

YAOYONG TIANRAN CHANWU

药用天然产物

徐筱杰 康文艺 主编



Chemical Industry Press

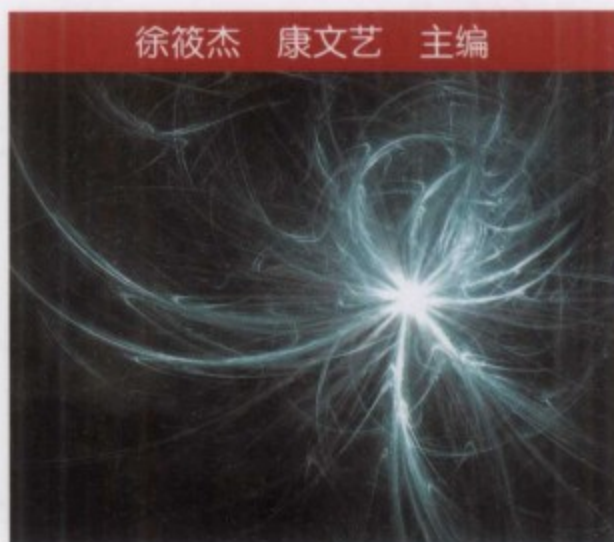


化学工业出版社

YAOYONG TIANRAN CHANWU

药用天然产物

徐筱杰 康文艺 主编



ISBN 978-7-122-06776-0



9 787122 067760 >

定价：66.00元

编写人员

(按姓氏音序排列)

陈丽蓉	侯廷军	黄 钦	康文艺
李昌勤	李旭东	李正全	骆宏鹏
乔学斌	谢建春	徐筱杰	章 威
朱丽荔	庄 艳		



前 言

天然产物作为药物已有悠久的历史，在 1970 年以前几乎所有的药物都是来自天然产物及其衍生物。虽然 20 世纪 90 年代由于组合化学及高通量筛选技术的发展使天然产物在药物研究与开发中的比重有所下降，但天然产物仍是创新药物研究及开发的重要资源。据统计，1981 年 1 月至 2006 年 9 月新增加药物中来源于天然产物及天然产物衍生物的占 28%。据统计，在全球范围内已发现的植物约 42 万种，其中药用植物约 5 万种以上，中药植物约 11000 种。虽然经过数百年的努力，但已经过植物化学及药理学研究的植物不足 10%。因此药用天然产物研究仍任重道远。

药用天然产物研究涉及生物、化学、医学、药学及食品科学等多种学科，是多学科交叉的重要研究领域。药用天然产物研究包括植物、动物及微生物中小分子的活性筛选、分离与纯化、结构鉴定、生物与化学转化及生物功能研究。这些研究有助于化学、生物及药学知识的丰富与深化，为新药的设计、研究及开发提供思路及科学基础，也有助于食品添加剂及保健品的研究与开发。药用植物体系是天然的组合化学库，具有丰富的结构及生物活性多样化。与合成的组合化学库比较，药用植物体系组成复杂，含量悬殊，而且许多成分是未知的。

传统的药用天然产物研究途径是选取一种植物，经系统的分离、纯化得到单一化合物，然后进行生物活性测定，获得活性化合物，或经过生物活性指导下的分离纯化得到活性化合物（从植物到活性分子）。传统方法是目前药用天然产物研究的重要途径，已取得重大的成效。但传统方法周期长，容易丢失微量的活性成分。因此如何从这个复杂体系中高通量筛选活性化合物及快速分离活性化合物一直是药用天然产物研究的瓶颈及挑战。多年来，我们根据分子识别原理（几何形状匹配、能量匹配、预组织、诱导匹配等概念）及化学基因组学两个基本假设（分享一定化学相似性的化合物也应分享同样的靶标，分享同样配体的靶具有相似的结合位点），发展了先发现活性分子，然后再从植物富集、分离得到该活性分子的反向途径（从活性分子到植物）。该反向途径包括计算机虚拟筛选、基于固定化靶酶或多克隆抗体的亲和前沿色谱技术及分子烙印聚合物固相萃取技术等。本书在编著者多年科研积累的基础上介绍传统方法及反向途径各种技术，包括方法的基本原理、基本知识、基本技能及主要进展。我们希望本书对从事天然产物、药物化学、食品化学、分析化学研究的科研工作者有所助益。欢迎读者在使用过程中对本书提出批评指正。

本书第 1 章、第 6~8 章由北京大学化学学院徐筱杰主编，第 2~5 章由河南大学药学院康文艺主编。本书取材于侯廷军、章威、乔学斌、黄钦、李旭东的计算机虚拟筛选的研究，北京大学化学学院陈丽蓉、庄艳、朱丽荔、李正全、谢建春、骆宏鹏等的实验筛选方法的研究。河南大学李昌勤参与了部分编写工作。

编著者
2009 年 9 月



本书系编者多年教学及科研的积累编写而成。全书共八章，内容可分为：药用天然产物的基础知识及传统研究方法（包括药用天然产物研究的历史及发展概况，天然产物的生物合成，化学成分的提取、分离与纯化，化学成分的分析及结构鉴定，成分的生物转化等五章内容）；编者根据分子识别原理发展的计算机虚拟筛选、活性化合物的实验快速分离及高通量筛选方法（包括计算药理学在药用天然产物研究中的应用，基于亲和前沿色谱-质谱联用技术的药用天然产物筛选方法，药用天然产物的筛选与分离一体化技术等三章内容）。这些新技术可为我国基于天然产物的创新药物研究及中药现代化研究提供新途径和新思路。

本书可作为天然产物、药物化学、食品化学、分析化学等相关学科的教材，也可供相关学科的科研人员学习参考。

图书在版编目（CIP）数据

药用天然产物/徐筱杰，康文艺主编. —北京：化学工业出版社，2009.12

ISBN 978-7-122-06776-0

I. 药… II. ①徐…②康… III. 生药学-研究 IV. R93

中国版本图书馆 CIP 数据核字（2009）第 179287 号

责任编辑：刘俊之

文字编辑：朱 恺

责任校对：洪雅姝

装帧设计：韩 飞

出版发行：化学工业出版社（北京市东城区青年湖南街 13 号 邮政编码 100011）

印 装：大厂聚鑫印刷有限责任公司

720mm×1000mm 1/16 印张 19 字数 367 千字 2010 年 1 月北京第 1 版第 1 次印刷

购书咨询：010-64518888（传真：010-64519686） 售后服务：010-64518899

网 址：<http://www.cip.com.cn>

凡购买本书，如有缺损质量问题，本社销售中心负责调换。

定 价：66.00 元

版权所有 违者必究

目 录

第 1 章 药用天然产物及其研究方法	1
1.1 药用天然产物	1
1.1.1 历史	1
1.1.2 天然产物在药物开发中的现状	2
1.1.3 天然产物在药物开发中的前景	5
1.2 天然产物的治疗领域	7
1.2.1 感染疾病	7
1.2.2 抗肿瘤药物	8
1.2.3 心血管及内分泌疾病	8
1.3 天然产物研究方法	8
1.3.1 传统方法	8
1.3.2 活性化合物的筛选新方法	14
1.4 药用天然产物的反向途径	16
参考文献	17
第 2 章 天然产物的生物合成	20
2.1 概述	20
2.2 天然产物的生物合成途径	21
2.2.1 一次代谢与二次代谢	21
2.2.2 生物合成假说的提出	22
2.2.3 主要的生物合成途径	24
2.3 天然产物生源合成研究应用	32
2.3.1 植物学领域	33
2.3.2 植物化学领域	36
2.3.3 生物化学与生物工程领域	38
参考文献	39
第 3 章 药用植物化学成分的提取、分离与纯化	40
3.1 天然产物化学成分的提取	40
3.1.1 药用天然产物经典提取方法	40
3.1.2 天然产物化学成分的现代提取方法	46
3.2 天然产物的分离	47

3.2.1 天然产物的经典分离方法	48
3.2.2 天然产物的现代分离方法	61
参考文献	65
第4章 天然产物化学成分分析及结构鉴定	66
4.1 天然产物化学成分结构鉴定方法学	66
4.1.1 核磁共振技术	67
4.1.2 有机质谱	71
4.1.3 其他波谱技术	71
4.2 天然产物化学成分的结构鉴定	72
4.2.1 萜类化合物	72
4.2.2 黄酮类化合物	77
4.2.3 苯丙素化合物	79
4.2.4 生物碱类化合物	86
4.2.5 甾体类化合物	88
参考文献	90
第5章 药用植物活性成分的生物转化	91
5.1 概述	91
5.2 药用植物萜类活性成分的生物转化	93
5.2.1 紫杉醇生物转化	93
5.2.2 青蒿素的生物转化	98
5.2.3 其他萜类化合物的生物转化	104
5.3 药用植物生物碱类活性成分的生物转化	112
5.3.1 长春花生物碱的生物转化	113
5.3.2 喜树碱的生物转化	117
5.3.3 其他类型生物碱的生物转化	120
5.4 药用植物甾体类活性化合物的生物转化	121
5.4.1 甾体化合物的生物合成	121
5.4.2 甾体类活性成分的生物转化	122
5.5 其他类型活性成分的生物转化	126
5.5.1 黄酮类化合物的生物转化	126
5.5.2 蒽醌类化合物	128
5.5.3 其他类型活性化合物的生物转化	130
参考文献	131
第6章 计算药理学在药用天然产物研究中的应用	136
6.1 计算机辅助药物分子设计 (CADD)	136

6.1.1	计算机辅助药物分子设计方法	136
6.1.2	CADD 中的几个难题	138
6.1.3	计算机辅助药物分子设计的前景	139
6.2	构建数据库	141
6.2.1	剑桥结构数据库 (CSD)	141
6.2.2	CrossFire Beilstein 数据库	142
6.2.3	MDL [®] 代谢数据库	142
6.2.4	MDL [®] 毒性数据库	143
6.2.5	NCI 数据库	143
6.2.6	天然产物数据库 (DNP)	144
6.2.7	中草药三维结构、性质数据库 (CHDD)	145
6.2.8	中国天然产物数据库 (CNPDP)	145
6.3	分子对接方法	147
6.3.1	分子对接方法的分类	147
6.3.2	分子对接方法中的重要问题	148
6.3.3	几种有代表性的分子对接方法	149
6.4	定量构效关系方法研究	155
6.4.1	二维定量构效关系方法	155
6.4.2	三维定量构效关系方法	156
6.4.3	建立定量构效关系的统计方法	157
6.5	结合自由能的计算	158
6.5.1	自由能计算方法	159
6.5.2	自由能微扰 (FEP) 方法	159
6.5.3	热力学积分 (TI) 方法	160
6.5.4	线性相互作用能 (LIE) 方法	161
6.5.5	基于 MM/PBSA 的自由能计算方法	163
6.6	计算机辅助的化合物药代动力学特征及毒性的预测	164
6.6.1	药代动力学特征的预测	164
6.6.2	化合物毒性预测方法	177
6.7	中草药有效成分化合物的化学空间分析	185
6.7.1	介绍	185
6.7.2	分子描述符的计算	186
6.7.3	结果与讨论	186
6.7.4	主成分分析	191
6.7.5	结论	191
6.8	血府逐瘀汤潜在协同作用及抑制多个靶酶的计算探讨	192

6.8.1	介绍	192
6.8.2	方法和预处理	192
6.8.3	血府逐瘀汤组分的化学空间分析	193
6.8.4	血府逐瘀汤与心血管系统靶酶的相互作用研究	196
6.8.5	血府逐瘀汤药物代谢动力学性质的研究	199
参考文献		200
第7章 基于亲和前沿色谱-质谱联用技术的药用天然产物筛选方法		202
7.1	亲和色谱及前沿色谱介绍	202
7.1.1	亲和色谱	202
7.1.2	亲和前沿色谱	203
7.2	亲和前沿色谱-质谱联用	205
7.2.1	原理与装置	205
7.2.2	FAC-MS 的应用	206
7.2.3	总结与展望	212
7.3	阻断 SARS 侵入细胞的抑制剂的固定化酶亲和前沿色谱-质谱联用法直接 筛选	214
7.3.1	严重急性呼吸道综合症病毒 (SARS-CoV)	214
7.3.2	阻断 SARS 侵入细胞的抑制剂的筛选	216
7.3.3	小结	219
7.4	固定化多克隆抗体亲和前沿色谱-质谱联用法直接筛选	220
7.4.1	多克隆抗体的应用背景	220
7.4.2	固定化多克隆抗体亲和前沿色谱-质谱联用法直接从鬼箭锦鸡儿中筛选 EGFR 抑制剂	225
7.4.3	抗 EGFR 药物的初步筛选	232
7.4.4	多克隆抗体的制备	233
7.4.5	多克隆抗体的纯化及其性质评价	235
7.4.6	多克隆抗体的固定化与亲和前沿色谱	238
7.4.7	多克隆抗体识别鬼箭锦鸡儿中活性成分	239
7.4.8	小结	242
参考文献		242
第8章 药用天然产物的筛选与分离一体化技术		245
8.1	分子烙印聚合物介绍	245
8.1.1	方法起源及概述	245
8.1.2	分子烙印聚合物的制备方法	246
8.1.3	分子印迹机理研究	258

8.1.4	分子烙印聚合物的应用范围	260
8.2	分子烙印聚合物的应用	260
8.2.1	固相萃取	260
8.2.2	液相色谱	262
8.2.3	传感器	263
8.2.4	模拟抗体	264
8.2.5	模拟酶催化	265
8.3	分子烙印固相萃取中药中特定药效组分——黄酮类	265
8.3.1	聚合物的制备	265
8.3.2	聚合物的选择性评价	266
8.3.3	聚合物固相萃取柱柱载量测定	268
8.3.4	固相萃取分离银杏叶中黄酮类药效成分	269
8.3.5	萃取分离沙棘草药粗提物中特定药效黄酮组分	270
8.4	分子烙印亲和分离在线鉴定中草药中有效成分	272
8.4.1	仪器和材料	272
8.4.2	聚合物制备	273
8.4.3	液相色谱评价聚合物对目标分子的选择性	274
8.4.4	前沿色谱方法测量结合位点数及离解常数	276
8.4.5	骆驼蓬样品的制备	278
8.4.6	MIP-LC-MS 在线分离鉴定	278
8.5	分子烙印识别中药中抗表皮生长因子受体 (EGFR) 的活性成分	283
8.5.1	聚合物制备	283
8.5.2	藏锦鸡儿粗提物的制备	284
8.5.3	槲皮素-MIP 柱固相萃取藏锦鸡儿中活性成分	284
8.5.4	白皮杉醇-MIP 柱识别藏锦鸡儿中活性成分	285
8.6	分子烙印在药物开发方面的应用前景	289
	参考文献	290

鄧平舫

PDG

第 1 章 药用天然产物及其研究方法

1.1 药用天然产物

1.1.1 历史

在现代西方医药学发展的数千年以前，人们就已经知道了植物以及其他生物含有生理活性物质，整个世界的各种文明都是围绕着如何将植物、动物、矿物作为治疗疾病的药物。天然产物（来自生物活体的化学物质）特别是来源于植物的天然产物作为药物使用的历史很悠久，最早的记载可追溯到公元前 2500 年，在美索不达米亚出土的数百页黏土板上已有详细的记载，在报道的药物中涉及上千种美索不达米亚植物，古巴比伦人用楔形文字记载了近千种来源于植物的物质，其中部分物质在今天仍用于咳嗽、感冒、寄生虫感染以及炎症的治疗。发现具有功效的植物主要通过偶然发现和仔细的观察，将两者结合的研究方法在当今药物化学的研究当中依然是非常普遍的。最早分离得到的天然产物是番木鳖碱 [图 1-1(a)]、吗啡 [图 1-1(b)]、阿托品 [图 1-1(c)]、秋水仙碱 [图 1-1(d)]，最早的天然产物抗生素是 β -内酰胺抗生素青霉素等。链霉素是第一个用于治疗肺结核的抗生素，天然产物抗病毒药物是海绵尿糖苷以及海绵胸苷，在此基础上发展了抗艾滋及其他抗病毒药物，可以认为它们是核苷类抗病毒药的先驱。

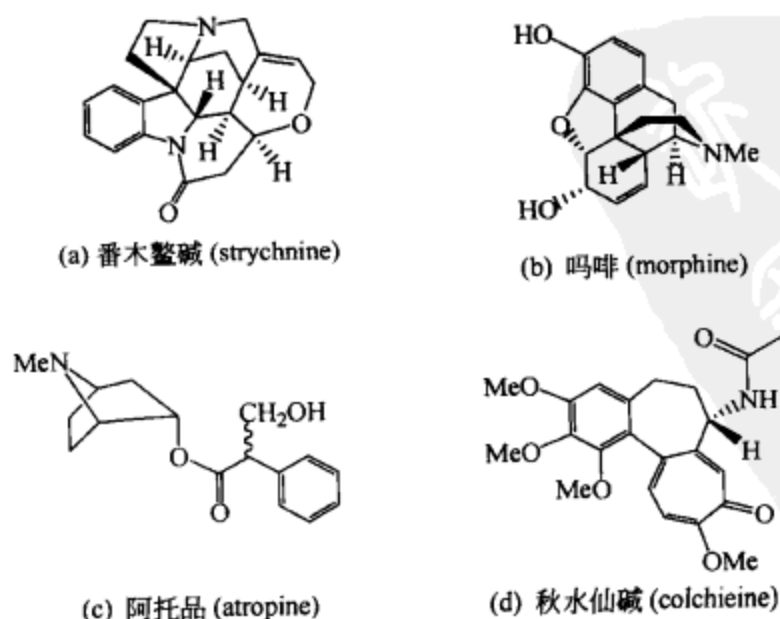


图 1-1 最早分离的几种天然产物

吗啡（morphine）是第一个商用纯天然药物，由默克公司（Merck Inc.）在 1826 年开始生产；阿司匹林（aspirin）则是第一个基于天然物的半合成药物，由拜尔公司（Bayer）在 1899 年开始生产。历史上，来源于植物的药物占据很重要的地位，尤其在药物开发早期，植物药举足轻重，其中在 1970 年以前几乎所有药物都是来于天然产物及其衍生物。在我国将植物、动物、矿物作为治疗疾病的药物已有数千年历史。

1.1.2 天然产物在药物开发中的现状

在 20 世纪 80 年代和 90 年代，许多制药公司对从天然产物开发新药减少了兴趣，其主要原因有两个。

① 20 世纪 70 年代末，随着组合化学快速发展，化学家可以在短时间内合成数万个化合物。

② 同时，随着自动化系统的高通量筛选（high-throughput screening, HTS）技术的发展，每天可以对数万样品进行筛选，为组合化学合成的样品提供了快速筛选方法。

由于 HTS 与以生物活性评价为指导进行天然生理活性物质分离发展不协调的缘故，从天然产物寻找新药被看作是一条既耗时又费钱且效率低的途径。

此外，天然活性物质分离的复杂化以及大量生物材料获得的问题，也是阻碍从天然产物开发新药的原因之一。因此，在 20 世纪 80 年代，制药行业对天然产物的研究跌到了谷底。然而，天然产物并没有完全失去竞争力，从 Cragg 等发表的文章可以看出，天然产物对人类所有疾病特别是癌症这样的特定领域产生了重要的影响。

图 1-2 是 1981 年 1 月～2006 年 6 月“New Chemical Entities”中新增的 1184

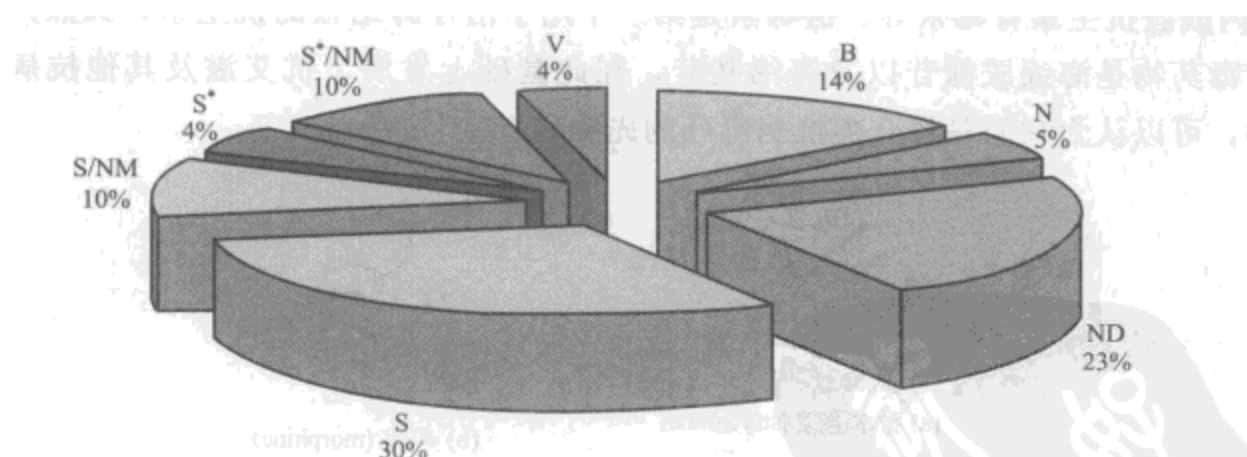


图 1-2 所有“New Chemical Entities”，1981 年 1 月～2006 年 6 月，
根据来源划分（总数为 1184）

B—生物制剂（疫苗、单克隆抗体等，来于哺乳动物），通常是大于 45 个残基的多肽或蛋白；

N—来源于未修改的天然产物；

ND—来源于天然产物，但经过修饰得到（如半合成物）；

S—全合成的药物，通常由随机筛选或由已知制剂修饰得到；

S*—全合成的药物，但药效团模型来自天然产物；

NM—天然产物模拟物；

V—疫苗

个化合物，它覆盖了所有疾病/国家/来源，其来源根据 Cragg 建立的 NM (natural product mimic) 分类方法来标明。其中来源于天然产物及天然产物改造的占 28%，并没有退出药物的舞台。再从图 1-3，可以看到来源于天然产物及天然产物改造的药物所占比例虽然越往后越有所下降，但其原因是生物制剂和疫苗的比例上升了，但即使如此，来源于天然产物及天然产物改造的药物依然占 20% 左右。

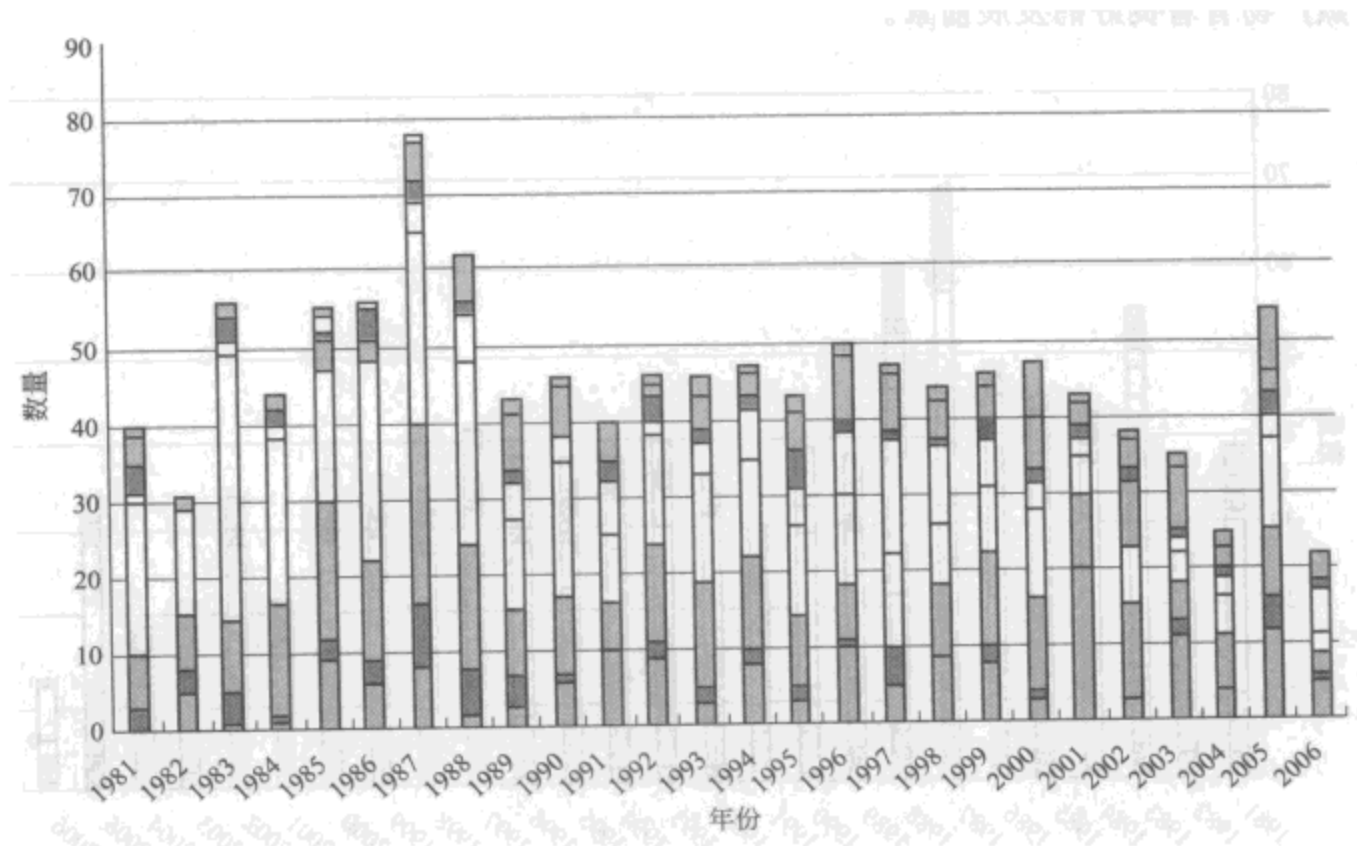


图 1-3 所有 “New Chemical Entities”，1981 年 1 月~2006 年 6 月，
根据来源和年份划分 (总数为 1184) (注：从下往上顺序为：
D, N, ND, S, S/NM, S*, S*/NM, V, 字母含义见图 1-2)

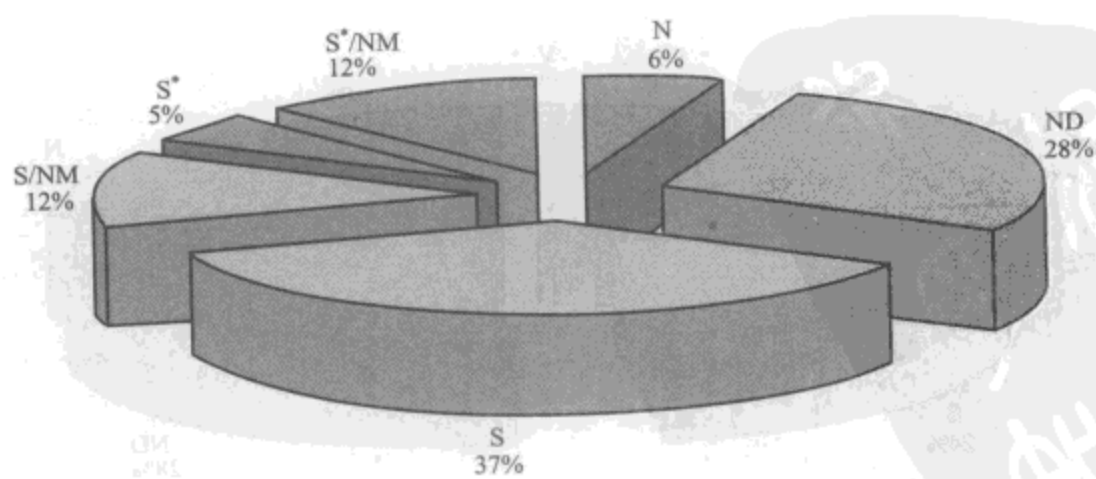


图 1-4 所有小分子 “New Chemical Entities”，1981 年 1 月~2006 年 6 月，
根据来源划分 (总数为 974) (注：图中字母含义见图 1-2)

从图 1-4 可以看出，如果不考虑生物制剂和疫苗，只考虑 25 年来新增小分子药物，那么，来源于天然产物及其改造的药物占到总数的 34%。从根据来源和年份划分的图 1-5 中可以看出来源于天然产物及天然产物改造的在新增小分子药物中所占的比例一直不低，即使在陷入谷底的 20 世纪 80~90 年代，也有很高的比例。再来看看图 1-6，在治疗癌症方面，来源于天然产物的药物分子占了 42%，未改造的天然产物分子占到了 14%，比总体的 5%要高出 9%，这说明在治疗癌症方面天然产物有着良好的发展前景。

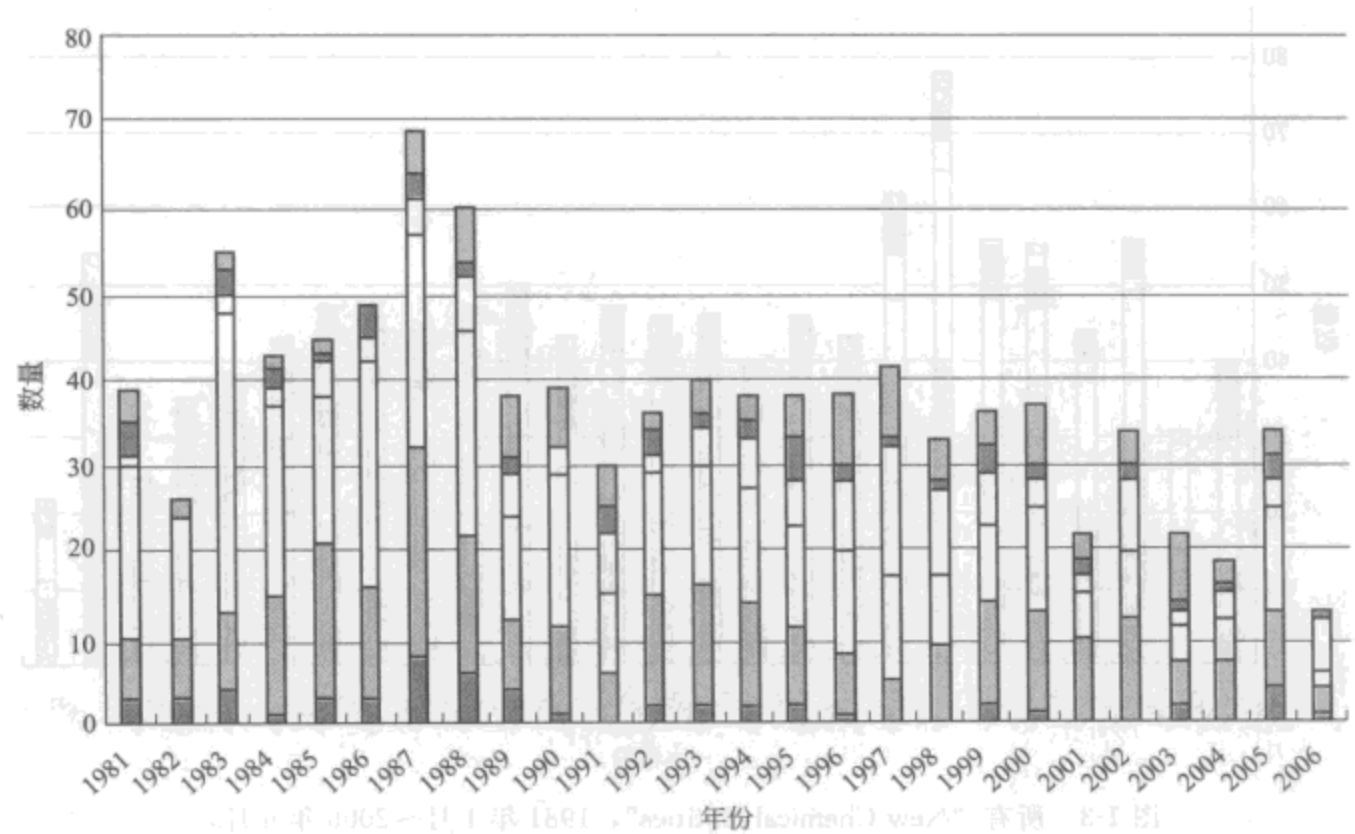


图 1-5 所有小分子 “New Chemical Entities”，1981 年 1 月~2006 年 6 月，根据来源和年份划分 (共 974) (注：从下往上顺序为：N，ND，S，S/NM，S*，S*/NM，字母含义见图 1-2)

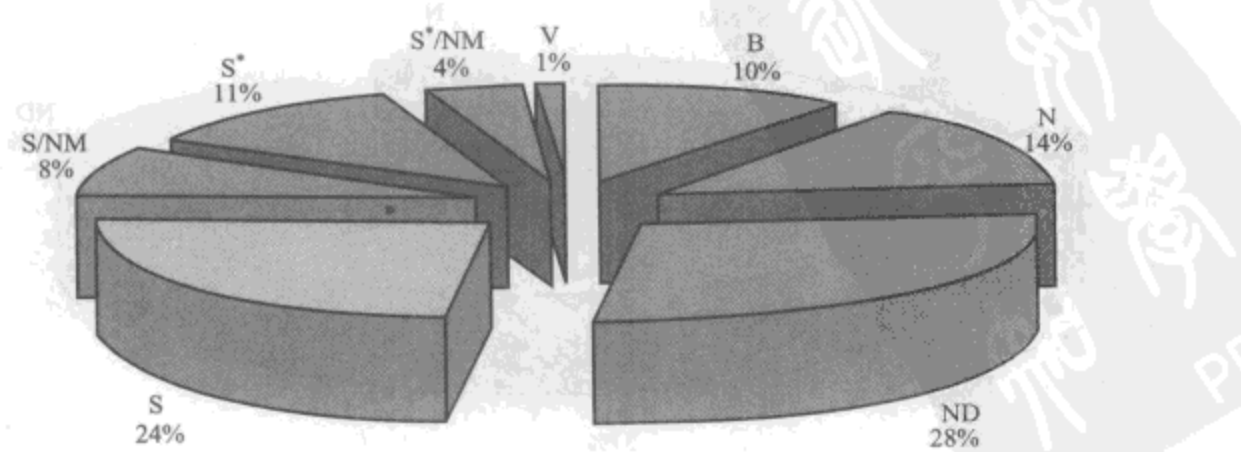


图 1-6 所有可用的抗癌药物分子，20 世纪 40 年代~2006 年 6 月，根据来源划分 (总数为 175) (注：图中字母含义见图 1-2)

1.1.3 天然产物在药物开发中的前景

目前超过 100 个与天然产物相关的化合物处于临床试验阶段，至少有 100 个相似的研究项目处于临床前研究阶段（见表 1-1），其中大多数来自植物及微生物。据统计从 2001~2008 年全球 312 药物研究项目中 30% 属于天然产物领域。目前除了应用纯的天然产物化合物外，人们对来自传统药物的提取物的兴趣也在不断增长，最近美国食品和药物管理局（FDA）批准了一个绿叶提取物作为新药物。

表 1-1 处于不同开发阶段的药物

开发阶段 (development stage)	植物 (plant)	细菌 (bacterial)	菌类 (fungal)	动物 (animal)	半合成 (semi-synthetic)	总数 (total)
临床前(preclinical)	46	12	7	7	27	99
一期(phase I)	14	5	0	3	8	30
二期(phase II)	41	4	0	10	11	66
三期(phase III)	5	4	0	4	13	26
注册前(pre-registration)	2	0	0	0	2	4
总数(total)	108	25	7	24	61	225

据世界卫生组织（WHO）估计，全球人口的 80% 目前仍在使用传统药物。在发展中国家，民族植物学及民族药理学仍是人们关注的学科，民族药为保障人民健康做出了巨大贡献，目前民族药已在一些国家进行大量的现代科学研究，并列入该国家的药典。但至今进行较系统研究的植物仍只占少数。如在巴西，大约有 5500 种植物，但已报道过的植物研究仅占 0.4%。目前已有 119 个从高等植物中提取的纯天然产物作为药物，其中 74% 化合物的相关来源植物早已作为药物使用。如玫瑰玉黍螺（长春花属，夹竹桃科），这种草本植物含有 75 个生物碱，其中两个对儿童白血病及霍奇金（Hodgkin）疾病有很高的成功率。这种植物在非洲马达加斯加地区已广泛用于治疗低血糖症。像长春花属植物一样，很多广泛使用的天然药物（如阿司匹林、麻黄素、管箭毒碱、麦角新碱、地高辛、利血平、阿托品等）都是从土著的植物药中经过生物科学研究得到的。因此，人类数千年使用传统药用植物的经验积累是我们开发药用天然产物的宝贵财富。中药在我国已有数千年的历史，对中华民族的繁衍及发展做出了巨大贡献，并形成独特的完整的医学理论，是我国以及全世界人民非常宝贵的财富。

天然产物是植物及微生物的二次代谢物，很多二次代谢物作为进化选择的结果，具有很好的药效及特殊的选择性。它们可能是信号分子或对抗食草动物的食用或病菌的侵蚀，是植物的防护体系。如植物能产生杀菌剂防护病菌，亚马孙河蛙能产生极毒的天然产物对抗其他动物的捕食。这些二次代谢物有很好的药用前景。

所有的机体需要转换及互换大量的有机化合物使其有生命力，能生长复制。它们需要通过形成 ATP 提供能量以及构筑它们自己的组织。为此需要一个整合的酶媒介的小心调控化学反应，称为代谢途径生物网络。这些分子中特别重要的是糖、

脂肪、蛋白质及核酸。机体具有广泛的能力合成或转化化学物质。例如植物非常有效地通过光合成从无机材料合成有机化合物，动物从它们的食物中合成有机化合物，因此代谢途径涉及降解从环境或食物中获取的有机物，转化或合成专一的分子。尽管生命机体的特征有很大的差别，但其合成糖、蛋白质及核酸的生物途径是相似的。这个途径称为一次代谢途径，所涉及的化合物称为一次代谢物。例如糖的降解及糖的产生，通过糖酵解途径以及克雷布斯（Krebs）循环通过氧化反应释放能量。同一次代谢相反，二次代谢在自然界是有许多限制的，二次代谢物仅存在于非常特殊的机体，许多二次代谢物产生的途径及功能至今仍然是未知的。如果所有

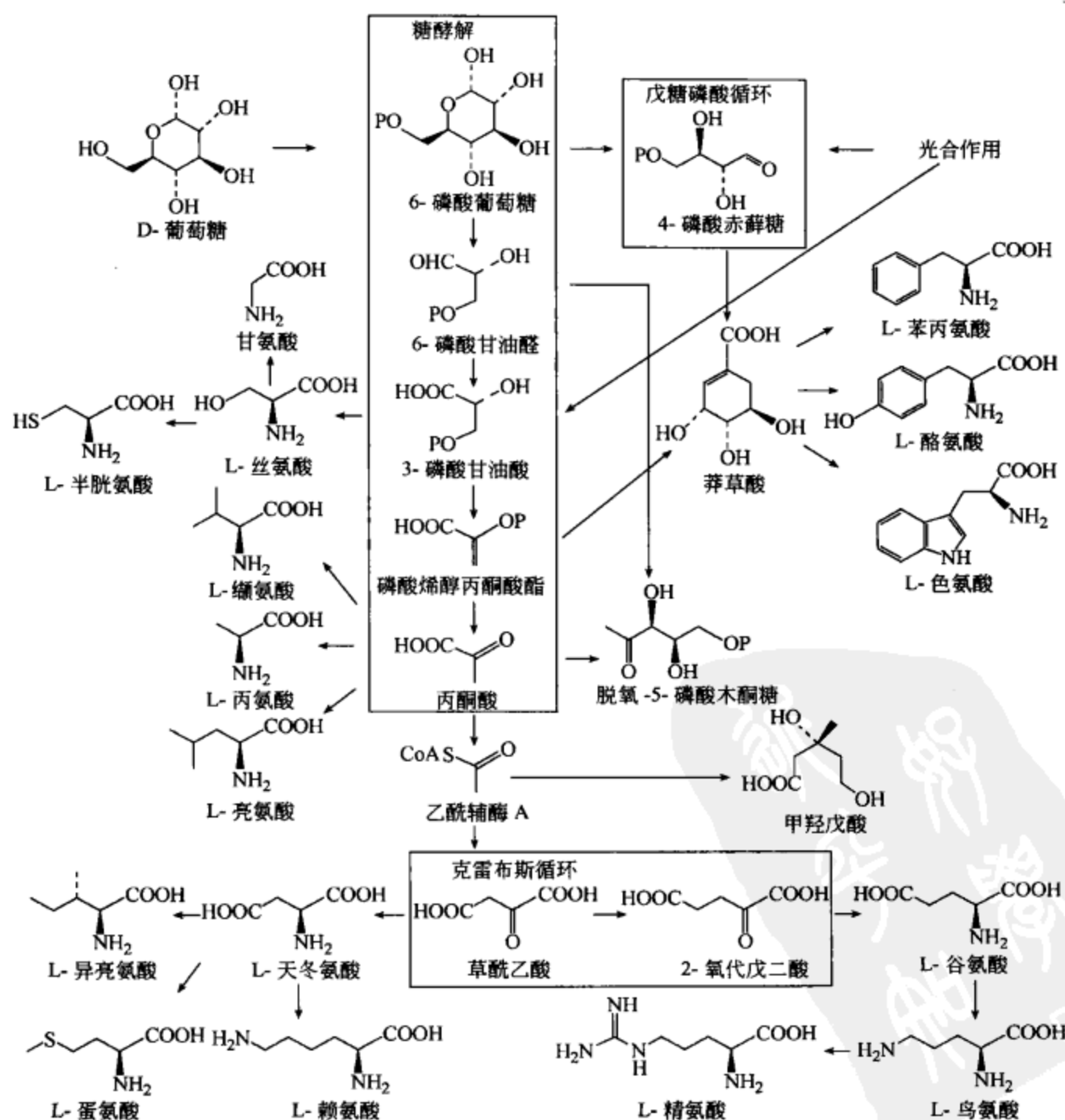


图 1-7 代谢途径

植物、动物及微生物产生相同的化合物的话，人的食物就会产生问题，二次代谢物的构成结构块 (building block) 来自一次代谢过程。如图 1-7 所示，该图概述了代谢物如何从能量产生过程中产生出光合成、糖酵解及克雷布斯循环等基本过程。提供生物合成的中间体，结构块的数目是非常少的，但像为数不多的积木块产生各种不同的积木建筑一样，产生各种不同的代谢物。这些最重要的二次代谢物结构块是来自中间体乙酰辅酶 A (acetyl-CoA)、莽草酸、甲戊二羟酸以及 1-脱氧木酮糖-5-磷酸，它们是通过乙酰、莽草酸酯、甲戊二羟酸酯以及脱氧木酮糖磷酸酯途径。如乙酰辅酶通过糖解途径的氧化脱羧作用产生焦磷酸，它也可通过脂肪酸的 β 氧化作用产生。酚类化合物、前列腺素、大环内酯抗生素等许多二次代谢物可以通过乙酰化途径产生。

一些结构块用于产生含碳及氮的天然产物如下。

C1: 含一个碳原子的最简单结构块，通常用于形成甲基基团，通常它附在 O 或 N 原子上，偶尔附到碳原子上。

C2: 两个碳原子单元可以由乙酰辅酶提供。它可以是简单的乙酰基团，又存在于酯及频繁出现在长的烷基链或芳香体系（如多酚化合物）。

C5: 支链 C5 异戊二烯单元是从甲戊二羟酸或脱氧木酮糖磷酸酯产生，可以由 3 个乙酰辅酶分子产生。

C6C3: 它涉及苯丙基单元，来自 L-苯丙氨酸或 L-酪氨酸。

C6C2N: 它来自 L-苯丙氨酸或 L-酪氨酸。

吲哚 C2N: L-色氨酸是第三个芳香氨基酸，含吲哚的体系经过脱羧作用的剩余物是吲哚 C2N 单元骨架。

C4N: C4N 单元通常是作为杂环吡咯烷体系，是由非蛋白质氨基酸 L-鸟氨酸形成的。

C5N: C5N 是通过相似于 C4N 产生途径形成的

以上 8 个结构块形成很多天然产物，因此可以看出，二次代谢物不仅具有很高药效及高选择性的优势，而且具有丰富的结构多样性。

1.2 天然产物的治疗领域

1.2.1 感染疾病

对于细菌疾病，天然产物是有效的“银弹” (silver bullet)。最早天然产物抗生素是 β -内酰胺抗生素青霉素 (penicillin) 等，另一个突破是首先用于治疗肺结核的抗生素链霉素 (streptomycin)，随后又发现了四环素 (tetracycline)、氯四环素 (aureomycin) 及大环内酯类化合物，其中最好的代表是红霉素 (erythromycin)。然而，在 20 世纪，随着组合化学及高通量技术的发展，一些制药公司放弃了天然产物库而致力于高通量筛选抗生素，但并没有获得有前景的抗生素，而仅获

得了为数有限的抗病毒候选化合物。由于许多病菌对现有抗生素的抗药性日益显著，发展新的抗生素已成为当务之急。据美国疾病控制与预防中心报道，分支杆菌肺结核种株对目前所有的治疗药物都有抗药性。最近 Merk 公司发现了很强的及广谱的革兰阳性抗菌活性化合物，新抗生素平板霉素 (platensimycin) 对万古球菌等没有抗药性。因此，从天然产物中筛选新的抗生素又引起人们更大的兴趣，并已取得进展。

目前仍需要发展新的抗真菌药物，天然产物不仅提供抗真菌药物的筛选资源，而且能够提供真菌生物学研究的探针。

天然产物抗病毒药物是海绵尿核苷 (spongouridine) 以及海绵胸苷 (spongothymidine)，在此基础上发展了抗艾滋病及其他抗病毒药物，可以认为它们是核苷类抗病毒药物的先驱。

1.2.2 抗肿瘤药物

癌症是人类最重要的杀手之一，抗癌植物药已有悠久的历史，早在公元 475~221 年，中国最早的医学巨著之一《黄帝内经》已提出肿瘤的概念。目前国际上已报道的具有抗癌活性的植物约有 3000 种。1945 年人们已发现足叶草毒素可用于生殖器癌症疾病，但毒性太大，经改造得到足叶乙苷以及表鬼臼毒噻吩糖苷 (teniposide)，可以防止 DNA 裂解。在 19 世纪 50 年代，人们发现了纯的抗癌天然产物长春碱、长春新碱、长春花碱等。在 1960 年美国 NIH 癌症研究所启动一个植物收集及筛选项目，发现了一批抗癌活性天然产物，如紫杉烷、喜树碱等。这些化合物至 20 世纪 90 年代才开发为药物。紫杉醇是第一个微管稳定剂，是防止它们聚合的药物，可以治疗肺癌、卵巢癌、乳癌等。2000 年销售额达 16 亿美元。

目前治疗多种癌症的药物大部分来自天然产物及天然产物的类似物，它们干扰细胞的分裂及增殖，包括烷基化试剂、生物碱、微管蛋白聚合剂及拓扑异构酶抑制剂等，例如紫杉醇、长春碱、蒽环霉素、逆诺霉素以及阿霉素等。

1.2.3 心血管及内分泌疾病

在我国利用植物及动物药治疗心血管及内分泌疾病已有悠久的历史，目前仍在广泛使用，如复方丹参滴丸，每年销售额达 10 余亿人民币，也已有一些纯的天然产物药物如芦丁、磷酸川芎嗪等。但相比于抗癌药物，来自天然产物的心血管及分泌病的纯天然产物药物相对较少。近年来，雷帕霉素用于再灌注局部缺血治疗，他汀类降血脂已受到广泛使用，如辉瑞公司的立普妥是销售额很大的药物之一。

1.3 天然产物研究方法

1.3.1 传统方法

近 50 年来，光谱方法结合色谱技术等天然产物的分离、纯化研究中已取得

显著的成效。采用各种合理的分离、纯化策略已经在生物活性化合物的分离及鉴定中取得成功，其中最重要的是生物活性指导下的分离纯化方法，是目前广泛采用的成果显著的方法。这些技术及方法不仅适用于植物活性成分的分离纯化，而且也已广泛应用于微生物及菌类代谢物的分离纯化。

冷萃取、热浸析、超临界萃取、索氏 (Soxhlet) 萃取等方法是目前常用的萃取方法。分离方法中重要的是色谱技术，薄层色谱方法是其中最快速、最广泛使用的方法之一，高效液相色谱也是目前广泛使用的有效方法之一。凝胶色谱、毛细管电泳等也是目前日益广泛使用的方法。

1.3.1.1 有效成分的概念

民族药特别是中草药的来源有动物、植物和矿物之分，其中以植物类为主。植物的化学成分很复杂，普遍含有蛋白质、糖类、淀粉、纤维素、树脂、叶绿素及无机盐等。有的植物还含有生物碱、苷类（如皂苷、强心苷、黄酮苷等）、挥发油、有机酸和氨基酸等一些其他化学成分。但是，一般来说，这些化学成分往往具有一定的生理活性，成为中草药的有效成分。目前，有些植物中的鞣质、多糖或蛋白质也可作为有效成分而逐渐被人们所重视。

一种中草药往往有多种临床用途，因此其有效成分可以有一个，也可以有好多个。例如中药阿片，其中的吗啡生物碱 (morphine) 具有显著的镇痛作用，罂粟碱 (papaverine) 对平滑肌有强力的抑制作用，可用来解除胃痉挛和支气管痉挛，而另一种生物碱可待因 (codeine) 具有显著的止咳作用。它们是阿片中 3 个具有不同临床用途的有效成分，分别部分代表阿片的临床用途。又如中药麻黄通常用作发汗、平喘、解热药，其中所含成分之一麻黄素 (ephedrine) 仅能部分代表麻黄的作用，用以治疗支气管性哮喘。

中草药一般具有广泛的临床用途，因此，在寻找它的有效成分时，首先应该确定寻找的目标，分别寻找其中某种疗效的有效成分，然后通过化学提取、分离纯化得到目标成分。有时代表一个中草药的某一特定疗效的有效成分往往有好几个，这是由于植物在形成有效成分时往往有一个生化演变的过程，因此很可能同时存在着结构相近的一系列化学成分，表现出相类似的药理和临床效用。例如中药常山中的抗疟有效成分就有常山甲碱、常山乙碱和常山丙碱 (α -dichroine、 β -dichroine、 γ -dichroine)，它们是 3 个异构体。又如洋金花和阿片中均含有好几个结构相近的有效生物碱。

这里所述的“有效成分”指的是化学上的单体化合物，即能用分子式和结构式表示并具有一定的物理常数。如果经纯度检查得到的是一个混合物，即使它在药理和临床上有效，能够代表或部分代表原来中草药的疗效，也只能称它为“有效部位”。只有在“有效部位”的基础上通过各种方法进一步分离纯化才能得到所需要的“有效成分”，即单体化合物。

1.3.1.2 常用提取方法

植物体内的成分较为复杂，对于如何着手进行中草药有效成分的提取分离，一般有以下两种情况：第一，对有效成分的性质一无所知；第二，从植物中提取已知的有效成分或已知的化学结构类型者，如从黄柏中提取生物碱黄连素等。两种情况不同，考虑其提取分离方法也不相同。

对于第一种情况，一般应在药理或临床实验配合下，经不同溶剂提取，以确定有效部位，然后再逐步划分，追踪有效成分最集中的部位，最后分得有效单体。通常采取不同溶剂，由低极性到高极性分步提取。将药材按以下次序提取：①石油醚或汽油或苯（所提取的物质是脂溶性大的化合物，如油脂、叶绿素、精油、甾体、三萜等中性物质等）；②乙醚（树脂及一些极性基团少的化合物，如甾体、某些生物碱、有机酸、黄酮体及香豆素苷元等）；③氯仿或乙酸乙酯（生物碱及许多中性成分）；④丙酮、乙醇或甲醇（提取极性化合物，如生物碱的盐、苷类、鞣质等）；⑤水（提取水溶性化合物，如氨基酸、糖类、无机盐等），进一步还可分为冷水、热水、酸水、碱水等步骤。这样，除纤维素等不溶物以外，植物中各类成分都能被提取出来。在确定哪一部位有效以后，再进一步分离。

对于第二种情况，一般先查阅有关资料，特别是工业生产的方法，搜集该成分的各种提取方法，再根据具体条件进行选择。所谓提取，就是用适当溶剂将植物中的化学成分从植物组织中抽提出来。提取可在室温下进行（即冷提），也可以加热提取（即热提）。一般来说，冷提杂质较少，而热提效率较高。有些有效成分冷时难溶热时易溶，则必须用热提。在不了解有效成分性质之前，一般采用温和的条件，不用酸碱，以免破坏。提取液浓缩一般在 60℃ 以下进行。对溶剂沸点超过 70℃ 的，在可能条件下应采用减压浓缩。这里要注意有时提取液浓缩至小体积，放置后，有固体或结晶析出，就可以将析出物滤出，供给进一步分离。滤液再浓缩，放置，观察有无固体再析出，直到母液蒸干为止。

现将常用提取方法概述如下。

（1）用水提取 水提取可分为水煎、水浸和水渗漉 3 种，也可以用酸水或碱水提取。以下物质用水提效果会比较好：①碱性化合物如生物碱，它们在植物中往往是与酸性成分成盐的形式存在，盐一般比较容易溶于水；②酸性化合物，羧基本身是一个极性基团，而且有时与金属成盐形式而存在，故溶于水的可能性很大；③中性物质如苷类，它们本身含有糖的部分，含有较多的羟基，是一些极性基团，在水中的溶解度增大，故常溶于大量的水中。

水提取液有时未及时处理，其中糖分、蛋白质等营养物质易于发霉，要注意防腐问题。可加入少量甲苯、甲醛或氯仿等作防腐剂。用水渗漉时，有时室温较高，在渗漉过程中生药即发酵，这时可用氯仿饱和水溶液进行渗漉。

（2）有机溶剂提取 凡脂溶性基团占优势的化合物，如甾体、萜类、蒽醌等，

用有机溶剂提取。有些化合物虽能溶于水，但为了使杂质尽量少带出来，也常用有机溶剂提取。用有机溶剂提取，可如前述将生药粉末用不同溶剂分步提取。药量少时可在索氏提取器中进行，量大时可在圆底烧瓶中进行。第一种溶剂提取完全后，取出药渣，待溶剂全部挥发完后，换用第二种溶剂提取。有时植物成分比较简单或有效成分含量较高，常可根据其极性大小或溶解性能，选择一种适当极性的溶剂把有效成分提取出来，而杂质留在植物里。举例如下。

① 用石油醚直接提取细辛素。细辛素为一种脂溶性化合物，以前用苯或酒精提取，由于含杂质多，因而手续较烦，得量也很低。后来改用石油醚回流，浓缩回流液，即析出细辛素结晶，方法简单，得量高。被石油醚一起提出的挥发油则留在母液中。

② 用苯提取苦楝素。苦楝素为苦楝皮中驱蛔有效成分。起初用酒精提取，但杂质较多。后改用苯提取，杂质较少。将药粉用苯回流，苯提取液浓缩后加入少量石油醚，即可获得粗制苦楝素。

③ 用丙酮提取鸦胆子苷和救必应苷。鸦胆子（种子）先用苯提取除去油脂，然后用丙酮回流，丙酮回流液浓缩即得白色粉末沉淀，为鸦胆子混合苷。救必应（树皮）先用苯提取除去树脂等杂质，再用丙酮回流，丙酮液浓缩即析出救必应总苷的白色沉淀。

④ 用酒精提取远志皂苷。远志先用石油醚脱脂，再用酒精提取，提取液浓缩后放入冰箱，即析出大量黄白色皂苷。

提取水溶性化合物时，有时为了避免用水提取造成的困难（如回收需高温，且容易产生大量泡沫等），而改用 70% 酒精提取。

（3）酸性、碱性有机溶剂提取 如果有效成分是酸性或碱性化合物，常可加入适当的碱和酸，再用有机溶剂提取。例如生物碱在植物体内一般与酸结合成盐，可在生药中加入适量的碱液，拌匀，使生物碱游离出来，再用有机溶剂（如苯、氯仿等）提取。同样，有机酸可加碱使其游离，然后用有机溶剂提取。反之，可以用酸性酒精提取弱碱性生物碱。

（4）水蒸气蒸馏 挥发油和某些挥发性成分能用水蒸气蒸馏法得到。如麻黄碱就可以用水蒸气蒸馏法从麻黄中直接蒸馏出来。这个方法主要用于挥发油的提取。

（5）酶解或抑制酶解 在提取苷类时，要注意生药中酶的分解作用。可先将植物、粉末加醇煮沸片刻，以抑制其中酶的分解作用，在提取蒽醌苷和氰苷时，尤其需注意这一点。相反地，有时又利用酶解作用，如在提取皂苷、强心苷时，有的苷分子中糖较多，极性大，难以分离，在提取分离前先将生药酶解，使苷分子中的糖部分水解，以获得生理作用相同的次生苷。

（6）化学处理 内酯化合物不溶于水，其内酯环遇碱水解成为羧酸盐而溶于水，再加酸酸化，可重新形成内酯环，回复原物而不溶于水，从而与其他杂质分

开。如从蛔蒿中提取驱蛔有效成分山道年就利用此性质。将生药粉末用石灰乳调匀，加热水提取，山道年成为山道年酸钙，被水提出，水提取液过滤浓缩后加酸酸化，则山道年沉淀析出，滤集，用醇重结晶可得纯品。

(7) 升华法 植物中有的成分在加热时直接变成气态，遇冷凝结成结晶。如茶叶加热时，咖啡因即升华出来。但一般可升华的植物成分很少。

1.3.1.3 常用的分离纯化方法

一般中草药按上述方法得到总提取物以后，要进行进一步的分离和纯化。常用方法如下。

(1) 溶剂处理 如果总提取物是水浸膏，可将水浸膏依次用石油醚（或苯）、乙醚、氯仿、乙酸乙酯、丙酮、酒精和水提取，或选择其中三四种不同溶剂提取。水浸膏常常为胶状物，可拌以适量惰性填充剂（如硅藻土或纤维粉等），然后低温干燥，使其成粉末状，这样提取比较完全，提取方法与前述生药用溶剂分步提取的方法相同。如果是水浓缩液，可以在分液漏斗中依次选择几种与水不相混溶的有机溶剂，如石油醚、苯、乙醚、氯仿、乙酸乙酯、正丁醇等抽提，分成若干部位。

(2) 酸碱处理 各种总提取物用酸水（一般用1%~2%盐酸）及碱水（一般用0.1mol/L氢氧化钠）处理，可以分为3个部分：溶于酸水的为碱性成分，溶于碱水的为酸性成分，酸碱不溶的为中性成分。如用酒精提取生物碱时，用酸水处理酒精浸膏，这时生物碱进入酸水。也可利用不同酸碱度进一步分离。如酸性化合物可分为强酸性、弱酸性和酚性3种，它们分别溶于碳酸氢钠、碳酸钠和氢氧化钠，借此可达到分离。有时生物碱也可利用不同pH值来分离。

(3) 制备衍生物 有时要分离原来的化合物有困难，可制成衍生物，但要注意制成的衍生物应能回复成原来的形态。如生物碱常常通过成盐而达到分离。制备生物碱的盐时，无机酸、有机酸都可以用。如果仲胺碱和叔胺碱在一起，可将仲胺碱制成亚硝基或乙酰基衍生物，借以与叔胺碱分开。此外，还有利用酯化来分离羟基化合物（如黄酮体）。用Girard试剂分离酮基化合物，使成水溶性腙，而其他有机物可用有机溶剂除去，水液酸化后又成原来的酮，再从氯仿等有机溶剂提出。

(4) 吸附法 这里分两种情况：一种是吸附杂质；另一种是吸附有效成分。所用吸附剂有氧化铝、硅胶、氧化镁、酸性白土、活性炭等。如在分离纯化皂苷时所采用的氧化镁法，就是利用氧化镁吸附杂质的。

(5) 沉淀法 最常用的是铅盐法，可以用于除去杂质，也可用于沉淀有效成分。铅盐法的原理是利用在水或稀醇溶液中加入中性醋酸铅或碱性醋酸铅，可与许多物质生成难溶的铅盐或络盐，用于分离植物成分。中性醋酸铅可以与酸性或酚性的物质结合成不溶性铅盐，因此可以沉淀有机酸、蛋白质、氨基酸、黏液质、鞣质、酸性皂苷、树脂、部分黄酮苷和花青苷等。碱式醋酸铅沉淀范围更广，除了上述能被中性醋酸铅沉淀的物质外，还可沉淀某些苷类、糖类及一些生物碱等碱性物

质。脱铅的方法可以有：①通硫化氢气体，使铅变成黑色硫化铅沉淀出来，至滤液再通硫化氢不产生沉淀为止，过滤，沉淀用水洗，洗滤液合并，置蒸发皿内，水浴上加热驱去过量的硫化氢；②加硫酸钠或磷酸钠饱和溶液，这时铅成为硫酸铅或磷酸铅沉淀出来，至沉淀完全，过滤即可；③加磷酸或稀硫酸，调节到 $\text{pH}=3$ ，滤去沉淀即可；④加强酸性阳离子交换树脂氢型，搅拌，即可脱去铅离子。

(6) 分馏法 如在分离毒芹总碱中的毒芹碱和羟基毒芹碱时，前者的沸点为 $166\sim 167^{\circ}\text{C}$ ，后者为 226°C ，彼此相差较远，可利用沸点不同进行分馏，大致先行分开，然后再精制纯化。

(7) 盐析法 水溶液加无机盐（常用氯化钠、硫酸钠、硫酸铵等）饱和时，可使有些物质溶解度降低，甚至成为沉淀析出，这种性质可用于分离纯化。如从滇三七中分离三七皂苷，三七粉先用戊醇提取，得到三七皂苷甲，药渣再用 98% 酒精提取，酒精提取液蒸干后加水溶解，滤去不溶物，加硫酸镁饱和，粗制三七皂苷乙即析出。

(8) 透析法 主要用于大分子的纯化。

1.3.1.4 层析分离方法

层析分离方法是近几十年来才发展起来的技术，正不断地趋于完善，目前已经成为中草药有效成分分离工作的重要手段之一。层析法的特点是能满意地分离一系列结构较为相似的成分，以达到一般分离方法难于达到的目的。它目前主要作为实验室方法，由于所用药品仪器要求较高，技术操作也较细致，操作周期较长，因此目前在工业生产上还很少应用。在中草药有效成分分离提取中，往往将经典分离方法和多种层析方法结合使用，以互补长短，通常先用一般分离方法分得有效部位，再用层析法分离有效成分。

层析法是利用下述现象：当各种成分在一支持剂上，随着另一溶剂通过时，由于各物质前进速度不同，而达到分离的目的。如按分离原理上大体可分为：①吸附层析（利用吸附能力的差别进行分离），常用吸附剂有氧化铝、硅胶、聚酰胺、活性炭等；②分配层析（利用在两种不互溶的溶剂中分配比不同进行分离），常用支持剂为硅胶、硅藻土、纤维粉等；③离子交换层析（利用离子解离强度不同进行分离），常用离子交换树脂有强酸性（磺酸型）、强碱性（季铵型）、弱酸性（羧酸型）、弱碱性（三级胺型）；④电泳（利用电流通过时离子趋电性不同），常用纸上电泳与琼脂电泳。此外，还有一些诸如气体层析（气-液分配层析）、分子筛层析等。按支持剂的形式可分为柱层析（支持剂放在玻璃柱内）、纸层析（以滤纸作为支持剂）、薄层层析（支持剂平铺在玻璃板上）。柱层析主要用于分离制备，薄层层析和纸层析主要用于分离鉴定。

层析用硅胶一般以 $\text{SiO}_2 \cdot x\text{H}_2\text{O}$ 表示，系多孔性的硅氧环 Siloxane 交链结构，由于其骨架表面具有很多硅醇基团，能吸着多量水分，此种水分几乎成游离状态存

在，因此当加热到 100℃左右能可逆地被除去。硅胶的活性与水分含量有关，含水量高则吸附力减弱，当游离水含量高达 17%以上，吸附能力极低，而可作为分配层析载体。

硅胶吸附层析目前使用比较广泛，既能用于非极性化合物，也能用于极性化合物，适用于芳香油、萜类、甾体、生物碱、强心苷、蒽醌类、酚类、酸性化合物、磷脂类、脂肪酸、氨基酸等。

一般由于硅胶为多孔性物质，装柱宜用湿法，即将硅胶混溶于装柱溶剂中，不断搅拌，待空气气泡除去后，连同溶剂一起倾入层析柱中，最好一次倾入，否则由于不同粒度大小的硅胶沉降速度不同，使硅胶柱有明显的分段，容易影响分段效果。对于硅胶吸附层析，样品与吸附剂的比例可为 (1:30)~(1:60)，但也并不是绝对的，根据分离工作的难度，对于较难分离者有时甚至可以用 (1:200)~(1:500)。一般需分离样品量较大时，需要装大柱，经验做法常采用干法装柱，也要求尽量一次全部倾入。另外，经常遇到样品难溶于所选用的分离溶剂条件，这时可以有两种做法：①如果样品可溶于比选择分离的溶剂极性略小的溶剂中，则用该溶剂条件（通常会选用混合溶剂）装柱，样品也用这种溶剂条件溶解，湿法装柱；②或者仍然采用湿法装柱，只是样品先用别的溶剂溶解，在挥干后用干法上样。

而凝胶柱只能选择用湿法装柱和湿法上样。凝胶的分离原理是根据分子量的大小不同来达到分离的目的，就像分子筛一样。分子量小的物质先出来，分子量大的物质后被洗出。要求用能溶解样品的溶剂洗脱即可。

相对来说，凝胶柱在操作上比硅胶柱容易得多，但考虑到价格的问题，一般是先用硅胶柱进行前期的分离，只有等到少量物质的精分时才考虑选择凝胶（一般只有 2~3 个成分时）。

1.3.2 活性化合物的筛选新方法

传统活性筛选方法及高通量筛选方法都可用于天然产物的活性筛选，但由于天然产物的分离纯化费时长、效率低，因此需要发展一些新的直接从植物粗提物中筛选活性化合物的方法。目前已报道的从植物粗提物中直接筛选活性化合物的方法有基于固定化酶多克隆抗体的前沿色谱方法、分子烙印高分子方法及微分智能筛选等。

1.3.2.1 固定化酶的筛选方法

该方法是将药物作用的靶酶固定化后制成亲和色谱柱作为固定相，选择合适的流动相并与 HPLC-MS 连接，利用前沿色谱的原理，实现天然产物复杂体系活性化合物的有效筛选；同时利用质谱在线检测到活性分子的分子量，基本上可以实现筛选与鉴定一体化。

1.3.2.2 多克隆抗体筛选方法

当药物的靶蛋白不清楚或者靶蛋白很难得到时，可以采用多克隆抗体作为模拟

物代替靶蛋白。它是将有抑制活性的化合物作为半抗原，与牛血清蛋白偶联后免疫兔子，从兔血清中得到带有这种半抗原特性的多克隆抗体，将它固定化后作为固定相，选择合适的流动相，结合前沿色谱实现中药活性成分的高通量筛选。

1.3.2.3 分子烙印聚合物 (MIP) 筛选方法

与前两种方法相比，MIP 筛选法方法更加稳定。它是将已知活性分子作为模板分子，先以非共价方式与高分子功能单体聚合，然后洗脱掉聚合体中的模板分子，MIP 即制作完成。将它填充色谱柱做成固定相，选择合适的流动相，结合前沿色谱实现中药活性成分的高通量筛选。MIP 的分子识别是来源于这聚合物中保留了模板分子的印迹，这个印迹包含了模板分子的空间结构、官能团（如氢键、静电和范德华引力等）。因此，MIP 的筛选过程同样是一个分子识别的过程，识别那些与模板分子在空间结构及官能团相类似的化合物。

1.3.2.4 计算药理学方法

计算机辅助药物设计已经有几十年的历史，已发展成为一套较为完整的算法技术及软件，包括各种数据库、定量构效关系、药效团模型、分子对接、ADME/T（吸收、分布、代谢、消除/毒性）预测、各种数据挖掘分析算法等。在这些方法与技术的基础上近年来发展了计算药理学（computational pharmacology）。

计算药理学的目标是在计算机上实现某些药理学实验，这样可以大大减少实验工作量，节省大量的人力、物力。随着系统生物学及生物网络的发展，人们已认识到以往的单一靶点单一药物的模式对于复杂疾病收效甚少，人们已逐渐注意到多靶点、多途径基于系统的药物开发模式。药用天然产物的特点是多组分、多靶点，因此，这些新进展为人们提供了开展基于天然产物的计算药理学研究的方法与途径。

基于天然产物的计算药理学研究包括传统的计算机辅助药物分子设计方法，各种来源的生物、化学、药物及医学数据的整合与挖掘，天然产物分子在化学空间及配体靶空间的分布，活性分子与多种蛋白质的相互作用谱，植物中多种天然产物与多种靶的相互作用谱，应用生物网络表达及分析天然产物的结构和与多靶的相互作用等。

1.3.2.5 微分智能筛选

微分智能筛选（differential smart screens）的基本原理是基于分子药理学，是一种适合于植物提取物的高通量药效筛选新方法。一种化合物具有对一组相似分子靶的作用谱，即对多种靶的结合亲和力谱。这种结合亲和力谱可作为该化合物的具有专一性的表征。如果植物提取物的对一组分子靶的结合亲和力谱与已知分子不同，则这种植物提取物值得进一步研究，有可能从中分离到新的活性化合物。

1.4 药用天然产物的反向途径

尽管全球范围药物研究与开发的投入逐年显著增加，但在过去 20~25 年被 FDA 批准的新药几乎维持在每年 20 个左右，按这个速度需要 300 多年才能使现有的药物数目增加一倍。造成药物种类增长缓慢的原因很多，重要因素之一是缺乏优质的先导化合物，即可谓“靶富有，先导化合物贫乏”。天然产物由于具有丰富的结构及活性多样性，是获得优质先导化合物的重要资源。

全球的高等植物有 250000~500000 种，仅 6% 经过活性筛选，15% 经过植化研究。据估计，全球药用植物 (2~4) 万种，而仅不到 10% 的药用植物经过详细的植物化学或药理研究。对于药用植物研究的传统模式 (从植物到分子) 是选用一种植物，通过系统分离、纯化得到单一化合物，然后对这些化合物进行活性鉴定或采用活性指导下的分离、纯化。传统的方法的缺点是费时、费力、周期长。因此，不可能在短时间内利用传统的途径对药用植物进行全面的研究。

北京大学化学学院徐筱杰课题组经过多年的努力探索出一条反向的途径 (从分子到植物)，即先发现活性分子，然后从植物中分离出该分子。反向途径包括利用计算机模拟虚拟筛选技术、活性分子的分子烙印高分子、固定化酶及固定化活性分子的多克隆抗体技术。这些技术的有效组合，使人们能在较短时间内从中药体系中发现及提取活性分子。这些方法不仅是一种高通量模式，可以大大加快发现活性化合物的速度、大大降低成本，而且可以避免假阳性，不易漏筛。

人类数百年药用天然产物研究已积累了丰富的数据、信息和知识。据理论推测，可能的天然产物可达数百万种，迄今已知的天然产物约 20 万种，其中来自植物的约 10 余万种，中药植物中已知的化合物估计在 2 万种以上。这些丰富的知识与信息积累是研究药用天然产物的重要财富。目前急需发展一条基于现有药用天然产物知识的快速发现活性化合物的新途径。这首先需要建立内容丰富的信息化系统以及发展有效的虚拟筛选技术。

药用植物体系是天然的组合化学库，直接从药用植物及中药复杂体系中高通量地筛选活性化合物以及直接从复杂体系中提取活性成分的方法是克服目前药用天然产物研究瓶颈及充分利用天然组合化学库的新途径。近年来，高通量筛选技术的发展，特别是前沿色谱方法、分子烙印聚合物的迅速发展，为人们寻找实现上述新途径的具体方法提供了可能。

药用植物体系特别是中药体系与化学合成库比较具有更丰富的结构多样性及活性多样性，但药用植物体系成分复杂，许多成分未知，而且含量悬殊。快速从中发现及提取活性化合物是一个高难度挑战。

北京大学化学学院徐筱杰课题组根据分子识别原理 (几何形状匹配、能量匹配、

预组织、诱导匹配等概念) 以及化学基因组学的两个基本假设 (分享一定化学相似性的化合物也应该分享同样的靶标, 分享同样配体的靶具有相似结合位点), 结合信息整理, 挖掘及融合技术, 发展了一系列理论计算与实验筛选、分离提取相结合的技术, 包括中药信息系统、快速虚拟识别筛选、高通量实验识别技术及筛选分离鉴定一体化实验技术等。

总体思路 (图 1-8) 是既可以从已知靶酶出发直接做成固定化酶层析柱, 也可以从已知的抑制剂出发做成多克隆抗体柱或分子烙印结合柱。根据分子识别原理, 酶与抑制剂的结合类似锁钥匹配模型。前者相当于锁, 最终的目的都是寻找适合于开锁的钥匙。这些钥匙存在于药用植物当中, 通过虚拟筛选结合数据库查询可以得到相关植物的信息, 使用这些植物的粗提取物, 通过固定化酶和多克隆抗体柱结合前沿色谱, 可以高通量地筛选出活性化合物的活性顺序, 从相应的分子烙印高分子层析柱结合固相萃取分离得到化合物实物, 再通过生物活性检测得到先导化合物。

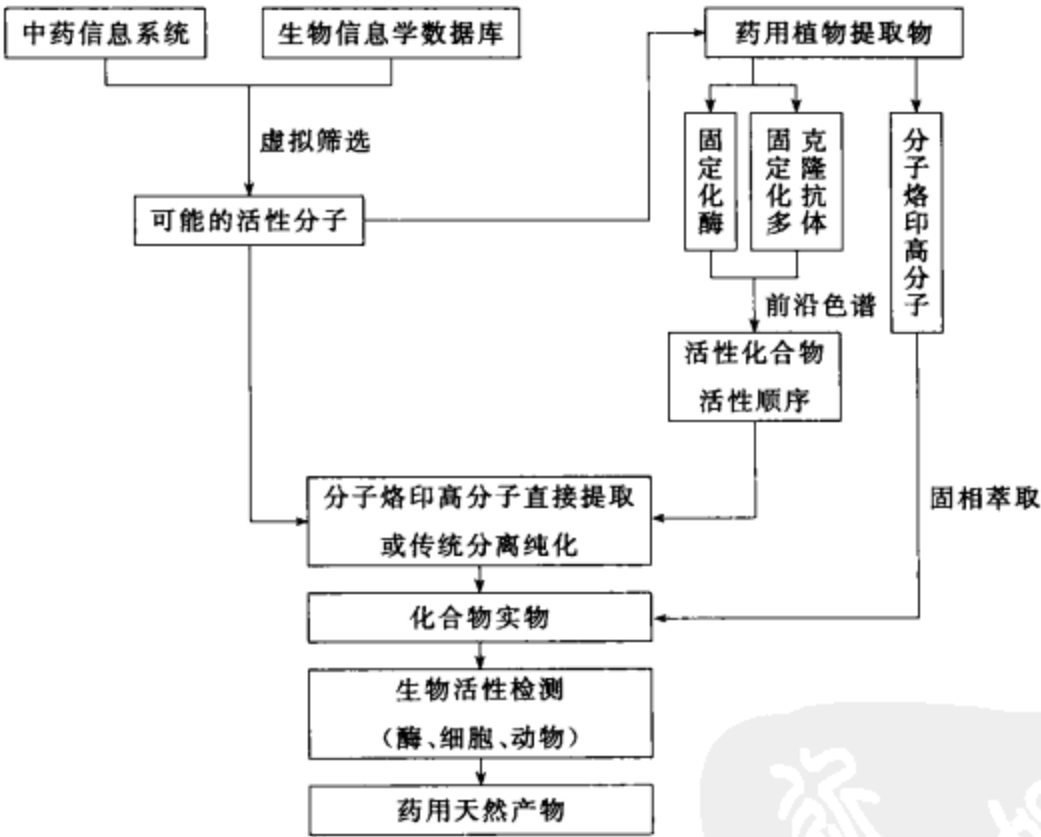


图 1-8 提取药用天然产物的思路

参 考 文 献

[1] Patrick G. Instant Notes in Medicinal Chemistry. BIOS Scientific Publishers, 2001. [中译本: 孙铁民、郭春译. 药物化学. 北京: 科学出版社, 2004.]

[2] Wermuth C G. (Ed.) The Practice of Medicinal Chemistry (2nd Edition). Academic Press, London, 2003. [中译本: 迟玉明等译. 创新药物化学. 广州: 广东世界图书出版公司, 2004.]

[3] Grabley S, Thiericke R. Bioactive agents from natural sources: trends in discovery and application.

Adv. Biochem. Eng. /Biotech, 1999, 64 (1): 104-154.

- [4] Cragg G M, Newman D J, Snader K M. Natural products in drug discovery and development. *J. Nat. Prod*, 1997, 60 (1): 52-60.
- [5] Newman D J, Cragg G M, Snader K M. Natural Products as Sources of New Drugs over the Period 1981-2002. *J. Nat. Prod*, 2003, 66 (7): 1022-1037.
- [6] Newman D J, Cragg GM. Natural Products as Sources of New Drugs over the Last 25 Years. *J. Nat. Prod*, 2007, 70 (3): 461-477.
- [7] Wilkinson B, Moss S J. Biosynthetic engineering of natural products for lead optimization and development. *Curr. Opin. in Discov. & Dev*, 2005, 8 (6): 748-756.
- [8] Pelaez F. The historical delivery of antibiotics from microbial natural products-Can history repeat? *Biochem. Pharm*, 2006, 71 (7): 981-990.
- [9] Xu X J. New concepts and approaches for drug discovery based on traditional Chinese medicine. *Drug Discov. Today: Tech*, 2006, 3 (3): 247-253.
- [10] Ameenah G F. Medicinal plants: Traditions of yesterday and drugs of tomorrow, *Molecular Aspects of Medicine*, 2006, 27: 1-93.
- [11] Ganesan A. The impact of natural products upon modern drug discovery, *Current Opinion in Chemical Biology*, 2008, 12: 306-317.
- [12] Knight V, Sanglier J J, DiTullio D, Braccili S, Bonner P, Waters J, Hughes D, Zhang L. Diversifying microbial natural products for drug discovery. *Appl. Microbio. & Biotech*, 2003, 62 (5-6): 446-458.
- [13] Blunt J W, Copp B R, Hu W P, Munro M H G, Northcote P T, Prinsep M R. Marine natural products. *Nat. Prot. Rep*, 2007, 24 (1): 31-86.
- [14] Morris J C, Nicholas G M, Phillips A J. Marine natural products: synthetic aspects. *Nat. Prot. Rep*, 2007, 24 (1): 87-108.
- [15] Klaus-Peter S. Extract from Ginkgo biloba leaves, its method of preparation and pharmaceuticals containing the extract. [P]. DE: 3940092, 1991-06-06.
- [16] Keith C T, Borisy A A, Stockwell B R. Multicomponent therapeutics for networked systems. *Nature Drug Discov*, 2005, 4 (1): 1-8.
- [17] 徐筱杰, 侯廷军, 乔学斌, 章威. 计算机辅助药物设计. 北京: 化学工业出版社, 2004.
- [18] Ulrich-Merzenich G, Zeitler H, Jobst D, Vetter H, Wagner H. Application of the "-Omic-" technologies in phytomedicine. *Phytomed*, 2007, 14 (1): 70-82.
- [19] Morphy R, Rankovic Z. Fragments, network biology and designing multiple ligands. *Drug Discov. Today*, 2007, 12 (3/4): 156-160.
- [20] Littleton J, Rogers T, Falcone D. Novel approaches to plant drug discovery based on high throughput pharmacological screening and genetic manipulation. *Life Science*, 2005, 78: 467-475.
- [21] Zhu, L L, Chen L R, Xu X J. Application of a molecularly imprinted polymer for the effective recognition of different anti-epidermal growth factor receptor inhibitors. *ANALYTICAL CHEMISTRY*, 2003, 75 (23): 6381-6387.
- [22] Zhu L L, Chen L R, Luo H P, Xu X J. Frontal affinity chromatography combined on-line with mass spectrometry: A tool for the binding study of different epidermal growth factor receptor inhibitors. *ANALYTICAL CHEMISTRY*, 2003, 75 (23): 6388-6393.

- [23] Luo H P, Chen L R, Li Z Q, Ding Z S, Xu X J. Frontal immunoaffinity chromatography with mass spectrometric detection: A method for finding active compounds from traditional Chinese herbs. *ANALYTICAL CHEMISTRY*, 2003, 75 (16): 3994-3998.
- [24] Yi L, Li Z Q, Yuan K H, Qu X X, Chen J, Wang G W, Zhang H, Luo H P, Zhu L L, Jiang P F, Chen L R, Shen Y, Luo M, Zuo G Y, Hu J H, Duan D L, Nie Y C, Shi X L, Wang W, Han Y, Li T S, Liu Y Q, Ding M X, Deng H K, Xu X J. Small molecules blocking the entry of severe acute respiratory syndrome coronavirus into host cells. *JOURNAL OF VIROLOGY*, 2004, 78 (20): 11334-11339.
- [25] Xu X J, Zhu L L, Chen L R. Separation and screening of compounds of biological origin using molecularly imprinted polymers. *JOURNAL OF CHROMATOGRAPHY B-ANALYTICAL TECHNOLOGIES IN THE BIOMEDICAL AND LIFE SCIENCES*, 2004, 804 (1): 61-69.
- [26] Mishra R C. A review of high throughput technology for screening of natural products. *Biomedicine and Pharmacotherapy*, 2008, 62: 94-98.
- [27] Beghyn T. Natural Compounds: Leads or Ideas? Bioinspired Molecules for Drug Discovery. *Chem Bio Drug Res*, 2008, 72: 3-15.



第2章 天然产物的生物合成

2.1 概述

天然产物的生物合成主要是探讨生物体中二次代谢产物的生源途径，研究从前体经各个中间体直至形成最终产物的历程以及有关的反应机制。该过程的研究涉及了植物学、植物化学、生物化学和生物工程等领域，包括研究各类成分在体内生物合成的途径、各种酶所起的作用以及所产生的各种中间产物的化学结构。这些天然产物在生物体内都有极其重要的适应性作用，例如协助植物抗食草动物、抵御微生物的侵染、吸引授粉昆虫和种子传播者以及作为种间抑制剂。

生源研究有多种设想与途径，因而也形成了多种学说，其中异戊二烯法则、醋酸学说等已普遍应用于研究药用植物有效成分的生物合成及其途径。随着同位素示踪技术和化学技术的发展，生源研究的进展也更为迅速，并在以下领域发挥作用。

(1) 药用活性产物的积累 了解了各类成分的生物合成途径，获知了某种成分的前体 (precursors)、合成诱导剂和各种中间产物后，就可以利用悬浮细胞培养或者愈伤组织有目的地进行药用活性产物的积累。如紫杉醇的细胞培养过程中，加入诱导剂甲基茉莉酮酸可以提高紫杉醇的产量，最高可以达到约占干重的 0.5%；异胡豆苷是长春碱和喜树碱生物合成的中间产物，在愈伤组织与细胞悬浮培养过程中，异胡豆苷的积累有利于长春碱和喜树碱的生成。

(2) 扩大生物活性物质的资源 如三萜类与许多甾体衍生物类在生源上具有密切关系，甾体衍生物类常具多种生物活性，三萜类成分在植物界分布广泛，故有可能通过三萜类成分来寻找具广泛生物活性的甾体类物质。

(3) 确定某类成分的结构类别 如四环三萜类成分原分类不属于三萜，通过生源关系的研究，才明确地将它们划在三萜范围内。

(4) 指导生产实践 了解某类成分在植物体内的原始状态与代谢途径后，就可以为进行植物成分的生物合成提供理论规律，这将能更好地对生产与实践（如生药的采收时间与部位、有效成分的合成等）起指导作用。

(5) 探讨化学成分特征与植物亲缘关系的相关性 依据天然产物类型及其合成途径，结合经典分类学及相关学科，共同研究植物类群特征、植物类群和植物系统

发育，探讨化学成分特征与植物亲缘系统的相关性。

2.2 天然产物的生物合成途径

2.2.1 一次代谢与二次代谢

二次代谢过程来源于一次代谢过程（图 2-1）。图 2-1 描绘了光合作用、糖酵解和三羧酸循环等一次代谢过程的产物，以及生物合成为二次代谢产物的途径。一次代谢中产生的基本结构单元数量虽然很少，但通过这些基本结构单元的累加可以形成各种类型的天然产物。目前已知在二次代谢产物的生物合成中出现最多的基本结构是乙酰辅酶 A、莽草酸、甲羟戊酸和 1-去氧木酮糖-5-磷酸，它们分别出现在乙酸酯、莽草酸酯、甲羟戊酸酯和去氧木酮糖-5-磷酸酯的生物合成途径中，这些途

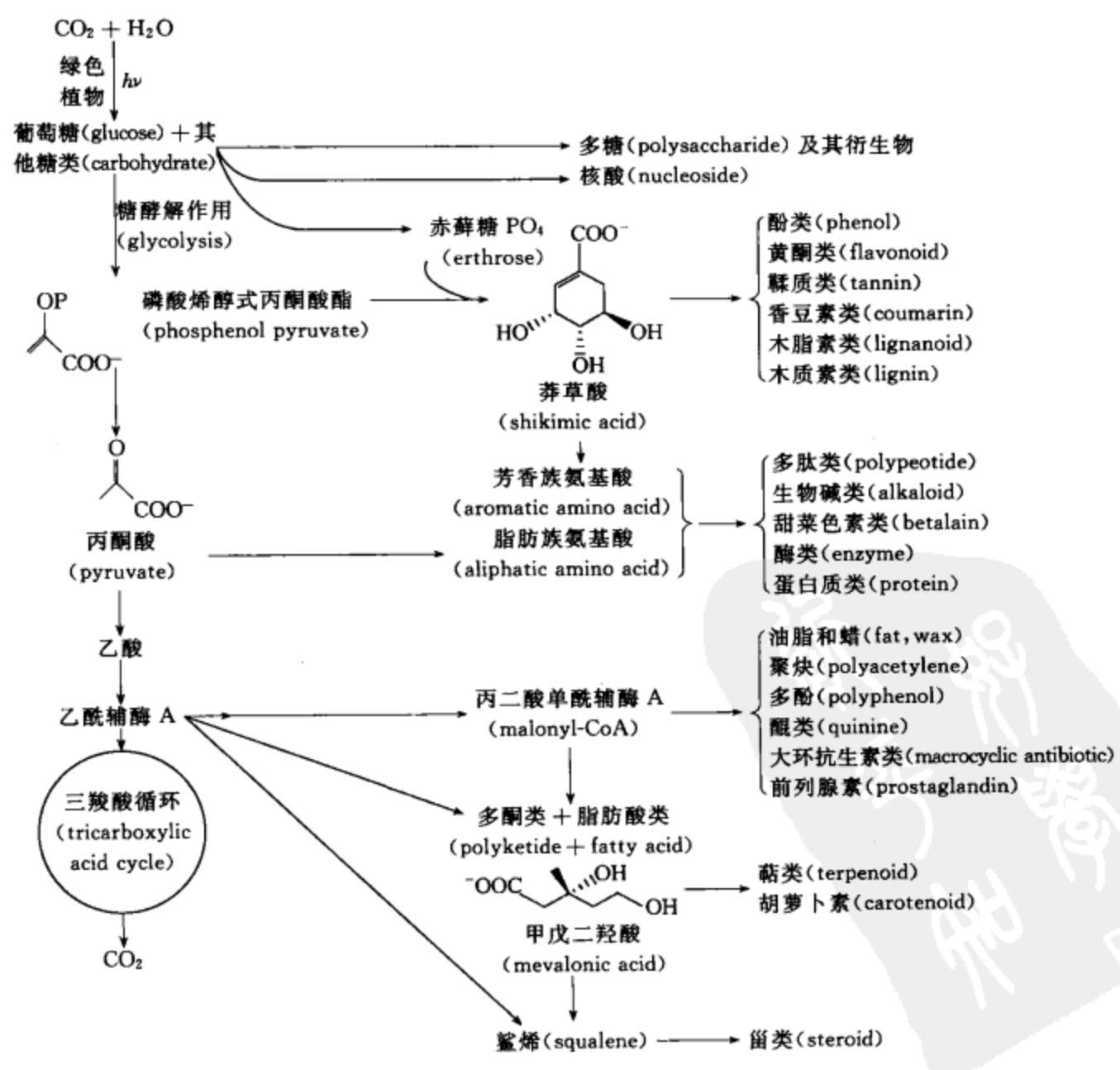


图 2-1 植物代谢产物生物合成途径

径是生物合成反应的基础。

糖酵解途径的产物丙酮酸经氧化脱羧可以得到乙酰辅酶 A，并可以通过脂肪酸的 β -氧化逆转乙酰辅酶 A 合成脂肪酸途径。在乙酸生物合成途径中产生的重要二次代谢产物有苯酚、前列腺素和大环内酯类的抗生素，还有各种脂肪酸和一次、二次代谢界面衍生物。莽草酸由磷酸烯醇式丙酮酸（糖酵解途径的中间体）和赤藓糖-4-磷酸（来自戊糖磷酸酯途径）结合而成。苯丙酸途径可以产生多种酚类、桂皮酸类衍生物、木脂素和生物碱。甲羟戊酸是由三分子乙酰辅酶 A 形成的，但甲羟戊酸生物合成途径与醋酸生物合成途径的产物不同。去氧木酮糖-5-磷酸酯是由糖酵解途径的两个中间体丙酮酸和 3-磷酸甘油醛结合而产生的。通过甲羟戊酸途径和去氧木酮糖磷酸酯途径可以合成多种萜类和甾类化合物。

氨基酸也可以作为重要原料和前体结构单元进行肽、蛋白质、生物碱和抗生素的生物合成。这些氨基酸有些是糖酵解途径和三羧酸循环产生的中间产物合成的，如鸟氨酸（一种非蛋白氨基酸）和它的同系物赖氨酸是生物碱的重要前体（来源于三羧酸循环的中间产物），但苯丙氨酸、酪氨酸和色氨酸等芳香氨基酸则是莽草酸生物合成途径的产物。

2.2.2 生物合成假说的提出

随着各种类型天然产物被不断分离和结构确定，对同类型的天然产物进行归属和分类，发现在大多数情况下结构相似的天然产物的生物合成具有同一起源。这一发现对寻找天然产物新的药用资源，以及在植物化学分类学、植物生态学和药用天然产物生物工程中起到了积极作用。例如，在比较茴香脑（anethole, 2-2）、丁香酚（eugenol, 2-3）等化合物结构时，发现它们具有相同的 C_6-C_3 骨架，推测它们具有相似的生物合成途径，即由一系列代谢产物酪氨酸（tyrosine, 2-1）及多巴（dopa, 2-4）代谢而来。又如，仔细剖析罂粟碱（papaverine, 2-5）、木兰碱（magnoline, 2-6）、全去甲劳丹诺苏林（2-9）、小檗碱（berberine, 2-10）等结构，发现它们的结构中也包含有多巴骨架，所以推测它们由多巴胺（2-7）经下列途径（图 2-2）与芳香醛（2-8）类化合物缩合而成。由于首次指出了生物碱与氨基酸之间的生源关系，对生物合成假说的提出具有重要的意义。

再如，早期认为只有符合异戊二烯法则的化合物才是萜类化合物（见图 2-3），如薄荷醇（menthol, $C_5 \times 2$, 2-11）、 α -山道年（ α -santonin, $C_5 \times 3$, 2-12）、海松酸（pimaric acid, $C_5 \times 4$, 2-13）、齐墩果酸（oleanolic acid, $C_5 \times 6$, 2-14）、 β -胡萝卜素（ β -carotene, $C_5 \times 8$, 2-15），分子中不断重复出现 C_5 骨架。

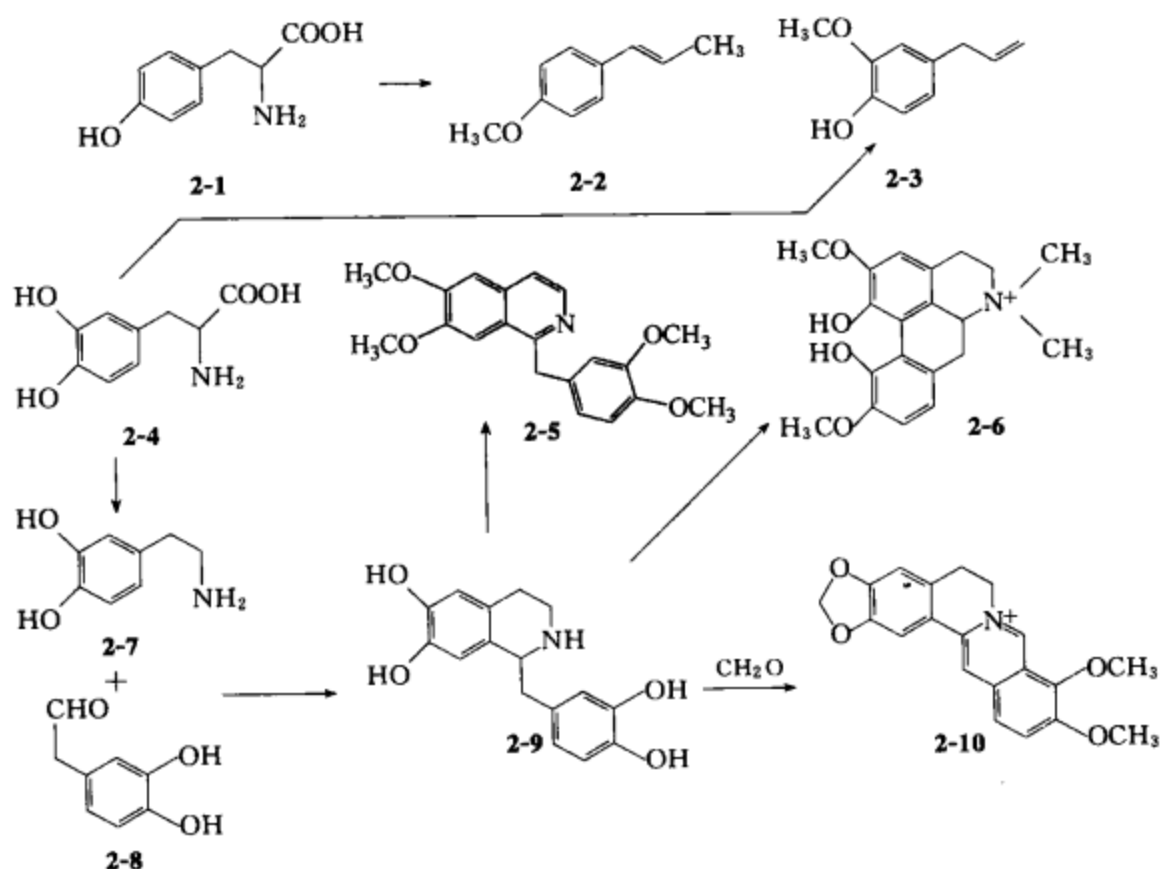


图 2-2 生物碱与氨基酸之间的生源关系

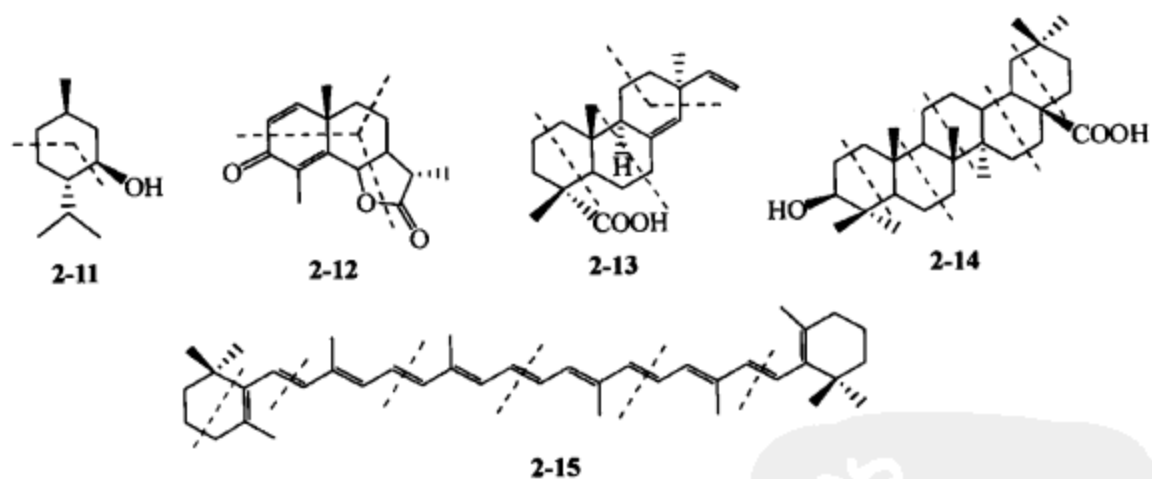


图 2-3 异戊二烯法则

后来，随着更多与异戊二烯法则不相符合的萜类化合物的出现，1953 年对上述法则做了修改。1956 年进一步发现在生物体内实际提供 C_5 单位的是甲戊二羧酸 (mevalonic acid, 2-16)。如在艾属植物中单萜化合物檀紫三烯 (santolinatriene, 2-19)、蒿酮 (artemisia ketone, 2-20)、艾醇 (yomogi alcohol, 2-21)、薰衣草醇 (lavandulol, 2-22) 和反式-5-羟基-2-异丙烯基-5-甲基-3-己烯-1-醇 (2-23) 就是经由焦磷酸二甲烯丙酯 (2-17) 生成焦磷酸菊酯 (2-18) 后，在 a、b 和 c 处断裂生成 (图 2-4)。

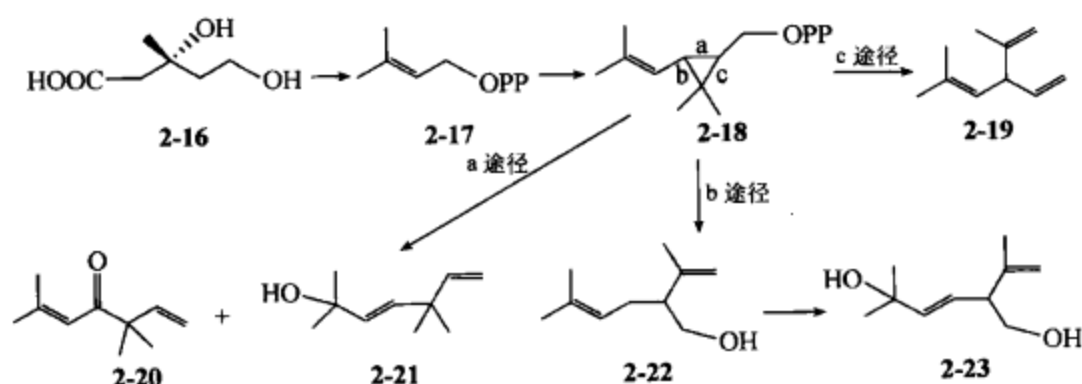


图 2-4 甲戊二羧酸途径生成萜类植物中单萜化合物

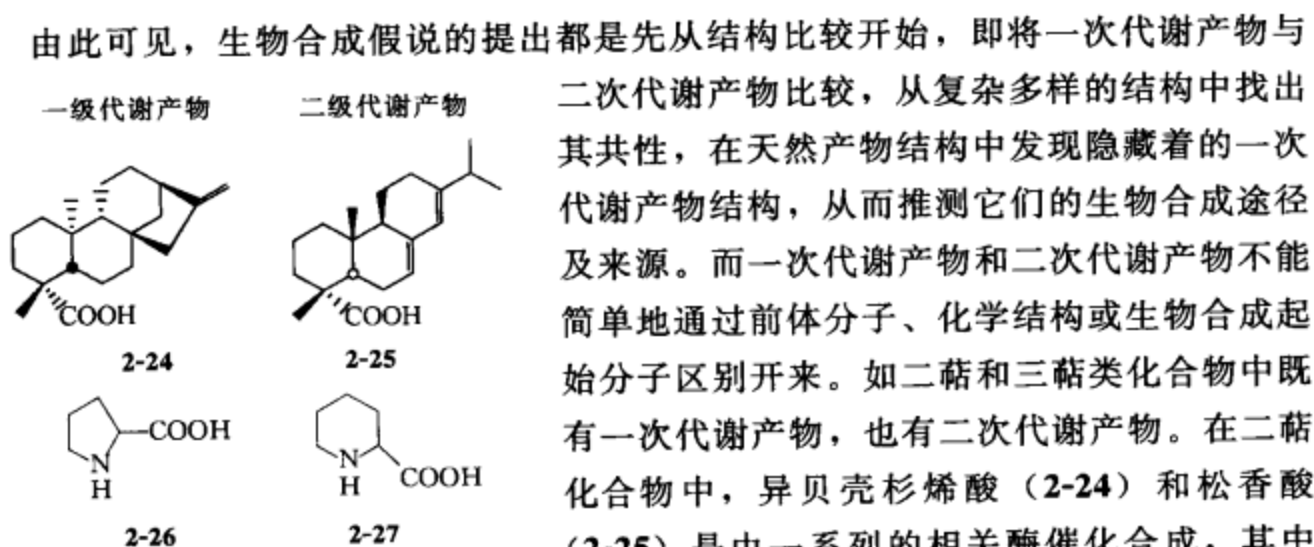


图 2-5 一次代谢产物与二次代谢产物区别

由此可见，生物合成假说的提出都是先从结构比较开始，即将一次代谢产物与二次代谢产物比较，从复杂多样的结构中找到其共性，在天然产物结构中发现隐藏着的一次代谢产物结构，从而推测它们的生物合成途径及来源。而一次代谢产物和二次代谢产物不能简单地通过前体分子、化学结构或生物合成起始分子区别开来。如二萜和三萜类化合物中既有一次代谢产物，也有二次代谢产物。在二萜化合物中，异贝壳杉烯酸（2-24）和松香酸（2-25）是由一系列的相关酶催化合成，其中 2-24 是赤霉素（一种植物生长激素）合成中的重要中间产物，属于一次代谢产物；2-25 是一种树脂组分，属于二次代谢产物，主要分布于豆科和松科植物中。再如，脯氨酸（2-26）是一种必需氨基酸，属于一次代谢产物，而其类似物六氢吡啶羧酸（2-27）是一种生物碱，因而被认为是二次代谢天然产物（图 2-5）。

2.2.3 主要的生物合成途径

虽然从自然界得到的化合物总数非常多，其结构也千变万化，非常复杂，但如果仔细加以分析，则不难看出它们均由一定的基本结构单位按不同方式组合而成。常见的基本结构有以下几种类型： C_2 单位（乙酸单位），如脂肪酸、酚类、苯醌等聚酮类（polyketide）化合物； C_5 单位（异戊烯单位），如萜类、甾类等； C_6 单位，如香豆素、木质素等苯丙素类化合物；氨基酸单位，如生物碱类化合物；复合单位，是由上述单位复合构成。主要生物合成途径如下。

2.2.3.1 乙酸-丙二酸途径（AA-MA 途径）

脂肪酸、酚类、萜类等化合物均由这一途径生成。

（1）脂肪酸类 碳链中碳原子数为偶数的脂肪酸生物合成的起始物质是乙酰辅酶 A（2-28），起延伸碳链作用的是丙二酸单酰辅酶 A（2-29），碳链的延伸由缩合

及还原两个步骤交替进行。碳链中碳原子数为奇数的脂肪酸，起始物质不是乙酰辅酶 A，而是丙酰辅酶 A (propionyl CoA)，支链脂肪酸的前体则为异丁酰辅酶 A (isobutyryl CoA)、 α -甲基丁酰辅酶 A (α -methylbutyryl CoA) 及甲基丙二酸单酰辅酶 A (methyl malonyl CoA) 等，但缩合及还原过程类似 (图 2-6)。

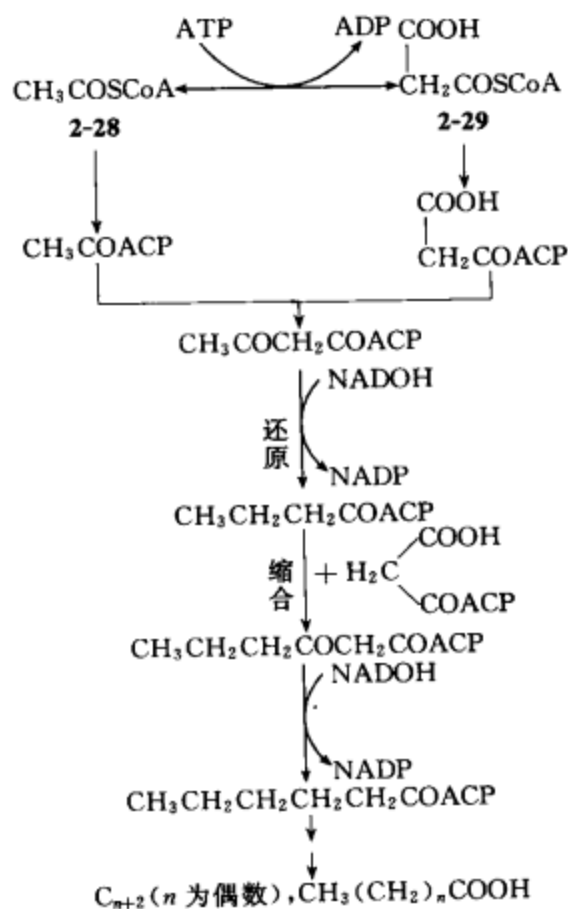


图 2-6 饱和脂肪酸的生物合成途径

自然界广泛分布的不饱和脂肪酸类，其生物合成可能有几条途径，但以先经生物氧化成羟基衍生物，再经脱水后生成为主 (图 2-7)。

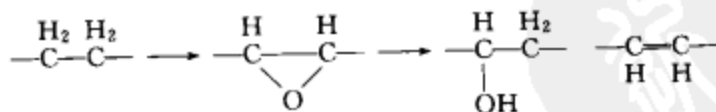


图 2-7 不饱和脂肪酸类的生物合成途径

(2) 酚类 天然酚类化合物的生物合成与脂肪酸的不同之处是，在由乙酰辅酶 A 起始后，延伸碳链过程中只有缩合过程，生成的聚酮类中间体 (2-30) 经不同途径环合而成萘酚类 (间苯二酚型，2-31)、乙酰间苯三酚 (间苯三酚型，2-32)、四乙酸内酯 (内酯型，2-33)，其特点是芳环上的含氧取代基 ($-\text{OH}$ ， $-\text{OCH}_3$) 多互为间位 (图 2-8)。

(3) 蒽酮类 以中药决明子为例，其中所含成分有红镰刀霉素 (rubrofusarin, 2-34)、决明子内酯 (toralactone, 2-35)、决明酮 (torachrysone, 2-36) 及大黄素

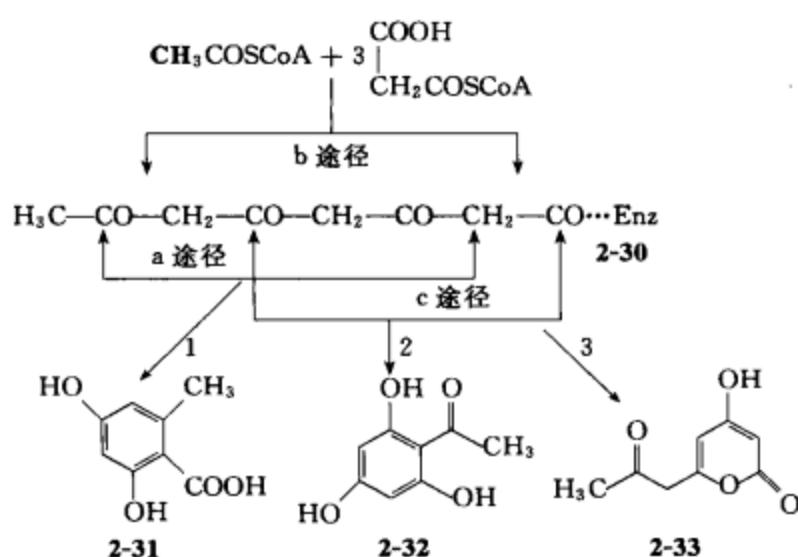


图 2-8 酚类的生物合成途径

1—经 a 途径生成 2-31；2—经 b 途径生成 2-32；3—经 c 途径生成 2-33

甲醚 (phycion, 2-37) 等。粗略看上去，似乎同一种植物中出现了几种结构类型不同的化合物。但就生物合成而言，其实它们均由同一途径即 AA-MA 途径生成，都可归入聚酮类化合物中 (图 2-9)。

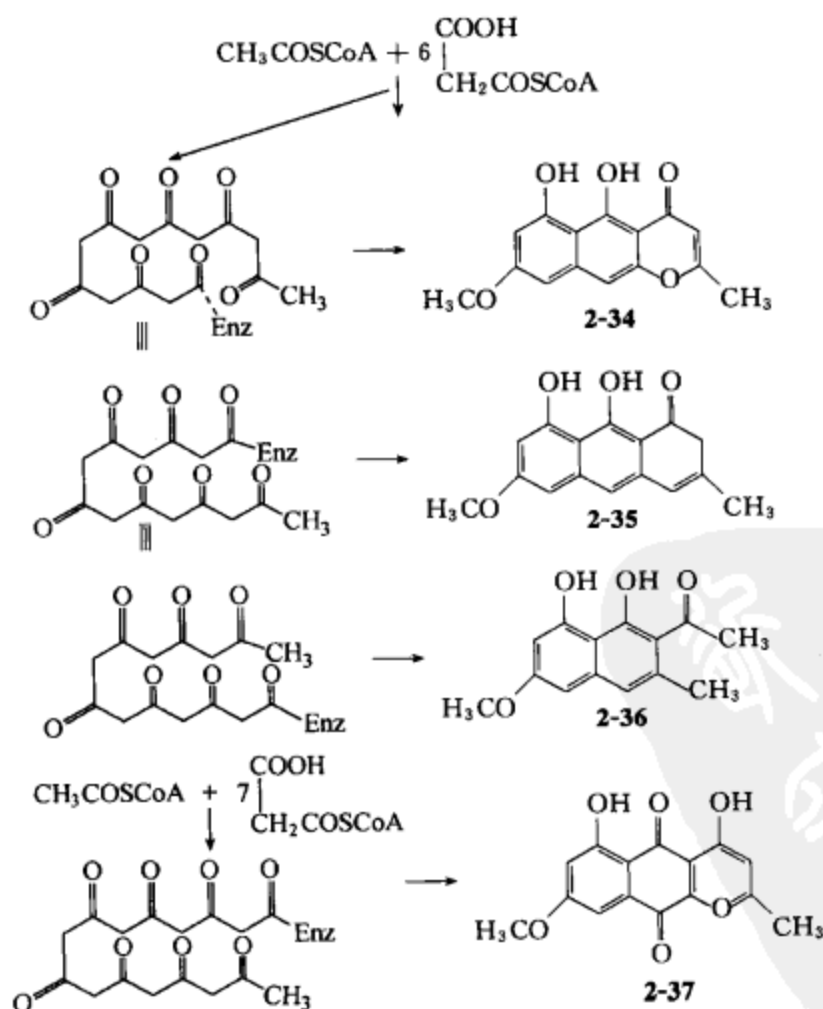


图 2-9 决明子的部分成分合成途径

2.2.3.2 甲戊二羟酸途径 (MVA 途径)

生物体内真正的异戊烯基单位为焦磷酸二甲烯丙酯 (DMAPP, 2-17) 及其异构体焦磷酸异戊烯酯 (IPP, 2-37), 它们均由 MVA (2-16) 变化而来, 通过 DMAPP 和 IPP 头尾连续加成的方式生成单萜 (C_{10})、倍半萜 (C_{15}) 和二萜 (C_{20}), 三萜 (C_{30}) 是由两个 C_{15} 单位头对头连接形成的, 四萜 (C_{40}) 是由两个 C_{20} 单位头对头连接形成的 (图 2-10)。

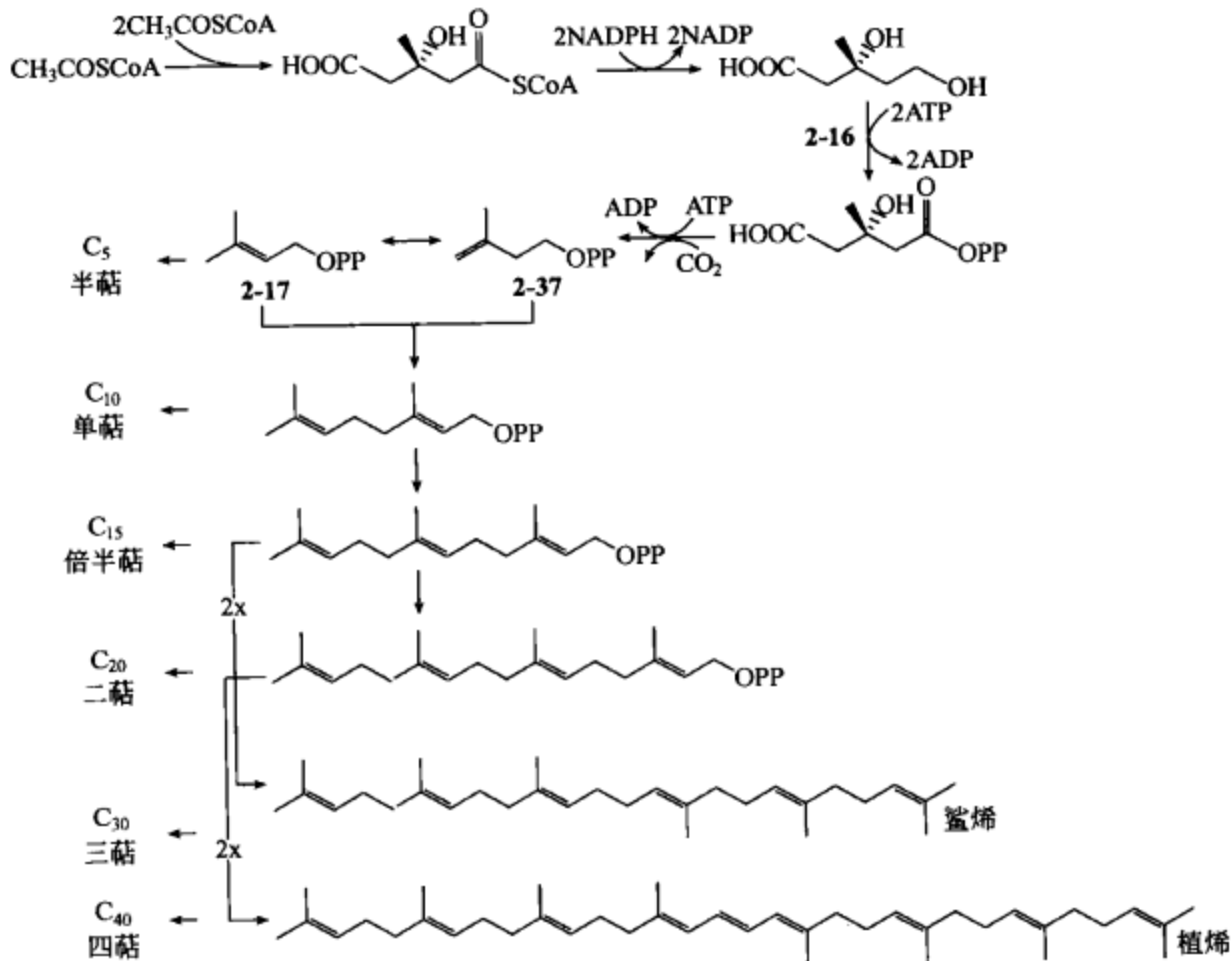


图 2-10 萜类成分甲戊二羟酸的合成途径

2.2.3.3 桂皮酸途径 (cinnamic acid pathway) 及莽草酸途径

天然产物中具有 C_6-C_3 骨架的苯丙素 (phenylpropanoids)、香豆素 (coumarins)、木脂素类 (lignans) 以及具有 $C_6-C_3-C_6$ 骨架的黄酮类化合物 (flavonoids) 极为多见。其中的 C_6-C_3 骨架均由苯丙氨酸 (phenylalanine, 2-38) 经苯丙氨酸脱氨酶 (phenylalanine ammonia-lyase, PAL) 脱去氨后生成的桂皮酸 (2-39) 而来。酪氨酸 (2-40) 经酪氨酸脱氨酶 (tyrosine ammonia-lyase, TAL) 脱氨后理论上虽也可生成对羟基桂皮酸 (2-41), 但 TAL 在高等植物中的分布远比 PAL 少 (仅见禾本科植物), 且几乎不存在使苯丙氨酸氧化成酪氨酸的酶, 所以这一途径实际上可以忽略不计。桂皮酸经氧化、甲基化、还原等反应转化为咖啡酸 (2-42)、阿魏酸 (2-43)、芥子酸 (2-44)、芥子醛 (2-45)、芥子醇 (2-46)、对香

[illegible]

1885 年从八角属植物 (*Illicium religiosum*) 果实中分离到一个当时未料想会如此重要的化合物——莽草酸 (2-60)。经过研究证明它是植物和微生物中芳香氨基酸 L-苯丙氨酸、L-酪氨酸或 L-色氨酸生物合成的一个关键中间体 (动物体不能进行这一生源途径合成), 这三个芳香氨基酸本身是许多重要二次代谢产物的前体, 从莽草酸到芳香族氨基酸的生物合成途径叫莽草酸途径。它本身是通过碳水化合物代谢产生, 第一步磷酸烯醇丙酮酸 (phosphoenol pyruvate, PEP, 2-55) 和 D-赤藓糖-4-磷酸盐 (D-erythrose-4-phosphate, 2-56) 进行立体专一性的醛醇缩合, 生

成 3-去氧-D-阿拉伯糖海泊酸-7-磷酸盐 (3-deoxy-D-binoheptulosonic acid-7-phosphate, DAHP, 2-57)。DAHP 的环合产生去氢奎尼酸 (dehydroquinic acid, 2-58)，然后经脱水产生去氢莽草酸 (dehydroshikimic acid, 2-59)，再经可逆性加氢产生莽草酸 (2-60) 及其 3-磷酸酯 (2-61)，2-61 与 PEP 缩合生成 2-62，再经过 1,4-消除转变到分支酸 (chorismic acid, 2-63)，然后形成芳香氨基酸。

分支酸的氨基化通过邻氨基苯甲酸 (anthranilic acid, 2-64) 形成色氨酸 (2-65)。苯丙氨酸 (2-67) 和酪氨酸 (2-40) 是通过柏利弗克酸 (prephenic acid, 2-66) 形成，而柏利弗克酸是从分支酸经过 Claisen 重排形成 (图 2-12)。

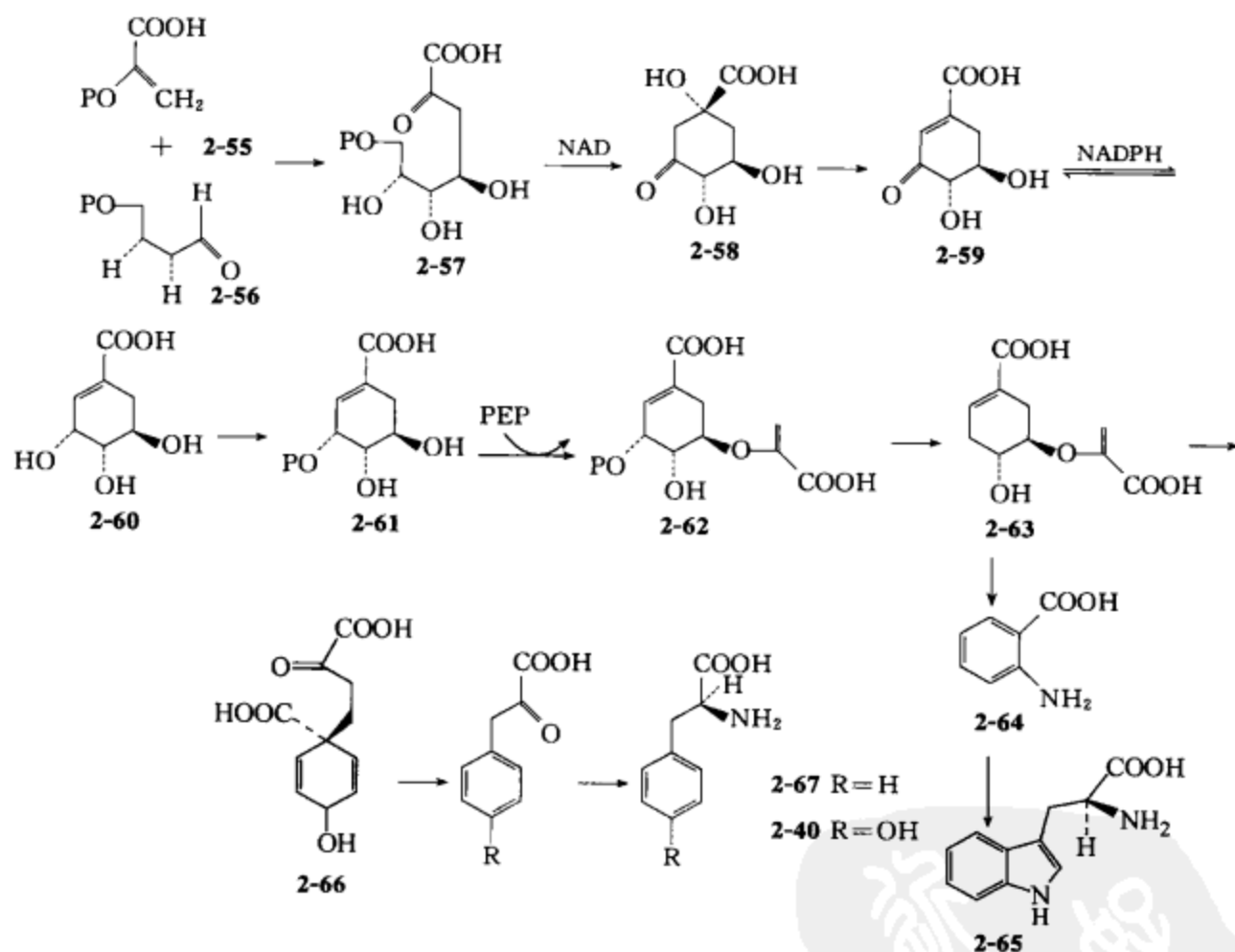


图 2-12 莽草酸途径

2.2.3.4 氨基酸途径 (amino acid pathway)

天然产物中的生物碱类成分均由氨基酸途径生成，如小檗碱 (2-10) 生物合成途径以 L-酪氨酸 (2-1) 经过酪氨酸脱羧酶和酚类氧化酶催化生成多巴胺 (2-7)。同时，L-酪氨酸经过转氨酶催化生成对羟苯丙酮酸 (2-68)，再经脱羧酶催化生成对羟苯乙醛。多巴和对羟苯乙醛经 (S)-去甲乌药碱合酶催化生成 (S)-去甲乌药碱 (2-69)，经过甲基转移酶、N-甲基转移酶、羟化酶等催化相继生成 (S)-乌药碱 (2-70)、(S)-N-甲基乌药碱 (2-71)、(S)-3'-羟基-N-甲基乌药碱 (2-72) 和 (S)-网状番荔枝碱 (2-73)。(S)-网状番荔枝碱经过小檗碱桥酶催化生成 (S)-金黄紫堇

碱 (2-74)，最后经过甲基转移酶、坎那定合酶和氧化酶催化生成小檗碱 (2-10) (图 2-13)。

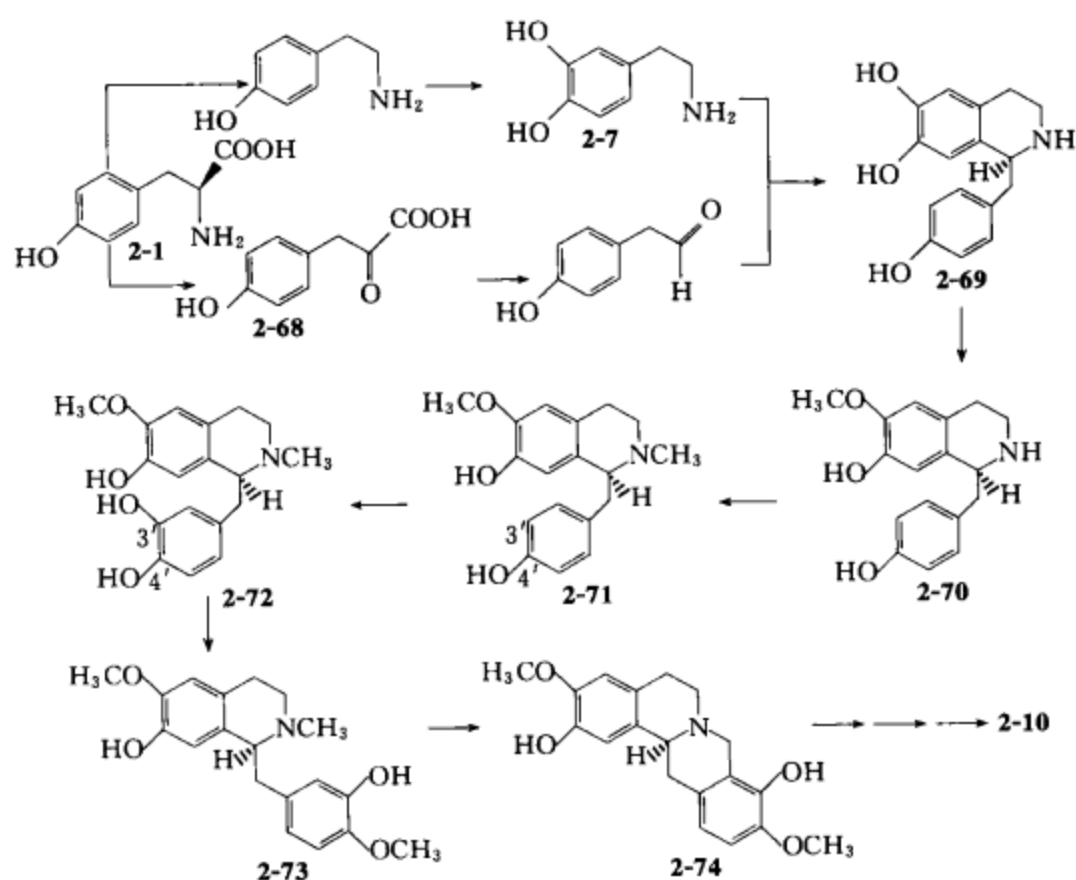


图 2-13 氨基酸生物合成途径

(S)-网状番荔枝碱也被称作是“化学千面手”，根据酶促反应前分子折叠和扭曲的程度，它可以形成众多不同结构的四氢苯基异喹啉类生物碱。如吗啡 (2-75)、前托品 (2-76)、血根碱 (2-77)、延胡索碱 (2-78)、小檗碱 (berberine, 2-10)、罂粟碱 (papaverine, 2-5)、那可汀 (noscapine) (2-79) 和可卡因 (2-80) 等 (图 2-14)。

并非所有的氨基酸都能转变为生物碱，已知作为生物碱前体的氨基酸，在脂肪族氨基酸中主要有鸟氨酸 (ornithine)、赖氨酸 (lysine)，在芳香族氨基酸中则有苯丙氨酸 (phenylalanine)、酪氨酸 (tyrosine) 及色氨酸 (tryptophan) 等。其中，芳香族氨基酸来自莽草酸途径，脂肪族氨基酸则基本上来自三羧酸循环中形成的 α -酮酸经还原氨化 (transamination) 生成。

2.2.3.5 复合途径

对于结构稍复杂的天然产物，其分子中各个部分可能来自不同生物合成途径。如大麻萜酚酸 (cannabigerolic acid, 2-81)、大麻二酚酸 (cannabidiolic acid, 2-82) 来自 MVA 和 AA-MA 途径，而查耳酮类 (chalcones, 2-83)、二氢黄酮类 (dihydroflavones, 2-84) 等化合物则来自 AA-MA 途径和氨基酸途径 (图 2-15)。

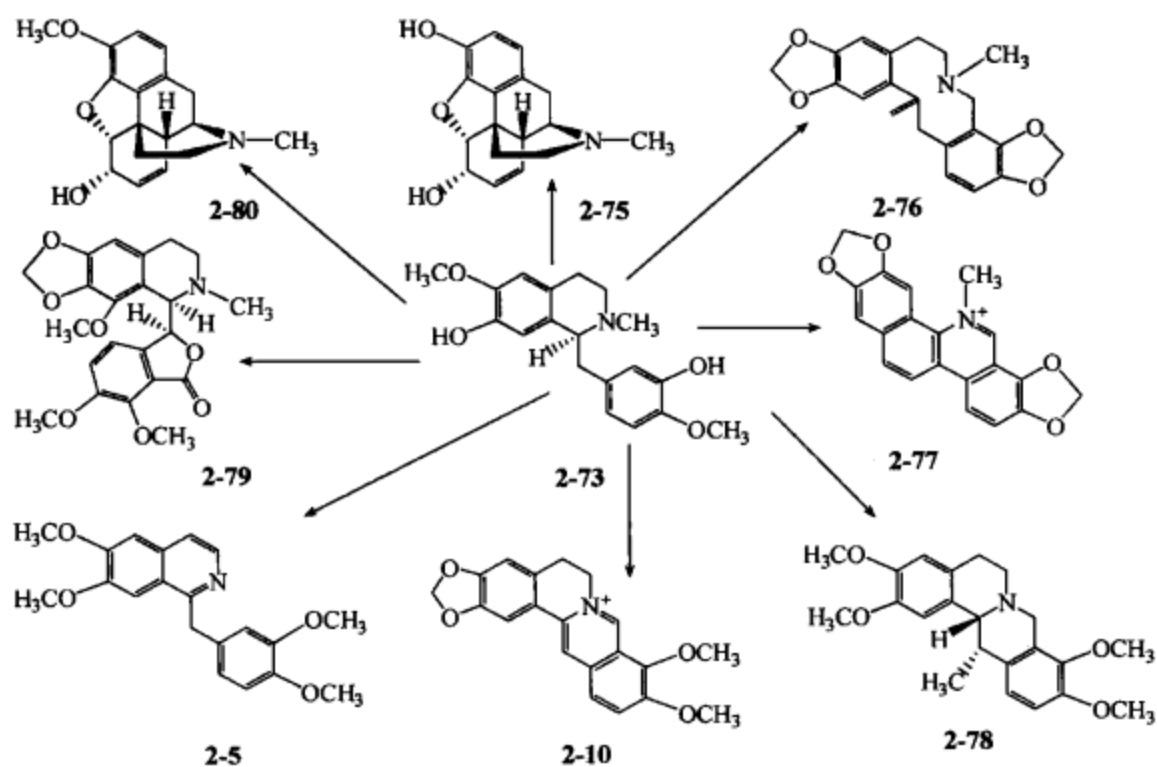


图 2-14 (S)-网状番荔枝碱与其他生物碱的转化

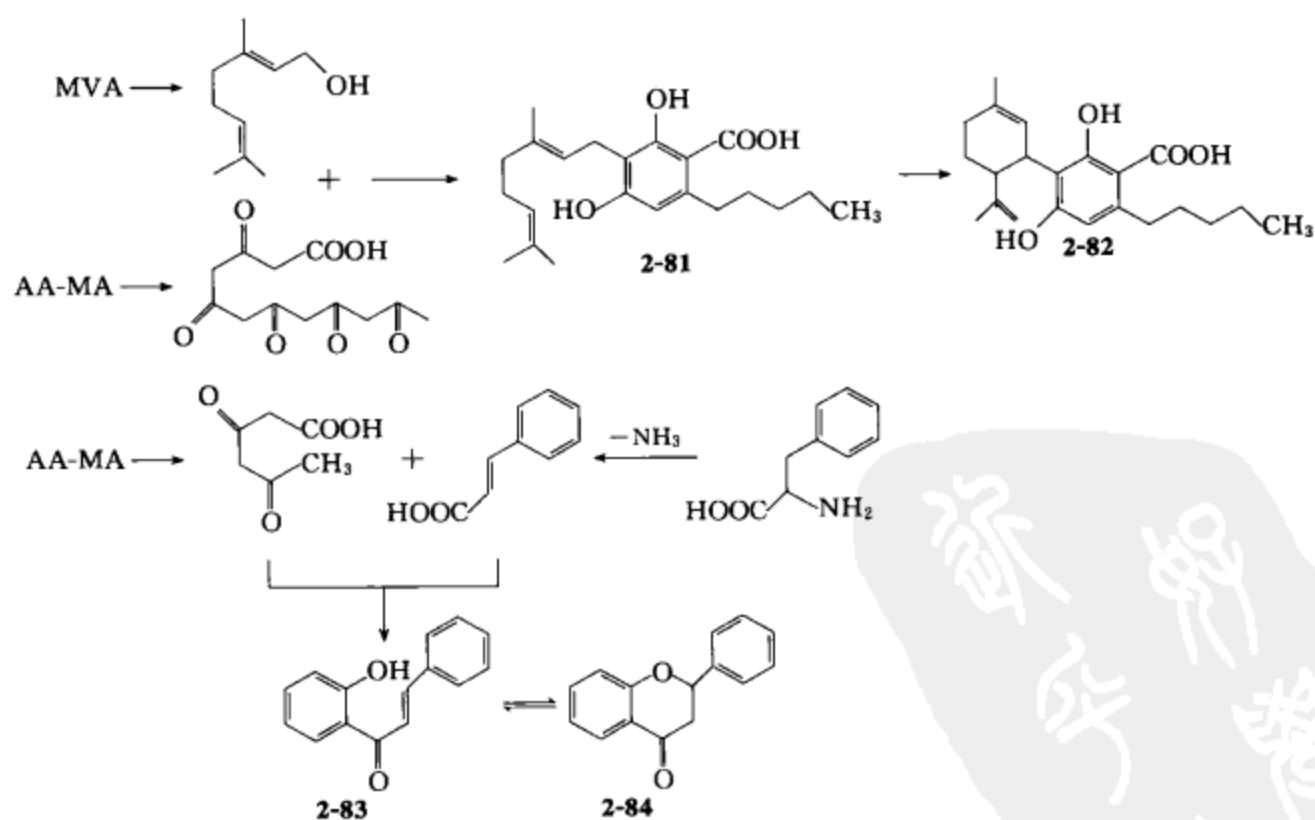


图 2-15 复合途径

常见的复合生物合成途径有以下几种：乙酸-丙二酸-莽草酸途径；乙酸-丙二酸-甲戊二羧酸途径；氨基酸-甲戊二羧酸途径；氨基酸-乙酸-丙二酸途径；氨基酸-莽草酸途径。

应当指出，许多天然产物均由上述特定的生物合成途径形成，但是也有少数例外。如植物界中广泛分布的没食子酸（gallic acid, **2-85**），在不同的植物中，或由莽草酸（**2-60**）直接生成（如老鹳草 *Gernium hyperacium*，图 2-16 途径 1），或由桂皮酸途径（**2-39**）生成（如漆树 *Rhus typhina*，图 2-16 途径 2），或由苔藓酸（**2-86**）途径苔黑酚（**2-87**）生成（如黑附球菌 *Epicoccum nigrum*，图 2-16 途径 3）。

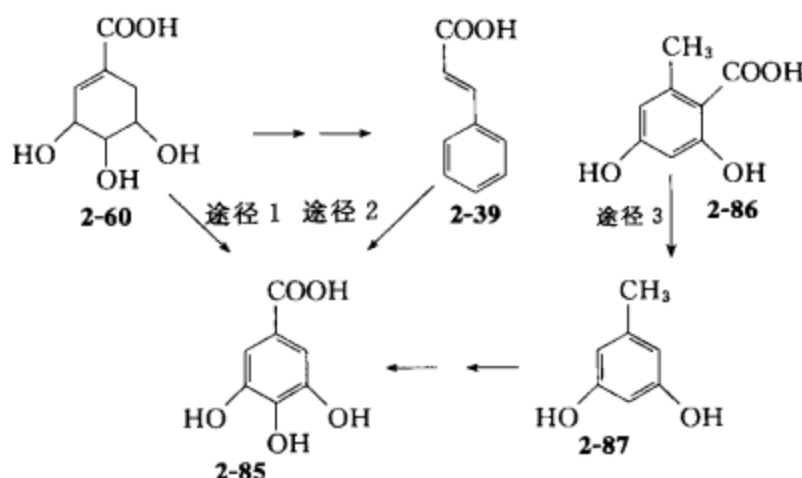


图 2-16 没食子酸途径

2.3 天然产物生源合成研究应用

尽管人们已经注意到天然产物（natural product）在化学结构和生物合成途径方面的复杂性，但早期的研究认为它们没有什么生物学意义，所以没有引起植物学家的足够重视。然而有机化学家却一直对这些奇异的天然产物很感兴趣，从 19 世纪 50 年代起就对它们的化学特性开展了广泛的研究。天然产物的研究促进了分离技术、结构分析的光谱方法以及合成方法等的发展，这些技术和方法奠定了当代有机化学的基础。对天然产物的研究，不仅仅是因为它们的学术价值，更主要是它们巨大的实用价值，如染料、多聚体、纤维、黏合剂、油、蜡、调味剂、香水和药物等天然产物在人类生产生活中起着重要的作用。

近年来，随着生命科学领域的相互渗透，对天然产物高度广泛的生物学效应有了越来越多的认识，使生物学家对这些分子在植物体内（特别是在生态相互作用的大环境中）可能扮演的角色重新进行评价，发现许多天然产物都有极其重要的适应性作用，可以协助植物抵抗食草动物、抵御微生物侵染、吸引授粉昆虫和种子传播者，以及作为种间抑制剂（种间感应化合物，影响植物间的相互作用竞争）。

同时，有机化学的发展也对天然产物的生物合成起到了推动作用。首先，生物合成理论有助于结构的推导，当推定结构有困难时，生物合成的假设是非常有帮助的。其次，核磁共振技术在天然产物结构解析中的应用，尤其是 X 射线衍射对结构的测定已普遍应用，可以对天然产物结构进行精确的解析。再者，许多植物和微

生物的生物合成途径已被确定,确定了结构的天然产物就可以依据已知的生物合成途径进行生源假说的推测,这种推测还可以进一步用试验确证。

正是沿着这样一种发展,天然产物生源合成研究已经渗透到了植物学、植物化学、生物化学和生物工程等领域。

2.3.1 植物学领域

2.3.1.1 天然产物与植物防御

在植物与环境病原相互作用中,植物对病原的侵袭信号能够感知、转导,并做出一系列有效的防御反应,以抵御病原微生物的侵袭。植物可以识别入侵的病原体,启动体内的防御反应,将入侵物限制在特定位点,从而使植物免受更大侵害。如早期对兰花 (*Orchis moris*) 和兰科植物 *Loroglossum hircinum* 的研究发现,当受到病菌丝核菌 (*Rhizoctonia*) 侵染后,会对该病原菌产生一定抗性,把受感染的组织放入培养基中进行真菌接种,结果发现真菌的生长受到抑制。1940年, Mueller 在对马铃薯 (*Solanum tuberosum*) 的研究中发现,感染了马铃薯疫病菌 (*Phytophthora infestans*) 的马铃薯会分泌一种化学成分抑制病菌生长,后来该成分被命名为植物抗病毒素 (phytoalexin)。

随着天然产物的分离和结构确证的数量增加,人们相继发现植物防卫素几乎涉及所有的天然产物结构类型,有的植物只产生某一类植物防卫素,而有的植物却产生多种不相关类型的植物防卫素,在双子叶植物中大约有 30 个科的植物在受真菌感染后会产生植物防卫素。从结构类型来看,可以归纳为四种主要类型:酚类、萜类、乙炔衍生物及含氮化合物。其中,酚类成分依据结构骨架不同,分为黄酮类、二苯乙烯类、苯丙素类衍生物、香豆素及简单酚性成分等,如豆科植物中的 6 α -羟基高丽槐素 (2-88)、白藜芦醇 (2-89)、花旗松 (*Pseudotsuga menziessi*) 根中的木脂素羟基马台树脂酚 (2-90)、葱属植物 *Allium cepa* 中的脂肪族二酮类化合物 4-己基-1,3-环戊二酮 (2-91)、芭蕉科植物 *Musa acuminata* 中的 irenolone (2-92)、萘二甲酸酐 (2-93),桑科植物 *Broussonetia* 中的 7-羟基黄烷、7,4'-二羟基黄烷和 7,4'-二羟基-8-甲基黄烷 (2-94);萜类化合物是仅次于酚类化合物的第二大类植物防卫素,如松树、云杉和冷杉挥发油中的单萜 α -蒎烯 (2-95) 和 β -蒎烯 (2-96) 具有抗树皮甲壳虫及其病原真菌作用,艾菊中的除虫菊酯 (2-97) 具有杀虫作用,禾本科植物 (*Gramineae*) 中存在的二萜稻壳酮 A (2-98)、稻叶素 S (2-99)、莴苣族 (*Lactuceae*) 中的倍半萜 lettucein A (2-100);炔类植物防卫素主要存在于菊科植物,如 mycosinol (2-101)、蚕豆酮 (wyeron, 2-102);含氮类植物防卫素主要存在于石竹科、十字花科和罂粟科,如芸薹宁 (brassinin, 2-103)、spirobrassinin (2-104)、芸薹抗毒素 (brassilexin, 2-105) (图 2-17)。

同时,一些天然产物,如 1,8-桉叶油素 (2-106) 可以诱导蜜蜂、蛾子和蝙蝠

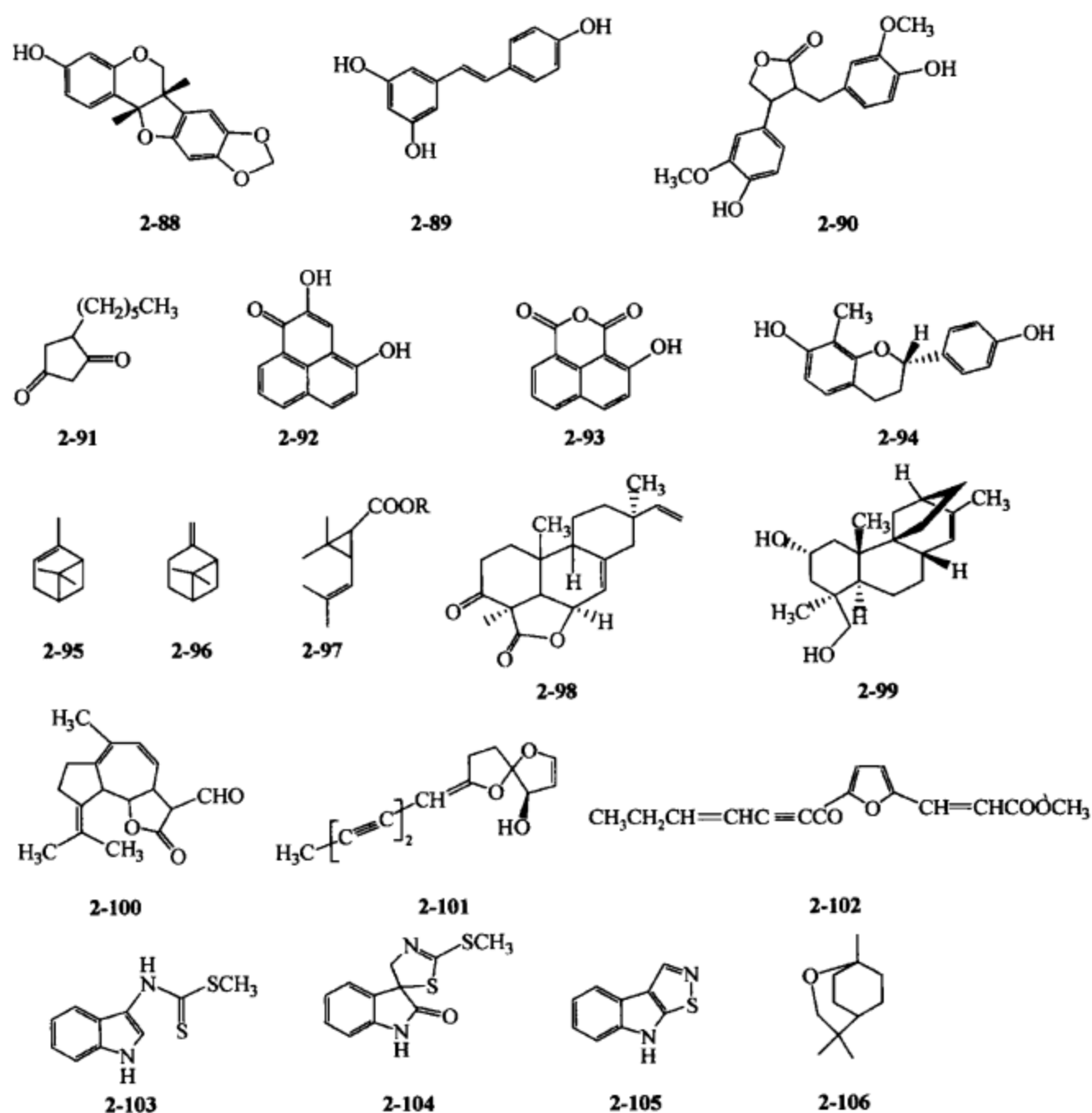


图 2-17 植物相关防卫素结构

等传粉动物，樟脑（2-107）、千里光碱（2-108）和白羽扇豆碱（2-109）可以作为食草动物的阻食剂，这些都是植物在生存或取食选择压力下进化出的化学防御系统（图 2-18）。

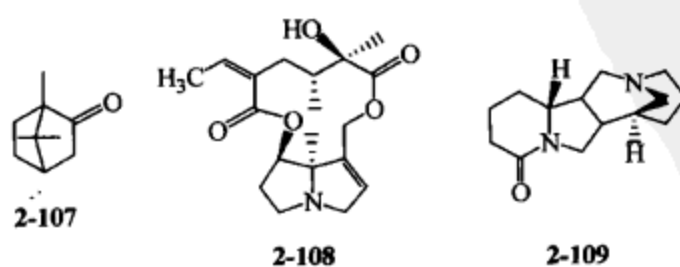


图 2-18 植物相关防御天然产物

2.3.1.2 天然产物生源合成与植物分类学

植物化学分类学 (plant chemotaxonomy) 是植物分类学与植物化学相互渗透、相互补充、互为借鉴而形成的一门新兴边缘学科。它以植物中分离得到的天然产物为依据, 以经典分类学为基础, 对植物加以分类和概述, 研究植物化学成分与植物类群间的关系, 探讨植物界的演化规律。

天然产物在植物体内的存在可以作为植物化学分类的依据, 但仅仅根据某类天然产物的有无作为分类的依据还是不够的。因为天然产物结构类型的多样性在一定程度上反映了生源及合成途径的复杂性。有的是同源殊途, 有的是异源同途, 还有平行演化等。如苍术属 *Atractylodes* 无叶柄组植物中的苍术炔 (2-110)、有柄组的二乙酰苍术素醇 (2-111) 是来源于同一个化合物, 无柄组在羟基转移酶和环合酶作用下形成苍术炔, 而有柄组缺少这两个酶直接生成直链的二乙酰苍术素醇 (图 2-19)。

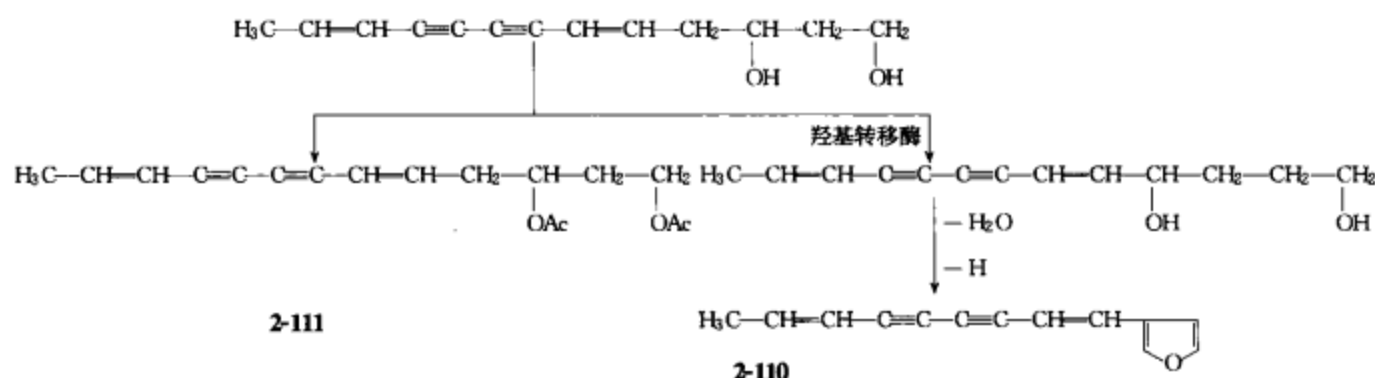


图 2-19 苍术炔生源合成

有些植物的天然产物类型相同, 结构相近, 似乎关系密切, 但分析它们的生物合成途径, 不仅起始物不同, 途径也不同, 或是产出化合物相似, 但它们之间的亲缘关系也并不密切。如分布在蓝雪科 (Plumbaginaceae)、胡桃科 (Juglandaceae)、鹿蹄草科 (Pyrolacaceae) 和紫草科 (Borraginaceae) 的萘醌类化合物石苺萘醌 (plumbagin, 2-112)、胡桃醌 (juglone, 2-113)、梅笠灵 (chimaphilin, 2-114) 和紫草素 (alkannin, 2-115) 的生物合成分别经过 AA-MA (途径 1)、邻琥珀酰苯甲酸 (路线 2)、尿黑酸 (途径 3) 和对羟基苯甲酸 (路线 4) 路线与途径生成 (图 2-20)。

化学分类学不仅重视生物合成的生源、途径及产物, 同时重视生物合成步骤的多少。巴西科学家 Gottlieb 提出, 只有同源化合物才有分类上的可比性。

另外, 不同学者基于同一科植物的植物系统学、植物化学和药理学等的研究报告, 利用植物系统关系-化学成分-疗效的相关性, 为植物分类学提供佐证。如毛茛科 (Ranunculaceae) 植物以木兰碱为代表的生物碱、三萜皂苷和毛茛苷为其主要化学特征和分类依据, 支持成立升麻亚科的观点。比较贝母属植物化学成分的变化

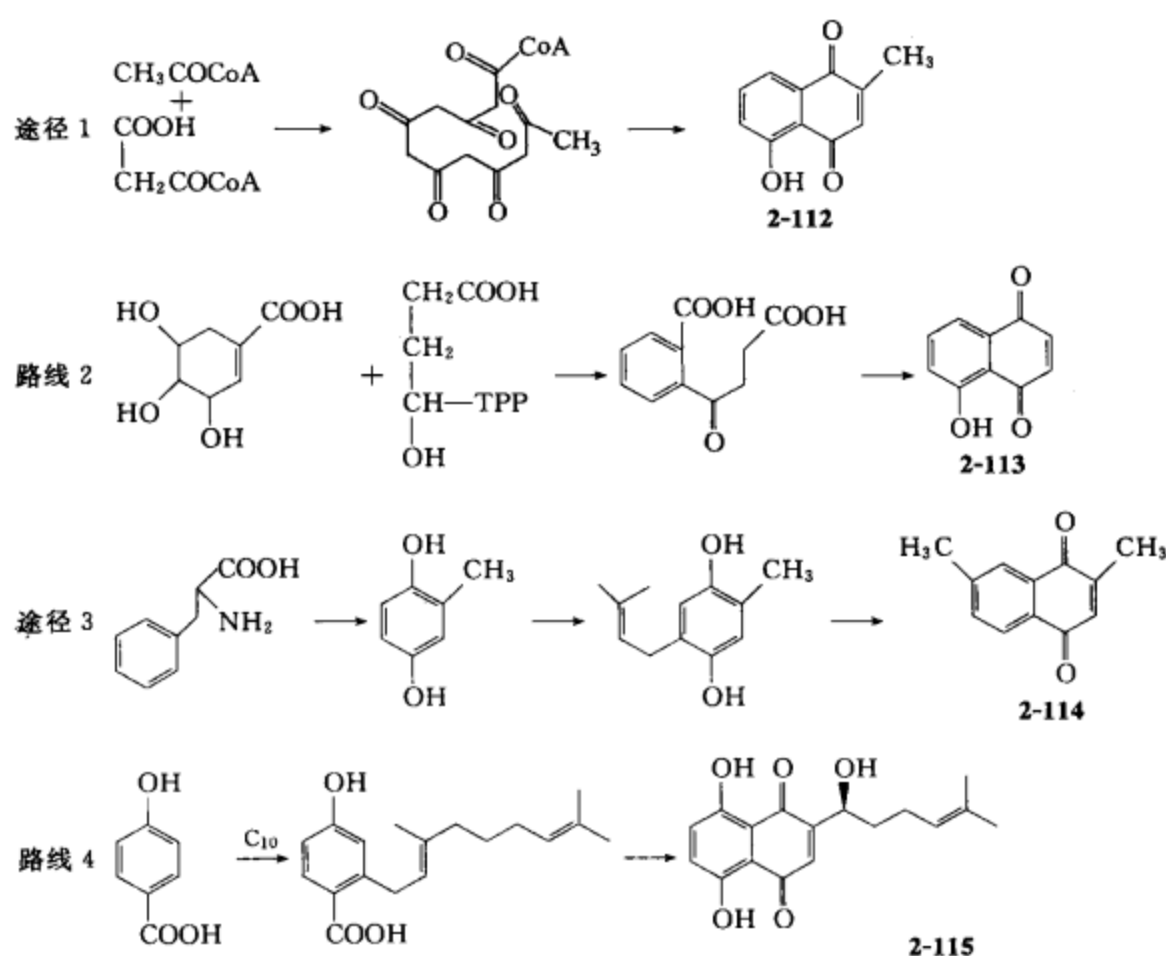


图 2-20 不同科植物萘醌生源途径

与形态学变化，发现有的化学成分比较一致但形态学的变化较大，如川贝母复合群，在植株形态和花部特征等方面表现出较大的差异，并且这种变化是连续的，从进化生物学角度考虑，川贝母复合群可能是一个比较年轻的并且正处于激烈分化的物种形成阶段；也有些化学成分相差较大而形态学变化不大，如湖北贝母和蒲圻贝母。

还有一些天然产物可能集中从单一科属中分离得到，并可能作为该科属特征性化合物。如喹诺酸类型的三萜化合物主要从茜草科与蓼科中得到。

2.3.2 植物化学领域

2.3.2.1 天然产物生源假说的提出

相对于植物系统的化学工作来说，一个植物种的特定天然产物结构生源假说是化学工作的一个延伸，是特定天然产物生源途径的一种推断。该部分工作涉及天然产物的提取、分离和结构确证，并需要依据现有的天然产物生物合成途径，将得到的系列天然产物结构犹如在儿童画中找到隐藏着的动物图案那样进行合理推测，具有一定的学术价值，并可以为生物化学或生物工程的研究提供一定的参考。如蝉翼藤 (*Securidaca inappendiculata* Hassk.) 中 3-羟基-4,6-二甲氧基吡酮 (2-119)、1,3,8-三羟基-4-甲氧基吡酮 (2-120) 和 2-甲氧基-3,4-亚甲二氧基二苯甲酮 (2-

121) 是通过莽草酸途径相继生成 4-羟基-3-甲氧基苯甲醛 (2-116)、4-羟基-3-甲氧基甲酸 (2-117) 和 2,3-二甲氧基-4-羟基二苯甲酮 (2-118) (图 2-21)。

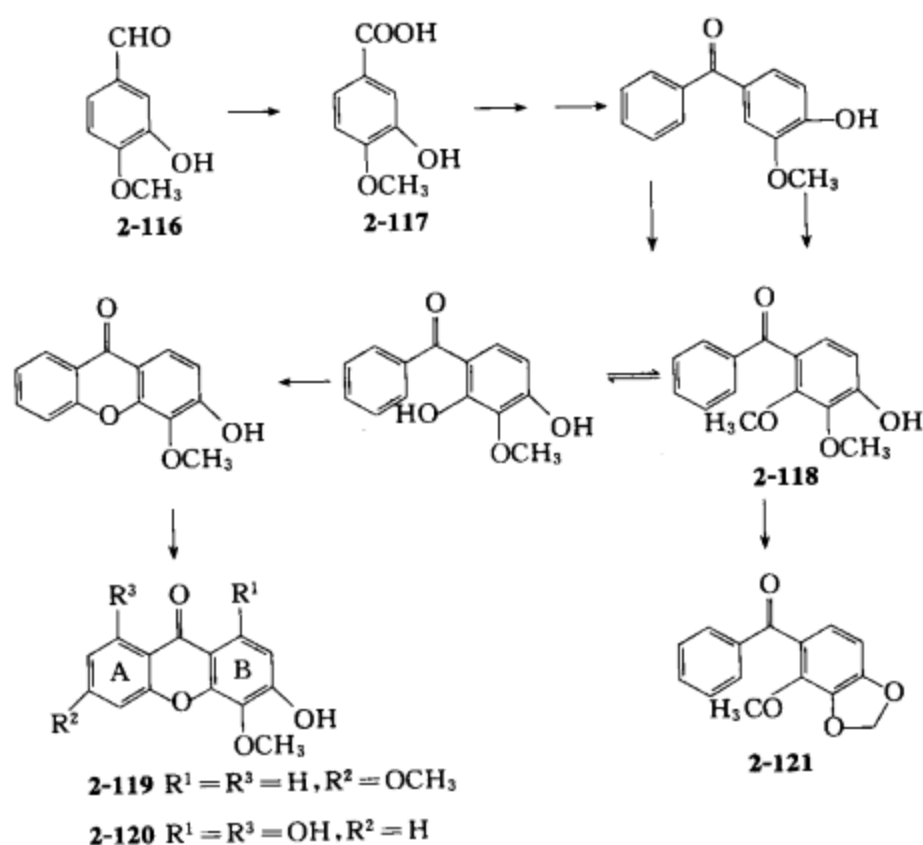


图 2-21 蝉翼藤中黄酮类生源合成途径

黄酮类化合物生源途径假说认为，氧化状态的该类型化合物 A 环可能是经过莽草酸途径 (shikimic acid pathway) 形成，B 环经过乙酸-丙二酸途径 (acetate-malonate pathway, AA-MA) 通过聚交酯 (polyketide) 作用形成，并且发现苯酮是该类化合物形成的一个重要中间体，而苯酮的来源主要有两个途径 (莽草酸途径和乙酸途径)。

对于植物化学工作者来说，一个类型化合物的提出是建立在分离得到一定数量的天然产物并确定其结构的基础上，而且幸运地得到了关键天然产物。如香茶菜属

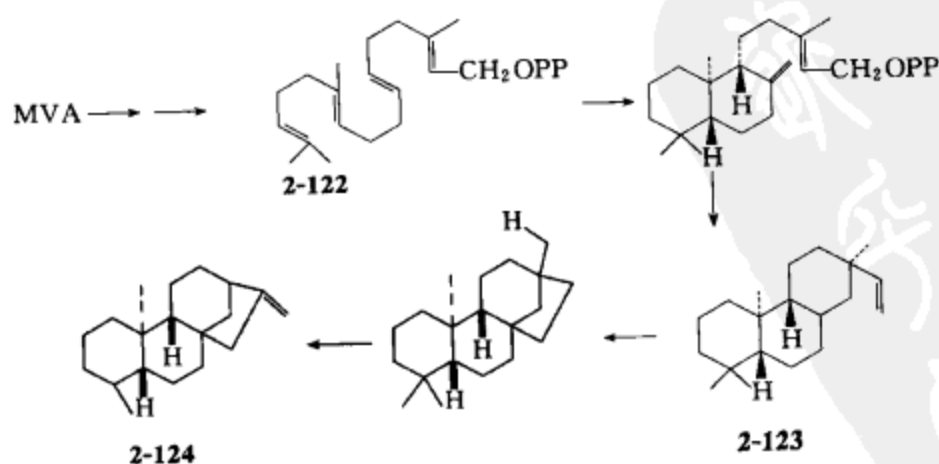


图 2-22 香茶菜属二萜生源合成途径

(*Isodon*) 对映贝壳杉烷型二萜 (ent-kaurane type diterpenoids) 生源途径的提出就是建立在这样的基础上。通过甲戊二羧酸途径 (mevalonic acid pathway, MVA) 形成前体焦磷酸香叶基香叶酯 (geranyl geranyl pyrophosphate, GGPP, 2-122), 再经过环合、脱羧和重排形成 (—)-copalol (2-123)、对映贝壳杉烯 (ent-kaurane, 2-124) (图 2-22)。

2.3.2.2 药用天然产物的动态积累

二次代谢产物在不同药用植物体内的合成和积累是药用植物在一定环境条件下长期生存选择的结果, 与产地的生态环境具有紧密的联系, 从而形成了“道地”药材的特性。外界生态环境条件的不同会导致相同品种药用植物体内的次生代谢过程的变化, 从而影响同一品种药材的内在质量。如对薄荷中黄酮类化合物香叶木素和蒙花苷含量的动态追踪, 发现蒙花苷含量在 7 月和 10 月中旬达到最高, 香叶木素含量在 7 月和 10 月中下旬达到最高; 毛茛酸模根中的白藜芦醇苷与生长环境中的钾呈现负相关, 其他土壤成分与其含量相关性弱; 不同产地连钱草中熊果酸和齐墩果酸含量存在差异, 其中熊果酸含量高于齐墩果酸; 同一产地连钱草中熊果酸含量高于齐墩果酸等。说明各种生态环境因素包括光、温度、土壤、空气以及生物因素均影响药用天然产物的累积。这些因素与二次代谢产物生物合成过程中的各种调控酶息息相关。

国外关于植物次生代谢与环境关系已开展了大量的研究, 但绝大多数研究都是基于探讨次生代谢产物抵御动物捕食和微生物侵袭、减弱不利生境条件危害、提高生存竞争能力为目的。所研究植物的次生代谢完全是在自然条件下进化产生, 其合成积累与胁迫生境形成了密切的关系。

2.3.3 生物化学与生物工程领域

生物学家通过天然产物生物合成关键酶和中间产物, 为天然产物生物合成提供直接的证据。如已经发现黄酮类化合物合成过程受到苯丙氨酸解氨酶、查耳酮合成酶、查耳酮异构酶、黄烷酮醇还原酶、异黄酮合成酶等系列酶的调控; 植物甾醇和三萜类皂苷生物合成过程中受到鲨烯合成酶、鲨烯氧化酶、氧化鲨烯环化酶和糖基转移酶等关键酶的催化。

目前天然产物及其药用价值研究领域正体现着学科交叉的许多最新进展, 如药用植物代谢工程对目标天然产物的植株进行细胞、组织和器官的培养, 富集药用产物。如名贵中药藏红花 (*Crocus sativus* L.) 具有很强的活血化瘀功效, 还具有良好的抗癌活性, 而且低毒、无副作用。藏红花培养细胞中产生的抗癌活性成分藏红花素 A (crocin A) 是人工栽培藏红花的 2~3 倍, 故可大大改善藏红花有效成分含量低的问题。以长春花叶和子房诱导的愈伤组织为实验材料, 对 4 种不同愈伤组织的生长, 吲哚生物碱组分、阿玛碱、长春酯碱的生物量进行分析比较, 发现用长

春花不同器官作为植物体外培养出的愈伤组织中吲哚生物碱的量有明显差异,从而为今后提高植物药用成分的量的研究提供了一条新途径。以野葛 [*Pueraria lobata* (Willd) Ohwi] 悬浮细胞为实验体系,进行了不同前体化合物和相关中间代谢产物的饲喂实验,发现对羟基苯丙酮酸和对羟基苯丙乳酸的处理同样有助于野葛培养细胞中异黄酮化合物的合成和积累,且单位细胞干重的黄酮产率有更大程度的提高;同时在野葛培养细胞中均可检测到对羟基苯丙酮酸还原酶和对羟基苯丙乳酸脱水酶的活性。因此,可推测对羟基苯丙酮酸和对羟基苯丙乳酸可能也是黄酮生物合成的前体化合物,并且植物黄酮类的生物合成可能存在另一条新的途径,即黄酮类生物合成的直接前体香豆酸是以预苯酸为前体化合物,越过了苯丙氨酸和桂皮酸(或肉桂酸)的中间代谢过程,经对羟基苯丙酮酸和对羟基苯丙乳酸而合成的。

另一个令人瞩目的代谢工程,就是通过转基因技术改变天然产物合成。如将编码长春花的色氨酸脱羧酶的 cDNA 转入油菜 (*Brassica napus*) 中,该酶催化单萜吲哚类生物碱的合成,将色氨酸合成方向由吲哚芥子油苷转向色胺,使油菜植物的成熟种子中基本不含吲哚芥子油苷,并且不积累色胺,解决了因为芥子油苷存在而影响油菜种子作为饲料,使它成为具有潜在经济价值的产品。

参 考 文 献

- [1] Dewick P M. Medicinal natural products. 2nd ed. Chippingham: John Wiley, 2002.
- [2] 周荣汉,段金廪. 植物化学分类学. 上海:上海科技出版社, 2005.
- [3] 吴立军. 天然药物化学. 第5版. 北京:人民卫生出版社, 2007.
- [4] 李绍顺. 天然产物全合成. 北京:化学工业出版社, 2004.
- [5] Sy L K, Brown G D. Phytochemistry, 2001, 58: 1159.
- [6] B B 布坎南著. 植物生物化学与分子生物学. 瞿礼嘉等译. 北京:科学出版社, 2004.
- [7] Harborne J B. Biochemical Systematics and Ecology, 1999, 27: 335.
- [8] 彭勇等. 中国中药杂志, 2006, 31: 1124.
- [9] 肖培根等. 植物分类学报, 2007, 45: 473.
- [10] Kang W Y, Hao X J. Biochemical Systematics and Ecology, 2006, 34: 585.
- [11] 臧鑫炎等. 天然产物研究与开发, 2007, 19: 356.
- [12] Kang W Y, et al. Helvetica Chimica Acta, 2005, 88: 2771.
- [13] Peres V, Nagem T J. Phytochemistry, 1997, 44: 191.
- [14] 孙汉董等. 香茶菜属植物二萜化合物. 北京:科学出版社, 2001.
- [15] 林彤等. 中药材, 2006, 29: 888.
- [16] 王振月等. 中国中药杂志, 2007, 32: 1984.
- [17] 王庆等. 南京中医药大学学报, 2006, 22: 44.
- [18] 段金廪等. 中国药学年鉴, 2007, 45.
- [19] 于荣敏. 中草药, 2006, 37: 1281.

第3章 药用植物化学成分的提取、分离与纯化

3.1 天然产物化学成分的提取

药用天然产物的提取方法很多，按形成的先后顺序和应用的普遍程度可分为经典提取方法和现代提取方法。经典提取方法是指使用比较普遍、技术比较成熟的传统提取方法，主要有溶剂提取法、水蒸气提取法和分子蒸馏法。其中，溶剂提取法是药用天然产物提取、分离、纯化和鉴定过程中最常应用的方法。按不同的研究习惯，溶剂提取法又可以分为浸渍法、渗漉法、煎煮法、回流法和连续回流提取法等。其中浸渍法、煎煮法和回流法不需要特殊的仪器，应用比较普遍。现代提取方法是以现代化的仪器为基础或是新发展起来的提取方法，主要有超临界流体萃取法、固相提取、固相微萃取法、微波提取技术、仿生提取技术和酶法提取技术等。

下面依据药用植物活性成分提取、分离、纯化和鉴定的研究方式，主要对经典方法中的溶剂提取法进行论述；现代提取方法多与仪器分析相关，故此，仅对超临界流体萃取法、固相提取和固相微萃取法进行简单介绍。

3.1.1 药用天然产物经典提取方法

如何进行天然产物的提取？一般有以下两种情况：第一种，从植物中提取分离已知的有效成分或已知化学结构类型者，如从槐米中提取芦丁、从葛根的干燥根中提取葛根素等；第二种，对有效成分的性质一无所知，如首次提取分离某一植物的化学成分。针对不同的情况，采取的提取分离方法也不相同：对于第一种情况，一般先查阅有关资料，搜集该成分的各种提取方法，再根据具体条件进行选择；对于第二种情况，一般应在临床试验或药理试验指导下，经不同溶剂提取，以确定有效部位，然后再逐步划分，追踪有效成分最集中的部位，最后分离得到有效单体。

3.1.1.1 溶剂提取法

溶剂提取法是利用相似相溶原理，选用与所需成分极性相似（即溶解度大）的溶剂将其从植物中溶解出来，而此溶剂对其他共存成分不溶解或溶解度小。将溶剂加入植物原料中，溶剂由于浸润、渗透作用逐渐透入到细胞内，溶解可溶性物质，造成细胞内外的浓度差，于是细胞内的高浓度溶液不断向外扩散，溶剂又不断进入药材组织细胞中，如此反复，直至细胞内外浓度达到动态平衡时，将浓溶液滤出，

原料中再分次加入新溶剂，就可以把所需要成分大部分溶出甚至近于完全溶出。

天然产物在溶剂中的溶解度与溶剂性质直接相关。溶剂可分为水、亲水性有机溶剂和亲脂性有机溶剂三类。溶剂极性的大小可以根据介电常数 (ϵ) 的大小来判断。常见溶剂的介电常数及极性排列顺序见表 3-1。

表 3-1 常用溶剂的介电常数及极性排列

溶剂	介电常数(ϵ)	极性
石油醚	1.8	弱 ↓ 强
苯	2.3	
乙醚	4.3	
三氯甲烷	5.2	
乙酸乙酯	6.1	
丙酮	21.5	
乙醇	26.0	
甲醇	31.2	
水	80.0	

提取药用植物中的有效成分，通常可采用不同极性溶剂依次提取的方法，由低极性到极性分步进行。具体操作是将原植物粉碎后按下列溶剂顺序依次进行提取：①石油醚（提取脂溶性大的化合物，如油脂、蜡、叶绿素、精油等）；②乙醚（提取树脂及一些极性基团少的化合物，如甾体、某些生物碱、有机酸、黄酮及香豆素苷元等）；③三氯甲烷或乙酸乙酯（提取生物碱及许多中等极性成分）；④丙酮、乙醇或甲醇（提取极性较大的化合物，如生物碱的盐、苷类、鞣质等）；⑤水（提取水溶性化合物，如氨基酸、糖类、无机盐等）。提取时可采用冷水、热水、酸水、碱水等不同方法。也可以将植物先用水或乙醇提取，所得总提取物拌以硅藻土等辅料，使成粉末后，再按上述次序逐步分离。

当用溶剂法提取天然产物时，如何确定适宜的溶剂是关键，溶剂选择合适就能顺利地把有效成分提取出来，或者尽可能减少干扰杂质的提出，从而降低对活性天然产物分离和纯化操作的影响。适宜的溶剂应符合以下要求：对目标成分溶解性大，对共存杂质溶解度小；不与目标产物发生化学反应；价廉、易得、浓缩方便，并且安全无毒。一般研究体系中常使用水、乙醇、甲醇和丙酮，或者是不同浓度的乙醇-水、甲醇-水和丙酮-水混合溶剂。提取溶剂体系中尽量控制使用三氯甲烷（对环境污染）、苯（毒性大）及乙醚（易燃）等溶剂。

(1) 有机溶剂提取法 有机溶剂提取法适用于脂溶性基团占优势的化合物，如甾体、萜类、蒽醌等；有些化合物虽然也能溶于水，但为了使杂质尽量少带出来，也常常使用有机溶剂提取。常用提取溶剂有乙醇-水、甲醇-水、丙酮-水。

按照操作的不同，有机溶剂提取法可以分为以下 4 种方式。

a. 浸渍法：将植物粉碎成适当大小的颗粒，装入容器中，加入溶剂，以能浸

透药材并稍有过量为宜，振荡或搅拌，放置一段时间，过滤，滤渣另加新溶剂。如此提取两次，滤液合并浓缩后得提取物。

b. 渗漉法：将植物粉末装入渗漉筒中，不断自上方添加新溶剂，使其向下渗透过药材，从渗漉器下部流出浸出液的一种浸出方法。

c. 回流提取法：采用回流加热装置，以免溶剂挥发损失。一般第一次加热回流保持沸腾约 1h，第二、三次分别约为 0.5h。

d. 连续回流提取法：实验室常用脂肪提取器（或称索氏提取器）进行连续提

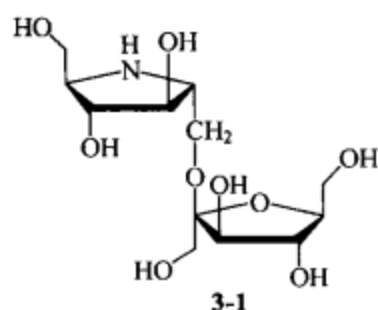


图 3-1 剑木豆 (*Baphia nitida* Lodd.) 中分离得到的新生物碱

取。如对毛茛科植物铁棒锤 (*Aconitum anthor-
ra*) 地上部分中黄酮类成分进行提取时，首先将原植物干燥粉末经环己烷脱脂，在索氏提取器中依次用氯仿、氯仿-甲醇 (9 : 1) 和甲醇提取，浸膏经反相高效液相 (RP-HPLC) 色谱柱，分别得到不同的黄酮类化合物。

① 以乙醇-水为溶剂的提取 这里介绍生物碱及生物碱苷的提取。取蝶形花科植物剑木豆 (*Baphia nitida* Lodd.) 干燥叶 8kg，用乙醇-水 (1 : 1，体积比) 提取，滤液经阳离子交换树脂 (2000mL)，使用 0.5mol/L 氨水交换洗脱，洗脱液浓缩得棕色糖浆状浸膏，再经多次离子交换柱纯化，即得 1 个新生物碱苷 (3-1) 和 6 个已知的生物碱及生物碱苷 (图 3-1)。

② 以甲醇及甲醇-水为溶剂的提取

a. 黄酮类 取玄参科植物毛泡桐 (*Paulownia tomentosa*) 新鲜果实 320g，用甲醇 (300mL) 浸泡提取，浓缩甲醇提取液，得干燥浸膏 5g。浸膏反复经硅胶色谱柱分离纯化，得 15 个二氢黄酮类化合物，其中 9 个 (3-2~3-10) 为新化合物 (图 3-2)。

b. 苯酚和环烯醚萜及苷 取马钱子科植物腋花马钱 (*Strychnos axillaris* Colebr.) 枝干干燥粉末 2kg，以甲醇为溶剂回流提取 1h，减压浓缩得浸膏。浸膏经萃取、反相制备型中压液相色谱 (MPLC)、制备型高效液相色谱 (HPLC) 和制备薄层色谱 (TLC) 等分离手段，分别分离纯化得到 5 个酚苷 (3-11~3-15) 和 1 个环烯醚萜苷类 (3-16) 化合物 (图 3-3)。

c. 生物碱 取梧桐科植物 *Waltheria douradinha* 的地上部分 5.1kg，干燥粉碎后用甲醇浸泡提取 3 次 (5 L×3，每次 3 天)。合并滤液，减压浓缩，浸膏经色谱柱分离纯化，得到 2 个新生物碱 (3-17, 3-18) (图 3-4)。

③ 以丙酮及丙酮-水为溶剂的提取

a. 萜类 取鼠尾草属植物 *Salvia palaestina* 地上部分 500g，干燥粉碎后依次用环己烷、丙酮在恒温容器内浸泡提取，得丙酮提取浸膏 6.0g，反复经硅胶色谱

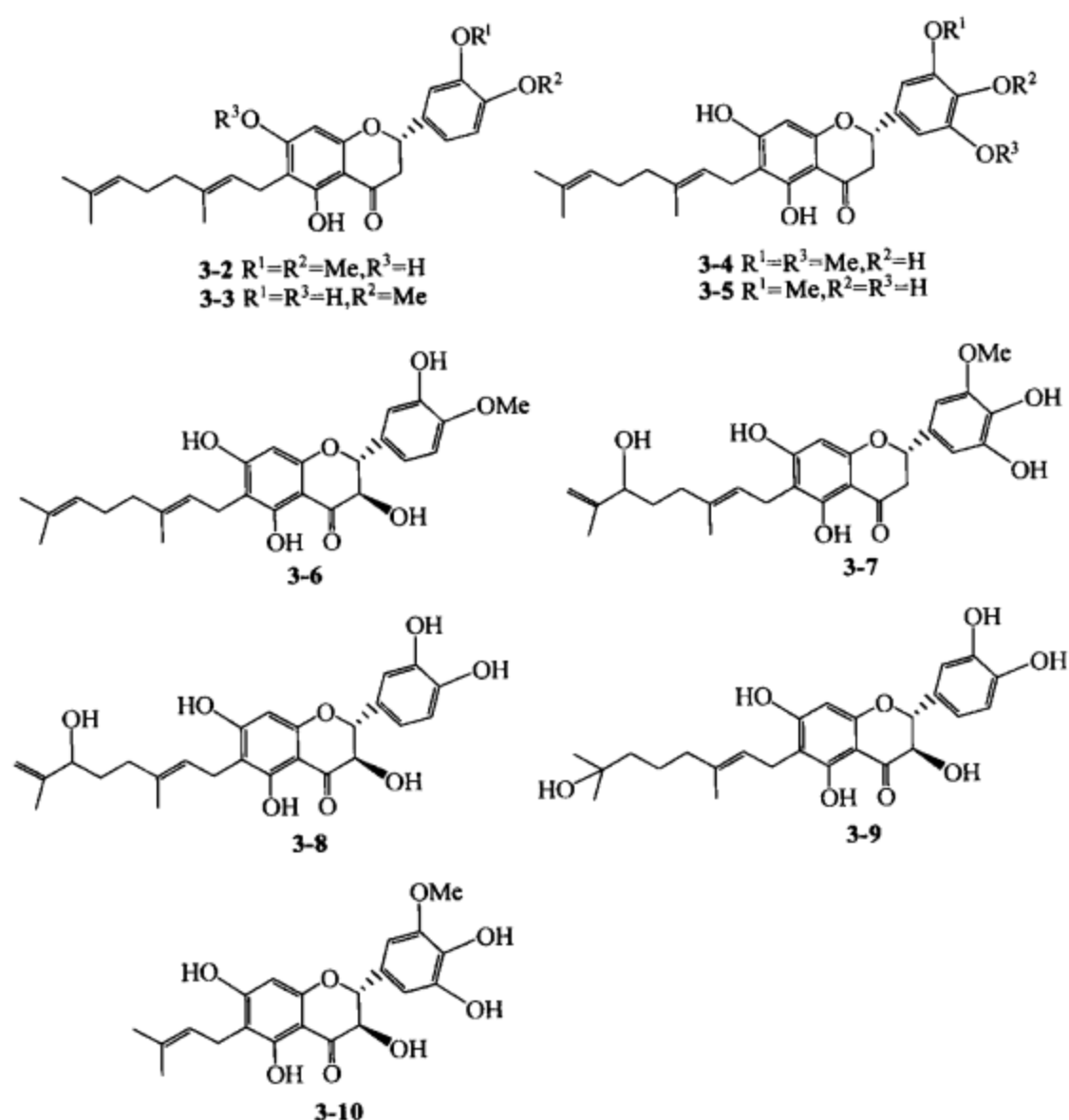


图 3-2 毛泡桐 (*Paulownia tomentosa*) 中分离得到的黄酮类化合物

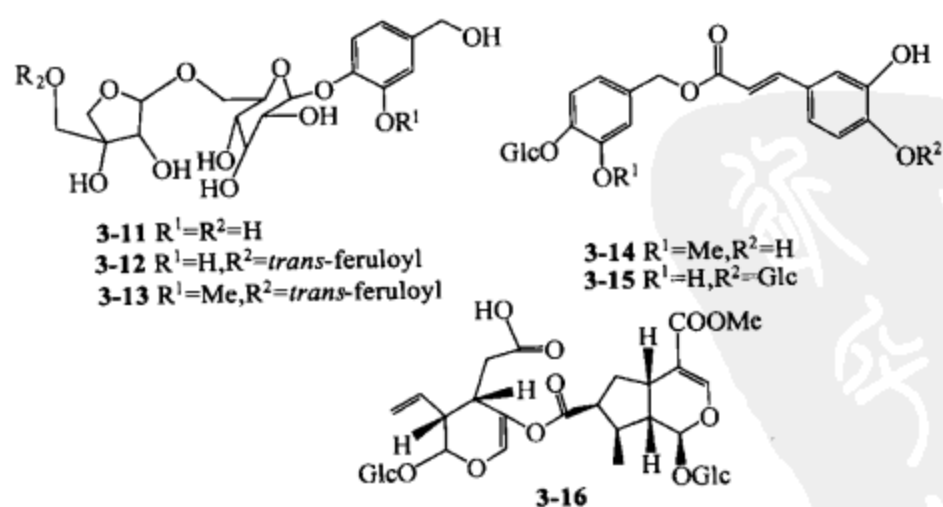


图 3-3 腋花马钱 (*Strychnos axillaris* Colebr.) 中分离得到的酚苷及环烯醚萜苷
 feruloyl—阿魏酰; Glc—葡萄糖

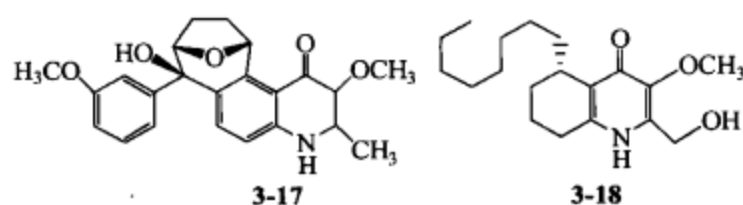


图 3-4 *Waltheria douradinha* 中分离得到的新生物碱

柱分离纯化，得 3 个萜类化合物 (3-19~3-21) (图 3-5)。

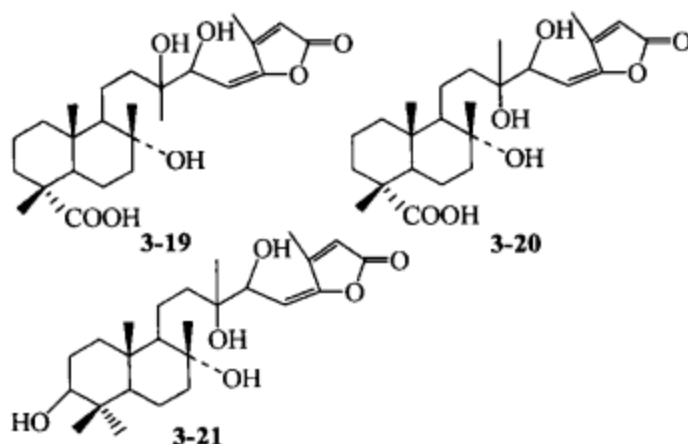


图 3-5 *Salvia palaestina* 中分离得到的萜类化合物

b. 二萜及糖苷 取蕨类植物芒萁 (*Dicranopteris dichotoma*) 的叶 32kg，干

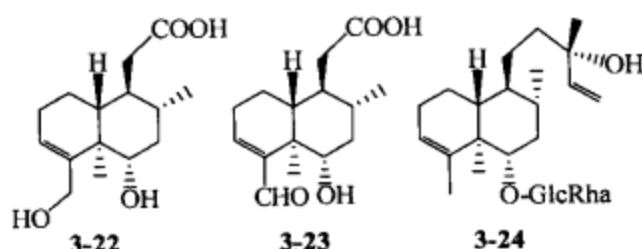


图 3-6 芒萁 (*Dicranopteris dichotoma*) 中分离得到的二萜

燥粉碎，以丙酮溶剂在恒温容器内提取 3 次 (30L×3, 24h)，提取液减压浓缩，得浸膏 2.3kg，反复经硅胶色谱柱纯化分离，得到 3 个二萜及糖苷化合物 (3-22~3-24) (图 3-6)。

(2) 水溶剂提取法 能够采用水提取的成分有如下几种：碱性化合物，如生物碱，它们在植物中往往是与酸性成分成盐而存在，盐一般在水中溶解度较大；酸性化合物，羧基本身是一个极性基团，而且有时与金属成盐而存在，故也能溶于水；苷类化合物，因糖含有较多的羟基，在水中的溶解性较大，故常溶于大量的水中。

常用水提取方法有水煎煮、水浸渍和水渗漉。现举例如下。

(a) 麻黄碱。麻黄切段后，用水浸煮 2 次，浸煮液用氢氧化钠碱化至 pH 值为 11，甲苯萃取，萃取液流经草酸溶液，加盐酸调节 pH 值至 6，析出盐酸麻黄碱结晶。另外，据文献报道，麻黄粗粉用酸性水浸泡或渗漉，渗漉液经强酸型离子交换树脂，可得到盐酸麻黄碱。

(b) 芦丁。荞麦叶粉用沸水提取 3 次，浓缩，冷却后加异丙酮，析出沉淀。滤

出沉淀，加沸水溶解，冷却后析出粗芦丁。

(c) 甘草酸。将甘草加冷水浸泡，所得水提取液加热煮沸，使蛋白质类物质变性后沉淀析出，过滤，滤液加硫酸酸化，即得粗甘草酸。

因水的沸点比较高，水提取液在减压浓缩时会产生大量泡沫，浓缩困难。通常在蒸馏瓶上装一防泡球加以克服。也有将水溶液直接喷雾干燥成粉末的。实验室中可以在水溶液中加入少量丁醇或戊醇克服气泡的产生。另外，水提法还有以下缺点。首先，提取物易发霉。水提取物含有较多的糖分、蛋白质等营养物质，易发霉，要及时处理，也可加入少量甲苯、甲醛或三氯甲烷等作防腐剂。其次，药材易发酵。用水渗漉或浸泡时，有时室温较高，在渗漉或浸泡的过程中药材容易发酵，这时可用三氯甲烷饱和的水溶液进行渗漉或浸泡。再次，易出现人工产物。植物水提取液成分复杂，提取温度较高，各成分之间容易相互作用，出现人工产物。

3.1.1.2 水蒸气提取法

水蒸气提取法适用于具有挥发性、能随着水蒸气蒸馏而不被破坏且难溶或不溶于水的成分提取。此类成分的沸点多在 100°C 以上，并在 100°C 左右有一定的蒸气压。水蒸气提取所依据的原理是基于两种互不相溶的液体共存时各组分的蒸气压和它们在纯粹状态时的蒸气压相等，而另一种液体的存在并不影响它的蒸气压，混合体系的总蒸气压等于两组分蒸气压之和，由于体系中的蒸气压比任何一组分的蒸气压为高，所以混合物的沸点要比任何一组分的沸点低。

水蒸气提取方法是目前从药用植物中提取挥发油常用的方法。

3.1.1.3 分子蒸馏技术

天然产物经溶剂提取得到的提取液，一般体积较大，有效成分浓度低，需要进行浓缩和提纯分离。常采用的浓缩方法有蒸发法、蒸馏法、精馏法和分子蒸馏法。

蒸馏和精馏是以液体混合物中各组分的挥发性差异作为分离依据。在通常的蒸馏中，同时存在由液相到气相和气相到液相的不同方向的蒸气分子流，前者大于后者的量就是蒸馏产物，因此只能实现液体混合物的粗分离，分离效率比较低，如果实现液相到气相的单一分子流流向而阻止返回液相的分子流，就是分子蒸馏。分子蒸馏减少了蒸气回流到液相表面的分子流，因此大大提高了蒸馏的效率，同时由于料液受热时间短，也能降低料液组分的热分解。

国外 20 世纪 30 年代出现分子蒸馏技术，并在 60 年代开始工业化应用。日本、美国、德国都设计制造出了各种式样的分子蒸馏装置，并不断对分子蒸馏设备进行改进和完善。20 世纪 80 年代中期，国内开始分子蒸馏技术的应用开发，并从国外引进了用于生产单甘酯的分子蒸馏装置。目前蒸馏已在油脂化学工业如单甘酯、双甘酯、长链脂肪酸、维生素 E、高碳醇、甾醇等浓缩和提取中得到了广泛的应用。

3.1.2 天然产物化学成分的现代提取方法

3.1.2.1 固相萃取法

固相萃取法 (solid-phase extraction, SPE) 是一种样品预处理技术,也是天然产物有效成分提纯的一种方法。其工作原理与液-固提取相同。常用两种方法:一种是将样品中的干扰杂质吸附于柱上,而所需的化合物被洗脱下来;另一种是将所需的化合物保留在柱上,而干扰性杂质被洗脱下来,通过更换溶剂将所需化合物洗脱下来。后者对化合物可起到浓缩作用。固相提取法很适合于自动化操作,尤其适用于对大量样品进行常规纯化。

3.1.2.2 固相微萃取法

固相微萃取法 (solid-phase microextraction, SPME) 是一项新型的无溶剂化样品前处理技术。固相微萃取法是在固相萃取的基础上发展起来的方法,此技术是在一根纤细的熔融石英纤维表面涂布一层聚合物,并将其作为萃取介质 (萃取头),采用直接浸没固相微萃取法 (direct-SPME, DI-SPME) 将萃取头直接浸入样品溶液,或采用顶空固相微萃取法 (headspace-SPME, HS-SPME) 采样,然后直接进行 GC 或 HPLC 分析。SPME 装置形状似一支微量注射器,由手柄和萃取头两部分构成 (图 3-7)。

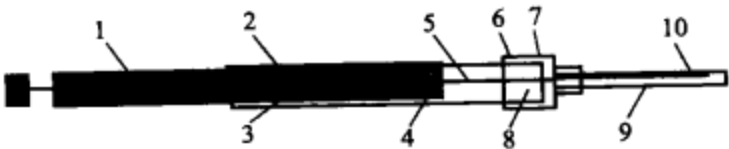


图 3-7 SPME 装置图

1—柱塞杆; 2—针筒; 3—固定螺丝; 4—色码螺旋套管; 5—接有石英纤维的钢针;
6—固定螺帽; 7—金属包头; 8—密封头; 9—针管; 10—石英纤维

固相微萃取法分为萃取过程和解吸过程两步 (图 3-8)。

李彩芳等采用固相微萃取法处理小叶女贞花及花蕾,分别称取粉碎后的小叶女贞花及花蕾各 1.0g,置于 15mL 专用采样瓶中,插入装有 65 μ m PDMS/DVB 纤维头的手动进样器,90 $^{\circ}$ C 下顶空萃取,保持 30min 取出,立即插入色谱仪进样口 (温度 250 $^{\circ}$ C) 中,脱附 1.5min,再进一步进行 GC-MS 分析。Kang 等采用同样的萃取方法分析国槐 (*Sophora japonica* L.) 的花和花蕾中的化学成分,发现了一个含量较高的物质 8-十七碳烯 (49.48%),而该物质在其他提取法中未能得到。

3.1.2.3 超临界流体萃取法

超临界流体萃取法是利用溶剂在超临界条件下流体的特殊性能对样品进行提取的一种提取方法。超临界流体具有较好的扩散性、较低的黏度、较高的扩散速度,提取不需要使用大量的有机溶剂,更加适合于对植物成分的提取。影响溶剂溶出能力的因素有温度、压力及改性溶剂。选择适当的温度与压力能提高提取溶剂的选择

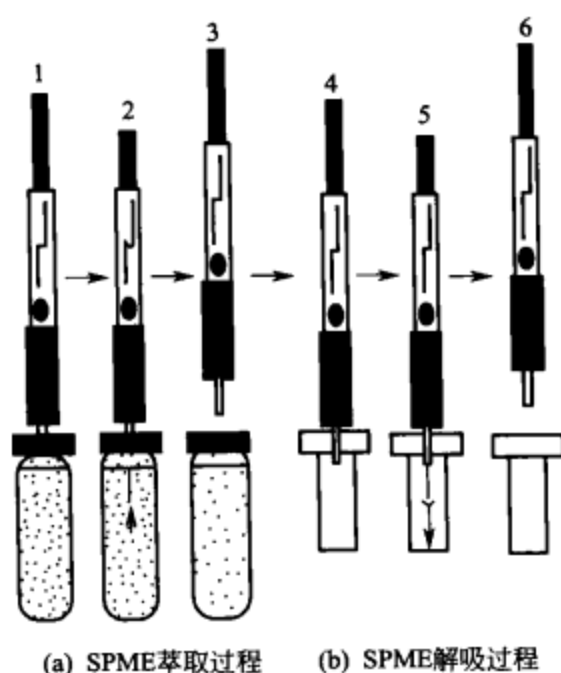


图 3-8 固相微萃取步骤

- 1—刺穿样品瓶盖；2—暴露出纤维、萃取；3—缩回纤维、拔出萃取器；
4—插入 GC 气化室；5—暴露出纤维、解吸；6—缩回纤维，拔出萃取器

性能，同时可以除去更多杂质，得到较为纯净的提取物。常用的超临界流体为二氧化碳，具有较好的安全性。超临界流体萃取法尤其适用于对热及化学不稳定的物质的提取，以及从混合物中提取低极性的组分（图 3-9）。

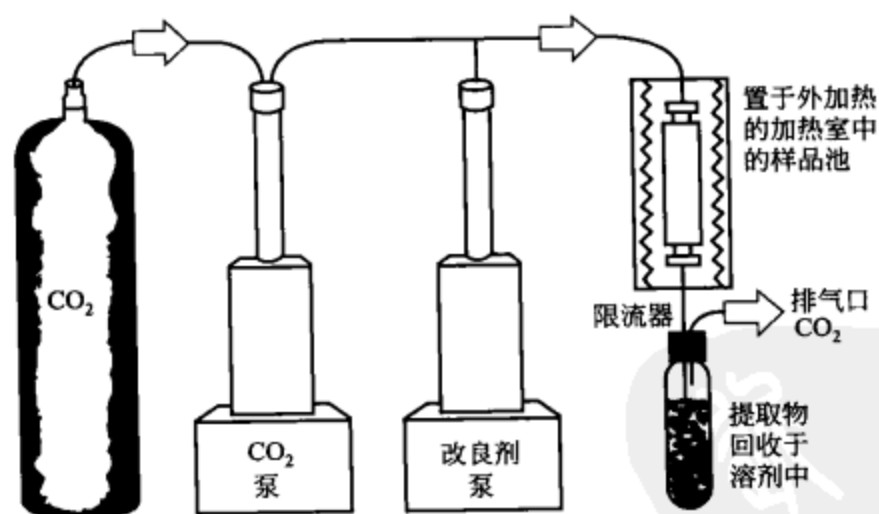


图 3-9 超临界流体萃取法装置图

3.2 天然产物的分离

药用天然产物经提取、浓缩得到的浸膏仍为混合物，含有较多的杂质，需要进一步分离纯化，除去杂质。常用分离方法较多，根据各类型化合物的性质选择合适

的分离方法，是保证分离成功的前提。常用的分离方法有系统溶剂分离法、两相溶剂萃取法和各种柱色谱法。

3.2.1 天然产物的经典分离方法

所谓的经典分离方法是指近年来使用较为广泛，操作较为简单，且无需使用复杂、昂贵的仪器的传统方法。

3.2.1.1 溶剂分离方法

溶剂分离方法就是根据目标物与所选择的溶剂相似相溶的原理，即极性目标物易溶于极性溶剂，非极性目标物易溶于非极性溶剂，结构相似或官能团相同的目标物互相溶解。自然界中天然产物的结构千差万别，分子结构中极性基团的极性大小、多少及取代位置等因素均影响着该物质的极性，进一步决定分离时溶剂的选择。常见天然产物的极性及适用的溶剂见表 3-2。

表 3-2 天然产物的极性及适用的萃取溶剂

天然产物的类型	天然产物的极性	适用的萃取溶剂
挥发油、脂肪油、甾醇、某些苷元、蜡	极性小(强亲脂性)	石油醚、正己烷
生物碱、醛、酮、醌、有机酸、树脂、苷元	极性较小(亲脂性)	乙醚、三氯甲烷
某些苷类(如强心甘等)	极性小	三氯甲烷-乙醇(2:1)
某些苷类(如黄酮苷等)	中等极性	乙酸乙酯
某些苷类(皂苷、蒽醌等)	极性大	正丁醇
糖类、氨基酸、某些生物碱、极性较大的苷	极性较大(亲水性)	丙酮、乙醇、甲醇
蛋白质、糖类、无机盐	极性大(强亲水性)	水

在对药用植物提取物进行分离的过程中，可以依据分离目标产物类型和性质，采取不同的处理方式。如依据目标产物在不同溶剂中的溶解度差异进行的结晶和重结晶、在一种溶剂中加入另一种溶剂改变目标产物溶解度进行沉淀和酸碱性不同目标产物的梯度分离等。

甘草酸是甘草中的主要活性成分，也称为甘草皂苷、甘草甜素，是甘草的甜味成分，在植物中常以钾盐、钙盐形式存在。在提取分离纯化甘草酸时，就是依据甘草酸在不同溶剂中的溶解性差异进行分离纯化。一般先采用氨-乙醇-水（0.5：10：89.5）超声提取 3 次，浓缩，将浸膏溶解于水中，以乙酸乙酯萃取 3 次，水相以 70%硫酸调节 pH 值至 2，离心，沉淀即甘草酸。本法就是利用甘草酸在水中和乙酸乙酯中的溶解性差异分离纯化。

而在分离纯化生物碱类成分时，依据生物碱在不同 pH 值条件下的溶解度不同，调节水相的酸碱性，再利用有机溶剂进行萃取，可使生物碱类成分得到富集，并可使强碱性生物碱与弱碱性生物碱得到初步分离（图 3-10）。

在分离纯化的过程中，采用溶剂分离法应当注意，有机溶剂可能会与天然产物发生化学反应，改变目标产物。如使用乙酸乙酯提取分离时，可能发生乙酰基转

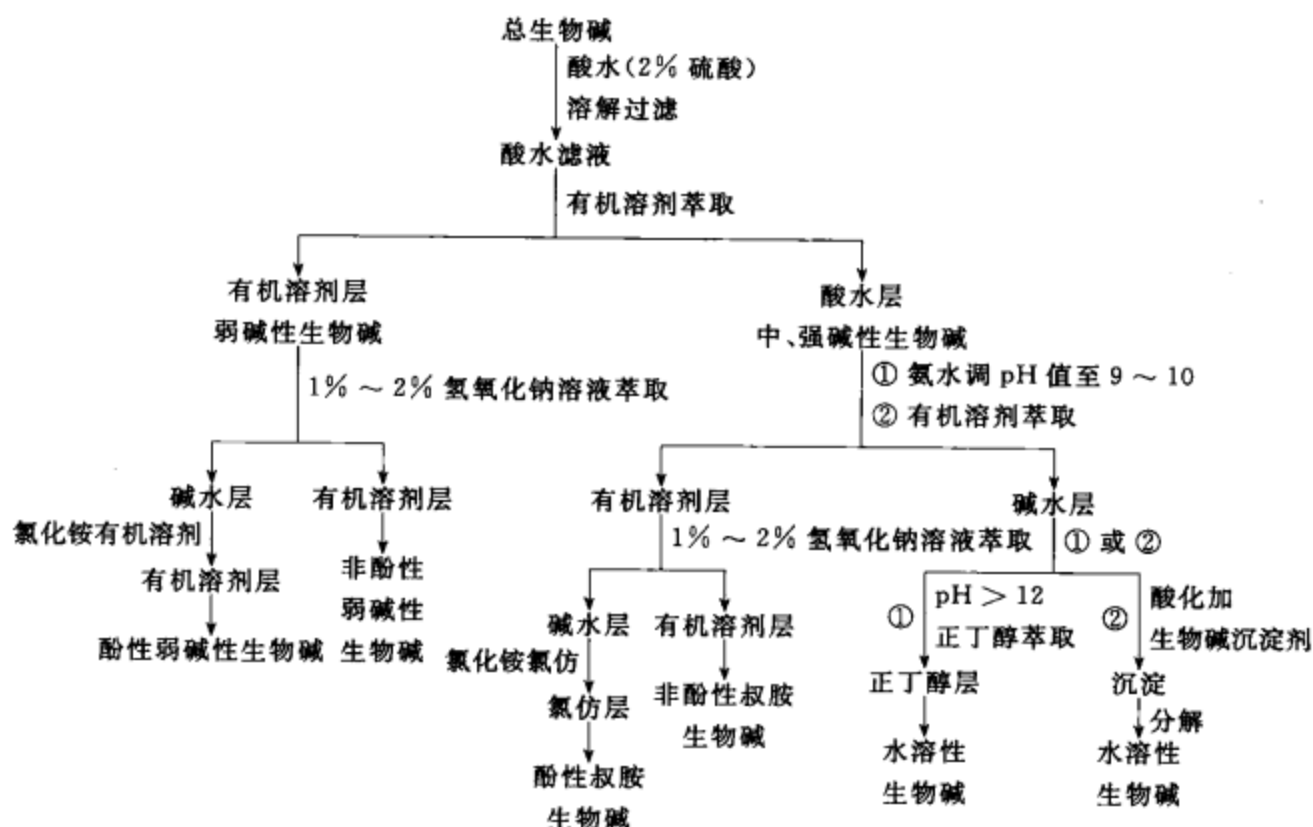


图 3-10 依据 pH 值对生物碱分离的流程

移；使用丙酮提取分离时，可能会与天然产物中的二醇基团形成缩酮结构；使用甲醇、乙醇或者正丁醇，有时会与天然产物中的羧基形成相应的酯。

3.2.1.2 色谱分离方法

1906 年，俄国植物学家茨维特 (Tsvet, 1872~1919) 首创了这种色谱分离方法，他把植物色素的石油醚浸渍液倒入一根装有碳酸钙吸附剂的细玻璃柱中，再加入纯石油醚，任其自由流下，原混合物被分离成不同颜色的谱带，且以不同的速度通过柱子，然后按谱带颜色对混合物进行鉴定分析。自此，色谱法 (chromatography) 被广泛用于天然产物的分离，成为分离多组分有机混合物的一种高效分离技术。

① 基本原理 色谱法是利用不同的成分在固定相与流动相中不同的分配系数进行分离的一种方法。在色谱分离中，试样混合物被载入固定相的支持体中，当流动相通过固定相时，试样各组分在固定相和流动相之间达到动态平衡，使各组分通过固定相的速率不同而得到分离。

② 色谱法的分类 色谱法按流动相的性质可分为：a. 有机溶剂系色谱，包括硅胶或氧化铝柱色谱和凝胶柱色谱；b. 水系色谱，包括大孔树脂柱色谱、聚酰胺柱色谱、离子交换树脂柱色谱、反相柱色谱等。

按分离原理大体可分为：a. 吸附色谱 (adsorption chromatography, AC)，利用吸附能力的差别进行分离，常用的吸附剂有氧化铝、硅胶、聚酰胺等；b. 分配色谱 (partition chromatography, PC)，利用在两种互不相溶的溶剂中分配比不同

进行分离，常用的支持剂为硅胶、硅藻土、纤维粉等；c. 离子交换色谱 (ion exchange chromatography)，利用离子解离强度不同进行分离，常用的离子交换树脂有强酸型（磺酸型）、强碱性（季铵型）、弱酸型（羧酸型）、弱碱性（三级胺型）；d. 排阻色谱，利用分子大小不同而阻滞作用不同进行分离。

按支持体的不同，可分为柱色谱、薄层色谱及纸色谱。按固定相和流动相的相对极性大小不同，可分为正相色谱和反相色谱。此外还有电泳亲和色谱。

研究体系中最常规的方法是采用有机溶剂系色谱法进行天然产物的分离与纯化。

(1) 有机溶剂系色谱法 有机溶剂系色谱是指流动相为有机溶剂或以有机溶剂为主的柱色谱，常用的有硅胶柱色谱、氧化铝柱色谱和凝胶柱色谱等。

① 硅胶柱色谱 以硅胶为填充材料的柱色谱统称为硅胶柱色谱。色谱硅胶为多孔性物质，可用通式 $\text{SiO}_2 \cdot n\text{H}_2\text{O}$ 表示，具有多孔的硅胶环及 $-\text{Si}-\text{O}-\text{Si}-$ 的交链结构，其表面的硅羟基 (silanol) 能够通过氢键与许多分子相互作用，同时能吸附多量的水分。硅胶的吸附性能与硅胶中硅羟基的数目及含水量有关，硅胶吸附的水分愈多，吸附其他化合物的能力愈弱，吸附活性愈低。若吸水量超过 17%，吸附能力极弱，不能用作吸附色谱，只可用作分配色谱的载体。硅胶的表面积、表面结构、微孔体积及微孔半径均直接影响色谱分离的效果。

掺入某些复合试剂，可改变硅胶的吸附能力，提高分离效果。通常用硝酸银处理的硅胶，对不饱和烃类有极好的分离作用。改良吸附剂制备一般用 1%~10% 复合试剂的水溶液或丙酮溶液，与硅胶混匀，待稍干后于 110℃ 干燥即可。硅胶柱色谱适用范围较广，能用于极性化合物及非极性化合物的分离，如挥发油、萜类、甾体、生物碱、强心苷、蒽醌、酸性或酚性化合物、磷脂类、脂肪类、氨基酸等物质的分离。

a. 硅胶柱色谱的基本原理 硅胶柱色谱具有吸附色谱及分配色谱的双重性质。吸附作用常指物理吸附，也叫表面吸附，最大特点是无选择性，吸附与解吸附过程可逆，且可快速进行，在实际工作中使用得较为广泛。此外还有化学吸附和半化学吸附。吸附作用的基本规律是相似者易于吸附；对极性物质有较强的吸附能力，极性强的物质优先被吸附；溶剂的极性愈弱，则吸附剂对溶质的吸附能力愈强，若溶剂的极性增强，则吸附剂对溶质的吸附能力减小；极性强弱是决定吸附过程的主要因素。化合物的极性则由分子中所含官能团的种类、数目及排列方式等综合因素决定的。常见官能团的极性见表 3-3。

分配色谱的分离原理是利用混合物的不同成分在固定相和流动相中分配系数的不同而达到分离。在洗脱的过程中，易溶于流动相的成分移动较快，易溶于固定相的成分移动较慢。

b. 硅胶色谱柱流动相的选择 硅胶吸附色谱一般是根据待分离物质的极性采

表 3-3 常见官能团的极性比较

R-COOH	大	$+NH_3-CHR-COO^-$	氨基酸
ArOH			
H ₂ O			
R-OH			
R-NH ₂ , R-NH-R ¹ , R ¹ -N(R ²)-R ³			
R-CO-NR ¹ R ²			葡萄糖
R-CHO			
R-CO-R ¹			
R-CO-OR ¹			
R-O-R ¹			硬脂酸
R-X			
R-H	小		

用相应的极性系统溶剂来洗脱。通常采用从低极性逐步递增压性的梯度洗脱方式。往往借助于硅胶薄层色谱的结果来摸索分离条件，在具体操作时吸附柱色谱所用的溶剂比薄层色谱展开剂极性略偏小。分配柱色谱可先用纸色谱方法探索分离条件。

c. 硅胶色谱柱的操作

(a) 装柱 一般采用湿法装柱，即将硅胶与装柱溶剂搅拌混匀，排尽气泡，连同溶剂一起倒入柱中，最好一次倾入，以防硅胶出现分段现象，影响分离效果。色谱柱中硅胶段直径与长度一般为 1:(20~30)。有时也可以用干法装柱，将所用硅胶一次倾倒入柱中，墩紧至硅胶高度不再下降。

(b) 加样 如样品能溶解于洗脱剂，则以尽可能少的洗脱剂溶解，完全溶解后从柱顶加入，使样品色带较狭窄，再进行洗脱。如样品难溶于洗脱剂，则使用适当的尽可能少的溶剂溶解，再与硅胶搅拌，待溶剂挥发尽后再上柱。并在上面覆盖一层纯净的砂或棉花。

(c) 洗脱 一般采用梯度洗脱，使吸附在硅胶上的不同化合物逐个洗脱，达到分离的目的。通常用混合洗脱剂洗脱，并通过调节比例以改变极性，达到梯度洗脱的目的。一般是逐渐增大较大极性溶剂的比例，直至全部为较大极性溶剂。实际操作中，往往根据薄层色谱探索柱色谱的条件。

d. 硅胶柱在天然产物成分分离中的应用

(a) 鱼藤酮类物质的提取纯化实例 鱼藤酮是一种广谱杀虫剂，对害虫高效且不易产生抗药性。在毒理学上是一种专属性很强的物质，对昆虫尤其是对菜粉蝶幼虫、小菜蛾和蚜虫具有强烈的触杀和胃毒两种作用。但是该类物质不溶于水，可溶于多种有机溶剂。易受日光、空气、高温的影响而分解，遇碱也会很快分解失效。

鱼藤酮物质多分布于豆科多年生藤本植物，它们结构极为相似，一般方法难以分离。用硅胶吸附色谱配合凝胶柱法，从豆科植物 *Millettia duchesnei* 中成功分离出 8 种鱼藤酮类成分 (3-25~3-32) (图 3-11、图 3-12)。

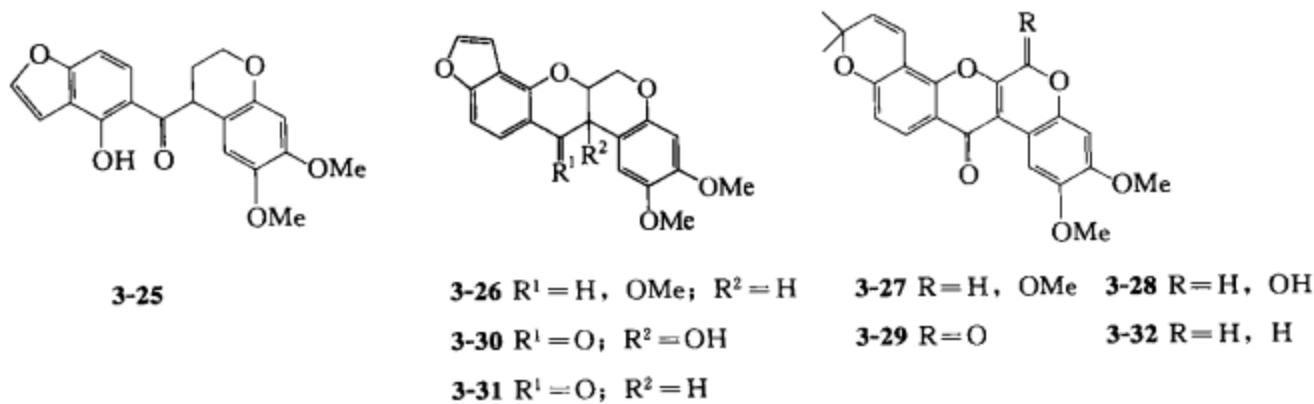


图 3-11 豆科植物 *Millettia duchesnei* 中分离的鱼藤酮类成分结构

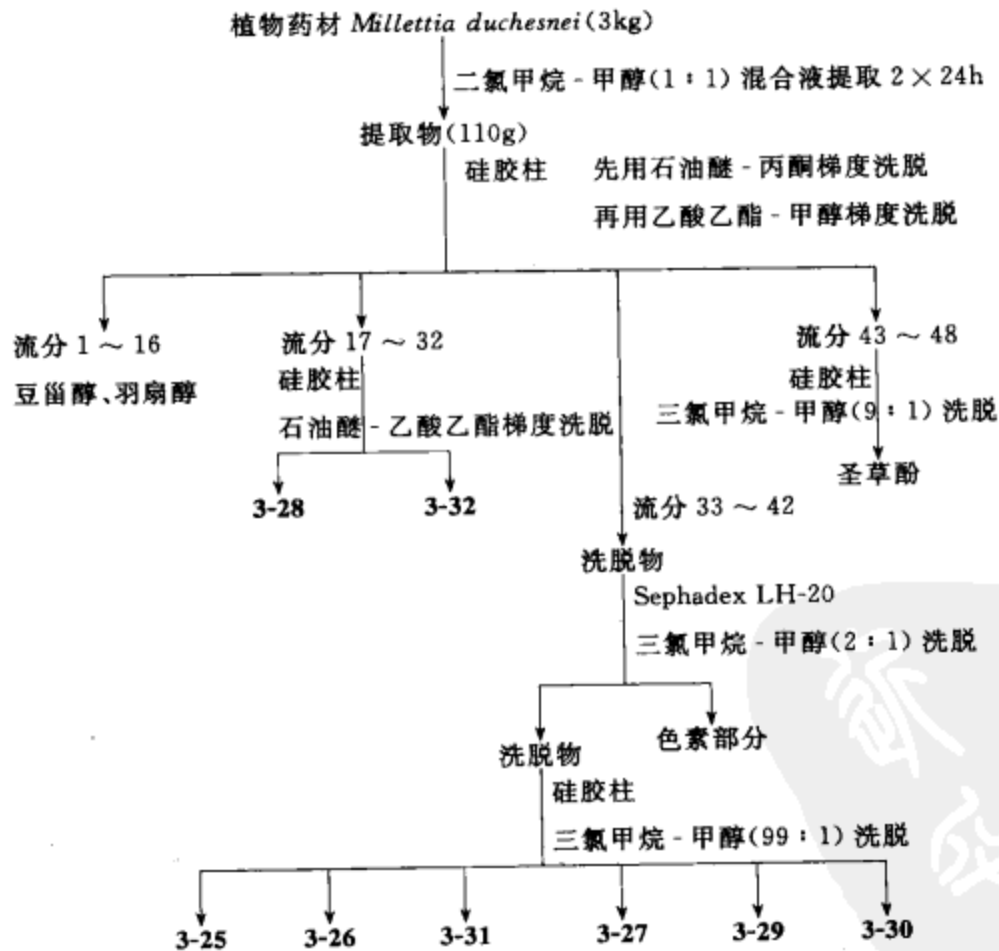


图 3-12 豆科植物 *Millettia duchesnei* 中鱼藤酮类成分的分离流程

(b) 柠檬苦素类化合物的提取纯化实例 楝科和茜草科中部分植物含有柠檬苦素类化合物，该类物质具有细胞毒素活性，是抗疟原虫的重要药物，在尼日利亚是治疗发热和疟疾的主要药物。

柠檬苦素类似物 (limonoid) 是以柠檬苦素的化学结构为基本单位，包括有苦

味的糖苷配基和无苦味的柠檬苦素 17 β -D-糖苷配体 (limonoids-17 β -D-glucopyranosides) 在内的一系列化合物, 属于含呋喃环的高度氧化的四环三萜类化合物。该类物质氧化度较高, 在提取分离过程中易发生结构的改变, 不易纯化。使用硅胶柱色谱, 从楝科植物大叶卡雅楝 (*Khaya grandifoliola*) 中分离出 6 个柠檬苦素衍生物 (3-33~3-38) (图 3-13、图 3-14)。

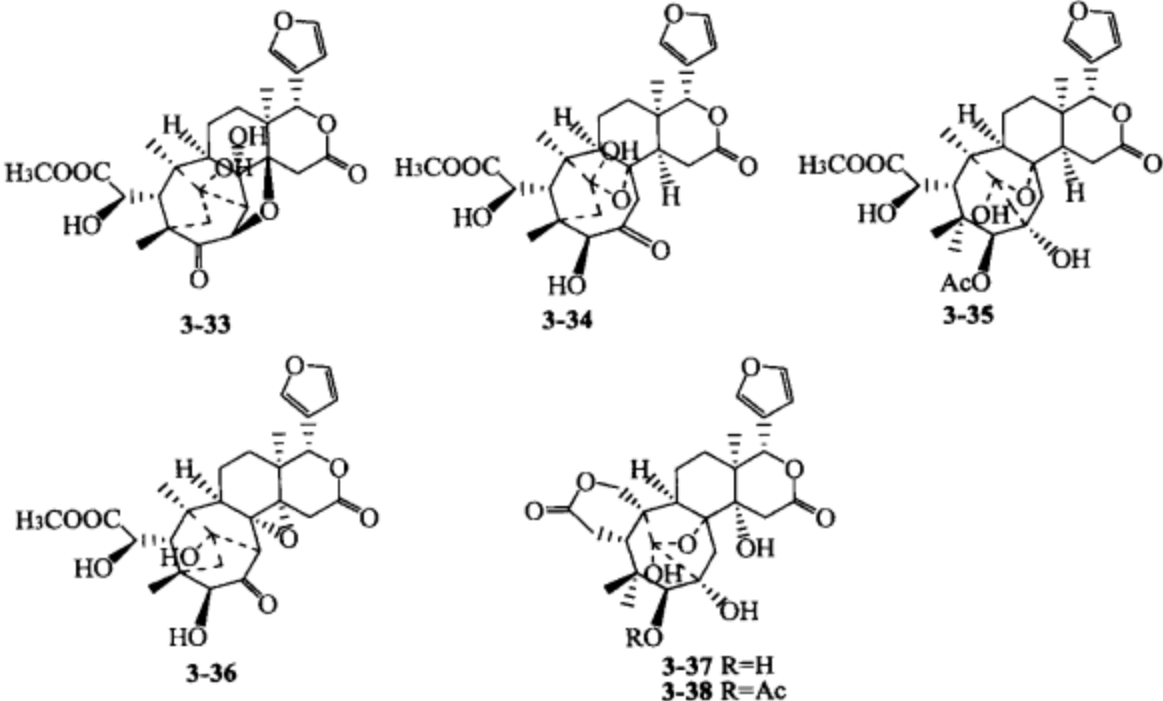


图 3-13 楝科植物大叶卡雅楝 (*Khaya grandifoliola*) 中分离的柠檬苦素衍生物结构

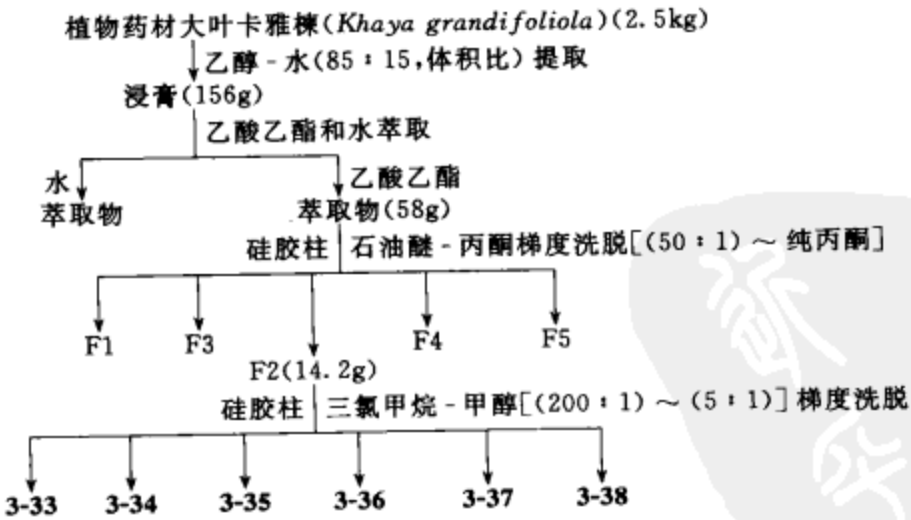


图 3-14 楝科植物大叶卡雅楝 (*Khaya grandifoliola*) 柠檬苦素衍生物分离流程

② 氧化铝柱色谱 氧化铝柱色谱分离法是以氧化铝为填料的层析分离法。填料氧化铝是由氢氧化铝直接在高温下脱水制得。由于制造关系使填料氧化铝带有微弱的碱性, 故对于分离植物中的碱性成分 (如生物碱) 较为理想, 但不宜用于醛、酮、酯和内酯等类型化合物的分离。

在氧化铝柱色谱中, 极性溶剂的洗脱能力较非极性溶剂的洗脱能力大, 故一般

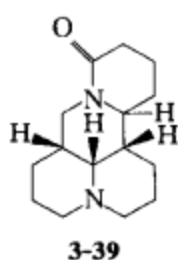


图 3-15 山豆根中分离的
苦参碱结构

采用混合溶剂进行梯度洗脱。氧化铝柱色谱分离法的原理、装柱、上样、洗脱和硅胶柱色谱法相同。

以山豆根粗粉中的苦参碱提取纯化为例。山豆根为豆科越南槐 (*Sophora tonkinensis* Gapnep) 的根, 具有清热解毒、消肿利咽的功效, 其主要成分为生物碱类化合物, 其中山豆根碱的含量最多。利用氧化铝柱色谱法, 可从该植物中分离出苦参碱 (3-39) (图 3-15、图 3-16)。

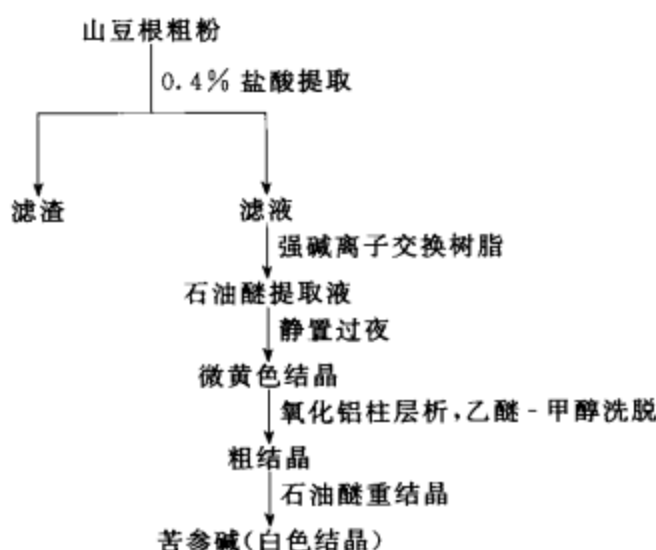
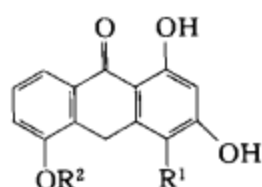


图 3-16 山豆根中苦参碱分离流程

③ 凝胶柱色谱 凝胶柱色谱法指固定相是凝胶的一种分离分析技术。凝胶是具有许多孔隙的网状结构的固体, 有分子筛的性质。凝胶色谱柱是利用分子排阻效应来发挥分离效能。当混合物通过凝胶相时, 比孔隙小的分子可以自由进入凝胶内部, 而比孔隙大的分子就不能进入, 因此移动速度就产生了差异。大分子不被迟滞而随溶液走在前面, 小分子由于向孔隙内扩散或移动而滞留, 落后于大分子而得到分离。凝胶柱色谱法的显著优点是分离效果较高、样品损失较少, 且可以过滤除去其他杂质。

a. 黄酮类成分凝胶柱色谱分离纯化实例 Di H 等采用凝胶柱色谱分离法成功地从藤黄科植物元宝草 (*Hypericum sampsonii*) 中分离纯化出两个新的吡啶磺酸盐 (3-40, 3-41) (图 3-17、图 3-18)。

b. 生物碱类成分凝胶柱色谱分离纯化实例 近年来, 由于交让木属中分离得到的生物碱结构上具有独特的复杂多环体系, 已成为全合成、生物合成及天然产物化学研究的焦点。目前已发现了 60 余个该类型的化合物, Mu 等采用硅胶柱色谱和凝胶柱色谱法顺利地交让木属植物虎皮楠 (*Daphniphyllum oldhami*) 中分离纯化得到 3 个该类型生物碱 (化合物 3-42~3-44) (图 3-19、图 3-20)。



3-40 $R^1 = \text{SO}_3\text{K}$, $R^2 = \text{CH}_3$

3-41 $R^1 = \text{SO}_3\text{K}$, $R^2 = \beta\text{-D-gluc}$

图 3-17 藤黄科植物元宝草 (*Hypericum sampsonii*) 中分离的吡啶磺酸盐结构

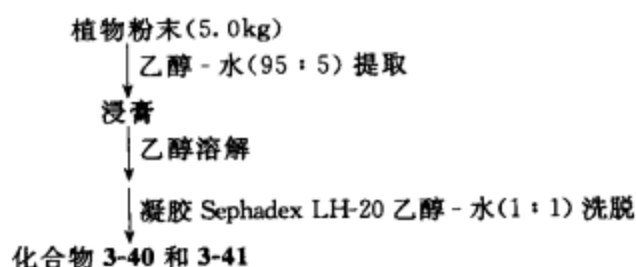
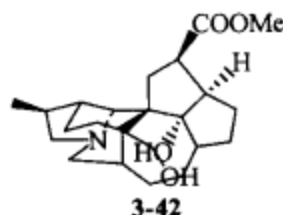
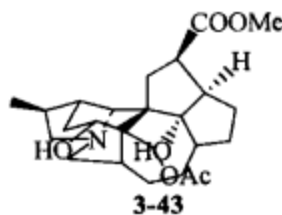


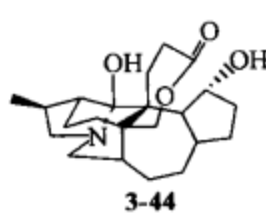
图 3-18 藤黄科植物元宝草 (*Hypericum sampsonii*) 中吡啶磺酸盐分离流程



3-42



3-43



3-44

图 3-19 交让木属植物虎皮楠 (*Daphniphyllum oldhami*) 中分离的生物碱结构

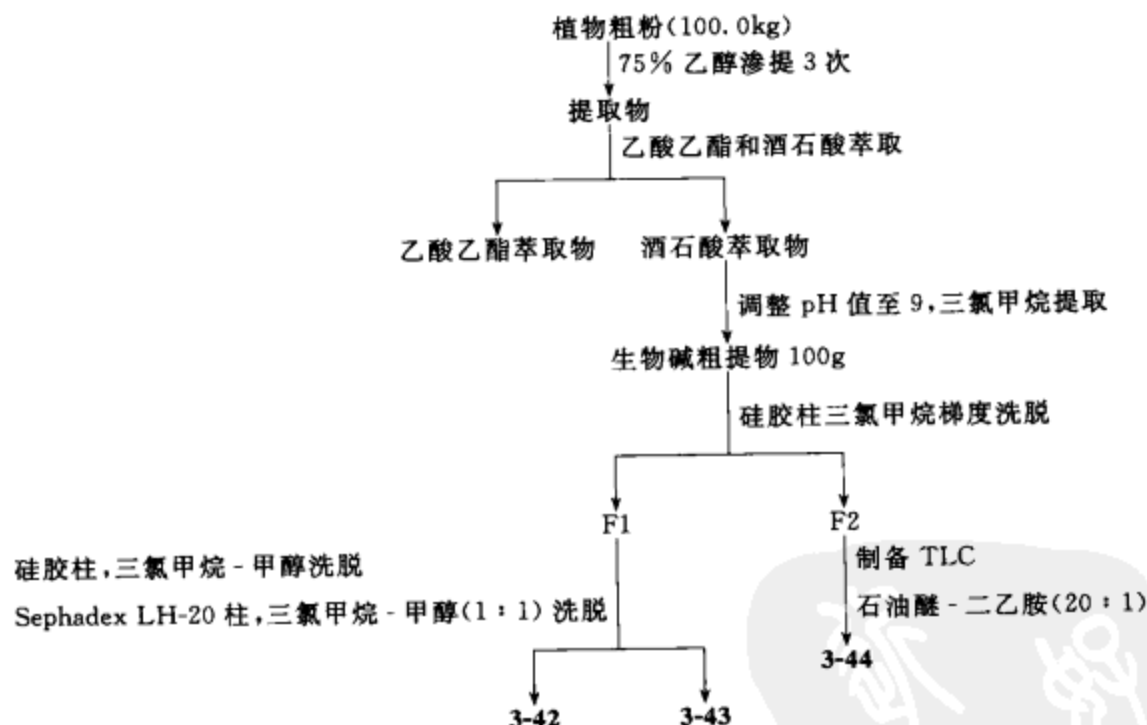


图 3-20 交让木属植物虎皮楠 (*Daphniphyllum oldhami*) 生物碱分离流程

(2) 水系色谱法 水系色谱法指以水或醇水溶液为流动相或洗脱剂的柱色谱法。常见的水系色谱法有大孔树脂柱色谱、聚酰胺柱色谱、离子交换树脂柱色谱和反相色谱等。

① 大孔树脂柱色谱 大孔吸附树脂 (macro-reticular resin) 是吸附与分子筛原理相结合的分离材料, 是不含有交换基团、具有大孔结构的高分子吸附剂, 是一种亲脂性物质。大孔树脂在水溶液中吸附力较强且具有良好的吸附选择性, 适用于在水溶液中分离低极性或非极性的化合物, 组分间极性差别越大, 分离效果越好。

混合物组分在大孔吸附树脂吸附后，一般依次用水、含水乙醇从 10%（体积分数）乙醇、20%（体积分数）乙醇直至高浓度乙醇或丙酮洗脱。

大孔树脂使用前，一般需要用甲醇、乙醇或丙酮连续洗涤数次，再用蒸馏水洗至无醇味即可。再生大孔树脂时，用甲醇或乙醇浸泡洗涤，必要时可用 1mol/L 盐酸或氢氧化钠溶液浸泡，然后用蒸馏水洗至中性，再浸泡在甲醇或乙醇中备用。一般以乙醇湿法装柱；样品液拌树脂上样；先用水洗，再用浓度递增的乙醇溶液进行洗涤。

李茂星等应用 XDA-1 型大孔吸附树脂对独一味水提取物中环烯醚萜苷类成分进行富集，以 50%（体积分数）乙醇溶液解吸附。结果富集后总环烯醚萜苷的含量达到 67.53%。

袁怀波等采用大孔树脂分离纯化木瓜总三萜酸，以 70%（体积分数）乙醇为洗脱剂，上样液 pH 值为 6 时，采用 AB-8 大孔吸附树脂对木瓜三萜酸进行纯化效果最优，产品的洗脱率达 89.33%，纯度达 36.51%。

陈乃富采用大孔树脂柱层析法纯化金樱子总黄酮，用 70%（体积分数）乙醇提取金樱子中黄酮类化合物，得到粗黄酮，利用分光光度法测定样品中的黄酮含量，以此分析对 AB-8 型大孔树脂动态吸附与解吸金樱子黄酮的各种因素的影响。结果最佳工艺为样品溶液浓度 0.3mg/mL，进样量为 120mg，以 1mL/min 的吸附速率进行吸附，5 BV 的 70%（体积分数）乙醇、1mL/min 的流速进行洗脱效果最佳。粗黄酮的黄酮含量为 25.2%，经 AB-8 型大孔吸附树脂纯化后的金樱子精黄酮中黄酮的含量达 42.2%。

② 聚酰胺柱色谱 聚酰胺（polyamide）是通过酰胺键聚合而成的一类高分子化合物，分子中富含酰胺基。酚类、酸类、醌类和硝基类化合物，能和酰胺基以氢键的形式结合被吸附，从而与不能形成氢键的化合物分离。聚酰胺既有亲水性又有亲脂性，性能较好，既可分离水溶性成分，又可分离脂溶性成分。聚酰胺吸附能力的强弱决定于各种化合物与之形成氢键缔合能力的强弱。通常认为化合物分子中酚羟基数目越多，吸附能力越强；化合物分子中芳香化程度越高，吸附能力越强；能形成分子内氢键的化合物，在聚酰胺上的吸附能力减小。

值得重视的是，聚酰胺色谱柱具有“双重色谱”性能，即指其流动相既可是含水系（如甲醇-水），又可是有机溶剂系（如三氯甲烷-甲醇）。因流动相的“两重性”，洗脱顺序也有两种可能。从聚酰胺柱上洗脱被吸附的化合物是一种溶剂分子取代酚性化合物的过程，即以新的氢键代替原有氢键的脱吸附完成的。如黄酮苷元与苷的分离过程中，若以水系为洗脱剂（如醇水），黄酮苷比苷元先洗脱下来；若以有机溶剂系为洗脱剂（如三氯甲烷-甲醇），则洗脱顺序恰恰相反，即黄酮苷元比黄酮苷先洗脱下来。

聚酰胺在天然产物有效成分的分离上应用较为广泛，既可分离纯化极性化合物又可分离纯化非极性化合物，常用于酚类、醌类、黄酮类等化合物的分离，也可用于萜类、甾体、生物碱及糖类等。

五味子抗 HCV 活性成分聚酰胺柱色谱分离实例。五味子具有广泛的药理活性，如收敛、抑菌、抗病毒、抑突变、癌化学预防、抗氧化等，临床上历来作为敛肺、涩肺药广泛用于呼吸系统疾病。段德良等使用聚酰胺柱色谱结合凝胶柱层析法，从五味子中分离得到 3 个抗 HCV 的活性成分 (3-44~3-46) (图 3-21、图 3-22)。

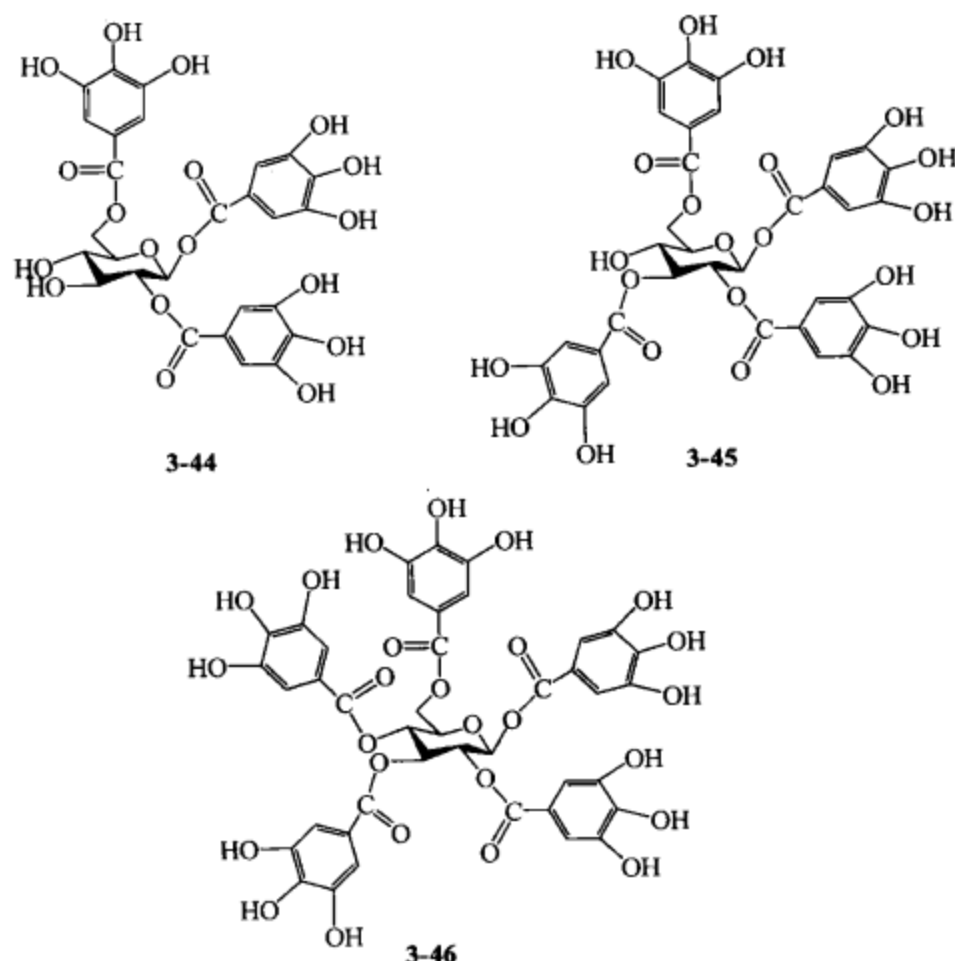


图 3-21 五味子中分离得到抗 HCV 的活性成分结构

③ 离子交换树脂柱色谱 离子交换柱色谱 (ion exchange chromatography, IEC) 是以离子交换树脂或表面涂有液体离子交换树脂的固体颗粒为固定相，装在色谱柱中，以水或含水溶剂为流动相。当样品溶液流经交换柱时，溶液中的中性分子及不与离子交换树脂发生交换的化合物将通过柱子流出，而具有可交换离子的化合物则与树脂上的交换基团进行离子交换并被吸附到柱上。当两种以上的成分被吸附在离子交换剂上，用洗脱液洗脱时，它们的被洗脱能力取决于各物质

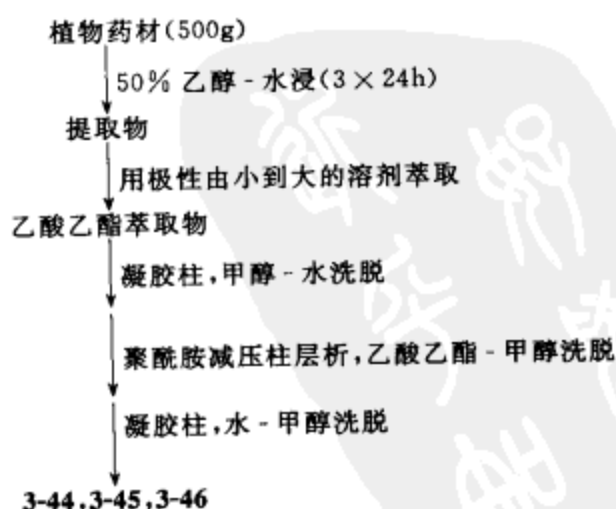


图 3-22 五味子中抗 HCV 的活性成分分离流程

洗脱反应的平衡常数，选择合适的洗脱剂即可实现不同组分物质的分离。

根据交换离子性质不同，离子交换树脂分为阳离子交换树脂和阴离子交换树脂，每种树脂根据它的解离性能大小又可分为强型、中型、弱型（图 3-23）。交换树脂由两部分构成，即交换基团和母核部分：母核为由苯乙烯通过二乙烯苯交联而成的大分子网状结构；常见的交换基团有 $-\text{SO}_3\text{H}$ 、 $-\text{COOH}$ 、 $-\text{NH}_2$ 、 $-\text{NH}$ 、 $-\text{N}^+(\text{CH}_3)_3\text{OH}^-$ 等基团。

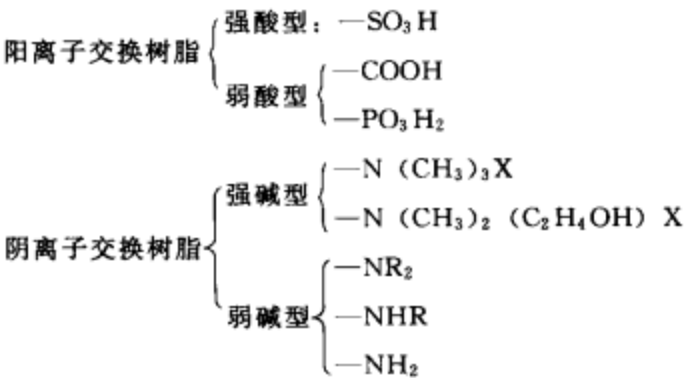


图 3-23 离子交换树脂的分类

根据离子交换树脂的分离原理、交换树脂的性质及分类可知，离子交换柱色谱法可用于天然药物有效部位的分离，使物质按酸碱性分成酸性、中性、碱性部位（图 3-24）。这种有效部位的初步分离方法有利于活性跟踪、临床试验或动物试验。

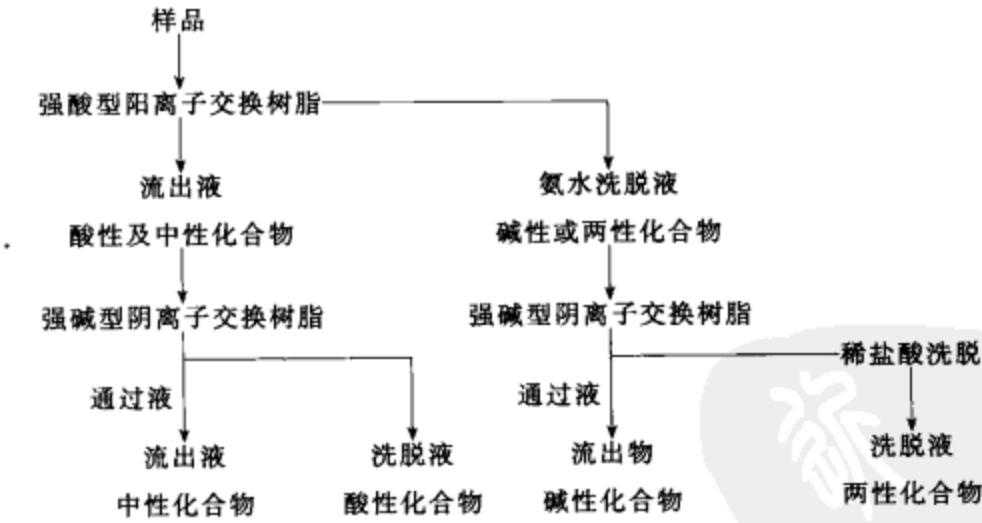


图 3-24 离子交换法分离天然产物有效部位流程

④ 反相色谱 反相色谱（reversed phase chromatography）是基于物质分子的疏水基团与色谱介质上的疏水基团相互作用，在非极性的固定相上，用极性较大的液体作为流动相，进行物质分离分析的一种色谱技术。在正相色谱和反相色谱中，极性物质的洗脱顺序恰恰相反：在正相色谱中，溶质极性越强，洗脱出越晚；在反相色谱中，溶质极性越强，洗脱出越早。反相色谱与正相色谱的比较见表 3-4。

表 3-4 反相色谱与正相色谱的比较

项 目	正相色谱	反相色谱
固定相极性	大	小
流动相极性	小、中	中、大
组分流出顺序	极性小的先流出	极性大的先流出
流动相极性增加	保留时间变小	保留时间变大
主要分离对象	极性化合物	极性较大的化合物

反相色谱分离介质骨架材料多为硅胶或聚苯乙烯,是将普通硅胶经硅醇基与有机氯硅烷或有机烷氧硅烷试剂反应,进行化学修饰,键合上长度不同的烷基(R),形成亲油表面而成。常见的有 RP-18、RP-8、RP-2,三者亲脂性顺序为 $RP-18 > RP-8 > RP-2$ 。反相色谱的流动相多为一定比例的醇或乙腈的水溶液,一般为梯度洗脱。

近年来,反相色谱已成为天然产物研究的常规手段,广泛应用于极性较大的物质或水溶性物质的分离纯化,如生物碱、萜类、苷类等化合物。在实际分离纯化过程中,往往结合硅胶柱色谱法、凝胶柱色谱法、反相色谱法及大孔树脂等多种分离法,以使分离纯化顺利进行。下面的例子就是反相色谱法与其他色谱法结合的典型。

a. 真菌灵芝中酚性成分的纯化 灵芝是一味著名的传统保健品和补药,化学成分研究主要涉及三萜和灵芝多糖。Niu 等采用硅胶柱色谱结合反相柱色谱及大孔树脂色谱法,从灵芝新鲜子实体中分离得到 3 个异戊烯化苯酚(3-47~3-49), (图 3-25、图 3-26)。

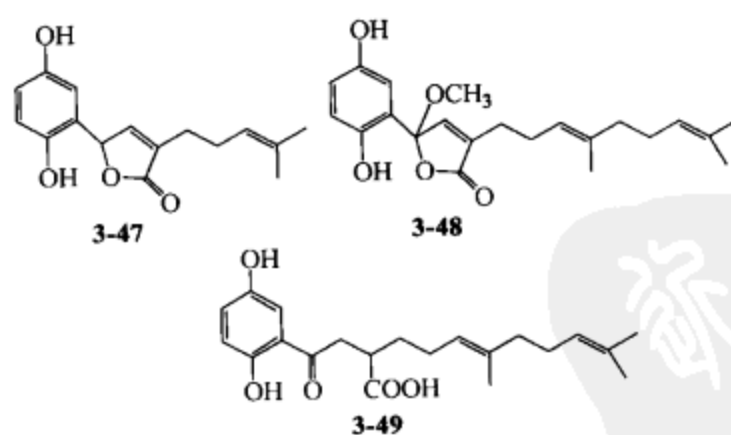


图 3-25 灵芝新鲜子实体中分离的异戊烯化苯酚结构

b. 唇型科植物二萜成分的分离纯化 唇型科香茶菜属包括 150 余种植物,该属因含有结构多变的二萜类成分和有意义的生物活性而受到极大重视。Sun 等为了寻找抗肿瘤活性成分,系统地研究了该属 50 多种植物,发现了 500 余个新二萜类成分,其中有一些二萜具有潜在的抗肿瘤作用,如冬凌草甲素、冬凌草乙素、毛萼

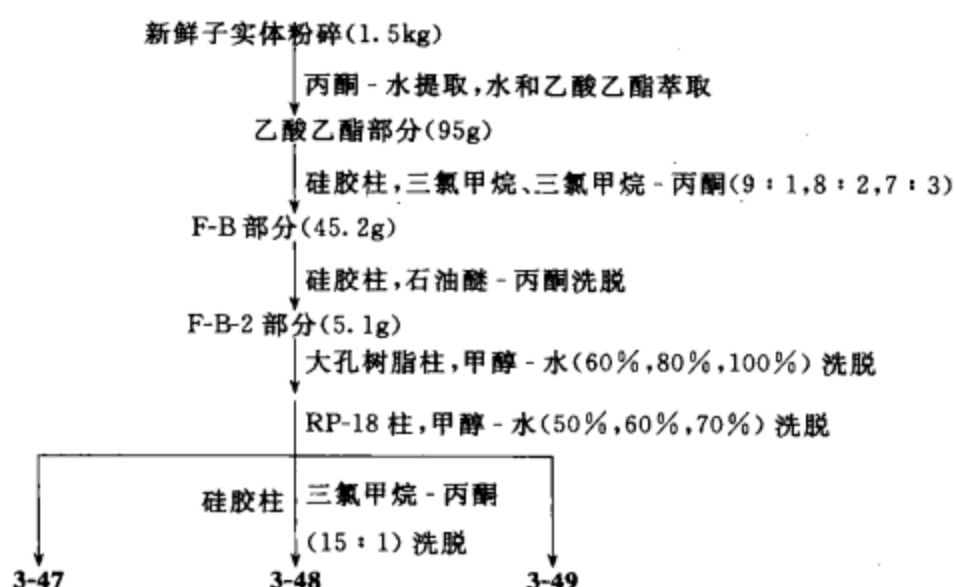


图 3-26 灵芝新鲜子实体中异戊烯化苯酚成分分离流程

乙素等。该课题组进一步寻找抗肿瘤成分,利用硅胶柱色谱和 RP 柱色谱法从该科香茶菜属植物白柔毛香茶菜 (*Isodon albopilosus*) 中得到 3 个阿替烷 (*ent*-Atisene) 型二萜 (3-50~3-52) (图 3-27、图 3-28)。

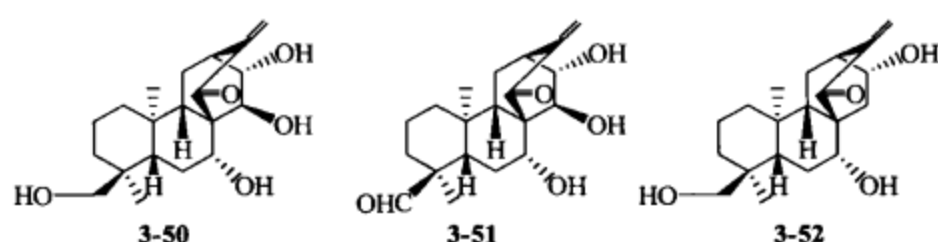


图 3-27 香茶菜属植物白柔毛香茶菜 (*Isodon albopilosus*) 中得到的二萜结构

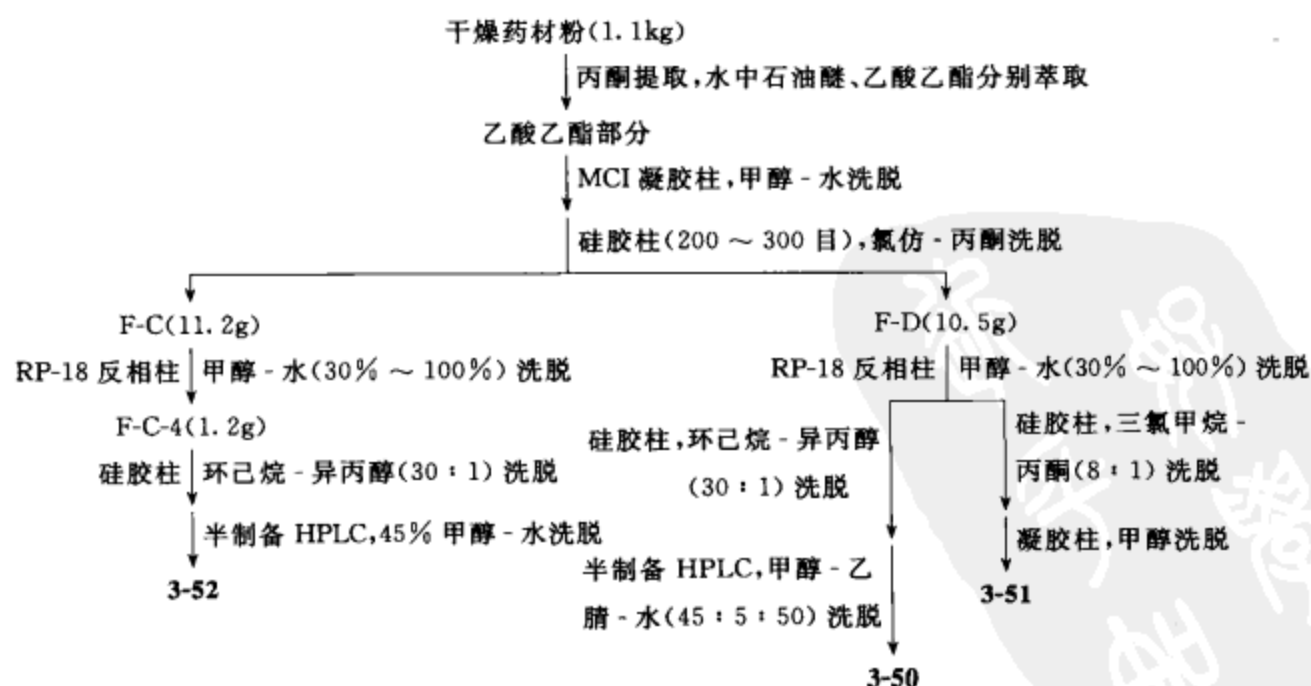


图 3-28 香茶菜属植物白柔毛香茶菜 (*Isodon albopilosus*) 中二萜的分离流程

3.2.2 天然产物的现代分离方法

3.2.2.1 制备薄层色谱

经典薄层色谱的流动相依靠毛细管作用力流经固定相，而制备薄层色谱的流动相是依靠外力强制流经固定相，如加压薄层色谱和离心薄层色谱。

(1) 传统制备薄层色谱 (PTLC) 传统制备薄层色谱一般配合柱色谱用于天然产物的分离，是最简单的制备型色谱分离方法之一。该法常用于分离纯化毫克级的样品。薄层板点样前最好先用溶剂展开一次，以减少化合物从吸附剂上洗脱下来时带入的杂质。上样时应使用易挥发性溶剂溶解，不易挥发性溶剂易使点样带变宽，降低分离效果。如果点样带较宽，应在展开前使用较大极性的溶剂先展开至点样带上 2cm 处，可以起到压缩点样带的作用。为了减少边缘效应，可以在展开槽内悬吊一滤纸，其下端接着展开剂，以促进展开剂蒸气饱和。PTLC 的展开剂的选择可以根据分析型的 TLC 来确定。

(2) 加压薄层色谱 (OPLC) 加压薄层色谱装置由一弹性的气垫覆盖在水平的薄层色谱板上，外力使薄层色谱在分离过程中没有气相存在，展开剂被强制流动，使样品展开速度加快。与 PTLC 相比，OPLC 由于采用更细颗粒的吸附剂及更长的色谱板，分离效果较好。OPLC 分离所需时间短，扩散效应相应减少。利用 OPLC 进行分离时，开始阶段干板一般处于不平衡系统中，使溶剂系统分层，从而导致样品色带形状变异，故需要先用溶剂对色谱板进行平衡，即可避免使用时溶剂分层的发生 (图 3-29)。

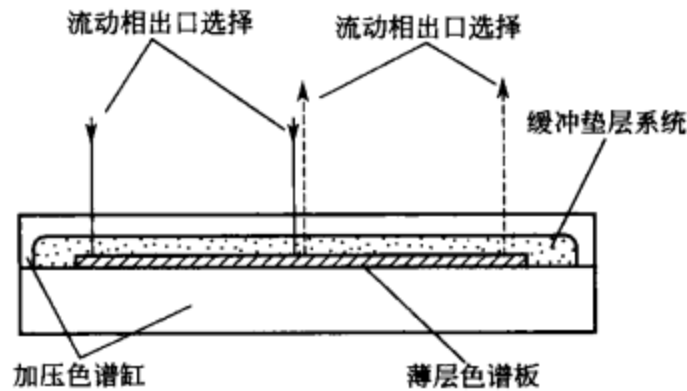


图 3-29 加压薄层色谱分离原理

(3) 离心薄层色谱 离心薄层色谱技术主要是在传统的 PTLC 基础上运用离心力，促进流动相加速流经固定相。洗脱过程就是不同组分成同心圆分布的过程，在离心力的作用下，当色谱带洗脱到色谱板边缘后，色带快速旋转脱离色谱板，并经色谱室内的流出管收集，同时利用薄层色谱对所收集的馏分进行分析，从而使不同成分分离 (图 3-30)。

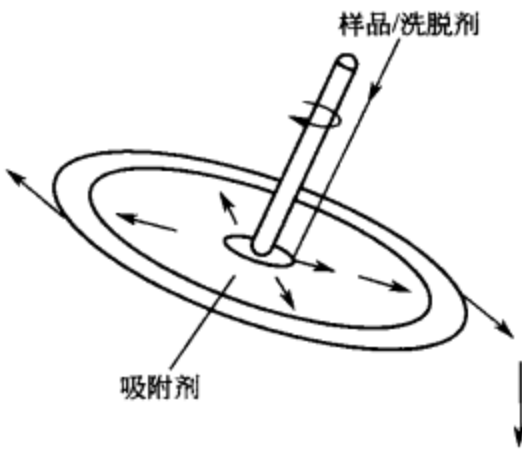


图 3-30 离心薄层色谱分离原理

离心力，促进流动相加速流经固定相。洗脱过程就是不同组分成同心圆分布的过程，在离心力的作用下，当色谱带洗脱到色谱板边缘后，色带快速旋转脱离色谱板，并经色谱室内的流出管收集，同时利用薄层色谱对所收集的馏分进行分析，从而使不同成分分离 (图 3-30)。

离心薄层色谱法与 PTLC 相比，主要优点是：上样方便；不需要刮板，直接收集洗脱液；可采用梯度洗脱，薄层经简单处理后可以反复使用。

3.2.2.2 制备型柱色谱

经典柱色谱是靠重力的作用使流动相流经固定相的色谱，操作简单，但分离速度慢，为保证有足够快的流动相流动速度，仅适合于颗粒度较大的固定相。为克服经典柱色谱的不足，人们不断制造出不同的制备型柱色谱。

(1) 低压柱色谱 目前使用最广泛的制备型低压柱色谱系统是 Lobar 系列产品，其中包括多种规格的预制柱。因色谱柱是由玻璃制成，可以看到其中的填料。色谱柱中的固定相被柱两端的烧结玻璃封闭。色谱柱两端连有聚四氟乙烯环的金属管，可与输液泵相

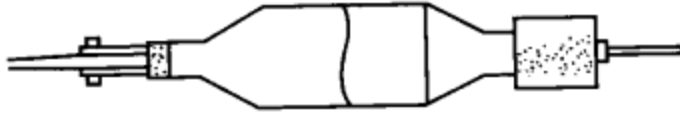


图 3-31 低压预制柱示意

连。有时可以将几根柱串联起来使用，以提高分辨率或增加上样量。Lobar 柱可以反复使用，使用寿命较长，且反相柱寿命长于正相柱。Lobar 柱洗脱速度较快，使用过程中需要输液泵，一般采用组成恒定的洗脱剂（图 3-31）。

(2) 中压柱色谱 中压柱色谱比低压柱色谱使用的填料颗粒度更小，分辨率更高，需要使用更大的压力来维持适当的流速，所需压力可由压缩空气或往返泵提供。中压柱色谱采用更长、具有更大内径的色谱柱，可以一次性分离更多样品。一般先用分析型 HPLC 选择适当的溶剂，再使用中压柱色谱分离。

Büchi 公司生产全套的中压液相色谱装置，其中包括各种规格的中压色谱柱，可用于分离 100mg~100g 的样品。活塞泵及可替换泵头可在最大压力为 4.0MPa 时，使流速在 3~160mL/min 之间调节，该系统也接有溶剂梯度形成装置，并可用于各种紫外检测仪器检测，其中包括可变波长检测器。

在制备色谱柱前连接一个前置柱，则在每次分离后，可将被污染的填料除去。利用甲醇—乙酸乙酯—己烷顺序冲洗柱，可使硅胶固定相进行再生，但在重复几次之后，应将固定相废弃。键合相色谱更易于清洗，并具有更长的使用寿命（图 3-32）。

(3) 高压液相柱色谱 高压液相色谱系统的色谱柱内装的固定相是粒度范围较窄的微小颗粒，为使流动相流出，需采用较高的压力。中压液相色谱系统填装较大颗粒的固定相，装柱较容易，柱的通透性较高，可采用更大的色谱柱和更经济的仪器，但分辨率也较低。而高压液相色谱系统中使用较小颗粒的固定相，系统的复杂性及成本都增大，但分辨率也得到较大的提高。

半制备型高压液相色谱的色谱柱直径在 8~10mm，内装粒度为 10 μ m 的固定相，可用于 1~100mg 混合物样品的分离。更大量样品的分离可通过重复进样完成。制备型高压液相色谱通常采用恒定的洗脱剂，以减少操作中可能出现的问题。对于难分离的样品，有时需要在分离过程中采用梯度洗脱。

(4) 减压液相色谱 减压液相色谱（vacuum liquid chromatography, VLC）是借助真空为动力使流动相加速流经固定相的色谱分离方法。减压液相色谱不同于

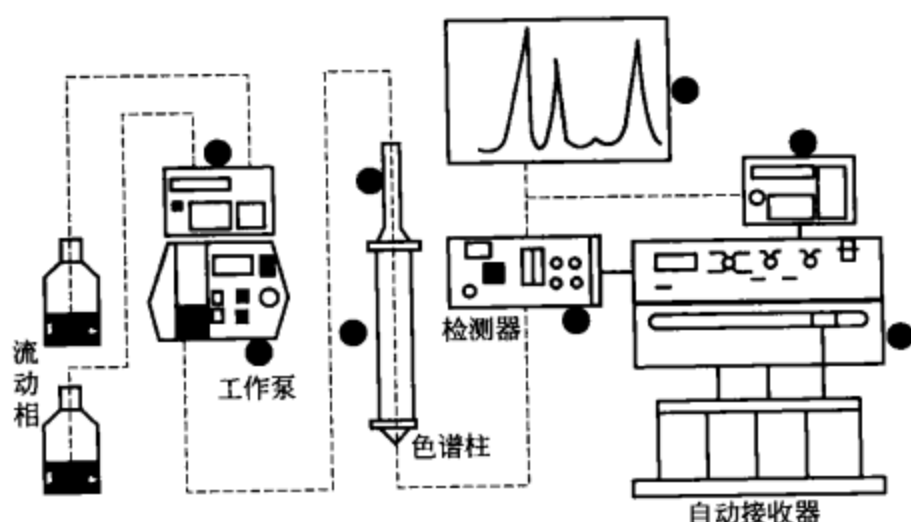


图 3-32 Büchi 公司 B-680A 中压液相色谱系统

快速柱色谱，因为该方法在收集每份流分后让色谱柱流干，类似于制备型薄层色谱，在完成一次展开，干燥后，再对其进行再次展开。减压柱色谱法具有设备简单、分离时间短、有较理想的分离度、分离容量大等特点。

减压液相色谱装置是在一装有熔融玻璃的短柱或布氏漏斗中加入吸附剂，轻击使其下沉，然后通过三通活塞抽真空，直至吸附剂变得坚硬。上样可以使用干法或湿法，流分的收集一般要求将柱体积抽干，从而减少流分之间的相互交叉。

3.2.2.3 逆流色谱

逆流色谱 (counter-current chromatography, CCC) 是一种基于某一样品在两个互不混溶的溶剂之间分配作用的分离技术，溶质中各组分在两相溶剂中的比例视它们的分配系数而定，是一种不用固态吸附剂的全液态的色谱方法。

逆流色谱装置实际上是由数百个单元部件组成，溶质在每一个单元的两个液相层之间进行分配，然后其中的一层转移到下一单元中去再分配。在逆流色谱中，两相层的平衡完成于转移之前。与液相色谱相比，逆流色谱是一个连续的非平衡过程，它的溶剂是连续流通的。与液-固色谱技术相比可以避免样品的不可逆吸附现象，并可避免样品因与某些固体固定相相互作用发生的破坏。常见的逆流色谱技术有液滴逆流色谱、离心分配色谱和旋转小室逆流色谱。

液滴逆流色谱 (DCCC) 是在一组垂直排列的分离管内充满液态固定相，使移动相成液滴通过作为固定相的液柱，促使溶质在两相间连续分配 (图 3-33)。

3.2.2.4 分子烙印技术

分子烙印技术 (molecular imprinting technique, MIT) 的基本思想源于对抗体-抗原及酶与底物结合的高度专一性的认识。应用分子烙印技术的前提是合成出具有分子识别能力的聚合材料——分子烙印聚合物 (MIP)。以目标化合物为模板，将具有结构互补的功能化聚合物单体分子通过共价键或非共价键的方式与模板分子

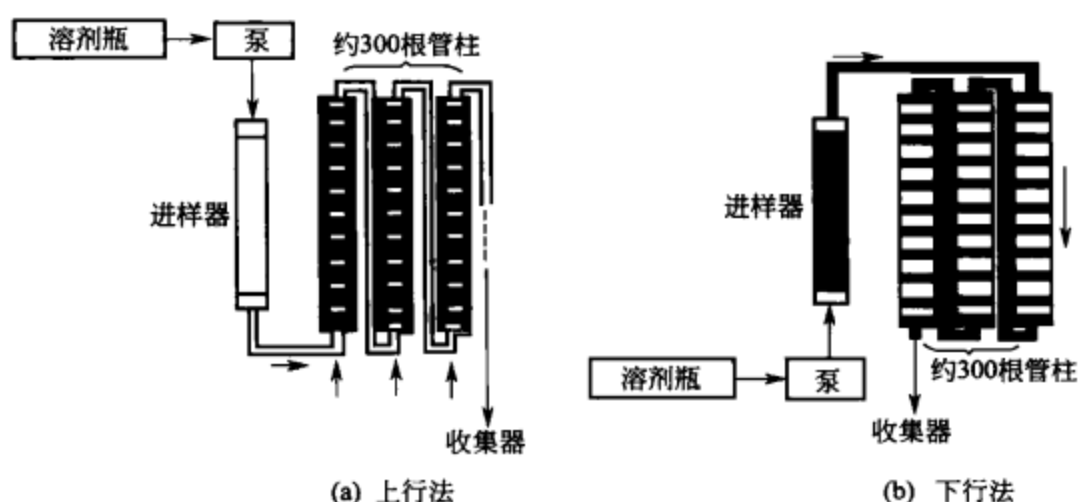


图 3-33 液滴逆流色谱分离示意

结合形成复合物，再加入交联剂，通过胶联聚合反应固化。反应完成后将模板分子除去，聚合物中即存在了一些大小、形状及功能基排列与模板分子互补的孔穴。该聚合物就是 MIP，孔穴就是模板分子的印记。MIP 对烙印分子的立体结构具有“记忆”功能。用 MIP 作固定相时，对烙印分子有较高的选择性，使保留时间较长，后出峰。另外，在更换洗脱剂时，固定相不会产生变形，具有较好的机械能和较长的使用寿命。详见第 8 章。

3.2.2.5 LC-MS/LC-NMR

在过去的 20 年间，高效液相色谱（HPLC）和核磁共振波谱技术（NMR）已经成为天然产物分离和鉴定的基本方法。各种色谱技术也能够分离纯化出 10mg 以下的样品，以用于质谱和核磁共振波谱来确定结构。但是这些经典的天然药物化学研究方法既浪费时间，又可能导致发现的化合物已经被报道或植物资源已经被研究。

已知化合物的快速鉴定及杂质的快速分离已经成为制约天然产物活性成分研究的关键。随着高通量活性筛选（high throughput screening, HTS）技术在天然产物活性成分筛选中的运用，进一步加剧了对已知化合物的快速鉴定及杂质的快速分离的要求。组合分离科学和结构鉴定技术应运而生，如 LC-MS 及 LC-NMR 技术。LC-MS 法仅能提供分子量而给结构鉴定带来困难，降低了鉴定结果的可信性，需要更大的信息来彻底确定结构。LC-NMR 法能够提供大量的关于结构的信息，包括结构中每个质子的化学环境及耦合情况。因此，LC-NMR 法能够很容易辨别相同分子量，甚至相同化学式的不同化合物；也可区分结构异构体及几何异构体。Steven C. Bobzina 等采用 LC-NMR 法，从海绵 *Aaptos* sp. 中分离得到生物碱阿普他明（3-51）（图 3-34）。

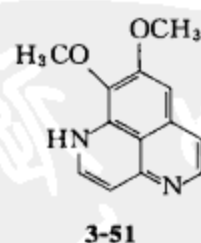


图 3-34 LC-NMR 法从海绵 *Aaptos* sp. 中得到的生物碱结构

参 考 文 献

- [1] 徐任生等. 天然产物化学导论. 北京: 科学出版社, 2006.
- [2] 吴立军等. 实用天然有机产物化学. 北京: 人民卫生出版社, 2007.
- [3] 徐任生等. 中草药有效成分提取与分离. 第2版. 上海: 上海科学技术出版社, 1983.
- [4] Mariani C, et al. *Phytochemistry*, 2008, 69: 1220.
- [5] Kato A, et al. *Phytochemistry*, 2008, 69: 1261.
- [6] Asai T, et al. *Phytochemistry*, 2008, 69: 1234.
- [7] Itoh A, et al. *Phytochemistry*, 2008, 69: 1208.
- [8] Gressler V, et al. *Phytochemistry*, 2008, 69: 994.
- [9] Cioffi G, et al. *Phytochemistry*, 2008, 69: 1005.
- [10] Li X L, et al. *J Nat Prod*, 2007, 70: 265.
- [11] 卢艳花. 中药有效成分提取分离实例. 北京: 化学工业出版社, 2007.
- [12] Li C F, et al. *Nat Prod Res Dev*, 2007, 19: 443.
- [13] Kang W Y, et al. *Modern Pharmaceutical Research*, 2008, 1: 22.
- [14] K 霍斯泰特曼著. 制备色谱技术. 赵维民, 张天佑译. 北京: 科学出版社, 2000: 155.
- [15] Ngandeu F, et al. *Phytochemistry*, 2008, 69: 258.
- [16] Zhang H, et al. *Phytochemistry*, 2008, 69: 271.
- [17] 祝经平等. 中成药, 1997, 19 (3): 36.
- [18] Hong D, et al. *Phytochemistry*, 2004, 65: 2595.
- [19] Mu S Z, et al. *J. Nat. Prod.* 2006, 69, 1065-1069.
- [20] 李茂星等. 中国中药杂志, 2007, 32: 1743.
- [21] 袁怀波等. 食品与机械, 2007, 23 (4): 79.
- [22] 陈乃富等. 中药材, 2007, 30 (8): 1013.
- [23] Duan D L, et al. *Bioorganic and Medicine Chemistry Letter*, 2004, 14: 6041.
- [24] Niu X M, et al. *J Nat Prod*, 2006, 69: 1364.
- [25] Huang S X, et al. *J Nat Prod*, 2007, 70: 1053.
- [26] 王夔. 中药研究现代方法学. 北京: 化学工业出版社, 2004.
- [27] Bobzin S C, et al. *J Chromatogr B*, 2000, 748: 259.

鄧子龍

PDG

第4章 天然产物化学成分分析及结构鉴定

4.1 天然产物化学成分结构鉴定方法学

核磁共振谱学 (nuclear magnetic resonance, NMR spectroscopy) 是 20 世纪中叶兴起的一门化学结构解析的学科。20 世纪 80 年代, 核磁技术的一维、二维及多维脉冲傅里叶变换核磁共振的相关理论的发展, 为脉冲傅里叶变换核磁共振技术的不断发展奠定了结构鉴定的理论基础。

以往天然产物的结构测定是以积累、渐进的方式进行的, 即依赖于前人积累的天然产物知识库。要确定一个具有全新骨架的天然产物的结构是十分不容易的, 往往要通过理化性质测定、波谱数据、化学降解、化学转化、半合成、全合成或晶体结构等手段, 经过数月乃至数年的努力才能完成。后人进行某种天然产物的结构研究往往是先要查阅文献, 收集有关该生物资源已知天然产物的信息。然后, 对分离纯化得到的化合物通过测定理化性质判断其属于何种类型化合物, 再进一步通过波谱数据判断与已知天然产物库中某个化合物的结构差异, 或通过化学变换与已知结构的天然产物相关联等方法才能确定其结构。二维 NMR 波谱技术的发展, 有可能对化合物的 ^1H 、 ^{13}C 、 ^{15}N 等核的性质以及它们间的连接关系进行详尽的研究, 并通过对这些性质及连接关系的解析, 对天然产物的结构进行系统的、全新的结构解析。因此, 天然产物的结构测定进入了跃进式发展的阶段: 具有新颖构造的天然化合物的结构测定往往只要数周乃至数日就可完成, 具有全新骨架的天然产物也层出不穷。

近年来, 核磁共振技术已经成为化学、医药、生物、物理等领域必不可少的研究手段, 同时, 核磁共振成像学及相关设备也成为医学领域无损伤检测疾病的重要工具。但是, 有关采用 NMR 波谱技术解析天然化合物产物结构的方法, 即根据 NMR 数据推导和解析分子结构的策略、技术和方法学的介绍书尚很少。

本章节对于其他教科书、专著中已有详细介绍的专业名词、技术参数和实验方法 (如化学位移、自旋-自旋偶合、偶合常数、信号多重性等) 不再重复, 而把重点放在介绍采用一维和二维 NMR 波谱进行结构解析策略和战术思想的运用, 提供大量的结构解析例子以帮助掌握 NMR 解析结构的技能, 希望对这些领域的后来者有所帮助。

4.1.1 核磁共振技术

在核磁共振技术中，氢 (^1H) 和碳 (^{13}C) 核磁共振技术 (NMR) 是天然产物结构测定中最为重要的工具。两者能够提供较为复杂分子的氢及碳的类型、数目、相互连接方式、周围化学环境、空间排列的结构信息。特别是二维核磁共振技术，对未知的有机化合物的结构类型甚至手性分子的确定发挥着重要的作用。

4.1.1.1 核磁共振氢谱

根据化学位移的定义，TMS (四甲基硅烷) 将会出现在右边的末端，并定为化学位移的起始点，具有较低的化学位移值，被称为高场；而处于左边，具有较高的化学位移的区域称为低场。

在 CW 谱 (连续谱) 中，信号的吸收与位于信号频率的质子数是成比例的，结果使吸收峰的面积与被检测到的质子数成正比。在 FT 谱 (傅里叶变换谱) 中则没有这么直接，为了使得到的积分可靠，在相继脉冲之间有必要使所有弛豫到平衡分布。在 ^1H 中，一般测量可以实现，但在 ^{13}C 谱中却并非如此。

影响化学位移的因素较多，可分为分子内因素和分子间因素。分子内因素包括诱导效应、化学键的各向异性。分子间因素有氢键、温度、溶剂等。各种因素影响化学位移的根本原因是：改变了原子或分子的外层电子分布，使电子云密度存在差异。

相邻的两个 (组) 磁性核之间的自旋偶合或自旋干扰就产生峰的裂分。两个裂分小峰间的距离就是自旋-自旋偶合常数，简称偶合常数 J 。在 ^1H -NMR 谱中影响最大的为氢核相互之间发生的自旋偶合，即同核偶合。

^1H -NMR 谱解析的一般程序如下。

- ① 检查图谱，包括内标物信号、基线平坦、溶剂峰位、信噪比等。
- ② 对照文献确定溶剂峰位。
- ③ 在高场寻找特征峰位，如甲基、甲氧基、乙酰基、双键甲基等信号，并以其峰面积核对确定其他信号相应氢核数。
- ④ 在低场区 δ (化学位移) 10~16 处，寻找出现的羧基。
- ⑤ 参考化学位移、自旋裂分情况，寻找低级偶合系统。
- ⑥ 结合其他图谱逐步确认。

4.1.1.2 核磁共振碳谱

由于 ^{13}C 的天然丰度仅 1.1%，灵敏度低，要获得一张碳谱需较多的样品和较长的测定时间；同时，在对碳谱的信号进行归属时，除了化学位移外，确定碳的性质 (多重性) 亦是十分关键的，虽然可用偏共振去偶法，但在复杂分子中偏共振去偶谱信号重叠，常常难于辨认。

不灵敏核的激化转移增益法 (INEPT) 通过极化转移，结合自旋回波，不仅可使 ^{13}C 核的信号强度增加四倍，更有益的是可以有效地区别伯碳、仲碳、叔碳和

季碳。

无畸变激化转移增益法 (DEPT) 是较 INEPT 优越的一种极化转移法。在 INEPT 法中, ^{13}C 的信号与 $^1J_{\text{CH}}$ 和延迟时间有关, 由于化合物分子各碳有差异, 难于进行最佳的延迟时间的设置, 因而会导致碳的信号强度和相位发生畸变。但在 DEPT 实验中, 脉冲间隔时间的选择对 $^1J_{\text{CH}}$ 的依赖不大, 碳信号的 3 种不同的多重性仅与脉冲的倾角有关。若分别采用 45° 、 90° 和 135° 3 个倾角发射脉冲进行 3 次实验, 就可以清楚地区别不同类型的碳。当倾角为 45° 时, 所有的 CH 、 CH_2 和 CH_3 均显正信号; 当倾角为 90° 时, 仅显示 CH 信号; 当倾角为 135° 时, CH_2 为负信号, 而 CH 和 CH_3 则为正信号; 季碳在 DEPT 谱中不出现共振信号。

4.1.1.3 二维核磁共振谱

二维核磁共振 (2D-NMR) 波谱技术对天然产物结构测定的方式已经产生了巨大的影响。2D-NMR 起源于傅里叶变换的概念, 即在时间范畴测得的 NMR 信号通过傅里叶变换就能得到频率范畴的 NMR 谱。如果把这个概念扩展到有两个时间变量的实验, 通过傅里叶变换就能得到二维频率的 NMR 谱 (2D-NMR)。2D-NMR 与 1D-NMR 一样, 都是探测核的磁化矢量在磁场中进动时的相互间的相干作用。这种相干作用有两种: 一种是通过共价键的标量耦合; 另一种是通过空间偶极的相互作用, 又叫 NOE 效应。检测通过化学键的标量相互耦合就能建立核与核通过化学键的连接关系。如果检测的是同种类核间的偶合作用, 得到的就叫同核相关谱 (COSY), 最常用的就是 H, H -COSY。如果检测的是异核之间的偶合关系, 得到的就叫异核相关谱, 最常用的就是 C, H -COSY。因此, 通过这些偶合相关信号的测定就能确定分子的共价键连接方式。另外, 探测核与核之间通过空间的偶极相互作用, 就能建立核与核通过空间的连接关系以及原子在空间的配置关系。可见, 前者能够补充有关分子平面结构的信息, 而后者能够提供有关分子的立体化学的信息, 是研究分子构型、构象和分子运动性的重要工具。

传统的 NMR 方法测定结构主要借助于化学位移 δ , 它包含有核化学还击的直接信息。在 NMR 技术应用的早期阶段特别强调 δ 的重要性, 常常把从未知物观测到的 δ 值与某一结构的标准数据通过经验规则加以比较, 并以此作出推测或判断。但是, 如果仅仅利用化学位移的信息进行结构解析, 这一技术就失去了真正的价值。在天然产物的结构解析中, δ 并不是 NMR 最重要的性质。NMR 技术能够提供组成分子的各种核之间的相关与连接信息, NMR 共振信号的完全归属能够在解析 2D-NMR 波谱数据的过程中自然而然地得到。

(1) H, H -COSY (质子间孪生、邻位和远程偶合相关) H, H -COSY 是最简单的二维 NMR 实验, 它关联分子中偶合质子的化学位移。在此二维谱中, 两个频率轴都以质子的化学位移作图, 因此谱图是方形对称的。其一维 ^1H -NMR 谱的投影出现于对角线上。对角线以外信号称交叉峰或相关信号, 它们处于 ^1H 的 1D-

NMR 信号的水平和垂直连线的交叉点上, 表明对应的两个质子间相互自旋偶合。因此, H,H-COSY 图谱揭示的是质子与质子间的自旋偶合关系。它在结构测定中可以替代自旋去偶实验, 快速确定分子中质子与质子间通过化学键的偶合关系, 包括质子间孪生 ($^2J_{HH}$)、邻位 ($^3J_{HH}$) 或远程。

H,H-COSY 谱图可按堆笈作图, 显示为三维图形, 信号强度以峰高表达在第三维。这种表达方式由于透视关系, 容易产生信号畸变和重叠, 也不便于作图和分析, 实际使用中很少采用。另外一种形式是等高线作图, 信号强度以一系列同心圆表示, 有些类似于军事地图。等高线作图显示明了, 有利于谱图分析。但是截面高度的选择可能会影响其提供的信息, 截面选择过高, 相关信号就弱甚至消失, 如果截面选择过低, 就会出现假峰和噪声信号, 干扰正常的谱图分析。实用中常常选择适当的等高线作图, 以得到包含最多的结构信息和最少的噪声的谱图, 或以不同等高线分别作图, 得到多个谱图, 相互参考以得到更多的信息, 并辨别和剔出假峰和噪声信号。

H,H-COSY 技术也有限制: ①因为交叉峰的截面图可能并未揭示所包含的偶合是属于孪生、相邻或远程偶合中的哪一种, 如果不考虑化学位移和可能的偶合常数, 解析就不一定能够得到结论性论点; ②图谱中重叠的信号无法用 H,H-COSY 分离。但是, 如果采用其他相敏的双量子滤波的 H,H-COSY 谱, 谱图的数据分辨率又足够高, 交叉峰的多重度的精细结构就能识别, 上述限制就能在很大程度上得到克服。

H,H-COSY 谱图揭示的是质子间孪生、相邻或远程质子间的标量偶合, 因此提供的是分子中原子与原子间通过键的连接。它们能够分别提供分子内结构片断的信息。尽管存在着种种限制, 由于其灵敏度高、结构信息丰富, 通过解析 H,H-COSY 方法, 接 H 碳骨架结构片断总能够能到, 加上从其他二维 NMR 实验得到互补信息, 就能得到完整的分子结构。因此, H,H-COSY 不失为最有效、最常用的二维 NMR 方法。

(2) H,H-NOESY 检测 H,H-NOESY 相关的脉冲序列也可以改变为检测质子间偶极相互作用, 这样 H,H-COSY 序列就转变为 H,H-NOESY 序列, 得到的就是 NOE (核极化效应) 信号强度的二维变化。NOESY 谱与 H,H-COSY 谱相似, 是方形的。对角线上有对角峰, 对角线外显示相关峰。但是相关峰揭示的是质子与质子间在空间的相互接近关系, 而非质子与质子通过键的标量偶合。NOESY 图谱在天然产物的结构解析中有两方面的功能。首先, 利用 NOESY 谱提供的相关信息, 也可以把分子中的原子或片断结构连接起来, 这样的连接称为通过空间的连接。当用 H,H-COSY 谱进行结构解析遇到困难时 (如相邻质子间二面角接近 90° , 偶合常数接近于零或受季碳原子阻隔), NOESY 常常能够提供有用的信息。因此, 在天然产物平面结构的解析中, 通过空间的连接是对通过键的连接的有力的

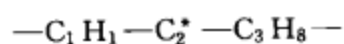
补充。NOESY 谱更重要的功能是能够通过分子内原子与原子在空间上的相互配置关系提供有关分子相对立体化学和溶液构象方面的重要信息，是研究分子构型、构象和运动性的重要工具。NOESY 谱揭示分子内所有的质子与质子间的通过空间的相互接近，因此效率比较高，特别适用于具有复杂结构的天然产物分子或生物大分子的研究。

(3) C,H-COSY C,H-COSY 通过单键 CH 偶合常数 $^1J_{CH}$ 把一维频率轴上的 ^{13}C 化学位移与另一维频率轴上的 1H 化学位移关联起来。所用的脉冲序列包含 1H - ^{13}C 极化转移，可提高灵敏度4倍。因此，C,H-COSY 一般并不比宽带去偶 ^{13}C -NMR 谱需要更多的样品。在二维 ^{13}C 和 1H 相关谱图中， ^{13}C 化学位移出现在 x 轴， 1H 化学位移现在 y 轴，交叉峰出现在相应的 ^{13}C 和 1H 信号的垂直和水平连线的交叉点上，提示相应的 1H 和 ^{13}C 核是通过 $^1J_{CH}$ 相互关联的。由于两个轴的频率不同，所以C,H-COSY 谱图中不出现对角峰。C,H-COSY 谱图的解析十分简单明了，但是所提供的信息非常可贵。首先，它能够提供C、H之间通过键的直接连接($^1J_{CH}$)，清楚地显示在一个C原子上连接几个H原子，提供了C共振信号的多重度的信息。同时揭示出哪个H原子接在哪个C原子上，为把质子与质子间的偶合相关信息转变成为C原子与C原子的连接提供了基础。不仅如此，在具有多环体系的复杂天然产物中，一些环上的 CH_2 结构单元的两个质子常常是磁不等价的，这种孪生质子间的 $^2J_{HH}$ 偶合与相邻质子间的 $^3J_{HH}$ 偶合混杂在一起，往往干扰质子偶合体系的分析。C,H-COSY 谱为非等价孪生质子的鉴定提供了有效的手段。此外，由于 ^{13}C -NMR 信号具有比 1H -NMR 信号化学位移散布宽的特点，一般在一维 1H -NMR 谱图中十分拥挤的区域中的质子信号在C,H-COSY 谱中随化学位移分布极宽的 ^{13}C -NMR 信号而散布开来，大大简化了 1H -NMR 信号的解析，这对于结构十分复杂的天然产物的结构解析和生物大分子溶液三维结构的计算十分有利。另外，还可以采用谱图编辑的手法，选择在相应的 ^{13}C -NMR 信号的位置作截面，得到相应的 1H -NMR 截面图，如果分辨率足够高，甚至可以从截面图得到相应质子信号多重度的精细结构信息。

(4) C,H-COLOC (孪生和相邻的CH相关) C,H-COSY 实验提供了C、H通过 $^1J_{CH}$ 偶合的连接，因此它只适用于解析连H的C原子，而不适用于解析不连H的季碳原子。但是C,H-COSY 实验的一个变种也可适用于解析季碳原子，就是在C,H-COSY 实验中把探测的C、H偶合常数调节到相当于 $^2J_{CH}$ 和 $^3J_{CH}$ 的比较大的数值(2~25Hz)，这个实验简称为C,H-COLOC。这种二维相关谱检测C、H通过双键或叁键的连接关系，同时抑制C、H间通过单键的偶合相关。这种谱图的图形和解析方法与C,H-COSY 十分类似，但是谱图交叉峰揭示的不是C、H之间隔单键的连接，而是间隔双键或叁键的连接关系。

C,H-COLOC 谱图中对应于季碳原子化学位移的交叉峰提示出与该季碳原子

相连接的分子部分构造的信息，如下所示：



式中， C_2^* 为季碳原子。

如果 H_1 显示与 C_2^* 和 C_3 的交叉峰或 H_3 显示与 C_2^* 和 C_1 的交叉峰，可以借此把 C_1 、 C_2^* 和 C_3 3 个 C 原子依次连接起来。因此，这个试验的实质是提供了 C 原子与 C 原子之间的连接。由于连 H 的 C 原子之间的连接可以通过 H, H-COSY 加上 C, H-COSY 得到，所以 COLOC 试验补充了把季碳原子接入分子骨架的手段，或者说通过季碳原子 C_2^* 把两侧的质子偶合链连接起来的手段，因此，COLOC 试验提供的信息十分重要。

充分利用 2D-NMR 提供的 C、H 隔双键和叁键的连接关系，把分子中的被季碳原子或杂原子隔断的部分构造连接起来，是采用近代 2D-NMR 技术解析天然产物结构的核心部分，也是结构解析取得成功的关键之一。

4.1.2 有机质谱

质谱 (mass spectrum, MS) 就是把带电荷的分子或经一定方式裂解形成的碎片离子按照质荷比 (m/z) 大小排列而成的图谱。由于分子离子和碎片离子往往带有一个电荷，因此，质谱所给出的信号又表示分子离子或碎片离子的质量。图谱中的每一个峰代表一个质量数。根据分子离子峰 (通常在最高处) 与碎片峰的质荷比推导出化合物的结构。高分辨质谱则可以直接给出相对分子质量的精确数字和分子式。这是解析化合物结构的重要基础。

质谱常用的离子源是电子轰击源 (EI)。电子轰击电离又称电子电离，是目前发展最成熟的电离方法。EI 过程在离子源中进行，呈气态的试样分子在较高真空和较高温度的电离室内，热阴极发射的电子经加速达到 70eV，与试样分子作用，外加一个辅助磁场，使电子导向性强，电子的运动轨迹呈螺旋性，可加大电子与试验分子的作用概率。

新技术电喷雾电离 (ESI) 与基质辅助激光解吸电离 (MALDI) 已日益普及，而联有飞行时间 (TOF) 的 MALDI-TOF-MS 与 ESI-TOF-MS 仪还可一次分析几千个生物大分子，从而大大促进了高效筛选药物的进程。质谱仪的发展的另一个热点是与液相色谱 (LC) 相联，目前较普及的有 LC-MS、LC-MS-MS、LC-ESI。

4.1.3 其他波谱技术

(1) 紫外吸收光谱 分子吸收波长范围在 200~400nm 区间的电磁波产生的吸收光谱称为紫外吸收光谱 (UV)。紫外光谱在许多领域里得到了广泛的应用。在天然产物的结构解析中，紫外光谱主要用于提供分子的芳香结构和共轭体系信息。二维核磁共振技术的发展已使紫外光谱成为次要的结构鉴定手段。常用的溶剂为甲醇或乙醇。当溶剂的极性增加时，极性键的跃迁受溶剂极性的影响变化如下： $\pi-\pi^*$

带发生红移, $n-\pi^*$ 带发生蓝移。

(2) 红外光谱 红外光谱 (IR) 是研究红外光与物质分子间相互作用, 物质吸收红外光后产生化学键振动而形成的吸收光谱。测定范围一般是 $\nu_{\max}$ 为 $500 \sim 4000\text{cm}^{-1}$ 。最方便的使用方法是样品与 KBr 混合压片, 它能处理少至 $5 \sim 10\mu\text{g}$ 的样品, 由于在压片过程中不可避免地要吸收水分, 因而在 3400cm^{-1} 和 1640cm^{-1} 附近常出现吸收峰。

近年来, 由于核磁共振谱和质谱的广泛应用, 红外光谱的应用范围大大缩小。目前, 它主要用于鉴别各种羰基、炔基、羟基等基团。

(3) 单晶衍射技术 晶体能对 X 射线、电子束和中子束产生衍射效应。利用这种衍射可以测定晶态物质在原子水平上的结构, 提供详尽的结构数据: 原子间的距离、键角、扭角, 分子的立体结构、绝对构型、热振动参数, 原子和分子的堆积、有序或无序的排列以及非计量的程度等。而获得这些数据的样品只需用一颗直径大约为 0.3mm 的单晶体。利用晶体的衍射效应测定晶体的结构, 关键的前提是培养出合用的单晶体。溶液中培养单晶体的一些共同的条件是: 结晶物质要纯净, 过饱和度要低, 溶液的组成和温度合适。常用的一些单晶体培养方法: 溶剂蒸发法、改变温度法、电化学制备法、蒸汽扩散法和凝胶法。

(4) 圆二色谱技术 圆二色谱 (CD) 是研究化合物在紫外光与可见光区所产生的椭圆偏振光的椭圆度与波长的关系的图谱。圆二色谱是透过介质后光的吸收曲线, 而旋光谱则是它的分散曲线。

由于一个已知其相对构型的化合物的绝对构型仅有两种可能, CD 的测定有助于推定出正确的结果。

4.2 天然产物化学成分的结构鉴定

4.2.1 萜类化合物

萜类是天然产物中的一组重要成分, 是由不同个数的异戊烯首尾相连构成的, 根据异戊烯数目可将萜类化合物分为单萜、倍半萜、二萜、三萜等。单萜化合物存在于挥发油内, 是一类具有挥发性、可随水蒸气蒸馏出来的油状液体, 大都有香气, 分布于植物界的唇形科、伞形科、菊科、芸香科、樟科和姜科等。倍半萜广泛存在于植物、微生物、海洋生物及某些昆虫中, 其中很多具有重要的生物功能和生理活性, 特别是倍半萜内酯, 具有诸如抗菌、抗肿瘤、抗病毒、细胞毒、免疫抑制、植物度、昆虫激素、昆虫拒食剂等活性。近年来从植物特别是海洋生物中分离得到一系列特殊结构的化合物, 为倍半萜的研究又增加了很多化学和生理作用方面的新奇内容。二萜和三萜均是自然界中广泛分布的一类天然化合物, 尤以双子叶植物中分布最多。

4.2.1.1 倍半萜及二萜

倍半萜类化合物骨架仅有 15 个碳原子，但是由于生物体内微妙的代谢过程而生成各种类型基本骨架的化合物，并且有复杂的立体结构。倍半萜类化合物基本骨架没有共同的紫外发色团，因此紫外光谱无特征吸收。红外光谱对倍半萜类化合物的羧基的性质确定极为重要。质谱能够提供分子离子峰信息，但是对于分子中含有 OH 或 OCOR 的物质，分子离子峰在电子轰击质谱中常不出现或者很弱，需用 CI（化学电离）、FAB（快速原子轰击）或 FD（场解析）质谱测定。

应用高兆赫磁共振解析立体化学问题在倍半萜结构研究中极为重要，该研究方法主要利用以下原理：①核磁共振谱裂分常数的大小与其相对的立体化学关系密切，主要利用质子间的两面角关系解析其立体配位；②官能团间不同的立体位置，屏蔽作用不同；③原子与原子间的 NOE。

下面是一个具有双键的杜松烷型倍半萜的结构鉴定。

从松柏类植物杉科日本柳杉的芳香油中分离得到一种倍半萜成分 1(10)-杜松烯 (cadin-1(10)-ene, **4-1**)，质谱显示 m/z 206，进行一系列核磁共振测试。结构确定如下。

$^1\text{H-NMR}$ 谱显示存在 4 个甲基信号，其中 3 个为双峰 [δ : 0.78 (3H, d, $J=6.5\text{Hz}$, H-12), 0.87 (3H, d, $J=6.0\text{Hz}$, H-15), 0.94 (3H, d, $J=6.5\text{Hz}$, H-13)], 1 个为宽单峰 [δ : 1.63 (3H, brs, H-14)]。甲基宽单峰信号显示甲基与双键连接。

$^{13}\text{C-NMR}$ 谱及 HSQC [heteronuclear single-quantum coherence, (检出 ^1H 的) 异核单量子相干] 谱显示有 15 个碳原子，其中有 4 个甲基信号 (δ : 17.0, 19.1, 22.0, 22.5)、5 个亚甲基信号 (δ : 21.3, 29.7, 31.8, 36.1, 43.1)、4 个次甲基信号 (δ : 27.5, 33.0, 40.8, 46.8)、2 个烯季碳原子信号 (δ : 124.2, 132.5)。

DQF-COSY 和 HMBC 谱确定 **4-1** 的基本骨架 (图 4-1)。在 H,H-COSY 谱中显示，C-15 甲基与 C-4 连接，C-12 和 13 甲基与 C-11 连接，并且构成异丙基结构。再综合利用 HMBC 异核相关确定了化合物的结构。利用 NOESY 试验，进一步确定其立体相应构型 (图 4-2)。详细数据见表 4-1。

4.2.1.2 三萜

三萜是由 30 个碳原子、6 个异戊烯结构单元组成的萜类化合物。三萜及其皂苷的结构测定主要依照生源关系并采用化学和波谱等方法。由于生源关系，同属植物常含有结构类似的化学成分，所以查阅同属植物的化学成分研究报道对确定所研究植物中的三萜及皂苷成分的结构会有很大帮助。由于三萜含有较多的甲基、亚甲基及次甲基碳原子，氢谱高场重叠较为严重，使信号的全归属更加困难。但是利用 2D NMR 技术也能完全解析三萜及皂苷成分。

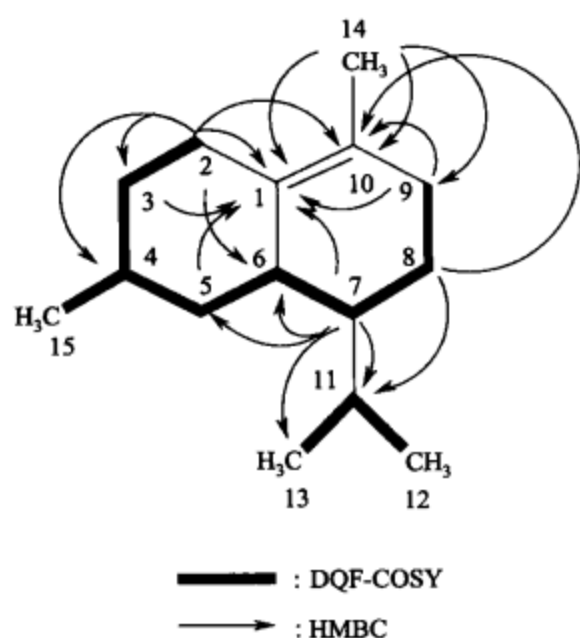


图 4-1 化合物 4-1 的 HMBC 相关性

DQF-COSY—double quantum filtered COSY

(双量子滤波 COSY); HMBC—heteronuclear

multiple-bond correlation (检出¹H 的) 异核多键相关

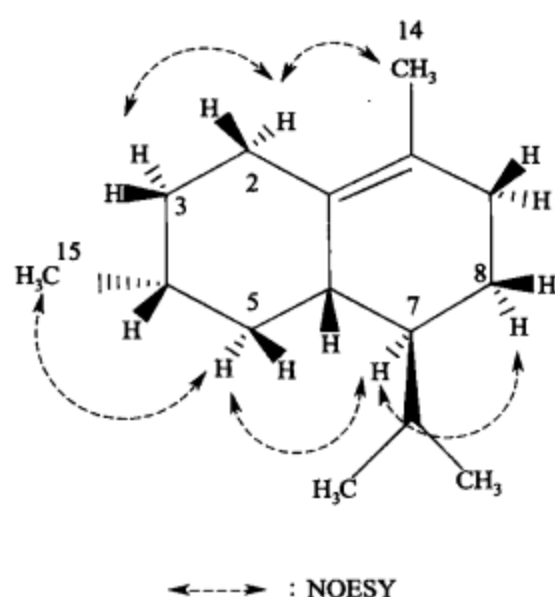


图 4-2 化合物 4-1 的 NOESY 相关性

表 4-1 化合物 4-1 的 NMR 数据 (500MHz, CDCl₃, JHz)

C	δ_C		δ_H	COSY	HMBC	NOESY
1	132.5	C	—	—	2a,3a,3b,5a,7,9,14	
2	29.7	CH ₂	a,2.69(1H,ddd, $J=3.5,3.5,13.5$) b,1.60(1H,m)	2b,3b 2a,3a,3b	3a,3b	2b,3a,14 2a
3	36.1	CH ₂	a,1.71(1H,m) b,0.85(1H,m)	2b,3b 2a,2b,3a,4	2a,4,5a,5b,15	2a,3b,4 3a
4	33.0	CH	1.51,m	3b,5a,15	2a,3a,3b,5a,5b,1b	3a,15
5	43.1	CH ₃	a,0.58(1H,ddd, $J=12.0,12.0,12.0$) b,1.91(1H,m)	4,5b,6 5a	3a,4,7,15	7,9,15 5a,6,11
6	40.8	CH	1.78(1H,m)	5a,7	2,5a,5b,7,8a,8b,11	11,12,13
7	46.8	CH	0.99(1H,m)	6,8b	5a,5b,8a,8b,9,11,12,13	5a,6,8a,8b,9
8	21.3	CH ₂	a, 1.55 (1H, $J = 3.5$), d b,1.16(1H,m)	8b,9 7,8a,9	7,11	7,8b,9,13 3a,7,9,12
9	31.8	CH ₂	1.90(2H,brs)	8a,8b	7,8a,8b,14	3a,5a,6,7,8b,14,15
10	124.2	C	—		2a,8b,9,14	
11	27.5	CH	1.84(1H,m)	12,13	7,8a,12,13	6,12,13
12	17.0	CH ₃	0.78(3H,d, $J=6.5$)	11	11,13	6,8b,11,13
13	22.0	CH ₃	0.94(3H,d, $J=6.5$)	11	7,11,12	6,8a,11,12,15
14	19.1	CH ₃	1.63(3H,brs)			2a,3b,6,9
15	22.5	CH ₃	0.87(3H,d, $J=6.0$)	4	3a,5a,5b	3a,4,5a,9

紫外光谱可用于判断齐墩果酸三萜化合物结构中双键类型。质谱在三萜类物质的结构确定中起着极为重要的作用,不仅可以提供分子离子峰信号,还可以根据裂解碎片的信号确定取代基的位置及类型。在核磁共振氢谱中可获得甲基质子、连氧的碳上质子、烯质子及糖的端基质子信号等重要信息。一般甲基质子信号在 $\delta=0.625\sim1.50$ 间。在 $^1\text{H-NMR}$ 谱的高场出现多个甲基单峰是三萜类化合物的最大特征。三萜化合物常有OH取代,连OH的碳上质子信号一般出现在 $\delta=3.2\sim4$ 。 $^{13}\text{C-NMR}$ 是确定三萜及其皂苷结构最有用的技术,由于分辨率高,一个三萜或皂苷的 $^{13}\text{C-NMR}$ 谱几乎可给出其中每一个碳原子的信号。在 $^{13}\text{C-NMR}$ 谱中角甲基一般出现在 $\delta=8.9\sim33.7$,其中 23-CH_3 和 29-CH_3 为 e 键甲基,出现在低场,化学位移依次为 $\delta=28$ 和 $\delta=33$ 左右。烯碳在 $\delta=109\sim160$,羰基碳为 $\delta=170\sim220$ 。

下面是一个新的三羟基齐墩果烷型三萜的结构鉴定。

木通为木通科(Lardizabalaceae)木通属植物木通 *Akebia quinata* (Thunb.) Decne. 的干燥藤茎。味苦性寒,具有清热利尿、活血通脉之功效,主治小便短赤、淋浊、水肿、胸中烦热、风湿痹痛、乳汁不通等症。木通中主要含有三萜和三萜皂苷类成分。从木通的80%乙醇提取物的正丁醇部分分离鉴定出一个新的三萜类化合物(4-2)。

$^1\text{H-NMR}$ (500MHz, CD_3OD) δ : 0.74, 0.86, 0.87, 0.97, 1.15 (each 3H, s) 为与季碳相连的甲基质子信号; δ : 3.13 (2H, s), 3.33 (1H, d, $J=11.5\text{Hz}$), 3.62 (1H, d, $J=11.5\text{Hz}$), 3.70 (1H, s) 为5个与羟基相连的质子; δ : 5.20 (1H, t, $J=3.0\text{Hz}$) 为烯氢质子。 $^{13}\text{C-NMR}$ (125MHz, CD_3OD) 给出28个碳信号,结合HMQC谱,显示有1个重叠峰、1个溶剂掩盖峰,共30个碳信号,其中羰基碳信号(δ : 180.8)和双键碳信号(δ : 145.1, 123.8)为齐墩果烷类三萜化合物(表4-2)。

表 4-2 化合物 4-2 的 NMR 数据 (500MHz, CD_3OD , JHz)

C	δ_{C}	δ_{H}	HMBC	C	δ_{C}	δ_{H}	HMBC
1	34.2	1.26(1H,m) 1.33(1H,m)	C10,C25	9	48.8	1.70(1H,m)	C1,C5,C8, C10,C14,C26
2	26.1	1.48(1H,m) 1.87(1H,m)	C10	10	38.0		
3	71.3	3.70(1H,brs)	C1,C4,C5	11	24.6	1.85(2H,m)	C8,C9,C12,C13
4	44.0			12	123.8	5.20(1H,t,J=3.0)	C9,C14
5	50.6	1.33(1H,m)	C3,C4,C6	13	145.1	—	—
6	19.5	1.33(1H,m) 1.48(1H,m)	C4	14	42.9		
7	34.3	1.23(1H,m) 1.46(1H,m)	C5	15	28.8	1.02(1H,m) 1.69(1H,m)	C27
8	40.7	—	C8,C26	16	24.1	1.54(1H,m) 1.96(1H,m)	C15,C28
			—	17	48.0		

续表

C	δ_c	δ_H	HMBC	C	δ_c	δ_H	HMBC
18	42.0	2.81(1H, dd, $J=4.0, 14.5$)	C12, C13, C14, C16, C17	24	66.3	3.33(1H, d, $J=11.5$) 3.62(1H, d, $J=11.5$)	C3, C4, C23 C23
19	41.4	1.02(1H, brs, 13.5); 1.75(1H, m)	C18, C20, C30	25	16.2	0.83(3H, s)	C1, C5, C9, C10
20	36.9			26	17.6	0.74(3H, s)	C7, C8, C9, C14
21	29.3	1.10(1H, m) 1.44(1H, m)	C17	27	26.5	1.15(3H, s)	C8, C13, C14, C15
22	33.1	1.56(1H, m) 1.71(1H, m)	C18, C20	28	181.8	—	—
23	22.8	0.97(3H, s)	C3, C4, C5, C24	29	74.4	3.13(2H, s)	C19, C20, C21, C30
				30	19.5	0.87(3H, s)	C19, C20, C21, C29

HMQC 谱显示, δ : 3.33, 3.62 为同碳 66.3 (C24) 上的质子, 这两个质子在 HMBC 谱上分别与 C23 (δ : 22.8)、C3(δ : 71.3)、C4(δ : 44.0) 远程相关, 表明 CH_2OH 连在 4 位, 且根据甲基 (δ : 22.8) 可判断 4 位含氧取代在 α 键, 即 24 位连有羟基, 为 β 型。

HMBC 谱显示, δ : 0.87 (3H, s, H30) 甲基质子分别与 C19 (δ : 41.4)、C20 (δ : 36.9)、C21 (δ : 29.3)、C29 (δ : 74.4) 远程相关; δ : 3.13 (2H, s, H29) 的质子分别与 C19 (δ : 41.4)、C20 (δ : 36.9)、C21 (δ : 29.3)、C30 (δ : 19.5) 远程相关。化学位移值表明 29 位 (e 键) 有羟基取代, 即 α 型 (图 4-3)。

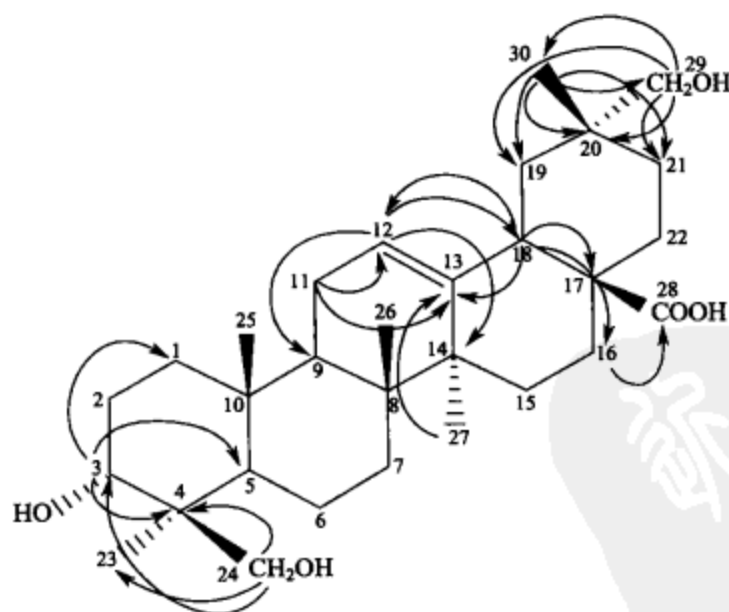


图 4-3 化合物 4-2 的 HMBC

由 ^1H , ^1H -COSY 谱显示邻偶相关关系为: C1 (δ : 1.26) 与 C2 (δ : 1.87), C2 (δ : 1.48) 与 C3 (δ : 3.70), C5 (δ : 1.33) 与 C6 (δ : 1.48), C6 (δ : 1.48) 与 C7 (δ : 1.23), C9 (δ : 1.70) 与 C11 (δ : 1.85), C11 (δ : 1.85) 与 C12 (δ :

5.20), C15 (δ : 1.02) 与 C16 (δ : 1.96), C18 (δ : 2.81) 与 C19 (δ : 1.75), C21 (δ : 1.10) 与 C22 (δ : 1.71) 分别相关。由此可推断化合物 **4-2** 的结构为 3 α ,24,29-三羟基齐墩果烷-12-烯-28-酸 (图 4-4)。

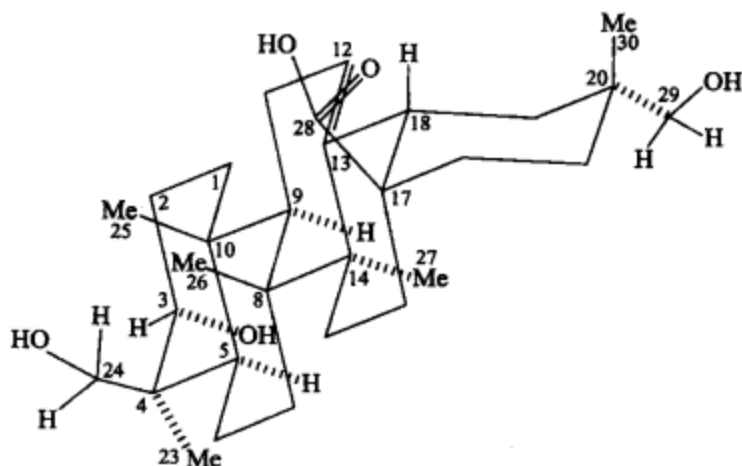


图 4-4 化合物 **4-2** 的立体构型

4.2.2 黄酮类化合物

黄酮类化合物是在植物中分布非常广泛的一类天然产物，其在植物体内大部分与糖结合成苷类，小部分以游离态的形式存在。绝大多数的植物体内都含有黄酮类化合物，黄酮类物质在植物的生长、发育、开花、结果以及防菌防病等方面起着重要的作用。

黄酮类化合物主要是指基本母核为 2-苯基色原酮类化合物，常指具有 2 个酚羟基的苯环通过中央三碳原子相互连接而成的一系列化合物。天然黄酮类化合物多为基本母核结构的衍生物，其常见的取代基有—OH、—OCH₃、—OCH₂O—、异戊烯基和咖啡酰基等。

对于已知黄酮类化合物，一般通过化学方法和波谱解析 (IR, UV, NMR 和 MS 等) 即可鉴定其结构。如果有已知物的标准品，经过 TLC (薄层色谱) 对照、混合熔点测定、红外光谱对照等更容易确定结构。即使是未知的黄酮类化合物，利用当前的波谱技术及一些辅助的化学方法，也很容易鉴定其结构，对于结构特殊的黄酮类化合物，则可通过全合成得到佐证。

质谱 (MS) 技术，尤其场解析质谱 (FD-MS) 与快速原子轰击质谱 (FAB-MS) 及串联质谱 (MS-MS) 的出现与应用，使其成为黄酮类化合物结构鉴定的重要手段之一。质谱技术只需要有微量的样品就可以获得有关整个分子结构及其主要碎片结构的重要信息。

紫外及可见分光光度法是鉴定黄酮类化合物的一种重要手段。常加入诊断试剂，以辅助确定黄酮及黄酮苷类化合物的结构。

圆二色谱 (CD) 和旋光谱 (ORD) 的经验规律可用于推定双氢黄酮类和双氢

黄酮醇类 C 环中手性碳的绝对构型。

目前利用单晶 X 射线衍射可以得到数千个独立衍射和可观察数据点，从而获取化合物分子中各个原子的坐标，确定化合物中各个原子的空间相对位置，最终为确定化合物的分子结构提供直观可靠的立体图形。

豆科甘草属植物刺果甘草为多年生草本，生于湿草地、河岸湿地及河谷坡地。种子可榨油。主要化学成分为黄酮类物质。农杆菌属根花菌 pRi 15834 诱导刺果甘草植物体产生化学成分分析，已经得到冠瘿碱、农杆菌碱、甘露碱等化学成分，从其根培养幼体的甲醇提取物中分离纯化得到两个新的异黄酮类物质（4-3 和 4-4）。

根据 EI-MS 和 NMR 数据可推定化合物 4-3 的分子式为 $C_{20}H_{16}O_5$ 。在 1H -NMR 谱中存在 δ : 8.17(s) 的峰信号，为异黄酮烯质子的特征吸收；AAXX 类型的芳香质子信号峰 δ : 7.30 (2H, d, $J=8.5\text{Hz}$) 和 7.75 (2H, d, $J=8.5\text{Hz}$)，为异黄酮 B 环质子吸收； δ : 6.55 (1H, s) 的峰信号，为 A 环芳香质子的吸收； δ : 1.75 (3H, brs, H_3-8'')，3.07 (1H, dd, $J=15.0, 7.5\text{Hz}$, H_a-4'')，3.39 (1H, dd, $J=15.0, 9.5\text{Hz}$, H_b-4'')，4.95 (1H, brs, H_a-7'')，5.17 (1H, brs, H_b-7'')，5.40 (1H, dd, $J=9.5, 7.5\text{Hz}$, $H-5''$) 的特征吸收峰指示异丙烯基二氢呋喃基团的存在。在 HMBC 试验中，二氢呋喃基团与 C-6 连接（图 4-5），圆二色谱进一步确定 C-5'' 绝对构型，在 470nm 处出现 Cotton 效应，指示 C-5'' 手性碳为 S 构型。结合 ^{13}C -NMR（表 4-3）及 H,H-COSY 确定化合物 4-3 的结构（图 4-6）。

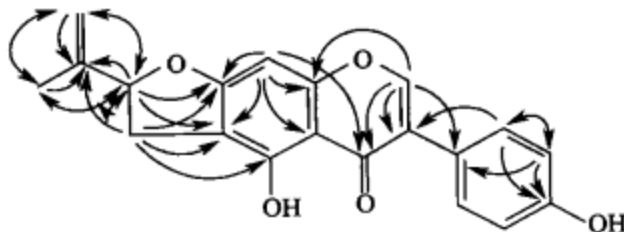


图 4-5 化合物 4-3 的 HMBC 的相关性

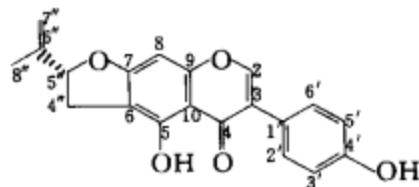


图 4-6 化合物 4-3 的结构

表 4-3 化合物 4-3 的 NMR 数据 (400MHz, pyridine- d_5 , J/Hz)

C	δ_H	δ_C	C	δ_H	δ_C
2	8.17(1H, s)	153.6	3'	7.30(2H, d, $J=8.5$)	116.5
3		123.7	4'		159.5
4		181.6	5'	7.30(2H, d, $J=8.5$)	116.5
5		157.9	6'	7.75(2H, d, $J=8.5$)	131.2
6		109.1	4''	3.07(1H, dd, $J=15.0, 7.5$)	31.0
7	6.55(1H, s)	166.7		3.39(1H, dd, $J=15.0, 9.5$)	88.5
8		89.3	5''	5.40(1H, dd, $J=9.5, 7.5$)	144.0
9		158.7	6''	4.95(1H, brs)	112.7
10		107.1	7''		
1'	7.75(2H, d, $J=8.5$)	122.3	8''	5.17(1H, brs)	17.2
2'		131.2		1.75(3H, brs)	

根据 HR-FAB-MS 和 NMR 数据可推定化合物 4-4 的分子式为 $C_{22}H_{20}O_{10}$ 。红外图谱在 1731cm^{-1} 吸收峰, 说明结构中有 δ -不饱和内酯基团。在 ^1H -NMR 谱中, 有 2 组 ABX 型芳香质子信号峰 [δ : 7.17 (1H, dd, $J=8.5, 2.0\text{Hz}$), 7.36 (1H, d, $J=2.0\text{Hz}$), 8.11 (1H, d, $J=8.5\text{Hz}$) 和 δ : 7.38 (1H, dd, $J=8.5, 2.0\text{Hz}$), 7.48 (1H, d, $J=2.0\text{Hz}$), 7.88 (1H, d, $J=8.5\text{Hz}$)], 1 个甲氧基质子信号 [δ : 3.79 (s)] 和 1 个糖端基质子信号 [δ : 5.81 (1H, d, $J=7.0\text{Hz}$)]. 根据糖质子的偶合常数可推测存在 β -D-吡喃葡萄糖。

根据 HMQC 和 HMBC 试验 (图 4-7), 完全归属 A 环和 B 环芳香质子之后, 还有 3 个芳香质子 (δ : 104.3, 158.2, 160.0) 在苷元中, 表明可能是拟雌内酯类物质。通过 HMBC 试验进一步确定甲氧基与 C-9 连接, 糖与 C-3 连接。结合 ^{13}C -NMR (表 4-4) 确定化合物 4-4 的结构 (图 4-8)。

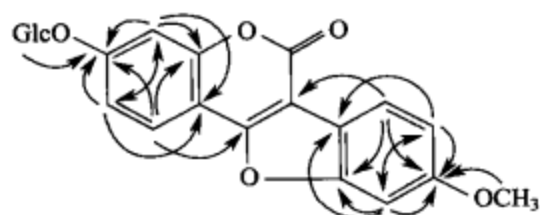


图 4-7 化合物 4-4 的 HMBC 的相关性

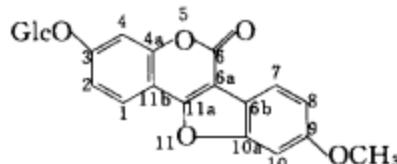


图 4-8 化合物 4-4 的结构

表 4-4 化合物 4-4 的 NMR 数据 (400MHz, C_5D_5N , J Hz)

C	δ_H	δ_C	C	δ_H	δ_C
1	7.88(1H,d, $J=8.5$)	121.8	7	8.11(1H,d, $J=8.5$)	123.0
2	7.38(1H,dd, $J=8.5,2.0$)	114.7	8	7.17(1H,dd, $J=8.5,2.0$)	114.1
3		161.4	9		160.0
4	7.48(1H,d, $J=2.0$)	105.2	10	7.36(1H,d, $J=2.0$)	97.6
4a		155.3	10a		157.0
5			11		
6		158.2	11a		160.0
6a		104.3	11b		107.4
6b		117.1	OCH ₃	3.79(3H, s)	56.0

4.2.3 苯丙素化合物

苯丙素类是一类含有一个或几个 C_6-C_3 单位的天然成分。从生物合成途径来看, 苯丙素类多数是由莽草酸通过苯丙氨酸和酪氨酸, 再经过脱氨、羟基化、偶合等反应形成的最终产物。

4.2.3.1 香豆素类

香豆素类是一类具有芳香性的天然产物, 其母核是由顺式邻羟基桂皮酸形成的内酯, 称为苯并[α]吡喃酮。香豆素类可以以游离或成苷的形式存在于植物界, 分布于 30 种以上的科, 其中以双子叶植物中的伞形科、芸香科和桑科含量最多。

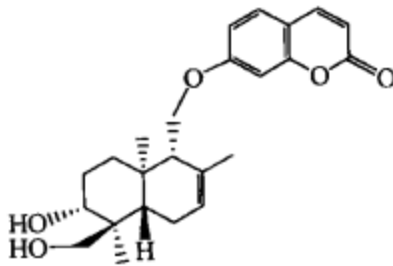
近年来,随着现代分离技术和波谱技术的应用和发展,发现了不少新类型的香豆素结构及一些罕见的结构,如香豆素硫酸酯、无含氧取代的3,4,7-三甲基香豆素和四氧取代的香豆素。

香豆素母核的紫外光谱主要有2个吸收带, $\lambda_{\max}(\lg\epsilon)$ 274nm(4.03)为苯环吸收,311nm(3.72)为吡喃环吸收,而且很容易与色酮相区别。香豆素在红外光谱中也有较为明显的吸收峰,主要是C—H伸缩振动、C=O伸缩振动、C=C骨架振动等类型的红外吸收。NMR谱在香豆素的结构鉴定中起着极为重要的作用。

下面是倍半萜型香豆素的结构鉴定。

伞形科阿魏属植物是亚欧大陆所独有的属,有130余种,广泛分布于地中海沿岸和中亚地区。该属的多数植物常作香料和药物制剂的辅料,植物树脂具有清热、通气等功效,有些植物在当地医疗上用于皮肤感染、亢奋等疾病的治疗。该属的化学成分主要为倍半萜及倍半萜香豆素。从*Ferula vesceritensis*的二氯甲烷-甲醇提

取物得到一个倍半萜香豆素(4-5)。



HR-FAB-MS (m/z 399.2165) 给出化合物4-5的分子式为 $C_{24}H_{30}O_5$,依据 1H -NMR和 ^{13}C -NMR数据(表4-5)确定结构(图4-9)。

在 1H -NMR谱中,有7-羟基香豆素质子共振信号 $[\delta: 6.18 (H-3), 7.60 (H-4), 7.31 (H-5), 6.77$

(H-6), 6.75 (H-8)];倍半萜部分质子共振信号有烯质子 $[\delta: 5.44 (H-7')]$,甲基质子 $[\delta: 1.60 (H-12')]$,6个质子的单峰信号 $[\delta: 0.87 (H-14', H-15')]$,连氧CH质子 $[\delta: 3.59 (H-3')]$,2个连氧 CH_2 质子 $[\delta: 3.94 (H-11'a), 4.10 (H-11'b), 3.30 (H-13'a), 3.56 (H-13'b)]$ 。

在 ^{13}C -NMR谱中,存在24个碳原子,其中有典型的7-羟基香豆素骨架的9个碳原子,剩余的15个碳原子为倍半萜部分。根据DEPT试验,化合物4-5有:3个甲基碳($\delta: 15.3, 11.2, 21.4$);3个亚甲基碳($\delta: 26.3, 23.2, 37.4$);2个乙醇碳($\delta: 66.8, 70.2$);7个次甲基碳 $[\delta: 113.2 (C-3), 143.6 (C-4), 128.7 (C-5), 112.2 (C-6), 101.3 (C-8), 43.4 (C-5'), 53.5 (C-9')]$,其中5个为7-羟基香豆素骨架上的碳。

在 1H -NMR有3个甲基信号 $[\delta: 1.60 (H-12'), 0.87 (H-14'), 0.87 (H-15')]$,另外化学位移值表明可能存在连接有羟基化的亚甲基($\delta: 3.30, 3.56$)。通过HMQC和HMBC试验,确定亚甲基的位置。在HMBC试验中,质子H-13'a($\delta: 3.56$)、H-13'b($\delta: 3.30$)与碳C-3'($\delta: 75.8$)、C-5'($\delta: 43.4$)、C-14'($\delta: 11.2$)有明显的相关性,而碳共振C-13'($\delta: 70.2$)与质子H-30($\delta: 3.60$)相关;另外还能观察到一个重要的相关,即碳原子C-1'($\delta: 37.4$)、C-9'($\delta: 53.5$)、C-10'($\delta: 35.7$)与质子H-15'($\delta: 0.87$)相关;碳原子C-7'($\delta:$

123.2)、C-8' (δ : 132.3)、C-9' (δ : 53.5) 与质子 H-12' (δ : 1.60) 间也存在相关; 碳原子 C-12' (δ : 21.4) 与质子 H-9' (δ : 2.01)、H-7' (δ : 5.44) 也有相关, 因此连接羟基的亚甲基应与 C-4' 连接 (图 4-10)。

在 NOE 谱中能观察到 H-3' 和 H-13'、H-5' 和 H-9'、H-5' 和 H-13'、H-11' 和 H-15'、H-14' 和 H-15 之间存在相关, 所以可以推定 H-3'、H-5'、H-9'、H-13' 为 β 构型, H-11'、H-14'、H-15' 为 α 构型。因此化合物 4-5 信号被完全归属确定结构。

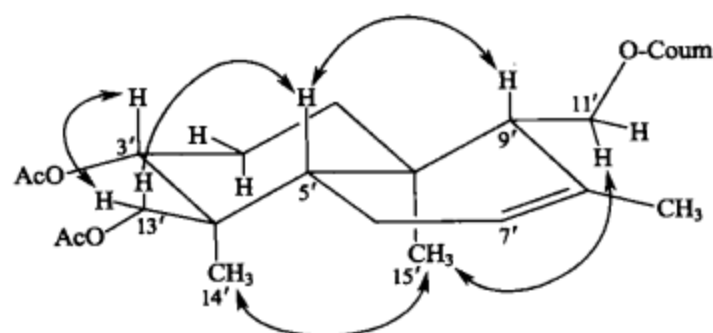


图 4-10 化合物 4-5 的 NOE

表 4-5 化合物 4-5 的 NMR 数据 (500MHz, CDCl_3 , J/Hz)

C	δ_{H}	δ_{C}	C	δ_{H}	δ_{C}
2		161.9	4'		41.8
3	6.18(1H,d,J=9.5)	113.2	5'	1.36(1H,dd,J=12.5,5.0)	43.4
4	7.60(1H,d,J=9.5)	143.6	6'	1.98(1H,m);1.81(1H,m)	23.2
5	7.31(1H,d,J=8.5)	128.7	7'	5.44(1H,brs)	123.2
6	6.77(1H,dd,J=8.5,2.5)	112.2	8'	—	132.3
7		161.6	9'	2.01(1H,dd,J=5.5,3.5)	53.5
8	6.75(1H,d,J=2.5)	101.3	10'		35.7
9		155.7	11'	4.10(1H,dd,J=10.0,3.5)	66.8
10		112.7		3.94(1H,dd,J=10.0,3.5)	
1'	1.95m	37.4	12'	1.60(3H,brs)	21.4
	1.24m		13'	3.56(1H,d,J=10.5)	70.2
2'	1.54(1H,m)	26.3		3.30(1H,d,J=10.5)	
	1.72(1H,m)		14'	0.87(3H,s)	11.2
3'	3.60(1H,dd,J=11.0,7.0)	75.8	15'	0.87(3H,s)	15.3

4.2.3.2 木脂素类

木脂素类是一类由苯丙素氧化聚合而成的天然产物, 通常所指的是二聚体, 少数为三聚体和四聚体等。木脂素类的苯丙素组成单元也可分成多种类型, 主要包括肉桂醇、肉桂酸、丙烯基酚等。木脂素类两个苯环上常有氧取代基, 并以 C₄、C₃、C₄ 或 C₃、C₄、C₅ 取代较多。侧链上含氧官能团也多有变化, 尤其是侧链末端碳原子上的含氧官能团, 如羟基、羧基、羰基等, 这些含氧基团相互脱水缩合, 又出现四氢呋喃、半缩醛、内酯等环状结构。这使木脂素类成分结构类型繁多。

近年来,木脂素类化合物的研究比较活跃,发现了许多具有重要生物活性的成分。从化学结构类型来看,木脂素并不是一类成分,而是具有 10 多种母核结构。

多数木脂素类化合物借助于波谱方法就能确定其化学结构,包括其立体结构。其中紫外光谱主要测定苯环的存在状态。多数木脂素的两个取代苯环是两个孤立的发色团,其 UV 吸收峰位置相似,吸收强度是二者的总和。木脂体一般呈现酚衍生物的吸收。木脂素的红外光谱,不仅在 $1500\sim 1600\text{cm}^{-1}$ 显示芳环吸收,还可以依据羰基的吸收位置确定木脂素类型。氢谱核磁共振、碳谱核磁共振以及它们的二维谱技术是测定木脂素类化合物结构的主要依据。

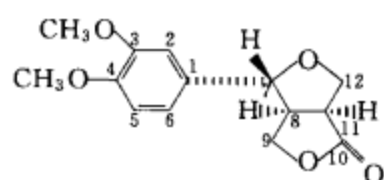


图 4-11 化合物 4-6
的绝对构型

下面是木脂素类化合物结构鉴定实例。

连翘为木犀科 (Oleaceae) 连翘属植物,分布于山西、河南、陕西、山东等地,《中华人民共和国药典》(1995 版一部) 收载作药用的为 *Forsythia suspense* (Thun. b) Vahl 的果实。连翘具有清热解毒、散结消肿之功效,用于痈疽、淋病、乳痛、丹毒、风热感冒、瘟病初起、温热入营、高热烦渴、神昏发斑、热淋尿闭等症。连翘属植物所含化学成分种类较多,包括苯乙醇咖啡酸苷类、 C_6-C_2 天然醇及其苷类、木脂素类、萜类、黄酮类等,其成分具有抗菌、抗炎等多种生理活性。对从连翘的 70% 乙醇提取物中分离得到的木脂素化学成分 (4-6) 进行结构鉴定 (图 4-11)。

EI-MS 给出分子离子峰为 m/z 264 $[\text{M}]^+$ 。元素分析给出分子式为 $\text{C}_{14}\text{H}_{16}\text{O}_5$ 。红外光谱在 1774cm^{-1} 处的吸收峰提示分子中含有五元内酯环结构, 1518cm^{-1} 和 1468cm^{-1} 处的吸收峰说明有苯环存在 (图 4-12)。

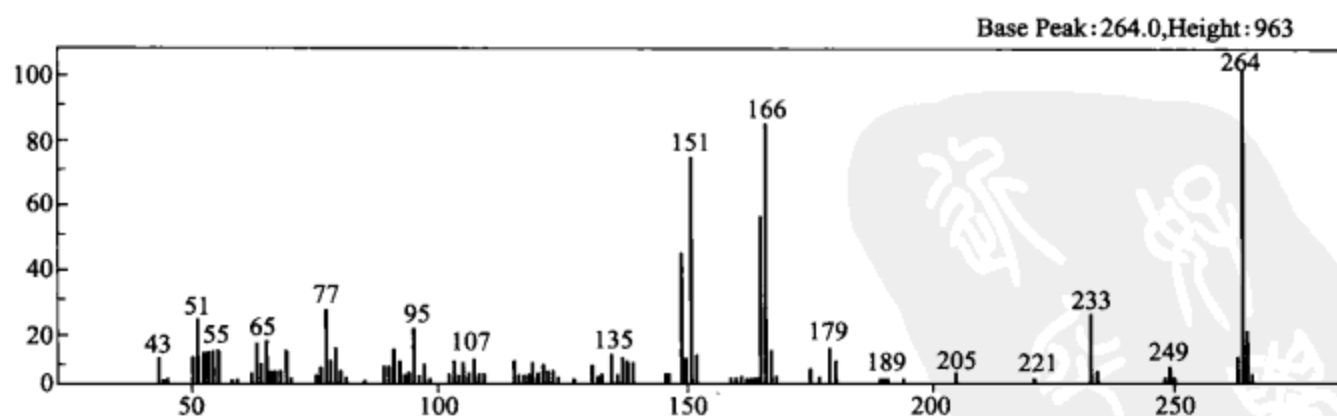


图 4-12 化合物 4-6 的 FAB-MS 谱

$^1\text{H-NMR}$ (表 4-6, 图 4-13) 存在 3 个芳香环质子信号 [δ : 6.80 (1H, d, $J=1.6\text{Hz}$), 6.75 (1H, dd, $J=8.2, 1.6\text{Hz}$), 6.87 (1H, d, $J=8.2\text{Hz}$)], 推测结构中可能含有 1,3,4-三取代苯环。 $^{13}\text{C-NMR}$ (表 4-6, 图 4-14) 和 DEPT (图 4-15) 显示分子中有 2 个 OCH_3 、2 个连氧 CH_2 、6 个 CH (其中 1 个连氧)、4 个

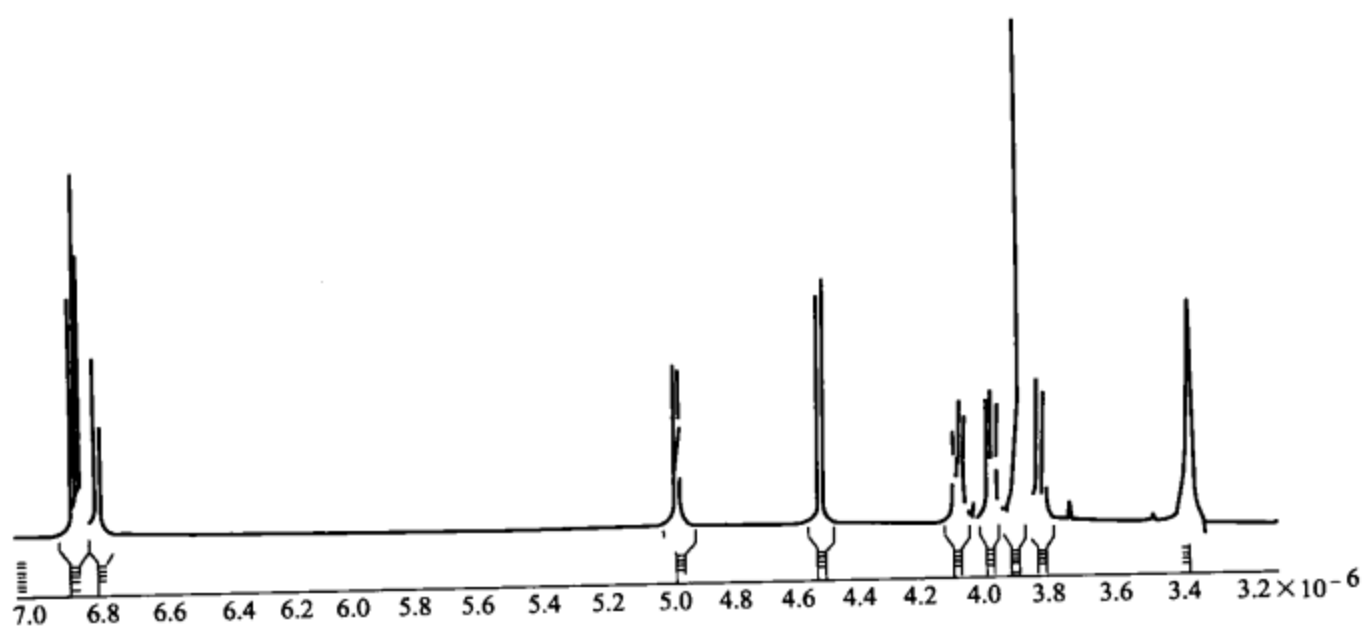


图 4-13 化合物 4-6 的 ^1H -NMR 谱

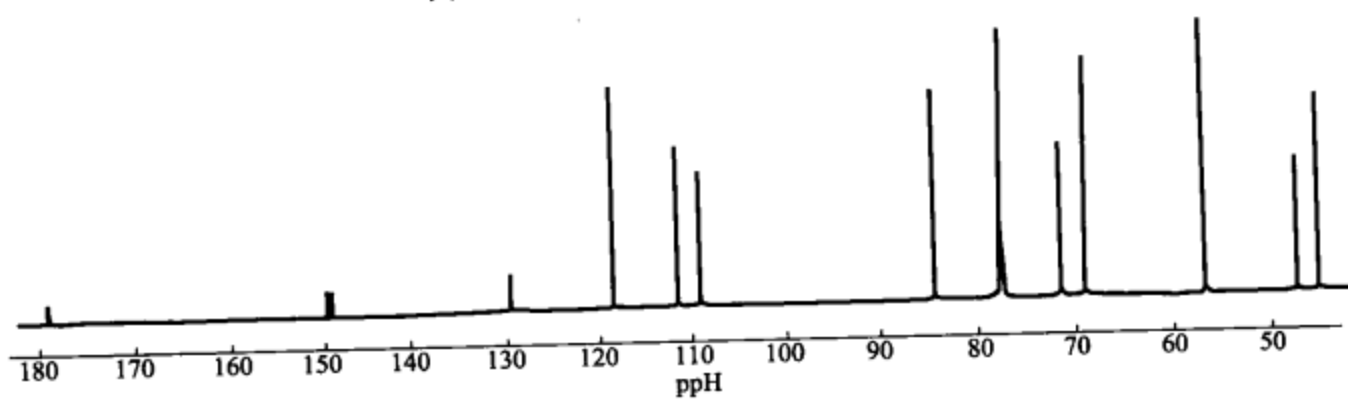


图 4-14 化合物 4-6 的 ^{13}C -NMR 谱

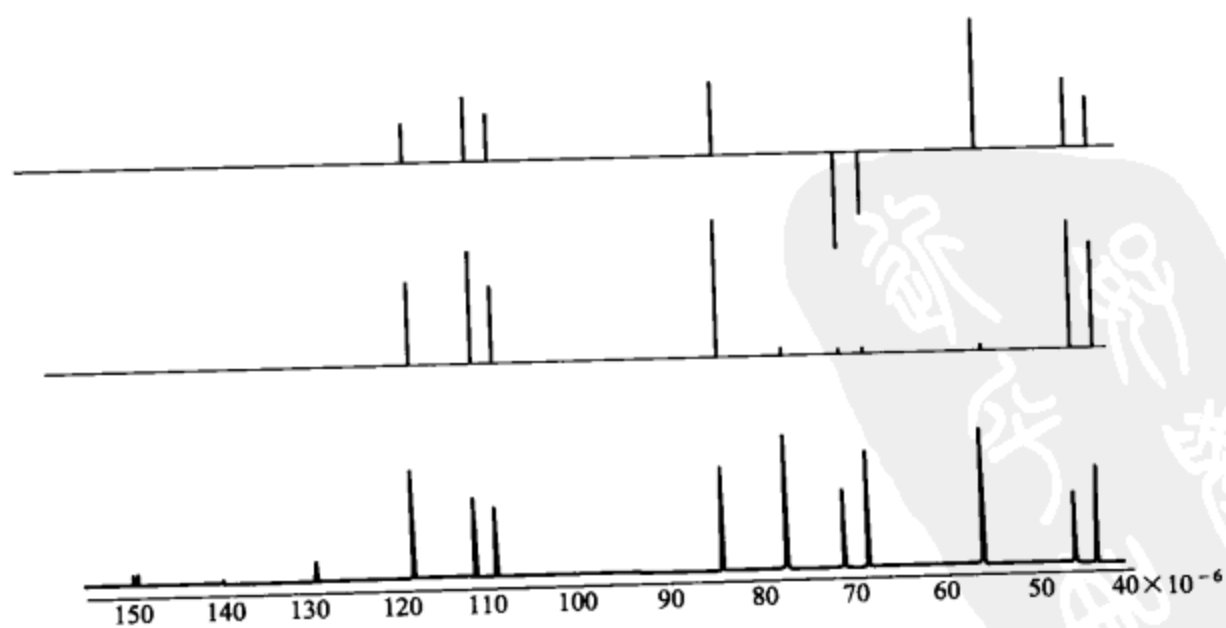


图 4-15 化合物 4-6 的 ^{13}C -NMR 谱 (DEPT)

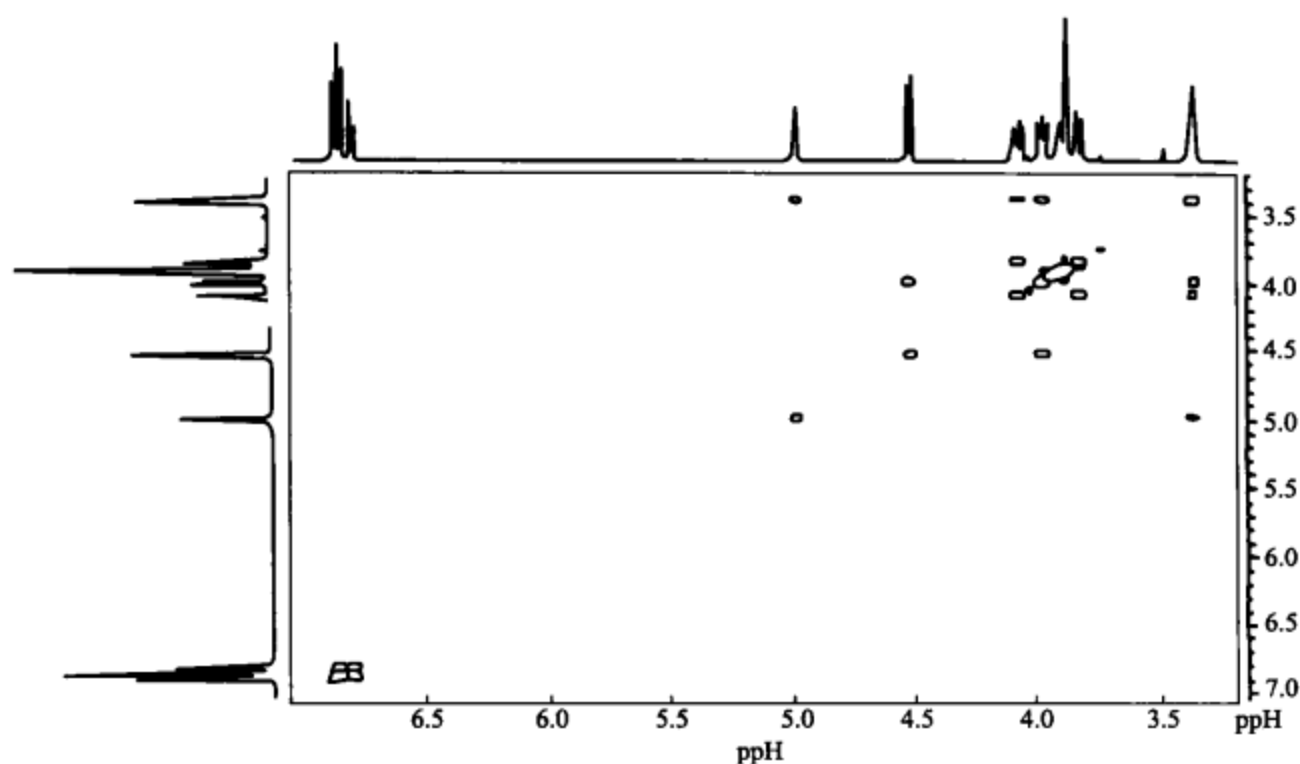


图 4-16 化合物 4-6 的 ^1H , ^1H -COSY 谱

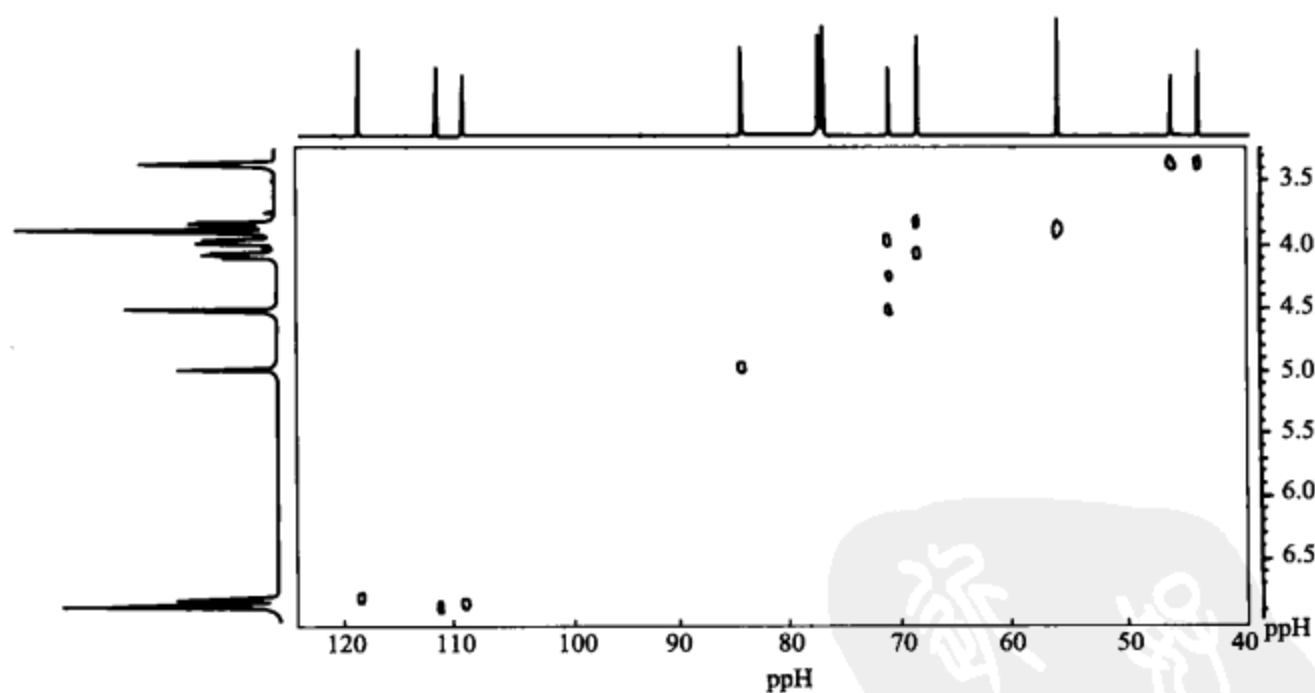


图 4-17 化合物 4-6 的 ^{13}C , ^1H -COSY 谱

季碳，其中 C-1 (δ : 129.0)、C-2 (δ : 109.0)、C-3 (δ : 149.2)、C-4 (δ : 148.6)、C-5 (δ : 111.2)、C-6 (δ : 118.1) 处的碳信号进一步证实了 1,3,4-三取代苯环的存在。 ^1H , ^1H -COSY 谱 (图 4-16) 显示 δ : 4.98 (1H, d, $J=5.4\text{Hz}$, H-7), 3.37 (1H, m, H-8), 3.95 (1H, dd, $J=9.5, 6.8\text{Hz}$, H-9a), 4.51 (1H, d, $J=9.5\text{Hz}$, H-9b) 处的信号依次相关; δ : 3.37 (1H, m, H-8), 3.36

(1H, m, H-11), 4.08 (1H, dd, $J = 8.2, 9.9\text{Hz}$, H-12a), 3.83 (1H, dd, $J = 9.9, 4.0\text{Hz}$, H-12b) 处的信号也依次相关。 ^{13}C , ^1H -COSY 谱 (图 4-17) 上与其相对应碳信号分别为 δ : 84.0, 43.6, 70.9, 4.60, 68.3。联系 ^{13}C -NMR 在 δ : 178.6 处的羰基信号, 推测分子中具有双并四氢呋喃环单内酯结构片段。由上推断化合物 4-6 的可能结构有几种 (图 4-18)。

用上述波谱数据难于准确判断化合物的结构, 通过 X 射线单晶衍射试验最后确定该化合物的结构。该化合物分子式为 $\text{C}_{14}\text{H}_{16}\text{O}_5$, 晶体属正交晶系, 空间群 $P2_12_12_1$, 每个不对称单位包含 1 分子 $\text{C}_{14}\text{H}_{16}\text{O}_5$, 晶胞内的分子数 $Z=4$ 。晶胞参数: $a=0.6663(2)\text{nm}$, $b=0.9126(3)\text{nm}$, $c=1.9213(6)\text{nm}$, $\alpha=\beta=\gamma=90^\circ$, $v=1.25361(8)\text{nm}^3$ 。MAC science DIP-2000 型面检测仪收集衍射数据, 辐射为铝靶 $\text{MoK}\alpha$ (0.071073nm), 由

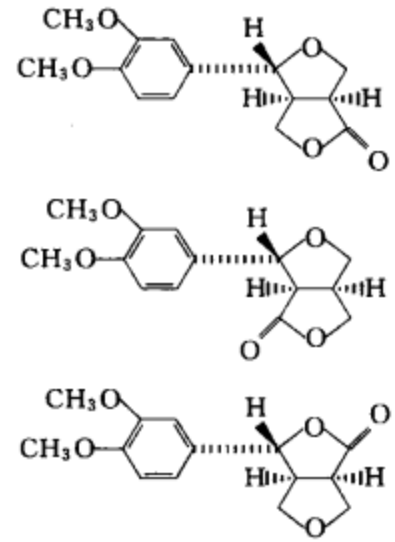


图 4-18 化合物 4-6 的可能结构有几种

石墨单色器滤波, 影像板到晶体距离为 100mm , 收集范围度 $0^\circ\sim 180^\circ$, 每隔 4° 收集一张衍射图, 共收集 45 个衍射面, 经 DNEZO 和 SCALEPACK 程序处理, $R_{\text{merge}}=0.032$, 数据完全度为 96.3%, 独立衍射点为 912, 其中可观察衍射点 906 [$|F_0| > 3\sigma(F_0)$], 各衍射指标的范围是 $0 < H < 8$, $0 < K < 11$, $-18 < L < 21$, 程序系统 CyrsnatG 解结构, 修正结构参数及确定氢原子的位置, 对 906 个可观察衍射的最终偏离因子 $R=0.0503$, $R_w=0.0493$ 。结构分析如图 4-19 所示。化合物的 CD 谱显示负性 Cotton 效应, 这是由于内酯环羰基的 $n\text{-}\pi^*$ 跃迁形成, 根据八区律规则

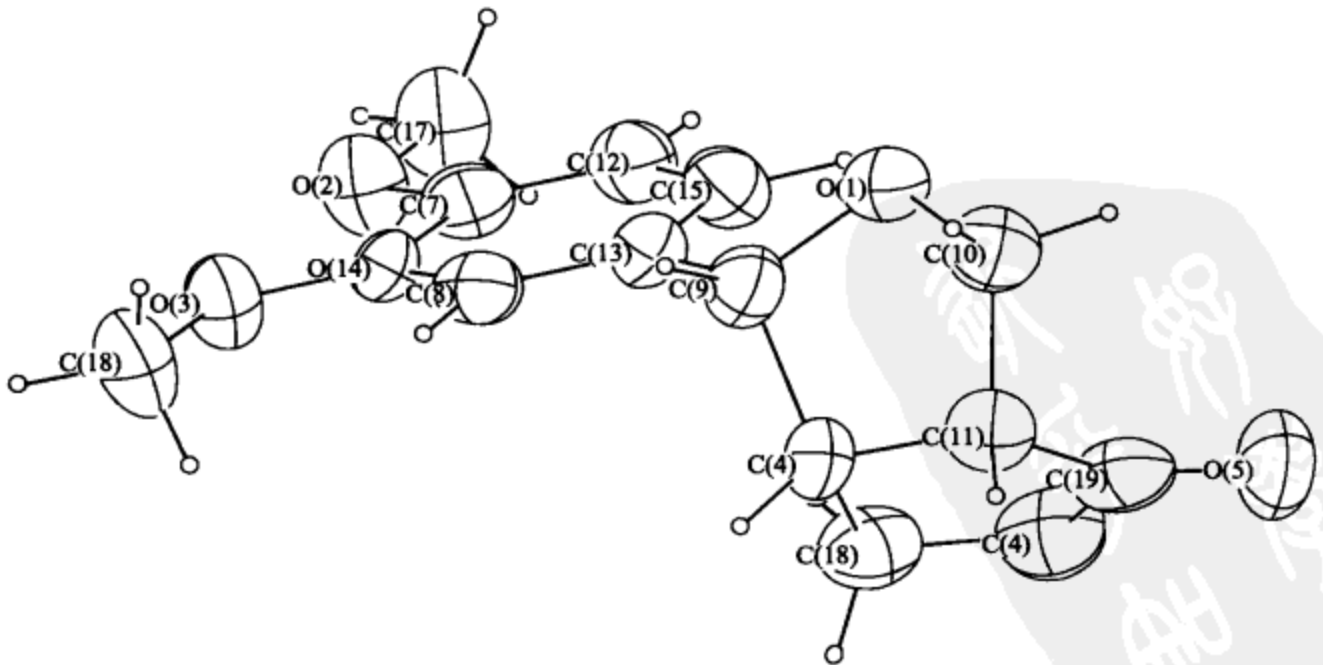


图 4-19 化合物 4-6 的 X 射线单晶衍射分析

推测双环氧连翘内酯的绝对构型（图 4-11）。

表 4-6 化合物 4-6 的 NMR 数据 (500MHz, CDCl₃, J/Hz)

C	δ_H	δ_C	C	δ_H	δ_C
1	—	129.0	9	4.51(1H,dd,J=9.5)	70.9
2	6.80(1H,d,J=1.6)	109.0		3.95(1H,dd,J=9.5,6.8)	
3	—	149.2	10	—	178.6
4	—	148.6	11	3.36(1H,m)	46.0
5	6.87(1H,d,J=8.2)	111.2		4.08(1H,dd,J=8.2,9.9)	
6	6.78(1H,dd,J=8.2,1.6)	118.1	12	3.83(1H,dd,J=9.9,4.0)	68.3
7	4.98(1H,d,J=5.4)	84.0	13	3.88(3H,s)	55.9
8	3.37(1H,m)	43.6	14	3.88(3H,s)	55.9

4.2.4 生物碱类化合物

生物碱一般是指植物中的含氮有机化合物（蛋白质、肽类、氨基酸及维生素 B 除外）。生物碱广泛分布于植物界，许多重要的植物药如鸦片、麻黄、金鸡纳、番木鳖、汉防己、莨菪、延胡索、苦参、洋金花、秋水仙、长春碱、三尖杉和乌头等主要含有生物碱成分。在许多情况下，含量最高的生物碱往往是主要的有效成分，但也有例外。如百部为止咳、杀虫中药，但百部生物碱只有杀虫作用，其止咳作用尚未见报道。又如乌头碱是乌头的主要成分，但它的强心止痛作用并不是乌头碱，而是含量极微的去甲乌头碱。长春新碱与美登木素也分别是长春花与美登木的微量成分。

由于同一植物中的生物碱往往来源于同一前体，因此它们的化学结构往往类似。同科同属中的生物碱也大多属于同一结构类型。因此研究生物碱结构时首先要鉴定植物品种并对同种同属植物中的生物碱情况有所了解，掌握有关它们的生源知识将有助于化合物的鉴定和结构推定。由于植物的科属是按形态分类的，而生物碱则是它们的内在成分，有时并不能仅仅依靠植物的科属预言它们所含的生物碱成分。例如，茜草科的吐根含吐根碱高达 2%，而同科的栀子、茜草则不含生物碱。

生物碱是天然产物中较大的一类化合物，各类生物碱的结构差别较大，给结构鉴定带来一定困难。生物碱的紫外光谱反映了基本骨架或分子中生色基团的结构特征，是生物碱结构测定的手段之一。红外光谱主要用于功能基的定性和与已知碱对照鉴定，主要有酮基 C=O 的伸缩振动吸收和 Bohlmann 吸收带。生物碱的质谱数据非常丰富，主要用于难于裂解或由取代基或侧链的裂解产生特征离子，主要裂解方式是以氮原子为中心的 α -裂解。

下面介绍生物碱结构鉴定实例。

茜草科乌檀属约 35 种植物，分布于热带亚洲、非洲和大洋洲，我国有海南、云南等地。片剂和注射剂用于治疗上呼吸道感染、急性扁桃体炎、支气管炎、泌尿系统感染等疾病。提取物具有抗菌和抗寄生物的作用。乌檀富含活性成分，已经分

离得到 40 多种吲哚类生物碱、黄酮类及酚类等化合物。进一步研究其化学成分，从茜草科植物乌檀的干燥植物体的甲醇提取物中分离得到一新的吲哚型生物碱(4-7) (图 4-20)。

HR-FAB-MS 给出 4-7 的分子式为 $C_{26}H_{26}N_2O_9$ 。NMR 谱 (表 4-7) 分析可知具有 4 个环 (环 A~D)。在 HMBC 谱中, 有 H-5 (δ : 4.51, 4.06) 与 C-22 (δ : 158.1)、H-14 (δ : 6.85) 与 C-16 (δ : 108.0)

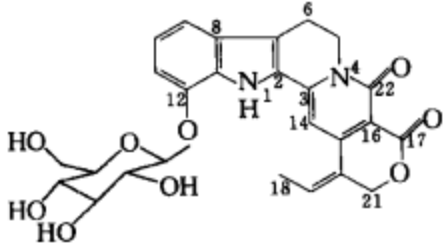


图 4-20 化合物 4-7 的结构

相关的存在, 推测 D 环可能是 δ -内酯环; 另外, 还有 H-14、H-18 (δ : 2.10) 与 C-20 (δ : 127.3), H-19 (δ : 6.33) 与 C-15 (δ : 150.2), H-21 (δ : 4.70) 与 C-17 (δ : 160.7)、C-15、C-20、C-19 (δ : 133.8) 的相关, 表明 D 环与 E 环相连。在 NOESY 谱中, 有 H-18 与 H-14、H-19 与 H-21 的相关, 可知 E 环具有一个 (E)-构型的三取代烯烃基团。

在 1H -NMR 谱中, 低场有 4 个芳香质子信号 (δ : 6.60~7.30), 1 个单峰 H-14 (δ : 6.85) 信号, 2 个双峰 H-9 (δ : 6.72) 和 H-11 (δ : 7.10) 信号, 1 个三重峰 H-10 (δ : 7.19) 信号。HMBC 谱中 H-5、H-9 与 C-7 (δ : 116.9), H-1 (δ : 12.09) 与 C-12 (δ : 152.9) 有相关性; C-9 (δ : 103.5) 与 C-10 (δ : 126.3)、C-10 与 C-11 (δ : 106.0) 相连接。在 ^{13}C -NMR 谱中, 存在着糖配基信号, 根据端基碳原子及偶合常数 (7.3Hz) 可判断为 β -葡萄糖。

在 HMBC 谱, 端基碳原子与 C-12 有相关, 可知糖部分与 C-12 相连。根据以上分析可知推测化合物 4-7 的结构。

表 4-7 化合物 4-7 的 NMR 数据 (400MHz, DMSO- d_6 , J/Hz)

C	δ_H	δ_C	C	δ_H	δ_C
1	12.09(1H,s)	—	15	—	150.2
2	—	125.9	16	—	108.0
3	—	142.2	17	—	160.7
5	4.51(1H,m) 4.06(1H,m)	40.3	18	2.10(3H,d,J=7.3)	15.3
6	3.51(1H,m) 3.20(1H,m)	20.4	19	6.33(1H,q,J=7.3)	133.8
7	—	116.9	20	—	127.3
8	—	116.3	21	4.70(2H,s)	71.2
9	6.72(1H,d,J=7.9)	103.5	22	—	158.1
10	7.19(1H,t,J=7.9)	126.3	1	5.01(1H,d,J=7.3)	100.4
11	7.10(1H,d,J=7.9)	106.0	2	3.23(1H,m)	69.7
12	—	152.9	3	3.30(1H,m)	76.7
13	—	140.5	4	3.33(overlapped)	73.4
14	6.85(1H,s)	97.0	5	3.32(overlapped)	77.1
			6	3.49(1H,m) 3.71(1H,m)	60.7

4.2.5 甾体类化合物

甾体化合物是天然广泛存在的一类化学成分，种类很多，但结构中都具有环戊烷并多氢菲的甾核。甾核的4个环可以有不同的稠合方式。甾核C-3位有羟基取代，可以与糖结合成苷。甾核的C-10和C-13位有角甲基取代，C-17位有侧链。根据侧链结构不同，天然甾类成分可分为许多类型。从生源观点来看，甾体化合物都是通过甲戊二羧酸的生物合成途径转化而来。

饱和甾体一般在200~400nm间无吸收，如果结构中引入孤立的双键、羰基、 α,β -不饱和酮基或共轭双键，则可产生吸收。含有螺缩酮结构侧链的甾体在红外光谱中几乎都显示 980cm^{-1} (A)、 920cm^{-1} (B)、 900cm^{-1} (C)和 860cm^{-1} (D)附近的4个特征吸收谱带，且A带最强。甾体由于分子中有螺甾烷侧链，其EI-MS谱中出现很强的 m/z 为139的基峰，中等强度的 m/z 为115的碎片离子峰及弱的 m/z 为126的辅助离子峰。

甾体在 $^1\text{H-NMR}$ 谱高场区有4个甲基（即18位甲基、19位甲基、21位甲基和27位甲基）质子的特征峰。其中18位甲基、19位甲基均为单峰，且前者处于较高场。21位甲基和27位甲基均为双峰，且后者处于较高场。如果25位有OH取代，则27位甲基成为单峰，并向低场移动。16位和26位上的氢为连氧碳上的质子，处于较低场。

下面是甾体化合物结构鉴定实例。

高等真菌已成为筛选药物先导化合物的重要生物资源，近年来的研究表明，高等真菌含有大量结构新颖和具有生物活性的次生代谢产物。大杯香菇（*Lentinus giganteus* Berk）为侧耳科香菇属真菌，为发生于炎热夏季的高温型菌类，可食用，分布于广东、海南、香港、福建和浙江等地。从大杯香菇子实体乙醇提取物的石油醚萃取部分分离得到过氧麦角甾醇类化合物（4-8），经波谱分析确定了它们的结构（图4-21）。

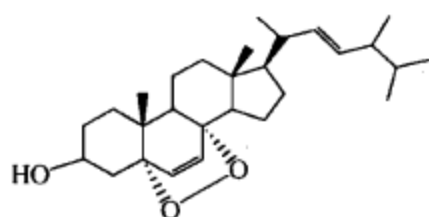


图4-21 化合物4-8的结构

IR光谱显示在 $\nu_{\max}\text{cm}^{-1}$: 3459(OH)、2958、2877、1655(C=C)、1456、1375、1050、1023、975、970、935、856有吸收峰信号。根据EI-MS谱，化合物（4-8）分子离子峰得出化合物的相对分子质量为428，结合 $^{13}\text{C-NMR}$ 和DEPT谱推出其分子式为 $\text{C}_{28}\text{H}_{44}\text{O}_3$ 。 $^{13}\text{C-NMR}$ (DEPT)谱显示有28个碳信号，其中有6个甲基信号、7个亚甲基信号、11个次甲基信号（1个为氧取代，另外有4个为双键碳信号）、4个季碳信号（2个为氧取代），初步推定该化合物为麦角甾醇类化合物（图4-22）。 $^1\text{H-NMR}$ (CDCl_3 , 500MHz) δ : 6.50(1H, d, $J=8.5\text{Hz}$, 7-H)、6.24(1H, d, $J=8.5\text{Hz}$, 6-H)、5.23(1H, dd, $J=7.2, 15.2\text{Hz}$, 23-H)、5.14(1H, d, $J=6.7\text{Hz}$, 22-H)、3.96(1H, tt, $J=5.0, 11.2\text{Hz}$, 3-H)、1.22

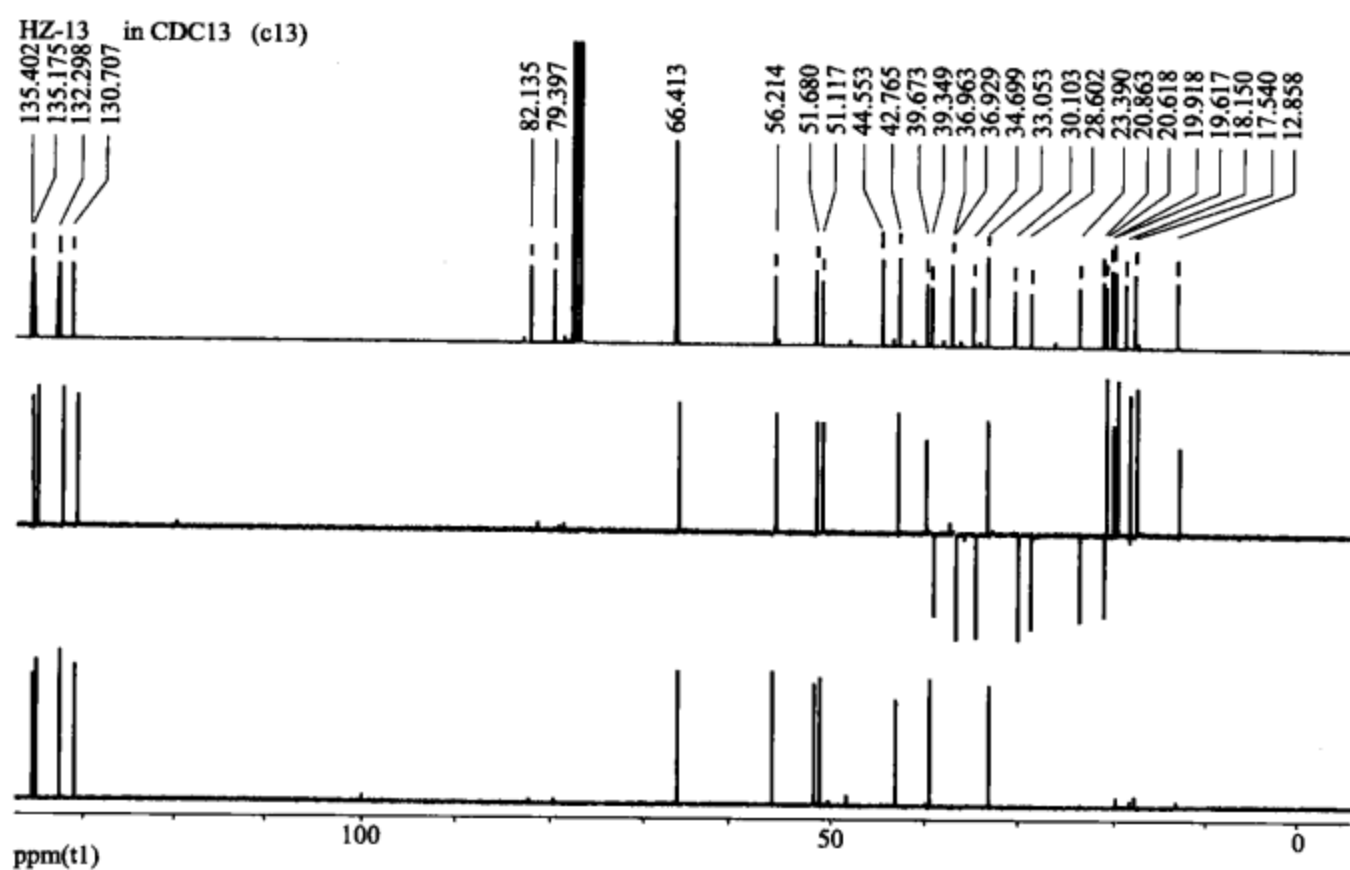
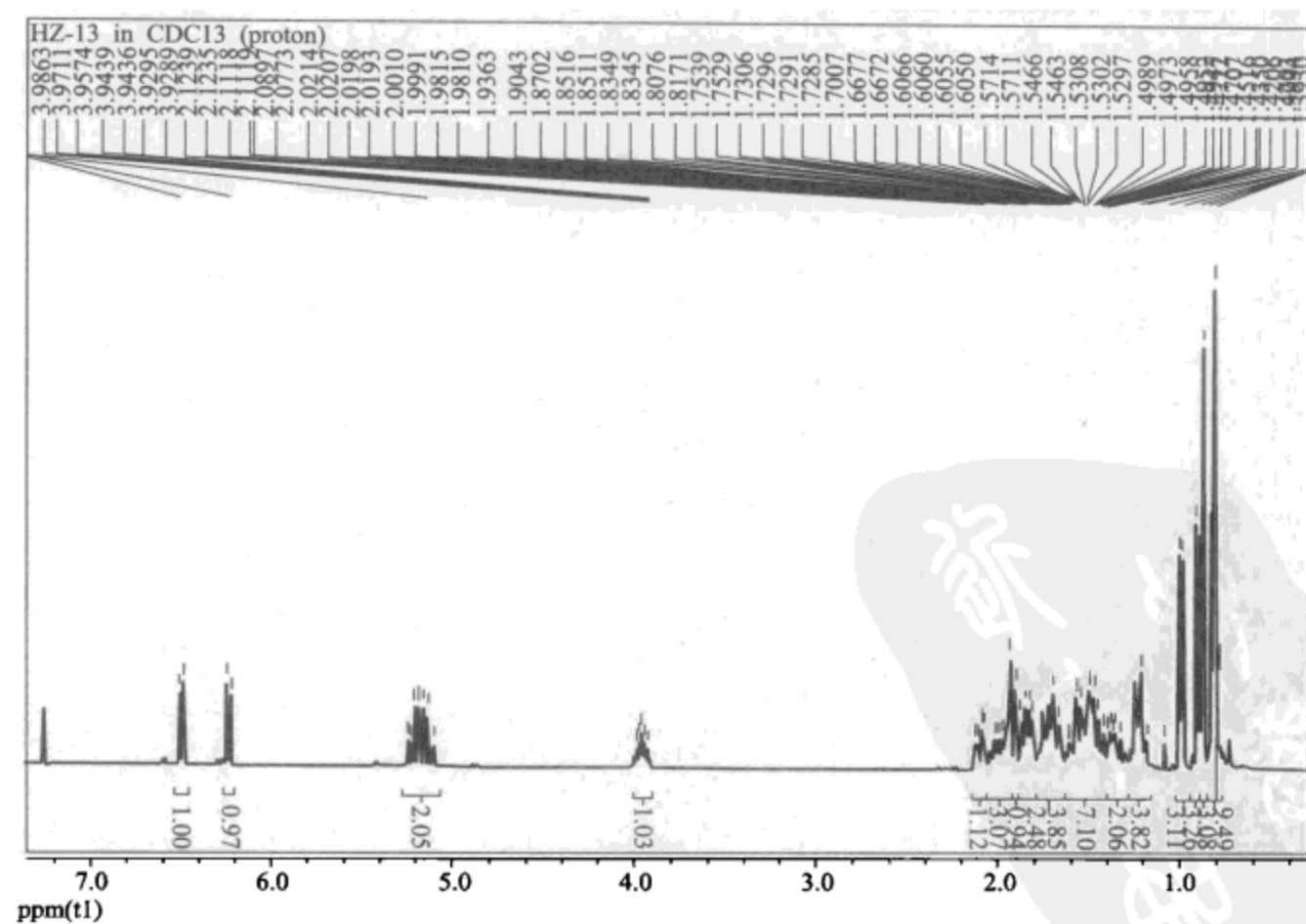


图 4-22 化合物 4-8 的¹³C-NMR 谱 (DEPT)



(3H, s, 19-H)、1.09 (3H, d, $J=6.7\text{Hz}$, 21-H)、0.99 (3H, d, $J=6.8\text{Hz}$, 28-H)、0.83 (3H, d, $J=6.5\text{Hz}$, 26-H)、0.81 (3H, d, $J=6.2\text{Hz}$, 27-H)、0.74 (3H, s, 18-H), 见图 4-23。综合上述信息, 确定化合物 **4-8** 为麦角甾醇过氧化物。

参 考 文 献

- [1] 徐任生等. 天然产物化学导论. 北京: 科学出版社, 2006.
- [2] 吴立军等. 天然药物化学. 第4版. 北京: 人民卫生出版社, 2004.
- [3] 龚运淮等. 天然产物核磁共振 ^{13}C -NMR 碳谱分析. 昆明: 云南科技出版社, 2004.
- [4] 姚新生. 有机化合物波谱分析. 北京: 中国医药科技出版社, 2004.
- [5] H W 德乐著. 有机化学中的光谱方法. 王剑波等译. 北京: 北京大学出版社, 2001.
- [6] 周公度等. 晶体和准晶体的衍射. 北京: 北京大学出版社, 1999.
- [7] Narita H, et al. *Phytochemistry*, 2007, 68: 587-590.
- [8] 刘桂艳等. 高等学校化学学报, 2006, 27: 2120-2122.
- [9] Li W, et al. *Phytochemistry*, 2001, 58: 595-598.
- [10] Ahmed A A, et al. *Phytochemistry*, 2007, 68: 680-686.
- [11] 明东升. 优秀硕博论文库, 1998, 49.
- [12] Sun J, et al. *Phytochemistry*, 2008, (in press).
- [13] 干玉娟等. 中国药物化学杂志, 2007, 17: 104.



第5章 药用植物活性成分的生物转化

5.1 概述

生物转化 (biotransformation) 是利用生物体系或其产生的酶对外源性化合物进行结构修饰的生物化学过程。就其本质而言,生物转化是生物体系对外源性底物的酶催化反应。可以进行传统有机合成不能或很难进行的化学反应,并具有反应选择性强 (立体选择性、位置选择性)、反应条件温和、副产物少、不造成环境污染和后处理简单等优点。生物转化与生物合成 (biosynthesis) 不同,后者是指利用整体细胞、器官和机体中简单的底物合成复杂化合物的过程。它们又和生物降解 (biodegradation) 不同,生物降解是将复杂的底物分解为简单物质。

生物转化的反应类型多种多样,常见的主要有羟基化 (hydroxylation)、糖苷化 (glycosylation)、氧化-还原 (oxido-reduction)、异构化 (isomerization)、甲基化 (methylation)、酯化 (esterification)、水解 (hydrolysis)、环氧化 (epoxidation) 以及重排反应 (rearrangement) 等。

用于转化研究的生物体系主要有真菌,细菌,藻类,植物离体培养的细胞、组织或器官,以及动物细胞等。其中,应用最多的是植物细胞悬浮培养体系 (plant cell suspension culture) 和微生物体系。

植物离体培养是指在无菌条件下,将离体的植物器官 (根、茎、叶、花、果实、种子)、组织 (如形成层、花药组织、胚乳、皮层等)、细胞 (体细胞和生殖细胞) 或原生质体,培养在人工培养基上,给予适当的营养条件,使其生长。

该技术在 Schleiden 和 Schwann 创立的细胞学说基础上,1902 年由德国生理学家 Haberlandt 提出,并对小野芝麻、风眼兰的叶肉组织、万年青属植物的表皮细胞等进行了培养,但没有成功。1934 年美国植物生理学家 White 成功进行了番茄根系培养,此技术在 20 世纪 70 年代末发展成熟,80 年代中期进入高速发展期。

我国植物组织培养研究工作开展较早。1931 年李继侗培养银杏胚;1935~1942 年罗宗洛进行了玉米根尖离体培养。其后,罗士韦进行了植物幼胚、根尖、茎尖和愈伤组织的培养。20 世纪 70 年代以来,我国在植物组织培养方面进行了大量研究,取得了一系列研究成果,该技术的普及与技术水平均居世界领先地位。与此同时,该技术开始应用于药用植物体系,生产培养一些药用有效成分,已经在工业上获得了成功。

据 20 世纪 80 年代末期的统计表明, 当时全世界有 40 多种资源植物的次生代谢细胞工程的研究获得成功, 悬浮培养体系中次生代谢物质的产量达到或超过整株植物的产量, 研究达到中试水平。其中培养紫草细胞生产紫草宁的成功比较令人瞩目, 1984 年日本的 Mitsui 石化公司利用紫草的细胞培养生产紫草宁, 规模达到 750L, 产物最终浓度达到 1400mg/L。20 世纪 90 年代至今, 利用植物细胞工程进行天然产物的生产进入了一个新的发展阶段, 全世界已经有 1000 多种植物进行过细胞培养方面的研究。利用植物细胞培养生产的植物次生代谢产物包括药用成分、香料、色素、杀菌剂等, 已经成功地培养出了人参皂苷、紫杉醇、阿马里新等一系列的药用植物活性成分。

作为生物催化剂, 植物细胞培养具有下列优点: ①可以在实验室中生长, 重现性好; ②细胞可以无限制地生长, 培养物可以大量累积, 从而源源不断地提供生物材料; ③生长循环周期短, 有利于实验的进行。

微生物转化是通过微生物细胞将复杂的底物进行结构修饰, 也就是利用微生物代谢过程中产生的某种或某一系列的酶对底物特定部位(基团)进行催化反应。微生物在生长过程中的合成酶、水解酶、异构酶、氧化酶、还原酶等都可以对加入其中的底物进行结构修饰。

微生物转化最早可追溯到我国 4000 年前的夏禹年代的酿酒和制醋技术, 及至 20 世纪 40 年代, 出现了工业发酵产品, 如乳酸、柠檬酸、甘油、葡萄糖酸、核黄素等, 开创了工业微生物的新纪元, 形成了第一代生物技术阶段。但这些产物均属一次代谢产物, 代谢过程简单, 代谢产物结构和原料要求不复杂。20 世纪 40 年代以后, 出现了以抗生素为代表的一类生物合成药物。它们形成途径复杂, 发酵周期长, 产物结构较原料复杂且不稳定, 但临床需要量很大, 如抗生素和维生素。所以 20 世纪 40~50 年代, 工业微生物生产进入了一个新阶段, 其里程碑是 1952 年美国普强药厂利用微生物黑根霉 (*Rhizopus nigricans*) 的羟化酶选择性进行的孕甾酮 (progesterone) 的 11 α -羟基化反应, 从而成功解决了皮质激素类药物合成过程中的难题。20 世纪 70~80 年代生物技术再度迅速发展, 形成了以基因工程为主导, 包括微生物工程、细胞工程和酶工程的新领域。

在人类已知的 30000 种天然产物中 80% 来源于植物。1985 年有 3500 种新的化学结构被鉴定, 其中 2600 种来源于高等植物。但药用植物的活性成分含量往往很低, 如紫杉醇、三尖杉酯碱、喜树碱、人参皂苷 Rh₂ 等在植物中的含量仅为万分之几或更低, 导致这些活性成分组成的药品售价不菲, 如 1g 抗癌药物紫杉醇价格可达到 5000 美元。而且, 直接从植物中提取这些活性成分, 因植物原料药需求量大, 生长周期长, 不但占用大片的耕地, 而且还可能造成一些珍贵的植物种类的灭绝。如红豆杉生长缓慢, 为了获得抗癌药物紫杉醇原始的做法是从红豆杉的树皮中提取, 提取 1kg 紫杉醇需要 1000 棵生长了 100 年的红豆杉, 按照这种提取方法红

豆杉很快就会从地球上消失。

为了解决现有的药用植物野生资源难以满足工业化生产，或者栽培品种因活性成分含量低难以实现产业化的现状，药用植物活性成分的生物转化技术为原料药人工资源的开发提供了有效途径。

目前，药用天然产物的研究，发现了大约有 10000 多种药用植物活性成分，结构类型包括酚类、黄酮类、香豆素、木脂素、生物碱、萜类、甾类、皂苷以及有机酸等。其中，萜类、生物碱类和黄酮类（酚性）化合物进行生物报道较多。

5.2 药用植物萜类活性成分的生物转化

萜类化合物是结构变化最大的一类天然产物，它们已经成为广泛应用的重要工业产品，如调味品、药品、香水、杀虫剂和抗菌剂等。在生态环境系统中，它们也同样发挥了重要作用，如植物与环境间相互作用、植物之间的信息传递、植物和昆虫及动物之间的相互作用等。除了一部分是一次代谢产物（如植醇、类胡萝卜素、植物激素赤霉素）外，大部分的萜类是典型的植物二次代谢产物，并具有非常明显的药用价值，如抗疟药物青蒿素（artemisinin）、抗肿瘤药物紫杉醇（taxol）和雷公藤甲素（triptolide）等，巨大的商业市场需求，使药用植物活性成分转化成为研究热点。对人参皂苷（ginsenosides）、白头翁皂苷（pulsatilloside）、甘草苷（glycyrrhizin）和白桦酸（betulinic acid）等三萜及其皂苷类也在不断进行各种类型生物转化研究。

5.2.1 紫杉醇生物转化

1971 年 Wani 与其同事分离鉴定了紫杉醇结构，1992 年 FDA 批准紫杉醇用于治疗卵巢癌，1994 年用于乳腺癌临床治疗。

在临床应用了一段时间后，发现作为一种药用植物的有效成分，紫杉醇药物的开发存在一系列的问题，如植物资源的匮乏导致原料药供应受到限制、紫杉醇提取工艺复杂、得率低、紫杉醇的疏水性引起的副作用等，使临床应用受到了很大的阻力。这一现象随着紫杉醇抗肿瘤作用机制的发现而改变。Wall（1995）和 Rowinsky（1997）相继发现了紫杉醇独特的抗癌机制，它可以诱导和促进 β -微管蛋白聚合，防止解聚，稳定微管，从而导致细胞在有丝分裂时不能形成正常的纺锤体和纺锤丝，抑制细胞分裂和增殖。此外，紫杉醇抑制有丝分裂所必须的微管网正常再生，防止有丝分裂纺锤体形成，导致染色体断裂并抑制细胞复制和转移，从而发挥抗肿瘤作用。在许多临床科学家的极力推动下，紫杉醇及其衍生物成为临床治疗药物，全球每年的销售额估计超过 30.0 亿美元。

目前，市场上紫杉醇（taxol, 5-1，俗名 paclitaxel）药物主要是细胞培养和

通过前体 [浆果赤霉素 (baccatin) III, 5-2] 进行半合成获得。如多烯紫杉醇 (taxoter, 5-3, 俗名 docetaxel) 是由 10-去乙酰基浆果赤霉素 III (5-4) 经过半合成而来, 并相继有一些全合成路线, 但这些方法均存在着产量低和成本高等问题。随着紫杉醇治疗癌症种类增多、治疗疾病范围的扩大、与其他抗肿瘤药如小红莓类 (anthracyclines) 和铂化合物联合用药等, 紫杉醇及其衍生物的市场扩大了三倍 (图 5-1)。

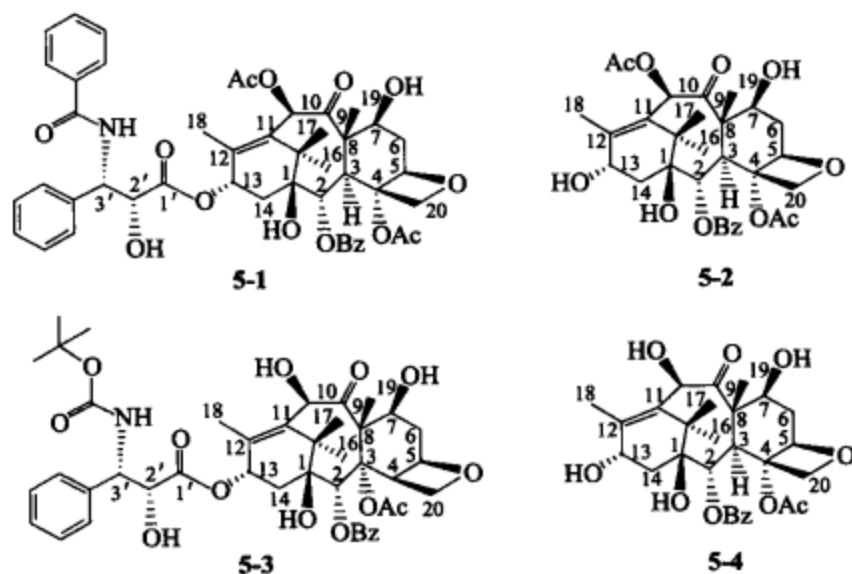


图 5-1 紫杉醇及相关半合成前体结构

紫杉醇工业化生产一直是各国学者研究的热点。全合成、半合成、植物细胞培养和微生物转化技术是目前尝试解决该成分产率低和实现产业化的关键。其中, 紫杉醇生物合成途径、酶催化反应的顺序、关键反应步骤和编码这些酶的基因是提高紫杉醇产量, 完成产业化的技术关键。在这些方法中, 细胞培养具有一定的潜在价值。

5.2.1.1 紫杉醇生物合成途径

目前, 从红豆杉属 (*Taxus*) 植物中分离得到约 400 个紫杉烷类化合物 (taxoids), 包括紫杉醇在内, 所有该类型化合物都具备紫杉烷骨架 (五甲基三环 [9.3.1.0]^{3,8}十五烷), 或者是经重排的骨架 (如 A/B, B/C 结合点的重排)。紫杉烷的核心具有 3 个立体中心, 紫杉醇具有 11 个立体中心, 这样的结构特点增加了紫杉醇类化合物全合成难度。

目前, 各种类型紫杉烷类化合物最合理的生物合成途径还不能确定, 但可以很清楚地知道, 已发现的 400 多个紫杉烷类化合物基本上来源于紫杉烷类的代谢产物或中间体。其中, 紫杉烷类化合物的合成从一次代谢产物开始, 经过 19 步反应生成前体焦磷酸香叶基香叶酯 (geranylgeranyl diphosphate, 5-5, GGPP)。随后, 紫杉烷骨架的环化涉及 8 个细胞色素 P-450 氧化酶参与的氧化反应, 3 个依赖于乙

酰 CoA 的转乙酰化, 包括羟化和乙酰化反应, 生成 4(5), 11(12)-紫杉二烯(5-6)、假想中间体 (5-7), 然后 C-9 位上的氧化成 D 环, 并生成浆果赤霉素 III (5-2) 和 10-去乙酰基浆果赤霉素 III (5-4)。该步骤可以通过化学半合成到达紫杉醇 (taxol, 5-1)。C-13 侧链的生成由另外 5 个反应步骤完成, 即 α -苯丙氨酸 (5-8) 在苯丙氨酸异构酶 (phenylalanine aminomutase, PAM) 的作用下生成 β -苯丙氨酸 (5-9), 再通过与乙酰 CoA 结合后转移浆果赤霉素 III 生成 β -苯丙氨酰浆果赤霉素 III (5-10), 2'-羟化(2'-hydroxylation) 和 N-苯甲酰反应 (N-benzoylation) 形成紫杉醇 (5-1) (图 5-2)。

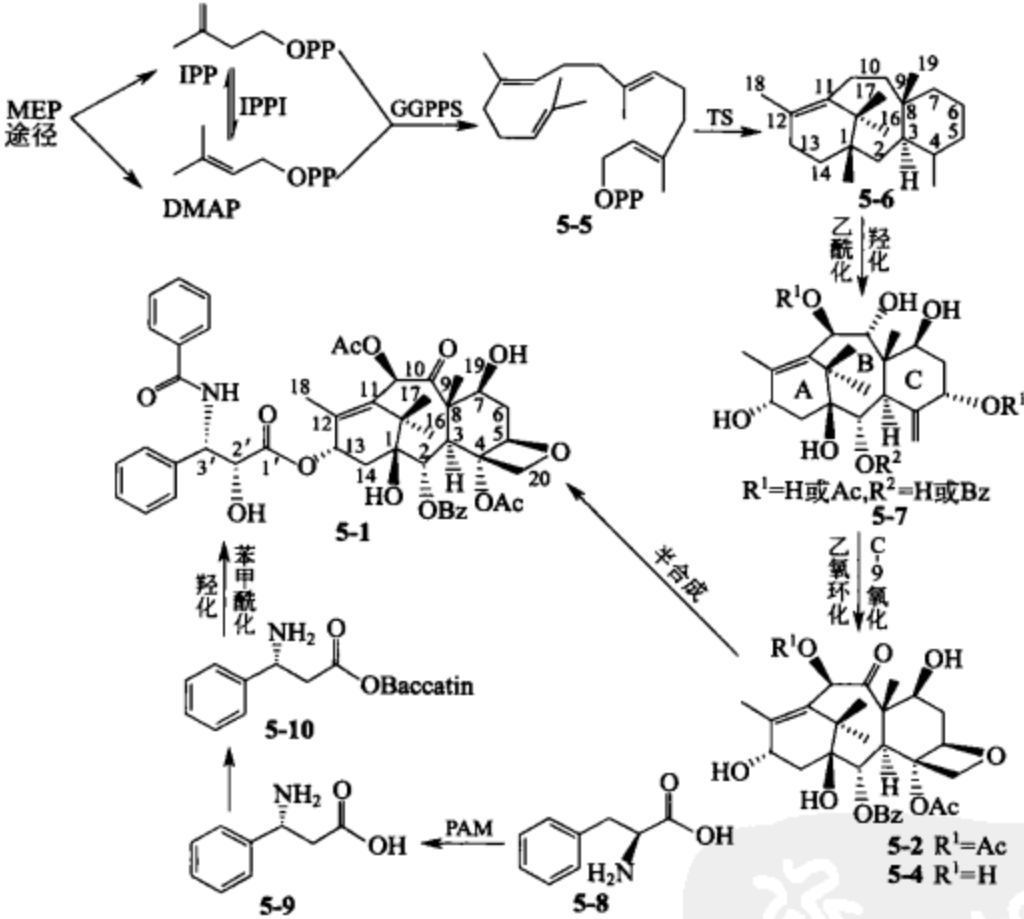


图 5-2 紫杉醇生物合成途径

有关紫杉醇生物合成的途径和假说还存在着争议, 但以上的生物合成途径都是建立在紫杉醇悬浮细胞饲喂培养试验基础上, 并为紫杉醇生物转化的产业化提供理论基础。

5.2.1.2 紫杉醇生物转化

(1) 植物细胞转化技术 细胞培养一直是紫杉醇生物转化的研究热点, 并被认为是有一种有前途的工业化生产手段。Christen 等 1989 年首次报道紫杉醇的细胞培养, 并于 1991 年获得美国专利 (专利号: 5019504) 至今, 在紫杉醇细胞培养过程中, 一直通过饲喂前体、糖和诱导剂如甲基茉莉酮酸等提高紫杉醇产量。这些研究

结果表明不同的处理方式可以增加紫杉醇产量。其中, Yukimune 报道, 加入诱导剂甲基茉莉酮酸可以提高紫杉醇的产量, 最高可以达到约为干重的 0.5%。Nguyen 报道了影响紫杉醇悬浮细胞稳定性因素和复苏。到目前为止, 已经有多种红豆杉属植物作为材料进行细胞培养生产紫杉醇, 但作为商业用途开发还存在着产量低、不稳定、产品收率小和生产成本高等缺陷。因此, 高产细胞株的筛选、快速培养工艺和降低生产成本仍是制约紫杉醇满足临床需要的瓶颈。

此外, 还有长春花 (*C. roseus*)、桔梗 (*P. grandiflorum*) 和银杏 (图 5-3) 将紫杉烷类化合物 sinenxan A (5-11) 经过悬浮细胞培养后生成系列转化产物 (5-12~5-15); 蛇根木 (*R. serpentina*) 细胞悬浮培养后, 将紫杉醇转化为 5-2、5-4 和 5-16 (图 5-4)。

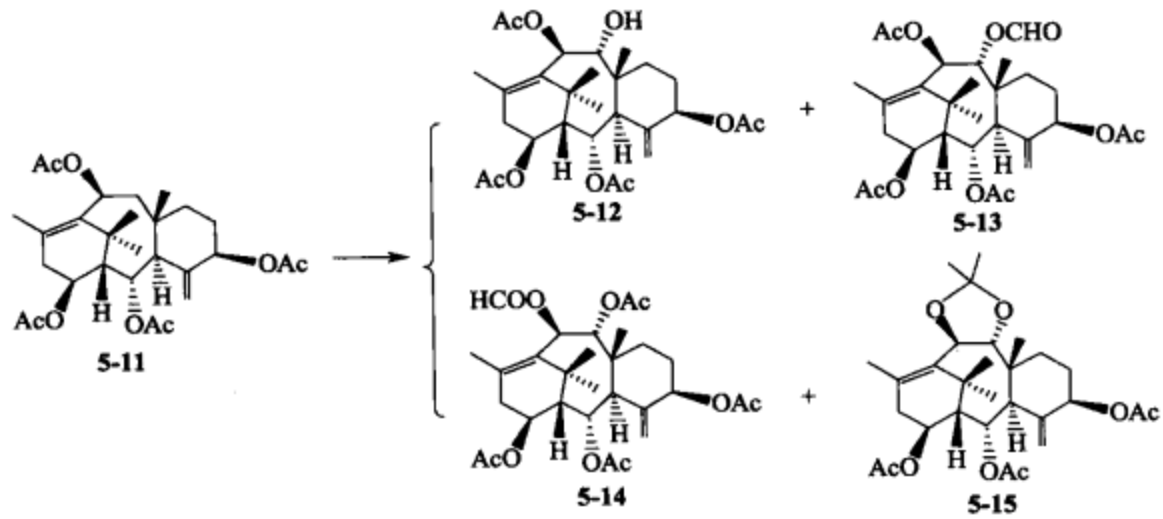


图 5-3 长春花、桔梗和银杏悬浮细胞对 sinenxan A 的转化

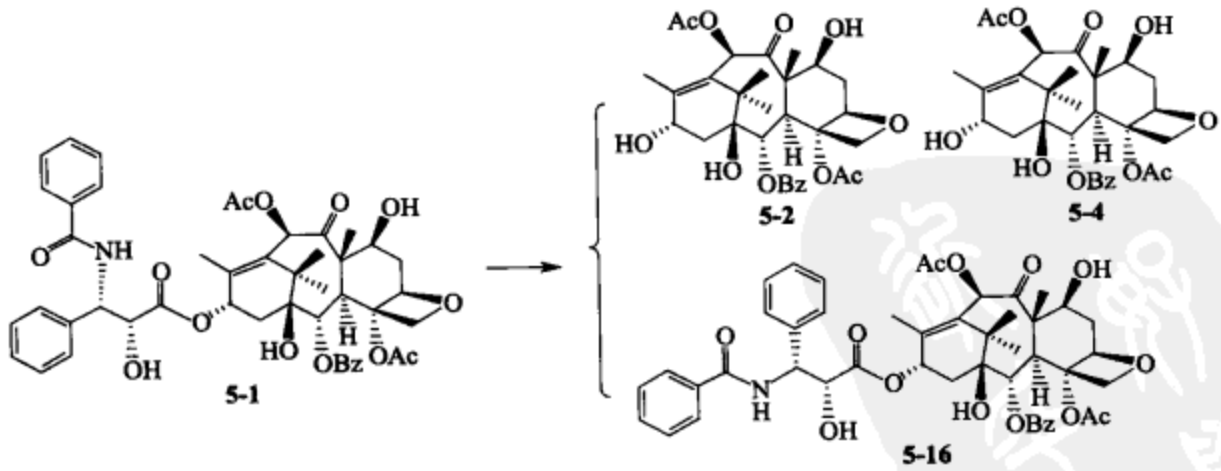


图 5-4 蛇根木悬浮细胞对紫杉醇的转化

(2) 微生物转化 1993 年 Stierle 等首次报道了产紫杉醇真菌, 但产量很低, 只有 24~50ng/L。尽管微生物发酵方法产生的紫杉醇含量很低, 但微生物的培养周期与红豆杉属植物的生长速度比较, 具有培养周期短、增长速度快

等优点，是一个值得期待的研究方向，引起了研究者的极大兴趣，并由此开始进入了紫杉醇微生物转化的研究领域，相继报道了产紫杉醇内生真菌的分离，证明非红豆杉属植物也可以生成紫杉醇，微生物转化有望成为紫杉醇产业化的方向。但研究同时发现，微生物发酵生产紫杉醇过程产量很低，并且产量不稳定。真菌发酵生成的紫杉醇产量为 $24\text{ng/L} \sim 70\mu\text{g/L}$ ，这也说明微生物转化的不稳定性。为解决这些问题，目前研究主要是寻找和分离稳定、高产的菌株和优化发酵条件等。

Chen 等报道从 500 株微生物中筛选出 *Streptomyces* sp. MA7065，对紫杉醇 (5-1) 及三尖杉宁碱 (cephalomannine, 5-17) 进行生物转化，分离得到 3 个转化产物 (5-18~5-20)，分别为侧链的苯环及甲基发生羟化反应 (图 5-5)。

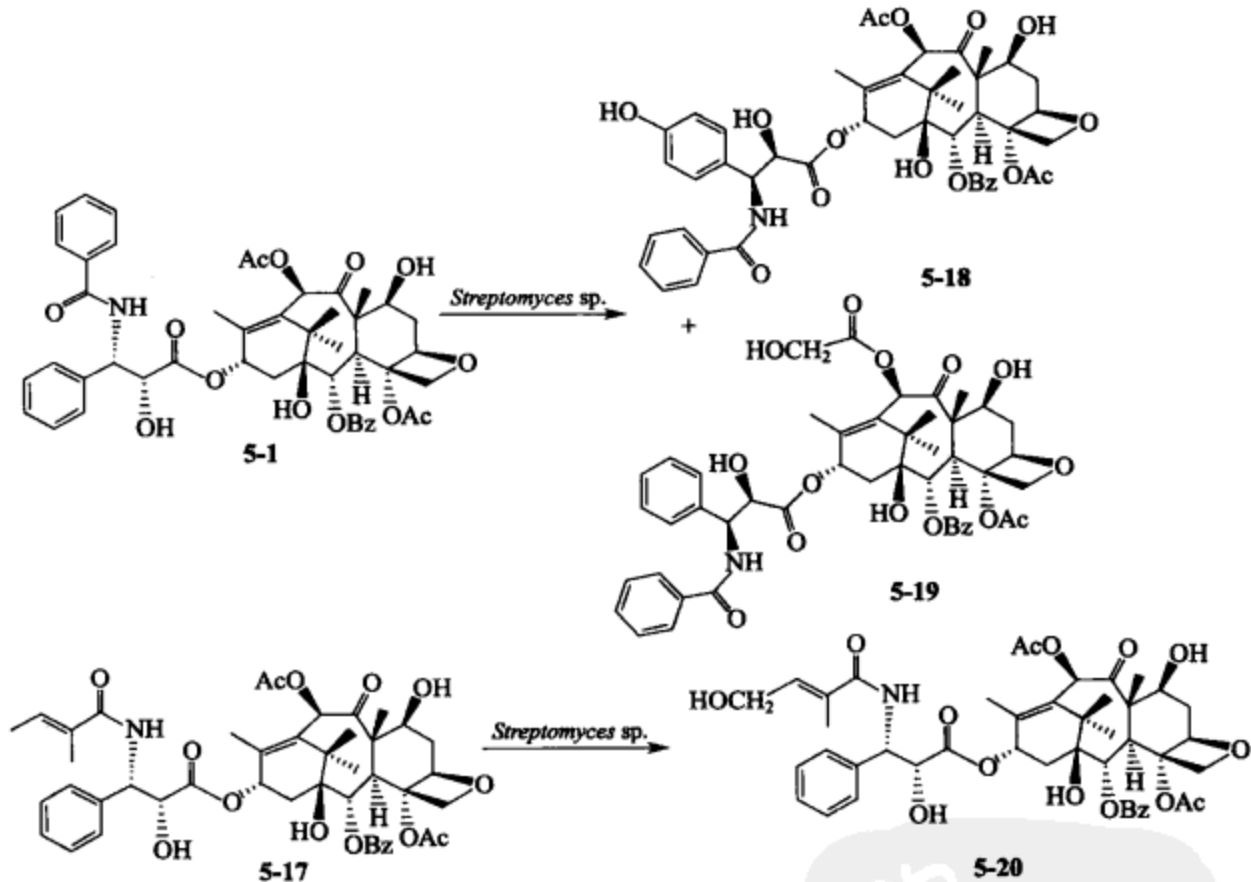


图 5-5 *Streptomyces* sp. 对紫杉醇和三尖杉宁碱的生物转化

占纪勋等从 32 株微生物中筛选出 4 株具有转化紫杉醇能力的菌株，并利用铜绿假单胞菌 *Pseudomonas aeruginosa* AS 1.860 对紫杉醇进行生物转化，得到 3 个侧链水解产物，分别为浆果赤霉素 III (5-2)、10-去乙酰基浆果赤霉素 III (5-4) 和浆果赤霉素 V (5-21) (图 5-6)。

为采用生物转化技术生产紫杉醇，美国 Bristol-Myers Squibb 公司做了大量研究，发现从 *Nocardioideis albus* SC 13911、*Nocardioideis luteus* SC 13912 和 *Moraxella* sp. SC 13963 等微生物中分离出来的 C-13 紫杉醇酶 (taxolase)、C-10

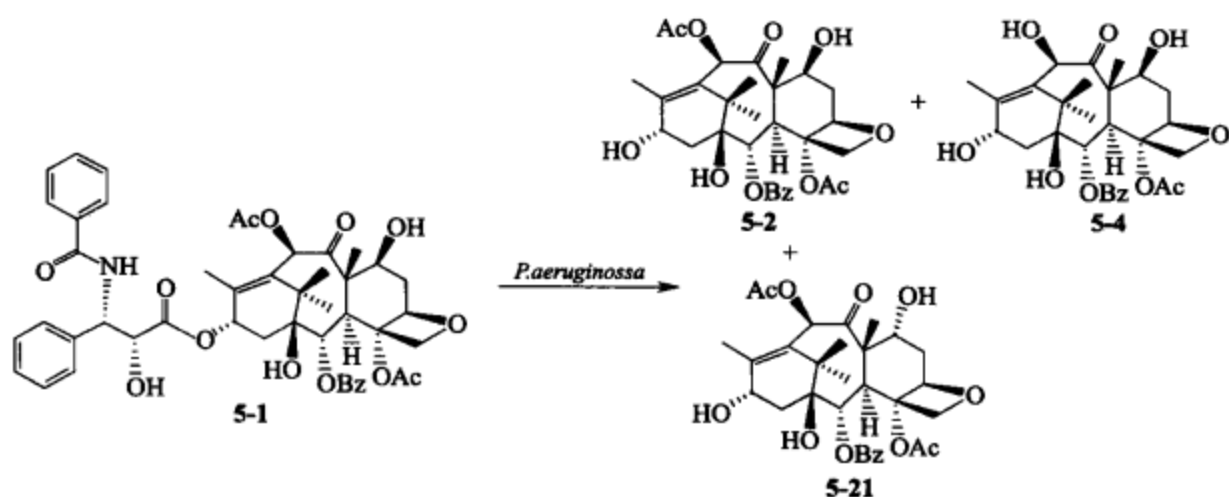


图 5-6 铜绿假单胞菌对紫杉醇的生物转化

脱乙酰酶 (deacetylase) 和 C-7 木糖苷酶 (xylosidase) 等酶能分别将红豆杉植物中的多种紫杉烷类化合物的 13-位侧链、10-位乙酰基和 7-位的木糖水解下来, 得到单一的 10-去乙酰基浆果赤霉素 III, 使得该产物产量提高 5.5~24 倍。10-去乙酰基浆果赤霉素 III 进一步发生化学反应可合成紫杉醇, 从而为解决紫杉醇来源问题提供了很好的方法和思路。

Zhang 等从云南红豆杉树皮内筛选出三株具有转化紫杉烷类能力的内生真菌, 其中 *Microspheeropsis onychiuri* 和 *Mucor* sp. 都可以将 10-去乙酰基-7-表紫杉醇 (5-22) 转化为 10-去乙酰基浆果赤霉素 V (5-23)、10-去乙酰基紫杉醇 (5-24) 和 10-去乙酰基浆果赤霉素 III (5-4), 而 *Alternaria alternata* 则将 1 β -羟基浆果赤霉素 I (5-25) 转化为 5-去乙酰基-1 β -羟基浆果赤霉素 I (5-26)、13-去乙酰基-1 β -羟基浆果赤霉素 I (5-27) 和 5,13-去乙酰基-1 β -羟基浆果赤霉素 I (5-28) (图 5-7)。

Hu 等利用刺孢小克银汉霉 (*Cunninghamella echinulata* AS 3.1990) 和雅致小克银汉霉 (*Cunninghamella elegans* AS 3.2033) 对 sinenxan A (5-11) 进行生物转化, 得到 9 个转化产物 (5-29~5-37) (图 5-8)。这两株菌株特异性地水解 5 位、10 位和 14 位的乙酰基, 并对 6 位进行选择性的羟化, 另外也可以将 4(20) 位的双键转化为三元氧环。

5.2.2 青蒿素的生物转化

20 世纪 60 年代后期, 中国中医研究院中药研究所、云南省药物研究所、山东中医药研究所等单位协作, 发现菊科植物青蒿 (*Artemisia annua* L.) 的中性提取物对鼠疟、猴疟有完全的原虫抑制作用, 临床应用取得了比较满意的疗效。1972 年分离鉴定出抗疟有效单体, 命名为青蒿素 (artemisinin)。到 1972 年底, 青蒿素及其衍生物已在全国的 10 多个地区治疗疟疾患者 6000 余人。1975 年, 中国科学院上海有机化学研究所、北京生物物理研究所等单位协作确定了青蒿素的化学结

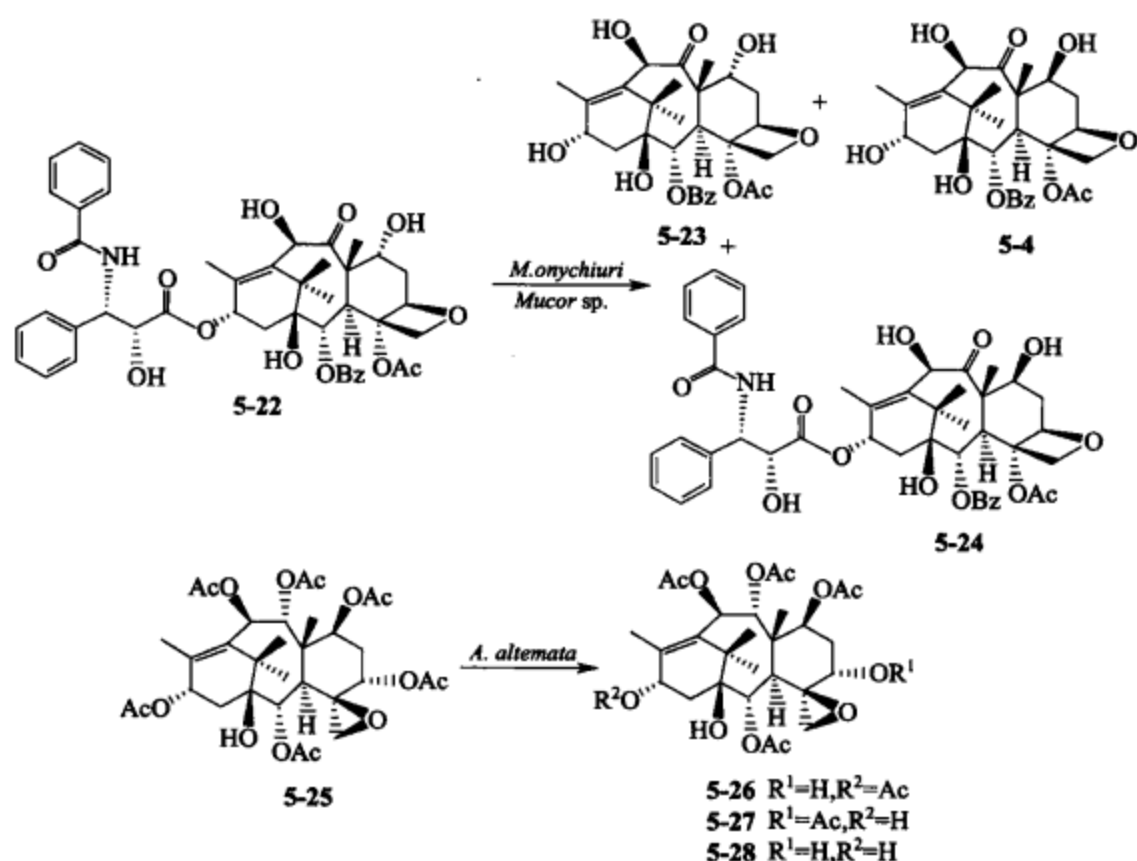


图 5-7 *Microsphaeropsis onychiuri* 和 *Mucor* sp. 的生物转化

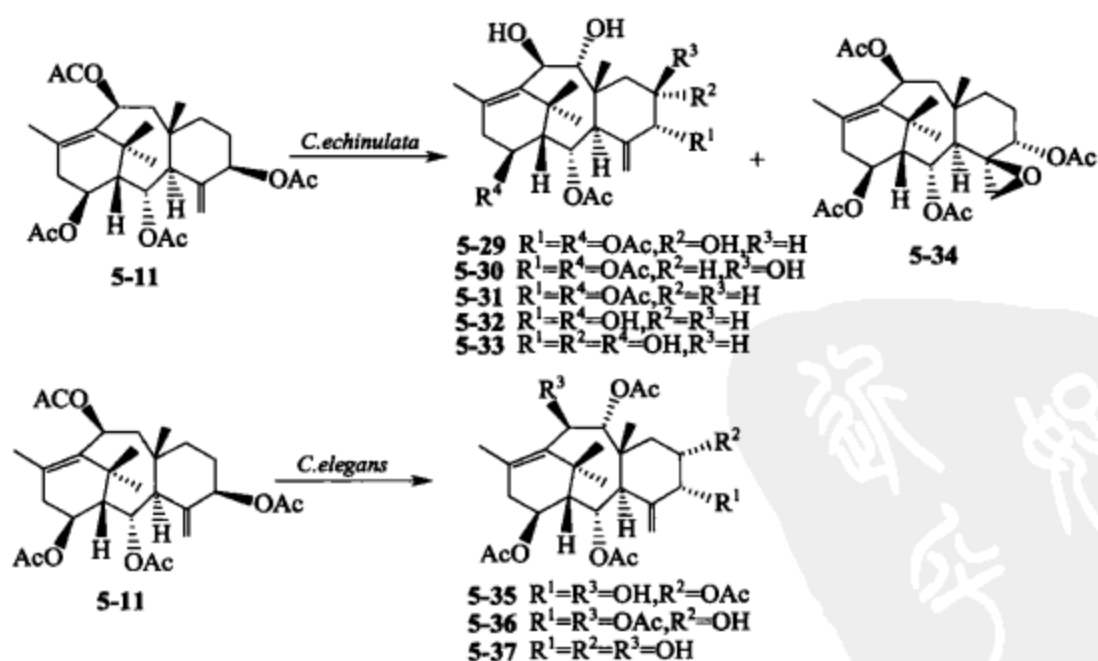


图 5-8 sinenxan A 的生物转化

构。在青蒿素化学结构基础上又开发了一系列衍生物，包括蒿甲醚、蒿乙醚、双氢青蒿素等。有报道，青蒿素衍生物青蒿酯钠及蒿甲醚的抗疟效果比青蒿素好，还原型青蒿素（内酯环上的羰基还原为羟基）的抗疟性也增强，二羟基乙酰化后的产物抗疟作用可以进一步提高。它们已成为世界抗疟药市场中抗脑性疟疾和抗氯喹恶性

疟疾的主要药物。另外，青蒿素还具有增强细胞免疫力、抗流感作用。随着研究的深入，发现青蒿素对白血病和结肠癌细胞具有很好的抑制作用，表明青蒿素还是一种潜在的抗癌药物。

目前药用青蒿素基本上是直接从青蒿植株中提取，但天然青蒿中青蒿素的含量极低（占干重的 0.1%~1%），且其提取环节多，费时费力，产量受环境和季节的限制，使青蒿素的生产成本高、产量低，难以满足市场需求。青蒿素虽已能化学合成，但因成本高、产量低、毒性大，也未能投入商业化生产。因此各国学者都在研究青蒿素生物转化的调控，探索提高青蒿素含量的途径。

5.2.2.1 青蒿素的生物合成途径

青蒿素是一种内过氧倍半萜内酯，其基本骨架来源于植物类异戊二烯代谢途径，倍半萜生物合成中的一些共有中间体也是青蒿素生物合成的中间体。通常情况下，单萜、倍半萜和二萜的生物合成都包括两个阶段进行：第一阶段是线性异戊烯前体生成，包括焦磷酸香叶酯（geranyl pyrophosphate, GPP）、焦磷酸金合欢酯（farnesyl pyrophosphate, FPP）和焦磷酸香叶基香叶酯（geranylgeranyl pyrophosphate, GGPP）；第二阶段是由萜类合酶/萜类环化酶将线性前体催化成环。随着反应酶体系不断地从植物体中分离，并在微生物中表达，有关单萜柠檬烯合酶、倍半萜紫穗槐-4,11-二烯（amorpha-4,11-diene）合酶和三萜紫杉二烯（taxadiene）合酶等萜类合成酶的报道不断增加。

萜类化合物的生物合成的重要前体分子异戊烯二磷酸（IPP）同时存在两个显著不同的途径：存在细胞质中的经典乙酸/甲戊二羧酸途径和质体中的甘油醛磷酸/丙酮酸途径。Towler 等采用水生培养的方式，分别添加乙酸/甲戊二羧酸途径 IPP 生成抑制剂洛伐他汀（mevinolin）和甘油醛磷酸/丙酮酸途径 IPP 生成抑制剂膦胺霉素（fosmidomycin）。在培养 7~14 天后，检测发现青蒿素含量明显低于对照组。结果表明，植物体内的青蒿素生物合成的前体物质 IPP 是由以上两种途径中形成。同时，在添加了甾醇类化合物生物合成前体抑制剂咪康唑（miconazole）后，发现青蒿素生物合成水平明显高于对照组，表明青蒿素的生物合成同时由细胞质和质体途径形成的 IPP 为前体，并可以通过调控，使 IPP 向有利于青蒿素合成的方向进行（图 5-9）。

Bouwmeester 等首次从青蒿叶中分离到青蒿素生物合成途径的重要倍半萜类中间产物紫穗槐-4,11-二烯（amorpha-4,11-diene）。而且另外一些研究通过体内和体外反应发现青蒿中含有比青蒿素多 8~10 倍以上的青蒿酸，因此青蒿酸可能是青蒿素生物合成途径中一个重要的中间体。所以青蒿素可能的合成途径是由第一阶段形成前体焦磷酸香叶基香叶酯（geranylgeranyl diphosphate, 5-5, GGPP）和焦磷酸金合欢酯（farnesyl pyrophosphate, 5-38, FPP）；第二阶段经由紫穗槐-4,11-二烯（5-39）、青蒿醇（artemisinic alcohol, 5-40）、青蒿醛（artemisinic aldehyde,

5-41)、青蒿酸 (artemisinic acid, 5-42)、青蒿酸过氧化氢 (artemisinic acid hydroperoxide, 5-43)、青蒿素 B (arteannuin B, 5-44)、青蒿烯 (artemisiten, 5-45)、二氢青蒿醇 (dihydroartemisinic alcohol, 5-46)、二氢青蒿醛 (dihydroartemisinic aldehyde, 5-47)、二氢青蒿酸 (dihydroartemisinic acid, 5-48) 和二氢青蒿酸过氧化氢 (dihydroartemisinic acid hydroperoxide, 5-49) 生成青蒿素 (artemisinin, 5-50) (图 5-10)。

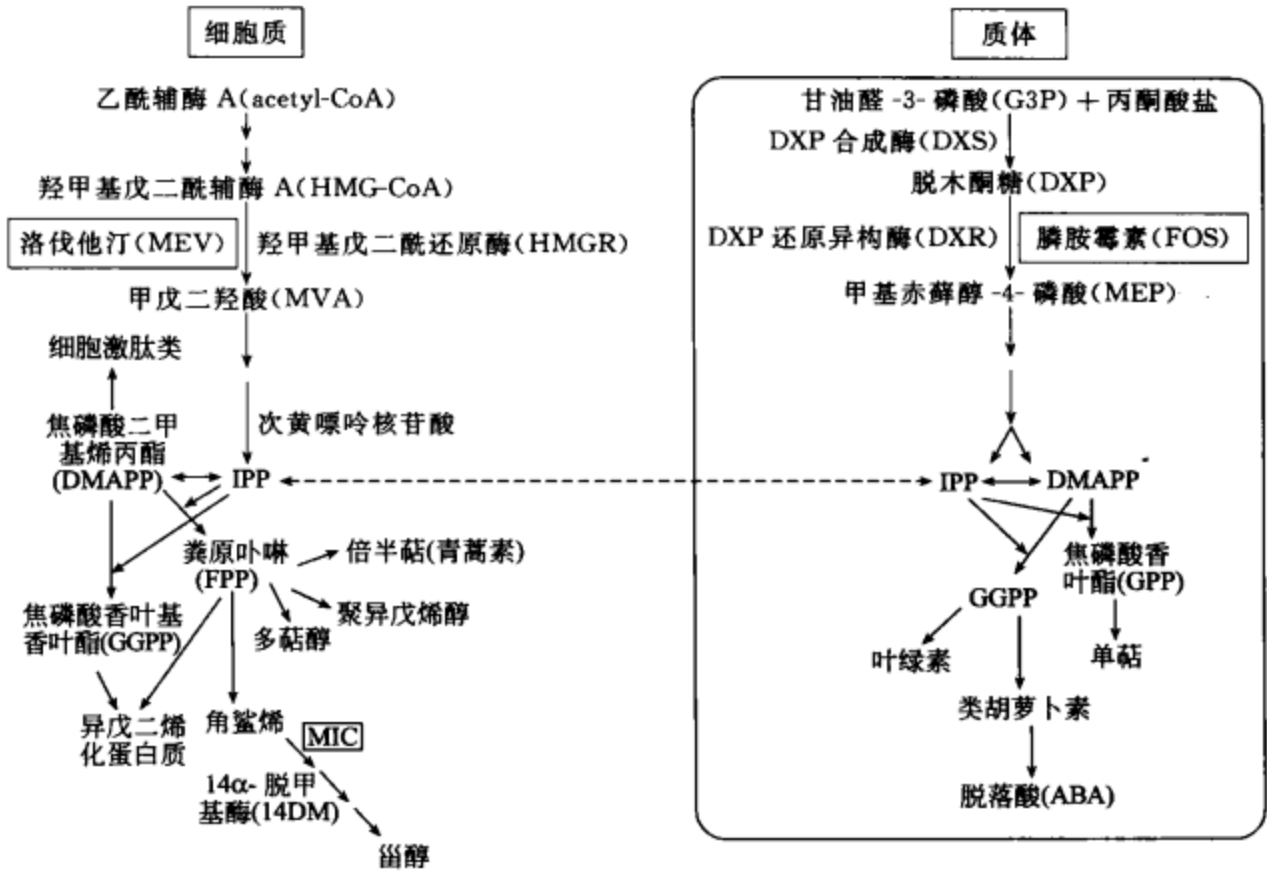


图 5-9 青蒿素前体 IPP 形成途径

5.2.2.2 青蒿素的生物转化

(1) 植物细胞转化技术 青蒿素的植物细胞转化主要集中在植物黄花蒿 (*Artemisia annua* L.) 毛状根培养系或者其他植物毛状根培养系中进行。

在对黄花蒿毛状根培养体系研究中, Liu 等发现有空气搅拌的反应器在 20 天培养后可使青蒿素的量达到 536mg/L; Putalun 等发现加入几丁质 (150mg/L) 可以使青蒿素的产量提高 6 倍, 达到 1.8μg/mg (培养物干重), 加入甲基茉莉酮酸酯 (0.2mmol/L) 和酵母提取物 (2mg/mL) 分别可以使青蒿素产量达到 1.5μg/mg 和 0.9μg/mg; Wang 等发现培养体系中加入 (22S, 23S)-表高油菜素内酯 (SSHB) 1μg/L 和不同浓度的氮源 ($\text{NO}_3^-/\text{NH}_4^+$) 可以使青蒿素产量增加 57%。

周立刚等利用露水草 (*Cyanotis arachnoidea*) 毛状根培养体系将青蒿素 (5-50) 选择性地还原为去氧青蒿素 (5-51); Li 等对大黄 (*Rheum palmatum*) 毛状根培养系的研究得到了同样的结论 (图 5-11)。

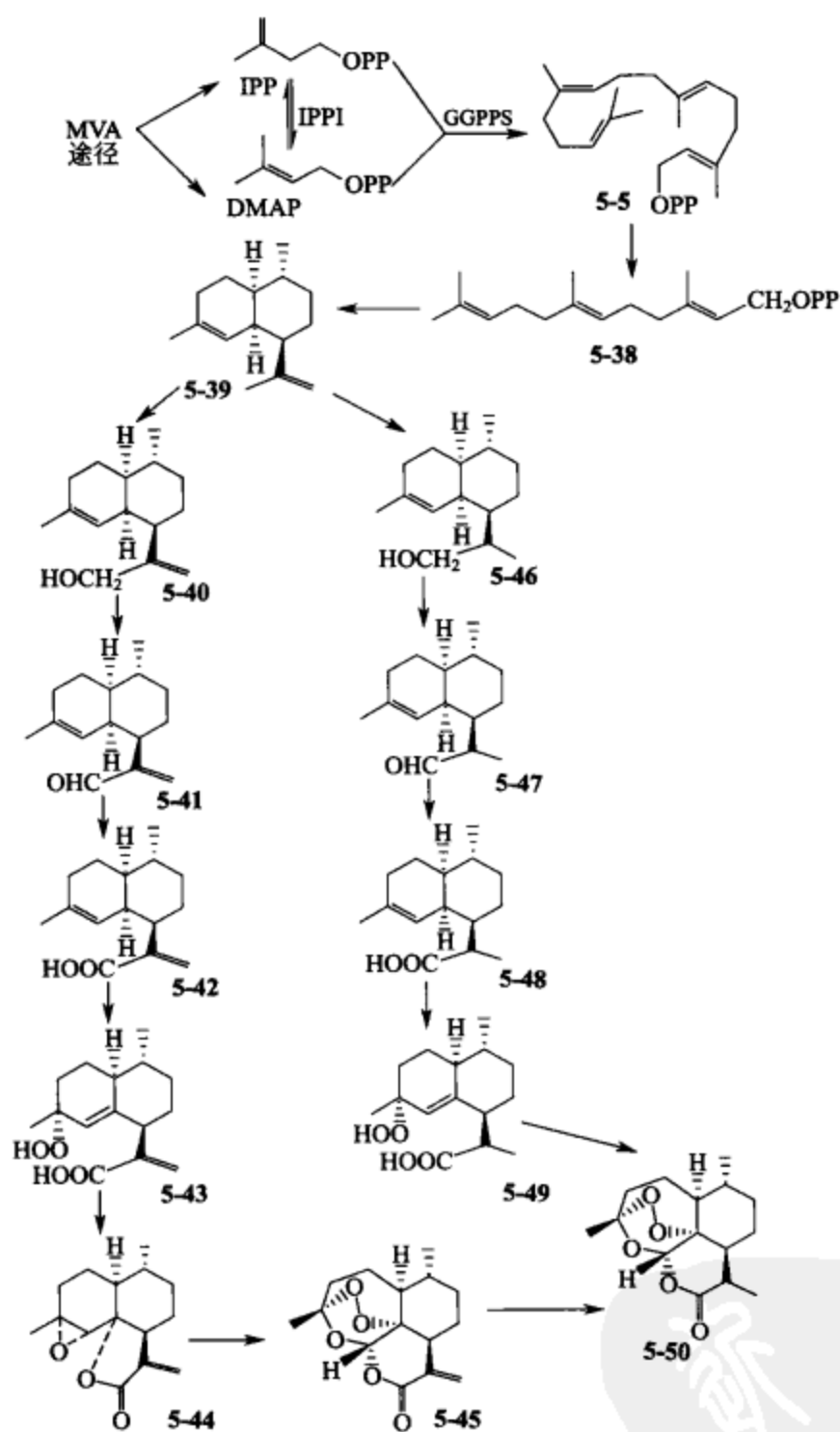


图 5-10 青蒿素生物合成途径

韩健等利用长春花及银杏植物悬浮细胞对青蒿素进行生物转化，发现两种植物的悬浮细胞均能将青蒿素转化成 3α -羟基去氧青蒿素 (5-52) (图 5-12)。

(2) 青蒿素及类似物的微生物转化 Lee 等利用 *Nocardia corallina* ATCC 19070 和 *Penicillium chrysogenum* ATCC 9480 对青蒿素进行微生物转化研究，发现利用 *N. corallina* 转化可以得到青蒿素，*P. chrysogenum* 则可以得到去氧青蒿素和 3α -羟基去氧青蒿素。陈有根等利用 *Streptomyces griseus* ATCC 13273 转化青蒿素得到了一个新化合物 9α -羟基青蒿素 (5-53) (图 5-13)。

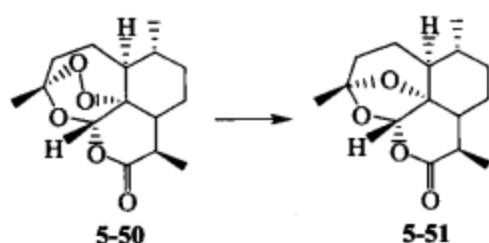


图 5-11 露水草和大黄毛状根
对青蒿素的转化

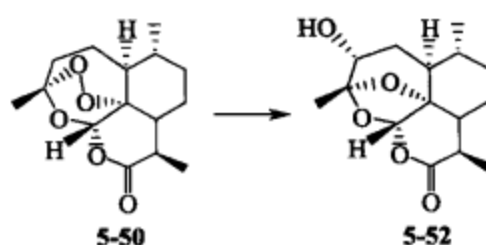


图 5-12 长春花及银杏悬浮培养
细胞对青蒿素的转化

Hufford 等利用 *Cunninghamella elegans* ATCC 9245 对蒿乙醚 (arteether, 5-54) 进行转化, 得到 1 α -羟基蒿乙醚 (5-55)、9 β -羟基蒿乙醚 (5-56)、3 α -羟基-11-表去氧二氢青蒿素 (5-57) 和一个重排产物 (5-58), 而采用 *Streptomyces lavendulae* L-105 转化体系则得到 9 α -羟基蒿乙醚 (5-59)、2 α -羟基蒿乙醚 (5-60)、14 α -羟基蒿乙醚 (5-61) 及 5-58 (图 5-14)。

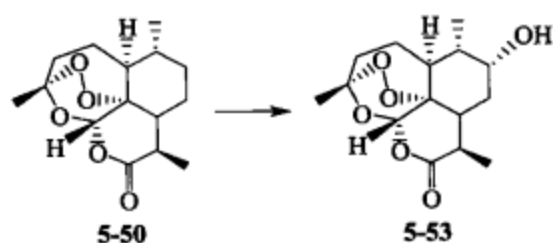


图 5-13 青蒿素 *Streptomyces griseus* 的转化

Abourashed 等利用同样的菌株对蒿甲醚 (5-62) 转化, 得到 9 β -羟基蒿甲醚

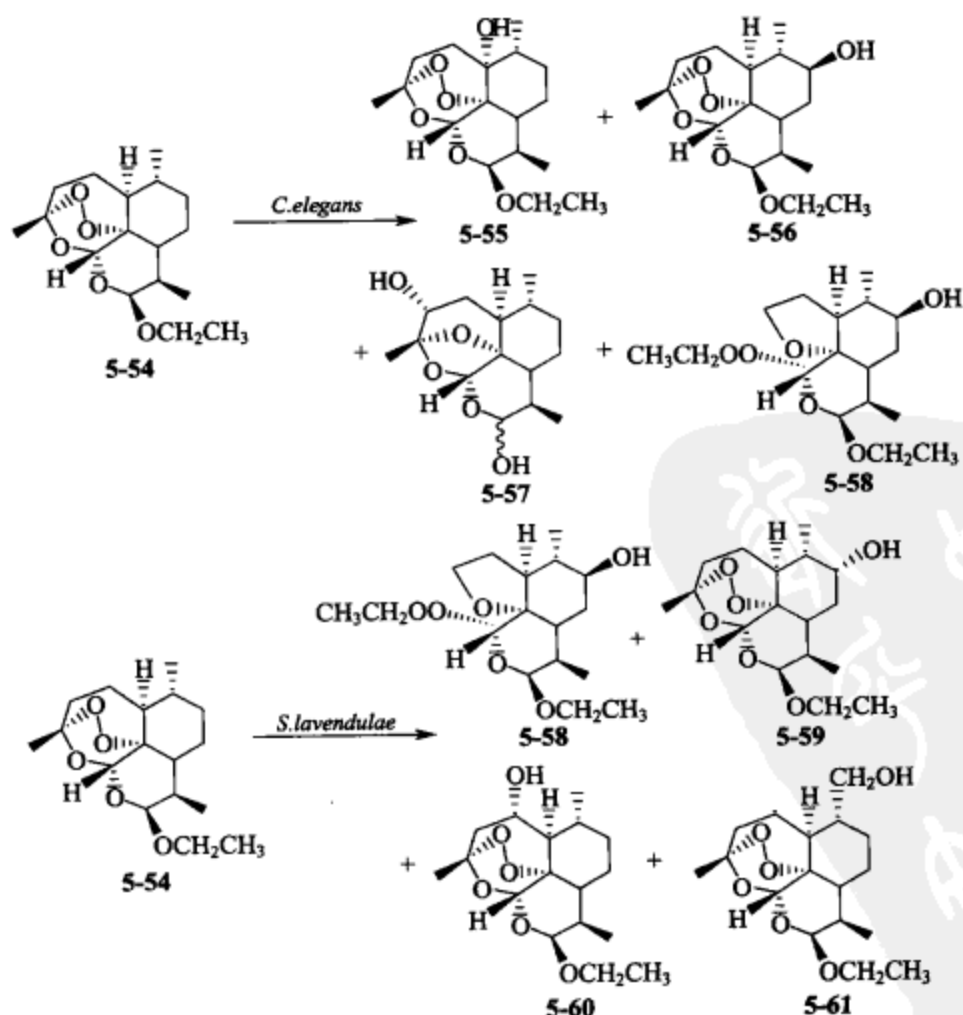


图 5-14 蒿乙醚的微生物转化

(5-63)、重排产物 (5-64)、3 α -羟基去氧蒿甲醚 (5-65)、9 α -羟基蒿甲醚 (5-66) 及 14-羟基蒿甲醚 (5-67) (图 5-15)。

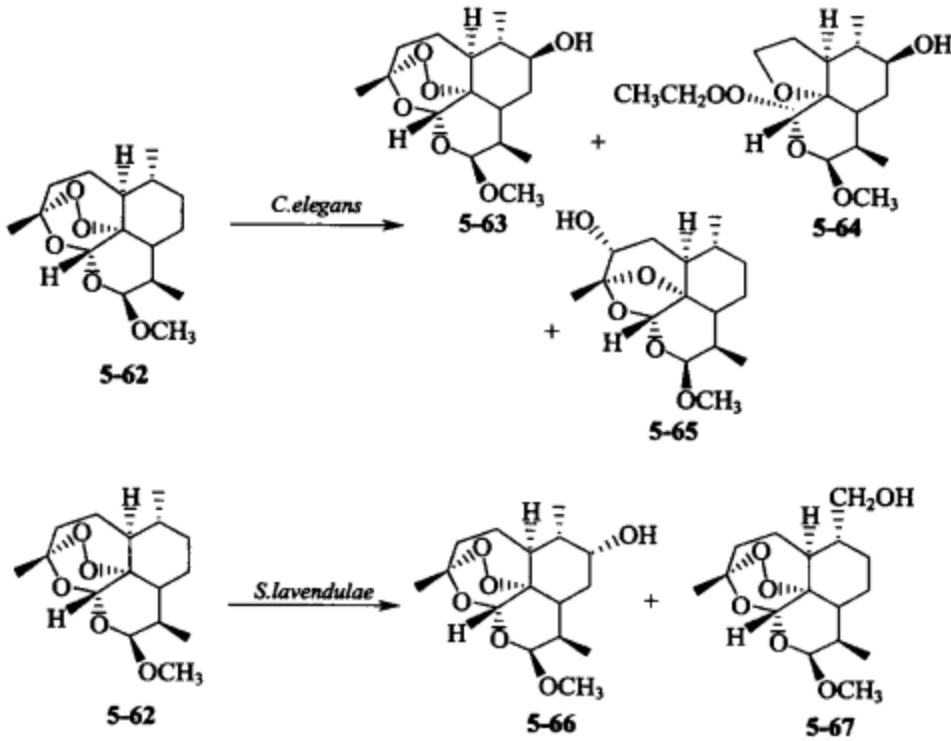


图 5-15 蒿甲醚的微生物转化

Elmarakby 等通过 *Aspergillus flavipes* ATCC 1030 和 *Beauveria bassiana* ATCC 7159 对青蒿素 B (5-44) 进行转化, 分别得到二氢青蒿素 B (5-68) 和 3 β -羟基青蒿素 B (5-69)、13-羟基-11-表二氢青蒿素 (5-70)。Orabi 等利用 *Aspergillus niger* NRRL 599 转化青蒿烯 (5-45), 得到 11-表青蒿素 (5-71)、9 β -羟基-11-表去氧青蒿素 (5-72) 和 9 β -羟基-11-表青蒿素 (5-73) (图 5-16)。

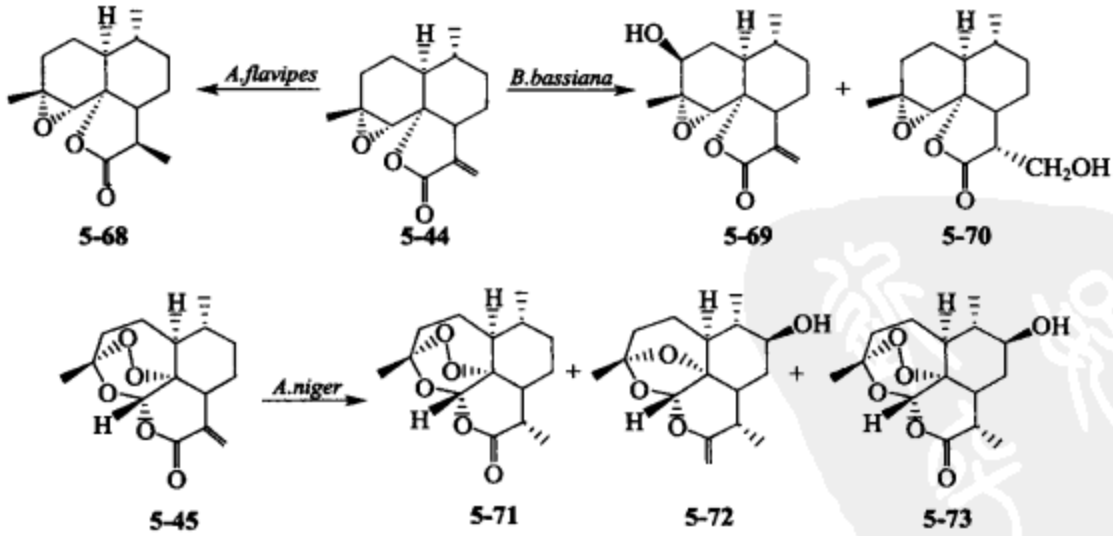


图 5-16 青蒿素 B 和青蒿烯的微生物转化

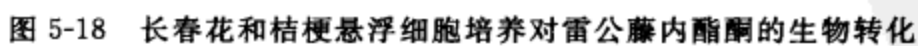
5.2.3 其他萜类化合物的生物转化

5.2.3.1 雷公藤二萜的生物转化

雷公藤二萜是卫矛科植物雷公藤 (*Tripterygium wilfordii*) 的主要有效成分,

Chemical reaction scheme showing the conversion of compound 5-74 to compound 5-75. Compound 5-74 is a complex polycyclic molecule with a carboxylate group, a ketone, and a hydroxyl group. Compound 5-75 is the product, which has a different stereochemistry at the hydroxyl group and the carboxylate group.





利用长春花和桔梗细胞悬浮体系对雷公藤内酯酮 (5-76) 进行生物转化, 得到 10 个产物, 分别是雷公藤甲素 (5-74)、雷公藤乙素 (5-77)、12 β ,13 α -二羟基雷公藤内酯酮 (5-78)、 $\Delta^{12,13}$ -雷公藤内酯酮 (5-79)、5 α -羟基雷公藤内酯酮 (5-80)、2 β -羟基雷公藤内酯酮 (5-81)、17-羟基雷公藤内酯酮 (5-82)、12 α ,13 α -二羟基雷公藤内酯酮 (5-83)、12 α ,13 α -表二羟基雷公藤内酯酮 (5-84) 和 14-O- β -D-葡萄糖表雷公藤甲素 (5-85)。长春花体系对雷公藤甲素及雷公藤内酯酮的转化反应均为选择性地使 12,13-位三元氧环打开。桔梗体系除此反应之外, 还具有多位点的羟基化作用, 另外还观察到糖苷化反应 (图 5-18)。

从 31 株微生物中筛选出短刺小克银汉霉 *Cunninghamella blakesleana* AS 3.970 对雷公藤甲素 (5-74) 进行生物转化, 得到 7 个产物, 分别是 5-羟基雷公藤甲素 (5-86)、19 α -羟基雷公藤甲素 (5-87)、19 β -羟基雷公藤甲素 (5-88)、雷公藤乙素 (2 β -羟基雷公藤甲素, 5-89)、16-羟基雷公藤甲素 (5-90)、雷醇内酯 (15-羟基雷公藤甲素, 5-91) 和 1 β -羟基雷公藤甲素 (5-92) (图 5-19)。

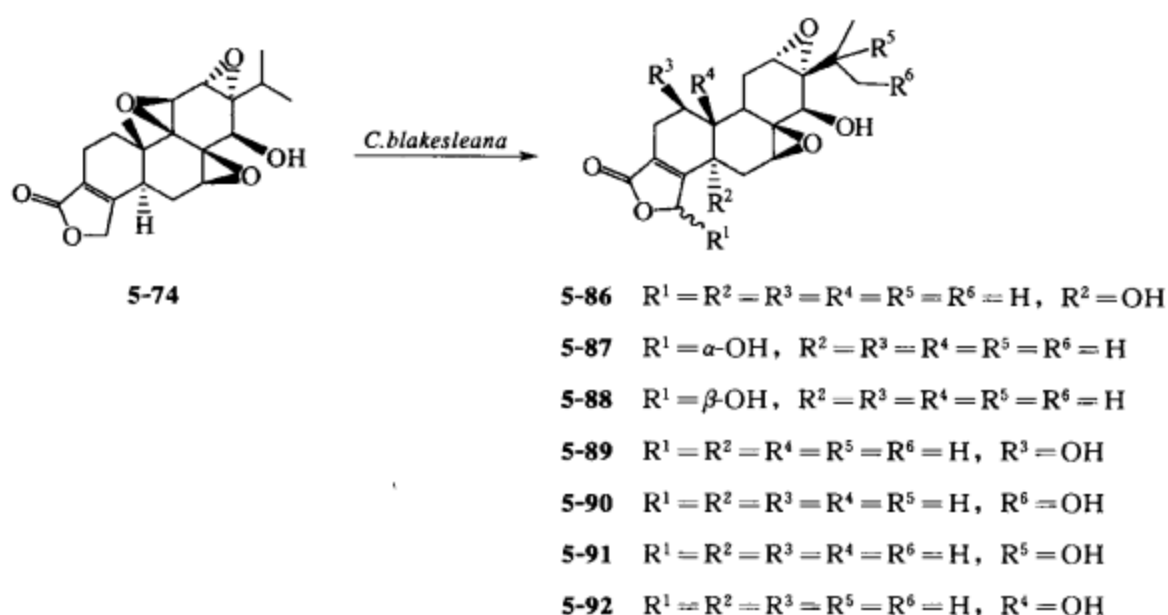


图 5-19 短刺小克银汉霉对雷公藤甲素的生物转化

黑曲霉 *Aspergillus niger* AS 3.739 对雷公藤内酯酮 (5-76) 的转化得 4 个产物, 分别是 17-羟基雷公藤内酯酮 (5-82)、16-羟基雷公藤内酯酮 (5-93)、5 α -羟基雷公藤内酯酮 (5-80)、雷公藤甲素 (5-74) (图 5-20)。

5.2.3.2 三萜类成分的生物转化

(1) 齐墩果烷 (oleanane) 型 该类化合物在植物界分布广泛, 最具代表性的化合物是齐墩果酸 (oleanolic acid, 5-94), 齐墩果酸在 *Nocardia* sp. 的作用下可以转化生成齐墩果酸甲酯 (5-95) (图 5-21)。由于五环三萜类化合物具有抗菌活性, C-27 或 C-28 位具有游离羧基的化合物活性更强, 在该反应中, 微生物可能通过酯化反应来降低底物的溶解性, 从而减少底物对微生物生长的抑制。

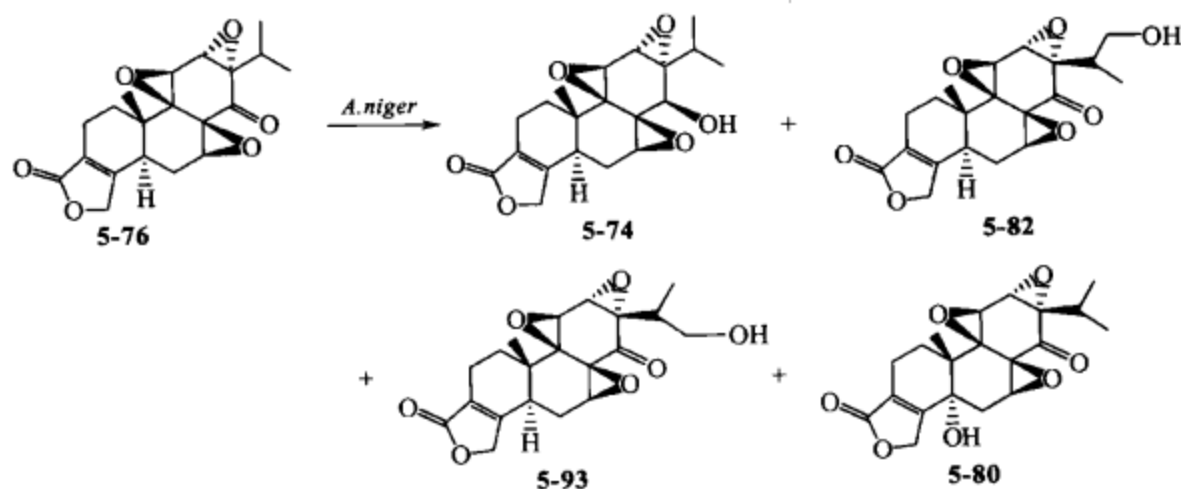


图 5-20 黑曲霉对雷公藤内酯酮的生物转化

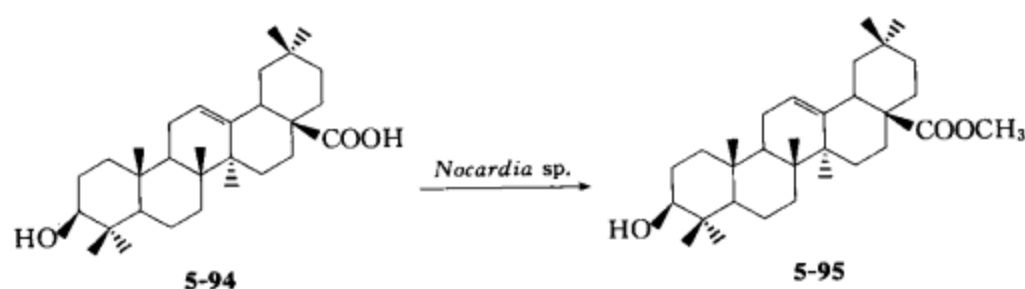


图 5-21 齐墩果酸的微生物转化

甘草酸 (glycyrrhizic acid, **5-96**) 是中药甘草 (*Glycyrrhiza uralensis*) 中重要的活性成分, 它是由甘草次酸 (glycyrrhetinic acid, **5-97**) 与两分子葡萄糖醛酸结合而成, 甘草酸和甘草次酸都有促肾上腺皮质激素 (ACTH) 样的生物活性。Kim 和 Alao 等研究发现, 肠道内厌氧菌 (*Streptococcus* LJ-22, *Ruminococcus* sp. Pol-3, *Eubacterium* sp. GLH 等) 能够将甘草酸转化生成 3-葡萄糖醛酸甘草次酸 (18 β -glycyrrhetinic acid 3-O- β -D-glucuronide, **5-98**), 然后进一步水解生成甘草次酸, 甘草次酸被氧化生成 3-氧代甘草次酸 (**5-99**), 通过可逆反应最后生成 3 α -羟基甘草次酸 (**5-100**)。体内外代谢结果相结合的研究阐明了甘草酸在体内主要是经大肠细菌代谢生成甘草次酸, 然后被吸收 (图 5-22)。

Shirane 等研究一种真菌 *Chaetomium longirostre* 对 3-氧代-12-齐墩果烯-28-酸 (3-oxo-olean-12-en-28-oic acid, **5-101**) 转化时发现, 该菌株可以将五环三萜母核上的 A 环开环生成化合物 3,4-断-12-齐墩果烯-4-醇-3,28-二酸 (3,4-seco-olean-12-en-4-ol-3,28-dioic acid, **5-102**)、21 β -羟基-3,4-断-12-齐墩果烯-4-醇-3,28-二酸 (**5-103**)。在破细胞的酶制剂 (cell-free culture) 条件下, 补充适量的 NADPH, **5-101** 可以被转化生成 A-同-3 α -氧杂-12-齐墩果烯-3-酮-28-酸 (A-homo-3 α -oxo-olean-12-en-3-one-28-oic acid, **5-104**) (图 5-23)。

七叶树皂苷-Ia (escins-Ia, **5-105**) 是传统中药娑罗子中的重要活性成分, 具有抗肿瘤活性。杨秀伟等报道了人肠内菌对该化合物的转化, 得到了 4 个代谢产

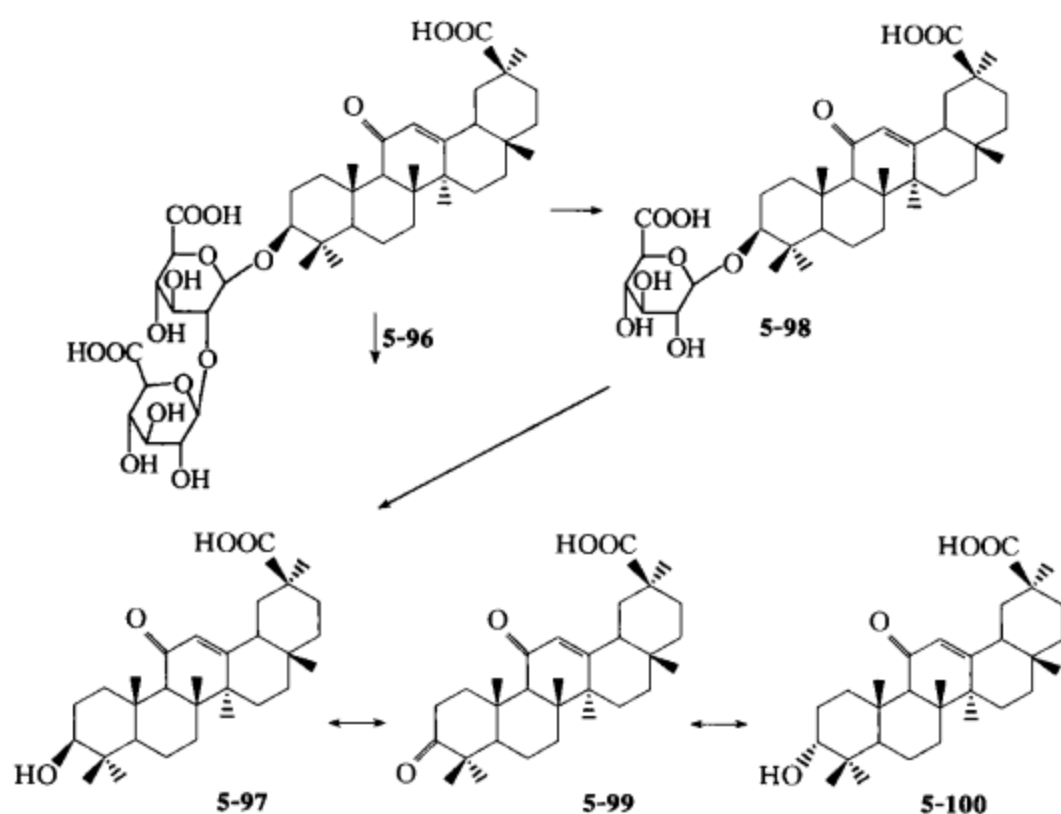


图 5-22 甘草酸的微生物转化

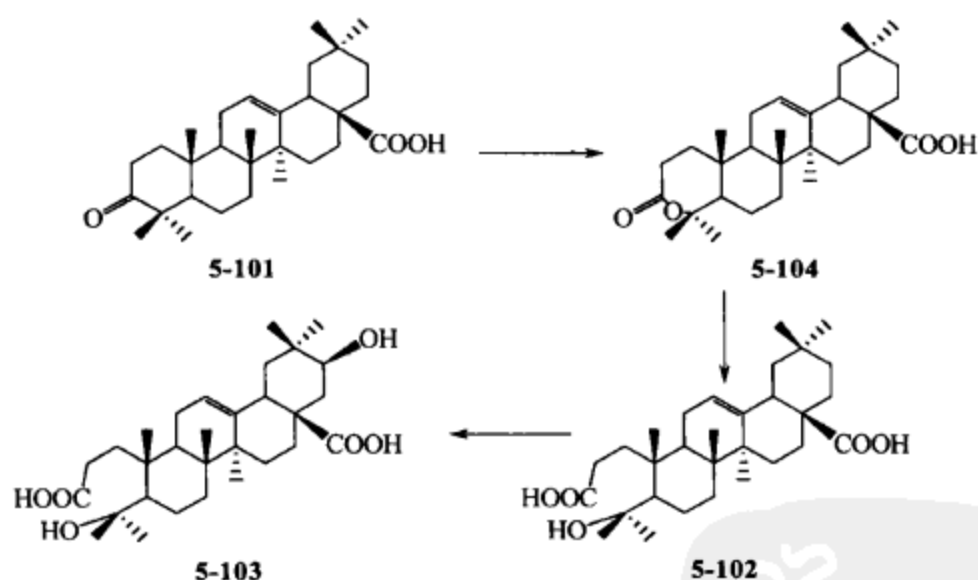


图 5-23 *Chaetomium longirostre* 对 3-氧代-12-齐墩果烯-28-酸的转化

物：desscylescin I (5-106)，isowscin I a (5-107)，原七叶树皂苷元 (protoaescigenin, 5-108) 和 21 β -O-当归酰基原七叶树皂苷元 (21 β -O-tigloylprotoaescigenin, 5-109) (图 5-24)，通过该方法模拟药物在人肠内的代谢，推断七叶皂苷生物利用度低的原因之一是经肠内细菌代谢转化，结构遭到破坏。

远志皂苷元 (senegenin, 5-110) 是远志皂苷在盐酸条件下水解生成的一种含氯的次生皂苷元，余伯阳等研究发现 *Nocardia* sp. 能够将该化合物转化得到远志酸甲酯 (senegenic acid methyl ester, 5-111) (图 5-25)。

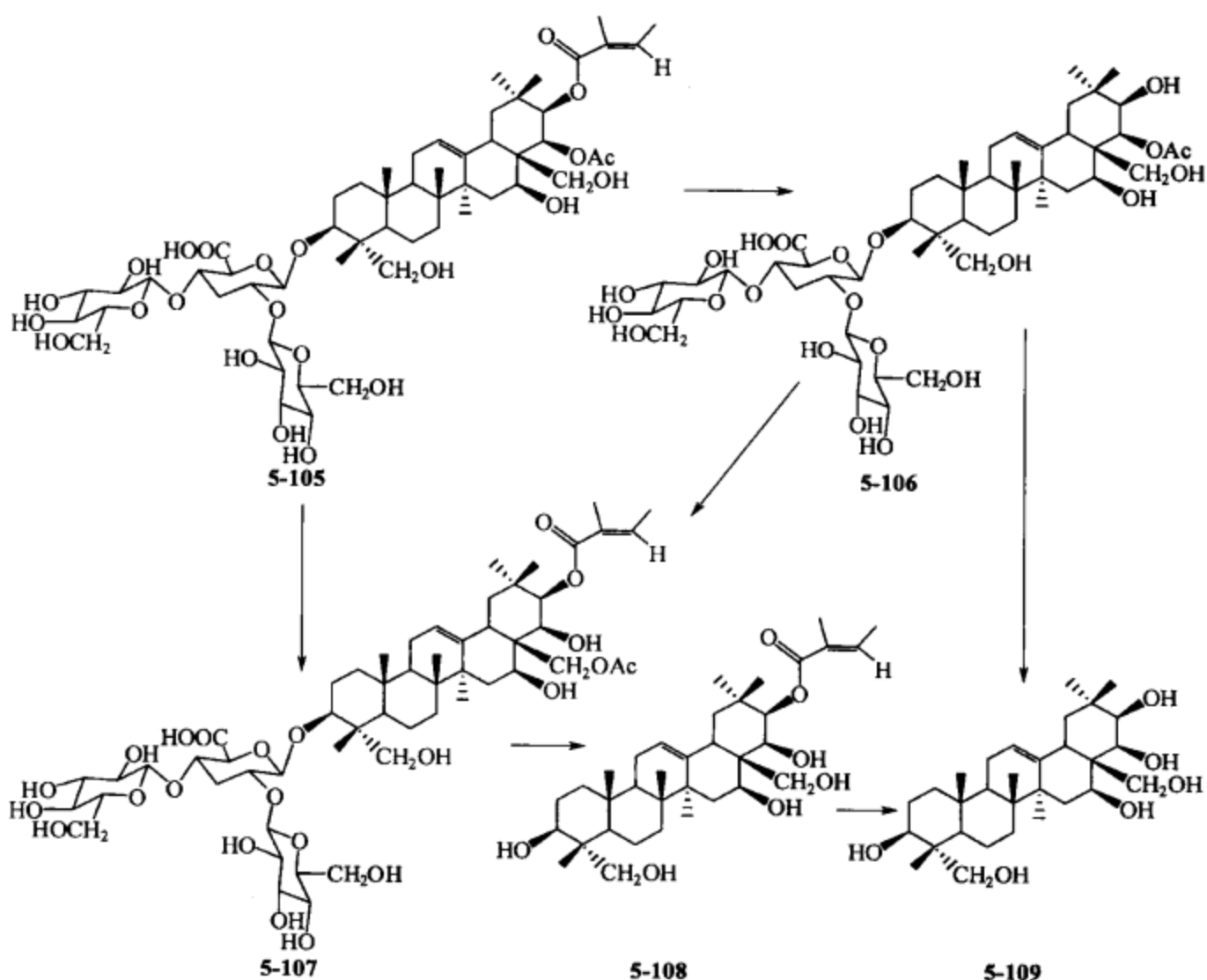


图 5-24 七叶树皂苷-Ia 的微生物转化

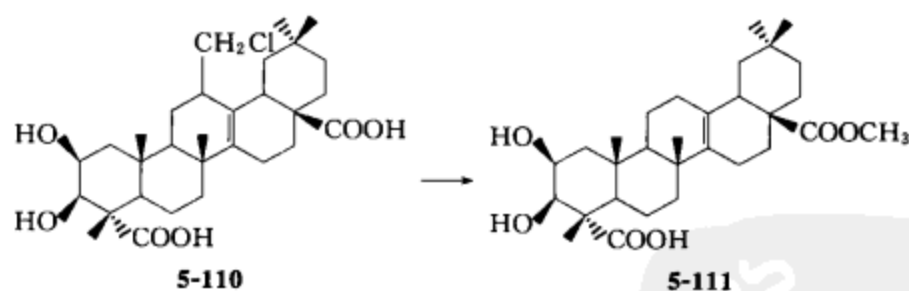


图 5-25 远志皂苷元的微生物转化

(2) 乌苏烷 (ursane) 型 此类化合物大多是熊果酸的衍生物, 余伯阳等研究 *Nocardia* sp. 对熊果酸 (ursolic acid, 5-112) 和 3-O-异鼠李糖奎诺酸 (quinovic acid 3-O-quinovoside, 5-113) 的生物转化时发现该菌株可以将乌苏烷型的五环三萜转化为齐墩果烷型的五环三萜。在乌苏烷类化合物的生源合成途径中, 齐墩果烷型的五环三萜是通过碳正离子重排形成乌苏烷型五环三萜。该研究首次发现了微生物体系所催化的乌苏烷型 (5-114) 向齐墩果烷型 (cincholic acid, 5-115) 转变的反应类型, 对于研究该类化合物的合成转变具有重要的意义 (图 5-26)。

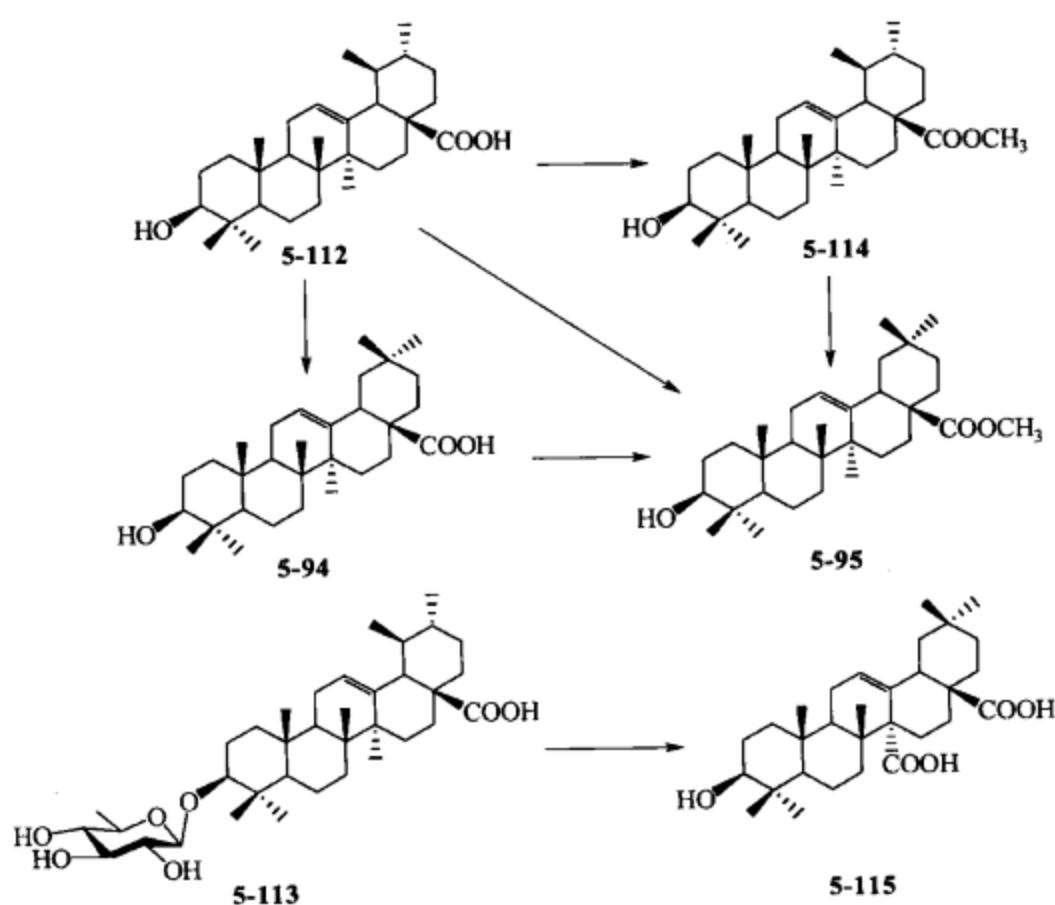


图 5-26 乌苏烷类型三萜与齐墩果烷类型三萜的微生物转化

Collin 等利用 *Mucor plumbers* ATCC 4740 转化熊果酸甲酯 (5-114) 得到了 1 个转化产物 $3\beta, 7\beta, 21\beta$ -三羟基-9(11), 12-乌萨二烯-28-酸甲酯 [methyl $3\beta, 7\beta, 21\beta$ -trihydroxyursa-9(11), 12-dien-28-oate, 5-116], 但在同样的条件下熊果醇 (uvaol) 却不能发生转化反应, 由此可见生物转化反应对底物的结构有较高的选择性 (图 5-27)。

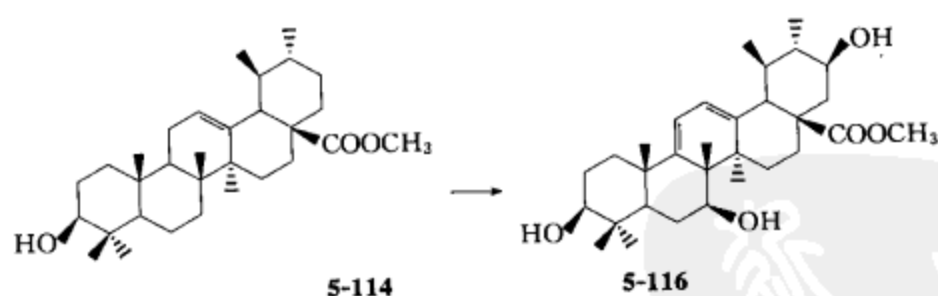


图 5-27 熊果酸甲酯的微生物转化

(3) 羽扇豆烷 (lupane) 型 Samir 等研究了 3 种微生物 *Bacillus megaterium* ATCC 14581、*Cunninghamella elegans* ATCC 9244、*Mucor mucedo* UI-4605 对白桦酸 (betulinic acid, 5-117) 的转化, 得到了 4 个转化产物: 白桦酮酸 (betulonic acid, 5-118), $3\beta, 7\beta$ -二羟基-20(29)-羽扇烯-28-酸 [$3\beta, 7\beta$ -dihydroxy-lup-20(29)-en-28-oic acid, 5-119], $3\beta, 6\alpha, 7\beta$ -三羟基-20(29)-羽扇烯-28-酸 (5-120) 和 $1\beta, 3\beta, 7\beta$ -三羟基-20(29)-羽扇烯-28-酸 (5-121)。同时还发现, 在一般的转化条件

下 *Bacillus megaterium* ATCC 14581 仅能转化产生 1 个产物 **5-120**，但在静态细胞悬浮液中加入肝药酶诱导剂苯巴比妥后可以诱导 *Bacillus megaterium* ATCC 14581 细胞中相关酶系的表达，转化产生 **5-118** 和 **5-119**。这说明微生物细胞中的酶系与哺乳动物细胞内的酶系有一定的类似，都具有一定的可诱导性。该发现对于寻求生物转化反应类型多样化和反应类型的可控性具有重要的研究价值（图 5-28）。

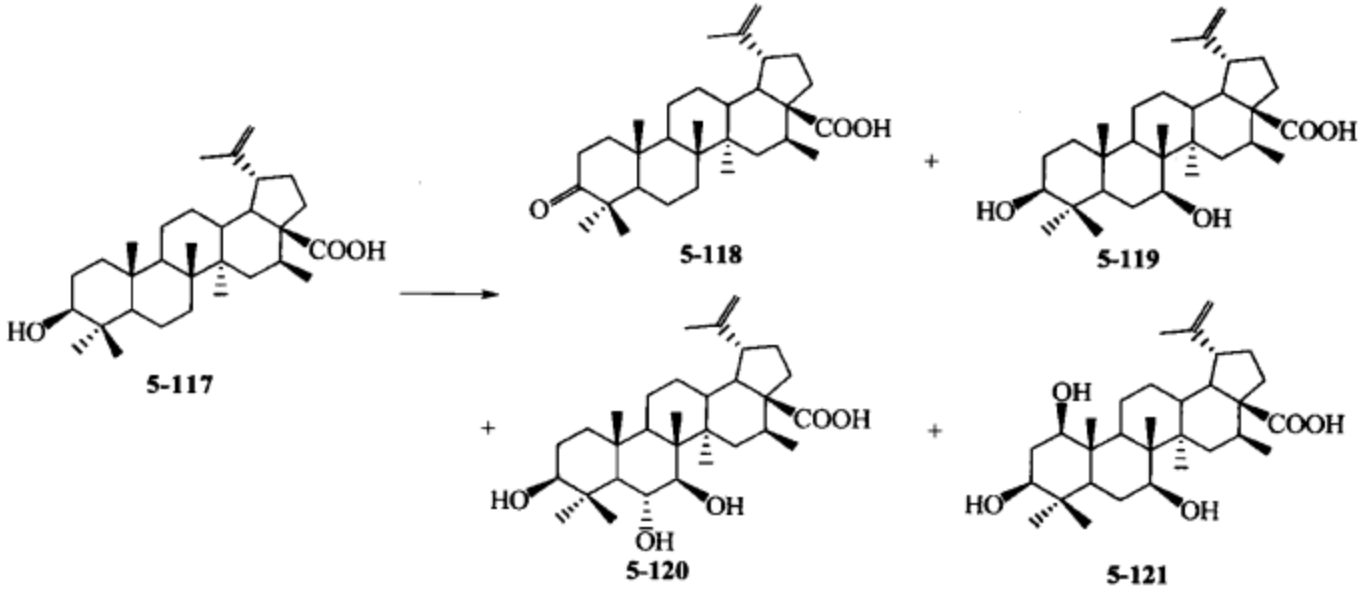


图 5-28 白桦酸的微生物转化

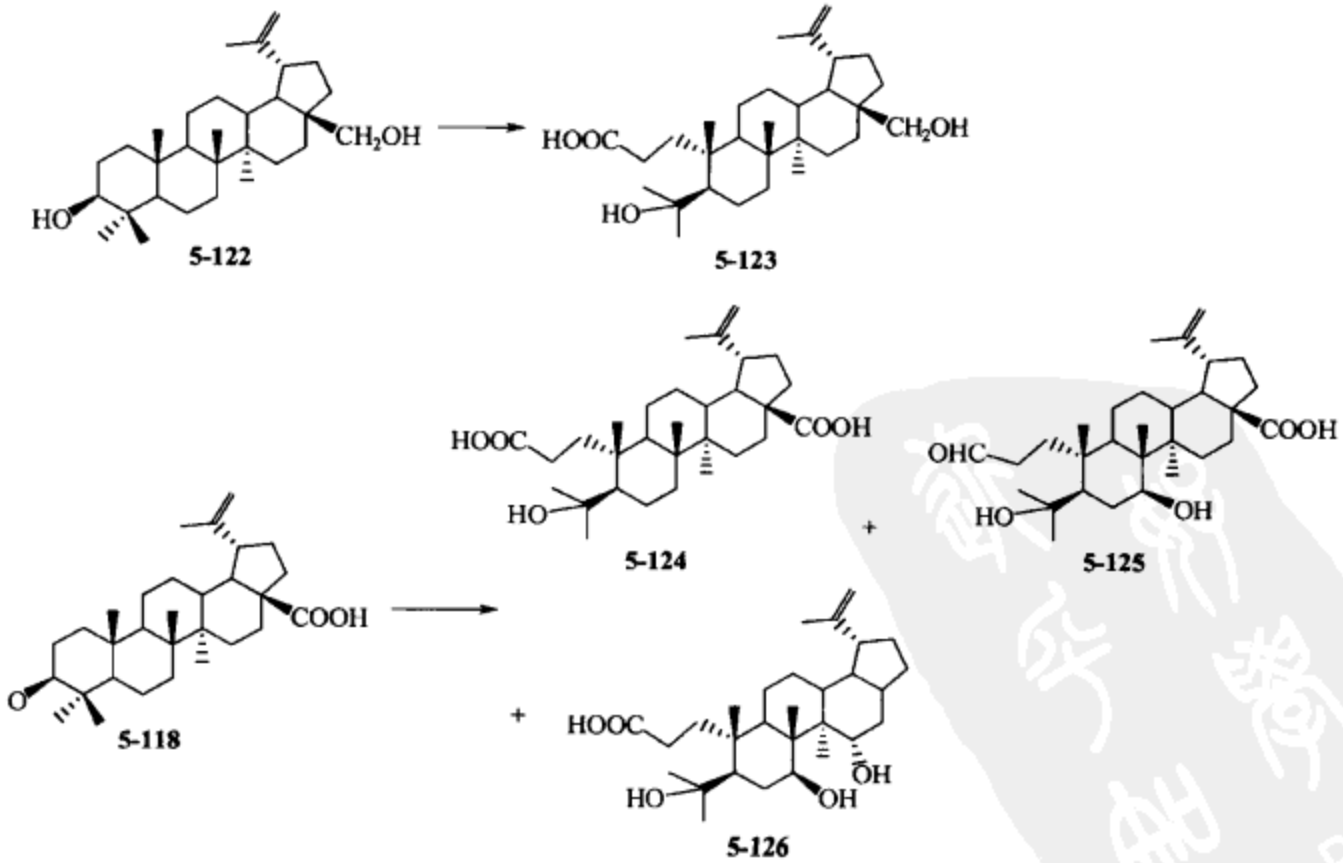


图 5-29 白桦醇和白桦酮酸的微生物转化

Akihisa 等报道了 *Chaetomium longirostre* 对白桦醇 (betulin, 5-122) 和白桦酮酸 (5-118) 的转化, 研究发现该菌株同样可以将羽扇豆烷型五环三萜的 A 环断裂, 5-122 被转化生成 4,28-二羟基-3,4-断-20(29)-羽扇烯-3-酸 [4,28-dihydroxy-3,4-seco-lup-20(29)-en-3-oic acid, 5-123], 化合物 5-118 转化生成 4-羟基-3,4-断-20(29)-羽扇烯-3,28-二酸 (5-124)、7 β ,15 α -二羟基-3-氧代-20(29)-羽扇烯-28-酸 (5-125) 和 4,7 β ,17-三羟基-3,4-断-28-去甲-20(29)-羽扇烯-3-酸 (5-126) (图 5-29)。

余伯阳等研究 *Nocardia* sp. 对白桦酸 (5-117) 和 3-羟基白桦酸 (5-127) 转化时发现该菌株可以转化生成白桦酸甲酯 (5-128)、3-羟基白桦酸甲酯 (5-129) (图 5-30)。

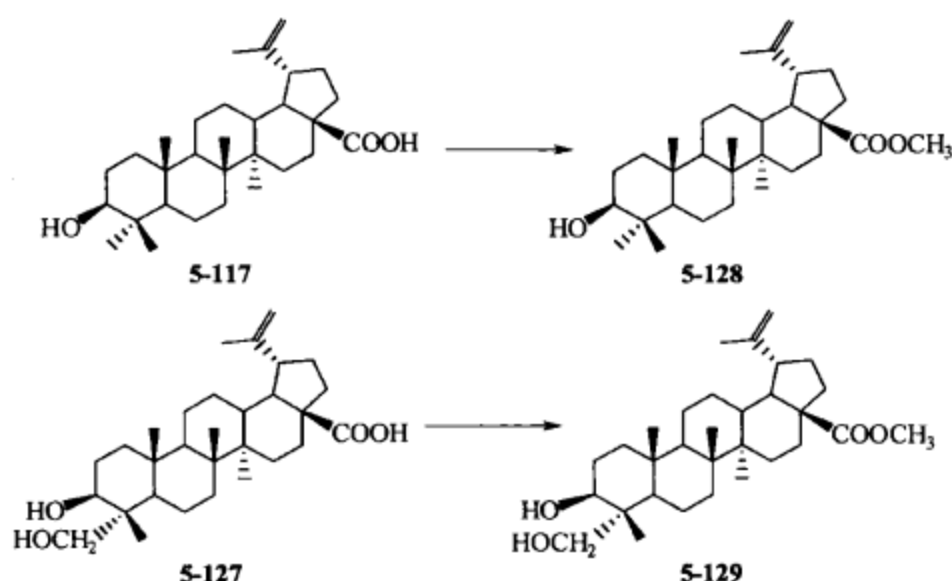


图 5-30 *Nocardia* sp. 对白桦酸和 3-羟基白桦酸的生物转化

5.3 药用植物生物碱类活性成分的生物转化

生物碱 (alkaloids) 是一类重要的天然产物。自 1086 年德国学者 F. W. Sertürner 从鸦片中分出吗啡碱 (morphine) 以后, 迄今已经从自然界分出了 16000 多种生物碱。其中许多重要的植物药如鸦片、麻黄、长春花、喜树、乌头、汉防己和延胡索等都主要含有生物碱成分, 并具有抗肿瘤、抗病毒等生物活性。这些药用植物的活性成分多数已经应用于临床治疗, 如长春花中的长春碱 (vinblastine) 和长春新碱 (vincristine) 已成为目前广泛应用的抗肿瘤药物, 喜树碱 (camptothecin, CPT) 是继紫杉醇后又一很有发展前途的植物性抗癌药, 麻黄素 (ephedrine) 临床上主要用于支气管哮喘、感冒、鼻黏膜肿胀以及低血压等疾病的治疗等。由于巨大的商业市场需求, 这几种类型的生物碱已经成为药用植物活性成分转化的研究热点。

5.3.1 长春花生物碱的生物转化

长春花 *Catharanthus roseus* (L.) G. Don 是夹竹桃科长春花属的一种重要的药用植物。它含有长春碱 (vinblastine)、长春新碱 (vincristine)、阿吗碱 (ajmalicine)、蛇根碱 (serpentine)、文多灵碱 (vindoline) 等 100 多种吲哚生物碱, 多数具有活性。其中长春碱和长春新碱可用于治疗恶性淋巴瘤、急性淋巴细胞型白血病、绒毛上皮细胞癌及其他癌症, 其硫酸盐已广泛应用于临床, 是目前应用最广泛的天然植物抗癌药物。但长春碱和长春新碱在植物体中的含量极其微小, 仅为百万分之几至十万分之几, 化学合成和半合成也不太理想。研究者从 20 世纪 60 年代就开始用组织培养生产长春花生物碱, 但目前只达到中试水平, 且只能生产用于降压和镇痛的吗啡碱、蛇根碱或长春质碱 (catharanthine)。

尽管对长春花生物碱的生物合成途径有了基本的认识, 然而对一些起到重要作用的关键酶的功能和调控机理仍然处在研究和推测阶段。

5.3.1.1 长春花生物碱的生物合成途径

吲哚生物碱的生物合成有如下两条途径组成。

(1) 甲戊二羟酸途径 (类萜途径) 经甲戊二羟酸途径 (MVA) 和甘油醛磷酸/丙酮酸途径 (MEP) 生成的香叶醇 (geraniol, 5-130) 在香叶醇 10-羟化酶 (G10H) 催化下生成 10-羟基香叶醇 (5-131), 再经过氧化形成 10-羟基香叶醛 (5-132)、10-氧代香叶醇 (5-133) 和 10-氧代香叶醛 (5-134)。在环化酶的作用下, 5-134 生成臭蚁二醛香叶醇 (5-135), 最后经多步反应生成 7-脱氧马钱子酸 (5-136)、马钱子酸 (5-137) 和马钱子苷 (loganin, 5-138), 最后生成裂环马钱子苷 (secologanin, 5-139) (图 5-31)。

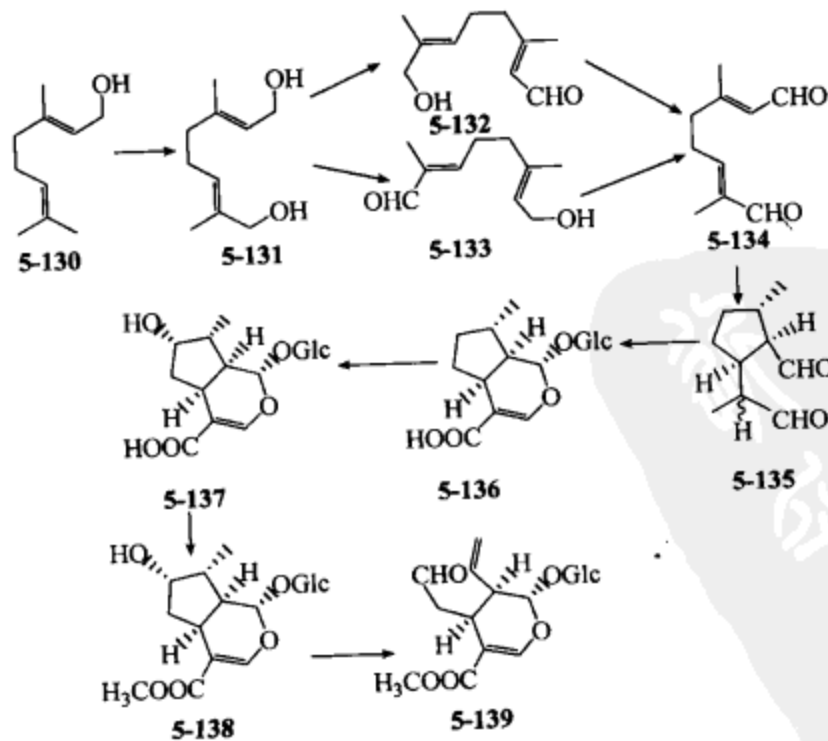


图 5-31 裂环马钱子苷的生物合成途径

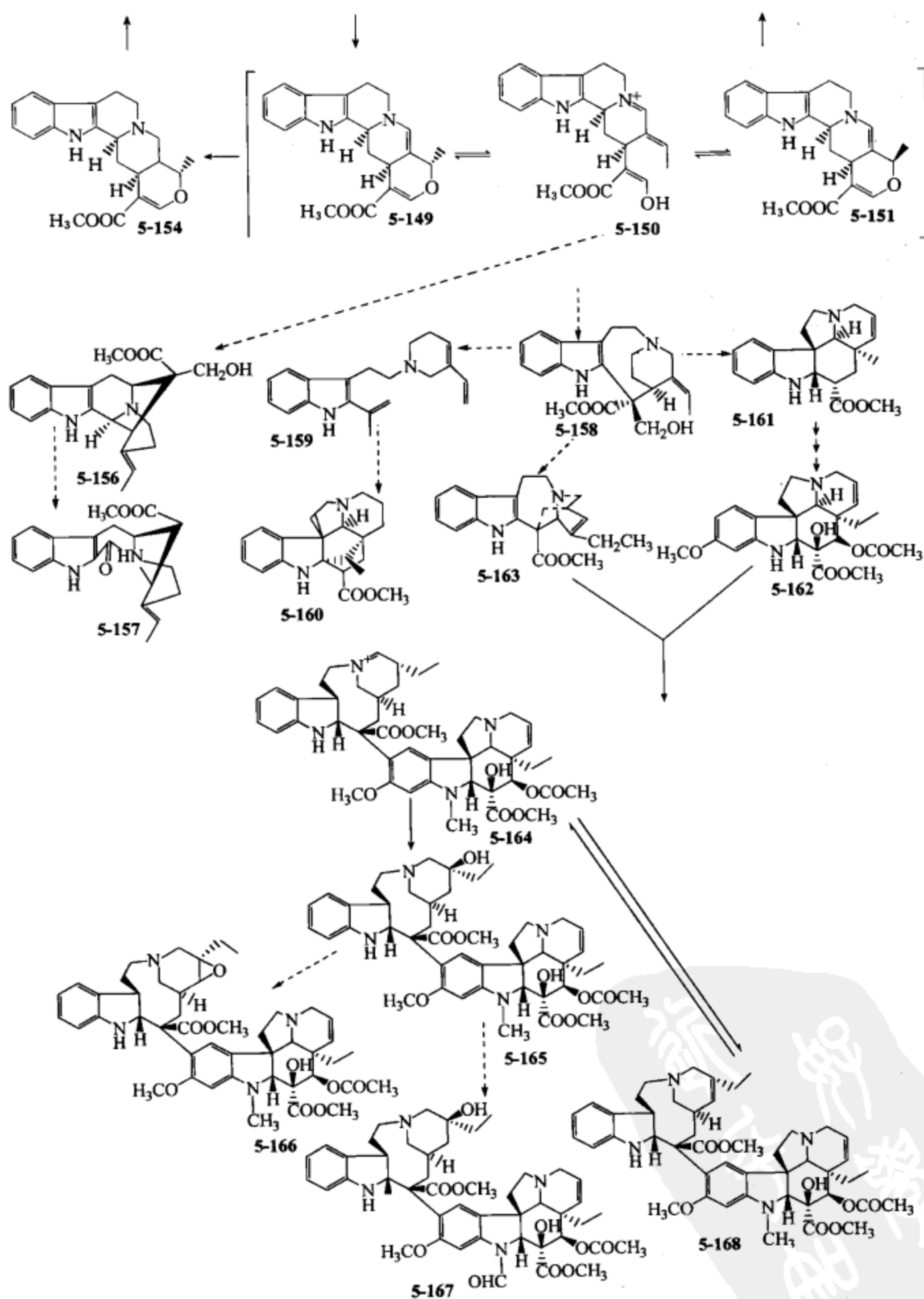


图 5-33 长春花中生物碱的生物合成途径

枯米定 (akuammidine, 5-156)、派利文碱 (perivine, 5-157)、花冠木碱 (stemmadenine, 5-158)、secodine (5-159)、长春里宁 (vindolinine, 5-160)、他波宁 (tabersonine, 5-161)、文朵灵 (vindoline, 5-162) 和长春质碱 (catharanthine, 5-163), 5-162 与 5-163 再经过 iminum (5-164) 生成长春花碱 (vinblastine, 5-165)、环氧长春碱 (leurosine, 5-166)、长春新碱 (vincristine, 5-167) 和 3',4'-脱水长春花碱 (3',4'-anhydrovinblastine, 5-168)。在此生物碱的生物合成途径中共有 30 多个步骤, 但只有 16 种酶得以鉴定 (图 5-33)。

5.3.1.2 长春花生物碱的转化

长春花生物碱的生物转化的研究主要集中在长春花生物碱的合成途径及其调控。其中涉及合成途径中的关键酶、细胞及毛根状培养条件的优化等, 并且长春花生物碱的转化研究基本上按照图 5-33 的生物合成途径确定相关的酶。

如 Furuya 等在长春花的悬浮细胞培养体系中饲喂他波宁 (5-161) 后, 发现生成了洛柯因 (lochnericine, 5-169) 和洛柯宁碱 (lochnerinine, 5-170), 检测不到文朵灵, 说明缺乏合成该化合物的关键酶 (图 5-34)。

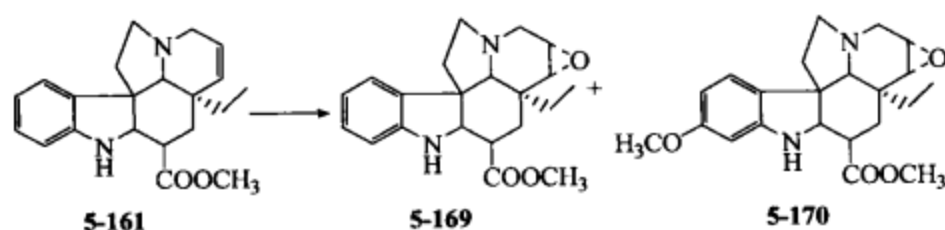
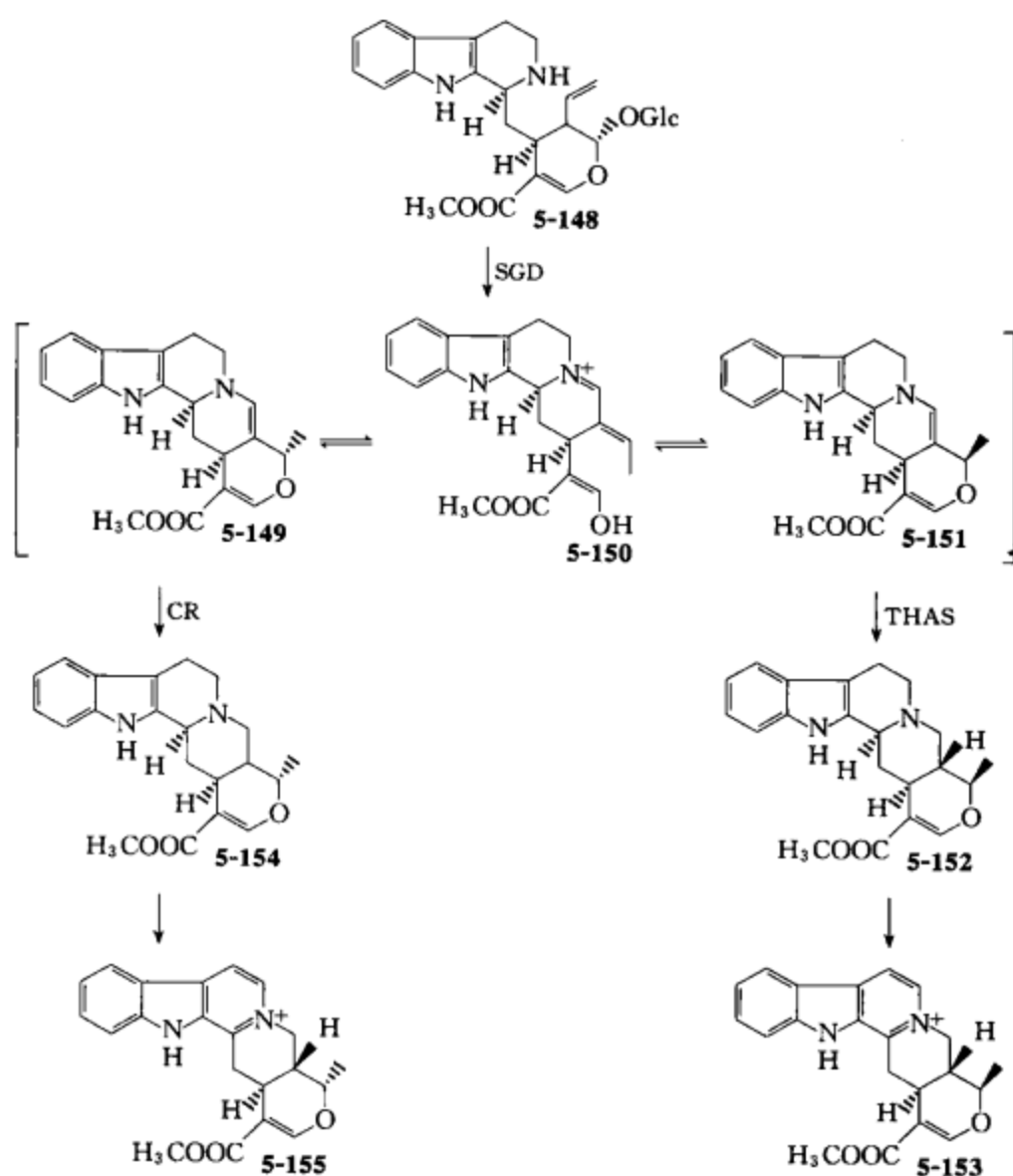


图 5-34 长春花悬浮培养对他波宁的转化

5-148 是长春花生物碱合成途径中的一个重要的中间产物, 异胡豆苷- β -D-葡萄糖苷酶 (strictosidine- β -D-glucosidase, SGD) 是吲哚类生物碱合成中非常重要的酶。该酶催化 5-148 脱去葡萄糖基生成 5-150。但 Stevens 等利用 ^1H NMR 技术进行检测, 发现中间产物是 5-149, 而不是 5-150, 说明这一步反应也是由糖苷酶催化的。5-149 在 cathenamine 还原酶 (CRs) 催化下转化成 5-154。Sierra 发现光照射长春花悬浮细胞后 5-155 的产量是暗培养的 20 倍, 因此认为 5-155 的积累与过氧化物酶 (peroxidase, POD) 活性呈正相关。四氢鸭脚木碱合成酶 (tetrahydroalstonine synthase, THAS) 催化 5-152 的转化。CR 和 THAS 都相继从长春花的悬浮细胞培养体系中被分离得到 (图 5-35)。

目前, 毛根状培养和细胞悬浮培养的瓶颈是缺乏 5-162 的积累, 从而导致 5-163 与 5-162 的偶合障碍而不能生成用于抗癌作用的长春碱和长春新碱。该合成途径支路共需要 6 种酶, 其中 5 种已鉴定 (图 5-36)。

5-162 的合成途径经过羟基化、O-甲基化、2,3-烯键羟基化、N-甲基化、4-羟基化和 4-O-乙酰化等过程。分别为: 第一步, 他波宁-16-羟化酶 (tabersonine-16-hydroxylase, T16H) 催化的羟基化反应生成 16-羟基他波宁 (16-hydroxytabersonine, 5-171); 第二步, O-甲基转移酶 (O-methyltransferase, OMT) 催化生成 16-



甲氧基他波宁 (16-methoxytabersonine, **5-172**); 第三步, 未知的羟化酶催化产生 16-甲氧基-2, 3-二氢化-2-氢-3-羟基他波宁 (16-methoxy-2, 3-dihydro-2-hydro-3-hydroxytabersonine, **5-173**); 第四步, *N*-甲基转移酶 (*N*-methyltransferase, NMT) 催化生成去乙酰基文朵灵 (desacetoxyvindoline, **5-174**); 第五步, 去乙酰基-4-羟化酶 (desacetoxyvindoline-4-hydroxylase, D4H) 催化生成脱乙酰文朵灵 (deacetylvindoline, **5-175**); 最后, 脱乙酰文朵灵-4-*O*-乙酰转移酶 (deacetylvin-doline-4-*O*-acetyltransferase, DAT) 催化生成 **5-162**。

喜树 (*Camptotheca acuminata* Decne) 是珙桐科 (Nyssaceae) 乔木, 为我国特有植物。1966 年, Wall 等从喜树茎的提取物中首次分离得到了有效成分喜树碱 (camptothecin, CPT) 及少量的 10-羟基喜树碱 (10-hydroxycamptothecin, 10-OH-CPT), 它们都具有较强的抗肿瘤和抗病毒活性。1985 年 Hsiang 等发现喜树

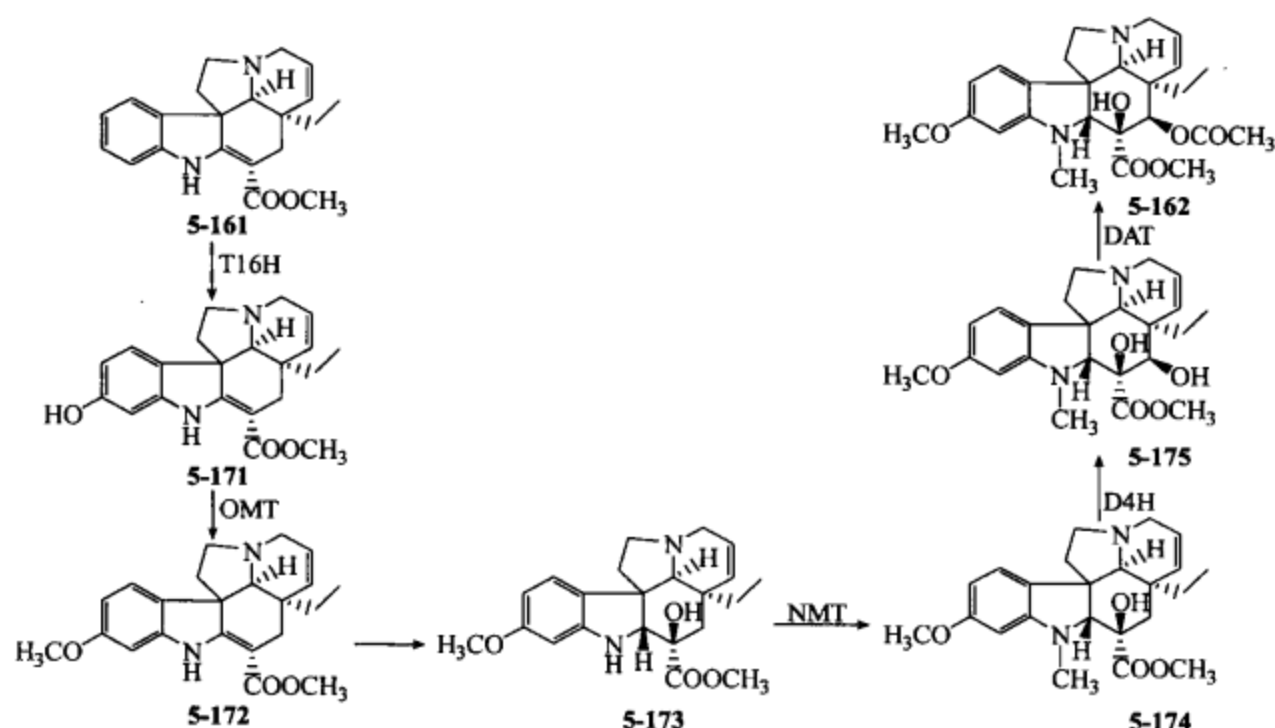


图 5-36 文朵灵的生物转化

碱及其衍生物是以拓扑异构酶 I 作为靶点抑制 DNA 的合成，从而对癌细胞产生致命的“打击”。研究表明，喜树碱及其衍生物对卵巢癌、膀胱癌、胃癌、结肠癌、直肠癌和白血病等多种恶性肿瘤均有一定的近期疗效，而 10-羟基喜树碱对多种动物肿瘤表现了更明显的抑制作用，且毒副作用更小。目前，半合成的水溶性喜树碱衍生物和美新 (topotecan, TPT) 和依列替康 (irinotecan, CPT-11) 已获美国 FDA (1992) 认证，是治疗卵巢癌和结肠癌的特效抗肿瘤药物。然而，由于喜树碱及其衍生物在植物体内含量甚微，且化学合成和半合成也不太理想，喜树碱的天然提取物价格昂贵且供应紧张，特别是喜树的天然资源有限且大量采伐涉及到未来的生态环境问题，所以应用生物技术对喜树进行组织细胞培养被认为是一条很有希望的替代途径而进行开发和利用。Sakato、Henge 等工作表明，通过细胞培养能够产生喜树碱，但由于细胞培养物中目标产物含量很小，因此至今未有实际生产的应用。

5.3.2.1 喜树碱的生物合成途径

喜树碱在结构上属于一种喹啉类生物碱，其合成来自单萜吲哚生物碱。吲哚生物碱的生物合成来源于两个不同途径：莽草酸途径和甲羟戊酸途径。由莽草酸途径来的色胺 (5-147) 与由甲羟戊酸途径经多步反应生成的裂环马钱子苷 (5-139) 结合形成异胡豆苷 (5-148)。异胡豆苷转变成异长春花苷内酰胺 strictosamide (strictosamide, 5-176)。之后的中间代谢步骤还没有被阐明，推测可能包括一个吲哚环的氧化体系，导致喜树碱 B 环和 C 环的重排，所形成的中间产物已证明是 3(S)-短小蛇根草苷 [3(S)-pumiloside, 5-177]。它最先是从小蛇根草 (*Ophiorrhiza pumila*) 的植物中提取并鉴定出来的 (这种植物不仅含有喜树碱，而

且还含有其他多种喜树碱衍生物),之后在喜树中也得到证实。这种中间产物之所以被命名为短小蛇根草苷(pumiloside),是为了表明其来自于裂环马钱子苷的葡萄糖单位在以后的过程中仍然保留,而其他的吲哚生物碱在异胡豆苷合成之后葡萄糖单位迅速脱去。接下来的可能中间体是 3(S)-脱氧短小蛇根草苷[3(S)-doxypumiloside, 5-178],迄今为止仅在短小蛇根草中有过报道。从脱氧短小蛇根草苷到喜树碱(5-179)之间约还有 3 个其他的中间体,但仍未最终确定(图 5-37)。

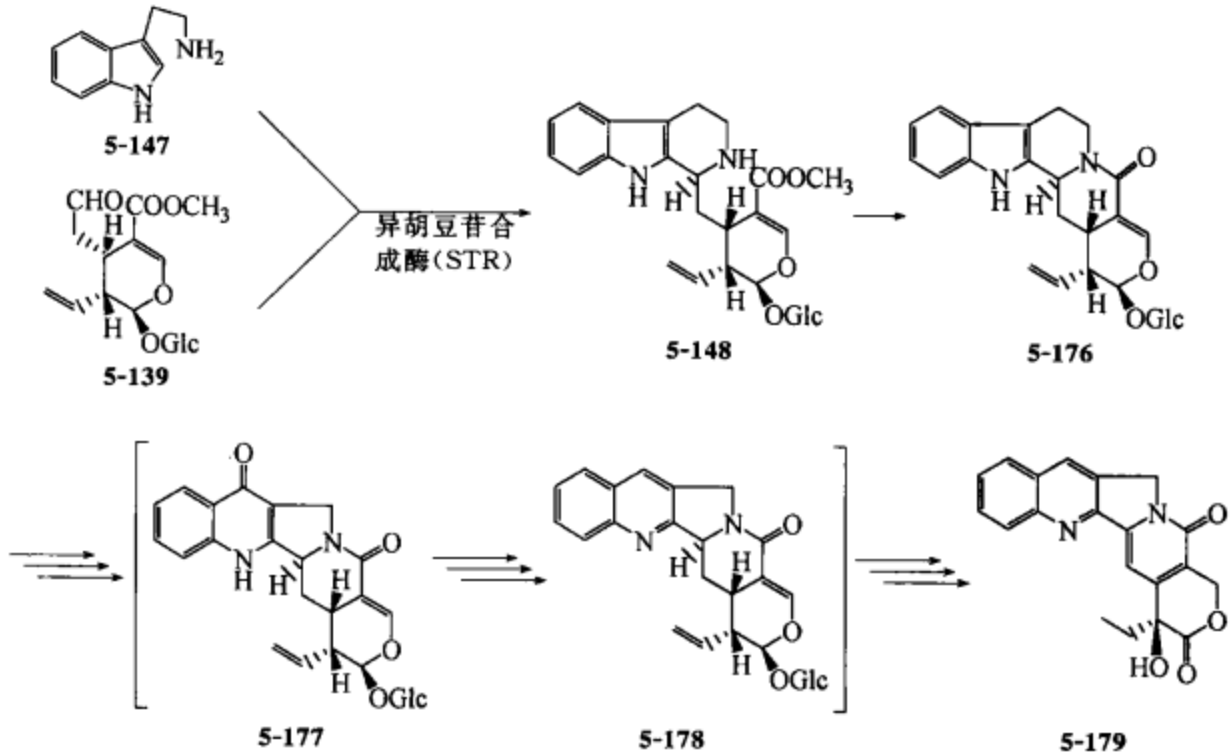


图 5-37 喜树碱的生物合成途径

5.3.2.2 喜树碱的生物转化

喜树的天然资源有限,喜树碱及其衍生物在植物体内的含量甚微。喜树碱的化学合成和半合成研究虽然在近几年来取得了一定进展,但成本仍然很高。因此,利用生物工程技术进行组织细胞培养被认为是一条很有希望的替代途径。研究表明,喜树碱可以由组织细胞培养产生。早在 1974 年,日本的 Sakato 等首次报道了利用组织培养的方法以喜树的茎为外植体诱导出愈伤组织,其喜树碱的含量为 0.0025%,是原植株的 1/20。1992 年, Hengel 等也从喜树的茎得到愈伤组织,并实现了细胞悬浮培养,悬浮培养生产的喜树碱可达到 0.0031%,达到了整株的水平。另外,韩国的 Song 等也报道了同样的结果,并发现添加酵母粉、茉莉酸以及茉莉酸甲酯也可显著增加喜树碱的含量。在研究中还发现,在愈伤组织中有根形成时,喜树碱的含量明显提高。这不仅表明细胞的分化对喜树碱的生物合成具有重要的影响,而且预示着研究次生代谢产物合成和积累的细胞和组织特异性对有效调控次生代谢产物的生产意义重大。应当指出,除了喜树之外,喜树碱及其衍生物在其

他的植物中也可形成，例如 *Ervatamia heyneana* (夹竹桃科)、*Nothapodytes foetida* (茶茺萸科)、*Merrilliodendron megacarpum* (茶茺萸科) 和 *Ophiorrhiza pumila* (茜草科) 等。这不仅为喜树碱类抗癌药物的生产提供了广阔的药源，而且也为喜树碱的生物合成研究提供了多样的种属材料。

喜树碱微生物转化研究最多的就是利用曲霉 T-36 和 T-419 将喜树碱 (5-179) 转化成羟基喜树碱 (5-180)，转化率在 50% 左右 (图 5-38)。

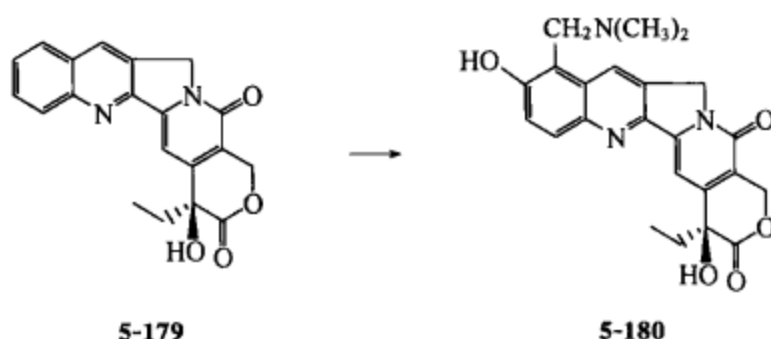


图 5-38 喜树碱的微生物转化

5.3.3 其他类型生物碱的生物转化

5.3.3.1 麻黄碱的生物转化

麻黄碱是从我国特有的药用植物麻黄中提取获得的一种生物碱。临床上用于支气管哮喘、感冒、鼻黏膜肿胀以及低血压等疾病的治疗，为一种常用中成药的原料药。近年来，这种药用植物因过度采挖，造成土地沙化，自然生态遭到破坏，使天然左旋麻黄碱原料严重不足。同时，天然左旋麻黄碱在麻黄中含量也有所下降。寻找新方法生产天然左旋麻黄碱已是当务之急。目前，利用生物合成技术与有机化学合成技术结合起来，成功地研究出一条不再依靠天然麻黄草资源来获取左旋麻黄碱的理想途径——左旋麻黄碱半生物合成技术。采用酵母细胞生物催化，将丙酮酸 (5-181) 和苯甲醛 (5-182) 缩合形成 L-苯基乙酰基甲醇 (L-phenylacetylcarbinol, 5-183, L-PAC)。L-PAC 是合成过程中的一个关键中间体，可经过甲胺还原胺化得到 L-麻黄碱 (5-184) (图 5-39)。

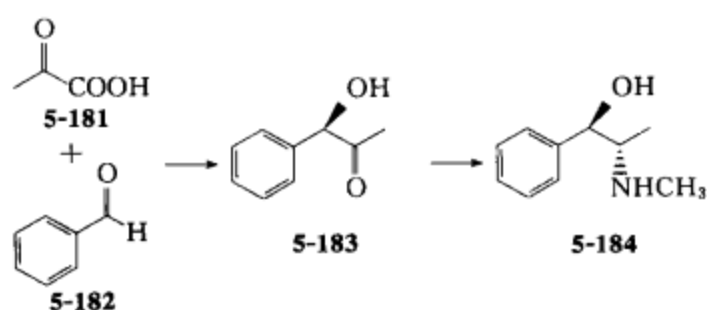


图 5-39 麻黄碱的生物转化

同时，天然左旋麻黄碱在麻黄中含量也有所下降。寻找新方法生产天然左旋麻黄碱已是当务之急。目前，利用生物合成技术与有机化学合成技术结合起来，成功地研究出一条不再依靠天然麻黄草资源来获取左旋麻黄碱的理想途径——左旋麻黄碱半生物合成技术。采用酵母细胞生物催化，将丙酮酸 (5-181) 和苯甲醛 (5-182) 缩合形成 L-苯基乙酰基甲醇 (L-phenylacetylcarbinol, 5-183, L-PAC)。L-PAC 是合成过程中的一个关键中间体，可经过甲胺还原胺化得到 L-麻黄碱 (5-184) (图 5-39)。

5.3.3.2 延胡索素的生物转化

延胡索素是罂粟科植物延胡索 (*Corydalis yanhusuo* W. T. Wang) 块茎中 20 多种生物碱的总称，主要有延胡索甲素 (*d*-corydaline)、延胡索乙素 (*dl*-tetrahydropalmatine, THP)、延胡索丙素 (protopine)、延胡索丁素 (*l*-tetrahydrocoptisine)、延胡索戊素 (*dl*-tetrahydrocoptisine)、延胡索己素 (*l*-tetrahydrocolum-

bamine) 等。延胡索素具有镇痛、镇静、安定及催眠的作用,对慢性持续性疼痛效果较好,对创伤及手术后疼痛的作用较差,有明显的镇静催眠作用。

张文伟等对链霉菌等 10 株菌株进行筛选,发现灰色链霉菌 (*Streptomyces griseus* YD-007) 可以转化延胡索素,将左旋延胡索乙素 (5-185) 转化为左旋紫堇达明 (L-corydalmin, 5-186) (图 5-40)。

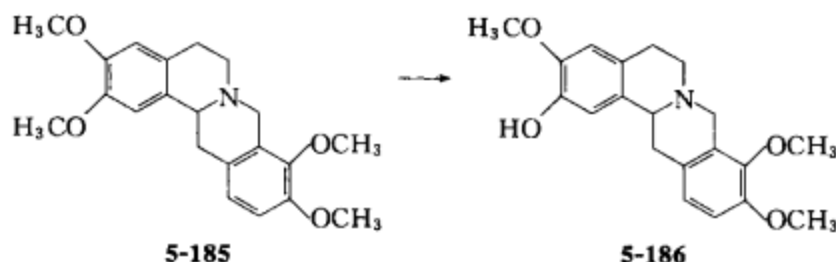


图 5-40 延胡索乙素的生物转化

5.4 药用植物甾体类活性化合物的生物转化

目前常用的甾体药物有 50 多种,在治疗风湿性关节炎、红斑狼疮以及计划生育等方面有着举足轻重的作用,是医药工业的一大门类药物。近年来其应用范围不断扩大,如用于心血管疾病、抗肿瘤、利尿、麻醉、肌松以及与内分泌有关的老年性疾病。此外,甾体激素亦被拓展应用于促进家畜繁殖生长及植物生长。国外甾体激素类药物主要是以薯蓣皂苷元为起始原料,我国除了生产大量的薯蓣皂苷元以供合成皮质激素外,还从番麻和剑麻的叶汁中提取番麻皂苷元。激素类药物若采用化学方法加工而成,合成步骤多、收率低、价格昂贵,而且某些关键反应用化学方法很难进行,例如将含氧的官能团导入甾体母核的 C-11 位特别困难,在天然资源中又找不到类似的甾体化合物,这曾一度影响了甾体药物的工业化生产。Pettersson 等于 1952 年应用黑根霉 (*Rhizopus nigricans*) 成功地把 11 α -羟基化导入孕甾酮的 C-11 位,解决了可的松生产过程中的最大困难。从此,微生物甾体转化技术引起了人们的高度重视,并得到不断完善和发展,其作用也日益广泛。目前,涉及甾体药物的微生物转化反应主要有 C-11 α -羟基化、C-11 β -羟基化及 C-16 α -羟基化、A 环的 C-1,2 位及 C-4,5 位脱氢、A 环的芳构化、C-3-OH 的脱氢以及甾醇的边链切除等。

5.4.1 甾体化合物的生物合成

药用植物中甾体类化合物的生物合成早期阶段是从甲戊二羟酸途径开始的,形成角鲨烯 (squalene, 5-187) 和 2,3-环氧角鲨烯 (2,3-oxidosqualene, 5-188), 然后由环阿屯醇 (cycloartenol, 5-189) 开始,经过 24(28)-亚甲基环阿屯醇 (5-190)、24-甲基链甾醇 (24-methyl-desmosterol, 5-191)、24 β -甲基胆甾醇 (24 β -methyl-cholesterol, 5-192)、4-甲基-24(28)-亚甲基-7-烯胆(甾)烷醇 [24(28)-methylene-

phenol, **5-193**] 和 4-甲基-24(28)-亚乙基-7-烯胆(甾)烷醇[24(28)-ethylidenelophenol, **5-194**], 生成谷甾醇 (sitosterol, **5-195**) (图 5-41)。

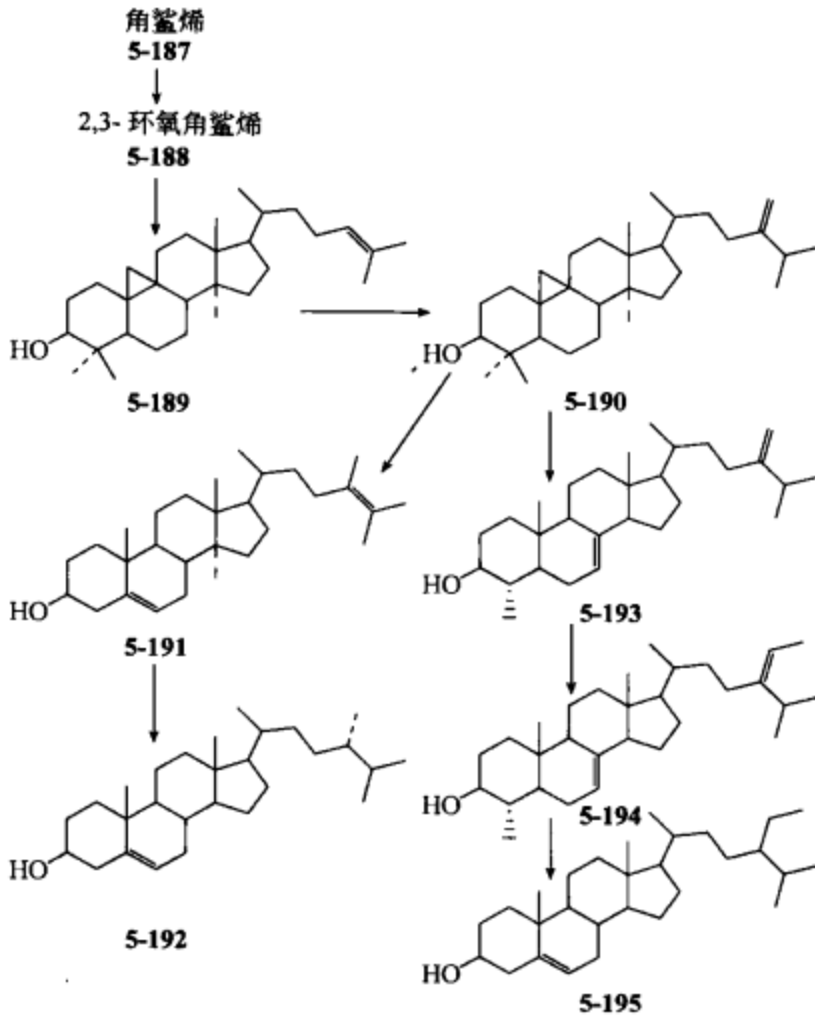


图 5-41 植物甾体类的生物合成途径

5.4.2 甾体类活性成分的生物转化

(1) 脱氢表雄酮类化合物的生物转化 Liu 等利用长春花 [*Catharanthus roseus* (L.) G. Don] 悬浮细胞培养体系对脱氢表雄酮 (dehydroepiandrosterone, **5-196**) 进行了生物转化, 生成 4-雄甾烯-3,17-二酮 (**5-197**), **5-197** 羟基化产物 6α -羟基-4-雄甾烯-3,17-二酮 (**5-198**)、 6β -羟基-4-雄甾烯-3,17-二酮 (**5-199**) 和 14α -羟基-4-雄甾烯-3,17-二酮 (**5-200**), 还原产物 17β -羟基-4-雄甾烯-3-酮 (**5-201**); **5-201** 再经过羟基化产生 $6\alpha, 17\beta$ -二羟基-4-雄甾烯-3-酮 (**5-202**)、 $15\alpha, 17\beta$ -二羟基-4-雄甾烯-3-酮 (**5-203**) 和 $15\beta, 17\beta$ -二羟基-4-雄甾烯-3-酮 (**5-204**), 羰基化生成 17β -羟基-4-雄甾烯-3,16-二酮 (**5-205**) (图 5-42)。

刘颖等利用唐古特山莨菪毛状根培养进行脱氢表雄酮生物转化, 得到氧化产物 4-雄甾烯-3,17-二酮 (**5-197**)。5-197 再经过还原生成 17β -羟基-4-雄甾烯-3-酮 (**5-201**); 6α -羟基化生成 6α -羟基-4-雄甾烯-3,17-二酮 (**5-198**)。5-198 氧化成 4-雄甾烯-3,6,17-三酮

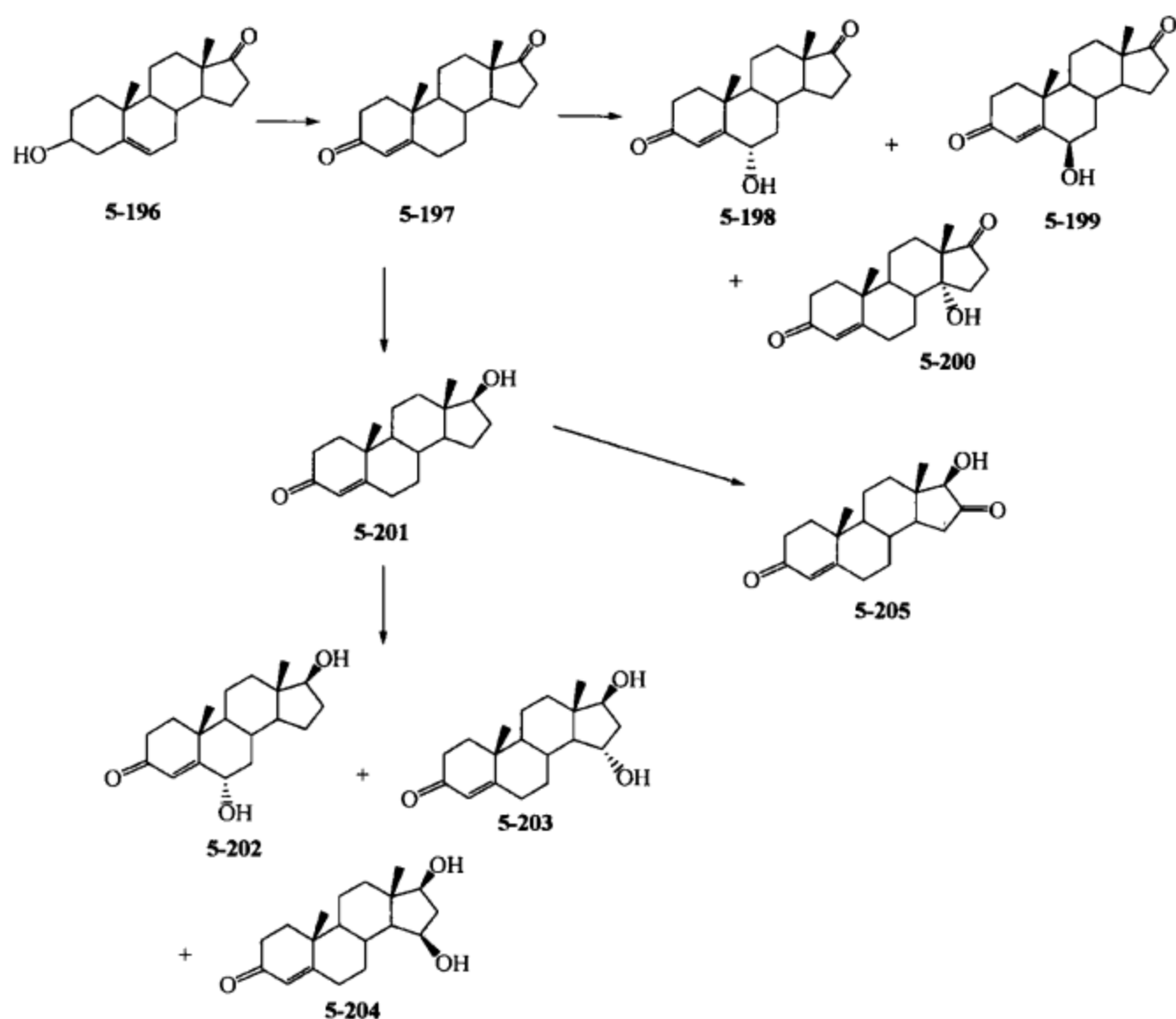


图 5-42 长春花悬浮细胞培养体系对脱氢表雄酮的生物转化

(5-206), 并可以还原成 $6\alpha, 17\beta$ -二羟基-4-雄甾烯-3-酮 (5-202) (图 5-43)。

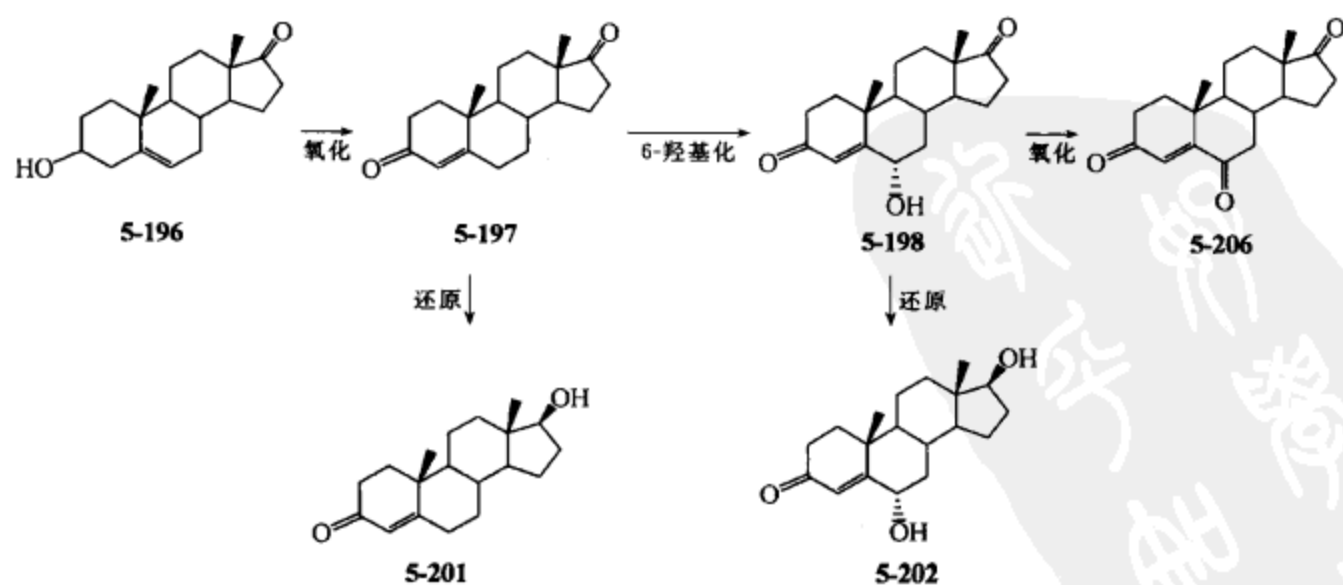


图 5-43 唐古特山茛菪毛状根培养进行脱氢表雄酮生物转化

葛文中等利用总枝状毛霉 (*Mucor racemosus*) 对脱氢表雄酮生物转化, 得到 7 α -羟基脱氢表雄酮 (5-207) 和 7 β -羟基脱氢表雄酮 (5-208) (图 5-44)。

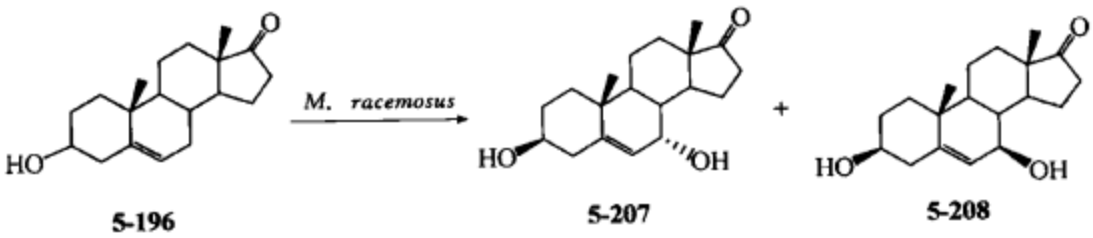


图 5-44 总枝状毛霉 (*Mucor racemosus*) 对脱氢表雄酮生物转化

(2) 洋地黄苷的生物转化 Doller 等应用洋地黄 (*D. lanata*) 悬浮细胞培养体

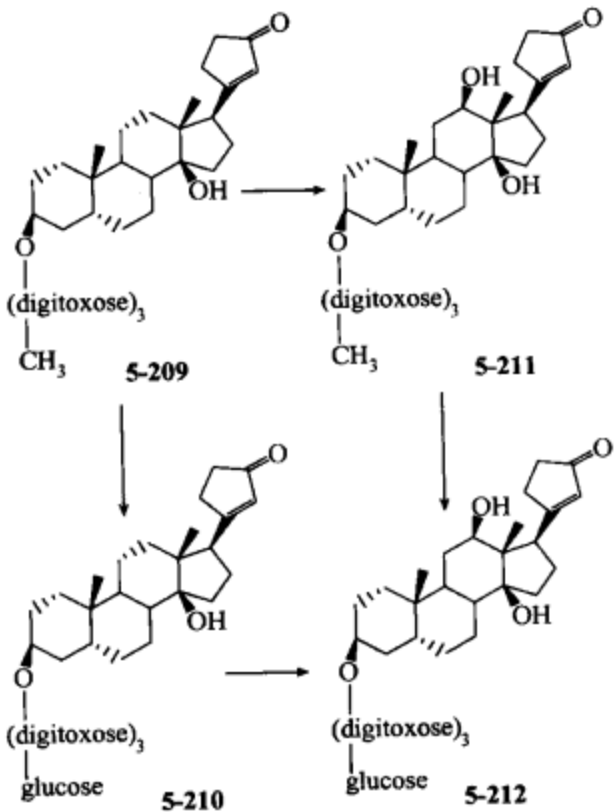


图 5-45 洋地黄悬浮细胞培养对 β -甲基洋地黄苷的生物转化

digitoxose—洋地黄毒糖; glucose—葡萄糖

系对 β -甲基洋地黄苷 (β -methyldigitoxin, 5-209) 进行转化, 生成紫花洋地黄苷 A (purpurea glycoside A, 5-210)。另外 5-209 和 5-210 还可以发生 12-羟基化, 生成 β -甲基地高辛 (β -methyldigoxin, 5-211) 和去乙酰毛花洋地黄苷 C (deacetyl-lanatoside C, 5-212) (图 5-45)。

Gengronella urolifera CBC 菌株可以将洋地黄苷 (digitoxin, 5-213) 的 C-7 位上导入羟基形成 7 β -羟基洋地黄苷 (5-214); 而 *Penicillium chrysoginicum* 能专一性地水解去掉洋地黄毒糖, 生成洋地黄苷元 (digitoxigenin, 5-215) (图 5-46)。

(3) 其他甾体类活性化合物的生物转化 冯冰等利用新月弯孢霉 [*Curvularia lunata* (Waldl) Boed] 对重楼皂苷进行转化, 发现新月弯孢霉 [*C. lunata* (Waldl)

Boed 3. 4381] 可选择性水解甾体皂苷 C-23 位糖链末端糖基的微生物, 将重楼皂苷 V (5-216) 和重楼皂苷 VI (5-217) 进行生物转化, 分别生成薯蓣皂苷元-3-O- β -D-吡喃葡萄糖苷即延龄草苷 (trillin, 5-218) 和偏诺皂苷元-3-O- β -D-吡喃葡萄糖苷 (5-219) (图 5-47)。

Quan 等利用米曲霉 (*Aspergillus oryzae*) 对原薯蓣皂苷 (5-220) 进行生物转化, 发现米曲霉可以在保留原薯蓣皂苷的 26 位葡萄糖情况下水解 3 位末端 1 $\rightarrow$ 2 鼠李糖生成次生呋甾皂苷 (5-221)。米曲霉也可以去掉原薯蓣皂苷的 22 位羟基生成 20(22)-双键的呋甾皂苷产物 (5-222, 5-223), 并且继续水解 26 位葡萄糖生成相

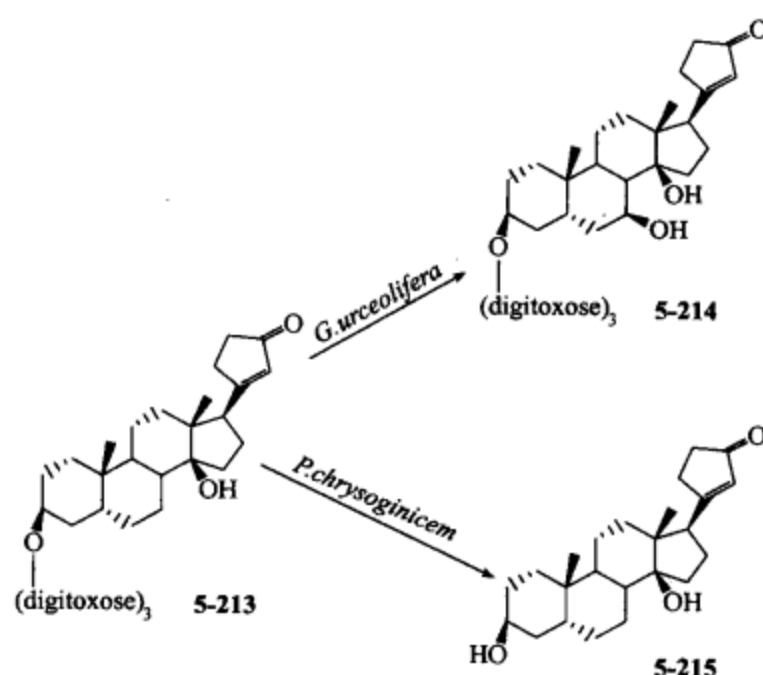


图 5-46 *Gengronella urolifera* 和 *Penicillium chrysoginicum* 对洋地黄苷的转化

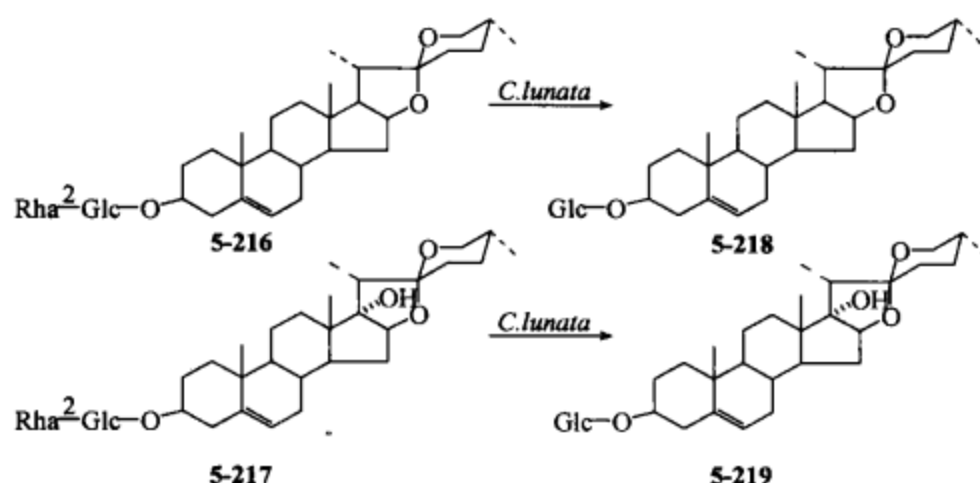


图 5-47 米曲霉 (*Aspergillus oryzae*) 对重楼皂苷的生物转化

应的螺甾皂苷 (5-224, 5-225) (图 5-48)。

葛文中等利用总枝状毛霉 (*Mucor racemosus*) 对 4-烯-3-酮类甾体衍生物进行生物转化, 转化条件为 27℃, 220r/min 振荡培养 4 天。得到黄体酮 (5-226) 生物转化产物 14α-羟基-4-孕甾烯-3,20-二酮 (5-227) 和 7α,14α-二羟基-4-孕甾烯-3,20-二酮 (5-228)。4-雄甾烯二酮 (5-229) 的转化产物是 14α-羟基-4-雄甾烯-3,17-二酮 (5-200)、14α,17β-二羟基-4-雄甾烯-3-酮 (5-230) 和 6α,17β-二羟基-4-雄甾烯-3-酮 (5-202) (图 5-49)。

甾体生物转化技术已经有几十年的发展历史, 人类在这方面积累了许多宝贵经验。转化的成功不仅需要优良菌种, 转化体系也至关重要。近年来, 对转化体系的研究是生物转化领域的一大热点, 但距离真正的工业化生产还有差距。

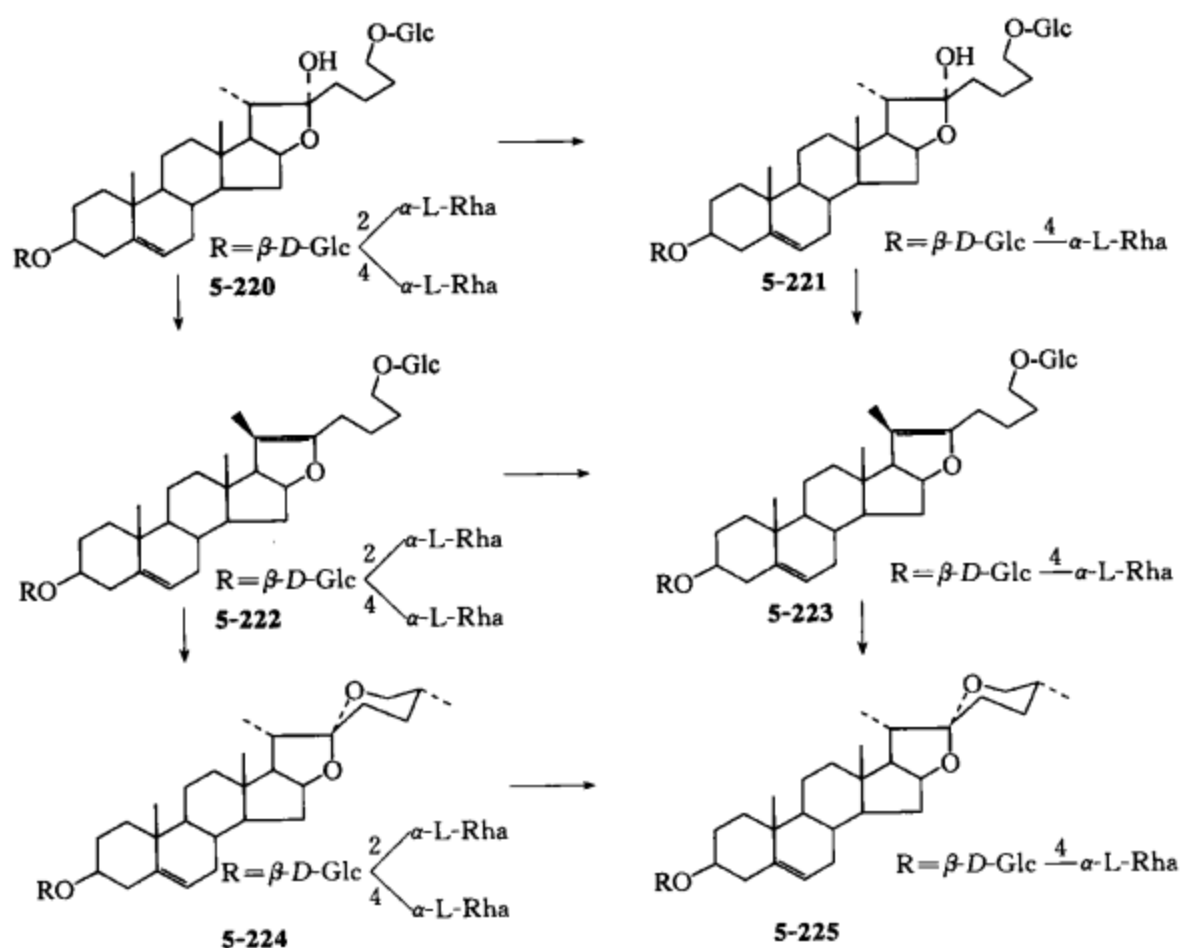


图 5-48 米曲霉 (*Aspergillus oryzae*) 对原薯蓣皂苷的生物转化

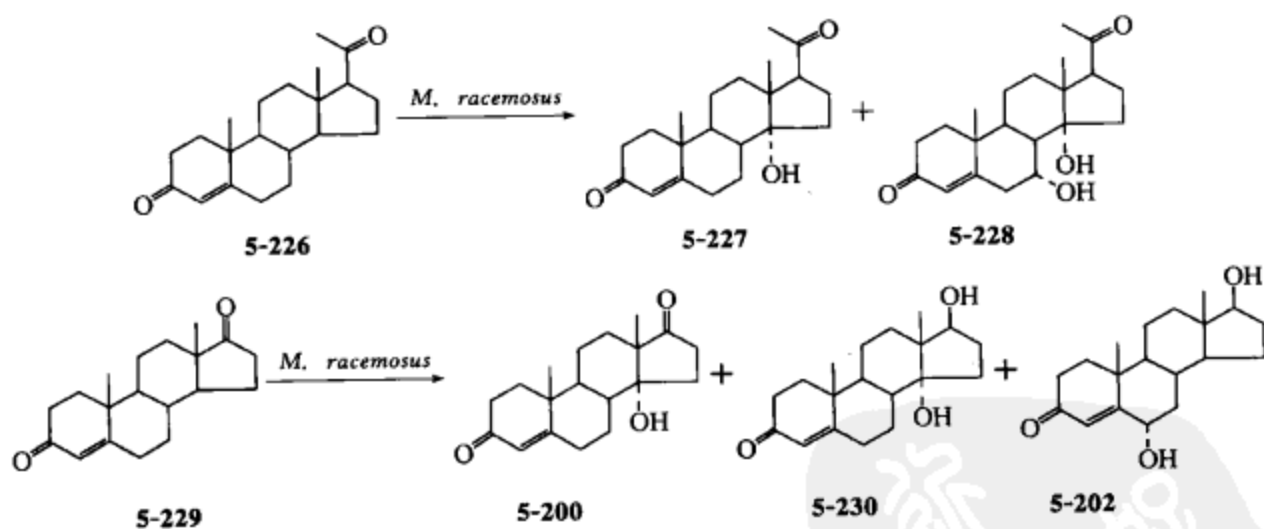


图 5-49 总枝状毛霉对 4-烯-3-酮类甾体衍生物的生物转化

5.5 其他类型活性成分的生物转化

5.5.1 黄酮类化合物的生物转化

截至 2000 年统计, 黄酮类化合物总数已超过 8000 个, 其结构类型之多是由于其每

一类别的羟基、甲基等取代基的数量和位置不同所致。黄酮类化合物通常是以糖苷的形式存在，也有游离体（苷元）；除 O-糖苷（O-glycosides）外，天然黄酮类化合物中还发现有 C-糖苷，如葛根素（puerarin）、葛根素木糖苷（puerarin xyloside）等。

黄酮类化合物广泛分布于植物界，具有多种多样的生物学功能，如调节植物生长、保护植物免受紫外线的损伤、在植物-微生物相互作用中的信号转导功能以及作为植物抗毒素（phytoalexin）和活性氧清除剂等。同时，黄酮类化合物也是药用植物的主要活性成分之一，相关药理和临床实验证实黄酮类化合物具有广泛的药理功能，如不少治疗冠心病有效的中草药或活血化瘀类中药中均含有黄酮类化合物，芦丁（rutin）、槲皮素（quercetin）、葛根素（puerarin）、人工合成的立可定（recordin）等均有明显的扩冠作用，并已用于临床；染料木素（genistein）、金雀花异黄酮（5,7-dihydroxy-4'-methoxyisoflavone）、大豆苷元（daidzein）等均有雌性激素样作用。还有许多黄酮类化合物被证明有抗癌和抗 HIV 病毒活性。由于该类化合物具有多种多样的生物活性和药理功能，且毒性较低，备受人们重视，成为研究和开发利用的重点。

Demetzos 等利用黑曲霉（*Aspergillus niger*）对银槲苷（tiliroside, 5-231）进行生物转化，得到银槲苷-7-甲基醚（5-232）（图 5-50）。

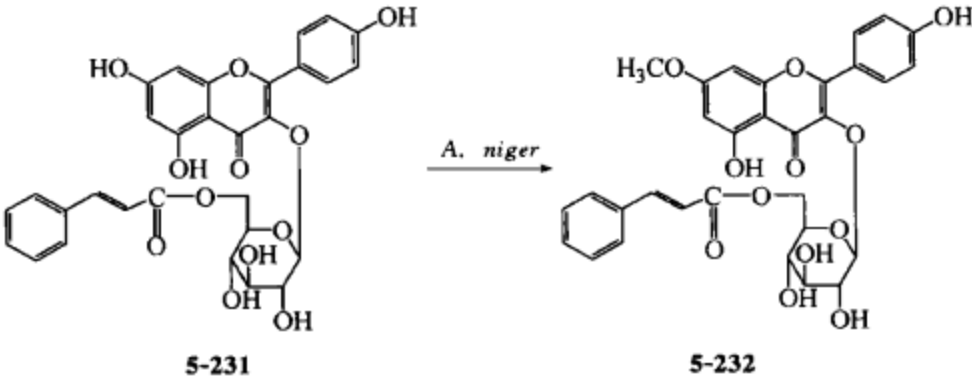


图 5-50 黑曲霉（*Aspergillus niger*）对银槲苷的生物转化

杨秀伟等利用人肠道厌氧菌对罗汉果（*Siraitia grosvenori*）中的山柰苷进行生物转化，得到了山柰酚-3-O- α -L-吡喃鼠李糖苷（阿福豆苷）、山柰酚-7-O- α -L-吡喃鼠李糖苷、山柰酚和对羟基苯甲酸。

谢明杰等从酒曲中分离得到 *Absidia* sp. 菌株，可以将染料木苷（genistin, 5-233）转化为染料木素（5-234）（图 5-51）。

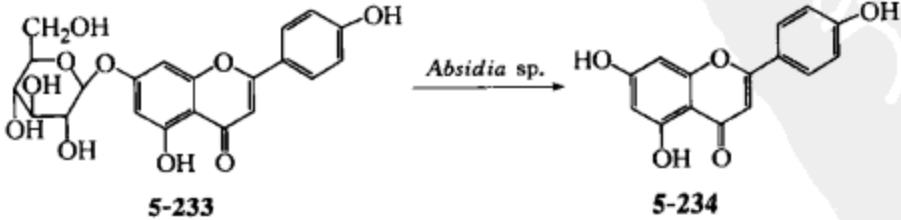


图 5-51 *Absidia* sp. 对染料木苷的生物转化

王园园等利用链霉菌蛋白酶 *Streptomyces griseus* ATCC 13273 对柚皮苷、橙皮苷、黄芩苷及木犀草素进行生物转化，分别得到柚皮素-7-O-葡萄糖苷、柚皮素、橙皮素、黄芩素、黄芩素-6-甲氧醚和柯伊利素。结果表明此转化过程中涉及到两类反应，即糖苷水解和甲基化反应，*S. griseus* 产生的 L-鼠李糖苷酶对底物有特异选择性， β -D-葡萄糖苷酶的选择性不强，甲基转移酶仅对黄酮类化合物的特异位点进行甲基化。

5.5.2 蒽醌类化合物

蒽醌类化合物是一类重要的天然产物，广泛分布在高等植物中，尤其是茜草科 (Rubiaceae)、鼠李科 (Rhamnaceae)、蓼科 (Polygonaceae) 和豆科 (Leguminosae) 等科植物。

Zhang 等利用刺囊毛霉 *Mucor spinosus* 对大黄酚 (5-235)、大黄素甲醚 (5-236) 和大黄素 (5-237) 进行转化，分别得到大黄酚-8-O- β -D-葡萄糖苷 (5-238)、大黄素甲醚-8-O- β -D-葡萄糖苷 (5-239)、大黄素甲醚-1-O- β -D-葡萄糖苷 (5-240) 和 ω -羟基大黄素 (5-241) 4 个产物 (图 5-52)。

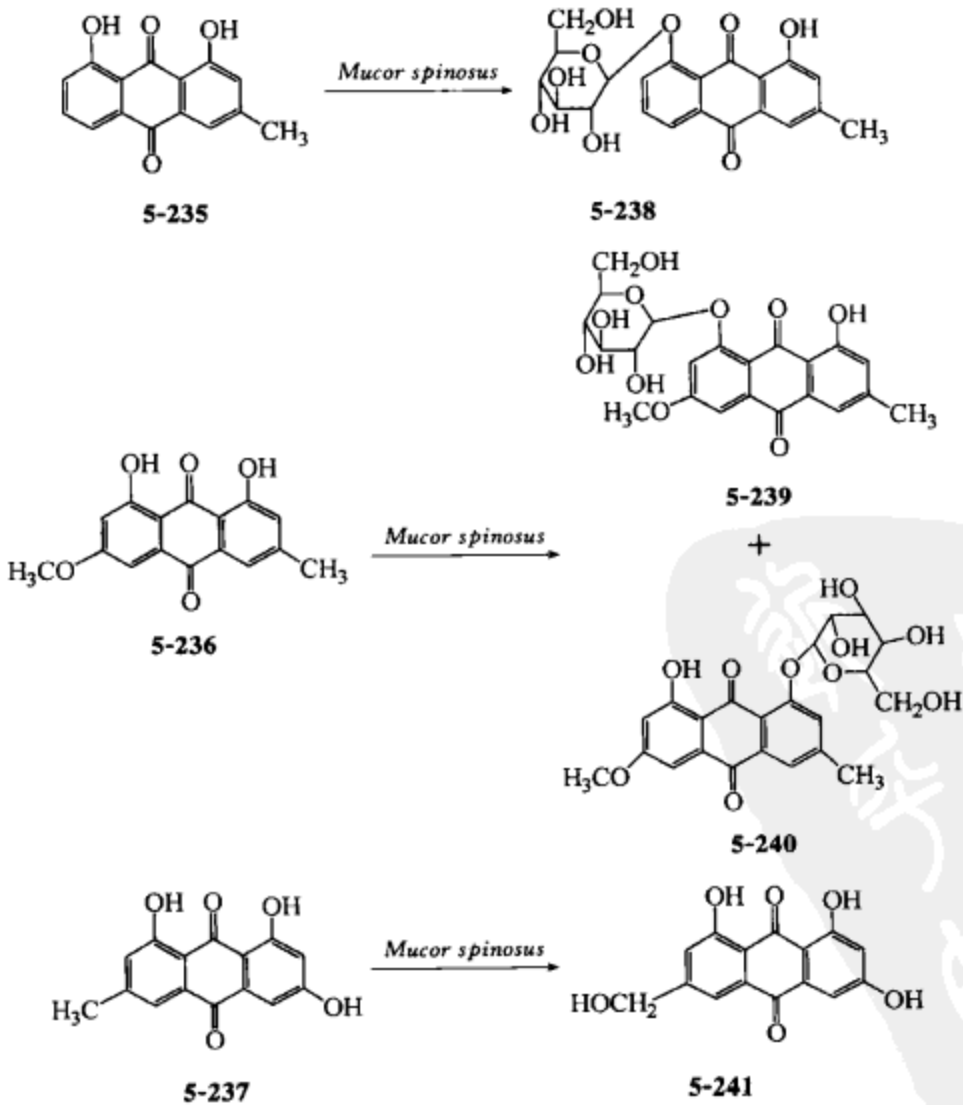
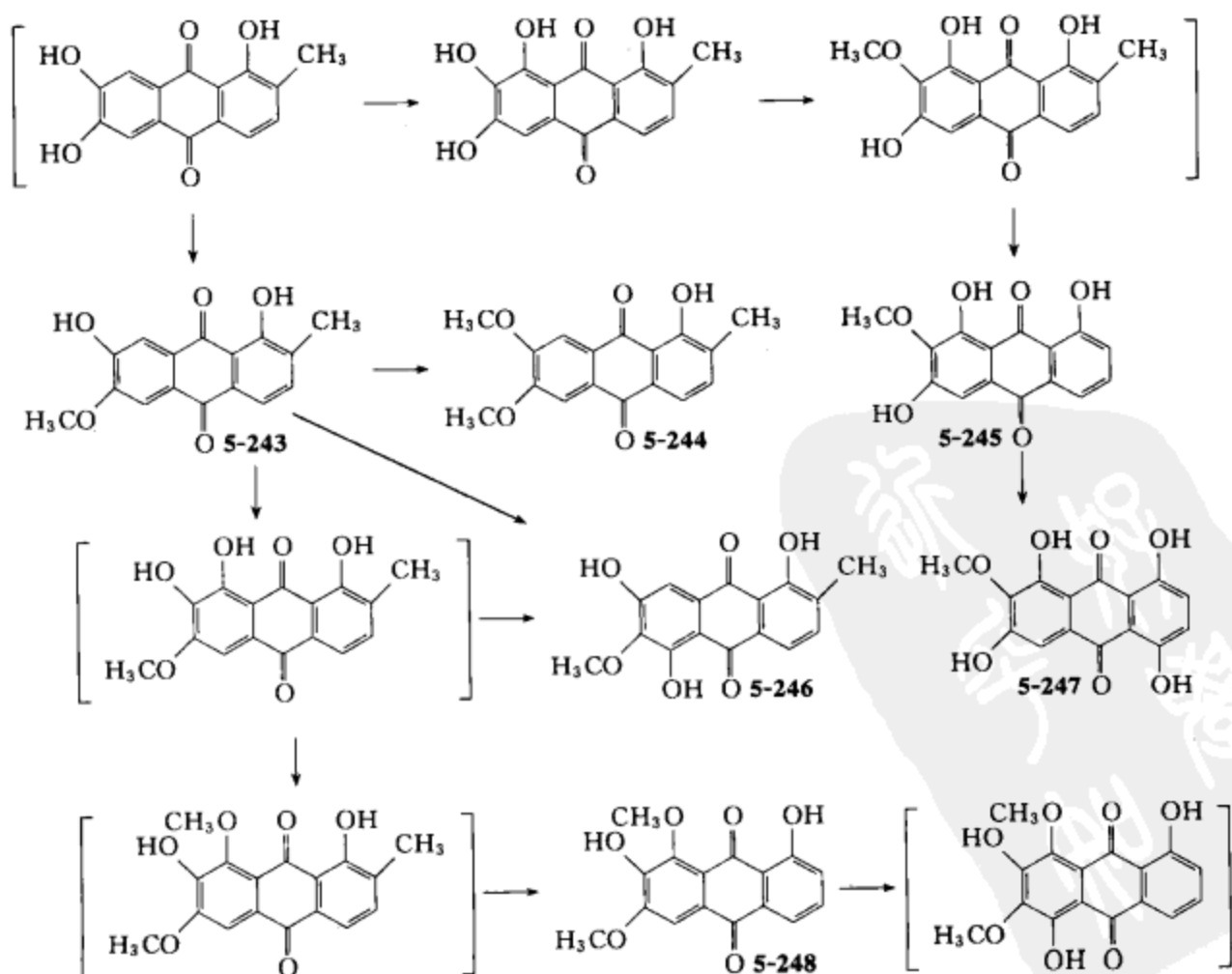


图 5-52 大黄酚、大黄素甲醚和大黄素的生物转化

金鸡纳属 *Cinchona robusta* How 细胞悬浮液在培养过程中并不产生蒽醌类化合物,但在培养过程中加入樟疫霉菌 (*Phytophthora cinnamomi*) 4 天后,培养体系颜色加深,经分离鉴定出 10 个蒽醌类化合物: robustaquinones D (5-243)、robustaquinones B (5-244)、1,3,8-三羟基-2-甲氧基蒽酯 (5-245)、茜网孔素 (copareolatin)-6-甲基醚 (5-246)、robustaquinones F (5-247)、robustaquinones C (5-248)、robustaquinones A (5-249)、robustaquinones E (5-250)、robustaquinones H (5-251) 和 robustaquinones G (5-252) (图 5-54)。



129

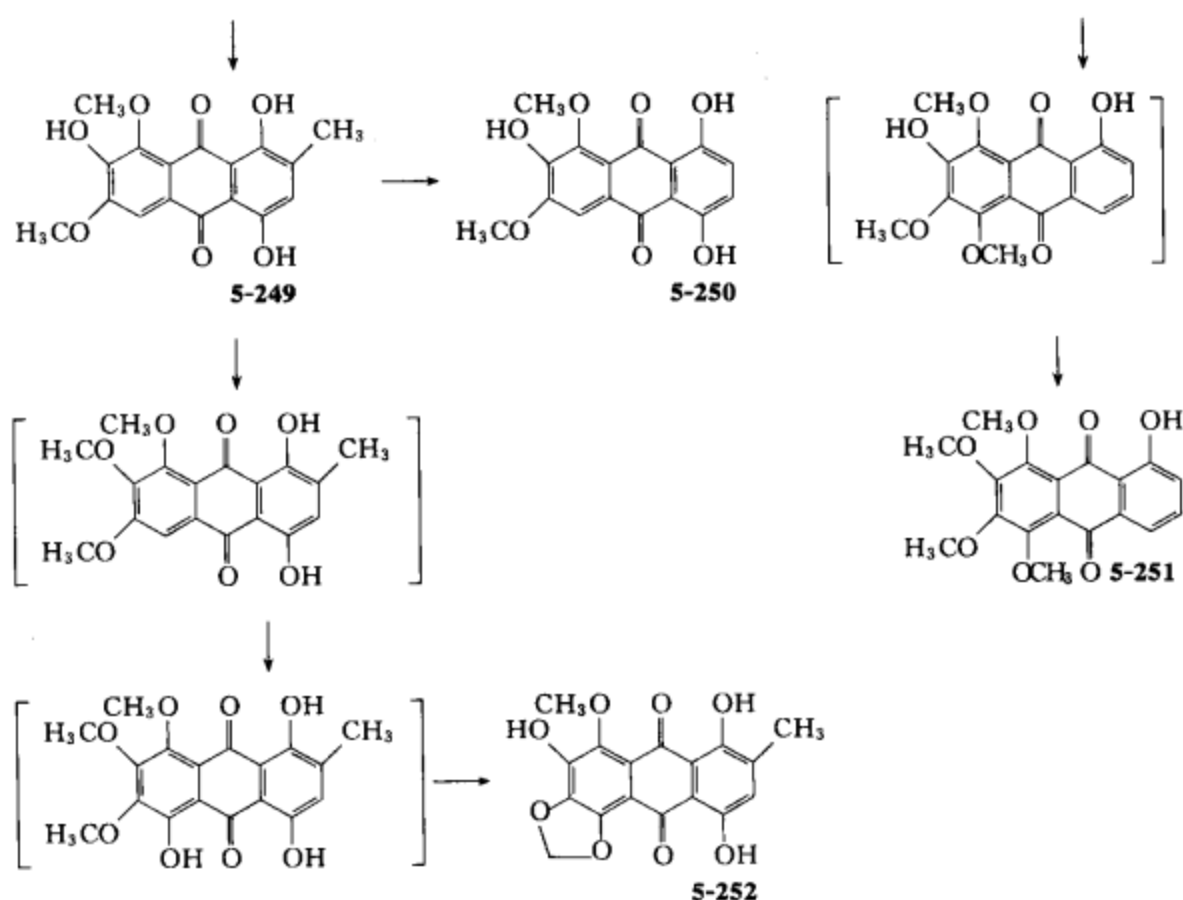


图 5-54 金鸡纳属 *Cinchona robusta* How 细胞悬浮液中蒽醌的转化

蒽醌类化合物尤其是茜草科中的蒽醌类化合物细胞悬浮培养研究比较集中，具体情况可参阅文献。

5.5.3 其他类型活性化合物的生物转化

Zhang 等利用刺囊毛霉 (*Mucor spinosus* 3.3450) 的培养液对天麻素 (5-253) 进行生物转化，得到其酶水解产物苷元对羟基苯甲醇 (5-254)，转化率接近 100% (图 5-55)。

朱宏莉等利用蓝色梨头霉 (*Abisidia coerulea* 3.3389) 进行了类似的研究工作，结果与 Zhang 等相似。

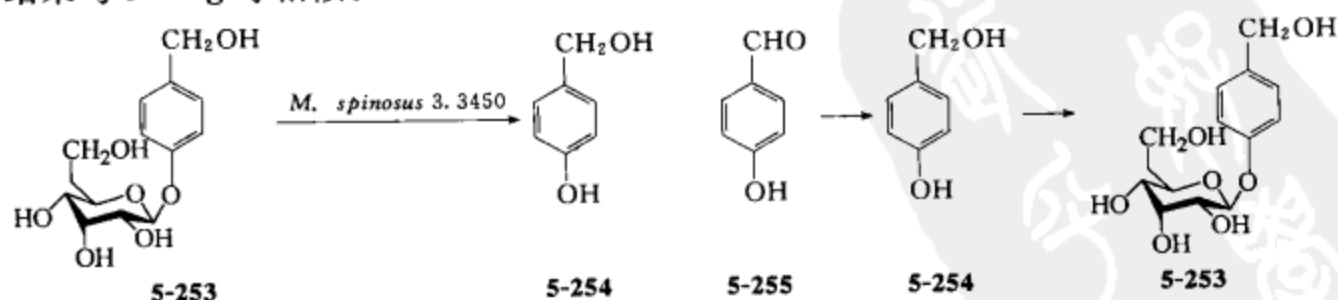


图 5-55 刺囊毛霉 (*Mucor spinosus* 3.3450) 对天麻素的生物转化

图 5-56 对羟基苯甲醛用白花曼陀罗和紫花曼陀罗悬浮细胞的生物转化

龚加顺等利用白花曼陀罗和紫花曼陀罗悬浮细胞分别将对羟基苯甲醛 (5-255)

转化为天麻素 (5-253) (图 5-56)。

戴均贵等利用长春花悬浮细胞对天麻素进行转化, 得到与 Zhang 等利用刺囊毛霉 (*Mucor spinosus* 3. 3450) 的培养液一样的结果。

孙敏等利用黏质沙雷菌菌株 *Serratia marcescens* AB 90027 将异丁香酚 (isoeugenol, 5-256) 转化成阿魏酸 (ferulic acid, 5-257) 和香兰素 (vanillin, 5-258) (图 5-57)。

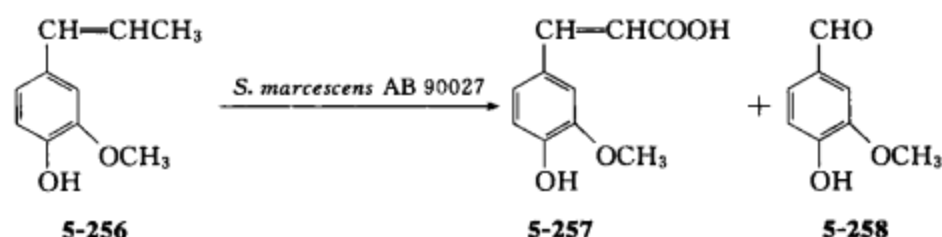


图 5-57 黏质沙雷菌菌株 *Serratia marcescens* AB 90027 对异丁香酚的生物转化

Overhage 等利用罗尔斯通菌菌株 *Ralstonia eutropha* H16 将丁香酚 (eugenol, 5-259) 转化为阿魏酸 (5-257) (图 5-58)。

董长颖和赵丽香利用链霉菌 *Streptomyces* L1936 对阿魏酸 (5-257) 进行转化, 生成香兰素 (5-258) (图 5-59)。

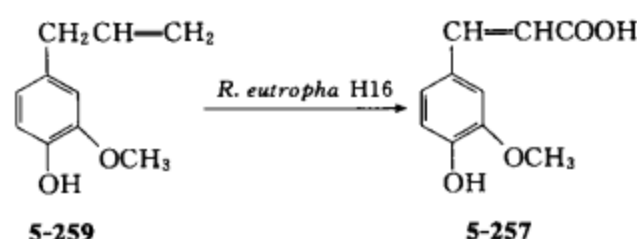


图 5-58 罗尔斯通菌菌株 *Ralstonia eutropha* H16 对丁香酚的生物转化

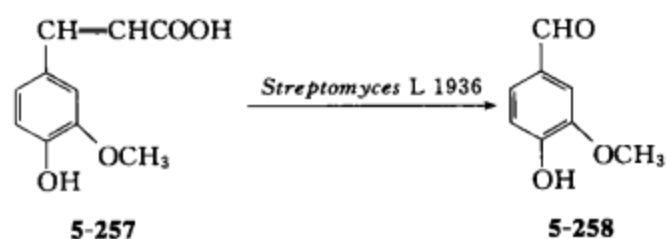


图 5-59 链霉菌 *Streptomyces* L 1936 对阿魏酸的生物转化

参 考 文 献

- [1] Loughlin W A. *Bioresource Technol*, 2000, 74: 49.
- [2] Rozzell J D. *Bioorg Med Chem*, 1999, 7: 2253.
- [3] 卢艳花. 天然药物的生物转化. 北京: 化学工业出版社, 2006.
- [4] 于德泉. 天然产物化学进展. 北京: 化学工业出版社, 2005.
- [5] 曹孜义, 刘国民. 实用植物组织培养技术教程. 兰州: 甘肃科技出版社, 1996.
- [6] Curtin M E. *Biol Tech*, 1983, 1: 649.
- [7] 周立刚, 郑光植. 生物工程进展, 1991, 11: 29.
- [8] Zhong J J, Zhu Q X. *Appl Biochem Biotechnol*, 1995, 55: 241.
- [9] Zhang Y H, Zhong J J, Yu J T. *J Biotechnol*, 1996, 46: 69.
- [10] Kingston D G I, Molinero A A, Rimoldi J M. *Progress in the Chemistry of Organic Natural Products*. New York: Springer-Verlag, 1993.

- [11] Wang H Q, Zhong J J, Yu J T. *Biotechnol Lett*, 1997, 19: 353 .
- [12] Ten Hoopen H J G, Vinke J L , Moreno P R H , et al. *Enzyme and Microbial Technology*, 2002, 30: 56.
- [13] Schlattmann J E, Vinke J L, Ten Hoopen HJ G, et al. *Biotechnol Bioeng*, 1995, 45: 435.
- [14] 徐任生. 天然产物化学. 第2版. 北京: 科学出版社, 2006.
- [15] Charlwood B V, Rhodesmjc. *Secondary products from plant tissue culture*. Oxford: Clarendon Press, 1990.
- [16] Payne G F, Bring V, Prince C, et al. *Plant cell and tissue culture in liquid systems*. Munich: Hanser Publ, 991.
- [17] Verpoorte R, Contin A, Memelink. *Phytochemistry Reviews*, 2002, 1: 13.
- [18] Kieran P M, Maccloughlin P F, Malone D M. *Journal of Biotechnology*, 1997, 59: 39
- [19] Martin VJJ, Pitera DJ, Withers ST, et al. *Nat Biotechnol*, 2003, 21: 1 .
- [20] Pichersky E, Gershenzon J. *Curr Opin Plant Biol*, 2002, 5: 237
- [21] Wani M C, Taylor H C, Wall M E, et al. *J Am Chem Soc*, 1971, 93: 2325
- [22] Croteau R, Ketchum R E B, Long R M, et al. *Phytochemistry Reviews*, 2006, 5: 75.
- [23] Wall M E, Wani M C. *Am Chem Soc Symp Ser*, 1995, 583: 18.
- [24] Rowinsky E K. *Ann Rev Med*, 1997, 48: 353.
- [25] Adams J D, Flora K P, Goldspiel B R, et al. *Monograph Natl Cancer Inst*, 1993, 15: 141.
- [26] Straubinger R M. *Biopharmaceutics of paclitaxel (Taxol): formulation, activity, and pharmacokinetics*. In: Suffness M (ed) *Taxol-Science and Applications*. Boca Raton: CRC Press, 1995, 237.
- [27] Tabata H, *Adv Biochem Engin Biotechnol*, 2004, 7: 1.
- [28] Guo B H, Kail G Y, Jin H B, et al. *African Journal of Biotechnology*, 2006, 5: 15.
- [29] Brown D T. *Preclinical and clinical studies of the taxanes* In: Itokawa H & Lee K-H (eds) *Taxus-The Genus Taxus*. London: Taylor & Francis, 2003, 387.
- [30] Baloglu E, Kingston D G I. *J Nat Prod*, 1999, 62: 1448.
- [31] Itokawa H. *Taxoids occuring in the genus Taxus* In: Itokawa H & Lee K-H (eds) *Taxus The Genus Taxus*. London: Taylor & Francis, 2003, 35.
- [32] Jennewein S, Wildung M R, Chau M, et al. *Proc Natl Acad Sci*, 2004, 101: 9149.
- [33] Chen Y Q, Wu Y Q, Hu Q, et al. *Acta Pharmacol Sin*, 1998, 33: 132.
- [34] Fett-Neto A G, Melanson S J, Nicholon S A, et al. *Biotechnol Bioeng*, 1994, 44: 967.
- [35] Furmanowa M, Oledzka H, Sykłowska-Baranek K, et al. *Biotechnol Lett*, 2000, 22: 1449.
- [36] Yukimune Y, Tabata H, Higashi Y, et al. *Natl Biotechnol*, 1999, 14: 1129.
- [37] Nguyen T, Eshraghi J, Gonyea G, et al. *J Chromatogr A*, 2001, 911: 55.
- [38] Dai J G, Cui J, Zhu W H, et al. *Planta Med*, 2002, 68: 1113.
- [39] Dai J G, Guo H Z, Ye M, et al. *J Asian Nat Prod Res*, 2003, 5: 5.
- [40] Dai J G, Ye M, Guo H Z, et al. *Tetrahedron*, 2002, 58: 5659.
- [41] Zhao Y, Yu R M, Schroeder C, et al. *Acta Botanica Sin*, 2004, 46: 1383.
- [42] Stierle A, Strobel G. , Stierle D. *Science*, 1993, 260: 214.
- [43] Li J Y, Sidhu R S, Ford E J, et al. *Journal of Industrial Microbiology & Biotechnology*, 1998, 20: 259.
- [44] Guo B H, Wang Y C, Zhou X W, et al. *African Journal of Biotechnology*, 2006, 5: 875.

- [45] Strobel G, Yang X S, Sears J, et al. *Microbiology*, 1996, 142: 435.
- [46] Chen T S, Li X, Bollag D, et al. *Tetrahedron Lett*, 2001, 42: 3787.
- [47] 占纪勋, 张元兴, 宁黎丽等. *应用与环境生物学报*, 2003, 9: 429.
- [48] Patel R N, *Annu Rev Microbiol*, 1998, 98: 361.
- [49] Zhang J, Zhang L, Wang X, et al. *J Nat Prod*, 1998, 61: 497.
- [50] Hu S, Tian X, Zhu W, et al. *Tetrahedron*, 1996, 52: 8739.
- [51] Hu S, Sun D, Tian X, et al. *Tetradron Let*, 1997, 38: 2721.
- [52] 刘彦. 青蒿素生物合成相关基因的克隆、大肠杆菌表达与分子分析: [学位论文]. 北京: 中国科学院研究生院, 2002.
- [53] Covello P S, Teoh K H, Polichuk D R, et al. *Phytochemistry*, 2007, 68: 1864.
- [54] Chang Y J, Song S H, Park S H, et al. *Arch Biochem Biophys*, 2000, 383: 178.
- [55] Towler M J, Weathers P J. *Plant Cell Rep*, 2007, 26: 2129.
- [56] Boumeester H J, Wallaart T E, Janssen M H, et al. *Phytochemistry*, 1999, 52: 843.
- [57] Zeng Q P, Qiu F, Yuan L. *Biotechnol Lett*, 2008, 30: 581.
- [58] Liu C Z, Zhao Y, Wang Y C. *Appl Microbiol Biotechnol*, 2006, 72: 11.
- [59] Liu C Z, Wang Y C, Ouyang F, et al. *Biotechnology Letters*, 1998, 20: 265.
- [60] Putalun W, Luealon W, De-Eknamkul W, et al. *Biotechnol Lett*, 2007, 29: 1143.
- [61] Wang J W, Zhang Z, Tan R X. *Biotechnology Letters* 2002, 24: 1573.
- [62] 周立刚, 阮德春, 贺震旦等. *云南植物研究*, 1998, 20: 229.
- [63] Li L X, Su Y F, Liu X F, et al. *Journal of Chinese Pharmaceutical Sciences*, 2002, 11: 122.
- [64] 韩健, 戴均贵, 崔亚君等. *中草药*, 2003, 34: 166.
- [65] Lee I S, Elsohly H N, Croom E M, et al. *J Nat Prod*, 1989, 52: 337.
- [66] 陈有根, 余伯阳. *药用生物技术*, 2001, 8: 90.
- [67] Hufford C D, Lee I S, Elsohly H N, et al. *Baker J K. Pharm Res*, 1990, 7: 923.
- [68] Hufford C D, Khalifa S I, Orabi K Y, et al. *J Nat Prod*, 1995, 58: 751.
- [69] Abourashed S A, Hufford C D. *J Nat Prod*, 1996, 59: 251.
- [70] Elmarakby S A, Elferaly F S, Elsohly H N, et al. *J Nat Prod*, 1987, 50: 903.
- [71] Orabi K Y, Galal A M, Ibrahim A R S, et al. *Phytochemistry*, 1999, 51: 257.
- [72] 叶敏, 宁黎丽, 占纪勋等. *北京大学学报(医学版)*, 2004, 36: 82.
- [73] Ning L L, Han J, Zhang X Y, et al. *J Asian Nat Prod Res*, 2004, 6: 93.
- [74] Ning L L, Han J, Jiang X, et al. *Pure Appl Chem*, 2003, 75: 389.
- [75] Ning L L, Zhang X Y, Qu G Q, et al. *Tetrahedron*, 2003, 59: 4209.
- [76] Ning L L, Qu G Q, Ye M, et al. *Planta Med*, 2003, 59: 4209.
- [77] 张剑, 程志红, 余伯阳. *中国天然药物*, 2004, 2: 376.
- [78] Kim D H, Lee S W, Han M J. *Biol Pharm Bull*, 1999, 22: 320.
- [79] Taiko A. *boil Pharm Bull*, 1999, 22: 787.
- [80] Noboru S, Yutaka H, Kazuo U, et al. *Phytochemistry*, 1996, 43: 99.
- [81] 郭维, 徐波, 杨秀伟等. *中国药理学通报*, 2003, 19: 351.
- [82] 杨秀伟, 赵静, 崔景荣等. *北京大学学报(医学版)*, 2004, 36: 31.
- [83] Zhang J, Cheng Z H, Yu BY, et al. *Tetrahedron Lett*, 2005, 46: 2337.
- [84] Cheng Z H, Yu BY, Cordell G A, et al. *Organic Lett*, 2004, 6: 3163.

- [85] Collins D O, Peter L D, Reese P B. *Phytochemistry*, 2002, 59: 479.
- [86] Samir A K, Parnali C, Pezzuto J M, et al. *J Nat Prod*, 2000, 63: 1653.
- [87] Akihisa T, Takamine Y, Yoshizumi K, et al. *J Nat Prod*, 2002, 65: 278.
- [88] 赵剑, 朱蔚华. *植物生理学通讯*, 1999, 35: 60.
- [89] Madyastha K M, Ridgway J E, Dwyer J G, et al. *J Cell Biol*, 1977, 72: 302.
- [90] Contin A, Collu G, Heijden R, et al. *Plant Physiol Biochem*, 1999, 37: 139.
- [91] Irmmler S, Schröder G, St-Pierre B, et al. *Plant J*. 2000, 24: 797.
- [92] Poulsen C, Verpoorte. *Phytochemistry*, 1991, 30: 377.
- [93] Jacobs D I, Gaspari M, Greef J, et al. *Planta*, 2005, 221: 690.
- [94] Ei-Sayed M, Verpoorte R. *Phytochem Rev*, 2007, 6: 277.
- [95] Furuya T, Sakamoto K, Itada K, et al. *Phytochemistry*, 1992, 31: 3065.
- [96] Hemscheidt T, Zenk M H. *FEBS Lett*, 1980;110: 187.
- [97] Stevens L H, Blom T J M, Verpoorter R. *Plant Cell Rep*, 1993, 12: 573.
- [98] St-Pierre B, De Luca V. *Plant Physiol*, 1995, 109: 131.
- [99] Lorence A, Nessler C L. *Phytochemistry*, 2004, 65: 2735.
- [100] Yamazaki Y, Urano A, Sudo H, et al. *Phytochemistry*, 2003, 62: 461.
- [101] Carté B K, Debrosse C, Eggelston D, et al. *Tetrahedron Letters*, 1990, 46: 2747.
- [102] Aimi N, Nishimura M, Miwa A, et al. *Tetrahedron Letters*, 1989, 30: 4991.
- [103] Kitajima M, Masumoto S, Takayama H, et al. *Tetrahedron Letters*, 1997, 38: 4255.
- [104] Kitajima M, Nakamura M, Takayama H, et al. *Tetrahedron Letters*, 1997, 38: 8997.
- [105] Sakato K, Tanaka H, Mukui N, et al. *Agirc Biol Chem*, 1974, 38: 217.
- [106] Hengel A J, Harkes M P, Wichers H J. *Plant Cell*, 1992, 28: 11.
- [107] Song S H, Byun S Y. *Microbiol, Biotechnol*, 1998, 8: 631.
- [108] Gunasekera S P, Badawi M M, Cordell G A, et al. *Nat Prod*, 1979, 42: 475.
- [109] Govindachari T R, Viswanathan N. *Phytochemistry*, 1972, 11: 3529.
- [110] Arisawa M, Gunasekera S P, Cordell G A, et al. *Planta Med*, 1981, 43: 404.
- [111] 朱关平, 林隆泽, 潘文君等. *科学通报* 1978, 23: 761.
- [112] 胥秀英, 郑一敏, 温寿祯等. *中国中药杂志*, 2001, 26: 119.
- [113] 胥秀英, 郑一敏, 温寿祯等. *生物学杂志*, 2001, 18: 18.
- [114] 张文伟, 余伯阳, 彭娟. *药物生物技术*, 2003, 10: 165.
- [115] 张春燕, 白宝星, 王明蓉. *国外医药抗生素分册*, 2007, 28: 210.
- [116] Song Z H, Nes W D. *Lipids*, 2007, 42: 15.
- [117] Liu Y, Cheng K D, Zhu P, et al. *Acta Botanica Sinica*, 2004, 46: 935.
- [118] 刘颖, 程克棣, 朱平等. *药学报*, 2004, 39: 445.
- [119] 葛文中, 李楠, 单丽红等. *微生物学杂志*, 2007, 27: 14.
- [120] Doller P C, Alfermann A W, Reinhard E. *Planta Med*, 1977, 31: 1.
- [121] Clayton R B J. *Lipid Res*. 1964, 5: 3.
- [122] 冯冰, 马百平, 康利平等. *中草药*, 2005, 36: 978.
- [123] Quan B, Ma B P, Feng B, et al. *Chin J Nat Med*, 2006, 4: 378.
- [124] 葛文中, 李楠, 单丽红等. *微生物学报*, 2007, 47: 54.
- [125] Havsteen B H. *Pharmacology & Therapeutics*, 2002, 96: 67.

- [126] Demetzos C, Magiatis P, Typas M A, et al. *Cell. mol. life sci*, 1997, 53: 587.
- [127] 杨秀伟, 张建业, 徐嵬等. *药学学报*, 2005, 40: 71.
- [128] 谢明杰, 石姗姗, 卢明春等. *食品与发酵工业*, 2004, 30: 21.
- [129] 王园园, 刘吉华, 余伯阳. *药物生物技术*, 2005, 12: 308.
- [130] Han Y S, Heijden R, Verpoorte R. *Plant Cell, Tissue and Organ Culture*, 2001, 67: 201.
- [131] Zhang W, Zhang J X, Cheng Y J, et al. *Chin J Nat Med*, 2003, 1: 219.
- [132] 牟淑慧, 刘颖, 张桂燕. *北京中医药大学学报*, 2007, 30: 478.
- [133] Schripsema J, Ramos-Valdivia A, Verpoorte R. *Phytochemistry*, 1999, 51: 55.
- [134] Zhang J X, Guo H Z, Dai J G, et al. *Journal of Chinese Pharmaceutical Sciences*, 2001, 10: 187.
- [135] 朱宏莉, 张嘉, 宋纪蓉等. *西北大学学报*, 2007, 37: 73.
- [136] 龚加顺, 马维鹏, 普俊学等. *药学学报*, 2006, 41: 963.
- [137] 龚加顺, 马维鹏, 普俊学等. *生物工程学报*, 2006, 22: 800.
- [138] Gong J S, Ma W P, Pu J X, et al. *Chin J Biotech*, 2006, 22: 800.
- [139] 戴均贵, 巩卓, 朱丹萌等. *植物学报*, 2002, 44: 377.
- [140] 孙敏, 姚日生, 高文霞. *生物加工过程*, 2006, 4: 33.
- [141] Overhage J, Steinbuchel A, Priefert H. *Applied and Environmental Microbiology*, 2002, 68: 4315.
- [142] 董长颖, 赵丽香. *吉林农业科技学院学报*, 2005, 14: 1.



第6章 计算药理学在药用天然产物研究中的应用

计算机辅助药物设计已经有几十年的历史，已发展成为一套较为完整的算法技术及软件，包括各种数据库、定量构效关系、药效团模型、分子对接、ADME/T（吸收、分布、代谢、消除/毒性）预测、各种数据挖掘分析算法等。在这些方法与技术的基础上近年来发展了计算药理学（computational pharmacology）。

计算药理学的目标是在计算机上实现某些药理学实验，利用计算机软件获取、分析及整合不同来源的生物、医学数据，建立各种预测模型。这样可以大大减少实验工作量，节省大量的人力、物力。随着系统生物学及生物网络的发展，人们已认识到以往的单一靶点单一药物的模式对于复杂疾病收效甚少，人们已逐渐注意到多靶点、多途径基于系统的药物开发模式。药用天然产物的特点是多组分、多靶点，因此，这些新进展为人们提供了开展基于天然产物的计算药理学研究的方法与途径。

基于天然产物的计算药理学研究包括传统的计算机辅助药物分子设计方法，各种来源的生物、化学、药物及医学数据的整合与挖掘，天然产物分子在化学空间及配体靶空间的分布，活性分子与多种蛋白质的相互作用谱，植物中多种天然产物与多种靶的相互作用谱，应用生物网络表达及分析天然产物的结构和与多靶的相互作用等。

6.1 计算机辅助药物分子设计（CADD）

近几年来，计算机辅助药物分子设计的方法已得到广泛的应用。许多制药公司也纷纷投资建立了相应的计算机辅助药物设计部门以带动新药的开发和研究。目前已有许多通过计算机辅助药物分子设计的药物获得成功，并且有多种通过 CADD 设计或 CADD 参与设计的新药已经上市。

图 6-1 显示了计算机辅助药物分子设计的基本方法。计算机辅助药物设计的方法大体可以分为 3 类。

6.1.1 计算机辅助药物分子设计方法

6.1.1.1 基于小分子的药物设计方法

这一类设计方法主要是针对未知受体结构的药物分子，主要包括定量构效关系方法和药效团模型方法。所谓定量构效关系（QSAR）就是借助于化合物的物化参数或结构参数，通过数学的模型描述建立起化合物的某些生理活性或性质和其物化

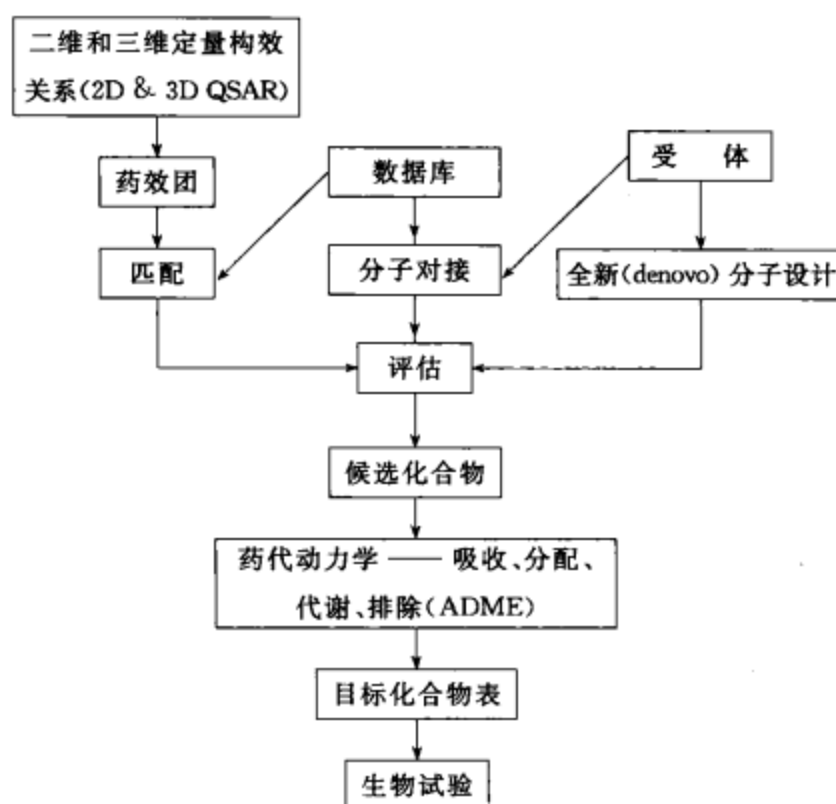


图 6-1 计算机辅助药物分子设计方法

参数和结构参数之间的定量关系。通过定量关系模型，可以预测化合物的某些性质或生理活性，也可以帮助人们进行化合物的改造和设计。以小分子的结构和活性为基础，常常也可以采用药效团模型方法找到化合物共同的药效作用模式。在药物分子设计中，所谓“药效团”（pharmacophore）指的是一系列活性化合物具有的一组共同特征，这些特征包括特定的化学基团、氢键给体或受体、正电基团或负电基团、疏水性基团等。一个三维的药效团模型给出了这些药效基团之间的特定三维排列方式。在得到的构效关系模型和药效团模型的基础上，也可以结合其他方法比如数据库搜索的方法来进行药物的设计。

6.1.1.2 基于受体结构的药物设计方法

随着分子生物学和结构生物学的发展，越来越多的生物大分子结构被解析。对于一些未知三维结构的受体（receptor）大分子，它们的结构常常也可以通过同源蛋白质模建的方法得到。在这种情况下，就可以采用基于受体结构的药物设计方法来寻找新的先导化合物。基于结构的药物设计方法的思路就是通过研究受体结构的特征以及受体和药物分子之间的相互作用方式来进行药物的设计。分子对接就是基于受体结构药物设计中的一类重要方法。所谓“分子对接”（docking）就是受体和药物分子之间通过几何匹配和能量匹配而相互识别的过程。在药物分子在药效发生的生物过程中，药物与受体相互结合，首先就是需要两个分子充分接近，采取合适的取向，从而使两者在必要的部位相互契合，进一步发生相互作用。因此，确定复合物中的两个分子正确的相对位置和取向，对于新药的设计非常关键。在基于受体

结构的药物设计中，另外一类重要的方法就是从头设计方法，它属于真正意义上的全新（de novo）的药物分子设计的方法。这个全新的概念比数据库搜索所指的先导化合物具有更广泛的含义。数据库搜索所提供的化合物并不是真正意义上的全新化合物，而且某些数据库可能会偏重于收集某些类的化合物，因此数据库搜索的方法在化合物的结构和种类方面会对设计的结果产生一些限制。从头设计方法是根据受体的三维结构产生新的先导化合物，设计过程中完全没有化合物种类和结果的预先限制，因此设计的结果可能是完全新型的。

6.1.1.3 计算组合化学方法

组合化学方法的建立和发展推动了计算机辅助药物设计方法的发展，随之也产生了计算组合化学方法。计算组合化学可以包括两部分的内容：第一个方面就是采用计算机技术来设计合成组合库的“构造块”（building block），通过计算机来生成包含足够分子多样性的虚拟组合库；第二个方面就是通过把得到的虚拟组合库和其他分子设计方法结合进行药物分子的设计。通过一些方法得到一些药效评价模型之后，比如构效关系模型、药效团模型、自由能预测模型等，就可以通过这些模型对虚拟组合库进行虚拟的高通量筛选。

当然，上面介绍的这些方法之间也并不是完全独立的。对于某个体系，常常会采用多种药物设计的方法来进行先导化合物的设计。比如知道了蛋白质和药物分子相互结合的复合物结构时，可以从蛋白质结构出发用基于分子对接的数据库搜索方法来进行先导化合物的寻找；从复合物的相互作用信息也可以得到药效团模型，然后从药效团模型出发进行药物设计；同时，也可以从蛋白质结构出发，用片断生长的方法进行全新药物分子的设计；此外，先导化合物得到以后，也可以采用计算组合化学的方法并结合定量构效关系来进行先导化合物的优化。

6.1.2 CADD 中的几个难题

6.1.2.1 受体和配体之间自由能的评估

分子活性的大小常常是由受体和药物分子之间的结合自由能来决定的，因此计算体系的自由能是非常重要的，但是自由能的计算不仅复杂，而且非常耗时。在自由能计算方法中，自由能微扰（FEP）和热力学积分（TI）的方法在理论上最为严谨，但这两种方法只能处理简单的体系。也有人提出了一些基于经验方程的结合自由能预测方法，但是这类方法一方面计算的精度不高，另一方面很难建立一个普适的评估手段，因此应用受到了一些限制。近年来又出现了基于线性响应近似的线性相互作用能（LIE）方法以及基于 MM/PBSA 计算的自由能计算方法（MM：分子力学，用于计算体系的焓变；PB：Poisson-Boltzmann 方程，用于计算溶剂效应中极性部分对自由能的贡献；SA：溶剂可及分子表面积，用于计算溶剂效应中非极性部分对自由能的贡献）。但总体来讲，结合自由能的评估还

需要做大量的工作。

6.1.2.2 大分子的构象问题

确定分子的构象有两种方法：实验方法和理论计算方法。常用的实验方法有 X 射线晶体学方法、多维核磁共振方法和圆二性色谱方法等。但是利用实验方法解析的大分子结构数目有限，目前蛋白质数据库中的蛋白质不过 10000 多个，但是已知的蛋白质序列却远远大于这个数目，而且还在迅速增加。构象分析理论方法主要包括系统搜索方法、距离几何方法、蒙特卡罗模拟方法、遗传算法优化方法以及分子动力学方法等。这些方法对于中小分子构象问题的解决具有一定的价值，但是对于大分子而言，单单依靠这些方法是很难确定出其优势构象的。

同源蛋白质预测方法仍然是确定大分子构象的最实际且最有效的方法。一般认为，模型蛋白质如果和参考蛋白质有 50% 以上的序列同源性，采用同源蛋白质预测的方法可以搭建出可靠的模型；对于序列同源性在 30%~50% 的蛋白质，所搭建的模型具有一定的可靠性；如果序列同源性低于 30%，则很难搭建出正确的模型。已经知道，两个蛋白质之间如果序列同源，那结构一定同源；但序列不同源，结构未必不同源。对于大量的序列，可能在蛋白质结构数据库中不能找到与其序列同源的蛋白质结构，但这并不意味着不存在与其结构同源的蛋白质。因此近年来，对于序列同源性较低的蛋白质，人们提出了折叠识别的方法，在考虑序列同源性的同时，还参考其二级结构的相似性确定其所属的折叠类型，从而正确地进行序列对比，提高了模型搭建的可靠性。但是应该说，如何从一级序列可靠地建立其三维模型这个问题还远没有解决。怎样发展新的蛋白质结构预测方法是计算化学中一个非常重要的课题。

6.1.2.3 溶剂效应的考察

如何处理溶剂问题是计算机辅助药物分子设计的一个难题。人们关心的大量生命过程都是在溶剂条件下进行的。在溶剂中和在真空中，分子的构象可能会存在较大的差别。在药物设计中，采用不考虑溶剂的简化模型可能会使计算的结果和实际情况之间产生很大的偏差。目前对溶剂效应的考察主要采用下面 3 种方法：水分子模型、连续介质模型以及基于分子表面的半经验模型。水分子模型需要大量的计算时间；而后两种方法在准确性上都存在问题。因此，发展快速而准确的溶剂效应计算的方法是计算机辅助药物分子设计中的一个重大挑战。

6.1.3 计算机辅助药物分子设计的前景

6.1.3.1 计算机技术的飞速发展

计算机技术的发展日新月异。目前已经出现了计算速度超过 1000 万亿次的超级计算机。计算机技术的迅猛发展必将引起计算化学、计算生物学以及计算机辅助药物分子设计等领域的革命性变化。计算机技术的发展为一些计算方法的应用提供

了一些新的契机，比如大分子的量子力学计算、大分子溶剂体系的长时间模拟以及大型数据库的搜索等。这样人们可以通过更加精确的计算、更加复杂的模拟来提高分子设计的成功率，同时在超级计算机的支持下发展新的适应复杂生物体系理论计算以及药物设计需求的新方法和新技术。

6.1.3.2 人类基因组和生物信息学的发展

据统计，目前治疗药物的作用靶点约为 483 个，其中受体占 45%，酶占 28%，激素和细胞因子占 11%，DNA 占 2%，核受体占 2%，离子通道占 5%，其他 7% 未知。随着人类基因组计划的完成和相关研究不断深入，大量与疾病相关的基因将被发现，这会促使药物作用的靶标分子急速增加，药物分子设计面临前所未有的广阔应用前景。

同时，近些年，生物信息学也取得了很大的进展，生物信息的存储、分析、处理以及利用等方面都取得了很大的突破。生物信息的利用以及和计算机辅助药物分子设计的方法的结合会为新药研制注入新的活力。

6.1.3.3 新的计算方法的发展

新的计算方法是随着计算手段的提高而发展起来的。一个好的药物分子除了要和靶点形成好的相互作用之外，还应该具有好的药代动力学特征以及尽量小的毒性和副作用。为了满足药物设计的要求，近年 CADD 在药代动力学以及毒性预测等方面也取得了较大的进展。CADD 正逐渐从基于结构的药物设计方法 (structure-based drug design) 向基于机理的药物设计方法 (mechanism-based drug design) 过渡。单一分子药物 (魔术子弹) 对于单一因素引起的疾病 (如慢性白血病) 有很好的疗效。近年来，随着科学的发展，特别是系统生物学的发展，人们已认识到多数人类疾病特别是癌症、心血管病、神经系统疾病等复杂疾病是与多种致病因素相关的。从分子水平发展到网络水平研究疾病的机理表明蛋白质之间是相互关联的，复杂疾病是由于多种基因变异引起的，作用于一种靶点的单分子药物不能有效地治疗复杂疾病，而且容易产生抗药性与副反应，因此开发一种分子能抑制多种蛋白质的药物 (混杂药物, promiscuous drug) 及组合药物已成为人们关注的热点之一。药物开发的模式已从基于单一靶点的药物开发转向基于网络或基于系统生物学的药物开发。

6.1.3.4 新的实验方法的发展

分子设计成功与否最终要由实验来验证。不仅如此，实验手段还为分子设计提供了理论依据和支持。例如人们对蛋白质的晶体结构数据进行了总结，得出了蛋白质折叠的部分规律，在此基础上发展了折叠识别蛋白质模型搭建方法；多维核磁共振技术结合分子动力学模拟大大提高了模型的可靠性。相信随着 X 射线晶体学的发展以及大功率多维核磁共振仪的出现，将有越来越多的蛋白质得到解析。又如组合化学方法的发展，它能迅速合成大量不同的化合物，在寻造新药发

面有巨大的应用前景，组合化学和 CADD 的结合必将极大地推动药物设计的发展。

6.2 构建数据库

过去，对理论和计算方法在药物发现过程中的地位作用，很多实验科学家可以说都有这样一个共同的印象：“总是在一个药物分子被发现之后不断地解释它为什么有效，但从来没有在这之前指出哪一个应该去值得研究。”正是这一误解不断激励理论计算工作者去探索和发展专门工具，能够根据分子的三维结构信息去预测其对特定受体的活性。最初这类软件工具就是基于数据库搜索，而要进行数据库搜索首先要构建数据库。下面介绍几个与药物开发相关的数据库。

6.2.1 剑桥结构数据库 (CSD)

剑桥结构数据库 (Cambridge Structural Database, CSD) 包含数十万个有机化合物、金属有机化合物以及金属配合物的晶体结构信息。所有这些晶体结构都是通过 X 射线或中子散射实验技术得到的。数据库中每一条记录包含所有来源文献资料以及简要的晶体结构和实验数据，分子的常规化学结构信息，原子座标、空间群对称性、共价键半径等信息 (图 6-2)。

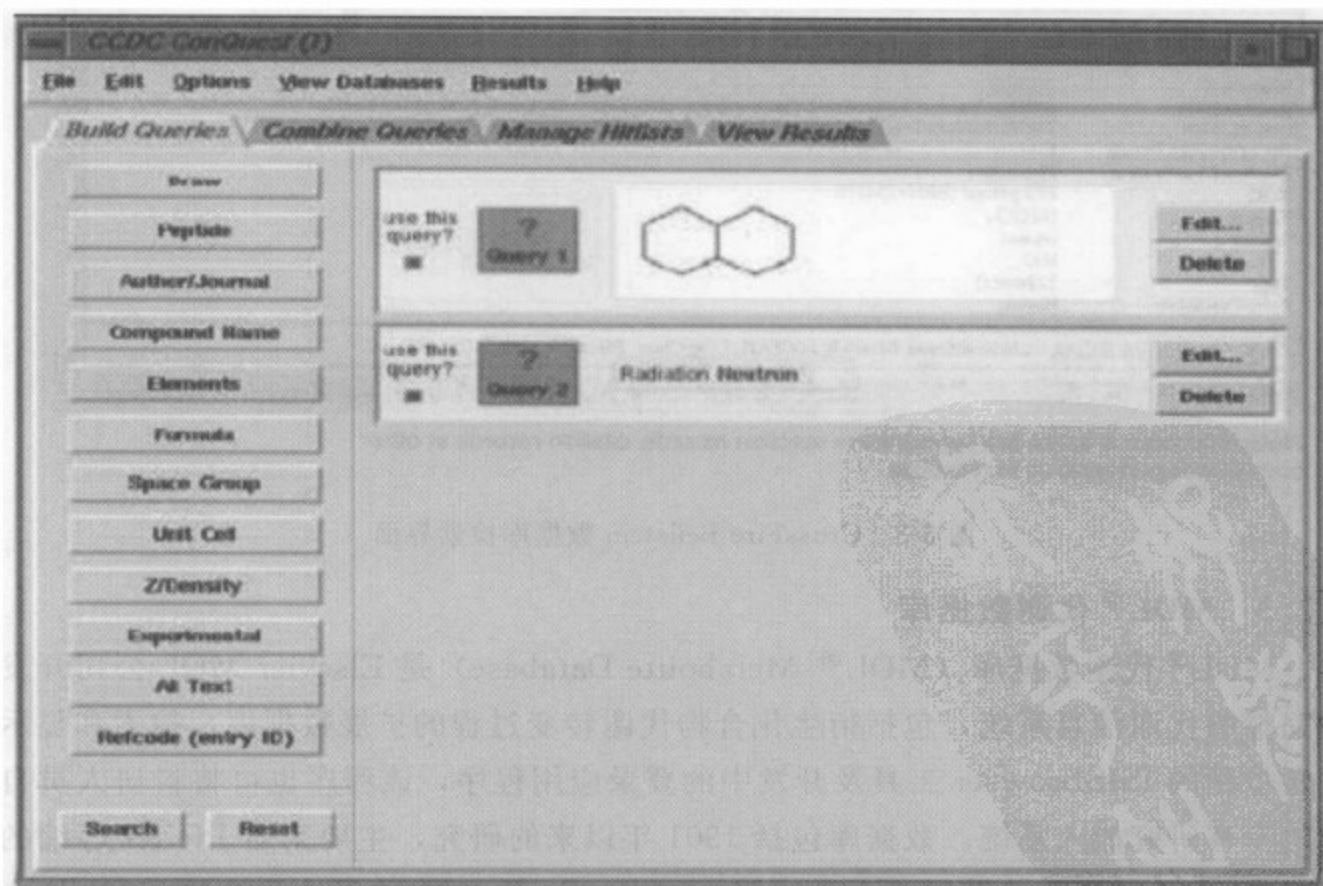


图 6-2 CSD 检索界面

6.2.2 CrossFire Beilstein 数据库

CrossFire Beilstein 数据库是世界上最大的有机化学综合数据库,以电子方式提供包含可供检索的化学结构和化学反应、相关的化学性质和物理性质,以及详细的药理学和生态学数据在内的最全面的信息资源。通过 CrossFire 强大的搜索引擎,科学家们可以很方便地查询、浏览已建立了内部关联的数据和结构信息。索引按时间顺序编排,包括:900 多万种包含结构和参考文献的化合物信息,数千万条相关的化学性质和生物活性信息,其中包括药理学、环境病毒学、转移、分布及最终结果等一系列数据;1000 多万种可供检索的化学反应信息;超过 80 万条摘要和标题,根据 1980 年以来的主要的有机化学文献进行索引(图 6-3)。

The screenshot displays the CrossFire Beilstein database search interface. The main window is titled "Display Hit: [Q02-B50034AE hit 1 of 3]". It features a menu bar (File, Edit, Task, View, Options, Window, Help) and a toolbar with various icons. The interface is divided into several sections:

- Chemical Data Section:** Lists properties for a specific compound.

Beilstein Preferred BRN	439.14.5
CAS Registry Number	439.14.5, 65854-76-4
Chemical Name	7-chloro-1-methyl-5-phenyl-1,3-dihydro-benzo[4,1,4]diazepin-2-one
Autotame	7-chloro-1-methyl-5-phenyl-1,3-dihydro-benzo[4,1,4]diazepin-2-one
Molecular Formula	C ₁₈ H ₁₅ ClN ₂ O
Molecular Weight	284.74
Lawson Number	26728, 2817
Compound Type	heterocycle
Constitution ID	691838
Tautomer ID	742046
Beilstein Reference	5-34-04-00300, 6-24
Entry Date	1983/1/28
Update Date	2000/10/21
- Chemical Structure:** A chemical structure diagram of the compound is shown in a separate window titled "Q02 Hit 1".
- Reaction Section:** Displays details for "Reaction 1 of 12".

Reaction ID	2027668
Reactant BRN	2152224 2-amino-N-[4-chloro-2-(hydroxyimino-phenyl-methyl)-phenyl]-N-methyl-acetamide
Product BRN	754371 7-chloro-1-methyl-5-phenyl-1,3-dihydro-benzo[4,1,4]diazepin-2-one
No. of Reaction Details	1
Reaction Classification	Preparation
Yield	89.3 percent (BRN=754371)
Reagent	NH ₄ SO ₃
Solvent	ethanol
Time	H ₂ O
Other Conditions	12 hour(s)
	Heating
- References:** A list of references is provided at the bottom, including "Ref: 1 5170288, J. Org. Chem., 45, 9, 1980, 1675-1681".

A note at the bottom states: "Note: Numbers in italics are hyperlinks to reaction records, citation records or other compounds participating in reactions".

图 6-3 CrossFire Beilstein 数据库检索界面

6.2.3 MDL® 代谢数据库

MDL® 代谢数据库 (MDL® Metabolite Database) 是 Elsevier MDL 公司开发的完整的代谢信息系统,包括衍生化合物代谢转变过程的扩展数据库、检索和显示代谢过程的 ISISBrowser 工具及开发中的登录应用程序,该程序也可将科研人员自身的专有研究加入系统。数据库包括 1901 年以来的研究,主要来自 Pfeifer 等编的《药物的生物转化》、《药代动力学》等。该库把关于 8590 种母体化合物、53373 个转化物及 34537 种分子的实验结果综合在一起,并划分为多分支体系以便查看。数

数据库包括种类、药物吸收和分泌途径、分析方法学、母体化合物分类、生理活性、酶的信息、生物利用度等内容。数据库每半年更新一次，最新资料的多少依赖于发表的最新研究数量，但通常每年大约有 3000 条转化信息加入到代谢数据库之中（图 6-4）。

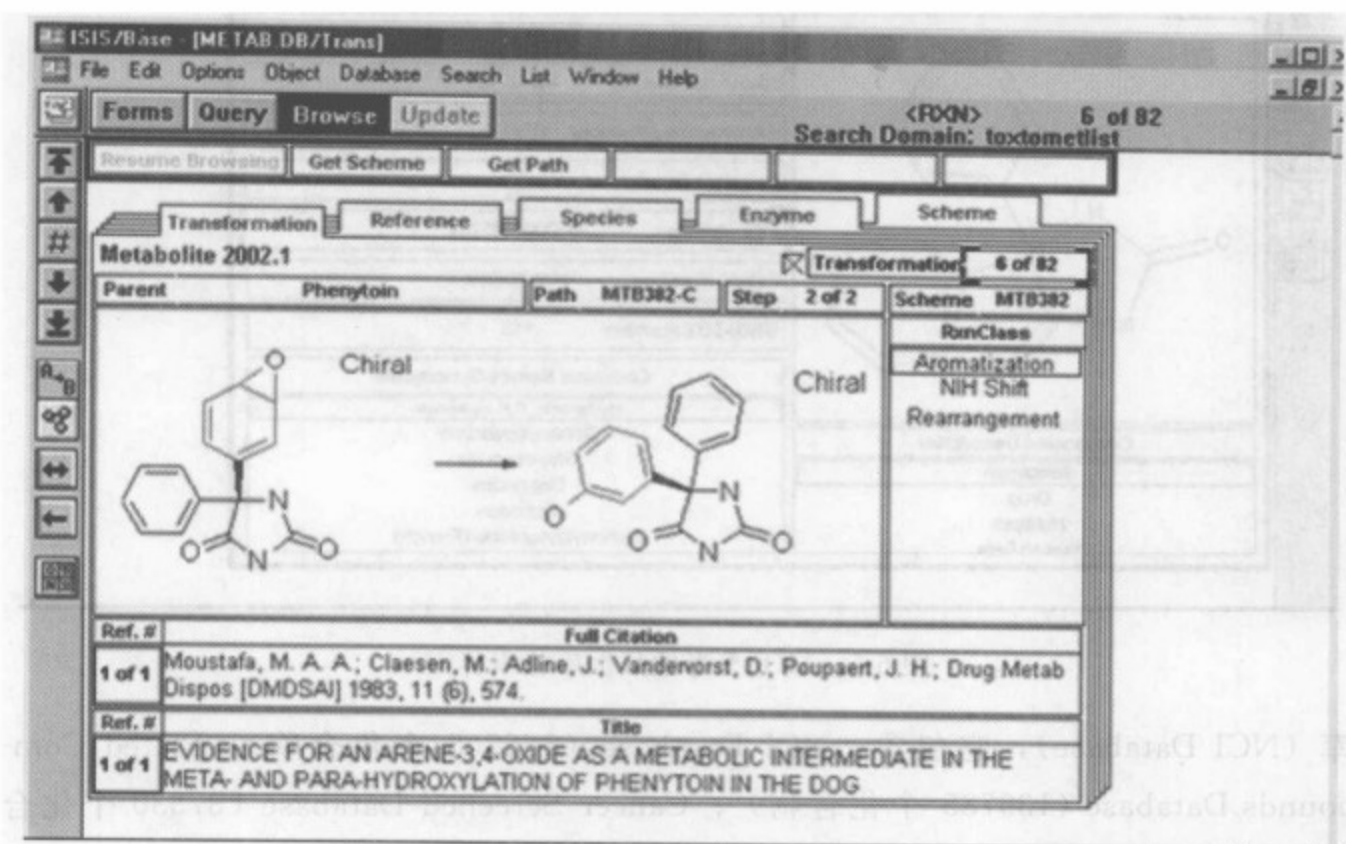


图 6-4 MDL[®] 代谢数据库检索界面

6.2.4 MDL[®] 毒性数据库

MDL[®] 毒性数据库 (MDL[®] Toxicity Database) 由 Elsevier MDL 公司开发，涵盖了现今报道过的化合物的毒性和毒理数据，包括急毒、致畸、对皮肤和眼睛的刺激、致癌及多剂量效应等体内和体外的实验结果。数据资源来自于美国国家职业安全与健康研究所 (NIOSH) 的化学物质毒性作用登记数据库 (RTECS)。这个数据库中包括超过 150000 种化学物质，其中 2/3 是药物和有希望成药的分子，其他的来自于其他领域，例如农业化学、石油化工、塑料制造业、香料和调味剂、动植物提取物、无机化合物和金属有机化合物。Toxicity 每个季度升级一次（图 6-5）。

6.2.5 NCI 数据库

美国国家癌症研究所 (National Cancer Institute, NCI) 自 1955 年以来一直致力于寻找具有抗癌疗效的合成化合物和天然提取物。NCI 早在项目的初期就使用计算机信息技术管理、跟踪整个项目过程中产生的大量数据，最终得到 NCI 数据

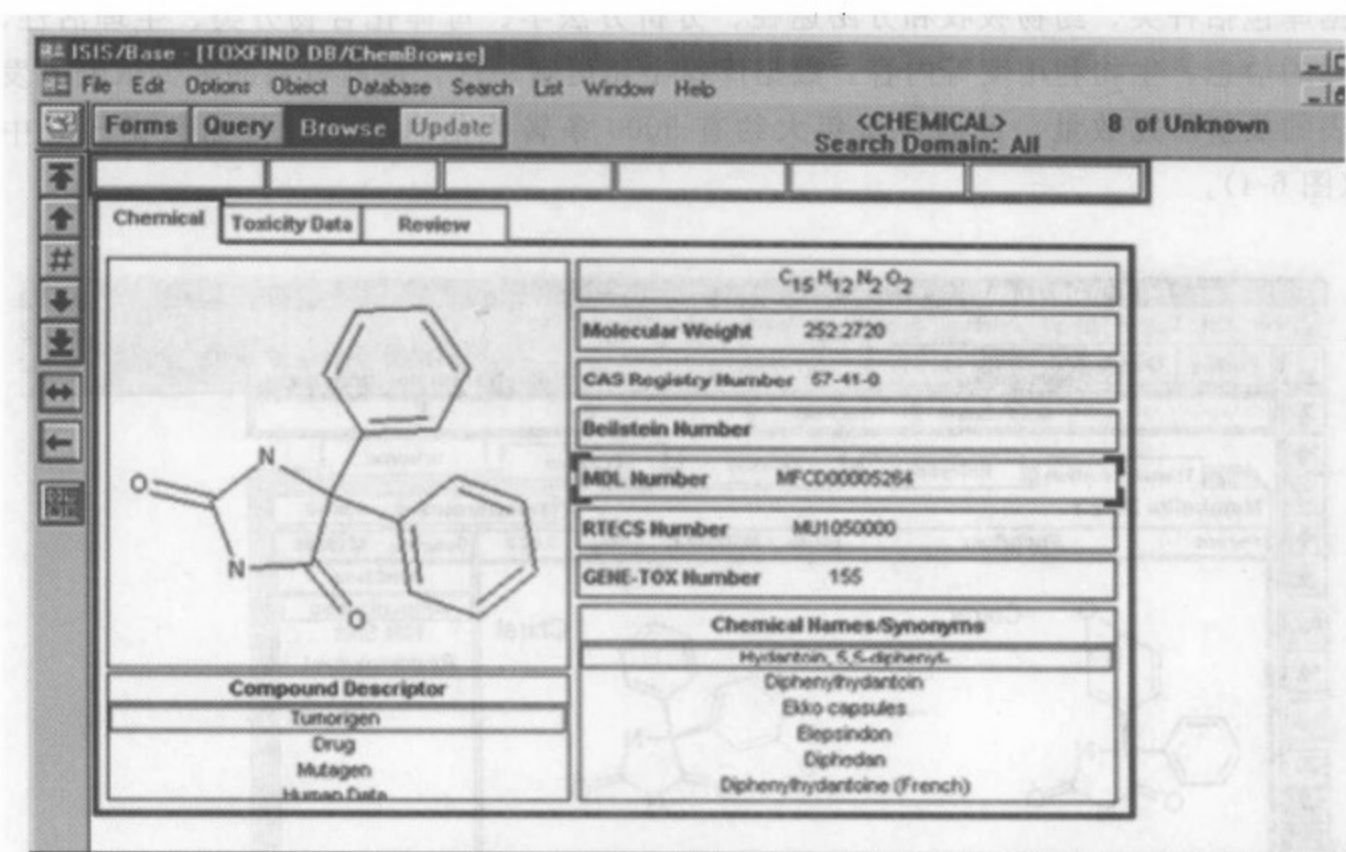


图 6-5 MDL® 毒性数据库检索界面

库 (NCI Database)，它包含：NCI Database (126705 个化合物)；Plated Compounds Database (139735 个化合物)；Cancer Screened Database (37330 个化合物)；AIDS Screened Database (42687 个化合物)。最近一个版本 NCI Databases 2001.1 共包括 213628 个化合物 (以及 212869 个 3D 模型)。NCI 数据库最大的特点是拥有与之相配套的对公众开放的实物库，通常定期或不定期地补充化合物的实物储备，一般情况下在 NCI 数据库中始终约有 60% 的化合物实物储备 (图 6-6)。

6.2.6 天然产物数据库 (DNP)

天然产物数据库 (Dictionary of Natural Products, DNP) 是一个天然产物源非常全面的化学数据库，这个数据库将从 1996 年发表的每个天然产物实际分离和报告的数据汇集起来。其电子版每 6 个月升级一次。DNP 提供大约 214500 个化合物的物理和化学的数据，这些数据主要来源于已有的出版物。每次查询结果包含以下信息：化学名称、系统命名和俗名；精确的高质量的结构图标；分子式和分子量；天然源数据库；使用和价值；重要的参考文献资料数据库。其电子版的快速、直观的搜索界面可以使用户通过多种手段查询：精确的结构、次级结构、新的查询、增加现有的查询、更新或者进一步精确现有的查询、增强的查询功能。DNP 有 7 卷，其电子版提供的文本内容同样适用于计算机按化学名称、分子式、CAS 登记号、化合物的类型和生物学种进行分类。

NCI Databases 2001.1 - MDL Information Systems, Inc.										
NCI: Y		Plated: Y		Cancer: Y		AIDS: Y		Formula:		
NSC#: 612049		CAS#: 63655-05-6		Mol. Wt.: 236.230		C ₁₀ H ₁₂ N ₄ O ₃				
C A N C E R										
	Panel Name	PN #	Cell Line Name	Cl #	Max. Tests	Cell Tests	Conc. Unit	Hi Conc. (log)	Result (-log)	Sensitivity (log)
G I S B	LNS	1	NCI-H23	1	1	1	M	-3.000	3.000	-0.659
	LNS	1	NCI-H23	1	1	1	M	-4.000	4.000	0.341
	LNS	1	NCI-H23	1	9	9	M	-4.000	4.000	0.341
T G I	LNS	1	NCI-H23	1	1	1	M	-3.000	3.000	-0.656
	LNS	1	NCI-H23	1	1	1	M	-4.000	4.000	0.344
	LNS	1	NCI-H23	1	9	9	M	-4.000	4.000	0.344
L C S B	LNS	1	NCI-H23	1	1	1	M	-3.000	3.000	-0.656
	LNS	1	NCI-H23	1	1	1	M	-4.000	4.000	0.344
	LNS	1	NCI-H23	1	9	9	M	-4.000	4.000	0.344
A I D S										
	# of Tests	Eff.	Conc. Unit	Hi Conc. (ug/ml)	Midpoint (ug/ml)	Hi Conc. (molar)	Midpoint (molar)			
L C S B	2	=	M	2.36E+000	6.54E-001	1.00E-005	2.77E-006			
	2	>	M	2.36E+000	2.04E+000	1.00E-005	0.64E-006			
	13	=	M	1.48E+002	3.38E+000	6.25E-004	1.43E-005			
I C S B	2	>	M	1.40E+002	1.40E+002	6.25E-004	6.25E-004			
	2	>	M	2.36E+000	2.36E+000	1.00E-005	1.00E-005			
	3	>	M	1.48E+002	1.48E+002	6.25E-004	6.25E-004			
Activity: CA [CI=Confirmed Inactive CA=Confirmed Active CM=Confirmed Moderately Active]								Names:		
								2',3'-Dideoxyinosine, hydrate		
								Inosine, 2',3'-dideoxy-		

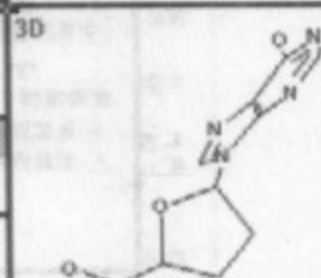
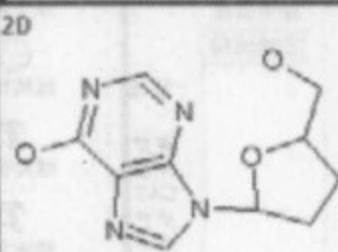


图 6-6 NCI 数据库检索界面

除了这些国外的数据库，我国也自行研发了一些有特色的数据库。

6.2.7 中草药三维结构、性质数据库 (CHDD)

中草药三维结构、性质数据库 (Chinese Herb Drug Database, CHDD) 是本书编著者所在的实验室自行开发的中草药数据库，目前已经发展了第二版，它主要收集、整理从中草药中提取的天然化合物的三维结构信息。目前收集了化合物记录 10564 条；有三维结构的 9820 条；来源植物记录 2037 条；涉及文献 10814 篇。存储分子结构信息的表存放所有与结构有关的信息，如三维结构、标准二维结构图、经过处理的连接表等；存储分子性质的表存放所有与结构无关的性质信息，如 CAS 登记号、来源植物、化合物的生理和药理作用、活性剂量、相关文献等等；存储库化合物基本信息的表用于存放化合物名称信息，如名称、IUPAC 名、俗名、别名等。提供了分子量、分子式等结构检索和名称、原植物、生理药理相关、CAS 登记号等性质检索 (图 6-7)。

6.2.8 中国天然产物数据库 (CNPD)

中国天然产物数据库 (Chinese Natural Product Database, CNPD) 是创腾科

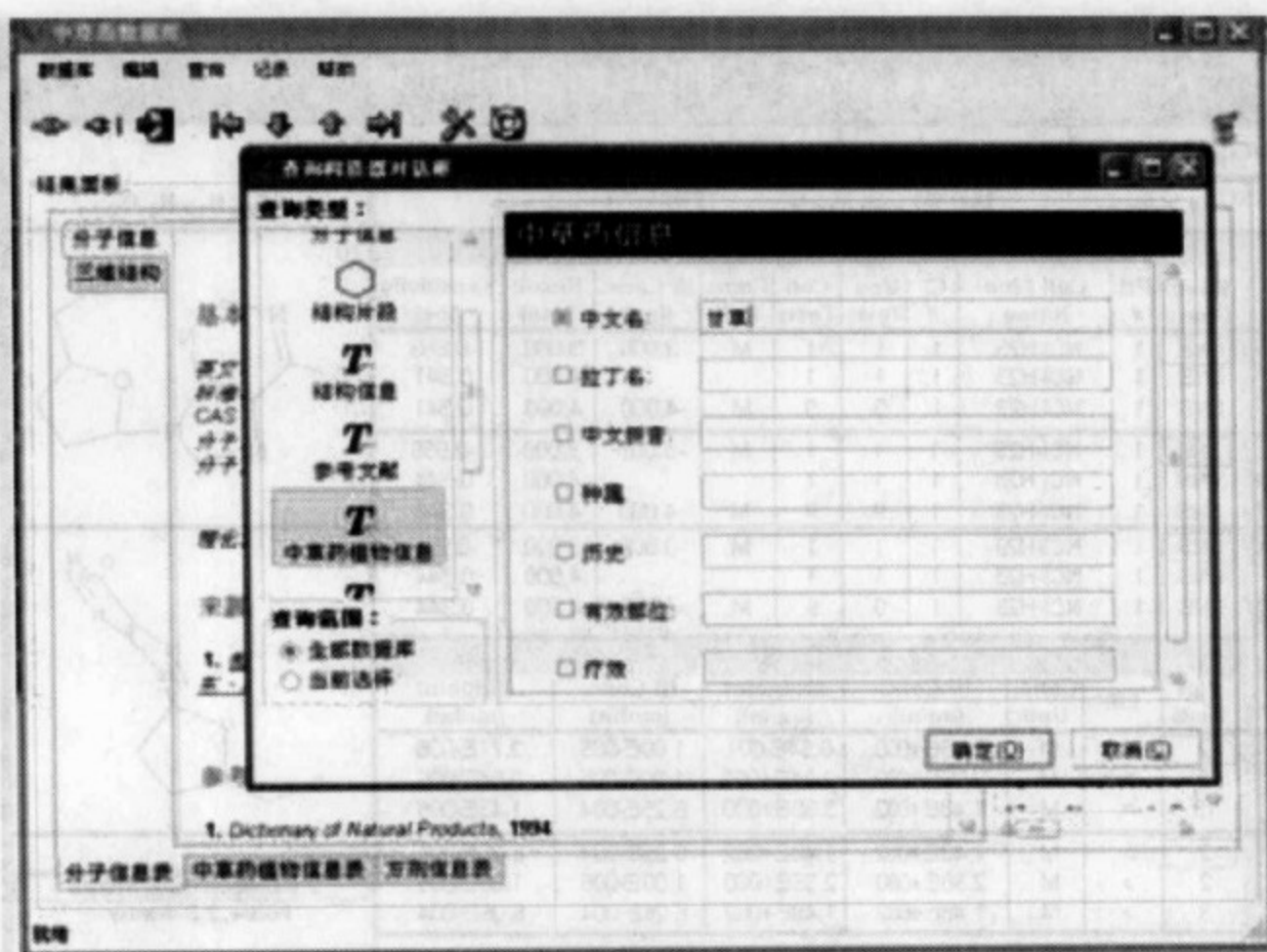


图 6-7 CHDD 检索界面

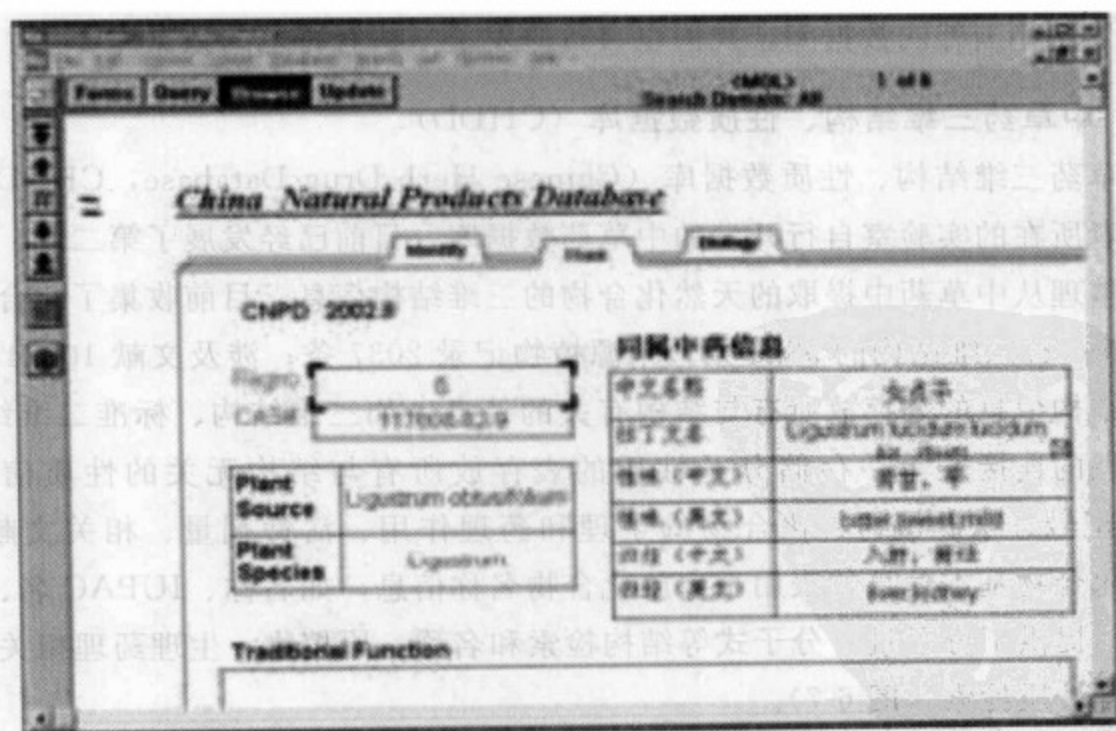


图 6-8 CNPD 检索界面

技有限公司和中国科学院上海药物研究所联合开发的综合性天然产物数据库。这一数据库是面向相关领域的科学家建立的易于使用的综合性天然产物数据库系统，该产品力图系统地收集、整理、分析从中国国产的植物中分离鉴定出的天然产物的物理性质、生物活性及化学结构等信息。CNPD 目前共收集了 57000 多个天然产物，涵盖天然产物的 37 个类别。而且类药性分析结果表明，CNPD 中有 70% 的分子是类药性分子（图 6-8）。

此外，对药物设计的一个重要的库是蛋白质数据库（Protein Data Bank, PDB）。截至 2007 年 4 月 3 日，PDB 库收集了 42627 个结构。

6.3 分子对接方法

分子对接（dock）的最初思想起源于 100 年前 E. Fisher 的“锁和钥匙模型”，即“一把钥匙开一把锁”。E. Fisher 认为，“锁和钥匙”互相识别的首要条件是它们在空间形状上要互相匹配。当然分子对接如药物分子与靶酶分子之间的识别要比“锁和钥匙”模型复杂得多。首先“锁”和“钥匙”是柔性的，即底物分子和靶酶分子的构象是变化的，而不是刚性的，底物分子和靶酶分子在对接过程中互相适应对方，从而达到更完美的匹配。分子对接和“锁和钥匙”模型的另一个不同之处是分子对接不仅要满足空间形状的匹配，还要满足能量的匹配，底物分子与靶酶分子能否结合以及结合的强度最终是由形成此复合物过程的结合自由能变化 ΔG_{bind} 决定的。

互补性（complementarity）和预组织（pre-organization）是决定分子对接过程的两个重要原则，前者决定识别过程的选择性，后者决定识别过程的键合能力。当然互补性包括空间结构的互补性和电学性质的互补性。1958 年 Koshland 提出了分子识别过程中的诱导配合（induced fit）概念，指出底物与受体互相结合时，受体将采取一个能同底物达到最佳结合的构象，这个过程被称为识别过程中的重组织。受体与底物分子在识别之前将受体中容纳底物的环境组织得愈好，其溶剂化能力愈低，则它们的识别效果愈佳，形成的复合物愈稳定，这就是分子识别的预组织原则。

6.3.1 分子对接方法的分类

分子对接方法根据不同的简化程度大致可以分为三类：刚体对接、半柔性对接以及柔性对接。刚性对接指在对接过程中研究体系的构象不发生变化，其中比较有代表性的方法就是 Jiang 等发展的软对接（soft dock）方法。半柔性对接指在对接过程中研究体系尤其是配体的构象允许在一定范围内变化，其中比较有代表性的方法有 Kuntz 等以及 Olson 等发展的分子对接方法。柔性对接是指在对接过程中研究

体系的构象基本是可以自由变化的，其中比较有代表性的方法有 Accelrys 公司发展的基于分子力学和分子动力学的分子对接方法。当然这只是一简单的分类方法，在很多分子对接程序方法中实际上采用了多种处理方法。如在 Accelrys 发展的 Affinity 程序中，作者实际上把半柔性和柔性的分子对接方法进行了结合。刚性对接适合考察比较大的体系，如蛋白质和蛋白质以及蛋白质和核酸之间的对接，它计算较为粗略，原理也相对简单。半柔性对接方法适合于处理小分子和大分子之间的对接，在对接的过程中小分子的构象一般是可以变化的，但大分子如靶酶则是刚性的，由于小分子相对较小，因此在一定程度考察柔性的基础上还可以保持较高的计算效率。在药物设计尤其在基于分子对接的数据库搜索中，一般采用半柔性的分子对接方法。柔性对接方法一般用于精确考察分子之间的识别情况，由于在计算过程中体系的构象是可以变化的，因此柔性对接需要耗费较长的计算时间。

6.3.2 分子对接方法中的重要问题

分子对接的目的是找到底物分子和受体分子之间最佳的结合位置。因此，分子对接会面对两个重要的问题：如何找到最佳的结合位置以及如何评价对接分子之间的结合强度。当然，这两个问题也是互相联系的。如何找到最佳的结合位置牵涉到优化的问题。底物分子和受体分子都是可以自由转动和平动的，同时两个分子自身构象也存在变化，因此它们之间可能的结合方式是非常复杂的。显然，不可能采用简单的系统搜索来评价每种可能的结合模式，而需要引入其他高效的优化方法。在分子对接中，一些智能化的优化算法已经得到了广泛的应用，包括遗传算法、模拟退火以及禁忌搜索等，在下面具体分子对接的方法介绍中将会提到这些方法。

分子对接中的第二个问题就是底物分子和受体分子之间结合能力的预测，它牵涉到结合自由能预测的问题。分子对接过程中结合自由能 (ΔG_{bind}) 的变化包括下面几个部分的贡献，如式(6-1)所示：

$$\begin{aligned}\Delta G_{\text{bind}} &= \Delta G_{\text{gas}} - \Delta G_{\text{solv}}^{\text{A}} - \Delta G_{\text{solv}}^{\text{B}} + \Delta G_{\text{solv}}^{\text{AB}} \\ &= \Delta H_{\text{gas}} - T\Delta S - \Delta G_{\text{PBSA}}^{\text{A}} - \Delta G_{\text{PBSA}}^{\text{B}} + \Delta G_{\text{PBSA}}^{\text{AB}}\end{aligned}\quad (6-1)$$

式中，A 和 B 分别表示受体分子和底物分子； ΔG_{gas} 为气态下分子对接过程的自由能变化，约为对接过程中 A 和 B 的焓变； $\Delta G_{\text{solv}}^{\text{A}}$ 、 $\Delta G_{\text{solv}}^{\text{B}}$ 、 $\Delta G_{\text{solv}}^{\text{AB}}$ 分别为受体分子、底物分子、复合物分子的溶剂化自由能； $\Delta G_{\text{PBSA}}^{\text{A}}$ 、 $\Delta G_{\text{PBSA}}^{\text{B}}$ 、 $\Delta G_{\text{PBSA}}^{\text{AB}}$ 分别为受体分子、底物分子、复合物分子用 PBSA 方法计算的自由能； $T\Delta S$ 表示对接过程中的熵变。在这几项中，气态下分子对接过程中的焓变可以通过分子力学的方法简单求算。去溶剂化自由能也是可以求算的，但去溶剂化自由能准确而快速的求算还存在一定的问题。熵变的求算可能是其中最大的问题，虽然可以通过正交模式分析 (normal-mode analysis)、准协波分析 (quasi-harmonic analysis) 以及准高斯方法 (quasi-Gaussian approach) 等方法求算熵变，但这些方法都需要耗费大量的时间。

在实际的药物设计过程中, 研究人员总是希望能快速地筛选成千上万的分子。由于受到目前计算能力的约束, 不可能采用严格的自由能计算方法来准确评价靶酶分子和每个底物分子之间的结合自由能, 而只能采用较为简单的自由能评价方法。因此在实际的分子对接计算中, 人们往往只考虑靶酶分子和底物分子之间的非键相互作用能, 或者用简单的基于分子表面的溶剂化计算方法来考虑对接过程中溶剂化能的变化。当然, 在一些分子对接程序中也采用了一些半经验的自由能计算方法。但总体来说, 目前分子对接程序中所采用的结合自由能计算方法还相对较为简单。

6.3.3 几种有代表性的分子对接方法

分子对接的方法有很多, 这里不一一介绍。下面主要介绍几种应用较为广泛的分子对接方法, 其中有两种方法已经实现商业化。

6.3.3.1 DOCK

DOCK 是 Kuntz 的研究小组发展的分子对接程序, 可能是目前应用最为广泛的分子对接程序之一。它能自动地模拟配体分子在受体活性位点的作用情况, 并把理论预测最佳的方式记录下来。而且该方法能够对配体的三维结构数据库进行自动搜索, 因此广泛应用于基于受体结构的数据库搜索的药物设计中, 并取得了巨大的成功。用 DOCK 进行药物设计以及数据库的搜索基本上可以分为下面几个步骤: 配体和受体相互作用位点的确定、评分系统的生成、DOCK 计算及 DOCK 结果的处理与分析。

活性位点的确定和表达是 DOCK 最重要的特点之一。活性位点特征的确定对于 DOCK 研究是非常重要的, 因为配体分子和受体相互作用过程的模拟主要就是参考几何位点的几何特征进行的。在 DOCK 中, 活性位点的确定通过 sphgen 程序来完成。DOCK 软件包中 sphgen 程序生成受体表面所有的凹陷的负像, 并对这些负像进行聚类分析。图 6-9 显示了一个与两个表面点相切的负像以及在活性口袋中互相叠合的多个负像。

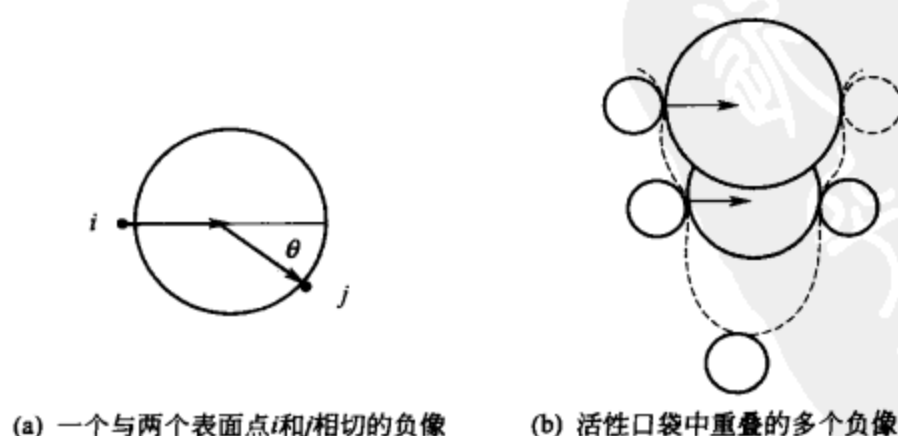


图 6-9 DOCK 程序生成的负像

在图 6-9(a) 中, 黑色的小球代表表面点。在 DOCK 程序中, 表面点采用了 Richards 提出的模型。在这些表面点的基础上, 采用 sphgen 程序生成了负像, 它实际上由一些与分子表面点相切的圆球叠加而成。

在生成的负像的基础上, 就可以进行配体分子和活性口袋之间的匹配。在这里, 配体也采用一组球集来表示, 与负像不同的是配体所用的球集表示配体所占的空间区域。如果一个配体分子能和活性口袋形成比较好的匹配, 那么配体的球集一定能和活性口袋中的负像形成好的叠合。配体分子和负像之间的匹配原则是基于配体和受体之间球集的内坐标的比较。按照匹配原则得到了配体和受体之间的匹配情况之后, 就要通过合理的得分函数来选择最优的结果。DOCK 提供了多种得分函数来评价配体和受体之间的结合情况, 包括原子接触得分以及能量得分。

DOCK 提供了简单的评价表面匹配的评价函数。这个得分函数为配体和受体之间接触重原子数的简单加和。DOCK 把配体和受体之间的非键相互作用能作为能量的匹配得分。能量可以用下式计算得到:

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left(\frac{A_{ij}}{r_{ij}^a} - \frac{B_{ij}}{r_{ij}^b} + 332 \frac{q_i q_j}{D r_{ij}} \right) \quad (6-2)$$

式中, E 表示配体和受体之间的相互作用能, r_{ij} 为原子 i 和原子 j 之间的距离, A_{ij} 、 B_{ij} 为范德华排斥和吸引参数, a 、 b 表示范德华吸引和排斥方次, q_i 、 q_j 为原子 i 和原子 j 上的部分电荷, D 为介电函数。

为了考察溶剂效应对分子对接的影响, Kuntz 的科研小组在式(6-2) 中引入了基于普适波恩模型的去溶剂化能量项。计算结果表明引入溶剂效应可以改善分子对接工作, 但这部分工作在最新的 DOCK 软件包中还没有体现。

DOCK 进行分子对接时, 配体分子可以是柔性的。对于柔性的分子, 其键长和键角保持不变, 但可旋转二面角是可以发生变化的。在 DOCK 中, 柔性分子的构象变化通过下面的操作实现: 首先是刚性片断的确定, 然后是构象搜索。构象搜索采用两种方法: 一种方法是锚优先搜索 (anchor-first search); 第二种方法是同时搜索 (simultaneous search)。

6.3.3.2 AUTODOCK

AUTODOCK 是 Scripps 的 Olson 科研小组开发的分子对接软件。AUTODOCK 采用模拟退火和遗传算法来寻找受体和配体最佳的结合位置, 用半经验的自由能计算方法来评价受体和配体之间的匹配情况。

在 AUTODOCK 中, 配体和受体之间结合能力采用能量匹配来评价。在 1.0 和 2.0 版本中, 能量匹配得分采用简单的基于 AMBER 力场的非键相互作用能。非键相互作用来自于 3 部分的贡献: 范德华相互作用、氢键相互作用以及静电相互作用。范德华相互作用采用 Lennard-Jones 12-6 函数形式:

$$E_{\text{vdw}} = \sum_{i < j, R_{ij} < R_{\text{cut}}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad (6-3)$$

氢键相互作用采用传统的 12-10 势函数形式：

$$E_{\text{H-bond}} = \sum_{i < j, R_{ij} < R_{\text{cut}}} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \quad (6-4)$$

$$C_{ij} = 5\epsilon_{XY} r_{\text{eqm}, XY}^{12}, \quad D_{ij} = 6\epsilon_{XY} r_{\text{eqm}, XY}^{10}$$

静电相互作用采用常用的库仑相互作用，见下式：

$$E_{\text{elec}} = \sum_{i < j, R_{ij} < R_{\text{cut}}} \left[\frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \right] \quad (6-5)$$

介电函数可以设为常数或者是与距离有关的函数，后者在 AUTODOCK 中采用 S 函数的形式：

$$\epsilon(r) = A + \frac{B}{1 + k e^{-\lambda B r}} \quad (6-6)$$

式中， $B = \epsilon_0 - A$ ； ϵ_0 为 78.4（水在 25° 时的介电常数）； $A = -8.5525$ ； $k = 7.7839$ ； $\lambda = 0.003627 \text{ \AA}^{-1}$ （ $1 \text{ \AA} = 0.1 \text{ nm}$ ）。

在 3.0 版中，AUTODOCK 提供了半经验的自由能计算方法来评价配体和受体之间的能量匹配，计算采用下面的函数形式，结合自由能来自于 5 部分的贡献：

$$\begin{aligned} \Delta G = & \Delta G_{\text{vdw}} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \Delta G_{\text{hbond}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + \\ & \Delta G_{\text{ele}} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + \Delta G_{\text{tor}} N_{\text{tor}} + \Delta G_{\text{sol}} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\delta^2)} \end{aligned} \quad (6-7)$$

式中， ΔG_{vdw} 、 ΔG_{hbond} 、 ΔG_{ele} 、 ΔG_{tor} 以及 ΔG_{sol} 都是半经验参数，通过拟合得到； N_{tor} 指配体在对接后被约束的可旋转键的数目。式中的静电相互作用、范德华相互作用、氢键相互作用的计算方法与 1.0 以及 2.0 版本中一样，只是每个计算出来的能量值需要乘上相应的权重系数。

在最早的 AUTODOCK 版本中，作者采用了模拟退火来优化配体和受体之间的结合。在 3.0 版本中，Morris 等发展了一种改良的遗传算法，即拉马克遗传算法（LGA）。测试结果表明，LGA 比传统的遗传算法和模拟退火具有更高的效率。在 LGA 方法中，作者把遗传算法和局部搜索（local search）结合在一起，遗传算法用于全局搜索，局部搜索用于能量优化。在 AUTODOCK 中，局部搜索方法是自适应的，它可以根据当前的能量调节步长大小。

LGA 算法引入了拉马克的遗传理论，这个操作过程可见图 6-10。LGA 最大的特点就是通过进化映射（developmental mapping）把基因型转化为表现型而实现局部搜索和遗传算法的结合。基因型空间通过遗传算子突变和交叉来定义；表现型通过问题的解来定义，这里表示体系的能量得分。

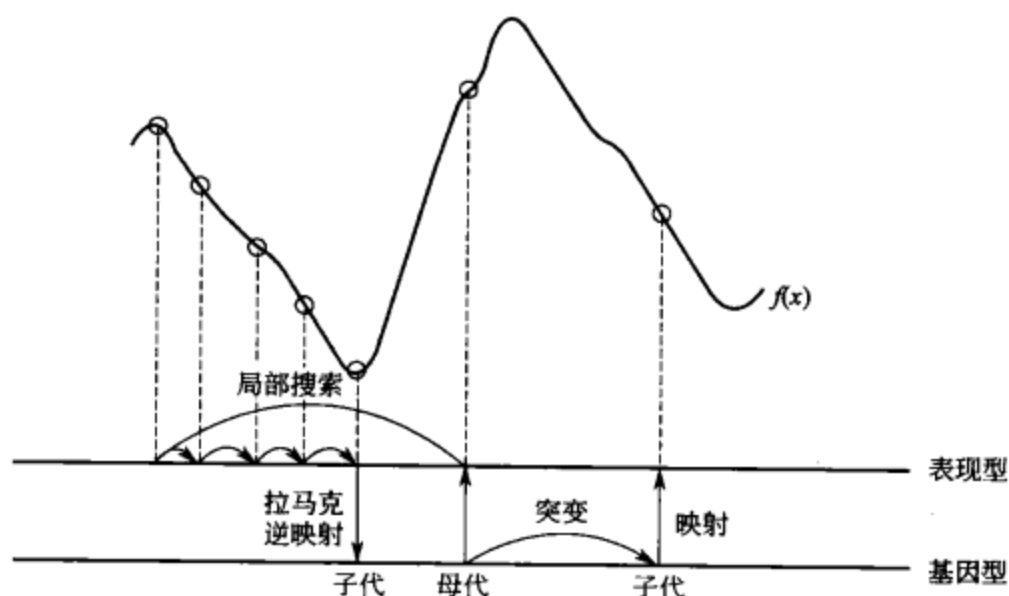


图 6-10 拉马克遗传算法的基本步骤

6.3.3.3 FlexX

FlexX 是德国国家信息技术研究中心生物信息学算法和科学计算研究室的 Matthias Rarey 等发展的分子对接方法，现在已经作为 SYBYL 分子模拟软件包中的一个模块实现了商业化。FlexX 中结合了多种药物设计的方法进行配体和受体之间的对接。

在 FlexX 中，配体和受体之间的结合情况的评价采用了类似 Böhm 提出的基于半经验方程的自由能评价方法，方程的形式见下式：

$$\Delta G = \Delta G_o + \Delta G_{rot} \times N_{rot} + \Delta G_{hb} \sum_{\text{neutral H-bonds}} f(\Delta R, \Delta \alpha) + \Delta G_{io} \sum_{\text{ionic int}} f(\Delta R, \Delta \alpha) + \Delta G_{aro} \sum_{\text{aroint}} f(\Delta R, \Delta \alpha) + \Delta G_{lipo} \sum_{\text{lipo. int.}} f^*(\Delta R) \quad (6-8)$$

式中， N_{rot} 是在结合过程中被冻结的可旋转键的数目； ΔG_{rot} 、 ΔG_{hb} 、 ΔG_{io} 以及 ΔG_o 是通过拟合得到的半经验参数； $f(\Delta R, \Delta \alpha)$ 为偏离最优构象的罚分函数； $\Delta R = R - R_o$ ，其中 R 为原子间的距离， R_o 为原子间的理想距离，一般为原子对的范德华半径和加上 0.6。这些参数以及 f 函数的形式直接采用 Böhm 报道的结果。此外，对于芳香基团的贡献，FlexX 采用了一个新的参数，即 $\Delta G_{aro} = -0.70 \text{ kcal/mol}$ 。式中最后一项表示亲脂原子的接触得分， ΔG_{lipo} 为亲脂原子接触得分的半经验参数，其中 $f^*(\Delta R)$ 在不同的 ΔR 下有不同的取值。

在 FlexX 中，分子对接的流程主要分为下面的步骤。

(1) 核心结构的确定 对接的第一步为核心片段 (base fragment) 的选择。核心片段是指能对配体和受体之间相互作用起决定作用的基团，而且核心片段的构象要尽量少一些。核心基团的正确选择对分子对接的计算结果有非常重要的影响。因为如果核心基团和受体之间不存在明显优势的相互作用，则很难预测正确的结合模

式。随着核心基团的增加，核心基团和受体之间的相互作用也会相应增强，那么准确预测结合模式的机会会大大增加。因此在选择核心基团时，核心基团包含的基团要尽量多一些，而且核心基团的构象数要尽量少一些。核心基团选定以后，就可以把配体分子划分为多个片断。

(2) 核心结构的放置 选择好了核心基团以后，就要把核心基团放置在活性位点的正确部位。在放置核心基团时，FlexX 采用了一种形态聚类算法 (pose-clustering algorithm)。在这个算法中，一个核心基团可以看作为一个具有明确相互作用点的刚性物体，受体分子的活性口袋也可以看作为一个具有明确相互作用点的刚性物体，把核心基团放置在活性口袋中的过程就相当于把配体中的 3 个相互作用点叠合在活性位点的 3 个相互作用点上 (假设这 3 个点不共线)。在匹配中，两个三角形的 3 个顶点所具有的相互作用特征应该是符合的，同时三角形对应边长的差别应该在一定的范围内。放置核心结构的第一步就是找出所有相匹配的这些三角形，而且对配体的位置进行坐标转化。所有的转化完成以后，检查配体是否和受体产生了碰撞，去掉一些不合理的核心结构取向。

对于得到的这些核心结构的可能位置，通过核心结构空间的位置均方根位移 (r. m. s. D) 进行聚类分析。对于那些 r. m. s. D 值小于一定阈值的空间位置进行归并，仅仅保留那些相差较大的空间位置。最后，检查这些核心结构和受体之间的相互作用情况，对结构进行简单的修正。

(3) 配体分子的生长 核心结构在活性口袋中的位置确定以后，配体分子的其他部分可以分为小的片断，依次“生长”在核心结构上。片断生长采用树形搜索 (tree search) 的方法，算法和 SYBYL 中的系统搜索所采用的算法基本类似，在搜索的过程中要尽量删除那些无用的分支。搜索树的第一层是核心结构在活性口袋中的不同放置位置。在下一层中，片断要采用尽可能多的形式连接到核心结构上。然后，按照树形结构，片断依次连接。如果片断能和受体形成氢键或盐桥，则优先连接，因为相互作用形式越明确，片断的几何定位越容易。在对接过程中，一个完全无遗漏的树形结构在操作上是很难的，因此对于树的每个节点，仅仅只考虑最佳的 k 种结果，这样会为下面的生长节约所要耗费的计算开支。 k 种最佳的结构采用上面介绍的得分函数来选择。

增加新的基团以后，如果发现新的相互作用或配体和受体之间存在重叠，则需要对配体的位置进行优化。FlexX 采用点的加权叠合法 (weighted superposition of points) 来进行优化。优化要使得下式得到的量最小：

$$E = \sum w_i (l_i - r_i)^2 \quad (6-9)$$

式中， l_i 表示配体的相互作用中心， r_i 为受体上的相互作用中心， w_i 为相互作用在最佳几何状态下的能量贡献。

配体生长结束以后,可以得到 k 个最佳的配体和受体的结合形式,从中用户可以选择需要的结果。在 FlexX 的分子对接过程中,配体的生长以及柔性的考察与 DOCK 中采用的锚优先方法基本类似。

FlexX 已经实现了商业化,而且程序可以自动地对数据库中的多个分子进行分子对接的计算,然后给出最佳的结果。同时 FlexX 提供了友好的图形界面,易于操作,因此可能在药物设计中具有较好的应用前景。

6.3.3.4 Affinity

Affinity 是 Accelrys (MSI) 和杜邦公司联合开发的分子对接方法,也是最早实现商业化的分子对接方法。Affinity 中提供了多种分子对接的策略,可以根据用户的需要提供多种方法的组合。在 Affinity 中,分子对接大致分为两个步骤:首先,通过蒙特卡罗或模拟退火计算来确定配体分子在受体活性口袋中可能的结合位置;然后,在第一步的基础上,采用分子力学或分子动力学方法进行进一步细致的分子对接。与其他分子对接方法比较, Affinity 具有自己的特色。首先, Affinity 中提供了多种对接方法的结合,比如蒙特卡罗方法和分子力学、分子动力学以及模拟退火方法的结合,这些方法结合的灵活性为多种分子对接问题提供了解决方案。第二,在 Affinity 中,不仅仅配体是柔性的,受体的重要部位如活性位点中的某些残基也可以定义为柔性的区域。第三, Affinity 提供了精确和快速计算配体与受体之间非键相互作用的两种有效方法:一种是基于格点的能量计算方法,另一种是单元多偶极 (cell multipole method) 方法。第四, Affinity 采用了 Stouten 提出的溶剂化模型来考察配体和受体在堆积过程中溶剂化能的变化。

Affinity 中,配体和受体之间匹配主要采用能量得分的评价方式。对于能量得分, Affinity 可以采用两种力场,即 CVFF 力场和 CFF 力场。Affinity 中提供了基于格点的能量计算方法以及不基于格点的能量计算方法。对于这两种不同的方法,其能量评分函数略有区别。在基于格点的能量计算中,可以考虑溶剂效应的影响。能量评价函数可以写为下面的形式:

$$\Delta G = \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \sum_{i,j} \frac{q_{ij}}{\epsilon(r_{ij}) r_{ij}} + \Delta G_{\text{hond}} - S_i f_j e\left(\frac{-r_{ij}^2}{2\sigma^2}\right) \quad (6-10)$$

式中,前两项为静电相互作用和范德华相互作用;第三项为键相互作用,在不同的力场中有不同的表达形式;第四项为去溶剂化自由能,表示一个原子从溶液环境到具有明确相邻原子环境的自由能变化,其中 S_i 为溶解参数 (solvation parameter), f_j 为原子 j 的分数体积 (fragmental volume), j 原子替代 i 原子周围溶剂分子的能力与 f_j 和高斯包络函数的乘积成正比。

在 Affinity 中,为了重点考虑对接过程中某些作用力的影响,可以在式(6-10)中加入某些约束力,包括氢键接触力、配体限制力以及位置约束力,这些约束力在基于格点的能量计算中不能使用。

Affinity 提供了一套多种方法相结合的分子对接流程，整个流程都是自动化的，不需要用户给予任何干预。Affinity 方法因为采用了多种方法的结合，为用户解决不同的问题提供了不同的解决方案。Affinity 适合对配体和受体之间的相互作用模式进行精细考察，但它不太适合对大量的配体分子进行基于分子对接的虚拟筛选，因为 Affinity 对配体和受体都采用了柔性的策略，需要消耗较大的计算量。

6.4 定量构效关系方法研究

定量构效关系 (quantitative structure-activity relationships, QSAR) 研究在现代药物设计方法中具有非常重要的地位，也是应用最为广泛的药物设计方法。所谓定量构效关系就是通过一些数理统计方法建立起一系列化合物的生理活性或某种性质与其物理化学性质之间的定量关系，通过这些定量关系可以预测化合物的生理活性或某些性质，可以指导人们设计出具有更高活性的化合物。

早在 1867 年，Crum-Brow 和 Fraser 就提出了构效关系的概念。后来，在 1900 年前后，Overton 和 Meyer 等提出了麻醉作用的类脂学说，即化学结构各异的麻醉剂的活性随着脂-水分配系数的增加而增加的现象。这可能是最早提出的化合物生理活性和物理化学性质之间的定量关系模型。但只是到了近几十年，尤其是 Hansch 法提出以后，随着计算机技术的发展和多变量解析技术的引入，定量构效关系方法才逐渐发展和应用起来。现在，定量构效关系的研究已经成为药物设计和开发中不可缺少的工具。

6.4.1 二维定量构效关系方法

传统的二维定量构效关系 (2D-QSAR) 方法很多，有 Hansch 法、模式识别 Free-Wilson 法、电子拓扑法等。其中最为著名、应用最为广泛的就是由 Hansch 和 Fujita 提出来的 Hansch 法，它假设同系列化合物某些生物活性的变化是与它们的某些可测量的物理化学性质的变化相联系的，这些可测量的特性包括疏水性、电性质、空间立体性质等，都有可能影响化合物的生物活性。Hansch 法假定这些因子是彼此孤立的，故采用多重自由能相关法，借助多重线性回归等统计方法就可得到定量构效关系模型。

Hansch 法最初可以表达为下面的式 (6-11)，即活性与疏水性参数 π 或 $\lg P$ 、电性参数 σ 以及立体参数 E_s 相关。

$$\lg(1/C) = a\lg P + b\sigma + cE_s + \cdots + \text{常数} \quad (6-11)$$

后来 Hansch 发现，药物要交替穿过水相和类脂构成的体系，其移动难易程度与 $\lg P$ 呈现函数关系。如果经过一定时间后药物在最末一相中为浓度 $\lg C$ ，以 $\lg C$ 对 $\lg P$ 作图，可以发现它们之间呈抛物线关系，因此式 (6-11) 就表达为式 (6-12)

的形式。实验表明,式(6-11)适用于体外活性数据,而式(6-12)适用于体内活性数据。

$$\lg(1/C)=a(\lg P)^2+b\lg P+c\sigma+cE_s+\cdots+\text{常数} \quad (6-12)$$

在 Hansch 和 Fujita 等最初采用的构效关系模型中仅仅使用了一些简单的分子参数。但对于一个分子来讲,可以用很多分子参数来表达分子的不同特征,如各种拓扑参数、各种热力学参数、各种量化计算得到的参数以及各种分子形状参数等。研究也证明采用这些参数往往能得到更好的结果。因此,在实际的应用中,总是尽量挑选最佳的参数来得到最有效的模型,而不是局限于 Hansch 和 Fujita 公式最初的形式。

此外,除了传统的线性回归方法以外,一些新的数理统计方法以及非数值算法也被应用到构效关系研究中,如偏最小二乘、人工神经网络以及遗传算法等。这些新方法的应用大大推动了 2D-QSAR 方法的发展。

6.4.2 三维定量构效关系方法

近些年来,随着构效关系理论和统计方法的进一步发展,又出现了一些三维的定量构效关系(3D-QSAR)方法,如分子形状分析(molecular shape analysis, MSA)、距离几何方法(distance geometry, DG)以及比较分子场分析方法(comparative molecular field analysis, CoMFA)等。与 2D-QSAR 比较,3D-QSAR 在物理化学意义上更为明确,能间接反映药物分子和靶点之间的非键相互作用特征,因此近十多年 3D-QSAR 方法得到了迅速的发展和广泛的应用。在 3D-QSAR 中,比较分子场分析方法(CoMFA)可能是目前最为成熟且应用最为广泛的方法。

按照 CoMFA 的基本原理,如果一组相似化合物以同样的方式作用于同一个靶点,那么它们的生物活性就取决于每个化合物周围分子场的差别,这种分子场可以反映药物分子和靶点之间的非键相互作用的特性。

CoMFA 计算可以简单地分为下面 3 个步骤。

首先,确定药物分子的活性构象,再按照一定的规则(一般为骨架叠合或场叠合)进行药物分子的叠合。

然后,在叠合好的分子周围定义分子场的空间范围,可以采用矩形或圆形的空间。把定义的空间按照一定的步长均匀划分产生格点,在每个格点上用一个探针离子来评价格点上的分子场特征(一般为静电场和立体场,有时也包括疏水场和氢键场)。

最后,通过偏最小二乘(PLS)方法建立化合物活性和分子场特征之间的关系,并给出各种分子场的等势面图。

近些年,研究人员对传统的 CoMFA 方法进行了大量的改进,其中涉及活性构象的确定、分子叠合规则、分子场势函数的定义以及分子场变量的选取等,在很大

程度上提高了 CoMFA 计算的成功率。其中,最具有代表性的可能就是比较分子相似性因子分析 (comparative molecular similarity indices analysis, CoMSIA) 方法和 4D-QSAR 方法的提出,以及遗传算法在 CoMFA 中的应用。

1994 年, Klebe 等报道了一种新的分子场分析方法,即 CoMSIA 方法。这种方法与 CoMFA 比较,最大的不同就是分子场的能量函数采用了与距离相关的高斯函数的形式,而不是传统的 Coulomb 和 Lennard-Jones 12-6 势函数的形式。在 CoMSIA 中,共定义了 5 种分子场的特征,包括立体场、静电场、疏水场以及氢键场(包括氢键受体场和氢键给体场)。这 5 种分子场由下式计算得到:

$$A_{FK}^q(j) = - \sum_{i=1}^n \omega_{\text{probe},k} \omega_{ik} e^{-\alpha \frac{r_{iq}^2}{r_0^2}} \quad (6-13)$$

式中, i 表示分子 j 中的原子序号, ω_{ik} 表示探针 (probe) 离子的某种分子场特征 (电荷为 +1, 原子半径为 0.1nm, 疏水性为 +1, 氢键给体和受体强度为 +1), α 为衰减因子 (attenuation factor), r_{iq} 表示探针离子在某格点上与分子中 i 原子之间的距离。

在 CoMSIA 中,由于采用了与距离相关的高斯函数的形式,可以有效地避免在传统 CoMFA 方法中由静电场和立体场的函数形式引起的缺陷。由于分子场能量在格点上迅速衰减,不再需要定义能量的截断 (cutoff) 值。人们对一些实际体系也进行了 CoMSIA 和 CoMFA 的比较分析。在计算中,采用了不同的格点取值,而且对计算体系均采用全空间搜索 (full space searching) 的策略。结果表明,CoMFA 计算对不同的格点大小取值以及叠合分子不同的空间取向非常敏感,采用不同的空间取向时交互验证回归系数 (q^2) 最大会相差 0.3 以上。而 CoMSIA 计算在不同格点大小取值以及分子空间取向得到的结果则稳定得多。在一般情况下,CoMSIA 计算会得到更加满意的 3D-QSAR 模型。

1997 年 Hopfinger 等提出了 4D-QSAR 的概念。研究者首次采用遗传算法来选择分子动力学产生的构象来产生最佳的构效关系模型。在这个方法中,研究者用每个格点对应的原子占有率 (atom occupancy) 作为 PLS 的变量,根据原子的不同特征定义了 7 种不同类型的原子类型。研究者考虑了药物分子的整个构象空间,而不是一个分子,而且采用了考察了多种原子叠合方式,因此在概念上比传统的 CoMFA 有了一定的进步。

6.4.3 建立定量构效关系的统计方法

在传统二维构效关系研究中,线性回归 (linear regression, LR) 是最为常见的统计方法。一个分子可以用很多分子参数来表达,但在建立线性回归模型时,为了避免过拟合 (overfitting),只能从这些物理化学参数中选择一部分参数来建立回归模型。一般来讲,同系物数目和所选取参数的数目比应大于 3~5,也有人提出

应大于 2^n (n 表示选取的参数个数), 怎样选择合适的参数一直是定量构效关系研究中的一个难题。而且对于线性回归来讲, 当体系噪声较强或干扰严重时, 有可能导致所得的模型失真。为了克服线性回归的不足, 在数学上可以采用主成分回归的方法。

所谓主成分回归就是采用主成分分析方法 (principle component analysis, PCA), 通过选取对活性影响最大的几个主成分来建立定量构效关系模型。所谓主成分, 为一组新变量, 该新变量是原来变量 x_{ij} 的线性组合。第一个主成分表征原来变量时所产生的平方误差最大, 第二个次之, 第三个再次之, 等等。也就是说, 主成分是一种线性组合, 用它来表征原来变量时所产生的平方误差最小。运用主成分分析, 原变量矩阵 X 可以表达为得分 (即主成分) 矩阵 T , 由在本征矢量上的投影所得。主成分与矩阵的本征矢量一一对应, 既 $XP=T$ 。主成分回归可以有效解决共线问题, 同时由于去掉了不太重要的主成分, 因而可以削弱噪声 (随机误差) 产生的影响。主成分回归可以分为两步: ①测定主成分数, 并由主成分分析将 X 矩阵降维; ②对于降维的 X 矩阵再做线性回归分析。但是, 由于主成分回归为二步法, 若在第一步中消去的是有用的主成分而保留的是噪声, 则第二步多元线性回归所得结果就将偏离真实的数学模型。

在主成分回归法的第一步中, 所处理的仅为 X 矩阵, 对于矩阵 Y 中的信息并未考虑。事实上, Y 中可能包含无用的信息。所以很自然的一种想法是, 在矩阵 X 因子的测试中应同时考虑矩阵 Y 的作用。偏最小二乘 (partial least square, PLS) 正是基于这种思想的一种回归方法。偏最小二乘在考虑自变量的同时也考虑了因变量的作用, 同时通过折衷各自空间内的因子使模型较好地同时描述自变量和因变量, 通过一系列变换可以有效地减少相关因素的影响。

当然, 在构效关系研究中还用到了很多其他的数理统计方法, 如逐步回归方法、非线性最小二乘方法以及一些模式识别方法等, 这里不一一详细论述。

6.5 结合自由能的计算

基于结构的药物分子设计的方法多种多样, 包括分子对接、数据库搜索以及从头设计等, 但是它们最终往往都面临一个同样的问题, 就是药物分子活性的评价。许多药物和其他生物分子的活性都是通过与受体大分子之间的相互作用表现出来的, 所以受体和配体之间的结合自由能 (binding affinity) 评价是基于结构的计算机辅助药物分子设计的核心问题。精确的自由能预测方法能够大大提高药物分子设计的效率。在过去的 20 年中, 随着受体-配体相互作用的理论研究以及计算机辅助药物分子设计方法的快速发展, 自由能预测方法的研究受到了越来越多的关注。

在药物分子设计中, 人们希望能对受体和药物分子之间的结合能力进行快速而

精确的评价。在一些分子对接的程序中,也采用了多种药物分子和受体之间结合的评价方法,如表面匹配、能量匹配(非键相互作用能)以及化学环境匹配等。虽然这些评价标准能部分反映底物和蛋白质的结合程度,但往往具有比较局限的应用范围,它们的评价能力非常有限。在这种情况下,自由能的计算就显得尤为重要了,因为结合自由能是评价底物和蛋白质结合程度的精确标准。在目前粗筛方法日臻完善的情况下,以自由能为代表的精确评价方法显得日益重要起来。

6.5.1 自由能计算方法

在配体(包括药物)和受体的相互作用过程中,牵涉到两类相互作用,即非键相互作用和共价相互作用。非键相互作用包括静电相互作用、范德华相互作用以及氢键相互作用等,这些相互作用都可以通过力场计算进行简单的表达;而共价相互作用则牵涉到化学键的断裂和生成,需要用量子力学的方法进行考察。在药物分子和受体的相互作用中,一般只牵涉到非键相互作用,共价相互作用只存在少数体系中。

根据热力学原理,结合自由能 ΔG^b 和结合常数 K_d 存在定量的关系:

$$\Delta G^b = RT \ln K_d \quad (6-14)$$

自由能的计算方法主要可以分为3类。第一类方法包括自由能微扰(free energy perturbation, FEP)和热力学积分(thermodynamic integration, TI)方法。这类方法是最为经典的方法,为大家广泛接受。这类方法在原理上比较严格,计算结果也较为精确,但需要长时间的数据采集,对计算体系有严格的限制,只能适合较为简单的情况,因此在药物分子设计中还没有达到实用化的阶段。第二类方法包括一系列基于经验方程的计算方法。这类方法把结合自由能分解为不同的相互作用能量项,通过一组训练集并利用统计方法来得到自由能计算的公式。这类方法取样简单、计算量小,在药物分子设计中应用较多。但这类方法得到的经验公式严重依赖训练集的选择,对不同的体系也具有不同的预测能力,而且这类方法不能很好地考虑体系的柔性以及溶剂效应,因此一般只能对一些体系具有有限的预测能力,也只能作为初筛的手段。第三类方法就是近几年发展起来的基于分子动力学采样的自由能预测方法,主要包括 LIE 方法和 MM/PBSA 方法。

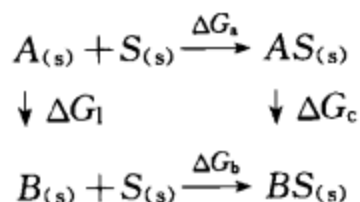
6.5.2 自由能微扰(FEP)方法

FEP 的基本思想是从一个已知体系出发,通过一系列微小的变化变到另一个体系,对每一个变化步骤做分子动力学模拟,把每一步的体系势能代入式(6-15)就可以得到两步之间的自由能变化,把所有的自由能变化加起来就得到两个体系之间的自由能变化。

$$G_{\lambda(i+1)} - G_{\lambda(i)} = -RT \ln \{ \exp - [V_{\lambda(i+1)} - V_{\lambda(i)}] / RT \}_{\lambda(i)} \quad (6-15)$$

$$\Delta G = G_1 - G_0 = \sum_i G_{\lambda(i+1)} - G_{\lambda(i)} \quad (6-16)$$

如果有两个配体分子 A 和 B, 它们和受体 S 形成复合物 AS 和 BS。为了求算这两个配体分子和受体之间结合自由能的差值, 可以采用下面的热力学循环。先求算配体分子 A 和 B 之间的自由能变 ΔG_1 , 然后再求算出复合物 AS 和 BS 之间的自由能变 ΔG_c 。 ΔG_1 和 ΔG_c 都可以采用自由能微扰的方法计算得到, 二者之差就是自由能变化 $\Delta(\Delta G)$ 。



在一般的自由能微扰的计算中, 求算的都是两个配体分子和受体之间结合能的差值。在有些应用中, 也可以采用这种方法来计算单个配体和受体之间的结合自由能。在这种情况下, 突变的过程不是一个配体到另外一个配体, 而是一个配体到相应的溶剂分子 (一般为水分子)。

FEP 方法有很多变种, 如定长窗口增长 (fixed wide window growth)、慢增长 (slow growth) 以及动态窗口增长 (dynamic modified window growth) 等, 这些方法的原理都是一样的, 不同在于计算过程中的步长 λ 的选取。

所谓定长窗口增长, 它的 λ 是一个固定的数值, 由变化步数 i 决定, 若步数为 20, 则 λ 为 0.05。

慢增长方法中, λ 是趋于零的一个数, 变化步数则趋于无穷, 因为步数多, 每步之间的变化小, 所以在每一步只做一步或几步动力学模拟。

动态窗口增长方法中, λ 是根据上一步计算自由能的结果变化的, 自由能变化若太大则 λ 小一点, 变化若太小则 λ 大一点, 在保证计算精度的同时又大大减少了计算的步骤, 可谓集二者所长。

6.5.3 热力学积分 (TI) 方法

TI 方法用下式计算两个状态之间的自由能变:

$$G_1 - G_0 = \int_0^1 \left(\frac{\partial V}{\partial \lambda} \right)_\lambda d\lambda \quad (6-17)$$

计算的步骤与 FEP 方法相似, 也是通过两次计算, 先算出两底物之间的自由能变, 再算出复合物之间的自由能变, 二者之差就是相对自由能。

与 FEP 方法一样, TI 方法也分为普通的热力学积分和动态热力学积分 (dynamic modified thermodynamic integration) 两种。

FEP 和 TI 是比较经典的自由能计算的方法, 它们的优点在于理论严格、逻辑清晰、具有普适性。但它们的缺陷也非常明显: 首先, 这两种方法计算量大, 耗时; 其次, 它们只能计算差别较小的两态之间的相对结合自由能, 当两态差别较大时很难指定变化的路径。由于这些缺陷, 这两种方法在药物分子设计中的应用受到

了很大的局限。

6.5.4 线性相互作用能 (LIE) 方法

线性相互作用能 (linear interaction energy, LIE) 的自由能计算方法是近几年发展起来的一种基于分子动力学采样的自由能计算方法, 把自由能分解为极性和非极性的贡献。它应该是介于前两类方法之间的一类方法。它也需要进行分子动力学的采样, 但与 FEP 比较, 这个方法仅仅需要对体系的始态和终态做动力学采样, 因此计算量比 FEP 大大减少。与半经验方程方法比较, 它的优势也非常明显: 首先, 它用分子动力学和蒙特卡罗模拟考虑了体系的柔性, 在正确预测复合物结构的基础上对体系的构象空间进行采样; 其次, 它采用在溶剂环境以及蛋白质环境两次分子动力学采样, 考察了去溶剂化效应对结合自由能的影响。

LIE 方法的理论基础来源于非平衡态统计物理学中的线性响应近似 (linear response approximation) 理论, 这是一个在各个学科领域中都有广泛应用的重要理论。

假想一下溶剂中简单的刚体溶质分子。溶质分子有两种状态 A 和 B, 在 A 状态下溶质分子没有电荷, 在 B 状态下溶质分子带部分电荷。这两种状态下, 体系的势能分别为:

$$V_A = V_{rr}^{gas} + V_{ss} + V_{rs}^{LJ} \quad (6-18)$$

$$V_B = V_{rr}^{gas} + V_{ss} + V_{rs}^{LJ} + V_{rs}^{el} \quad (6-19)$$

上两式中, V_{rr}^{gas} 表示溶质分子之间的相互作用, V_{ss} 表示溶剂分子之间的相互作用, V_{rs}^{LJ} 表示溶质分子和溶剂分子之间的范德华相互作用, V_{rs}^{el} 表示溶质分子和溶剂分子之间的静电相互作用。这两个态之间的能量变化就是 V_{rs}^{el} 。

这两态之间的自由能变化可以表示为

$$\Delta F_{el} = -\beta^{-1} \ln[e^{-\beta(V_B - V_A)}]_A$$

$$\beta = (k_B T)^{-1}$$

对 A 状态进行展开, 可以得到:

$$\begin{aligned} \Delta F_{el} = & (\Delta V)_A - \frac{\beta}{2} [\Delta V - (\Delta V)_A^2]_A + \frac{\beta^2}{6} [\Delta V - (\Delta V)_A^3]_A \\ & - \frac{\beta^3}{24} \{ [\Delta V - (\Delta V)_A^4]_A - 3[\Delta V - (\Delta V)_A^2]_A^2 \} + \dots \end{aligned} \quad (6-20)$$

同样, 对 B 状态进行展开:

$$\begin{aligned} \Delta F_{el} = & (\Delta V)_B - \frac{\beta}{2} [\Delta V - (\Delta V)_B^2]_B + \frac{\beta^2}{6} [\Delta V - (\Delta V)_B^3]_B \\ & - \frac{\beta^3}{24} \{ [\Delta V - (\Delta V)_B^4]_B - 3[\Delta V - (\Delta V)_B^2]_B^2 \} + \dots \end{aligned} \quad (6-21)$$

把式(6-20) 和式(6-21) 相加, 就可以得到下式:

$$\Delta F_{el} = \frac{1}{2}[(\Delta V)_A + (\Delta V)_B] - \frac{\beta}{4}\{[\Delta V - (\Delta V)_A^2]_A - 3[\Delta V - (\Delta V)_B^2]_B\} + \dots \quad (6-22)$$

把上式中的二次扩展项忽略，就可以得到下式，而且下式中的两项都可以通过势能的简单平均得到：

$$\Delta F_{el} \approx \frac{1}{2}[(\Delta V)_A + (\Delta V)_B] \quad (6-23)$$

在 A 状态下，溶质没有电荷的中心分布，因此近似可以认为 $(\Delta V)_A = (V_{rs}^{el})_A = 0$ ，就可以得到下面的简单关系：

$$\Delta F_{el} \approx \frac{1}{2}(V_{rs}^{el})_B \quad (6-24)$$

当然上式也可以通过其他的途径得到，这里不具体论述。

线性响应近似理论用于简单离子的水合作用取得了一些成功。例如 1990 年 Åqvist 用自由能微扰的方法模拟了 Na^+ 和 Ca^{2+} 的水溶液体系，计算结果与式(6-24)符合得很好。1996 年 Åqvist 又对多个体系进行了进一步的系统研究，研究表明大部分体系都能较好地符合式(6-24)，但是对一些体系尤其是包含较多氢键给体或受体的体系产生了一些偏离。

对药物分子和受体之间的结合自由能中的极性部分，可以通过线性响应近似来处理。但对于非极性相互作用和疏水效应对自由能的影响，处理起来就困难一些。对于一些长链烷烃，实验已经证明其溶剂化自由能和烷烃的链长呈线性的关系。通过计算也已经证明，对长链烷烃，其在水中和溶剂之间的平均非键相互作用能和烷烃的链长有很好的线性关系。因此，不难想到自由能中的非极性贡献和溶质与溶剂之间平均范德华相互作用能存在一定的线性关系，即 $\Delta G_{\text{bind}}^{\text{vdw}} = \alpha \Delta V_{i-s}^{\text{vdw}}$ 。从上面的推导，Åqvist 于 1994 年提出了基于线性响应近似理论的自由能计算方法，并给出了如下的公式：

$$\begin{aligned} \Delta G_{\text{bind}} &= \alpha(V_i^{\text{el}} - V_s^{\text{el}}) + \beta(V_i^{\text{vdw}} - V_s^{\text{vdw}}) \\ &= \frac{1}{2}(V_i^{\text{el}} - V_s^{\text{el}}) + \beta(V_i^{\text{vdw}} - V_s^{\text{vdw}}) \end{aligned} \quad (6-25)$$

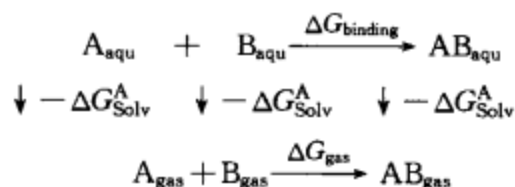
式中， ΔG_{bind} 是结合自由能， V_i^{el} 是底物和溶剂分子之间的静电相互作用能， V_i^{vdw} 是底物和溶剂分子之间的范德华相互作用能， V_s^{el} 、 V_s^{vdw} 分别是在复合物中底物和环境（包括蛋白质和溶剂）之间的静电相互作用能、范德华相互作用能， β 是非极性相互作用对结合自由能的响应系数。

对于研究的体系，需要进行两次分子动力学采样：一次在溶剂环境下；一次在蛋白质环境下。从收集得到的轨迹中可以得到 ΔV_s^{el} 和 V_s^{vdw} ，通过线性拟合就可以得到 β 参数。用得到的模型就可以进行新化合物的结合自由能预测。

6.5.5 基于 MM/PBSA 的自由能计算方法

这种方法也是近年得到广泛关注的一种自由能计算方法，它的思路实际上也很简单。与 LIE 方法不同，这种方法的自由能模型不需要拟合得到的经验参数，因此这种方法在药物分子设计中是一种非常有前景的方法。

在这种方法中，对于一个反应 $A+B \longrightarrow AB$ ，其自由能变化可以通过下面的热力学循环计算得到：



复合物中蛋白质和抑制剂间的结合自由能应该为：

$$\begin{aligned} \Delta G_{\text{bind}} &= \Delta G_{\text{gas}} - \Delta G_{\text{solv}}^A - \Delta G_{\text{solv}}^B + \Delta G_{\text{solv}}^{AB} \\ &= \Delta H_{\text{gas}} - T\Delta S - \Delta G_{\text{PBSA}}^A - \Delta G_{\text{PBSA}}^B + \Delta G_{\text{PBSA}}^{AB} \quad (6-26) \\ &= \Delta H_{\text{gas}} - T\Delta S + \Delta(\Delta G_{\text{PB}}) + \Delta(\Delta G_{\text{SA}}) \end{aligned}$$

$$\Delta H \approx \Delta E_{\text{gas}} = \Delta E_{\text{intra}} + \Delta E_{\text{electrostatic}} + \Delta E_{\text{vdW}} \quad (6-27)$$

$$\Delta(\Delta G_{\text{PB}}) = \Delta G_{\text{PB}}^{AB} - (\Delta G_{\text{PB}}^A + \Delta G_{\text{PB}}^B) \quad (6-28)$$

$$\Delta(\Delta G_{\text{SA}}) = \Delta G_{\text{SA}}^{AB} - (\Delta G_{\text{SA}}^A + \Delta G_{\text{SA}}^B) \quad (6-29)$$

式中， ΔG_{gas} 是气相中 A 和 B 之间的相互作用能，这部分可以通过分子力学 (MM) 方法计算得到； ΔG_{solv}^A 、 ΔG_{solv}^B 、 $\Delta G_{\text{solv}}^{AB}$ 分别是 A、B、AB 的溶剂化自由能，通常可以通过有限差分求解 Poisson-Boltzmann (PB) 方程计算以及溶剂可及化表面积 (SA) 的计算得到。

经典的 PB 方程为：

$$\nabla \cdot [\epsilon(r) \nabla \cdot \phi(r)] - \epsilon(r) \kappa(r)^2 \sinh[\phi(r)] + \frac{4\pi\rho^f(r)}{K_B T} = 0 \quad (6-30)$$

式中， $\phi(r)$ 为 r 点的电势； $\epsilon(r)$ 为介电常数； ρ^f 为单位电荷密度； κ 为德拜-休克尔常数， $\kappa^2 = 1/\lambda^2 = 8\pi q^2 I / e k_B T$ ， λ 为德拜-休克尔长度， I 为离子强度。变量 ϕ 、 ϵ 、 κ 、 ρ 都为 r 的函数。

以 $k_B T/q$ (k_B 为 Boltzmann 常数， T 为绝对温度， q 为单位电荷) 为电势单位，用有限差分法求解 PB 方程，得到蛋白质分子内 o 点的电势 ϕ_o 为：

$$\phi_o = \frac{\sum_i^6 \epsilon_i \phi_i + \frac{4\pi q_o}{\kappa_B T h}}{\sum_i \epsilon_i + \kappa^2 h^2} \quad (6-31)$$

式中， ϕ_i 为与 o 点相邻的 6 个网格点的电势， ϵ_i 为与 6 个正方形表面中心相关联的介电常数， h 为线性网格间隔。

溶剂效应中非极性对自由能的影响通过下式可以得到：

$$\Delta G_{PB} = 0.00542SAS + 0.92$$

(6-32)

式中, SAS 为分子的溶剂可及化表面。

式(6-26)中 S 指的是体系的熵。当一组同系列化合物和同一受体相互作用时, 考虑到它们在结合过程中的熵变对自由能的贡献基本相同, 因此熵效应可以作为一个常数项, 在这种情况下可以不考虑体系的熵效应对结合自由能的影响。熵效应的计算一般有下面几种方法: 正则模分析 (normal-mode analysis)、准协波分析 (quasi-harmonic analysis) 以及准高斯方法 (quasi-Gaussian approach)。前面两种方法可能比较适合于生物体系, 它们的原理基本相同, 不同之处在于准协波分析的原子波动矩阵不来自正交分析的计算, 而来自分子动力学的样本。考虑到正则模分析已经成功用于多个生物体系, 因此在本文中采用正则模分析来计算体系熵效应对结合自由能的影响。

6.6 计算机辅助的化合物药代动力学特征及毒性的预测

在传统的药物设计流程中, 对于候选化合物, 最为关心的往往是其对于特定靶点的活性以及对于相关靶点的选择性。至于这个化合物能否在肠道中溶解、能否通过肠道的黏膜细胞进入循环系统、能否通过屏蔽异型生物质的肝脏、能否通过血液循环到达有效部位等问题, 人们知之甚少。此外, 该化合物是否能够在体内对靶点保持足够长的药效而不被肝脏和肾脏降解或排除也是值得关注的问题。对于一个先导化合物或候选药物分子而言, 除了它的活性以外, 其药代动力学性质和毒性 (toxicity) 也是决定它能否成为一个成功药物的关键因素。西欧从 20 世纪 70 年代到 90 年代末的很长时间内, 研制出的候选药物能够最终上市的比例始终稳定在 1/10 左右。即使某个新药能够成功上市, 也有可能因为某些当时尚未了解的对人体隐形的长期危害导致其最终的失败, 这种例证不胜枚举。可见, 尽量在药物研制的早期阶段对化合物的药代动力学特征和毒性进行比较全面的评估显得尤为重要。

本节着重介绍化合物药代动力学特征和毒性的计算机辅助预测方法, 并对几种目前比较成功的软件进行简要的评述。

6.6.1 药代动力学特征的预测

在药物化学研究中, 药物的药代动力学性质是有关药物的吸收 (absorption)、分配 (distribution)、代谢 (metabolism) 和排除 (excretion) (ADME) 的性质及机理的研究。药代动力学牵涉到的研究方向很多, 目前理论研究牵涉到的部分主要包括脂水分配系数、水溶性、肠吸收、脑血穿透、生物利用度、药物-药物相互作用、主动扩散、异型生物质的代谢等。在这些理论的研究方向中, 研究较为充分的主要包括脂水分配系数、水溶性、肠吸收以及脑血穿透。

从方法学上讲,药代动力学的理论研究在很大程度上属于构效关系研究的范畴。研究首先以一组已知某种药代动力学特征的化合物作为起点,然后通过一些统计方法建立特征的预测或分类模型,研究所用到的统计方法主要包括多重线性回归、主成分分析、偏最小二乘、模式识别以及一些非数值算法等。目前,有效的药代动力学数据的获取还存在较大的困难,其原因是多方面的。首先,某些药代动力学测定实验对外界条件非常敏感,在不同的条件下很难得到稳定的数据;其次,部分药代动力学特征的实验数据还非常缺乏;最后,某些药代动力学过程很可能与多种生理学机制相关,如生物利用度、代谢稳定性等,这也增加了药代动力学理论研究的难度。

1997年,Lipinski对2287个药物分子的结构特征进行了分析,这2287个分子基本上通过了一期临床的实验。分析结构表明,如果一个药物分子具有好的吸收和穿透特性,应该符合下面的规则:

- ① 氢键给体(连在N和O上的氢原子数)少于5个;
- ② 相对分子质量小于500;
- ③ 脂水分配系数 $\lg P$ 小于5(CLOGP)或4.15(MLOGP);
- ④ 氢键受体(N和O的数目)少于5个。

一般把Lipinski得到的规则称为“五规则”(rule of 5),这个规则已经广泛用于数据库的初筛中。Lipinski等提出的“五规则”可能是最为简单的药代动力学特征评估原则,因为它所揭示的分子的结构特征实际上是与分子的通透性和渗透性密切相关的。

1999年,Ghose等对CMC(Comprehensive Medicinal Chemistry,综合医药化学)数据库中的6454个药物分子进行了统计分析,分析结果表明其中80%的分子的脂水分配系数(ALOGP)在 $-0.4 \sim 5.6$ 之间,相对分子质量在160~480之间,摩尔折射率在40~130之间,原子数在20~70之间。

6.6.1.1 脂水分配系数

脂水分配系数是构效关系研究中的一个重要的物理化学参数,同时脂水分配系数在药代动力学研究中也具有非常重要的地位。许多药代动力学特征,包括水溶性、脑血穿透以及肠吸收等,都与脂水分配系数有着非常密切的关系,因此脂水分配系数也常常被认为是一个重要的药代动力学特征。

1962年,Hansch等提出了脂水分配系数的概念。脂水分配系数最初主要应用在构效关系研究中,但随着研究工作的不断深入,它的应用扩展到药物研究和药效科学中的其他很多方面,包括药代动力学特征的预测以及分子间相互作用的研究。

Hansch等把分配系数的自然对数定义为“疏水性”($\lg P$),分配系数为一种物质在有机相和水相中浓度的比值:

$$P = \frac{c_{\text{有机相}}}{c_{\text{水}}} \quad (6-33)$$

当然，采用不同的有机相得到的分配系数的含义和数值是不同的。在药物设计中，脂水分配系数一般指物质在正丁醇和水中平衡浓度的比值。

脂水分配系数的计算方法有很多，大致可以分为 4 种：片段加和法、原子加和法、基于分子参数的计算方法以及基于转移自由能的计算方法。在这 4 种方法中，基于转移自由能 $\Delta G_{\text{transfer}}$ 的计算方法在原理上最为严格，因为转移自由能和 $\lg P$ 之间存在严格的定量关系： $\lg P = -\Delta G_{\text{transfer}} / (2.303RT)$ 。转移自由能的计算可以基于分子动力学模拟、蒙特卡罗模拟或连续介质模型，但这 3 种方法在转移自由能的计算精度上都存在很大的问题，因此基于转移自由能的方法虽然在理论研究上有一定的价值，但在实际体系的研究中基本上没有得到应用。表 6-1 列出了常用的脂水分配系数计算方法。

表 6-1 常用的脂水分配系数计算方法

名 称	计算方法	作 者	注 释
CLOGP	片段加和	Leo/Hansch	商业化, http://www.daylight.com
KLOGP	片段加和	Klopman	商业化, http://www.multicase.com
CHEMICALC	片段加和	Suzuki	
KowWin	原子和片段加和	Meylan	商业化, http://esc.syrres.com
SMILOGP	原子加和	Dubost	
ALOGPS	原子加和	Klebe	
ALOGP	原子加和	Ghose/Crippen	商业化, 用于 Cerius 和 TSAR 软件中
SLOGP	基于原子表面的原子加和	Hou	http://www.cadd.chem.pku.edu.cn
XLOGP	原子加和	Wang	
PROLOGP	片段加和	Rekker	
HINT	原子加和、疏水势	Abraham/Kellog	商业化, http://www.tripos.com
BLOGP	基于分子参数的线性回归	Bodor	

(1) 片段加和法 片段加和法是最早发展起来的一类脂水分配系数计算方法，其中具有代表性的方法有 Rekker 法、CLOGP 法、KLOGP 法以及 CHEMICALC 法等。片段加和法的基本假设是整个分子的分配系数等于组成该分子的基团或片段分配系数的加和，如下式所示：

$$\lg P = \sum_{i=1}^n a_i f_i + B \quad (6-34)$$

式中， n 为所有片段的种类数， a_i 为第 i 种片段在分子中的个数， f_i 为第 i 种片段的疏水常数， B 为常数。

式(6-34)中每种片段的疏水常数可以通过一组已知脂水分配系数的分子通过多重线性回归的方法得到。基于式(6-34)，大部分有机小分子的脂水分配可以得到很好的预测，但某些分子的计算值却和实验值有较大的偏差。偏差主要来源于分子

内相互作用、溶液中溶质分子之间的相互作用以及溶质分子和溶剂之间的相互作用的影响，如碳氢化合物的聚集以及分子内的氢键等。在加和法中，这些作用的考察主要存在两种方法：一种方法是通过引入新的片段类型，把存在相互作用的片段包含在一个新的片段之内；另一种方法是引入校正因子（correction factor），通过校正因子来修正某种特定相互作用对分配系数的影响。

如果考虑校正因子，分配系数的计算就如下式所示：

$$\lg P = \sum_{i=1}^n a_i f_i + \sum_{j=1}^m b_j F_j + B \quad (6-35)$$

式中， m 为校正因子的种类数， F_j 为第 j 种校正因子在分子中出现的频率， b_j 为第 j 种校正因子对分配系数的贡献值。

(2) 原子加和法 与片段加和法不同，原子加和法的基本加和单位为原子，而不是官能团，因此原子加和法在参数化以及计算方面都较为简单。基于原子加和的脂水分配系数的计算方法最早由 Broto 等于 1984 年提出。在 Broto 提出的方法中，把分子中的原子按照不同的化学环境分为 222 种原子类型。Broto 提出的模型没有考虑氢原子，即氢原子的贡献都包含于与之相连的重原子中。但 Broto 采用的某些原子类型在训练集中出现的频率很低，因此这些原子类型的疏水参数可能会出现过拟合的可能。

在原子加和法中，具有代表性的方法包括 ALOGP、XLOGP 以及 SLOGP 等。这些方法的基本思路相同，不同之处在于原子类型的定义。这里以侯廷军等发展的 SLOGP 方法为例说明原子加和法的基本原理。

SLOGP 基于 1850 个有机分子组成的训练集来进行模型的参数化。SLOGP 与其他原子加和法比较，最大的特色就是加和的基本单位不是某种类型原子的个数，而是某种类型原子具有的溶剂可及化表面积，如下式所示：

$$\lg P = \sum_i b_i s_i \sum_j c_j b_j \quad (6-36)$$

式中， i 为原子类型的序号， s_i 为第 i 种原子类型具有的溶剂可及化表面积， b_i 为第 i 种原子类型对分配系数的贡献值， j 为校正因子的序号， b_j 为第 j 种校正因子在分子中出现的频率， c_j 为第 j 种校正因子对分配系数的贡献值。

SLOGP 共定义了 100 种原子类型。原子类型的定义基于下面四个原则：原子的元素类型、原子的杂化形式、连接原子的类型以及芳香原子所处的芳香环境。在原子加和法中，原子类型的定义是原子加和法的关键，原子类型并不是越多越好，关键是能否充分反映不同的化学环境。

为了考察分子间和分子内的相互作用对脂水分配系数的影响，SLOGP 采用了两种校正因子，即疏水碳（hydrophobic carbon）和分子内氢键（intramolecular hydrogen）。疏水碳原子用于考察分子间由于疏水相互作用和范德华相互作用引起

的聚集。与疏水碳原子三重键相连的所有原子都必须是碳原子。

采用 SLOGP 方法预测了 138 个化合物组成的测试集的脂水分配系数，实验值和预测值能够很好地吻合 ($r=0.974$, $SD=0.327$)。此外，采用同样的测试集还比较了 15 种脂水分配系数预测方法的预测能力，结果表明 SLOGP 能得到比其他方法更为精确的结果。

6.6.1.2 脑血分配系数

大约在 100 年前科学家就已经发现，当给动物注入活性染料时，全身组织都会染上色，唯独脑组织不被染色。但如果把染料直接注入蛛网膜下腔，则脑组织迅速被染色。此后，大量的科学研究表明：有些物质完全不能由血液进入脑组织间液；有些物质进入很缓慢；而有些物质进入则较为迅速。总之，血液和脑组织之间也存在着类似的屏障，可限制物质在血液和脑组织之间的自由交换，称为血脑屏障 (BBB)。血脑屏障的功能在于保证脑的内部环境保持高度的稳定性，以利于中枢神经的活动，同时阻止异物的侵入。

脑部的毛细血管和其他部位的毛细血管在结构上存在很大的不同。第一，脑部毛细血管内皮细胞间连接比较紧密，不像其他组织毛细血管壁那样有较大的缝隙。第二，毛细血管间内皮细胞外的基底膜 (basement membrane) 是连续的。第三，毛细血管壁外表面积的 85% 被神经胶质细胞的终足包绕，因此物质由血液进入脑组织间液需要穿越较多的层次，这就更加强了脑毛细血管壁的屏障功能。此屏障对许多大分子或极性较高的药物是一个障碍，但脂溶性较高的药物仍能穿透血脑屏障而进入组织。例如极性高的季铵类 (氯化铵等) 不易进入中枢神经系统，而脂溶性高的全身麻醉药则容易进入。

图 6-11 显示了物质从血液进入脑组织间液的途径。物质透过 BBB 主要有两条途径：小分子的被动扩散 (passive diffusion) 以及载体转运 (carrier-mediated transport)。被动扩散也称简单扩散，脂溶性小分子药物可以通过被动扩散通过 BBB 入脑。由于 BBB 是一类脂质膜，脑毛细血管内皮细胞的“有效”孔径为 1.4~1.8nm。凡小分子药物其直径小于 1.8nm 者均可通过 BBB，而这些可以透过 BBB 的小分子药物在性质上至少有部分脂溶性，因此决定小分子药物是否通过 BBB 的主要因素是生理 pH 值时其脂水分配系数及分子的电荷分布情况。大多数碱性的脂溶性药物很容易进入脑内，具有这类特性的药物有强效镇痛剂、苯噻嗪类、三环抗抑郁剂以及一些抗胆碱和抗组织胺类药物。另外，许多挥发性麻醉药，如环丙烷和氧化亚氮以及静脉麻醉药硫喷妥钠等，均可非常迅速地进入脑内。 β -受体阻断剂心得安在脑内的浓度也较高，季铵类化合物和胍乙啶在生理 pH 值下也具有较高的解离度。而未解离药物的脂水分配系数又很低，故不易透过 BBB。被动扩散是最重要的，也是理论研究最多的一类脑血穿透机制。氨基酸、葡萄糖以及离子等需要载体才能进行转运，目前已知的载体系统有己糖载体、中型氨基酸载体、碱性氨基酸载

体以及短链单羧酸载体等。

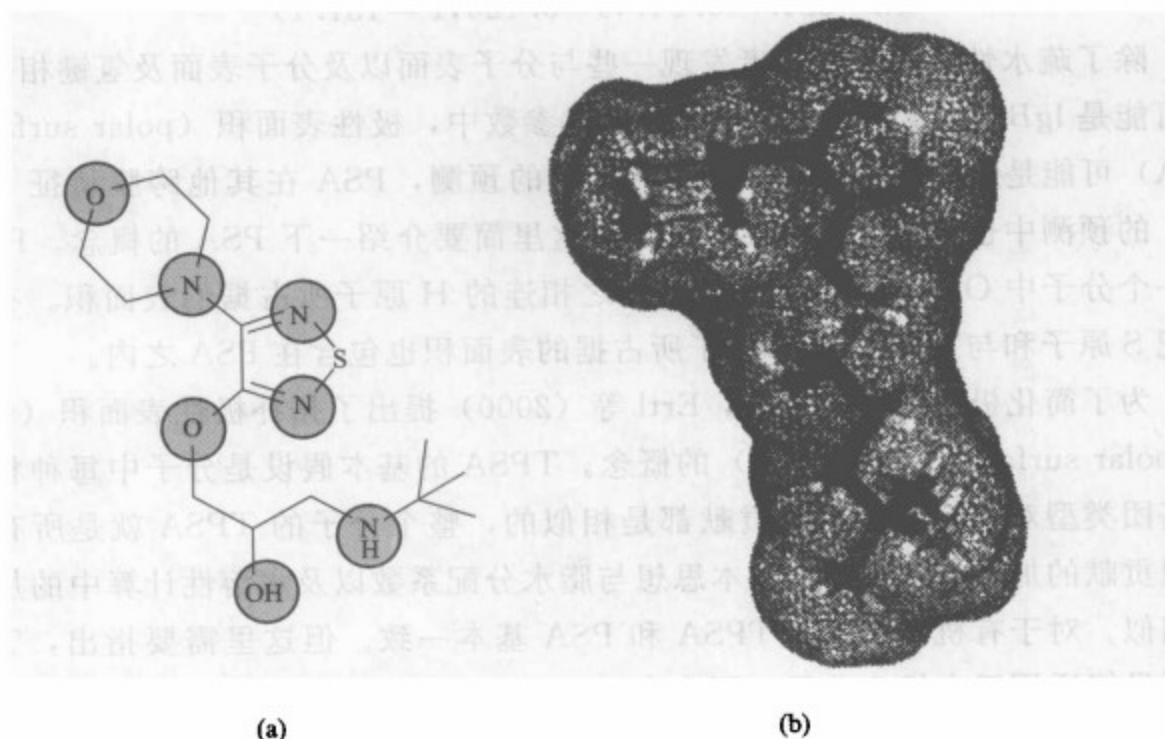


图 6-11 物质从血液进入脑组织间液的途径

实验上，一般采用脑血分配系数 BB 的自然对数来表征物质的脑血穿透能力：

$$\lg BB = \lg(c_{\text{brain}}/c_{\text{blood}}) \quad (6-37)$$

式中， c_{brain} 、 c_{blood} 分别为物质在脑组织间液和血液中的平衡浓度。

目前已经建立了多种体外和体内的 $\lg BB$ 的测定方法，但这些方法花费都很大，而且需要耗费较长的时间，同时 $\lg BB$ 的测定需要足够量的化合物纯品。目前，对 $\lg BB$ 的实验测定还不能实现高通量的过程。由此可见，发展 $\lg BB$ 预测的理论计算模型对于药物设计具有非常重要的意义。

(1) 线性回归模型 到目前为止，研究者提出了大量 $\lg BB$ 的理论预测模型，这些模型的建立大多采用了 Young 等发现的 $\lg BB$ 和 $\Delta(\lg P)$ 具有较好的线性关系。

$$\lg BB = 1.889 - 0.485 \Delta(\lg P) \quad (6-38)$$

$$(n=20, r=0.831, s=0.439, F=40.23)$$

式中， $\Delta(\lg P)$ 是 $\lg P_{\text{ow}}$ 和 $\lg P_{\text{cyclo}}$ 之间的差值， P_{ow} 、 P_{cyclo} 是辛醇-水和环己烷-水的分配系数。

但是，在大部分情况下， $\lg P$ 和 $\lg BB$ 不具有好的相关性。大量研究结果也显示仅采用与分配系数相关的参数不能得到好的 $\lg BB$ 预测模型。

Kaliszan 也建立了一个与 $\lg P$ 相关的模型，但在这个模型中引入了相对分子质量 M_m ，作为另外一个分子参数 [式(6-39)]。在回归中引入一个与分子尺寸有关的参数可以更好地考察分子的非特异性分散特性。

$$\lg BB = -0.088 + 0.272\Delta(\lg P) - 0.00112M_m \quad (6-39)$$

$$(n=33, r=0.947, s=0.126, F=131.1)$$

除了疏水性参数, 研究者发现一些与分子表面以及分子表面及氢键相关的参数也可能是 $\lg BB$ 重要的影响因子。在这些参数中, 极性表面积 (polar surface area, PSA) 可能是最重要的参数。除了 $\lg BB$ 的预测, PSA 在其他跨膜特征 (如肠吸收) 的预测中也有着非常重要的意义。这里简要介绍一下 PSA 的概念。PSA 一般指一个分子中 O 原子、N 原子以及与之相连的 H 原子所占据的表面积。有些研究者把 S 原子和与之相连的 H 原子所占据的表面积也包含在 PSA 之内。

为了简化极性表面的计算, Ertl 等 (2000) 提出了拓扑极性表面积 (topological polar surface area, TPSA) 的概念。TPSA 的基本假设是分子中每种相同原子或基团类型对极性表面积的贡献都是相似的, 整个分子的 TPSA 就是所有原子或基团贡献的加和。TPSA 的基本思想与脂水分配系数以及水溶性计算中的加和法基本相似。对于有机小分子, TPSA 和 PSA 基本一致。但这里需要指出, TPSA 的计算只能适用于有机小分子, 因为有机小分子中所有原子都暴露在分子的外部; 如果分子自身发生了折叠或分子的一部分被其他部分覆盖, TPSA 和 PSA 之间就会存在很大的差别。Ertl 在 TPSA 的参数化过程中采用了分子的范德华表面积。本书编著者采用溶剂可及化表面积重新对 TPSA 进行了参数化。

最近, 侯廷军等也采用多重线性回归的方法得到了具有较好预测能力的 $\lg BB$ 预测模型 [式(6-40)]。为了得到可靠的模型, 共采用了 115 分子, 其中 78 个分子构成训练集, 而剩下的 37 个分子构成预测集。

$$\lg BB = 0.00845 + 0.197\lg P - 0.0135\text{HCPSA} - 0.0140(M_w - 360) \quad (6-40)$$

$$(n=78, r=0.876, s=364, F=81.5)$$

式(6-40)中包含 3 个参数: 脂水分配系数、高电荷极性表面积以及相对分子质量。脂水分配系数的计算采用了侯廷军等提出的 SLOGP 方法。HCPSA 和 PSA 比较, 区别在于 HCPSA 仅仅考虑了绝对净电荷大于 $0.1|e|$ 的 N、O、S 原子和与之相连的 H 原子对分子表面积的贡献。

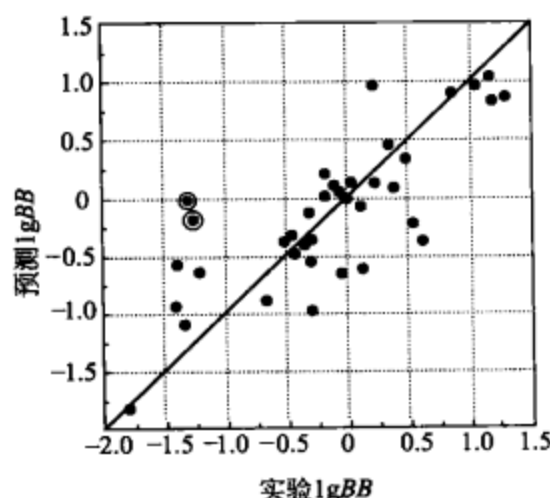


图 6-12 预测集中的分子实验 (点) 和预测 $\lg BB$ (线) 之间的线性相关关系

作者分别采用 HCPSA 和 PSA 对一系列化合物的 $\lg BB$ 进行了相关分析, 计算结果表明 HCPSA 能得到更好的回归模型。在式(6-40)中, 对于相对分子质量采用了截断模型, 其最佳截断值为 360, 这就意味着相对分子质量小于 360 时其对 $\lg BB$ 没有影响, 当相对分子质量大于 360 时其影响才开始线性增长。研究者采用这个模

型进行了 37 个分子的 $\lg BB$ 预测，取得了很好的效果。图 6-12 显示了预测集中的分子实验和预测 $\lg BB$ 之间的线性相关图。

考虑到相对分子质量和分子表面积以及分子体积一般存在一定的相关性，用体积 V 以及表面积 $SASA$ 替代相对分子质量 M_w ，可以得到下面两个模型：

$$\lg BB = -0.00740 + 0.207 \lg P - 0.0135 \text{HCPSA} - 0.0166(V - 290) \quad (6-41)$$

$$(n=78, r=0.872, s=0.370, F=78.4)$$

$$\lg BB = 0.291 + 0.138 \lg P - 0.0098 \text{HCPSA} - 0.00969(SASA - 450) \quad (6-42)$$

$$(n=78, r=0.867, s=0.377, F=74.4)$$

侯廷军等编写了 Drug-BB 程序来进行 $\lg BB$ 的预测。采用 Drug-BB 程序，分子的读入、原子类型的确认、Gasteiger 电荷的计算、TPSA 的计算以及 $\lg P$ 的计算可以程序化地进行（图 6-13）。

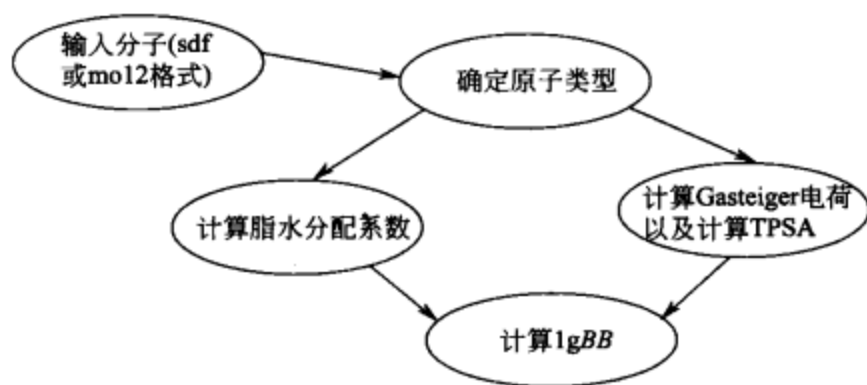


图 6-13 Drug-BB 程序的计算流程

(2) 偏最小二乘回归模型 前面介绍的这些模型都是采用了多重线性回归的方法。除了多重线性回归，偏最小二乘（PLS）也广泛应用于 $\lg BB$ 的预测中。

1999 年，Luco 采用 PLS 进行了一组包含 61 个化合物的 $\lg BB$ 预测。在 PLS 分子中，Luco 采用了 25 个与分子拓扑性质、分子结构以及分子氢键性质相关的分子参数。排除 3 个非正常点以后，包含 3 个主成分的 PLS 模型具有较好的回归和预测能力。但 Luco 提出的 PLS 模型需要采用较多的分子参数，这为其他研究者使用或评价这个模型带来了很大的不便。

表 6-2 MolSurf 参数

序号	参数描述	符号表示	序号	参数描述	符号表示
1	表面积	$\lg P$	8	氧原子氢键受体的数目	# HBA _O
2	脂水分配系数		9	氮原子氢键受体的强度	HBA _N
3	极化率		10	氮原子氢键受体的数目	# HBA _N
4	极性		11	氢键个体的强度	HBD
5	Lewis 碱	LB	12	氢键个体的数目	# HBD
6	Lewis 酸	NR	13	氢键个体和氢键受体的总数	ΣHB
7	氧原子氢键受体的强度	HBA _O	14	碳原子电子转移特性值的总和	ΣCT

1998 年, Norinder 采用 MolSurf 参数以及 PLS 方法分析了 56 个化合物的 $\lg BB$ 。计算结果显示 PLS 模型具有很好的回归能力以及预测能力。表 6-2 和图 6-14 分别显示了 MolSurf 参数和相应的计算流程。

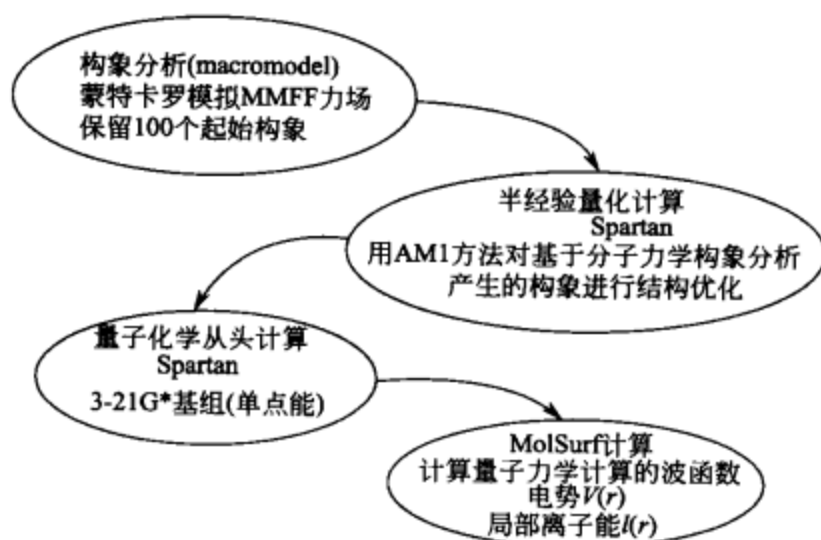


图 6-14 MolSurf 的参数化过程

Cruciani 等也提出了基于分子参数和 PLS 分析的药代动力学性质预测方法, 即 VolSurf 方法。VolSurf 首先产生一系列分子表面性质参数和其他分子参数, 即 VolSurf 参数, 然后通过 PLS 分子得到预测模型。在 VolSurf 参数计算过程中, 首先采用 GRID 程序进行探针和分子之间的相互作用能计算, 然后把分子表面积按照不同的能量区间进行分类, 不同区间的分子表面积作为不同的参数。在 GRID 计算中, 选用不同的探针类型可以得到不同的分子参数。在 VolSurf 参数计算中, 常用的探针类型包括水分子 (H_2O) 探针和疏水 (DRY) 探针。Ooms 等采用 VolSurf 参数和 PLS 方法建立了 79 个化合物的 $\lg BB$ 预测模型, 最佳的 PLS 回归模型具有较好的回归能力 ($r^2=0.76$, $q^2=0.65$)。

6.6.1.3 肠通透性

在药物研发的过程中, 一旦化合物具有了理想的药效, 接下来要考虑的问题就是该药物分子是否有好的细胞通透性 (intestinal permeability), 不利的通透性和吸收对药物的进一步研发是非常不利的。

一个药物分子从肠道通过肠壁进入血液可能存在多种机制, 其中最主要的就是由浓度梯度导致的被动扩散 (passive diffusion)。被动扩散可以发生在两个细胞之间的连接处或一个细胞内。亲脂性的分子, 如睾丸酮, 一般可以从一个细胞的一端通过细胞质到达细胞的另外一端; 亲水性的分子, 如甘露醇, 一般通过两个细胞之间的连接处从肠道到达血管。除了被动扩散, 有一些物质, 如氨基酸和葡萄糖, 其肠吸收需要一个特定传送物质。由传送物质支配的肠吸收过程是需要能量消耗的。此外, 某些分子还可能出现由 P-糖蛋白 (P-glycoprotein, P-gp) 引起的溶出过程

(efflux process)。如果某些分子是 P-gp 的底物，如长春花碱 (vinblastine)，当该分子通过被动扩散进入细胞后，会被与之结合的 P-gp 重新传送到肠道内。因此，抑制 P-gp 可能会改善某些药物分子的吸收。图 6-15 显示了药物通过肠上皮组织进入血液的多种机制。在肠吸收的理论模型中，一般都是对药物分子的被动扩散进行预测，而没有牵涉到其他的扩散过程。

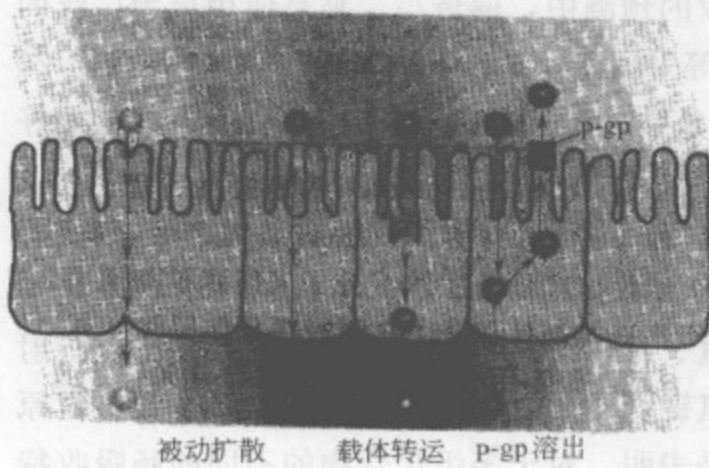


图 6-15 药物通过肠上皮组织的多种机制

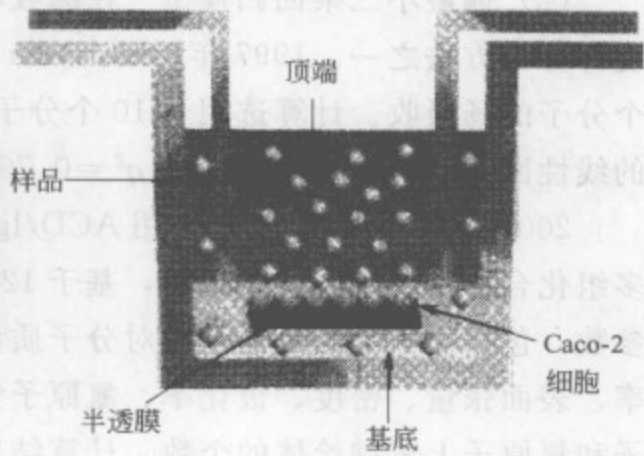


图 6-16 Caco-2 穿透检测实验的示意

药物分子的肠吸收特征一般采用 Caco-2 穿透检测 (Caco-2 permeability assay) 评价。图 6-16 显示了 Caco-2 穿透检测实验的示意图。首先把药物分子放置在半透膜上部 Caco-2 细胞层上，在某一时刻检测半透膜下部的药物分子的总数，则分子的肠通透能力采用表观穿透系数 (apparent permeability coefficient, P_{app}) 表示：

$$P_{app}(\text{cm/s}) = \frac{\text{穿透总量}}{\text{面积} \times \text{初始浓度} \times \text{时间}} \quad (6-43)$$

脑血分配以及肠通透在机制上都属于跨膜行为，因此这两种药代动力学特征的预测方法非常相似。一些重要的分子参数对于二者的预测都是非常重要的，如脂水分配系数、极性表面积等。考虑到脑血分配以及肠通透在机制上的相似性，对肠道透性的预测方法不再详细介绍，具体的讨论可以参考相应的综述性论文。

(1) 线性回归模型 与 $\lg BB$ 的预测模型一样，分子的极性表面积 (PSA) 和 P_{app} 之间表现了很好的相关性。图 6-17 显示了 13 个结构差异性较大的药物分子的空肠穿透以及 14 个化合物的 Caco-2 细胞穿透与 PSA_d 之间的分布关系，图中 PSA_d 为动态平均的极性表面积，指

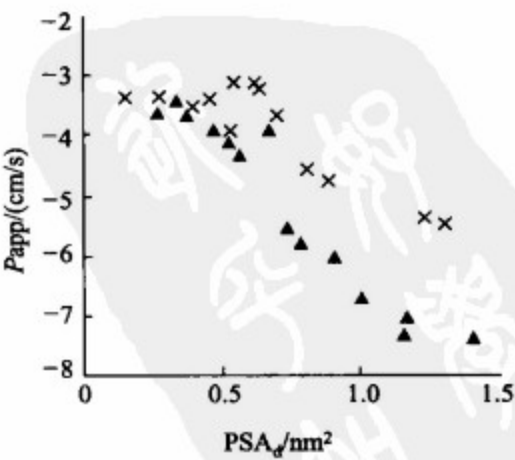


图 6-17 个结构差异性较大的药物分子的空肠穿透 (x) 以及 14 个化合物的 Caco-2 细胞穿透 (▲) 与 PSA_d 之间的分布关系

通过分子动力学或蒙特卡罗模拟后得到的相空间中所有低能构象的极性表面积统计平均值。由图上可以看到,当分子的 $PSA_d < 0.6 \text{ nm}^2$ 时,分子能够被完全吸收;当分子的 $PSA_d > 1.4 \text{ nm}^2$ 时,分子的通透性很差;而 PSA_d 在 $0.6 \sim 1.4 \text{ nm}^2$ 之间时,分子的通透性随极性表面积的增加而逐渐减小。文献多次报道了肠吸收和极性表面积之间的相关研究,这些文献报道的结果基本吻合。

(2) 偏最小二乘回归模型 在肠吸收的预测中,偏最小二乘是使用最为广泛的统计分析方法之一。1997年, Norinder 等采用 MolSurf 参数和 PLS 分析预测了 20 个分子的肠吸收。计算选用了 10 个分子参数,PLS 的计算结果表明模型具有很好的线性回归能力 ($r^2 = 0.915$, $q^2 = 0.798$, $SD = 0.493$)。

2001 年, Österberg 等采用 ACD/lgP 和 ACD/ChemSketch 参数以及 PLS 建立多组化合物的肠吸收预测模型,基于 12 个 ACD/lgP、ACD/ChemSketch 以及氢键参数,包括脂水分配系数、相对分子质量、摩尔折射、分子体积、等张比容、折射率、表面张量、密度、极化率、氮原子氢键受体个数、氧原子氢键受体个数、氮原子和氧原子上氢键给体的个数。计算结果表明,对于多组化合物的不同的肠吸收特征,PLS 模型都能很好地预测。

2000 年, Cruciani 等采用 VolSurf 参数和 PLS 分析建立了 11 个化合物的肠吸收预测模型。结果表明实验值和预测值能很好地吻合。

(3) 片段加和模型 2002 年, Klopman 等提出了一种基于片段加和的肠吸收预测模型。作者首先把分子分为不同的片段,并假定每个片段对肠吸收具有一定的贡献,然后采用 CASE 分析确定对肠吸收具有较大贡献的片段。CASE 分析显示共有 36 个结构片段对肠吸收具有较大的贡献,因此,在后续的多重线性回归中只考虑这 36 个结构片段。此外,还发现氢键给体的数目对肠吸收具有一定的贡献,因此,这个参数也包含在多重线性回归中。最后,通过多重线性回归得到这 36 个结构片段和氢键个体数目对肠吸收的贡献。从统计结果上来看,模型具有较好的线性回归能力 ($r^2 = 0.79$, $SD = 12.34$) 和实际预测能力 (测试集 $r^2 = 0.79$, 测试集 $SD = 12.34$)。

与其他的几种预测模型相比,基于片段加和的预测模型能给出更为明确的分子结构和肠吸收之间的相关信息。这种信息在分子结构的改造方面是非常有意义的。但 Klopman 采用的训练集在规模上还比较小,加上回归所用的参数较多,可能会存在过拟合的现象。

6.6.1.4 水溶性

一个药物分子要通过生物膜,必须具有足够的水溶性。如果一个药物分子的水溶性很差或溶解速度很慢,那么这个分子很难从胃肠道进入心血管系统。一个分子的水溶性一般表达为 $\lg S$, S 为分子在饱和溶液中的浓度 (mol/L)。85% 的药物分子的水溶性在 $-5 \sim -1$ 之间。当化合物的水溶性大于 -1 时,化合物的透膜能力会

较差, 因为此类化合物的极性很大, 如糖类和小肽。

水溶性的理论预测方法可以包括两类: 基于分子参数的预测模型以及基于片段加和法的预测模型。

(1) 基于分子参数的预测模型 这类预测模型是最为常见的一类水溶液预测模型。早在 1968 年, Hansch 等就观察到液态有机化合物的水溶性和脂水分配系数之间呈现很好的线性关系:

$$\lg S = -1.339 \lg P + 0.987 \quad (6-44)$$
$$(n=156, \text{rms}=0.473, r^2=0.874)$$

这个单参数模型可以较好地预测液态有机化合物的水溶性, 但对于固态有机化合物水溶性的预测不太合适。Yalkowsky 通过引入熔点对这个单参数模型进行了扩展 [式(6-45)], 使之适合于固态有机化合物水溶性的预测。

$$\lg S = 0.5 - \lg P - 0.01(t_m - 25) \quad (6-45)$$

式中, t_m 表示熔点。

虽然 $\lg P$ 可以采用各种理论模型进行预测, 但熔点的预测还没有有效的理论方法, 还需要采用实验的手段进行测定, 这就大大限制了该模型的应用。

为了发展更有应用价值的 $\lg S$ 预测方法, 研究者在线性回归中采用了各种基于分子结构的理论计算参数, 而不是与实验有关的参数。

1999 年, Abraham 等提出了包含 6 个参数的水溶性的预测模型, 如下式:

$$\lg S = 0.510 - 1.020R_2 + 0.813\pi_2^H + 2.124\sum\alpha_2^H$$
$$+ 4.187\sum\beta_2^H - 3.337\sum\alpha_2^H \times \sum\beta_2^H - 3.986V_x$$
$$(n=594, \text{rms}=0.56, r^2=0.92) \quad (6-46)$$

式中, R_2 为摩尔折射率, π_2^H 为两性极性, $\sum\alpha_2^H$ 、 $\sum\beta_2^H$ 分别为氢键的酸性和碱性, V_x 为分子体积。

2000 年, Meylan 等报道了包含 2 个分子参数和 15 个校正因子的线性回归模型。在 Meylan 等的模型中, 定义了 15 个校正因子来考察特定官能团对水溶性的影响。

$$\lg S = 0.796 - 0.854 \lg P - 0.00728M_w + \sum_i f_i \quad (6-47)$$
$$(n=3000, \text{rms}=0.90, r^2=0.84)$$

上式的拟合采用了 3000 个化合物组成的训练集。在这 3000 个化合物中, 817 个化合物的熔点已经通过实验测得。对于这 817 个化合物, 引入熔点作为第三个分子参数, 对式(6-46)和式(6-47)重新进行拟合, 得到了下式:

$$\lg S = 0.693 - 0.960 \lg P - 0.00314M_w + \sum_i f_i - 0.0092(t_m - 25) \quad (6-48)$$
$$(n=3000, \text{rms}=0.62, r^2=0.90)$$

采用熔点以后, 模型的回归能力得到了极大的改善, 这再次证明熔点在水溶性

的预测中是一个非常重要的参数。

2000 年, Jorgensen 提出了基于 6 个参数的水溶性的预测模型, 计算采用的训练集包括 150 个有机小分子。首先对训练集中的分子进行水环境下的蒙特卡罗模拟, 然后基于蒙特卡罗模拟的采样得到多个分子参数的统计平均值, 这些分子参数包括分子和溶液之间的平均非键相互作用 (ESXL)、体积、平均溶剂可及化表面积 (SASA)、氢键受体的个数 (HBDN) 以及氢键给体的数目 (HBAC) 等。基于这些参数, 得到了下面的模型:

$$\begin{aligned}\lg S = & 0.32\text{ESXL} + 0.65\text{HBAC} + 2.19 \# \text{ amine} - 1.76 \# \text{ nitro} \\ & - 162\text{HBAC} \times \text{HBDN}^{1/2} \text{SASA} + 1.18 \\ & (n=150, \text{rms}=0.72, r^2=0.88)\end{aligned}\quad (6-49)$$

式中, # amine 为非共轭胺基团数, # nitro 为硝基数。

2000 年, Huuskone 等采用电拓扑指数进行了 $\lg S$ 的预测, 预测模型中包含 24 个电拓扑指数以及 6 个拓扑参数。模型拟合共采用了 1297 个分子, 其中 884 个分子组成训练集, 剩余的 413 个分子组成预测集。当采用多重线性回归时, 得到的模型如下式所示:

$$\lg S = \sum (a_i S_i) - 1.350 \quad (6-50)$$

$$(n=884, \text{rms}=0.67, r^2=0.89, F=227.31, r_{cv}^2=0.88, \text{rms}_{cv}=0.71)$$

考虑到预测模型中采用的电拓扑指数或拓扑指数对水溶性的贡献可能不是线性的, 它们之间会存在一定的相互影响, 采用神经网络的方法进行了模型的拟合。在拟合过程中, 预测集中的分子用来控制神经网络的收敛, 当预测集中分子的预测精度不再提高时训练结束。采用神经网络时, 模型的统计质量得到了明显的提高, 训练集分子的预测标准偏差从 0.67 减小到 0.47, 预测集分子的预测标准偏差从 0.71 减小到 0.60。

2001 年, Tetko 等也采用电拓扑指数进行了水溶性的预测, 但 Tetko 等仅仅采用了电拓扑指数进行模型的拟合, 而没有采用其他分子参数。在 Tetko 提出的模型中共包含 40 个电拓扑指数。Tetko 等的计算结果也表明, 采用神经网络得到的预测模型要优于采用多重线性回归得到的模型。从模型的预测能力上看, Tetko 提出的模型和 Huuskone 等提出的模型差别不大。

(2) 基于片段加和的方法 从原理上讲, 水溶液的片段加和法和脂水分配系数中的加和法是完全相同的, 其基本假设就是整个分子的水溶性是组成这个分子的所有片段贡献的加和, 如式(6-51) 所示:

$$\lg S = \sum_{i=1}^n a_i f_i + B \quad (6-51)$$

式中, n 为所有片段的种类数, a_i 为第 i 种片段在分子中的个数, f_i 为第 i 种片段的水溶性常数, B 为常数。

在水合自由能的预测中, 采用片段加和的模型很多, 包括 Nirmalakhandan 模

型、Suzuki 模型、Kühne 模型、Lee 模型、Klopman 模型以及 drug-LOGS 模型。但其中只有 Klopman 提出的模型是真正意义上的片段加和模型，而其他方法都采用了其他的一些分子参数。Klopman 等最早在 1992 年就提出了水溶性的片段加和模型，2001 年对原有的模型进行了扩展。在扩展的 Klopman 片段加和模型中，有机分子中的常见的片段分为了 118 种基本类型。

2003 年，侯廷军等提出了一种新的基于原子加和的水溶性预测模型，即 drug-LOGS 模型。在 drug-LOGS 中，常见有机分子中的原子分为了 76 种不同的原子类型。

为了比较 drug-LOGS 模型和 Klopman 模型的预测能力，侯廷军等对 Klopman 采用的预测集中的化合物进行了预测，预测集中包含 120 个化合物，预测集中所有的化合物都没有包含在用于 drug-LOGS 参数化的训练集中。图 6-18 显示了采用 drug-LOGS 得到的 120 个化合物的预测水溶性和实验水溶性之间的线性相关图。从统计结果上看，drug-LOGS 要明显优于 Klopman 模型的预测结果。对于预测集，drug-LOGS 得到的平均绝对误差为 0.57，而 Klopman 模型得到的平均绝对误差为 0.70。

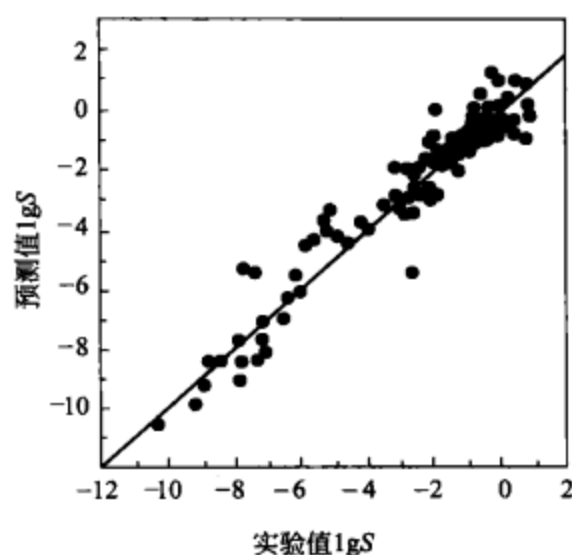


图 6-18 预测集中化合物水溶性实验值（点）和预测值（线）之间的线性相关性

从原理上看，片段加和法应该优于原子加和法，因为片段加和法可以更好地考虑片段中不同组成部分之间的相互作用。但片段

加和法中片段的划分比原子加和法中原子类型的划分复杂得多，不合理的片段划分反而不能得到合理的结果。此外，由于分子中的片段类型很多，片段加和法在定义片段时往往会遗漏某些片段，这就造成不能对包含此种片段的分子作出合理的预测，Klopman 模型就存在这种问题；但在原子加和法中，由于采用了原子作为最小单元，就完全可以避免这种问题。

6.6.2 化合物毒性预测方法

化合物毒性研究的特点是致毒机理种类繁多且与生物体代谢过程紧密关联，因此不可能使用简单、一致的理论模型予以描述。在早期，对化合物毒性的理论评估主要是通过对已知毒物的分子结构特点进行人工归纳及分类，然后根据专家的经验对未知化合物的可能毒性进行粗略估计。随着 QSAR 的普及，对同系列结构相似的化合物的毒性预测有了相对可靠和准确的方法。这一时期出现了大量针对各种系列结构化合物的各种致毒机理的 QSAR 建模研究，而且结果也相当令人满意。但

是 QSAR 的局限也非常明显，它只能适用于同系列化合物毒性的预测。然而，同一种类的有毒物分子结构上的差异可能很大。随着人工智能方法的发展，各种模式识别技术逐渐崭露头角。它们共同的特点是使用某个已知的化合物毒性值（如 LD_{50} ， LC_{50} ）训练集对模型进行训练，然后将该模型用于未知化合物的毒性评估。模式识别的技术可以有效地克服 QSAR 方法的局限，但是其代价是对毒性大小预测的准确程度远不如 QSAR。此外，使用分子对接方法进行化合物毒性预测也开始出现。

6.6.2.1 定量构效关系

定量构效关系（QSAR）是使用数学模型来描述分子结构与分子的某种生物活性之间的定量关系。其基本假设是化合物的分子结构包含了决定其物理、化学及生物等方面性质的信息，这些理化性质进一步决定了该化合物的生物活性。进而，化合物的分子结构性质数据与其生物活性也应该存在某种程度上的相关。就化合物毒性预测领域来说，QSAR 就是指使用统计建模方法寻找可以描述分子结构性质与化合物毒性数据之间相关性的最佳数学模型。由于 QSAR 方法只对同系物有效，因此在使用 QSAR 方法之前必须对化合物进行分类，以确定其结构类型以及性质描述符的取值范围，这样才能保证模型的可靠性。在毒性 QSAR 方面的文献非常多，可以参见相关的综述。

6.6.2.2 模式识别方法

在计算机辅助药物分子设计技术中，模式识别（pattern recognition）是一类应用非常广泛的技术。其背后所依据的假设很简单，即如果一系列化合物具有相似的生物活性，那么它们的分子结构之间也必定具有某些共同的特征或性质。模式识别方法就是要设法找出这些共性，进而对未知化合物进行分析并考察前述共性是否出现，如果出现则认为它具有与该系列化合物相似的生物活性。寻找共同特征的方法多种多样，但从基本思路上看，多数采取对比训练集中有活性和无活性两类化合物各种特征出现的频率差别，进而找出有活性化合物的共同特征。在这个意义上，它们都属于统计学习范畴。在化合物毒性预测中使用的模式识别方法一般分成两步：首先，需要对训练集中的分子结构进行分析和比较，以确定有活性和无活性类化合物中各种特征的统计分布，从而找出有活性化合物的共同特征；然后，根据得到的统计分布，使用合适的统计学习方法对某些特征与生物活性数据之间的关系进行建模，并使用测试集来检验模型的预测能力。

在药物分子设计中，“药效团”指的是一系列活性化合物具有的一组共同药效特征在空间中的三维排列。其中共同药效特征可以包括特定的化学基团，也可以包括特定的化学环境，如氢键的给体和受体、正电基团和负电基团、疏水基团和亲水基团等。毒效团（toxicophore）可以看作是对药效团模型的推广。从广义上讲，在分子结构中所有可以影响生物活性的一组特征及其分布都可以称作是活性团

(biophore)。如果这些特征可以产生有益的生物活性，则可称之为药效团；反之，如果这些特征会导致有毒或有害的生物活性，则可称之为毒效团。图 6-19 列出了由 DEREK 软件识别出的几个简单毒效团。

在很多情况下，化合物毒性的评估可以看作是由一系列具有层次结构的简单判定〔通常为单选的提问-解答形式(question-answer)〕组成。在每一步中只需进行二选一或多选一的判断（例如是否具有某个毒效团，或者从口服、皮肤接触式及吸入式中选择一种人体对有毒物的接触方式等），以决定下一步具体执行哪一个后继判定。整个过程可以看作是使用判定树进行推理。关于化合物毒性的相关知识存储在各个节点当中，随着判定过程的逐步深入，可能答案的数目不断减少，最终到达某个叶子节点时就得到了最终答案。这类问题求解的关键在于构造适合的判定树，即将毒性评估问题的解集及其中间步骤转化成层次结构。

片段特征概率是指分子中包含的子结构片段具有某种特征（生物活性或毒性）的概率。根据分子中所包含的每一个子结构片段特征概率，使用概率统计方法可以得出整个分子具有该类特征的概率。HarzardExpert 软件就是使用类似的方法评估化合物毒性。这类概率方法依据的基本假设是分子中包含的子结构片段的活性特征概率具有某种加和性。也就是说，如果某个化合物具有两个特征且它们对分子毒性贡献的概率相同，那么该化合物的毒性将两倍于仅具有一个特征的化合物。这类方法的优点是计算比较简单。不过，由于绝大多数决定化合物是否表现某种毒性的因素并不是完全独立、随机的，因此，使用概率来描述子结构片段可能具有的毒性不是完全可靠，就更不用说整个分子所具毒性的概率。

传统 QSAR 方法只适用于结构非常相似的同系物，每一种不同类型的分子结构都要求不同的 QSAR 模型。因此，使用传统 QSAR 方法进行毒性预测之前，必须首先按结构类型将化合物进行划分，以确定合适的 QSAR 模型。如果有两个以上的模型适用于该化合物，那么还需要某种仲裁手段来确定最佳模型，而实际使用的仲裁方法往往具有相当大的随意性。使用人工神经网络 (artificial neural network) 方法在一定程度上可以克服这一局限。

6.6.2.3 专家系统

与其他生物活性的建模工作相比，化合物毒性研究要复杂得多，因为不但化合物毒性种类、致毒机理种类繁多，而且其分子结构类型分布也很广泛。传统的 QSAR 方法在使用之前必须对化合物的类型进行划分以确定使用何种 QSAR 模型

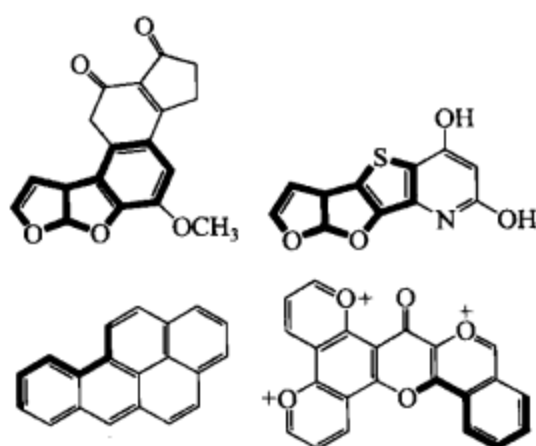


图 6-19 DEREK 软件识别的毒效团（黑体部分）示意

进行计算，否则结果的准确性非常差。使用神经网络及其他统计建模方法可以在一定程度上克服分子结构类型上的限制，但这是以预测的准确程度和模型的明晰性为代价。同时，上面两类方法得到的模型都需要与单一、确定的致毒机理相对应。然而，实际中对于未知分子的致毒机理往往不太清楚，甚至毫无头绪。因此，专家的知识与经验在毒性评估工作中发挥了非常重要的作用。

“任何能够帮助使用者对化合物毒性进行合理预测的、形式化的系统”都可以被认为是毒性预测领域的专家系统 (expert system)。根据这个定义专家系统不一定必须是基于计算机的，不过目前绝大多数成功的系统都使用计算机进行处理。目前已知的各类专家系统根据其设计理念可以大体分成两类 (图 6-20)：基于规则 (rule-based) 的系统以及基于统计 (statistical-based) 的系统。

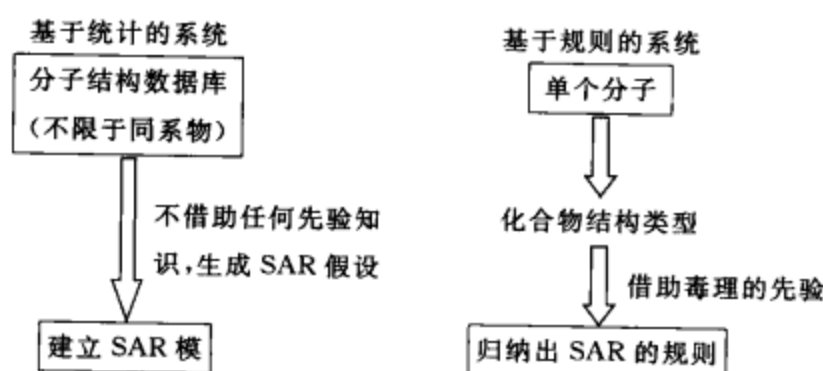


图 6-20 两类专家系统设计思路的示意

专家系统的底层一般包含 3 部分：知识库 (rule base)、推理机 (inference engine) 以及新知识获取。知识库用于存储所有已知的知识。推理机的作用是对已知的知识进行逻辑推理，以找到最终答案。最后，新知识获取部分用于将新知识方便地添加进知识库。基于规则的专家系统的设计思路是以单个分子为起点的，首先借助先验知识、专家判断或其他启发式策略对单个化合物进行分析，以总结出各种分子结构类型，进而产生具体规则，然后根据得到的规则运用逻辑推理手段对分子可能的毒性进行预测。在整个系统设计中，将专家的知识转化为规则的过程不但是最重要的一步，也是最困难的一步。如果规则本身没有正确地表达专家知识，那么再有效的推理算法也无济于事。对于未知的化合物，人类专家最常使用的办法是寻找与已知毒性分子之间的相似性。不过由于相似性定义本身具有很大的主观随意性，因此，专家们一般倾向于通过直接观察在分子结构中是否存在特定的活性基团来确定该分子是否具有某种毒性。如何把专家脑子中隐含而抽象的知识转化为明确的具体化的规则，现在还没有什么行之有效的办法。相对比较可靠的做法之一是与专家进行面谈 (interview)。Jelovsek 等在 1990 年发表的文献中给出了一个具体的例子。

理论上，如果具有可以覆盖问题空间的完备规则集，那么对该空间内的所有问

题都应该可以找到答案。推理机的任务就是对规则进行检查,结合特定的逻辑推理方法和编程技巧搜寻所有规则之间可能的联系。本质上讲,这实际上是某种形式的图遍历,因为问题空间内所有规则之间的关系可以用图来表述。如果将各规则之间所有可能的联系看作是一个集合,那么当节点数目较大时集合的规模将异常地大。因此,必须选择适合的搜索策略才能使问题在可接受的时间内得到合理的结果。

目前比较成功的基于规则的专家系统包括 HazardExpert、DEREK 以及 ON-COLOGIC 等。这些系统的应用都限于已有的知识和特定的毒性机制模型。它们的主要差别表现在建模所用的毒性终点、模型建立和预测时使用的信息、存储及表示通用化学结构的方法、反馈的信息、用户在添加新知识过程的参与程度等。

与基于规则的专家系统不同,基于统计的专家系统将已知毒性的化合物数据库看作一个整体,通过使用各类统计分析方法(多重线性回归、主成分分析、偏最小二乘以及判别分析等)或其他自动化算法对已知数据进行分析并从中抽象出合理的构效关系(SAR)模型,然后根据得到的 SAR 模型预测未知分子的可能毒性。这类系统的优点是几乎不依赖任何与机理或结构类型有关的先验知识。基于统计的专家系统有 TOPKAT、CASE/Multi CASE 等。

6.6.2.4 常见化合物毒性预测软件

(1) COMPACT COMPACT (Computer-Optimized Molecular Parametric Analysis of Chemical Toxicity, 化学毒性计算机优化分子参数分析)是由英国 Surrey 大学的 Lewis 等研究开发的化合物毒性预测系统。COMPACT 软件主要针对与 P450 酶家族有关(直接或间接)的受体参与毒性预测。

Lewis 等首先假定分子中的某些结构特点可以决定其与受体活性位点进行结合的能力,然后根据 P450 酶或其诱发因子的已知数据建立能够与受体进行特异性结合的结构特征模型,最后通过判断分子结构与结构特征模型之间的符合程度来判断分子是否具有潜在的毒性。COMPACT 软件的优点是结合了有关反应机理方面的信息,这些机理信息受分子建模研究支持,因此理论基础相对牢固。此外,COMPACT 的计算是基于整个分子而不是片段,因而不但简单易行,而且可以处理大到 150 个重原子的体系。COMPACT 的局限在于它只针对 P450 酶家族参与的代谢过程中有关的毒性,因而不能作为独立的化合物毒性预测系统。

(2) DEREK DEREK (Deductive Estimation of Risk from Existing Knowledge, 从已有知识推理风险评估)软件是一种基于规则的专家系统。该软件的作用原理是使用已知的毒性构效关系(SAR)模型对未知化合物的毒性进行定性评估。这些已知的知识(指毒性有关的构效关系模型)主要从专家经验中抽象而来,并经过协作专家组的鉴定。参与鉴定的专家来自商业、教育、非盈利组织等不同领域,因而得到的知识可靠性较高。DEREK 可以对相当广泛的化合物毒性类型进行评估,但该软件对基因突变、致癌以及致敏效应等毒性预测的能力相对突出。

系统的知识库中不仅保存毒效团的详细信息，还包括一些能够增强或削弱毒性效应的结构片段特征的信息。因此，DEREK 不但可以对结构中出现的毒效团进行识别，而且在一定程度上还可以考察它们之间的相互影响。在较新的版本（4.01 版本及以上）中，为提高对致敏毒性预测的准确程度，软件引入了评价分子皮肤渗透性的指标（主要依据化合物的 $\lg P$ 值）。这些数值信息与一般的非数值性规则通过 DEREK 专有的推理引擎结合在一起，共同对输入事件做出响应。

尽管 DEREK 自身并不包含有关代谢途径评估的模块，但是作为一个基于规则的专家系统，DEREK 可以方便地通过外部程序调用或加入新规则等方式将这些功能集成进知识库。不过，这些规则将限于那些机理非常清楚的化合物结构类。但是，在最新的版本中，DEREK 可以通过与 METEOR 的联用来解决有关化合物代谢激活/去活途径的问题。为了提高系统的灵活性，DEREK 包含图形界面的规则编辑器，用户可以方便地通过规则编辑器将新规则引入知识库。系统另一个非常有用的特点是输出结果中包含结论推理过程的简要信息，即系统依据何种规则得出最终结论。这也是所有基于规则的专家系统的共同特征之一。此外，DEREK 的输出结果还可能包括样例化合物的毒性实验数据以及相关文献信息。DEREK 软件的局限在于它不能给出化合物毒性的定量数据。

(3) HazardExpert HazardExpert 是 CompuDrug 公司开发的专门用于化合物毒性预测的基于规则的专家系统软件。它使用毒效团模型对未知化合物进行毒性预测。毒效团模型主要通过分析已知的 QSAR 模型、美国 EPA (Environment Protection Agency, 环境保护局) 以及 ITC (Interagency Testing Committee, 机构间检验委员会) 发表的专题文献得到。HazardExpert 将目标化合物与知识库中已知的毒效团模型进行比较。一旦目标结构中的毒效团被识别出来，系统将根据已有的规则对化合物可能的毒性进行估计，如致癌、致畸、神经毒性、皮肤过敏、眼刺激等。

与其他已知的专家系统相比，HazardExpert 具有一个与众不同的特点，即在用户输入目标化合物时不但要求输入具体的结构信息，而且要求指定物种、剂量、暴露途径、持续时间等参数。通过各类性质描述符，如 $\lg P$ 、 pK_a 的计算，配合适当的 QSAR 模型，系统可以模拟分子在生物体内的生物利用度和生物性聚集效应，因而能够比较全面地评估目标化合物在生物体内的毒性作用。

根据系统评估的结果，HazardExpert 给出目标化合物在几种不同毒性类型下的分类数据，以及 IARC (International Agency for Research on Cancer, 国际癌症研究机构) 人体致癌风险的百分数。此外，用户还可以使用关键词在系统收集的期刊 “Quantitative Structure-Activity Relationship” (定量构效关系) 文摘库中查询支持毒性预测的证据的有关文献。另外，与 DEREK 类似，HazardExpert 也有一个相关软件 MetabolExpert 用于解决代谢中间体的问题。

(4) CASE/Multi-CASE CASE (Computer-Automated Structure Evaluation, 计算机自动化结构评估) 是由 MULTICASE 公司出品的软件系统。该软件的特点是不依赖任何先验知识, 而是借助统计分析手段由非同系物的化合物训练集中自动产生出预测模型。从这个意义上讲, CASE 应属于基于统计的专家系统。CASE 的计算结果包括使用自定义单位 (CASE unit) 标度的活性/毒性值及其置信度数据。例如, 计算值在 10~19 之间属于无活性, 20~29 之间属于弱活性, 30~99 之间属于有活性。

CASE 软件的功能并不单独针对化合物毒性预测。相反, 它主要用于对给定化合物的生物活性进行评估。其原理是首先对给定的化合物训练集生成所有可能的结构片段 (重原子数在 4~10 之间), 然后使用统计分析手段将这些片段分成有活性 (biophores) 和无活性 (biophobes) 两类。其中有活性的片段, 包括药效团或毒效团, 是指在统计意义上与活性相关的片段。反之, 无活性的片段是指在统计意义上与活性无关的片段。如果有具体的活性/毒性数据, 那么 CASE 将使用线性回归方法对活性/毒性的值进行预测。此外, CASE 还可以为用户指定的训练集产生预测模型。Multi-CASE 是 MULTICASE 公司的另一套软件, 与 CASE 相比, 它不但可以识别对活性片段具有修饰 (增强或削弱) 作用的结构片段, 而且可以给出一些可用于 QSAR 的计算描述符。

由于完全依赖统计分析方法产生预测模型, 因此, 与其他统计方法一样, CASE 不但要求训练集中的化合物结构具有很好的多样性以覆盖尽可能多的化学空间, 而且从多个来源得到的活性/毒性数据的划分必须保持一致。也就是说, CASE 很容易受训练集结构代表性差或活性数据划分错误的影响。

(5) TOPKAT TOPKAT (Toxicity Prediction by Komputer-Assisted Technology, 计算机辅助毒性预测技术) 是由 Health Design 公司设计开发的一套面向 PC 的专家系统软件。该软件目前由 Accelrys 公司负责开发和销售, 可能是目前应用最为广泛的毒性理论预测软件系统。TOPKAT 可以预测各种急性毒性和慢性毒性终点, 包括啮齿动物致癌、Ames 测试、皮肤过敏以及小鼠口服 LD₅₀ 值等十多个类型, 而且每个毒性类型都作为模块独立销售。

TOPKAT 使用模式识别和拓扑分析的方法对用户指定的化合物进行毒性预测。软件首先对输入的结构进行分析, 以确保目标分子被指定模块的所谓“最优预测空间” (optimum prediction space, OPS) 覆盖, 然后识别所有特征子结构并计算或查找模型所需的各类子结构性质参数, 将之代入预测模型得到预测结果。最终结果不但包括定量的毒性终点数据, 而且给出统计显著性等可靠性信息。

其毒性预测模型采用标准的统计回归方法。模型的自变量选择基于拓扑的描述符, 系统开发者认为使用这些描述符能够与传统 QSAR 方法中常用的分子轨道类描述符具有相近的性能。模型建立的方法可以分成两大类: 对于一般连续型的毒性

数据（如 LD₅₀ 值），系统采用多重回归方法建立预测模型；对于那些以定性分类（如阳性/阴性）形式给出的实验数据，系统采用线性判别分析函数建立预测模型，且结果以判别概率的形式给出。

此外，TOPKAT 独有的特点是自带若干已知化合物的毒性数据库。用户可以通过相似性搜索在数据库中查找与目标分子类似的结构来辅助对预测结果可靠性的判断。不过 TOPKAT 使用的相似性定义是基于前面提到的拓扑描述符而不是更直观的结构描述符，因此在实际使用过程中其价值可能会打折扣，因为经常搜索到结构不同但描述符数值相近的分子，用户通过观察这些命中结构并不能直接增强对目标分子是否有毒的信心。

TOPKAT 软件的局限主要表现在两个方面。首先，TOPKAT 的模式识别技术只假定分子结构中的子结构各自独立地对整体毒性产生贡献，而没有考虑各结构片段在空间的分布及其相互作用。事实上，这一假设并不总能成立，因而导致预测结果可靠性的降低。甚至有人认为该系统只对同系物有效。这只能靠借助真正的毒效团模型来解决。其次，使用判别概率描述定性毒性数据具有某种误导性。一些用户经常会误将此概率直接作为毒性大小排序的指标，而此处概率仅仅是判别概率，并不直接与毒性大小相关。

(6) 几种专家系统的性能比较 前面介绍的几种专家系统都各具特色，在此仅用于对比。然而究竟哪一种系统性能最好当然是研究人员最关心的问题。一项对各实验室 Salmonella Assay（沙门菌试验，一种诱变性能测试方法）结果重复性之间的对比研究表明，该实验结果的正确率在 85% 左右，也就是说对该类型毒性数据的统计建模工作来说预测准确度超过 85% 是没有意义的。美国 NTP（National Toxicity Program，国家毒理学规划处）在 20 世纪 90 年代初对当时几个著名专家系统的性能进行比较研究。NTP 选择 44 个化合物，首先要求包括 DEREK、TOPKAT、Multi-CASE 以及 COMPACT 在内的几个专家系统对给化合物集合进行毒性预测，并将结果公开发表，当实验数据基本完成之后将各专家系统预测的结果与生物鉴定实验的结果相对比。对于新化合物的预测检验，令人惊讶的是人类专家（Ashby 和 Tennant）预测的准确率要远高于任何一个测试的专家系统。对于这次测试的结果众说纷纭，见仁见智。但有一点可以肯定，参与测试的几个专家系统在对未知结构类型的化合物毒性预测方面还存在很大局限。在这之后，还进行了多次的性能对比研究，但结果差异不大。

6.6.2.5 小结和展望

目前化合物毒性预测的计算机辅助方法已经有了长足的进步，特别是各类毒性预测的专家系统更是受到学术界、政府以及企业的广泛关注。由于化合物毒性在毒性类型、致毒机理以及受代谢途径影响等诸方面的多样性且不可能有统一的模型进行描述，因此使用专家系统对化合物毒性进行预测就显得非常必要。相比之下，

QSAR 虽然能给出较准确的定量数据,但由于方法自身的局限以及缺乏对致毒机理和代谢过程影响的考虑,因此其使用范围非常有限。

不过,与其他专业领域中比较成熟的专家系统相比,化合物毒性预测的专家系统的使用必须小心谨慎,尤其需要该领域专家的指导。其中的问题主要集中在两个方面。首先,在开发专家系统的过程中缺乏准确、可靠的实验数据。实际中,常常出现在同种或不同种毒性化验数据之间不一致甚至矛盾的现象。对任何一个专家系统,高质量实验数据都至关重要,基于规则的专家系统尤为如此。因此,在专家系统的开发过程中对实验数据的选取要求依据严格一致的标准,而这有赖于与本领域的专家紧密合作,开发的进程中专家介入得越早越好。其次,基于统计的专家系统只适合于与生物体的相互作用可以用单一模型描述的化合物系列,这是由于其模式识别和统计方法原理自身的局限造成的。因此,使用这类专家系统必须保证目标化合物在训练集化合物结构性特征空间所覆盖的范围以内。但随着高质量实验数据的不断累积以及各种智能化统计方法的引入,毒性的理论预测模型的精度会大大提高。

6.7 中草药有效成分化合物的化学空间分析

6.7.1 介绍

高通量筛选 (HTS) 和虚拟筛选是药物设计领域中的重要方法,其基本问题是搜索巨大的化学空间,从众多化合物中选择出合适的化合物组成一个子集数据库,对于高通量筛选和虚拟筛选都至关重要。目前,数据库以及数据库系统在药物设计领域中具有广泛的应用,成为一个存储、检索、管理化学信息以及研究分子结构、活性,计算受体与底物之间相互作用等的平台。中药具有数千年历史,具有很好的临床效果,如何有效地利用中药信息来指导药物开发以及中药治疗的机理研究是一个关键问题。针对这些问题,徐筱杰课题组在 2002 年开发了一个中草药有效成分三维结构与性质数据库 (Chinese Herbal Drugs Database, CHDD),目前已经发展了第二版。CHDD 主要收集、整理从中草药中提取的天然化合物的三维结构信息,目前收集了化合物记录 10564 条,有三维结构的 9820 条,来源植物记录 2037 条,涉及文献 10814 篇。相似的化合物具有相似的属性,这一观点是定量构效关系以及分子相似性分析的基础,而根据这一观点推广出来化学空间这一概念。

化学空间是通过每一个分子计算一系列分子描述符,并用这些数值作为一个个点构成的多维空间来定义的。一个化合物数据库相当于一个分子的子集,希望所构建的子集能代表巨大的化学空间的多样性,以确保该子集相关的子空间有确定先导化合物的最大机会。因此,研究中药的治病机制,研究中药分子的多样性很重要,尤其是寻找这些分子同类药分子的相似和差别。

在这里, 为了说明 CHDD 具有足够的多样性并与药物分子库具有足够的多样性, 对 CHDD 包含的分子进行了多样性分析, 并与药物分子数据库的数据进行比较。

6.7.2 分子描述符的计算

分子描述符是一些表征分子属性的数值, 它们有的反映分子的物理化学性质, 有的是通过对分子结构应用算法推导出来的数值, 许多不同的描述符已经被描述并用于广泛的用途, 而根据数据库的分子描述符数值点构成的化学空间可以用来描述数据库的多样性。

用于多样性分析的分子描述符主要包括分子结构指纹、二维描述符或三维描述符、分子的表面性质、拓扑描述符等。在这里, 针对 CHDD 考虑的描述符包括结构描述符 [分子量、手性中心 (chiral center) 数目、可旋转键 (rotatable bond) 数目、氢键受体 (H bond acceptor) 数目、氢键给体 (Hbond donor) 数目], 空间描述符 [V_m (molecular volume, 分子接触体积)、Area (molecular surface area, 分子表面面积)、TPSA (total polar surface area, 总极性表面面积)、THSA (total hydrophobic surface area, 总疏水表面面积)、 r_{gyr} (radius of gyration, 分子回旋半径)], 拓扑描述符 [柔性 (flexibility) 指标、Wiener 指标、Zagreb 指标、Hosoya 指标、CHI (12 个)、SC (6 个)、Kappa 指标 (6 个)、Balaban 指标 (J_x)], 脂水分配系数 AlogP98 等 40 个描述符。全部计算在 Cerius2 的 QSAR+ 模块中进行。

6.7.3 结果与讨论

对计算出的分子描述符的数值进行了分析, 一些重要的描述符的最大值、最小值、平均值和中位值列在表 6-3 中。此外, 还对一些重要的分子描述符作了它们的分布柱状图。

表 6-3 CHDD 的一些通用分子描述符的最大值、最小值、平均值和中位值

分子描述符	最大值	最小值	平均值	中位值
分子量	2812	48.1	449.6	391.0
手性中心数目	57	0	6.2	5
可旋转键数目	74	0	8.9	7
柔性指标	59.3	0	6.9	5.6
氢键受体数目	78	0	7.5	6
氢键给体数目	43	0	3.5	2
Area	2549	73.7	518.9	452.2
TPSA	1964.5	0	238.4	181.4
Kappa-1	164.6	2	25.2	22.0
Kappa-2	68.9	0	10.3	8.5
Kappa-3	60.0	0	5.4	4.2
V_m	2156	47.4	418.4	361.7
AlogP98	26.6	-13.0	2.6	2.5

从表 6-3 中的数据可以看出, CHDD 所含的分子的分子量的平均值与中位值为 449.6/391.0。Feher 和 Schmidt 对来源于 Chapman and Hall 的“Dictionary of Drugs”(药物词典)的 10968 个药物分子以及来自 BioSPECS 天然产物数据库、ChemDiv 天然产物数据库、Interbioscreen IBS2001N 和 HTS-NC 数据库的 3287 个天然产物分子、27338 个半天然产物分子进行了分析, 这些药物分子的平均值/中位值为 340/312, 天然产物分子的平均值/中位值为 414/362, 半天然产物分子的平均值/中位值为 381/134。CHDD 的分子量的平均值/中位值和天然产物的更为接近, 而与药物分子的差别最大, 这同 CHDD 属于天然产物的一个特殊子集有关。

再看看分子量分布图(图 6-21), CHDD 的分子量分布有一个好的峰型分布, 它的最高峰在区间 [300,400] 上, 而在 Feher 和 Schmidt 的研究中天然产物分子与半天然产物分子的最高峰在区间 [200,300] 上, 均有所差别, 但与 Feher 和 Schmidt 的研究中的天然产物分子分布相似的是在最高峰以后渐次下降, 而在分子量大于 1000 处还有一个比较高的峰, 这也是造成中草药分子和天然产物分子的分子量平均值相对药物分子较大的原因。

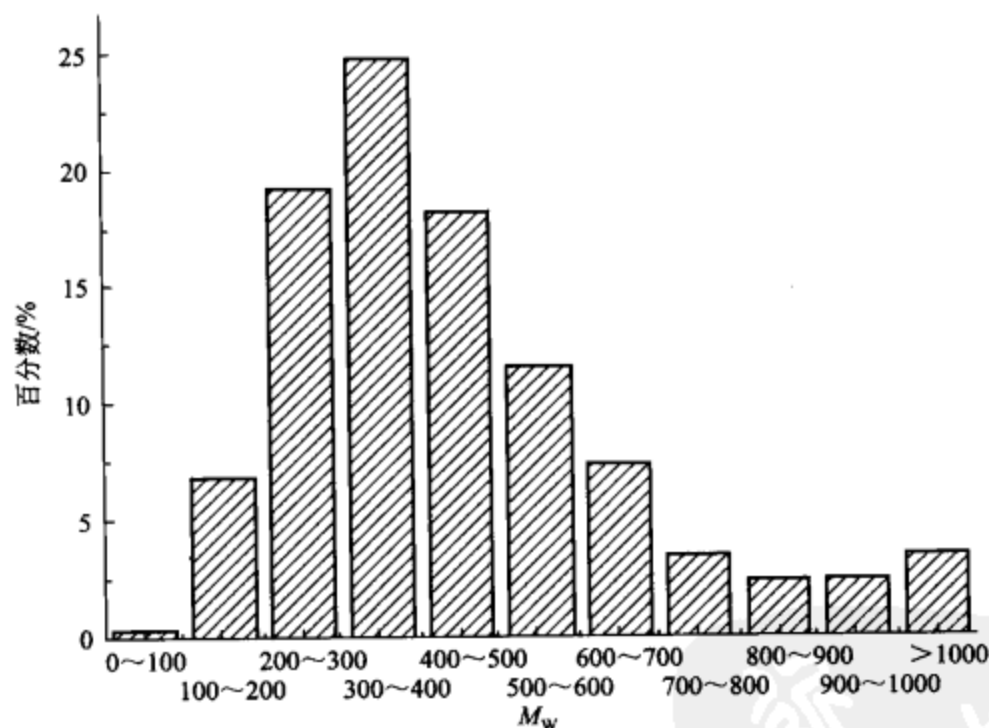


图 6-21 CHDD 的分子量分布

根据 Lipinski 的“五规则”, 当分子的分子量大于 500 时, 将影响分子的口服吸收性能, 然而天然产物中分子的分子量大于 500 的占了很大的比例, 这是因为许多天然产物分子含有多糖环, 这直接导致分子的分子量增大。当然, 多糖环很有可能被水解或代谢形成二次代谢产物, 其中的机制比较复杂, 这也直接影响了对药用天然产物的研究。而中药是药用天然产物的子集, 因此也有同样的困难。而由于 CHDD 的最高峰处于分子量更大的区间, 也造成了 CHDD 分子的分子量平均值比天然产物更大。

再来看看分子的手性中心数目，CHDD 中分子的平均值/中位值为 6.2/5。而 Feher 和 Schmidt 的工作给出的天然产物分子的平均值/中位值为 6.2/4，半天然产物分子的平均值/中位值为 2.2/1，药物分子的平均值/中位值为 2.3/1。CHDD 的平均值与 Feher 和 Schmidt 得出的天然产物的平均值一样，它们的分布也有相似之处（图 6-22）。而由天然产物改造出来的半天然产物的手性中心数目的平均值和中位值均有较大的降低，与药物分子更为接近，这是因为对天然产物的改造导致了手性中心的缺失，药物分子来源于天然产物、半天然产物、合成分子，因此它的手性中心与半天然产物接近。而组合化学库分子的手性中心数目则更低，导致组合化学库分子的手性中心很少的可能的原因是手性分离的代价高。这也表明了天然产物分子有着多样性更高的骨架，对天然产物以及中药分子的利用是有前景的。

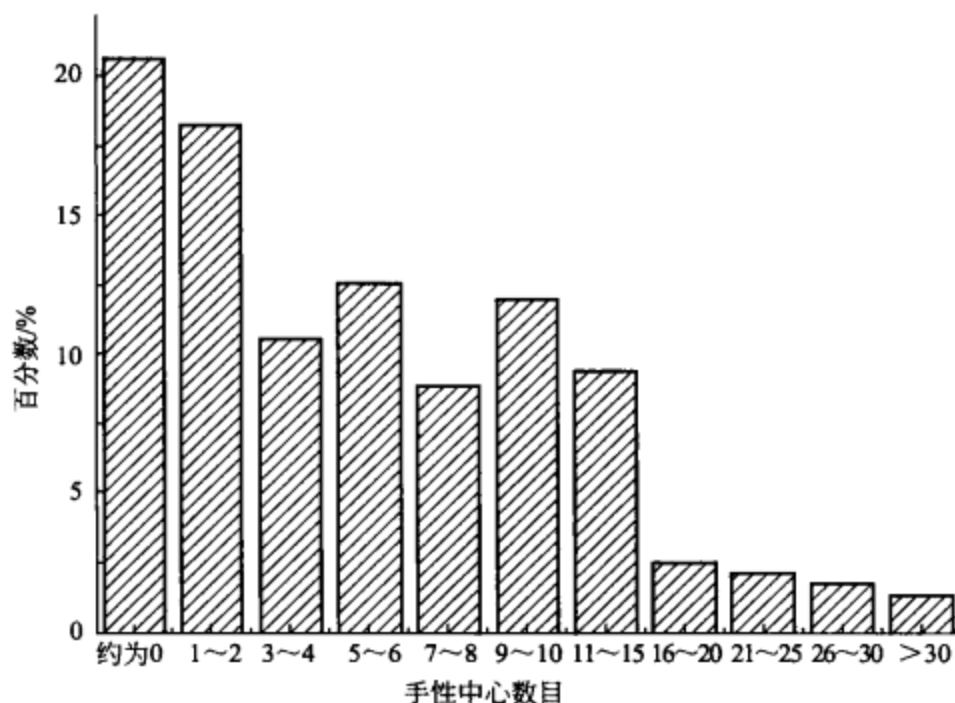


图 6-22 CHDD 的手性中心数目分布

CHDD 的分子的可旋转键数目的平均值和中位值分别是 8.9 和 7，它有更广的分布，可旋转键数目为 0 的分子不到 4%，这与 Feher 和 Schmidt 的工作给出的 0 个可旋转键分子占接近 30% 有很大的差别，而在可旋转键数目大于 10 的比例中 CHDD 中也有更多分子（图 6-23）。可旋转键数目表示分子的柔性性质，天然产物分子以及 CHDD 从分布图（图 6-24，图 6-25）中可以看到，CHDD 的分子能形成较多的氢键受体和给体，氢键受体数目分布的最高峰在区间 3~4 上，区间 5~6 上的分布峰次高，与最高峰很接近，它们各自都有超过 20% 的分子，这与 Feher 和 Schmidt 的工作给出的结果相似。不过 Feher 和 Schmidt 的工作给出的结果中，区间 3~4 上的分子所占比例约 30%，区间 5~6 上的分子比例近 25%，而 CHDD 有更多的分子氢键受体数目大于 6，这也表现在平均值/中位值上，Feher 和 Schmidt 给出的天然产物的是 6.8/5，稍小于 CHDD 的 7.5/6。图 6-25 显示的 CHDD 的氢键给体数目分布图与 Feher 和

Schmidt 的工作给出的天然产物分子氢键给体数目分布图的形状比较相似，氢键给体数目为 0、1、2 的分子均占 20% 左右，然后分布峰渐次下降。不过 CHDD 的平均值/中位值为 3.5/2，天然产物的为 2.6/2，有一些差别，这主要是由于 CHDD 在氢键给体 0 处只有约 18% 的分子，而天然产物分子的分布中氢键给体数目为 0 的有近 25% 造成的。虽然氢键给体受体的数目越多与受体结合的强度也越大，但根据 Lipinski 的“五规则”，数目过多，对药物的吸收未必是有利的；此外，在虚拟筛选当中，化合物的氢键受体给体的数目过多，也容易引起假阳性的问题。

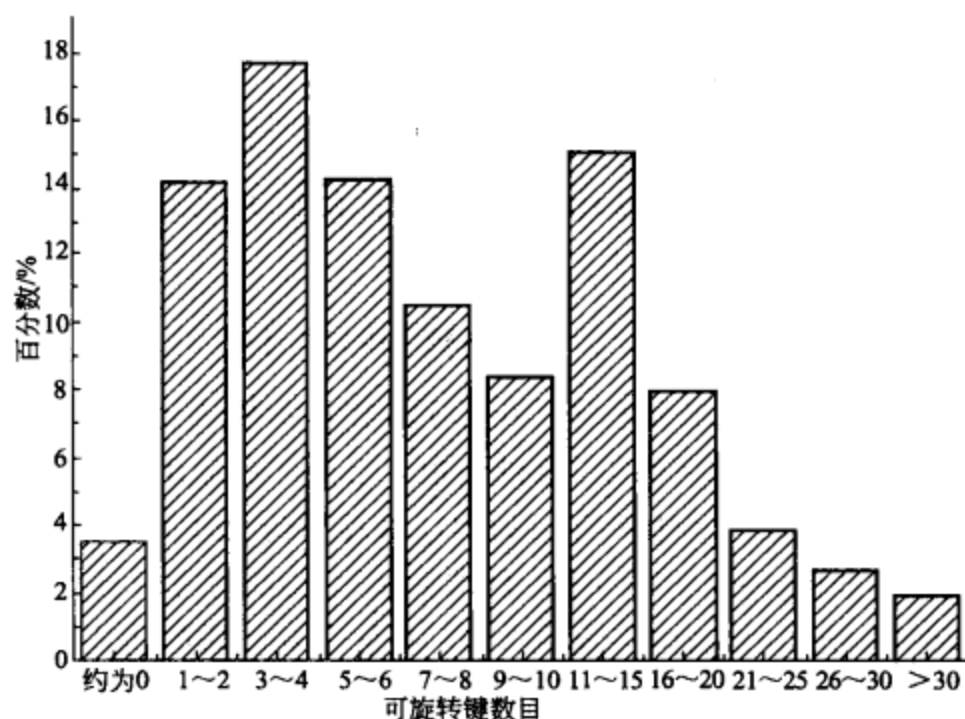


图 6-23 CHDD 的可旋转键数目分布

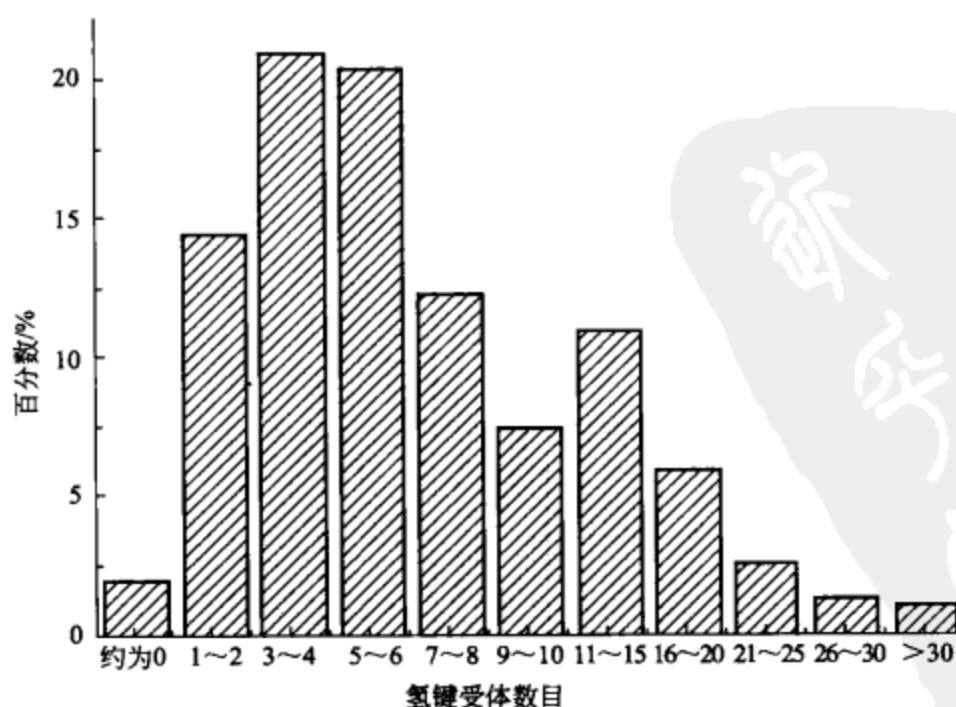


图 6-24 CHDD 的氢键受体数目分布

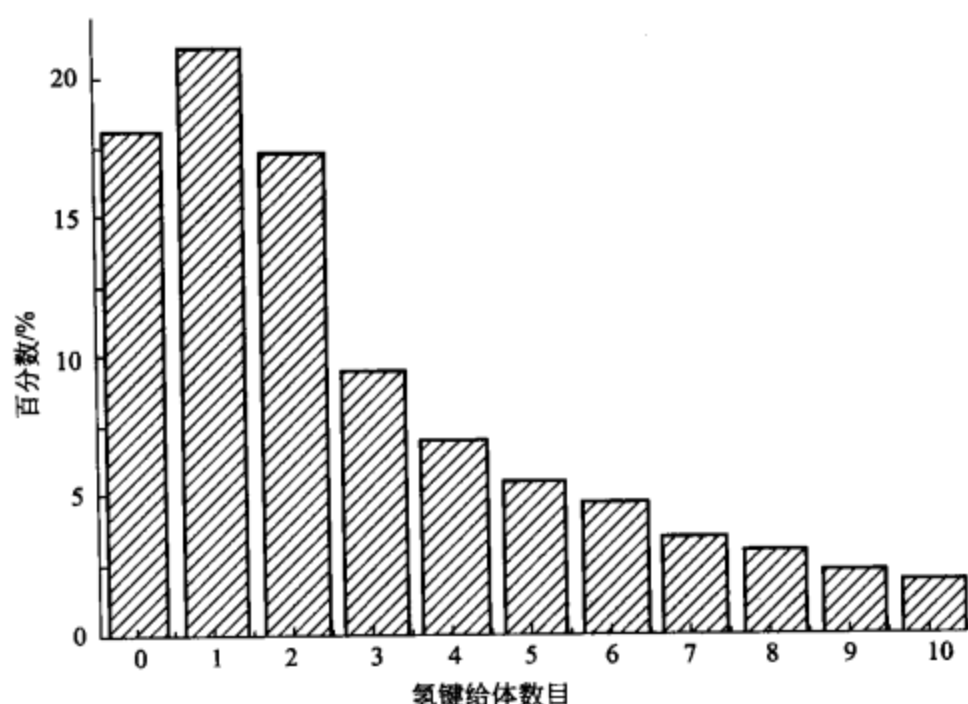


图 6-25 CHDD 的氢键给体数目分布

脂水分配系数 $\lg P$ 是用来衡量分子在正丁醇相和水相的分配系数，分子的亲脂性在分子的药代动力学上扮演着决定性的角色，已经广泛应用于评价膜的渗入和通透性中，包括胃肠道吸收、脑血屏障 (BBB) 以及与药代动力学性质的相互关系。这里计算了描述符 AlogP98 ， AlogP98 是 Ghose 等实现的基于原子类型的脂水分布系数，经验表明 AlogP98 处在区间 $[0, 6]$ 的分子有较好的药代动力学性质。从 CHDD 的 AlogP98 值的分布图上可以看出 (图 6-26)，CHDD 的分子大多数都落在区间 $[0, 6]$ 上，说明 CHDD 的分子大多具有好的药代动力学性质。

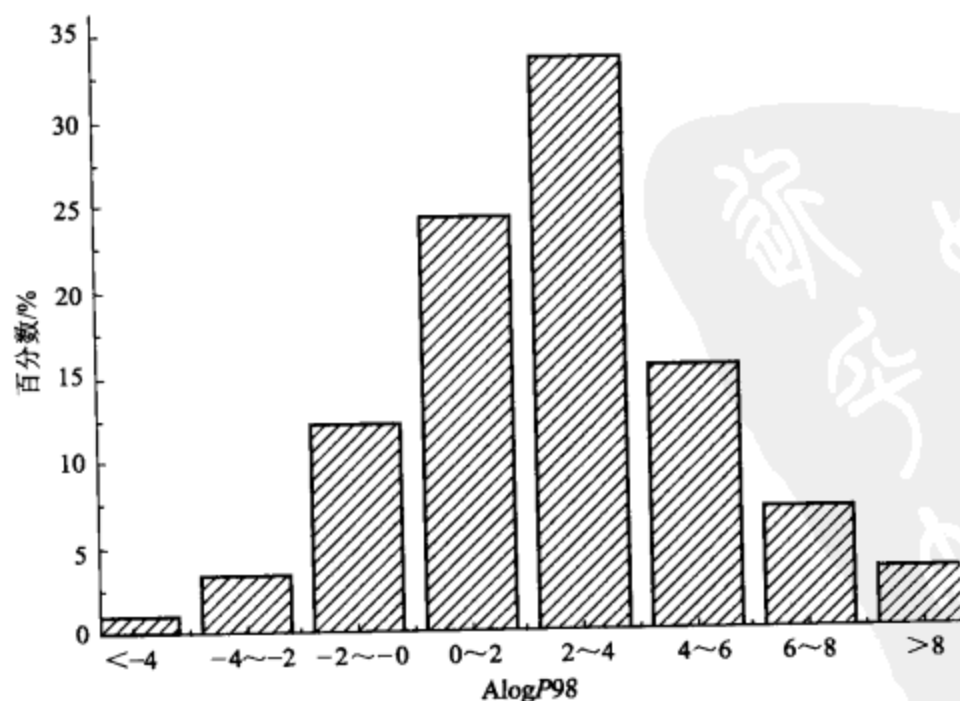


图 6-26 CHDD 的 AlogP98 数值分布

6.7.4 主成分分析

为了从多维描述符数据中提取合理而直观的信息, 在这里对所计算的 40 个描述符进行了主成分分析 (PCA)。主成分分析是一种将多变量数据向二维平面映射的方法, 这个方法首先计算相关矩阵, 然后对其进行对角化, 得到特征值和特征向量, 最后用特征向量矩阵作为一个变换矩阵, 将原始数据变换为新的数据。将变换后的数据把最大部分信息带进相关矩阵的新成分中的两个成分进行作图, 就可以完成映射过程。这个分析也是在 Cerius2 的 QSAR+ 模块中进行的。通过主成分分析, 多维描述符的数据可以有效地映射到二维图上 (图 6-27)。结果显示, CHDD 的前两个主成分的二维图具有较好的分散性, 这说明 CHDD 所包含的分子具有良好的多样性。

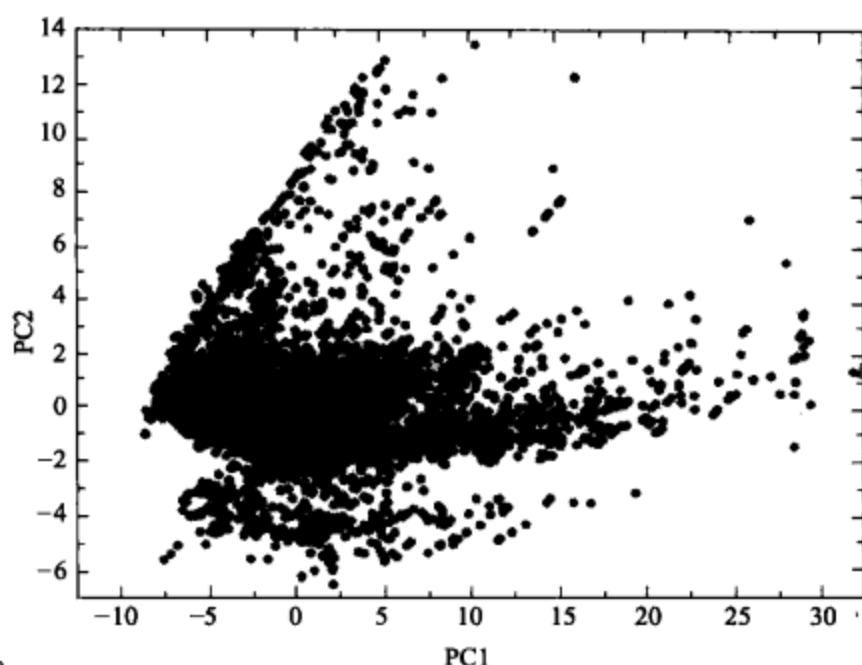


图 6-27 CHDD 的 PC2-PC1 图

6.7.5 结论

中医药具有悠久的历史, 如何最大程度地利用数千年来积累下来的中医药资源是一个重要的课题, 为此徐筱杰课题组在 2002 年开发了中草药有效成分三维结构与性质数据库 (CHDD), 根据以上对这个数据库的化学空间的分析, 并与前人对一些药物分子以及天然产物分子、半天然产物分子的分析结果作比较, 发现 CHDD 的分子在一些性质上与药物分子以及天然产物分子、半天然产物分子有许多相似之处, 但也有一定的差别, 这反映了 CHDD 包含的中草药有效成分化合物既有类药性质, 也是天然产物的一个特殊的子集。而从 CHDD 的前两个主成分的二维图可以看出, CHDD 具有良好的多样性, 可以作为一个好的存储、检索、管理中药的化学信息以及研究中药分子结构、活性, 计算受体与底物之间相互作用等的平台, 有助于进一步对中药治疗机理进行研究。

6.8 血府逐瘀汤潜在协同作用及抑制多个靶酶的计算探讨

6.8.1 介绍

长久以来,天然产物一直都是药物的主要来源,至今仍有超过 25% 的上市药物来源于天然产物。利用各地传统医学积累下来的医药信息,在从天然产物中寻找新型药物方面,可以给药物化学家们提供许多有用的线索。中药的使用已有数千年历史,积累了许多完备的药方及临床经验,这为开发现代治疗药物提供了许多直接的信息,也为对中药进行更有效的利用提供了可能。另一方面,在“回归自然”的世界性的趋势下,中药也引起了全世界越来越多的关注,在 2005 年,全球中药市场销售额近 50 亿美元,而中国加入 WTO,也在客观上刺激着中国中药产业的发展。如何更加有效地利用中药资源,并对现代化中药进行开发,成了一个紧迫的问题。

对中药通常的研发主要集中于对中药原植物活性成分进行分离、提取并检测,这方面的努力已经取得了一些重要的成果,如青蒿素 (artemisinin) 和麻黄素 (ephedrine) 的应用。然而,许多记载中具有临床疗效的中药方剂,在对它们的研究当中,并没有发现它们的组成植物中的化学成分对已知的靶点蛋白有明显的活性,或者有的虽然含有有活性的化合物,但这些化合物的含量偏低。通常认为,中药方剂有更复杂的机制,尚无法以现有的西方医药理论来解释中医药机制。

6.8.2 方法和预处理

这里主要应用化学信息学方法,对中药方剂血府逐瘀汤的组成进行药效和药物代谢动力学方面的探讨。血府逐瘀汤出自清朝医师王清任的《医林改错》,它是王清任医师根据前人的经验总结出的一个去血化瘀的药方,由当归、生地、桃仁、红花、枳壳、赤芍、柴胡、甘草、桔梗、川芎、牛膝等 11 种药物组成。这个方剂主要应用于心血管系统等方面,文献报道了血府逐瘀汤和丸制剂的临床疗效,而且同其他中药方剂一样,它也是一种广谱药剂,对许多疾病的临床疗效同样有文献报道。因此选择这个方剂着手对中药的治疗机制进行初步的探讨。

这里对血府逐瘀汤的成分进行了一系列的研究,主要考察分子多样性属性、ADME/Tox 属性以及分子同多个心血管疾病的靶酶的相互作用,以展示血府逐瘀汤的潜在的协同作用。计算在 Cerius2 的 QSAR+ 模块中进行,在 IBM IntelliStation Z Pro 工作站上进行。首先,根据血府逐瘀汤的植物组成,从 CHDD、《中药原植物化学成分手册》以及 15 篇文献收集到其中含有的 501 种分子,在分子计算软件 Cerius2 上搭建分子模型,分子的构象在 CAS 上进行校验,选用 MMFF94 力场,设置能量阈值为 $0.001\text{kcal}/(\text{\AA} \cdot \text{mol})$,对 501 个分子模型的结构进行优化。

分子式和来源植物见表 6-4。

在现代药物筛选中, 根据分子性质尤其是药物代谢动力学性质对类药分子进行预筛选(如使用“五规则”)是一件很重要的工作。在这里, 为了全面阐述血府逐瘀汤组成的整体性质, 并没有将一些不符合预筛选要求的化合物预先排除, 仍对这些性质进行了阐述。

已经知道, 化合物的化学结构决定其化学性质, 相似的化合物具有相似的性质, 根据这个想法发展出了一个更广泛的概念——化学空间(chemical space)。化学空间的概念是基于化合物可以通过一系列描述符来刻画其特性, 就可以计算出分子的一系列描述符, 然后用这些描述符的值作为需要的化学空间的点, 通过这个化学空间来对比分子的相似性, 进而推测其性质是否具有相似性。Feher 等已经对天然产物的分子与药物分子以及组合库化学分子的化学空间进行了比较。在此对血府逐瘀汤包含的化合物与一些治疗心血管疾病的药物分子及类药分子进行了比较。治疗心血管疾病的药物分子及类药分子收集自《心血管药理》。

6.8.3 血府逐瘀汤组分的化学空间分析

考虑了多种分子的性质, 计算了包括片段描述符、电性描述符、拓扑描述符、空间描述符、结构描述符、热力学描述符在内的共 34 个描述符: 分子量、C 原子数、O 原子数、N 原子数、其他重原子数、总重原子数、环(ring)数目、可旋转键数目、氢键受体数目、氢键给体数目、分子接触体积、分子表面面积、总极性表面面积(TPSA)、总疏水表面面积(THSA)、相对极性表面面积(relative polar surface area, RPSA)、相对疏水表面面积(relative hydrophobic surface area, RHSA)、总分子溶剂可及表面面积(total molecular solvent-accessible surface area)、AlogP98、手性中心数目、柔性指标、Balaban 指标(J_X)、分子回旋半径(r_{gyr})、Kappa 形状指标和修正形状指标(共 6 个)、分子偶极矩(dipole moment)、原子极化率总和(sum of atomic polarizabilities)、Wiener 指标、Hosoya 指标(lgZ)、Zagreb 指标、主惯性矩(principal moment of inertia)。部分计算结果见表 6-4。

表 6-4 血府逐瘀汤所含组分和药物分子/类药分子的一些分子描述符的最大值、最小值、中位值和平均值

项 目	血府逐瘀汤(501 个分子)				药物分子/类药分子(489 个分子)			
	最大值	最小值	中位值	平均值	最大值	最小值	中位值	平均值
分子量	1549.6	59.11	352	416.3	2180.3	116.12	390.52	417.45
C 原子数	69	3	19	21.93	98	3	20	21.56
O 原子数	38	0	5	7.45	33	0	4	4.14
N 原子数	7	0	0	0.13	24	0	2	2.60
重原子数	107	4	25	29.52	155	8	28	29.14
环数目	12	0	3	3.32	9	0	2	2.68

续表

项 目	血府逐瘀汤(501 个分子)				药物分子/类药分子(489 个分子)			
	最大值	最小值	中位值	平均值	最大值	最小值	中位值	平均值
可旋转键数目	57	0	7	9.9	93	0	8	8.79
氢键受体数目	38	0	5	7.5	35	1	5	5.74
氢键给体数目	23	0	2	4.1	31	0	2	2.26
分子接触体积	1351.3	74.44	319.47	385.11	1950.6	106.94	361.45	379.88
分子表面面积	1589.8	112.42	403.52	476.77	2473.9	140.41	456.96	480.84
总极性表面面积	863.78	0	161.96	214.07	955.55	0	170.70	177.26
总疏水表面面积	1506.5	66.95	386.15	435.15	1356.3	0	452.97	423.18
AlogP98	18.41	-9.55	2.2	1.96	10.96	-7.58	3.02	3.00
手性中心数目	41	0	3	6.61	27	0	1	2.1
柔性指标	45.79	0.86	5.36	7.266	53.36	1.18	6.39	7.05

从表 6-4 中的数据看,很明显,血府逐瘀汤所含分子与药物分子/类药分子的分子量的平均值接近,然而,药物分子/类药分子的分子量的中位值大于血府逐瘀汤所含分子。再看看其分布图(图 6-28),可以看出,两类分子在 0~100 区间的分子都比较少,而血府逐瘀汤所含分子的峰值处于区间 [100,200] 处,以后渐渐下

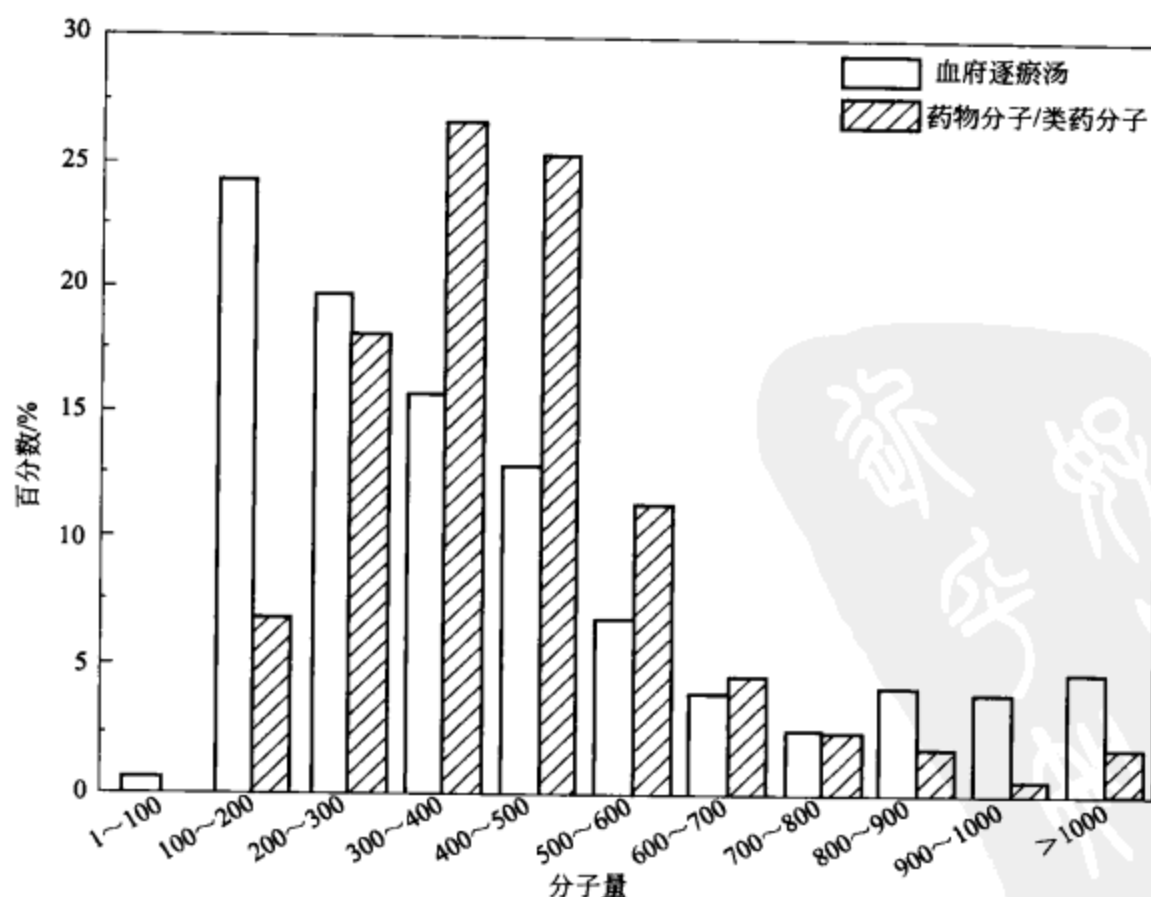


图 6-28 血府逐瘀汤所含组分和药物分子/类药分子的分子量分布图

降，然而到 800 以后又有一个尾巴，而药物分子/类药分子渐次上升，在区间 [300,400] 处达到峰值，然后渐次下降，形成一个明显的峰形，其分布显得更狭窄。大多数血府逐瘀汤所含分子的分子量符合 Lipinski 的“五规则”。然而也可以看到，分子量大于 500 的有近 30%，这些分子大多包含一个或多个葡萄糖环，这可能令这些分子难以被人体肠道吸收，或者也可能在进入人体吸收系统前被水解或代谢，以令这些分子以更小的分子形式被人体吸收。此外，血府逐瘀汤所含分子的碳原子数同药物分子/类药分子的中位数和平均值均接近，但是两类分子的氧原子数和氮原子数目有很大的差别，其中血府逐瘀汤所含分子的平均氧原子数接近药物分子/类药分子的 2 倍，这直接反映到两类分子的氢键给体数目上。

从它们的氧原子数分布图看（图 6-29），血府逐瘀汤所含分子显示了更大的复杂性。在氧原子数方面，约 85% 的药物分子/类药分子所含氧原子数目为 0~6 个，峰值处于 [3,4] 的区间上，血府逐瘀汤所含分子的氧原子数虽然峰值在 [1,2] 的区间上，然而它的氧原子数有个长长的尾巴，有超过 20% 的分子的氧原子数大于 10 个。在可旋转键数目上，药物分子/类药分子的峰值在 [7,8] 的区间上有近 20%，而血府逐瘀汤所含分子的峰值在 [3,4] 的区间上有 18%，然而药物分子/类药分子的可旋转键数目渐次下降，只有大概 2% 的分子有超过 20 个可旋转键，对血府逐瘀汤所含分子这个比例超过 10%。而在环数目方面，两类分子的峰值在环数为 2 的地方，然而药物分子/类药分子有更高的峰值和更明显的峰形，而血府逐

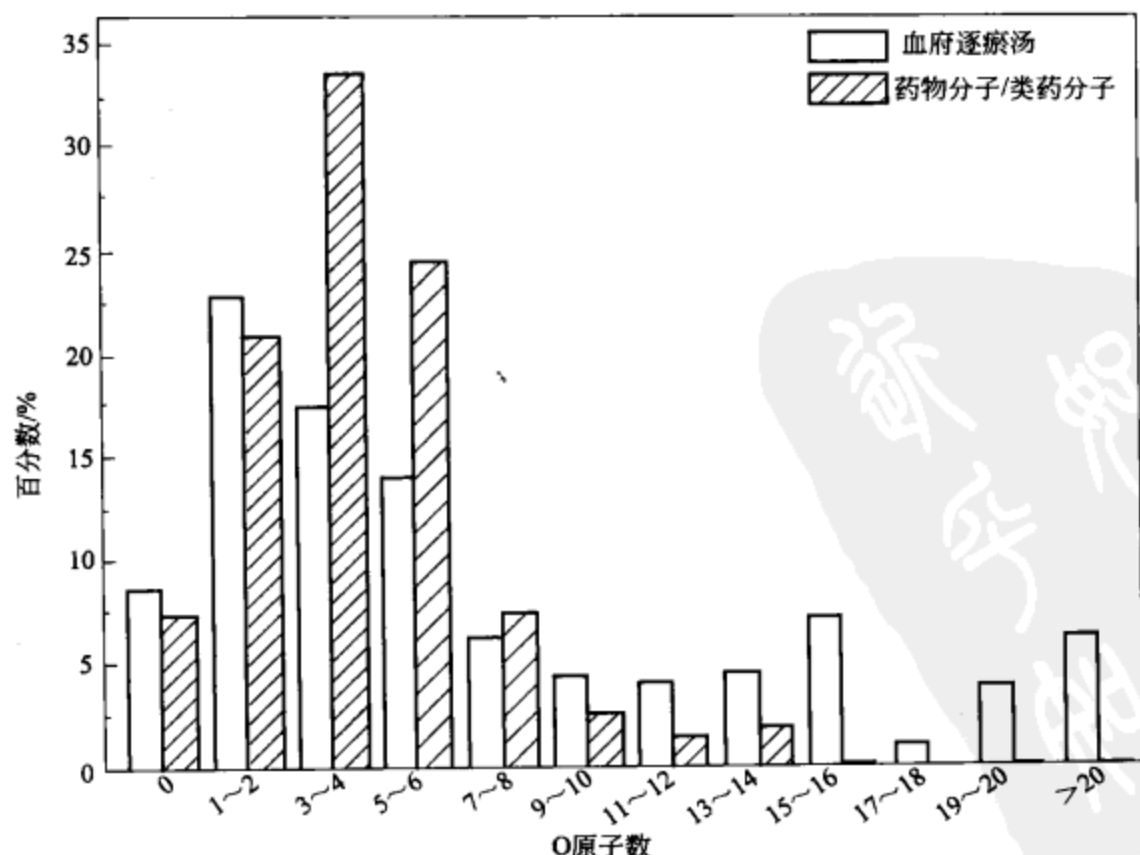


图 6-29 血府逐瘀汤所含组分和药物分子/类药分子的氧原子数分布图

瘀汤所含分子在 0~5 个环的所占的比例相近, 在 12%~17% 之间, 药物分子/类药分子则集中于 1~4 个环, 属于这区间的分子所占比例有 89%。

对药物的吸收来说, 其脂水分布系数 $\lg P$ 是一个很重要的指标, 广泛应用于评价膜的渗入和通透性, 包括胃肠道吸收、血脑屏障 (BBB) 以及与药代动力学性质的相互关系。这里预测了 AlogP98, 由图 6-30 可以看到, 两类分子的 AlogP98 的峰形相近。绝大多数分子都处在区间 $[0, 6]$ 之间, AlogP98 位于这个区间的化合物有较好的成药的药代动力学性质。

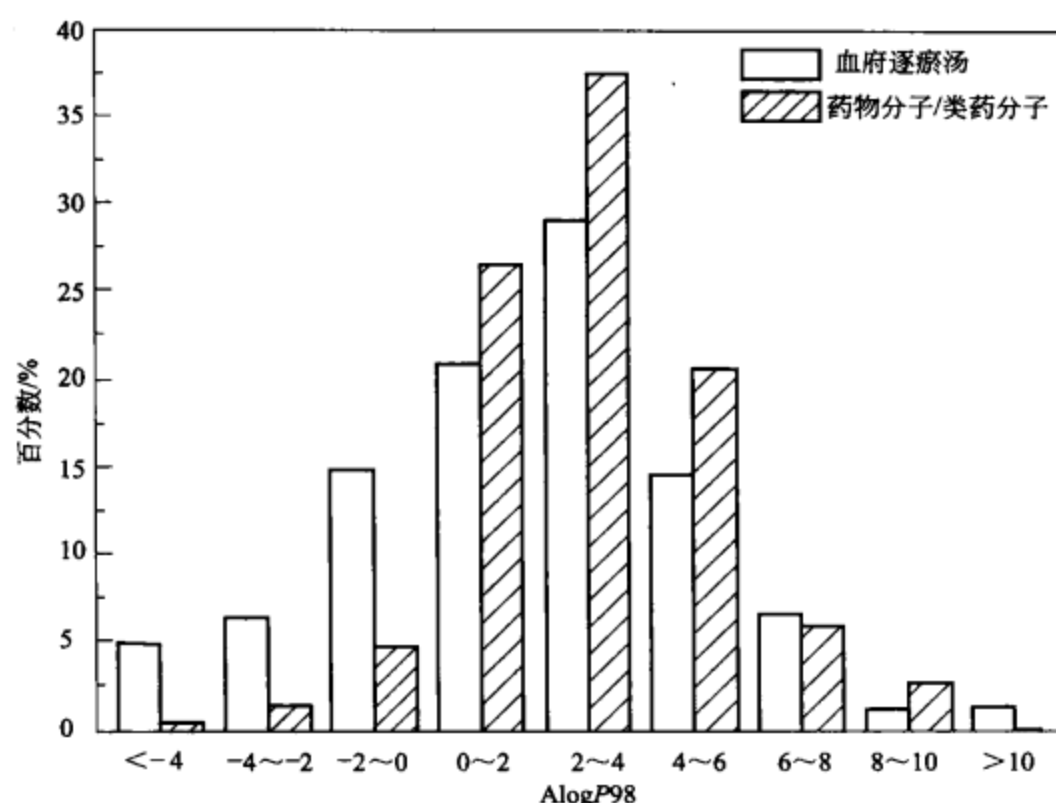


图 6-30 血府逐瘀汤所含组分和药物分子/类药分子的 AlogP98 分布图

这些结果同 Feler 等对类药分子以及天然产物的结果同样有点相近。但这个选择的小集合的特殊性使得结果有些不同, 如两类分子的氮原子数目的结果。根据以上的结果, 看到了血府逐瘀汤与治疗心血管疾病的药物分子/类药分子的异同。为了更加直观地显示两类分子的差别, 从上述计算的多元描述符中降低复杂度, 将数据进行简化, 从中提取合理化信息, 对这些化合物进行了主成分分析, 结果如图 6-31、图 6-32 所示。

结果显示血府逐瘀汤所含分子的多样性优于药物分子/类药分子, 药物分子/类药分子的分布更集中, 虽然有一两个离群点。中药方剂体系可以视为一个小型的天然组合化学库, 这提示了血府逐瘀汤所含分子可能比药物分子/类药分子对更广谱蛋白靶点有相互作用。

6.8.4 血府逐瘀汤与心血管系统靶酶的相互作用研究

将血府逐瘀汤所含的 501 个分子与一些重要靶酶进行对接, 以原配体的得分作

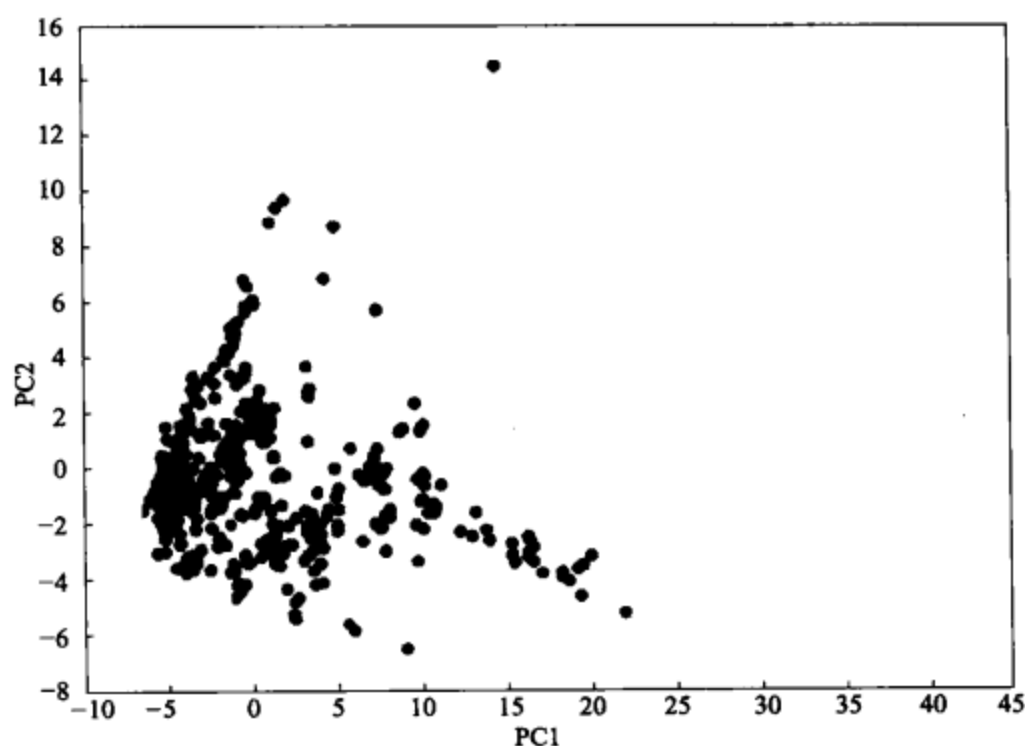


图 6-31 血府逐瘀汤所含分子的 PC2-PC1 图

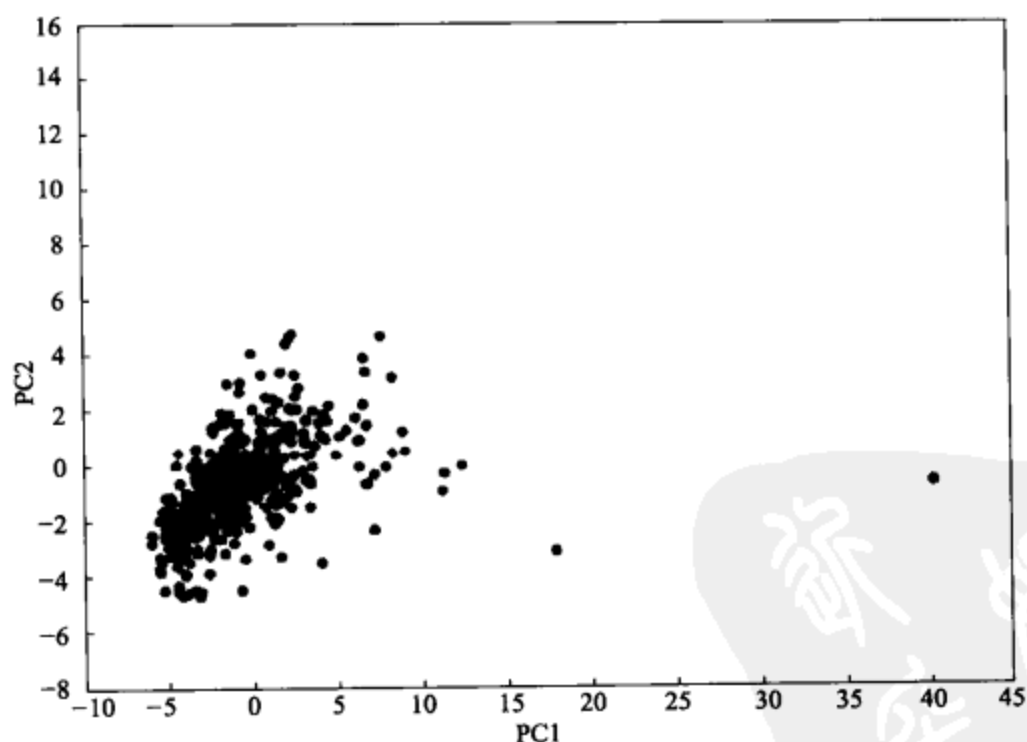


图 6-32 药物分子/类药分子的 PC2-PC1 图

为一个阈值，来获得与这些靶酶相应的有更高得分的化合物数目，用这种方法突出血府逐瘀汤所含化合物与心血管系统重要靶酶可能的活性。首先考虑了药物对肾素血管紧张素系统（renin-angiotensin system, RAS）的作用，这个系统是调节心血管生理功能的重要体液系统，由肾素（renin）、血管紧张素原（angiotensinogen）、血管紧张素转化酶（angiotensin-converting enzyme, ACE）I 和 II、血管紧张素

(angiotensin, Ang) 及相应的受体构成。在多种因素的作用下, 其中的 ACE I 在 RAS 中起着重要的作用, 它的抑制剂对高血压和心力衰竭有很好的作用, 并能起到抗心肌缺血、保护血管内皮细胞、纠正血脂紊乱以及抗动脉硬化等作用。最终肾素释放增加, 作用于血管紧张素原, 使其生成十肽化合物血管紧张素 I (Ang I), Ang I 在 ACE 的作用下转化为 Ang II。Ang I 转化为 Ang II 除通过 ACE 的作用途径外, 也可经糜酶 (chymase, Chy)、组织蛋白酶 G (cathepsin G) 的作用途径。Ang II 与效应器细胞膜上的特异性受体结合产生生物学效应。此外, 心血管组织中的 RAS 在高血压、心血管重构、动脉粥样硬化及再狭窄等的发生和发展过程中起重要作用。

对 RAS 的 3 个重要的靶蛋白 ACE I、ACE II、Renin 进行了分子对接, 结果见表 6-5。

表 6-5 血府逐瘀汤原植物所含化合物与多个靶点对接命中数目

植物	化合物数目	renin	ACE I	ACE II	VGFR	HMG-CoA, 位点 1	HMG-CoA, 位点 2	Pgp, ATP 位	Pgp, ADP 位
总数	501 ^①	10	128 ^②	无	183 ^③	64	34	10	13
牛膝	14	—	3	—	2	—	—	—	—
当归	77	—	11	—	55	—	—	—	—
柴胡	74	1	3	—	30	—	1	—	1
红花	24	—	7	—	5	4	3	3	2
枳壳	23	3	2	—	10	1	4	1	—
甘草	104	3	35	—	4	25	14	1	7
川芎	78	—	23	—	47	3	1	—	—
赤芍	19	1	8	—	4	7	7	2	1
桔梗	30	—	1	—	—	—	—	—	—
桃仁	6	—	6	—	3	4	—	—	—
地黄	82	1	33	—	34	20	4	3	2

① 有的化合物同时包含于不同植物。

② 有的化合物在不同植物中重复, 其中: 1 个在牛膝和川芎; 2 个在当归和川芎; 1 个在红花和地黄。

③ 有的化合物在不同植物中重复, 其中: 2 个在牛膝和柴胡, 3 个在当归和川芎, 2 个在柴胡和地黄, 1 个在红花和地黄, 1 个在当归和地黄, 1 个在当归、红花、地黄。

注: 对接所用靶点的 PDB 代码为: renin (1BIL), ACE (1J37), ACE2 (1R42), HMG-CoA (1HW8), Pgp (1MV5)。

同时, 另外一些对心血管疾病重要的靶酶也被考虑, 如血管内皮生长因子 (VGFR) 以及实现调节去肾上腺素和肾上腺素对体内器官功能的肾上腺素受体。从表 6-5 的结果可以看到, 血府逐瘀汤所含化合物没有一个分子对 ACE II 有较好的匹配, 而 ACE II 在 SARS 的病理途径中起了作用。

另外, 许多分子对一些重要的靶酶有较好的对接匹配, 有的化合物与不止一个酶有好的结果, 如甘露三糖 (manninotriose)。许多科学家喜欢使用“激酶属性”

(kinaseprofile) 来代替“抑制剂”(inhibitor) 这个术语, 这可以用来描述使用单个化合物在一个协同方案中调节几个并协发生和相关通路的观念, 这种潜在的协同方案有助于解释血府逐瘀汤在治疗心血管疾病中的良好疗效。当然, 进一步的解释需要进一步的实验支持。

6.8.5 血府逐瘀汤药物代谢动力学性质的研究

血府逐瘀汤是一个口服汤剂, 在前边的讨论中看到了血府逐瘀汤可能是心血管疾病众多途径同时作用的类药分子或类先导化合物的集合, 有的化合物在实验中也表现出了较好的活性, 然而这些化合物其实在原植物当中的含量并不高。这里对血府逐瘀汤的药物代谢动力学性质(ADME/T)进行了预测, 其结果见图 6-33。需要指出的是, 对 ADME/T 的预测并不足够准确, 尤其是对 P4502D6 抑制的预测和肝毒性的预测, 不过结果仍有指导意义。

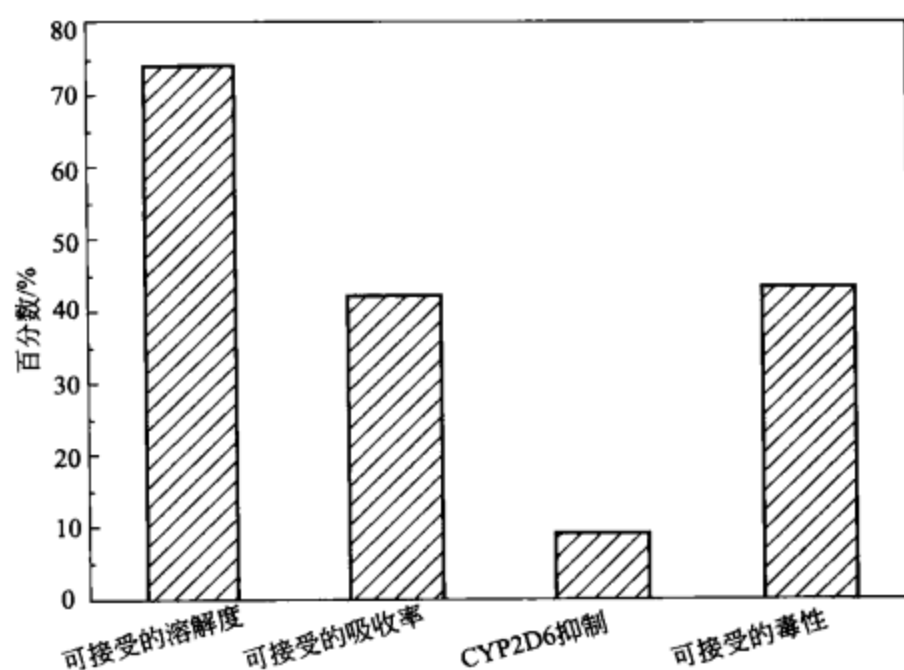


图 6-33 血府逐瘀汤的可接受的溶解度、吸收率、代谢、毒性图

从图 6-33 中可以看到, 超过 70% 的化合物有可以接受的水溶性, 但其肠吸收、肝毒性、P4502D6 抑制并没有期待的那样好, 尤其是超过半数分子具有坏的肝毒性, 这与以前的印象不大一致。除了预测的误差之外, 这也提出了一个要求, 要完善中药的毒性实验。当然, 实验表明血府逐瘀汤毒性很小, 对这种原因的解釋可能有点: ①有毒性的化合物含量过低; ②通过人体内二次代谢。这些问题值得关注, 需通过实验检验。

另外, 吸收差也会影响药剂的疗效, 对单个化合物吸收不好的情况在混杂情形中会得到改善。药物化学家们在一些领域研究过天然产物的协同作用, Keung 的研究表明提取的黄豆苷(daidzin)的剂效 EC50 是纯黄豆苷的 10 倍。近年来, ATP 结合盒转运载体蛋白质对药物体内过程的影响已被广泛研究。P-糖蛋

白 (P-gp) 是其中最大的一个亚系。研究发现, P-gp 在许多组织中有分布, 是一种 ATP 依赖性膜转运体, 作为药物转运子, 其作用类似于排出泵, 可将药物从细胞内外排而使胞内药物浓度降低, 从而降低药效。因此, P-gp 与底物及调节子之间的相互作用能影响药物的吸收、分布、代谢、排泄。它虽然不是心血管疾病的靶蛋白, 但是它在药物吸收转运时充当转运载体, 具有极其重要的作用。将这个酶同血府逐瘀汤所含分子进行对接, 结果见表 6-5。有一些分子可以与 P-gp 结合, 这有助于考虑协同作用对中药 ADME/T 性质的帮助。

参 考 文 献

- [1] Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British Journal of Pharmacology*, 2007; 1-12.
- [2] Christopher Lipnski, Andrew Hopkins. Navigating chemical space for biology and medicine. *Nature*, 2004, 432: 855-861.
- [3] Cristopher M Dobson. Chemical space and biology. *Nature*, 2004, 432: 824-828.
- [4] 徐筱杰, 侯廷军, 乔学兵, 章威编著. 计算机辅助药物分子设计. 北京: 化学工业出版社, 2004.
- [5] 徐筱杰编著. 超分子建筑. 北京: 科学文献出版社, 2000.
- [6] Gasteiger J, ed. *Handbook of Chemoinformatics*. Weinheim: Wiley-VCH, 2003.
- [7] Oprea T I, ed. *Chemoinformics in Drug Discovery*. Weinheim: Wiley-VCH, 2005.
- [8] Hou T J, Xu X J. Recent development and application of virtual screening in drug discovery: An overview//*CURRENT PHARMACEUTICAL DESIGN*, 2004, 10 (9): 1011-1033.
- [9] Hou T J, Wang J M, Zhang W, Xu X J. ADME evaluation in drug discovery. 6. Can oral bioavailability in humans be effectively predicted by simple molecular property-based rules? *Journal of Chemical Information and Modeling*, 2007, 47 (2): 460-463.
- [10] Hou T J, Wang J M, Zhang W, Xu X J. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *Journal of Chemical Information and Modeling*, 2007, 47 (1): 208-218.
- [11] Hou T J, Zhang W, Huang Q, Xu X J. An extended aqueous solvation model based on atom-weighted solvent accessible surface areas; SAWSA v2.0 model. *Journal of Molecular Modeling*, 2005, 11 (1): 26-40.
- [12] Hou T J, Zhang W, Xia K, Qiao X B, Xu X J. ADME evaluation in drug discovery. 5. Correlation of Caco-2 permeation with simple molecular properties. *Journal of Chemical Information and Computer Sciences*, 2004, 44 (5): 1585-1600.
- [13] Hou T J, Xu X J. ADME evaluation in drug discovery. 2. Prediction of partition coefficient by atom-additive approach based on atom-weighted solvent accessible surface areas (vol 43, pg 1058, 2003). *Journal of Chemical Information and Computer Sciences*, 2004, 44 (4): 1516-1516.
- [14] Hou T J, Xu X J. Calculations of free energies based on the combination of molecular dynamics simulations and continuum solvation model. *Progress in Chemistry*, 2004, 16 (2): 153-158.
- [15] Hou T J, Qiao X B, Zhang W, Xu X J. Empirical aqueous solvation models based on accessible surface areas with implicit electrostatics. *Journal of Physical Chemistry B*, 2002, 106 (43): 11295-11304.

- [16] Hou T J, Xu X J. A new molecular simulation software package-Peking University Drug Design System (PKUDDS) for structure-based drug design. *Journal of Molecular Graphics & Modelling*, 2001, 19 (5): 455.
- [17] Smithing M P, Darvas F. HazardExpert: an expert system for predicting chemical toxicity, in: Finlay J W, Robinson S F, Armstrong D J, eds. *Food Safety Assessment*; Washington D C: American Chemical Society, 1992: 191-200.
- [18] William L Jorgensen. The Many roles of Computation in Drug Discovery. *Science*, 2004, 303: 1813-1818.
- [19] Keith C T, Borisy A A. Multicomponent therapeutics for networked systems. *Nature Review Drug Discovery*, 2005, 4: 71-78.
- [20] Zimmermann G R, Lehar J, Keith C T. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discovery Today*, 2007, 12 (1-2): 34-42.
- [21] Ekachai Jenwitheesuk, et al. Novel paradigms for drug discovery: computational multitarget screening trends in pharmacological. *Science*, 2007, 29 (2): 62-71.
- [22] Maris Lapinsh, et al. Improved approach for proteochemometrics modeling: application to organic compound-amine G protein-coupled receptor interactions. *Bioinformatics*, 2005, 21 (23): 4289-4296.
- [23] Xiaojie Xu. New concept and approaches for drug discovery based on Traditional Chinese Medicine. *Drug Discovery Today: Tech*, 2006, 3 (3): 247-253.
- [24] Qin huang, et al. Potential synergism and inhibitors to multiple target enzymes of Xuefu Zhuyu Decoction in cardiac disease therapeutics: A Computational Approach. *Bioorganic & Medicinal Chemistry Letters*, 2007, 17: 1779-1781.
- [25] Li X D, Zhang W, Qiao X B, Xu X J. Prediction of binding for a kind of non-peptic HCVNS3 serine protease inhibitors from plants by molecular docking and MM-PBSA method. *Bioorganic & Medicinal Chemistry*, 2007, 15 (1): 220-226.
- [26] Qiao X B, Hou T J, Zhang W, Guo S L, Xu S J. A 3D structure database of components from Chinese traditional medicinal herbs. *Journal of Chemical information and Computer Sciences*, 2002, 42 (3): 481-489.
- [27] Hou T J, Li Y Y, Liao N, Xu X J. Three-dimension quantitative structure-activity relationship analysis of some cinnamamides using comparative molecular similarity indices analysis (CoMSIA). *Journal of Molecular Modeling*, 2000, 6 (5): 438-445.
- [28] Hou T J, Wang J M, Chen L R, Xu X J. Automated docking of peptides and proteins by using a genetic algorithm combined with a tabu search. *Protein Engineering*, 1999, 12 (8): 639-647.
- [29] Houck K A, Kavlock R J. Understanding mechanisms of toxicity: Insights from drug discovery research. *Toxicology and Applied Pharmacology*, 2008, 277: 163-178.
- [30] Rajman I. PK/PD modelling and simulations: utility in drug development. *Drug Discovery Today*, 2008, 13: 341-346.
- [31] Hellebrecht A, Klebe G. Use of QSAR models for database screening: A feasibility study. *J. Chem. Inf. Model*, 2008, 48: 384-396.

第7章 基于亲和前沿色谱-质谱联用技术的药用天然产物筛选方法

7.1 亲和色谱及前沿色谱介绍

7.1.1 亲和色谱

亲和色谱 (affinity chromatography) 也称为亲和层析, 是 20 世纪 60 年代末发展起来的一项分离纯化技术, 利用生物分子间具有的专一性亲和力进行分离, 已得到广泛的应用。亲和色谱的原理即利用抗原和抗体、酶和底物、配体和受体等特异性反应物之间的亲和力, 把被识别的分子 (底物、配体或抗原) 以共价键结合到含有活化基团的固相载体上, 然后使含有欲分离的大分子化合物的混合液从这个载体滤过, 绝大部分对配体没有亲和力的化合物均顺利通过载体而不滞留, 欲分离的大分子能识别配体 (有亲和力) 并与其结合而滞留在色谱柱上, 待所有杂质从色谱柱上流走后, 即第一个色谱峰出现后, 改变洗脱条件, 促使结合在配基上的大分子化合物解离下来并形成第二个色谱峰, 原来混合液中欲分离的大分子化合物则以高度纯化的形式在流出液中出现, 得到需要分离纯化的物质。亲和色谱通常用来分离纯化各种生物样品, 几乎适用于任何可以纯化的大分子化合物, 如酶、抗体、激素、药物、核酸、维生素结合蛋白、转运蛋白、阻抑蛋白、载脂蛋白和其他调控成分等。外界条件如 pH 值、离子强度、缓冲液组成、化学环境往往对亲和固定相有较大的影响。相对来说, 亲和色谱的适用范围较窄, 如不能使用强酸、强碱、重金属, 不能耐高温、高压等, 但它对某些生物样品的分析分离具有其他色谱手段无法比拟的优势。

化合物要对靶蛋白产生作用, 首先要有一定的亲和力, 亲和力越强, 越有可能成为药物先导化合物。利用这一性质, 就能用靶蛋白将少数有活性的化合物从大量无活性的化合物中直接分离出来。将化合物库与药物作用靶在一起孵育, 所有有亲和力的化合物 (有活性的化合物) 将与靶结合, 而无活性的化合物仍为游离状态, 这样, 结合一定的手段很容易实现分离, 然后只对活性化合物进行鉴定, 从而省去了针对大量无活性化合物的工作量。亲和筛选法适于针对特定药物作用靶的筛选, 并且需要纯化的靶分子。

亲和筛选法的优点有如下几点。

① 可以获得构效关系的信息。在一次测定中, 所有有亲和活性的化合物均与

靶结合，通过测定不同的解离速度可确定化合物的活性顺序。

② 不会漏筛。因为所有有亲和活性的化合物均与靶结合，并且活性最强的亲和力和力最强，所以不会漏筛。

③ 对浓度要求不高。在靶大量过量的情况下，结合灵敏的检测手段，微量的活性物质也能检出，也能适用于化合物种类和浓度均未知的天然产物的筛选。

④ 测定快速。以将靶固定在色谱填料上的亲和色谱为例，活性化合物结合到柱子上之后，在洗脱条件下，活性由弱到强依次流出柱子，通过一个灵敏的检测器，如质谱，在一个色谱过程中就能获得所有活性化合物的活性、浓度和结构的信息。

亲和色谱操作过程简述如下。

(1) 载体的选择 选择亲和色谱所用的载体必须符合的条件如下：非特异性吸附低；良好的流动特性；在 pH 离子强度和变性剂浓度较大改变的情况下化学和物理等理化性能稳定；具备大量能被活化的化学基团；高度有效的多孔结构。

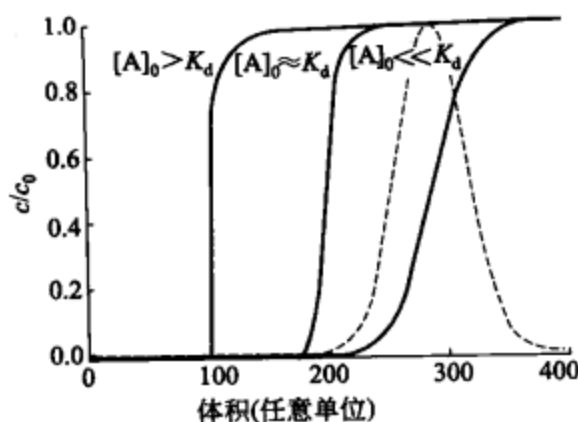
(2) 配体（配基）的选择 亲和色谱应根据被分离纯化的物质的理化性质和生物学性质选择配体。一般考虑选择的条件是：与纯化的物质有较强的亲和力；具有与基质共价结合的基团。

(3) 配体与载体的联结 配体与载体的联结方法有物理和化学的两类。最常用的方法是化学法，配体与载体共价联结的方法包括载体功能基团的活化；配体与活化基团的联结。使二者进行联结的化学反应必须足够温和，使配体和载体均可耐受而不发生变性。常用的化学试剂有溴化氰（CNBr）、环氧氯丙烷、双环氧乙烯、丁二烯砜和二乙醇胺等。

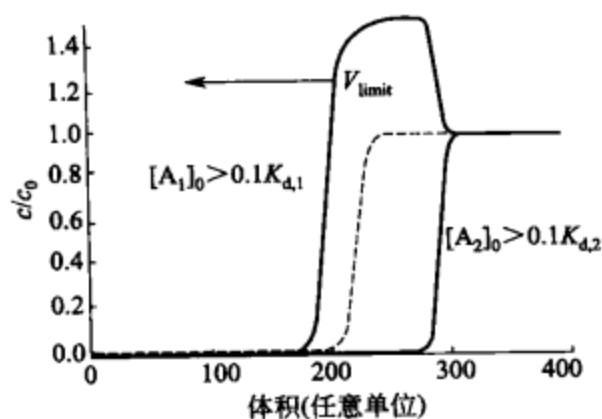
目前，亲和色谱也用于固定生物大分子，具备亲和能力的蛋白、酶、抗体、DNA、氨基酸等常与载体连接后用做亲和色谱固定相，反过来寻找与之结合的小分子化合物。其操作过程与固定化小分子配体的差异不大。下文所述除指明外均指基于固定化生物大分子的亲和色谱。

7.1.2 亲和前沿色谱

近年来，亲和前沿色谱（FAC）正迅速发展成一种定量分析生物分子之间相互作用的方法，也是一种很有前途的筛选方法，正成为一个新的研究热点。亲和前沿色谱与标准亲和色谱的差异在于样品是连续不断地上样的。将被分析物样品溶液以一定的浓度连续上样，直到亲和柱达到饱和吸附，此时开始出现样品的出峰前沿，同时以相应的空白柱扣除非特异性吸附作用（也得到相应的前沿峰），通过计算前沿时间（或者前沿体积）来考察固定相对被分析物的特异性亲和能力。图 7-1 显示了亲和前沿色谱与亲和色谱的色谱行为差别及多配体存在时亲和前沿色谱的色谱行为。



(a) 随给定配体浓度提高出峰加快 (实线, 从右至左)。在线性色谱条件下 ($[A]_0 \ll K_d$) 带状色谱存在一直接的关系, 穿透曲线的保留时间同带状峰 (虚线) 一致



(b) 多配体洗脱操作下非线性色谱条件的影响。在配体 A_2 存在下, 配体 A_1 给出期待的穿透曲线 (虚线)。在配体 A_2 存在下, 配体 A_1 出峰有所加快, 而且由于配体置换引起瞬时过浓度

图 7-1 从线性到非线性色谱行为
 c/c_0 —同属配体浓度与进样浓度之比

Ken-Ichi Kasai 等发展了前沿色谱理论, 提出了计算结合力的公式。利用亲和前沿色谱来定量计算柱容及解离常数的公式如下:

$$V - V_0 = \frac{B_t}{c + K_d} \quad (7-1)$$

式中, V 代表样品的前沿体积, V_0 代表死体积 (通常表示样品的非特异性吸附能力), B_t 表示亲和柱的柱容, c 代表样品浓度, K_d 为解离常数 (表征样品与固定相之间相互作用的强弱)。

式(7-1) 常转换成另一种形式表现, 如下:

$$\frac{1}{c(V - V_0)} = \frac{1}{B_t} + \frac{K_d}{B_t} \times \frac{1}{c} \quad (7-2)$$

测量一组不同浓度下的前沿体积, 由 $1/c(V - V_0)$ 对 $1/c$ 作图, 由拟合的直线的截距和斜率即可求出 B_t 和 K_d 。

当样品浓度 $c \ll K_d$ 时, 样品的前沿体积不再受样品浓度的影响, 而只与柱容及解离常数 (K_d) 有关, 式(7-1) 可以简化如下:

$$V - V_0 = \frac{B_t}{K_d} \quad (7-3)$$

在同样浓度下, 固定相对被分析物的亲和能力越强, 则样品在此亲和柱上的保留时间越长, 因此其前沿时间越长, 相应的前沿体积越大。亲和前沿色谱使得定量分析固定相与被分析物之间的相互作用 (常用解离常数 K_d 表示) 及测定亲和柱的柱容成为可能, 从而可以通过计算不同分析物的出峰前沿 (保留时间或保留体积) 定量比较它们与亲和固定相之间的相互作用强弱。

当样品是混合物进样时，与固定相不结合的化合物在死体积流出，配体与柱上的受体结合而被保留，当受体饱和或配体被更强的配体替代后，配体从柱子中流出。在一个色谱过程中，所有配体依它们与受体的亲和力作用大小由弱到强依次流出，反映它们亲和力的保留时间或保留体积也可同时测定。

使用亲和前沿色谱的优势主要在于：第一，它可以通过定量的方式直接地反映出分析物与亲和受体之间的作用强弱，前沿时间（保留时间）越长（相应的前沿体积也越大），说明固定化亲和受体对分析物的亲和作用越强；第二，因为亲和前沿色谱的显著特点就在于将一定浓度的样品连续上样并连续检测，反映出来的是样品的原始浓度（不像一般的色谱检测是通过脉冲的方式上样，样品在检测过程中被稀释），因此检测的灵敏度大大提高。

应当指出的是，在前沿亲和色谱中，流出成分仍为混合物，并没有实现真正意义上的分离。

7.2 亲和前沿色谱-质谱联用

1998 年，D. C. Schriemer 和 O. Hindsgaul 首次将亲和前沿色谱（FAC）与质谱（MS）联用。亲和前沿色谱可直接测定生物大分子与配体之间的结合，质谱可以检测混合物及确定化合物，这两者的联用为研究生物分子间的亲和力提供了有力的工具。下面详细介绍亲和前沿色谱-质谱联用（FAC-MS）方法及其应用。

7.2.1 原理与装置

FAC-MS 原理可概述如下（图 7-2）：样品持续进样，通过固定化的蛋白，与

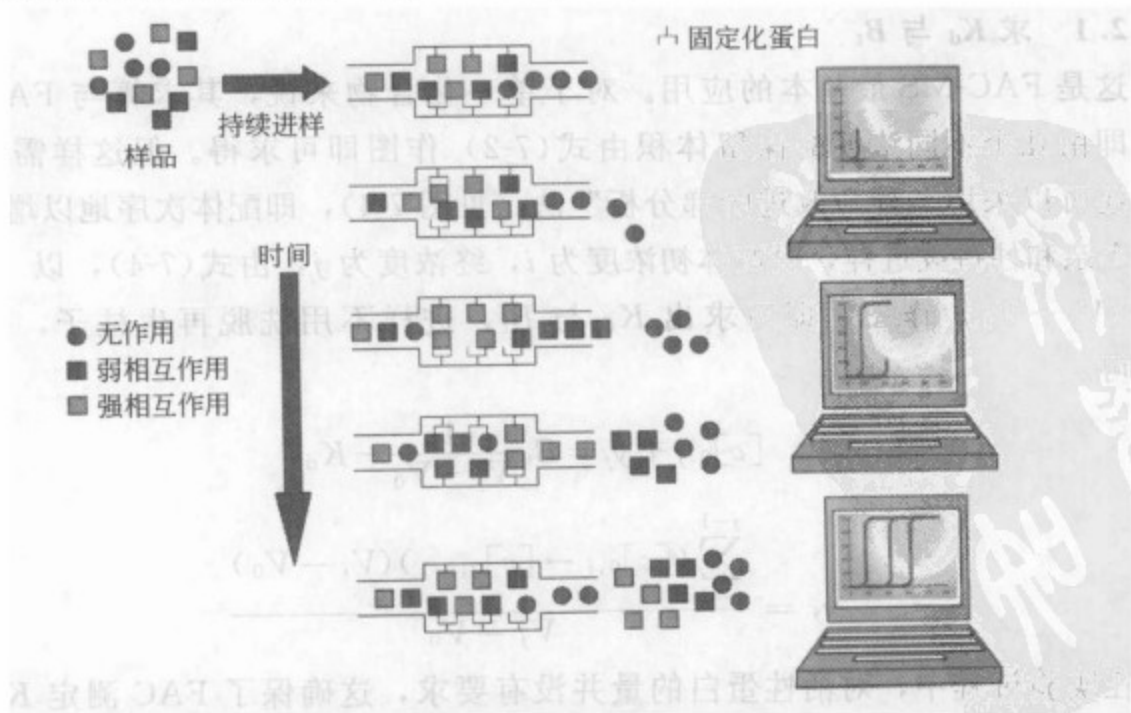


图 7-2 FAC-MS 原理示意（样品持续进样，依亲和力由弱到强流出被质谱检测）

蛋白相互作用，样品与蛋白有不同的亲和力，与蛋白没有作用的化合物将最早流出柱子并被质谱检测到，与蛋白有弱的相互作用的化合物将稍晚流出柱子，结合强的化合物将最晚流出柱子。最终流出柱子的溶液浓度等于样品的上样浓度。质谱记录的是总离子流图（total-ion chromatogram, TIC）和每一时刻的质荷比图，从特异的荷质比峰，可以从 TIC 提取得到选择离子流图（selected ion chromatogram, SIC），从而得到对应化合物的保留时间，与流速相乘即得到保留体积。

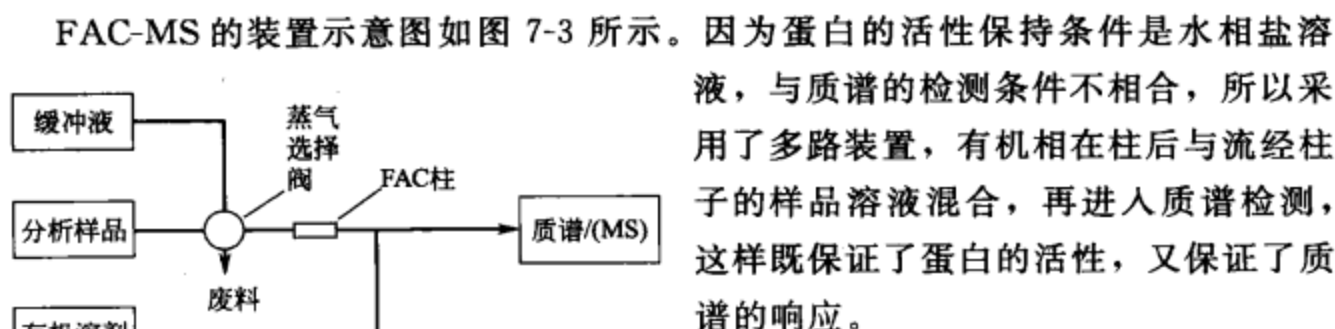


图 7-3 FAC-MS 系统示意

在 FAC-MS 方法中，为减少蛋白及配体的用量，一般使用微型亲和柱，

每一个柱子体积仅为 2~40L，固定在载体上的蛋白量为 10~1000pmol。质谱使配体的检测非常灵敏。更重要的是，质谱提供了多一维的信息，即质荷比 (m/z)。通常的 FAC 柱常采用单波长的紫外/可见 (UV/Vis) 光谱仪检测，在一个时间点所有的组分都对测量的响应有贡献，这样痕量组分难以从过量的其他化合物的背景中分辨出来。而使用质谱检测则提供了二维的信息，只要质荷比不同，即使痕量的组分也可被检测。

7.2.2 FAC-MS 的应用

7.2.2.1 求 K_d 与 B_t

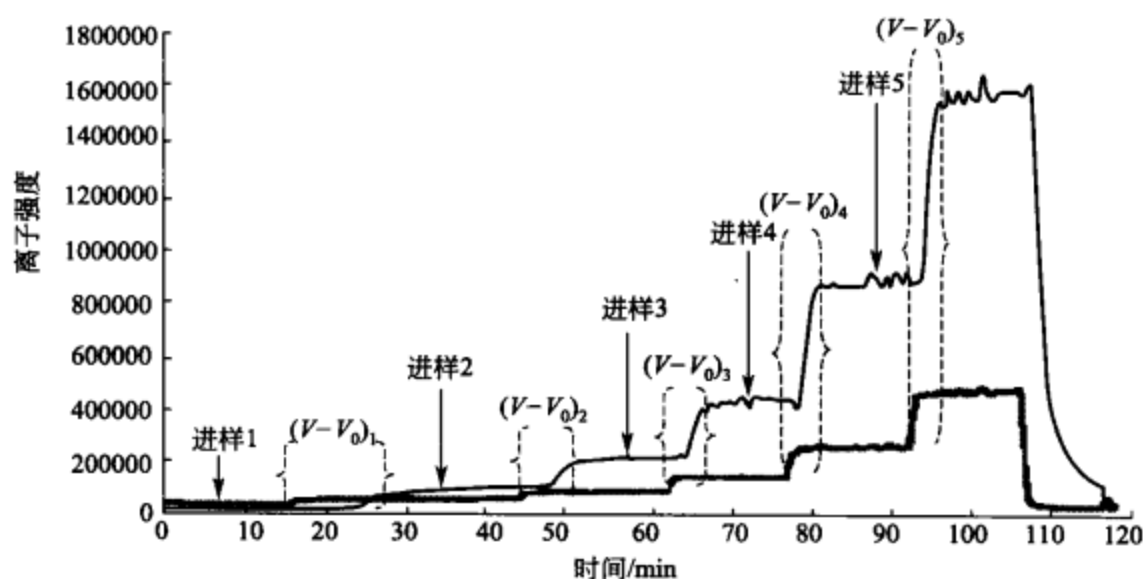
这是 FAC-MS 最基本的应用。对于单一化合物来说，其求算与 FAC 没有差别，即由几个不同浓度的保留体积由式(7-2)作图即可求得。但这样需要测定多次。还可以采用一种“改进阶梯分析”法（见图 7-4），即配体次序地以增大的浓度对一个亲和柱持续进样，设配体初浓度为 i ，终浓度为 j ，由式(7-4)，以 $[C]_0 + y_j$ 对 $1/(V_j - V_0)$ 作图，即可求出 K_d 与 B_t ，进样不用洗脱再生柱子，做一次实验即可。

$$[c]_{0,j} + y_j = B_t \frac{1}{V_j - V_0} - K_d \quad (7-4)$$

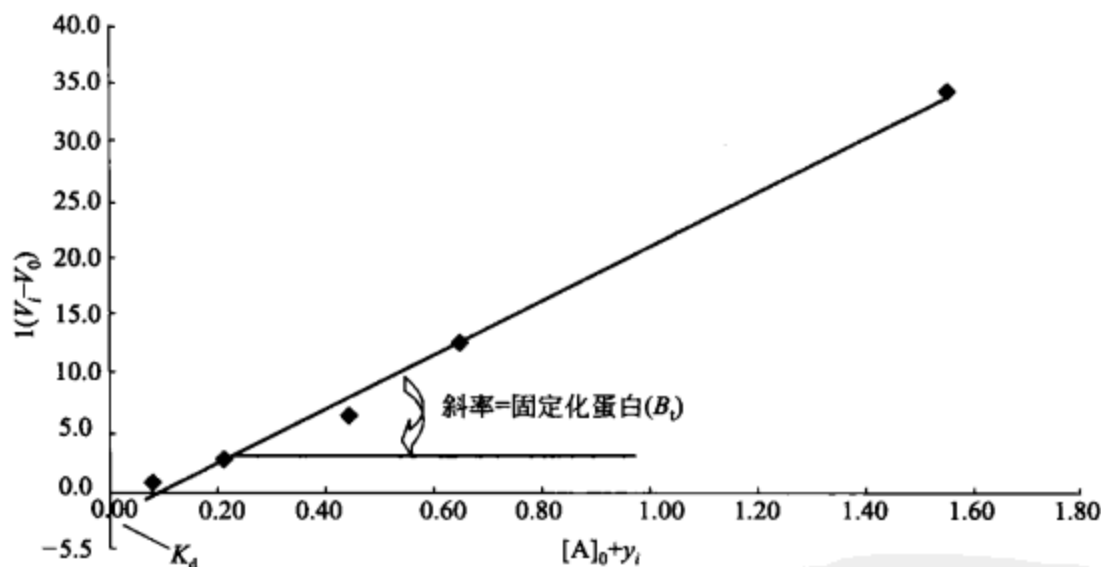
$$y_j = \frac{\sum_{i=1}^{j-1} ([c]_{0,i} - [c]_{0,i-1})(V_i - V_0)}{V_j - V_0}$$

其中

在实验过程中，对活性蛋白的量并没有要求，这确保了 FAC 测定 K_d 值的准确性，因为在制备蛋白的过程中很难保证蛋白有 100% 的活性。



(a) 配体（浅色线，箭头所指）以逐渐增大的浓度多次进样。
空白（深色线）随配体一起进样



(b) 保留体积的倒数 $1/(V_j - V_0)$ 对总配体浓度 $[A]_0 + y_j$ 作图，
为线性关系。由直线 Y 轴的截距求得 K_d ，斜率求得 B_i

图 7-4 “改进阶梯分析”法示意

7.2.2.2 按亲和力分级

在筛选过程中，FAC-MS 可依配体的不同保留时间确定出保留时间长的配体。如果是对先导化合物进行优化过程，混合物中存在多个活性配体时，FAC-MS 能否正确地对配体进行分级呢？N. W. C Chan 等固定了 SDH（sorbitol dehydrogenase，山梨糖醇脱氢酶），选择了 7 个配体，其活性 IC_{50} （为单一化合物的酶法分析的测定值）范围从 nmol/L 到 mol/L 级。这 7 个配体与空白标记物一起进样，可在一次实验中对不同亲和力的配体实现分级，见图 7-5。虽然混合物使配体的保留时间压缩，且存在配体的替代（displacement）现象 [参见图 7-1(b) 和图 7-6]，这

些并不影响配体的分级，这在先导化合物的优化中对不同配体的比较是非常有用的。

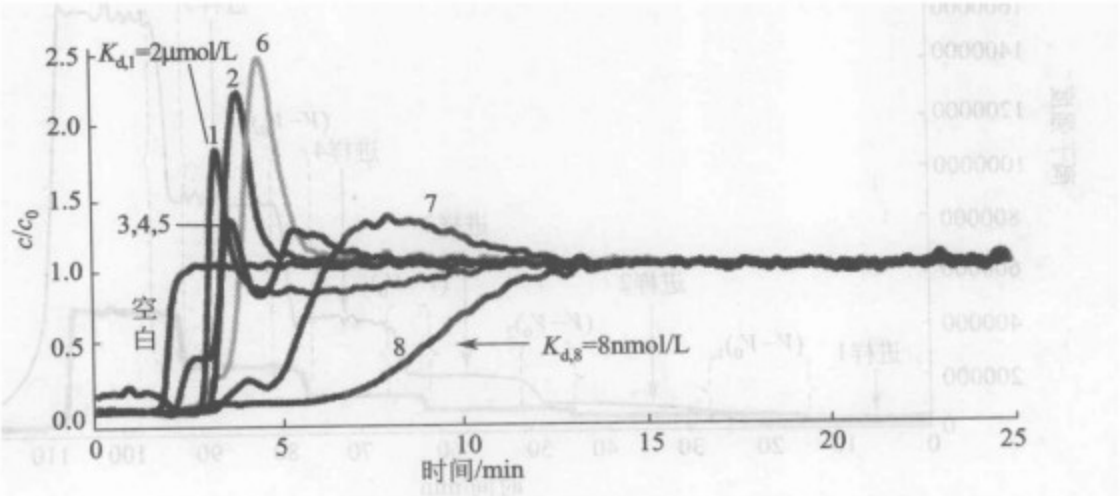


图 7-5 FAC-MS 对配体的分级实验 (线 1~8 分别对应样品 1~8)
SDH 柱, 8 个样品混合进样, 浓度均为 1 mol/L。7 个配体活性 IC_{50} 范围
从 nmol/L 到 mol/L 级, 其流出柱子的顺序符合其与 SDH 的亲合力顺序,
即亲和力弱的先流出, 亲和力强的后流出

7.2.2.3 使用指示分子做竞争性分析

直接地测定化合物并不总是能够做到的, 尤其是当不同配体要求不同的质谱测定条件时, 优化质谱条件, 使所有的化合物都能够被检测到有时是一件困难的工作。

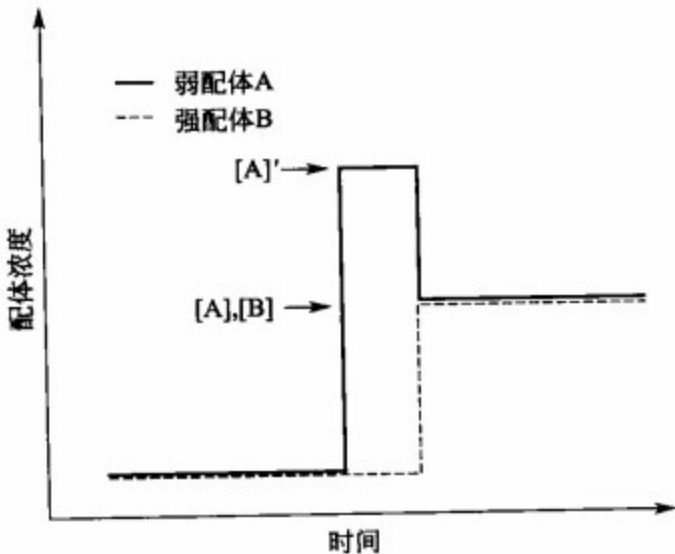
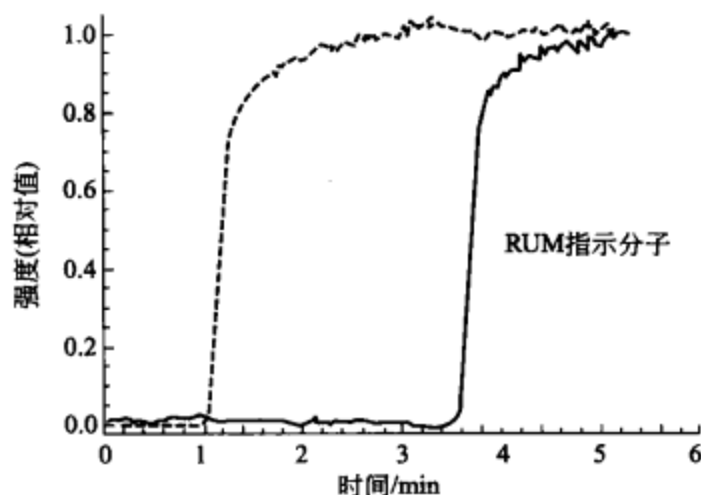


图 7-6 理想的两配体混合物的竞争前沿色谱图
(瞬时过浓度 $[A]'$ 将在强配体 B
流出的同时回到浓度 $[A]$)

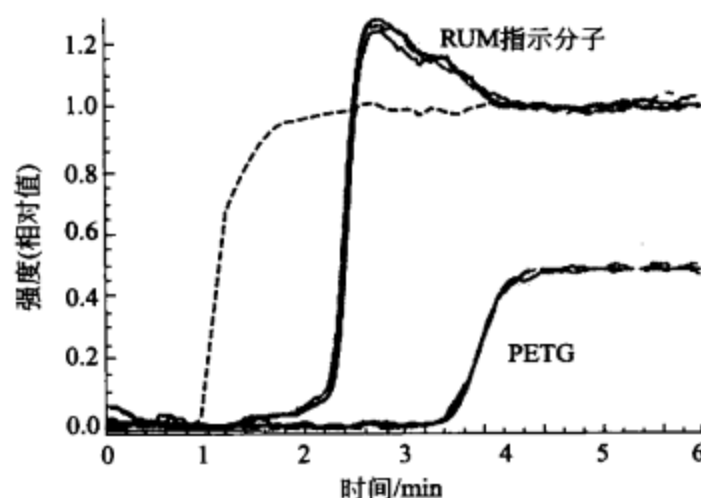
因此, 发展通用的非配体依赖性的操作条件是必要的。这可以采用竞争性的策略来解决, 即使用与待测配体竞争同一活性位点的配体作为指示分子 (indicator), 将指示分子与待测配体混合进样。弱的配体比强的配体更快地通过柱子, 当配体共流出时, 强的配体前沿经过弱配体已经通过的区域时将发生强配体对弱配体的置换, 这种置换增加了弱配体的流出浓度, 引起弱配体暂时性的过浓度, 这种现象被形象地称为 “roll-up” [图 7-1(b) 和图 7-6]。因此, 如果使用一种指示分子, 只有当混合物中存在至少一个配体有更低的 K_d 值时,

才会引起置换, 即 “roll-up”; 多个更弱的配体则不会引起置换, 只会使指示分子的前沿时间缩短。这种方法被称为 RUM 法。N. W. C. Chan 等固定化了 β -半乳糖苷酶 (β -galactosidase), 采用指示分子 (1S, 4S)-4-羟基-2,2-二甲基环戊烷-1-基-

1-硫代- β -D-吡喃半乳糖苷 [(1S, 4S)-4-hydroxy-2,2-dimethylcyclopent-1-yl-1-thio- β -D-galactopyranoside] 验证了这种方法 (图 7-7)。当混合样品中存在强的配体 PETG 时, 指示分子出现了 “roll-up”。



(a) RUM 指示分子与空白指示分子过柱图



(b) 同样的指示分子与空白指示分子和含 PETG 的混合物混合进样。

仅监测指示分子、空白指示分子和 PETG

图 7-7 前沿色谱图

这种方法有两个重要的目的：一是可以减少由于非特异性吸附造成的假阳性；二是减少由于化合物的检测问题造成的假阴性。只要优化指示分子的质谱响应条件就可以了，这对于大量化合物的筛选是非常重要的。因只检测指示分子，快速可靠，且只有存在更强的配体时才需对所筛选的组进行解析，节省了大量时间。当然，选择合适 K_d 值的指示分子也是重要的，在实验中一般采用的指示分子的 K_d 值为 0.5~50mol/L。

7.2.2.4 区分不同的位点

N. W. C. Chan 等制备了 β -丰乳糖苷酶和 GSIB4 lectin 双蛋白柱。采用双指示

分子可以区分与两种蛋白的结合位点结合的化合物。在图 7-8 中，混合物 5 的指示分子 (β -丰乳糖苷酶) 前沿色谱图的“roll-up”显示混合物 5 中存在强的 β -丰乳糖苷酶配体，而没有类似的“roll-up”在其他混合物中或在 GSIB4 lectin 的指示分子前沿色谱图中出现。

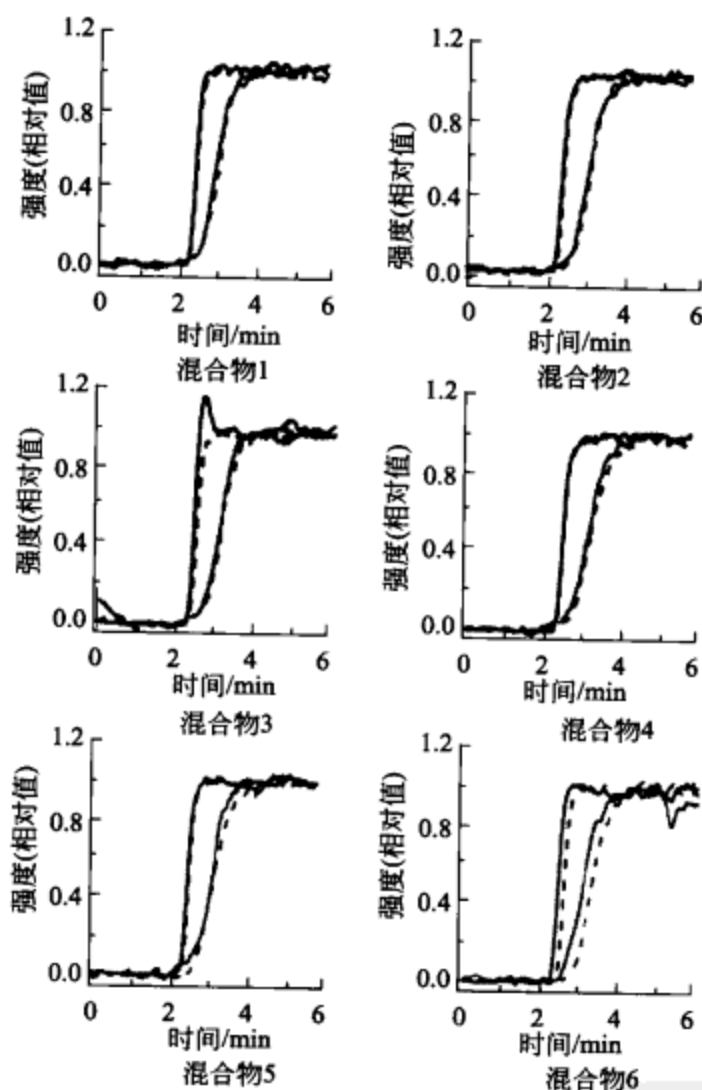


图 7-8 6 根同样的双蛋白柱 (β -丰乳糖苷酶, GSIB4 lectin) 筛选 90 种化合物的混合物。对 β -丰乳糖苷酶的指示分子 (深色线) 和 GSIB4 lectin 的指示分子 (浅色线) 同时记录。点线代表混合物进样前的保留曲线，实线代表混合物进样时的保留曲线。

J. J. Slon-Usakiewicz 固定化 PKC，同时监测蛋白激酶的 ATP 位点和底物结合位点。采用针对这两个位点的指示分子可以区分与这两个位点结合的化合物，如图 7-9 所示。

7.2.2.5 非活性构象

在细胞中，蛋白激酶的活性受到非常复杂的调控。大多数激酶抑制剂的靶标是活性的激酶，目标在于抑制异常的激酶活性。然而，即使在管理良好的系统中也存在活性激酶与非活性激酶的平衡。新型抗癌药格列维克 [Gleevec, 又译“给你活”，商标名为甲磺酸伊马替尼 (imatinib mesylate)] 的作用机制即与此有关，它

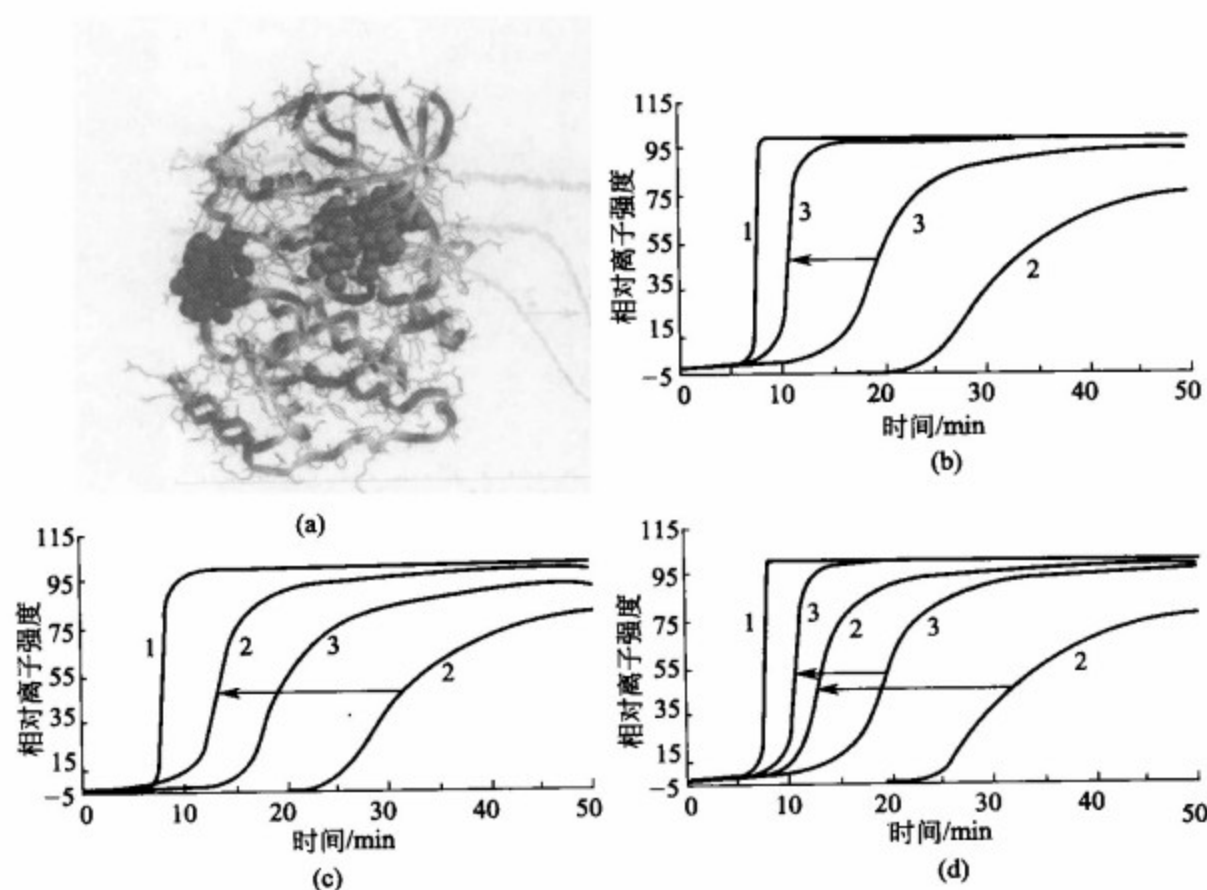


图 7-9 PKC α 双指示分子的 FAC-MS 色谱图

线 1 代表空白指示分子；线 2 相应于底物位点指示分子氯化白屈菜红碱 (chelerythrine chloride) ($1\mu\text{mol/L}$)，线 3 相应于 ATP 位点指示分子 PD153035 ($1\mu\text{mol/L}$) [图 (a)]。箭头显示相关的位移。图 (b) 显示在进样液中加入 WHI-P180 ($5\mu\text{mol/L}$) 后的效果，只有 ATP 位点指示分子 PD153035 从 19min 前移到 10min。图 (c) 显示在进样液中加入 PKC 抑制剂多肽 19-36 ($5\mu\text{mol/L}$) 后的效果，只有底物位点指示分子氯化白屈菜红碱 ($1\mu\text{mol/L}$) 从 30min 前移到 12min。图 (d) 显示在进样液中加入 WHI-P180 ($5\mu\text{mol/L}$) 和 PKC 抑制剂多肽 19-36 ($5\mu\text{mol/L}$) 后的效果，两个指示分子位移的程度与它们单独进样时相同

与靶标 Abl 的结合稳定了 Abl 的非活性构象，从而废除了活化的 Bcr-Abl 蛋白的异常激酶活性。但是在当前的筛选方法都依赖于酶的活性的条件下，对于如何寻找与非活性构象结合的抑制剂，FAC-MS 提供了独到的分析工具。J. J. Slon-Usakiewicz 等测定了与 EphB2 (Eph 为 erythropoietin-producing hepatocellular, 促红细胞生成素产生肝细胞) 受体酪氨酸激酶 (receptor tyrosine kinase) 的活性与非活性形式 (Y604/610F) 激酶的结合。对于活性与非活性形式蛋白，它们与 ATP 结合位点结合的四个化合物有相似的保留时间，这表明活性与非活性激酶的 ATP 结合口袋是相似的，是 Y604/610F 的阳端基环的突变导致了激酶的功能性失活。这是首次确定功能性失活的激酶的结合配体，如图 7-10 所示。

7.2.2.6 药用天然产物的筛选

亲和前沿色谱与高灵敏度高分辨率质谱联用，通过质谱的选择离子流图分析数据，大大提高了检测限 (可以达到质谱的检测限 fmol 数量级) 和灵敏度，可以针

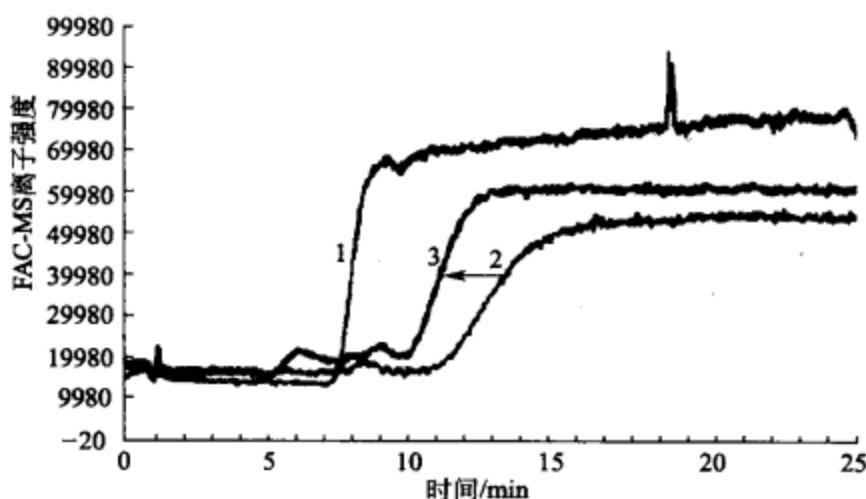


图 7-10 Y604/610F EphB2 对混合物与指示分子 WHI-P180 的前沿色谱图

线 1 相应于空白指示分子。线 2 相应于 WHI-P180 单独进样，

线 3 相应于 WHI-P180 与混合物同时进样。箭头显示位移，表明混合物中至少存在一种化合物与 WHI-P180 竞争这个非活性激酶的位点

对中草药中的每一个组分（包括痕量成分）进行数据分析，从而非常有效地解决了中草药中各个组分之间含量差别大、有些活性成分含量低微的问题。

朱丽荔首次将多克隆抗体亲和色谱应用于中草药体系筛选抗 EGFR 活性成分，将活性化合物白皮杉醇（piceatannol）进行简单的结构改造之后偶联到载体蛋白牛血清白蛋白（BSA）上用于免疫，对得到的多克隆抗体进行纯化，再将其连接到载体 Sepharose CL-4B 上用作亲和色谱固定相，用多克隆抗体模拟靶酶 EGFR，直接从藏药鬼箭锦鸡儿中筛选出 6 个抗 EGFR 的活性成分，保留时间越长的化合物活性越高。骆宏鹏以抗丙肝药物 RD3-4078 为半抗原，连接到载体 BSA 上后免疫兔子，得到的多克隆抗体用 FAC-MS 方法用于中药叶下珠的筛选，各化合物在多克隆抗体柱上的保留与其用蛋白酶直接测活的结果一致。

李正全等以 SARS-CoV 的 S₂ 蛋白为亲和固定相，用 FAC-MS 方法对大量药材粗提物进行了筛选，在短时间内即发现了 2 个能够与 S₂ 蛋白结合，从而抑制病毒感染细胞的抑制剂。

7.2.3 总结与展望

7.2.3.1 FAC-MS 与其他基于直接结合的方法的比较

相对于功能分析方法和基于标记分子的间接结合检测方法，FAC-MS 方法是基于直接结合检测的方法。同时，还有其他一些基于直接结合的检测方法，如表面等离子体共振（surface plasmon resonance, SPR）、核磁共振（NMR）、毛细管电泳（capillary electrophoresis, CE）和自动化配体确认系统（automated ligand identification system, ALIS）。这些方法用于药物筛选时都各有优缺点，其总结见表 7-1。

表 7-1 基于直接结合的筛选平台的比较

能力	FAC-MS (Protana 公司)	NMR	ALIS (NeoGenesis 公司)	CE-Assay TM (Cetek 公司)	SPR (Biacore 公司)
处理量	高(Q1 筛选) 中(指示分子)	中	高	高	低
蛋白质固定化	有	无	无	无	有
化合物分级	有	有	无	有	有
测定 K_d	有	无	有	无	有
可用化学库	无限制	有	受限(要求质量编码)	受限(参数必须优化)	无限制
直接读出(实时)	是	否	否(MS 测定前要求色谱分离)	有	有
其他潜在的筛查应用					
非活性蛋白	有	有	有	有	有
区分多个结合位点	有	有	无	无	无
蛋白-蛋白相互影响	有	有	无	无	有
单个试验中多靶点	有	有	无	无	无

7.2.3.2 FAC-LC/MS

样品的检测对于 FAC-MS 来说是个重要的问题。即使在初级筛选中可以使用指示分子,但在找到活性化合物后仍需确定化合物,而且有些靶点并不存在合适的指示分子,更不用说那些还没有发现已知的结合化合物的靶点了,所以使所有样品都能被检测是很重要的。E. S. M. Ng 等在 FAC-MS 之间增加了 LC(液相色谱)过程,使样品的检测灵敏度和检测条件大大改善。他们的实验装置图如图 7-11 所示。他们固定了 GnT-V,从三糖库中筛选酶抑制剂。FAC 过程采取离线方式,其流出液分部收集、蒸干、重溶于乙腈中,然后这些部分再依次进入 LC 系统,梯度洗脱进入 MS 检测。检测过程以空白的不含酶的 FAC 柱作为参照,并采用半自动的方式,通过比较相应的样品与空白的多维数据集确定结合的化合物。在筛选的 81 个化合物的 A 库中,确定了质荷比为 795 的化合物为强的结合化合物。并且,此方法在筛选 1000 个化合物的 B 库时,对于预先加入的 A 库中的 15 个化合物中的配体,均在筛选中得到了确证。这表明这种方法可适用于大量化合物的筛选,单一的 LC-MS 系统每天可筛选超过 5000 个化合物。

FAC-LC/MS 联用对于天然产物提取物的筛选将是非常有效的,可以进行准确的分子量测定,直接给出化合物的分子式。而且,可进一步使用多级质谱,对未知化合物的结构进行研究。这样,就有可能在一步过程中完成活性化合物的确认和结构解析工作。

7.2.3.3 固定化酶反应器色谱-质谱筛选

Richard J. Hodgson 等采用固定化酶反应器进行基于酶功能的酶抑制剂的筛

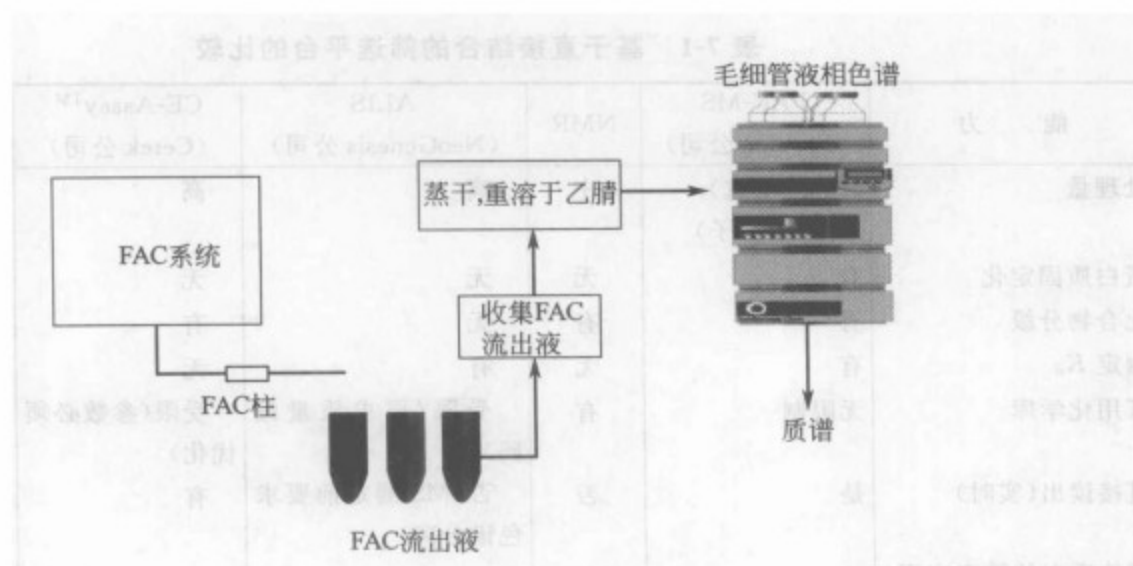


图 7-11 FAC-LC/MS 装置示意

FAC 过程采取离线方式，分部收集、蒸干、重溶。这些部分再进入 LC 系统，
梯度洗脱进入 MS 检测

选。他们用溶胶-凝胶 (Sol-Gel) 法固定了腺苷脱氨酶 (adenosine deaminase)，持续进样底物或底物-抑制剂混合物，直接用质谱检测产物和底物的变化，从而可以计算 IC_{50} 和 K_i 值。这种方法提供了基于酶的功能的酶抑制评估，并且可以直接用于混合物中酶抑制剂的筛选。

7.3 阻断 SARS 侵入细胞的抑制剂的固定化酶亲和前沿色谱-质谱联用法直接筛选

严重急性呼吸道综合征 (SARS) 在 2003 年由世界卫生组织 (WHO) 命名，是一种流行广泛、发病急、传播快、严重威胁患者生命的“非典型肺炎”。据统计，共有 32 个国家和地区报告了 SARS 病例。自 2002 年 11 月爆发开始至 2003 年 5 月 WHO 监测的各国发病总数已经达到 6903 例，其中死亡 495 例，病死率由最初的 3.70% 上升到 7.16%。研究表明，SARS 是一种全新的传染病，SARS 病毒的病原体是冠状病毒的家族变种。到目前为止，治疗 SARS 并无特效药物。北京大学化学学院徐筱杰教授与生命科学学院邓宏魁教授合作，以 SARS-CoV 的 S 蛋白为固定相，采用固定化酶亲和前沿色谱-质谱联用法直接筛选中药粗提物，以期快速找到其有效抑制剂进行抗 SARS 病毒研究。

7.3.1 严重急性呼吸道综合症病毒 (SARS-CoV)

在病毒形态上，SARS-CoV 形态与经典冠状病毒相似，是有包膜的病毒，直径 60~120nm，包膜上花瓣样或纤毛样突起呈放射状排列，长约 20nm 或更长，其底窄，形似王冠。成熟与未成熟的病毒体的大小和形态有很大差异，成熟的病毒呈

圆球形、椭圆形，未成熟病毒可出现很多形态，如肾形、马蹄形、铃铛形、鼓槌形等，在大小上病毒颗粒可从开始的 400nm 减少到成熟后期的 60~120nm。病毒在细胞质内繁殖，利用细胞材料，由 RNA 基因编码的多聚酶进行 RNA 复制和蛋白质合成，组装成新病毒，并出芽分泌到细胞外。

7.3.1.1 SARS-CoV 基因结构

SARS-CoV 为正义单链 RNA，其全长大约 30kb 个核苷酸。不同的病毒株长度有所不同。基因组结构与其他冠状病毒相似：5'端有甲基帽，3'端有 poly (A) 尾，有开放阅读框架 (open reading frame, ORF) (图 7-12)。基因排列顺序为 5'-复制酶 (rep)、刺突蛋白 (spike glycoprotein, S)、包膜蛋白 (envelope protein, E)、膜蛋白 (membrane protein, M)、核衣壳蛋白 (nucleocapsid protein, N)-3'，此外在 5'和 3'末端有小片段的非转录区。SARS-CoV 复制酶基因约占整个基因组的 2/3，由两个 ORF (ORF1a 和 ORF1b) 组成，编码的两个多聚蛋白经裂解后形成 RNA 依赖的 RNA 聚合酶 (POL) 和螺旋酶 (HEL)。在 rep 的下游即后 1/3 的区域有 4 个 ORF，与其他已知的冠状病毒相似，预期可编码结构蛋白，按基因组上的排列顺序为 S 蛋白、E 蛋白、M 蛋白、N 蛋白，未发现 HE (haemagglutinin-esterase, 凝血素酯酶) 蛋白编码序列。冠状病毒还编码一些位于 S 和 E 之间、M 和 N 之间或 N 下游的非结构蛋白，这些非结构蛋白功能不详，并不是病毒复制时必不可少的蛋白。

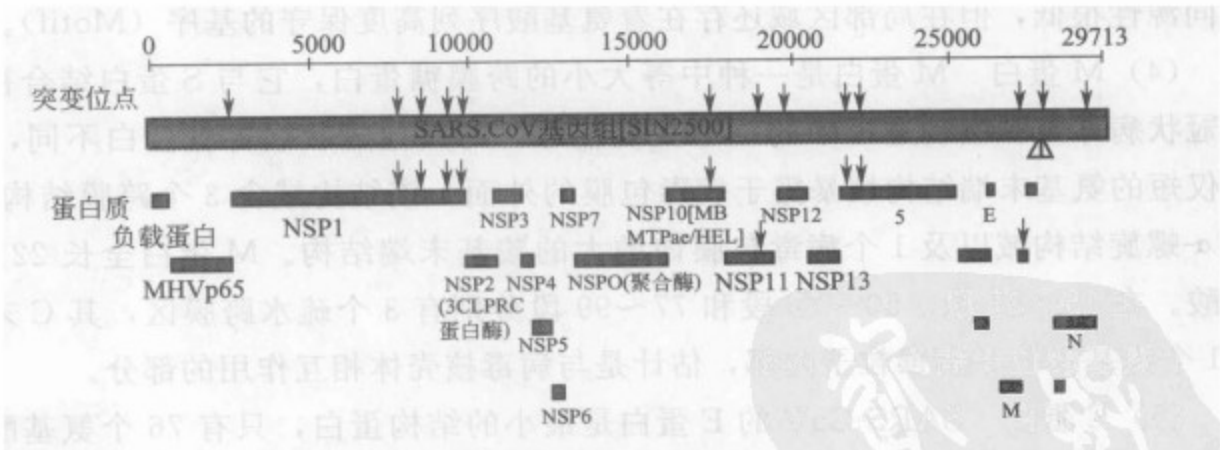


图 7-12 SARS-CoV 的基因结构

7.3.1.2 SARS-CoV 编码的蛋白

SARS-CoV 基因组 5 个主要开放阅读框架 (ORF) 分别编码 RNA 聚合酶蛋白、S 蛋白、E 蛋白、M 蛋白和 N 蛋白。

(1) RNA 聚合酶基因区编码产物 SARS-CoV RNA 聚合酶编码区长 21.2kb，约占全基因组的 2/3。依据已知冠状病毒结构特点，推测该区域有可能表达一个长约 7000 个氨基酸的复合蛋白。在自身催化作用下，经水解可得 14 个未知蛋白 (putative protein)。该复合蛋白前 2/3 与后 1/3 同源性差异显著，推测该复合蛋白

的编码区由两个 ORF 组成, 即 ORF1a 和 ORF1b。其中, 在 13392~13398 位发生了-1 位翻译移框, 因而被分割表达为两个多肽段。在这些 ORF 编码的复合蛋白中, 除了有 RNA 聚合酶核心外, 可能还有与鼠肝炎病毒类似的胰磷乳蛋白酶样结构域、磷酸酯酶结构域和一些疏水结构域。

(2) S 蛋白 S 蛋白即为在病毒粒子表面形成的大的花瓣样突起, 是弱酸性跨膜糖蛋白。作为冠状病毒表面最重要的膜蛋白, S 蛋白全长为 1256 个氨基酸, 由两个结构域 (domain) 组成。S1 部分靠近 N 端, 形成一个球状结构域。S2 部分近 C 端, 形成一个穿膜的棒状结构, 包含一个 N 螺旋 (helix)、一个 M 螺旋 (helix)、一个 C 螺旋 (helix) 和一个穿膜部分。S1 和 S2 之间通过分子间作用力相互结合。S2 的穿膜部分能把整个 S 蛋白固定在病毒外壳膜上。S 蛋白的两个结构域各自具有独特的功能, S1 与病毒和受体的识别结合有关, S2 与病毒外壳膜和宿主细胞膜融合有关。

(3) N 蛋白 N 蛋白是 SARS-CoV 中另一个重要的结构蛋白, 全长 123 个氨基酸。N 蛋白是一种磷酸化蛋白, 处于病毒颗粒的核心部分, 通过与病毒组的 RNA 结合, 可能对 RNA 的转录和病毒形态发生重要作用, 但其确切的生物学功能及结构不清楚。通过对鼠肝炎病毒的研究发现 N 蛋白对病毒基因组 RNA 特征性序列的识别及与其他结构蛋白的相互作用对于病毒颗粒的准确组装具有重要意义。各研究组对 SARS-CoV 和其他冠状病毒的氨基酸序列同源性分析发现, 虽然 N 蛋白整体氨基酸同源性很低, 但在局部区域还存在着氨基酸序列高度保守的基序 (Motif)。

(4) M 蛋白 M 蛋白是一种中等大小的跨膜糖蛋白, 它与 S 蛋白结合被认为是冠状病毒颗粒组装过程中的一个关键信号。与其他冠状病毒糖蛋白不同, M 蛋白仅短的氨基末端结构域暴露于病毒包膜的外面。该结构域含 3 个跨膜结构域、1 个 α -螺旋结构域以及 1 个病毒包膜内的大的羧基末端结构。M 蛋白全长 221 个氨基酸。在 15~37 段、50~72 段和 77~99 段分别有 3 个疏水跨膜区, 其 C 末端的 121 个氨基酸位于病毒粒子内部, 估计是与病毒核壳体相互作用的部分。

(5) E 蛋白 SARS-CoV 的 E 蛋白是最小的结构蛋白, 只有 76 个氨基酸, 主要分布于病毒包膜上, 通过与 M 蛋白起协同作用决定出芽的位置。分析显示其 C 末端和一个 46 个氨基酸的疏水区域位于病毒包膜外。

7.3.2 阻断 SARS 侵入细胞的抑制剂的筛选

由于 SARS 严重威胁患者的生命健康, 快速筛选出抗 SARS 抑制剂对其药物的发展具有重要意义。北京大学化学学院徐筱杰教授与生命科学院邓宏魁教授合作, 以中药材为研究对象, 以 SARS-CoV 的 S 蛋白酶为靶蛋白, 采用亲和筛选和活性测定两步法筛选具有阻断 SARS 侵入细胞活性的抑制剂, 从而阻断病毒对细胞的感染, 达到抑制病毒生长的目的。

7.3.2.1 聚合载体的合成

在可密封的试管中加入多功能单体甲基丙烯酸 (MAA, 4.7mmol, 即 100 μ L)、交联剂三羟甲基丙烷三甲基丙烯酸酯 (TRIM, 24mmol, 2mL) 以及引发剂偶氮 (二) 异丁腈 (AIBN, 20mg), 超声混匀, 除气后充氮气 5min, 密封。在 60 $^{\circ}$ C 恒温水浴下聚合 24h。将聚合物磨碎, 过 500 目 (30 μ m) 筛子, 在丙酮中沉降, 使颗粒尽可能均匀。沉降好的颗粒挥干丙酮, 然后真空干燥。

7.3.2.2 S₂ 蛋白的纯化

S₂ 蛋白为 3.9mL 1.12mg/mL 蛋白溶液, 储存在 0.05mol/L NaHCO₃ 缓冲液 (pH 8.5) 中。用 0.01mol/L NaHCO₃ 溶液 (pH 8.5) 在 4 $^{\circ}$ C 透析 5h。

7.3.2.3 S₂ 蛋白的固相化

载体上的自由的羧基基团可以与蛋白上的氨基反应, 从而使蛋白酶固定在载体上。1.137g 聚合载体用 3mL NaHCO₃ 缓冲溶液 (10mmol/L, pH 8.5) 浸润, 加入 4.8mg EDCI 交联剂, 然后加入 3.9mL S₂ 蛋白, 在室温反应 3.5h (100 μ L 蛋白原液留作浓度测定用)。再用 Tris-HCl (pH 7.4, 25mmol/L) 将没有结合的蛋白洗下来, 得上清液 10ml。测定结合前后的蛋白浓度可以计算出蛋白的结合量。处理好的亲和固定相保存在 4 $^{\circ}$ C 冰箱中, 长期保存则应加入适量叠氮化钠。

用绿色的 PEEK 管 (ϕ 0.75mm $\times$ 50mm) 和两个接头以及不锈钢网 (孔径 10 μ m) 组成体积为 20 μ L 的自组装柱。用注射器把亲和固定相注入其中, 使其填满, 然后用 2mmol/L NH₄Ac 溶液 (pH 7.4) 平衡, 成为可以使用的亲和柱, 即可进行随后的实验。不用时保存在 4 $^{\circ}$ C 冰箱中。

7.3.2.4 样品的准备

121 种中药采用 85% 浸提 2 周, 然后蒸除溶剂, 得到浸膏, 再用 DMSO 配成 10mg/mL 的溶液。每 5 个样品混合, 然后用 2mmol/L NH₄Ac 溶液 (pH 7.4) 稀释成 50 μ g/mL (每个样品 10 μ g/mL)。部分药材的分段部分单独进样。

亲和筛选的装置与图 7-3 相同, 即样品通过亲和柱后对每个化合物计算其亲和前沿时间, 通过其前沿时间比较其与亲和固定相的结合能力。

7.3.2.5 结果和讨论

(1) 蛋白含量测定 表 7-2 为工作曲线数据, 图 7-13 为蛋白浓度测定工作曲线。

表 7-2 工作曲线数据

V(1mg/mL BSA)/ μ L	20	40	60	80
A _{595nm} (吸光度)	0.102	0.207	0.294	0.366

工作曲线 $Y=BX$, 其中 B 等于 0.00477, 而 R 为 0.996。取 40 μ L 蛋白原液,

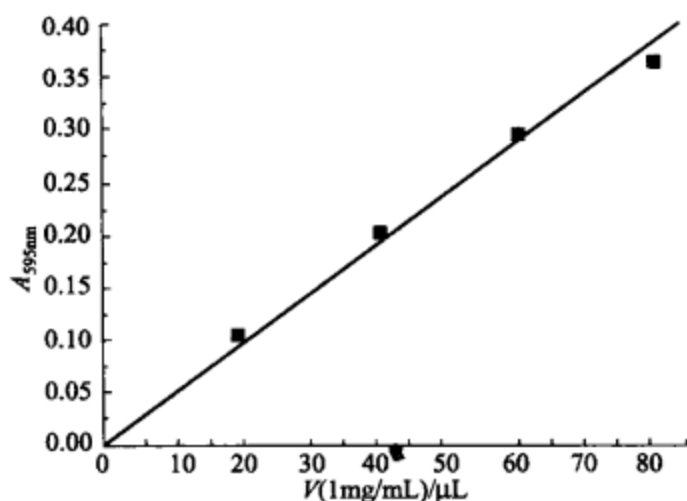


图 7-13 蛋白浓度测定工作曲线

测得其 $A_{595\text{nm}}$ 为 0.219; 100 μL 上清液, 其 $A_{595\text{nm}}$ 为 0.087, 计算表明其结合量为 2.3mg/g 载体。

(2) 亲和筛选 以 SARS 的 S_2 蛋白为亲和固定相, 通过对近百种药材的筛选, 寻找具有较长保留时间的化合物。图 7-14 为五倍子正丁醇部分中 10 个化合物的质谱图, 图 7-15 为其对应的亲和前沿色谱图。用与以前同样的方法算出其前沿保留体积 (见表 7-3)。发现五倍子中

有化合物 10 具有很长的保留时间和保留体积, 说明其与蛋白的结合能力也强。通过常规方法得到了化合物 10 的单一组分。对化合物 10 进行测活, 首先测定了其假病毒活性, 发现有明显的抑制病毒侵入细胞的活性, EC_{50} 为 2.86 $\mu\text{mol/L}$ 。而抗野生型病毒活性显示化合物 10 同样具有较高的活性, 其 EC_{50} 为 4.5 $\mu\text{mol/L}$ (见表 7-4)。

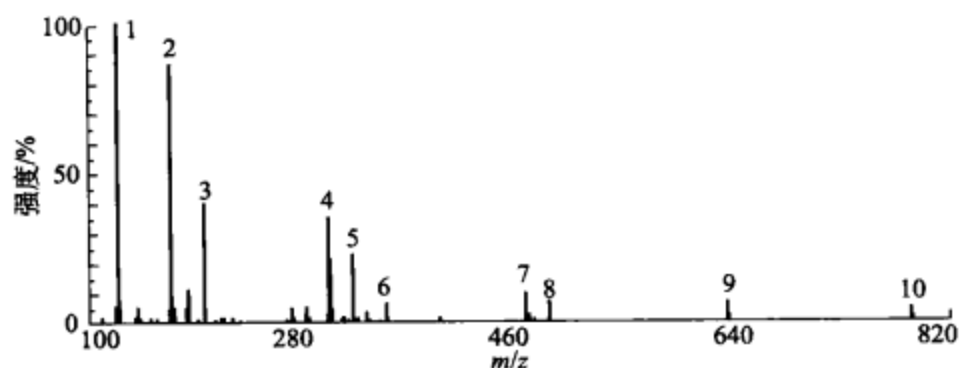


图 7-14 五倍子正丁醇部分质谱图

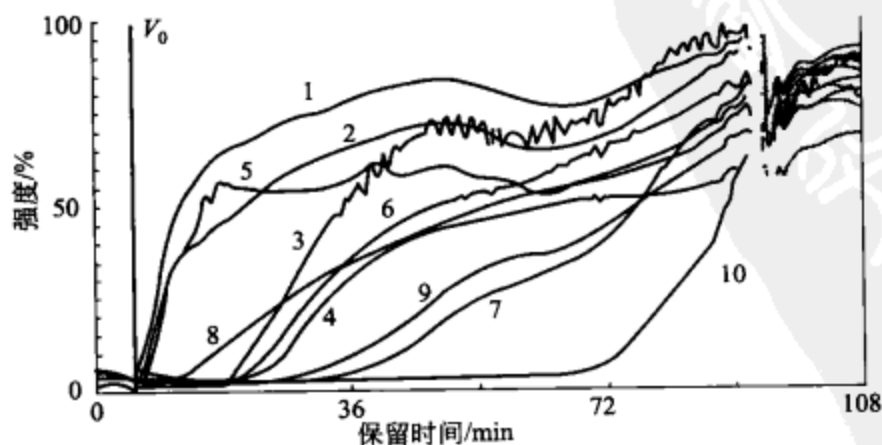


图 7-15 图 7-14 中 10 个化合物的相应的亲和前沿色谱图

表 7-3 五倍子中 10 个化合物的亲和前沿体积

成分	前沿体积/ μL	成分	前沿体积/ μL
1	74	6	233.5
2	101.5	7	345.5
3	183	8	180
4	243	9	340.5
5	49	10	432.5

表 7-4 木樨草素和四 (O-没食子酸) - β -D-葡萄糖苷 (TGG) 抗野生型 SARS 感染的活性数据

化合物	$\text{EC}_{50}/(\mu\text{mol/L})$	SI	$\text{CC}_{50}/(\text{mmol/L})$	$\text{LD}_{50}/(\text{mg/kg})$
TGG	4.5(1.96~5.8)	240.0	1.08	232.2
木樨草素	10.6(2.64~3.5)	14.62	0.155	456

表 7-5 不同浓度化合物 10 在 S_2 亲和柱上的保留体积

$c/(\mu\text{g/mL})$	$c/(\mu\text{mol/L})$	$c/(\mu\text{mol/L})^{-1}$	$(V-V_0)/\mu\text{L}$	$c(V-V_0)/\mu\text{mol}^{-1}$
20	25.4	0.0394	191.5	206
10	12.7	0.0787	322.5	244
5	6.35	0.157	540.0	291

采用类似的方法还发现木樨草素 (luteonlin) 同样具有很好的抗病毒活性, 假病毒活性为 $9.02\mu\text{mol/L}$, 抗野生型 SARS 感染的 EC_{50} 为 $10.6\mu\text{mol/L}$ (见表 7-4)。利用亲和前沿色谱的方法测定了化合物 10 与 S_2 蛋白的结合能, 见表 7-5 和图 7-16, 测得的 K_d 值为 $3.87\mu\text{mol/L}$ 。用同样的方法还测定了木樨草素与 S_2 蛋白的 K_d 值为 $6.88\mu\text{mol/L}$ 。化合物 10 与木樨草素的结构见图 7-17。

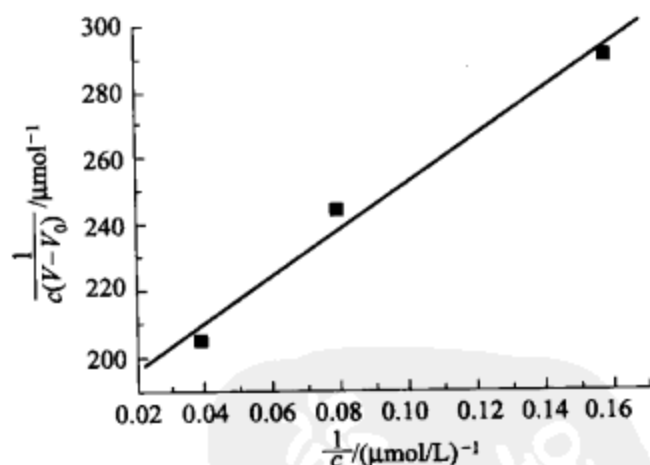


图 7-16 利用 FAC-MS 测定柱容和四 (O-没食子酸)- β -D-葡萄糖苷 (TGG) 的配体解离常数, 对实验数据进行线性回归分析: $Y=A+BX$,

$$A=182\pm 9, B=705\pm 91, R=0.992,$$

$$K_d=B/A=3.87\mu\text{mol/L}$$

对这两个化合物的细胞毒性和动物毒性均进行了评价, 表明这两个化合物可以使用比较高的剂量而不会引起细胞中毒和动物死亡。与报道的甘草甜素相比, 这两个化

合物在体外抑制病毒感染活性要高两个数量级。由于结构的类似性, 对槲皮素的假病毒活性也进行了测定, 其 EC_{50} 为 $83.4\mu\text{mol/L}$ 。

7.3.3 小结

以 SARS 的 S_2 蛋白为亲和固定相, 首先采用亲和色谱-质谱联用法对大量药材

或化合物进行筛选，找到与 S₂ 蛋白具有很好结合的化合物，然后进行假病毒活性测定，进而进行野生型病毒感染的测试，在一两个月的时间内即发现了 3 个能够与 SARS 的 S₂ 蛋白结合从而能抑制病毒感染细胞的抑制剂，具有很好的抗病毒活性。这也从另一个方面说明亲和前沿色谱用于筛选中药体系中是可行的。

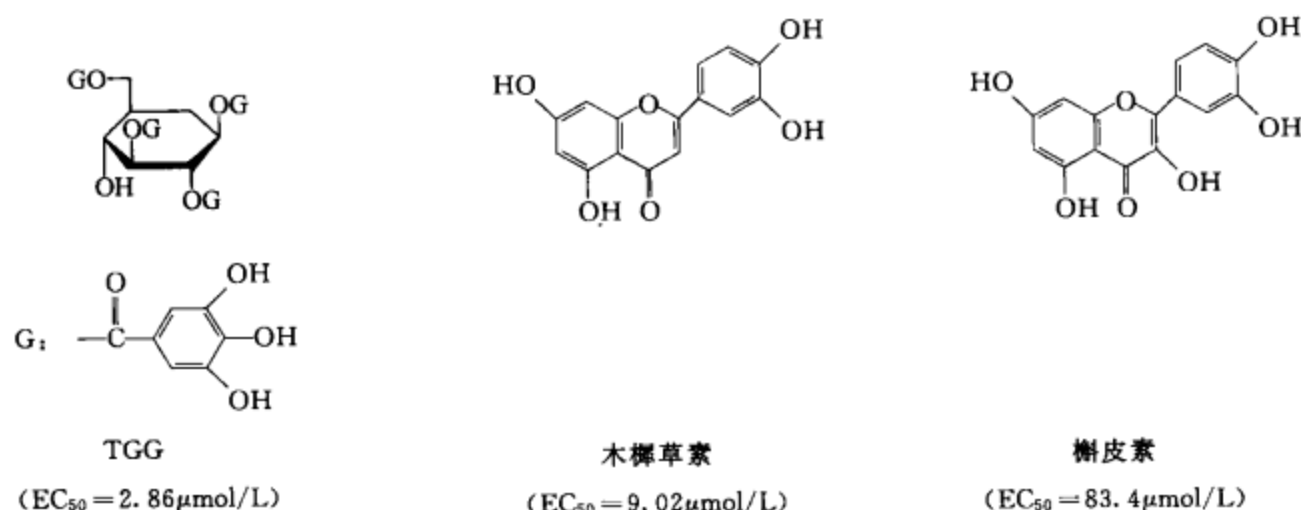


图 7-17 3 种具有抗病毒活性的化合物的结构

7.4 固定化多克隆抗体亲和前沿色谱-质谱联用法直接筛选

7.4.1 多克隆抗体的应用背景

根据对小白鼠的研究推测动物体内含有几亿个 B 细胞，在免疫原的刺激下，可产生几千万个具有免疫活性的 B 细胞系（克隆型）。其中，平均每 10000 个 B 细胞系中约有 1 个能产生针对特定抗原决定簇的抗体，因此每免疫一种动物可有几万个细胞系能产生针对各种抗原决定簇的抗体。像这样通过直接免疫动物，然后从血清（哺乳类）和卵黄（禽类）中分离得到的抗体，称为多克隆抗体（多抗，PcAb）。与单克隆抗体（单抗，McAb）比较，多克隆抗体的理化性质比较多样化，特异性没有单抗强，且能与同等纯度免疫原发生交叉反应。这里用具有抗 EGFR 的活性分子作为半抗原制备其相应的多抗，根据构效关系可知一个针对某种酶的抑制剂与酶之间有强的分子间相互作用，包括静电相互作用、范德华相互作用和疏水相互作用等，因此该活性分子（有效抑制剂）的抗原决定簇能部分反映酶的“活性口袋”（binding pocket）特征，从而可以利用多克隆抗体来模拟酶，直接从植物的粗提物里识别半抗原及其结构类似物。多克隆抗体的制备比较简单，周期也较短。

7.4.1.1 抗原与抗体

能刺激机体免疫系统产生特异性免疫应答，并能与相应的免疫应答产物即抗体和致敏淋巴细胞在体内外发生特异性反应的物质，称为抗原（antigen, Ag）或免疫原（immunogen）。前一种性能称为免疫原性（immunogenicity）或抗原性（an-

tigenicity), 后一种性能称为反应原性 (reactogenicity) 或免疫反应性 (immunoreactivity)。凡具有这两种性质的物质称为完全抗原 (complete antigen), 只具有反应原性而无免疫原性的物质称为半抗原 (hapten) 或不完全抗原 (incomplete antigen)。各种抗原刺激机体后, 可产生针对这些抗原的各种特异性抗体。虽然抗体种类很多, 但其化学本质是一致的, 均为球状糖蛋白, 即 γ -球蛋白 (gamma-globulin, γ G)。 γ G 的分子量在 150000~1000000 之间。由于一部分 γ G 不具抗体特性, 为了便于区分, 通常将具有抗体活性的 γ G 称为免疫球蛋白 (immunoglobulin, Ig)。

Ig 可从抗血清中分离得到, 而且从各种抗原的抗血清中分离得到的 Ig 化学性质均相同。这表明同是 Ig 但可能具有不同的抗体性质, 可与不同的抗原特异性结合。Ig 有多种类型, 但其基本结构 (单价) 均由两条重链 (heavy chain, H 链) 和两条轻链 (light chain, L 链) 组成, 经二硫键连接而呈 Y 形, 其结构示意图可用图 7-18 表示。

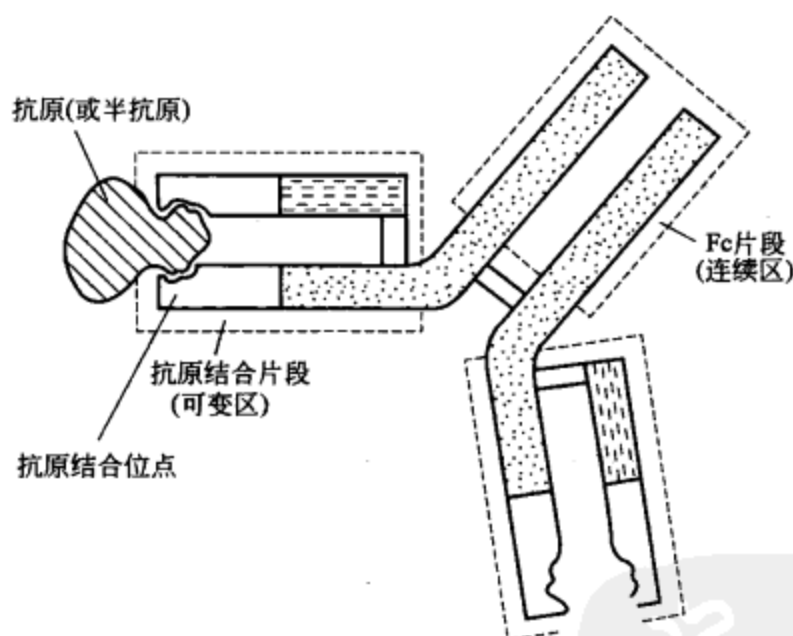


图 7-18 免疫球蛋白 (Ig) 的结构

重链和轻链的可变区 (variable region, Fab fragment) 构成抗体分子的抗原结合位点, 称为配位或互补位 (paratope)。抗体的配位具有针对相应半抗原或抗原决定簇 (表位, epitope) 的独特氨基酸序列和空间构型, 配位和表位在结构和性质上有高度互补性, 如电子云分布、构象等, 犹如“锁与钥匙”的关系, 构成了抗原与抗体特异性结合的分子基础。

7.4.1.2 完全抗原与半抗原

小分子化合物 ($M_w < 2500$) 一般不具备免疫原性, 即不能刺激动物的免疫应答反应而产生抗体, 因而通常称不具有免疫原性的小分子为半抗原。但小分子化合

物具有反应原性，即能与相应的抗体发生特异性结合反应。抗原与抗体的结合主要依靠抗原与抗体间结合位点相关空间结构的精密契合和范德华力等的维系。这同酶与抑制剂的结合机理类似，故对酪氨酸蛋白激酶——表皮生长因子受体（EGFR）具有较高抑制活性的小分子所产生的抗体很大程度上反应了酶活性位点的信息。但是目前发现的抗 EGFR 的抑制剂分子量均小于 1000，不具有免疫原性，必须与相应的载体结合才能产生免疫原，即完全抗原，用于制备抗体。

7.4.1.3 载体

蛋白质是一种大分子，其结构复杂，并具有胶体的特质，是一种良好的半抗原载体。不同动物的蛋白可以刺激另一种动物的免疫系统，产生强烈的免疫反应。牛血清白蛋白（bovine serum albumin, BSA）、卵清白蛋白（OVA）、肌红蛋白、钥孔血蓝蛋白（keyhole limpet hemocyanin, KLH）以及破伤风类毒素等是常用的载体蛋白。这些载体的免疫原性较强，且易获得。一般分子量大一些、来源于免疫动物的亲缘关系远一些的蛋白免疫效果较好。常用的这几种载体蛋白之间并无交叉反应性。用一种蛋白作为载体免疫动物所产生的抗体如果能够识别用另一种蛋白作为载体的抗原，则可以判断产生了针对小分子半抗原的特异性抗体。

7.4.1.4 半抗原与载体的连接

动植物体内含有的生物活性物质，除一部分为蛋白质外，大多数为肽类（氨基酸）、酯类及其衍生物、黄酮类、异黄酮类、萜类、甾体类等小分子物质（半抗原）。这些小分子物质尽管具有反应原性，可与其特异性抗体结合形成半抗原-抗体复合物，但是小分子本身并不具有免疫原性，不能直接刺激动物机体产生特异性抗体。因此，为了用从中草药中直接提取的活性化合物制备其多克隆抗体，必须将半抗原与大分子物质偶联。可选择牛血清白蛋白（BSA）为载体。

原则上用于合成肽连接的偶联剂及连接方法亦可用于半抗原与载体的连接。常用的偶联方法有戊二醛法、碳二亚胺法、活泼酯法、亚胺酸酯法和卤代硝基苯法等。这些偶联方法使半抗原与载体在 $-COOH$ 、 $-NH_2$ 、 $-SH$ 等基团部位发生结合。对于不具备 $-COOH$ 、 $-NH_2$ 、 $-SH$ 等基团的半抗原，需要加以改造，使其转变为带有这些基团的衍生物后，方可偶联制备完全抗原。

7.4.1.5 完全抗原的评价

完全抗原的评价包括两个方面：一方面是蛋白的浓度，它反应抗原制备过程中蛋白的损失情况；另一方面是一个载体分子结合上多少个小分子半抗原。蛋白的浓度一般用紫外分光光度法估测，将抗原稀释到适当浓度，在 260nm 和 280nm 分别测定紫外吸收的 OD 值，按下式计算蛋白浓度：

$$\text{蛋白浓度 (mg/mL)} = 1.45A_{280} - 0.74A_{260}$$

一般判断肽与载体蛋白是否结合、结合比是多少，是通过氨基酸组分的分析结果计算的，方法较麻烦。载体结合了小分子之后，分子量必然增加，通过测定大分

子的分子量可以直接判断连接效果。本实验首次采用先进的质谱技术，直接对反应前后大分子的分子量进行测定，可以快速地判断偶联效果。如果小分子成功偶联到大分子上，在 MALDI-TOF 质谱图上表现在两个方面：一方面是分子量增加；另一方面是分子量的分布加宽，这是由于一个大分子可与多个小分子结合，能结合上小分子的个数是不确定的，导致分子量的分布加宽。如 BSA 有 59 个氨基，如果平均与 10 个小肽连接，那么有的蛋白分子结合 10 个，也有的结合 7 个、8 个、9 个或 11 个等，用 MALDI-TOF 测定反应后 BSA 的分子量按峰高最高点计算。半抗原与载体蛋白的结合比（每分子载体上连接半抗原的数目）按下式计算：

$$\text{结合比} = \frac{\text{反应后 BSA 的分子量} - \text{反应前 BSA 的分子量}}{\text{每结合一个半抗原分子所增加的质量}} \quad (7-5)$$

偶联后载体与半抗原的摩尔比以 1 : (10~25) (BSA) 为宜，结合太多的半抗原会导致免疫耐受。

7.4.1.6 抗体的制备

目前人工制备的特异性抗体分为 3 种类型，即多克隆抗体 (PcAb)、单克隆抗体 (McAb) 和基因工程抗体 (GeAb)。

(1) 多克隆抗体 (PcAb) 的制备 多克隆抗体是动物血液中多个淋巴细胞系分泌的不同选择性和亲和性抗体的混合物，可由抗原免疫动物（多为家兔）获得。多克隆抗体容易获得，均一性较差。动物机体接触到抗原刺激后，要经过致敏阶段、反应阶段和效应阶段产生抗体。机体初次接触某种抗原后，需经一定的潜伏期才能产生少量的抗体。为了提高抗体的效价，必须用相同的抗原再次免疫动物。再次免疫的抗原激活初次免疫产生的记忆细胞，产生大量的抗体，且所需的时间缩短。适当地使用佐剂可以提高动物机体对免疫原刺激的应答能力。制备高免血清常用的佐剂为福氏佐剂 (Freund's adjuvant)，能引起机体局部形成肉芽肿，起到富集巨嗜细胞和免疫活性细胞的作用。初次免疫使用的福氏完全佐剂除了含有不完全佐剂的石蜡油和羊毛脂外，还含有灭活的结核分枝杆菌来进一步刺激机体免疫系统。加强免疫使用不完全佐剂，以防止动物产生大量针对结核杆菌的抗体。

(2) 单克隆抗体 (McAb) 的制备 单克隆抗体是由单一的 B 淋巴细胞系（称为一个克隆）分泌的抗体，针对一个抗原决定簇或半抗原表位，其在选择性、亲和性及理化性质方面都是高度均一的。制备单克隆抗体动物多选用 Balb/c 小鼠，骨髓瘤细胞亦来自 Balb/c 小鼠。免疫动物的脾细胞（含许多 B 淋巴细胞克隆）可分泌特异性的抗体，但在体外培养中不能无限繁殖而很快死亡。骨髓瘤是已失去分泌能力的 B 淋巴细胞，具有体外培养中无限繁殖的能力，将脾细胞与骨髓瘤细胞融合形成杂交瘤细胞，兼具有分泌抗体和体外无限繁殖的能力，在适当的选择性培养条件下将分泌特异性抗体的杂交瘤细胞加以筛选和克隆，再把适宜的克隆扩增后用来进行单克隆抗体的生产。

(3) 抗体库技术 (antibody library technique) 抗体库技术的主导思想是将某种动物的所有抗体可变区基因克隆在质粒或噬菌体中表达, 利用不同的抗原筛选出携带特异性抗体的克隆, 从而获得相应的特异性抗体。从 Ig 种系基因的多样性、种系基因重排和体细胞突变等来推算, 一个动物个体可以产生 $10^6 \sim 10^8$ 种针对不同抗原的抗体。抗体库的最基本的一步就是把所有这些抗体基因扩增出来, 尽量保持这种多样性, 然后通过错配 PCR、半合成的互补决定性区 (complementary determinant region, CDR) 等方法模拟动物体内亲和力成熟过程, 获得高亲和力抗体。

在获得抗体的 3 种途径中, 多克隆抗体获得较快, 操作成熟简便, 但由于均一性差, 在利用抗体的交叉抗性进行药物筛选时可能出现重现性差、干扰信息多等现象。因此, 抗体的纯化很重要, 一般的饱和硫酸铵沉淀法和 DEAE 纤维素柱纯化法不能满足要求, 有必要进行亲和色谱纯化。尽管单克隆抗体的制备初期要比多克隆抗体复杂得多, 但重现性和可比性要好得多, 有利于制备高容量的亲和色谱柱, 为药物筛选中首选的方法。与单克隆抗体和多克隆抗体相比, 基因工程抗体除了具有单克隆抗体的优点外, 还有一个突出的优点是无须免疫, 利用抗原即可直接从非免疫动物抗体库中筛选出特异性抗体。这就避免了由半抗原制备完全抗原时半抗原的构象可能造成的改变, 所产生的抗体更能反映小分子整个分子的特征, 使得所产生的抗体更接近酶的活性位点。对那些缺乏合适的基团用于与载体蛋白连接的小分子和需要选择性保护活性基团的小分子, 基因工程抗体技术可以省去烦琐的小分子改造过程。

7.4.1.7 抗体的鉴定

表征抗体效价常用的方法有琼脂双向扩散法和酶联免疫吸附测定法 (ELISA)。当抗原与抗体相遇, 并达到一定的比例时, 可以形成大的复合物, 产生沉淀反应。这一反应可以在琼脂板上进行, 将抗原加到中间孔中, 抗体倍比稀释加到边缘的孔中, 抗原-抗体在琼脂中可自由扩散, 相遇后产生白色沉淀线, 可以根据在多大稀释倍数不再产生沉淀线判断抗体的效价。该方法经典, 操作容易, 但灵敏度偏低, 可用于免疫动物后粗略估计抗体产生的情况。

ELISA 灵敏度极高, 是常用的评价抗体的方法。ELISA 法有很多种形式, 间接 ELISA 是测定抗体效价的主要形式。将抗原包被到 96 孔微量滴定板的孔中, 洗掉多余抗原, 将抗体或抗血清系列稀释, 与抗原反应, 洗掉多余抗体, 加入酶标二抗与抗原抗体复合物结合, 洗掉多余二抗, 加入酶的底物, 测定酶促反应后的颜色变化。颜色越深, 说明抗体的含量越多。在检测小分子抗体时要注意, 动物产生的抗体有一部分是针对载体的, 如果用免疫原作为包被抗原, 所测出的抗体效价并不直接反映识别小分子的情况。如果小分子是多肽 (通常为含多于 8 个氨基酸的肽), 可以直接包被; 如果是分子量较小的多肽或有机小分子, 通常将小分子连接到与载体蛋白没有交叉反应性的其他蛋白上作为检测抗原。如 BSA 作为载体与小分子相联免疫动物, 可以用 OVA 与该小分子相联作为包被抗原。

7.4.1.8 抗体的纯化

抗血清中含有大量的杂蛋白，需纯化和浓缩才能用于亲和色谱的研究。对于兔血清中 IgG 的纯化，一般用饱和硫酸铵（SAS）对 IgG 进行粗提和浓缩。向抗血清中加入 SAS，使其达到一定的饱和度，由于盐析的作用，一些蛋白可以产生沉淀，从而与不形成沉淀的蛋白分开。IgG 在 SAS 大于 35% 时可以沉淀出来，可以与其他蛋白分开，粗提物经 PBS 透析后可用离子交换法进一步纯化。

用亲和色谱进行药物筛选时，是利用抗体的交叉反应性。即使是纯净的 IgG，也含有大量的无关抗体，尤其是抗 BSA 的抗体。为了避免无关抗体带来的干扰信息和提高柱容量，还需进行亲和色谱纯化。免疫吸附剂由抗原与基质（matrices）交联而成。通常抗原与基质连接需要一定的间隔臂以防止立体结构的干扰，尤其是抗原的分子量很小时。用于纯化抗体的免疫吸附剂有两类，一类是针对抗体的 Fc 端，如葡萄球菌 A 蛋白（SPA）能与多种哺乳动物 IgG 分子的 Fc 段结合，将 SPA 连到载体 Sepharose CL-4B 上，可将血清中的 IgG 纯化。但是，这种纯化出的 IgG 抗体可能是针对多种抗原，如不能去除抗载体的抗体。亲和纯化的另一种形式是用抗原本身直接提纯抗体，这种形式得到的抗体针对性强、纯度高，缺点是要每次制备抗原和载体连接物，且抗原不同连接方法也可能不同。有些很小的分子即使加上连接臂也不能纯化抗体，或小分子的结构特殊，很难连到柱上，这时可以将载体连到柱上，选择性地去去除抗载体的抗体。

7.4.2 固定化多克隆抗体亲和前沿色谱-质谱联用法直接从鬼箭锦鸡儿中筛选 EGFR 抑制剂

7.4.2.1 表皮生长因子受体研究进展

表皮生长因子受体（epidermal growth factor receptor, EGFR）是一种跨膜糖蛋白，由细胞膜外的配基（表皮生长因子）结合区、疏水跨膜结构域和细胞膜内的激酶区三部分组成，总分子量约为 170000。EGFR 作为酪氨酸蛋白激酶家族中的一员，具有蛋白激酶的共有特性。它与细胞外部的生长因子（EGF）结合后导致细胞膜内的激酶区被激活，可以将 ATP 上的 γ -P 转移到底物上，从而使自身磷酸化或催化外源性底物的磷酸化。这种酪氨酸磷酸化作用与许多细胞调节过程有关，包括 T 细胞和 B 细胞的激活、有丝分裂、分化和发育，对外界刺激的反应，细胞的存活及生长状况的调节等。因此，不难理解，一旦这种酪氨酸激酶的功能受阻，将会导致各种各样的疾病。研究表明，当 EGFR 表达受阻或者过度表达时，细胞的生长调节失控，始终处于增生状态，从而可能发展成为恶性肿瘤。许多临床研究表明，在一些乳腺癌、卵巢癌、膀胱癌、肺部和口腔的鳞状细胞癌以及食道癌中都有高水平的表皮生长因子受体，而且该受体与病理变化和临床表现均有一定联系，这为用抗 EGFR 的抑制剂治疗某些癌症提供了实验依据。

到目前为止,酪氨酸蛋白激酶及其参与的信号传递过程被作为相关药物设计的重要靶点广泛研究。酪氨酸蛋白激酶 (protein tyrosine kinase, PTK) 是很庞大的一类蛋白质超家族,包括不带受体 (non-receptor tyrosine kinase) 的 PTK 和带有跨膜受体 (receptor tyrosine kinase) 的 RTK。在 RTK 中,研究最为广泛一类就是 EGFR 或 erbB,文献已经报道了多类 EGFR 的抑制剂。在过去的几年中,其他几种 PTK 也引起了人们的重视,其中包括 PDGFR (platelet-derived growth factor receptor, 血小板衍生生长因子受体)、FGFR (fibroblast growth factor receptor, 成纤维细胞生长因子受体)、VEGFR (vascular endothelial growth factor receptor, 血管内皮生长因子受体) 以及 IGF-1R (insulin-like growth factor-1 receptor, 胰岛素样生长因子 I 型受体)。在 PTK 中,研究最为广泛的就是 c-Src [cytoplasmic protein (细胞浆蛋白) p60^{c-Src}]。图 7-19 显示了几种有代表性的激酶的结构域。对于非受体的 PTK,细胞质信息传导区,如 SH2 区,是这一类蛋白激酶的保守性较强的结构域。而对于包含跨膜受体的 RTK,其结构域可以分为 3 部分:配体结合的膜外区、疏水跨膜区以及细胞内激酶区。尽管这些不同的激酶的一级结构不尽相同,但是它们的活性催化区域或者激酶区是高度保守的。激酶区可以分为 12 个结构子域,折叠成为两个区域:较小的 N 端部分由 I ~ IV 子域组成,包含 Mg-ATP 结合位点;较大的 C 端由 VI ~ XI 子域组成,包含蛋白质或多肽底物结合位点。这两个部分通过子域 V 相连。

表皮生长因子受体属于 RTK I 型 4 种酪氨酸蛋白激酶中一种,这 4 种酪氨酸蛋白激酶包括 EGFR 激酶 (c-erbB-1 的基因产物)、p185 (c-erbB-2 的基因产物) 以及两种最近确定的 c-erbB-3 和 c-erbB-4 的基因产物。它们通常通过二聚变成具有活性的酪氨酸激酶。这种聚合包括自身之间的同类二聚以及和其他家族成员之间的杂交二聚。EGF 以及 TGF- α (transforming growth factor- α , 转化生长因子 α) 是 EGFR 最常见的配体。其他的配体还包括 AR (amphiregulin, 双调蛋白)、EPR (epiregulin, 上皮调节蛋白)、cripto 蛋白、BTC (betacellulin, β -细胞调节素) 以及 HRG (heregulin, 调蛋白)/NRG (neuregulin, 神经调节蛋白)。在正常细胞环境下,这些配体可以导致 EGFR-EGFR 以及 c-erbB-4-c-erbB-4 的同类二聚以及其他 6 种可能的杂交二聚。

EGFR 为跨膜糖蛋白,根据它在细胞内外的分布及功能,可将它分为细胞外、跨膜和细胞内 3 个不同的结构域,现分别讨论。

从氨基端开始的 621 个氨基酸残基,与 N-连接的糖链一起,组成了对 EGF 的结合部位,位于细胞外。该结构域有下面两个明显的特点。

① 它含有 12~16 个 N-连接的糖基化部位,其中多数已被修饰。这些寡聚糖链约有 1/3 为甘露糖。糖基化的结果是使受体的分子量增至 170000,而由氨基酸组成计算所得分子量只有 134300。高含量甘露糖糖链的作用,可能是通过影响受

用于形成二硫键。但是，在无还原剂的条件下进行的有限蛋白质水解实验表明，两个区段之间无二硫桥键。明确地说，EGFR 有两个细胞外结构域，各自通过域内的二硫键折叠并形成高度稳定的结构（图 7-20）。

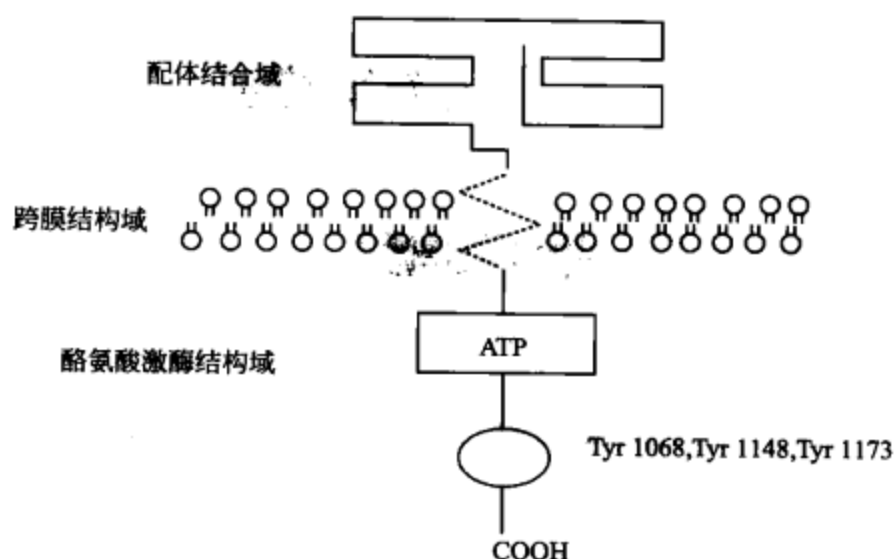


图 7-20 表皮生长因子受体 (EGFR) 的结构

EGFR 的跨膜区由 23 个氨基酸残基组成，细胞内结构域由 542 个氨基酸残基组成。按照功能划分，EGFR 的胞浆内结构域也可分为几个功能不同的区段。在细胞质膜的内侧，有 8 个精氨酸残基和 1 个赖氨酸残基，组成一个“传递停止”信号。在质膜与胞浆交界的第 654 位氨基酸处为一苏氨酸残基，该残基可被蛋白激酶磷酸化，从而降低受体对配基的亲合力，并能稍微减弱 EGF 激活的激酶活性。靠近 C-末端的第 690~940 位残基形成了另一功能区段。它们与 *src* 基因总体的顺序有相当高的同源性。该区段实际上是受体的激酶结构域，其中第 695~700 位的 Gly-X-Gly-X-X-Gly 和第 719~721 位的 Ala-X-Lys 顺序在所有具有酪氨酸蛋白激酶活性的分子中都是保守的，这一段顺序负责与细胞内的底物蛋白特异的相互作用。在细胞内结构域的 C-末端处有 3 个 EGF 诱导的自身磷酸化位点：第 1068、1148、1173 位的 3 个酪氨酸残基。由分别含有这 3 个位点的合成肽诱导产生的抗体能抑制 EGF 诱导的受体自身磷酸化作用，但对受体磷酸化外源底物的能力没有影响。在这 3 个位点中，越接近 C-末端者，自身磷酸化的程度越高。

氨基酸顺序分析表明，EGFR 胞浆内结构域与鸟类成红细胞增多症病毒 (avian erythroblastosis virus, AEV) 的转化蛋白 *erbB* 的顺序有高度同源性：在所定序的表皮生长因子受体的 83 个氨基酸残基中，有 74 个与由 AEV *erbB* 癌基因编码的转化蛋白相同。实验表明，*erbB* 癌基因仅仅编码表皮生长因子受体的跨膜区和与酪氨酸激酶有关的结构域。这些都提示 *erbB* 可能源于为 EGFR 编码的基因，并通过失控的 EGFR 的表达产生转化作用。

7.4.2.2 表皮生长因子受体、癌基因与癌

由上述 EGFR 的分子结构可知，EGFR 胞浆结构域与 erbB 基因产物的氨基酸顺序有很大一部分是相同的，它们之间有非常密切的联系。

erbB 癌基因是 30 年代初从一鸟类逆转录病毒中分离鉴定到的，它主要引起成红细胞增多症和纤维肉瘤。最近，又从一些逆转录病毒中发现了这种癌基因，并证明它除能导致成红细胞增多症外，还能引起血管肉瘤。对该基因的结构分析表明，它具有酪氨酸蛋白激酶的顺序，属于跨膜糖蛋白之列。人们很自然地把它与生长因子基因联系起来，并很快证实了它与人 EGFR 基因的同源性，进而确定它就是来自鸡的 EGF 受体基因。

然而，虽然已经证实某些生长因子受体基因可作为癌基因，但是对何以会出现这种情况人们却了解很少。图 7-21 是 erbB 癌基因产物同 EGFR 结构的比较。可见，erbB 蛋白实际上是 EGFR 经“掐头去尾”后的残留形式。它的 N-末端只有 61 个氨基酸残基，与 EGFR 细胞外结构域的 621 个残基相比少多了，而这种缺失无疑是将 EGFR 的 EGF 结合部位去掉了。它的 C-末端即细胞内结构域的缺失似乎不像 N-末端那么多，但是由于除去了 EGFR 主要的自身磷酸化部位——第 1173 位残基，使其自身磷酸化能力大为减弱，亦即失去了 EGF 借以调节激酶活性的主要区段。很可能正是由于这些变化，使 EGFR 基因转变为癌基因。

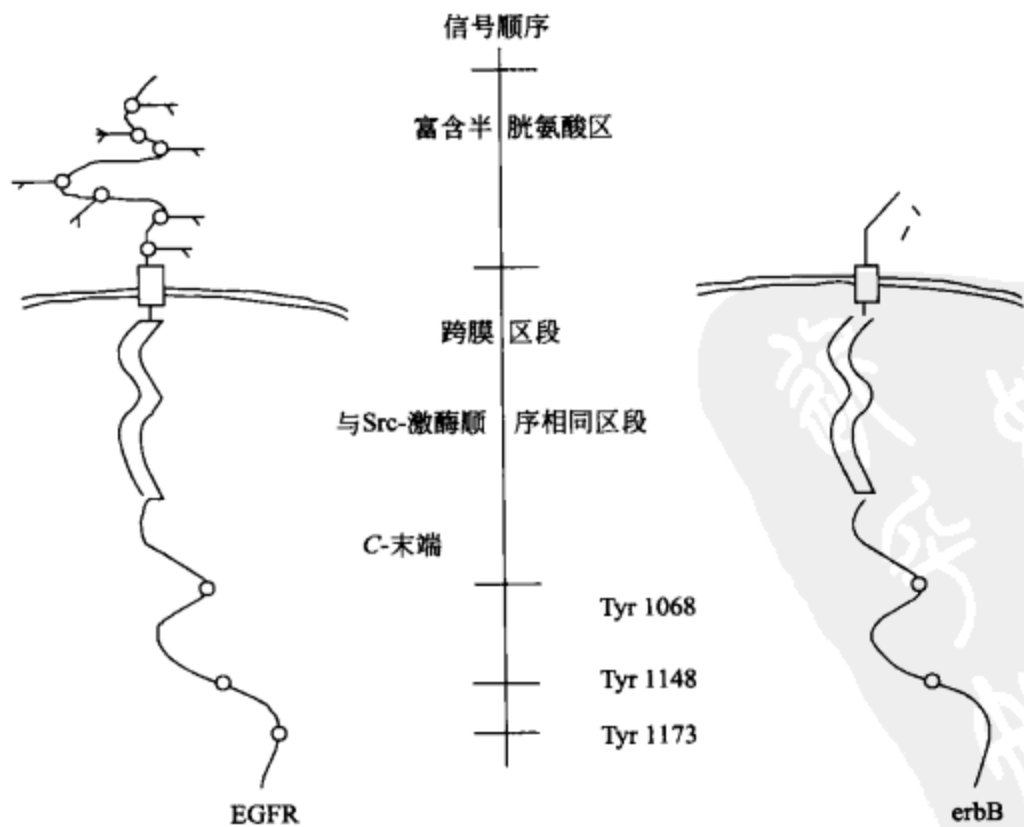


图 7-21 erbB 癌基因产物同 EGFR 结构的比较

但是，这种从分子水平比较作出的推断是否符合实际呢？这就要了解 erbB 基因与临床上一些涉及 EGFR 的癌瘤之间的关系。实践证明，很多关于人的 EGFR 的研究是以人上皮癌细胞系 A431 为对象进行的。深入研究发现，在这种细胞中 EGFR 基因的表达扩增了约 50 倍。在其他一些鳞状上皮癌（主要为头部和颈部）细胞中，以及某些脑肿瘤（主要来源于胶质细胞）中，也都发现了表皮生长因子受体扩增现象。

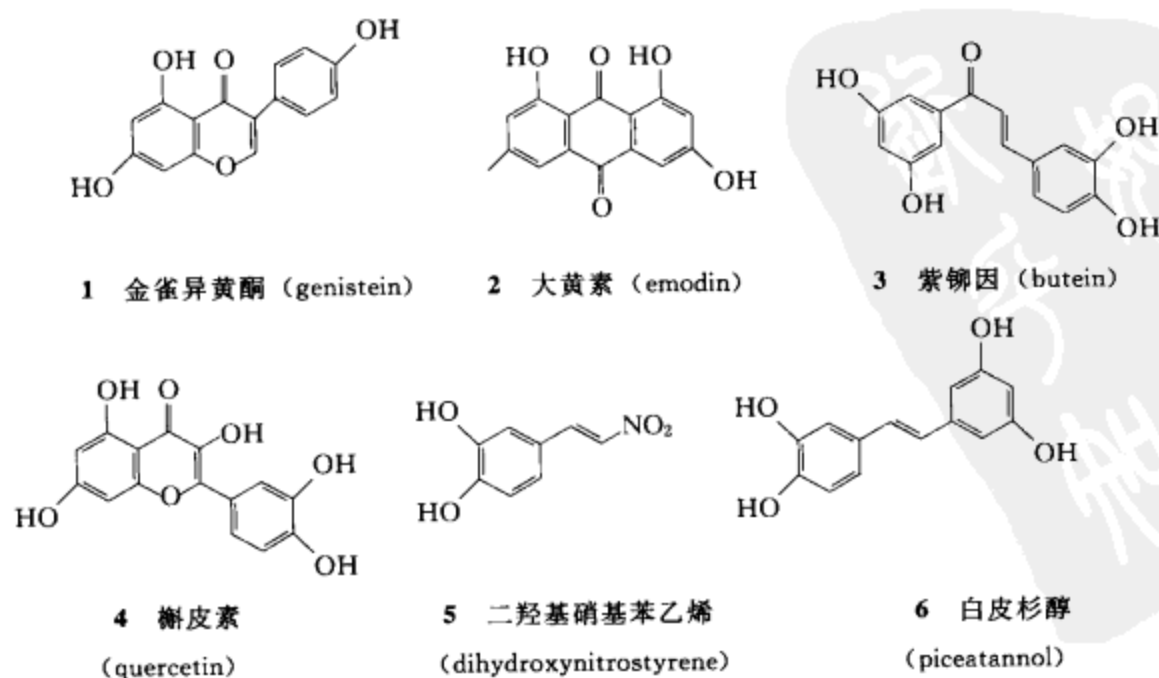
在临床研究中，应用放射性同位素标记的抗体定量 EGFR 数目，发现在鳞状上皮癌、某些胶质细胞瘤、乳腺癌、膀胱癌以及食道癌中都有高水平的 EGFR，且与病理变化和临床表现有一定关系。M. S. Berger 等报道了对 109 例原发肺癌进行免疫化学和 DNA-DNA 印记分析结果，发现鳞状细胞癌的 EGFR 远高于其他受体类型。

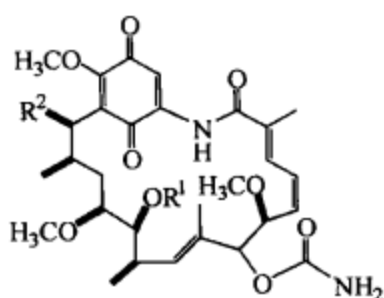
由于高水平的表皮生长因子受体与病理变化和临床表现相关，这就为用抗 EGFR 抑制剂治疗某些癌症提供了一定的实验依据。

7.4.2.3 天然产物类抑制剂

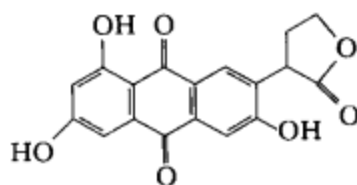
通过活性追踪的方法从天然产物中分离提取先导化合物一直是药物开发的常规思路。EGFR 的抑制剂最初也是从天然产物着手的，主要用癌细胞体系进行筛选。EGFR 的抑制机理大致可分为两种：一种为与受体竞争性的抑制剂（在细胞膜外作用）；另一种为与 ATP 竞争性的抑制剂（在细胞膜内的激酶区作用）。此外，有些分子既是受体的竞争性抑制剂又是 ATP 的竞争性抑制剂，而有些分子既非受体的竞争性抑制剂又非 ATP 的竞争性抑制剂，却可以抑制 A431 细胞的生长，其作用机理并不是特别清楚。这里将迄今为止文献报道的天然产物类抑制剂总结如下（图 7-22）。

ATP 竞争性抑制剂（1~9）





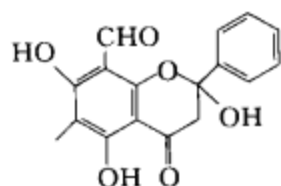
7 $R^1=R^2=H$: 格尔德霉素 (geldanamycin)



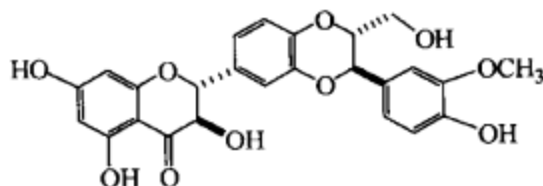
9 paecilquinone A

8 $R^1=CH_3$, $R^2=OCH_3$: 除莠霉素 A (herbimycin A)

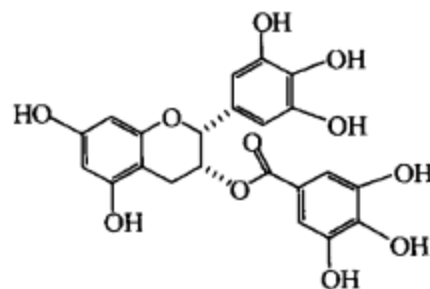
受体竞争性抑制剂 (10~12)



10 Desmal

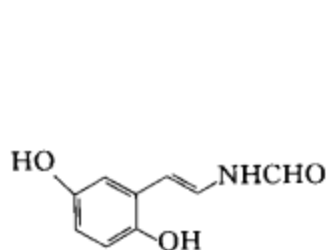


11 水飞蓟素 (silymarin)

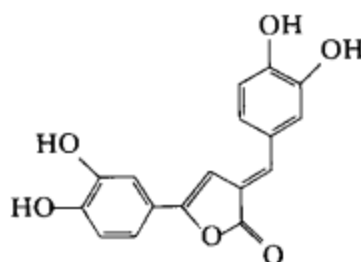


12

受体竞争性及 ATP 竞争性抑制剂 (13,14)

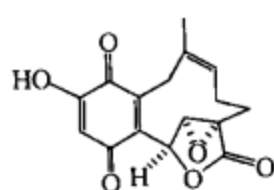


13 制表菌素 (erbstatin)

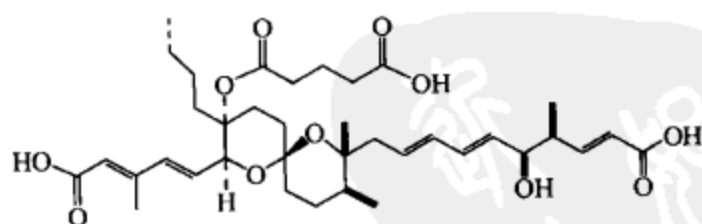


14 BE-23372M

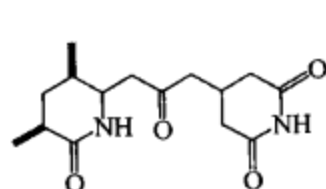
非受体竞争性及非 ATP 竞争性抑制剂 (细胞增殖抑制剂) (15~24)



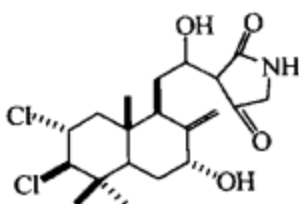
15 Clavilactone CD



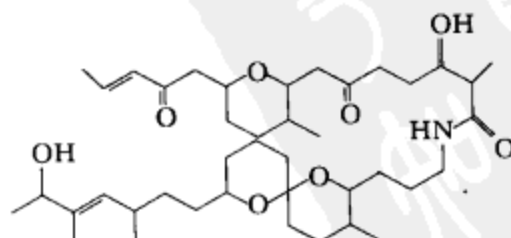
16 Reveromycin A



17 抑表皮素 (epiderstatin)



18 dichlorolissoclimide



19 bistramide A

图 7-22

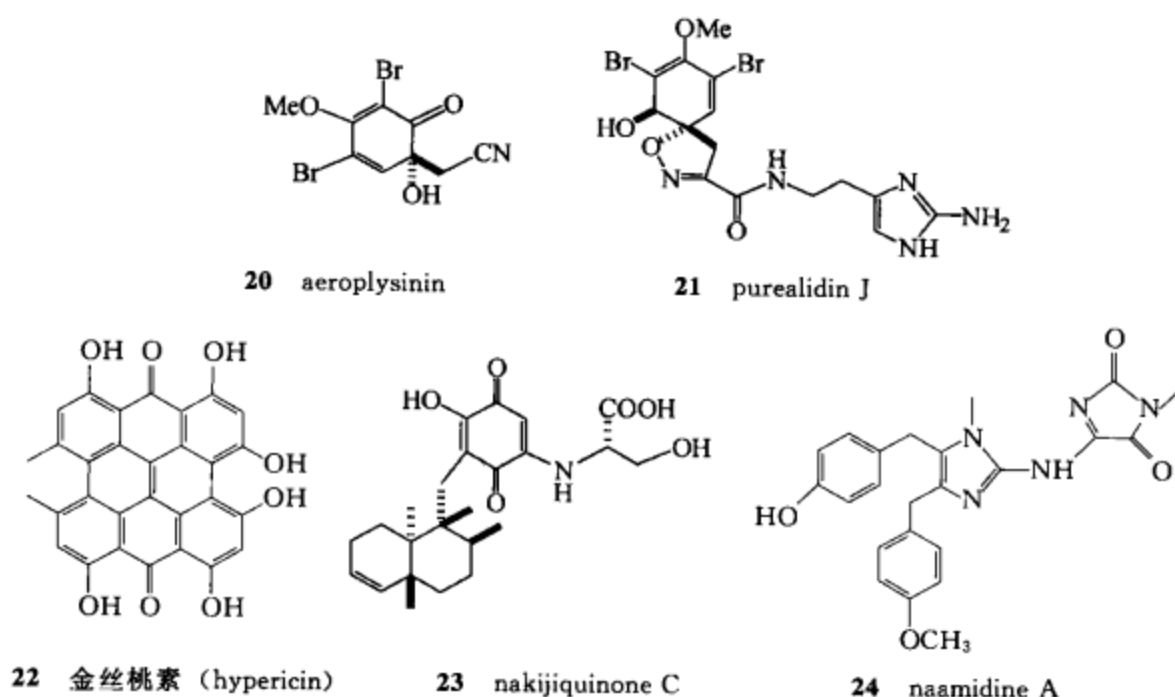


图 7-22 EGFR 的各种天然产物类抑制剂

早期发现的 EGFR 抑制剂具有比较简单的结构，如一些黄酮（4）、异黄酮类（1）抑制剂作为其中的典型代表，直到现在仍然被研究。实验证明，这些抑制剂通过竞争性抑制酶 ATP 结合位点起到抑制作用。天然产物类抑制剂通常具有较低的毒性，但其特异性也往往较差，如以上所示的小分子抑制剂大多数对 HER-2 也具有一定的抑制作用。另外，好的天然产物抑制剂主要为药物的开发提供了一个分子构型，在此基础上进行改造或者全新的组合化学合成，从大量的结构类似物中选择出高效、高特异性与低毒性的先导化合物。

7.4.3 抗 EGFR 药物的初步筛选

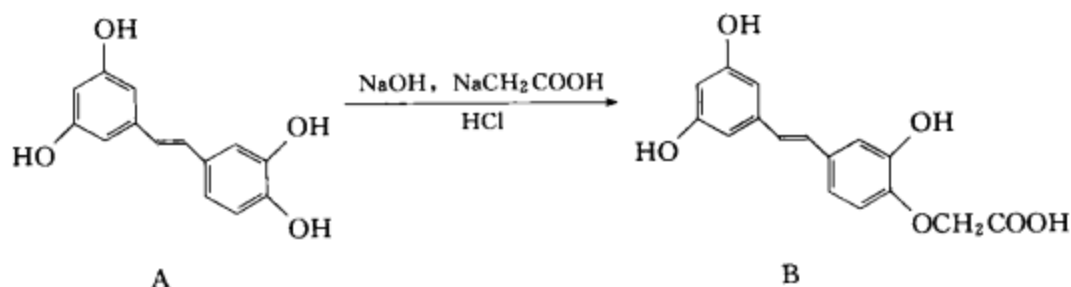
为了寻找抗 EGFR 的有效药物，对近 50 种中草药的粗提物进行了活性筛选，包括冬凌草、石蒜、创当（音译，为贵阳本地药材）、藏药鬼箭锦鸡儿、索罗玛保等，从中选择了具有较高活性的藏药鬼箭锦鸡儿作为最后的研究对象。

在活性跟踪指导下用传统的分离提取手段对藏药鬼箭锦鸡儿（*Caragana jubata*）中的有效成分进行了分析分离。先通过萃取的方式将乙醇粗提物分段，根据溶剂极性由小到大的顺序将粗提物分为 5 段，分别为石油醚部分、氯仿部分、乙酸乙酯部分、正丁醇部分、水层部分。活性检测结果表明其中乙酸乙酯部分具有相对较高的活性，因此下面主要对此部分进行进一步的分离。通过一次完整的硅胶柱色谱的过程将乙酸乙酯部分分为 24 段，然后对每一段测活，根据活性筛选结果选择其中两段具有一定活性的分段产物进一步细分，主要仍然通过硅胶柱色谱和凝胶柱色谱的方法进行分离，最终得到 10 个单一化合物，通过质谱和核磁手段鉴定分子结构，其中化合物 6 和 8 的结构尚待鉴定。活性测定结果表明化合物 5、9、10 具有较高的抗 EGFR 活性，分别为槲皮素、紫柳花素和 3,4,3',5'-四羟基二苯乙烯。

7.4.4 多克隆抗体的制备

7.4.4.1 半抗原的改造

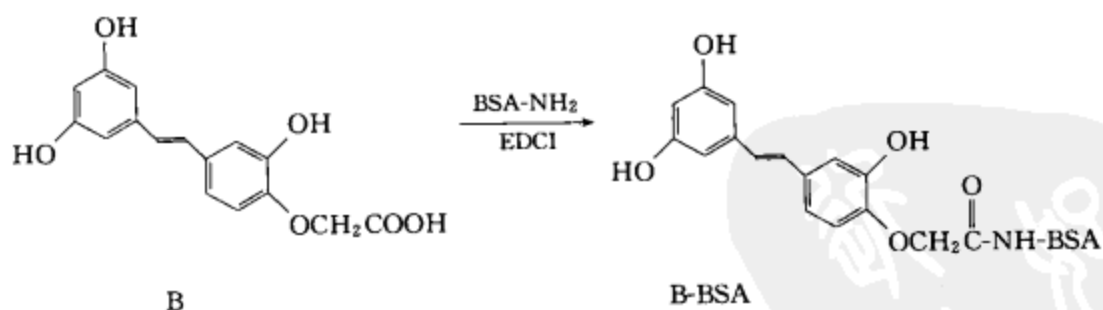
如上所述,这里要将白皮杉醇(piceatannol)上的酚羟基改造成羧基衍生物,其反应方程式如下:



(注:该反应过程中没有羟基选择性,即4个位置的羟基都有可能发生反应,只能通过控制原料量来控制发生反应的羟基个数,B分子写成上述形式只是示意。)

将73.2mg (0.3mmol) 白皮杉醇(A)溶解于1mL干燥的二甲亚砜(DMSO)中,往该溶液中加入40mg NaOH粉末(加入强碱后溶液颜色变深,由无色变为黄色,直至深褐色),搅拌使其充分溶解并反应10min后,加入28.0mg (0.24mmol) ClCH₂COONa,搅拌溶解,继续反应2h后,加入50mL冰水终止反应。得到的产物用2mol/L盐酸酸化,放置一段时间,溶液中出现褐色沉淀,过滤得到沉淀,并将其用水洗至中性,干燥后得到褐色粗品,称重约为40mg,产率为55%。将得到的粗品用CH₃OH-CHCl₃重结晶的方法纯化,用甲醇溶解粗品,在回流状态下缓慢滴加甲醇,使之全溶,然后滴加适量氯仿,使产品处于析出的临界状态。将溶液置于烧杯慢慢冷却,析出褐色晶体。

7.4.4.2 全抗原的合成



在反应瓶中加入3.02mg (10 μ mol) B、2.8mg (15 μ mol) EDCI,再加入1mL干燥的DMSO溶解反应物后,搅拌反应2h。同时,将20mg (30mmol) BSA溶解于2.5mL NaHCO₃ (0.1mol/L, pH 7.0) 溶液中。将活化后的B反应液缓慢滴加入BSA溶液中,边加边轻轻搅拌,同时维持反应液的pH值在7.0以上。反应完成后,蛋白质溶液用0.01mol/L pH 7.4的PBS透析,除去未反应完全的小分子,得到的产品用聚乙二醇浓缩(测得终浓度为8mg/mL)后,测偶联率。

7.4.4.3 B-BSA中偶联率的测定

这里用MALDI-TOF测定B-BSA的质量分布图,如图7-23所示。

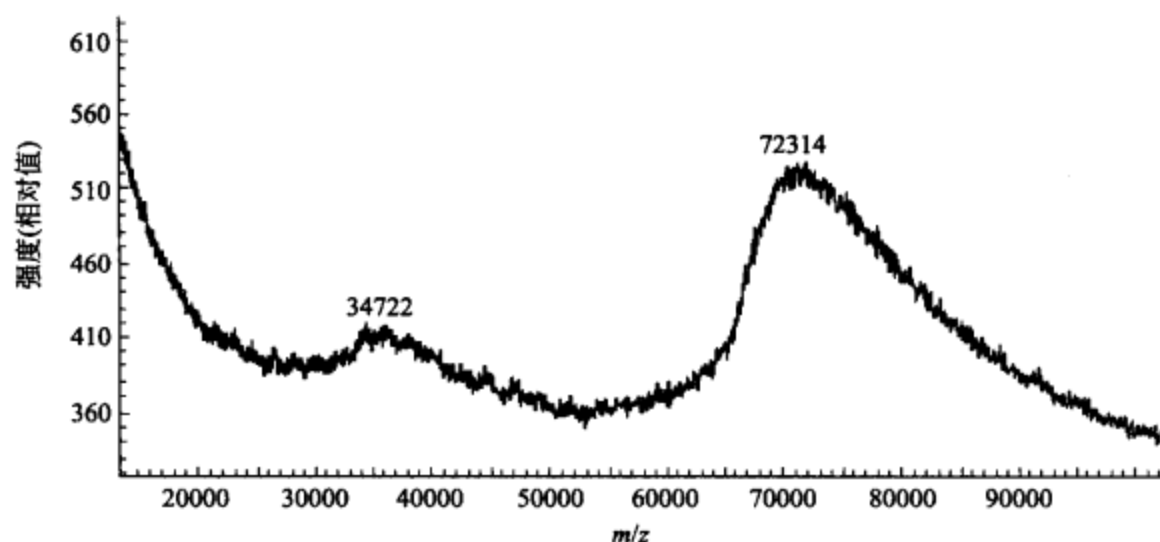


图 7-23 用 MALDI-TOF 测定 B-BSA 的质量分布图

得到 B-BSA 中 B 的偶联率为 18。

计算公式如下：

$$\text{偶联率 (B : BSA)} = \frac{72314 - 67000}{302} = 18$$

7.4.4.4 实验动物的免疫

选择 4~5kg 重雄性健康青紫蓝兔 2 只，免疫前观察 1 周，并采血制备阴性血清。

免疫程序如下。

首免：每只兔子免疫剂量为 1mg。将上述合成的 B-BSA 母液 (pH 7.4 的 PBS 缓冲液) 用生理盐水稀释至 2mg/mL，取 1.5mL 该溶液与等量的福氏完全佐剂充分混匀 (B-BSA 的终浓度为 1mg/mL)，用注射器来回混合使之成为油包水的乳浊液。免疫脚掌，每只兔子 1mL (每只脚掌 0.5mL)。

Ⅱ免：隔 10 天，剂量为每只兔 0.5mg，佐剂为福氏不完全佐剂，注射部位为背部多点注射，其他同首免。

Ⅲ免：隔 11 天，注射部位为踝窝肿大的淋巴结，其他同Ⅱ免。13 天后耳缘静脉采血测抗体。

Ⅳ免：隔 14 天，注射部位为 4 个肩胛骨缝，其他同Ⅱ免。10 天后耳缘静脉采血测抗体。

Ⅳ免过 10 天后从耳缘静脉采血测抗体效价为 10^5 ，因而Ⅳ免过 11 天后从兔颈动脉放血。

7.4.4.5 抗血清的制备

颈动脉放血可收集到兔子总血量 80%~90% 的血液，因此是一次性放血比较常用的方法。将兔子绑定，纵向切开颈动脉皮肤，分离肌肉，暴露颈动脉。用止血

钳夹住动脉近心端，另一端（相隔 2~3cm）用绳子结扎，将颈动脉朝近心端剪一 V 形小口。将在肝素溶液中预先浸泡过的塑料细管插入动脉，结扎，出口接离心管，松开止血钳，血液很快流出。开始血流速度较快，以后逐渐减慢，可逐渐将家兔后躯垫高些。所得血液在 4℃ 冰箱中静置过夜。离心取上清液，必要时用刀划开血块。收集上清液，分装 3mL/管。保存于 -20℃ 冰箱中。

7.4.5 多克隆抗体的纯化及其性质评价

7.4.5.1 抗体的纯化

(1) 饱和硫酸铵盐析沉淀法 取经初步沉降后的血清 20mL 与等体积的生理盐水混合后，于搅拌下缓慢滴加 40mL 饱和硫酸铵溶液 (SAS)，于 4℃ 沉淀 3h 或过夜，使蛋白充分沉淀 (50% 饱和度)。10000g、4℃ 离心 30min，弃上清液，用 12mL 生理盐水充分溶解沉淀，再滴加 8mL 饱和硫酸铵溶液，4℃ 放置 1h (40% 饱和度)。离心 (10000g，4℃) 30min，弃上清液，用 13.3mL 生理盐水溶解沉淀，滴加 6.6mL 饱和硫酸铵溶液，4℃ 放置 1h (33% 饱和度)。同上述离心过程 (10000g，4℃，30min)，离心所得沉淀物用少许 0.01mol/L pH 7.4 的 PBS 溶解，装入透析袋。置 4℃ 冰箱中，用 1000mL pH 7.4 的 PBS 透析，更换 PBS 数次。直至用奈氏试剂测透析外液无黄色为止。

(2) DEAE-纤维素 (diethyl-aminoethyl-cellulose) 提取法 称取 DEAE 30g 加入 500mL 烧杯中，加蒸馏水浮去细小颗粒，加 0.1mol/L NaOH 溶液 300mL 浸泡 10min，倾出 NaOH 溶液，重复浸泡一次，用水洗至中性，加 0.1mol/L 盐酸 300mL 浸泡 10min，用水洗至中性，再用 NaOH 溶液 300mL 浸泡 10min，加水洗至中性，浸泡在 PBS 中，用前以 pH 8.6 的 0.01mol/L Tris-HCl 替换 PBS。

将层析柱固定于滴定架上，出口接一细塑料管，并关闭出水口。将浸泡于 0.01mol/L pH 8.6 的 Tris-HCl 中的 DE52 倒入柱中，待 DE52 自然沉降 3~5cm 高时松开出水口螺旋夹，控制流速 1~2mL/min，同时连续加入 DE52 至所需高度。松开出水口螺旋夹，以 0.01mol/L pH 8.6 的 Tris-HCl 平衡，控制流速为 15 滴/min，待流出液与洗脱液的 pH 值达到一致时停止平衡。将经 SAS 沉淀法粗提的抗体 1mL 加入柱中，继续用 0.01mol/L pH 8.6 的 Tris-HCl 洗脱，紫外检测器 280nm 检测流出液，收集第一个峰。装入透析袋中，用 0.01mol/L pH 7.4 的 PBS 透析，并用聚乙二醇 20000 浓缩。

(3) 亲和色谱纯化 因为经改造后的小分子与 BSA 偶联后得到的抗原中含有不同偶联度的完全抗原 (18 为其平均偶联度)，并且可能含有少量没有偶联上小分子的 BSA，则免疫后可能产生抗 BSA 的抗体。为了排除抗 BSA 的抗体的干扰，制备了 BSA 的亲和色谱柱用来分离所得多克隆抗体中抗 BSA 的成分，并经过 ELISA 检测判断经过纯化后所得抗体抗 BSA 的效价确实下降了一个数量级 (见表 7-5)。

① 亲和色谱柱的制备

a. 载体——Sephacrose CL-4B 的活化 载体 (Sephacrose) 活化的方法通常有环氧氯丙烷活化法和溴化氰活化法, 这里采用了前者。其过程简单介绍如下。

将适量的 Sephacrose CL-4B 于玻璃烧结漏斗抽干, 称取 8g (湿重) Sephacrose CL-4B beads, 用 50mL 0.1mol/L NaCl 溶液淋洗, 再用蒸馏水洗, 抽干, 转移到小三角瓶内, 然后加入 1.6mL 2mol/L NaOH 溶液、0.375mL 环氧氯丙烷、3.75mL 56% 1,4-二噁烷, 在 40℃ 的恒温水浴摇床活化 2h。然后将活化的 Sephacrose CL-4B beads 转移至玻璃漏斗中, 用蒸馏水洗去未反应完的残留试剂, 再用 10mL 0.2mol/L, pH 9.5 的 Na_2CO_3 缓冲液洗涤, 抽干后立即偶联。

b. 载体与 BSA 的偶联 将上述活化好的 Sephacrose CL-4B 装入一个小三角瓶内。用 10mL 0.2mol/L pH 9.5 的 Na_2CO_3 缓冲液将 80mg BSA 溶解。取出 0.1mL 溶液, 稀释后用紫外分光光度计检测蛋白含量。剩余溶液加入到装有 Sephacrose CL-4B 的瓶内, 与之充分混合, 把三角瓶放入 40℃ 的恒温水浴摇床, 振荡 24h 使之偶联。

偶联终止后, 将 Sephacrose CL-4B 转移到玻璃漏斗中, 用 100mL 0.5mol/L NaCl 溶液洗去未被偶联的蛋白 (收集滤液, 量取体积, 用紫外分光光度计测定未被偶联的蛋白含量)。结果表明未被偶联的蛋白量为 30mg, 由此推算 8g Sephacrose CL-4B 上偶联的抗体为 50mg。再用 100mL 蒸馏水洗, 转移到小烧杯内, 用亲和平衡液 (0.5mol/L NaCl + 0.1mol/L pH 7.8 的 Tris-HCl) 浸泡 15min, 脱气后装柱。偶联后的 Sephacrose CL-4B 需保存在 4℃。

c. BSA 亲和柱的填装 用 PerSeptive Biosystem 的自动装柱器 (ID Self Pack) 填柱。每根亲和柱 ($\phi 2.1\text{mm} \times 30\text{mm}$) 大约需要上述偶联抗体后的 Sephacrose CL-4B beads 100 μL 。此亲和柱按如下程序进行洗涤即可重复利用: 0.5mL 2mmol/L (pH 6.7) NH_4Ac 溶液, 1mL (0.1mol/L CH_3COOH + 0.5mol/L NaCl) 溶液, 2.5mL 2mmol/L (pH 6.7) NH_4Ac 溶液。

② 亲和色谱纯化

(a) 亲和层析柱用 PBS 洗至平衡。

(b) 将 DEAE 纤维素柱纯化后的抗体用 PBS 稀释 5 倍, 缓慢通过亲和层析柱, 流出的样品再次上柱, 反复 3 次。

(c) 加 PBS 5mL 洗柱, 合并洗出液, 并用聚乙二醇 20000 浓缩。

亲和色谱柱用 pH 2.7 的甘氨酸-盐酸溶液洗脱再生, 用 PBS 平衡后, 在 4℃ 保存于含 0.02% 叠氮钠的 PBS 中。

7.4.5.2 抗体效价的测定 (ELISA 方法)

(1) ELISA 方法中所用溶液 详述如下。

包被液: 0.05mol/L pH 9.6 的 NaHCO_3 - Na_2CO_3 缓冲液。

封闭液：0.7%明胶的 PBS 溶液。

洗涤液 (PBST)：0.05% Tween-20 的 PBS 溶液。

底物液 (OPD 显色液)：临用前取 A 液 4.86mL 与 B 液 5.14mL 混合，加入 OPD 10mg，待充分溶解后加入 30% (体积分数) H_2O_2 50 μL ，即得底物显色液。其中 A 液与 B 液如下。

A 液 (0.1mol/L 柠檬酸溶液)：柠檬酸 1.92g，加蒸馏水至 100mL。

B 液 (0.2mol/L Na_2HPO_4 溶液)： $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 7.17g，加蒸馏水至 100mL。

终止液：2mol/L H_2SO_4 。

(2) ELISA 方法测定抗体效价 这里用两种底物包被，分别为与 OVA 偶联的小分子 (白皮杉醇) (称为 B-OVA) 以及 BSA，其中半抗原白皮杉醇与 OVA (B-OVA) 偶联的方法同白皮杉醇与 BSA 偶联的方法相同。ELISA 方法的操作步骤如下。

① 包被抗原。将抗原用包被液稀释至 10~20 $\mu\text{g}/\text{mL}$ ，加入到 96 孔微量滴定板中，每孔 100 μL 。然后放入 4℃ 冰箱中过夜，使抗原附着于孔底。倒掉包被液，用 PBST 洗 3 次，每次加入洗涤液后摇床振荡 5min，然后甩干反应板。

② 封闭。每孔加封闭液 200 μL ，于 37℃ 放置 2h，倒掉包被液，用 PBST 洗 3 次。

③ 加入系列稀释的待测血清，每孔 100 μL ，于 37℃ 放置 1h，用 PBST 洗 3 次。

④ 加入 5000 倍 PBS 稀释的辣根过氧化物酶标记羊抗兔 IgG (H+L) (二抗)，每孔 100 μL ，于 37℃ 放置 1h，用 PBST 洗 3 次，然后用重蒸水洗 2 次。

⑤ 显色。加入现配的底物液 100 μL ，将微量滴定板置于室温暗处反应 10min，当阳性对照孔颜色变黄时，每孔加入 50 μL 终止反应液。

⑥ 在酶标检测仪上 OD_{490 nm} 读数。

ELISA 方法测得抗体纯化前后的效价见表 7-6。从表 7-6 数据可以得出如下结论：抗体经纯化后有效效价 (抗小分子的抗体效价) 降低约 3 倍 (从 27000 降低到 10000)，但因为相应的蛋白含量降低了约 5 倍，所以可以认为单位蛋白含量的有效效价经纯化后提高。而且经过亲和色谱纯化后抗体抗 BSA 的效价大大下降 (10 倍)，单位蛋白的抗体效价也有所下降，因此认为经纯化处理后该抗体的有效效价提高。

表 7-6 PcAb 对不同抗原的效价

项 目	B-OVA	BSA	蛋白含量/(mg/3mL)
纯化前	27000	80000	65.5
纯化后	10000	8000	13.0

7.4.6 多克隆抗体的固定化与亲和前沿色谱

7.4.6.1 多克隆抗体的固定化（亲和吸附剂的合成）

（1）载体 Sepharose CL-4B 的活化 载体活化的方法仍然采用环氧氯丙烷活化法，其过程介绍见 7.4.5.1 节。

（2）载体与抗体的偶联 将上述活化好的 Sepharose CL-4B 装入一个小三角瓶内。另取已经过纯化的抗体（原血清 5mL，蛋白含量为 65.5mg，经纯化后体积为 3mL，蛋白含量为 13mg，见表 7-5），先用 1000mL 0.2mol/L pH 9.5 的 Na_2CO_3 缓冲液透析 4h，再加入到装有 Sepharose CL-4B 的瓶内与之充分混合，把三角瓶放入 40℃ 的恒温水浴，摇床振荡 24h 使之偶联。

偶联终止后，将 Sepharose CL-4B 转移到玻璃漏斗中，用 100mL 0.5mol/L NaCl 溶液洗去未被偶联的蛋白（用考马斯亮蓝法测得未被偶联的蛋白量为 5mg，由此推算 2g Sepharose CL-4B 上偶联的抗体为 8mg），再用 100mL 蒸馏水洗，转移到小烧杯内，用亲和平衡液（0.5mol/L NaCl+0.1mol/L pH 7.8 的 Tris-HCl）浸泡 15min，脱气后装柱（最终偶联了抗体后的 Sepharose CL-4B 体积约为 3.5mL）。偶联后的 Sepharose CL-4B 需保存在 4℃。

（3）抗体亲和柱的填装 用 PerSeptive Biosystem 的自动装柱器（ID Self Pack）填柱。每根亲和柱（ $\phi 2.1\text{mm} \times 30\text{mm}$ ）大约需要上述偶联抗体后的 Sepharose CL-4B 100 μL 。此亲和柱用来测柱容及解离常数时需按如下程序进行洗涤即可重复利用：0.5mL 2mmol/L (pH 6.7) NH_4Ac 溶液，1mL (0.1mol/L CH_3COOH +0.5 mol/L NaCl) 溶液，2.5mL 2mmol/L (pH 6.7) NH_4Ac 溶液。但用来识别中草药中活性成分时，因为其成分比较复杂，亲和柱一般不回收利用。

7.4.6.2 亲和前沿色谱

采用亲和前沿色谱-质谱联用的手段，有效地实现了中草药中活性成分的筛选与鉴定一体化的技术。实验装置为徐筱杰课题组根据所用仪器的特点自行设计，如图 7-24 所示。

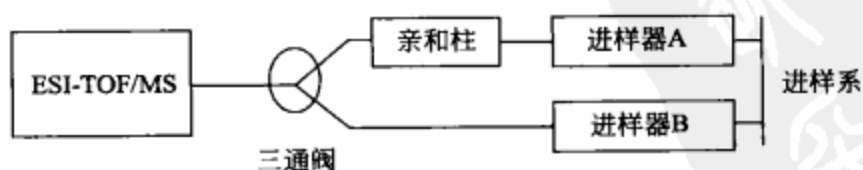


图 7-24 FAC-MS 联用系统示意

图 7-24 中，A 和 B 为两个容量为 2.5mL 的进样器。实验过程中，将待分析样品装入进样器 A 中（将样品溶解并稀释在 2mmol/L NH_4Ac 溶液中），进样器 B 中装有同样体积的溶剂甲醇以辅助样品在质谱中的分析。在进样泵推动下，样品经过亲和柱后与溶剂甲醇在三通阀中混合，然后进入质谱分析。质谱分析采用电喷雾离子化、飞行时间检测方式（ESI-TOF）。

对于目标分子，通过其选择离子的亲和前沿色谱图进行分析。

7.4.7 多克隆抗体识别鬼箭锦鸡儿中活性成分

本研究内容主要用多克隆抗体制成的亲和色谱柱识别藏药鬼箭锦鸡儿中的活性成分。为了表征多克隆抗体的性质，先检测多抗亲和柱对半抗原（白皮杉醇）的亲和能力及柱容。

7.4.7.1 亲和柱的柱容（ L_t ）及解离常数（ K_d ）的测定

利用亲和前沿色谱-质谱联用测定亲和柱的柱容（ L_t ）及解离常数（ K_d ）。这里用一组不同浓度的半抗原（白皮杉醇）进行测量。将一定量的半抗原溶于甲醇，浓度为 5mg/mL，再用 2mmol/L NH_4Ac 溶液依次稀释为 20 $\mu\text{g/mL}$ 、10 $\mu\text{g/mL}$ 、5 $\mu\text{g/mL}$ 和 2.5 $\mu\text{g/mL}$ 。将样品溶液装入进样针 A（2.5mL），溶剂甲醇装入进样针 B，以同样的流速（10 $\mu\text{L/min}$ ）推动进样针，样品溶液经过多克隆抗体的亲和色谱柱后与溶剂甲醇于三通阀中混合，再经过质谱检测其亲和前沿质谱信息，仪器自动收集谱图，得到样品的总离子流图（TIC），由样品的分子量选择出样品的选择离子流图（XIC），其结果如图 7-25 所示。通过质谱图计算出每一浓度下对应的前沿时间，并由流速计算出相应的前沿体积（表 7-7）。通过不同浓度对应的不同前沿体积由式(7-2) 计算出该亲和柱的柱容（ L_t ）及解离常数（ K_d ），分别为 44.90 $\mu\text{g/mL}$ 和 17.14 $\mu\text{mol/L}$ （图 7-26）。

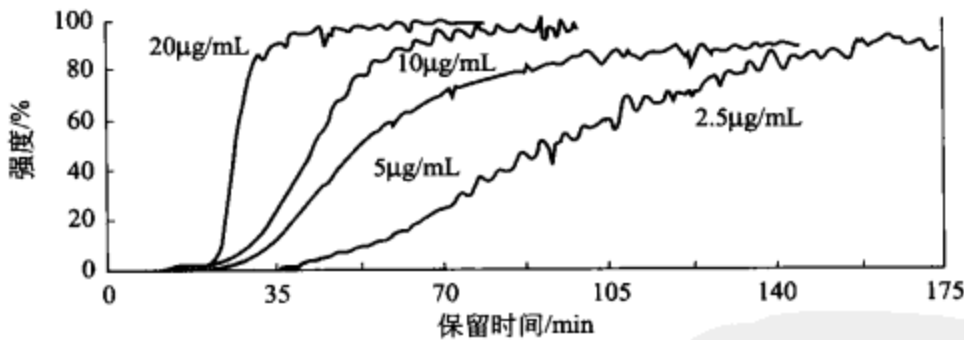


图 7-25 不同浓度待分析样品（白皮杉醇）的 FAC 图

表 7-7 白皮杉醇的 FAC 测定结果

$c/(\mu\text{g/mL})$	$c(\mu\text{mol/L})$	$\frac{1}{c}(\mu\text{mol/L})^{-1}$	$(V-V_0^{\oplus})/\mu\text{L}$	$\frac{1}{c(V-V_0^{\oplus})}/\mu\text{mol}^{-1}$
20.0	82.0	0.0122	188.0	64.9
10.0	41.0	0.0244	314.0	77.7
5.0	20.5	0.0488	486.0	100.4
2.5	10.2	0.0976	676.0	144.9

① V_0 是使用 Sepharose CL-4B 载体用空白亲和柱测定。

7.4.7.2 利用多克隆抗体亲和柱识别鬼箭锦鸡儿中活性成分

主要目的是研究将多克隆抗体的亲和色谱柱直接应用于中草药的粗提物体系来识别不同的活性成分，因此这里对研究对象——鬼箭锦鸡儿的乙酸乙酯粗提物直接

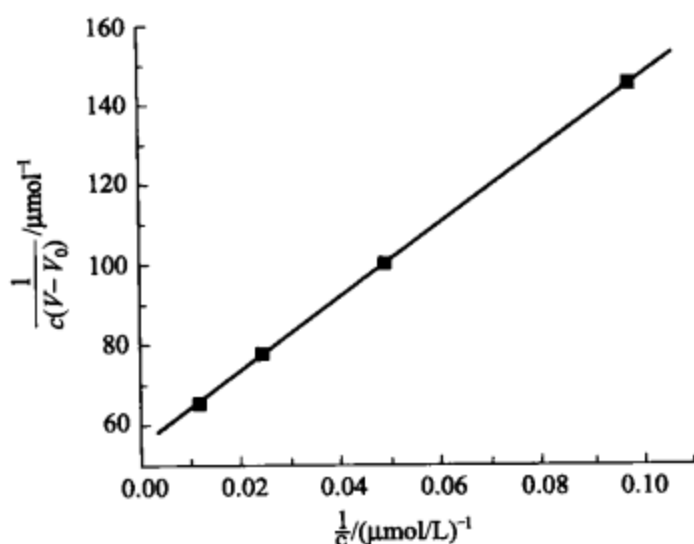


图 7-26 PcAb 的柱容 (V_t) 及解离常数 (K_d) 的测定

进行分析。

将鬼箭锦鸡儿的乙酸乙酯部分旋干后溶于甲醇 (初浓度为 10 mg/mL), 再用 2mmol/L NH_4Ac 溶液将样品稀释到 100 $\mu\text{g/mL}$ (如果样品能溶于 NH_4Ac 溶液则不需用甲醇先溶解, 可以直接溶在 2mmol/L NH_4Ac 溶液中)。将 2.5mL 稀释后的样品装入进样器 A, 同样将 2.5mL 溶剂甲醇装入进样器 B, 样品在经过亲和色谱柱后与辅助溶剂甲醇混合, 进入到质谱仪中进行检测, 得到粗提物中各组分的总离子流图 (TIC), 再根据各个

组分的分子量选择出其相应的选择离子流图 (XIC)。整个过程是连续进样连续检测, 各个组分在亲和柱上根据其保留能力不同而先后达到饱和吸附, 得到相应的亲和前沿色谱图。相应的质谱图及亲和前沿色谱图如图 7-27、图 7-28 所示。

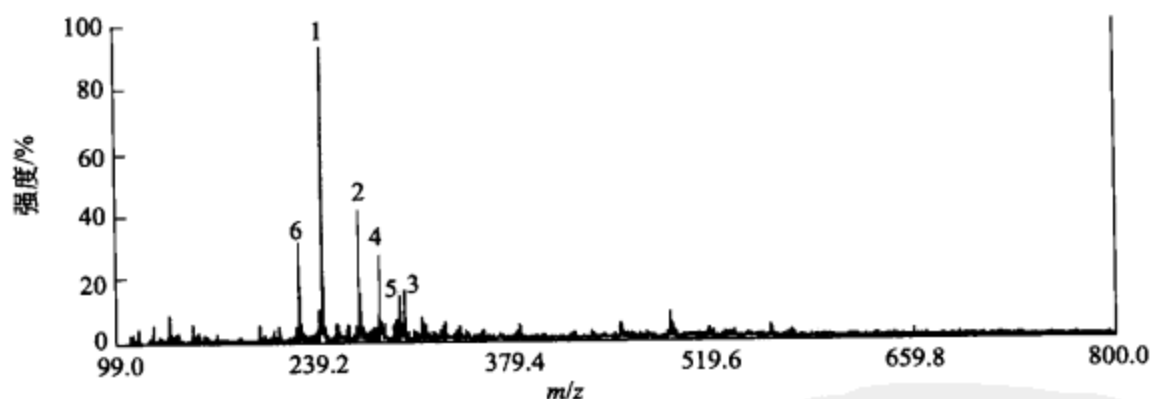


图 7-27 鬼箭锦鸡儿提取物的 MS 图 (图谱中的 6 个主峰对应于图 7-29 中的 6 个化合物)

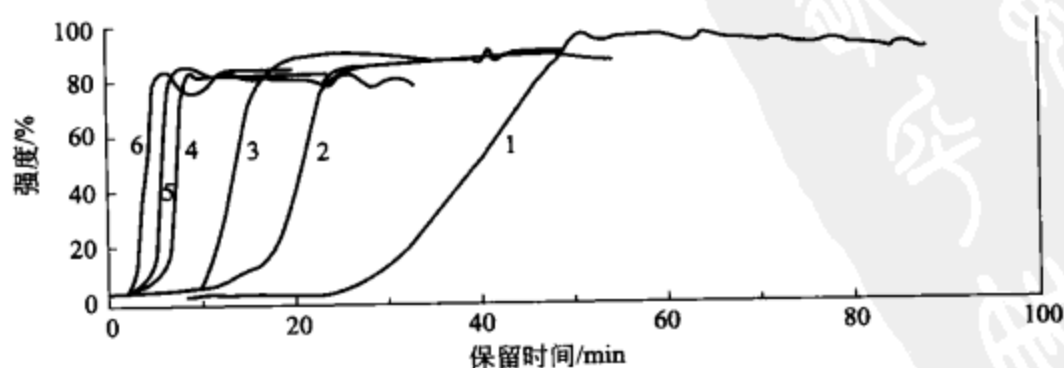


图 7-28 鬼箭锦鸡儿提取物中 6 个未知化合物的 FAC 图
(图谱中的 1~6 对应于图 7-29 中的 6 个化合物)

结果显示鬼箭锦鸡儿中有 6 个主要的活性成分 (图 7-29)。

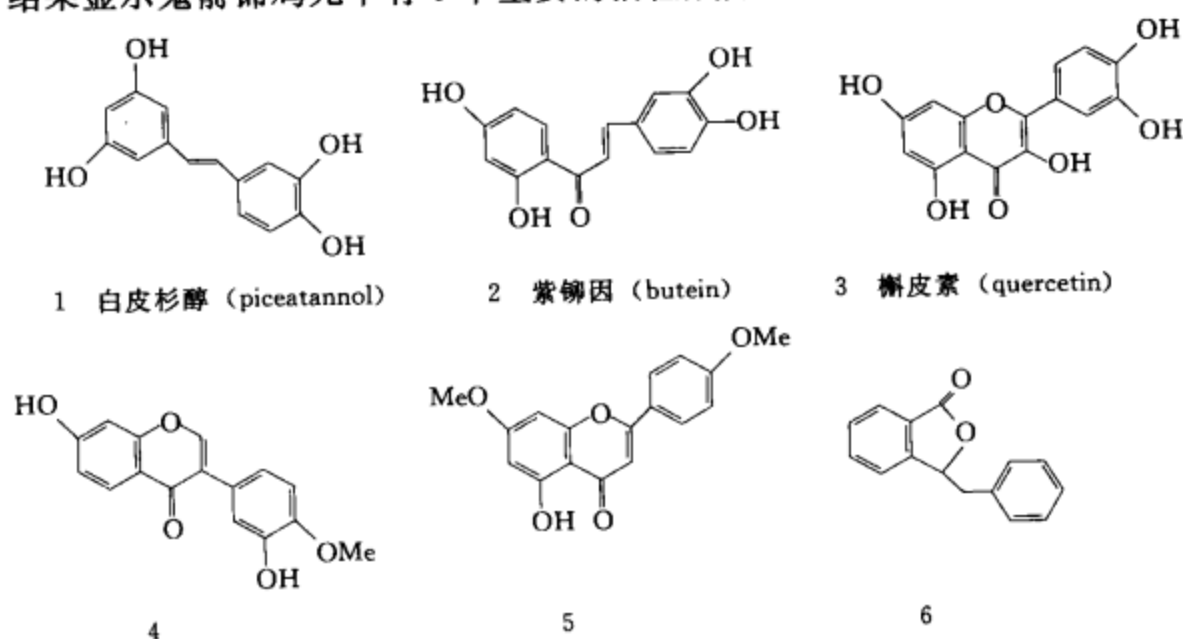


图 7-29 与图 7-28 中的 6 个峰对应的 6 个化合物的化学结构

同样可以计算出这 6 个分子的前沿时间与前沿体积, 与其相应的活性数据对比见表 7-8。从数据可以分析, 这 6 个主要组分在抗体亲和色谱柱上的保留顺序基本与其活性顺序相一致: 活性越高的分子 (IC_{50} 值越小) 在抗体柱上的保留时间越长 (前沿时间越大), 初步证明了用一种活性分子白皮杉醇为半抗原制备的多克隆抗体亲和色谱柱可以将藏药鬼箭锦鸡儿中的其他几种活性成分有效地识别开来。这里主要体现在对 3 种有一定活性的分子具有较强的特异性亲和能力并能根据其活性强弱对它们有效地识别 [即白皮杉醇 (piceatannol)、紫柳因 (butein) 与槲皮素 (quercetin)], 而对其他活性较低的分子则相对保留较弱, 如对于活性很低的分子 6 则几乎没有保留作用 (前沿时间为 5min, 即该分子在亲和柱上很快达到饱和吸附)。

表 7-8 FAC 测定结果及提取物中 6 个未知化合物的活性

化合物	t/min	$V/\mu\text{L}$	$(V-V_0)/\mu\text{L}$	$IC_{50}/(\text{mol/L})$
1	45.3	453.0	409.0	4.9×10^{-6}
2	21.9	219.0	175.0	1.0×10^{-5}
3	19.6	196.0	152.0	1.5×10^{-5}
4	8.3	83.0	39.0	1.1×10^{-3}
5	7.2	72.0	28.0	1.2×10^{-3}
6	5.0	50.0	6.0	5.9×10^{-2}

当然, 正如前面所讨论过的一样, 这里各个组分的含量 (浓度) 对前沿体积也有一定的影响 [见式 (7-2)], 但因为这 6 个组分是质谱检测到的几个主要组分, 从质谱定量可以得到这 6 个化合物在乙酸乙酯粗提物部分的含量在 2%~8% 之间变化。表 7-9 给出了根据用质谱外标法定量所得的各个组分的大致含量, 并根据式

(7-2) 计算各个组分所对应的解离常数 K_d ，这里以半抗原白皮杉醇为参考物，将它与抗体之间的相互作用强度定义为 1，则可以计算出其他分子与抗体之间的相对作用强度，从而进一步定量地说明抗体柱对各个组分的亲和作用强弱。

表 7-9 PcAb 强度对比

化合物	大致含量/%	K_d (估计)/(mol/L)	相对结合强度	IC_{50} /(mol/L)
1	7.6	1.7×10^{-5}	1.00	4.9×10^{-6}
2	5.4	8.1×10^{-5}	0.21	1.0×10^{-5}
3	2.5	9.7×10^{-5}	0.18	1.5×10^{-5}
4	2.7	4.2×10^{-4}	0.04	1.1×10^{-3}
5	2	5.7×10^{-4}	0.03	1.2×10^{-3}
6	3	3.6×10^{-3}	0.005	5.9×10^{-2}

所得多克隆抗体体现出对半抗原最强的亲和作用，对于活性稍低的另外两种分子（紫铆因和槲皮素）的相对亲和强度则为 21% 和 18%，而对于活性更低的化合物 4 和 5 的相对亲和强度则降为 4% 和 3%，对于基本无活性的化合物 6 体现出极弱的亲和能力（0.5%）。从而得出初步结论，认为该多抗可以将中草药中的活性成分从大量的干扰信息中有效地筛选出来，并且可以根据各化合物活性高低顺序对它们进行识别：活性越高的分子在亲和柱上的保留能力越强，活性较低的分子的保留能力会减弱。

7.4.8 小结

本章主要通过前面从藏药鬼箭锦鸡儿中分离得到的一种抗 EGFR 抑制剂——3,4,3',5'-四羟基二苯乙烯（白皮杉醇）为半抗原与载体蛋白牛血清白蛋白（BSA）偶联后制备其多克隆抗体，利用该多克隆抗体来模拟靶酶——抗表皮生长因子受体（EGFR），直接从鬼箭锦鸡儿的乙酸乙酯部分将几种主要的抗 EGFR 活性成分识别出来。研究过程采用了徐筱杰课题组自行开发的亲和前沿色谱-质谱联用技术进行检测，大大提高了检测的灵敏度与方便性。实验结果表明，利用多抗的亲和前沿色谱柱来识别中草药粗提物中的不同活性成分基本可行的：抗体柱不仅能将活性成分从大量的非活性干扰成分中筛选出来，而且能根据其活性高低对它们进行有效的识别，活性越高的分子在亲和柱上的保留能力越强。

参 考 文 献

- [1] Chan N W C, Lewis D F, Rosner P J, Kelly M A, et al. Frontal affinity chromatography-mass spectrometry assay technology for multiple stages of drug discovery applications of a chromatographic biosensor. *Anal. Biochem.*, **2003**, 319: 1-12.
- [2] Wang J, Zhang B, Fang J, Sujino K, et al. Frontal affinity chromatography coupled to mass spectrometry: An effective method for k_d determination and screening of α -gal derivatives binding to anti-gal antibodies (igg). *J. Carbohydr. Chem.*, **2003**, 22: 347-376.

- [3] Schriemer D C. Biosensor alternative: Frontal affinity chromatography. *Anal. Chem.*, **2004**, 76: 440A-448A.
- [4] Chan N W C, Lewis D F, Hewko S, Hindsgaul O, et al. Frontal affinity chromatography for the screening of mixtures. *Comb. Chem. High Throughput Screen.*, **2002**, 5: 395-406.
- [5] Slon-Usakiewicz J J, Dai J-R, Ng W, Foster J E, et al. Global kinase screening. Applications of frontal affinity chromatography coupled to mass spectrometry in drug discovery. *Anal. Chem.*, **2005**, 77: 1268-1274.
- [6] Noble M E M, Endicott J A, Johnson L N. Protein kinase inhibitors: Insights into drug design from structure. *Science*, **2004**, 303: 1800-1805.
- [7] Zhu L, Chen L, Luo H, Xu X. Frontal affinity chromatography combined on-line with mass spectrometry: A tool for the binding study of different epidermal growth factor receptor inhibitors. *Anal. Chem.*, **2003**, 75: 6388-6393.
- [8] Luo H, Chen L, Li Z, Ding Z, et al. Frontal immunoaffinity chromatography with mass spectrometric detection: A method for finding active compounds from traditional chinese herbs. *Anal. Chem.*, **2003**, 75: 3994-3998.
- [9] Li Y, Li Z, Yuan K, Qu X, et al. Small molecules blocking the entry of severe acute respiratory syndrome coronavirus into host cells. *J. Virol.*, **2004**, 78: 11334-11339.
- [10] Makara G M, Athanasopoulos J. Improving success rates for lead generation using affinity binding technologies. *Curr. Opin. Biotechnol.*, **2005**, 16: 666-673.
- [11] Ng E S M, Yang F, Kameyama A, M. Palcic M, et al. High-throughput screening for enzyme inhibitors using frontal affinity chromatography with liquid chromatography and mass spectrometry. *Anal. Chem.*, **2005**, 77: 6125-6133.
- [12] Hodgson R J, Besanger T R, Brook M A, Brennan J D. Inhibitor screening using immobilized enzyme reactor chromatography/mass spectrometry. *Anal. Chem.*, **2005**, 77: 7512-7519.
- [13] WHO. Cumulative Number of Reported Probable Cases of Severe Acute Respiratory Syndrome (SARS). <http://www.who.int/csr/sarscountry/2003-05-07/2004-05-03>.
- [14] 卫生部. 通报全国内地非典型肺炎最新疫情 (附表) [R]. 2003-05-08.
- [15] Drosten C, Gunther S, Preiser W, van der Werf S, Brodt H R, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier R A, Berger A, Burguiere A M, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra J C, Muller S, Rickerts V, Sturmer M, Vieth S, Klenk H D, Osterhaus A D, Schmitz H, Doerr H W. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med.*, 2003, **348**: 1967-1976.
- [16] Ksiazek T G, Erdman D, Goldsmith C S, Zaki S R, Peret T, Emery S, Tong S, Urbani C, Comer J A, Lim W, Rollin P E, Dowell S F, Ling A E, Humphrey C D, Shieh W J, Guarner J, Paddock C D, Rota P, Fields B, DeRisi J, Yang J Y, Cox N, Hughes J M, LeDuc J W, Bellini W J, Anderson L J, SARS Working Group. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med.*, 2003, **348**: 1953-1966.
- [17] 杨新艳, 徐, 钟南山. SARS 冠状病毒的分子生物学研究进展. 国外医学呼吸系统分册, 2005. **25** (1): 12-15.
- [18] Rota P A, Oberste M S, Monroe S S, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, 2003, **300** (5624): 1394-1399.

- [19] Marra M A, Jones S J, Astell C R, et al. *The genome sequence of the SARS-associated coronavirus*. Science, 2003. **300** (5624): 1399-1404.
- [20] Qin E d, ZHU Qingyu, YU Man, et al. *A complete sequence and comparative analysis of a SARS-associated virus (isolate BJ01)*. Chinese Science Bulletin, 2003, **10** (48): 941-948.
- [21] Krueger D K, Kelly S M, Lewicki D N, et al. *Variations in disparate regions of the murine coronavirus spike protein impact the initiation of membrane fusion*. Virology, 2001, **75** (6): 2792-2802.
- [22] Gallagher T M B, M J. *Coronavirus spike proteins in viral entry and pathogenesis*. Virology, 2001, **279** (2): 371-374.
- [23] Nelson G W, Stohlman S A, Tahara S M, et al. *High affinity interaction between nucleocapsid protein and leader/ intergenic sequence of mouse hepatitis virus RNA*. General virology, 2000, **81** (Pt1): 181-188.
- [24] Chen H, Wurm T, Britton P, et al. *Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell*. Virology, 2002, **76** (10): 5233-5250.
- [25] 徐庆, 刘卫兵, 罗伟生. 抗 SARS 药物的研究现状与展望. 华夏医学, 2004, **17**: 635-637.
- [26] 赵景民, 周光德, 孙艳玲等. 1 例地方传染性非典型肺炎病例病理及病原学发现. 解放军医学杂志, 2003, **28** (6): 14- 17.
- [27] Cinatl J, Morgenstern B, Baner G, et al. *Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus*. Lancet, 2003. **361**: 2045- 2046.
- [28] Wang Z M, Zhu X L, Cui X L, et al. *Screening of traditional Chinese remedies for SARS treatment*. Proceeding of international science synposium on SARS. Beijing China, 2003: 205.
- [29] 王重庆, 李云兰, 李德昌, 陈劲秋, 周先碗, 郝富英, 廖助荣, 袁洪生编. 高级生物化学实验教程. 北京: 北京大学出版社, 1994.



第 8 章 药用天然产物的筛选与分离一体化技术

亲和前沿色谱-质谱 (FAC-MS) 联用可以快速筛选出中药中的活性物质, 但有些靶点蛋白是很难获得的, 而且此种方法只适用于筛选, 并不适于分离物质。即使确定了活性化合物, 想要得到实物, 往往还是要采用传统的分离提取方法, 目前使用的常规分离材料 (正相、反相或离子交换型) 的选择性都不是很理想, 致使分离效率较低、耗时较长。许多研究工作曾致力于提高分离材料的选择性。为分子量身定做模板, 即分子烙印 (molecular imprinting, 又称分子印迹), 是近年发展起来的能够制备出对目标分子具有特定识别能力的聚合物材料分子烙印聚合物 (molecularly imprinted polymer, MIP, 又称分子印迹聚合物) 的一种技术。分子烙印聚合物又称为模拟受体, 对模板分子及其结构类似物有选择性的识别能力, 可以选择性地分离目标产物。MIP 制备简单, 性质稳定, 耐酸、碱、有机溶剂, 可以重复多次使用, 是一种很有发展前途的新型仿生材料, 已经成为化学工作者的研究热点之一。

8.1 分子烙印聚合物介绍

8.1.1 方法起源及概述

分子识别在生物进化中起着特别重要的作用, 选择性是分子识别的重要特征。人类免疫系统中的抗体就可识别抗原。在 20 世纪 40 年代, 著名的诺贝尔奖获得者 Linus Pauling 在研究抗原-抗体的作用时提出了以抗原为模板合成抗体的思想。虽然 Pauling 的理论被后来的“克隆选择理论”推翻, 但却为分子烙印的发展奠定了一定的理论基础。1972 年, 德国 G. Wulfd 研究小组首次报道了人工合成分子烙印聚合物 (MIP), 成功地用于葡萄糖类衍生物的手性分离。1984 年, 瑞典学者 K. Mosbach 成功地将模板分子与功能单体的非共价作用引入 MIP 的合成, 使得 MIP 的合成简单容易。特别是 1993 年, K. Mosbach 在“Nature”杂志上发表了有关茶碱分子烙印聚合物的文章后, 分子印迹技术 (molecular imprinting technique, MIT) 的研究迅速发展。从 1994 年开始有关分子烙印的论文数急剧上升。1997 年在瑞典 Lund 大学成立了国际性的分子印迹协会 (Society for Molecular Imprinting, SMI)。根据 SMI 的统计, 全世界至少有 100 个以上的学术机构和企事业团体在从事 MIT 的研究及开发工作。由于其优越的分子识别性能, MIP 已成为科研工作者的研究热点之一。

MIP 的分子识别来源于其内部存在的与原模板分子的空间结构互补、官能团相互作用（氢键、静电力或范德华力等）的孔穴（图 8-1）。目前，许多分子（如氨基酸及其衍生物、药物分子、糖类化合物及糖类的衍生物）的 MIP 已被成功地制备和研究。

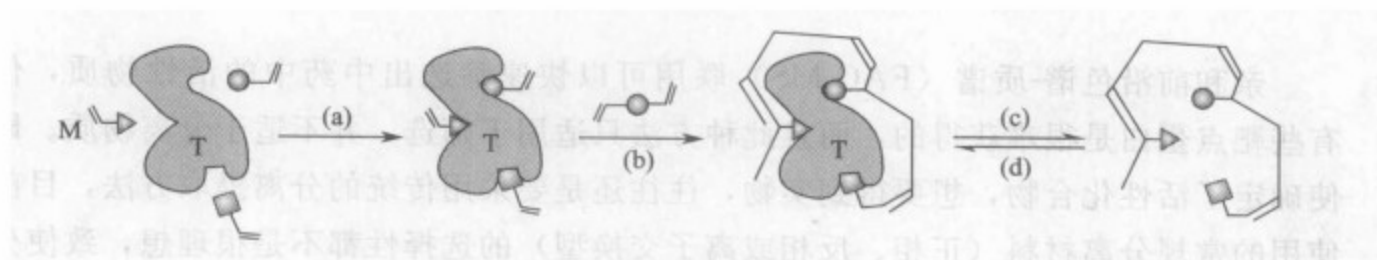


图 8-1 MIP 孔穴形成示意

- (a) 模板分子 (T) 和功能单体 (M) 形成预聚复合物；(b) 与过量的交联单体共聚；
(c) 模板分子从印迹结合位点脱离；(d) 模板分子在印迹结合位点再结合

8.1.2 分子烙印聚合物的制备方法

要想形成模板分子的印迹，需要功能单体与模板分子结合，然后在一定的致孔剂中加入交联剂引发聚合反应，使模板分子与功能单体形成的复合物固化，再将加入的模板分子除去，即得到模板分子的烙印聚合物。

下面从模板分子与功能单体的作用方式、分子烙印聚合物制备的相关因素、MIP 的制备技术 3 个方面对 MIP 的制备做详细介绍。

8.1.2.1 模板分子与功能单体的作用方式

MIP 的分子识别来源于 MIP 与模板分子在化学官能团、空间结构上的相互匹配，在官能团上的相互匹配是由功能单体达到的。模板分子与功能单体的作用方式见图 8-2，总体上可分为两类：共价分子印迹和非共价分子印迹。共价分子印迹中，模板分子与功能单体以共价作用结合形成复合物，这种共价键必须是可逆的（如酯、缩酮、席夫碱等），以便能将模板分子从聚合物中除去。非共价分子印迹中模板与功能单体通过氢键、静电作用、疏水作用和金属离子的络合等非共价键形成复合物。另外还可综合共价与非共价作用方式。

共价分子印迹又称为预组织法（preorganization），由 G. Wulff 等最早提出，其过程可用图 8-3 表示。利用 4-乙烯基苯硼酸与单糖苷相邻两个羟基形成硼酸酯的反应形成 2:1 的共价化合物，然后用交联剂将此化合物固定在聚合物中，最后用稀酸水解聚合物中的硼酸酯键，将模板分子除去。得到的聚合物可将模板分子（一种糖异构体）与其光学异构体分开。利用席夫碱和缩酮，G. Wulff 还制备了氨基酸衍生物及酮的分子烙印聚合物。

M. J. Whitecombe 等成功地用可逆共价方法制备了单官能团分子胆固醇的分子烙印聚合物。胆固醇分子的羟基与碳酸 4-乙烯基苯酯官能团共价结合，交联固化在聚合物中，然后水解切断碳酸 4-乙烯基苯酯官能团与胆固醇分子之间的化学键，

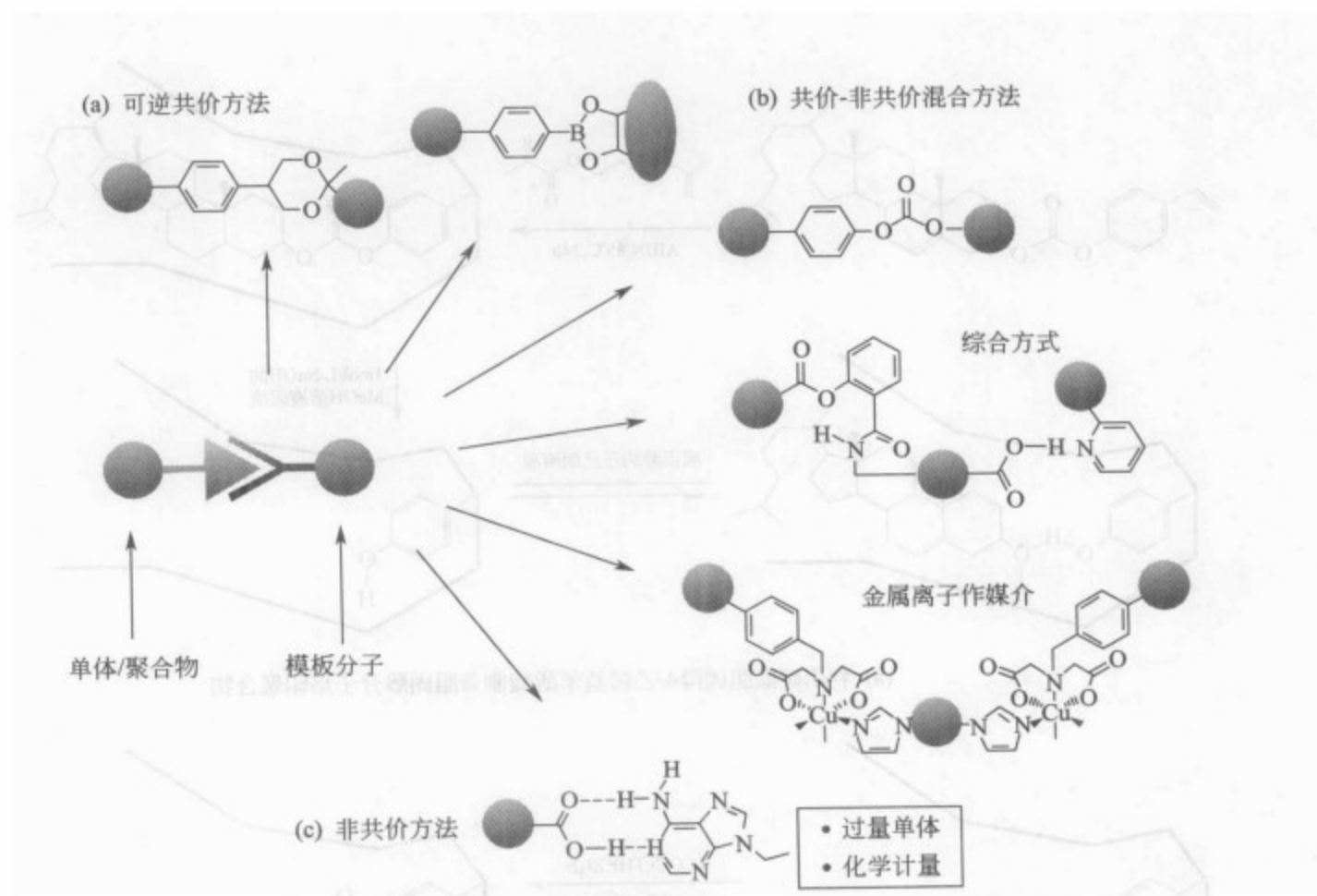


图 8-2 把结合或催化官能团放置在体型分子印迹位点的限定位置策略

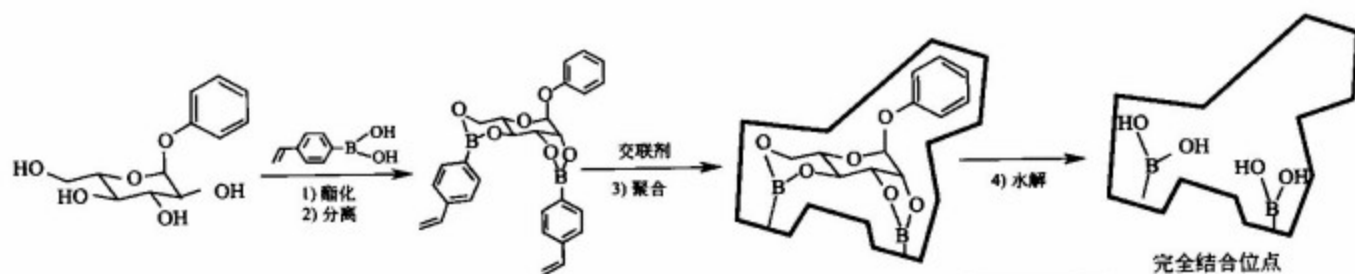
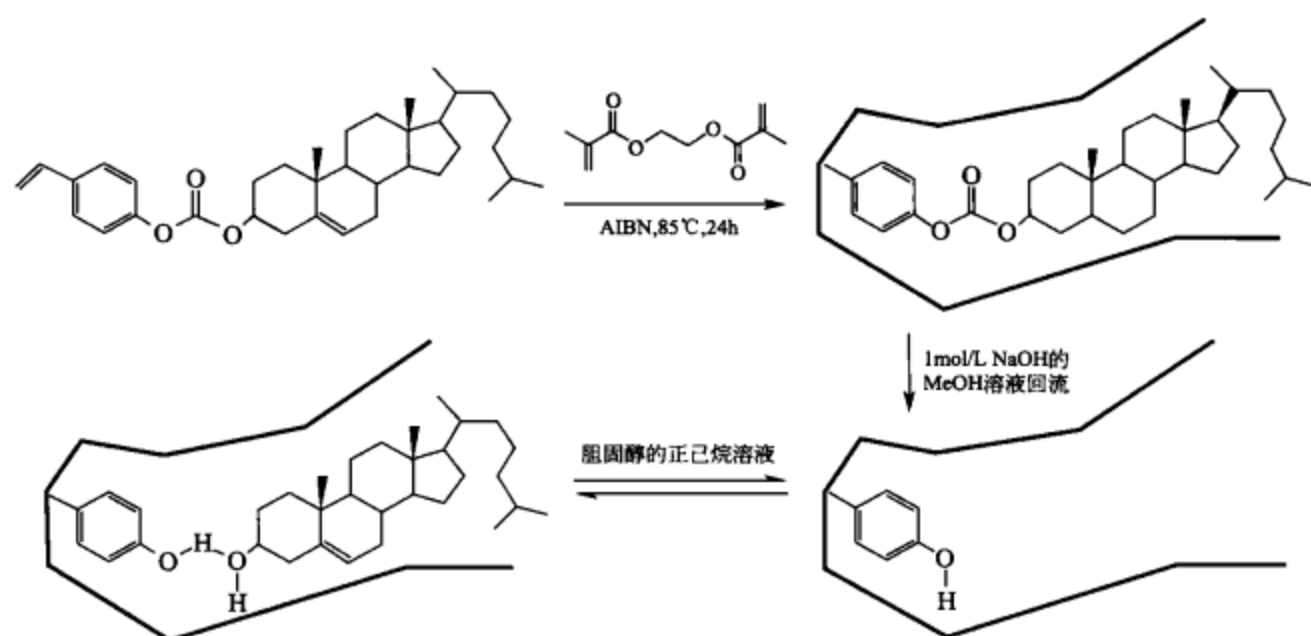


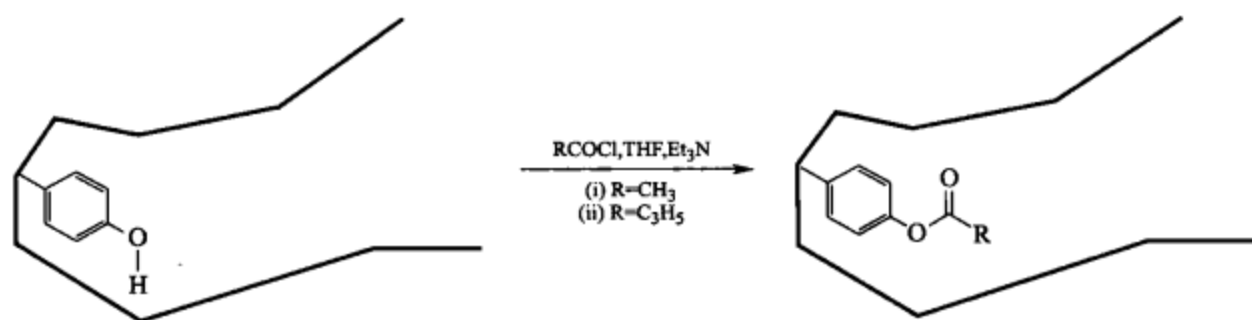
图 8-3 基于共价结合的糖分子烙印聚合物

除去模板胆固醇分子并失去一分子二氧化碳，得到含有酚羟基官能团的聚合物（图 8-4）。此分子烙印聚合物中含有可识别胆固醇分子的孔穴，聚合物的酚羟基官能团以氢键作用识别原来的模板分子胆固醇。在聚合时是共价反应，但识别时是非共价作用。Scatchard 分析表明，所制备的聚合物具有单一的离解常数 ($K_d = 0.59 \text{ mmol/L}$) 和可观的结合位点数 ($114 \mu\text{mol/g}$ 聚合物)，此聚合物的识别特性完全类似于真正的生物受体或其他模拟受体。

Mohammad A. Khasawneh 等报道了共价法制备药物分子去甲替林 (nortriptyline) 的分子烙印聚合物，模板分子功能化变为碳酸 4-乙烯基苯酯。MIP 对化合物的识别性能与化合物和模板分子的结构相似度相关，其 HPLC 毛细管柱可以将去甲替林从一个由去甲替林和其结构类似物组成的化合物库中筛选出来。



(a) 利用碳酸胆固醇4-乙烯基苯酚酯制备胆固醇分子烙印聚合物



(b) 酰化作用识别位点的化学修饰

图 8-4 胆固醇分子烙印聚合物的制备

共价法的优点是聚合中可获得在空间精确固定排列的结合基团，结合位点具有特异性及均一性。但寻找合适的可逆共价作用比较困难，化学合成步骤也较为复杂。

目前大多数分子烙印聚合物是用非共价法合成的，非共价法又叫自组装法 (self-assembling)。非共价分子印迹最早由 K. Mosbach 等提出，他们以 L-苯丙氨酸的衍生物苯丙氨酰苯胺为模板分子、甲基丙烯酸为功能单体，合成了对 L-苯丙氨酰苯胺有选择性的 MIP，证明用非共价作用也可制备出具有高选择性的 MIP。

非共价方法制备比较简单，它利用模板分子与功能单体之间的氢键作用、静电作用或疏水作用等形成复合物，分子间的各种非共价作用可组合产生协同效果。这一分子自组装过程模拟了生物体系中的分子间相互作用过程，功能单体将成为 MIP (受体) 识别位点部分。MIP 的迅猛发展也得益于 K. Mosbach 提出的非共价分子印迹法。

当然，非共价方法也具有一定的局限性。非共价印迹中，由于模板分子与功能单体之间的作用较弱，往往需加入过量的功能单体以利于功能单体与模板分子形成

复合物，而由于模板分子与功能单体的相互作用处于动态平衡之中，导致所制备的聚合物形成了对模板分子具有不同亲和力的结合位点。非共价分子烙印聚合物作用位点的非均一性与抗体的“多抗性”（polyclonality）相似，所以也称为 MIP 的“多抗性”。当 MIP 用作色谱固定相时，样品的色谱峰有严重的扩展或拖尾现象，其原因也与 MIP 的“多抗性”有关。

Jens U. Klein 等混合采用共价和非共价法，烙印了三肽 Lys-Trp-Asp。采用酰氯与氨基共价反应，而吡啶与羧基非共价结合（图 8-5），合成的 MIP 可以区分非常相似的氨基酸序列，其特异性类似于抗体和其他生物受体。

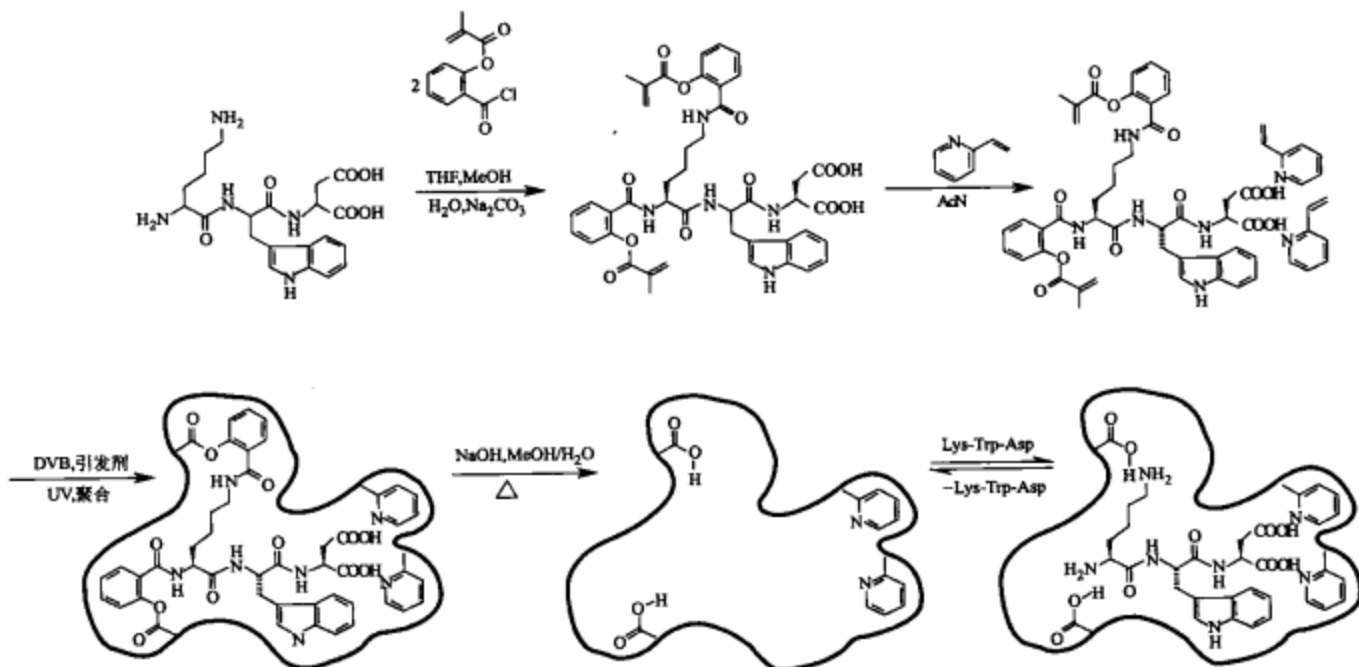


图 8-5 利用 2-甲基丙烯酸氧基苯甲酰氯共价修饰 Lys-Trp-Asp 及制备分子烙印聚合物

8.1.2.2 MIP 制备相关因素

这里的制备相关因素是指较通用的一些条件。

(1) 功能单体 一般认为增强功能单体与模板分子的作用有利于分子印迹的形成，有利于 MIP 的识别与结合能力，因此应依据模板分子选择合适的功能单体。模板分子与功能单体相互作用的位点一般为 2~3 个，过多的结合位点并不合适。在非共价分子印迹中，由于单体与模板分子处于动态平衡的自组装过程，为了使其形成稳定的复合物，往往要加入过量的功能单体。但单体量并不是越多越好，单体量过多时，MIP 的非选择性位点增加，其选择性将下降。一般功能单体的用量按其模板分子的结合位点数的化学计量关系过量 1 倍即可。

图 8-6 列出了使用过的一些功能单体。其中，甲基丙烯酸（MAA）是使用较多的功能单体，因为 MAA 的羧基既能接受电子又能给出电子，可与氨基、羧基、酰胺和羟基等多种功能团形成静电或氢键作用。取代乙烯基吡啶（2-VPy 和 4-VPy），丙烯酰胺（AA）也用得较多。可以通过设计合成新的功能单体提高 MIP

的分子识别能力。一些新的功能单体可参阅 A. J. Hall 等的综述文章。事实上，目前用于分子烙印的功能单体种类还很少，只是一些具有特殊结构和作用基团的分子能够直接作为模板分子用于 MIP 的制备。一般而言，如果目标分子结构不具有一定的刚性，作用基团很少或极性很大而不能溶于弱极性致孔溶剂，那么这个化合物的 MIP 是很难制备的。

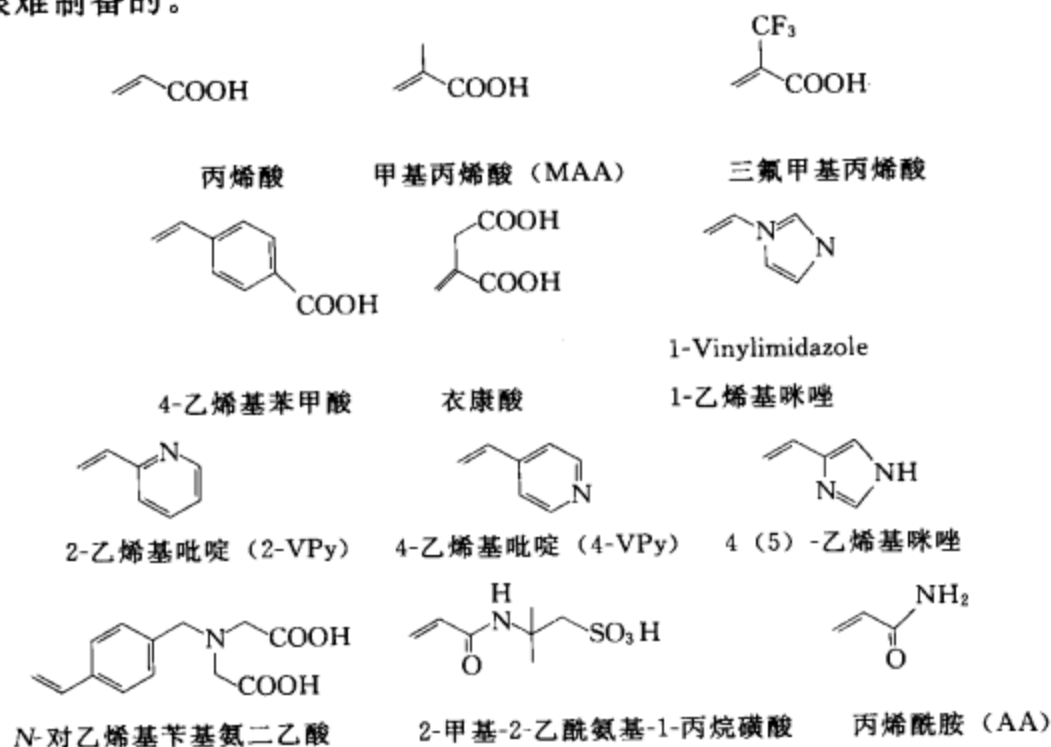


图 8-6 一些功能单体的化学结构

(2) 交联剂 在 MIP 制备过程中，为了将模板分子固定在聚合物中并保持聚合物孔穴的空间构型，要求聚合物有较高的硬度。但当聚合物在溶液中进行分子识别时，为了便于目标分子进入孔穴，又要求聚合物有一定的柔性，所以交联剂用量不能太大。另外，交联剂用量过大时，模板分子的比例相对降低，MIP 的载量也会相应减小。在非共价分子印迹中，交联剂的用量比共价分子印迹高一些，一般功能单体与交联剂的摩尔比为 1 : 5。

一些交联剂的化学结构见图 8-7。常用的双功能交联剂有乙二醇二甲基丙烯酸酯 (EDMA)、二乙烯基苯 (divinylbenzene, DVB)。还有一些含两个以上丙烯酸基团的交联剂，最常用的为二羟甲基丁醇三甲基丙烯酸酯 (TRIM)。

(3) 致孔剂及溶剂 在非共价分子印迹中，模板分子与功能单体之间靠静电和氢键等作用形成复合物，复合物的形成受溶剂影响较大。虽然有使用极性溶剂（如四氢呋喃、甲醇）的报道，但极性溶剂不利于复合物的形成，常用的溶剂为中等极性的乙腈和弱极性的氯仿、甲苯。受乙腈和氯仿溶解能力限制，很多化合物不能直接进行分子印迹。研究水或极性环境的分子印迹必将极大扩展 MIP 的应用范围。

MIP 的识别受溶剂影响很大，若要使 MIP 表现较好的选择性，往往需对所用溶剂进行优化。一般来说，致孔剂对 MIP 的识别有影响，在与致孔剂有关的溶剂

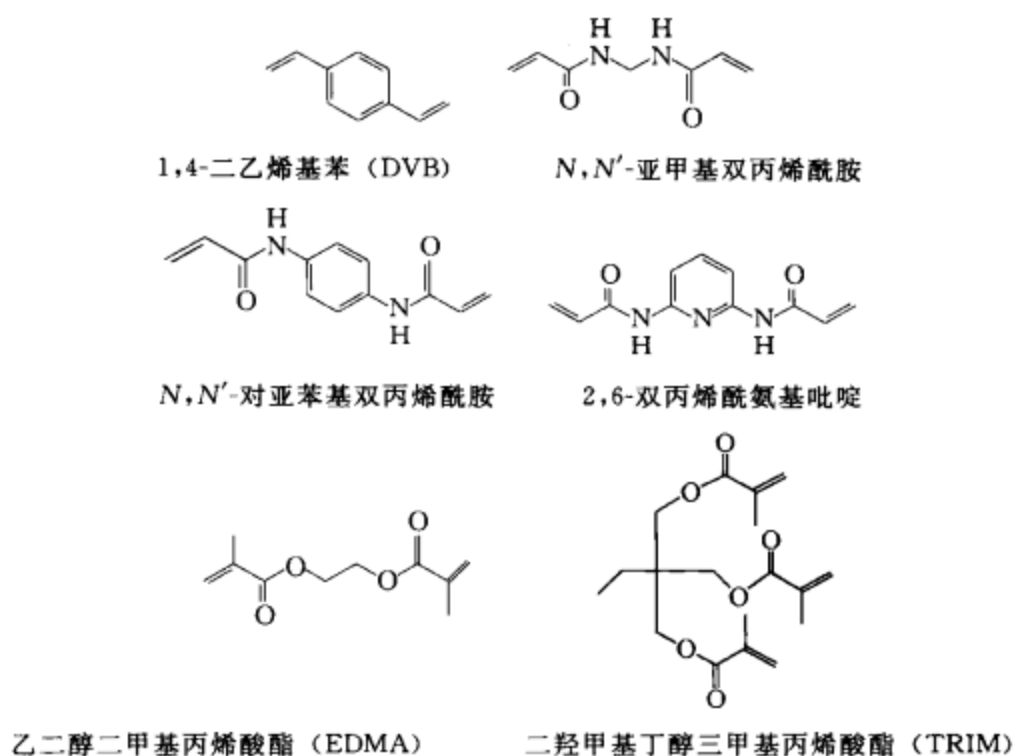


图 8-7 一些交联剂的化学结构

中 MIP 的选择性更好一些。

(4) 聚合温度 分子烙印聚合反应一般是自由基聚合，可在偶氮二异丁腈引发下进行热聚合，聚合温度 60~120℃。也可以采用低温光引发聚合。低温有利于功能单体与模板分子复合物的生成，因此低温光引发聚合制备的 MIP 表现更好的选择性，而且低温光引发聚合可用于制备热不稳定化合物的 MIP。但低温光引发聚合有时聚合物内部反应会不完全，造成聚合物结构不均匀，这需要反应后的热处理来解决。

(5) 模板分子的去除 聚合反应后，必须用甲醇-乙酸等强极性溶剂或更强的溶剂洗脱除去模板分子，留出空穴，才能使聚合物具有分子识别能力。模板分子的有效去除有利于分子识别，也避免了之后的应用中模板分子漏出 (leaking) 造成的干扰。

8.1.2.3 MIP 制备技术

分子烙印聚合物制备方法的研究主要是对材料的性能加以改善。以下对各种 MIP 制备技术做一介绍。

(1) 本体聚合 (bulk polymerization) 本体聚合法是最原始简单也最常用的方法。一般将模板分子、功能单体和交联剂按一定的比例加入致孔剂 (有机溶剂) 中，在加热或紫外光照的条件下引发自由基聚合反应，得到高度交联的块状聚合物，经破碎、过筛、浮选、干燥得到合适大小的聚合物颗粒，再用适当的方法除去模板分子，即可进行聚合物吸附或分离等性质的研究 (图 8-8)。该方法操作条件容易控制，实验装置简单，但聚合物磨碎过筛的过程烦琐费时，且得到的聚合物颗粒收率低 (约 50%)、形状不规则。尽管本体聚合法制备的 MIP 可在许多方面应

用，但用于色谱固定相时柱效往往较低，而且不能用于一些需要特定聚合物形态的场合。

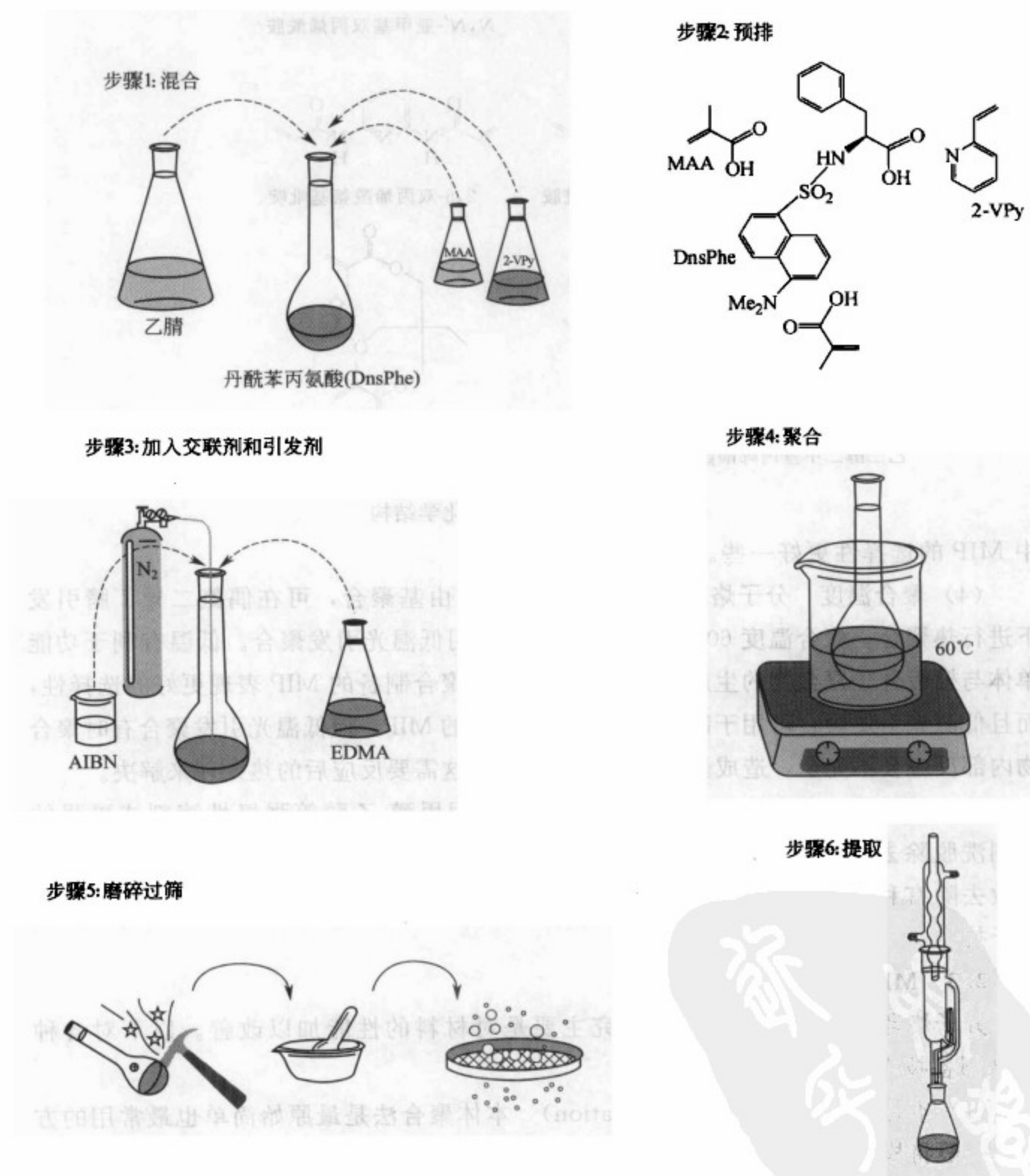


图 8-8 本体聚合法合成 MIP 步骤示意

功能单体：模板分子 (M : T) 的比大致是功能单体与模板分子作用基团化学计量比的两倍。但 Ecevit Yilmaz 等制备茶碱 (theophylline) 的 MIP 时，发现 M : T 即使高达 500 : 1，烙印聚合物仍然要比非烙印聚合物结合更多的模板分子。

这使昂贵的、难以大量得到的或溶解性差的分子的 MIP 的合成成为可能。

(2) 沉降聚合 (precipitation polymerization) 沉降聚合是将预聚合液加入过量的溶剂中聚合, 随着聚合反应的进行, 纳米级或微米级聚合物粒子即沉降出来 (图 8-9)。

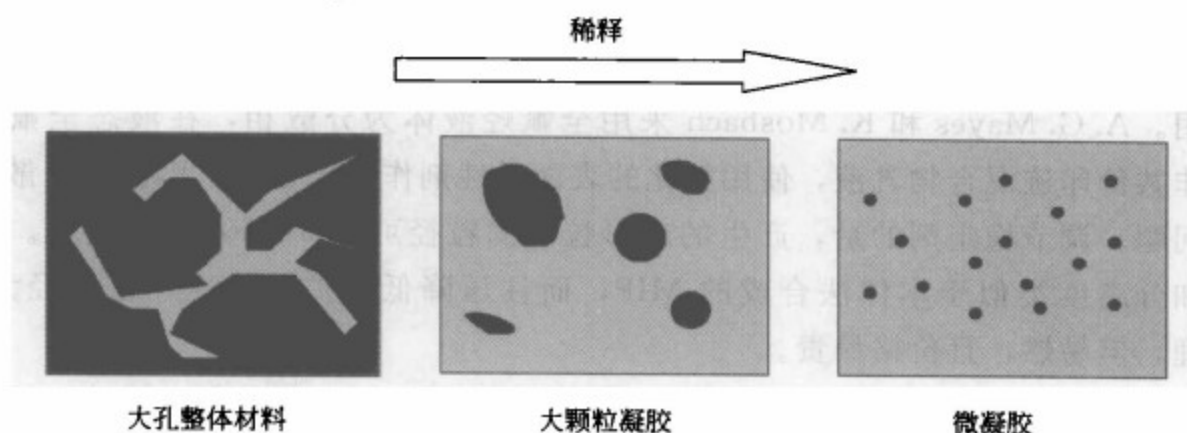


图 8-9 反应溶液稀释时聚合物结构从大孔整体材料到不连续微凝胶的变化

L. Ye 等用沉降聚合法制备了茶碱 (theophylline)、咖啡因 (caffeine) 和 17β -雌甾二醇 (17β -estradiol) 的 MIP 颗粒, 产生单分散的球形粒子, 方法快速、干净且产率高。以杀虫剂 2,4-滴为模板沉降聚合法制备的 MIP 可在水相溶液中选择性地结合模板分子。S. Boonpangrak 等将沉降聚合用于胆固醇分子和碳酸 4-乙烯基苯酯的半共价法, 合成的微球 MIP 的性能要优于本体共价法合成的 MIP (图 8-10)。沉降聚合法可以直接得到球形的 MIP 粒子, 但是由于是在大量溶剂中进行反应, 需要使用更多的模板分子。

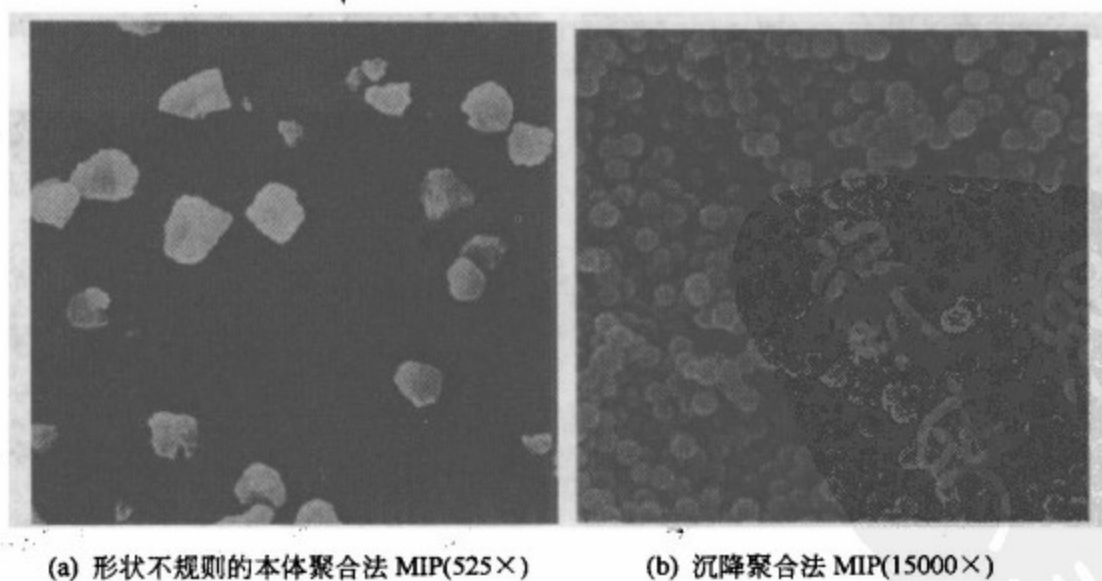


图 8-10 不同方法制备的 MIP 的扫描电镜 (SEM) 照片

(3) 乳液聚合 (emulsion polymerization) 乳液聚合是在加入表面活性剂的油-水两相体系中进行高分子聚合反应, 最后得到球形聚合物颗粒, 也可用于 MIP 的合成。N. Pérez 等在聚苯乙烯核外的聚二乙烯基苯壳层用乳液聚合法制备了胆固

醇的 MIP，以 12-对乙烯基苄氧羰基十二烷基硫酸吡啶鎓作为可聚合的表面活性剂分子，以 12-胆甾烯基氧羰基氧基十二烷基硫酸吡啶鎓为模板，这样一来，形成的 MIP 识别位点将位于聚合物颗粒的表面。

(4) 悬浮聚合 (suspension polymerization) 在通常的产生球形粒子的悬浮聚合中，水常作为分散介质，但水的强极性会极大地降低功能单体与印迹分子间的相互作用。A. G. Mayes 和 K. Mosbach 采用全氟烃液体为分散相，在液态全氟烃中形成非共价印迹混合物乳液，使用氟化的表面活性剂作为稳定剂，解决了分散相的干扰问题。调节稳定剂的量，产生的球形粒子的粒径可在 $5 \sim 50 \mu\text{m}$ 间变化。其选择性和分离度类似于本体法合成的 MIP，而柱压降低、扩散快速。全氟烃无毒、易处理，但易燃，且价格昂贵。

M. Zourob 等使用螺旋形微通道反应器，连续相使用全氟烃或矿物油，采用悬浮聚合法合成了粒径均一（变异系数低于 2%）的 MIP 微球（图 8-11）。

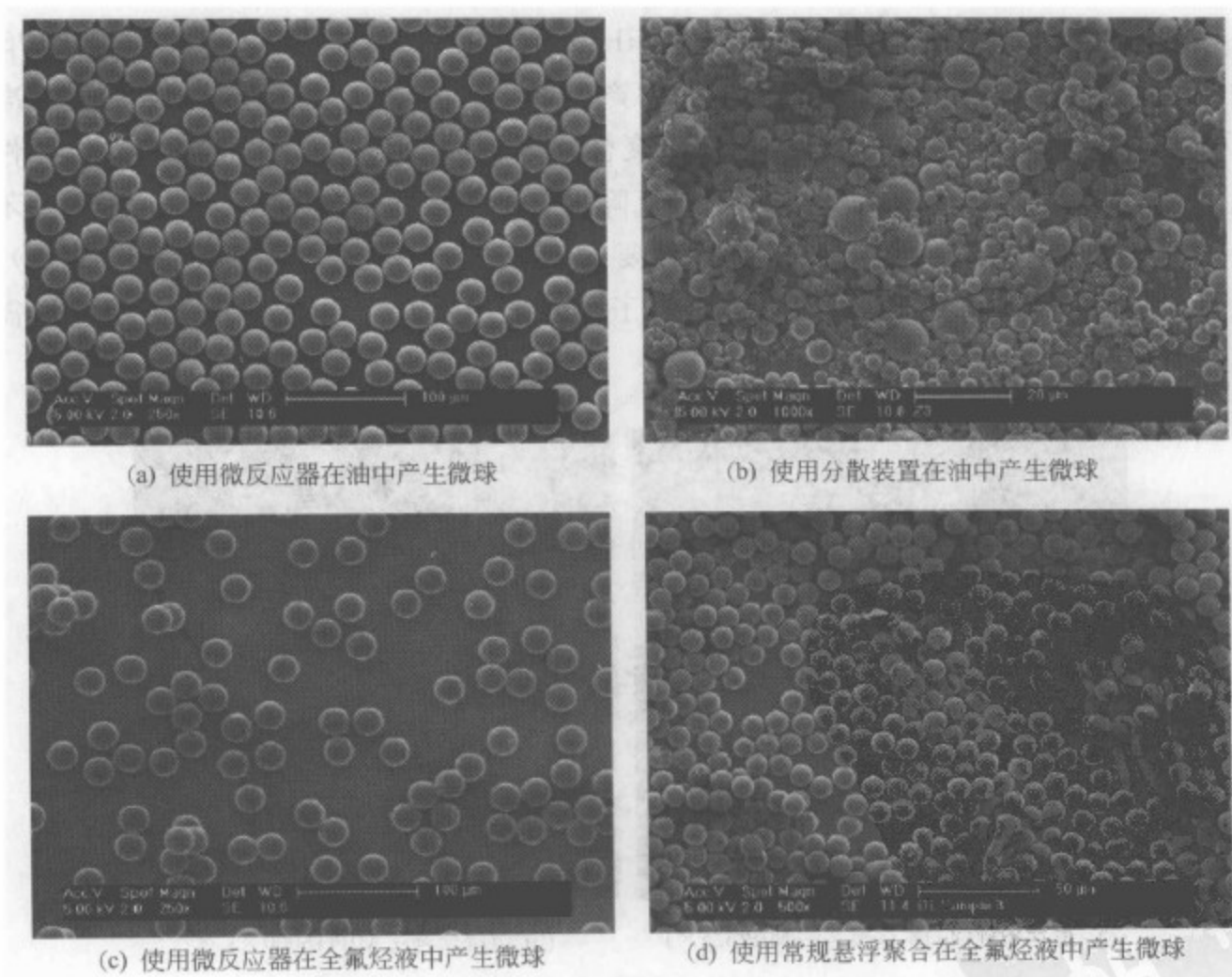


图 8-11 典型 MIP 微球的扫描电镜 (SEM) 照片

K. Hosoya 等采用两步膨胀法 (two-step swelling) 在水相中制备 MIP 微球，得到大孔体积的多孔穴粒子，可用于分离手性化合物。但其制备过程烦琐。

(5) 原位聚合 (in-situ polymerisation) J. Matsui 等在 1993 年将原位聚合技术用于分子烙印聚合物的制备, 即将聚合液装入不锈钢色谱柱中, 直接在柱中加热发生聚合反应。这样, 省去了破碎、过筛、装柱等步骤, 合成非常简单。之后其应用也较多。

(6) 表面印迹 通过对球形粒子表面进行修饰制备分子烙印聚合物材料是一个好的方法, 可以利用粒子本身的物理性质, 而不必过分关注聚合物的孔径、形态和其他结构特性。

这可以通过接枝 (graft) 技术达到。在多孔球形硅胶粒子表面进行硅烷化反应, 连接上双键, 然后再聚合的方法是接枝到 (grafting to) 表面方法 (图 8-12, Si-PMA)。但这种方法由于引发剂存在于溶液中, 溶液中也会发生聚合, 不能保证完全的表面聚合。C. Sulitzky 等将引发剂共价连接或非共价吸附于硅胶表面, 这样由于引发剂存在于硅胶表面, 聚合反应将主要在表面进行, 而且 MIP 层的厚度还可以调控。这种方法称为从表面接枝 (grafting from) 方法。图 8-13 为用此方法合成的含碳量不同的聚合物的扫描电镜照片。

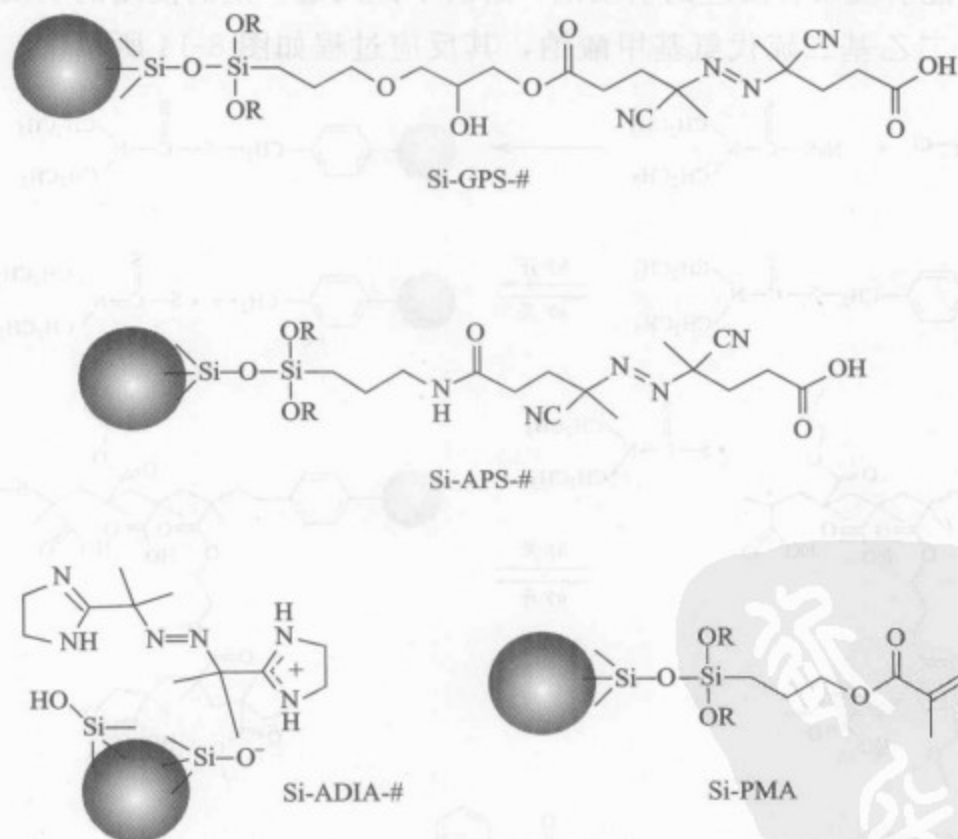


图 8-12 表面修饰。硅胶载体在偶氮引发剂 ACPA 作用下通过两步用 GPS 或 APS 共价修饰 (Si-GPS-# 或 Si-APS-#), 或在联脲偶氮引发剂 ADIA 作用下通过一步非共价修饰。Si-PMA 使用的是接枝到表面方法

通过使用固定化引发剂, 合成的 MIP 复合材料动力学性能得到改善。但由于引发剂只有一点连接于硅胶表面, 另一半解离的活性自由基处于溶液中, 溶液的聚合和凝

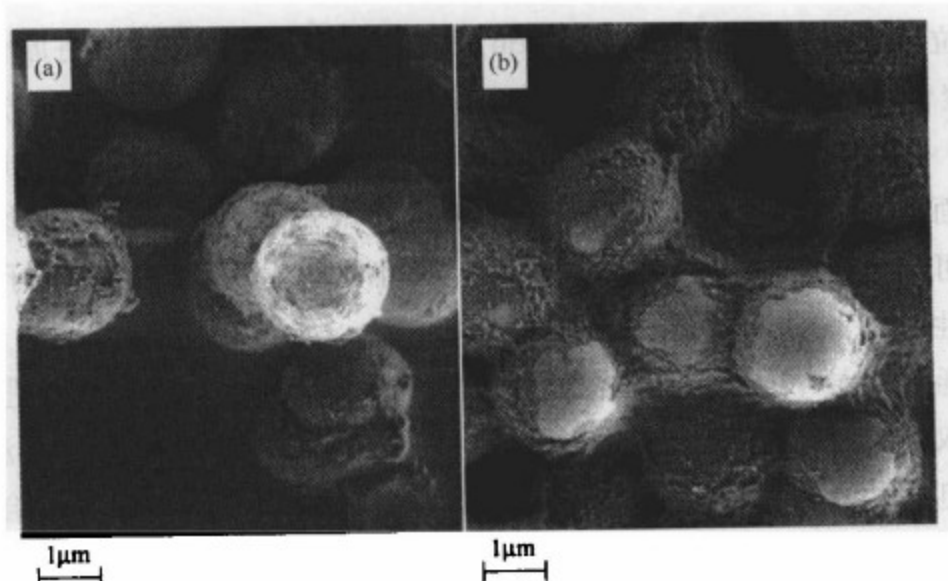


图 8-13 含碳量不同的接枝聚合物 Si-APS-3a 的扫描电镜 (SEM) 照片 (a) 含碳 3.6%；(b) 含碳 16%

胶化仍然难以避免。B. Bückert 等通过使用引发-转移-终止剂 (iniferter)，即解离的其中一个自由基不能引发聚合反应的引发剂，解决了此问题。他们使用的引发-转移-终止剂是 *N,N*-二乙基二硫代氨基甲酸钠，其反应过程如图 8-14 所示。

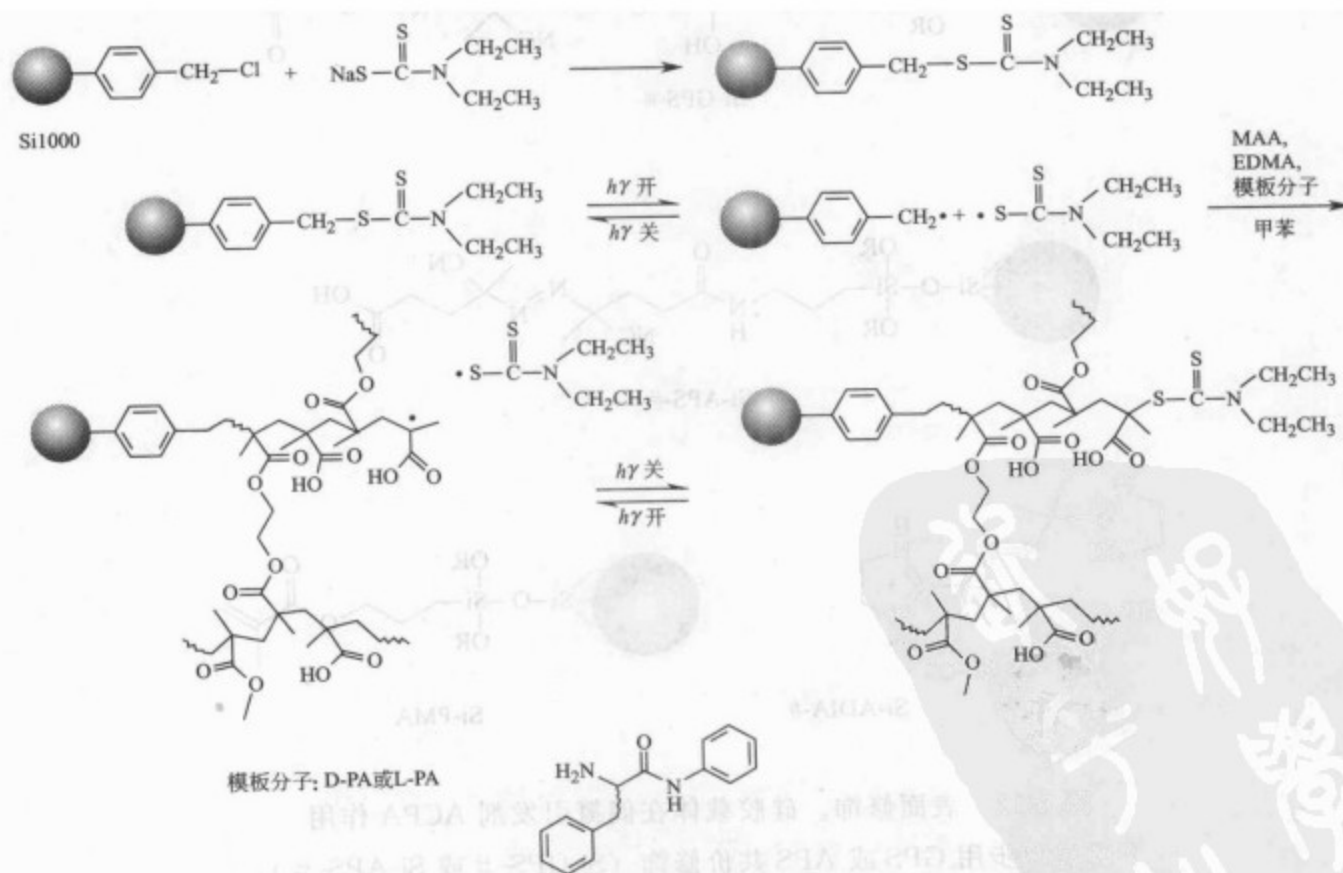


图 8-14 修饰负载的硅胶 (例如 Si1000) 或聚苯乙烯 (例如 Merrifield 树脂) 的反应。使用二硫代氨基甲酸盐作为引发-转移-终止剂，利用 *D*-苯丙氨酰苯胺 (D-PA) 或 *L*-苯丙氨酰苯胺 (L-PA) 作为模板分子，使甲基丙烯酸 (MAA) 与乙二醇二甲基丙烯酸酯 (EDMA) 接枝共聚，生成分子烙印共聚物

通过使用引发—转移—终止剂，还可能以任意顺序连续接枝使用两种不同模板分子的两层聚合物，或一层印迹、一层非印迹（图 8-15）。

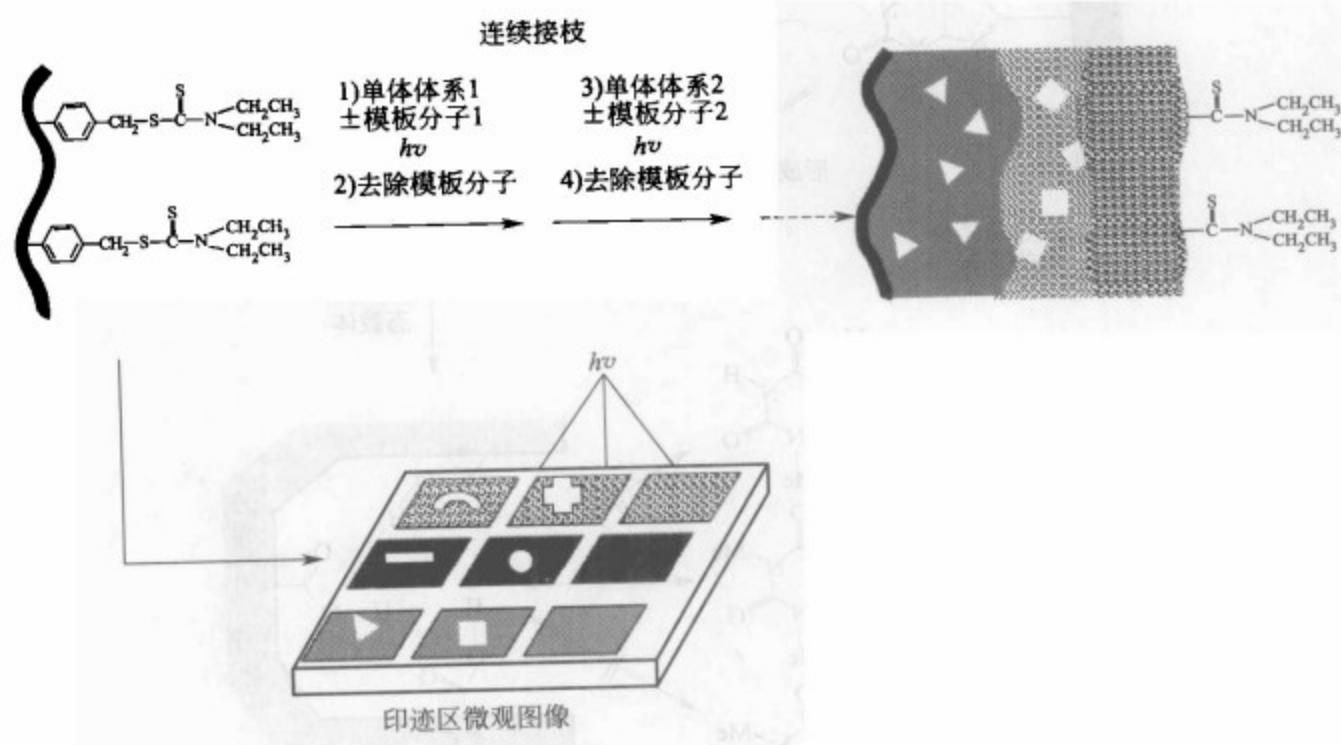


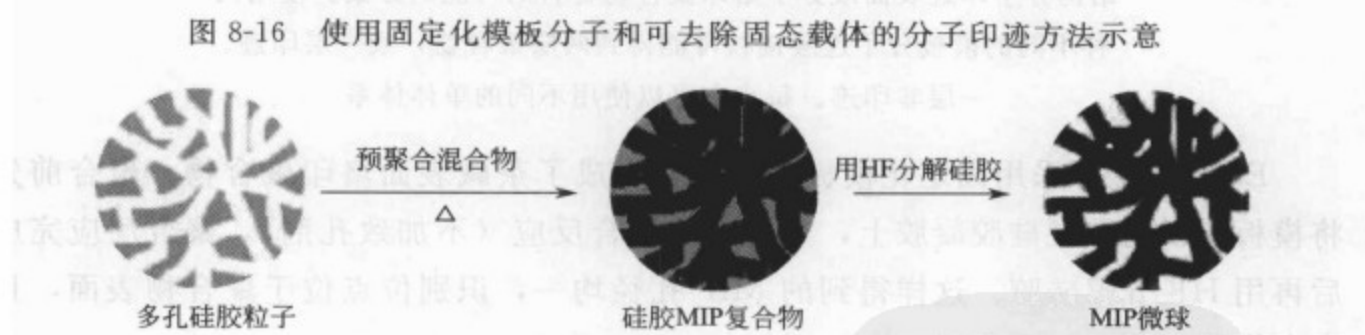
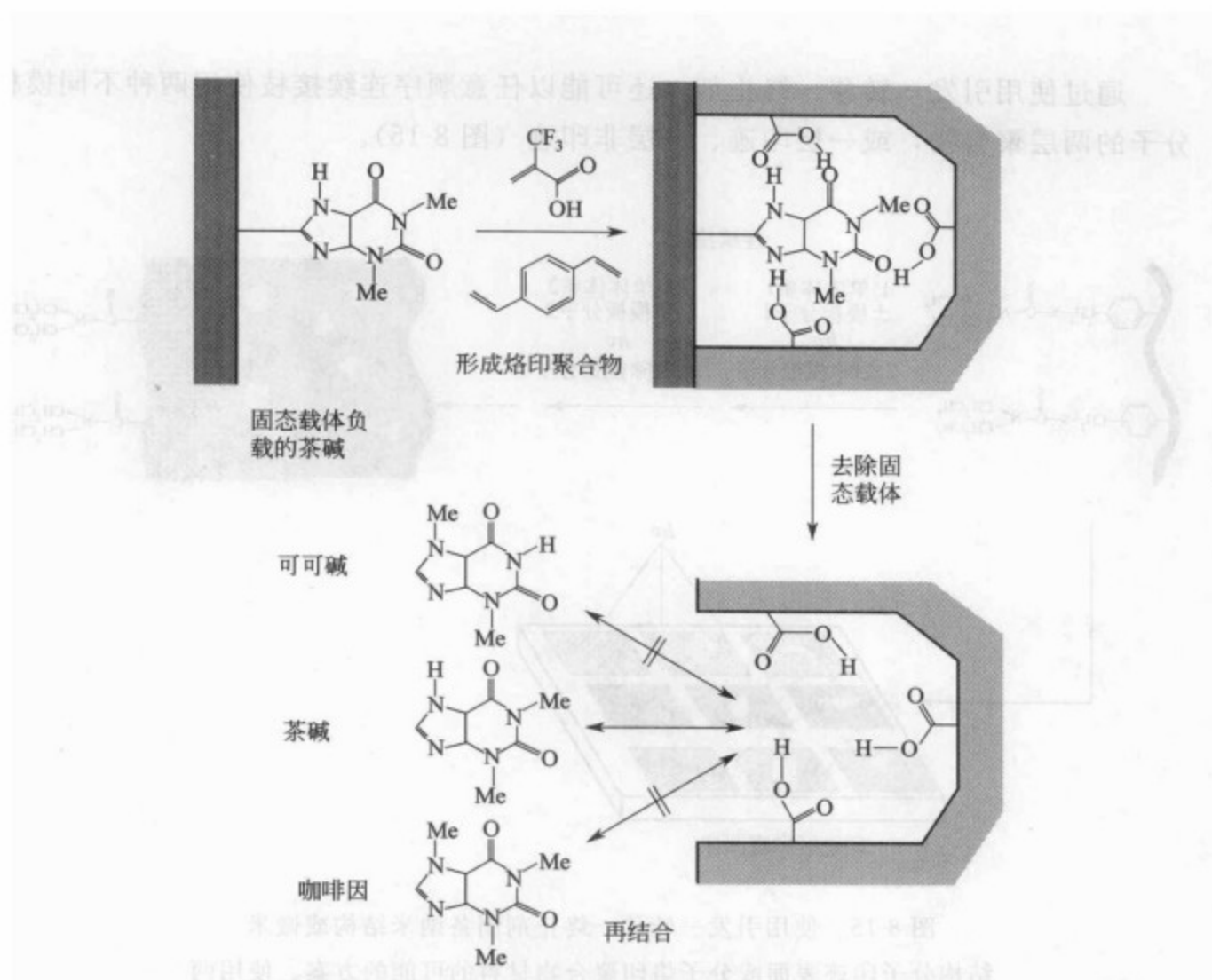
图 8-15 使用引发—转移—终止剂制备纳米结构或微米结构分子印迹表面或分子烙印聚合物材料的可能的方案。使用两种不同的模板分子连续接枝可能得到两层聚合物，或一层印迹、一层非印迹。每步中可以使用不同的单体体系

E. Yilmaz 等采用固定化模板分子法，合成了茶碱表面烙印聚合物。聚合前先将模板分子固定在硅胶凝胶上，然后进行聚合反应（不加致孔剂），聚合反应完成后再用 HF 分解硅胶。这样得到的 MIP 孔径均一，识别位点位于聚合物表面，且定向排列（图 8-16）。

E. Yilmaz 等采用了一种更简单的方法来合成硅胶-MIP 复合粒子。他们使用表面为 C_4 相的硅胶粒子，不对表面进行任何处理，直接加入预聚合液进行聚合反应。但为了获得单分散的硅胶-聚合物粒子，阻止粒子聚集，加入的预聚合液的量要小于硅胶的孔体积。合成的硅胶-MIP 复合粒子也可用 HF 溶去硅胶，得到 MIP 微球（图 8-17）。

对平面的表面印迹则形成 MIP 膜。MIP 膜的制备对于 MIP 在许多领域的应用（如传感器、膜分离）有重大意义。MIP 膜的制备方法一般是在一个模具内聚合，或将反应混合液涂在平的表面上聚合，得到膜形聚合物材料。

(7) 其他 MIP 还有其他一些制备方法，如使用功能单体化的有机硅烷、溶胶-凝胶 (sol-gel) 法等。



8.1.3 分子印迹机理研究

I. A. Nicholls 在分子印迹热力学理论及定量描述方面进行过有益的尝试。分子印迹聚合物与模板分子（配体）之间的识别受热力学定律控制，配体与受体之间形成复合物的自由能改变可由下式计算：

$$\Delta G_{\text{bind}} = \Delta G_{\text{t+r}} + \Delta G_{\text{r}} + \Delta G_{\text{h}} + \Delta G_{\text{vib}} + \sum \Delta G_{\text{p}} + \Delta G_{\text{conf}} + \Delta G_{\text{vdw}} \quad (8-1)$$

式中 ΔG_{bind} ——形成单体-模板分子复合物时 Gibbs 自由能的改变；

$\Delta G_{\text{t+r}}$ ——分子平动和转动引起的自由能的改变；

ΔG_{r} ——分子内旋被冻结时引起的自由能改变；

ΔG_{h} ——疏水相互作用引起的自由能改变；

ΔG_{vib} ——基团振动引起的自由能改变；

ΔG_{p} ——偶极基团相互作用的总自由能改变；

ΔG_{conf} ——构象变化引起的自由能改变；

ΔG_{vdw} ——非范德华力引起的自由能改变。

这个公式相对完善，但形式过于复杂，不便应用。基于 MIP 的实际情况，可简化为：

$$\Delta G_{\text{bind}} = \Delta G_{\text{t+f}} + \Delta G_{\text{r}} + \Delta G_{\text{vib}} + \sum \Delta G_{\text{p}} + (\Delta G_{\text{h}}) \quad (8-2)$$

I. A. Nicholls 等对功能单体与模板分子作用的焓和熵对自由能的贡献进行讨论，并设计选择了功能单体。M. J. Whitcombe 等用 NMR 测定了非共价印迹的亲数和常数，不同亲和力的配体和 MIP 的选择性可以联系起来。研究功能单体与模板的作用最常用的方法是 NMR、UV 谱学方法，但 FT-IR、X 射线晶体学方法也已经使用。用 UV 光谱甚至可以研究大量交联剂存在时预聚合液中复合物的形成。分子动力学模拟结合 NMR 和 IR 实验研究得到的限制条件，可确定不同溶剂中的 2,4-滴/4-乙烯基吡啶复合物的主要相互作用类型，并由实验得到证实，因此分子动力学模拟或许可以作为有用的工具来预测相关成分的相互作用。

对 MIP 结合性质的分析基于抗体和其他生物受体的评价方法，如模拟等温结合曲线的单 Langmuir 或多 Langmuir 模型是常用的方法。但 Langmuir 等温线并不能很精确地描绘非共价分子印迹聚合物的结合行为，因此其他的结合模型如 Freundlich 等温线和 Langmuir-Freundlich 等温线也被应用。化学计量学方法也用于描述和预测模板分子与 MIP 结合的程度，对不同溶剂在不同温度下的平衡结合

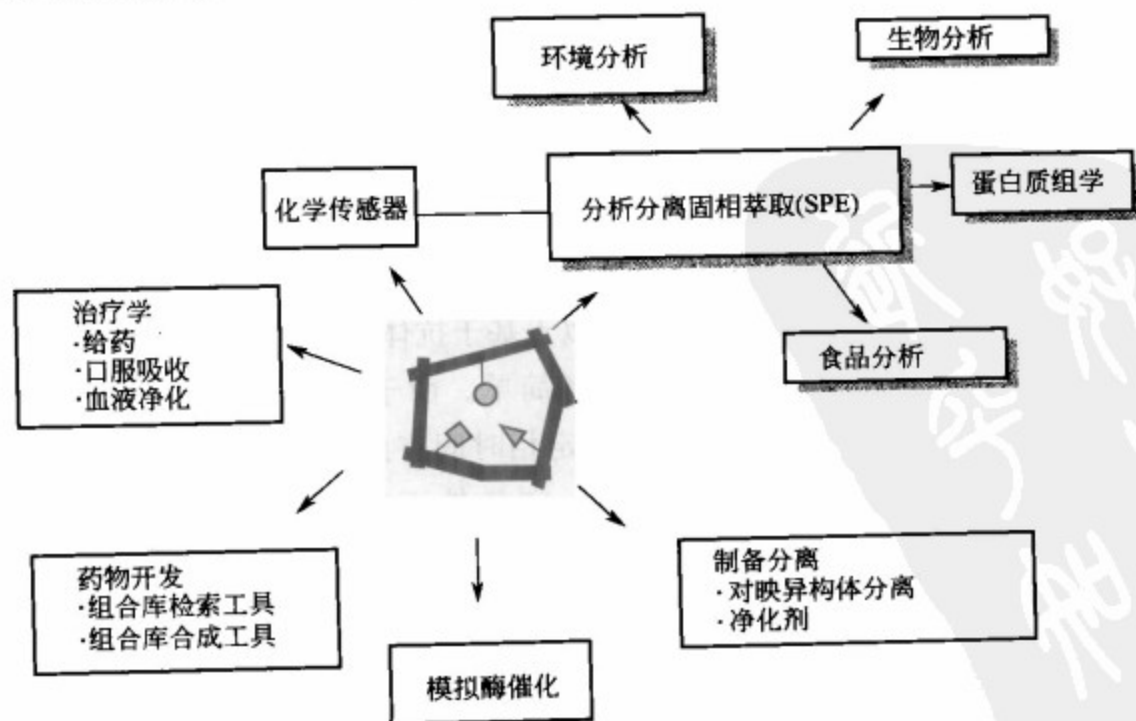


图 8-18 分子烙印聚合物 (MIP) 的主要应用

实验数据用偏最小二乘回归 (partial least-squares regression, PLSR) 法拟合, 可产生化学计量模型来描述模板分子的结合。

8.1.4 分子烙印聚合物的应用范围

分子烙印聚合物的应用范围示于图 8-18。

8.2 分子烙印聚合物的应用

分子烙印聚合物 (MIP) 制备简单, 所得聚合物在有机溶剂体系中更能够表现出其分子识别能力, 它不仅对原来的模板分子具有很高的亲和性, 而且对与其结构类似的化合物也表现出较高的结合能力, 而对于与模板分子结构相差较远的分子只表现出较弱的表面吸附作用。目前, MIP 作为分子识别材料, 除了用于分离 (手性分离)、免疫分析、传感器、模拟酶催化以外, 在固相萃取 (SPE) 方面的应用也越来越引起人们的注意。MIP 作为选择性固相萃取材料 (MISPE) 曾用于萃取生物流体中的药物分子, 萃取口香糖和烟草中的尼古丁, 萃取牛肝或水中的三嗪农药等。但至今尚未见到有关 MISPE 直接萃取中药中有效成分的报道。以下简单介绍一下关于 MIP 在分析化学领域中的一些应用。

8.2.1 固相萃取

分子烙印聚合物在分离领域中最实际和广泛的应用可能就是固相萃取 (solid phase extraction, 简称为 SPE, MIP 在这方面的应用可以简称为 MISPE)。国际上有多多个小组正从事这方面的研究工作, 主要是将 MISPE 应用于生物及环境样品中目标产物的分离, 据估计该技术可能在不远的将来能够用于实际工业生产中。SPE 技术作为一种有效的分析前样品预浓缩或纯化的方法, 已经广泛应用于医药、食品及环境等领域。与液相提取方法相比, 固相萃取方法处理样品快速, 重现性好, 且处理后的样品含杂质相对较少。另外, SPE 方法可以节省溶剂, 能实现自动化。MIP 是一类很有潜力的固相萃取材料, 与传统的分离材料相比, MIP 比 C18、离子交换或其他分离材料选择性好, 而且对目标分子的保留类似于基于抗体技术制备的亲和色谱介质, 但 MIP 比一般生物亲和材料稳定、价廉、制备简单。由于 MISPE 是基于亲和色谱分离机制, 因此不存在如同 MIP 在用于色谱固定相时固有的因色谱峰扩展而导致分离因子小的问题。而且, 还可以利用 MIP 在一定条件下存在的“交叉反应性” (cross reactivity), 将结构相似的一类化合物从一些复杂的体系中有效地萃取出来。

MISPE 最开始用于尿液中喷他脒 (pentamidine) 的在线富集, 到现在为止已经用于多种样品的分析分离 (见表 8-1), 如稀释过的或者经过蛋白沉淀后的生物样品, 水溶液体系中有机的分离, 一些生物组织或体液, 口香糖、烟草或发酵后产物中目标分子的分析分离等。

表 8-1 MIP 在固相萃取方面的应用概述

分 析 物	示 例
喷他脒(pentamidine)	稀释的尿液
莠去津(atrazine)	牛肝匀浆的氯仿提取液
灭草松(bentazone)	标准水溶液
达非那新(darifenacin)	血浆-乙腈(体积比 1:1)
羟基香豆素(hydroxycoumarin)	尿液
吲哚乙酸(indoleacetic acid)	标准氯仿溶液
红曲色素(monascus pigments)	发酵液的乙酸乙酯提取液
尼古丁(nicotine)	口香糖的乙酸乙酯提取液; 烟草的甲醇提取液
普萘洛尔(propranolol)	狗血、鼠胆汁和人尿; 标准水溶液
西玛津(simazine)	标准水溶液
沙美利定(sameridine)	人血提取物的庚烷-乙醇(体积比 9:1)溶液
他莫昔芬(tamoxifen)	人血和人尿
茶碱(theophylline)	人血清的氯仿提取液; 标准氯仿溶液
三嗪类农药(triazine herbicides)	自来水的 C ₁₈ 柱提取物的有机溶液; 尿液、苹果提取物的 C ₁₈ 柱提取物的乙腈溶液

MIP 在有机相中往往表现出更好的识别能力，而对于水溶液体系中的识别则需要更仔细地选择条件。已经知道，为了有利于功能单体与模板分子之间的正相极性作用如氢键作用，烙印过程一般是在弱极性有机溶剂中完成的，在含水体系中这种极性作用会减弱，憎水作用增强，从而使 MIP 的选择性降低。因此直接固相萃取生物流体，对分子烙印技术而言是一个比较难处理的问题。Sellergren 在用 MIP 萃取尿液中的喷他脒时，是以含 20%（体积分数）磷酸盐缓冲液的乙腈溶液为吸附和洗脱液，采取改变缓冲液 pH 值的方法使喷他脒被保留和识别，当缓冲液 pH 值为 5 时样品吸附，在缓冲液 pH 值为 3 时洗脱。Matsui 等研究萃取水溶液中三嗪除草剂的过程，如图 8-19 所示：①样品水溶液直接上样，样品中三嗪分子以反相机制保留（憎水作用）；②以二氯甲烷为溶剂冲洗 MISPE 柱，使 MIP 的亲水性及

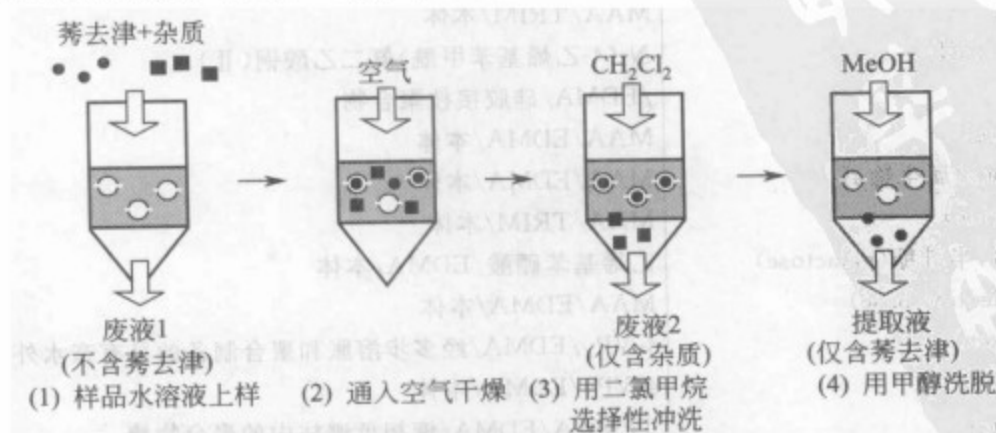


图 8-19 固相萃取 (SPE) 的过程

选择性恢复，结构不相关的杂质分子被洗掉；③以甲醇为溶剂洗脱，破坏氢键作用，回收目标分子三嗪化合物。经过 MISPE 柱后，西玛津（simazine，一种常用的除草剂）从混合物溶液中被选择性浓缩了 50 倍，回收率大于 90%。以上介绍的两种在水溶液体系中进行 MISPE 的方法被后来的许多研究者采用。

目前，固相萃取已被认为是分子烙印技术最有应用前景的领域，MISPE 已逐步作为样品前处理手段，与一些分析技术联用。

8.2.2 液相色谱

分子烙印聚合物作为液相色谱固定相的应用也是近年来研究非常多的热点之一。MIP 作为一种亲和色谱固定相，主要用来分离一种或一类目标产物，尤其在分离手性化合物中的应用具有一般材料所无法比拟的优势，有些甚至可以与商品化的手性柱相媲美。在天然或合成的药物中许多是手性化合物，其对映体的药效及毒性往往有很大差别，有时甚至相反，因此手性分离非常重要。分子烙印过程中，尽管所用功能单体一般没有手性，但用手性化合物作为模板，通过烙印过程即能得到可识别手性分子的聚合物材料。与传统的手性固定相不同的是，这种烙印聚合物中存在对模板分子“量身定做”的识别位点。例如，烙印时用 L 型异构体作为模板，则此 MIP 的色谱柱对 L 型异构体的保留显然大于对 D 型异构体的保留，相反亦然。但如果用消旋体做模板，得到的 MIP 将不具有手性识别能力。事实上，MIP 的立体选择性是由于手性分子的模板效应造成的。但 MIP 要求有一定量的单一化合物作为模板分子来合成，这对于有些手性化合物来说可能有一定困难（如价格较贵、难以得到等）。

表 8-2 列出了一些 MIP 用于液相色谱固定相分离一些天然产物、药物分子、环境检测样品及临床药物的例子。

表 8-2 MIP 用于色谱分离天然产物、药物分子、环境检测样品及临床药物实例

分 析 物	聚合物体系(单体/交联剂/聚合物形式)
	液 相 色 谱
氨基酸衍生物	MAA/TRIM/本体
氨基酸(amino acid)	N-(4-乙烯基苯甲酰)氮二乙酸铜(Ⅱ) /EDMA/硅胶接枝聚合物
莠去津(atrazine)	MAA/EDMA/本体
金鸡纳(cinchona)属生物碱	MAA/EDMA/本体
麻黄素(ephedrine)	MAA/TRIM/本体
果糖(fructose), 半乳糖(galactose)	乙烯基苯硼酸/EDMA/本体
核苷酸碱(nucleotide base)	MAA/EDMA/本体
萘普生(naproxen)	4-VPy/EDMA/经多步溶胀和聚合制备的具有亲水外层的微球; 4-VPy/EDMA/本体
尼古丁(nicotine)	TFMAA/EDMA/液相色谱柱中的聚合物棒
唾液酸(sialic acid)	乙烯基苯硼酸/EDMA/本体

分 析 物	聚合物体系(单体/交联剂/聚合物形式)
类固醇(steroid)	MAA/EDMA/本体
合成雌激素(synthetic estrogen)	甲基丙烯酸二乙氨基乙酯/TRIM/本体
睾丸酮(testosterone)	MAA/EDMA/本体
噻吗洛尔(timolo)	衣康酸/EDMA/本体 薄层色谱
麻黄素(ephedrine)及其类似物	MAA/EDMA/本体 毛细管电泳
普萘洛尔(propranolol)	N-丙烯酰苯丙氨酸/EDMA/本体 毛细管电色谱
氨基酸(amino acid)	MAA/EDMA/本体
丹酰苯丙氨酸(dansyl-phenylalanine)	MAA+2-VPy/EDMA 或 TRIM/毛细管中薄膜
普萘洛尔(propranolol)	MAA/TRIM/毛细管中超大孔连续材料
局部麻醉药	MAA/TRIM/毛细管中超大孔连续材料
喷他肽(pentamidine)	MAA/EDMA/原位制备分散型聚合物

注: MAA—甲基丙烯酸; TRIM—三羟甲基丙烷三甲基丙烯酸酯; EDMA—乙二醇二甲基丙烯酸酯; VPy—乙烯基吡啶; TFMAA—三氟甲基丙烯酸。

尽管 MIP 作为亲和色谱固定相迄今为止已经在多个领域得到了广泛的应用, 如氨基酸、多肽、糖类、药物分子、核苷酸碱基等作为模板分子均能得到具有良好分离效果的 MIP, 但是由于 MIP 的固有属性造成其色谱峰的宽化、拖尾一直是限制 MIP 进一步推广应用的一个主要因素。造成色谱峰拖尾的原因较多, 主要包括烙印过程中结合位点的亲和能力的不均一性以及结合与解析过程中的可变动力学问题等。亲和位点的不均一性往往导致样品超载, 从而引起 MIP 的柱容量降低。改善这种情况的一种办法是改善颗粒的形状和大小, 使之尽量均一化。悬浮聚合等方法可以改善颗粒形状和大小的不均一性, 仔细地筛分、缩小采用的颗粒粒径范围将有助于提高柱效。另外, 色谱条件的优化, 如选择合适的溶剂、柱温与洗脱条件等, 也可以改善峰形和分离效果。

8.2.3 传感器

分子识别在传感器技术中起着关键的作用。化学传感器一般包括三大部分: 分子识别部分、信号转换部分、信号放大和处理部分。其中分子识别部分决定传感器的选择性。许多用于环境检测、生化测试、食品分析的传感器是以生物大分子如抗体和酶为分子识别元件, 但这些生物大分子的化学和物理稳定性较差。MIP 相对来说物理性质和化学性质都比较稳定, 又具有良好的分子识别能力, 所以非常适合作为传感器的识别元件。分子烙印技术使得人们可以根据分析对象进行分子设计, 并制备出相应的分子识别元件, 提高传感器对分析物的选择性。近年来已经有许多关于 MIP 在传感器方面应用的报道。

例如, Mosbach 小组以 L-苯丙氨酰苯胺-MIP 膜为场效应器件, L-苯丙氨酰苯

胺和其结构类似物酪氨酸苯胺产生的电容变化相近,而结构差异比较大的苯丙胺醇只产生较小的电容变化。另外,他们还以吗啡的 MIP 制得对吗啡有选择性的电流传感器。该方法分为两个步骤:首先将吗啡的 MIP 固定在铂电极上,测定时,吗啡选择性地结合到传感器的膜上,然后加入电活性竞争物可待因,并用电流法测定释放出来的吗啡。检测的浓度范围为 $0.1 \sim 10 \mu\text{g/L}$ 。

有些模板分子有荧光,当它用于合成相应的 MIP 时能使聚合物带荧光,荧光的强度与分析物的浓度有定量的关系。例如,丹酰-L-苯丙氨酸苯胺能产生荧光,将它的 MIP 置于光纤的顶端,就可制成对丹酰-L-苯丙氨酸苯胺有选择性的光纤传感器,但此传感器的响应时间较长,需要约 4h。有些模板分子自身虽然不能发荧光,但如果在制备 MIP 的过程中加入一种带荧光的单体,利用模板分子对这种单体的荧光增强或淬灭效应也可制成荧光传感器。如 Powell 等以反式-4-对二甲氨基苯乙烯基-N-乙烯基-N-苄基吡啶鎓(DMASVBP)和甲基丙烯酸 2-羟乙酯(HEMA)为功能单体,以腺苷-3',5'-环单磷酸(cAMP)为模板分子合成 MIP,DMASVBP 在此同时起识别和产生信号的作用。在含水的甲醇溶液中(3:7,体积比),利用 cAMP 对 DMASVBP 的荧光淬灭效应,可以测定 cAMP 的含量。Piletsky 等还报道了一种基于唾液酸的 MIP 的荧光传感器,他们将荧光试剂邻苯二醛加入唾液酸溶液,聚合物吸附唾液酸时也吸附部分邻苯二醛,使 MIP 的荧光增强。荧光增强的多少与唾液酸的吸附量成正比。

MIP 在石英微天平和表面声波传感器上的应用也有报道。最早,石英微天平主要用于气体的分析。例如 Karube 小组以 2-甲基异冰片-MIP 修饰石英晶体,其对模板分子的响应比空白聚合物修饰的石英晶体产生的响应高 1.1~1.3 倍,而其他同类添加剂在 MIP 和 NIP(非分子烙印聚合物,即空白)上的响应几乎相同。

总体说来,分子烙印在传感器中的应用可以分为 3 类:①基于模板分子固有性质产生信号的传感器;②基于标记化合物产生信号的传感器;③基于 MIP 与模板分子结合产生性质变化的传感器。虽然 MIP 在传感器上的应用有巨大的潜力,但目前大部分已报道的传感器的响应时间较长,这对于 MIP 的实际应用是一个亟待解决的问题。预计超薄膜、纳米 MIP 以及其他技术的发展将会有利地推动 MIP 在传感器中的实际应用。

8.2.4 模拟抗体

MIP 根据分子大小、形状和官能团进行分子识别的过程与抗体类似,同时 MIP 比较稳定,耐有机溶剂、酸、碱、高温等,因此 MIP 可作为抗体模拟物用于免疫分析。例如茶碱和地西泮(diazepam)烙印的聚合物表现出非常好的分子识别能力,以这种聚合物作为人工抗体用于茶碱和地西泮的放免分析,结果表明,它们对相关结构类似物没有活性或活性大大低于模板化合物,对这些药物

的交叉活性与单克隆抗体一致。用茶碱的 MIP 测定病人血清中茶碱的浓度, 结果与酶联免疫分析一致。

总之, MIP 作为模拟抗体在免疫分析中的应用是基于被分析物与标记物的竞争吸附。被分析物与一定量的标记物竞争吸附于 MIP 上, 未被吸附的标记物与加入的被分析物呈现一定的比例关系。

8.2.5 模拟酶催化

MIP 作为酶的模拟物不仅具有分子识别功能, 而且还可以模拟酶的催化作用。通过选用适当的模板分子和功能单体, 可以在 MIP 的孔穴内形成类似酶活性中心的结构。MIP 作催化剂时, 许多情况与酶作催化剂相似。

Efendiev 等首先把分子烙印技术应用于催化剂的研究。Sarhan 等研究了含 4-乙烯基吡啶单元的 MIP 对扁桃酸的光学异构体的催化再消旋化反应; Andersson 等证明了含吡哆醛基的 MIP 可以加速氨基酸衍生物的质子交换作用; Leonhard 等以乙烯基咪唑为功能单体制得的 MIP 可以增加模板分子的酯解活性。但是, 大部分 MIP 的催化活性不是很高, 因此, 许多研究者试图用反应过渡态的类似物作为模板分子合成 MIP, 以提高 MIP 的催化效率。例如, Mosbach 等以甲基膦酸对硝基苯酯为模板制备出能催化乙酸对硝基苯酯水解的 MIP, 这是由于甲基膦酸对硝基苯酯与乙酸对硝基苯酯水解过渡态结构类似, MIP 的烙印孔穴优先结合水解反应的过渡态, 从而诱导乙酸对硝基苯酯水解速度增加。另外, 活性基团的引入也能显著地提高 MIP 的催化活性。

分子烙印聚合物具有构效预定性 (predetermination)、特异识别性 (specific recognition), 而且制备简单, 物理化学性能稳定 (抗机械作用、高温、高压, 耐酸碱、有机溶剂、金属离子), 使用寿命长, 因此具有广泛的实用性 (practicability)。

8.3 分子烙印固相萃取中药中特定药效组分——黄酮类

黄酮类化合物是广泛存在于中药中的一大类化合物。黄酮类化合物的药物活性是多种多样的。例如槲皮素、槲皮苷等能增强心脏收缩, 减少心脏搏动次数; 杜鹃素、紫花杜鹃素等有止咳祛痰的作用; 黄芩苷、木犀草素等有抗菌消炎的作用; 牡荆素、汉黄芩素等具有抑制肿瘤细胞的作用; 水飞蓟素具有保肝作用。为了探索 MIP 直接选择性萃取分离中药中具有特定药效结构组分的可行性, 以较易得到的黄酮化合物——槲皮素作模板制备了分子烙印聚合物, 研究了槲皮素聚合物的分子识别特点, 在此基础上探讨了分子烙印直接固相萃取分离中药中特定药效结构黄酮化合物的可行性。

8.3.1 聚合物的制备

采用封管聚合方法。将模板分子槲皮素 (quercetin)、功能单体丙烯酰胺

(AA) 或甲基丙烯酸 (MAA) 或 4-乙烯基吡啶 (4-VPy)、引发剂偶氮二异丁腈 (AIBN)、致孔剂四氢呋喃 (THF) 以及交联剂乙二醇二甲基丙烯酸酯 (EDMA) 超声波混合后, 液氮冷却封管, 60℃ 水浴聚合 24h。得到的棒状聚合物磨碎过筛 (30 μ m), 颗粒小于 30 μ m 部分用丙酮反复沉降, 以除去细末, 将得到的聚合物真空干燥。空白聚合物 (BP) 的制备方法同上, 只是未加模板分子。聚合物用甲醇-乙酸(9:1, 体积比) 索氏提取洗 48h 后, 干法装柱, 填充柱为 ϕ 4.6mm $\times$ 150mm 不锈钢柱。

由于模板分子与功能单体形成稳定的复合物是分子烙印的重要条件, 在其他条件相同的情况下, 首先分别用 MAA、AA 或 4-VIP 作功能单体制备了槲皮素模板分子的聚合物 (表 8-3 和表 8-4), 并用液相色谱评价了不同单体的分子烙印效率 (imprinting efficiency, IE)。另外, 由于预聚合混合物经过了较长时间的通氮气除氧处理, 在聚合反应中只发现缓聚, 而没有观察到酚类化合物对自由基反应的阻聚现象。这也可能是由于预聚合混合物中加入的功能单体是大量的, 在槲皮素分子与功能单体作用后, 槲皮素分子所含的酚羟基被功能单体 AA “包围” 所致。

表 8-3 用不同单体合成的聚合物

聚合物	单体	交联剂	引发剂	溶剂	性状
MIP1 (BP1)	MAA	EDMA 40mmol	AIBN 30mg	THF 9mL	坚硬, 不透明
MIP2 (BP2)	AA	EDMA 40mmol	AIBN 30mg	THF 9mL	坚硬, 不透明
MIP3 (BP3)	4-VPy	EDMA 40mmol	AIBN 30mg	THF 9mL	坚硬, 不透明

表 8-4 3 种聚合物的色谱参数

参 数	MIP1	MIP2	MIP3
容量因子 k' (MIP)	3.02	7.14	5.22
容量因子 k' (BP)	1.50	2.0	2.3
分子烙印效率 (IE)	2.0	3.57	2.27

由于研究目的主要是探索分子烙印技术直接选择性萃取分离中草药中黄酮化合物的可行性, 为此, 对分子烙印效果较好的 MIP2 的选择性及亲和性又做了进一步研究。在后面的讨论中所指的 MIP 即为 MIP2。

8.3.2 聚合物的选择性评价

在选择性实验中, 使用了结构与槲皮素类似的黄酮组分及一些非黄酮组分作为参考对象 (图 8-20)。尽管所制备的 MIP 在乙腈中保留很强, 但由于 MIP 将用于中草药体系中直接提取黄酮组分, 中药中的各种组分一般能够很好地溶解在甲醇中, 因此在 MIP 的选择性实验中甲醇作为流动相。

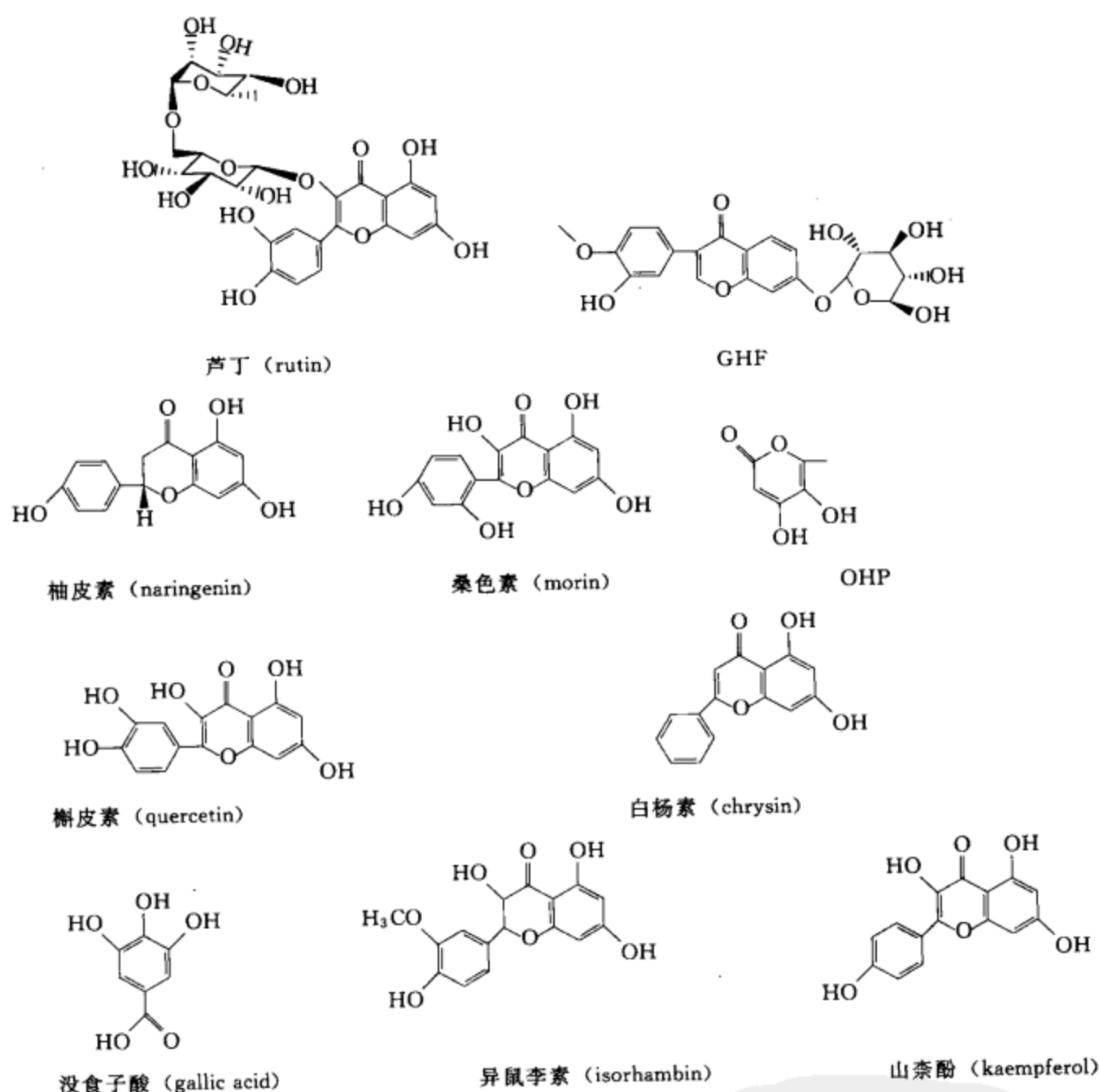


图 8-20 槲皮素及其类似物的分子结构

从分子烙印效率 IE 值也可以看出 (表 8-5), 模板分子槲皮素分子的烙印效应最强。从分离因子值可以看到, MIP 柱对模板分子槲皮素的选择性也是最强。另外所考察的化合物在 MIP 柱上的保留远大于它们在空白聚合物 (BP) 上的保留, 这表明分子烙印聚合物柱容量大于空白聚合物柱。更重要的是 MIP 对黄酮类化合物的保留远大于对非黄酮类化合物的保留, 即 MIP 对黄酮类化合物存在“交叉反应性”。

事实上, 由于 MIP 对底物的识别靠的是分子的大小和形状的匹配及官能团的相互作用, MIP 除了能够对于模板分子表现出很高的亲和性以外, 还能够对结构与模板相类似的化合物具有较强的结合, 底物分子的结构越是与模板分子结构类似, 则 MIP 对于这个底物的保留越强。MIP 存在的对黄酮化合物的“交叉反应性”是由于 MIP 是用典型的黄酮化合物槲皮素作为模板分子制备的。异鼠李素与槲皮

素的结构非常相似,因此在 MIP 上也表现出很强的保留。表 8-5 中未列出山柰酚的数据,但可以推测 MIP 应对山柰酚表现较强的“交叉反应”。

表 8-5 槲皮素和其他样品的色谱参数

参 数	槲皮素	白杨素	柚皮素	异鼠李素	桑色素	芦丁	OHP	GHF	没食子酸
容量因子 k' (MIP)	7.14	2.24	2.02	5.20	2.10	0.49	0.10	0.32	0.29
容量因子 k' (BP)	2.00	2.77	1.18	1.85	1.05	0.63	0.10	0.40	0.63
分离因子 α (MIP)	1	3.19	3.53	1.37	3.40	14.6	71.4	22.0	24.6
分离因子 α (BP)	1	0.72	1.69	1.08	1.90	3.17	20.0	5.0	3.17
分子烙印效率(IE)	3.57	0.81	1.71	2.81	2.0	0.78	1	0.8	0.46

值得注意的是,尽管桑色素分子结构与槲皮素分子结构也非常相似,但在表 8-5 的实验条件下并没有发现桑色素在 MIP 上表现相对更高的“交叉反应性”。但当进样量进一步减少时,发现桑色素在 MIP 上的保留增强。这很可能是在表 8-5 的实验条件下桑色素存在分子内氢键,这时桑色素的化学结构与槲皮素模板分子结构差异较大;而当进样量减少后,这种分子内氢键减弱或消失,从而使 MIP 对桑色素的“交叉反应性”增大。

MIP 的分子识别能力还可以从表 8-5 的结果进一步证实。像没食子酸、芦丁及 OHP、GHF 这样的强极性多酚羟基化合物在 MIP 柱上的分离因子(α)值很大,表明这 4 个化合物在 MIP 的保留主要靠非特性吸附作用。这可以从 MIP 对分子的大小和形状识别角度很好地解释。由于没食子酸和 OHP 比槲皮素分子小很多,不能很好地与用槲皮素作模板形成的烙印空穴相匹配,从而 MIP 对这两个分子的保留小。同样,由于芦丁和 GHF 分子空间结构很臃肿,难于进入分子空间结构比其小的槲皮素模板分子孔穴中,使得这两个分子在 MIP 上的保留也很小。

由上可知,MIP 除了对槲皮素分子表现出很高的选择性外,还对于结构与槲皮素类似的其他黄酮化合物表现出较高的“交叉反应性”。这一特点使得它有可能作为分离材料,在某种条件下选择性提取中草药中的具有特定药效结构的黄酮化合物。

8.3.3 聚合物固相萃取柱柱载量测定

在 MIP 分离中复杂的中草药样品之前,首先用纯槲皮素甲醇溶液按照 SPE 过程考察了柱子的容量及操作条件。在分子烙印固相萃取研究中,对于分子烙印固相萃取柱的容量还没有一个公认的测定方法。将 0.3517g 空白聚合物(BP)或分子烙印聚合物(MIP)装入 6mL 空聚丙烯柱中,再将配好的 0.8 μ g/mL 槲皮素甲醇溶液按每次 1mL 加入到聚丙烯 MIP 柱上。收集上样时的流出液,并用液相色谱检测槲皮素分子漏出情况,当检测到漏液时停止加样品溶液,并计算出已经加入柱子

的样品量。然后再用甲醇洗柱（每次用 1mL 甲醇洗柱），以洗去非特异吸附部分，洗柱溶液收集后也用液相色谱检测，检测到样品露出时停止洗柱。最后用甲醇-乙酸（9：1，体积比）洗脱，计算出实际吸附量。同时做空白实验。

液相色谱分析使用一个 Zobax Eclipse XDB-C8 ($\phi 4.6\text{mm} \times 150\text{mm}$) 柱，并配上预保护柱（DIKMA）。38% 甲醇/62% 磷酸液→55% 甲醇/45% 磷酸液，梯度洗脱 30min。检测波长 365nm。进样量 20 μL ，流速 1mL/min。柱温室温。

表 8-6 柱容量测定中槲皮素吸附和洗脱的量

阶段	槲皮素的量/ μg	
	BP	MIP
上样	2.4	6.0
吸附	2.0	5.8
洗脱	0.6	0.4
萃取	1.2	5.4

由表 8-6 可以看出，虽然空白聚合物的上样量少于 MIP 的上样量，但空白聚合物上样时仍有槲皮素没有被保留。可见，由于底物在空白聚合物上的保留是靠非特异性吸附作用，结合作用较弱，空白聚合物容量较小。这还可以从对洗柱液的检查中进一步证明，在洗柱时空白聚合物上有更多的槲皮素分子被洗出来。

经过上面的洗柱过程，在 MIP 上最后的回收量为 5.4 μg ，回收率可达到 90%。而在空白聚合物上的回收量才为 1.2 μg ，回收率是 50%。由于槲皮素在 MIP 上的吸附包含了特异性结合与非特异性结合两部分，如果认为在空白聚合物上的吸附量 1.2 μg 相当于 MIP 上的非特异性吸附部分，则在 MIP 上的特异性吸附部分为 4.2 μg 。这还可以折合成 MIP 的特异性吸附量为 12 $\mu\text{g/g}$ 聚合物。这种方法得到的柱子容量数值与采用的固相萃取过程及萃取条件很有关系。这个容量实验得出的结论只供后来的进一步固相萃取中草药复杂样品参考。

8.3.4 固相萃取分离银杏叶中黄酮类药效成分

(1) 银杏叶粗提取物样品的制备 将 1g 干燥的银杏叶细粉加入 180mL 甲醇，索氏提取约 2.5h，旋干。加入 30mL 甲醇、5mL 25% 盐酸，80℃ 水浴中回流 1h。加入无水甲醇，反复旋干，再定容至 150mL。使用前，提取液用 Bond-Elut-C18 (Varian) 处理，除去脂溶性物质。样品中槲皮素浓度及山柰酚的浓度用槲皮素作为外标测定。液相色谱分析条件同以上所述。

(2) MIP 直接固相萃取分离银杏叶的甲醇粗提物 所用的分离柱子与以上所述相同，使用前用 10mL 甲醇平衡柱子。将所制备的银杏叶提取液上样，每次 0.2mL，共上样 1mL。然后再用甲醇洗柱，按每次 0.2mL，共用 3mL 甲醇洗柱。最后用 3mL 甲醇-乙酸（9：1，体积比）洗脱柱子。收集每次上样时的流出液、每次

洗柱时的洗柱液及最后的洗脱液，进行色谱分析。

在分离实验进行前，首先采用液相色谱方法分析了所制备样品中的槲皮素和山柰酚含量。通过用标准品对照确认色谱图（图 8-21）中 a 为槲皮素，通过与文献色谱图对照确认 b 为山柰酚。槲皮素和山柰酚的含量是用槲皮素作为标准采用外标方法计算得到的。对于色谱图中其他小峰对应的组分，本研究没有讨论。

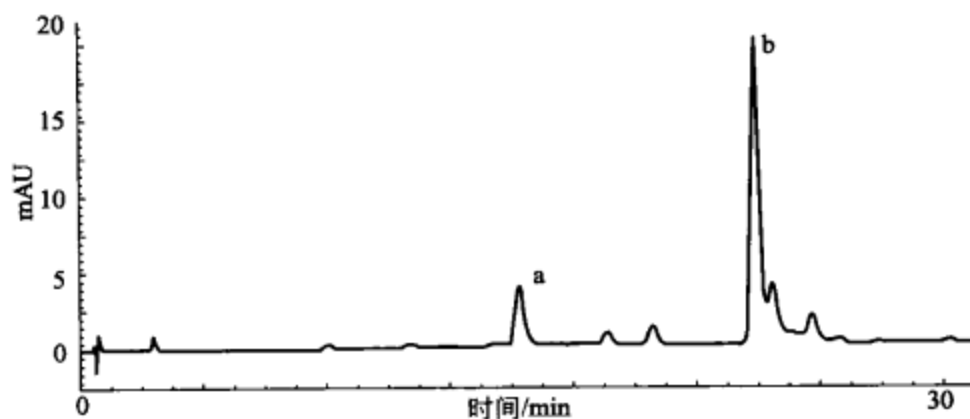


图 8-21 银杏叶粗提物的液相色谱图
mAU—毫吸光度单位

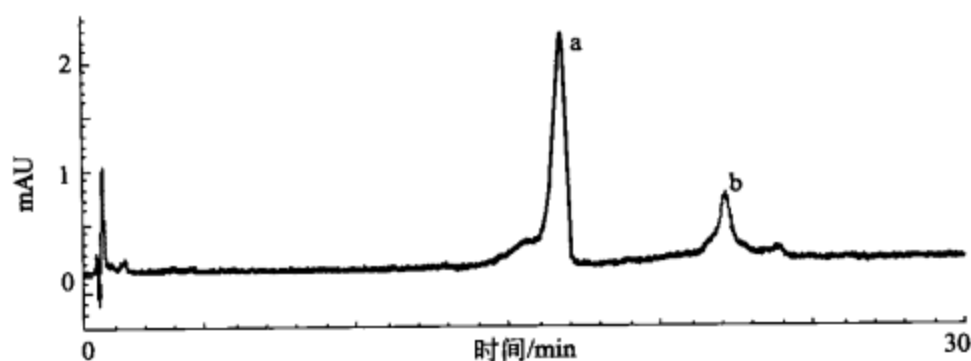


图 8-22 银杏叶粗提物经过 MIP 柱后的液相色谱图
mAU—毫吸光度单位

由图 8-21 及图 8-22 可以看出，对于 MIP 柱子，最后的洗脱液中许多杂质组分已在洗柱过程中被洗掉，谱图简化，洗脱液中主要是槲皮素及山柰酚组分。

8.3.5 萃取分离沙棘草药粗提物中特定药效黄酮组分

在上面萃取银杏叶中黄酮组分的基础上，又尝试用 MIP-SPE 萃取沙棘草药中的黄酮组分（图 8-23、图 8-24）。萃取方法及思路与上面描述的相同。由于异鼠李素是槲皮素的结构类似物，MIP 对于异鼠李素表现“交叉反应性”，因此 MIP 能够直接固相萃取出沙棘草药中的槲皮素及异鼠李素组分。图 8-23 和图 8-24 中，峰 a 为槲皮素，峰 b 为异鼠李素（通过与标准品对照辨认）。峰 c 无法确认，但推测其很可能是另一个黄酮化合物。

（1）粗提物样品的制备 称取 0.45g 沙棘，溶于 50mL 甲醇，加 10mL 25%盐

酸，加热回流 1h，旋转蒸发至干，甲醇定容至 250mL。待用。

(2) SPE 分离沙棘粗提物 称取 MIP2 0.45g，装入容量为 6mL 的空聚丙烯柱中，装填均匀后用 10mL 甲醇湿润该提取柱。将所制备的沙棘水解物溶液 1mL（每次 100 μ L，分 10 次加样）缓慢地加到提取柱上。再用甲醇 1.6 mL（每次 200 μ L，分 8 次冲洗）缓慢冲洗提取柱。用 6mL 甲醇-乙酸（9：1，体积比）洗脱提取柱 6 次，收集每次洗柱液和洗脱液，并分别合并，旋转蒸发至干，再溶于流动相 1mL 中进行 HPLC 分析。

色谱分析条件为 Zobax Eclipse XDB-C8(ϕ 4.6mm $\times$ 150mm) 柱，并配上预保护柱 (DIKMA)。甲醇 45%/(0.3%) H_3PO_4 55% 等强度洗脱，流速 1mL/min；柱温为室温；进样量 20 μ L；检测波长 365nm。粗提物经过 MIP 柱前后的色谱图如图 8-23、图 8-24 所示。

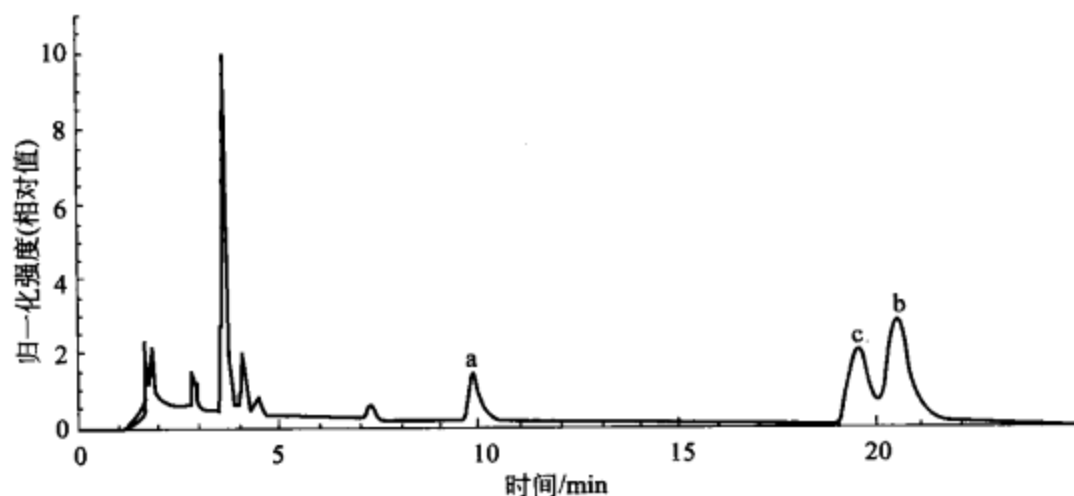


图 8-23 沙棘粗提物的液相色谱图

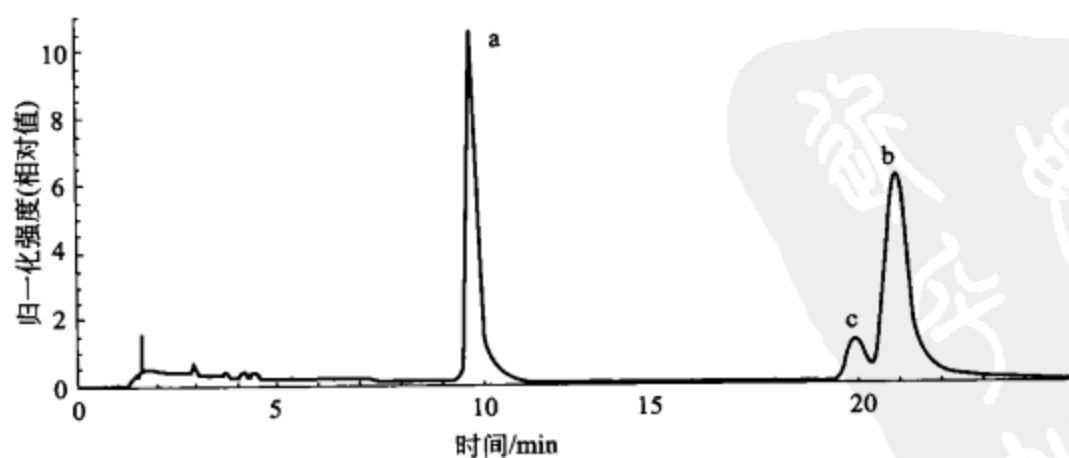


图 8-24 沙棘粗提物经过 MIP 柱后的液相色谱图

萃取分离沙棘草药粗提物的实验结果从方法上进一步证明了用槲皮素模板分子烙印聚合物直接固相萃取中药中具有特定药效结构的黄酮成分是可行的。

8.4 分子烙印亲和分离在线鉴定中草药中有效成分

到目前为止, MIP 的几个应用领域已经得到了深入的研究。MIP 用于 HPLC 亲和色谱固定相手性分离药物分子, 一直是个众人关注的研究课题。而且曾有研究者用分子烙印方法获得了非常吸引人的分离及手性拆分结果。另外有关将分子烙印与薄层色谱 (TLC) 相结合的分离技术研究也有报道。目前, 国内外很多工作研究 MIP 在固相萃取 (SPE) 方面的应用。在这些例子中, MIP 固相萃取 (MIP-SPE) 多数是分析前用于从生物体液或环境样品中预浓缩或预纯化模板分子或结构上与模板分子相似的一类化合物。但 MIP 作为 HPLC 固定相直接亲和分离多组分样品中的目标组分 (MIP-LC) 或将 MIP 直接作为 HPLC 固定相与质谱联用 (MIP-LC-MS) 实现亲和分离及在线鉴定多组分样品中的目标组分方面的研究目前还未见报道。

LC-MS(液相色谱-质谱联用) 已逐渐被看作一种准确、敏感、高效的手段; 用在分离鉴定中草药有效成分上。由于质谱本身特有的专一性和高分辨能力, 一方面不需要标准品即可对中草药中的目标组分进行确认, 另一方面还可从质谱图上出现的分子离子峰对中草药中含有的其他未知组分进一步分析, 这对于从中草药中发现先导化合物是很有利的。

在 LC-MS 中, 常用的 LC 固定相一般是反相填料或正相填料, 由于分子在正相或反相填料上的保留属于非特异性表面吸附, LC 分离条件往往需要根据目标分子或样品中其他组分的性质仔细确定。另外 LC-MS 中选用的 LC 流动相一方面要有利于 LC 固定相的稳定存在, 另一方面还要适于质谱检测。分子烙印材料是一种对目标分子具有特定亲和性的聚合物材料, 它稳定耐用, 而且由于其对化合物的保留是亲和机制, LC-MS 分离过程中 LC 流动相容易确定。另外, MIP 色谱柱使用的流动相一般是有机溶剂体系, 有利于质谱检测。因此, 将分子烙印聚合物液相色谱柱与质谱联用应是一种有效的能够实现的分离技术。

这里采用抗肿瘤成分哈尔明及哈马灵的结构类似物哈尔满作为模板分子合成了对哈尔明及哈马灵具有亲和性的 MIP, 并将此 MIP 作为液相色谱固定相用于 LC-MS 系统, 在线分离鉴定骆驼蓬草药中的抗肿瘤组分。证明了用 MIP 作为 HPLC 固定相与 MS 联用, 实现中草药有效成分的亲和分离及在线鉴定一体化, 加快中草药有效成分的分离鉴定过程是可行的。

8.4.1 仪器和材料

液相色谱-质谱联用系统 (LC-MS) 由一个 Mariner Time-of-Flight (TOF) 质谱 (PE PerSeptive Biosystems, Framingham, MA, USA) 和一个 HP1100 高效

液相系统 (Hewlett Packard, Palo Alto, CA, USA) 组成。液相色谱含一个四元泵、一个波长可调紫外检测器、一个真空在线脱气机, 配置一个 20 μ L 与一个 2mL 进样环的手动六通阀。质谱部分配置一个 ESI 源和一个 APCI(大气压电离离子化) 源。当使用 ESI(电喷雾离子化) 源时, 采用正离子检测, 条件为: 载气和帘气为 N₂, 流速分别为 0.5mL/min、1.5mL/min; 喷嘴和四极温度均是 140℃; 喷嘴梢和喷嘴电压分别为 5000V 和 100V; 收谱速度为每秒 3 张谱。当使用 APCI 源时, 载气、帘气和辅助气为 N₂, 流速分别为 0.70 mL/min、0.90 mL/min 和 2.5mL/min; 喷嘴、四极和喷室温度分别是 170℃、140℃和 475℃; 其他条件与用 ESI 源时相同。

哈尔满 (harmane, HMA)、哈尔明 (harmine, HMI)、乙二醇二甲基丙烯酸酯 (EDMA)、甲基丙烯酸 (MAA)、偶氮二异丁腈 (AIBN) 从百灵威购买; 哈马灵 (harmaline, HML) 从 Sigma 购买。在使用前 EDMA 用下面的方法处理: 10%氢氧化钠水溶液洗后, 用无水硫酸镁干燥, 然后再减压重新蒸馏。甲苯溶剂在使用前, 用钠干燥处理后重新蒸馏。甲醇、乙腈和四氢呋喃为 HPLC 级。乙酸为分析级。超纯水由北京大学中级仪器实验室提供。

8.4.2 聚合物制备

将哈尔满 1mmol 溶于不同的致孔剂中 (见表 8-7), 加入甲基丙烯酸 4mmol, 超声混合。然后再加入乙二醇二甲基丙烯酸酯 30mmol、偶氮二异丁腈 41.7mg, 通氮气除氧 10min, 抽真空, 液氮冷却封管。60℃水浴聚合 24 h 后, 磨碎过筛 (35 μ m), 颗粒小于 35 μ m 部分用丙酮反复沉降除去细末, 最后再将聚合物真空干燥。聚合物粒子干法装入不锈钢柱 (ϕ 4.6mm $\times$ 150mm) 中, 然后用甲醇-乙酸 (8:2, 体积比)、甲醇在线洗脱, 直到得到稳定的基线, 以保证除去模板分子。作为对照, 空白聚合物 (BP) 制备方法同上, 只是不加模板分子。

研究的目的是尝试用 MIP 作亲和色谱固定相与质谱联用, 达到亲和分离与在线鉴定中草药中活性组分的目的。在骆驼蓬种子中有两种抗肿瘤组分哈马灵及哈尔明, 哈马灵及哈尔明均为含氮碱性化合物, 且都不溶于乙腈或氯仿等极性较弱的有机溶剂中。因此, 以甲基丙烯酸作为功能单体, 采用现有的分子烙印方法很难直接得到哈马灵及哈尔明的分子烙印聚合物。为了得到对于目标组分哈马灵及哈尔明具有亲和性与选择性的分子烙印聚合物, 采用哈马灵及哈尔明分子的结构类似物哈尔满分子为模板、甲基丙烯酸为功能单体制备了哈尔满的分子烙印聚合物。

由于在不同的致孔剂中得到的 MIP 具有不同的保留特性, 而且致孔剂还对于 MIP 的聚合物网络有影响, 因此, 分别采用 3 种致孔剂进行了 MIP 的制备, 以优化实验结果。表 8-7 列出了所制备的聚合物 (P 表示所制备的聚合物, BP 表示空白聚合物)。

表 8-7 使用不同致孔剂合成的聚合物

聚合物	P1,BP1	P2,BP2	P3,BP3
致孔剂(10mL)	甲苯-乙腈(体积比 1:1)	乙腈	四氢呋喃
性状	松脆,透明	坚硬,不透明	坚硬,不透明

从表 8-7 可以看出,用不同的致孔剂聚合反应后,得到的棒状聚合物具有不同的特点。在混合溶剂中得到的聚合物 P1 较脆且透明,在乙腈中和四氢呋喃中得到的聚合物 P2 和 P3 较硬而且不透明。在用丙酮浮选聚合物粒子时,可能由于 P1 聚合物在丙酮溶剂中有些溶胀, P1 丢失得较多。为了有效地除去模板分子,将制备好的聚合物颗粒装入液相色谱空柱,使用了甲醇-乙酸(8:2, 体积比)强极性溶剂在线洗脱柱子大约 48h。

8.4.3 液相色谱评价聚合物对目标分子的选择性

用 HPLC 研究所制备的 MIP 的亲和性及选择性。在室温下进行。选用的流动相为甲醇,流速为 0.5mL/min,进样体积为 20 μ L,柱温为室温。不同的进样量是通过固定进样体积,改变进样浓度实现的。紫外检测波长为 254nm。容量因子 k' 是通过方程 $k' = (t_R - t_0)/t_0$ 计算的,其中 t_R 是样品的保留时间, t_0 是用丙酮记录的死时间。分子烙印效率 $IE = k'_{MIP} / k'_{BP}$, 其中 k'_{MIP} 是样品在分子烙印柱上的容量因子, k'_{BP} 是样品在空白柱上的容量因子。

目标化合物哈马灵及哈尔明对哈尔满模板分子烙印聚合物 MIP1、MIP2 和 MIP3 的亲和性与选择性通过液相色谱方法进行了检验。由于 P1、P2 和 P3 在乙腈中表现出很强的保留(进样量为 0.1mmol/L 时, 3h 没能将哈马灵及哈尔明洗脱出来),液相色谱评价是在甲醇流动相中进行的。采用甲醇流动相是出于下面的考虑:所选用的流动相将用于分离骆驼蓬提取物,甲醇能充分溶解骆驼蓬提取物中各组分。另外,在甲醇中抗肿瘤组分哈马灵及哈尔明仍以碱性形式存在,这将有利于最后拿到所要的组分。图 8-25 所示为所研究分子的结构式。

图 8-26 表示出了液相色谱评价结果。P1、P2 和 P3 除了对模板分子哈尔满具有较强的亲和性外,对于哈尔满的结构类似物哈马灵及哈尔明也表现出“交叉反应性”。但这些分子烙印聚合物对结构上与哈尔满相差较大的其他含氮碱性化合物保留较小,这一特点使得所制备的分子烙印聚合物选择性分离骆驼蓬提取物中抗肿瘤组分哈马灵及哈尔明成为可能。另外,所有的样品分子在分子烙印聚合物上的容量因子均大于它们在空白聚合物上的容量因子,说明用分子烙印方法得到的聚合物容量更大一些。

在同一分子烙印聚合物上,容量因子按照哈尔满、哈尔明和哈马灵的顺序增大,这是由于对于用 MAA 作功能单体的 MIP,底物的碱性在分子识别中也起到很大的作用。但是,由于 P1、P2 和 P3 是用哈尔满分子为模板制备的,在 P1、P2 或

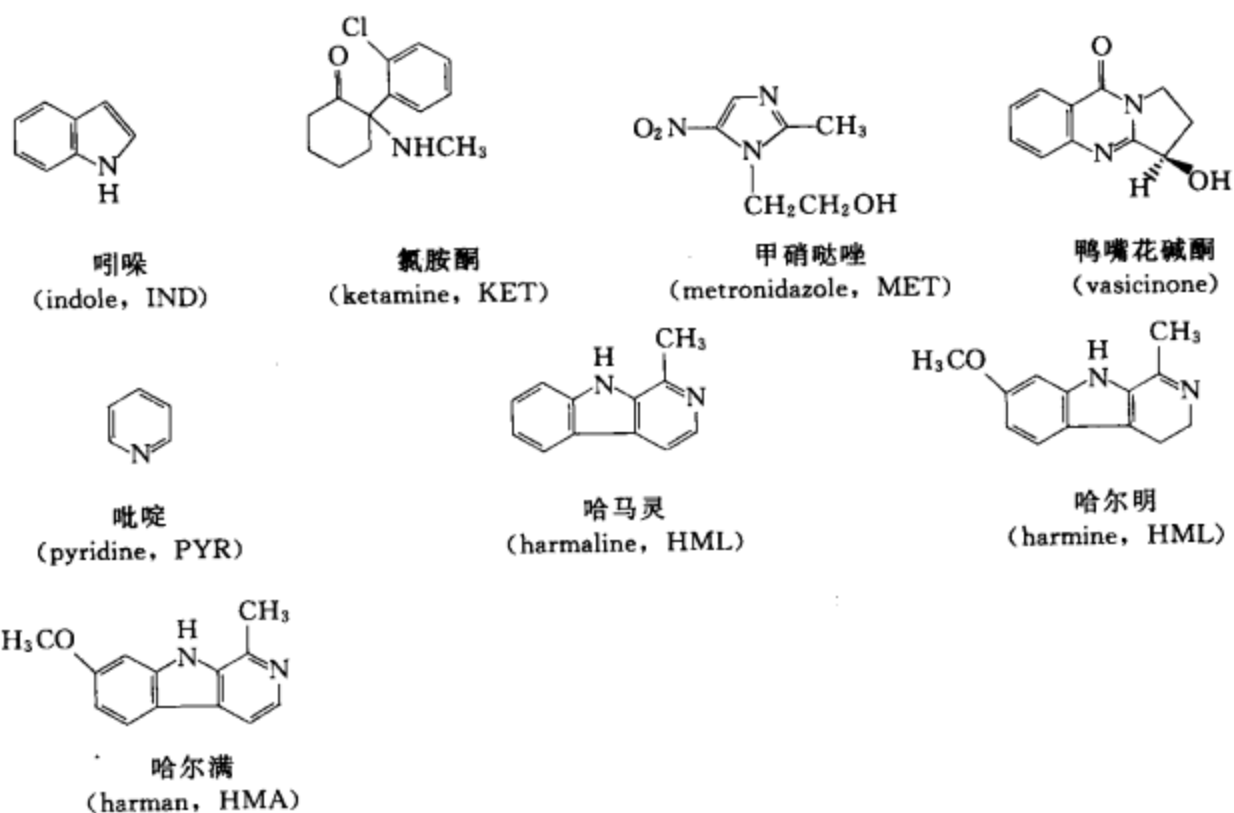


图 8-25 所研究分子的结构式

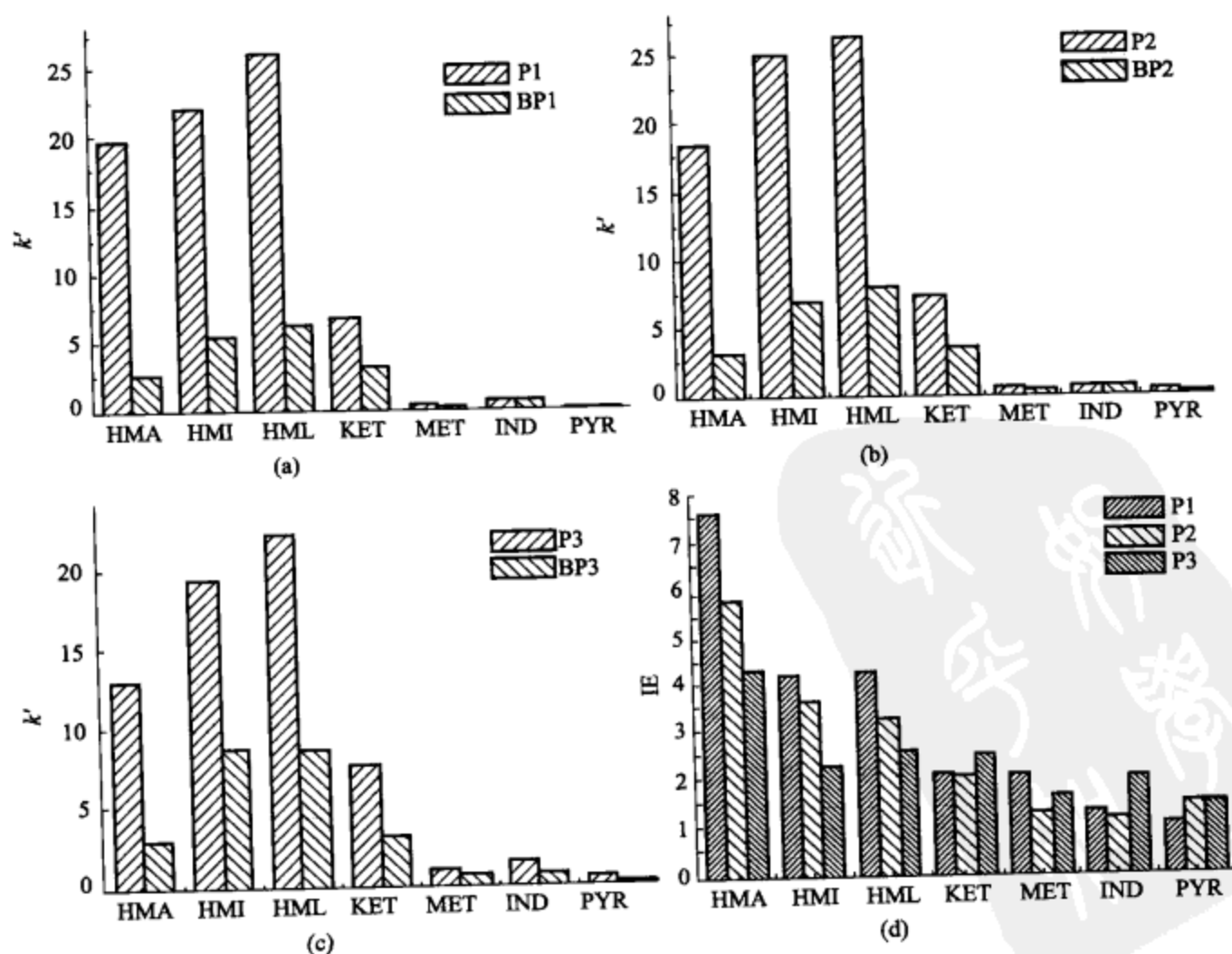


图 8-26 液相色谱评价聚合物对不同分子的选择性与亲和能力

P3 上, 哈尔满的分子烙印效率比哈尔明或哈马灵都大。

由图 8-26 还可以看出所用的致孔剂对 MIP 的亲和性及选择性的影响。根据分子烙印理论, 模板分子与功能单体之间复合物的形成是影响分子烙印效率的主要因素, 所用的致孔溶剂形成氢键能力越弱, 越有利于复合物的形成, 将来制备出的分子烙印聚合物对目标分子的亲和性及选择性越强。由于烙印过程所用的致孔溶剂形成氢键能力顺序是甲苯+乙腈<乙腈<四氢呋喃, 因此同一分子(模板分子哈尔满或目标化合物哈马灵、哈尔明)在所制备的分子烙印聚合物上的容量因子和分离因子按照 P1(甲苯+乙腈)、P2(乙腈)、P3(四氢呋喃)的顺序减小。由图 8-26 可进一步看出, 由于 P3 用形成氢键能力较强的溶剂四氢呋喃作致孔剂, 分子烙印效应值在 P3 柱上很小。这也表明在 P3 柱上分子烙印基本没有形成, 样品分子在 P3 柱上的保留主要是靠非特异性的表面吸附作用。

另外, 由于所制备的 MIP 是为了用于 HPLC 固定相分离中草药中目标有效成分, 因此有必要进一步探索在不同的进样量时所制备的 MIP 的保留特点。由图 8-27 可以看出, 随着进样量的逐渐增大, 模板分子哈尔满、目标化合物哈马灵和哈尔明在所制备的 3 个分子烙印聚合物 P1、P2 或 P3 上的容量因子及分子烙印效应值减小。另外, 在进样量较小(2nmol)时看到的是一个很宽的色谱峰, 当进样量增大后发现色谱峰变窄。这种色谱保留对于上样量的依赖关系很可能是由于分子烙印聚合物中结合位点的“不均一性”造成的。由于靠非共价作用形成的模板分子-功能单体复合物处于一种动态平衡中, 最后得到的分子烙印聚合物中除了含有少量的高选择性结合位点外, 还含有大量的选择性较弱的结合位点。当进样量增多时, 高亲和力位点就会饱和, 从而有更多的分子与分子烙印聚合物内的弱结合位点作用。因此, 当进样量增多时, 容量因子减小, 同时也造成了其分子烙印效应值降低。

8.4.4 前沿色谱方法测量结合位点数及离解常数

在上面的液相色谱研究基础上, 进一步用前沿色谱方法测定了 P1 和 P2 的结合容量, 结果见表 8-8。所制备的分子烙印聚合物的离解常数和结合位点数是通过前沿色谱方法得到的。将聚合物粒子湿法装入 $\phi 10\text{mm} \times 2.1\text{mm}$ 的 HPLC 柱中, 流动相使用甲醇, 流速为 $0.2\text{mL}/\text{min}$, 柱温为室温。不同浓度($1\text{mmol}/\text{L}$, $2\text{mmol}/\text{L}$, $3\text{mmol}/\text{L}$, $4\text{mmol}/\text{L}$, $5\text{mmol}/\text{L}$)的哈尔满、哈马灵、

表 8-8 用前沿色谱测定 P₁ 和 P₂ 的结合容量

化合物	结合位点数/($\mu\text{mol}/\text{g}$ 聚合物)		离解常数/(mmol/L)	
	P1	P2	P1	P2
HMA	10.0	7.1	4.4	5.5
HMI	14.2	13.0	1.1	4.8
HML	14.4	13.2	0.8	4.3

哈尔明甲醇溶液分别进样。此实验进样使用一个 2mL 的 loop 阀。每次进样量为 2mL。结合位点数和离解常数所测得的数值与文献报道的对其他分子烙印测定的数值在一个数量级。前沿色谱分析结果与前面的液相色谱分析是一致的。在图 8-27 中, 哈尔满的容量因子比哈马灵和哈尔明的容量因子小, 前沿色谱

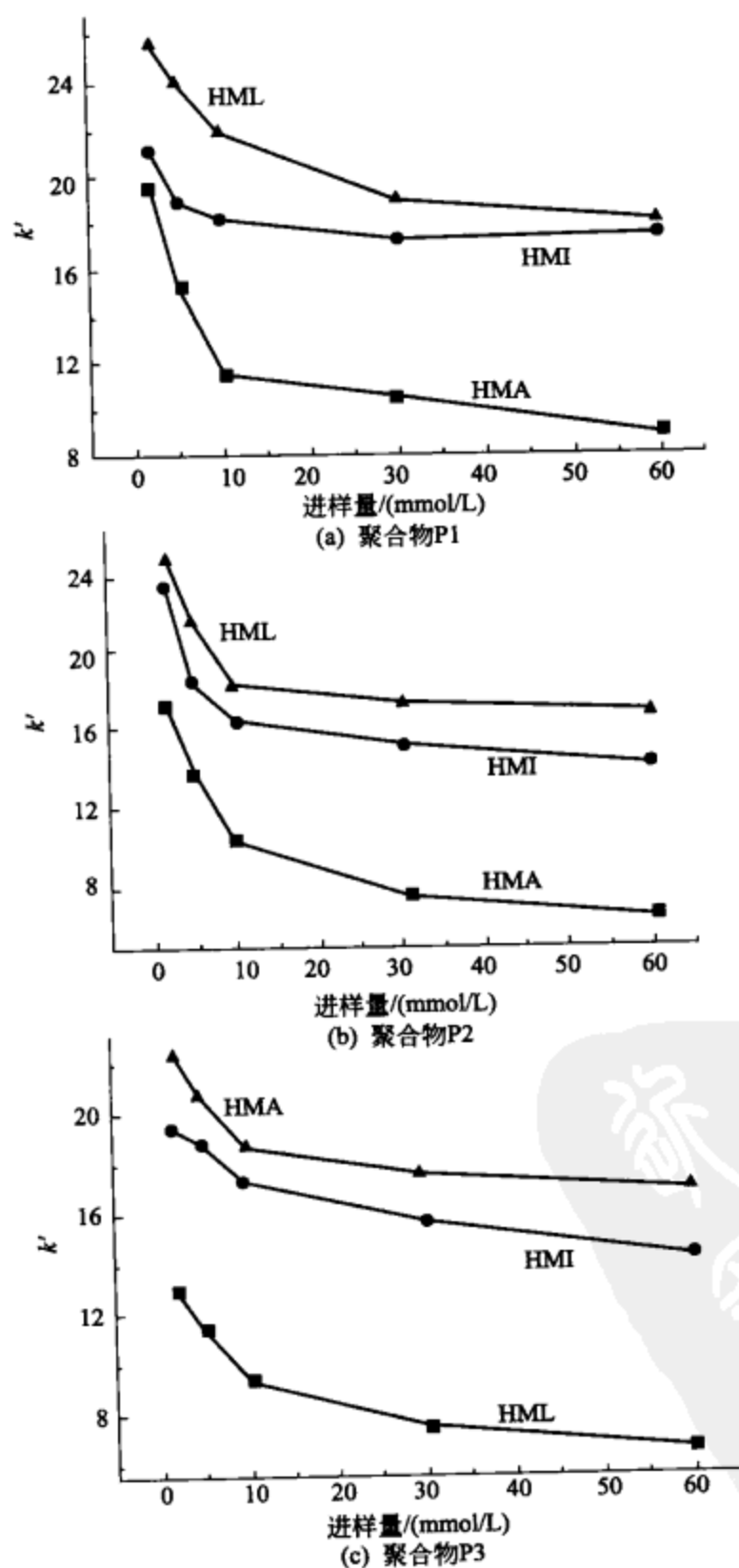


图 8-27 不同进样量时聚合物对 3 种目标分子的亲和能力

测得的哈尔满的结合位点数值也小于哈尔明及哈马灵，而且离解常数大于哈尔明及哈马灵的离解常数。但前沿色谱法测得的数值反映的是底物与聚合物之间的各种作用（特异性和非特异性）的共同结果，因此虽然在液相色谱上同一个化合物在 P1 与 P2 上的容量因子值差异较大，但在 P1 与 P2 上的结合位点数值差异并不明显。

8.4.5 骆驼蓬样品的制备

将 2g 骆驼蓬药草种子放在索式提取器上，用 40mL 乙醇溶液回流，直到溶液为淡粉色。溶液旋转蒸发至干后，再重新溶于 200mL 甲醇中。得到的溶液用 0.45 μ m 膜过滤，待用。

分别用哈尔满、哈马灵作为外标，液相色谱方法测定哈马灵及哈尔明在骆驼蓬溶液中的浓度。液相色谱流动相为甲醇/8mmol/L 三乙胺（1：1，体积比）溶液，流速为 1mL/min。进样量为 20 μ L。检测在 336nm 下进行。柱温为室温。

哈马灵及哈尔明在骆驼蓬溶液中的存在，还用正离子 ESI-TOF-MS 进行了确认。

8.4.6 MIP-LC-MS 在线分离鉴定

将装有分子烙印聚合物的液相色谱柱子出口直接与质谱仪相连，组成一个 LC-MS 系统。质谱使用一个 APCI 离子源，使用正离子检测。使用的流动相为甲醇。流速为 0.5 mL/min。对于 P1 柱进样量为 20 μ L，对于 P2 柱进样量为 13 μ L。柱温为室温。Bond-Elut C18 处理样品的方法如下：先将固相萃取柱（6mL Bond-Elut C18）用甲醇溶剂 5mL 活化，再用甲醇/8mmol/L 三乙胺（2：8，体积比）溶液平衡两次，每次 5mL。然后再将所制备的骆驼蓬甲醇溶液 0.5mL 缓慢加入到处理好的 Bond-Elut C18 小柱上。使用 9mL 甲醇分 3 次洗脱柱子。上样时流出液及洗脱液合并后旋干，再溶于 0.5mL 甲醇，待用。

（1）骆驼蓬提取液的分析测定结果 骆驼蓬提取液中哈马灵及哈尔明两种抗肿瘤成分的存在使用 ESI-MS（电喷雾离子化质谱）进行了确认。在图 8-28 中， m/z 213 为哈尔明化合物， m/z 215 为哈马灵化合物。 m/z 183 峰不存在，说明所制备的骆驼蓬提取液中基本不含哈尔满分子。液相色谱测出哈尔明及哈马灵的含量分别为 150 μ g/mL 和 82 μ g/mL。测定过程中使用哈尔明及哈马灵的标准溶液作为外标化合物，用峰面积进行定量（哈尔明的标准曲线为： $y=271.5x+3.83$ ， $n=5$ ，回归系数 $r=0.9995$ 。哈马灵的标准曲线为： $y=405.05x+3.68$ ， $n=5$ ，回归系数 $r=0.9986$ ）。

（2）MIP-LC-MS 分离实验 目前，在样品的萃取分离应用上 MIP 多数被离线（off-line）使用，即样品首先经过 MIP-SPE 纯化处理，再进行 HPLC 或 MS 的分析鉴定。目前，MIP 除了用于亲和色谱固定相用于手性化合物分离外，直接用作

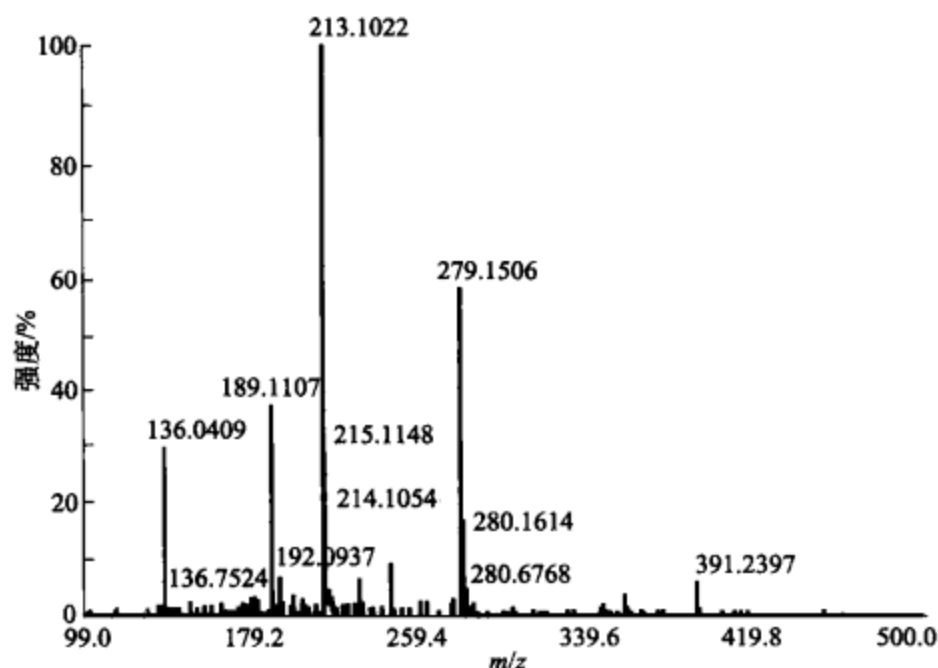


图 8-28 骆驼蓬提取液的质谱图 (ESI-MS)

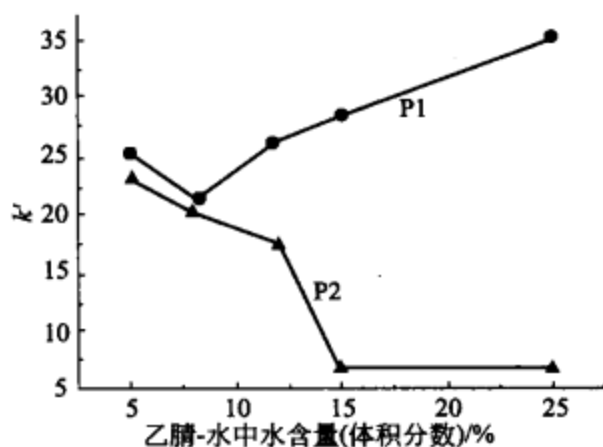
HPLC 固定相在线亲和分离复杂体系中的目标化合物尚未见报道。本研究尝试用分子烙印效率较高的 P1 和 P2 聚合物作为 HPLC 固定相，直接分离中草药提取物样品，并与质谱联用，实现亲和分离与在线鉴定的一体化 (MIP-LC-MS)。

首先尝试了将粗提物直接进样进行 MIP-LC-MS 实验，但由于粗提物中的未知组分干扰，分离效果较差。对于 P2 柱，从总离子流图上看到的哈马灵或哈尔明组分峰严重拖尾。但对于 P1 柱，从总离子流图上基本上观察不到哈马灵组分峰。另外，通过实验发现分子烙印聚合物 P1 和 P2 的稳定性和重现性均较好。在粗提液进样后，将 P2 柱再用乙酸-甲醇 (8:2, 体积比) 强极性流动相活化，用哈尔满单独进样，发现哈尔满在 P2 柱上的保留时间与原来基本相同。P1 柱直接分离粗提物后，HPLC 柱压很高，但可采用下面的方法使 P1 柱恢复：将聚合物倒在烧杯中，用乙酸-甲醇 (8:2, 体积比)、甲醇溶剂反复洗后再装入 HPLC 柱中。

从上面的实验结果可以看出，粗提物中的未知组分对 P1 和 P2 柱的干扰程度是不一样的。P1 和 P2 的这种不同表现，很可能与 P1 和 P2 制备时所用的致孔剂有关。一方面，不同的致孔剂在聚合过程中溶解聚合物链的程度不一样，从而在不同的致孔剂中聚合得到的交联聚合物网络不同。另一方面，在不同的致孔剂中模板分子与功能单体结合程度不同，从而使聚合过程中过量的功能单体在聚合物中的随机分布不同。因此在不同致孔剂中得到的聚合物的交联网络结构或极性不同，这很可能造成在直接分离粗提物时聚合物 P1 和 P2 对骆驼蓬粗提物中所含的一些未知组分表现不同的非特异性吸附，从而在液相色谱分离过程中 P1 和 P2 两个柱子受到的干扰程度不同。由于目前对于分子烙印聚合物的孔结构及表面特性还没有直接、有效的表征方法，上面的解释还只能是一种推测。

然而, 经过 C18 填料处理, 将粗提物中的极性弱的亲脂性组分除去后, 在 MIP-LC-MS 实验中, 骆驼蓬提取液中的哈马灵和哈尔明组分被选择性分离并鉴定出来 (见后文的图 8-30~图 8-38)。由此可知, 直接分离粗提物时的干扰物质很可能是一些弱极性化合物。P1 和 P2 两个聚合物的不同疏水性特点可以在某种程度上证明这种推测。

由图 8-29 可知, 当乙腈-水流动相中水的含量逐渐增加时, 由于水能破坏模板



分子哈尔满与聚合物上羧基的结合作用, 哈尔满分子在 P1 和 P2 柱上的保留均减小。但当水的含量很高时, 哈尔满在 P2 柱上的保留减小, 在 P1 柱上的保留反而增强, 说明 P1 具有明显的疏水性。一般而言, 固定相的疏水性越强, 其对脂类等弱极性化合物的吸附越强。因此 P1 和 P2 两个柱子对弱极性物质的不同程度吸附很可能是由于 P1 和 P2 聚合物的不同疏水性造成的。

图 8-29 哈尔满在 P1 和 P2 柱上的容量

因子随乙腈-水中水含量的变化

图 8-30~图 8-38 是 MIP-LC-MS 分离

预处理后的骆驼蓬提取液 (经过 C18 填料

处理) 的结果。由于 P1 的分子烙印效率较高, 柱子的容量较大, 在 P1 柱上的进

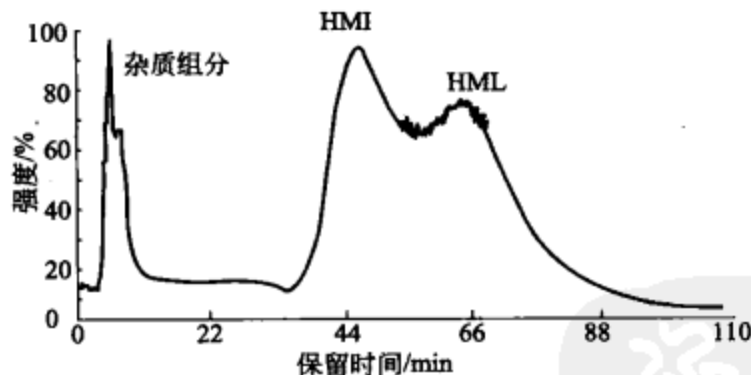


图 8-30 P1 柱上粗提物的总离子流图

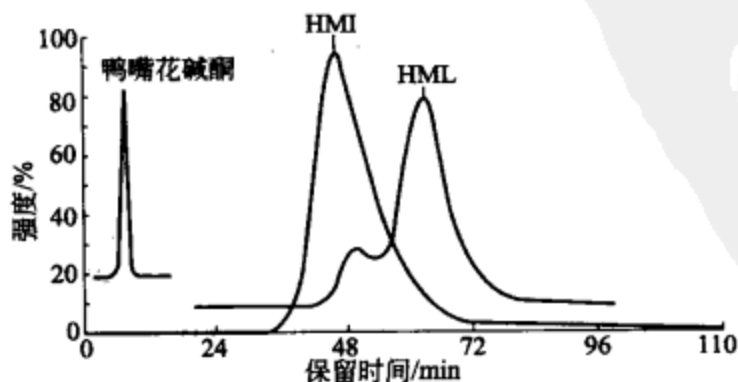


图 8-31 粗提物经过 P1 柱 LC-MS 的选择离子流图

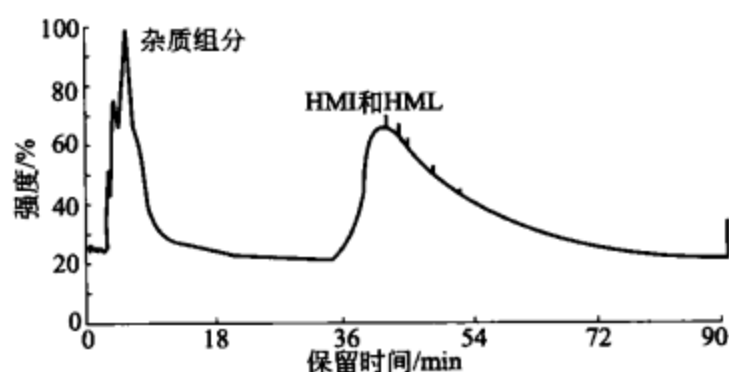


图 8-32 P2 柱上粗提物的总离子流图

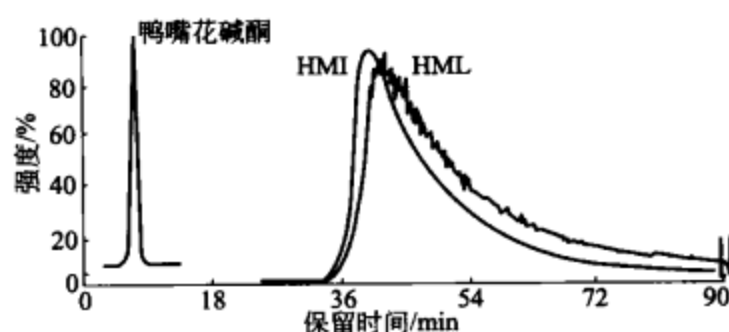


图 8-33 粗提物经过 P2 柱 LC-MS 的选择离子流图

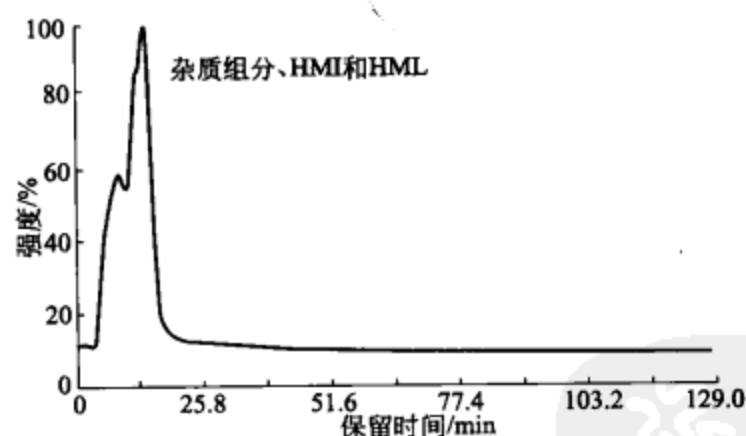


图 8-34 BP1 柱上粗提物的总离子流图

样量 ($20\mu\text{L}$) 多于在 P2 柱上的进样量 ($13\mu\text{L}$)。由图 8-30 与图 8-32 可以看出, 由于在分子烙印聚合物柱子 P1 或 P2 上的分离是亲和分离过程, 整个总离子色谱图上只包含一个杂质组分峰和一个目标组分峰。由于杂质组分在聚合物上的保留是靠非特异的表面吸附作用, 无论在 P1 上还是 P2 上, 骆驼蓬提取液中所含杂质组分在不到 20min 内在一个峰中被洗脱出来, 而亲和性目标化合物哈马灵和哈尔明保留很强, 使哈马灵和哈尔明抗肿瘤组分与杂质组分很好地分离开。另外, 由于 P1 的分子烙印效率更高, 选择性更好, 哈尔明与哈马灵两个化合物在 P1 柱上得到了部分分离 (图 8-31 和图 8-33), 而在 P2 柱上两个化合物没有分开。由总离子流

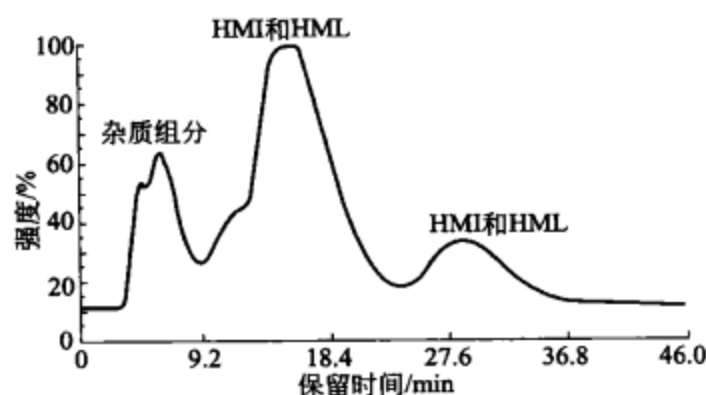


图 8-35 BP2 柱上粗提物的总离子流图

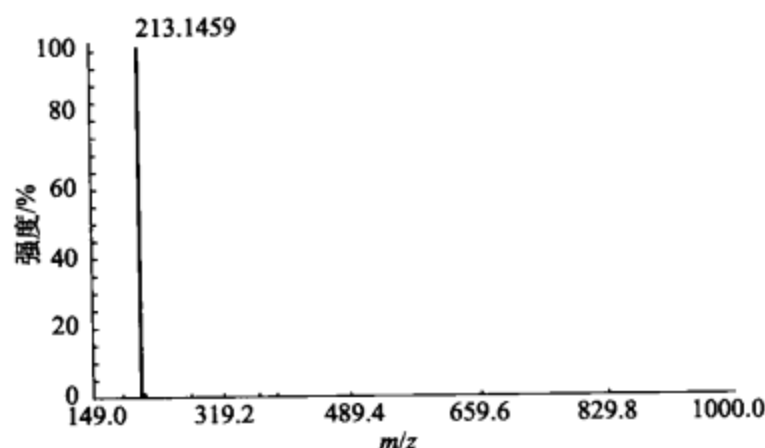


图 8-36 哈尔明对应的选择离子质谱图

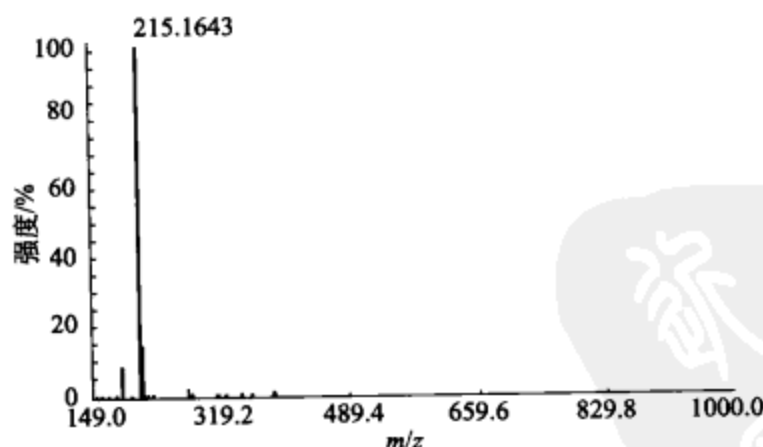


图 8-37 哈马灵对应的选择离子质谱图

色谱图还可以看出, 在分离多组分样品 (骆驼蓬提取液) 时, 由于存在其他组分的非特异吸附, 哈马灵和哈尔明的容量因子小于它们单独进样时的容量因子。

在分离骆驼蓬提取液这一复杂体系时, 分子烙印聚合物的分子识别能力还可以从 LC-MS 选择离子流图进一步证实。由图 8-31 和图 8-33 可以看出, 骆驼蓬提取液中的一个含氮碱性化合物鸭嘴花碱酮 (vasicinone) (图 8-38, m/z 203) 是在杂

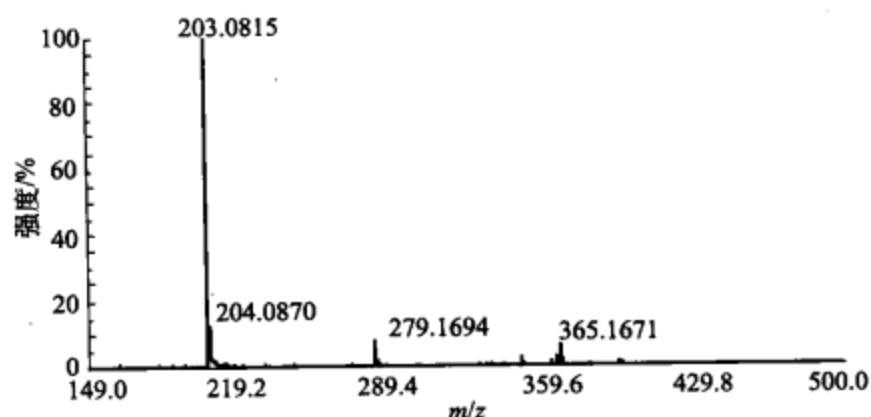


图 8-38 鸭嘴花碱酮对应的选择离子质谱图

质峰中被洗脱出来的。从前面的结果可知，由于提取液中所含的其他组分干扰，在 ESI-MS 分析骆驼蓬提取液时，并没有看到鸭嘴花碱酮这一化合物的质谱峰。可见，对于复杂体系，在线 LC-MS 分离鉴定更具有优越性。

用 P1 或 P2 对应的空白聚合物柱子 BP1 或 BP2 柱分离骆驼蓬提取液并没有得到上面所述的分离结果。在 BP1 柱上，骆驼蓬提取液中的杂质组分与目标组分哈尔明与哈马灵一起洗脱出来（图 8-34）。在 BP2 柱上，骆驼蓬提取液中的杂质组分与目标组分哈尔明与哈马灵没有很好地分开（图 8-35）。从而可以得出结论，P1 或 P2 是靠特异性的分子烙印作用而不是表面吸附作用，将骆驼蓬提取液中的哈尔明与哈马灵抗肿瘤组分选择性地分离出来的。

8.5 分子烙印识别中药中抗表皮生长因子受体（EGFR）的活性成分

EGFR 已被国际上认为是一种非常重要的药物设计靶点，从中药中寻找抗 EGFR 的抑制剂作为抗癌药物的替代物或先导化合物也是目前研究的一个热点。

8.5.1 聚合物制备

(1) 以槲皮素 (quercetin) 为模板分子制备聚合物 将槲皮素 1mmol 溶于致孔剂 (THF : CH₃CN = 3 : 1) 中，加入丙烯酰胺 4mmol，超声混合。然后再加入乙二醇二甲基丙烯酸酯 30mmol、偶氮二异丁腈 41.7mg，通氮气除氧 10min，抽真空，液氮冷却封管。60℃ 水浴聚合 24 h 后，磨碎过筛 (35μm)，颗粒小于 35μm 部分用丙酮反复沉降除去细末，最后再将聚合物真空干燥。聚合物粒子干法装入不锈钢柱 (φ4.6mm×150mm) 中，然后用甲醇-乙酸 (9 : 1，体积比)、甲醇在线洗脱，直到得到稳定的基线，以保证除去模板分子。作为对照，空白聚合物 (BP) 制备方法同上，只是不加模板分子。

(2) 以白皮杉醇 (piceatannol) 为模板分子制备聚合物 制备过程同上。

8.5.2 藏锦鸡儿粗提物的制备

通过对大量中药粗提物的活性筛选,选择一种藏药——藏锦鸡儿(*Caragana jubata*)作为研究对象。将2g藏锦鸡儿的根部粉碎后,用85%甲醇-水溶液浸泡1周(室温),然后将得到的粗提物旋转蒸发至干,再分别用极性由小到大的溶剂进行萃取,依次为石油醚、氯仿、乙酸乙酯、正丁醇、水。然后分别测定各分段产物的抗EGFR活性,结果表明其乙酸乙酯部分具有较高的活性,因此以下主要研究该部分粗提物。先把乙酸乙酯部分萃取物旋转蒸发至干,再将5mg该粗提物溶于5mL乙腈中,配成1mg/mL溶液,待用。

8.5.3 槲皮素-MIP柱固相萃取藏锦鸡儿中活性成分

一般固相萃取分为主要3部分进行,包括进样、冲洗和洗脱。通过对不同溶剂(如甲醇、乙腈、氯仿、二氯甲烷、四氢呋喃等)进行大量的筛选,最后选定的条件为:进样和冲洗溶剂为乙腈,洗脱溶剂为甲醇:乙酸=9:1。

MIP的固相萃取柱先用10mL乙腈活化,然后将400 μ L藏锦鸡儿的乙酸乙酯粗提物(溶解于乙腈,相当于400 μ g)直接进样到固相萃取柱上,等到吸附完全后,先用10mL乙腈冲洗,分3次进行,依次为3mL、4mL、3mL,之后再用8mL甲醇-乙酸(9:1,体积比)进行洗脱,共分8次进行,每次1mL。将冲洗后的乙腈部分及洗脱下来的甲醇-乙酸部分分别旋干,溶解于100 μ L甲醇中,分别用质谱定性、色谱定量。空白聚合物的固相萃取柱也用同样的过程处理。表8-9列出了MIP和NIP(空白聚合物)柱固相萃取之后的结果。

表 8-9 中药中抗 EGFR 抑制剂白皮杉醇和紫柳因的 MIP 以及 NIP 的固相萃取 (SPE) 结果

SPE 步骤		抑制剂(白皮杉醇和紫柳因)的量/ μ g			
		MISPE		BSPE (空白 SPE)	
		紫柳因	白皮杉醇	紫柳因	白皮杉醇
进样(全部)		400(含杂质组分)			
冲洗	1~3mL	— ^①	—	—	23
	4~7mL	—	0.3	16.7	1.8
	8~10mL	—	0.2	1.6	0.9
洗脱	1mL	—	1.5	0.3	—
	2mL	—	3.4	—	—
	3mL	0.6	17.5	—	—
	4mL	0.4	2.0	—	—
	5mL	11	—	—	—
	6mL	2.0	—	—	—
	7mL	1.9	—	—	—
	8mL	0.3	—	—	—
	合计	16.2	24.4	18.6	25.7

① 由于杂质的干扰,无法测定分析物。

从上面固相萃取的过程可以得到如下结果。

① 藏锦鸡儿的乙酸乙酯粗提物部分经过 MIP 的固相萃取柱后, 将其冲洗和洗脱下来的部分分别进行活性测定, 发现洗脱下来的部分比上柱前的粗提物活性有很大提高, 说明经过 MIP 的固相萃取柱后活性成分得到了富集。而经过 NIP 柱却不能达到这种效果: 冲洗下来的部分活性与粗提物活性相当 (表 8-9 中只列出了两个主要分析物的数据, 实际上冲洗过程中还有许多其他的无活性或低活性成分被冲洗下来, 它们与聚合物之间是一种非特异性的吸附); 而洗脱下来的部分也没有活性。也就是说通过 NIP 的固相萃取柱并不能达到活性富集的效果。用 10mL 乙腈冲洗固相萃取柱的过程中, 两种分析物 (白皮杉醇和紫铆因) 在 NIP 柱上几乎完全被洗下来, 而在 MIP 柱上保留较强, 要用甲醇-乙酸才能洗脱下来, 显然这两种分析物与槲皮素聚合物之间的作用是一种特异性的相互作用。对甲醇-乙酸洗脱 MIP 柱后的产物 (白皮杉醇和紫铆因) 分别进行活性测定, 得到它们抗 EGFR 的活性分别为: 白皮杉醇的 $IC_{50} = 4.9 \mu\text{mol/L}$, 紫铆因的 $IC_{50} = 10 \mu\text{mol/L}$ 。这两种成分是首次从藏锦鸡儿中被提取出来。

② 从上述的固相萃取过程可以看出, 经过 MIP 固相萃取柱后共有 $16.2 \mu\text{g}$ 紫铆因和 $24.4 \mu\text{g}$ 白皮杉醇从 $400 \mu\text{g}$ 乙酸乙酯粗提物中萃取出来, 从而计算出这两种化合物的含量分别为 4.1% 和 6.1%, 而用这两种化合物的纯品做标准品, 通过色谱峰面积定量, 可以计算出来在 $400 \mu\text{g}$ 乙酸乙酯粗提物中这两种化合物的实际含量分别为 $21.6 \mu\text{g}$ 和 $30.5 \mu\text{g}$, 也就是说它们的实际含量分别为 5.4% 和 7.6%, 从而得知 MIP 固相萃取柱的回收率分别为 76% (紫铆因) 和 80% (白皮杉醇)。损失部分主要来自 MIP 柱的不完全吸附。

③ 从上述洗脱过程可以看出, 在前面 3mL 洗脱液 (甲醇-乙酸, 9:1) 中主要是白皮杉醇被洗脱下来, 而紫铆因主要在后面 3mL 洗脱液中被洗脱下来, 这两种化合物能通过固相萃取过程分开, 而且显然它们两者在 MIP 柱上的保留能力是有区别的。实际在选择溶剂过程中, 发现用甲醇也能将白皮杉醇洗脱下来, 而紫铆因的保留能力较强, 只有加入少量的乙酸 (甲醇-乙酸, 9:1) 才能将它洗脱下来。

8.5.4 白皮杉醇-MIP 柱识别藏锦鸡儿中活性成分

如上所述, 通过槲皮素-MIP 柱固相萃取过程可以将藏锦鸡儿中两种抗 EGFR 的活性成分白皮杉醇和紫铆因直接从其粗提物中提取出来, 其中白皮杉醇具有相对较高的活性。从理论上讲, 活性较高的抑制剂具有更好的模拟酶 (这里指 EGFR) 的能力, 由于活性较高的分子与酶之间的相互作用更强, 从分子大小、形状及与酶之间相互作用的功能基团来看, 以该分子为模板合成出来的聚合物所形成的空穴能更准确地模拟与之有很好相互作用的酶的活性口袋, 因此预计用活性较高、更具有代表性的化合物作为模板得到的聚合物能更好地识别针对这种酶的抑制剂。这里又

用活性相对较高的分子白皮杉醇作为模板合成了一个分子烙印聚合物，研究该聚合物对藏锦鸡儿粗提物中抗 EGFR 抑制剂的识别能力。

(1) 聚合物选择能力的色谱表征 为了考察聚合物对模板分子及其他活性分子的选择能力，使用一组与模板分子白皮杉醇共存于藏锦鸡儿中的化合物（包含主要 3 种抑制剂）作为研究对象，测定了聚合物对这一组化合物的保留能力，图 8-39 显示了所研究化合物的结构式。MIP 的色谱表征参数见表 8-10。

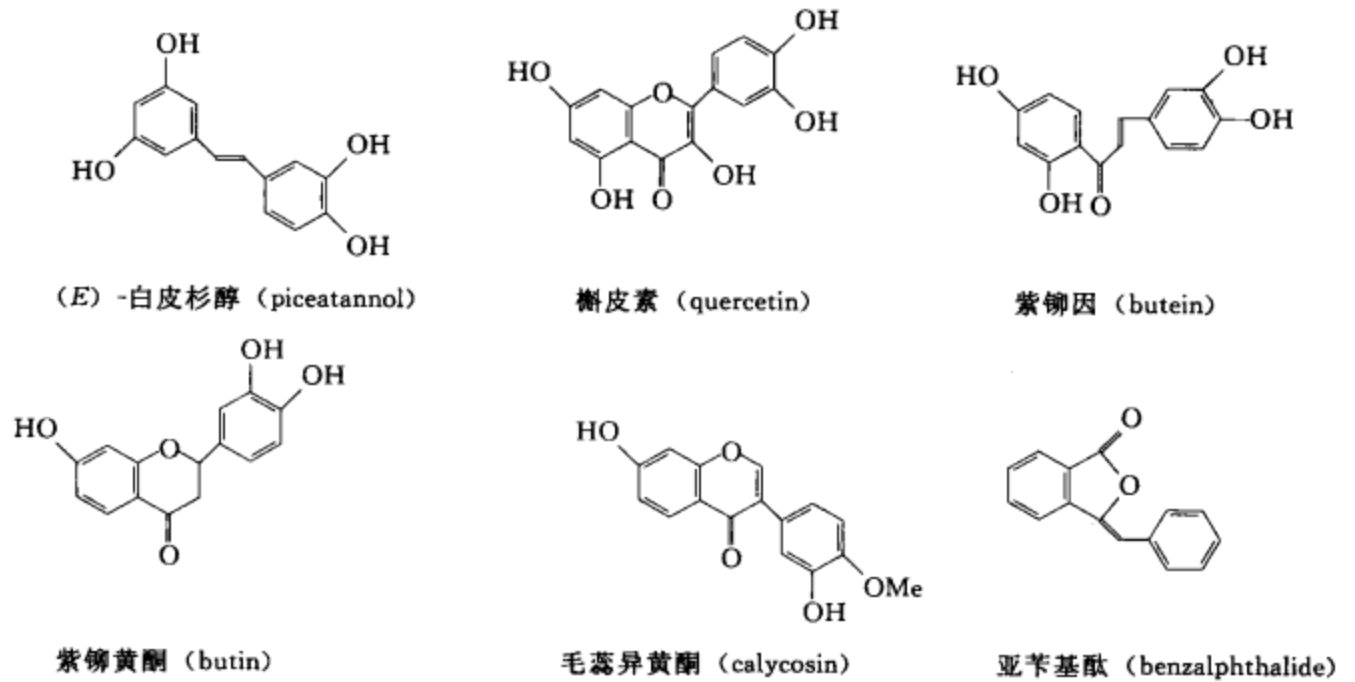


图 8-39 研究对象的分子结构示意图

表 8-10 白皮杉醇和其他分析物的色谱参数

参数	(E)-白皮杉醇 (piceatannol)	紫柳因 (butein)	槲皮素 (quercetin)	紫柳黄酮 (butin)	毛蕊异黄酮 (calycosin)	亚苄基酞 (benzaldehyde)
容量因子 k' (MIP)	9.71	5.57	4.14	0.45	0.27	0.14
容量因子 k' (NIP)	2.05	2.07	2.01	0.49	0.33	0.28
分离因子 α (MIP)	1	1.74	2.35	21.6	36.0	69.4
分离因子 α (NIP)	1	0.99	1.02	4.18	6.21	7.32
分子烙印效率 (IE)	4.74	2.69	2.05	0.92	0.82	0.50

从表 8-10 可以看出来，MIP 对模板分子具有最强的保留能力，此外对另两种化合物（紫柳因和槲皮素）也表现出相对较强的保留，而其他 3 种化合物在 MIP 上的保留非常弱。从结构及功能基团的角度方面可以对此作出解释：从结构相似性来看，紫柳因与模板分子比较相近，而亚苄基酞与模板分子结构相差最远；从功能基团来看，模板分子与功能单体起作用的可能是酚羟基，紫柳因和槲皮素分子中带有较多可以与功能单体相作用的酚羟基，而毛蕊异黄酮上该功能基团较少，亚苄基酞上没有。因而 MIP 体现出来对紫柳因有较强的保留，而对亚苄基酞的保留效果

最差。以上色谱常数是以前以甲醇为溶剂检测得到，当以乙腈为溶剂时，MIP 对各化合物的保留时间相对延长（其中模板分子的保留时间长达 90min），对它们的保留效果差异更显著。这可能因为以甲醇这种质子型溶剂检测 MIP 柱的效果时比乙腈能更强地破坏模板分子与功能单体之间的非键相互作用（主要是氢键相互作用）。

(2) 利用聚合物识别藏锦鸡儿粗提物中的活性组分 这里依然以藏锦鸡儿的乙酸乙酯部分为研究对象，将该部分粗提物直接进样，用 HPLC 检测 MIP 柱对粗提物各组分的识别能力。溶剂为甲醇，流速为 0.5mL/min；紫外检测（波长为 254nm）；进样量为 20 μ L，浓度为 10 μ g/mL；柱温为室温。主要组分的鉴定用高分辨质谱（ESI-TOF），质谱的工作条件如下所示：电离条件为负离子；喷嘴电压为 4500V，喷嘴电压为 100V，探测器电压为 1950V；喷嘴温度为 140 $^{\circ}$ C，四极温度为 140 $^{\circ}$ C；其他参数设置均为缺省值。图 8-40 显示了藏锦鸡儿的乙酸乙酯部分经过 MIP 柱的液相色谱图。

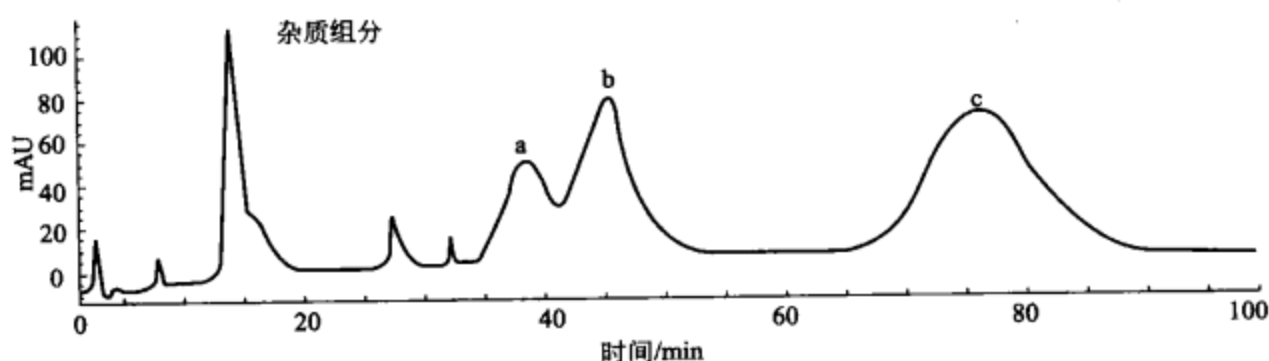


图 8-40 藏锦鸡儿的乙酸乙酯粗提物在 MIP 柱上的液相色谱图
mAU—毫吸光度单位

从图 8-40 中可以看出，粗提物中主要有 3 个组分 a、b、c 被 MIP 柱特异性地识别出来，而大量其他组分在上柱后不久就被洗脱下来。主要成分 a、b、c 通过质谱以离线（off-line）的方式进行检测，即将与 3 个色谱峰对应的组分接收下来，浓缩后用质谱检测，图 8-41~图 8-43 为 a、b、c 对应的质谱图。

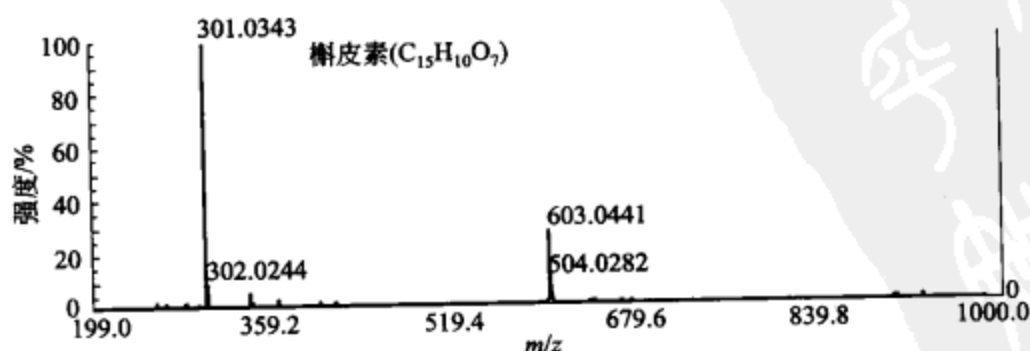


图 8-41 图 8-40 中峰 a 对应的分子质谱图

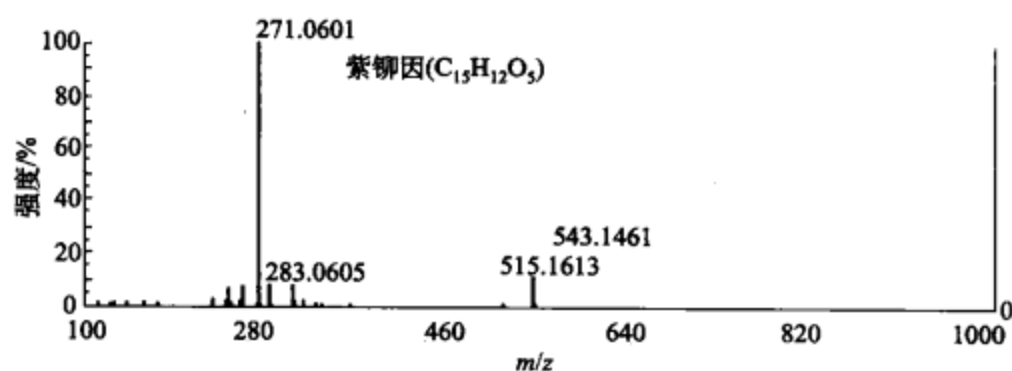


图 8-42 图 8-40 中峰 b 对应的分子质谱图

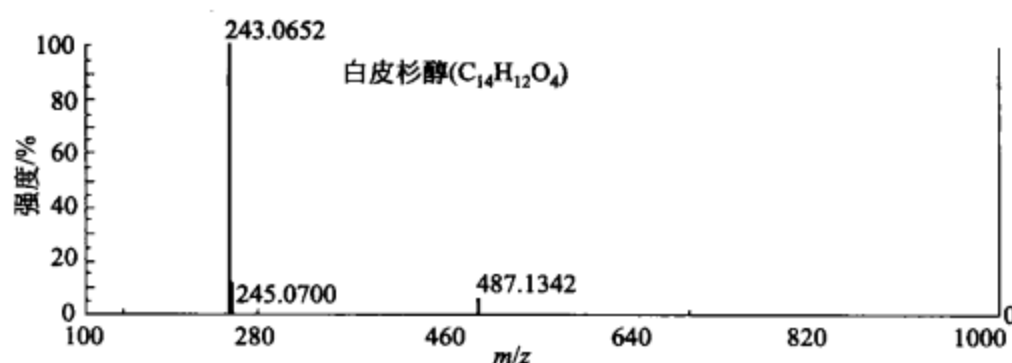


图 8-43 图 8-40 中峰 c 对应的分子质谱图

通过质谱鉴定，a、b、c 3 个组分分别为槲皮素、紫柳因和白皮杉醇，其中模板分子在 MIP 柱上保留时间最长，其次是紫柳因，然后是槲皮素，这与以上的色谱表征结果一致。从质谱图可以看出，所得到的组分基本没有杂质峰的存在，也就是说这 3 种活性分子可以通过 MIP 柱被选择性地纯化出来。因此利用 MIP 可以避免传统的分离提取烦琐而冗长的过程，而直接从粗提物中将活性成分提取出来。从图 8-40 中可以看出，以甲醇为溶剂时，MIP 柱对 a、b 两个组分并不能做到基线分离，因为它们两者在 MIP 柱上的保留性差别不大（表 8-10 也说明了这个问题）。但若以乙腈为溶剂，则这两个组分可以基线分离，只是整体保留时间延长。

(3) 聚合物模拟酶的能力评价 这里主要考察以白皮杉醇为模板分子合成聚合物模拟酶 (EGFR) 的能力，因此将聚合物对几个主要组分的识别能力（以烙印效应因子 IE 值表征）与这些化合物的活性结果（以 IC_{50} 值表征）进行对照，其中 IE 值越大说明聚合物对其保留能力越强， IC_{50} 值越小说明该化合物的活性越高。结果见表 8-11。为更加清楚地显示各组分在 MIP 上的保留效果与活性之间的关系，以 IF 对 $-\ln(IC_{50})$ 作图，如图 8-44 所示。

显然，活性越高的分子在 MIP 柱上的保留时间越长，也就是说活性越高的分子与 MIP 之间的相互作用越强。以白皮杉醇为模板分子的聚合物能选择性地

表 8-11 6 个分析物

化 合 物	分子烙印效率(IE)	抑制活性(IC ₅₀)/(μmol/L)
(E)-白皮杉醇(piceatannol)	4.74	4.9
紫柳因(butein)	2.69	10
槲皮素(queracetin)	2.05	15
紫柳黄酮(butin)	0.92	1.1
毛蕊异黄酮(calycosin)	0.82	1.2
亚苄基酞(benzalphthalide)	0.50	59.0

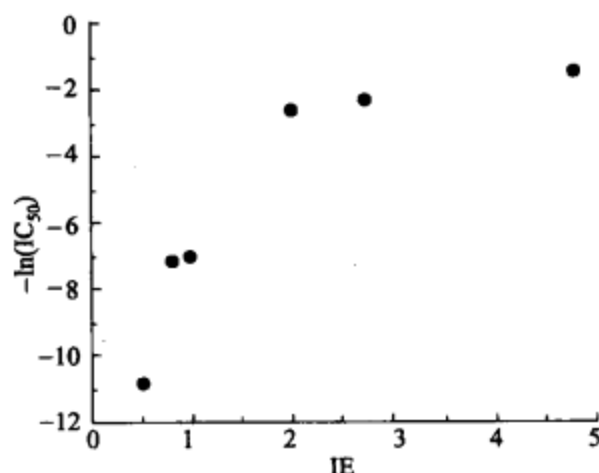


图 8-44 各化合物在 MIP 柱上的保留能力与其活性关系图

锦鸡儿粗提物中的几种主要抗 EGFR 抑制剂识别出来。而比较以槲皮素为模板分子的 MIP 选择性结果可以发现, 白皮杉醇比紫柳因活性高, 但在槲皮素-MIP 柱上的保留却相对较差。可以预测, 若以活性更高、更具有代表性的抑制剂为模板合成相应的 MIP, 应该具有更好的识别能力, 能广泛应用到其他更多的中药体系。这也正是下一步进行研究的课题之一。

8.6 分子烙印在药物开发方面的应用前景

上面以几种不同结构的分子为模板, 分别合成了其相应的分子烙印聚合物, 并首次将其应用到中药系统, 直接从粗提物中将模板分子及其结构类似物提取出来, 或者将其他活性成分提取出来, 达到了快速有效地分离目标成分的目的。分子烙印作为传统的分离提取的一种补充手段, 在药物开发方面有极大的应用前景, 总结如下。

(1) 寻找先导化合物 以一种已知的活性化合物为模板合成相应的分子烙印聚合物, 可以将其用于合成的组合化学库或天然组合化学库——中药体系(包括复方), 直接从构型多样的库中提取出来其他结构类似的先导化合物, 从而避免了传统分离提取的非特异性和低效性。

(2) 寻找已知药物替代品 高活性的抑制剂在成为药物之前, 必须经过药理研究、毒性筛选等过程, 许多有效的抑制剂往往因为高毒性或副作用或在体内吸收不好而无法最终成药。而中药经过中国人民几千年的积累和体内实验, 一般具有较低的毒性。若以某种高效高毒性的分子作为模板分子烙印出来聚合物, 可以直接从这种天然组合化学库中筛选出来其他有效且低毒性的化合物作为替代品。或者某种高效无毒的药物分子因为合成困难而导致非常昂贵, 也可以用分子烙印的方法寻找其他成本低廉、容易得到的替代品。

(3) 高效提取微量(或痕量)成分 在中药的传统分离提取过程中, 许多有效成分往往因为含量低微而被漏筛或难以纯化出来, 利用分子烙印的强特异性和高选择性可以将一些含量很低的化合物直接从粗提物中提取出来。

(4) 制定中药质量标准 中药在中国的应用已经有几千年的历史, 但往往由于其成分的复杂性和有效成分不够明确而难以在国际市场上被认可。如果能够利用分子烙印的方法对某些有效成分的含量进行监控(可以用色谱或质谱的手段检测目标产物的存在并确定其含量), 制定一套质量标准, 将有利于中药的进一步应用推广。

参 考 文 献

- [1] 姜忠义, 吴洪. 分子印迹技术. 北京: 化学工业出版社, 2003.
- [2] Alexander C, Andersson H S, Andersson L I, Ansell R J, et al. *J. Mol. Recognit.* **2006**, 19: 106-180.
- [3] Vlatakis G, Anderson L I, Muëller R, Mosbach K. *Nature*, **1993**, 361: 645-647.
- [4] Wulff G. *Angew. Chem. Int. Ed.*, **1995**, 34: 1812-1832.
- [5] Batra D, Shea K J. *Curr. Opin. Chem. Biol.*, **2003**, 7: 434-442.
- [6] J. Hall A, Emgenbroich M, Sellergren B. Imprinted polymers. *Top. Curr. Chem.*, **2005**, 249: 317-349.
- [7] Wulff G, Schauhoff S. *J. Org. Chem.*, **1991**, 56: 395-400.
- [8] Boonpangrak S, Prachayasittikul V, Bülow L, Ye L. *J. Appl. Polym. Sci.*, **2006**, 99: 1390-1398.
- [9] Zourob M, Mohr S, Mayes A G, Macaskill A, et al. *Lab Chip* **2006**, 6: 296-301.
- [10] Hui Li, Yingju Liu, Zhaohui Zhang, Haiping Liao, et al. *J. Chromatogr. A*, **2005**, 1098: 66-74.
- [11] Caro E, Marcé R M, Borrull F, Cormack P A G, et al. *Trends Anal. Chem.*, **2006**, 25: 143-154.
- [12] Hsieh R Y, Tsaib H A, Syub M J. *Biomaterials*, **2006**, 27: 2083-2089.
- [13] Karakhanov E A, Karapetyan L M, Kardasheva Y S, Maksimov A L, et al. *Macromol. Symp.*, **2006**, 235: 39-51.
- [14] Zhang H, Piacham T, Drew M, Patek M, et al. *J. Am. Chem. Soc.*, **2006**, 128: 4178-4179.
- [15] Potyrailo R A. *Angew. Chem. Int. Ed.*, **2006**, 45: 702-723.
- [16] Rathbone D L. *Adv. Drug Deliv. Rev.*, **2005**, 57: 1854-1874.
- [17] Mosbach K, Yu Y H, Andersch J, Ye L. *J. Am. Chem. Soc.*, **2001**, 123: 12420-12421.
- [18] Xie J, Zhu L, Xu X. *Anal. Chem.*, **2002**, 74: 2352-2360.
- [19] Zhu L, Xu X. *J. Chromatogr. A*, **2003**, 991: 151-158.

- [20] Zhu L, Chen L, Xu X. *Anal. Chem.*, **2003**, 75: 6381-6387.
- [21] Zhuang Y, Luo H, Duan D, Chen L, et al. *Anal. Bioanal. Chem.*, **2007**, 389: 1177-1183.
- [22] Zhou L, Xie J C, Ge Y F, Xu X J. *ACTA PHYSICO-CHIMICA SINICA*, **2002**, 18 (9): 808-811.
- [23] Xie J C, Zhu L L, Xu X J. *ACTA CHIMICA SINICA*, **2002**, 60 (3): 385-388.
- [24] Xie J C, Luo H P, Zhu L L, Zhou L, Li C X, Xu X J. *ACTA PHYSICO-CHIMICA SINICA*, **2001**, 17 (7): 582-585.
- [25] Xie J C, Chen L R, Li C X, Xu X J. *JOURNAL OF CHROMATOGRAPHY B-ANALYTICAL TECHNOLOGIES IN THE BIOMEDICAL AND LIFE SCIENCES*, **2003**, 788 (2): 233-242.
- [26] Liu S, Yi L Z, Liang Y Z. *Journal of Separation Science*, 2008, 31 (11): 2113-2137.
- [27] Lemos V A, Teixeira L S G, Bozerra M D. *Applied Spectroscopy reviews*, 2008, 43 (4): 303-334.
- [28] Sellergran B. *Anal. Chem.*, **1994**, 66: 1578.
- [29] Muldon M T, Stanker L H. *Anal. Chem.*, **1997**, 69: 803.
- [30] Baggiani C, Trotta F, Giraudi G, Giovannoli C, Vanni A. *Anal. Commun.*, **1999**, 36: 263.
- [31] Venn R F, Goody R J. *Methodological Surveys in Bioanalysis of Drugs*, **1998**, 25: 13.
- [32] Walshe M, Howarth J, Kelly M T, O'Kennedy R, Smyth M R. *J. Pharm. Biomed. Anal.*, **1997**, 16: 319.
- [33] Kugimiya A, Takeuchi T. *Anal. Chim. Acta*, **1999**, 395: 251.
- [34] Ju J Y, Shin C S, Whitcombe M J, Vulfson E N. *Biotechnol. Bioeng.*, **1999**, 64: 232.
- [35] Zander A, Findlay P, Renner T, Sellergren B, Swietlow A. *Anal. Chem.*, **1998**, 70: 3304.
- [36] Mullett W M, Lai E P C, Sellergren B. *Anal. Commun.*, **1999**, 36: 217.
- [37] Martin P, Wilson I D, Morgan D E, Jones G R, Jones K. *Anal. Commun*, **1997**, 34: 45.
- [38] Olsen J, Martin P, Wilson I D, Jones G R. *Analyst*, **1999**, 124: 467.
- [39] Matsui J, Okada M, Tsuruoka M, Takeuchi T. *Anal. Commun.*, **1997**, 34: 85.
- [40] Andersson L I, Paprica A, Arvidsson T. *Chromatographia*, **1997**, 46: 57.
- [41] Rashid B A, Briggs R J, Hay J N, Stevenson D. *Anal. Commun.*, **1997**, 34: 303.
- [42] Mullett W M, Lai E P C. *Anal. Chem.*, **1998**, 70: 3636.
- [43] Mullett W M, Lai E P C. *Microchem. J.*, **1999**, 61: 143.
- [44] Ferrer I, Barcelo D. *Trends Anal. Chem.*, **1999**, 18: 180.
- [45] Bjarnason B, Chimuka L, Ramstrom O. *Anal. Chem.*, **1999**, 71: 2152.
- [46] Kempe M. *Anal. Chem.*, **1996**, 68: 1948.
- [47] Vidyasankar S, Ru M, Arnold F H. *J. Chromatogr. A*, **1997**, 775: 51.
- [48] Matsui J, Miyoshi Y, Doblhoff-Dier O, Takeuchi T. *Anal. Chem.*, **1995**, 67: 4404.
- [49] Matsui J, Nicholls I A, Takeuchi T. *Tetrahedron: Asymmetry*, **1996**, 7: 1357.
- [50] Ramstrom O, Yu C, Mosbach K. *J. Mol. Recogn*, **1996**, 9: 691.
- [51] Wulff G, Schauhoff S. *J. Org. Chem.*, **1991**, 56: 395.
- [52] Spivak D A, Shea K J. *Macromolecules*, **1998**, 31: 2160.
- [53] Haginaka J, Takehira H, Hosoya K, Tanaka N. *J. Chromatogr. A*, **1999**, 849: 331.
- [54] Kempe M, Mosbach K. *J. Chromatogr. A*, **1994**, 664: 276.
- [55] Matsui J, Takeuchi T. *Anal. Commun.*, **1997**, 34: 199.
- [56] Matsui J, Kato T, Takeuchi T, Suzuki M, Yokoyama K, Muguruma E, Elgersma A V, Karube I. *Anal. Lett.*, **1995**, 28: 2317.

- [57] Ramström O, Ye L, Krook M, Mosbach K. *Anal. Commun.*, **1998**, 35: 9.
- [58] Tarbin J A, Sharman M. *Anal. Commun.*, **1999**, 36: 105.
- [59] Cheong S H, McNiven S, Rachkov A, Levi R, Yano K, Karube I. *Macromolecules*, **1997**, 30: 1317.
- [60] Fischer L, Müller R, Ekberg B, Mdsbach K. *J. Am. Chem. Soc.*, **1991**, 113: 9358.
- [61] Suedee R, Songkram C, Petmoreekul A, Sangkunakup S, Sankasa S, Kongyarit N. *J. Pharm. Biomed. Anal.*, **1999**, 19: 519.
- [62] Walshe M, Garcia E, Howarth J, Smyth M R, Kelly M T. *Anal. Commun.*, **1997**, 34: 119.
- [63] Lin J M, Nakagama T, Ucgutana K, Hobo T. *Chromatographia*, **1996**, 43: 585.
- [64] Tan Z J, Remcho V T. *Electrophordsis*, **1998**, 19: 2055.
- [65] Schweitz L, Andersson L I, Nilsson S. *Anal. Chem.*, **1997**, 69: 1179.
- [66] Schweitz L, Andersson L I, Nilsson S. *J. Chromatogr. A*, **1997**, 792: 401.
- [67] Nilsson K, Lindell J, Norrlov O, Sellergren B. *J. Chromatogr. A*, **1994**, 680: 57.