

药物分子设计研究进展*

罗小民 蒋华良 沈建华 陈凯先

(上海生命科学研究院 上海 201203)

摘要 综述了药物分子设计研究取得的一些进展,对今后的发展趋势进行了展望,对我院在药物设计方面的发展提出了一些建议。

关键词 药物设计,虚拟筛选,分子对策,全新药物设计,定量构效关系,药效团模型



1 学科意义

创新药物研究具有重大的社会效益和经济效益。然而,新药的寻找是一件耗资巨大而效率很低的工作。据国际上的统计,研制成功一种新药,平均需要花费 10—12 年的时间,

筛选 1.5—2 万种化合物,耗资 3.0—5.0 亿美元。造成这种状况的重要原因之一,就是新药的发现还缺乏成熟的理论,新药的研制仍主要依赖反复试验。这种状况迫切需要通过发展、应用新的理论方法和技术予以改变。

药物设计一直是人们的一个梦想,经过科学家多年的努力探索,药物分子设计方法已经越来越成熟,在许多药物的研究中发挥了重要的作用。

2 国内外研究现状

药物设计方法主要分为两种:一种是从小分子结构出发,通过改造、修饰等方法得到活性更好、毒性更低的新化合物,称为间接药物设计;另一种是从生物靶标大分子结构出发,寻找、设计能够与它发生相互作用并调节其功能的小分子,称为直接药物设计。间接药物设计包括定量构效关系(QSAR)

和三维药效基团模型方法,直接药物设计分为分子对接和全新药物设计两种方法。

2.1 间接药物设计

2.1.1 定量构效关系

定量构效关系方法在药物和农药研究中起着重要的作用。所有的定量构效关系方法都假定被研究的分子结合在同一个靶标生物大分子的相同部位,但不同方法采用不同的结构性质以及不同的确定构效关系^[1]。

传统的定量构效关系(Hansch-Fujita 方法)利用分子和基团的物理化学性质作为结构参数,虽然很有用,但常常因为缺乏实验数据而无法应用(那些性质通常难以通过计算获得)。而建立在分子拓扑连接性质基础上的二维定量构效关系方法则一方面容易获取参数,另一方面也不存在三维定量构效关系方法中的构象问题,是目前 LogP、吸收、分布等性质研究中最常见的方法。

利用小分子三维结构作为参数的三维定量构效关系方法在预测小分子与生物大分子的相互作用时非常有用,各种在化合物三维结构基础上进行三维定量构效关系研究的方法,在药物研究中已经越来越广泛地应用。较重要的方法有:距离几何(Distance Geometry)、比较分子力场分析(CoMFA)、比较分子相似性指数分析(CoMSIA)。其中,CoMFA 方法应用最广,它采用化合物周围的静电场、范德华力场等的空间分布作为化合物结构描

* 收稿日期:2003 年 6 月 20 日

述变量,通过最小二乘法建立化合物的生物活性与化合物周围各种力场空间分布之间关系的模型。根据这一模型和相应的计算机处理,显示出应当如何进行结构的改造,以提高化合物的生物活性。

小分子的构象对三维定量构效关系影响非常大。而在早期,由于缺乏生物大分子的结构信息,无法确定小分子与生物大分子的结合构象,只能采取构象搜寻、模板叠合的方法,存在较大的局限性。建立多个构象叠合模型,选择其中预测能力最好的模型进行三维构效关系研究是一种可取的方法。随着蛋白质结构数据库(PDB)中的蛋白质结构数据的不断增加,研究人员较多地利用分子对接方法得到小分子与生物大分子的结合构象,在此基础上,采用 CoMFA、CoMSIA 等方法进行三维定量构效关系研究。这种方法所得小分子结合构象更为合理、可靠。另一方面,将三维构效关系模型与生物大分子三维结构进行映射,可以更好地解释分子的结构与活性的关系。

定量构效关系是最主要的先导化合物优化方法,而 CoMFA 等三维定量构效关系由于其预测能力强,模型形象、直观,已成为最常用的药物设计方法之一。

2.1.2 三维药效基团模型

药效基团模型方法是另一种重要的间接药物设计方法,可用于先导化合物的发现。三维药效基团通常是指那些可以与受体结合位点形成氢键相互作用、静电相互作用、范德华相互作用和疏水相互作用的原子或官能团以及它们之间特定的空间排列方式。对一组具有生物活性的化合物进行化学结构分析和比较,找出其共同的特征结构,即可建立药效基团的模型。另外,还可以通过全新药物设计得到针对某个特定靶标生物大分子的三维药效基团模型。得到药效基团模型后,即可以此为提问结构的模板,搜寻现有的小分子数据库,“筛选”出符合药效基团要求的其它分子,进行药理测试。

三维药效基团搜寻速度比较快,因此,在分子对接前,可以用三维药效基团搜寻方法对化合物数据库进行预处理。这样在减少分子对接计算量的同时,提高了发现对靶标生物大分子有活性的小分子

的概率。

2.2 直接药物设计

2.2.1 分子对接

在现代化学史上,有机化学家已经合成或从天然原料中分离到超过 1 600 万种有机化合物,其中,不到 1%的化合物曾被用于药物筛选,利用这种方法发现了约 7 000 种药物。很可能仍然有大量的潜在的活性化合物被埋藏在这个有机化合物的“矿藏”中,未被人们发现。因此,充分利用这些已知化合物的结构信息,是寻找活性化合物的一种重要手段。分子对接方法通过将化合物三维结构数据库中的分子逐一与靶标分子进行“对接”,通过不断优化小分子化合物的位置、方向以及构象,寻找小分子与靶标生物大分子作用的最佳构象,计算其与生物大分子的相互作用能。利用分子对接对化合物数据库中所有的分子排序,即可从中找出可能与靶标分子结合的分子。分子对接方法的优势在于化合物数据库中的分子均是已知化合物,可以向试剂公司订购,或者按照其合成方法合成,因而可以较快地进行后续的药理测试。

并行化的分子对接方法又被称为高通量虚拟筛选,可以在几天内完成对包含数十万,甚至数百万个化合物的数据库的筛选。分子对接已成为一种与高通量筛选互为补充的寻找先导化合物的方法。研究者比较了随机的高通量筛选和分子对接方法在寻找蛋白质酪氨酸磷酸酯酶(一种 II 型糖尿病的靶标)抑制剂的结果,表明分子对接对活性化合物的富集率是随机筛选的 1 700 倍^[2]。其他研究人员采用分子对接的方法对 ACD-SC 数据库(当时包括 110 万种化合物,现在增加到 220 万种)进行筛选,寻找雌激素受体的配体。在分子对接找到的 37 种化合物中,21 种化合物活性高于 300nM,14 种化合物活性高于 100nM,2 种化合物活性高于 10nM。

虽然分子对接方法取得了极大的成功,但仍有许多需要改进之处,因此,新的分子对接方法不断出现。一般认为,分子对接需要改进的地方包括打分函数、生物大分子柔性和溶剂化效应三个方面。

现在应用于分子对接的打分函数有三种:基于回归的参数、基于知识的打分函数和基于力场参数

的打分函数。不同的打分函数对于不同的体系其效果也不一样,各种打分函数各有不同的优缺点。对于虚拟筛选目的的打分函数,能够区分活性化合物和无活性化合物是最重要的,因此,并无必要正确地对活性化合物排序。将不同的打分函数组合起来(consensus score)可以降低假阳性的比率^[3]。打分函数的严重局限是不考虑溶剂化效应,或是先产生真空构象,然后用一个包含溶剂化效应项的打分函数进行排序。这样使得分子对接倾向于找到小分子与生物大分子在真空中结合的构象^[4]。

小分子与生物大分子结构的对接算法包括分而治之的算法(生长构造法)、模拟退火算法、遗传算法、禁忌搜索等,后面的几种算法统称为随机算法。研究发现在分子对接中,分而治之的算法具有更好的鲁棒性,特别是在进行虚拟筛选时,一般分而治之的算法具有更快的速度^[5]。

在早期的分子对接中,小分子和大分子的柔性都没有考虑,只考虑了小分子的 3 个平动自由度和 3 个转动自由度。现在大部分分子对接程序都考虑小分子的柔性,而把生物大分子当作一个刚体。单一的生物大分子结构会限制药物设计的小分子类型,基于某种小分子化合物与生物大分子晶体结构,也许只能设计出与之类似的结构^[3]。核磁共振(NMR)结构,或者多个晶体结构能够为药物设计提供更丰富的信息^[5]。考虑生物大分子环区(loop)、结构域(domain)柔性的方法可能在不久的将来发展起来,并用于药物设计。将柔性包括进来的主要问题是:结果不是变得更好,而是更糟糕。现有的结果表明,只考虑生物大分子的侧链的柔性对结果没有改善^[5]。利用生物大分子不同的静态结构来考虑其柔性可能会成为一个主要方法,特别是那些引起生物大分子不同结合构象的小分子。

2.2.2 全新药物设计

全新药物设计方法根据靶标生物大分子的活性位点(“结合口袋”)的几何形状和化学特征,设计出与其相匹配的具有新颖结构的药物分子。全新药物设计的方法目前主要有两种。一种方法称为碎片连接法,该方法首先根据靶标分子活性部位的特征,在其“结合口袋”空腔中的相应位点上放置若干

与靶标分子相匹配的基团或原子,然后用合适的连接片段(linker)将其连接成一个完整的分子。另一种方法称为碎片生长法,该方法首先从靶标分子的结合空腔的一端开始,逐渐“延伸”药物分子的结构。在“延伸”过程中,每一步都要对其延伸片段(基团或原子)的种类及其方位进行计算比较,选择最优的结果,再向下一步延伸,直至完成。

全新药物设计方法倾向于得到一个与生物大分子结合最佳的小分子,而化学家们一般对单个分子不感兴趣,因为预测精度低(包括假阳性和假阴性),吸收、分布代谢、代谢、清除和毒性等性质不佳,化学合成困难,可能造成设计的小分子化合物难以成药。基团生长法的缺点是最后的结果严重依赖最初的种子原子放置位置,而最后得到的分子可能难以合成。将来可能在分子生长时将考虑更多的化学规则。另一个发展趋势是通过全新药物设计产生三维药效基团,然后进行数据库搜寻^[6]。全新药物设计的一个优点是分子动力学、蒙特卡罗模拟结合更紧密,因此,在处理大分子柔性方面比分子对接方法更简便。

3 展望

从 1894 年 Emil Fischer 提出药物作用的“锁钥原理”开始,药物分子设计方法日益成熟,技术不断丰富。通过与实验紧密结合,药物设计在药物研究中正发挥越来越重要的作用,已成为药物研究的主要技术之一。药物设计方法的不断丰富、完善必将为建立成熟的药物研究理论奠定基础。随着与药物相关的化学、物理学、生命科学、计算机和信息科学等学科的发展、交叉和综合,药物设计理论、方法不断成熟,药物设计的梦想正在逐步实现。

3.1 从配体设计到药物设计

小分子化合物能够与靶标生物大分子作用并调节其功能是小分子成为药物的必要条件之一,小分子要发展成药物还必须具有良好的吸收、分布、代谢和排泄性质(ADME)以及较低的毒性(Toxicity)。药物研究过程一般是先找到在分子和细胞水平活性较好的化合物,然后进行 ADME/Tox 实验。早期的药物设计偏重于提高小分子活性的设

计,并取得一定的成果,药物设计方法为先导化合物的发现和优化提供了有效的工具。研究人员认识到越早把那些 ADME/Tox 性质较差的化合物淘汰,就越能更有效地减少人力、物力的浪费,因此,发展、利用各种计算方法来预测化合物的 ADME/Tox 性质将成为一个研究热点。目前发展的方法包括定量构效关系、数据库方法和分子模拟。

3.2 计算方法从粗糙到精确

药物设计方法,特别是分子对接方法,已经在很多药物的研究中发挥了重要作用,提供了大量先导化合物结构。然而,药物设计中的假阳性和假阴性一直是一个困扰研究人员的问题。假阳性可能使合成人员花费大量时间去合成没有活性的化合物,而假阴性又可能使研究人员失去发现好的药物先导化合物的机会。造成假阳性和假阴性的原因可能有:打分函数问题、没有考虑生物大分子的柔性以及溶剂化效应。利用几种打分函数组合的方法可以提高打分的稳健性,另外,也可能使一些系统误差发生抵消。考虑生物大分子柔性的方法目前还不成熟,其计算结果没有实质性的提高。一般用广义波恩模型和溶剂可及性表面(GB/SA)处理溶剂化效应,现已有一些分子对接程序采用了这些方法。不过以上方法从本质上来说,还是静态的,只考虑了一个或几个能量最优的构象。要进一步提高计算精度,只能采用分子动力学模拟方法,分子动力学模拟方法又分为自由能微扰和线性作用能方法^[7]。在不同阶段采用精度不同的计算方法可能是今后药物设计的一种策略。

4 我院的学科优势与发展建议

我院自 20 世纪 70 年代末开始建立药物设计学科,经过 20 多年的发展,已经建成了硬件软件方面都居国际领先的计算机辅助药物设计实验室。

基于我院在结构的药物设计方面具有的积累,尤其在虚拟筛选方面,与江南计算机研究所合作在国产超级计算机上并行化分子对接程序,使计算速度大大加快。采用这种方法,我们从商用化合物数据库 MDL/ACD、MDL/MDDR、ACD-SC 以及中国天然产物数据库(CNPD)中找到了一系列对酪氨酸

激酶、钾离子通道、 β -分泌酶、MMP 以及 PPAR 等靶标具有生物活性的化合物。针对 PPAR 的配体结合域,我们利用并行化的 DOCK 对 ACD-SC(包括 220 万种化合物)进行筛选,找到了 23 种活性在 M 以上的化合物,其中 5 种化合物的结合活性高于所有已经上市的 PPAR 激动剂。采用虚拟筛选方法,我们仅仅用了 6 个月就完成了以上工作。利用活性最好的化合物作为骨架进行组合化学集中库研究,以及复合物的晶体结构测定工作正在进行中。

我院在天然产物数据库的构建方面也有优势,上海生命科学研究院药物研究所建立了包括 7 万多个天然产物三维结构、植物来源和中药药性等信息。过程工程研究所建立了包括 1 万多海洋天然产物的数据库。

我院还开展了计算机辅助组合化学库的设计以及计算机虚拟筛选的研究,利用小分子与受体的结合能、类药性(包括 LogP、分子量、氢键给体数、氢键受体数以及柔性键数)为打分函数,对虚拟组合化学库进行评价,并利用遗传算法对虚拟组合化学库进行优化。

建议继续加强对虚拟筛选以及天然产物数据库的研究和开发,为我国的新药研究提供服务。虚拟筛选是发现新药先导化合物的重要途径,加强对虚拟筛选的研究,减少误筛和漏筛,是提高药物研发效率的重要手段。天然产物是药物先导化合物的主要来源之一,50%左右的药物的先导结构来自天然产物,或其本身就是天然产物。因此,加强天然产物数据库的构建,充分利用已知天然产物的结构信息和生物活性信息,将可加速药物的研究与开发。

建议加强对化合物 ADME/Tox 性质预测方法研究。发展 ADME/Tox 性质预测方法,并应用于药物设计研究,将 ADME/Tox 性质不好的化合物尽早淘汰,或结合构效定量关系研究对药物结构进行改造,设计 ADME/Tox 性质更好的化合物,以减少药物在开发后期失败的风险,对药物研发具有重要意义。

建议加强对小分子与生物大分子结合自由能精确计算方法的研究。结合自由能精确算法是基于结构的药物设计的核心,与药物设计的其它方法密切相关,一旦在这方面有所突破,必将带动整个

药物设计方法的发展。

主要参考文献

1 Kellogg G E, Semus S F. 3D QSAR in modern drug design. EXS, 2003, (93): 223-241.

2 Doman T N, McGovern S L, Witherbee B J et al. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. J. Med. Chem., 2002, 45: 2 213-2 221.

3 Abagyan R, Totrov M. High-throughput docking for lead generation. Curr. Opin. Chem. Biol., 2001, 5(4): 375-382.

4 Taylor R D, Jewsbury P J, Essex J W. A review of protein-

small molecule docking methods. J Comput Aided Mol Des., 2002, 16(3): 151-166.

5 Gane P J, Dean P M. Recent advances in structure-based rational drug design. Curr. Opin. Struct. Biol., 2000, 10(4): 401-404.

6 Joseph-McCarthy D. Computational approaches to structure-based ligand design. Pharmacol Ther., 1999 , 84(2): 179-191.

7 Aqvist J, Luzhkov V B, Brandsdal B O. Ligand binding affinities from MD simulations. Acc Chem Res., 2002, 35(6): 358-365.

The Progress of Drug Design

Xiaomin Luo Hualiang Jiang Jianhua Shen Kaixian Chen
(Shanghai Institute of Materia Medica, SIBS, CAS, 201203 Shanghai)

The progress of drug design in recent years is reviewed in this paper. The prospect of drug design methods in the future is also discussed. Some suggestions are given for the development of drug design in CAS.

Keywords Drug design, Virtual screening, Molecular docking, de novo drug design, Quantitative structure-activity relationship, Pharmacophore model

罗小民 上海生命科学研究院药物研究所副研究员。2002 年毕业于药物研究所有机化学专业,获理学博士学位。从事计算机辅助药物设计研究,利用分子对接法进行酪氨酸激酶、β-分泌酶、钾离子通道和基质金属蛋白酶的抑制剂设计,发现了多个有活性的化合物。发表论文 10 余篇。