



本文经编委遴选,英文版将通过 ScienceDirect 全球发行。

常用药用辅料数据库的设计与构建*

□丁振浩 (北京中医药大学中药学院 北京 100102)

戴幸星 (首都医科大学中医药学院 北京 100069)

史新元** 乔延江** (北京中医药大学中药学院 北京 100102)

摘要: 采用 MDL ISIS/Base 数据库管理系统建立了常用药用辅料数据库, 该数据库共收录了 270 种辅料。数据库由辅料基本信息、理化信息和结构信息 3 部分内容构成。其中, 基本信息主要包括辅料的中英文名称、来源与制法、管理状态等 17 个字段; 理化信息主要包括辅料的理化性质, 如熔点、溶解性等 11 个字段; 结构信息主要是由材料计算平台 Materials Studio 4.1 中的 Synthia 模块计算得到的部分聚合物辅料性质数据。该数据库为药用辅料信息的查询提供了极大便利, 同时也为基于数据库的数据挖掘提供了研究平台。

关键词: 药用辅料 数据库 MDL ISIS/Base 数据库管理系统

doi: 10.3969/j.issn.1674-3849.2011.04.005

辅料是药物制剂生产必不可少的重要组成部分。它不仅赋予药物一定剂型, 且对药物作用的速度、生物利用度、毒副作用、稳定性及药效发挥等都有很大影响^[1]。辅料在中药制剂研发、生产的各个环节都发挥着非常重要的作用。因此, 研究开发、合理应用辅料, 对提高中药制剂质量和开发新剂型、新品种具有重要意义。目前国内外多家组织机构建立起了具有不同特色的药用辅料数据库。欧美几家大型制药企业联合开发了只对内部开放的药用辅料数据库[△]; FDA 建立了包括辅料的非活性药物数据

库^{△△}; 国际药用辅料网^{△△△}根据《中国药用辅料》部分内容, 收集和整理了目前常用的药用辅料, 按“通用名、别名、化学名和 CAS 注册号、分子式和分子量、制备工艺、类别、制剂应用、性状、稳定性和贮藏条件、配伍禁忌、安全性”等项目对资料进行了总结归类, 可按辅料功能类别或辅料名称进行检索, 支持模糊查询。这些数据库对现有辅料信息进行了系统整理, 可用于辅料性质的查询, 为制剂过程辅料的应用提供有力支持。

本文初步构建了规范化的、有利于辅料信息数据挖掘的药用辅料数据库, 为后期基于数据库的数

收稿日期: 2010-11-29

修回日期: 2010-12-04

* 国家自然科学基金资助面上项目(81073058/H2806): 基于计算机模拟的中药皂苷类成分复方配伍增溶机理研究, 负责人: 史新元; 北京中医药大学青年教师资助项目(2009JYB22-JS036): 基于计算机模拟的增溶性辅料筛选方法研究, 负责人: 史新元。

** 通讯作者: 史新元, 副教授, 硕士生导师, 主要研究方向: 中药信息工程, E-mail: xyshi@126.com; 乔延江, 本刊编委, 教授, 博士生导师, 主要研究方向: 中药信息工程、中药新药研发, E-mail: yjqiao@263.net。

△ https://www.lhasalimited.org/resources/FELASA_2007.pdf.

△△ FDA 官网 <http://www.accessdata.fda.gov/scripts/cder/iig/index.cfm>.

△△△ 国际药用辅料网 <http://www.phexcom.cn/index.asp>.

据挖掘提供研究平台。目前,该药用辅料数据库共收录了 270 种药用辅料。

一、药用辅料数据库的构建

常用药用辅料数据库是将辅料的基本信息、理化信息和结构信息数据融合,面向药学领域的研究人员,建立起易于使用的药用辅料数据库系统。本文系统地收集、整理常用药用辅料的化学结构、物理性质等信息,采用 ISIS/BASE 数据库管理软件建立药用辅料数据库系统,为药物的研发与生产及有关部门的决策和监管提供有力支持。

1. 数据来源

为保证辅料数据库数据的权威性,针对入选文献制定了相应的遴选原则:一是国家法定标准、法规;二是著名学术团体、机构编写的专著。主要包括《中华人民共和国药典》(二部)^[2]、《中国药用辅料》^[3]、《Handbook of Pharmaceutical Excipients》(6th Edition)^[4]。部分聚合物辅料的计算性质通过 Accelry 公司材料计算平台 Materials Studio 4.1 中 Synthia 模块计算得到。

2. 数据库结构设计

数据库采用美国信息系统公司开发的 ISIS/Base 数据库管理系统^[5],它可在单机上运行,也可作为客户端使用。

该数据库能够结合 ISIS/Host 软件建立客户机/服务器式(Client/Server,C/S)结构的数据库系统。在 C/S 中,客户端将数据传送到服务器,服务器处理后,将结果返回客户端(见图 1),从而减少了网络上的数据传输量,提高系统的性能和负载能力^[6]。

本文采用 ISIS/Base 作为数据库开发工具,建立了对话框、控件和后台数据库。数据库包括药用辅料基本信息、基本性质和计算性质 3 张表格,以辅料的 ID 为主键(见图 2),可通过后台参数“Change form button”命令实现不同信息表间的切换,见图 3。基本信息表格共 17 个属性,包含中文名称、英文名称、别名、CAS 号、来源与制法、形状、应用实例、安全性、配伍禁忌、作用与用途、管理状态、包装与贮藏等信息(见图 4)。基本性质表格共 11 个属性,包含熔点、溶解性、密度、闪点、酸碱度、临界胶束浓度等信息(见图 5)。计算性质主要包括 6 大类:(1)结构性质,如重复单元

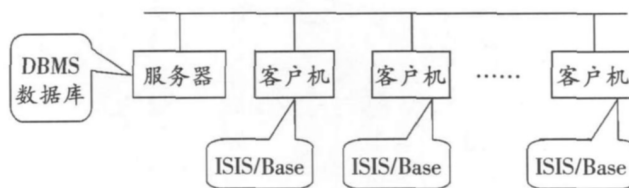


图 1 客户机/服务器(Client/server,C/S)结构

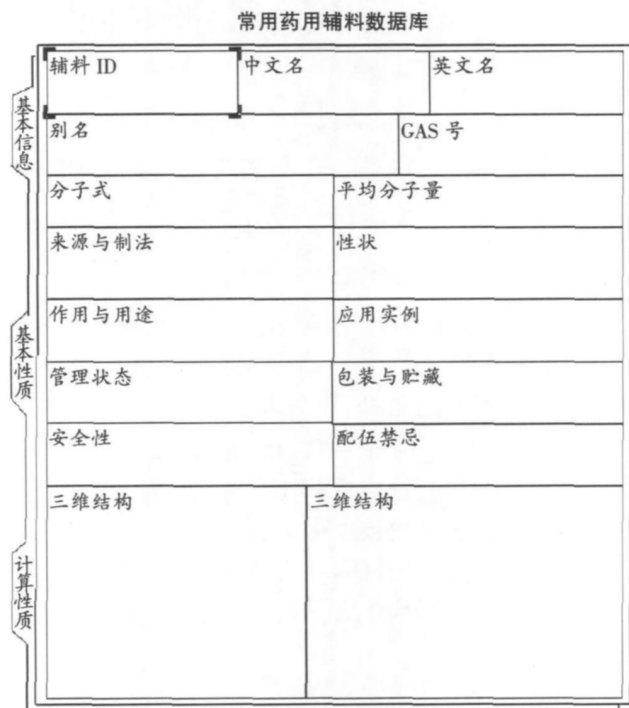


图 2 数据表格设计图

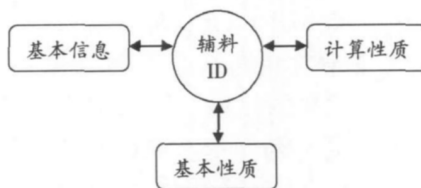


图 3 药用辅料数据库实体关系图

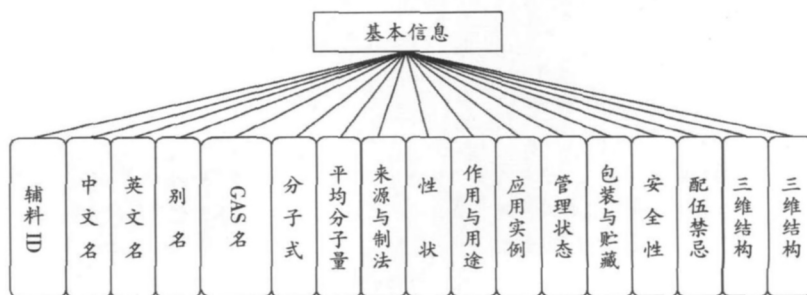


图 4 基本信息表格示意图

的分子量、重复单元的长度、连接性指数等；(2)热力学性质,如 VdW 体积、玻璃转化温度 T_g、体积热膨胀系数、摩尔体积、内聚能、溶解度参数、表面张力等；(3)电学、光学、磁学性质,如折射率、介电常数等；(4)机械性质,如体积模数、杨系数、泊松比等；(5)链的伸缩和弯曲,如特征比、摩尔直链函数、卷曲分子量等；(6)传输性质,如粘滞流动的激活能、氧渗透率、氮渗透率等,见表 1。

3. 数据库的规范化方法

(1)量纲的规范化处理。

同一物理量有不同的单位,为了方便查看比较,所有物理量尽可能采用法定单位,如表面张力的单位有牛顿/米(N/m)和达因/厘米(dyn/cm),本数据库统一采用法定单位牛顿/米(N/m)。

(2)基本信息规范化处理。

辅料包含信息面范围较广,为确保数据库的简洁性,本数据库中只选取具有代表性的信息。如数据库溶解度属性中只录入辅料在较为常用的溶剂(水、甲醇、乙醇等)中的溶解度数据;对辅料性状信息进行精炼,数据库只对辅料的色、香、味及存在状态进行了描述。

4. 化学结构的生成

采用 ISIS/Draw 绘图软件绘制辅料的二维结构式并存储为 mol 格式,再由 3D 结构转换软件 Corina 软件转换得到辅料的三维结构式。数据库中化学结构可导出成 mol 格式,直接用于分子模拟和分子设计软件的虚拟筛选,为根据辅料化学结构预测其理化性质提供载体。

二、数据库的运行环境

Intel P 或更高性能的微处理器 ;VGA 或更高性能的显示器 ;Windows 环境 ;64MB 以上内存 ;20MB 以上硬盘空间 ;ISIS/Base 数据库管理系统^[5]。

三、数据库的功能和特点

本数据库具有强大的数据存储、管理功能,可对化学结构和文本数据进行灵活检

索,对每个字段均可进行精确或模糊查询。可利用 Query Builder 功能自定义检索式,通过任意组合字段构建联合查询,提供独特的查询功能。

数据库检索结果以表格的形式呈现给用户预览,

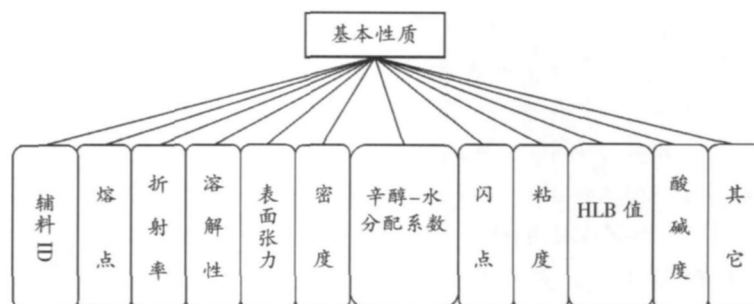


图 5 基本性质表格示意图

表 1 计算性质表格包含的属性

属性类别	属性名称	属性类别	属性名称
结构性质	重复单元分子量	机械性质	体积模数
	重复单元长度		剪切弹性模数
	重复单元原子数		杨系数
	重复单元非氢原子数		泊松比
	连接性指数 ⁰ χ		剪切屈服应力
	连接性指数 ⁰ χ'		脆性断裂应力
	连接性指数 ¹ χ		
	连接性指数 ¹ χ'		
热力学性质	范德华体积	高分子链伸缩与弯曲性质	空间位阻参数
	摩尔体积		特征比
	热膨胀系数		摩尔直链函数
	密度		卷曲分子量
	固体摩尔热容		卷曲分子长度
	液体摩尔热容		临界分子量
	玻璃转化温度		
	T _g 下的摩尔热容变化值		
	半分解温度		
	内聚能		
电学、光学、磁学性质	溶解度参数	传输性质	黏流活化能
	表面张力		氧渗透率
	导热系数		氮渗透率
	折射率		二氧化碳渗透率
	摩尔折射率		零剪切粘滞
	介电常数		
	体积电阻率		
	抗磁化率		

同时用户可将所需信息进行各种格式的输出,以建立自己的研究报告。用户可将 ISIS/Base 管理的数据,利用 MDL 公司所提供的另一个产品 ISIS for Excel,输出至微软的 Excel,以建立结构与活性和结构与性质关系表,并对研究结果进行统计和分析,还可直接将药用辅料的化学结构、理化性质数据导入 QSAR 软件包,计算其分子描述符,通过化学计量学方法建立辅料结构-性质关系模型。

本数据库还可以利用通用的接口与其他系统,如 Tripos 和 Accelrys 分子模拟和分子设计软件配合使用,将数据库中的化学结构直接导出用于介观动力学模拟及分子动力学模拟等。如本课题组刘南岑等^[7]从辅料数据库中查询泊洛沙姆 188 分子结构,导入 Accelrys 公司的 Materials studio 模拟软件,通过介观模拟方法考察泊洛沙姆 188 与胆酸的聚集形态。

该数据库具备良好的可扩展性,能较为简便的扩充数据库;还可非常简便地向关系型数据库迁移*,从而具备更加良好的通用性及与 Internet 通讯的能力。它可使用 SQL 语句在服务器和客户机之间传送请求,把数据库放在服务器端,在客户机上的 SQL 操作指令通过数据库管理系统发送给服务器,服务器的数据库管理系统解释并执行这些命令,然后返回给客户机,通过 Web 服务器、应用服务器和数据库实现基于 Internet 应用的数据库信息系统。

四、数据库的应用实例

为了探讨辅料分子结构与其增溶性能间的关系,本课题组从数据库中导出具有表面活性的辅料化学结构及增溶性信息,文献报道可作为增溶剂且溶于水的分子归为阳性分子(即具有增溶性能分子),文献报道不溶于水的分子归为阴性分子(即没有增溶性能分子)。计算表征分子结构特征分子描述符,采用支持向量机方法建立了预测辅料增

溶性能的判别模型。该模型具有良好的预测能力,能够为难溶性药物制剂开发过程中增溶剂的筛选提供指导。

五、总结

本文利用 ISIS/Base 建立了常用药用辅料数据库,内容权威,界面友好,针对性强,易于查询。该数据库对辅料信息进行了规范化处理,可直接用于数据挖掘研究,从而能够借助数据挖掘方法在海量数据信息中挖掘潜在的、有用的信息。

目前药用辅料的标准化建设还存在许多问题,其中又以中药辅料的标准化最为复杂。药用辅料数据库的构建将为中药辅料标准化建设提供有用信息,也为中药制剂的研发与生产提供有力支持,进一步推进中药制剂的现代化。构建药用辅料数据库,对长期累积的数据进行信息化、数字化和知识化研究,将对我国制药工业发展产生重大意义。目前,该药用辅料数据库收录的辅料种类尚不完善,有待对数据库进一步扩充,最终建立一个具备权威性、完备性、新颖性以及界面友好性的药用辅料数据库。

参考文献

- 1 张志芬. 从现代药剂学看药用辅料的发展. 药物分析杂志, 2005, 25(12):1576~1580.
- 2 罗明生, 高天慧, 宋民宪. 中国药用辅料. 北京: 化学工业出版社, 2006.
- 3 国家药典委员会. 中华人民共和国药典(二部). 北京: 化学工业出版社, 2005.
- 4 Raymond C., J. Paul, Sian C. Handbook of pharmaceutical excipients. Washington: American Pharmaceutical Association (APhA), 2009.
- 5 刘颖, 王耘, 乔延江. 常用活血化瘀类中药组分数数据库的构建. 计算机与应用化学, 2006, 23(11):1130.
- 6 刘南岑, 史新元, 乔延江. 增溶性辅料泊洛沙姆 188 与胆酸聚集形态的介观模拟. 中国科学 B 辑, 化学, 2010, 41(3):500~508.
- 7 雷静. ISIS Base 化学数据库向关系数据库的迁移. 计算机与应用化学, 2005, 22(11):1044~1046.

Design and Development of Pharmaceutical Excipients Database

Ding Zhenhao¹, Dai Xingxing², Shi Xinyuan¹, Qiao Yanjiang¹

(1. School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 100102, China;

* ISIS/Desktop 2.1.4 手册. 美国 MDL 信息系统有限公司.

2. *School of Traditional Medicine, Capital Medical University, Beijing 100069, China*)

Abstract: Pharmaceutical excipients database (PED) was established on the basis of MDL ISIS/Base Database Management System (DBMS). Currently, 270 pharmaceutical excipients were included. The pharmaceutical excipients database consists of three tables. The first table is basic information that contains 17 attributes in total, such as Chinese name and English name of excipients and so on. The second table is the basic property table which contains physical and chemical property and so on. There are 11 attributes in total. The third table is the computed property that was calculated by Synthia module of Materials Studio 4.1. The PED not only provides great convenience for the query of pharmaceutical excipients, but also a data mining platform on the pharmaceutical excipients.

Keywords: Pharmaceutical excipients, database, MDL ISIS/Base

(责任编辑 李沙沙, 责任译审 王 晶)