

计算机辅助药物设计

理论及应用

胡建平 等 编著



科学出版社

计算机辅助药物设计理论及应用

胡建平 等 编著

科学出版社

北 京

内 容 简 介

计算机辅助药物设计以化学和生物理论为基础,通过分子模拟、预测药物与靶点之间的分子识别,设计和优化先导化合物,在新药研发中发挥着不可或缺的作用。本书针对学生在计算机辅助药物设计学习过程中可能遇到的各种问题,介绍了数据库、定量构效关系、药效团、分子对接、虚拟筛选、同源模建、分子动力学模拟、自由能计算理论以及药物设计过程中常用软件和数值计算方法。本书覆盖全面,简单实用,加入了大量应用实例,穿插操作方法和技巧;每个章节自成体系,从易到难,层次分明。

本书不仅可以作为高等院校药学、计算化学、生物物理学及其他相关专业教材,也可以为科研院所教师、研究人员和研究生参考使用。

图书在版编目(CIP)数据

计算机辅助药物设计理论及应用/胡建平等编著. —北京:科学出版社, 2019.4

ISBN 978-7-03-058530-1

I. ①计… II. ①胡… III. ①药物—计算机辅助设计—高等学校—教材

IV. ①R914.2—39

中国版本图书馆CIP数据核字(2018)第186688号

责任编辑:郑述方 / 责任校对:彭 映

责任印制:罗 科 / 封面设计:墨创文化

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

成都锦瑞印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2019年4月第 一 版 开本:787×1092 1/16

2019年4月第一次印刷 印张:23 1/2 插页:2

字数:570 000 字

定价:98.00 元

(如有印装质量问题,我社负责调换)

《计算机辅助药物设计理论及应用》 编委会

编委会主任：胡建平

编委会副主任：苟小军 石虎兵 唐典勇 常 珊

编委(按姓氏汉语拼音排序)：

安海龙 高雪峰 龚新奇 郭锋彪 何 杨

焦 雄 李春华 卢本卓 罗亚飞 苏计国

田元新 万 华 许 磊 张大为

编委会主任简介



胡建平，1978年生，湖南永州人，博士，教授，药物化学硕士生导师，四川省学术和技术带头人后备人选，成都大学学术带头人。1999年毕业于吉林大学生命科学学院生物制药专业；2008年毕业于北京工业大学生命科学与生物工程学院生物医学工程专业，获得北京工业大学优秀博士学位论文奖(排名第3，获奖率<8%)，指导老师是王存新教授。2013年6月作为学术带头人调入成都大学药学与生物工程学院工作。2015年3月~2016年3月在美国匹兹堡大学药学院做访问学者。从2000年至今，一直从事基于化学和生物信息学的分子模拟研究。目前的研究方向为药物靶点-配体的相互作用以及计算机辅助药物设计，筛选和生物活性评价等。主讲课程有药物合成反应、药物设计学、制药工艺学、计算机辅助药物分子设计、生物信息学和计算机在化学中应用等。截至2019年3月，在国内外期刊及会议上发表学术论文105篇，SCI检索论文60篇，其中第一作者及通讯作者SCI检索论文42篇，引用500余次；主编教材2部，参编2部；主持完成2项国家级、2项省部级和10余项市厅级课题。目前是“药食同源植物资源开发四川省高校重点实验室”的主要成员。

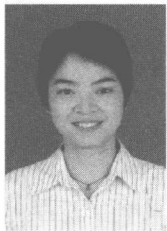
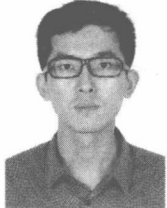
编委会副主任简介

	苟小军，教授，硕士生导师，现任成都大学四川抗菌素工业研究所所长，成都大学药食同源植物资源开发四川省高校重点实验室主任，担任四川省科技青年联合会理事，四川省分析测试协会环境与食品安全专委会副主任委员，全国生化检测标准化技术委员会委员，成都市科学技术顾问团特聘专家，四川省第九批学术和技术带头人后备人选，成都市特殊津贴专家。长期从事糖生物化学相关的研究。历年来，主持和参与国家和省部级项目近 20 项；出版专著 3 部；发表论文 130 余篇(SCI 收录 28 篇)；制定国家标准 1 项；获发明专利 18 项，申报 5 项；获四川省科学技术进步奖等省部级奖励 9 项。
	石虎兵，1977 年生，教授，博士生导师，2007 年毕业于清华大学，获得博士学位；后在美国加州大学洛杉矶分校从事肿瘤分子靶向治疗及耐药性机理研究。2015 年入选国家“青年千人”计划。2016 年成为国家“十三五”重点研发计划“精准医学研究”项目负责人。主要利用高通量交叉组学技术，从事以 MAPK 分子靶向和 PD-1-PD-L1 免疫检测点抑制的精准个体化治疗机理研究。研究期间共发表研究论文 20 篇，影响因子共计 358.2，被引用 4620 次，撰写专著 1 部，以第一作者或共同第一作者身份共发表论文 9 篇，包括：<i>Nature</i>, <i>Cell</i>, <i>Cancer Cell</i>, <i>Nature Communications</i>, <i>Cancer Discovery</i>, <i>Blood</i>, <i>Cancer Research</i>，影响因子共计 186.8，被引用 2970 次。
	唐典勇，1978 年生，博士，重庆文理学院创新靶向药物国际研究院教授，硕士生导师，创新靶向药物重庆市工程实验室副主任，四川省学术和技术带头人后备人选。主要从事药物设计和催化反应理论研究，主持国家自然科学基金面上项目 1 项，主持和主研完成 10 余项国家级科研项目。主持、参与了省级教改和质量工程项目 5 项，出版教材 2 部。近年来发表 SCI 研究论文 70 余篇，其中包括在 <i>Nano Letters</i>、<i>ACS Catalysis</i>、<i>Journal of Physical Chemistry Letters</i> 等国际著名科技期刊上发表论文 20 余篇。
	常珊，1982 年生，湖南岳阳人，博士，教授，硕士生导师，美国密苏里大学哥伦比亚分校博士后，现任江苏理工学院生物信息与医药工程研究所所长，江苏省生物信息学专业委员会委员。主要研究方向：生物信息学和智能计算方法。在国内外期刊及会议上发表学术论文 70 余篇，被 SCI 检索 40 余篇，引用超过 400 次，H 指数为 12，获得软件著作权 11 项，主持国家自然科学基金等各类项目十余项。获广东省首批优秀青年教师，广东省第八批“千百十”人才省级培养对象，江苏省“六大人才高峰”和江苏省“双创博士”。

编委(按姓氏汉语拼音排序)简介

	安海龙，博士，教授，博士生导师，河北工业大学元光学者特聘教授。河北省政府特殊津贴专家，河北省学科评议组成员，天津市五一劳动奖章获得者。获得河北省杰出青年科学基金资助，入选河北省高校“百名优秀人才创新支持计划”和河北省“三三三人才工程”三层次人选。现任河北工业大学国家大学科技园管理中心副主任兼科学技术研究院副院长，生物物理学研究所常务副所长。河北省物理学会、天津市物理学会理事，<i>Scientific Reports</i> 编委，<i>PLOS ONE</i>、《中国物理快报》《北京工业大学学报》审稿人。主要研究领域为离子通道结构与功能关系、离子通道与重大疾病以及离子通道靶向药物分子设计。发表学术论文 50 余篇，被 SCI 收录 40 余篇(总影响因子：126.41，被引用 200 余次)。主持、承担国家级、省部级项目 18 项。获得河北省优秀博士论文奖、河北省自然科学二等奖(第二名)、三等奖(第四名)各 1 项。
	高雪峰，副教授，1977 年 2 月生，1999 年 6 月毕业于吉林大学生物化学专业。2000 年 9 月~2005 年 6 月就读于吉林大学理论化学研究所，并获得博士学位。2006 年 3 月起在吉林大学生命科学学院工作，为高年级本科生及研究生讲授生物信息学理论。近年来主要从事酶的结构和功能以及代谢物组学研究，发表第一责任作者 SCI 收录文章 20 余篇，并主持翻译了 D. R. 韦斯特海德等的《生物信息学》先导版。
	龚新奇，博士，中国人民大学数学科学研究院 PI，Tenure-track 副教授，从事计算生物学和生物数学的教学科研工作。已经发表 SCI 学术论文 30 余篇，其中计算方法和生物应用各占一半，发表的杂志包括 <i>Nature</i>、<i>Science</i>、<i>PNAS</i>、<i>Cell Research</i> 和 <i>Proteins</i> 等顶级学术刊物，论文被 SCI 引用次数已超过 500 次。
	郭锋彪，电子科技大学教授，博导，教育部新世纪优秀人才。中国细胞生物学会功能基因组信息学与系统生物学专委会委员，中国运筹学会计算系统生物学专委会委员，中国生物物理学会生物信息学分会理事。<i>Scientific Reports</i> 编委，SCI 刊物 <i>Current Bioinformatics</i> 编委。至今已在 <i>Molecular Biology and Evolution</i>、<i>Nucleic Acids Research</i>、<i>Briefings in Bioinformatics</i>、<i>Bioinformatics</i> 等杂志发表 SCI 论文数十篇，被引用 1000 余次。2001 年至今一直围绕微生物基因组展开生物信息学的研究工作。

	何杨, 四川大学华西医院研究员。四川大学理学学士、中科院成都生物所硕士、德国奥斯纳布吕克大学生物有机化学博士、德国马克斯普朗克学会斯图加特智能体系研究所博士后。研究方向为核酸化学, 围绕核苷与核酸相关化合物的设计、合成、生物活性以及在纳米生物医学材料和分子示踪剂方面的应用。曾完成德国科学基金项目(平行链 DNA), 参与欧盟第七框架项目(抗病毒核苷类似物), 主持并参与多项国家自然科学基金项目与省市科研项目。申请国家发明专利 16 项, 已获得授权 9 项, PCT 国际专利 5 项, 美国专利 1 项, 总计发表 SCI 论文 40 余篇。
	焦雄, 博士, 教授, 2007 年博士毕业于北京工业大学生命科学与生物工程学院。目前主要从事细胞分子生物力学研究、生物大分子的计算机模拟、蛋白质结构功能关系研究、蛋白质氨基酸网络的计算分析等。主持一项国家自然科学基金面上项目, 一项博士后科学基金特别资助项目, 一项博士后基金面上项目, 一项山西省青年基金项目。近年来在国内外核心期刊以及学术会议发表论文 20 余篇。2010 年作为主要参与人获得山西省科学技术奖(自然科学类)二等奖, 2012 年当选山西省高等学校优秀青年学术带头人。
	李春华, 北京工业大学生命学院教授, 博士生导师。2003 年博士毕业, 获全国百篇优秀博士学位论文提名奖, 导师王存新教授。目前从事蛋白质-配体相互作用、生物大分子变构和药物设计方面的研究。曾在美国密歇根大学 Yang Zhang 实验室访学一年。已入选北京市优秀人才培养计划和北京市科技新星计划。主持国家自然科学基金项目 4 项和北京市自然科学基金项目 2 项等, 参加国家自然科学基金重大项目子课题、科技部国际科技合作交流项目(中意)。近年来, 在国内外核心学术期刊, 如 <i>Proteins</i>, <i>Physical Review E</i>, <i>Biophysical J</i>, <i>the Journal of Chemical Physics</i>, <i>Journal of Physical Chemistry</i> 等上发表学术论文 70 余篇。
	卢本卓, 中国科学院数学与系统科学研究院“百人计划”研究员。中国仿真学会集成微系统建模与仿真专业委员会副主任委员。2002 年毕业于中国科技大学, 获得分子生物与生物化学博士。2003~2008 年分别在美国加州大学圣地亚哥分校及霍华德休斯医学研究所做博士后和研究人员。从事计算化学/生物/器件/数学的交叉研究。在生物分子静电和电扩散过程模型、计算方法和软件方面取得了系统的、有国际影响的成果。担任几个国际国内期刊编委/编委顾问。
	罗亚飞, 理学硕士, 四川乐山人。2013 年毕业于乐山师范学院化学学院, 同年考取西南大学物理化学专业硕士研究生。2016 年 7 月进入重庆文理学院创新靶向药物国际研究院工作至今。主要从事 OLED 磷光材料分子设计及光失活机制、表面催化反应、计算机辅助设计方面的理论研究。目前发表 SCI 论文 15 余篇。
	苏计国, 博士, 教授, 2000 年本科毕业于兰州大学金属物理专业, 2011 年获北京工业大学生物医学工程专业博士学位。目前在燕山大学理学院应用物理系任教, 主要从事蛋白质折叠和结构-功能关系研究, 基于物理学理论, 发表了多个有效的理论模型和方法, 成功用于蛋白质折叠机理、蛋白质变构效应以及功能位点识别研究。发表 SCI 论文 40 余篇, 主持国家自然科学基金 2 项, 荣获首批“河北省青年拔尖人才”称号。

	田元新，女，南方医科大学药学院副教授，主要从事计算机辅助药物设计的教学科研。博士毕业于中山大学分析化学专业。2014 年获国家留学基金委资助，赴新加坡国立大学药学系访学一年。主持和参与多项各级课题，在 <i>Journal of Chemical Information and Modeling</i>、<i>Journal of Physical Chemistry</i>、<i>European Journal of Medicinal Chemistry</i> 等杂志上发表论文 10 余篇，授权专利多项。主持药物设计学教学改革项目，获南方医科大学 2017 年度教学成果二等奖。
	万华，女，1978 年生，副教授，硕士生导师。2003 年 6 月于华南师范大学计算机专业获得硕士学位，2003 年 7 月加入华南农业大学数学与信息学院。科研方面目前主要从事分子模拟和机器学习等方面研究，在国内外期刊发表 SCI 论文十余篇。主持国家自然科学基金青年基金 1 项、广东省科技计划项目 1 项、国家自然科学基金委员会—广东省政府联合基金(第二期)超级计算科学应用研究专项资助 1 项、国家星火计划引导项目 1 项。2013 年被列为“华南农业大学青年骨干教师”。教学方面获得第六届华南农业大学校级青年教师教学观摩比赛二等奖，主持华南农业大学校级教改项目 1 项，院级教改项目 1 项。
	许磊，山东临朐人，博士，江苏理工学院/生物信息与医药工程研究所副教授，硕士研究生导师。2013 年在苏州大学获得博士学位，随后在浙江大学从事两年博士后研究，2016 年获得江苏省高校自然科学成果三等奖 1 项(医药类，1/3)，2017 年入选江苏省高校“青蓝工程”优秀青年骨干教师培养对象。2012 年至今在 <i>Advanced Drug Delivery Reviews</i>、<i>Journal of Medicinal Chemistry</i>、<i>Drug Discovery Today</i>、<i>Journal of Chemical Information and Modeling</i> 等国外著名期刊发表 SCI 论文 20 余篇，引用超过 500 次。长期从事计算机辅助药物分子设计方法学、生物信息学和应用研究。
	张大为，1985 年生，中国医学科学院北京协和医学院博士。目前工作于江苏理工学院生物信息与医药工程研究所，硕士研究生导师，研究方向：基于蛋白-蛋白相互作用的抗艾滋病小分子抑制剂的发现；老药新用在抗病毒领域的研究；基于噬菌体展示库的多肽或抗体药物发现。已在 <i>Scientific Reports</i>、<i>Analytical Biochemistry</i>、《药学报》等国内外杂志上发表学术论文 20 余篇；近 5 年承担国家自然科学基金 1 项，参与 3 项；主持江苏省自然科学基金 1 项；获得江苏省“双创博士”(世界名校类〈创新〉)一项。

前 言

随着人类基因组计划的完成,蛋白组学的迅猛发展,以及大量与人类疾病相关基因的发现,药物作用的靶标分子急剧增加。在计算机技术的推动下,计算机辅助药物设计的方法孕育而生,并在近几年取得了巨大的进展。计算机辅助药物设计是生命科学与计算机、药学、化学、物理、数学等全方位交叉的学科。主要研究内容包括研究靶点与药物的分子识别,靶点及药物的结构功能关系,先导化合物的发现及优化等。计算机辅助药物设计技术的飞速发展生物医药业开辟了广阔的前景,极大地推动了新药研发的速度。因此,世界各国都把计算机辅助药物设计确定为未来科技发展的关键技术和开发新药的重要手段。

目前,有关计算机辅助药物设计的书籍较少,内容的编排不甚合理。各个高校的计算机辅助药物设计课程的内容侧重点不同,有的立足于基础理论,但晦涩难懂,初学者难以接受;有的侧重于实践操作,但没有深入地阐明基本原理;有的虽然理论与操作兼顾,但表述得不够详细。总之,课程的覆盖面、侧重点不统一,且缺少细节的阐述。计算机辅助药物设计教学必须与科研实践相结合。

本书主编从2000年硕士研究生开始涉足分子模拟和计算机辅助药物设计领域,针对国家重大需求与国际学术发展趋势,结合承担的一些国家级、省部级项目,获得了一系列研究成果。经与科学出版社协商,出版一本关于计算机辅助药物设计方面的技术指导性专著,希望对从事计算机化学、药物研发科研工作者、药物化学和药学相关专业低年级研究生有所帮助。实际上,在编写本书初稿后,我们用PPT、讲义,对药学专业研究生进行了试用,取得了良好的教学和科研效果。

本书按照计算机辅助药物设计科学所需理论以及技术知识顺序而编排,总共分为四个部分,十个章节。第一部分是计算机辅助药物设计的知识背景和交叉学科基础,包括第一、二章,主要围绕药物研发、生命科学及药学基本理论、Linux操作系统展开,内容涉及生物化学、药物化学、计算机科学等领域。该部分的主要负责人是苟小军,撰写人包括苟小军、石虎兵、郭锋彪和罗亚飞。第二部分给出了在药物设计研究中的常用数据库计算和软件,即第三章内容,详细介绍了药物小分子、靶点、生物学通路、毒理性质等数据库的使用流程,并详细讲述了ChemDraw、PyMol、EndNote等软件的操作方法。该部分主要负责人是唐典勇,撰写人包括唐典勇和高雪峰。第三部分是科学工作者必备的计算机辅助药物设计方法和内容,分布在本书第四、五、六、七、八、九章。知识点覆盖了计算化学、生物物理学、物理化学、药物化学和数学等学科,详细介绍了计算机辅助药物设计中的基本原理和软件模块,包括定量构效关系、药效团、分子对接、同源建模、分子动力学模拟、结合自由能、粗粒化模型、自由能曲面、聚类分析等。该部分负责人有胡建平和石虎兵,参与编写人员有:胡建平、石虎兵、安海龙、何杨、龚新奇、李春华、苏计国、

焦雄、张大为、许磊、田元新、万华等。第四部分是药物设计常用的数值分析，涉及回归分析、ROC 曲线、贝叶斯分类等方法，主要是第十章内容。该部分负责人是常珊，撰写人有：常珊、卢本卓。

总之，本书内容编排新颖、语言简单易懂，每个章节彼此有一定联系，但也可独立学习。并且在介绍基本原理之后加入了已经发表在 SCI 杂志上的科研应用实例，具有严谨、实用、可操作和指导性等特点。本书由近二十位编者共同完成。本书的编者均是科研第一线从事计算机辅助药物设计或计算化学的相关研究人员，对于他们在百忙之中完成本书的撰写并促成本书的出版表示深切的感谢。本书的应用实例得到了课题组老师梁立、刘嵬、何钢、甘亚和研究生杜文义、左柯、孙歆、严潇、吴志祥、谢涛、段怀川、罗青的帮助，在本书的编写过程中他们共同参与查阅文献和相关书籍，在此谨对他们的付出一并表示衷心的感谢。

本书的出版得到了成都大学 2016 年教材建设项目、成都大学 2017 年研究生教育教学改革项目(cdjgy2017036)、国家“十三五”重点研发计划“精准医学研究”项目(No. 2016YFC0906000)、国家自然科学基金(No. 11247018, 31600591)、四川省科技厅国际合作项目(2016HH0012)、四川省科技厅项目(2015JY0117)、四川省教育厅重点项目(17ZA0194, 17ZA0092)、四川省中医药管理局中医药科研专项(2018KF006)、成都市科技局项目(2016-XT00-00023-GX)、乐山市科技局项目(14GZD022)、四川省抗生素研究和再评价重点实验室项目(ARRLKF17-11)、四川省天然产物化学小分子催化重点实验室项目(No. TRCWYXFZCH 2016015, TRCWYXFZCH2016014)、NSFC-广东联合基金(第二期)超级计算科学应用研究专项、江苏省六大人才高峰资助项目(2016-XYDXXJS-020)、江苏省产学研前瞻资助项目(BY2016030-06)等的大力支持，在此表示感谢。

随着生物技术和计算机技术的快速发展，有关新药研发的信息层出不穷，各类计算机辅助药物设计方面的文献报道推动着本课程的发展，这也给该门课程的教学和研究提出了更高的要求。由于我们的水平有限，且时间紧迫，本书存在一些纰漏在所难免，请广大读者见谅。如果您有什么宝贵的意见或建议，请您发邮件至 hjpcdu@163.com。

胡建平

2018 年 2 月于成都大学

目 录

第一章	绪 论	(1)
第一节	新药开发的基本途径和方法	(1)
第二节	计算机辅助药物设计的常用研究方法	(10)
第三节	计算机辅助药物设计的常用软件	(14)
第四节	近年来 CADD 在药物设计中的应用实例	(19)
	参考文献	(21)
第二章	计算机辅助药物设计的相关基础	(23)
第一节	计算机辅助药物设计的生命科学基础	(23)
第二节	计算机辅助药物设计的药学基础	(40)
第三节	计算机辅助药物设计的 Linux 基础	(48)
	参考文献	(62)
第三章	常用数据库和基础软件的使用	(64)
第一节	药物设计常用数据库	(64)
第二节	药物和靶点的三维结构显示及绘制软件	(79)
第三节	EndNote 文献管理	(85)
	参考文献	(91)
第四章	基于配体的药物分子设计	(92)
第一节	定量构效关系(QSAR)	(92)
第二节	药效团模型	(119)
	参考文献	(125)
第五章	基于受体的药物分子设计	(129)
第一节	分子对接的原理和分类	(129)
第二节	分子对接的搜索算法	(130)
第三节	分子对接的打分函数	(134)
第四节	分子对接的操作流程	(138)
第五节	分子对接实例	(144)
第六节	虚拟筛选实例	(147)
	参考文献	(153)

第六章	蛋白质结构预测及全新药物设计	(158)
第一节	同源建模	(158)
第二节	同源模建实例	(171)
第三节	蛋白质结合位点预测	(178)
第四节	全新药物设计方法	(182)
参考文献		(186)
第七章	分子动力学模拟	(190)
第一节	分子动力学模拟的原理	(190)
第二节	恒温恒压分子动力学模拟与溶剂模型	(196)
第三节	分子动力学模拟的主要步骤	(201)
第四节	分子动力学模拟应用实例	(210)
参考文献		(231)
第八章	结合自由能的计算	(234)
第一节	生物大分子体系结合自由能研究的背景及意义	(234)
第二节	结合自由能计算方法	(238)
第三节	基于 MM/PBSA 自由能算法的操作实例	(241)
第四节	自由能计算实例	(247)
参考文献		(258)
第九章	药物靶点的构象分析	(260)
第一节	药物靶点基础	(260)
第二节	粗粒化模型	(262)
第三节	自由能曲面	(275)
第四节	聚类分析	(279)
第五节	自由能曲面和聚类分析实例	(282)
第六节	ProDy 使用	(286)
参考文献		(291)
第十章	药物设计常用的数值分析	(295)
第一节	相关和回归分析	(295)
第二节	ROC 曲线	(305)
第三节	基于贝叶斯分类的类药性预测	(313)
第四节	Matlab 在药物设计中的应用	(322)
参考文献		(337)
附 录	常用 Linux 计算软件的安装	(339)
中英文对照索引		(347)
附 图		(363)

第一章 绪 论

药物设计是新药研究的中心环节,是人工预建可与机体重要功能分子(蛋白质、核酸、酶、离子通道等)发生作用的化学物质的过程。具体来说,药物分子设计是依据生物化学、酶学、分子生物学、遗传学等生命科学的研究成果,针对这些基础研究中所揭示的药物作用靶位,再参考其内源性配体或天然底物的化学结构特征设计合理的药物分子^[1]。近年来,生命科学和计算机科学的进展,使药物分子设计趋于定向化和合理化^[2]。

基于配体与受体间相互作用原理的不断完善,出现了开发新药的新方法——计算机辅助药物设计(computer aided drug design, CADD)^[3],该方法以计算机为工具,利用有关药物及其生物大分子靶标的结构知识,通过理论模拟、计算和预测,来指导和辅助新型药物分子的设计和发现^[4]。随着计算机化学、分子图形学、X 射线晶体衍射、多维核磁共振以及各种分子模拟技术的发展,计算机辅助药物设计已渐渐成为药物化学家研究的热点。

CADD 是一个多学科互相渗透的新兴研究领域。在 CADD 过程中,首先利用晶体学确定受体部位以及药物和靶点之间的相互关系,并利用计算机图形学和分子模拟技术来分析这些相互关系的全过程,设计出符合要求的新化合物分子,最后利用化学或生物合成的方法得到这些活性化合物。

第一节 新药开发的基本途径和方法

一、新药研发到上市的主要流程

药物研发从无到有再到最后上市的流程主要包括四个阶段:临床前研究、I~III 期临床研究、药物批准上市、IV 期临床研究。在临床前研究中,首先需要确认药物的作用靶点,这是所有工作的开始;随后进行化合物的合成,包括新化合物合成和现有化合物的改造和优化,不是所有合成出来的化合物都能有理想的活性,在这个阶段需要通过生物实验手段筛选出初步有活性的先导化合物用作备选;继续进行先导化合物的改造和优化,得到活性更高的化合物;评估药物的药理活性、安全性与毒性、药物的吸收、分布、代谢和排泄情况;最后是制剂开发设计。临床前研究一般耗时 4~6 年。

临床研究包括临床 I 期研究,其间将新药第一次用于健康人体以研究新药性质的试验,得到药代动力学信息;临床 II 期给药于少数病人志愿者,然后重新评价药物的药代动力学和排泄情况,对新药的有效性和安全性作出初步评价;临床 III 期中将试验药物用于更大范围的病人志愿者身上,进行扩大的多中心临床试验,进一步评价药物的有效性和耐受性(或安全性)。临床研究耗时一般为 6~7 年。完成临床 I、II、III 期研究后,药物被批

准上市, 审批手续大约耗时 0.5 年。药物上市之后, 还要长期进行药物上市后监测(即临床Ⅳ期研究), 主要关注药物在大范围人群应用后的疗效和不良反应监测。从图 1-1 可以看出, 从新药开发到上市, 历时 12 年左右, 其中有一个较为重要的阶段: 先导化合物的发现和优化。而计算机辅助药物设计的主要任务就是提供一种先导化合物发现与优化的策略, 能有效缩短获取候选化合物的时间(节约 2~3 年), 降低药物研发成本。

综上, 创新药物研发具有周期长、费用高的特征。一般而言, 从疾病靶标(target)的确定、先导化合物的发现、临床前药理学、药物代谢动力学(pharmacokinetics)及安全性评价研究到药物成功上市, 一般需要花费 10~15 年。一般来说, 研发一种新药的平均费用为 13 亿美元^[5]。对于药物企业而言, 提高研发效率、降低研发成本、缩短新药发现和早期开发时间是保持市场竞争实力的关键^[6]。

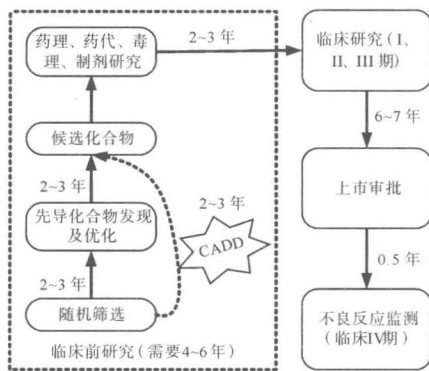


图 1-1 新药开发的基本流程

二、先导化合物的发现

先导化合物又称原形物, 是通过各种途径或方法得到的具有某种生物活性的化学实体。先导化合物未必是可使用的优良药物, 可能由于药效不强、特异性不高或毒性较大等缺点, 不能直接药用, 但作为新的结构类型, 可对其进行进一步的结构修饰和改造, 即先导化合物的优化。先导化合物的发现是创新药物研究的关键环节之一。先导化合物通常是通过合理的药物设计、组合化学、高通量筛选、虚拟筛选等方法来发现, 或是药物化学专家将来源于天然产物或微生物代谢物中的化学成分提取出来, 应用各种动物模型进行筛选, 从中发现的新的功能性化合物。

先导化合物的主要来源有几个途径, 第一个途径为天然产物: 从动物、植物、微生物以及海洋生物体内分离出具有生物活性的物质, 是先导物甚至是药物的主要来源。青蒿素是我国学者屠呦呦在 1971 年自黄花蒿中分离出的倍半萜类化合物, 具有强效抗疟作用。经醚化得蒿甲醚, 毒性比青蒿素低, 对疟原虫有较强的杀灭作用。二氢青蒿素进行酯化得青蒿素琥珀酸酯, 可制成粉针静注, 适用于抢救脑疟和危重昏迷的疟疾病人。基于此, 2015 年 10 月, 屠呦呦获得诺贝尔生理学或医学奖。图 1-2 给出了青蒿素系列化合物的结构。

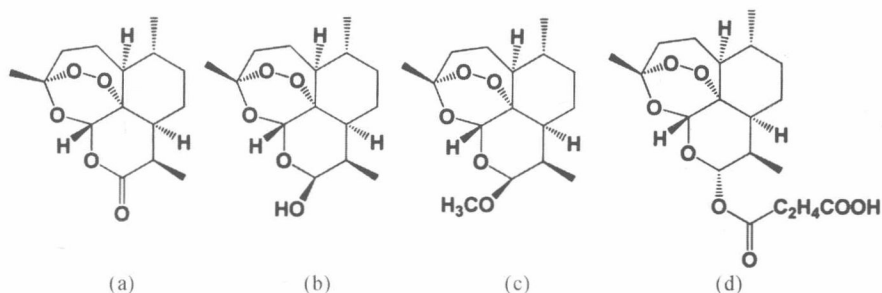


图 1-2 青蒿素系列化合物结构

(a)青蒿素；(b)二氢青蒿素；(c)蒿甲醚；(d)青蒿素琥珀酸酯

另外，紫杉醇别名泰素、紫素、特素，化学名称 $5\beta, 20$ -环氧- $1, 2\alpha, 4, 7\beta, 10\beta, 13\alpha$ -六羟基紫杉烷-11-烯-9-酮-4, 10-二乙酸酯-2-苯甲酸酯-13[($2'R, 3'S$)- N -苯甲酰-3-苯基异丝氨酸酯]，分子量 853.92，分子式 $C_{47}H_{51}NO_{14}$ 。紫杉醇能促进微管聚合和稳定已聚合的微管，能使细胞分裂停止于有丝分裂期，阻断细胞的正常分裂。紫杉醇主要适用于卵巢癌和乳腺癌，对肺癌、大肠癌、黑色素瘤、头颈部癌、淋巴瘤、脑瘤也有一定疗效。紫杉醇最初就是从国家一级保护植物紫杉(又叫红豆杉)的皮、根、枝叶中提炼出来的(结构式详见图 1-3)。

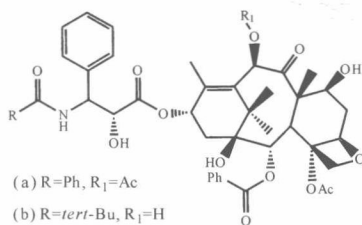


图 1-3 紫杉醇系列化合物结构

(a)紫杉醇；(b)多西紫杉醇

第二个途径是以现有药物作为先导物，比如异丙嗪、氯丙嗪；可以由药物副作用发现先导化合物，再比如治疗糖尿病的药物磺胺异丙噻二唑、甲苯磺丁脲等；可以通过药物代谢研究得到先导物，比如保泰松和羟基保泰松。还可以以现有突破性药物作先导，比如基于西咪替丁研发出“me too”药物，雷尼替丁和法莫替丁等。相关分子结构式见图 1-4。

第三个途径是用活性内源性物质作先导化合物，比如抗肿瘤药氟尿嘧啶，其内源性活性物质为尿嘧啶；以及基于炎症介质 5-羟色胺为先导化合物研发出抗炎药吲哚美辛等。相关分子结构式见图 1-5。

第四个途径为利用组合化学和高通量筛选(high throughput screening, HTS)得到先导化合物。具体来说，由于组合化学技术的发展，化合物的合成速度显著提高，能够更多更快地发现先导化合物或功能性化合物，对于创新药物研究是一项挑战，也是缩短新药发现时间的关键。高通量筛选技术采用自动化的操作系统，可以进行大规模的化合物筛选，但是这种方法筛选设备复杂，需要培养大量的靶酶或靶细胞，阳性率低，并且需要大量资金支持，因此仅仅采用高通量筛选的方法寻找先导化合物成本高，效率低。随着计算机技术的更新以及大数据技术的发展，应用虚拟筛选策略发现先导化合物逐渐成为主流。这种策略通过各种算法对大量化合物库进行搜索来获得有功能的化合物分子，其中应用分子对

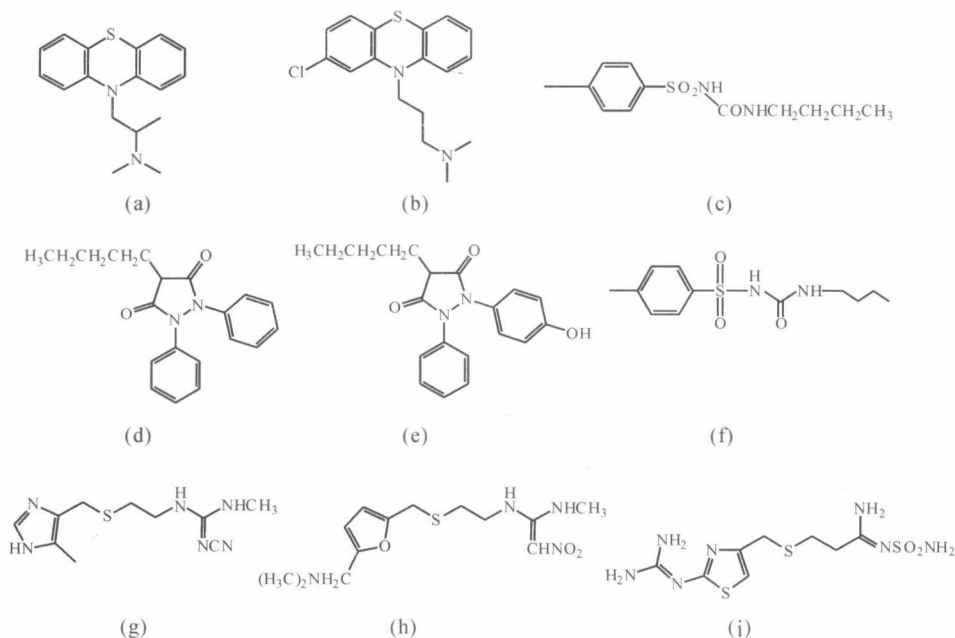


图 1-4 基于药物代谢得到的先导化合物相关分子

(a) 异丙嗪; (b) 氯丙嗪; (c) 磺胺异丙噻二唑; (d) 保泰松; (e) 羟基保泰松; (f) 甲苯磺丁脲;
(g) 西咪替丁; (h) 雷尼替丁; (i) 法莫替丁

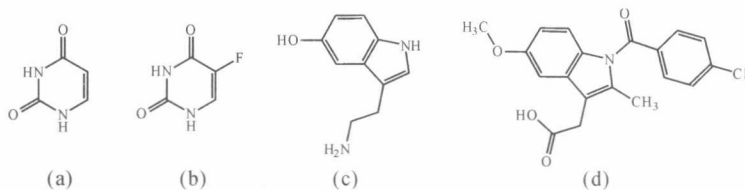


图 1-5 基于活性内源性物质得到先导化合物

(a) 尿嘧啶; (b) 氟尿嘧啶; (c) 5-羟色胺; (d) 褪黑素

接方法进行药物或功能化合物发现是一项有效的筛选技术, 该技术通过计算的方法将靶蛋白和小分子化合物进行一对一的对接, 从大量的化合物库中筛选出与靶蛋白有作用的小分子, 从而发现先导化合物。与高通量筛选相比, 虚拟筛选方法可以富集活性化合物, 降低筛选成本, 提高药物筛选的可行性, 因此应用虚拟筛选技术进行药物发现已成为新药发现的重要方法。

一直以来我们国家发展新药的口号是“以仿为主, 仿创结合”。真正由我国自己研制的全新结构药物寥寥无几。但随着我国药品专利法的实施和进入 WTO, 必须尽快完成由仿制向创新的转轨。面临创新药物研究方面人才缺乏、资金不足、技术理论储备不够的困境, 药学工作者肩上的担子很重, 我们也正面临着严峻的挑战。

三、先导化合物的优化

先导化合物结构优化是新药研发的关键环节。由于先导化合物只提供一种具有特定药理作用的新结构类型, 往往由于在药效学、药代动力学等方面的缺点而不能直接用于临床。因此, 需要对先导化合物进行进一步化学结构改造或修饰, 以期优化上述特性。生物