计算机辅助药物分子设计

徐筱杰　侯廷军　乔学斌　章威　编著

化学工业出版社
现代生物技术与医药科技出版中心
·北 京·

需要全本请访问 www.ertongbook.com

（京）新登字 039 号

图书在版编目（CIP）数据

计算机辅助药物分子设计/徐筱杰等编著. —北京：化学工业出版社，2004.4
ISBN 7-5025-5520-X

Ⅰ.计… Ⅱ.徐… Ⅲ.药物-化学结构-计算机辅助设计-高等学校-教材 Ⅳ.R962

中国版本图书馆 CIP 数据核字（2004）第 043844 号

计算机辅助药物分子设计

徐筱杰 侯廷军 乔学斌 章威 编著

责任编辑：杜进祥

文字编辑：焦欣渝

责任校对：洪雅姝

封面设计：郑小红

*

化 学 工 业 出 版 社
现代生物技术与医药科技出版中心 出版发行

（北京市朝阳区惠新里 3 号 邮政编码 100029）

发行电话：（010）64982530

http://www.cip.com.cn

*

新华书店北京发行所经销

化学工业出版社印刷厂印刷

三河市前程装订厂装订

开本 787mm×1092mm 1/16 印张 34¾ 字数 864 千字

2004 年 7 月第 1 版 2004 年 7 月北京第 1 次印刷

ISBN 7-5025-5520-X/R·220

定 价：75.00 元

版权所有 违者必究

该书如有缺页、倒页、脱页者，本社发行部负责退换

前　言

从20世纪90年代开始，随着各种学科的发展，药物设计进入了一个崭新的阶段，即合理药物设计的阶段。所谓合理药物设计，是依据生物化学、酶学、分子生物学、遗传学、信息学以及计算化学等学科的研究成果，针对这些基础研究中所揭示的包括酶、受体、离子通道以及核酸等潜在的药物设计靶点，并参考其他类源性配体或天然产物底物的化学结构特征设计出合理的药物分子，以发现选择性作用于某种靶点的新药，这是目前新药发现的主要方向。合理药物设计取得重大突破的重要原因之一就是计算机辅助药物设计（CADD）方法的发展。从20世纪60年代构效关系方法提出以后，经过40多年的探索和努力，尤其是20世纪90年代以后，随着多种新的方法的出现以及计算机技术的高速发展，CADD已经从基础理论的研究开始过渡到实际应用的阶段，各种CADD参与设计的药物已经相继上市或进入临床研究阶段。

从20世纪70年代至今，我国的CADD方法研究和应用取得了很大的进展，但与国际先进水平相比，我国仍有较大的差距。另外，我国介绍CADD方法以及相关理论计算方法的专著还比较少，为了使我国读者更系统地了解计算机辅助药物分子设计及相关理论方法，促进我国在这一领域中的技术进步和发展，作者根据掌握的国内外最新进展和多年研究工作的积累编写了本书。本书的内容主要分为三个部分：第一部分介绍了计算机辅助药物设计的发展概况与计算机辅助药物分子设计有关的统计和理论计算方法，包括数值和非数值优化方法、分子力学方法、分子动力学方法、构象分析方法、溶剂效应以及自由能计算方法；第二部分介绍了计算机辅助药物分子设计中的主要方法，包括构效关系方法、药效团模型方法、从头设计方法、分子对接方法、数据库搜索方法、药代动力学以及毒性的预测方法；第三部分则选择了两个具体的实例，介绍了计算机辅助药物分子设计方法在药物先导化合物发现和优化方面的应用。附录一和附录二则简要介绍了CADD方法中几个常用的概念及列出了CADD中常用的软件和程序。在介绍CADD方法和应用的基础上，本书还用了大量篇幅介绍CADD中涉及的相关理论和计算方法，这部分内容涉及了较多的线性代数、高等数学、分析力学、统计力学以及量子力学等学科的相关内容。这些理论计算方法在CADD中应用非常广泛或是某些CADD方法的重要组成部分，对其系统的了解可以帮助读者更加深入地理解后续章节中介绍的CADD方法。当然，这

几章的内容也具有相对的独立性，读者完全可以跳过这部分内容而直接阅读CADD方法和应用的相关章节。本书没有单独分出章节介绍量子化学方法，主要原因是这部分涉及的内容非常广泛，在很小的篇幅内很难对其做出系统及有效的阐述，此外，国内外关于这方面的专著也很多，有需要的作者可以参考有关的著作。本书力求介绍CADD方法研究和应用领域的最新进展，但CADD是国际上非常活跃的研究领域，发展速度很快，涉及的内容非常广泛，加上作者学识有限，在写作的过程难免有遗漏和错误，恳请读者批评和指正。

本书的第一章由徐筱杰教授执笔，第二章～第三章由乔学斌博士执笔，第四章由侯廷军博士和章威执笔，第五章～第六章由章威和侯廷军博士执笔，第七章～第十三章由侯廷军博士执笔，第十四章由侯廷军博士和乔学斌博士执笔，第十五章～第十六章由侯廷军博士执笔，附录由侯廷军博士、乔学斌博士和章威共同完成。最后全书由徐筱杰教授和侯廷军博士修改定稿。

本书是徐筱杰教授课题组多年研究工作的结晶，是组里一届又一届本科生、研究生研究工作的经验总结和积累。多年来我们的工作得到国内同行们的许多支持与帮助，也得到养生堂药业有限公司、创腾科技有限公司及凯鹏公司的大力支持与帮助。作者特向他们致以衷心的感谢。书中引用了大量国内外同行发表的一些宝贵资料，在此一并致谢。

编者

2004年3月

目　　录

第一章　药物研究及计算机辅助药物分子设计…… 1
第一节　药物研究和开发的历史及现状 …… 1
第二节　现代药物研究的四大技术支柱…… 3
一、分子生物学、基因组学及蛋白质组学…… 3
二、组合化学…… 5
三、高通量筛选…… 6
四、与药物研究相关的信息科学及技术…… 7
第三节　计算机辅助药物分子设计 …… 11
一、概述 …… 11
二、CADD 方法的分类 …… 12
参考文献 …… 15

第二章　计算化学中的最优化方法 …… 17
第一节　引论…… 17
一、最优化问题概述 …… 17
二、数学预备知识 …… 18
三、最优化条件 …… 26
第二节　数值最优化方法 …… 30
一、最优化算法的基本结构 …… 30
二、无约束问题的最优化方法 …… 33
参考文献 …… 53

第三章　计算化学中的非数值最优化方法 …… 54
第一节　引论 …… 54
一、计算复杂性 …… 54
二、局部搜索算法 …… 56
三、组合优化问题算法设计的思路 …… 58
第二节　模拟退火 …… 63
一、算法的基本结构 …… 63
二、冷却进度表的选择 …… 66
三、几种改进算法 …… 73
第三节　遗传算法 …… 77
一、算法概述 …… 78
二、遗传算法的数学基础 …… 83
三、算法的改进形式 …… 86

需要全本请访问 www.ertongbook.com

四、常规遗传算法的变形 …… 92
第四节 神经网络 …… 94
一、人工神经网络基本模型简介 …… 94
二、用于优化计算的网络模型——连续型 Hopfield 网络 …… 96
三、自组织网 …… 100
参考文献 …… 103

第四章 分子力场和力场参数化 …… 106
第一节 分子力场的势函数形式 …… 108
一、键伸缩能 …… 109
二、键角弯折能 …… 110
三、二面角扭转能 …… 111
四、交叉相互作用项 …… 112
五、范德华相互作用能 …… 112
六、静电相互作用 …… 113
七、共轭体系 …… 115
第二节 分子力场的分类 …… 115
一、传统力场 …… 115
二、第二代力场 …… 118
三、通用力场 …… 120
四、其他力场 …… 121
第三节 力场参数的拟合 …… 121
一、基于振动频率的力场参数拟合 …… 123
二、基于实验值的范德华参数拟合 …… 129
三、基于量化计算的力场参数拟合 …… 131
参考文献 …… 135

第五章 构象分析方法 …… 139
第一节 小分子的构象分析方法 …… 139
一、系统搜索方法 …… 139
二、片段连接方法 …… 142
三、随机搜索方法 …… 143
四、距离几何方法 …… 143
五、分子动力学方法 …… 146
六、基于遗传算法的构象分析方法 …… 148
第二节 蛋白质结构的预测 …… 149
一、同源模建 …… 150
二、Modeler ——一个简单而有效的蛋白质模建方法 …… 159
三、反相折叠方法 …… 160
四、从头折叠方法 …… 161

参考文献…… 162

第六章　分子动力学…… 164
第一节　积分方法…… 165
第二节　初始化…… 167
第三节　粒子受力的求算 …… 169
一、键伸缩能…… 169
二、键角弯折能…… 170
三、二面角扭转能…… 171
四、非键相互作用能…… 172
第四节　边界条件…… 173
一、周期性边界条件…… 173
二、非周期性边界条件…… 174
第五节　非键相互作用能的处理…… 175
一、截断值方法…… 175
二、FMM 方法 …… 177
三、Barnes-Huts 算法 …… 180
四、周期性边界条件方法…… 181
第六节　约束条件动力学…… 186
第七节　恒温和恒压分子动力学…… 188
一、恒温分子动力学…… 188
二、恒压分子动力学…… 188
第八节　分子动力学轨迹分析…… 189
一、静态热力学性质…… 190
二、相关函数…… 194
三、迁移特性…… 195
参考文献…… 195

第七章　溶剂效应…… 197
第一节　显示溶剂模型…… 198
第二节　连续介质模型…… 199
一、表面加和模型…… 199
二、泊松-玻耳兹曼模型 …… 202
三、GB/SA 模型 …… 208
四、模型多级矩展开方法…… 213
五、极化的连续介质模型…… 213
六、SMx 系列溶剂化模型 …… 215
参考文献…… 216

第八章　结合自由能计算…… 218

第一节　引言 …… 218
一、熵增加原理和 Gibbs 自由能 …… 218
二、受体-配体结合的亲和力 …… 220
三、自由能计算方法的分类 …… 223
第二节　自由能微扰和热力学积分方法 …… 223
一、自由能微扰方法 …… 224
二、热力学积分方法 …… 226
三、FEP 和 TI 在药物设计中的应用 …… 227
第三节　基于主方程的自由能计算方法 …… 229
一、MM/PBSA 方法的原理 …… 229
二、MM/PBSA 方法的应用 …… 230
三、MM/PBSA 方法的发展和完善 …… 232
第四节　基于经验方程的自由能预测 …… 234
一、Böhm 的自由能预测模型 …… 234
二、Eldredge 的自由能预测模型 …… 235
三、Head 的自由能预测模型 …… 236
第五节　LIE 方法 …… 237
一、LIE 方法的原理 …… 237
二、LIE 方法的应用 …… 239
三、LIE 方法的前景和缺陷 …… 241
参考文献 …… 242

第九章　定量构效关系方法研究 …… 246
第一节　二维定量构效关系方法 …… 246
一、Hansch 法 …… 246
二、Free-Wilson 法 …… 250
三、各种参数 …… 250
第二节　建立定量构效关系模型的统计方法 …… 263
一、回归分析 …… 263
二、遗传算法 …… 265
三、人工神经网络 …… 267
第三节　三维定量构效关系方法 …… 268
一、距离几何 3D-QSAR …… 268
二、分子形状分析 …… 270
三、比较分子场分析方法 …… 272
四、虚拟受体方法 …… 277
第四节　QSAR 的应用 …… 281
一、2D-QSAR 的应用 …… 281
二、3D-QSAR 的应用 …… 287
参考文献 …… 289

第十章　药效团模型方法 …… 295
第一节　药效团模型的表达 …… 296
一、药效特征元素 …… 296
二、几何约束 …… 299
第二节　药效团模型的识别 …… 300
一、药效团识别的基本步骤 …… 301
二、分子叠合和活性构象 …… 301
第三节　基于药效团模型的数据库搜索 …… 303
一、基于药效团的数据库搜索 …… 303
二、柔性构象搜索 …… 304
第四节　药效团识别系统 …… 305
一、Receptor …… 305
二、Apex-3D …… 306
三、DISCO …… 307
四、GASP …… 308
五、CATALYST …… 309
六、几种药效团识别系统的性能比较 …… 311
第五节　药效团模型方法的应用 …… 314
一、Muscarinic M_3 受体拮抗剂的设计 …… 315
二、5-HT3 受体拮抗剂的设计 …… 317
三、MC 增生抑制剂的设计 …… 318
四、PKC 抑制剂的设计 …… 319
五、HIV-1 整合酶抑制剂的设计 …… 320
参考文献 …… 321

第十一章　分子对接方法 …… 325
第一节　分子对接的原理 …… 325
一、分子对接的理论基础 …… 325
二、分子对接方法的分类 …… 326
三、分子对接方法中的重要问题 …… 326
第二节　几种有代表性的分子对接方法 …… 327
一、DOCK …… 328
二、AUTODOCK …… 332
三、FlexX …… 335
四、Affinity …… 337
五、LigandFit …… 340
六、SFDOCK …… 341
第三节　虚拟筛选的策略 …… 343
第四节　分子对接在药物设计中的应用 …… 345
一、胸苷酸合成酶抑制剂的设计 …… 348

二、DHFR 抑制剂的设计 …… 349
三、EGFR 和抑制剂间相互作用模式的研究 …… 349
参考文献 …… 350

第十二章 从头设计方法 …… 354
第一节 从头设计方法的分类 …… 354
一、片段定位法 …… 355
二、位点连接法 …… 356
三、片段连接法 …… 357
四、逐步生长法 …… 360
第二节 几种重要的从头设计方法 …… 362
一、GRID …… 362
二、MCSS …… 363
三、HINT …… 365
四、LUDI …… 366
五、Leapfrog …… 369
第三节 从头设计在药物开发中的应用 …… 371
一、fⅩa 抑制剂的设计 …… 372
二、全新 FKBP-12 配体的设计 …… 373
参考文献 …… 374

第十三章 分子三维结构数据库和虚拟组合化学 …… 377
第一节 三维结构数据库 …… 378
一、剑桥结构数据库 …… 379
二、国家癌症研究所 …… 380
三、Available Chemicals Directory 3D（ACD-3D） …… 380
四、Available Chemicals Directory-Screening（ACD-SC） …… 381
五、MDL Drug Data Report 3D（MDDR-3D） …… 381
六、Comprehensive Medicinal Chemistry（CMC） …… 381
七、类似物设计的结构数据库 BIOSTER …… 382
八、Chapman & Hall Dictionary of Natural Product（DNP） …… 382
九、Metabolite …… 382
十、Toxicity …… 382
十一、中草药三维结构数据库 …… 383
第二节 分子结构的拓扑表达和结构转化 …… 385
一、分子结构的拓扑表达 …… 386
二、分子三维结构的转化 …… 387
第三节 数据库搜索技术 …… 388
一、子结构匹配 …… 388
二、相似性搜索 …… 392

第四节　重要的三维结构搜索系统…… 395
一、MDL/Base …… 395
二、Unity …… 396
第五节　计算组合化学方法…… 396
一、组合化学…… 396
二、组合合成设计…… 397
三、虚拟组合库的筛选…… 400
参考文献…… 403

第十四章　药代动力学特征和毒性的预测…… 405
第一节　药代动力学特征的预测…… 406
一、脂水分配系数…… 406
二、脑血分配系数…… 414
三、肠通透性…… 420
四、水溶性…… 422
第二节　分子毒性的预测…… 425
一、引论…… 425
二、计算机辅助的化合物毒性预测方法…… 427
三、常见化合物毒性预测软件…… 436
参考文献…… 442

第十五章　EGF-R 和抑制剂间相互作用模式的研究 …… 447
第一节　酪氨酸蛋白激酶…… 447
一、PTK 的结构特征 …… 447
二、EGF-R 的结构特征 …… 448
三、EGF-R 抑制剂的分类 …… 450
第二节　苯胺喹唑啉类抑制剂的三维构效关系…… 454
一、计算方法…… 455
二、结果与讨论…… 461
第三节　EGF-R 和喹唑啉类抑制剂结合模式的预测 …… 466
一、计算方法…… 467
二、结果与讨论…… 469
参考文献…… 473

第十六章　HIV-1 蛋白酶抑制剂的设计 …… 476
第一节　HIV-1 病毒的结构和侵袭靶细胞的机制 …… 476
第二节　基于 HIV 蛋白酶的抑制剂设计 …… 480
一、HIV 蛋白酶的结构 …… 480
二、HIV 蛋白酶抑制剂 …… 481
三、HIV 蛋白酶抑制剂的计算机辅助设计实例 …… 488

需要全本请访问 www.ertongbook.com

参考文献…… 492

附录一　分子模拟方法中常用概念和名词…… 495
一、分子坐标表示…… 495
二、分子表面…… 498
三、分子图形显示模型…… 500
四、量化计算常见术语简介…… 502

附录二　药物分子设计中常用软件列表…… 505
1. 大型分子模拟软件系统 …… 505
2. 小型分子模拟软件系统 …… 507
3. 小型药物设计软件系统 …… 508
4. 常用量化计算软件 …… 509
5. 分子力学、分子动力学和蒙特卡罗模拟软件 …… 510
6. QSAR 和分子参数计算软件 …… 513
7. 药效团模拟软件 …… 518
8. 分子对接软件 …… 519
9. 从头设计方法 …… 522
10. 分子相似性和差异性分析以及组合库设计 …… 524
11. 药代动力学和毒性预测软件 …… 526
12. 蛋白质三维结构模建、结构评估和活性位点预测软件 …… 529
13. 数据库搜索软件 …… 532
14. 分子叠合 …… 532
15. 二维转三维和文件格式转换软件 …… 533
16. 分子显示软件 …… 534
17. 求解 PB 方程的软件 …… 536
18. 化学软件开发包 …… 536
19. 国内主要软件代理商的相关信息 …… 537

附录三　常见毒性数据库简介…… 538
1. RTECS（the Registry of Toxic Effects of Chemical Substances） …… 538
2. HSDB（the Hazardous Substances Data Bank） …… 538
3. TOXLINE …… 538
4. MICROMEDEX …… 538
5. Toxicity …… 538

第一章　药物研究及计算机辅助药物分子设计

第一节　药物研究和开发的历史及现状

自古以来，疾病一直威胁着人类的生存和发展，人类在实践中在不断地寻找征服疾病的方法与手段，药物治疗始终是其中最重要的方法与手段。在古代，人类就开始从自然界的植物、矿物及动物中寻找药物。中国、印度、南美等国家和地区都有非常久远的使用天然药物的历史。相传5000多年前，我国《神农本草》就记载了许多药用植物，其中常山是作为治疗发烧的药物，其活性成分生物碱在现在仍是治疗疾病的药物；另一种植物麻黄可以作为心脏的刺激剂、发汗剂，并具有减轻咳嗽的功效，它的有效成分麻黄碱仍是目前使用的药物。在国外，公元3世纪就已经把含有吗啡的植物作为止痛剂。虽然植物药已有数千年的使用历史，但对其活性成分的分离、纯化及深入研究还是近百年的事。至今从天然产物或天然产物的衍生物中发现药物仍是药物研究及开发中的主要途径之一。

对生理及生物化学过程中的意外结果进行追踪，也是发现药物的途径之一。青霉素（penicillin）的发现是一个最典型的例子。1928年，Alexander Fleming注意到葡萄球菌的培养中的绿色菌的生长，虽然其他人也观察到这个现象，但只有Fleming对这个现象进行追踪，进而发现了青霉素（Richard 1992）。最初，Fleming建议青霉素可用作防腐剂，但他没有在治疗感染疾病方面取得成功。到20世纪40年代，牛津大学的Howard Florey重新研究了青霉素，相继通过其他多个研究组的努力才把青霉素开发成了药物（Richard 1992）。

随机筛选即普筛可能是常规药物设计中应用最为广泛的手段。传统发现新药的主要途径就是基于大量化合物（包括动物、植物以及微生物来源的天然产物和合成化学品）在各种活性测定模型下通过广泛筛选获得。从20世纪30年代开始普筛就成为发现药物的主要方法，如抗疟药甲氟喹是从大量的合成化合物中筛选得到的。在20世纪40年代及20世纪50年代许多制药公司企图从土壤中通过随机筛选来寻找抗生素，结果发现了链霉素（streptomycin）及四环素两种重要的药物（Martin 1997）。但这个方法显然需要耗费大量的人力和物力，而且盲目性较大，因此命中率低。但鉴于疾病病理的复杂性，至今人们还是不能完全预测和控制化合物的生物活性，故普筛仍然在药物发现上发挥着重要的作用，同时也是先导化合物发现的重要途径之一。

通过化学和生物的方法对各种途径获得的具有特定生物活性的先导化合物进行修饰和结构改造，以提高化合物的活性和减少副作用是新药寻找的常用方法。因为先导化合物的基本结构已经得到，因此成功率较普筛会大大提高。如头孢菌素、二氢吡啶类抗钙拮抗剂都是通过对先导化合物的改造而获得的。以中药有效成分作为先导化合物进行结构改造是我国的优势，青蒿素衍生物蒿甲醚及青蒿琥酯是成功的例子。但这种方法只能得到先导化合物的衍生物及类似物。因此，一般来讲，新化合物的活性和副作用方面会有所改善，但对作用本质不可能有重大突破。

基于疾病或药理的作用机理及临床观察也是发现新药的一条途径。一种候选药物通过动物实验或临床观察可能发现另外的活性及用途。1947 年，一种抗组胺剂在 Johns Hopkins 大学作临床实验时被发现它还具有治疗晕车、晕船及晕飞机的作用，现在已广泛用于治疗各种运动眩晕药物（Martin 1997）。

从 20 世纪 90 年代开始，随着各种学科的发展，药物设计进入了一个新的阶段，即合理药物设计的阶段。所谓合理药物设计，是依据生物化学、酶学、分子生物学、遗传学、信息学以及计算化学等学科的研究成果，针对这些基础研究中所揭示的包括酶、受体、离子通道以及核酸等潜在的药物设计靶点，并参考其他类源性配体或天然产物底物的化学结构特征设计出合理的药物分子，以发现选择性作用于某种靶点的新药，这是目前新药发现的主要方向。合理药物设计取得重大突破的重要原因之一就是计算机辅助药物设计方法（CADD）的发展。从 20 世纪 60 年代构效关系方法提出以后，经过 40 多年的探索和努力，尤其是 20 世纪 90 年代以后，随着多种新的方法的出现，CADD 已经发展为一门完善和新兴的研究领域。CADD 方法的出现，大大提高了药物开发的效率，也为人们攻克一些顽症提供了崭新的思路和成功的希望。

现代药物研究可以归纳为如下几个阶段：

20 世纪 40～60 年代，主要药物研究是从大量合成的化合物或天然产物中筛选药物，一般的概率是从几万个化合物中筛选出一个药物。

20 世纪 70～80 年代，随着受体生物化学及计算机辅助药物设计的，除了随机筛选外，逐步向比较理性有效的方式寻找先导化合物。如开展基于作用机理及基于受体及靶酶结构的药物设计。

20 世纪 90 年代～现在，药物研究主要依赖于分子生物学特别是基因组、蛋白质组、组合化学、高通量筛选及生物信息学与药物设计。

药物工业是全球销售额最大发展最快的产业之一。据统计，2001 年全球的销售额为 3920 亿美元，其中 70 个国家的销售额为 3642 亿，比 2000 年增长 12%，美国是世界最大的药物市场，1992 年占全球的市场的 34%，2002 年占 50%，预计 2006 年将达 3300 亿美元的销售额。药物市场中畅销药占很大的比例，其中前 10 种药约占全球市场的 30%。最畅销的是降胆固醇（cholesterol）、降甘油三酯（triglyceride）口服减肥药，年增长达到 20%。药物市场的另一特点是大公司占据市场的很大份额，如与 Pharmacia 公司合并后，Pfizer 公司是全球最大的药物公司，占市场份额的 11%，前 10 位的公司占市场销售额的 47%。

药物研究与开发是一个多学科交叉、长周期、高投入的系统工程，其开发流程见图 1.1 所示。

几个重要步骤包括：

（1）新先导化合物的发现　筛选与鉴定与某一疾病有关的靶标；克隆、表达、分离、纯化相关的酶/受体，表征及测序；获得相关的酶/受体的结构信息（NMR、X 射线、结构预测）；发展活性测定方法；随机筛选或计算机辅助药物设计后合成并测定活性。

（2）先导化合物的优化　建立专一性测定方法；建立细胞水平的活性及专一性的检测方法；安全及低水平动物实验；计算机辅助药物设计改造先导化合物；组合化学优化先导化合物；结构-活性关系研究设计新化合物。

（3）临床前的研究　动物安全研究；药代动力学特征以及毒性的研究；计量及制剂研究。

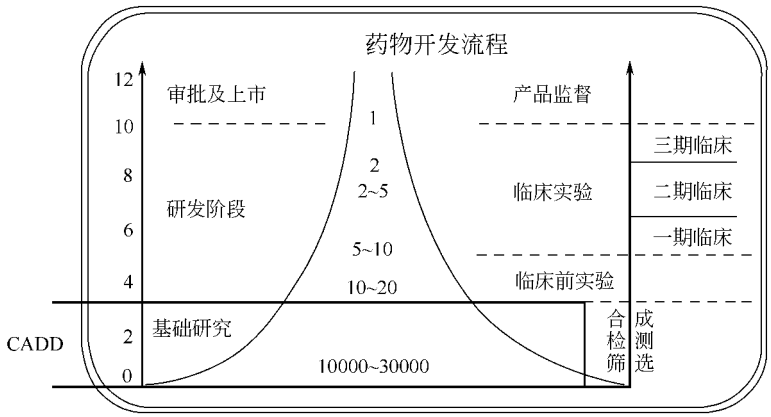

图 1.1　药物研究与开发流程

（4）临床研究　准备申请报告；一期临床；二期临床；三期临床。

（5）进入市场。

一个新药的研究与开发过程涉及多学科的交叉与配合，包括生物化学、分子生物学、生物物理学、免疫学、微生物学、合成化学、结构化学、分析化学、药物化学、临床医学、统计学以及计算机科学等多方面学科。

药物研究与开发不仅周期长，过程涉及多学科并且投资费用大。早年开发一个新药要投几千万至 1 亿美元，现在一般要几亿美元（有人估计为 8 亿美元）。各大制药公司的研究与开发的费用在大幅度的增长。如全球最大的制药公司 Pfizer 公司 2002 年的研究与开发的费用为 48 亿美元，2003 年将达到 70 亿美元。

在 20 世纪 90 年代各制药公司花了 300 亿美元用于基因研究。如此巨大的投资并没有得到相应的回报。在 2001 年仅 37 个新化学物质问世是 20 年来最低。2002 年 9 月底前欧洲的医药产品评估机构 EMEA（European Ageney for Evaluation of Medicinal Products）仅收到 25 个新药申请，而 2001 年为 58 个。美国 FDA 2002 年收到 16 个新药申请，而 2001 年收到 24 个新药申请。由此可见，药物工业也是一个高风险的产业。虽然如此，药物产业仍是一个非常有前景及高速发展的产业，这是因为社会需求在不断增长。美国的处方药的销售目前为 1406 亿美元，据估计 2012 年将增至 4459 亿美元。另外一个原因就是相关科学技术的迅速发展将大大促进药物研究与开发的发展。分子生物学（特别是基因组学、蛋白组学及代谢组学）、组合化学、高通量筛选、生物信息学与计算机辅助药物分子设计的快速发展已成为现代药物研究与开发的四大技术支柱，有可能缩短药物开发的周期，提高药物开发的效率（Class 2002）。

第二节　现代药物研究的四大技术支柱

一、分子生物学、基因组学及蛋白质组学

分子生物学（特别是基因组学及蛋白质学）的发展不仅为药物研究提供新的靶标，而且为疾病的机理提供新的认识。受体仍是当前药物研究的核心，虽然分子生物学在受体研究中起着重要作用，人类基因组学及蛋白质组学的研究正逐渐起着重要的主导作用。人类基因包括 4 万个基因，所编码的蛋白质的数目就更多了，如激酶类的蛋白估计有 500～2000 个，这些基因及

蛋白质都有可能成为有用的药物标靶。据 1996 年统计（Draws 2002），目前在药物研究中心药物标靶仅 483 个，其中受体占 45%、酶占 28%、激素及生长因子占 11%、核酸受体占 2%、离子通道占 3%、未知的标靶占 7%。虽然目前基因组及蛋白质组的研究提供药物标靶的数目不多，但基因组及蛋白质组的进展将大大增加药物标靶的数目，可开发出大量的新药。

新加坡大学 Chen X. 年开发了一个 TTD 数据库（therapeutic target database）包括 433 个核酸及蛋白靶标及有关资料（包括靶标与疾病的关系、与配体的相互作用及条件）（Chem 2002）。通过与相关的数据库连接可获得这些靶标的序列、三维结构及功能的信息。

蛋白质组的概念是在 1995 年提出的，旨在研究一个基因组、细胞或组织中蛋白质的完全表达，近年来已迅速发展为蛋白质组学（proteomics）（Burbaum 2002）。可以通过比较疾病与正常样品的差别发现药物靶标。从蛋白组学研究到药物靶标发现的过程是多步骤的复杂过程，包括如下几个步骤：

① 细胞表达及检测；

② 基因克隆；

③ 蛋白分离、纯化及鉴定；

④ 蛋白质的功能表征及活性位点的确定；

⑤ 蛋白质的结构测定；

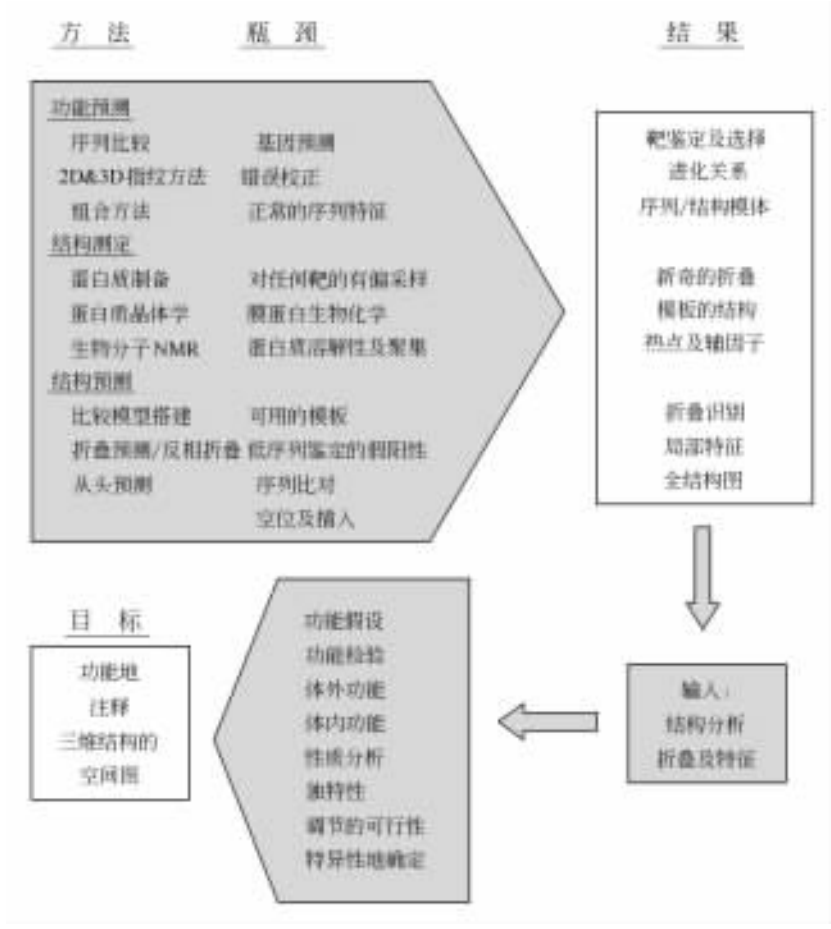

图 1.2 结构蛋白质组学的示意

⑥ 与治疗有关的实验证实；

⑦ 活性检测方法的确定。

随着蛋白组学的进展，结构蛋白质组学也应运而生（Yanagida 2002），它包括蛋白质的功能预测、蛋白质的结构测定及预测，如图 1.2 所示。

二、组合化学

组合化学（combinatorial chemistry）是用巧妙的战略和程序，从某些单体或基本化合物出发，快速合成大量化合物的方法。组合化学的概念是匈牙利科学家 Arpad Furka 于 1982 年首次提出（Furka 1982），1988 年他在一个学术会议上作了报告引起巨大的反响，20 世纪 90 年代组合化学得到迅速发展并广泛用于药物先导化合物的发现及优化。

组合化学的原理可描述如下：①A＋B→AB；②如 A，B 各有 n 个衍生物可组合得到 A^nB^n 个化合物，如图 1.3 所示。

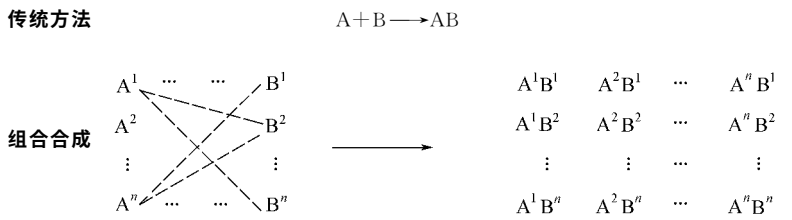

图 1.3　传统方法与组合合成比较

组合化学合成方法包括同步合成技术（parallel approach）、混合-均分技术（mixture-split synthesis）及 selectide 技术、编码合成技术、解卷积法，以及基于 DNA 的“噬菌体”展示技术等。

目前，组合化学的概念已有较大的扩展，“库”的概念已很广泛。如微生物的代谢产物、蛋白质的酶解液，以及植物的提取液等都可称为库。只要有相当数量的一群分子混合物，且能用于功能活性的筛选，都称为一种库。各种“库”可分为单一化合物组成的库及混合物库（见表 1.1）。

表 1.1　混合物库和单一化合物库

混　合　物　库	单一化合物库
最大的筛选通量	容易质量控制
库要求很大的化学工作及假阳性干扰	增加生物活性数据的可靠性
库要求用解卷积或茶袋技术	高的筛选费用

在组合化学合成多采用固相合成技术，但液相合成也有它的优点。固相合成可以通过混合-均分法合成大数目的化合物，产品可以通过物理方法分离，容易自动化，可以合成长序列化合物。但固相合成方法试剂及溶剂消耗大，受化学反应条件的限制，反应监测困难，同时得到产物的量不大（通常小于 50mg）。液相合成方法反应类型及反应条件范围较宽，且产物的量可以比较大，但分离纯化要比较长的时间。

在药物研究中，组合化学可用于发现全新先导化合物（大数目及丰富多样性的库）、先导化合物的优化（小的比较集中的库）、现有先导化合物的结构-活性关系研究以及扩展专利范围等。图 1.4 为组合化学用于药物研究的流程。

需要全本请访问 www.ertongbook.com