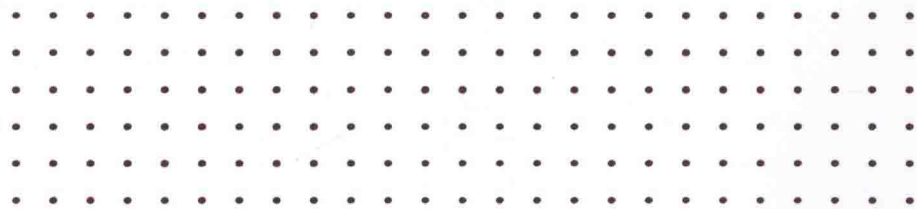


Progress in Medicinal
Chemistry

药物化学进展

9



-
- 主 编 彭司勋
 - 副主编 赵守训 廖清江 尤启冬

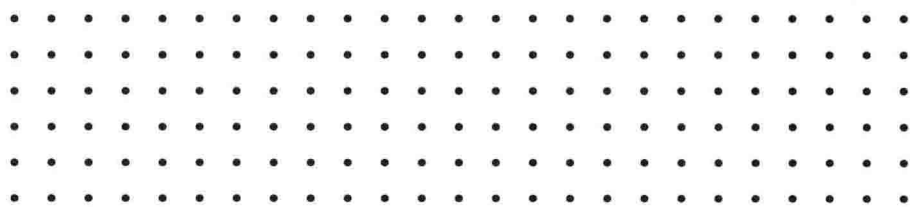


化学工业出版社

Progress in Medicinal
Chemistry

药物化学进展

9



-
- 主 编 彭司勋
 - 副主编 赵守训 廖清江 尤启冬



化学工业出版社

本书是“药物化学进展”系列图书的第9卷,该系列图书邀请国内药物化学和有关学科的专家从各自研究或熟悉的领域撰写文章,多侧重于新药的新理论、新靶点、新成果和新方法的介绍,旨在便于读者了解新药研究的前沿动态,有助于扩展视野,开拓思路。

本书收载15篇综述文章,其中关于药物设计、新合成方法和干扰RNA方面的综述3篇,抗肿瘤药物研究进展4篇,抗HIV病毒药物研究进展2篇,天然药物及其衍生物的研究进展3篇,抗高血压药物、神经保护剂和抗菌药物的研究进展3篇。

本书适合从事新药研发的人员阅读参考,同时也可作为相关专业高年级本科生、研究生的教学参考书。

图书在版编目(CIP)数据

药物化学进展.9/彭司勋主编. —北京:化学工业出版社,2014.4
ISBN 978-7-122-18942-4

I. ①药… II. ①彭… III. ①药物化学-进展-世界
IV. ①R914-11

中国版本图书馆CIP数据核字(2013)第265168号

责任编辑:杨燕玲 孙小芳
责任校对:王素芹

文字编辑:向东
装帧设计:张辉

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)

印刷:北京永鑫印刷有限责任公司

装订:三河市万龙印装有限公司

787mm×1092mm 1/1 印张20¼ 彩插1 字数521千字 2014年4月北京第1版第1次印刷

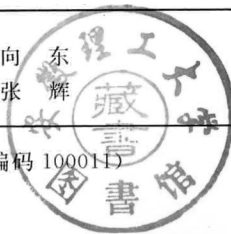
购书咨询:010-64518888(传真:010-64519686) 售后服务:010-64518899

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:98.00元

版权所有 违者必究



前 言

本卷共刊载 15 篇综述文章，前 3 篇是关于药物设计、新合成方法和干扰 RNA 方面的综述，其次是抗肿瘤药物研究进展 4 篇，抗 HIV 病毒药物研究进展 2 篇，天然药物及其衍生物的研究进展 3 篇，其余 3 篇分别为抗高血压药物、神经保护剂和抗菌药物的研究进展。

“药效团模型方法与应用研究进展”对基于受体和配体的药效团模型构建原理、方法及相关计算机辅助药物设计软件进行了比较全面的介绍。同时阐述了药效团模型在虚拟筛选和全新药物设计中的应用。对从事药物设计者具有参考价值。

“高效合成技术在药物化学中的应用”介绍了下述高效合成技术及其在药物化学中的应用：微波化学、点击化学、组合化学技术、串联反应和多组反应技术、DOS 和 pDOS 技术。上述合成技术为快速合成复杂性、多样性的活性化合物提供了技术支撑。

“干扰 RNA (siRNA) 药物研究进展”综述了 siRNA 的发展过程、作用机制、化学修饰，以及目前在临床研究阶段的 siRNA 药物，此类药物具有广泛应用前景，深受广大研究者的关注。为研究 siRNA 药物这一较新领域提供了重要参考。

“共价抑制剂在肿瘤化学生物学中的应用”综述了共价抑制剂的作用机制、作用特点，以及在肿瘤化学生物学中的应用。共价抑制剂是抗肿瘤的新领域，虽然尚不够成熟，如对其结构进行合理设计和优化可能为肿瘤治疗提供新希望，其研究前景值得关注。

“酪氨酸激酶 c-Met 小分子抑制剂的研究进展”综述了通过阻断该激酶的信号转导发挥抗肿瘤作用，总结了近年来该类抑制剂的研究概况，并着重分析了 c-Met 抑制剂先导化合物的发现和/或优化过程，提出研发抗肿瘤活性好、选择性强的新抑制剂的思路。

“小分子 p53-MDM2 结合抑制剂的研究进展”全面论述了 p53-MDM2 的生理功能，及其在肿瘤发生发展过程中的重要作用，并重点综述此类抑制剂的研究进展。对多种表达野生型 p53 的肿瘤细胞有效，与不同作用机制的抗肿瘤药物联合使用有协同作用，副作用较小。

“Nrf2 信号转导通路介导的化学预防：抗氧化应激结合抗炎作用”综述了 Keap1-Nrf2 信号通路及其在癌症治疗中的潜在作用。Keap1-Nrf2-sMaf-ARE 转导通路，是亲电子和化学预防剂产生的主要信号感应器，是引发抗氧化反应的关键信号通路。该通路可能是潜在的化学预防药物的作用部位，为抗氧化应激及恶性炎症的治疗带来希望。

“HIV 进入抑制剂的研究进展”针对艾滋病感染过程中涉及的多种蛋白为靶点的 HIV 进入抑制剂，分别综述了小分子进入抑制剂、以 gp41 胞外区为靶点的融合抑制剂、趋化因子辅助受体 CCR5 或 CXCR4 抑制剂。上述抑制剂针对病毒未侵入人体细胞完成永久感染的初期发挥作用，提供了可能优于其它抗 HIV 的治疗方法。

“靶向 HIV-1 整合酶的抗逆转录病毒药物研究”介绍 HIV-1 整合酶的结构和功能及其抑制剂的结构类型和作用机理，重点介绍整合酶链转移抑制剂和 LEDGF/p75-整合酶抑制剂的研究进展，对该领域的研究有所启迪。

“寻找抗高血压药物作用新靶点：辣椒素受体”对辣椒素受体的结构、表达与分布，及

其介导的生物学效应等作了较为全面的综述。辣椒素受体是调节舒血管物质降钙素基因相关肽 (CGRP) 合成与释放的受体, 此文对抗高血压药物的研究提供了新的视野。

“神经保护剂研究进展”针对中枢神经系统疾病, 如: 神经退行性疾病、缺血性脑卒中、创伤性脑损伤等, 从此类疾病所导致的神经元的死亡、脑组织梗死和中枢神经系统功能损伤等角度, 综述了神经保护剂的研究进展, 可为研发更有效的治疗药物提供参考。

“AI-2 型细菌群体感应抑制剂的研究进展”, 细菌群体感应 (QS) 是指细菌通过分泌一种称为自诱导剂 (AI) 的信号分子来感应细菌种群密度及周围环境变化, 进而调控相关基因的表达, 协调细菌生物功能的一种信息交流机制。AI-2 是目前发现的在革兰阳性菌和革兰阴性菌中都存在的自诱导剂。此文从抑制自诱导剂的合成和抑制受体蛋白的作用机制介绍了 AI-2 抑制剂的研究进展, 可能对抗菌药物的研究提供新的思路。

“灵菌红素生物碱及 Roseophilin 的研究进展”系统综述了灵菌红素类生物碱的生物活性及其全合成方面的研究, 为此类生物碱用于抗肿瘤、免疫抑制的进一步研究提供了很有意义的参考。

“天然黄酮碳苷类化合物的研究进展”, 黄酮碳苷广泛分布于植物界, 具有清除氧自由基和抗氧化作用、保护大鼠肝损伤作用、抗艾滋病毒和降血糖作用。此文对该类化合物的进一步研究和应用具有参考意义。

“含植物环肽的药典中药的物质基础及其活性研究进展”对 2010 年版《中华人民共和国药典》一部收录的 24 种含植物环肽中药的化学成分和生物活性及药用研究的主要进展进行了详细介绍。植物环肽具有多方面生物活性, 包括抗肿瘤、抗 HIV、抗菌、抗疟、镇静、抑制血小板聚集等。此文为含环肽类中药的研究和开发提供了参考。

本卷书主要作者均为在各自研究领域有所成就的专家、教授, 对他们百忙中为本卷撰稿表示衷心的感谢。

彭司勋

中国药科大学
2013 年 12 月

目 录

1	药效团模型方法与应用研究进展	1
	The Research Progress of Pharmacophore Modeling Methods and Applications	
	汪文静 杨胜勇	
2	高效合成技术在药物化学中的应用	15
	Application of Efficient Synthetic Techniques in Drug Research	
	周宇 张磊 栗增柳 红	
3	干扰 RNA (siRNA) 药物研究进展	37
	Development of siRNAs as Drug Candidates	
	陈玥 黄野 杨振军	
4	共价抑制剂在肿瘤化学生物学中的应用	59
	The Application of Covalent Inhibitors in Cancer Chemical Biology	
	赵亚雪 周虎臣	
5	酪氨酸激酶 c-Met 小分子抑制剂的研究进展	76
	Advances in the Development of Small Molecules Targeting Tyrosine Kinase c-Met	
	王元相 宋子兰 张 翱	
6	小分子 p53-MDM2 结合抑制剂的研究进展	105
	Progress of Small Molecule Inhibitors of the p53-MDM2 Interaction	
	王味思 胡纯琦 胡永洲	
7	Nrf2 信号转导通路介导的化学预防: 抗氧化应激结合抗炎	130
	Chemoprevention through Nrf2 Singaling Pathway: Combination Anti-oxidation and Anti-inflammation	
	叶 鸣 姜正羽 钱 晨 王亚菁 胡容	
8	HIV 进入抑制剂的研究进展	152
	Current Status and Trends of HIV Enter Inhibitors	
	贾启燕 蔡利锋 史卫国 刘克良	
9	靶向 HIV-1 整合酶的抗逆转录病毒药物研究	171
	The Antiretroviral Drug Development Targeting HIV-1 Integrase	
	张凤华 林建平 龙亚秋	

10	寻找抗高血压药物作用新靶点：辣椒素受体·····	189
	Seeking Novel Targets for Antihypertensive Drugs: Capsaicin Receptor	
	彭 军 胡高云 李元建	
11	神经保护剂研究进展·····	201
	The Research Progress of Neuroprotectants	
	钱晓丹 罗春霞 朱东亚	
12	AI-2 型细菌群体感应抑制剂的研究进展·····	213
	Progress on AI-2 Bacterial Quorum Sensing Inhibitors	
	朱 鹏 姜天宇 李敏勇	
13	灵菌红素生物碱及 roseophilin 的研究进展 ·····	232
	Progress in the Research of Prodigiosin Alkaloids and Roseophilin	
	宋传君 常俊标	
14	天然黄酮碳苷类化合物的研究进展·····	255
	Progress on Natural Flavonoid C-glycosides	
	徐方方 张晓琦 叶文才	
15	含植物环肽的药典中药的物质基础及其活性研究进展·····	297
	Progresses on Chemistry and Bioactivity of Cyclopeptide-containing Traditional Medicines Cited in the Chinese Pharmacopoeia	
	谭宁华 赵思蒙	

1 药效团模型方法与应用研究进展

The Research Progress of Pharmacophore Modeling Methods and Applications

汪文静 杨胜勇

1.1 引言	1
1.2 药效团模型的构建	3
1.2.1 基于配体的药效团模型构建方法及最新进展	3
1.2.2 基于结构的药效团模型构建方法及最新进展	4
1.3 药效团模型方法的应用	7
1.3.1 基于药效团模型的虚拟筛选	7
1.3.2 基于药效团模型的从头设计	8
1.3.3 基于药效团模型的靶标预测	10
1.3.4 基于药效团模型的药代动力学性质和毒性预测	10
1.4 问题与展望	10
参考文献	11

1.1 引言

药效团 (pharmacophore) 是药物化学中最重要的基本概念之一^[1,2]。按照国际纯粹与应用化学联合会 (IUPAC) 对药效团的标准定义, 药效团是“一组必需的可以确保与特定的生物靶标有很好相互作用, 并能触发 (或阻断) 特定生物反应的立体和电子结构特征组成的系综”。而更通俗地讲, 药效团是指在生物活性分子中对活性起重要作用的“药效特征元素”及其空间排列形式。药效团模型中的药效特征元素不仅指某些特定的原子或原子团, 还包括一些抽象化的化学功能结构, 如氢键供体和受体、疏水中心、芳环中心以及正、负电荷中心等。

药效团概念的出现虽然较早, 但长期以来在药物研发中并没有得到很好的应用, 直到 20 世纪 80 年代末, 随着计算机科学、计算化学、生物化学、生物信息学等学科的发展, 计算机辅助药物分子设计进入了一个崭新的快速发展和应用阶段, 与此同时, 药效团模型方法及其应用也得到了迅速的发展, 这可以从近 20 年来在国际学术期刊上发表的与药效团相关的论文数量的增加情况上得到体现 (见图 1-1)。目前药效团方法已成为一种主要的药物分子设计方法, 在药物发现中获得了广泛的应用。已有大量应用药效团方法的成功实例, 例如, Singh 等人^[3]使用基于药效团模型的虚拟筛选方法, 获得了 12 个 $\alpha_4\beta_1$ 拮抗剂, 其中活性最高的化合物的 IC_{50} 达到了 1 nmol/L; Brust 等人^[4]使用药效团模型方法针对去甲肾上腺素转运体 (norepinephrine transporter) 开展了虚拟筛选, 获得多个高活性的化合物, 经过进一步优化最终得到了候选药物 Xen2174,

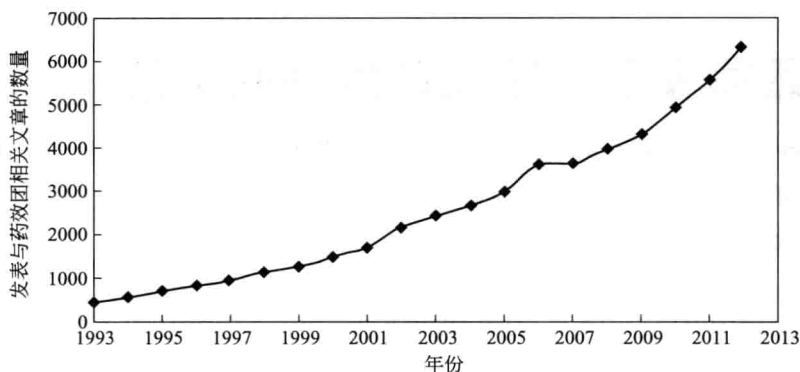


图 1-1 近 20 年来在国际学术期刊上发表的与药效团相关的论文数量统计
(统计源数据来自 <http://scholar.google.com>)

目前该候选药物已进入 II 期临床试验；Barreca 等人^[5]利用药效团模型方法对 HIV-1 逆转录酶抑制剂进行优化，获得了三个高活性的 HIV-1 逆转录酶抑制剂。更多的应用药效团模型方法的实例参见参考文献 [6]。

在药物分子设计中，药效团方法包括两个层面（见图 1-2），即药效团模型构建（又称药效团模型识别）与药效团模型的应用。药效团模型的构建方法又可分为基于配体小分子和基于结构（又称基于受体大分子）两大类。药效团方法的传统应用是虚拟筛选，近年来，其应用范围已拓展到了从头设计、靶标预测、药代动力学性质和毒性预测等领域。虽然药效团模型方法已取得了重要进展，并在药物研发中获得了广泛的应用，但仍然存在许多问题和挑战，导致基于药效团的药物设计方法准确度较低。为克服这些问题，近年来多个课题组开展了深入的研究，取得了良好的进展。本章将针对近年来药效团模型方法和应用研究方面的最新进展进行综述。

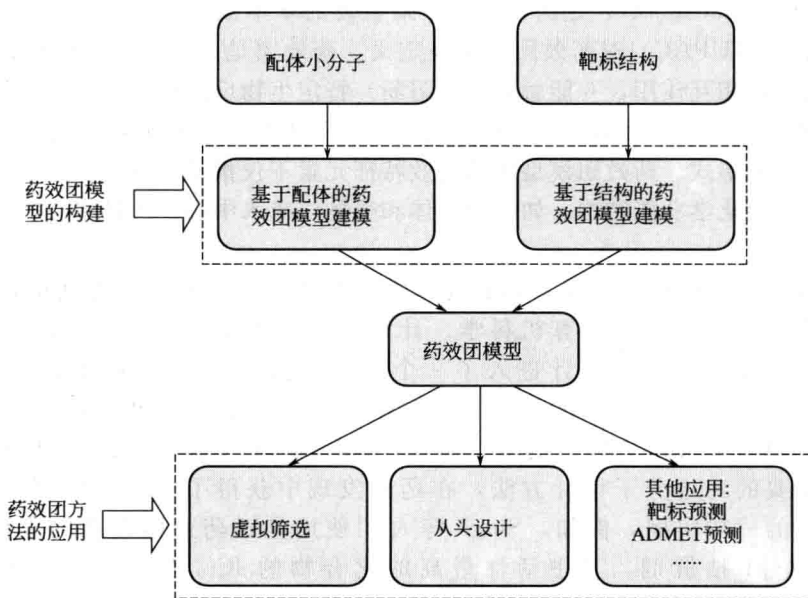


图 1-2 药效团方法的整体框架（模型构建与应用）

1.2 药效团模型的构建

1.2.1 基于配体的药效团模型构建方法及最新进展

1.2.1.1 基于配体的药效团模型构建方法概述

基于配体的药效团模型构建方法是应用最早的一种药效团模型构建方法。该方法从一系列已知活性的化合物的结构特征和活性数据出发,通过构象分析、分子叠合等方法,模拟并发现对化合物活性起重要作用的药效特征基团,得到它们之间的空间位置约束参数,从而建立药效团模型。基于配体的药效团模型可分为定性药效团模型和定量药效团模型。前者对配体数量要求不高,只需要几个活性分子,所得模型一般仅包含共有药效特征元素;后者对配体数量及结构多样性有一定要求,所得模型具有明确的定量构效关系,可预测未知化合物针对特定药物靶标的生物活性。构建药效团模型的基本步骤主要包括:①选择用于药效团模型构建的活性小分子化合物组成训练集(training set),并定义药效特征元素。一般要求所选择的活性化合物具有结构多样性。药效特征元素的定义视不同的药效团模型建模方法不同而有所差异,常见的药效特征元素包括氢键给体、氢键受体、疏水中心、芳环中心、正电荷中心、负电荷中心等。②构象分析。配体小分子一般有多种可能的构象,其中与受体结合的构象称为活性构象。鉴于一般情况下人们事先并不知道哪一个构象为活性构象,构象分析的目的需要产生所有可能的活性构象。许多软件例如 Catalyst^[7]等,在默认情况下,构象分析会保留和最低能量构象相差 20kcal[●]/mol 之内的所有构象。③分子叠合和药效团映射。得到药效团特征元素以及构象以后,就需要对这些构象进行叠合以发现对活性起重要作用的药效特征元素,并定义这些药效特征元素之间的空间位置约束参数,从而建立药效团模型。④药效团模型的验证。药效团模型可以通过测试集方法或其它方法进行验证,一个好的药效团模型一般应与活性较高的化合物具有较好的匹配,而与活性较低的化合物匹配较差。

目前,已有多种基于配体的药效团模型构建软件,包括商业化的软件如 HipHop^[8], HypoGen^[9], HypoRefine^[10], DISCO^[11], GASP^[12], GALAHAD^[13], PHASE^[14] 和 MOE^[15]等,以及一些非商业化的学术软件,如 DANTE^[16], GAMMA^[17]等。这些软件各自的特点与优势参见文献 [18]。

基于配体的药效团模型构建方法虽然发展较早,技术也相对成熟,但仍然存在许多问题和挑战。其中最重要的两个问题是配体分子的构象分析(conformational sampling)和分子叠合,关于构象分析和分子叠合的常见方法及进展参见文献 [19]。而另一个最常遇到的问题是训练集的选择,此问题看似简单,但常常令使用者困惑,即使对于经验丰富的使用者。研究表明,训练集中分子的种类、训练集的大小以及训练集中分子的多样性都对药效团模型有重要影响。即使使用相同软件相同算法,不同的训练集也可能产生完全不一样的药效团模型。例如,Hecker 等人^[20]、Toba 等人^[21]和 Vadivelan 等人^[22]分别独立地构建了三个细胞周期素依赖激酶 2 (cyclin-dependent kinase 2, CDK2) 抑制剂的药效团,他们使用了同一个药效团建模软件 Catalyst,但由于使用了不同的训练集,最后产生了完全不一样的药效团模型。Zou 等人^[23]最近的研究发现,这些不同训练集产生的药效团模型实际上均为全特征药

● 1kcal=4.1840kJ。

效团模型的一个子集，即它们均可以从一个侧面反映配体-受体相互作用。此研究结果提示，解决训练集问题的最简单有效的方法是通过将训练集分组，分别建立不同的药效团模型，所有这些药效团模型构成一个药效团模型组，在后续的应用研究中，使用该药效团模型组而非单一的某个药效团模型。

1.2.1.2 基于 GRID 的药效团模型方法——FLAPpharm

传统的基于配体的药效团模型构建方法主要是寻找共同的药效特征，这样产生的药效团模型往往只有少数几个药效基团。一方面，这些药效基团过于简单，所包含的信息太少；另一方面，由这些少数的几个药效基团组成的药效团模型也不足以准确反映配体-受体之间的复杂相互作用。为解决此问题，Cross 等人最近发展了一种基于 GRID 分子相互作用场 [GRID molecular interaction fields (MIFs)]^[24,25] 和 FLAP 方法^[26] 的药效团模型方法——FLAPpharm^[27]。FLAPpharm 首先将多个配体进行叠合，分别按照 MIFs、虚拟场 (pseudofields) 和原子中心点，提取共同相互作用特征，所有这些特征的组合形成一个公共药效团模型，该药效团模型实际为一个包含多重信息的药效团虚拟分子。测试结果表明，基于 FLAPpharm 产生的药效团模型的虚拟筛选，具有很好的筛选准确度^[27]。

1.2.2 基于结构的药效团模型构建方法及最新进展

1.2.2.1 基于结构的药效团模型构建方法概述

基于结构的药效团模型构建方法是根据靶标大分子的三维结构来构建药效团模型。该方法通过探测受体活性位点的关键基团的互补基团与其空间约束关系来产生大量的药效特征元素，然后选取其中最重要的药效特征元素组成药效团模型。基于结构的药效团模型构建方法的基本步骤如下：①受体分子的准备。例如去除受体中的水分子及其它杂原子等。②确定活性位点。一般将已知配体小分子占据的位点确定为活性位点。③药效特征的确定。一般将活性位点按一定的步长均匀划分产生格点，在每个格点上使用一个探针来评价格点上的分子场特征，例如氢键场、疏水场、静电场等。通过对这些格点的分子场特征进行聚类分析而确定药效特征。④选取其中最重要的药效特征元素组成药效团模型。

基于结构的药效团模型方法可以被进一步分为两类：基于受体-配体复合物与基于受体 (不含配体) 的方法。基于受体-配体复合物的方法能方便快速地确定受体的配体结合位点，并且还能确定受体大分子与配体之间的重要相互作用位点。目前已有多个基于受体-配体复合物的药效团模型构建软件，包括 LigandScout^[28]、Pocket v. 2^[29] 和 GBPM (GRID-based pharmacophore model)^[30] 等。在仅知道受体结构而无受体-配体复合物的情况下，则可采用基于受体的药效团模型构建方法。Discovery Studio 中的 SBP 模块就是一种典型的基于受体的药效团模型构建软件^[31]。SBP 从受体位点的特性直接得到相互作用位点图，并且将这些信息转化成 Catalyst 的药效基团，包括氢键受体、氢键供体和疏水相互作用等。这些药效基团可进行编辑、分类，以保留重要的信息形成可用于虚拟筛选的药效团模型。同时 SBP 中还可考虑受体活性位点处氨基酸残基的空间排布信息，可在药效团模型中加入排斥球。

基于结构的药效团模型构建方法常常会产生多个药效特征，太多的药效特征组成的药效团模型 (例如，大于 7 个)，一般不适合于实际的虚拟筛选应用。因此，常常选择有限数量的药效特征元素 (通常在 3~7 个之间) 来组成一个具有实用性的药效团模型。但如何选择药效特征元素 (包括种类和数量) 组成药效团模型仍然是一个尚待解决的问题。另外，基

于结构的药效团模型建模方法不涉及配体活性数据, 因此所产生的药效团模型常常不能很好地反映定量的结构活性关系。为了克服这些问题, 近年来多个课题分别发展了不同的方法。

1.2.2.2 高频特征药效团模型构建方法

Zou 等人最近发展了一种基于多复合物的高频特征药效团模型 (The most-frequent-feature pharmacophore model) 建模方法^[23]。以 CDK2 抑制剂的药效团模型构建为例, 在该方法中, 作者收集了 CDK2 的 124 个受体-配体复合物晶体结构, 对于每一个复合物分别在其活性口袋探测所有的化学特征, 然后将 124 个复合物进行叠合, 将相近位置的相同化学特征进行聚类, 形成了多个药效特征元素。最后选择了在所有 CDK2 受体-配体复合物中出现频率较高的 7 个药效特征元素组成一个高频特征药效团模型。经验证, 该模型可以很好地区分已知的 CDK2 抑制剂与非活性分子 (见图 1-3)。

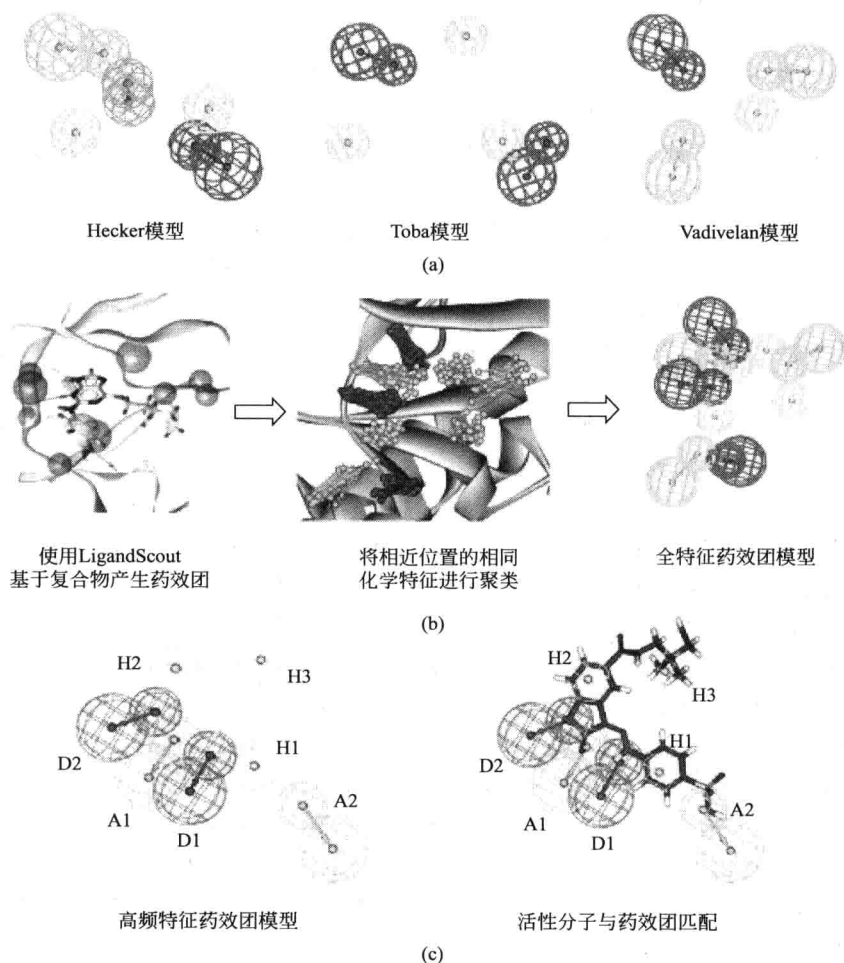


图 1-3 CDK2 抑制剂的药效团模型 (见文后彩图)

(a) 由 Hecker 等人、Toba 等人 and Vadivelan 等人构建的 CDK2 抑制剂的药效团模型; (b) 基于多个 CDK2 复合物构建的全特征药效团模型; (c) 左为高频特征药效团模型, 右为活性分子 ($IC_{50} = 6.8 \text{ nmol/L}$) 与高频特征药效团模型的匹配。药效基团的颜色: 绿色, 氢键受体; 品红, 氢键供体; 浅蓝, 疏水基团; 橙色, 芳香环

1.2.2.3 基于能量优化的药效团模型方法

近年发展起来的基于能量优化的药效团模型建模方法 (e-pharmacophore), 也可以有效地确定药效特征的数目和权重^[32,33]。其基本思想如下: 针对配体-受体复合物, 首先使用 Glide XP 对配体-受体相互作用能进行打分, 将这些能量打分函数值映射到配体分子的每个原子; 然后产生多个药效区域, 将包含在每个区域的原子所对应的能量值进行累加, 并将该能量累加值赋予对应的药效区域; 最后按照能量值的大小将这些药效区域进行排序, 选取能量值较高的区域作为药效特征, 这些药效特征及其空间位置形成了药效团模型。使用这种方法所建药效团模型, 称为基于能量优化的药效团模型 (e-pharmacophore)。对多个复合物晶体结构的测试结果表明, 基于所建的 e-pharmacophore 的虚拟筛选, 可以显著提高活性化合物的富集率。该方法已在一些例子中获得了成功的应用^[34~36]。

1.2.2.4 水合位点限制的药效团模型方法 (HSRP 方法)

Hu 等人最近发展了一种新的称为水合位点限制的药效团模型建模方法 (hydration site restricted pharmacophore modeling method, HSRP 方法)^[37]。该方法的基本思想如下: 第一步, 基于靶蛋白的全特征药效团模型的建立。将靶蛋白活性口袋划分为三维网格, 计算每个格点的势能, 然后通过聚类分析产生多个药效特征, 所有这些药效特征形成一个全特征药效团模型。第二步, 水合位点的确定。在溶剂水存在的条件下, 对靶蛋白进行动力学模拟。根据计算的分子动力学轨迹, 采用聚类分析确定靶蛋白活性位点内的水合位点。第三步, 将第一步产生的全特征药效团模型与第二步产生的水合位点图进行叠合, 从而产生水合位点限制的药效团模型。测试结果表明, 在虚拟筛选时, 虽然 HSRP 模型比起全特征药效团模型在富集率上并没有优势, 但筛选速度得到了很大的提高。

1.2.2.5 针对 A 类 G 蛋白偶联受体 (GPCR) 的药效团模型建模专用工具——Snooker

GPCR 是一类非常重要的药物靶标, 据统计 50% 以上的临床用药以及正在研发中的药物的靶标为 GPCR。由于 GPCR 是跨膜蛋白, 晶体结构解析十分困难, 目前仅有为数很少的几个 GPCR 结构被解析, 因此, 对这些大量结构未知的 GPCR, 要采用基于结构的药效团模型建模方法建立其配体的药效团模型困难较大。GPCR 共有 A、B、C、D、E、F 六类, 其中 A 类是研究最多的 GPCR。Sanders 等人发展了一个针对 A 类 GPCR 的药效团模型建模专用工具 Snooker^[38]。针对一个结构未知的 A 类 GPCR, 其药效团模型建模的基本过程简述如下: 首先选择结构已知的 GPCR 作为模板进行同源建模, 同时针对每个残基, 根据已知的“ α 螺旋特异性旋转异构体数据库 (α helix specific rotamer library)”中的对应残基的旋转异构体建立系统, 并记录每个旋转异构体的概率。然后根据 C α 和旋转异构体的侧链的位置定义整个蛋白和结合口袋区域。接着针对每个残基, 根据其在 GPCR 亚家族和整个 GPCR 家族中的保守情况进行打分, 此打分值可以反映该残基位置在与配体相互作用时的重要性。然后 Snooker 根据结合口袋里的重要残基的性质, 产生与其互补的药效基团中心位点 (在这一步中, 对每一个旋转异构体都产生相应的药效基团中心位点)。最后将药效基团中心位点密度 (与旋转异构体的概率相关) 高的位置定义为对应的药效特征, 并最终产生药效团模型。这是一个半自动化的基于生物信息学的药效团模型建模方法。为了验证该方法, 作者将其应用于 15 个 A 类 GPCR 不同亚家族成员。结果显示, 全部 15 个药效团模型都能很好地反映配体与受体之间的相互作用, 并且能有效地区别活性分子与非活性分子, 特别是通过筛选得到的分子具有靶点特异性。

1.3 药效团模型方法的应用

1.3.1 基于药效团模型的虚拟筛选

1.3.1.1 基于药效团模型的虚拟筛选概述

虚拟筛选是药效团模型的最主要的应用之一，它以药效团模型作为模板，通过对 3D 化合物库的搜索，寻找满足药效团模型要求的化合物分子。基于药效团的虚拟筛选与基于对接的虚拟筛选是目前应用最广泛的两种虚拟筛选方法。基于对接的虚拟筛选方法虽然应用广泛，但存在着许多至今尚无有效解决办法的问题，包括蛋白质的柔性、打分函数、溶剂水分子等。而基于药效团模型的虚拟筛选则避开了分子对接中存在的这些棘手问题。

基于药效团模型的虚拟筛选过程主要包含两个关键环节，同时也是两个主要的技术难点：一是分子构象分析；二是小分子与药效团模型的匹配。构象分析的应用主要包括两种类型，即前列举法和实时方法。前列举法是在虚拟筛选前，即对虚拟化合物库中的所有分子进行构象分析，生成多种构象并存放于数据库中，在实际筛选时直接从数据库中取出即可；而实时方法则是在虚拟筛选时实时进行小分子的构象分析。采用前列举法的优点是虚拟筛选速度较快，但需要大容量的存储空间；而实时方法则相反，对存储空间要求不高，但速度相对较慢。小分子与药效团模型的匹配，即检查一个用于查询的药效团模型是否与小分子构象匹配。一般采用基于图论 (graph theory) 的子结构搜索方法，常见的方法包括：Ullmann^[39]、回溯算法 (the backtracking algorithm)^[40] 和 GMA 算法^[41] 等。目前常用的药效团模型高通量筛选软件包括 Catalyst、LigandScout、Phase 和 MOE 等，关于这些软件的详细情况可参见文献 [42]。

基于药效团模型的虚拟筛选方法面临的最大问题是假阳性 (false positive) 和假阴性 (false negative)。导致假阳性率高的原因很多，其中一个主要原因是药效团模型本身的质量不高，不能正确反映配体-受体相互作用以及 QSAR 信息，对药效团模型进行综合有效的验证和优化是解决这一问题的重要手段。目前已经有多种验证方法，如交叉验证法和测试集法等^[43,44]。另一个导致基于药效团模型的虚拟筛选假阳性率高的原因是缺乏对受体信息的考虑。在基于药效团模型的虚拟筛选中往往只考虑对药效团模型的匹配，而没有考虑这些分子能否进入靶分子的活性口袋，以及这些分子与其靶分子相互作用时是否会发生碰撞等，这些都可能导致筛选的化合物为假阳性。为解决此问题，目前广泛采用的方法是在药效团模型中引入排斥球，以模拟靶分子的空间限制，此方法可在一定程度上降低假阳性率。

基于药效团模型的虚拟筛选的另一个问题是假阴性率较高。与前述一样，药效团模型本身的质量是导致假阴性的重要原因，这可以通过提高药效团模型的建模质量以及对药效团模型进行优化而得以改进。除此之外，由于被用于搜索的药效团模型通常是全特征药效团模型图的一个子图，因此采用单一模型筛选仅能得到满足该模型的化合物，而满足其它药效团模型的化合物则被漏掉，这可能是导致假阴性率高的另一个重要原因。采用基于多药效团模型的虚拟筛选方法是解决此类假阴性率高问题的一种有效方法。

1.3.1.2 基于药效团模型的快速筛选方法

相对于分子对接，基于药效团模型的虚拟筛选方法一般具有较快的筛选速度，但对于大型数据库，例如上千万个化合物组成的数据库，或者数据库包含大量分子柔性较大的化合物时，其筛选速度仍然较慢，这也是目前基于药效团的虚拟筛选面临的主要问题之一。为此，Koes 等人开发了一个基于药效团模型的快速筛选程序——Pharmer^[45]。Pharmer 的基本思

想如下。①将用于筛选的药效团模型转化为一种称为 Bloom 指纹 (Bloom fingerprints) 的指纹信息, Bloom 指纹的优点在于所需空间小且能表征药效团的空间信息。②将化合物库中的每个分子也转化为指纹, 每个分子的指纹信息包含以下两部分: 一是将小分子首先用药效团表征, 并据此产生的 Bloom 指纹信息; 二是将小分子本身的信息 (例如分子量、可旋转键等) 直接转化成的指纹信息。然后采用一种称为 KDR 树的方法对化合物数据库进行索引排序。Bloom 指纹以及 KDR 树的采用使得 Pharmer 可以对大型数据库进行快速筛选。测试结果表明, Pharmer 在筛选速度上远远超过现有的其它方法, 例如筛选 200 万个分子也仅需要不到 1min。基于 Pharmer 的优越性能, Koes 等人又开发了针对 ZINC 数据库的在线 Pharmer 筛选工具 ZINCPharmer (<http://zincpharmer.csb.pitt.edu>)^[46]。ZINCPharmer 采用交互界面, 允许使用者使用药效团模型方法对 ZINC 数据库进行筛选。使用者可以直接在 ZINCPharmer 上建立药效团并对其进行优化, 并能用产生的药效团模型直接进行搜索, 搜索得到的结果可在网上浏览也可下载。

1.3.1.3 多步分层虚拟筛选方法

基于药效团模型的虚拟筛选方法在筛选大型数据库时, 除速度慢外, 其筛选的命中率和富集率也较低。一种可行的办法是使用多步分层筛选方法 (multistep hierarchical virtual screening approach)^[47], 首先应用筛选速度最快的方法对数据库进行初步筛选, 针对筛选结果再应用筛选速度较慢的方法进行筛选。这种策略不仅提高了筛选速度, 同时多种筛选方法的应用, 有助于提高最终筛选结果的命中率和富集率。例如本课题组最近利用三种虚拟筛选方法, 即基于支持向量机的虚拟筛选 (SVM-based VS, SB-VS)、基于药效团的虚拟筛选 (pharmacophore-based VS, PB-VS) 以及基于对接的虚拟筛选 (docking-based VS, DB-VS), 采用了一种多步分层虚拟筛选策略, 针对 PubChem (18831686 个化合物), Specs (202408 个化合物), Enamine (980000 个化合物) 以及实验室内部数据库等大型数据库筛选 Pim-1 抑制剂^[47]。首先应用筛选速度最快的 SB-VS 对整个数据库进行筛选, 针对 SB-VS 的筛选结果, 再使用速度稍慢的 PB-VS 进行筛选, 针对 PB-VS 的筛选结果, 最后使用筛选速度最慢的 DB-VS 进行筛选。测试结果表明, 三种方法的组合 (SB/PB/DB-VS) 大大提高了筛选结果的命中率和富集率 (相对于单用 PB-VS, SB/PB/DB-VS 的命中率和富集率均提高了 100 倍以上), 就筛选速度而言, 筛选 2000 多万化合物也仅用了 19h (见表 1-1), 远远低于单用 DB-VS 的 2240.5h。在本项研究中, 选择了 47 个分子进行活性测验, 其中 15 个化合物具有较好的活性, 且其中 4 个具有新型骨架结构。另外, 也将类似的多步分层筛选方法应用于 BTK^[48]、CK2^[49]、mGluR1^[50] 和 C5aR^[51] 等靶点的抑制剂的虚拟筛选, 均取得了良好的效果。

表 1-1 基于 SB-VS、PB-VS 和 DB-VS 的多步分层虚拟筛选方法 (SB/PB/DB-VS) 与单独使用这些方法的性能比较 [该项测试为针对 PubChem (18831686 个化合物), Specs (202408 个化合物), Enamine (980000 个化合物) 以及实验室内部数据库等大型数据库筛选 Pim-1 抑制剂^[47]]

方法	命中率/%	富集因子	耗时/h
SB-VS	11.90	69.19	0.017
PB-VS	0.40	2.30	6
DS-VS	0.67	3.89	2240.5
SB/PB/DB-VS	41.91	243.78	19

1.3.2 基于药效团模型的从头设计

从头设计是药效团模型的另一种重要应用。与虚拟筛选相比, 基于药效团模型的从头设

计方法可以产生符合药效团模型要求的具有全新结构的化合物。由 Tschinke 和 Cohen 开发的 NEWLEAD 程序^[52]，是第一个基于药效团的从头设计软件。但 NEWLEAD 只能处理包含具体原子或原子团的药效团模型，而不能处理包含抽象化学功能特征（如氢键受体、氢键供体以及疏水特征等）的药效团模型，这极大限制了该软件的使用范围。其次，NEWLEAD 也不能处理包含排斥球的药效团模型。此外，NEWLEAD 还面临着所有传统从头设计方法都面临的一个共同问题，即产生的化合物的可合成性问题。

为了克服 NEWLEAD 的缺点，笔者的课题组近年来发展了一种新的基于药效团模型的从头设计算法和软件——PhDD (Pharmacophore-based De novo Design)^[53]。给定一个药效团模型，PhDD 能自动产生完全符合该药效团模型要求的类药性分子。图 1-4 显示了 PhDD 的基本流程。PhDD 以输入的药效团模型为模板，针对药效团模型中的每个药效特征元素，分别从对应的分子片段库（这些分子片段库事先已建立）中取出分子片段放置于相应的药效特征位置，通过旋转调整其取向，然后使用连接子将这些独立的分子片段连接成一个完整的分子。然后 PhDD 对产生的这些分子进行评价，包括使用利平斯基五规则进行类药性评价、使用 fit 值（即化合物分子与药效团模型的匹配程度）预测其生物活性以及使用基于复杂度的方法进行化合物可合成性评价等。PhDD 具有如下特点：①使用的药效团模型的格式为 CHM 格式，应用广泛，可通过多种方法获得；②可处理包含多种抽象药效特征的药效团模型，包括氢键供体、氢键受体、正电中心、负电中心、芳香环、一般疏水、芳香性疏水和脂肪性疏水等；③可处理包含排斥球的药效团模型；④包含了数量较大的药效基团分子片段数据库以及连接子数据库；⑤可对所生成的分子进行类药性评价、生物活性预测以及可合成性评价，使得最终输出的化合物分子不仅具有类药性，也易于合成。

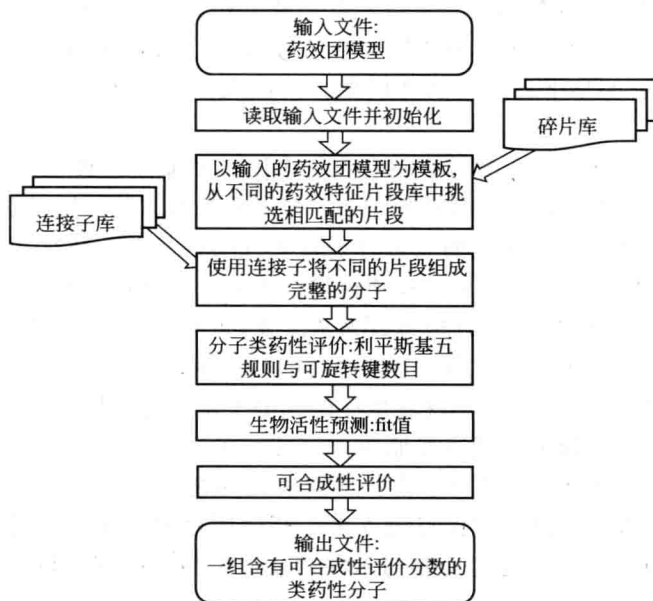


图 1-4 PhDD 的基本流程图

另外，Lippert 等人^[54]也发展了一种新的基于药效团模型的从头设计方法 Qsearch。Qsearch 与 PhDD 一样，需要一个药效团模型和一个片段库。与 PhDD 不同，Qsearch 产生新分子的基本过程如下：首先利用片段库中的片段，按照一定的规则随机产生大量分子，将产生的分子与药效团模型进行匹配；然后对与药效团模型基本匹配的分子进行循环优化，即

更换分子的某一片段并对其与药效团模型的匹配程度进行打分，依据打分结果决定是否接受这个修改，直到产生的分子能很好地匹配药效团模型为止。

1.3.3 基于药效团模型的靶标预测

靶标预测是近年来发展起来的药效团模型方法的一种新的应用。例如，上海药物研究所蒋华良研究组与华东理工大学药学院李洪林课题组合作，针对不同的靶标，构建了药效团数据库——PharmTargetDB，并建立了基于“反向药效团匹配方法”的靶标预测服务器 PharmMapper (<http://59.78.96.61/pharmmapper>)^[55]。该系统具有预测准确度高、速度快、全自动化、使用方便、可自由定制等优点，在国内外得到了广泛的应用。另外，最近笔者课题组将基于药效团模型的靶标预测方法（反向药效团匹配方法）与基于分子对接的靶标预测方法（反向对接方法）进行整合，发展了一个新型的基于融合概率的综合药物靶标预测系统^[56]。该系统可以同时发挥反向药效团匹配方法和反向对接方法在靶标预测方面的优势，克服单个方法由于自身局限导致的靶标预测准确度较低的问题。另外，该系统可给出每个靶标预测正确的概率，可帮助用户选择靶标进行实验验证。

1.3.4 基于药效团模型的药代动力学性质和毒性预测

药效团模型方法除应用于虚拟筛选和从头设计外，近年来也被用于药代动力学性质和毒性预测^[57]。统计表明，近一半的药物研发失败主要原因是其吸收（absorption）、分布（distribution）、代谢（metabolism）、排泄（excretion）等药代动力学性质较差，以及毒副作用（toxicity）较大（这些药代动力学性质和毒性简称 ADMET）。因此，将先导化合物的后期 ADMET 评价提前到药物的发现阶段，可以降低药物开发的后期风险。而计算机辅助药物分子设计方法是一种快速廉价的药物早期 ADMET 性质预测方法。目前已发展了多种基于统计分析方法的 ADMET 性质预测模型。而药效团方法近年来也被应用于 ADMET 预测。

药效团模型方法应用于 ADMET 预测存在许多困难和障碍，其一，药物的吸收、分布、代谢、排泄和毒性发生的过程和机制极其复杂，与多个基因、酶、代谢途径、信号通路等相关；其二，与 ADMET 性质相关的酶大多分子较大且结构复杂，药物与其相互作用模式多样，很难对这些酶的底物建立特定的药效团模型。近年来，已有多个课题组建立了与代谢相关的细胞色素 P450 同工酶底物的药效团模型，例如 CYP1A、2B6、2C9、2C19、2D6、2E1、3A4，以及其它与代谢相关的酶，例如 CYP19、CYP51、PXR 和体内清除相关的酶的药效团模型。但对同一种酶，不同的课题组所建模型千差万别，甚至同一个课题组使用不同的训练集所建模型也差别较大。其主要原因是这些酶的活性位点很大，并且具有较大的柔性，这些可导致小分子具有多个结合位点。为解决此问题，Güner 等人^[57]最近建议对每一种同工酶，针对不同的结合模式和位点建立多个药效团模型，从而形成一个完整而系统的 P450 相互作用药效团模型谱，通过药物与这些药效团模型的匹配分析，推演出可能的代谢途径。

1.4 问题与展望

近年来，药效团方法已成为计算机辅助药物设计的一种重要方法，在药物研发中获得了广泛的应用。虽然如此，在药效团模型建模方面仍然存在诸多问题和挑战，而基于药效团模