A Web-Based 3D-Database Pharmacophore Searching Tool for Drug Discovery

Xueliang Fang and Shaomeng Wang*

Departments of Internal Medicine and Medicinal Chemistry, University of Michigan, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-0934

Received August 14, 2001

Three-Dimensional (3D) structural database pharmacophore searching has become a very effective approach for discovery of novel lead compounds in drug discovery. Although several commercial programs are available, these commercial programs are primarily used as a stand alone and require a local database. In recent years, the Internet has become the main medium of choice for multiuser application program distribution. Herein, we describe our development of a Web-based 3D-database pharmacophore-searching tool based on the server-client Web architecture. Both rigid and conformationally flexible searching methods are implemented. Our results show that for a typical three-center rigid pharmacophore search, the run time for searching 50 000 compounds is less than three minutes, and for four-center pharmacophore searching, the run time is less than 10 minutes on a desktop computer. For a flexible 3D-pharmacophore search, the run time for searching 50 000 compounds generally takes between one and several hours. The search results are comparable to those obtained using a commercial program. We expect that this Web-based tool will be very useful for scientists who are interested in 3D-database pharmacophore searching via the Internet.

1. INTRODUCTION

In recent years, three-dimensional (3D) structural database searching has been widely used in drug discovery and design.[1−4] Gund[5] was probably the first who described that functional groups (pharmacophores) could be used for searching databases to identify molecules that may share the same structural features. This has led to the successful development and application of 3D-database pharmacophore searching for discovering novel drug lead compounds in drug discovery. Indeed, several commercial programs are now available for 3D structural database searching, including 3D-pharmacophore searching. However, most currently available commercial programs are primarily used as standalone program packages and require local databases.[6−11]

In recent years, the Internet has become the main medium of choice for multiuser application program distribution. The major advantages for Web-based database searching are that users can search the most recent data without regularly downloading and updating the whole database to their local disks. Also, there is no need to install additional software on the user's computer except a standard Web browser. Herein we describe our development of a Web-based tool for carrying out 3D database pharmacophore searching.

Since 1955, the National Cancer Institute has collected approximately 500 000 compounds for its anticancer drug discovery efforts. Out of the 500 000 compounds, approximately 250 000 are classified as "open" compounds, whose structural and biological data are accessible to the general public. We use the NCI "open" database to demonstrate the use of our Web-based tool for carrying out 3D-database pharmacophore searching in this paper.

* Corresponding author phone: (734)615-0362; fax: (734)647-9647; e-mail: shaomeng@umich.edu.

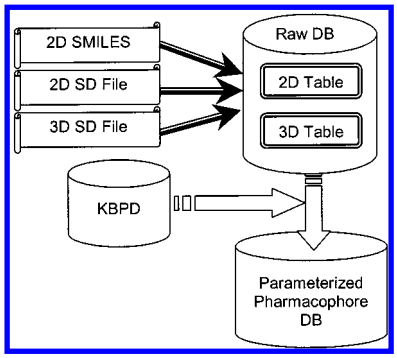

Figure 1. Flowchart of pharmacophore database construction.

2. DATABASE CONSTRUCTION

We choose to use the ORACLE relational database and the Internet Server Application Programming Interface (ISAPI) technology to perform both 2D and 3D database searching via the Internet.

The construction of the database is crucial, both for its content and its search speed. The NCI 2D database contains the SMILES[12−14] codes and 2D coordinates, while the NCI 3D database includes spatial 3D coordinates of each atom, and the connectivity of each bond for every compound, as generated by the Corina program.[15,16] For the purpose of searching efficiency, we have generated a parametrized pharmacophore database[17,18] (PPDB) from the original databases (see Figure 1) by applying a knowledge based conversion program. This parametrized database is a relational database management system (RDBMS) and is managed by ORACLE on a Windows NT server. This kind of database is very suitable for pharmacophore-based searching for multiple users via the Internet. The table space of this parametrized NCI "open" database of 250 000 compounds is about 5 gigabytes. The most important parameters in the

10.1021/ci010083i CCC: $22.00
Published on Web 02/26/2002

parametrized database include functional group centers and class definitions. These parameters play a role for screening out structures and improving the searching efficiency during 2D structure searching. For this purpose, we introduce the knowledge-based pharmacophore definition (KBPD).

2.1. Pharmacophores Description. A pharmacophore is a description of the essential functional groups in a molecule for a specific biological target of interest, together with their 3D geometric relationship. Often the centers of these essential groups are used for specifying the geometry of a pharmacophore. Typically, six standard types of groups are defined in the context of 3D pharmacophore searching. The six groups are hydrogen bond acceptor, hydrogen bond donor, positive charge center, negative charge center, hydrophobic group, and aromatic ring center. Note that these groups are not mutually exclusive. For example, an aromatic ring center can also be a hydrophobic center. Negatively charged centers are sometimes incorporated as part of the hydrogen bond acceptors.

There are two additional center types for the ACID and BASE special atom types. There are also subsets for the donor and acceptor centers. The definition of these pharmacophore center types are described in more detail in previous publications by other investigators.[19−22]

With the exception of hydrophobic centers, these center types can be automatically derived from the atom types in the molecule. Correct atom type assignments are therefore crucial. To be able to define the hydrophobic centers in a molecule, the hydrophobic regions need to be determined. First, the molecular structure is split into many possible fragments by cutting off all the bonds attached to hydrophilic atoms, such as O, N, and other specified elements.[22] The logarithm of the partition coefficient between *n*-octanol and water (log*P*) has been widely used to assess of the hydrophobicity of a molecule or a molecular fragment in drug design. Thus, the log*P* is calculated for each fragment to assess the hydrophobicity of the fragment.[23] We use the XLOGP[24] program to calculate molecular hydrophobicity of a molecular fragment. The XLOGP program is a simple, atom-based method with reasonable accuracy and efficient speed.

Generally, the center of a hydrogen bond donor or acceptor is defined as an actual atom such as an oxygen or a nitrogen atom as well as a charge center. However, for a hydrophobic region or an aromatic ring, the center is defined as the centroid of the group. The geometric coordinates (x_c, y_c,z_c) of the centroid of a group is calculated as follows

$$x_c = \frac{\sum_{i=1}^{N} x_i M_i}{\sum_{i=1}^{N} M_i},\ y_c = \frac{\sum_{i=1}^{N} y_i M_i}{\sum_{i=1}^{N} M_i},\ z_c = \frac{\sum_{i=1}^{N} z_i M_i}{\sum_{i=1}^{N} M_i}$$

where N is the number of atoms, (x_i, y_i, z_i) are spatial coordinates of atom i.

We use three different definitions for M_i as follows: (1) $M_i = 1$ for aromatic rings. Only geometric coordinates of atoms are considered; (2) for hydrophobic regions, M_i is the atomic hydrophobic value of atom i, taken from log P handbooks; and (3) for other function groups, $M_i = r_i^3$, r_i is the vdW radius of atom i.

Table 1. Information Stored in the Database

information description
identification serial numbers
NSC serial numbers
molecular weights
connectivity, including bond order and type
atom types and names
atom serial and group numbers
possible pharmacophore center definitions
2D SMILES strings
2D and 3D coordinates

So the pharmacophore distance between each pair of centers (i, j) is expressed as

$$d_{i,j} = (x_{ci} - x_{cj})^2 + (y_{ci} - y_{cj})^2 + (z_{ci} - z_{cj})^2$$

In some cases, dihedral angles, torsion angles, angles between planes or vectors, vector magnitudes, and distances of points above a plane are also used to define a pharmacophore and are used in pharmacophore searching.

2.2. Database Configuration. The database comprises a series of tables, indexes, and clusters. The main information stored in the database is summarized in Table 1. The pharmacophores are stored in the database using a serial of bit strings to describe the pharmacophoric types, group centers/centroids, and other information. In this database, unique keys (identification serial number and NSC serial number) are the entry ID used for locating the compounds quickly. Each table contains these fields.

3. DESCRIPTION OF SEARCHING STRATEGY

The pharmacophore searching process consists of the following steps: (1) 1D screening; (2) 2D substructure matching; (3) 3D geometric searching (rigid searching); and (4) conformational searching (flexible searching). If flexible searching is performed, conformations are generated using the stored structure in the database when a rigid search fails to meet the spatial restrains in the query.

3.1. Rigid Searching. The following 1D screening is first performed in pharmacophore searching: (1) number of each pharmacophore presented in a compound; (2) molecular weight; and (3) atom types and bond connectivity.

Through 1D searching, we can often reduce the number of compounds from 250 000 to a subset of 20 000 compounds within one second.

2D searching compares pharmacophore type and atom-by-atom substructure, so topological information contained in the SMILES code is sufficient. Since the database contains detailed information about predefined pharmacophores and most pharmacophores contain only a few atoms and bonds, this procedure is also rapid. The 2D searching results in a much smaller subdatabase.

After the 2D searching, 3D spatial searching is subsequently carried out. The actual geometry of the query is used to define the query constrains explicitly. 3D searching checks whether the pharmacophores embedded in the compounds from the 2D searching meet the geometric constrains (e.g. distance, angel and dihedral) within certain tolerances specified in the pharmacophore query definition. However, since only one low-energy conformation for each compound

Table 2. Number of Rotatable Bonds in NCI Database

rotatable bonds	number of molecules	possible conformations[a]
0−5	107 319	~7776
6−15	123 289	$\sim 10^{11}$
16−50	16 301	$\sim 10^{38}$
51−100	298	$\sim 10^{77}$
101−165	18	$\sim 10^{128}$

[a] The rotation step is set to 6; a flexible bond must satisfy all of the following conditions: (1) single bond, (2) acyclic, and (3) not in a termination.

is stored in the database and searched, the rigid search is also very fast.

3.2. Flexible Searching. A flexible compound can have many low-energy conformations. When the compound binds to its target, it often changes its conformation. It is unlikely that the single conformation stored in the database for the compound is the bound conformation to the target.[30] For this reason, conformationally flexible searching is essential to identify the compounds that contain the pharmacophore in their low-energy conformations. Two generally used approaches are storing multiple conformations in the database or generating multiple conformations dynamically (on the fly) when performing the pharmacophore search.

In our implementation, we chose to generate multiconformations on the fly. For this purpose, the flexible bond information was stored in a relational table with the 3D pharmacophore database. We have analyzed the number of flexible bonds in NCI database. The results are provided in Table 2.

There are approximately 17 000 compounds having more than 15 rotatable bonds in the NCI database, and the most flexible compound contains 165 rotatable bonds (Table 2). Therefore, it is not possible to generate all conformations for these highly flexible compounds. In our implementation, we choose to generate conformations dynamically during the pharmacophore searching process. This strategy not only employs the usual rule-based sampling, which selects the low-energy conformations, but also applies any conformationally dependent geometric criteria as rejections during the regeneration. For example, flexible bonds that are closer to pharmacophore centers always have the priority to be rotated at first, because they often have a large effect on the geometric parameters in the pharmacophore. A systematic search procedure is employed.[25] The conformation generation will be stopped when a newly generated conformation is found which meets geometric parameters in the pharmacophore, or the maximal number set for generated conformations is reached (default value is set as 20 000). Furthermore, to avoid yielding unreasonable high-energy conformations, vdW energy for each generated structure is calculated and compared with the reference conformation stored in the database. The vdW energy is calculated using the L−J 6−12 potential[27] as follows

$$E_{vdw} = \sum_{i=1}^{N_{atoms}} \sum_{j>1} E_{ij} \left[\frac{1.0}{a_{ij}^{12}} - \frac{2.0}{a_{ij}^{6}} \right]$$

where

$$E_{ij} = \sqrt{E_i E_j},\ a_{ij} = r_{ij}/(R_i + R_j)$$

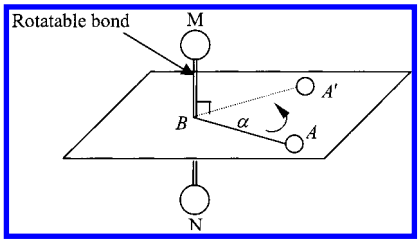

Figure 2. 3D coordinates conversion in conformation generation.

r_{ij} is the distance between atoms i and j. E_i is the vdW constant (in kcal/mol), and R_i is vdW radius (in Å) of ith atom.

To ensure the hit conformations are in low energy, the cutoff value for the energy difference between the candidate and the stored conformer is set to 5 kcal/mol. The vdW interactions between atom pairs that are separated by one, two, and three bonds are ignored since these are bonded terms in force field calculation. The emendatory coefficient for vdW interaction between 1 and 4 pair is set to 0.50.

The process of generating 3D coordinates for new conformations is illustrated in Figure 2. Chirality is not considered in conformation generation. Given the original coordinates A (x_0,y_0,z_0) of any atom A in the former conformation structure, coordinates of the two nodes (M, N) of flexible bond are expressed as M (x_1,y_1,z_1) and N (x_2,y_2,z_2). After rotation, the target coordinates A' (x,y,z) in the new conformation can be calculated using the following equations. MN is the vertical of plane ABA'.

The coordinates of the point of intersection can be calculated as B $(\bar{x},\bar{y},\bar{z})$, and

$$\begin{cases} \bar{x} = x_1 + k(x_2 - x_1) \\ \bar{y} = y_1 + k(y_2 - y_1) \\ \bar{z} = z_1 + k(z_2 - z_1) \end{cases}$$

where

$$k = \frac{(x_0 - x_1)(x_2 - x_1) + (y_0 - y_1)(y_2 - y_1) + (z_0 - z_1)(z_2 - z_1)}{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

Therefore, we can derive the following equations:

$$\begin{cases} (x - \bar{x})(x_2 - x_1) + (y - \bar{y}) + (y_2 - y_1) + (z - \bar{z}) + (z_2 - z_1) = 0 & (1) \\ (x - \bar{x})(x_0 - \bar{x}) + (y - \bar{y}) + (y_0 - \bar{y}) + (z - \bar{z}) + (z_0 - \bar{z}) = \rho \cos \alpha & (2) \\ (x - \bar{x})^2 + (y - \bar{y})^2 + (z - \bar{z})^2 = \rho & (3) \end{cases}$$

where $\rho = (x_0 - \bar{x})^2 + (y_0 - \bar{y})^2 + (z_0 - \bar{z})^2$, and α is the torsion angle. The target coordinates of A' can be calculated by resolving the quadratic equations.

If the rotation is anticlockwise (as described in Figure 1), eq 2 can be changed to

$$(x - \bar{x})(x_0 - \bar{x}) + (y - \bar{y}) + (y_0 - \bar{y}) + (z - \bar{z}) + (z_0 - \bar{z}) = \rho \cos \alpha \quad (2a)$$

Otherwise, if the rotation is clockwise, eq 2 can be changed to

$$(x - \bar{x})(x_0 - \bar{x}) + (y - \bar{y})(y_0 - \bar{y}) + (z - \bar{z})(z_0 - \bar{z}) = -\rho\cos\alpha \quad (2b)$$

In the present work, the rotation angles are set as (0, 60, 120, 180, −60, −120), for the main purpose is only focused on pharmacophore search, which usually involves some distance tolerances. A positive rotation angle always means an anticlockwise rotation and *vice versa*.

3.3. System Architecture Description. The system utilizes client-server Web architecture. Users interact with the system through their Web browser. Internet Server Application Programming Interface (ISAPI) is used for transferring data between the client and the server, while the ORACLE database is linked to this application by Oracle Objects for OLE (OO4O). (See Figure 3.) An authorized user can visit the user input interface on client browser via the Internet and specify the necessary information for pharmacophore searching, including pharmacophore types, geometrical constrains, corresponding tolerances, and dataset range.

The NCI "open" database is split into five smaller databases for searching efficiency and each subdatabase contains about 50 000 compounds. When a query form is submitted, the request is sent to the server. ISAPI opens a connection using OO4O. The ISAPI is programmed in C++, which makes it very fast and convenient to access the database. The search results are returned to the client. If users want to display the 2D or 3D structure of a hit, the Oracle database will be contacted using ActiveX Database Object (ADO), which is programmed in JAVA scripts. The Web server for this searching tool is Internet Information Service (IIS) 4.0 for Windows NT. The data are managed by Oracle 8.16, which is a relational database management system (RDBMS), so it is very flexible to access 2D, 3D, and pharmacophore tables using a serial of primary keys. Since the database and ISAPI reside on the server, client users do not need to install any additional software on their computers.

4. EXAMPLE OF PHARMOCOPHORE SEARCHING

Previously, 3D-database pharmacophore searching has been performed on the NCI 3D-database using the program Chem-X to discover novel HIV-1 protease inhibitors.[3] We wish to demonstrate the use of our Web-based pharmacophore searching tool through this example. The simple HIV-1 pharmacophore model used in the 3D search is shown in Figure 4. In this pharmacophore model, the carbonyl is defined as a hydrogen bond acceptor, and the two hydroxyls are defined as hydrogen bond donors.

The pharmacophore search of the NCI "open" database of 206 876 compounds using this pharmacophore model yielded 2368 hits.[3] Employing additional criteria, 50 compounds were ultimately submitted for biological testing, and 15 compounds were found to have ID_{50} values below 50 μM.

The pharmacophore is drawn using Java Molecular Editor[24] and translated into SMILES[12−14] for 2D substructure comparison.[25,26] The pharmacophore types for each center is specified in the model. After setting the geometrical constrains and tolerances, the query is built and then submitted to the server for pharmacophore searching.

4.1. Rigid Search Example. For pharmacophores of HIV-1 PR inhibitor defined in Figure 4, we discovered 2315 hits, while the Chem-X program found 2271 hits. (See Table 4.) The 2D and 3D structures of three rigid hit examples (NSC 115497, NSC 117027, and NSC 251156) are shown as Figure 5. The 2D structures are drawn using JAVA applet, and the 3D structures are plotted using VRML (Virtual Reality Modeling Language), and the distances between each two pharmacophoric centroids are automatically calculated and displayed.

We examined why there are certain discrepancies between the Chem-X results and our results. The NCI database has

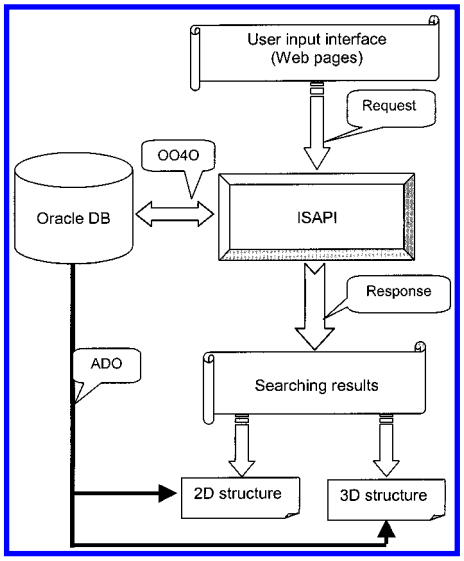

Figure 3. Architecture chart of Web-based pharmacophore searching tool.

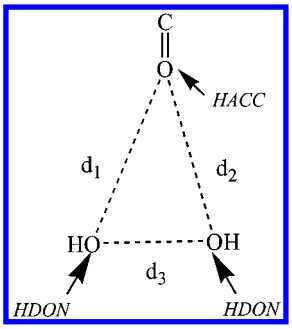

Figure 4. A pharmacophore model used for HIV-1 protease inhibitor searching ($d_1 = 5.4 \pm 1.0$ Å, $d_2 = 5.1 \pm 1.0$ Å, and $d_3 = 2.8 \pm 1.0$ Å).

Table 3. Activities[3] of HIV-1 Protease Inhibitor Discovered in Rigid and Flexible 3D Pharmacophore Search

NSC (active)	ID_{50} (μM)	rigid hit	flexible hit
7576	27.0	√	√
9152	2.0		√
20143	17.0		√
32180	0.32		√
40328	32.0		√
41234	21.0		√
109412	23.0		√
115497	26.0	√	√
122500	16.0	√	√
117027	0.75	√	√
117272	24.0	√	√
158393	1.7	√	√
251156	6.0	√	√
373937	30.0		√
382025	35.0		√
percentage		47%	100%

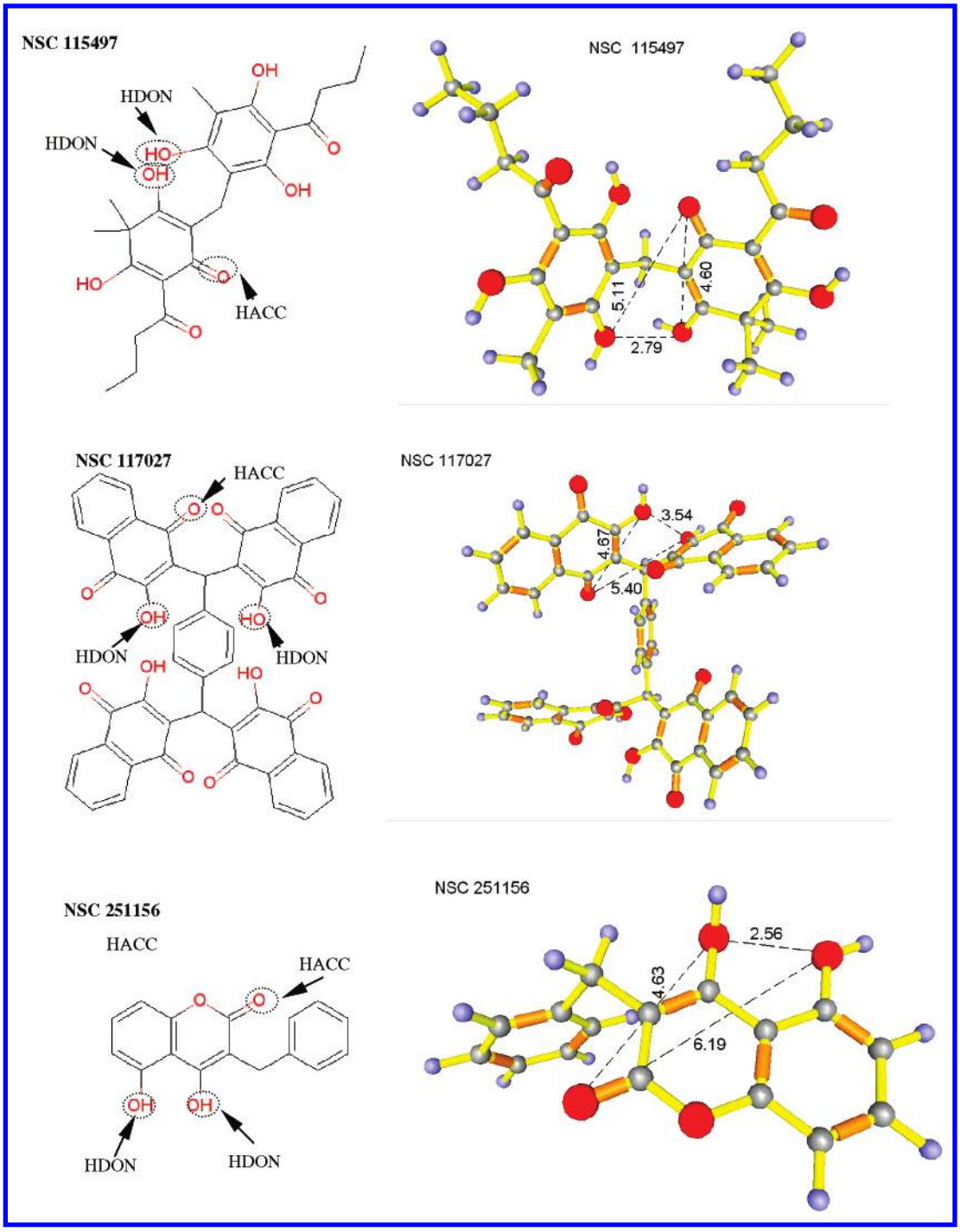

Figure 5. Structures of hits for HIV-1 protease inhibitors discovered in rigid search.

Table 4. Search Results for HIV-1 PR Inhibitor in Rigid Search

	rigid hits			time used (s)	
NSC range	Chem-X	present work	overlap (%)	Chem-X	present work
1−60 040	268	273	98	90	112
60 041−131 742	321	325	99	106	186
131 743−240 407	504	510	99	110	195
240 408−506 238	567	570	99	118	213
506 239−706 805	611	637	96	183	280

some compounds that contain more than one segment. The Chem-X program failed to process these compounds. Our Web-based searching tool can process these structures correctly and are able to identify a number of hits such as NSC 3384, NSC 31661, and NSC 35658. It is of note that the Chem-X program is somewhat faster than our Web-based searching tool.

4.2. Flexible Search Example. The flexible search results for HIV-1 PR inhibitor are summarized in Table 5. As expected, the flexible search takes a much longer time than the rigid search. While it takes approximately 1−3 h for the Chem-X program to search through 50 000 compounds, it takes 2−6 h for our Web-based tool. There is an good agreement between the search results obtained by these two programs (97−99%). The slight discrepancies between them may be primarily due to how the conformations are generated.

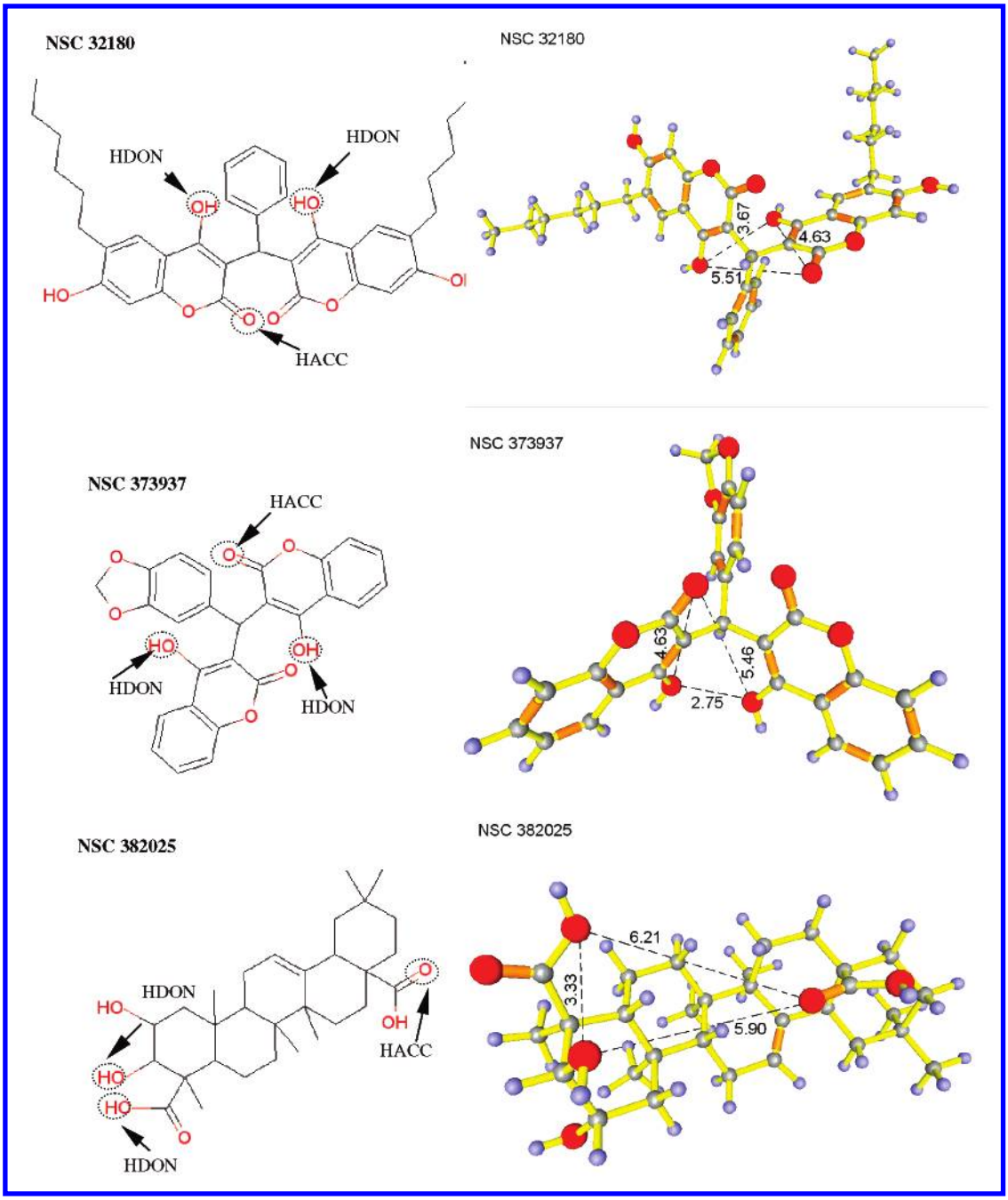

Figure 6. Structures of hits for HIV-1 protease inhibitors discovered in flexible search.

Table 5. Flexible Search Result for HIV-1 PR Inhibitor

	flexible hits			time used (min)	
NSC range	Chem-X	present work	overlap (%)	Chem-X	present work
1−60 040	720	713	99	65	121
60 041−131 742	815	821	98	75	146
131 743−240 407	1063	1040	97	92	181
240 408−506 238	1549	1570	98	98	193
506 239−706 805	1640	1637	99	180	305

Table 6 shows the distance variations from the stored structures in the database to the generated conformations, which meet the pharmacophore requirements. The 2D and 3D structures of three active compounds, NSC 32180, 373937, and 382025, identified from flexible searching are shown as Figure 6.

Table 6. Distances of Stored Structure and Generated Conformations Discovered in a Flexible Search for HIV-1 Protease Inhibitor

NSC	distances of stored structure (d_1,d_2,d_3 in Å)	distances of generated conformation (d_1,d_2,d_3 in Å)
20143	3.94, 3.91, 2.93	5.73, 4.84, 2.76
32180	4.63, 5.47, 3.85	4.63, 5.51, 3.67
40328	4.70, 4.48, 4.23	4.50, 4.31, 2.67
109412	3.91, 3.91, 2.89	5.13, 4.12, 2.78
373937	5.47, 4.63, 3.85	5.46, 4.63, 2.75
382025	8.85, 6.52, 3.21	6.21, 5.90, 3.33

Although the flexible search takes more time to complete, it is worthwhile to note that the flexible searching allows the discovery of more novel lead compounds than the rigid searching. For example, several active compounds, including

the most active compound, NSC 32180, were not identified through the rigid searching.

5. SUMMARY

In this paper, we present a Web-based tool for 3D pharmacophore searching. This Web-based tool is based on a server-client Web architecture. A Gateway E-5200 PC (PIII 700MHZ CPU, 1GB RAM, 36GB hard disk) running Windows NT server acts as the Oracle database server and Web server. In general, a three-center pharmacophore search time (for searching 50 000 compounds) will generally be less than three minutes. For searching the entire database (250 000 compounds), the search time is approximately 10 min for a three-center pharmacophore model. When the database becomes larger, more system resources are required. For a flexible pharmacophore search, the search time is increased to several hours, and the process cannot be interactive. Currently, the results are written to a formatted file and sent back to the user with a HTTP link once the search is done.

Given the extensive use of the 3D-pharmacophore searching in drug discovery, we expect that our Web-based pharmacophore searching system described in this paper will become a useful tool for scientists who like to perform the pharmacophore search of large chemical databases via the Internet.

ACKNOWLEDGMENT

The authors would like to thank Dr. Renxiao Wang[23] from our group for offering his X-LOGP program for calculating the partition coefficient between *n*-octanol and water (Log*P*). We appreciate Dr. Peter Ertl[24] from Novartis Crop Protection AG for making his Java Molecular Editor publicly available for use in this work. The authors would like to thank Dr. Marc Nicklaus at the National Cancer Institute, NIH for careful reading of the manuscript and many useful suggestions. The excellent editorial assistance of Ms. Linda Hemingway is greatly appreciated.

REFERENCES AND NOTES

(1) Greenidge, P. A.; Carlsson, B.; Bladh, L.; Gillner, M. Pharmacophores Incorporating Numerous Excluded Volumes Defined by X-ray Crystallographic Structure in Three-Dimensional database Searching: Application to the Thyroid Hormone Receptor. *J. Med. Chem*. **1998**, *41*, 2503−2512.

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

(3) Wang, S.; Milne, G. W. A.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of Novel, Nonpeptide HIV-1 Protease Inhibitors by Pharmacophore Searching. *J. Med. Chem*. **1996**, *39*, 2047−2054.

(4) Milne, G. W. A.; Nicklaus, M. C.; Driscoll, J. S.; Wang, S.; Zaharevitz, D. National Cancer Institute Drug Information System 3D database. *J. Chem. Inf. Comput. Sci*. **1994**, *34*, 1219−1224.

(5) Gund, P. Three-dimensional Pharmacophore Pattern Searching. *Prog. Mol. Subcell. Biol*. **1977**, *5*, 117−143.

(6) Van Drie, J. H.; Weininger, D.; Martin, Y. C. ALLADIN: An Integrated Tool for Computer-assisted Molecular Design and Pharmacophore Recognition from Geometric, Steric, and Substructural Searching of Three-dimensional Molecular Structures. *J. Comp-Aided Mol. Design* **1989**, *3*, 255−251

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

(8) UNITY 4.2; Tripos Inc.: 1699 South Hanley Rd., St. Louis, MO, 63144.

(9) Murrall, N. W.; Davies, E. K. Conformational Freedom in 3-D Database. 1. Technology. *J. Chem. Inf. Comput. Sci*. **1990**, *30*, 312−316.

(10) Allen, F. H.; Davies, J. E.; Galloy, J. J.; Johnson, O.; Kennard, O.; Macrae, C. F.; Mitchell, E. M.; Mitchell, G. F.; Smith, J. M.; Watson, D. G. The Development of Version 3 and 4 of the Cambridge Structural Database System. *J. Chem. Inf. Comput. Sci*. **1992**, *31*, 187−204.

(11) 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.

(12) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci*. **1988**, *28*, 31−36.

(13) Weininger, D.; Weininger, A.; Weininger, J. L. SMILES 2. Algorithm for Generation of Unique SMILES Notation. *J. Chem. Inf. Comput. Sci*. **1989**, *29*, 97−101.

(14) Weininger, D. SMILES, Depict. 3. Graphical Depiction of Chemical Structures. *J. Chem. Inf. Comput. Sci*. **1990**, *30*, 237−243.

(15) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic Generation 3D Atomic Coordinates for Organic Molecule. *Tetrahedron Comput. Method* **1990**, *3*, 537−547.

(16) Sadowski, J.; Gasteiger, J. From Atoms and Bonds to Three-Dimensional Atomic Coordinates: Automatic Model Builders. *Chem. Rev.* **1993**, *93*, 2567−2581.

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

(18) Wang, X.; Wang, J. T. L. Fast Similarity Search in Three-Dimensional Structure Databases. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 442−451.

(19) Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol*. **1995**, *245*, 43−53.

(20) Pickett, S. D.; MasonJ. S.; McLay, I. M. Diversity Profiling and Design Using 3D Pharmacophores: Pharmacophore-Derived Queries (PDQ). *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 1214−1223.

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

(22) Wang, T.; Zhou, J. 3DFS: A New Flexible Searching System for Use in Drug Design. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 71−77.

(23) Brown, R. D.; Martin, Y. C. The Information Content of 2D and 3D Structural Descriptors Relevant to Ligand−Receptor Binding. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1−9.

(24) Wang, R.; Gao, Y.; Lai, L. Calculating Partition Coefficient by Atom-additive Methodol. *Perspect. Drug Dis. Des*. **2000**, *19*, 47−66

(25) Ertl, P.; Jacob, O. WWW-based Chemical Information System. *J. Mol. Struct. (THEOCHEM)* **1997**, *419*, 113−120.

(26) Willett, P. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 983−996.

(27) Brown, R. D.; Martin, Y. C. Use of Structure−Activity Data to Compare Structure-Based Clustering Methods and Descriptors for Use in Compound Selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572−584.

(28) SYBYL/Force Field, Molecular modeling software; Tripos Inc.: 1699 S. Hanley Road, St Louis, MO 63144−2913.

(29) Beusen, D. D.; Shands, E. F. B.; Karasek, S. F.; Marshall, G. R.; Dammkoehler, R. A. Systematic Search in Conformational Analysis. *J. Mol. Struct. (Theochem.)* **1996**, *370*, 157−171.

(30) Nicklaus, M. C.; Wang, S.; Driscoll, J.; Milne, G. W. A. Conformational changes of small molecules binding to proteins. *Bioorg. Med. Chem.* **1995**, *3*, 411−428.

CI010083I