3DFS: A New 3D Flexible Searching System for Use in Drug Design

Ting Wang* and Jiaju Zhou

Laboratory of Computer Chemistry (LCC), Institute of Chemical Metallurgy, Chinese Academy of Sciences, P.O. Box 353, Beijing 100080, China

Received August 30, 1997[⊗]

This paper describes a new 3D flexible searching system 3DFS which supports two types of query definitions: simple atom-based definition and generalized function-based definition. The simple and practical definitions of hydrogen bond acceptors/donors, charge centers, aromatic ring centers, and a rapid hydrophobe recognition algorithm are described in detail. 3DFS adopts a four-step searching strategy: a rapid ID screening, an exact 2D substructure searching using the GMA algorithm, a rigid 3D searching, and a conformationally flexible 3D searching using POWELL method. The utility of 3DFS is illustrated by several typical searching examples.

1. INTRODUCTION

With the advent of pharmacophore concept, three-dimensional (3D) searching for a pharmacophore in a 3D database emerged as a useful tool for exploiting 3D database for potential active compounds.[1−4] However, 3D searching technology was not widely practiced until a variety of large 3D databases were generated by 2D-3D conversion programs such as CONCORD,[5] Chem-X,[6] or Corina.[7]

3D searching retrieves compounds, called hits, which match the pharmacophore query. Of these hits, some are similar to known active compounds, but some others may be entirely new classes of compounds. Thus 3D searching provides a prompt approach to the discovery of novel lead compounds. Moreover because the hits do exist and are synthetically accessible, even available for testing, the process of lead discovery can be further accelerated. Since the first 3D searching program[8] appeared 20 yeas ago, various 3D searching systems have been developed.[9−12] 3DSEARCH[9] and ALADDIN[10a] are two early 3D searching systems, subsequently a number of commercial systems appeared, primarily including MACCS-II/3D (MDL Information System, Inc.), ISIS/3D (MDL Information System, Inc.), Chem-X (ChemDBS-3D) (Chemical Design Ltd.), SYBYL (3DB Unity) (Tripos Associates), and Catalyst (BioCad Corp). The abundance of 3D searching systems is a testament to its interest in drug design. Recently, 3D searching has been successfully applied to the discovery of novel protein kinase C agonists,[13] HIV-1 protease inhibitors,[14,15] and angiotensin II antagonist.[16]

The efficiency (i.e., computational requirement) and the effectiveness (i.e., number of hits) of a 3D searching system are primarily determined by its query definition and searching algorithms (including 1D, 2D, and 3D searching algorithms). Besides efficiency and effectiveness, a good 3D searching system should also exhibit high selectivity (the ratio of active hits to all hits) which can be influenced by query definition. In query definition, 3D searching is undergoing the development of atom-based query definition to function-based query definition. The notion of chemical environment in ALADDIN[10b] is an early effort in this direction. However, of commercial systems hitherto, only Catalyst supports strictly the function-based query definition. In searching strategy, a significant improvement is that confirmational flexibility of a database structure is taken into account,[17−20] i.e., the development from rigid searching to flexible searching. A number of approaches have been suggested to address conformational flexibility problem; for example, ChemDBS-3D uses the approach of conformational analysis at registration and search time, and SYBYL-3DB Unity uses the directed tweak approach. However, though with different flexible searching approaches, most 3D searching systems adopt almost the same searching strategy to reduce the number of candidate compounds before flexible searchings, i.e., a prescreen based on a key scheme and a geometric search using the Ullmann algorithm.[21]

We also developed a 3D searching system-3DFS in our laboratory. Besides simple atom-based query, 3DFS supports generalized function-based query. In addition, 3DFS uses a set of new searching strategy different from those in other 3D searching systems. The query definition, query file, and searching strategy in 3DFS will be described in detail in this paper.

2. QUERY DEFINITION

In most 3D searching systems, the elements of a query pharmacophore are defined as actual atoms or function groups such as a nitrogen atom or a carbonyl group. As known to us, a basic nitrogen atom usually acts as a positive charge center, and a carbonyl oxygen usually plays a role of hydrogen bond acceptor in the ligand−receptor interaction. If a positive charge center or hydrogen bond acceptor is used as the query element instead of nitrogen atom or carbonyl group, 3D searching will retrieve more functionally equivalent but structurally diverse hits because a positive charge center or hydrogen bond acceptor can represent much more groups which serve the same binding function to a nitrogen

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1997.

S0095-2338(97)00070-X CCC: $15.00
Published on Web 01/19/1998

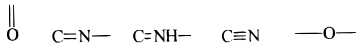


Figure 1. Definition of hydrogen bond acceptors.

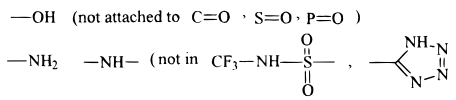


Figure 2. Definition of hydrogen bond donors.

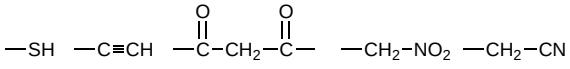
or carbonyl group. So, it is necessary to introduce the generalized binding function definition into query.

3DFS considers four important ligand−receptor interactions, hydrogen bonding, charge interaction, hydrophobic interaction, and $\pi-\pi$ interaction, and thus defines six binding sites as the superelements of query. They are hydrogen bond acceptor, hydrogen bond donor, positive charge center, negative charge center, hydrophobic region, and aromatic ring center.

Greene[22] has proposed the detailed definition for hydrogen bond acceptors/donors, charge centers, and hydrophobe. On the basis of Greene's studies, we propose simpler and more practical definitions for the above superelements, especially for hydrophobe.

2.1. Hydrogen Bond Acceptor and Donor. Generally, any nitrogen, oxygen, fluorine, or sulfur atom with at least one nondelocalized lone pair can be an acceptor, but a too generalized definition could result in an overload of hits, which may reduce the selectivity of hits. Accordingly, 3DFS only considers some nitrogen or oxygen atoms which often appear in bioactive molecules as acceptors. As shown in Figure 1, they include 1. oxygen with double bond, 2. nitrogen attached to carbon with double bond or triple bond, and 3. ether oxygen.

Hydrogen bond donor atoms are primarily oxygen or nitrogen atom with one or two electropositive hydrogen atoms. As shown in Figure 2, they include: 1. hydroxyls not attached to C═O, S═O, or P═O group and 2. aminos that are not part of a trifluoromethylsulfonamide or tetrazole moiety. Weaker groups such as



can also be added to the definition if necessary.

2.2. Charge Center. An atom with a formal charge is a charge center, but a neutral moiety might also be considered as charge center if it would be ionized at physiological pH. For example, an aliphatic amine can be protonated to be a positive charge center, whereas a carboxylic acid can be deprotonated to be a negative charge center at physiological pH. π-Delocalized systems such as carboxylate, guanidine, and amidine and even their substitutes can also make charge centers.

The positive charge centers (Figure 3) include 1. an atom with a formal positive charge, 2. the nitrogen in aliphatic amines, 3.the imino nitrogen of N,N-disubstituted amidines or N,N,N,N-tetrasubstituted guanidines, and 4. the centroid of nitrogens in guanidines bearing at least one hydrogen on each amino nitrogen or in amidines bearing at least one hydrogen on amino nitrogen.

The negative charge centers (Figure 4) include 1. an atom with a formal negative charge, 2. the nitrogen in trifluoromethylsulfomamide, 3. the centroid of the oxo and

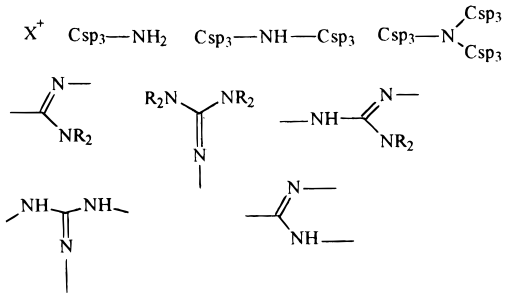


Figure 3. Definition of positive charge centers.

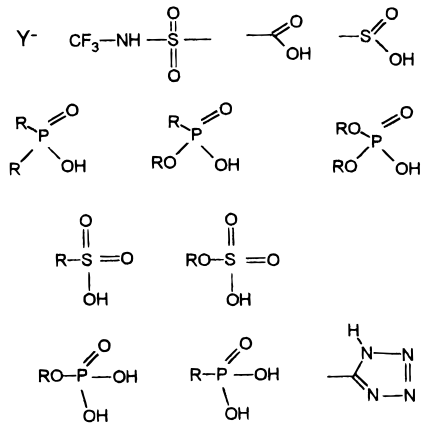


Figure 4. Definition of negative charge centers.

hydroxyl oxygens in carboxylic, sulfinic, or phosphinic acid, 4. the centroid of the oxo and hydroxyl oxygens in phosphoric diesters or phosphonic ester, 5. the centroid of the two oxo oxygens and the hydroxyl oxygen in sulfuric or sulfonic acid, 6. the centroid of the oxo oxygen and the two hydroxyl oxygens in phosphoric monoester or phosphonic acid, and 7. the amino nitrogen in a non-N-substituted tetrazole.

2.3. Aromatic Ring Center. Aromatic ring centers are primarily the centroids of five- or six-membered aromatic rings such as triophene and benzene ring. These aromatic rings participate in $\pi-\pi$ interaction with the π systems of a protein receptor. 3DFS system can automatically identify the aromatic rings in a database structure by its 2D substructure matching algorithm with such a rule that all carbon atoms in an aromatic ring must have aromatic, double, or triple bonds, while heteroatoms may have any bond order.

2.4. Hydrophobic Region. 3DFS uses a special algorithm to identify hydrophobic regions in a database structure, since it is almost impossible to list all hydrophobic fragments just as we do for hydrogen bond acceptors/donors and charge centers. We employ the assumption that the hydrophobicity of a hydrophobic region is the sum of atomic contributions[23] and classify hydrophilic atoms into ten types (Table 1). Our hydrophobe recognition algorithm is as follows.

Firstly, we mark all the hydrophilic atoms according to Table 1 and then cut off all the bonds connected to these atoms. If we regard the database structure as a graph, then at this time the graph is cut into many subgraphs disconnected to each other. In traversing every subgraph to numerate its non-hydrogen atoms, if the number of atoms is not less than three, which ensures that a hydrophobic region has enough surface area, then the subgraph is considered as a hydrophobic region. The position of a hydrophobe is

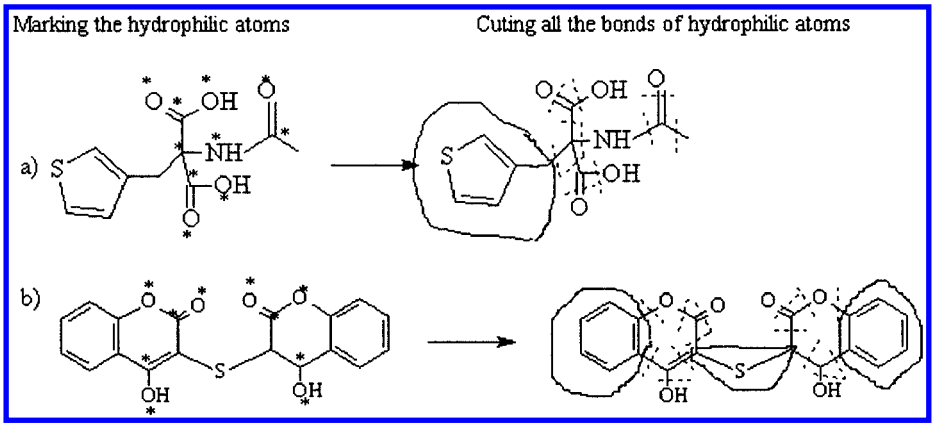


Figure 5. Two Examples of recognizing hydrophobic region. The atoms with "* are hydrophilic atoms and the regions circled by solid curve are hydrophobes.

Table 1. Hydrophilic Atom Types

type		description
1		N, O
2		S in SH
3		S with double bond
4		≤1 bond away from charged atom
5		≤1 bond away from OH, NH, or NH2 with no delocalized electrons
6		≤1 bond away from SH with no delocalized electrons
7		≤1 bond away from O with double bond
8		≤1 bond away from S with valence > 2
9		>one neighboring O or N with no delocalized electrons
	A	two bonds away from O with double bond
	B	one bond away from S with valence > 2
10		two or more intances of any of the previous two (A, B) conditions

Table 2. Query File of HIV-1 PR Inhibitor Shown in Figure 6

```
>ATOMS   4
   1   OH
   2   OH
   3   C
   4   O
>BONDS   1
 3   4   2
>DISCON   3
>START    3
 1
 2
 3
>DISTANCE CONSTRAINT   3
 1   4   5.4   1.0
 2   4   5.1   1.0
 1   2   2.8   1.0
```

represented by the geometric center of atoms in the hydrophobe.

This method is simple and rapid. Figure 5 shows two examples of identifying the hydrophobic regions. The atoms with an "* are hydrophilic atoms according to Table 1, and structure a) has one hydrophobic region (circled by solid curve) and structure b) has three hydrophobic regions (circled by solid curve).

2.5. Spatial Constraints in Query. 3DFS supports three spatial constraints commonly appearing in pharmaphore patterns: distance constraint, angle constraint, and dihedral constraint with an allowed tolerance.

2.6. Query File. A pharmacophore usually consists of several disconnected fragments and their spatial relationships, and a fragment may contain a single atom or some atoms connected with each other by bonds. 3DFS allows at most six disconnected fragments in a query pharmacophore.

Hence, a query file can be divided into two parts: connection table block and spatial constraint block, including nine items as follows:

>ATOMS	Provides the number of atoms and atom list. One can specify more than one atom type per atom, and one must specify the implicated hydrogen if the implicated hydrogen of an atom must be matched.
>BONDS	Provides the number of bonds and bond list.
>CENTER	Provides the number of cycle centers and the ring atom lists if there is any cycle center in pharmacophore.
>DISCON	Specifies the disconnectivity of pharmacophore (the number of disconnected fragments.)
>START	Specifies the starting atom in every disconnected fragment.
>DISTANCE CONSTRAINT	Specifies the number of distance constraints and constraint list.
>ANGLE CONSTRAINT	Specifies the number of angle constraints and constraint list.
>LINE_ANGLE CONSTRAINT	Specifies the number of angle constraints between planes or direction vectors and constraint list.
>DIHEDRAL CONSTRAINT	Specifies the number of dihedral constraints and constraint list.

In order to conveniently represent the superelements in query file, 3DFS gives each a symbol just like the element symbol of an actual atom: Hr, Hd, Pc, Nc, Hy, Pi, representing hydrogen bond acceptor, hydrogen bond donor, positive charge center, negative charge center, hydrophobic region, and aromatic ring center, respectively.

For example, Table 2 lists the query file of HIV-1 protease inhibitor pharmacophore shown in Figure 6 . If we regard two hydroxyl groups as hydrogen bond donors and carbonyl group as a hydrogen bond acceptor, the function-based

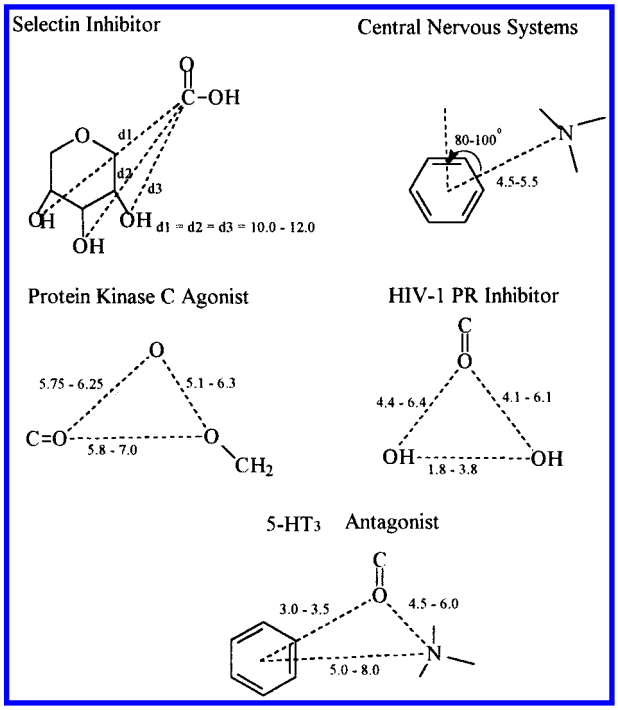


Figure 6. Queries used for performance measurements. Distances are specified in Å.

Table 3. Function-Based Query File of HIV-1 PR Inhibitor

```
>ATOMS   3
   1   Hd
   2   Hd
   3   Hr
>BONDS   0
>DISCON   3
>START   3
 1
 2
 3
>DISTANCE CONSTRAINT   3
 1   3   5.4   1.0
 2   3   5.1   1.0
 1   2   2.8   1.0
```

definition query file can be represented as shown in Table 3.

3. SEARCHING STRATEGY

After the query file is input into the system, the search starts. The search consists of four steps: 1. 1D screening, 2. 2D substructure searching, 3. rigid 3D searching, and 4. conformationally flexible searching. Each of these steps will be described in detail.

3.1. 1D Screening. 1D screening is a rapid prescreen to eliminate database structures which cannot possibly satisfy the query. This step compares only types and numbers of atoms in a database structure with that in the query, so are called 1D screening. For example, if a query contains two oxygen atoms and two nitrogen atoms, then only the database structures which contain at least two oxygen atoms and two nitrogen atoms can pass the 1D screening. For the super-element in the query, the process is similar: 1. a hydrogen bond acceptor/donor means an oxygen or nitrogen atom, 2. a positive charge center means an atom bearing a formal positive charge or a nitrogen atom, 3. a negative charge center means an atom bearing a formal negative charge or an oxygen or nitrogen atom, 4. a hydrophobic region means at least one carbon atom, 5. an aromatic ring center means any atom type.

The 1D screening in 3DFS is much rougher than the key screening used in other 3D searching systems, but we think it is enough, for two reasons: 1. most pharmacophores contain little quantity of atoms and simple structures, it is unnecessary to constructing complex key set (for example, ISIS/3D uses 962 keys in key screening), and 2. the importance of prescreen has been decreased because subsequent atom-by-atom substructure searching has obtained a significant increase in speed by the use of effective algorithm and powerful computer.

3.2. 2D Substructure Searching. The database structure passing the 1D screening needs an exact substructure searching to check whether its atoms interrelate as defined in query, i.e., whether the query is a substructure of the database structure.

Most 3D searching system use Ullmann algorithm[21] for substructure searching, but the 3DFS system uses the GMA algorithm proposed by Xu[24,25] for substructure searching.

The GMA algorithm is a partial-ordering-based back-tracking substructure searching algorithm. It consists of two steps.

Step 1. Reorder the query graph (QG) by a depth-first traversal procedure to get a Partial Order Set (POS) of QG, i.e., a traversal route on QG.

Step 2. Use POS as an instruction set to walk on target graph (TG), this is a constrained back-tracking procedure. If the node in TG matches the route node in POS, the walk continues, otherwise, it tracks back to the last matched node and selects an alternative walk direction on TG. If all route nodes in POS find matched nodes in TG, i.e., the walk is complete, then QG and TG are homomorphic or isomorphic.

Because the match procedure is carried out under the direction of POS, the GMA algorithm is also called directed match algorithm. The match route in the GMA algorithm is clearer and more perceptual than that in the Ullmann algorithm used in many commercial substructure searching systems. Moreover the GMAs computing complexity is much less than the factorial computing complexity. (More details are provided by ref 25.)

Once a 2D match mapping is found, then the match mapping is submitted to subsequent rigid 3D searching.

3.3. Rigid 3D Searching. This step checks whether the atoms in the match mapping meets the spatial constrains (distance, angle, dihedral) in query. The word "rigid" means that only the stored conformation of the database structure is checked.

If all the differentials between constraint values and measured values are within the corresponding tolerances, the structure is declared as a hit, otherwise, the database structure must undergo a further, final conformationally flexible searching.

3.4. Conformationally Flexible Searching. In general, the database stores only a single low-energy conformation (or a limited number of such conformations) for each molecule. Accordingly, a rigid searching is likely to fail to identify a large number of matching molecules that can adopt a conformation containing the query pattern but that are represented in the database by a low-energy conformation that does not contain this pattern.

Table 4. Searching Results for Five Typical Queries Shown in Figure 6

	Selectin		CNS		PKC		HIV-1		5-HT3	
	hits	time (s)	hits	time (min)	hits	time (min)	hits	time (min)	hits	time (min)
rigid	0	482	2379	14	7	1 l	571	9	24	17
flexible[a]	17	505	9809	89	505	55	2754	151	974	128
flexible[b]	3	587	5230	220	232	102	1735	369	343	222
					Adding Hydrophobe					
rigid					3	12	379	12		
flexible[a]					394	52	2102	152		
					Function-Based Query					
rigid									179	294
flexible[a]									3628	12[c]

[a] Without bump check. [b] With bump check. [c] This was measured in hours not minutes.

The conformational space is the torsional space of the rotatable bonds in a flexible molecule. We define a rotatable bond as an acyclic single bond except for the terminal single bond and use the conventional POWELL[26] optimization method to search the torsional space.

The optimization is to minimize the root mean square (RMS) deviation between constraint values in query and fitted-model measured values by means of rotating the rotatable bonds in the database structure. During this process, if each differential between constraint value and measured value is within the corresponding tolerance, the RMS is set to zero, and the optimization completes successfully, then the structure is declared as a hit. Otherwise, the optimization continues until the maximal iteration number MAX_ITERS is reached.

If the optimization fails, the search will return to 2D substructure searching for an alternative 2D match mapping and then repeat rigid 3D searching or flexible searching, until a match mapping satisfying the spatial constraints is found, or no alternative 2D match mapping exits, or maximal match mapping number MAX_MAPPINGS is reached. In order to try as many match mappings as possible, all match mappings must be nonredundant to each other.

In the process of optimization, a van der Waals energy calculation can be optionally included in order to void hit conformations with unfavorable steric interactions. The vdw energy is calculated using the Lennard-Jones 6-12 potential, and the vdw interactions between atoms which are separated by one, two, or three bonds are ignored. If the energy difference between the candidate conformation and the reference conformation stored in the database exceeds CUTOFF, a user-specified parameter, then the conformation is penalized by a prespecified amount. The addition of a vdw energy calculation ensures that hit conformations not only match the spatial constraints but also are of low steric energy; however, in the mean time, the vdw calculation can result in considerable increase in CPU time for the search. This can be seen in the latter searching examples.

Though a directed tweak method was concluded by Clark *et al.*, the algorithm of choice for flexible searching after the detailed comparison of five different algorithms (systematic search, random search, distance geometry, genetic algorithm, and directed tweak), this method has an inherent disadvantage of being suitable only for distance constraints not for other constraints (e.g., angle, dihedral constraints). The POWELL method is by far the most effective nonderivative optimization method. Moreover, the method is easily engineered to accommodate new types of constraints and evaluation functions.

4. PERFORMANCE MEASUREMENT

The 3DFS system is programmed in C and run on an IBM PC Pentium/90MHZ. We select five typical pharmacophore examples shown in Figure 6 to illustrate the utility of 3DFS system. The five queries are selectin inhibitor pharmacophore,[27] central nervous system pharmacophore,[19] HIV-1 protease inhibitor pharmacophore,[15] protein kinase C agonist pharmacophore,[13] and 5-HT_3 antagonist pharmacophore,[28] respectively. All the searches are conducted on the NCI-3D database (a set of 126 705 compounds collected at the National Cancer Institute). The two important parameters which affect the search time and search yield, the maximal match mapping number MAX_MAPPINGS and the maximal iteration number MAX_ITERS in the POWELL method, are set to 10 and 5, respectively. If bump-checking is desired, the CUTOFF is set to 5 kcal/mol.

The results of searches are shown in Table 4. For the five queries, the times of rigid searchings are 8 (482 s), 14, 9, 11, and 17 min, respectively. Considering that it takes about 458 s (about 8 min) only reading the 126 705 compounds in database and computing the implicated hydrogen of each atom (since this information is not stored in the 3D database), the searches are very fast. The search performance varies widely with the query. Generally, the more complex the query, the few the hits. Of the five queries, the selectin inhibitor pharmacophore is most complex, so the hit number is least. For each query, flexible searching identifies many more hits than rigid searching but takes more time. In the PK-C example, the search time for flexible searching takes 5 times as long as rigid searching, returning 72 times as many hits. In the selectin inhibitor example, the increase of search time is least, only 23 s, while the number of hits increases from 0 to 17.

The following is the further investigations and discussions for the latter two queries.

Protein Kinase C Agonist Query. Wang and coworkers[13] discovered five compounds possessing PK-C binding affinity in the low micromolar range by flexible search over the open NCI database of 206 876 (including the 126 705 compounds in the database used by 3DFS) with the query shown in Figure 6. Their search took 40 h on a Silicon Graphics IRIS 3000 and found 535 hits (without bump check) and the searching software is Chem-X.

However, Table 4 shows that our search for this query takes only 55 min and obtains 505 hits. When we increase the MAX_MAPPINGS to 50, the search time increases to 100 min and the hit number increases to 554. This illustrates that the 3DFS system has higher efficiency and effectiveness than Chem-X.

In addition, any hit which has no hydrophobic moiety—this being requisite for PK-C affinity, can be discarded automatically and conveniently by adding a hydrophobe (Hy) to the query in 3DFS system. The modified query yields 3 rigid hits and 394 flexible hits in the search time of 12 and 52 min (Table 4, without bump check), respectively. However, Wang and a co-worker had to do this manually because Chem-X does not support generalized binding function definition at that time.

The addition of generalized hydrophobic region definition not only enhances the selectivity of search results but also finds structurally diverse hits with hydrophobic property. According to Wang's studies,[15] a hydrophobe is also requisite for HIV-1 PR inhibitors, so, a hydrophobe is added into the HIV-1 PR query, and the search result is also shown in Table 4.

5-HT$_3$ Antagonist Query. Hibert et al.[28] suggest that the carbonyl oxygen is serving as a hydrogen bond acceptor, that the basic nitrogen is serving as a positive charge center, and the phenyl ring is serving as an aromatic ring center. Accordingly, a function-based query can be defined by substituting a hydrogen bond acceptor, a positive charge center, and an aromatic ring center for the oxygen, nitrogen, and phenyl ring, respectively. Searching the same database with this query yields 179 rigid searching hits and 3628 flexible searching hits (Table 4). The function-based query finds 155 additional rigid hits and 2654 additional flexible hits missed by the atom-based query because of the generality of function-based query. The hits resulted from atom-based query contain some common substructures, at least one phenyl ring, one carbonyl group, and one nitrogen atom, though their backbones may be different, while the hits resulted from the function-based query may be completely different structures without any common substructures. Accordingly, a function-based query can give a greater chance to find structurally novel compounds.

5. CONCLUSION

3DFS is a computer program which searches a 3D database for compounds matching a given pharmacophore query. Its characteristics lie in two aspects: 1. using a set of practical binding site definitions and a rapid hydrophobe recognition algorithm for the function-based query and 2. using a set of effective searching algorithm different from those used in other 3D searching systems. By virtue of both of above, 3DFS demonstrates high utility.

ACKNOWLEDGMENT

The authors would like to acknowledge the National Cancer Institute for providing 3D database used in this work.

REFERENCES AND NOTES

(1) Martin, Y. C. 3D Database Searching in Drug Design. *J. Med. Chem.* **1992**, *35*, 2145−2154.

(2) Pearlman, R. S. In Kubinyi, H., Ed.; 3D QSAR in Drug Design. Theory, Methods and Applications; Leiden: ESCOM Science Publishers, 1993; p 41.

(3) Willett, P. Searching for Pharmacophoric Patterns in Database of Three-Dimensional Chemical Structures. *J. Mol. Recogn.* **1995**, *8*, 290−303.

(4) Finn, P. W. Computer-based Screening of Compound Databases for the Identification of Novel Leads. *DDT* **1996**, *1*, 363−370.

(5) Rusinko, III, A.; Skell J. M.; Balducci, R.; McGarity C. M.; Pearlman, R. S. Concord: A Program for the Rapid Generation of High Quality Approximate 3-Dimensional Molecular Structures; The University of Texas at Austin and Tripos Associates: St. Louis, MO, 1988.

(6) Davies, K.; Dunn, D.; Upton, R. An Algorithm to Generate 3D Structures from 2D Connection Tables; Poster on the 5th Molecular Modeling Workshop; Darmstadt, 1991.

(7) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic Generation of 3D Atomic Coordinates for Organic Molecules. *Tetrahedron Comput. Method.* **1992**, *3*, 537−547.

(8) Gund, P. Three-dimensional pharmacophore pattern searching. *Prog Mol. Subcell. Biol.* **1977**, *5*, 117−143.

(9) Sheridan, R.; Nilakantan, R.; Rusinko, III, A.; Bauman, N.; Haraki, K. S.; Venkataraghavan, R. 3DSEARCH: A System for Three-Dimensional Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 255−260.

(10) Van Drie, J.; Weininger, D.; Martin, Y. ALADDIN: An Integrated Tool for Computer-Assisted Molecular Design and Pharmacophore Recognition from Geometric, Steric and Substructure Searching of Three-Dimensional Molecular Structures. *J. Comp.-Aided Mol. Design* **1989**, *3*, 225−251. (b) Though the query defintion in ALADDIN system is superior to the simple atom-based query defintion, it is significantly different from the notion of function-based query in our paper, at most it is an early effort in the development direction of atom-based query to function-based query.

(11) Fisanick, W.; Cross, K. P.; Forman, J. C.; Rusinko, III, A. Experimental System for Similarity and 3D Searching of CAS Registry Substances. 1. 3D Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 548−559.

(12) Bures, M. G.; Martin, Y. C.; Willett, P. Searching Techniques for Databases of Three-Dimensional Chemical Structures. *Topics Stereochem.* **1994**, *21*, 467−511.

(13) Wang, S.; Zaharevitz, D. W.; Sharma, R.; et al. The Discovery of Novel Structurally Diverse Protein Kinase C Agonists Through Computer 3D Database Pharmacophore Search. Molecular Modeling Studies. *J. Med. Chem.* **1994**, *37*, 4479−4489.

(14) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; et al. Rational Design of Potent, Bioavailable, Nonpeptide Cyclic Ureas as HIV Protease Inhibitors. *Science* **1994**, *263*, 380−384.

(15) Wang, S.; Milne, G. W. A.; Yan, X.; et al. Discovery of Novel, Non-Peptide HIV-1 Protease Inhibitors by Pharmacophore Searching. *J. Med. Chem.* **1996**, *38*, 2047−2054

(16) Kiyama, R.; Homma, T.; Hayashi, K.; et al. Novel Angiotensin II Receptor Antagonists: Design, Synthesis and In Vitro Evalution of Dibenzo[a,d]cycloheptene and Dibenzo[b,f]oxepin Derivatives. Searching for Bioisosteres of Biphenylyltetrazole using a Three-Dimensional Search Technique. *J. Med. Chem.* **1995**, *38*, 2728−2741.

(17) Murral, N. W.; Davies, E. K. Conformational Freedom in 3D Databases. 1. Techniques. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 312−316.

(18) Hurst, T. Flexible 3D searching: The Directed Tweak Technique. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 190−196.

(19) Moock, T. E.; Henry, D. R.; Ozkabak, A. G.; Alamgir, M. Conformational Searching in ISIS/3D Databases. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 184−189.

(20) Clark, D. E.; Jones, G.; Willett, P. Pharmacophoric Pattern Matching in Files of Three-Dimensional Chemical Structures: Comparison of Conformational-Searching Algorithms for Flexible Searching. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 197−206.

(21) Ullmann, J. R. An Algorithm for Subgraph Isomorphism. *J. Assoc. Comput. Machinery* **1976**, *23*, 31−42.

(22) Greene, J.; Kahm, S.; Hamid, S.; Sprague, P.; Teig, S. Chemical Function Queries for 3D Database Search. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1297−1308.

(23) Ghose, A.; Crippen, G. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships 1. Partition Coefficient as a Measure of Hydrophobicity. *J. Comput. Chem.* **1986**, *319*, 199−203.

(24) Xu, J.; Zhang, M. HBA: New algorithm for structural match and applications. *Tetrahedron Comput. Methodol.* **1989**, *2*, 75−83.

(25) Xu, J. GMA: A Generic Match Algorithm for Structural Homomorphism, Isomorphism, and Maximal Common Substructure Match and Its Applications. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 25−34.

(26) Powell, M. J. D. An Efficient Method for Finding the Mininum of a Function of Several Variables without Calculating Derivatives. *Comput. J.* **1964**, *7*, 155−162.
(27) Rao Narasinga, B. N.; Anderson, M. B.; et al. Sialyl Lewis X Mimics Derived from a Pharmacophore Search are Selectin Inhibitors with Anti-inflammatory Activity. *J. Bio. Chem.* **1994**, *26*, 19663−19666.
(28) Hibert, M.; Hoffmann, R.; Miller, R.; Carr, A. Conformation-Activity Relationship Study of 5-HT3 Receptor Antagonists and a Definition of a Model for this Receptor Site. *J. Med. Chem.* **1990**, *33*, 1594−1599.

CI970070Y