Journal of Computer-Aided Molecular Design, **13:** 513–532, 1999.
KLUWER/ESCOM

DREAM++: Flexible docking program for virtual combinatorial libraries

Shingo Makino[a,b], Todd J.A. Ewing[a] & Irwin D. Kuntz[a,*]
[a]*Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143–0446, U.S.A.*
[b] *Central Research Laboratory, Ajinomoto Co. Inc., 1–1 Suzuki-cho, Kawasaki-ku, Kawasaki-shi 210, Japan*

Received 4 August 1997; Accepted 14 December 1998

Key words: backtrack algorithm, conformation, de novo, diversity, incremental construction algorithm

Summary

We present a set of programs, DREAM++ (Docking and Reaction programs using Efficient seArch Methods written in C++), for docking computationally generated ligands into macromolecular binding sites. DREAM++ is composed of three programs: ORIENT++, REACT++ and SEARCH++. The program ORIENT++ positions molecules in a binding site with the DOCK algorithm [1, 2]. Its output can be used as input to REACT++ and SEARCH++. The program REACT++ performs user-specified chemical reactions on a docked molecule, so that reaction products can be evaluated for three dimensional complementarity with the macromolecular site. The program SEARCH++ performs an efficient conformation search on the reaction products using a hybrid backtrack [3, 4] and incremental construction [5, 6] algorithm. We have applied the programs to HIV protease–inhibitor complexes as test systems. We found that we can differentiate high-affinity ligands based on several measures: interaction energies, occupancy of protein subsites and the number of successfully docked conformations for each product. Encouraged by the results in the test case, we applied the programs to propose novel inhibitors of HIV protease. These inhibitors can be generated by organic reactions using commercially available reagents. They are alternatives to the inhibitors synthesized by Glaxo [7, 8].

Introduction

Combinatorial chemistry [9] is one of the most promising new techniques in medicinal chemistry. It enables the synthesis of a large number of molecules in a short period of time. However, the number of synthetically accessible molecules can be enormous. Under these circumstances, it may be advantageous to use computational techniques to assist in the design of the library. Methods have been proposed to increase library diversity [10, 11], and to design libraries for structure-based targets [12–14].

Previous efforts have largely focused on assembling molecular fragments in predefined ways, primarily using the computer to construct three-dimensional (3D) coordinates for the final molecules. Here, we explore the concept that computer programs can make use of explicit chemical reactions to design libraries for combinatorial chemistry according to receptor site information. Previous programs [13, 14] addressed combinatorial library design for the case in which the ligands can be treated as variable fragments attached at multiple sites onto a core scaffold. The fragments are evaluated in such a way that the search time grows linearly rather than exponentially. In this paper, we address library design in which variable fragments are joined consecutively. Such ligands often do not have a core scaffold.

Our program incorporates techniques from de novo ligand design and flexible docking. The existing methods for de novo ligand design are: (1) growing one atom at a time (Legend [15], Genstar [16], MCDNLG [17], CONCEPTS [18]); (2) using single atoms and fragments to grow molecules (GrowMol [19]); (3) using organic fragments to grow or to connect molecules (GroupBuild [20], SPROUT [21, 22], GROW [23], LUDI [24, 25], BUILDER [26], SMoG [27],

*To whom correspondence should be addressed.

CONCERTS [28]). A general problem with these approaches is the evaluation of *synthetic accessibility*: molecules are built up without regard for the principles of organic chemistry. The programs GROW [23] and one mode of LUDI [29] use amide bond formation to produce peptide or peptide mimetic molecules that are expected to be synthesized by standard chemical procedures. However, these programs currently only support this one reaction. To extend this idea to include a much more diverse set of compounds, we decided to generate molecules using well-studied organic reactions, especially those that can easily be applied to combinatorial chemistry on solid support. This approach is particularly attractive if one can build putative inhibitors in the 3D context of a receptor site because it is possible to discard a large number of molecules based on the properties of the intermediate products. This protocol helps to avoid the large size of virtual combinatorial libraries. To model the 3D properties of combinatorial molecules, we must consider the conformational flexibility. There are two approaches to address flexibility: (a) using multiple conformations generated with pre-calculation for each compound; and (b) run-time generation of conformations as needed during a search. The precalculation of conformations is employed by the programs LUDI, GROW and FLEXIBASES/FLOG [30, 31]. The merit of this method is that generation of conformations only has to be done, once avoiding time-consuming run-time generation of conformations. However, it is difficult to store enough conformations for flexible docking. In one example, 5000 conformations for nine flexible bonds were not sufficient to regenerate the X-ray structure [23]. Additionally, it is difficult to prune conformations efficiently based on partial structure energies. Therefore, we are interested in the run-time generation of conformations. The approaches to be considered include: (1) docking fragments independently and fusing subsequently [32]; (2) simulated annealing (AutoDock [33, 34]); (3) distance geometry [35]; (4) a complete combinatorial search (ADAM & EVE [36]); (5) genetic algorithm [37, 38]; (6) Monte Carlo search [39]; (7) the backtrack algorithm [3, 4]; and (8) the incremental construction algorithm (FLEXX [5], hammerhead [6]). Among these, the most attractive approaches for us are backtrack (depth-first) and incremental construction (breadth-first) algorithms. Both methods begin by selecting and docking an anchor placement; then the conformations of the rest of the fragments are incrementally searched either by the depth- or the breadth-first algorithm. Although the backtrack method can explore many conformations very rapidly, the number of conformations to be searched grows rapidly as the amount of flexibility increases. The incremental construction algorithm can avoid this explosion of conformational space, restricting the number of conformations in each step. Although other incremental construction search programs consider one flexible bond at a time, we found that our implementation did not regenerate X-ray complex structures accurately, as we will discuss later. Therefore, we developed a hybrid algorithm, combining incremental construction and the backtrack algorithm, taking advantage of both methods. In our search algorithm, ligands are divided into fragments where each fragment has several flexible bonds. Conformations of each fragment are searched by the backtrack algorithm; then the number of sampled conformations are reduced, based on the scores and conformational diversity, and used for the next backtrack search. Another new feature is the inheritance of conformations. After searching for the conformation of an anchor fragment, the multiple conformations of the anchor are inherited by product compounds generated by the chemical reactions. Thus, the conformational search for the anchor fragment is done only once even though this anchor fragment appears many times in product compounds. In this paper, we describe the basic algorithms and test this method using known HIV protease–inhibitor systems. Finally, the program is applied to generate putative new inhibitors.

General methods

In general, molecules are constructed in a binding site by the following procedure. After positioning the anchor parts of a first set of molecules into the binding site using a DOCK algorithm (ORIENT++), the conformations of these docked molecules are searched and minimized. A set of molecules are selected according to certain criteria (scores, binding modes, etc.). Then, these selected molecules (reactants) are virtually reacted with a set of reagents to produce products in the binding sites. The conformations of the additional part of each product are searched for the next selection of partial molecules. These reactions, conformational searches and selections of molecules are repeated until producing entire molecules designed by users. We describe these steps in detail in the following order: scoring; conformational search and minimization; and chemical reactions algorithms.

Scoring

We have adopted an AMBER-type potential function [40, 41] to rank the conformations of ligands. The intermolecular interaction energy is calculated using the grid-based force field [42] and the receptors are assumed to be rigid. The intramolecular energy calculation is done somewhat differently from the usual way, in which all the energy between atoms with more than 1–3 relationships is calculated. Since the conformational freedom of ligands is only determined by torsion angles, the interatom energy within each rigid fragment is constant. It is wasteful of computational time to calculate such constant energy repeatedly to compare the energy of the different conformations within each molecule. Therefore, the calculation of the interatom energy within a rigid fragment is avoided (Figure 1), reducing computation time by 15%.

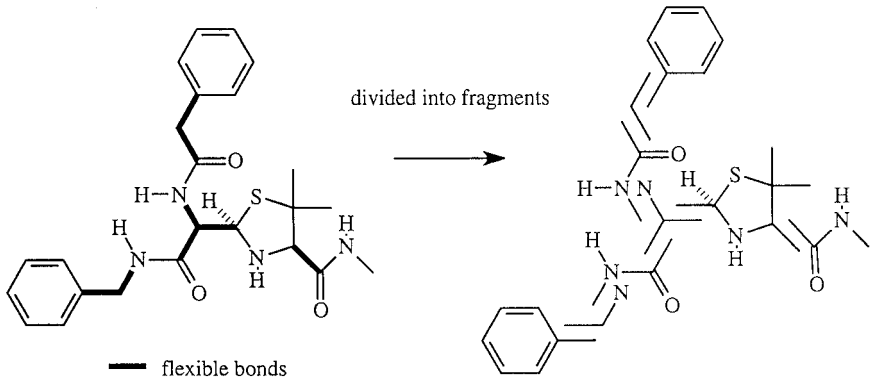
Additionally, we have adopted the approximation introduced by Glen and co-workers [43] to calculate the energy between atoms only if the distance of two atoms is less than the sum of the radius of the two atoms. This approximation makes the calculation about 3 times faster, and the quality of the results is about the same in our experience.

Hybrid conformational search and minimization

We have developed the hybrid conformational search method, which is a compromise between the depth-first algorithm (backtrack) and the breadth-first (incremental) algorithm, as the following. We start the conformational search at the *i*th flexible bond, and the program scans to the $i + c$ flexible bond, where c is a constant number specified by the user. The scan stops if the program finds either a fragment which contains five or more heavy atoms (usually a ring system) or the end of the molecule. Flexible bond and ring detection are done as described previously [44]. Then the backtrack search is performed from i to the terminating fragment. If no terminating condition is found, only the *i*th flexible bond is searched. The results presented in this paper are obtained by setting c equal to 3, so each fragment has up to four flexible bonds. Before searching the next conformational iteration, the number of current conformations is reduced based on scores. In addition, the current round of conformations is clustered according to pairwise root mean square deviation (rmsd) of different conformations.

The clustering algorithm is shown in Figure 2. Numbers 1–7 represent the conformations to be clustered and are sorted according to the score of each molecule (Figure 2a). The conformation with the best score is selected; then all the other conformations which are within the rmsd limit from the selected conformation are placed in the first cluster (Figure 2c). In this case, compound 5 was found to be similar to 1 and put into cluster 1 together with compound 1. Then, the next best scoring conformation which does not belong to any cluster is chosen (Figure 2d), and the same cycle is repeated until all the conformations are clustered. The number of clusters is controlled by the rmsd criterion as follows. Starting from a large value for rmsd (iteration 1), the distances are reduced gradually until the number of clusters approaches a preset value. The elements of the distance matrix have to be calculated once through all the iterations. Since the most time-consuming part in this clustering is the calculation of rmsd, we can significantly reduce computation time by avoiding unnecessary comparisons (the elements marked by the star symbol). If three clusters are appropriate, the calculations for 10 out of 21 elements would be avoided (Figure 2). The calculation of the rmsd is done not by atom index but by all the possible atom alignments based on the symmetries of ligands to avoid placing similar conformations in the different clusters. Such symmetries are automatically identified by SEARCH++. Moreover, the flexible bond symmetries are identified according to atom type, connectivity, relative position of atoms and chirality to avoid sampling the same conformations repeatedly. To optimize each local conformation, the simplex method [45] is used. During minimization, only the bonds currently being searched are included, fixing orientations and flexible bonds previously searched.

We employed this hybrid conformational search method for the following reasons. The advantage of the backtrack method is that it can explore many conformations rapidly, avoiding recalculation of the rotation matrix and vectors. However, the number of conformations increases exponentially, which is prohibitive for highly flexible molecules. The advantage of the incremental method is that it is easy to control the conformational space or the computation time by limiting the number of conformations which become the seeds for the next level of conformational exploration. Although it is fast, in our experience it can fail to regenerate the binding modes for fragments which must penetrate narrow pockets in the receptor site. Thus, it is better to take advantage of both search methods.



1) The intramolecular energy between the atoms in each fragment are not calculated.

2) The intramolecular energies between the different fragments are calculated if the atoms in the unfragmented molecule are connected via more than 3 bonds.

Figure 1. Special intramolecular energy calculation method to reduce the amount of computation time, where the calculation of intramolecular energy in each rigid fragment is avoided.

Reactions as a tool to generate molecules

The programs REACT++ and SEARCH++ work in concert as a de novo generation program. To ensure that the resulting molecules are synthetically accessible, the molecular assembly protocol is restricted to well-characterized organic reactions. In contrast to existing de novo design programs, we incorporate several reaction types in addition to amide bond formation. We have selected reactions that may be utilized by combinatorial chemistry on solid support [46], to improve the ease of synthesis of computer-generated compounds. The reactions implemented in the program REACT++ are the following: (1) amide bond formation (including sulfonamide); (2) urea formation (isocyanate or oxycarbonation + amination); (3) reductive amination; (4) alkylation (N- and O-alkylations); (5) ester formation (condensation); and (6) Mitsunobu-type ester and ether formation (alkyl phosphite, phenol, carboxylic acid). Functional groups for reactions are assigned by other programs in advance; then such information is written in mol2 formatted files using the 'ATOM NAME' field. The special feature of the program REACT++ is its ability to deal with conformational flexibility as described in the following section.

Inheritance of conformations through reactions

The compounds generated by the systematic combination of fragments have many partial structures in common. It is a waste of computer time to dock or search these common fragments repeatedly. The program REACT++ preserves the information about these partial conformations and avoids any redundant conformation search. The products generated by the program REACT++ have inherited partial multiple conformations after a reaction, so only the conformations of the newly added fragments are searched. Since structures and conformations of molecules are stored in different files, these inheritances of conformations through virtual reactions can be performed simply by passing file names and pointer addresses of conformational information to their products. Thus, the memory requirements for these inheritances of conformations are negligible.

Calculation of the partial volume of a ligand inside a receptor

To help discriminate between productive and nonproductive docking, we have developed a new way to calculate the volume of the part of a ligand that is inside a receptor. A cubic grid is established in the binding pocket. The center of a dodecahedron [47] is put at each grid point. Vectors from the center and passing through the corners of the dodecahedron are checked for intersection with the vdW surface of a receptor (Figure 3a). The number of vectors which make contact with a receptor surface are counted and stored on each grid point; thus, the maximum number on each grid point is 20. 'Buried' ligand volume is estimated by counting the number of cubes centered on the grid points that overlap a particular ligand conformation and for which 75% of the dodecahedron

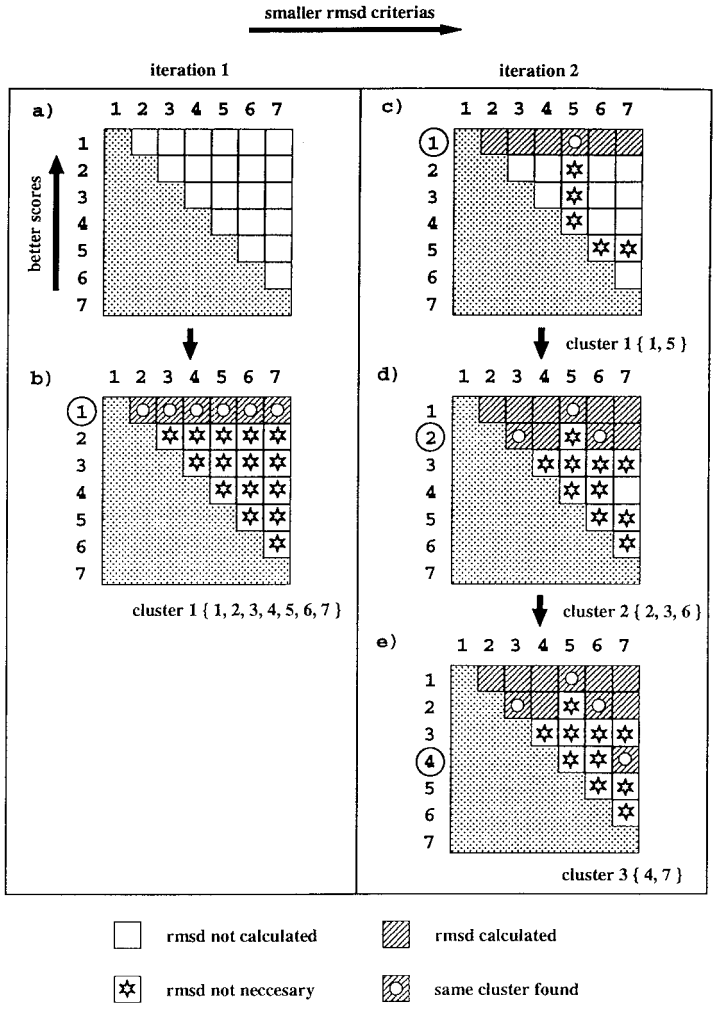


Figure 2. Method for clustering conformations in DREAM++, where unnecessary rmsd calculations are avoided to speed up calculations. Iteration 1: The selection of the large rmsd criterion produces only one cluster. The star mark indicates that these conformations are already clustered and the rmsd calculations for these elements are not necessary. Iteration 2: The rmsd criterion is changed to produce the proper size of clusters.

vectors intersect the surface of a receptor (Figure 3b). This volume could be roughly related to the number of relatively fixed water molecules released by the binding of the ligand and might be proportional to the entropical factor related to the releasing of water molecules. This quantity differs from the buried contact surface (Figure 3c), which does not distinguish between the outside and inside of the receptor pocket.

Programming and resource usage

This program is written in C++. The Standard Template Library (STL), which provides generic container classes (e.g. list, queue, binary trees), is especially useful to provide a simple interface for complicated data structures. All the programs are compiled by GNU gcc-2.7.2. All calculations were performed on Silicon Graphics Indigo2 workstations with 250 MHz R4400 processors and 128 Mb RAM.

Results

All the molecular modelings were performed with the program SYBYL [48] from Tripos using the

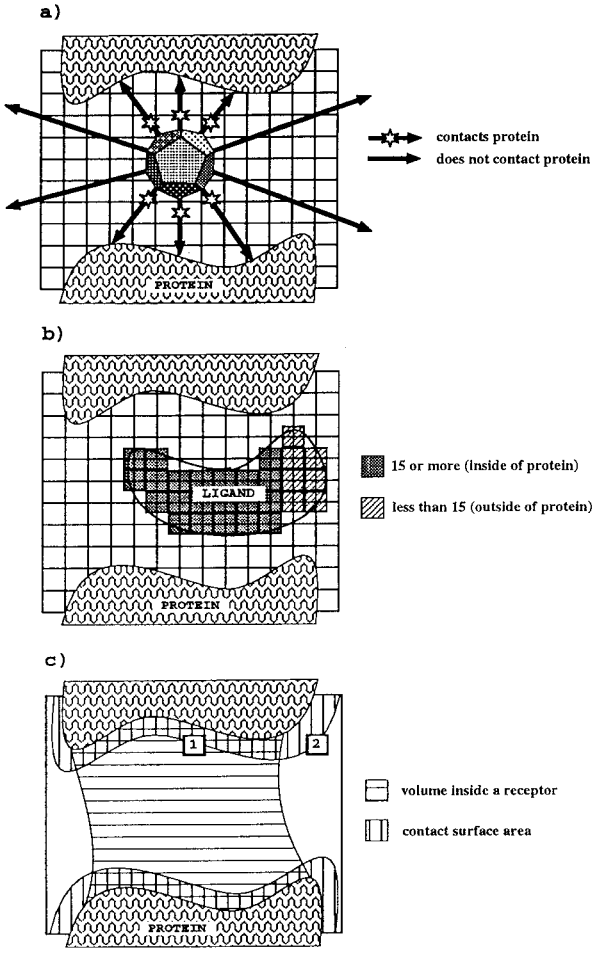


Figure 3. Method to calculate the partial volume of a ligand inside a receptor: (a) definition of the depth of the space inside a receptor; (b) volume of a ligand inside a receptor; (c) difference between the volume inside a receptor and the contact surface area.

Gasteiger–Marsili charges [49, 50]. We use a modified mol2 file format with conformation and reaction information for the input and output of molecules. This file format can be read by the program SYBYL. For the molecules in our test set, all starting conformations were obtained using SYBYL, except that the ring conformation of thiazolidine was taken from an X-ray complex structure (1htg). The energy grid spacing was set as 0.3 Å. Molecules were divided into fragments with a maximum of four flexible bonds. The search angles were 15.0° for the first fragments and 20° for the additional fragments. If the inter- and intramolecular energy from a partial fragment exceeded absolute 10 kcal/mol regardless of the fragment, the conformation was pruned. The top scoring 500 conformations were kept. Then, the number of conformations was reduced to ∼100 according to diversity of conformations.

Test system 1: Comparison of programs GROW and DREAM++

In this section, we compare the two conformational search techniques (precalculated search versus runtime search). The program GROW, which uses precalculated conformations, failed to produce any bound conformations of a reduced peptide inhibitor in rhizopuspepsin [51]. The reason why it could not produce any desired results was that the fragment Frf was so flexible that 5000 conformations of this fragment in the multiple conformation database did not contain any conformations that have appropriate interactions with the protein. After including the X-ray conformation of the fragment Frf in the precalculated conformations, the binding mode of the partial ligand structure (Ac-Pro-Phe-His-Frf-Val-NHMe) was successfully regenerated. The actual reduced peptide was (D) His-Pro-Phe-His-Frf-Val-Tyr; however, the terminal (D) His and Tyr were not seen in the X-ray structure in 3apr. We have tested our program in a similar system (Figure 4).

After the anchor part of the internal His fragment was superimposed on the X-ray orientation, the conformations for the rest of the part were explored. This test is equivalent to the GROW test case. The additional fragment (Frf in this case) was then added by REACT++ and the search was continued. The rmsd of the top scoring conformation was 1.54 Å for Pro-Phe-His-Frf-Val and 1.00 Å for Pro-Phe-His-Frf. This was an impressive result, considering that the number of flexible bonds was 21. The best scoring conformation of each stage of the product is compared with the X-ray conformation in Figures 5a–g.

We noticed that the calculated conformations became more similar to the X-ray conformations as the molecules grew. For example, the conformations of His, Phe and Pro at the first stages were very different from the X-ray conformations. We used AcOH at the N-terminal and $MeNH_2$ at the C-terminal of the ligand instead of (D) His and Tyr, respectively, because they were missing from the X-ray structure. We thought that the reason why the calculated conformations of the Val residue were different from the X-ray conformations was because $MeNH_2$ did not represent the Tyr residue well. In our calculated conformations, the Me group of $MeNH_2$ is in the site where the Val side chain is supposed to interact. Therefore, we performed the

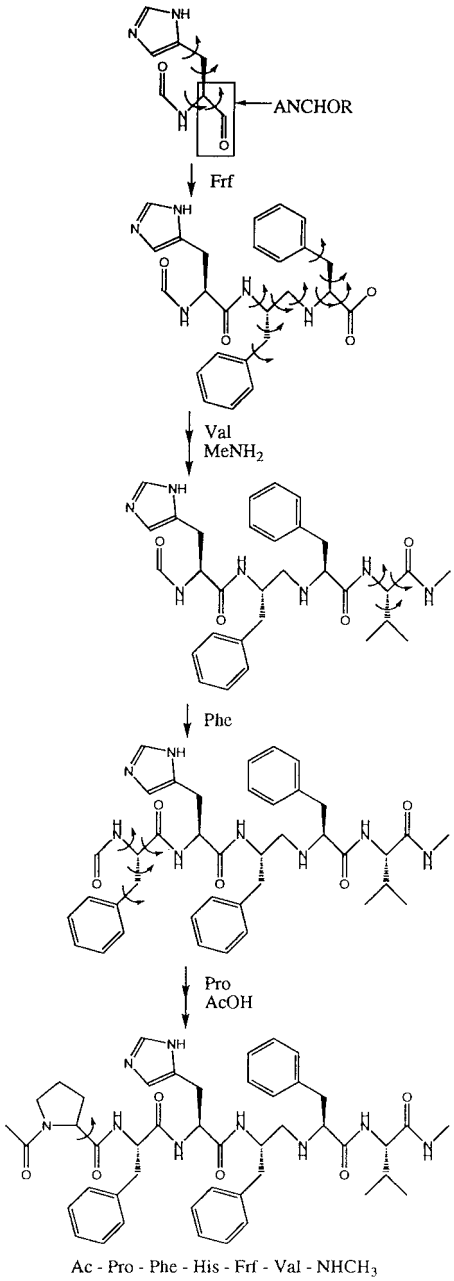


Figure 4. Order of reactions and flexible docking of the reduced peptide inhibitor used in DREAM++.

same simulation using Tyr instead of $MeNH_2$. The top scoring conformation of the calculated structure gave an rmsd of 1.07 for Pro-Phe-His-Frf-Val (the rmsd was 1.54 when $MeNH_2$ was used), yielding a conformation of Val more similar to the X-ray conformation (Figure 5h).

Since it was not possible by GROW to produce any conformations in the receptor without including the X-ray conformation of Frf, we think the run-time generation of conformations is the better method to explore the conformational space of a ligand in a receptor. Especially, it is computationally efficient in the run-time generation of conformations to be able to prune most of the conformations as the partial structures of the ligand are generated. This implies that most of the precalculated conformations are never explored during the run-time calculation.

Test system 2: HIV protease inhibitors

We have applied DREAM++ to HIV protease inhibitors that have been developed by Glaxo because there are several HIV protease–inhibitor complex structures available [7]. Before applying DREAM++ to these inhibitors developed by Glaxo, we analyze the binding modes of a different set of ligands in X-ray complex structures by examining the complex structures in the PDB identification 1aaq [52], 1hih [53], 1hii [53], 1hiv [54], 1hos [55], 1hpv [56], 1hpx [57], 1hvc [58], 1hvi [59], 1hvr [60], 4hvp [61], 7hvp [62], 8hvp [63] and 9hvp [64]. Then, the compounds developed by Glaxo are docked and the correlation between the affinities and the scores of these inhibitors is discussed.

Analyzing of binding modes

In Table 1, we report the affinities and the volume inside the receptor for the chosen ligands.

Among the HIV protease inhibitors that we examined, all ligands except 1hvr have hydrogen bonds between the carbonyl groups of ligands and water 301, and the orientations of this carbonyl group are very similar when the coordinates are superimposed by the protein backbone coordinates. There are no such hydrogen bonds in the case of 1hvr, because the oxygen atom within the inhibitor XK263 replaces water 301. Such hydrogen bonds are very important for tight binding [65]. We looked at the binding modes of other HIV protease inhibitors. All the high-affinity inhibitors showed hydrophobic interactions with the S1 pockets of the protease and hydrophobic or hydrophilic interactions with the S2 pockets of the protease [65]. However, the HIV protease inhibitors in 1hpv and 1hvr do not have any side chains that bind pockets other than the S1 and S2 pockets, suggesting that the binding to the S3 pocket is not as important as to the S1 and S2 pockets.

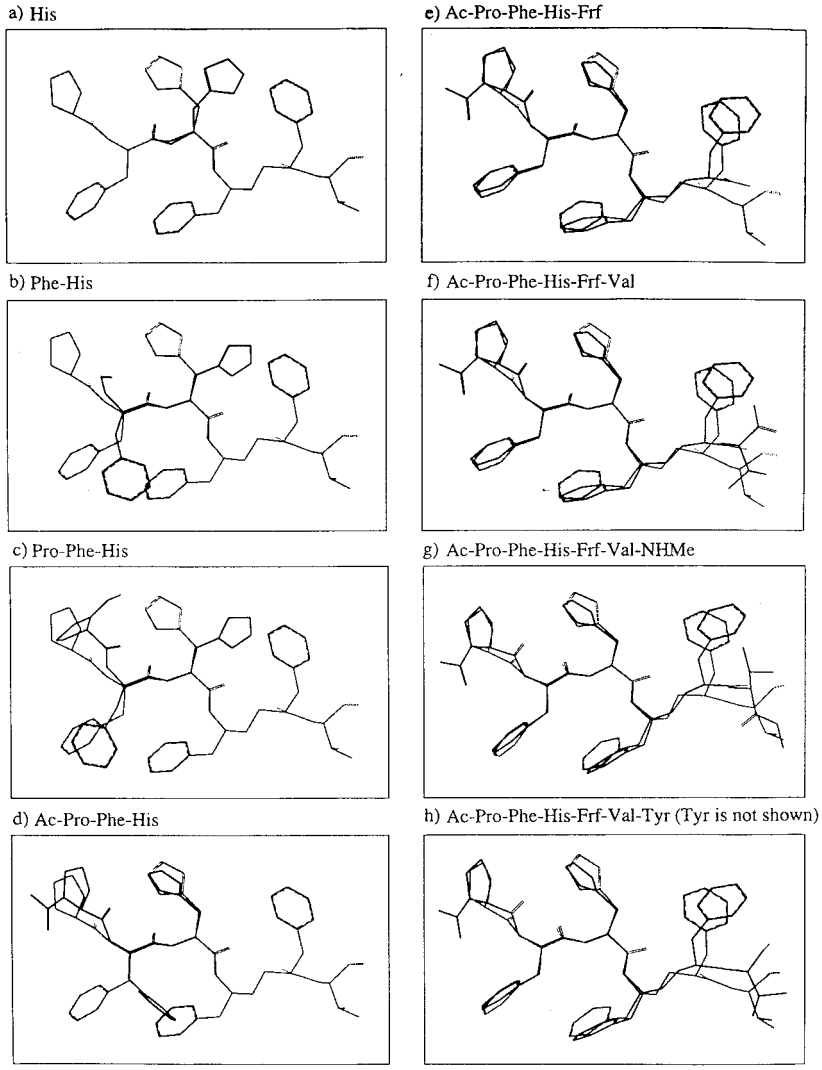


Figure 5. Top scoring conformation of each stage in the process of synthesizing Ac-Pro-Phe-His-Frf-Val-NHMe shown with the X-ray complex structure of the inhibitor (the receptor structure is not shown). The top conformation becomes similar to the X-ray complex structure as the molecule grows.

Correlation between affinity and score for penicillin-derived HIV protease inhibitors

We next applied DREAM++ to penicillin-derived HIV protease inhibitors. We did not use a DOCK algorithm (the program ORIENT++) for this part of the study. Instead, orientations of an anchor fragment were derived from a reference orientation. Moreover, since the tested ligands were symmetric, we developed a special technique for this study. The following two conditions were assumed: (1) The binding modes of ligands are C_2-symmetric. (2) The oxygen atom from the carbonyl group has the hydrogen bond with water 301. Assumption 1 was deduced from the fact that the C_2-symmetric ligands in 1hos, 1hvc, 1hvi, 9hvp and 1hvr had C_2-symmetric binding modes. Assumption 2 was deduced from the fact that many of the inhibitors for HIV protease had such hydrogen bonds (previous section). The penicillin-derived inhibitor to be docked is shown in Figure 6a. According to assumption 1, we cut the symmetric molecules in half to make a model compound for the docking study (Figure 6b).

Then, the orientation of the carbonyl group was searched around the reference carbonyl group as shown in Figure 6b. That reference carbonyl orientation was taken from 9hvp (Figure 7).

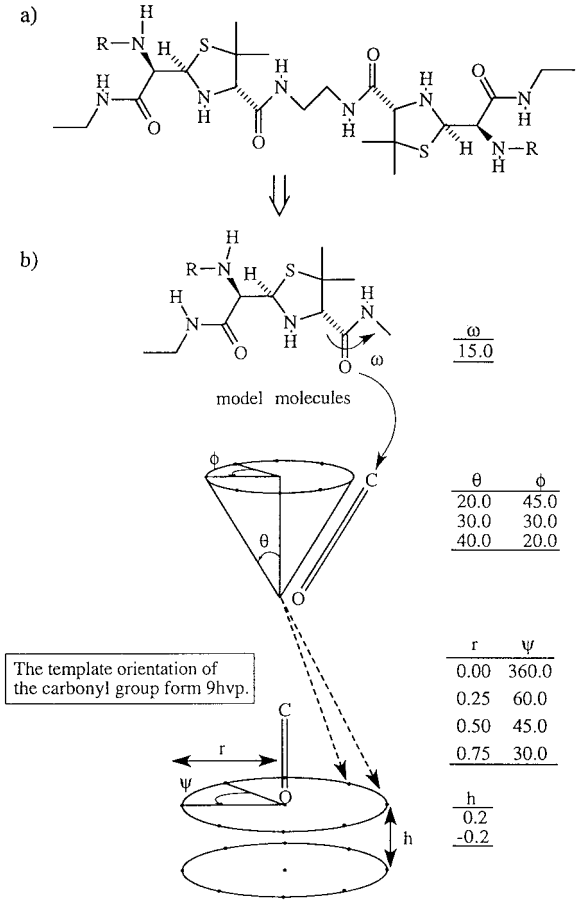


Figure 6. Method to generate the anchor orientation and to search conformations for penicillin-derived HIV protease inhibitors. (a) The original penicillin-derived inhibitor. (b) Half of the original molecule to be used for docking and conformational search with the symmetric constraint (see text). The orientation of the carbonyl group is sampled around the reference orientation of the carbonyl group from 9hvp.

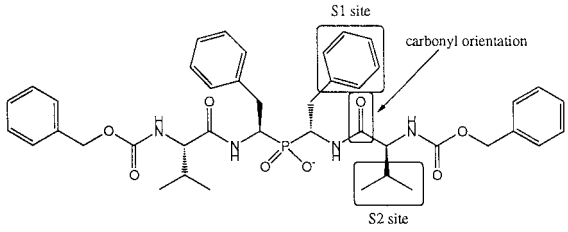


Figure 7. Reference coordinates used in the experiment for carbonyl orientation and the definition of the S1 and S2 sites.

Table 1. Affinities and the volume of ligand inside the receptor used as reference for docking

PDB ID	Ligand name	K_i(nM)	Volume (Å^3)	Reference
1aaq	Hydroxyethylene isostere	0.4	571	52
1hih	CGP53820	9 (IC_{50})	702	53
1hii	CGP53820	9 (IC_{50})	702	53
1hiv	U75875	<1	785	54, 65
1hos	SB204144	2.8	706	55
1hpv	VX-478	0.6	584	56
1hpx	KNI-272	0.0055	639	57
1hvc	A76928	0.011	849	58, 59
1hvi	A77003	0.15	833	59, 65
1hvr	XK-263	2	696	60
4hvp	MVT-101	760	778	61, 65
7hvp	JG-365	0.24	726	62, 65
8hvp	U-85548E	<1	708	63, 65
9hvp	A-74704	4.5	696	64, 65

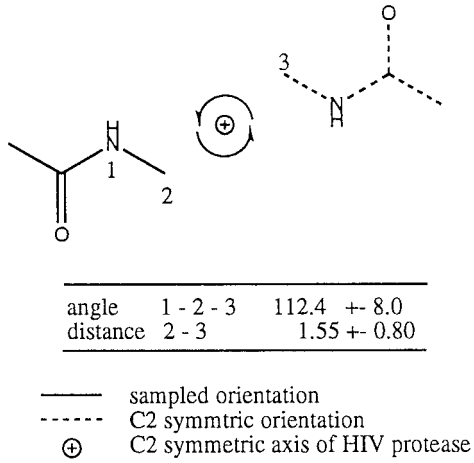


Figure 8. Orientations for the anchor fragment of penicillin-derived inhibitors are symmetrically constrained so that entire molecules can be reproduced by generating the coordinates of the symmetrical positions of the docked molecules.

Many orientations deviated from the original orientation of 9hvp as explained in Figure 6. The orientations of this rigid fragment were pruned according to the force field scoring and were further restricted by the constraint of C_2-symmetry (Figure 8), where the axis of the C_2-symmetry of the HIV protease was calculated based on the α-carbon from the protease. The C_2-symmetric orientation of this fragment was generated according to this axis; then a distance and an angle were restricted to make it possible to connect the

original fragment to the fragment generated according to C_2 symmetry by using an ethylene linker.

We think this orientation sampling method explored most of the possible symmetric orientations of the thiazolidine fragment, which possessed the specific hydrogen bond. The additional conformational search and reactions were performed in the same manner as test case 1, except that all conformations that did not have the critical interactions with the S1 and S2 pockets were removed after the attachment of the first fragment R. The S1 and the S2 pockets were defined by the center of the phenyl group of Phe and all the atoms of the side chain of Val from the ligand in 9hvp (Figure 7). Thus, compound **14**, which did not have an appropriate fragment to interact with the S1 and S2 sites, was always discarded at this stage. We first applied this search method to inhibitors shown in Figure 9 to examine the correlation between affinity and score.

The affinities of these ligands are known experimentally [8]. After the orientations and conformations were searched, the top scoring conformation for each ligand was taken and analyzed (Table 2).

The ligands were divided into two groups: those with high affinities ($<$20 nM) and those with low affinities ($>$70 nM). Several scores were evaluated to see if it was possible to distinguish high affinities from low affinities. Such scores were: (1) intermolecular + intramolecular force field energy; (2) intermolecular force field energy; (3) the volume of the part of the ligand inside a receptor; and (4) the numbers of conformations generated by SEARCH++.

Ligands **5**, **9**, **10**, **11**, **12** and **14** could not have any conformations, mainly because we have selected the binding mode that has interactions with both the S1 and S2 pockets. Of these ligands, **9**, **10**, **11**, **12** and **14** do not have high affinities. Among the ligands which have appropriate conformations (having interactions with both the S1 and S2 pockets), only compound **13** was not active. We measured the correlation between scores and affinities by calculating Z_{score} (the rank-sum test [66]). Since only the order of scores is considered and the scores themselves are not included in the calculation of Z_{score}, Z_{score} is an appropriate metric even if some of the compounds have no conformations and therefore no energies. If the prediction of the order of affinities is perfect, Z_{score} will have a maximum value of 3.1. As shown in the last line of Table 2, the Z_{score} value for the number of conformations was the largest. Thus, we further examined the number of conformations. In the original tests, we used the orientation of the carbonyl group of 9hvp as a reference. To avoid bias, we carried out the same conformational search using carbonyl orientations from the other inhibitors in the X-ray-elucidated complex structures. The distances and angles of these carbonyl groups from the X-ray-elucidated penicillin-derived inhibitor (1htg) are shown in Table 3.

The number of conformations of each molecule accepted by DREAM++ were examined (Table 4). The top six molecules with the maximum number of conformations for each system are marked by underlines. In all cases except for 1hpv and 1hpx, Z_{score} tended to have large values ($\sim$2.6). Moreover, through all the references except for 1hpv and 1hpx, the calculated conformations of compounds **1**, **2**, **3**, **4**, **6** and **7** had the same binding mode as that of similar compounds elucidated by X-ray (1hte, 1htf, 1htg and a previous paper [7]), having a thiazolin ring and fragments R in the S1 and the S2 sites, respectively.

We next examined 1hpv and 1hpx to find the reason why these two results were not as good as the others. These two carbonyl orientations were the most different, leading to a difficulty in regenerating the binding modes. The reason why DREAM++ failed to pick compound **5** as active was that it was difficult to sample the specific binding mode due to tight geometric constraints. We hope to return to this problem in the future. Although DREAM++ could not find any conformation for compound **5**, in general, we think a compound with 0 conformations in our calculation indicates that it does not have strong affinity to its receptor since such a compound cannot have plausible docked conformations.

Overall, it was possible to identify most of the active compounds simply by the number of possible conformations. We think this number of conformations might be related to an entropical factor for ligand binding.

The computation time for docking 14 compounds was 848 s (61 s per compound).

Application: Finding new types of putative inhibitors

We next applied DREAM++ to find new types of putative inhibitors for HIV protease. Since the purpose in finding putative inhibitors here was simply to show the ability of DREAM++ to design synthetically accessible putative inhibitors, we limited the number of molecules by focusing on compounds similar to penicillin-derived inhibitors. The conditions we set in this trial were as follows: (1) Compounds have to be C_2-symmetric connected by the ethylene diamine

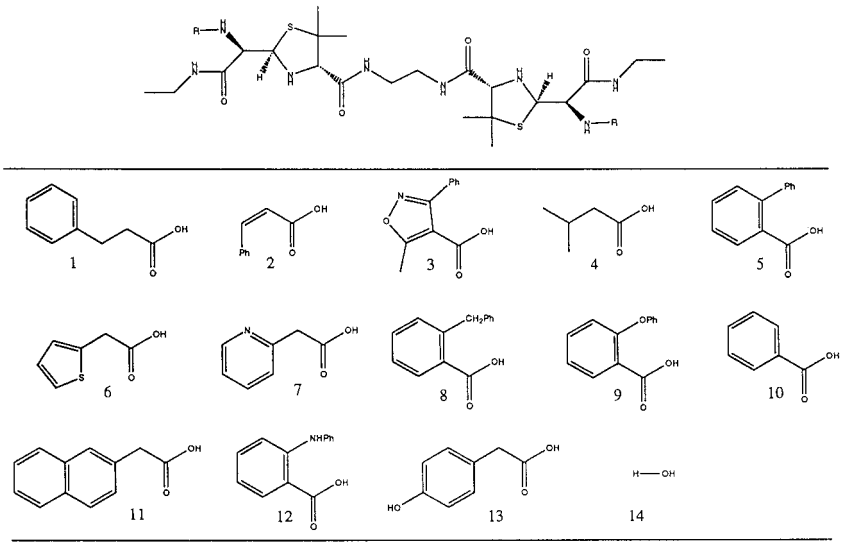


Figure 9. Penicillin-derived compounds used to examine the reliability of the DREAM++ prediction of affinities. The reagents in this figure were reacted with the amine functions of reactant molecules. The affinities of these compounds are known experimentally, as shown in Table 2.

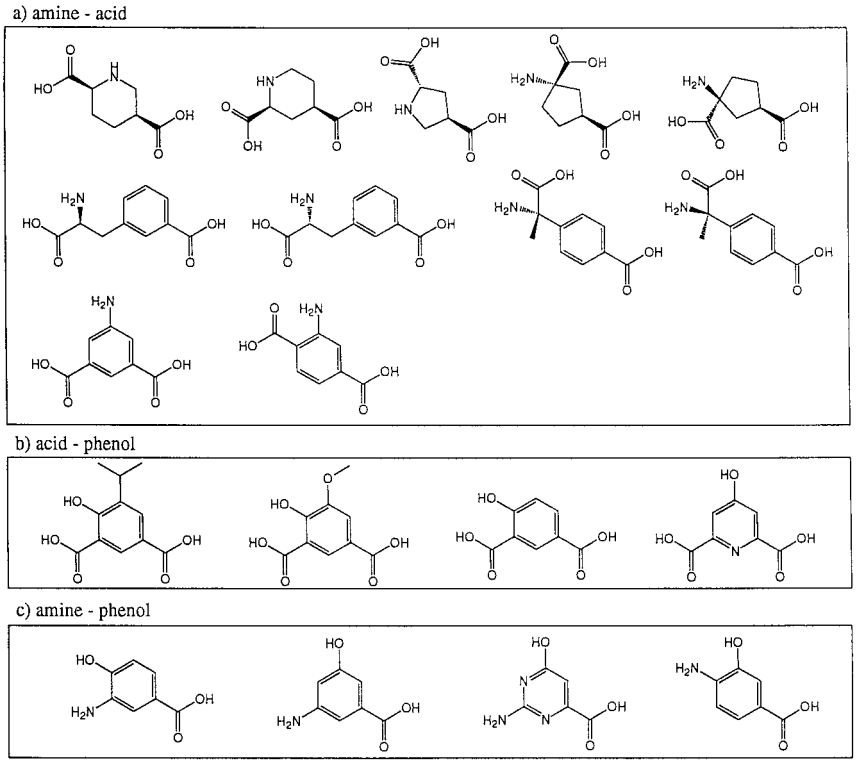


Figure 10. List of first fragments to be used to generate virtual combinatorial libraries (carboxylic acid plus two other functionalities).

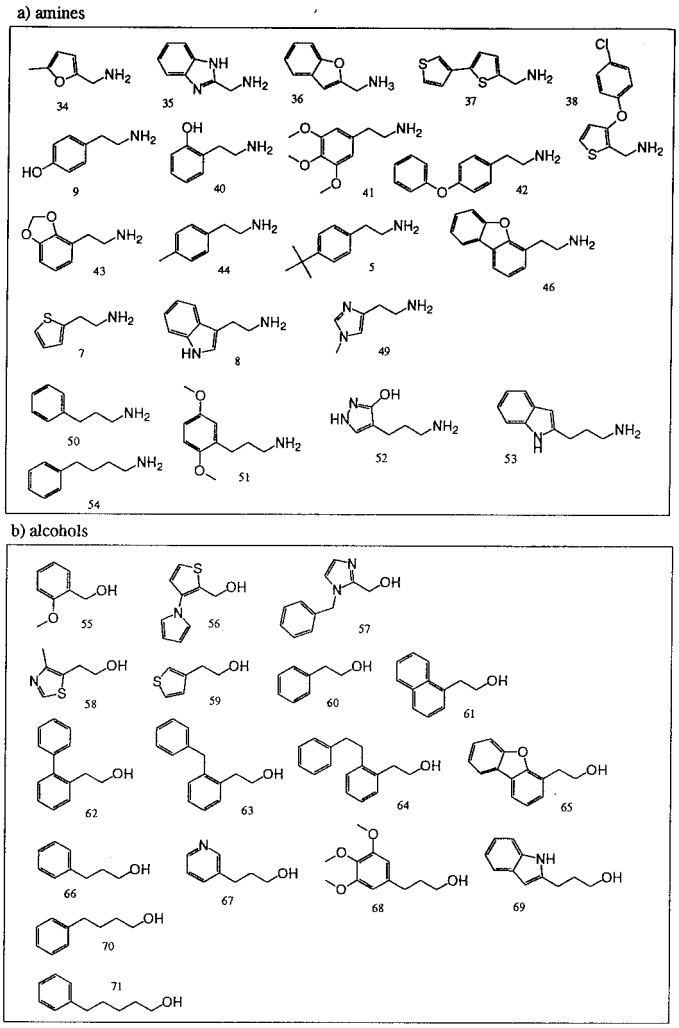


Figure 11. List of second fragments to be used to generate virtual combinatorial libraries (amines and alcohols). These fragments have to be reacted with appropriate chemical functional groups as described in Figure 12.

linker. (2) Compounds have to be composed of three different fragments to express chemical diversity. (3) Compounds have a hydrogen bond with water 301 similar to many HIV protease inhibitors.

The first fragments were selected to satisfy the above three conditions as follows: (1) The carboxylic acid group is directly connected to a ring system. (2) A fragment has to have two additional functional groups to react with two other different fragments. According to these conditions, the Available Chemicals Directory (ACD 95.1) was searched by ISIS3D. The results were further visually inspected to choose the first position fragments as shown in Figure 10.

According to the chemical nature of the first fragments, we selected the second fragments: amines, acids and alcohols connected to ring systems by methylene chains from ACD by ISIS3D (Figure 11).

The fragments in Figure 10 were used as acid fragments. By using these fragments, we built up the libraries shown in Figure 12a. The molecule depicted in Figure 12b represents half of the molecules that are actually docked considering the symmetry of molecules.

The methods to generate orientations and to search out conformations were the same as the search methods for penicillin-derived inhibitors. Third fragments were chosen to be simple, matching the chemical natures to be connected properly ($EtNH_2$, EtCOOH, EtOH) in order to screen the sets of first and second fragments. The conformations of intermediates that

Table 2. Results of docking studies of the penicillin-derived inhibitors using the carbonyl group from 9hvp

Compound	IC_{50} (nM)	No. of flexible bonds[a]	Inter + intra energy[b]	Inter energy[c]	Volume[d] (Å^3)	No. of conformations
1	2.4	3	−28.0	−36.5	376	36
2	2.9	2	−31.7	−38.3	378	35
3	5.4	2	−25.2	−30.7	289	40
4	8.7	2	−20.3	−29.9	348	100
5	9.9	2	–[e]	–	–	0
6	15	2	−29.0	−36.2	345	91
7	19	2	−28.7	−36.3	351	78
8	72	3	–	–	–	0
9	160	3	–	–	–	0
10	390	1	–	–	–	0
11	820	2	–	–	–	0
12	4800	3	–	–	–	0
13	>12000	3	−24.1	−36.5	351	19
14	120000	0	–	–	–	0
Z_{score}			2.49	2.04	2.30	2.62

[a]Number of flexible bonds in the fragment R.
[b]The lowest force field energy (inter + intra) of each compound is shown.
[c]The inter force field energy of the conformation that gives the lowest inter + intra energy.
[d]Volume of a ligand inside the receptor.
[e]No appropriate conformation was generated.

Table 3. Relationships between various carbonyl orientations of inhibitors in X-ray structures[a]

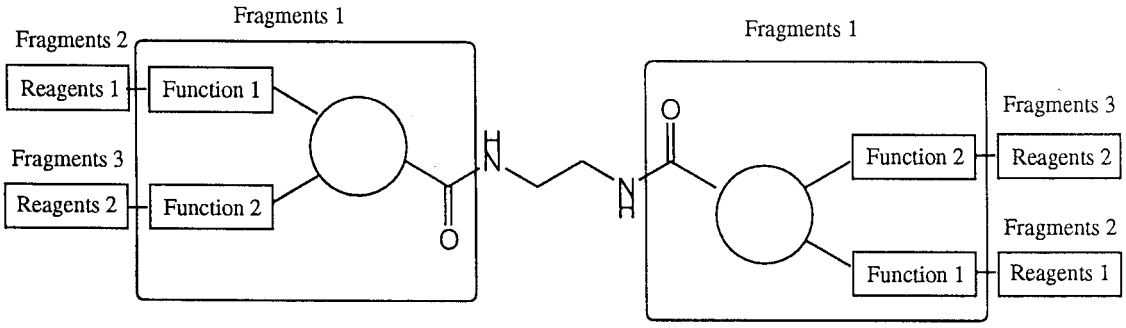
PDB ID	Distance (Å)[b]	Angle (deg)
1aaq	0.98	26.2
1hih	0.93	24.5
1hii	0.76	23.1
1hiv	0.55	16.2
1hos	0.42	32.6
1hpv	1.14	17.9
1hpx	1.01	34.7
1htg	0.00	0.0
1hvc	0.52	29.7
1hvi	0.75	26.0
4hvp	0.56	11.3
7hvp	0.92	17.4
8hvp	0.73	7.7
9hvp	0.59	30.9

[a]The X-ray structures are superimposed using the main chain of HIV protease.
[b]Distances between the oxygen atoms

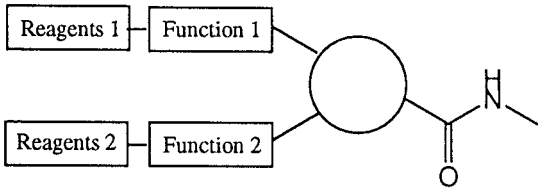
did not have interactions with either S1 and S2 pockets were discarded in the same way as in the search for penicillin-based inhibitors. The results of the searches are shown with the numbers of compounds in each category of molecules (Table 5).

The generated molecules (a total of 137 compounds) with conformations assumed to be active were graphically examined. It was found that many of the conformations had a very small amount of volume overlap with the S1 and S2 pockets because the attached fragments were docked outside the receptors (Figure 13).

While some of these conformations had favorable force field scores, we found that the ligands with unfavorable conformations tended to have small volumes inside the receptor according to our definition. Since the ligands in the PDB used for references had a range of volumes from 571–745 Å^3, conformations with more than 285 Å^3 were selected. As a result, conformations that have interactions with the S1 and S2 pockets were successfully extracted, where it was not possible for force field scoring to distinguish such binding modes from those that have interaction mainly with the outside of the receptor. We think this new metric is very useful as a complement of force field scoring. Although contact areas were also evaluated, the binding modes that have interactions with the S1 and S2 pockets could not be distinguished from those that have interactions mainly with the outside of these



a) The chemical library designed as the alternative inhibitor of HIV protease inhibitor.
The fragments 1 have two of the following function groups: amines, carboxylic acids and phenols.



b) The molecules that are actually docked to HIV protease considering symmetory of molecules and the protease.
Proper sets of reagents are selected for the virtual reactions and the following sets of molecules are generated:

i) fragments 1 (carboxylic acid and amine) + fragments2 (carboxylic acid) + $EtNH_2$
ii) fragments 1 (carboxylic acid and amine) + fragments2 (amine) + EtCOOH
iii) fragments 1 (carboxylic acid and phenol) + fragments2 (amine) + EtOH
iv) fragments 1 (carboxylic acid and phenol) + fragments2 (alcohol) + $EtNH_2$
v) fragments 1 (amine and phenol) + fragments2 (carboxylic acid) + EtOH
vi) fragments 1 (amine and phenol) + fragments2 (alcohol) + EtCOOH

Figure 12. Libraries to find new types of putative inhibitors for HIV protease.

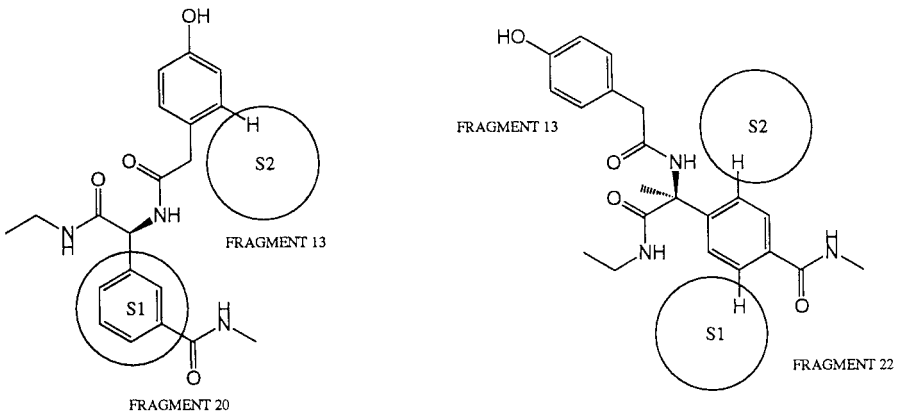


Figure 13. Unfavorable conformations of virtually generated compounds where the S1 or S2 binding site is not occupied sufficiently.

Table 4. Number of possible conformations of compounds using different carbonyl reference orientations

Compound	1aaq	1hih	1hii	1hiv	1hos	1hpv	1hpx	1htg	1hvc	1hvi	4hvp	7hvp	8hvp	9hvp
1	17	38	25	36	30	13	6	52	35	32	31	19	40	36
2	22	42	31	51	44	11	30	38	50	60	58	30	54	35
3	22	94	32	100	48	0	0	100	100	100	100	59	79	40
4	87	49	100	95	96	54	100	97	100	85	91	60	64	100
5	0	0	0	0	0	0	0	0	0	3	2	0	3	0
6	83	73	71	88	96	25	52	96	100	84	92	60	61	91
7	72	65	73	71	81	2	31	93	93	77	81	40	62	78
8	0	6	0	0	0	0	0	12	0	0	2	0	0	0
9	0	0	0	6	0	0	0	9	0	15	12	0	32	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	1	1	3	2	24	9	0	0	3	4	0	3	0
12	0	3	0	0	0	0	0	18	0	24	21	0	31	0
13	7	11	15	11	15	0	9	14	22	23	18	12	11	19
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Z_{score}	2.62	2.43	2.56	2.49	2.56	1.73	1.73	2.43	2.62	2.68	2.56	2.62	2.68	2.62

The top six molecules with the maximum number of possible conformations for each system (column) are underlined.

pockets. The reduction of conformations based on volume not only lessened the number of compounds (a total of 83 compounds, Table 5) but also reduced the number of conformations for each compound, which in turn affected the further reduction of compounds according to the number of conformations. Compounds with less than 25 conformations were discarded, leaving 10 compounds as a total, composed of only three different fragment groups (Table 5). All the 10 remaining compounds are listed in Table 6 with scores and numbers of conformations.

As a representative, the compound with the best energy is shown together with a penicillin-derived inhibitor binding to the protease (Figure 14).

The putative inhibitor occupies the S1 and S2 pockets very nicely. Of course, we do not know if any of these proposed compounds are active, but we introduce them to show that this procedure can lead to very plausible candidates. The energies, number of conformations and buried volume for these compounds compare well with the known high-affinity inhibitors.

The total computation time of calculation for Table 5 was 1.3×10^5 s, screening 646 virtual compounds (203 s per compound).

Discussion

Screening a database by a docking program is an efficient strategy to find experimentally active compounds if compounds in the database are readily available. The problem is whether or not a database is large enough to contain active compounds. One advantage of synthetically reliable de novo ligand design is that it is possible to focus on certain classes of chemical structures based on fragments which fit to a receptor either experimentally or theoretically. A second advantage is that it is easy to modify the structures of known hits, because we can think about future possible modification before building up a virtual library. To modify hit compounds from a pre-existing database, one has to spend a large amount of time to find synthetically reliable intermediates from which a great number of compounds containing critical features can be synthesized easily.

Before starting our search for new putative inhibitors, we were concerned that our evaluation procedures might be nonselective. However, this screening identified only three out of 18 carefully selected first fragments as parts of the active compounds. This is encouraging considering the purpose of a screening, but further experimental tests for more diversified compounds are necessary to estimate the quality of results.

These studies raise several issues for conformational search. The truncation procedure at each stage of conformational search is particularly critical in the case of incremental search. The number of conformations in each stage is shown in Figure 15 schematically.

Table 5. Number of compounds that meet 'activity criteria' in the libraries

Second fragments	Third fragments	First fragments (index)	Number of compounds ($N > 0$)[a]	Number of compounds ($N > 0$; volume > 285)[a]	Number of compounds ($N \geq 25$; volume > 285)
Acid	Amine	20	12	12	5
		21	9	9	0
		22	1	1	0
		25	3	3	0
Amine	Acid	18	3	3	0
		20	6	2	0
		21	19	18	4
		23	20	0	0
Amine	Alcohol	26	6	0	0
		27	20	11	0
		28	2	2	0
		29	1	1	0
Alcohol	Amine	26	1	0	0
		27	15	3	0
Acid	Alcohol	30	3	2	0
		31	5	4	0
		33	9	0	0
Alcohol	Acid	31	1	1	0
		33	16	11	1

[a]The number of conformations for each molecule.

Table 6. Each compound that meets 'activity criteria' in the libraries

Fragment 1 (index)	Fragment 3	Fragment 2 (index)	Inter + intra energy	Volume (Å^3)	No. of conformations
		1	−22.5	334	31
20	$EtNH_2$	6	−24.0	328	61
		7	−25.1	335	59
		9	−30.2	370	25
		13	−27.4	347	25
		34	−22.3	322	37
21	EtCOOH	47	−21.7	303	36
		51	−26.7	345	67
		52	−23.4	290	27
33	EtCOOH	59	−22.4	296	26

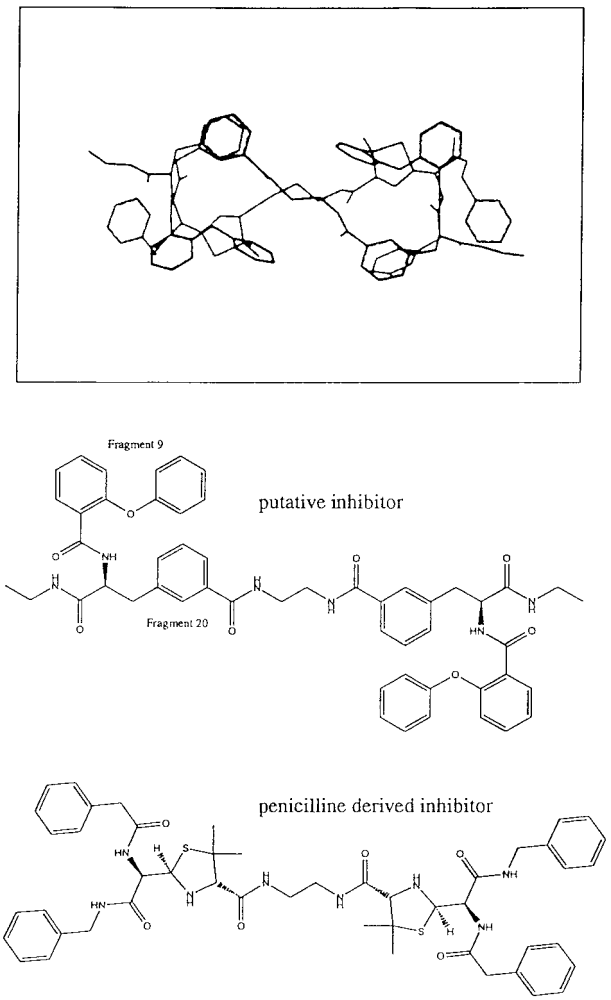


Figure 14. The top scoring putative inhibitor shown with the X-ray complex of a penicillin-derived inhibitor (the receptor is not shown).

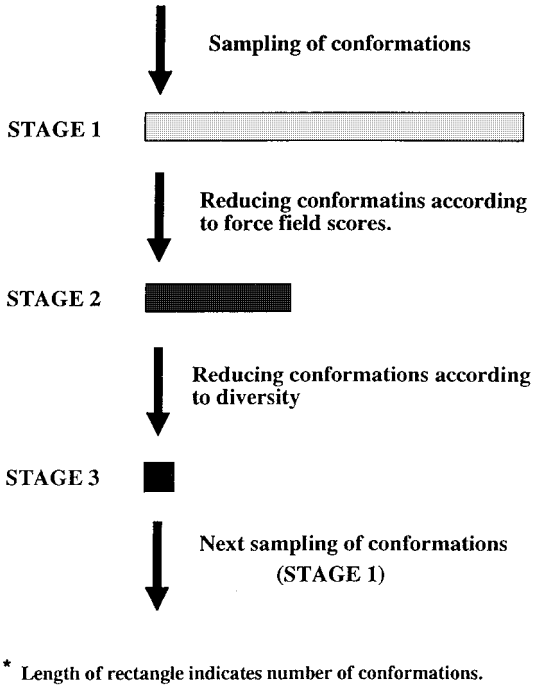


Figure 15. Number of conformations to be considered in each conformational search stage. The sizes and the number of conformations are critical to produce good results.

The sizes and ratios of the number of conformations in each stage are very important in obtaining good results for the following reasons. The number of conformations stored in stages 2 and 3 are fixed within SEARCH++. (The numbers we used in this article for stages 2 and 3 were 500 and 100, respectively. The numbers can be specified by users.) Thus, if the number of conformations sampled is increased by changing other parameters (the number of allowed flexible bonds within the divided fragments, the number of degrees of each torsion angle set in order to scan for conformations, etc.), all the increased conformations are discarded based on force field energies. This reduction of conformations might cause a loss of conformational diversity in stage 2. If the number in stage 3 is too small, the critical conformations needed for obtaining appropriate results might be lost. However, if the number in stage 3 becomes larger, the number of conformations in the next sampling stage will be increased, losing conformational diversity through the reduction of conformations according to force field energies, as we have mentioned above. Therefore, the ratios among the number of conformations at each stage as well as the actual numbers themselves are important. Although the larger numbers may be better if the ratios are the same, more computation times and memory are necessary. Thus, the size of numbers depends on the size of the library to be searched. One of the merits of this hybrid algorithm is that the search space can also be controlled based on a different aspect as follows. If the maximum flexible bond in each fragment for backtrack search is set large enough, simple backtrack search will be performed. The larger number of the maximum flexible bonds in each fragment requires more computation time in general. However, the larger conformational space and computation time do not always mean that the quality of results is better, because of the problem of ratios that we have discussed above.

The successful reduction of computation time was achieved by combining docking programs

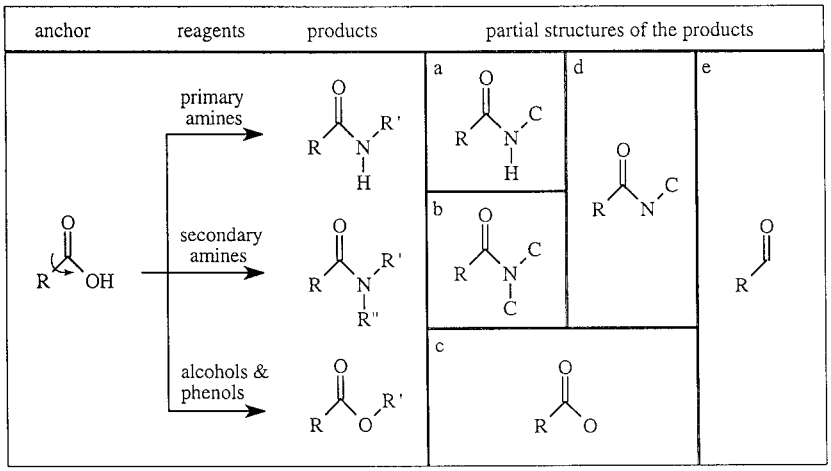


Figure 16. Design of programs for docking and reactions in concert. The virtual reaction has to be performed considering the future chemical modification of other functional groups to make the conformational search of the intermediate representative of the entire molecule. (i) The usage of the structures **a**, **b**, **c** or **d** in the conformational search of carboxylic acid restricts the reagents to be used in the next reaction as follows: **a** – primary amines, **b** – secondary amines, **c** – alcohols, **d** – amines). (ii) The usage of the structure **e** does not restrict the next reagents. However, this structure cannot screen the conformations as efficiently as **a**, **b**, **c** or **d**.

(ORIENT++ and SEARCH++) and a reaction program (REACT++). These conceptually different programs were designed to work together, making it possible to inherit conformations through reactions. For example, if inheritance of conformation was not used, it would take 1869 s instead of 848 s for the calculation of Table 3; thus, 45% of the computation time was saved because of the inheritance. The amount of reduction of calculation time depends on how long it takes to calculate conformations to be inherited and how many times such conformational information is used.

Obviously, the reaction generator and the conformational search program are coupled together. This coupling should be tightened further. For example, we searched the conformations of carboxylic acids, knowing that they were going to form peptide bonds by reaction with the primary amines. It would be better to search the partial structures of future products such as **a** in Figure 16 by substituting the partial structures of future products for carboxylic acid.

However, carboxylic acids can also react with secondary amines, alcohols or phenols. Thus, the substitution of the partial structures of different future products for the same carboxylic acids such as **a**, **b** and **c** should be considered. In this case, a reaction program has to recognize three different chemical patterns as a carboxylic acid functionality, making the reaction program complicated. If **a** and **c**, or even **e**, is used as the partial structures of the products, the reaction program will be simpler than having three different patterns **a**, **b** and **c**. However, the efficiency of a search will be lessened. Although the representation of the partial structures for future products is closely related to conformational searches, it is advantageous to avoid changing the search program according to the modifications of these representations. In this article, we simply used **c** to represent all the products (without substituting for carboxylic acid) without complications; we think the above approach should further improve the results of conformational searching.

We would like to emphasize that the reactions implemented in the program REACT++ are common; however, the virtual reactions we used are not realistic from the point of view of organic chemistry, where functional groups have to be appropriately protected to react selectively and to help purifications. In the program REACT++, no protection groups are necessary unless there is more than one possible reaction center in reactants and reagents. In such cases, functional groups can be protected easily on the computer but not so easily in the laboratory. Moreover, the order of reactions should be fixed to make it possible to search conformations incrementally. Thus, we think it is important not to allow reaction programs to automatically decide what kind of reaction should take place, but to let users specify them.

Scoring procedures are also a serious issue. The molecular force field contains no entropy contributions and requires knowledge about bound water

molecules to work properly. Therefore, we used the volume-based scoring method to screen conformations (the screening of conformations by certain volume as in Table 6) as one way to express the effect of water molecules implicitly. More work is needed to establish the usefulness of this procedure.

Conclusions

We have developed the set of programs DREAM++ to design combinatorial libraries effectively, introducing new techniques: (1) a hybrid algorithm between backtrack and incremental construction algorithms; and (2) the inheritance of conformations through reactions. DREAM++ successfully regenerated the X-ray binding modes of HIV protease inhibitors, showing the reliability of these algorithms. The correlation between affinity and the number of conformations was examined to find the best way to screen virtually generated molecules. Finally, we have demonstrated the applicability of DREAM++ by proposing putative symmetric HIV protease inhibitors.

Acknowledgements

Support for this work was provided by grants GM-31497 and GM-39552 (C. Craik, Principal Investigator). We thank Greg Couch of the UCSF Computer Graphics Laboratory for many suggestions about programming in C++. We also thank Dr. Takashi Tuji, Dr. Mitsuo Takahashi, Masataka Shoji and Nobuyasu Suzuki of Ajinomoto Co. Inc. for helpful comments and discussions. Tripos Associates provided the SYBYL program, and Molecular Design Limited provided the Available Chemicals Directory, for which we express our appreciation.

References

1. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T., J. Mol. Biol., 161 (1982) 269.
2. Ewing, T.J.A. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1175.
3. Leach, A.R. and Kuntz, I.D., J. Comput. Chem., 13 (1992) 730.
4. Makino, S. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1812.
5. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., J. Mol. Biol., 261 (1996) 470.
6. Welch, W., Ruppert, J. and Jain, A., Chem. Biol., 3 (1996) 449.
7. Wonacott, A., Cooke, R., Hayes, F.R., Hann, M.M., Jhoti, H., McMeekin, P., Mistry, A., Murray-Rust, P., Singh, O.M. and Weir, M.P., J. Med. Chem., 36 (1993) 3113.
8. Humber, D.C., Bamford, M.J., Bethell, R.C., Cammack, N., Cobley, K., Evans, D.N., Gray, N.M., Hann, M.M., Orr, D.C., Saunders, J., Shenoy, B.E.V., Storer, R., Weingarten, G.G. and Wyatt, P.G., J. Med. Chem., 36 (1993) 3120.
9. Thompson, L.A. and Ellman, J.A., Chem. Rev., 96 (1996) 555.
10. Martin, E.J., Blaney, J.M., Siani, M.A., Spellmeyer, D.C., Wong, A.K. and Moos, W.H., J. Med. Chem., 38 (1995) 1431.
11. Pearlman, R.S., Lead Generation & Optimization, Conference by Strategic Research Institute in Princeton, NJ, 18 September 1996.
12. Kick, E.K., Roe, D.C., Skillman, A.G., Liu, G.C., Ewing, T.J.A., Sun, Y., Kuntz, I.D. and Ellman, J.A., Chem. Biol., 4 (1997) 297.
13. Sun, Y., Ewing, T.J.A., Skillman, A.G. and Kuntz, I.D., J. Comput.-Aided Mol. Design, 12 (1998) 597.
14. Murray, C.W., Clark, D.E., Auton, T.R., Firth, M.A., Li, J., Sykes, R.A., Waszkowycz, B., Westhead, D.R. and Young, S.C., J. Comput.-Aided Mol. Design, 11 (1997) 193.
15. Nishibata, Y. and Itai, A., J. Med. Chem., 36 (1989) 2921.
16. Rotstein, S.H. and Murcko, M.A., J. Comput.-Aided Mol. Design, 7 (1993) 23.
17. Gehlhaar, D.K., Moerder, K.E., Zichi, D., Sherman, C.J., Ogden, R.C. and Freer, S.T., J. Med. Chem., 38 (1995) 466.
18. Pearlman, D.A. and Murcko, M.A., J. Comput. Chem., 14 (1993) 1184.
19. Bohacek, R.S. and McMartin, C., J. Am. Chem. Soc., 116 (1994) 5560.
20. Rotstein, S.H. and Murcko, M.A., J. Med. Chem., 36 (1993) 1700.
21. Gillet, V.J., Johnson, P., Mata, P., Sike, S. and Williams, P., J. Comput.-Aided Mol. Design, 7 (1993) 127.
22. Gillet, V.J., Newell, W., Mata, P., Sike, S., Zsoldos, Z. and Johnson, A.P., J. Chem. Inf. Comput. Sci., 34 (1994) 207.
23. Moon, J.B. and Howe, W.J., Proteins, 11 (1991) 314.
24. Böhm, H.-J., J. Comput.-Aided Mol. Design, 6 (1992) 61.
25. Böhm, H.-J., J. Comput.-Aided Mol. Design, 6 (1992) 593.
26. Roe, D.C. and Kuntz, I.D., J. Comput.-Aided Mol. Design, 9 (1995) 269.
27. DeWitte, R.S. and Shakhnovich, E.I., J. Am. Chem. Soc., 118 (1996) 11733.
28. Pearlman, D.A. and Murcko, M.A., J. Med. Chem., 39 (1996) 1651.
29. Böhm, H.-J., J. Comput.-Aided Mol. Design, 10 (1996) 265.
30. Kearsley, S.K., Underwood, D.J., Sheridan, R.P. and Miller, M.D., J. Comput.-Aided Mol. Design, 8 (1994) 565.
31. Miller, M.D., Kearsley, S.K., Underwood, D.J. and Sheridan, R.P., J. Comput.-Aided Mol. Design, 8 (1994) 153.
32. DesJarlais, R.L., Sheridan, R.P., Dixon, J.S., Kuntz, I.D. and Venkataraghavan, R., J. Med. Chem., 29 (1986) 2149.
33. Goodsell, D.S. and Olson, A.J., Proteins, 8 (1990) 195.
34. Morris, G.M., Goodsell, D.S., Huey, R. and Olson, A.J., J. Comput.-Aided Mol. Design, 10 (1996) 293.
35. Ghose, A.K. and Crippen, G.M., J. Comput. Chem., 6 (1985) 350.
36. Mizutani, M.Y., Tomioka, N. and Itai, A., J. Mol. Biol., 243 (1994) 310.
37. Oshiro, C.M., Kuntz, I.D. and Dixon, J.S., J. Comput.-Aided Mol. Design, 9 (1995) 113.
38. Clark, K.P. and Ajay, J. Comput. Chem., 16 (1995) 1210.
39. Chang, G., Guida, W.C. and Still, W.C., J. Am. Chem. Soc., 111 (1989) 4379.

40. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta Jr., S. and Weiner, P., J. Am. Chem. Soc., 106 (1984) 765.
41. Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A., J. Comput. Chem., 7 (1986) 230.
42. Meng, E.C., Shoichet, B.K. and Kuntz, I.D., J. Comput. Chem., 13 (1992) 505.
43. Jones, G., Willett, P. and Glen, R.C., J. Comput.-Aided Mol. Design, 9 (1995) 532.
44. Ewing, T.J.A., Ph.D. Thesis, University of California at San Francisco, CA, 1997.
45. Nelder, J.A. and Mead, R., J. Comput., 7 (1965) 308.
46. Hermkens, P.H.H., Ottenheijm, H.C.J. and Rees, D., Tetrahedron, 52 (1996) 4527.
47. Perry, N.C. and van Geerestein, V.J., J. Chem. Inf. Comput. Sci., 32 (1992) 607.
48. SYBYL, v. 6.0.2, Tripos Associates, St. Louis, MO, 1993.
49. Marsili, M. and Gasteiger, J., Croat. Chim. Acta, 52 (1980) 601.
50. Gasteiger, J. and Marsili, M., Tetrahedron, 36 (1980) 3210.
51. Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D. and Davies, D.R., Proc. Natl. Acad. Sci. USA, 84 (1987) 7009.
52. Dreyer, G.B., Lambert, D.M., Meek, T.D., Carr, T.J., Tomaszek, T.A., Fernandez, A.V., Bartus, H., Cacciavillani, E., Hassell, A.M., Minnich, M., Petteway, S.R. and Metcalf, B.W., Biochemistry, 31 (1992) 6646.
53. Priestle, J.P., Fassler, A., Rosel, J., Tintelnot-Blomley, M., Strop, P. and Gruetter, M.G., Structure, 3 (1995) 381.
54. Thanki, N., Rao, J.K.M., Foundling, S.I., Howe, W.J., Tomasselli, A.G., Heinrikson, R.L., Thaisrivongs, S. and Wlodawer, A., Protein Sci., 1 (1992) 1061.
55. Abdel-Meguid, S., Zhao, B., Murthy, K.H.M., Winborne, E., Choi, J.-K., DesJarlais, R.L., Minnich, M.D., Culp, J.S., Debouchk, C., Tomaszek Jr., T.A., Meek, T.D. and Dreyer, G.B., Biochemistry, 32 (1993) 7972.
56. Kim, E.E., Baker, C.T., Dwyer, M.D., Murcko, M.A., Rao, B.G., Tung, R.D. and Navia, M.A., J. Am. Chem. Soc., 117 (1995) 1181.
57. Baldwin, E.T., Bhat, T.N., Gulnik, S., Liu, B., Topol, I.A., Kiso, Y., Mimoto, T., Mitsuya, H. and Erickson, J.W., Structure, 3 (1995) 581.
58. Bhat, T.N., Baldwin, E.T., Liu, B., Cheng, Y.-S.E. and Erickson, J.W., Nat. Struct. Biol., 1 (1994) 552.
59. Hosur, M.V., Bhat, T.N., Kempf, D., Baldwin, E.T., Liu, B., Gulnik, S., Wideburg, N.E., Norbeck, D.W., Appelt, K. and Erickson, J.W., J. Am. Chem. Soc., 116 (1994) 847.
60. Lam, P., Jadhav, P., Eyermann, C., Hodge, N., Ru, Y., Bacheler, L., Meek, J., Otto, M., Rayner, M., Wong, Y., Chang, C.-H., Weber, P., Jackson, D., Sharpe, T. and Erickson-Viitanen, S., Science, 263 (1994) 380.
61. Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B.H. and Wlodawer, A., Science, 246 (1989) 1149.
62. Swain, A.L., Miller, M.M., Green, J., Rich, D.H., Schneider, J., Kent, S.B.H. and Wlodawer, A., Proc. Natl. Acad. Sci. USA, 87 (1990) 8805.
63. Jaskolski, M., Tomasselli, A.G., Sawyer, T.K., Staples, D.G., Heinrikson, T.L., Schneider, J., Kent, S.B.H. and Wlodawer, A., Biochemistry, 30 (1991) 1600.
64. Erickson, J., Neidhart, D.J., Vandrie, J., Kempf, D.J., Wang, X.C., Norbeck, D.W., Plattner, J.J., Rittenhouse, J.W., Turon, M., Wideburg, N., Kohlbrenner, W.E., Simmer, R., Helfrich, R., Paul, D.A. and Knigge, M., Science, 249 (1990) 527.
65. Wlodawer, A. and Erickson, J.W., Annu. Rev. Biochem., 62 (1993) 543.
66. Glenberg, A.M., Data, Harcourt Brace Jovanovich, Inc., 1988.