

基于蛋白质结构知识的合理药物分子设计*

来鲁华 骆兆文 徐筱杰

(北京大学化学系, 北京 100871)

提 要

利用受体的结构进行合理药物设计是目前国际上药物设计的主要发展方向之一, 综述了有关合理药物设计及全新药物设计的方法, 并以 HIV-1 蛋白水解抑制剂的 design 为例介绍了该方法在新药设计中的成功应用。

关键词 合理药物设计, 蛋白质结构, HIV-1 蛋白水解酶抑制剂

1 药物分子设计的发展

如何研制更加有效的抵抗疾病的药物, 一直是人们面临的巨大挑战之一, 传统的药物开发方式一般是基于对传统药用植物中提取的活性物质的深入研究, 对于现代生理或生化研究中异常发现的追踪 (青霉素就是最著名的例子), 大量的筛选过程, 或者是基于对已知关键性代谢途径或生化过程的了解。显然这几种方法所面临的共同困难在于其作用对象的不可见性, 相当于针对黑箱模型进行研究^[1]。关于受体的有关知识将有助于引导药物设计走向理性化的过程。

由于对相当数量的药物靶蛋白或其它生物大分子 (如 DNA) 三维结构的精确了解, 已经使得基于蛋白质或核酸结构的药物设计成为可能。国际上已经由此发展出来一种新的药物设计方法: 合理药物设计 (rational drug design)^[2]。合理药物设计不仅得助于结构测定及研究手段的发展, 如 X-射线晶体学, 以及近年来发展起来的多维核磁共振技术、计算机图形学及分子力学、分子动力学计算; 也由于基因重组、蛋白测序、分离纯化等基因技术和蛋白质工程的迅速发展, 使得有可能为结构研究提供一定数量生物大分子的纯样品^[2]。

合理药物分子设计是与蛋白质工程相伴而

生的, 蛋白质工程的发展为合理药物设计提供了可能, 而合理药物分子设计则拓宽了蛋白质工程的研究范围。由于蛋白类药物自身的缺点, 也使得相当大的注意力从蛋白质工程开始转向非蛋白质工程 (non-protein engineering)^[3], 其基本的出发点还是利用蛋白质工程的技术来开发非蛋白类的小分子药物。可以说药物设计与蛋白质工程已经出现了一种融合的趋势。

由于药物作用机制的复杂性, 目前与 (设计→合成→具有设计药效) 一次成功的目标还有相当的距离, 令人欣慰的是合理药物设计可以大大减少进行实验药物筛选的数量^[4], 特别是最近在 HIV-1 蛋白水解酶抑制剂设计中的成功应用更加显示了合理药物设计的威力^[5,6]。

2 合理药物设计概况

一般进行合理药物设计首先需要已知药物作用靶蛋白或 DNA 的三维结构, 这可以由 X-射线晶体学、多维核磁共振或同源蛋白质结构预测等计算机模型方法得到; 其次利用实验手段测定或用计算方法研究药物与受体的作用模型; 在此基础上才能进行合理药物设计。如图 1 所示^[2]。

*“863” 高技术项目资助。

收稿日期: 1991-12-30

修回日期: 1992-02-24

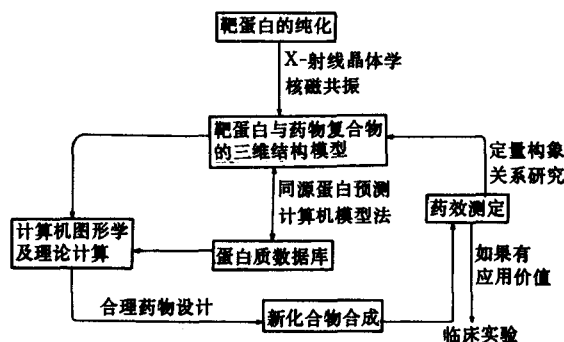


图 1 以蛋白为受体的合理药物设计循环

从图 1 可以看出合理药物设计的关键在于研究药物与受体作用的结构模型及机制。除了可以利用 X-射线晶体学直接测定药物与受体复合物的晶体结构外，也发展了多种研究受体与药物相互作用模型的算法与程序。但是这种合理药物设计的过程同样未能突破药物设计中的一大限制：由于基于研究已知药物与受体的作用，或者本身就是已知的作用于该受体的药物，或者是由与该受体相关的蛋白或其它生物分子的抑制剂转推而来的，所以往往受已知药物模式的限制，较难实现突破而找到完全新的药物。有关根据受体结构模型进行新型药物设计的方法将在下文加以介绍。

合理药物设计往往不是一次循环就可以成功的，常需要 10—15 次循环的努力才有可能达到目的^[4]。除了考虑药物与受体的作用，还需要考虑药物的溶解性、毒性、代谢过程等因素的影响，有关前药物（prodrug）的研究也表明了这一点^[7]。以下是 Abbott 公司利用合理药物设计方法设计抗高血压药物——renin 抑制剂的一个实例^[3]。

renin 抑制剂的设计：

已知条件：一个血管紧张素肽可以与活性 renin 相结合。

未知条件：renin 结构。

a. 利用已知的同源蛋白结构数据预测 renin 空间结构。

b. 检查模型中何处为可能的结合位点。

c. 同时进行合成工作。

d. 对合成的化合物在计算机上进行预筛选，发现有 2/3 不可能与 renin 模型相匹配。

e. 取剩余的 1/3 化合物进行药效实验。

f. 进一步的改造，增加溶解性。

g. 初步的临床实验。

3 基于受体结构模型的全新药物设计

如何突破现有药物框架的限制为基于受体结构模型的合理药物设计提出了一个挑战性问题。由于药物与受体空间结构及其物理化学性质上的互补性，一种合理的设想就是从受体的三维结构反推出药物的形状及化学式来。由于算法上的困难，目前成功的方法并不多。下面仅就作者了解到的两种解决方法加以介绍。

3.1 自动定位药物设计^[8]

自动定位药物设计采取利用某种平面型空间骨架与受体中的可能配位点相匹配，裁剪出合适的分子骨架，再配以匹配的化学基团的方式，直接从靶蛋白的结构自动构建出药物的结构与化学式。

3.1.1 受体表面配位区域和结合位点的确定

以寻找氢键区域为主。利用蛋白质的结构数据自动寻找出蛋白质中的氢键形成基团，确定其中的氢键区，将可形成内氢键的区域去除。然后再确定溶剂可及表面，最后得到蛋白质表面的可能配位点。

3.1.2 建立空间骨架（spacer skeleton）

考虑到计算的方便性，选用平面型的空间

骨架, 骨架由点和边组成, 其大小可以容纳多个配位点, 可以同时模拟多个环系统, 并能代表多个不同的结构见图 2。

3.1.3 分子模板的产生

首先求得由配位点构成的最小二乘平面, 在此平面内对空间骨架与配位点进行拟合, 然后将伸入可及表面的顶点剪裁掉, 即可得到分子模板。

3.1.4 分子的构成

分子模板只是一个具有正确几何构象能与受体配位的分子模型, 还不是真正的分子, 需要按一定的化学规则接上化学基团。

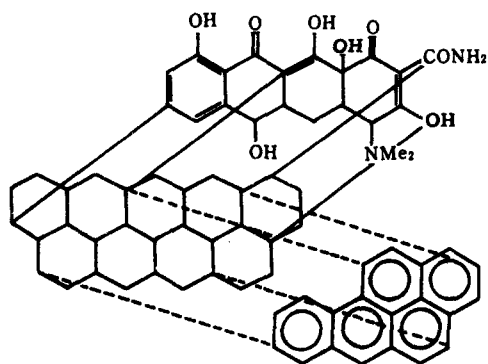


图 2 由六边形构成的空间骨架
可以产生上、下两种分子

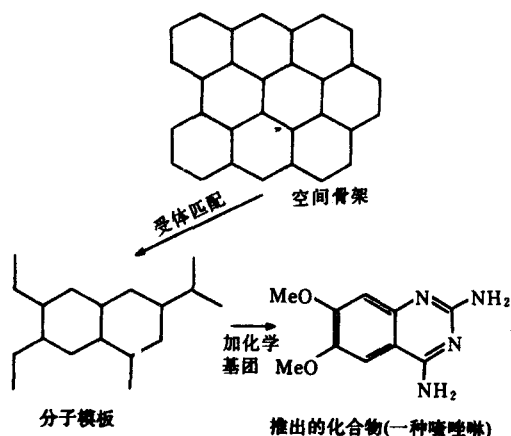


图 3 由空间骨架到最终化合物的过程

自动定位药物设计利用受体的结构进行直接设计, 具有直接准确的特点。但目前的方法步骤多、计算量大, 且只能设计出平面型化合

物。在二氢叶酸还原酶抑制剂的设计中得到的分子模板与已知的药物很接近, 但到目前为止尚未见到在实际药物设计中应用的例子。

3.2 数据库算法^[9-11]

直接从受体结构反推出药物分子从计算方法来看还存在着相当大的困难。数据库算法就是一种绕过这个障碍而达到目的的方法。其最大的特点是借助于剑桥晶体结构数据库(CSD)中收集的小分子晶体的数据逐一与受体进行匹配试验, 筛选出几何形状较好的系列分子, 再进行物理化学性质改造, 从而设计出药物分子, 其基本假定是 CSD 中已经包含有足够的分子形状。

该方法的原理于 1982 年^[9]就提出来了, 但直到其在 HIV-1 蛋白水解酶抑制剂设计中的成功应用才引起了人们的足够重视^[6]。数据库算法大致分为以下几个步骤:

- 根据已知靶蛋白的空间结构产生出活性部位的互补结构。
- 利用小分子晶体结构数据库(CSD)组成分子形状数据库(一般包括 10 000 个分子)。
- 将互补结构与形状数据库中的所有分子进行结构对比、打分, 按分数高低排列。
- 对分数位于前数位的分子按照物理化学性质进行筛选。
- 进一步的改造。

数据库算法的不足之处在于尚不能很好地对形状合适的分子进行调整改造, 使之在疏水作用、氢键形成、静电作用等诸方面与受体完全匹配, 还有待于进一步的改进与完善。

4 未知受体结构的合理药物设计

合理药物分子设计必须在已知受体结构模型的前提下才能进行。但到目前为止许多已知药物作用的受体结构是未知的。在未知受体结构时应用合理药物设计的原理和概念开始药物设计也有了不少的尝试, 这方面的研究大致可分为两类: 探索系列小分子药物三维结构与活性的关系——3D-QSAR (quantitative structure activity relationship)^[12]; 根据已知药物结构反

推受体结构模型,再行合理药物设计。

4.1 三维构效关系研究 (3D-QSAR)

从对药物与受体相互作用的研究可以知道药物的作用是依赖自身空间形状的,其与受体的作用一般为非共价性质。虽然在未知受体结构时无法进行常规意义上的合理药物设计,但可以在对已知药物研究的基础上进行受体形状推测 (receptor-mapping)^[13],将与药物本身形状有关的参数引入到定量构效关系 (QSAR) 中,称之为 3D-QSAR。3D-QSAR 有分子形状分析^[14]、距离几何方法^[15]、分子场比较分析 (CoMFA)^[16]等,其中方法新颖极有应用价值的 CoMFA 方法倍受注目。

CoMFA 的主要原理是: a. 先对系列药物进行活性构象分析,按一定规则进行分子叠合; b. 在分子上加上足以环绕所有化合物的三维格点,格点间距一般为 2.0 Å; c. 以 sp^3 杂化的带正电荷的碳原子为探针 (probe atom) 在各格点上分别逐一求算探针与药物分子的范氏作用能及静电作用能; d. 以线性方程方式列出生物活性与这数个参数间的方程; e. 利用部分最小二乘方法 (partial least-squares method) 对这个变量数大于已知条件的方程进行分析; f. 结果的图象表达。

CoMFA 方法在预测未知化合物的活性方面取得了成功。为在未知受体结构的条件下,利用合理药物设计的概念,进行三维定量构效关系研究和药物设计提供了可行的方法。如果可以将 CoMFA 方法与前面介绍的数据库算法相结合,以 CoMFA 中分子叠合出的形状为引导在数据库中进行搜寻,再根据 CoMFA 的结果对分子进行改造,可以而拓宽这两种方法的应用范围。

4.2 蛋白质准受体 (protein pseudo-receptor)

既然受体的结构知识可以对药物设计工作起重要的指导作用,一种大胆的设想就是可以根据药物的结构人工构建出蛋白质的受体——当然只能称之为准受体^[17]。

构建准受体的做法一般是先用聚丙氨酸多

肽链折叠出一个与底物结合的口袋;然后对底物周围的残基进行替换,使得突变后的氨基酸侧链能与底物形成合适的疏水作用、氢键或静电作用等。由此建立起一个蛋白质的准受体。

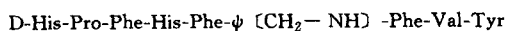
虽然这种方法无实验证明成功的先例,但由于全新蛋白质设计的成功^[18],人工设计的氨基酸序列已经可以按预期的方式折叠起来,而且成功地进行了表达,这使得人工设计蛋白质准受体完全有可能被合成或表达出来,从而在实验上制造出来人工的蛋白类药物受体。可以预见此类蛋白准受体可能会具有某些实用价值。

5 HIV-1 蛋白水解酶抑制剂的合理设计

人体获得性免疫缺损病毒 (HIV) 是引起爱滋病的主要因素, HIV 的基因产物是很好的药物作用靶子, HIV-1 蛋白水解酶就是其中之一。HIV-1 蛋白水解酶是由两个含 99 个氨基酸的单体构成的天冬氨酸水解酶,其晶体结构已经被测定。有关 HIV-1 蛋白水解酶抑制剂设计的工作很多^[5],多数是在其它天冬氨酸蛋白水解酶抑制剂的基础上进行改造。在此我们介绍两个合理药物设计在 HIV 蛋白水解酶抑制剂设计中成功应用的实例。

5.1 一种具有 C2 对称性的抑制剂^[19]

由于 HIV-1 蛋白水解酶的活性部位具有二重轴对称性,合理的设想是有同样对称性的抑制剂将会有高抑制活性。现有的天冬氨酸蛋白水解酶抑制剂一般为多肽,但多肽类的抑制剂由于其本身的缺点很难开发成有效的药物,因此 Abbott 公司的 Erickson 小组的研究人员设想开发一种对称性的多肽类似物,从 RSV 蛋白水解酶抑制剂出发在 Phe 上还原的 CH_2 处切断,将 N 端的分子片断按二重轴转 180 度,构成一个对称的分子,简化为 A-74702。



↓



实验测定 A-74702 抑制活性不高。从抑制剂与酶作用的情况看,抑制剂的 P_1 , P_2 位置位

于酶活性口袋内, P_3 就有可能暴露在外边了(见图 4). 将 A-74702 加上 Val 扩展为 A-74704, 实验测定结果 A-74704 的抑制活性很

高, 而且对 HIV-1 蛋白水解酶的选择性比其它相关的酶要高 10 000 倍.

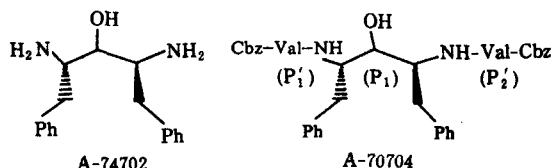


图 4 化合物 A-74702 及 A-74704 的结构式

5.2 数据库搜寻法在 HIV-1 蛋白水解酶抑制剂设计中的应用^[20]

为了寻找非肽类的抑制剂, 美国加州旧金山分校药化系的 I. D. Kuntz 等研究人员利用前面介绍的数据库算法进行了新型 HIV-1 蛋白水解酶抑制剂的设计. 利用 HIV-1 蛋白水解酶的晶体结构数据推出该酶活性部位的互补结构, 在由 CSD 构成的分子形状数据库中进行逐一对比、叠合、打分, 按分数高低排列, 然后对前 200 位的化合物按以下三条规则进行筛选: a. 分子中至少有一个原子在酶的 Asp-25 或 Asp-25' 侧链的羰基氧原子附近 (4 Å 以内); b. 与蛋白质内表面成氢键的能力; c. 较易于合成.

筛选后从中选出了 bromperidol, 其 OH 恰好可以与酶活性部位的天冬氨酸作用, 与 penicillopepsin 与 statine 类抑制剂在复合物晶体结构中的作用方式类似.

利用商品化的 haloperidol (HAL) 及还原的羟基-HAL 进行了抑制剂活性测定, 结果表明 HAL 对 HIV-1 蛋白水解酶有抑制活性, 且具有很高的选择性. 但该化合物只有在很高浓度时才有抑制活性, 尚需要进一步的改造. 该实例表明数据库方法在寻找先导化合物方面是成功的, 当然进一步的改造工作是必不可少的. 从 HIV-1 蛋白水解酶抑制剂设计这两个成功的实例可以看出合理药物设计在药物的设计与开发, 特别是在新药的开发中极具应用价值, 可以期待合理药物设计方法将会在药物工业中起

到越来越重要的作用.

参 考 文 献

- Hol W G J. Protein crystallography and drug design. *NA-TO ASI Ser Ser A*, 1987; **126**: 223
- Hol W G J. Protein crystallography and computer graphics-toward rational drug design. *Angew Chem Int Ed Engl*, 1986; **25**: 767
- Knight P. Non-protein engineering; small drug design. *Biotechnology*, 1990; **8**: 101
- Hodgson J. Data-directed drug design. *Biotechnology*, 1991; **9**: 19
- Blundell T L, Lapatto R, Wilderspin A F *et al.* The 3D-structure of HIV-1 proteinase and the design of antiviral agents for the treatment of aids. *TIBS*, 1990; **15**: 425
- Stewart K. New shapes in HIV protease inhibitors. *Protein Engineering*, 1990; **4**: 1
- Hiller S. Prodrugs against cancer. *Science* 1991; **253**: 1095
- Danziger D J, Dem P M. Automated site-directed drug design. *Proc R Soc Lond*, 1989; **B236**: 101, **B236**: 115; **B236**: 141
- Kuntz I D, Blaney J M, Oatley S J *et al.* A geometric approach to macromolecule-ligand interactions. *J Mol Biol*, 1982; **161**: 269
- DesJarlais R L, Sheridan R P, Dixon J S *et al.* Docking Flexible ligands to macromolecular receptors by molecular shape. *J Med Chem*, 1986; **29**: 2149
- DesJarlais R L, Sheridan R P, Seibel G L *et al.* Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J Med Chem*, 1988; **31**: 722
- Martin Y C, Danaher E B, Man C S *et al.* Strategies in drug design based on 3D-structures of ligands. In: Fauchere I *et al.* eds, *QSAR: Quantitative structure-activ-*

- ity relationships in drug design, New York; Alan R Liss INC, 1989; 177
- 13 Ohta M, Koga H. Three-dimensional structure-activity relationships and receptor mapping of N_1 -substituents of quinolone antibacterials. *J Med Chem*, 1991; **34**: 131
- 14 Hopfinger A J. A QSAR investigation of dihydrofolate reductase inhibition by baker triazines based upon molecular shape analysis. *J Am Chem Soc*, 1980; **102**: 7196
- 15 Sheridan R P, Nilakantan R, Dixon J S *et al.* The ensemble approach to distance geometry; application to the nicotinic pharmacophore. *J Med Chem*, 1986; **29**: 899
- 16 Cramer I R D, Patterson D E, Bunce J D. Comparative molecular field analysis (CoMFA). 1. effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc*, 1988; **110**: 5959
- 17 Momany F A, Pitha R, Klimkowski V J *et al.* Drug design using a protein Pseudoreceptor. *ACS Symp Ser Expert Syst Appl Chem*, 1989; **408**: 82
- 18 Regan L, Degradó W F. Characterization of a helical protein Designed from first principles. *Science* 1988; **241**: 976
- 19 Erickson J, Neidhart D J, VanDrie J *et al.* Design, activity, and 2.8 Å crystal structure of 2 C₂ symmetric inhibitor complexed to HIV-1 protease. *Science*, 1990; **249**: 527
- 20 DesJarlais R J, Seibel G L, Kuntz I D *et al.* Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proc Natl Acad Sci USA*, 1990; **87**: 6644

蛋白质和多肽 C 端氨基酸序列研究进展

屠红旻 夏其昌

(中国科学院上海生物化学研究所, 上海 200031)

提 要

C 端序列测定长期以来是蛋白质化学领域中的一大难题, 但是生物化学与分子生物学中的许多研究表明了 C 端序列的重要性. 文中介绍了 C 端测定的一些最新技术, 尤其侧重于化学降解法, 特别是运用商品化的 N 端分析仪器分析 C 端序列, 增强了仪器的通用性. 文中还推荐了几种 C 端标记进而分离 C 端肽段, 然后用 Edman 化学法测定 C 端序列的简便方法, 有一定的可行性.

关键词 C 端序列分析, C 端标记, 硫氰酸盐

随着分子生物学研究的深入, 蛋白质及多肽氨基酸序列研究的地位越来越重要. 人们往往要从几十乃至几微克蛋白质样品中获得蛋白质一级结构信息, 以研究其结构功能关系, 或更多的是由蛋白质氨基酸序列制备 DNA 探针, 筛选其 cDNA, 进而进行基因重组工作或表达蛋白质产物. 另外, cDNA 预测蛋白质序列不能完全代替经典的蛋白质分析方法, 尤其对于经过磷酸化、糖基化和酰胺化等后加工的蛋白质. 因此蛋白质的微量序列测定一直是重要研究课题.

N 端氨基酸序列测定应用 Edman 化学降解法已经能达到高度自动化、微量 (pmol) 水平, 可以分析蛋白质 N 端 20—40 或更多氨基酸残基, 小于 50 个氨基酸残基的多肽则可分析完全. 但有不少蛋白质尤其在哺乳动物及高等植物中 N 端被修饰而封闭, 且封闭的方式也不尽相同, 如焦谷氨酰化、甲酰化、乙酰化或脂肪酰化等, 给蛋白质 N 端序列测定带来了困难, 而 C 端羧基被修饰的情况除酰胺化外, 很