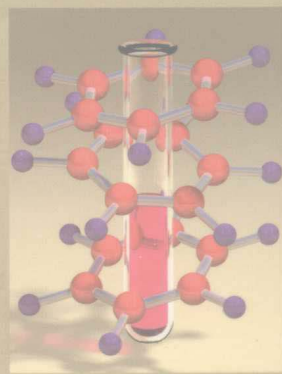
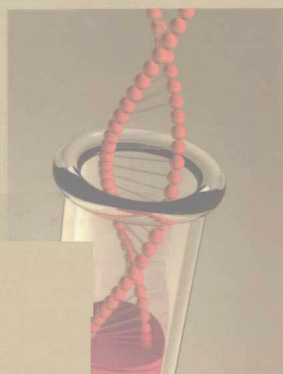
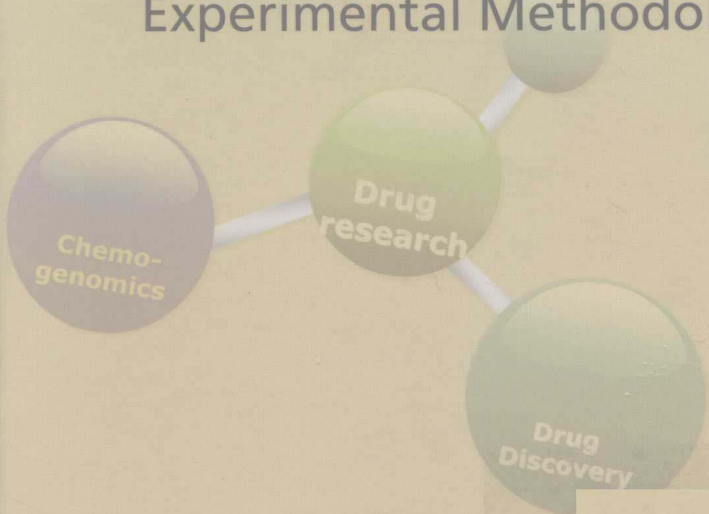


药物化学实验方法学

Experimental Methodology of Medicinal Chemistry

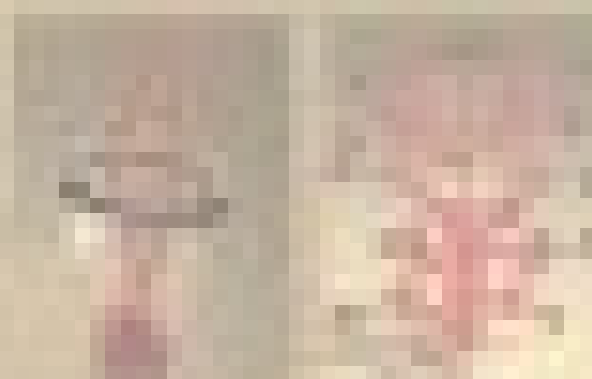
主编 徐文方



药物化学实验方法学

Experimental Methods in Medicinal Chemistry

主编 张永健



中国医药出版社

药物化学实验 方法学

主 编 徐文方

编 者 (以姓氏笔画为序)

方 浩 (山东大学药学院)

李敏勇 (山东大学药学院)

张亮仁 (北京大学药学院)

胡永洲 (浙江大学药学院)

徐云根 (中国药科大学)

徐文方 (山东大学药学院)

人民卫生出版社

图书在版编目 (CIP) 数据

药物化学实验方法学/徐文方主编. —北京:
人民卫生出版社, 2010. 6
ISBN 978-7-117-12466-9

I. ①药… II. ①徐… III. ①药物化学-化学实验-
实验方法 IV. ①R914-33

中国版本图书馆 CIP 数据核字(2009)第 227735 号

门户网: www.pmph.com	出版物查询、网上书店
卫人网: www.ipmph.com	护士、医师、药师、中医 师、卫生资格考试培训

版权所有, 侵权必究!

药物化学实验方法学

主 编: 徐文方
出版发行: 人民卫生出版社 (中继线 010-59780011)
地 址: 北京市朝阳区潘家园南里 19 号
邮 编: 100021
E - mail: pmph@pmph.com
购书热线: 010-67605754 010-65264830
010-59787586 010-59787592
印 刷: 尚艺印装有限公司
经 销: 新华书店
开 本: 850×1168 1/16 印张: 16
字 数: 463 千字
版 次: 2010 年 6 月第 1 版 2010 年 6 月第 1 版第 1 次印刷
标准书号: ISBN 978-7-117-12466-9/R · 12467
定价 (含光盘): 56.00 元
打击盗版举报电话: 010-59787491 E-mail: WQ@pmph.com
(凡属印装质量问题请与本社销售中心)

前言

药物化学作为药学专业的龙头学科,是化学小分子药物开发的源头。随着科学技术的发展和多学科的交叉融合,药物化学的内涵正在由传统的化学模式向化学生物学模式转变。纵览世界药物化学最权威的学术期刊《Journal of Medicinal Chemistry》,便可发现当前药物化学的研究内容和技术手段所发生的重大变化。因此,从培养研究生的角度考虑,一个完整的药物化学研究课题应当首先从生物信息学入手,基于药物靶点的结构与功能,利用虚拟筛选和构建药效团等技术手段进行合理的药物分子设计,以确定先导化合物的结构类型;然后从化学信息学入手,应用有机合成化学的基本知识和技术手段进行目标化合物的合成、分离纯化与结构确证工作;下一步围绕药物靶点的作用机制,在分子(酶或受体)水平、细胞水平和整体动物水平上进行目标化合物的初步体内药效学评价;最后应用日趋成熟的 CoMFA 和 CoMSIA 技术进行 3D-QSAR 总结。

由于药物化学研究的领域日趋扩大和多学科交叉,在客观上也要求我们必须不断学习更多的实验技术和方法。但对于那些刚刚从事药物化学研究的科研人员和研究生来说,他们还难以在短时间内掌握涉及这些多学科的实验技术。因此本书写作的初衷,就是帮助他们尽快熟悉本领域研究的基本过程,特别是掌握如何选题并顺利开展相关实验研究工作。

本书在写作中突出了实验方法的实用性,写作素材尽量取材于研究生培养过程中常见的问题。在编写过程中,力图通过必要的基本原理介绍,并结合实际工作中总结出来的各种注意事项,有针对性地帮助研究生解决实际工作中可能会遇到的具体问题,从而使学生掌握严格的基本实验技能训练,以适应今后科研工作中可能会遇到的各种挑战。

另一方面,本书内容还涉及许多的学科前沿知识,这些内容在本科教学中很少提及或几乎没有接触,如计算机辅助药物设计软件操作和微波合成等方法。它们代表着学科发展的前沿,在发展过程中也出现了由不同公司开发的技术设备和操作规程。其实,它们的基本原理都是相同的,本书在编写中将首先简要介绍其基本原理,然后选择代表学科发展且日趋成熟的主流软件或仪器设备进行相关的基本操作介绍。

本书的编者来自北京大学、浙江大学、中国药科大学和山东大学四所国内著名高校,长期工作在教学科研第一线,有着多年从事科学研究和指导研究生工作的经历。但由于药物化学学科及其相关研究领域发展很快,许多内容涉及多学科交叉,编写过程中存在疏漏之处在所难免,敬请广大师生和读者提出宝贵意见和建议,以便修订时加以完善。

编者

2009 年 10 月

—目—录—

第一章 基于靶点的药物发现	1
第一节 基本理论计算基础.....	1
一、量子化学.....	1
二、分子力学.....	2
三、分子动力学.....	4
四、分子动力学实验方法.....	5
第二节 分子对接技术.....	7
一、实验方法及操作步骤.....	8
二、分子对接进行药物设计举例	17
第三节 蛋白质同源模建技术	19
一、实验方法及操作步骤	21
二、同源模建应用举例	24
第二章 基于药效团的药物发现	28
第一节 药效团技术基本原理	28
一、药效团的基本概念	28
二、药效特征元素的定义	29
第二节 药效团模型的构建方法	32
一、基于配体活性小分子共同特征的药效团	32
二、基于受体结构的药效团	33
第三节 药效团实验方法和操作步骤	33
一、基于配体共同特征的药效团	33
二、基于化合物构效关系的药效团构建	36
三、基于受体结构的药效团	38
第四节 药效团方法运用案例	41
一、HIV-1 整合酶抑制剂的发现	41
二、HIV-1 整合酶抑制剂的结构优化	41
三、基于受体结构的药效团	42
第三章 类药化合物的合成路线设计与优选	47
第一节 类药化合物合成路线的设计策略	47
一、有机合成设计理论与逆合成分析法	47
二、类型反应法	48
三、模拟文献法	48
第二节 逆合成分析法	48
一、逆合成分析的有关概念	49

二、逆合成分析的常见切断方式	51
三、逆合成分析理论的应用	57
四、逆合成分析方法的局限性	61
第三节 药物合成反应的文献调研	62
一、常用的化学类三次文献信息资源	63
二、美国化学文摘	65
三、贝尔斯坦有机化学大全	70
第四节 类药化合物合成路线选择与评价原则	74
一、原料和试剂的来源	74
二、合成步骤、操作与收率	75
三、单元反应的次序安排	75
四、技术条件与实验安全	75
 第四章 药物合成实验研究的思路与常用方法	77
第一节 药物合成实验室工作规范	77
一、实验室安全是首要责任	77
二、学术道德及实验记录规范	78
三、实验记录模板	80
第二节 影响药物合成反应的常见因素	81
一、反应物料的配比与浓度	81
二、温度	83
三、压力	84
四、加料方式与加料顺序	84
五、溶剂	85
六、酸碱度(pH)	86
七、搅拌	86
八、杂质与原料的纯度	86
第三节 药物合成工艺研究的常用方法	87
一、应用新试剂或新反应	87
二、后处理方法的改进	89
三、正交设计	90
四、微波合成反应	93
第四节 药物合成路线探索的一般步骤	98
一、如何研究单元反应	98
二、药物合成工艺路线探索实例	98
 第五章 类药化合物合成的实验技术	102
第一节 药物合成反应基本操作	102
一、基本操作简介	102
二、加热、回流	105
三、冷却	106
四、搅拌	107
五、通气	109

六、无水无氧反应	111
七、半微量操作	118
第二节 反应终点的监测	118
一、TLC 简介及基本操作	118
二、HPLC 监测反应进程的基本操作	125
三、GC 监测反应进程的基本操作	128
四、其他方法	130
第六章 类药化合物的分离纯化与结构确证	132
第一节 类药化合物制备的常用分离与纯化技术	132
一、蒸馏	132
二、重结晶	137
三、柱层析	141
四、快速柱层析	143
五、制备薄层色谱	144
六、制备高效液相	145
第二节 类药化合物的结构解析	146
一、化合物纯度鉴定的常用方法	146
二、化学鉴定	147
三、元素分析	155
四、波谱解析	156
第三节 药物合成产品的结构确证范例	172
一、元素分析	172
二、紫外吸收光谱	173
三、红外吸收光谱	173
四、核磁共振氢谱(^1H NMR)	175
五、核磁共振碳谱(^{13}C NMR)	178
六、质谱	181
七、热分析	181
八、热重量分析	182
九、X-粉末衍射	182
十、综合解析	183
第七章 类药化合物的初步活性评价	185
第一节 常用药理活性评价方法简介	185
一、分子水平的活性评价	185
二、细胞水平的活性评价	187
三、动物水平的活性评价	187
第二节 类药化合物在分子水平上的活性评价	188
一、基于酶活性的先导化合物评价方法	188
二、基于受体活性的先导化合物评价方法	193
三、基于离子通道的先导化合物活性评价	197
第三节 类药化合物在细胞水平上的活性评价方法	199

一、抗肿瘤药物细胞水平的活性评价方法	199
二、抗菌药物的活性评价方法	202
三、抗病毒药物的活性评价方法	206
第四节 基本统计学方法与试验数据的处理	211
一、统计学的基本概念	211
二、常用的量反应资料统计方法	212
三、质反应的显著性检验	215
第五节 应用举例	218
一、目标化合物抑制氨肽酶 N 的活性试验(In vitro)	218
二、目标化合物抑制明胶酶的活性试验(In vitro)	218
三、目标化合物对白血病细胞生长抑制试验	218
四、目标化合物对人卵巢黏液囊腺上皮癌细胞的生长抑制以及明胶酶谱分析试验	219
五、目标化合物抑制荷肝癌 H22 小鼠血道转移以及碳粒廓清实验	220
第八章 QSAR 研究方法及结构优化	223
第一节 QSAR 的基本原理与方法	223
一、二维定量构效关系	223
二、三维定量构效关系	229
三、四维和五维定量构效关系	234
第二节 CoMFA 和 Topomer CoMFA 研究的一般步骤及注意事项	236
一、实验流程	236
二、实验软硬件基础	236
三、应用 CoMFA 方法建立定量构效关系模型	236
四、Topomer CoMFA 与 Topomer Search	238
第三节 定量构效关系在新药研究中的应用实例	239
一、CoMFA 模型的应用	239
二、Topomer CoMFA 的应用	239
索引	242

第一章

基于靶点的药物发现

第一节 基本理论计算基础

合理的药物分子设计是新药研究和开发过程中的重要环节,它对于缩短成药周期、降低药物成本具有重要的意义。随着理论计算技术以及计算机硬件和软件的发展,计算机辅助药物分子设计(computer aided drug design, CADD)手段越来越成熟,已广泛地应用到新药研发过程中,且发挥着越来越重要的作用。它以计算机为工具,充分利用有关药物及其生物大分子靶标的知识,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,以缩短药物的开发周期。

计算机辅助药物设计是以各种分子模拟技术和数理统计方法为基础的,所基于的理论计算大体有三种,即量子化学、分子力学和分子动力学。每种方法有各自的优点和局限性,很多时候需要结合使用,取长补短。

一、量子化学

量子化学(quantum chemistry)是理论化学的一个分支学科,是应用量子力学的基本原理和方法,研究和解决化学问题的一门基础科学。1927年海特勒和伦敦用量子力学基本原理讨论氢分子结构问题,说明了两个氢原子能够结合成一个稳定的氢分子的原因,并且利用相当近似的计算方法,算出其结合能。由此,使人们认识到可以用量子力学原理讨论分子结构问题,从而逐渐形成了量子化学这一分支学科。量子化学的发展历史可分两个阶段:第一个阶段是1927年到20世纪50年代末,为创建时期,其主要标志是三种化学键理论的建立和发展,包括价键理论、分子轨道理论和配位场理论;第二个阶段是20世纪60年代以后,主要标志是量子化学计算方法的研究,其中严格计算的从头算方法、半经验计算的全略微分重叠和半略微分重叠等方法的出现,扩大了量子化学的应用范围,提高了计算精度。

随着计算机科学和量子化学计算方法不断发展,计算量以及计算速度不断提高,所计算的体系越来越复杂,现在可以计算有机分子甚至分子量较大的生物分子。运用量子化学计算方法,可以计算得到分子的几乎一切性质,如稳定结构、能量、轨道能级、原子电荷、静电势、偶极矩、电离能、电子亲和、电子密度、过渡态和反应途径等。

量子化学研究的电子-原子核体系可用相应的薛定谔方程解的波函数来描述。原则上,薛定谔方程的全部解保证了多电子体系中电子结构与相互作用的全面描述。然而,由于数学处理的复杂性,在实践中,总希望发展和运用量子力学的近似方法,从而无需进行很繁杂的计算就可以说明复杂原子体系的主要特性,这就必须在原始量子化学方程中引进一些重要的简化,以便得到一定程度的近似解。量子化学发展到现在,根据为解薛定谔方程而引入近似程度的不同,大致可分为以下几种方法:

1. 密度泛函理论(density functional theory, DFT)

1964年, Kohn提出了密度泛函理论, 其以 Hohenberg-Kohn 定理为基础, 指出电子密度决定分子的一切性质, 体系的能量是电子密度的泛函。各种密度泛函理论差别在于选择交换泛函和相关泛函的不同, 例如: 纯密度泛函包含一个相关泛函和一个交换泛函, 如 BDW、BLYP 等; 而杂密度泛函则包含一个相关泛函和多个交换泛函, 例如 B3LYP、BHandHLYP 等。

由于密度泛函计算结果精确, 计算速度快, DFT 以无可比拟的优越性成为当前国际研究的主流方向, 与分子动力学结合的分子模拟, 更是当前理论化学研究反映动态过程的有力工具, 成为计算材料科学的重要基础和核心技术。但 DFT 也并不是适合所有的体系, 研究表明, 其对共价键体系计算结果精确, 氢键体系次之, 范德华键体系再次之。

2. 从头算方法(ab initio method)

在量子化学中, 从头算法是指基于 Born-Oppenheimer、独立电子(hartree)和非相对论三大近似, 利用电子质量、plank 常数和电量三个基本物理量及原子系数, 对分子的全部积分严格进行计算, 不借助任何经验参数来求解薛定谔方程的计算方法。不同的从头算法考虑了不同的相关能项, 如: HF 方法只考虑了同电子自旋的相关(交换相关)问题, 而没有考虑相反自旋的电子相关问题和瞬时电子相关的问题; MPn 方法给体系考虑了微扰项, 而更为精确的计算应包含更多的相关能项, 如组态相互作用方法(CIS, CISD)和耦合簇方法(CASSCF)等。

由于从头算法在理论上的严格性和计算结果的精确性、可靠性, 从小分子体系到大分子体系、从静态性质到动态性质, 各方面都有从头算法的应用。对过渡金属配合物, 金属原子簇合物等大分子化合物的研究也迅速增加。但基于计算精度和计算资源的矛盾的考虑, 从头算法主要应用于小分子体系的高精度计算, 对中等分子量的小体系进行的定量计算, 对大分子体系的定性计算三个方面。

3. 半经验方法(semi-empirical method)

从头算法虽然有严谨的理论支持, 能得到较好的计算结果, 但是当遇到诸如酶、聚合物、蛋白质等大分子体系时, 计算很耗时, 其计算代价无法承受。为了在计算时间和计算精度上找到一个平衡点, 科学家们以从头算法为基础, 忽略一些计算量极大, 但是对结果影响极小的积分, 或者引用一些来自实验的参数, 从而近似求解薛定谔方程, 就诞生了半经验算法。其最核心的近似方法是忽略双原子轨道微分重叠的 NDDO 近似(neglect of diatomic differential overlap)。在此基础上又有了所谓全略微分重叠近似的 CNDO(complete neglect of differential overlap), 间略微分重叠近似的 INDO(intermediate neglect of differential overlap)和改进的间略微分重叠近似 MINDO(modified intermediate neglect of differential overlap)等。以后的半经验方法, 是在 MINDO 方法的基础上进行改进。目前常用的改进后的方法有 ZINDO(Zerner INDO), AM1 和 PM3 等。

半经验方法理论上没有从头算法那么严谨, 因而在处理复杂体系的中间体、过渡态时会遇到一定的困难, 其计算的结果只带有定性和半定量的特性。主要应用于非常大的体系的计算或处理大体系的第一步, 或为了得到一些分子的初步研究结果。

量子化学为研究药物的电子结构提供了一个有效的方法, 能求出一系列分子的各种参数, 从中找出与药物活性相关的参数, 由此得知有关受体的结构、构象及反应性等方面的信息, 揭示药物结构和生物活性的关系, 也能研究药物和受体的相互作用, 为合理的药物设计提供依据。目前常用的量子化学计算软件有 GAUSSIAN、GAMESS、MOLPRO 等。

根据量子化学计算出的药物构象为分子在真空状态下的构象, 而药物在体液中呈现活性, 目前还没有准确处理药物分子的较为实用的方法, 量子化学计算一般也没有考虑到反应的熵效应, 对于大分子的药物, 计算量大而不实用。

二、分子力学

随着计算机科学的迅速发展和研究体系的尺度的增加, 分子力学(molecular mechanics, MM)已

成为当前化学家经常使用的研究手段,也是分子模拟方法的重要组成部分。其基本用途是寻求分子的平衡构型及能量。它是分子中原子间存在化学键,非键原子之间有范德华及静电相互作用这一经典理论的自然扩展。

分子力学的基本思想可以追溯到 1930 年安德鲁斯(Andrews)提出的经典力学模型:分子中的化学键具有“自然”键长,键角,并由这些键长和键角调节构象,给出核位置的最佳分布,即分子的平衡构型。基于 Born-Oppenheimer 近似,分子力学计算不包含对电子的处理。简单说来,分子力学着眼于核的运动,把化学键形成的力场类比为弹簧,即将分子内各原子之间的键合作用看作是符合 Hooke 定律的谐振运动。分子力学计算是一种试图通过基于原子坐标的势能函数来理解分子系统的物理性质的尝试,这方面的应用最早是和 Allinger 的有机小分子的构象分析相联系的,它主要用于能量优化、简振分析、分子动力学模拟、蒙特卡罗模拟等领域。

分子力学方法基于力场或称经验力场(force field),忽略分子中电子的运动,把系统的能量仅作为各原子核的函数即经验势函数,利用力场方法模拟分子结构,计算分子的性质。因此所计算的体系可以包含大量数目的原子,通常可达 5000 以上。在某些情况下,使用力场进行计算所得的结果甚至可以和最高级别的量子力学计算的结果相媲美,而计算时间却只需几分之一甚至几十分之一。但是,从本质上说,分子力学是对真实势能函数的近似拟合,是唯象的。即它不但在将物理概念转化为公式的同时牺牲了物理概念本身,而且仅在平衡态有效,在电子状态及过渡态却不适用。由于忽略了电子的运动,分子力学无法获得那些取决于分子中电子分布的性质如态密度、过渡态、电子结合能等,但它确实让计算机“认识”了势函数,使化学和计算机的结合成为现实。对于多数有机和高分子系统,分子力学得到的热力学、动力学、平衡构象等性质是合理的、可靠的。

分子力学以经典牛顿力学为基础,能快速计算分子的构象、能量和多种性质。分子力学将分子看作由一组靠弹性力(或谐振力)维系在一起的原子的集合,其中每个化学键(相当于弹簧)都有“自然”的键长和键角标准值,分子要调整它的几何形状(构象),必须使其键长和键角尽可能接近标准值,同时也使非键作用能处于最小。如果原子或基团在空间上过于靠近,便会产生较大排斥力;如果过于远离又会使键长伸长或键角发生弯曲,也将引起相应的能量升高。因此每个真实分子结构是相互制约各作用力的折中表现。在计算中,分子总能量 E_{total} 可表示为:

$$E_{\text{total}} = E_c + E_b + E_t + E_v + E_h + E_e + E_d \quad (\text{式 1-1})$$

式中, E_c 是键伸缩能, E_b 是键角弯曲能, E_t 是二面角扭转能, E_v 为 van der Waals 作用能, E_h 为氢键作用能, E_e 为静电作用能, E_d 为偶极作用能。每一能量项均由一定势能函数形式和力场参数构成。分子力学研究已成为当今药物分子设计的发展方向之一,在天然产物结构或构象平衡、糖类、核酸、多肽以及蛋白质的构象计算或模拟方面正在开展广泛的研究。

量子力学计算方法能够准确而又细致地为我们提供有关分子结构、能量变化,以及反应活性等方面的信息。然而由于该方法在体系大小方面受到限制,对于复杂的生物体系,目前用量子化学直接计算仍相当困难。为此,科研人员往往不得不采用较为粗糙的、纯属经验性质的分子力学方法来处理大型的化学体系。这些分子力学方法虽然能够提供一些有关分子构象以及静电势分布的信息,但计算精度很差,而且它不能有效地描述化学键的形成和断裂过程。为了解决上述矛盾,科学家们将量子力学从头算方法与分子力学方法有机地结合起来,对于体系中重要的、牵涉到化学键变化的部分,也就是活性中心,使用量子力学从头算方法来描述;而对于体系中次重要的、不直接牵涉到化学键变化的部分,也就是活性中心以外的原子,使用分子力学方法来描述。这种方法就是量子力学和分子力学联用方法,简称 QM/MM 方法。在目前的条件下使用 QM/MM 方法处理含有上万个非氢原子的体系是完全可行的。最早有关于 QM/MM 方法的研究是由 Warshel 等人在 20 世纪 70 年代开展。经过随后近三十年的改进,QM/MM 方法已经从理论探讨逐渐走上实践应用,为配体-受体相互作用和酶催化反应机制的研究提供了新的计算途径。

三、分子动力学

分子动力学(molecular dynamics, MD)是在原子、分子水平上求解多体问题的重要计算模拟方法。自1970年起,由于分子力学的迅速发展,力场的不断开发,随之建立起许多适用于生化分子体系、聚合物、金属与非金属材料的力场体系,使得计算复杂系统的结构与一些热力学与光谱性质的能力及准确性大为提高。分子动力学模拟就是应用这些力场及根据牛顿力学原理所发展起来的计算方法,目前已成为结构生物学、分子生物学、生物物理、计算化学、药物化学、流体力学、材料和纳米研究领域中的重要理论工具。

分子动力学由Alder等在1957年发明,1977年McCammon首次将其应用到蛋白质模拟研究领域。此后生物大分子的动力学模拟经历了迅速发展,普遍应用于蛋白质、核酸、糖类生物大分子以及相关复合物结构、功能的研究,能够获取分子体系的平衡热力学性质和结构性质;通过构象搜索为核磁共振、X射线晶体衍射的结构进行精细优化,可以研究生物大分子的动力学过程,例如生物膜传输、受体-配体的相互作用、大幅度的蛋白质构象变化、蛋白质折叠与解折叠、离子通道的门控、分子马达的运动等。由于能够广泛应用于生物大分子的各个领域,目前已经发展了许多适合生物大分子体系的分子动力学软件,如AMBER、NAMD、CHARMM、GROMACS和GROMOS等。

1. 基本原理

在经典分子动力学模拟中,假定原子运动由牛顿方程所决定,亦即表明原子运动按特定轨迹进行。分子动力学忽略核运动的量子效应,假设绝热近似严格成立,即假设每一时刻电子均处在相应原子结构的基态。在一定力场下进行分子动力学模拟,首先由经验势能函数通过能量极小化得到坐标 r ,势能(E_p)对坐标的一阶导数的负值就是力 $F = -\partial E_p / \partial r$,再按牛顿第二定律得到加速度。一旦知道了某时刻 t 的 r 、 F 和 E_p ,就可以知道另一时刻 $t + \Delta t$ 的新的力,再由新的力求得新的速度,用新的力和速度得出新的位置。于是位置随时间向前移动,通常称之为模拟。分子动力学模拟既能得到原子的运动轨迹,进而基于轨迹计算得到所需各种性质,还能像做实验一样进行各种观察,可作为对理论和实验的有效补充。对于平衡体系,计算所得是一定时间内某物理量的统计平均值;对于非平衡体系,获得的是一定时间内物理现象的直接模拟。许多与原子有关的细节,在实验中无法获得,而在分子动力学模拟中可方便地求得。

2. 分子动力学模拟的系统

采用分子动力学模拟,必须在一定的系统下进行,经常用到的系统包括微正则系统、正则系统、等温等压系统和等压等焓系统。

(1) 微正则系统(NVE):是孤立的、保守的系统的系统,在这种系统中,系统沿着相空间中的恒定能量轨道演化,在分子动力学模拟的过程中,系统中的原子数(N)、体积(V)和能量(E)都保持不变。

(2) 正则系统(NVT):NVT保持不变,并且总动量为零。恒温下,系统的总能量不是一个守恒量,系统要与外界发生能量交换。保持系统的温度不变,通常运用的方法是让系统与外界的热浴处于热平衡状态。由于温度与系统的动能有直接的关系,通常的做法是把系统的动能固定在一个给定值上。

(3) 等温等压系统(NPT):NPT保持不变,这种系统是我们常见的系统,许多分子动力学模拟都要在此系统下进行。这时,要保证系统的温度恒定,还要保持它的压力恒定。温度恒定和以前一样,是通过调节系统的速度或加一约束力来实现的。而对压力进行调节,就比较复杂。由于系统的压力 P 与其体积 V 是共轭量,要调节压力值可以通过改变系统的体积来实现。

(4) 等压等焓系统(NPH):NPH保持不变,焓值通过 $H = E + PV$ 得到。在系统下进行模拟时要保持压力和焓值为固定值。

3. 周期性边界条件

正确处理周期性边界和边界效应对模拟方法是至关重要的,因为它从模拟相对较少的原子来

计算物质的宏观性质的。为了减小有限尺寸的影响,经典的方法是使用周期性边界条件。

周期性边界条件使得可以用相对少的粒子数目来真实的模拟大块的体系,使得粒子仿佛处在一个完整的体系中。假设一个用来模拟的立方单胞,使这个单胞在各个方向上都不断重复,看起来像有周期一样。在二维图像中,每一个单胞都被其他的 8 个单胞所包围;在三维方向上,每一个单胞就会被 26 个单胞所包围。

所有的单胞中的粒子的坐标都可以通过一个整数而得到。当模拟的单胞中的一个粒子由于力的作用而离开这个单胞的时候,就会有另一个和它对应的粒子运动到这个单胞中来,这样,模拟的整个体系的粒子数就会保持不变。

在分子动力学模拟中经常采用的边界条件有:矩形盒子周期性边界条件;立方体盒子周期性边界条件;六方柱盒子周期性边界条件;截顶八面体盒子周期性边界条件。

4. 分子动力学模拟结果分析与方法

分子动力学模拟的时间能够达到几百皮秒或纳秒级,甚至更长。在运行分子动力学模拟时,体系的速度和坐标被保存下来。在分析时可以对热力学参数进行计算。热力学参数随时间演化的特性可以用图形显示,一个坐标对应时间,另外一个坐标是感兴趣的物理量,如能量、均方根偏差、原子位置的涨落,此外还能计算出平均结构与实验数据进行对照,这些都有助于在原子水平上直观地理解构象的变化。

5. 分子动力学常用的分子力场

(1) 第一代经典的分子力场。主要是 AMBER、CHARMM 以及 CVFF。其最初的参数化主要基于实验结果。其特点是函数形式简单;应用的范围比较特定(大部分适合于生物分子);优化力场参数的方法多,结果也比较好;能合理地预测分子结构,构象性质,凝聚态性质。

(2) 第二代分子力场。CFF91、CFF、PCFF、COMPASS 以及 MMFF。这一类力场的共同特点是函数形式较复杂、附加项多、适用范围宽,特别是大而复杂的分子模型的多能量极小和势垒、优化振动频率,构象性质。

(3) 针对周期表中的所有元素的力场。所有的力学参数基于元素,杂化和化合的规则而产生,并为许多结构形式所证实。这类力场有 ESFF、UFF、Dreiding FF。特点是:函数形式简单,适用的范围广,能比较合理地预测分子结构,根据某些理论或经验的规则使力场参数化。Universal 力场(UFF)是针对整个周期表的分子力学和动态模拟的力场,力场参数是依靠元素、元素的杂化及化合性而计算出来的,与电荷平衡算法相结合。这类立场由于在预测精度上存在一些问题,因此在分子模拟中并没有得到广泛的应用。

(4) 用于特定目的的分子力场。如用于研究无机氧化玻璃体和离子体系的 Glass 力场,用于聚偏二氟乙烯和相关高分子的 MSXX 力场等。

四、分子动力学实验方法

分子动力学已成为当前最广为采用的计算庞大复杂体系的方法,可广泛适用于多种体系和各类性质的研究。下面简要介绍一下使用 GROMACS 软件进行分子动力学模拟研究的操作过程。GROMACS 最早是由荷兰的 Groningen 大学开发的免费分子动力学模拟软件,主要用于生物大分子的模拟研究。后来由瑞典的 Uppsala 和 Stockholm 大学及德国马普学会高分子研究所进一步的发展和完善,应用领域越来越广,与其他动力学软件相比,速度快是它的特点。

具体的操作步骤可能会因为模拟目标的不同而存在差异,这里只是做基本的流程示范:

1. 准备文件

(1) 准备 gromacs 需要的坐标文件和拓扑文件

命令: `pdb2gmx-ignh-f xxx.pdb-o xxx.gro-p topol.top`

注:①该命令之后,程序会提醒选择力场,选择 gromacs forcefield 力场,当然也可以根据自己的

体系需要选择别的力场。②gro 文件是坐标文件,所有的 gro 文件都可以由 pdb 文件所替代,top 文件是拓扑文件。

(2) 构建模拟所需要的水盒子

命令:editconf-f xxx.gro-o xxx_box.gro-d 1

注:构建的盒子为 xxx_box.gro 其中水盒子中的蛋白与盒子边界是 1Å 的距离,该距离可以根据自己的体系调节,最小不能低于 0.85Å。

(3) 给水盒子填充水溶液

命令:genbox-cp xxx_box.gro-cs spc216.gro-p topol.top-o xxx_sol.gro。

注:gromacs 默认的水是 SPC 水模型,也可以选择别的模型,如 spce、tip3p、tip4p 等。此时,一个有盒子,一个有溶质(蛋白质)、有溶剂(水)的坐标文件产生了,即为 xxx_sol.gro。

(4) 给体系分组

命令:make_ndx-f xxx_sol.gro-o yyy_sol.ndx

注:分组根据自己的体系需要。若只是跑空蛋白,此步骤可以省略,gromacs 有默认的分组。

(5) 做限制文件,即 itp 文件

1) 命令:genpr-f xxx_sol.gro-n-xxx_M.itp 选择 MainChain

注:该 itp 文件是用于限制蛋白的主链。

2) genpr-f xxx_sol.gro-n-xxx_C.itp 选择 C-alpha

注:该 itp 文件是用于限制蛋白的。

(6) 给体系加抗衡离子

命令:grompp-f sd.mdp-c xxx_sol.gro-p topol.top-o preion.tpr

genion-s preion.tpr-o ion.gro-nn/-np n

注:nn 表示需要加入负离子作为抗衡离子,gromacs 默认为 Cl^- ;np 表示需要加入正离子作为抗衡离子,gromacs 默认为 Na^+ ;n 表示所填入的抗衡离子的数目。

给体系添加了抗衡离子要相应地修改拓扑文件。

2. 能量优化和动力学模拟

(1) 最陡下降法,能量最小化

命令:grompp-f sd.mdp-c ion.gro-p topol.top-o sd.tpr

mdrun-s sd.tpr-o sd.trr-c sd_after.gro-e sd.edr-g sd.log-nice 0

(2) 共轭梯度法,能量最小化

命令:grompp-f cg.mdp-c sd_after.gro-p topol.top-o cg.tpr

mdrun-s cg.tpr-o cg.trr-c cg_after.gro-e cg.edr-g cg.log v-nice 0

(3) 主链限制的动力学模拟

命令:grompp-f pr1.mdp-c cg_after.gro-p topol.top-o pr1.tpr

mdrun-s pr1.tpr-o pr1.trr-c pr1_after.gro-e pr1.edr-g pr1.log-nice 0

(4) Ca 限制的动力学模拟

命令:grompp-f pr2.mdp-c pr1_after.gro-p topol.top-o pr2.tpr

mdrun-s pr2.tpr-o pr2.trr-c pr2_after.gro-e pr2.edr-g pr2.log-nice 0

(5) 没有限制的动力学模拟

命令:grompp-f full.mdp-c full_after.gro-p topol.top-o full.tpr

mdrun-s full.tpr-o full.trr-c full_after.gro-e full.edr-g full.log-nice 0

mdp 文件为 gromacs 的参数文件。可以理解为 molecular dynamics parameter files。

tpr 文件作为动力学的输入文件包含了多方面的信息即可以认为:tpr=top+gro+mdp+ndx。

trr 文件为轨迹文件,但是该文件通常比较大,可以用 trjconv 命令将 trr 文件压缩为 xtc 文件。

3. 动力学结果分析

(1) 压缩轨迹文件

命令: `trjconv-f xxx. trr-o xxx. xtc`

注:此时可以用 VMD 来查看轨迹文件。

(2) 计算蛋白骨架 RMSD 值

命令: `g_rms-s xxx. tpr-f xxx. xtc-o bkbone_rmsd. xvg` 选择 bakebone 选项

注:所有的 xvg 文件都可以由 xgramce 显示。

(3) 叠合初始状态的蛋白和动力学模拟后的蛋白

命令: `g_confrms-f1 xxx_original. pdb-f2 xxx_final. gro-o fit. pdb`

注:得出的 fit. pdb 是两个蛋白的叠合构象。可以根据自己的需要选择某两帧的蛋白进行叠合。

(4) 取平衡后的平均结构(假设取平衡 1ns 中最后的 200ps)

命令: `g_covar-f xxx. xtc-s full. tpr-b 801-e 1000-av traj_avg. pdb`

注:平均结构比较粗糙,需要做一个能量优化。

(5) 分析能量并作图

命令: `g_energy-f full.edr-o xxx_pe. xvg` 选择 Potential 选项,给势能作图,从曲线图中可以观察到每一帧构象的势能变化,可以选择低能构象进行下一步的研究。

(6) 计算分子或是分组之间的氢键数

命令: `g_hbond-f xxx. trr-s full. tpr-num xxx_hnum. xvg`

第二节 分子对接技术

分子对接是应用锁匙原理的 CADD 方法。分子对接的功能包括预测结合模式与结合能力以及进行虚拟筛选。

在准确预测受体与配体相互作用的基础上,分子对接方法有一项重要应用——虚拟筛选(图 1-1)。由于进行药物的实体高通量筛选需要构建大规模的化合物库,提取或培养大量实验必需的靶酶或者靶细胞,并且需要复杂的设备支持,因此进行实体的药物筛选要投入巨额的资金。虚拟药物筛选是将药物筛选的过程在计算机上进行模拟,对化合物对某药物靶点的生物活性作出预测,进而对比较有可能成为药物的化合物进行有针对性的实体筛选,从而可以极大地减少药物开发成本。

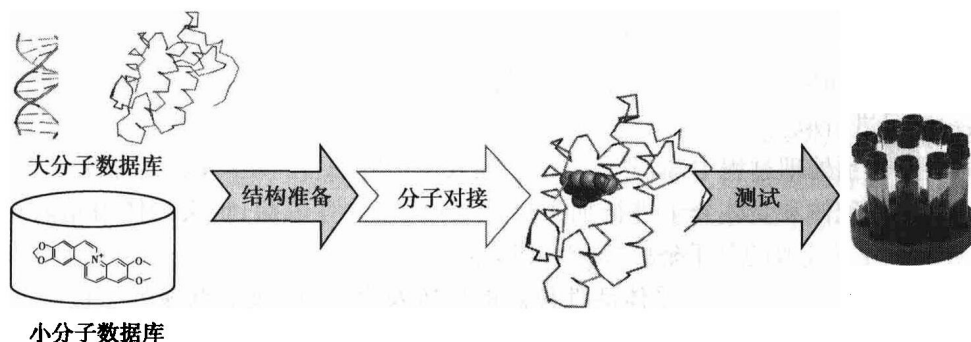


图 1-1 虚拟筛选流程

目前常用的分子对接软件有很多,包括:DOCK、AutoDock、FlexX、Gold、Glide、FRED、Surflex Dock、3D-DOCK 等。本节主要介绍 AutoDock 软件和 Sybyl 软件包中的 Surflex Dock 模块。

一、实验方法及操作步骤

(一) AutoDock

AutoDock 是由美国 Scripps 研究院(The Scripps Research Institute, TSRI)开发的一款免费的分子对接软件,其 4.0 版本遵循 GNU 通用公共许可证(GPL)协议。作为经典的分子对接程序组件,AutoDock 有着对接结果准确,速度较快的优点,其引用次数居于同类软件前列。AutoDock 的具体应用包括 X-射线晶体学研究、基于结构的药物设计、先导物优化、虚拟筛选、组合化学库设计、蛋白-蛋白对接以及化学动力学研究。本书主要从实验方法学的角度介绍该软件基本的分子对接方法,感兴趣的读者可参考本部分最后的进一步阅读建议,从而更加深入地掌握 AutoDock。

AutoDock 程序套件由 AutoGrid 和 AutoDock 两个程序组成。前者用于在配体-受体结合位点附近进行格点计算,而后者将小分子对接入前者预先计算好的格点。二者的输入和输出均以对文本文件的操作实现,使用者可以方便地对各种参数进行调整。为了方便用户使用,Scripps 研究院的分子图形实验室还开发了配套的 AutoDockTools(ADT)作为配套软件。其图形用户界面(graphic user interface, GUI)可实现预处理、对接操作以及对接后结果分析的图形化操作,而其集成的脚本套件则可帮助使用者轻松地实现大量配体的自动化虚拟筛选。ADT 作为 MGLTools 软件包(molecular graphic laboratory tools)的主要组成部分,可在 AutoDock 页面上免费下载。其最新稳定版本为 1.5.2。

本实验以雌激素受体 ER- α 与其配体雌二醇的分子对接为例,讲述 AutoDockTools 和 AutoDock 的一般使用方法以及 AutoDock 对接结果的一般分析方法。

1. 实验流程

运用 AutoDock 进行对接操作的大致步骤为:获取受体和配体的坐标、预处理、格点计算、对接操作、结果分析、参数优化(并进行新的对接操作)。作为一本实验方法学指导书籍,本书只有限地介绍与 AutoDock 对接操作直接相关的操作基本原理。对 AutoDock 对接原理感兴趣的读者,可参阅 AutoDock 的官方网站以及软件开发人员 Morris G. M. 等发表的相关文献。

(1) 获取受体和配体的坐标文件:在进行对接操作之前,首先要获得受体和配体的坐标文件。本例从 RCSB-Protein Data Bank(蛋白结构数据库)获得雌激素及其受体的复合物坐标文件,并使用普通的文本编辑器完成初步的修改:

- A. 去除原晶体坐标文件 6 个复合物中的 5 个;
- B. 去除原晶体坐标文件中所有的水分子;
- C. 分别保存配体以及受体的坐标。

读者需要注意的是,这里的操作只代表了典型的操作流程。坐标文件中多个相同分子的去除、水分子的去除都要视具体体系情况,尤其是受体活性位点位置来确定,不可一概而论。读者需要根据具体所研究课题情况进行取舍。

(2) 预处理:在预处理过程中,需要进行各种输入文件的准备。主要操作包括:

- A. 为待对接的配体和受体分子进行加入极性氢原子、计算电荷、定义配体可扭转键等预处理,生成 AutoDock 4 可以读取的分子坐标“.pdbqt"文件;
- B. 根据不同配体的原子类型、受体活性位点的位置及大小,生成格点参数文件“.gpf”(grid parameter file);
- C. 根据配体的复杂程度以及受体活性位点大小等因素,生成对接参数文件“.dpf”(docking parameter file)。

(3) 格点计算:在预处理的基础上,AutoGrid 针对每一种原子类型,如脂肪、芳香型碳原子,脂肪、芳香型氧原子等,在蛋白质以及核酸等大分子中进行格点计算。这一步计算中,AutoGrid 负责生成不同原子类型在格点处与受体的相互作用能表,供对接时由 AutoDock 调用。这张相互作用能表