

# 中草药化学成分的晶体结构图谱库系统\*

吕 扬 郑启泰 伍伯牧\*\*

(中国医学科学院药物研究所,北京 100050; \*\*中国科学院生物物理研究所,北京 100101)

**摘要** 中草药化学成分的晶体结构图谱库系统是我国建立的第一个单晶结构数据库,汇集了近250个中草药化学成分的晶体结构。该库除了通常的数据库含义外,尚兼备多功能检索、绘图、计算等功能,是一个应用广泛的工具库。全部软件是用 dbase, FORTRAN 及汇编语言写成。该软件可在286, 386, 486及其兼容的微机上运行。

**关键词** 中草药;晶体结构;图谱库

应用 X-衍射分析测定有机分子晶体结构始于本世纪二十年代。历经七十余年的发展,迄今, X-衍射单晶结构分析中的两个主要方法(重原子法和直接法)已成为解析有机分子晶体结构的最有效的方法,现在每年可解析约五千个左右的未知有机化合物晶体结构。J. Karle 与 H. Hauptman 积近四十载之研究,在直接法中取得了突破性成就,并因此荣获1985年诺贝尔化学奖,至此, X-衍射单晶结构分析已具备发展为测定有机分子晶体结构的一种常规解析的物理方法。同时,化学家、药物学家、生物化学家在深入探讨有机分子的结构与功能关系,药物分子结构—活性与受体大分子关系等工作中所使用的计算化学方法和分子图形学方法等,均以有机分子的立体化学结构数据作为重要的实验基础。八十年代前后各种有机晶体结构数据库应运而生,诸如:普通有机与金属有机化合物库(Cambridge crystallographic center 英)、蛋白质库(protein data bank 美)和晶体结构库(NBS crystal data center 美)等。

以中草药的化学成分为研究目标的药物化学工作,已成为我国寻找新药的重要研究途径,并已作为植物化学、分子生物学、分子药理学、有机化学、分子力学、量子化学、计算机图形学和晶体学等诸多学科的汇集点,其共同的特征是要在分子、原子乃至电子水平上揭示药物分子立体结构与生物活性的关系。如何获取来自中草药中新化合物的立体结构信息?如何得到相应化合物中分子的结构数据?如何绘制分子的各种立体结构图象?如何计算出相应的立体化学结构参数值等,是实现上述多学科研究目的的一个重要基础。适时地建立一个具有我国特色的、以中草药化学成分为对象的晶体结构图谱库,并使之能在业已广泛进入化学实验室的微机上运行,这将会推进上述化学工作的深入研究。

中草药化学成分的晶体结构图谱库系统,集世界上普通有机与金属有机库(英国)、蛋白质库(美国)、晶体结构库(美国)中的分立功能于一体,并首次列入生物活性,它将成为从事旨在揭示药物分子结构与活性关系的多学科研究中共享的第一个图谱库。

“中草药化学成分的晶体结构图谱库系统”(以下简称为“图谱库”)是我国建立的第一个(单)晶体结构数据库<sup>(1,2)</sup>。它聚集了我国晶体学实验室三十年来完成的近250个中草药化学成分的晶体结构。“图谱库”除了通常的数据库含义之外,尚兼具多功能检索、绘图、计算等功能,

本文于1992年9月2日收到。

\*国家自然科学基金资助项目

因此,它也是一个应用性广泛的工具库。它可以快速完成化合物分类、化合物名称、分子式、化合物中指定元素类型、生物活性、作者、发表刊物、晶胞参数等检索;可以调用指定化合物的原子坐标参数并进行正交化计算,计算任意原子间的距离,成键原子间的键长、键角、扭转角值,分子内或分子间的氢键值,分子平面性及二面角值;也可以绘制分子不同形式的立体结构图(线状图、球棒图、热椭球图、体视图、范德华堆积图等)和反映晶态下分子间关系的沿晶胞不同方向的分子排列图。全部软件是用 Dbase, FORTRAN 及汇编语言写成,绘图与计算程序取自于 NOMCSDP 程序包<sup>(3,4)</sup>,“图谱库”的软件可在286,386,486及其兼容的微机上运行。

### 运行条件

盘操作系统: MSDOS 或 PC DOS

内存: >或=640K

硬盘: >或=40M

显示系统: 高分辨彩显, TVGA, VGA, EGA.

输出: 打印机或笔绘仪

### 录入信息

- 1 天然有机化合物的分类 按中草药化学成分的化合物类型归属,如:甾体类(51),单萜类(52),倍半萜类(53),二萜类(54),三萜类(55)和生物碱类(58)等,为便于国际交流,采用英国剑桥数据中心的分类序号;化合物名称,分子式及生物活性;论文作者,发表刊物(书),年、卷、页,见表1。
- 2 晶体学数据,包括空间群,晶胞参数( $a, b, c, \alpha, \beta, \gamma$ ),晶胞内分子数( $Z$ ),原子坐标参数( $XYZ$ ),见表2。

Tab 1 Example of information from atlas data base (I)

---

Number: 73\*

Classification: 10\*\*

Molecular formula:  $C_{10}H_{19}NO_3$

Compound name: neoclausenamide

Parameters of unit cell:

a=13.895      b=6.886      c=15.9452.

Alpha=90.00    Beta=91.12    Gamma=90.00

Space group: P 21/c

Molecular number in unit cell: 4

Activity: liver protection

Author: JC Cheng

Author: JJ Yang

Author: CH He

Author: QT Zheng

Publication: Chinese Science Bulletin

Volume and page: 34(11), 957

Year: 1989

---

\*Number of atlas data base

\*\*Classification number of cambridge crystallographic center

Tab 2 Example of information from atlas data base (II)——atomic coordinate and deviation

Name	X/a	Y/b	Z/c	$\Delta X/a$	$\Delta Y/b$	$\Delta Z/c$
C 1	0.4726	0.7095	0.4126	0.0002	0.0005	0.0002
C 3	0.3915	0.8045	0.4427	0.0002	0.0004	0.0002
C 4	0.3226	0.6916	0.3884	0.0002	0.0004	0.0002
C 5	0.3652	0.4959	0.3629	0.0002	0.0005	0.0002
C 6	0.5194	0.4126	0.3311	0.0002	0.0005	0.0002
C 7	0.3387	0.4085	0.2659	0.0002	0.0004	0.0002
C 8	0.2443	0.3778	0.2597	0.0002	0.0004	0.0002
C 9	0.2054	0.2537	0.3248	0.0002	0.0005	0.0002
C10	0.1194	0.2311	0.3219	0.0002	0.0005	0.0002
C11	0.0706	0.3273	0.2536	0.0002	0.0005	0.0003
C12	0.1088	0.4450	0.1879	0.0002	0.0006	0.0003
C13	0.1952	0.4720	0.1908	0.0002	0.0005	0.0002
C14	0.2426	0.6665	0.4439	0.0002	0.0004	0.0002
C15	0.2428	0.5636	0.5305	0.0002	0.0005	0.0002
C16	0.1696	0.5367	0.5803	0.0002	0.0005	0.0003
C17	0.0951	0.6116	0.5447	0.0002	0.0006	0.0003
C18	0.0941	0.7151	0.4596	0.0002	0.0006	0.0003
C19	0.1675	0.7426	0.4098	0.0002	0.0005	0.0002
H 3	0.3844	0.7998	0.5111	0.0002	0.0004	0.0002
H 4	0.3028	0.7607	0.3324	0.0002	0.0004	0.0002
H 5	0.3491	0.3954	0.4067	0.0002	0.0005	0.0002
H6A	0.5042	0.2818	0.3471	0.0002	0.0005	0.0002
H6B	0.5728	0.4432	0.3605	0.0002	0.0005	0.0002
H6C	0.5230	0.4250	0.2625	0.0002	0.0005	0.0002
H 7	0.3638	0.2827	0.2589	0.0002	0.0004	0.0002
H 9	0.2387	0.1838	0.3716	0.0002	0.0005	0.0002
H10	0.0931	0.1477	0.3677	0.0002	0.0005	0.0002
H11	0.0108	0.3121	0.2523	0.0002	0.0005	0.0003
H12	0.0754	0.5095	0.1393	0.0002	0.0006	0.0003
H13	0.2209	0.5559	0.1448	0.0002	0.0005	0.0002
H15	0.2944	0.5110	0.5559	0.0002	0.0005	0.0002
H16	0.1706	0.4658	0.6398	0.0002	0.0006	0.0003
H17	0.0441	0.5919	0.5790	0.0002	0.0006	0.0003
H18	0.0423	0.7682	0.4348	0.0002	0.0006	0.0003
H19	0.1663	0.8155	0.3509	0.0002	0.0005	0.0002
H2'	0.4121	1.0677	0.4682	0.0023	0.0058	0.0026
H3'	0.3829	0.4851	0.1530	0.0026	0.0065	0.0029
O 1	0.5447	0.7752	0.4284	0.0001	0.0003	0.0002
O 2	0.3898	1.0054	0.4188	0.0001	0.0003	0.0002
O 3	0.3671	0.5328	0.1913	0.0001	0.0004	0.0002
N 1	0.4551	0.5434	0.3673	0.0001	0.0004	0.0002

## 图谱库功能

### 1 功能

- 1.1. 检索 可从七个方面进行检索:化合物名称(全名或局部名称)、分子式(全部或分子式中指定元素)、作者、发表刊物、生物活性、晶胞参数(全部或部分准确值或近似值)、分类号。

1. 2. 立体化学结构参数计算 可实现为进行分子力学计算所需的晶体学坐标系(三斜、单斜、三方及六方晶系)的正交化计算。计算键长、键角、二面角、分子内(分子间)氢键值,指定分子(全部或局部)的平面性计算和分子平面间的夹角计算。

1. 3. 绘图 提供分子的立体结构图;线图、球棒图、晶胞内分子密堆积图等。

## 2 示例

2. 1. 检索 以化合物名称、分子式为例:

### Compound name index

Neoclausenamide		
Number: 73		
Classification: 10		
Molecular formula: $C_{18}H_{19}NO_3$		
Compound name: neoclausenamide		
Parameters of unit cell:		
a=13.895	b=6.886	c=15.9452.
Alpha=90.00	Beta=91.12	Gamma=90.00
Space group: P 21/c		
Molecular number in unit cell: 4		
Activity: liver protection		
Author: JC Cheng		
Author: JJ Yang		
Author: CH He		
Author: QT Zheng		
Publication: Chinese Science Bulletin		
Volume and page: 34(11), 957		
Year: 1989		

### Molecular formula index

$C_{18}H_{19}NO_3$		
Number: 73		
Classification: 10		
Molecular formula: $C_{18}H_{19}NO_3$		
Compound name: neoclausenamide		
Parameters of unit cell:		
a=13.895	b=6.886	c=15.9452.
Alpha=90.00	Beta=91.12	Gamma=90.00
Space group: P 21/c		
Molecular number in unit cell: 4		
Activity: liver protection		
Author: JC Cheng		
Author: JJ Yang		
Author: CH He		
Author: QT Zheng		
Publication: Chinese Science Bulletin		
Volume and page: 34(11), 957		
Year: 1989		

2. 2. 计算 依次计算键长、键角、最小二乘平面与扭角值。

坐标 (从略)

键长

C(1)—C(3)	1.3961(2)	C(8)—C(9)	1.4566(2)
C(1)—O(1)	1.1233(1)	C(8)—C(13)	1.4366(6)
C(1)—N(1)	1.3719(1)	C(9)—C(10)	1.2051(3)
C(3)—C(4)	1.4960(1)	C(10)—C(11)	1.4337(5)
C(3)—O(2)	1.4350(2)	C(11)—C(12)	1.4345(2)
C(4)—C(5)	1.5300(1)	C(12)—C(13)	1.2148(3)
C(4)—C(14)	1.4446(2)	C(14)—C(15)	1.5520(3)
C(5)—C(7)	1.6934(4)	C(14)—C(19)	1.2794(9)
C(5)—N(1)	1.2919(3)	C(15)—C(16)	1.3156(7)
C(6)—N(1)	1.4010(1)	C(16)—C(17)	1.2793(9)
C(7)—C(8)	1.3304(3)	C(17)—C(18)	1.5325(3)
C(7)—O(3)	1.5236(2)	C(18)—C(19)	1.3184(2)

## 键角

C(3)-C(1)-O(1)	117.16(3)	C(9)-C(8)-C(13)	129.23(2)
C(3)-C(1)-N(1)	115.82(9)	C(8)-C(9)-C(10)	115.46(6)
O(1)-C(1)-N(1)	127.00(1)	C(9)-C(10)-C(11)	115.04(7)
C(1)-C(3)-C(4)	93.97(4)	C(10)-C(11)-C(12)	129.75(2)
C(1)-C(3)-O(2)	111.69(7)	C(11)-C(12)-C(13)	115.98(6)
C(4)-C(3)-O(2)	109.89(9)	C(8)-C(13)-C(12)	114.46(7)
C(3)-C(4)-C(5)	111.40(5)	C(4)-C(14)-C(15)	127.70(9)
C(3)-C(4)-C(14)	101.47(6)	C(4)-C(14)-C(19)	108.76(5)
C(5)-C(4)-C(14)	111.32(7)	C(15)-C(14)-C(19)	123.52(7)
C(4)-C(5)-C(7)	118.50(2)	C(14)-C(15)-C(16)	127.82(9)
C(4)-C(5)-N(1)	98.12(9)	C(15)-C(16)-C(17)	107.68(6)
C(7)-C(5)-N(1)	109.00(5)	C(16)-C(17)-C(18)	125.00(7)
C(5)-C(7)-C(8)	108.70(6)	C(17)-C(18)-C(19)	127.32(9)
C(5)-C(7)-O(3)	117.28(4)	C(14)-C(19)-C(18)	108.62(5)
C(8)-C(7)-O(3)	107.53(5)	C(1)-N(1)-C(5)	113.60(9)
C(7)-C(8)-C(9)	114.79(6)	C(1)-N(1)-C(6)	130.16(9)
C(7)-C(8)-C(13)	115.96(6)	C(5)-N(1)-C(6)	116.01(6)

## 最小二乘平面

Distances(A) to the least-squares planes

Plane no. 1

Equation of the plane :  $2.69(3)X + 5.986(5)Y + 7.188(21)Z = 7.838(8)$ 

Distances(A) to the plane from the atoms in the plane.

C14	-.004(4)	C15	.003(4)
C16	.003(6)	C17	-.005(6)
C18	.000(6)	C19	.004(4)

Chi squared for this plane 3.501

Plane no. 2

Equation of the plane :  $-1.94(3)X + 5.569(6)Y + 9.152(19)Z = 3.997(7)$ 

Distances(A) to the plane from the atoms in the plane.

C 8	.009(4)	C 9	-.011(4)
C10	.004(4)	C11	.009(6)
C12	-.011(6)	C13	-.002(4)

Chi squared for this plane 18.815

## 二面角

Dihedral angle between planes A and B

A	B	Angle(deg)
1	2	20.65(15)

扭角

O 1	C 1	C 3	C 4	167.1(4)	O 1	C 1	C 3	O 2	53.9(2)
N 1	C 1	C 3	C 4	-13.6(2)	N 1	C 1	C 3	O 2	-126.8(3)
C 3	C 1	N 1	C 5	-3.0(2)	C 3	C 1	N 1	C 6	-177.8(4)
O 1	C 1	N 1	C 5	176.2(4)	O 1	C 1	N 1	C 6	1.4(2)
C 1	C 3	C 4	C 5	24.0(2)	C 1	C 3	C 4	C14	142.6(4)
O 2	C 3	C 4	C 5	138.8(3)	O 2	C 3	C 4	C14	-102.6(3)
C 3	C 4	C 5	C 7	-144.3(3)	C 3	C 4	C 5	N 1	-26.9(2)
C14	C 4	C 5	C 7	103.2(3)	C14	C 4	C 5	N 1	-139.4(4)
C 3	C 4	C14	C15	-59.9(2)	C 3	C 4	C14	C19	120.9(3)
C 5	C 4	C14	C15	58.8(2)	C 5	C 4	C14	C19	-120.5(3)
C 4	C 5	C 7	C 8	-62.9(2)	C 4	C 5	C 7	O 3	59.3(2)
N 1	C 5	C 7	C 8	-174.1(4)	N 1	C 5	C 7	O 3	-51.9(2)
C 4	C 5	N 1	C 1	17.4(2)	C 4	C 5	N 1	C 6	-167.0(3)
C 7	C 5	N 1	C 1	141.6(3)	C 7	C 5	N 1	C 6	-42.8(2)
C 5	C 7	C 8	C 9	-54.5(2)	C 5	C 7	C 8	C13	125.1(3)
O 3	C 7	C 8	C 9	177.6(3)	O 3	C 7	C 8	C13	-2.9(2)
C 7	C 8	C 9	C10	177.1(4)	C13	C 8	C 9	C10	-2.3(2)
C 7	C 8	C13	C12	-178.4(4)	C 9	C 8	C13	C12	1.1(2)
C 8	C 9	C10	C11	1.3(2)	C 9	C10	C11	C12	.6(2)
C10	C11	C12	C13	-1.8(2)	C11	C12	C13	C 8	.8(2)
C 4	C14	C15	C16	-178.3(4)	C19	C14	C15	C16	.8(2)
C 4	C14	C19	C18	178.4(4)	C15	C14	C19	C18	-.9(2)
C14	C15	C16	C17	.0(2)	C15	C16	C17	C18	-.6(2)
C16	C17	C18	C19	.5(2)	C17	C18	C19	C14	.4(2)

2.3. 绘图 见图1~3.

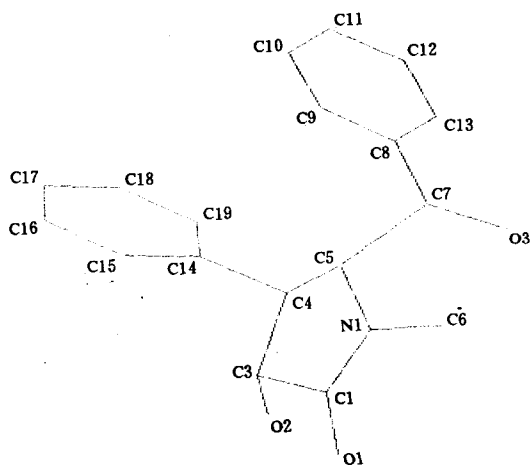


Fig 1 Drawing diagram.

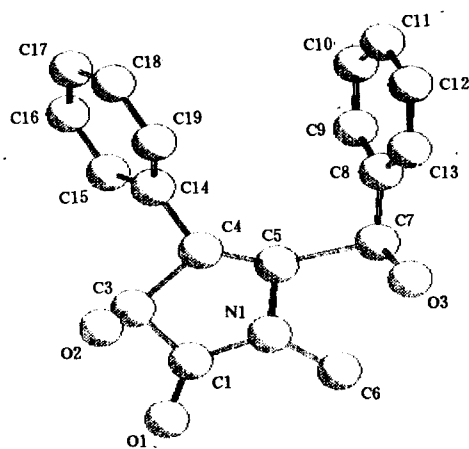


Fig 2 Setredrawing.

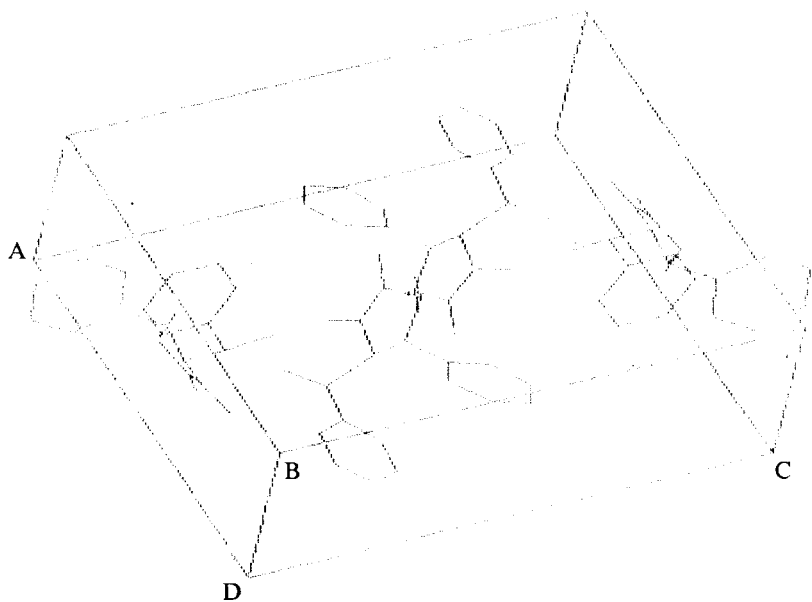


Fig 3 The molecular packing in unit cell.

## 结 语

1. “图谱库”主要供药物化学、植物化学实验室使用。我国的中草药化学成分晶体结构工作历经三十余年的发展,由于缺少晶体学专门刊物,特别是一部分早期工作,数据不够完善(如缺少氢原子位置,缺少标准偏差等),刊印漏误较多,生物活性记录不详。因此,目前提出的图谱库尚存在不足之处,希望通过药化、植化实验室的使用得到完善,以期成为开展新药研究的一个基本数据库,并逐步能与有可能建立的常见中药化学成分数据(库)联享共用

2. “图谱库”将在实现 X-衍射成为化学实验室常规分析方法的历史性转折阶段中,伴随化学家进入 X-衍射分析领域,并成为他们研究新药的得力助手。

3. 中草药是我国发现新药的重要来源。积国人数千年之耕耘,汇集了诸如《神农本草经》、《本草纲目》等各种药学珍籍。世界之药学研究源于天然药物,中国药学研究之优势仍在中草药。在以现代分析方法(如波谱分析和 X-衍射分析)与现代计算手段(微机乃至工作站)窥视药物结构与活性,并在分子水平上深入研究其机理以期创造出源于自然又超于自然的新药时,“图谱库”以原子、分子水平展现出药物分子立体图象及定量立体结构参数,这将为药学家在创造、发现新的活性药物分子的里程中,立下一个小小的路标。倘若在对常见的数百种中草药的化学成分系统研究的同时,从生态学与植物化学两个系统分类汇集其化学成分、生物活性之数据(包含各种波谱数据),这将是中草药学的“大典”。

4. 汇集了约250个晶体结构,已编为《中草药化学成分的晶体结构图谱库系统》(1960~1990),可以作为本库的“使用指南”,内中按分类列出化合物分子式、化合物名称、生物活性、作者、刊物、晶体学参数、原子坐标参数与分子结构立体图等信息;计算与绘图功能将以软件方式提供。进一步的工作欲将世界上已测定的中草药化学成分晶体结构数据汇编入本库,以扩展其应用。

“图谱库”的全部内容写入 CFTD(Chinese folk and tradition database)软件。使用者可按需要

选择“使用指南”或“CFTD”软件。由于 CFTD 软件可运行于 286, 386, 486 及其兼容的微机上, 故而便于推广, 各类化学实验室均可配置。

**致谢** 田之悦、管彦、杨晶、于德洁等同志参加了数据录入工作。鲍光宏、王奇光、潘德济教授等提供了所在实验室的晶体学数据。

### 参 考 文 献

- 1 吕扬, 等. “天然产物晶体结构微机系统(II)——天然有机分子晶体结构数据库”第四届全国天然有机化学学术讨论会论文摘要集(上) 4 1990, 11, 上海.
- 2 Lu Y, et al. The crystal structure database for active molecules of Chinese medicine. *International Symposium on New Drug Research and Development, Collection of Papers and Abstracts*, 237, Oct. 22~25, 1991 Beijing, China.
- 3 Lu Y and Wu BM. A microcomputer analytic system for crystal structure used in chemical laboratories. *Chin Chem Lett* 1992;3:637.
- 4 伍伯牧、吕扬. 晶体结构微机解析系统——NOMCSDP 程序包. *药学期刊* 1992;27:522.

## THE CRYSTAL STRUCTURE ATLAS DATA BASE SYSTEM FOR CHEMICAL CONSTITUENTS OF CHINESE TRADITIONAL AND FOLK MEDICINE

Y Lu, QT Zheng and BM Wu\*

(*Institute of Materia Medica, Chinese Academy of Medical Sciences, Beijing 100050;*

*\* Institute of Biophysics, Academia Sinica, Beijing 100101)*

**ABSTRACT** The atlas data base system for chemical constituents of Chinese traditional and folk medicine is the first one for single crystal structure in China. It includes about 250 crystal structures of chemical ingredients of Chinese traditional and folk medicine. It has searching, plotting and computing functions. It is a useful reference base. All softwares are written in dbaseIII, FORTRAN and Assembler Languages. They can be run on PC—286, 386, 486 and their compatible microcomputers.

**Key words** Chinese traditional and folk medicine; Crystal structure; Atlas data base