

A 3D Structure Database of Components from Chinese Traditional Medicinal Herbs

Xuebin Qiao, Tingjun Hou, Wei Zhang, SenLi Guo, and Xiaojie Xu*

College of Chemical and Molecular Engineering, Beida Yang ShengTang Joint Laboratory for Natural Product, Peking University, Beijing 100871, China

Received November 6, 2001

This article described a 3D structure database of components extracted from Chinese Traditional Medicinal (CTM) herbs. It offers not only basic molecular properties and optimized 3D structure of the compounds but also detailed information on their herbal origin, including basic herbal category (e.g. English name, Latin name, and family), effective parts, and clinical effects. An easy to use, interactive GUI browser allows users to perform various searches via complex logical query builder. Combined with the latest network database engine (MySQL), it can achieve excellent performance under both a local network and an Internet environment. We have tested it on the design of inhibitors of NS3-NS4A protease. Results show that the structure database of components extracted from Chinese medicinal herbs can be a rich source in searching the lead compound.

INTRODUCTION

Phytomedicine is a part of health care systems around the world. The World Health Organization (WHO) estimates that 80% of the world's people rely on herbs for their primary health care needs.¹ Chinese traditional phytotherapy has been developed and incorporated the therapeutic use of herb for more than four thousand years. Recently, Chinese phytomedicine is gaining more and more attention in clinical practices throughout the world,^{2,3} due to the global trends of using natural products in prevention and diagnosis of physical and mental diseases.

However, the challenges in incorporating Chinese herbs into clinical practices are broad. In addition to creating a resource base and clarifying liability issues, keeping quality and consistency of herbal products, identifying primary bioactivity components, and describing drug pharmacological interaction in modern scientific dialect are all issues to be addressed when practicing in a conventional setting at regions outside China, especially in western developed countries.

The methodology of structure based drug design used in modern chemistry gave rise to pharmacology and the development of synthetic drugs. Unfortunately, the drug discovery process is such a money and time-consuming process that it becomes more and more difficult for the industry to develop new drugs. To organize the vast quantities of biochemical data and speed up drug discovery, various structure databases have emerged and have become the common tools for the drug designer. Among these, the 3-D structure database of compounds derived from natural product gains wide interest, since natural products could be seen as an abundant source of candidate drugs. A growing number of successful drug design cases based on structural database searches have been reported.^{4–7}

In this paper, we reported a 3D structure database developed in our lab. The system currently enables the following:

- Easy retrieval molecular information from a remote database
 - Simple query as well as complex logic query on almost every field of a record
 - Incremental query upon previous search
 - Easy substructure query
 - Various display modes with many fine-tuned options for 3D structure visualization, e.g. wire frame, ball-and-stick, and space fill
 - Calculation of bond distance, torsion angle, and bond angle on a 3D structure
 - Batch 3D structures output to integrate with other commercial molecular simulation or drug design software
- This structure database consists of chemical compounds isolated from Chinese traditional herbs. Although there are a few successful natural product related databases now,^{8,9} they all have their own limits, e.g. the NAPRALERT⁸ lacks structural information and the DNP⁹ lacks herbal details. Our database focuses on providing both structural and herbal information specialized to the CTM field. Using this kind of database as an supplement for current drug design tools, researchers can take full advantage of our knowledge of Chinese traditional herbal medicine while avoiding the obstacles in incorporating Chinese herbs directly into clinical practices addressed above. Although currently the scale of commercial natural product databases is far beyond ours, their data has been collected without respect to the therapeutic effects. Naturally, it is expected that our database could, more or less, be more efficient in drug design than that of the commercial natural products since all compounds presented here are stemmed from medicinal botanies reputed in ancient Chinese folklore. At last, a test case to discover potential inhibitors of NS3-NS4A protease is shown to prove the efficiency of using this database.

METHODS

Philosophy Behind the Database Design. The primarily applied form of Chinese Traditional Drugs (CTDs) in clinical

* Corresponding author phone: (086)-010-62757456; fax: (086)-010-6275; e-mail: xiaojie@chem.pku.edu.cn.

practice is an herbal mixture. In many instances, the use of the entire plant or leaf is more clinically effective than isolating the primary active compounds. For example, feverfew (*Tanacetum parthenium*) for migraine prophylaxis did not demonstrate significant effects when isolated to the active ingredient in a recent randomized controlled trial.¹⁰ Some studies concluded that the therapeutic effects might be a complementary action of natural constituents in the whole plant form. Even though we do not have a good understanding about how the traditional medicines work, an essential point is that the basic principle of CTDs and Western drugs (WDs) are the same, that is, the interactions between the medicinally active molecules of CTDs or WDs and their biological targets at the molecular levels provide therapeutic functions for diseases. Thus, in practical applications, a reduction of the scientific model of CTDs is to isolate certain compounds and document their effects. Controlled experiments were conducted to measure curative effects under specific conditions. These chemical compounds were broken down and analyzed to establish weight parameters.¹¹ We consider that the latter model is more practical in current drug design, since the herbal mixtures in conventional herbal therapy are too complex to trace their complementary action. Our database design adopts this reduced model.

There are two kinds of audiences that our database is primarily aimed for. The first kind are those drug designers who want to speed up the searching process of lead compounds and improve the shooting accuracy, because the compounds in our database come from Chinese medicinal herbs having definite curative effects. Moreover, many of compounds here have not listed in any commercial database by now since their original papers were published in Chinese. The second kind are those phytomedicine researchers who may use it in a more advanced manner. For example, one can set up a new pharmacophore model to guide new drug discovery through finding the common structural characteristic among bioactivity components extracted from medicinal herbs having the same curative effects to specific disease. Of course, a researcher may also take it as an auxiliary tool when he does data statistic work on linking the Chinese traditional herbal medicine theory with modern biochemical theory.

According to the design objectives above, all the compounds in the database were carefully chosen from those being reported to isolate from the Chinese traditional medicinal herbs. To keep our data as comprehensive as possible while not digressing away from the drug design purpose, properties from three categories were selected to form a complete record. The first category is basic molecular information, including common name, standard name, formula, weight, CAS entry number, and physical properties. The second category is structural information, including 2D structure, optimized 3D structure, and various molecular structural properties useful in drug design, especially in QSAR (see Table 1). All the 3D structures are optimized via the molecular mechanics module in Cerius2¹² using the MMFF force field.¹³ The last category is about the information on source herbs, including Latin name, family, effective parts, and curative effects. The herbal information is not quite as comprehensive but is enough for a drug designer. Partial resources used in data collection are listed in Appendix A.

Table 1. List of Structural Descriptors Stored in the Database

descriptor name	brief description
area	molecular surface area
Vm	molecular volume
MR	Ghose and Crippen molar refractivity
Apol	sum of atomic polarizabilities
RadOfGyration	radius of gyration
Shadow Sxy	area of the molecular shadow in the XY plane
Shadow Sxz	area of the molecular shadow in the XZ plane
Shadow Syz	area of the molecular shadow in the YZ plane
Shadow (Sxy, f)	fraction of area of molecular shadow in the XY plane over area of enclosing rectangle
Shadow (Sxz, f)	fraction of area of molecular shadow in the XZ plane over area of enclosing rectangle
Shadow (Syz, f)	fraction of area of molecular shadow in the YZ plane over area of enclosing rectangle
Shadow η	ratio of largest to smallest dimension
Shadow Lx	length of molecule in the X dimension
Shadow Ly	length of molecule in the Y dimension
Shadow Lz	length of molecule in the Z dimension
MW	molecular weight
density	density
PMI	principal moment of inertia
AlogP98	Ghose and Crippen octanol/water partition coefficient
Fh2o	desolvation free energy for water
Foct	desolvation free energy for octanol
JX	Balaban indices
$\kappa 1$	Kier's shape indices 1
$\kappa 2$	Kier's shape indices 2
$\kappa 3$	Kier's shape indices 3
$\kappa \alpha 1$	Kier's alpha-modified shape indices 1
$\kappa \alpha 2$	Kier's alpha-modified shape indices 2
$\kappa \alpha 3$	Kier's alpha-modified shape indices 3
Wiener	Wiener index
LogZ	logarithm of Hosoya index
Zagreb	Zagreb index

Current Scales. By now, there are 10 564 herbal component records and 2073 CTM herbs records from 296 families having been collected. The records in the database contain bibliographic information, factual molecular, and herbal data as well as computational descriptors on the components from CTM herbs. There are over 90% components having optimized 3D structures and 80% components having a CAS registry number. The database contains records from 1950 to the present. Approximately 80% of the records are from systematic survey of the literatures from 1980 to the present. The remaining records were obtained by selective retrospective dating back to 1950.

Database Engine Configuration. The system employs client/server architecture. We choose **MySQL** as our back-end database engine in order to achieve scalability, flexibility, and high performance with minimum costs. **MySQL** is a fast, multithreaded, multiuser, platform-independent, and robust SQL (Structured Query Language) database server being widely used on the Internet. Thanks to the power of client/server architecture and **MySQL** engine, our database back-end may be installed onto not only the same desktop PC with the client browser but also a remote database server on network supporting TCP/IP or NETBIOS protocols. For personal use, the easiest way is to install the database and the client browser on the same computer. For group-wise application, it is highly recommended to install the database on a dedicated server to improve the overall performance and simplify the maintenance efforts. There are no limits in selection of operating system running on the server, since almost all modern systems work well with **MySQL**.

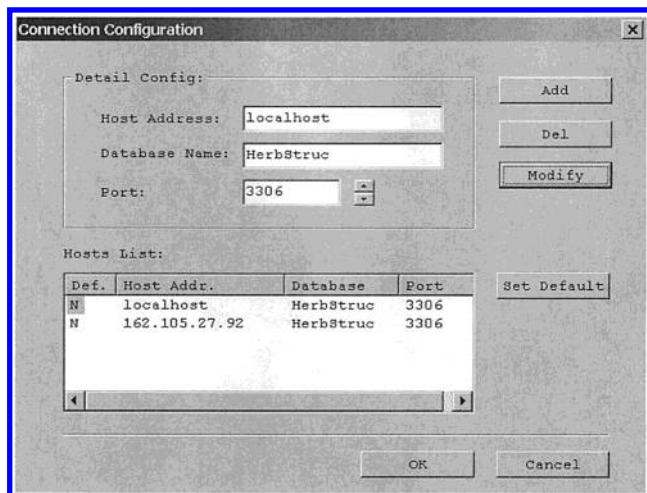


Figure 1. Database connection configuration interface.

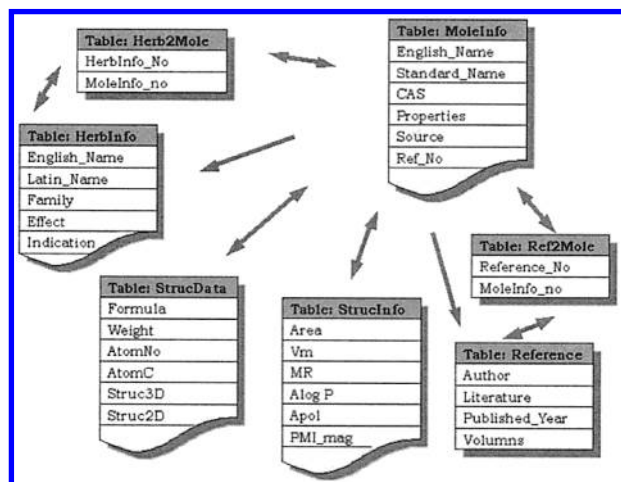


Figure 2. Table layout and internal linkage of the database. Partial important field names in each table are listed under the table name. Table-pair linked with a one-way arrow means that their records do not have the one-to-one mapping characteristics and thus result in the reverse searching which proves to be a time-consuming process. An intermediary index table (e.g. table [Ref2Mole]) has to be introduced to speed up the query.

A very flexible and secure host-based verification system allows better data protection when the database opens to Internet. At client peer, a user can add, modify, or delete a specific database, local or remote, from database configuration list and select a default database to connect to at start up (see Figure 1). The combination of host address (either IP address or host domain name), database name, and port number specifies a unique database configuration. Successful connection is based on whether the password supplied by a user can be verified.

At the front-end, graphical input interfaces and user-friendly search forms provide an easy access to the data without requiring knowledge of detailed command syntax. At present, only MS Windows version of client running on IBM compatible PC is available. A Unix version will be released in the future.

A multitable layout is designed to store all property fields of data (see Figure 2). There are five tables designed to store compound information and two auxiliary index link tables to speed up query. Two tables store basic molecular information, including common name, standard name, CAS number, formula, weight, original source herb, and literature.

Two tables store structural information, including 2D structure, optimized 3D structure, individual element's count, and more than 30 useful molecular structural properties (see Table 1). Another table stores basic herbal information, including English name, Latin name, family, effective parts, pharmacological effect, and clinic indication. The last table stores standard references information. Due to the multi-to-multi mapping nature among the compound properties, it is not wise to simply put all information into a single table. For example, a compound may be isolated from multiple herbs and an herb may have multiple ingredients. The same holds true for compound's literature field. This multitable layout may not only significantly reduce the storage costs but also have balanced performance under concurrent queries.

Supported Query Type. The query method adopted in present implementation is very similar to the one in the Cambridge Database.¹⁴ That is, all kinds of retrieval options, simple or complex, are supported via a special query builder.

The following retrieval options are available:

- Compound name and chemical name segments
- Formula segments and weight range
- CAS registry numbers
- Individual elements count from C, N, O, S, P
- Substructural search based on GMA algorithm suggested by Jun Xu et al.¹⁵

•Numeric value or range query for various molecular structural properties, such as

- Area: surface area
- Vm: molar volume
- MR: molar refractivity
- AlogP: logarithm of the octanol–water partition coefficient

- Bibliographical search term
- Herb name, Latin name, family segments
- Effective parts, pharmacological effect, and clinic indication segments of specific herb
- Complex Boolean expressions
- Incremental search upon previous query result

A typical query procedure consists of two steps. For example, if we want to retrieve all the structures that MUST have 1–10 oxygen atoms and MUST NOT have “HYDROXY” substring in their compound's standard name, we could do it in two steps. In the first step, we can create each query operand in a respective query dialogue box (see Figure 3). The second step is to dispatch each operand in logical operand pool into specific logical operator's (AND, OR, and NOT) processing queue (see Figure 4). In addition, one can select whether to search the records from the whole database or the previous search result set. Finally, a user may execute a real query to retrieve the concerned information.

Logical Query Formula. To simplify the composite logical query formula, we chose the fix formula as follows

$$(Q_{a1} \text{ AND } Q_{a2} \text{ AND } \dots Q_{an}) \text{ AND}$$

$$(\text{NOT } (Q_{n1} \text{ AND } Q_{n2} \text{ AND } \dots Q_{nn})) \text{ OR}$$

$$(Q_{r1} \text{ OR } Q_{r2} \text{ OR } \dots Q_{rn})$$

where Q_a represents query operands in queue of logical AND,

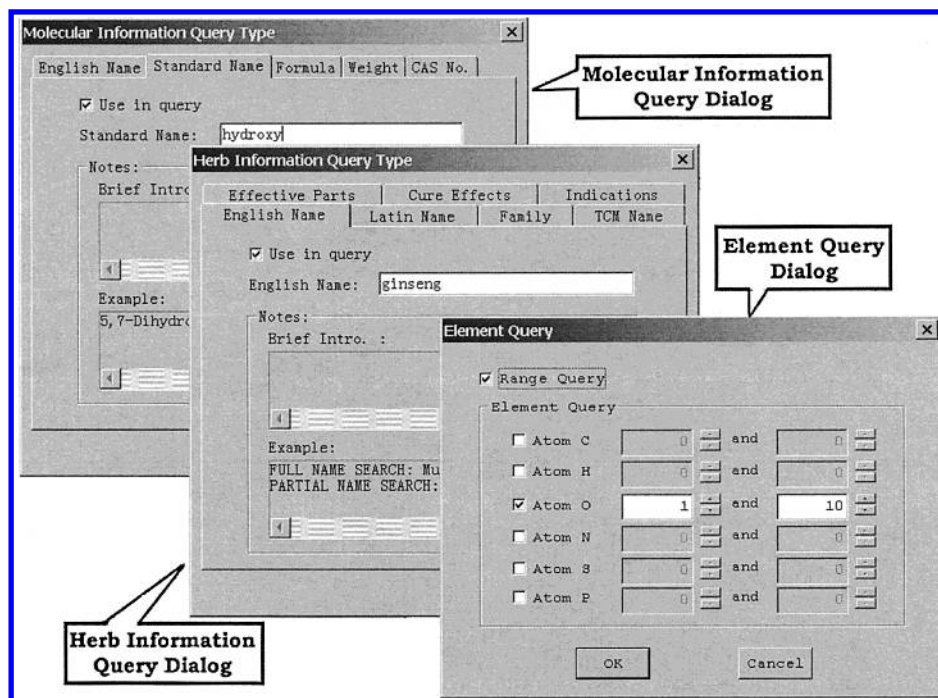


Figure 3. Three types of query boxes used in the example query.

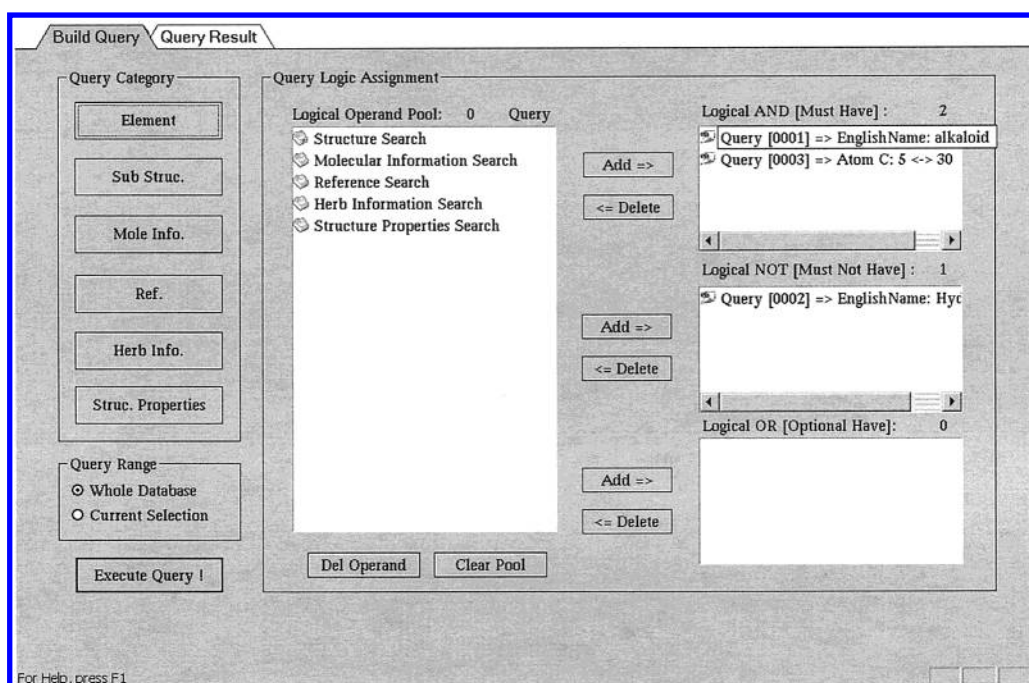


Figure 4. The query builder view of the database client. The example operands have been dispatched into respective logical operator processing queue.

Q_n represents query operands in queue of logical NOT, and Q_r represents query operands in queue of logical OR.

Although this logical formula cannot cover all possible composite logical query patterns, it is a quite convenient way to express the common query pattern. For example, the most common query pattern that a researcher uses is to find the set which MUST HAVE certain properties and MUST NOT HAVE certain properties OR MAY HAVE certain properties. More over, almost every complex logical formula may be constructed when combined with the technique of incremental query.

Structure Edit and Substructure Query. We developed a convenient molecular construction tool to aid the building

of substructure query. This module enables easy construction of organic molecules by adding/connecting atoms, bonds, rings, and functional groups. A user may choose the desired action from the menu or left toolbar and then picks the appropriate place on the drawing area. After the structure fragment construction, a user can transfer it to the query operand pool by clicking [Store] command under the file menu and then quit the construction tool.

Since the structural match problem is a well-known example of NP-complete problem¹⁷ and a traditional algorithm has factorial of factorial computing complexity, an efficient algorithm for structure query is indispensable to any structure database. At present, we implemented the GMA

The screenshot displays two overlapping windows from a web application. The 'Herb Information Page' (top) contains the following fields: English Name: Oriental Stephania, Latin Name: Stephania cepharantha Hayata, TCM Name: bai yao zi, Family: Menispermaceae, Parts: tuberous root., Effects: To clear heat and, and Indications: Sore throat and th. The 'Molecular Information Page' (bottom) contains: SerialNo: 9527, English Name: Alkaloid FK 3000, Standard Name: Alkaloid FK 3000; FK 3000; 6,7-Di-O-acetylsinococuline, Formula: C22H27NO7, Weight: 417.4578, CAS No: N/A, and Properties: Mp 185-186 deg C; Optical Rotation: -163 at 22 deg C (c, ...). Both windows have scroll bars and a search icon in the top right corner.

Figure 5. The sample molecular and herbal info page in query result view.

algorithm proposed by Jun Xu.¹⁵ GMA is a partial-ordering-set-based backtracking algorithm, which has exponential computing complexity. The distinctive concept of GMA is that it considers a query structure as a “program”, which will be executed on the queried structure (or substructure). Because the substructure query algorithm in our database requests just an answer of “Yes” or “No”, the GMA algorithm does not have to start searching every node on target structure. It may stop at any starting node when a mapping has been found. Thus in our implementation, we ignore other possible mappings and get partial order set information from a query structure only one time at the beginning. It can save much time in a large database search.

Due to the fact that the UDF (User Defined Function) mechanism is not currently supported in the **Windows** version of **MySQL**, we cannot further improve the substructure query efficiency through defining a UDF routine for a substructure query at the server side on the **Windows** platform. On the other hand, creating a native function for **MySQL** is not a good choice either, since it needs to modify the source of the **MySQL** engine. If then, our database would have to be severely dependent on the modified version of the **MySQL** engine, meanwhile it would be difficult to integrate our database into users' existed **MySQL** engine. So we have to make a tradeoff between efficiency and portability. In the present implementation, the substructure query routine is put to the client side, and all the structures have to be retrieved to the client side before running it. When the UDF functions are supported by the **MySQL** engine on both the **Windows** and the **UNIX** platform, the substructure query routine will be moved to the server side as soon as possible.

Result Record Display. The properties of a record are displayed in six different property pages according to their

category (see Figures 5 and 6). The [Basic Info.] page is the basic molecular information, such as name, formula, weight, CAS number, and physical properties. The [Herb Info.] page contains the information on each source herb, including English name, Latin name, family, effective parts, indications, and curative effects. The [Reference] page stores the standard literature description. The [2D structure] page and [3D structure] page display the 2D and 3D structure. In addition to the standard wire-frame view, various modes have been introduced to aid the visualization of 3D structure (see Figure 7). The advantage of CPK (space filling) images is the possibility to see the overall space requirement of the molecule. In ball-and-stick models, on the other hand, the atomic connectivity is visible. Moreover, the system also allows the calculation of bond length, angle, and torsion on the fly like Rasmol. The [Structure Info.] page shows more than 30 structure properties of a molecule (see Table 1). If the result set is too large, it may take several minutes to retrieve all the data at one time. In view of this, a simplest fixed size cache algorithm is introduced to accelerate the display, reduce the client-peer storage, and balance the server payload under concurrent connections. The main idea behind it is that a user usually views the record in sequential manner. Thus we can retrieve only nearby records of the target, say five records before and after the target, if it is not hit in the cache. And any old records in the cache are simply discarded. This little trick can significantly improve the client response on slow-speed network, especially on the Internet.

To demonstrate the record display, the results of a sample query for alkaloid, an interesting kind of secondary metabolite, are shown in Figures 5 and 6. There are 23 components founded. The 3-D structure of one hit record, Alkaloid FK3000, is shown in Figure 6. To let the user have a better idea of the result representation, the content of two other

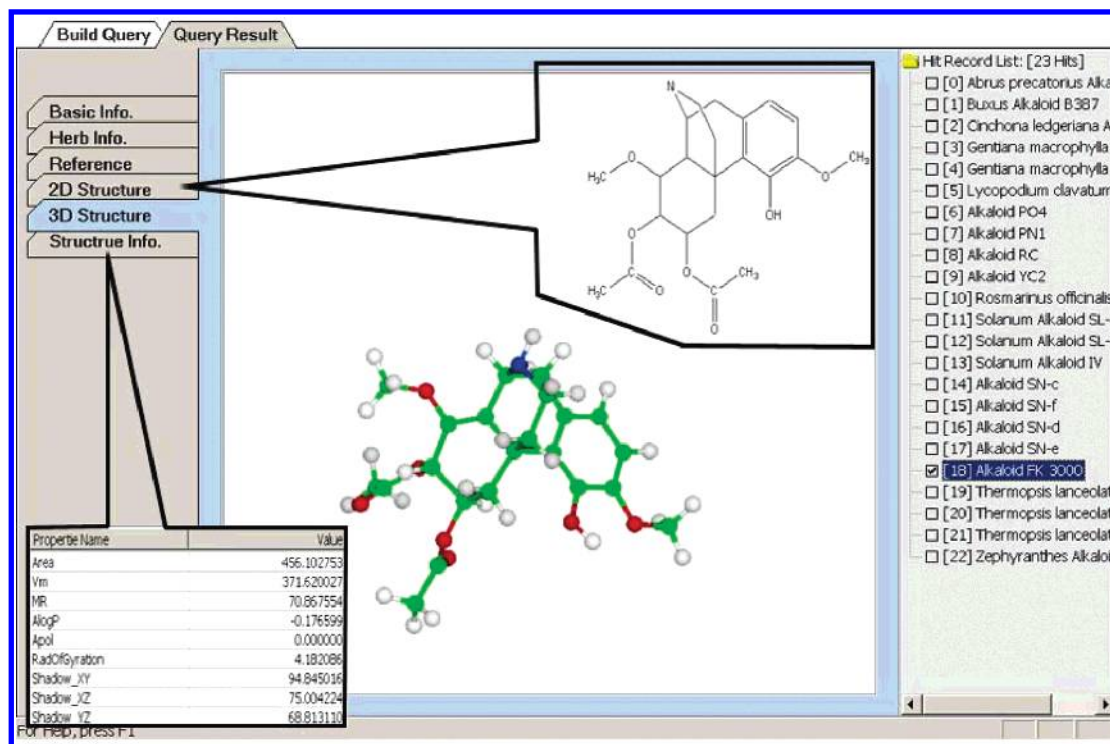


Figure 6. The 3D structure property page in result view of the database client. The partial content of two other pages, [2D Structure] and [Structure Info], are shown on the same page.

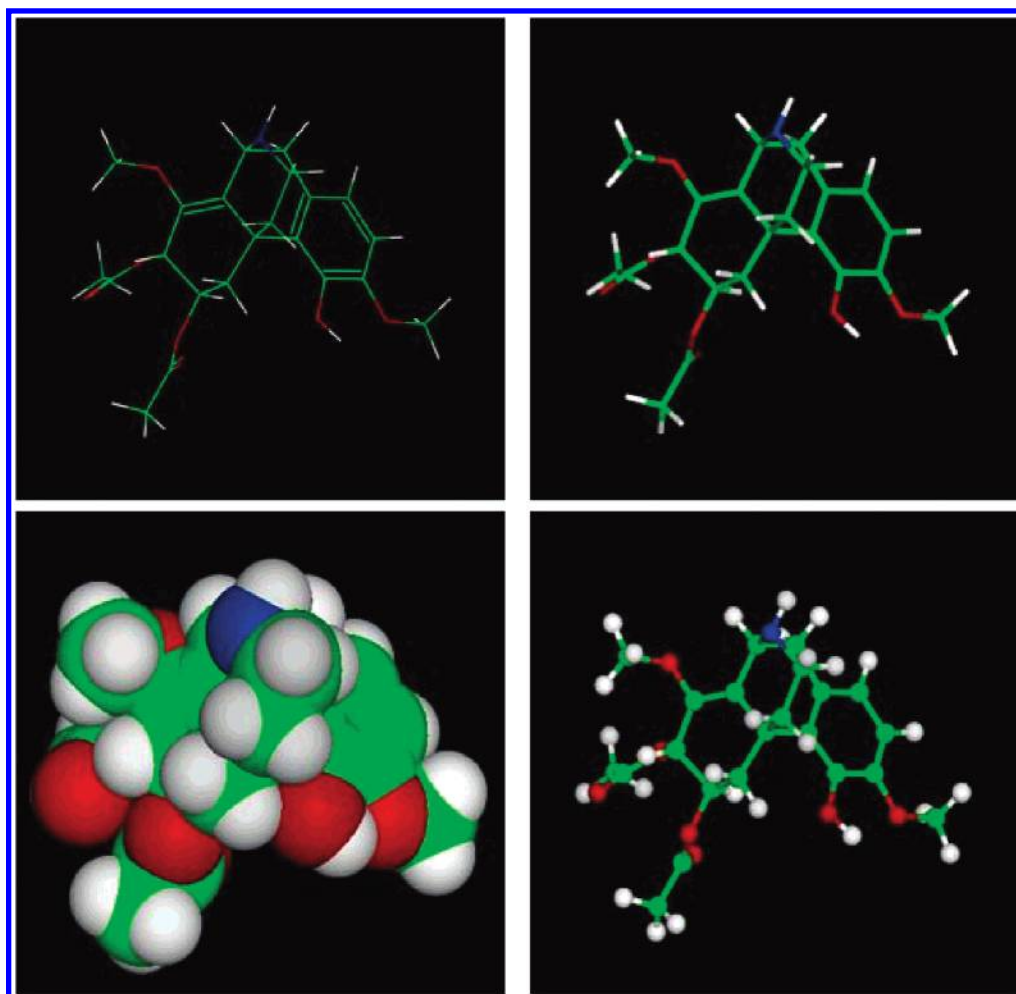


Figure 7. Various visualization modes supported in the result view: space-fill (lower-left), stick (upper-right), ball-stick (lower-right), and wire-frame (upper-left).

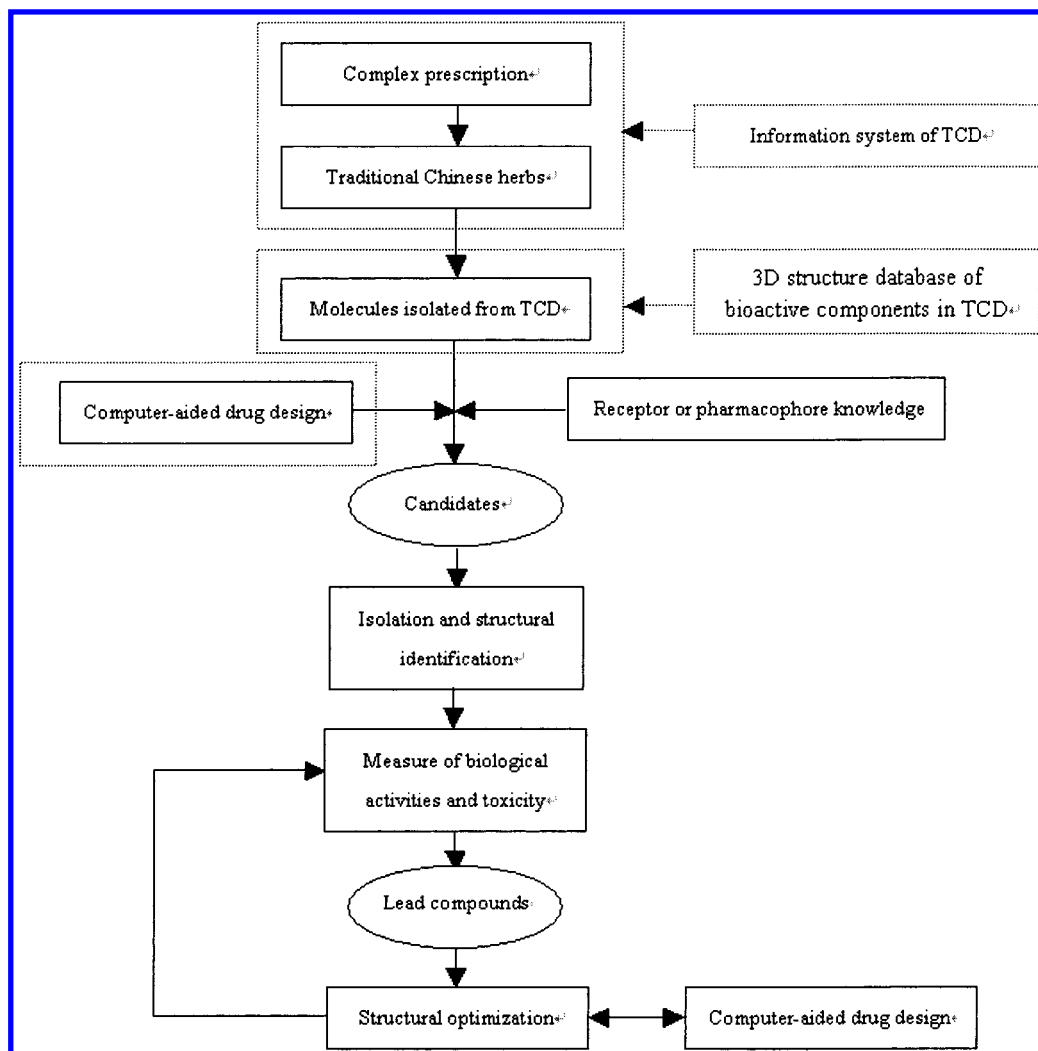


Figure 8. The basic scheme of computer-aided drug design based on TCDs.

pages, [2D Structure] and [Structure Info], is shown on the same page.

In addition, the result structures can be stored in MDL SD file and imported directly into the most popular commercial molecular design software.

DISCUSSION

The methodology used in modern chemistry gave rise to pharmacology and the development of synthetic drugs. Various structural databases have been widely used to accelerate the discovery of new drug. Although there are a few Chinese natural products database currently in use, all of them are not particularly schemed to cater to the specific needs in computer aided drug design. Our database construction minds the distinct requirements in drug discovery from the very beginning.

In practical or clinical applications of CTDs, the complex prescription (herb mixture) is the most widely-used form in therapy. So, in CTDs, the diseases are directly connected with complex prescriptions, not molecules, sometimes not even herbs. So we think that complex prescription is the starting point in the computer-aided drug design based on CTDs.

To utilize the wealth of information from the Chinese medicinal herb better, we suggest an efficient way for

searching the lead compound based on the database search of components from CTM herbs. The start of this method is a complex prescription (Chinese herb mixture). One can retrieve the information of herb mixtures related to specific diseases from certain Chinese herbal medicine databases. In our lab, we have constructed a Traditional Chinese Medicine Drugs Information System based on networks.²¹ This system is based on four databases, including Chinese Herb Database, Chinese Medical Recipe Database, Chinese Patent Medicine Database, and Chinese Medical Dietotherapy Prescription Database. Using this information system developed in our lab, we can directly get the information about herbs related to specific diseases. Then, a 3-D structure subset of ingredients related to specific diseases can be constructed through searching all structures origin from plants consisting of the herb mixture. Next, this structure subset is filtered via virtual screening. There are several screening methods in modern computer-aided drug design. For example, molecular docking can be used to study the interaction between bioactive ingredient and target enzyme when the target 3-D structure is known.^{16–18} Pharmacophore searching can be used to search candidate compounds with specific pharmacophoric groups. Further, candidate compounds may be extracted from respective herbs and be subjected to bioactivity and toxicity test. Finally, a researcher may further

modify the structures of those bioactive compounds to improve their activity. Figure 8 shows the basic scheme of computer-aided drug design based on TCDs.

We have applied the above method into the work of discovering new potential inhibitors of NS3-NS4A protease. HCV is the major pathogenic virus infecting hepatitis after blood transfusion. HCV DNA encodes two proteins, i.e., NS2-NS3 metal protease and NS3 serine protease. The former cleaves NS2-NS3 in multiprotein cleaving. The later catalyzes NS3 cleaving at all cleaving locations. Both of them are vital to the process of HCV duplication. Kim and Love have identified the structure of the NS3-NS4A crystal^{19,20} successfully. We selected the NS3-NS4A protease as a design target. At first, we retrieved an herb mixture and herbal plant, which have therapeutic effects for hepatitis, from a Chinese herbal medicine network information system (CHMNIS)²¹ established by our group. Then we constructed an 1100 compound subset found in the 3-D structure database of bioactive ingredients from medicinal herbs. For each of the molecules in the subset, we studied the interaction between the compound and target protease via a molecular docking module also developed by our group. Then several high score candidate compounds were selected according to an energy and geometry match. Finally, these compounds were extracted from respective Chinese medicinal herbs and subjected to NS3-NS4A protease bioactivity test. This project is in progress now. The results are very encouraging that several compounds exhibit distinct bioactivity against the NS3-NS4A protease.

ACKNOWLEDGMENT

The database structure, retrieval, and display features have been developed over years and are the result of extensive and very valuable work of many colleagues in our group. Therefore the authors would like to express their sincere thanks to Lili Zhu, Ning Liao, Huidong Yu, Dong Zhang, Wei Wang, Haige Luo, and Changmin Gu for their stimulating discussions and valuable contributions. This project was supported by YangShengTang China Ltd. Co. and Peking University Modern Research Center for Chinese Medicine.

APPENDIX A: PARTIAL BIBLIOGRAPHIC LISTS FOR DATA COLLECTION

1. *Acta Botanica Yunnanica*; 1980–2001
2. *Acta Pharmaceutica Sinica*; 1980–2001
3. *Chinese Pharmaceutical Journal*; 1980–2001
4. *Chinese Traditional and Herbal Drugs*; 1980–2001
5. Jiangsu New Medical College. In *Chinese Medicine Dictionary*; Shanghai Science and Technology Press: Shanghai, China, 1979; in Chinese.
6. Crude Drug Company of China. In *A Complete Collection of Chinese Drug Sources in China*; Zhang, H. Y., Ed.; Science Press: Beijing, China, 1994; in Chinese.
7. Editing Group of the Handbook of Bioactivity Components from Medicinal Plants. In *Handbook of Bioactivity Components from Medicinal Plants*; The People's Medical Publishing House: China, 1986; in Chinese.
8. Qi, Ch. In *The Methodology of Studying the Pharmacology of Chinese Drugs*; Qi, Ch., Ed.; The People's Medical Publishing House: Beijing, China, 1993; in Chinese.

9. Jian, Y.; Guo, L. G In *Modern Study of Chinese Drugs and Clinical Applications*; Jian, Y., Ed.; Xueyuan Press: China, 1993; Vol. 1–2, in Chinese.

10. Jian, Y.; In *Modern Study of Chinese Drugs and Clinical Applications*; Jian, Y., Ed.; Ancient Book Press of Chinese Medicine: China, 1997; Vol. 3, in Chinese.

11. Institute of Materia Medica. In *Modern Studies of Chinese Herbal Medicine*; Chinese Academy of Medical Sciences: China, 1996; Vol. 1–3, in Chinese.

12. Ji, Y. B. In *Pharmacological Action and Application of Available Composition of Traditional Chinese Medicine*; Ji, Y. B., Ed.; Heilongjiang Science and Technology Press: Heilongjiang, China, 1995; in Chinese.

13. Ji, Y. B.; Zhang, G. M. In *Pharmacological Action and Application of Available Antitumor Composition of Traditional Chinese Medicine*; Zhang, G. M., Ed.; Heilongjiang Science and Technology Press: Heilongjiang, China, 1998; in Chinese.

REFERENCES AND NOTES

- (1) Akerele, O. In *Summary of WHO guidelines for the assessment of herbal medicine*; *Herbalgram* **1993**, 28, 13–19.
- (2) Kenner, D.; Requena, Y. In *Botanical medicine: a European professional perspective*; Paradigm Publishing: Brookline, MA, 1996.
- (3) Castleman, M. In *The healing herb*; Rodale Press: Emmaus, PA, 1991.
- (4) Wang, S. M.; Zaharevitz, D. W.; Sharma, R.; Marquez, V. E.; Lewin, N. E.; Du, L.; Blumberg, P. M.; Milne, G. W. A. *J. Med. Chem.* **1994**, 37, 4479–4489.
- (5) Wang, S. M.; Milne, G. W. A.; Yan, X. J.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of novel, non-peptide HIV-1 protease inhibitors by pharmacophore searching. *J. Med. Chem.* **1996**, 39, 2047–2054.
- (6) Neamati, N.; Hong, H. X.; Mazumder, A.; Wang, S. M.; Sunder, S.; Nicklaus, M. C.; Milne, G. W. A.; Proksa, B.; Pommier, Y. Depsides and depsidones as inhibitors of HIV-1 integrase: Discovery of novel inhibitors through 3D database searching. *J. Med. Chem.* **1997**, 40, 942–951.
- (7) Kiyama, R.; Honma, T.; Hayashi, K.; Ogawa, M.; Hara, M.; Fujimoto, M.; Fujishita, T. Novel Angiotensin-II Receptor Antagonists – Design, Synthesis, and In-Vitro Evaluation of Dibenzo[A, D] cycloheptene and Dibenzo[B, F]oxepin Derivatives – Searching for Bioisosteres of Biphenyltetrazole Using a 3-Dimensional Search Technique. *J. Med. Chem.* **1995**, 38, 2728–2741.
- (8) Loub, J. *Chem. Inf. Comput. Sci.* **1985**, 25, 99–103.
- (9) Tripos Inc. *Chapman & Hall Chemical Database, Dictionary of Natural Products 6.2*; Tripos Inc.: 1998.
- (10) De Weerd, T. Herbal medicines in migraine prevention: randomized double-blind placebo controlled trial of feverfew preparation. *Phytotherapy* **1996**, 3, 225–255.
- (11) Weiss, R. F. In *Herbal medicine*; Beaconsfield Publishing: Beaconsfield, UK, 1991.
- (12) Molecular Simulation, Inc. In *Cerius2 User Guide*; San Deigo, U.S.A., 1998.
- (13) Halgren, T. A. Merck molecular force field .1. Basis, form, scope, parametrization, and performance of MMFF94. *J. Comput. Chem.* **1996**, 17, 490–519.
- (14) Cambridge Crystallographic Data Centre (CDC). In *CSD User Guide*; Cambridge, U.K., 2001.
- (15) 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.
- (16) Hou, T. J.; Wang, J. M.; Xu, X. J. Automated docking of peptides and proteins by using a genetic algorithm combined with a tabu search. *Protein Eng.* **1999**, 12, 639–647.
- (17) Wang, J. M.; Hou, T. J.; Xu, X. J. Automated docking of peptides and proteins by genetic algorithm. *Chemometr. Intell. Lab.* **1999**, 45, 281–286.
- (18) Hou, T. J.; Xu, X. J. A new molecular simulation software package – Peking University Drug Design System (PKUDDS) for structure-based drug design. *J. Mol. Graph Model.* **2001**, 19, 455–465.
- (19) Kim, J. L.; Morgenstern, K. A.; Lin, C.; Fox, T.; Dwyer, M. D.; Landro, J. A.; Chambers, S. P.; Markland, W.; Lepre, C. A.; Omalley,

- E. T.; Harbeson, S. L.; Rice, C. M.; Murcko, M. A.; Caron, P. R.; Thomson, J. A. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **1996**, 87, 343–355.
- (20) Love, R. A.; Parge, H. E.; Wickershan, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Hostomska, Z. The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **1996**, 87, 331–342.
- (21) Qiao, X. B.; Yu, H. D.; Hou, T. J. Research and development of Traditional Chinese Medicine Drugs Information System based on networks. *Acta Phys.-Chim. Sin.* Submitted for publication.

CI010113H