Kuntz, I. D. Matching chemistry and shape in molecular docking. *Protein Eng.* **1993**, *6*, 723−732.
(14) Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P.; Miller, M. D. Flexibases: A way to enhance the use of molecule docking methods. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 565−582.
(15) Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Modell.* **2002**, *20*, 281−295.
(16) *ICM-Pro*, version 3.0; Molsoft, LLC: LaJolla, CA.

(17) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.
(18) Berman, H. M.; Westbrook, J.; Feng, J.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(19) Pan, Y.; Huang, N.; Cho, S.; MacKerell, A. D. J. Consideration of molecular weight during compound selection in virtual target-based database screening *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 267−272.
(20) Hoffmann, D.; Kramer, B.; Washio, T.; Steinmetzer, T.; Rarey, M.; Lengauer, T. Two-stage method for protein−ligand docking. *J. Med. Chem.* **1999**, *42*, 4422−4433.
(21) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11−26.
(22) *Cerius*^{2}, version 4.6; Accelrys Inc.: San Diego, CA. http://www.accelrys.com.

CI050089Y