

·综述·

CADD 的新进展——基于受体结构的药物设计

杨洁 华维一 彭司勋

(药物化学研究室)

Computer-Aided Drug Design in Prospect: Receptor's Structure-Based Drug Design

Yang Jie, Hua Weiyi, Peng Siuxun

摘要 较详细介绍了计算机辅助药物设计的一种新方法——基于受体结构的药物设计(SB-DD),即将某些生物大分子(如核酸、酶和蛋白质)的结构信息,与计算机技术结合起来进行药物设计。通常,新药研究包括先导化合物的产生与优化两个方面。目前,计算化学和分子模拟技术已成功地应用于先导化合物的优化,本文则主要叙述基于受体结构的药物设计思想在新药研究中的应用。

关键词 基于受体结构的药物设计;三维结构;先导化合物;计算机图形学;CADD;反义RNA

1 前言

计算机辅助药物设计(Computed-Aided Drug Design, CADD)是一门边缘性很强的新兴学科,它利用电子计算机技术在分子水平上研究药物结构与生物活性之间的关系,并指导合成药物或修饰已知药物,以提高活性。

新药研究主要包括两个环节,即先导化合物的获取与优化。按照对受体结构的了解情况,先导化合物设计可分为直接设计和间接设计^[1](图1)。在已知受体三维结构的情况下,运用定向设计原理,根据受体结构要求,通过计算机图形学的研究直接设计与受体相适应的化合物分子。例如,作用于二氢叶酸还原酶(DHFR)的抗菌药物和作用于血红素的抗镰状细胞贫血化合物的设计,可谓直接设计。间接设计是在受体结构未知的情况下,通过分析已知受体配基的构效关系,假设药物分子的可能作用机理,推导先导化合物。因为与受体结合的配基化合物,其结构可能差异很大,但可通过计算机找出其共同的三维结构形状,即药效构象^[2](Pharmacophore)。以药效构象为

基础可设计出先导化合物。例如,从5-HT_{1A}受体的不同配基化合物(1~4)中找出共同的三维结构(即Pharmacophore,图2),再以此为基础设计出新的5-HT_{1A}受体拮抗剂(5,6)。

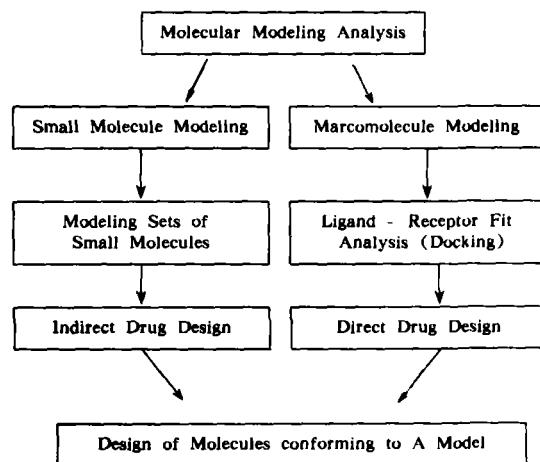


Fig 1. The cycle of steps in drug design^[1]

在先导化合物优化方面,计算化学和分子模拟方法^[3]的应用已有成功的先例,如安定(7)和

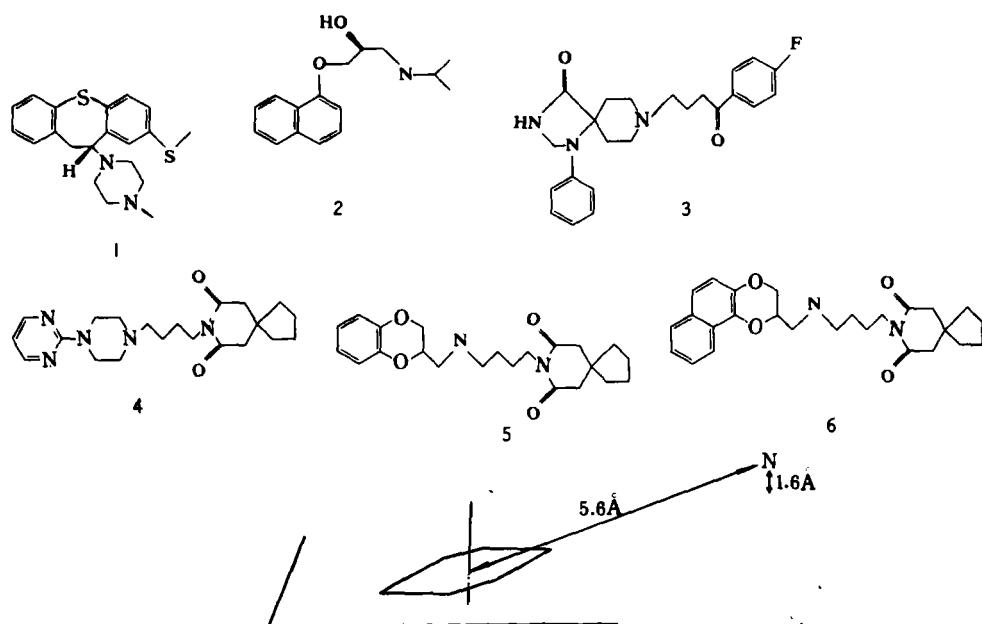
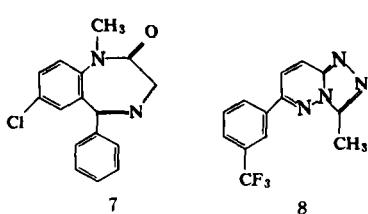


Fig. 2. Basic pharmacophore of the 5-HT_{1A} antagonist recognition site^[2]

CL218872(8),虽然它们的结构差异很大,但通过计算它们的表面静电势非常相似,后者可望成为一类新型抗焦虑药物。这些技术既能装配、分析和贮存复杂分子的结构模型,并具可视性,而且可进一步阐明化合物的构效关系,从而指导新化合物的设计。

基于受体结构的药物设计属直接设计范畴,与前者不同的是从一个崭新的角度——受体的结构来研究新药。这个新思想正在基于受体结构药物设计中发挥出巨大的作用。下面以其研究方法和应用两方面加以综述。



一般情况下,仅仅知道受体的结构,这对判断配基与受体的作用方式仍然十分困难。例如,有时与配基结合时蛋白质的构象会发生变化。因此,仅根据蛋白质的结构进行药物设计会导致错误的结论。现时,有下列研究方法用于探索配基与受体作用方式。

通常采用结合能计算法研究配基与受体结合方式。

2 基于受体结构药物设计的研究方法

基于受体结构药物设计是将某些关键的生物大分子(受体)的结构信息与计算机图形学结合起来进行药物设计,强调受体与药物之间的立体互补性和化学互补性,立体互补性是基于受体结构药物设计的理论基础。而受体-配基复合物结构的确定是基于受体结构药物设计的关键环节(图3,图4)^[4,5]。

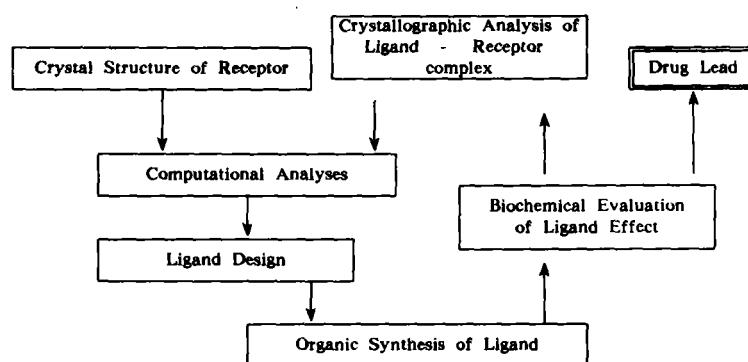
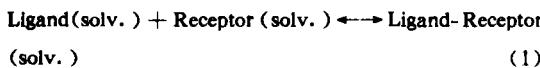


Fig. 3. An iterative cycle for the discovery and elaboration of lead compound^[4]

如方程(2)所示,配基与受体的结合能是由溶液中配基和受体的总能量与溶液中配基-受体复合物的能量差而计算出来的。由于大部分配基的结合能较小,而体系(包括溶剂)较大,并具有构型复杂性,由统计方法处理,因此,常会引起误差。

为了减少误差,目前广泛采用的是自由能微扰法^[3],该法试



$$\Delta G(\text{binding}) = \Delta G(\text{Ligand-Receptor}) - [\Delta G(\text{Ligand}) + \Delta G(\text{Receptor})] \quad (2)$$

图计算相似配基结合能的变化值($\Delta\Delta G$),而不是计算其绝对结合能(图5)。该法已成功地用于计算结构差异较小的配基结合能,而且,计算的结合能变化值($\Delta\Delta G$)与其实验值十分接近。由此可见,根据 $\Delta\Delta G$ 值可大致推测蛋白质结构的变化程度。

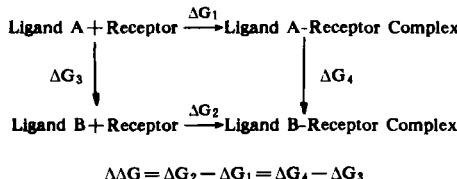


Fig 5. Free-energy perturbation thermodynamic cycle^[5]

亦可采用“探针法”(probe method)、“模板法”(template method)和“筑块法”(building blocks method)来探求先导化合物。“探针法”把不同的探针(如CH₃或OH)放在受体中的不同位置上,通过计算它与受体的结合能找出“热点”,用键连接产生分子的结构形状,并保证分子处于能量最低状态。“模板法”将数据库中已知分子结构与受体位点空间结构进行比较,根据立体互补性寻找到有效的化合物。“筑块法”则试图让分子在受体域内“生长”(类似于种子发芽生长),通过逐个递增原子建筑化合物分子,推导新化合物^[8]。通常,这些方法是相辅相承的,并具有一定应用价值。

对于受体等生物大分子结构的测定比较困难,通常采用X-衍射结晶学技术来确定,并利用NMR中NOE技术推测溶液中受体的构象。NOESY谱中的交叉峰提供了分子中质子的偶极交叉驰豫信息,反映出一些原子之间空间距离的远近程度。而且,前者是指晶体结构,与生物活性状态下的结构可能不一

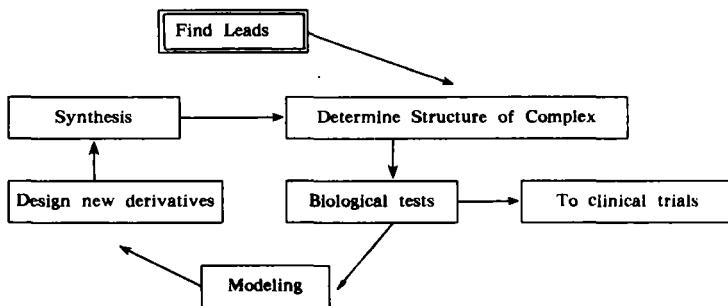


Fig 4. The cycle of steps in structure-based drug design strategy^[5]

样,后者测得的结果代表生物环境下的结构,因此,后者比前者更为有用。上述两种技术不仅能够提供受体分子的结构信息,而且还可能提供受体-配基复合物的结构信息。

近来已创立了所谓用神经网络方法^[7]来预测蛋白质的二级结构。该方法模拟神经元的功能接收信息并输出结果,根据蛋白质的氨基酸序列直接预测其二级结构,精确度可达63.5%,其α-螺旋、β片层和卷曲(coil)的相关系数C_α、C_β和C_{coil}分别为0.40、0.30和0.40。神经网络是一种建立在人类大脑理论模型基础上的计算系统,适合处理那些不能简单或确切定义的数据关系。在神经网络中,很多类似神经元的简单处理单元相互连接,形成高度并行的结构。而以反向传播算法为基础的多层网络^[8]是目前运用较广的神经网络模型,可用于分析构效关系以及预测未知化合物的活性等方面。

3 基于受体结构药物设计的应用

基于受体结构的药物设计离不开计算机的帮助。例如,在有关研究的基础上,将QSAR与计算机图形学相结合的三维QSAR及上述有关方法等。最近,已开发出DOCK软件包^[9],对阐明受体-配基结合模型十分有利,主要在药物设计的三个方面,即产生受体靶域的虚拟构象,模拟配基进入靶位点过程,以及根据受体与配基的吻合程度设计新化合物等。DOCK不仅适用于蛋白酶的结构研究,而且也可推广应用到核酸结构、病毒衣壳蛋白以及蛋白质-蛋白质相互作用方式等研究领域^[10]。

3.1 以蛋白质为受体的药物设计

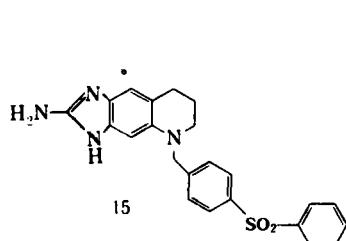
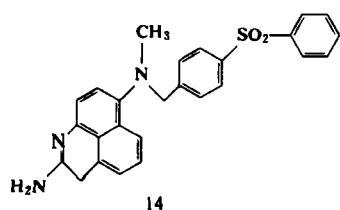
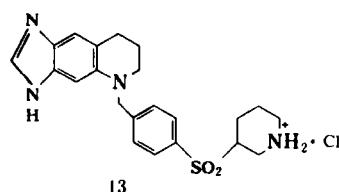
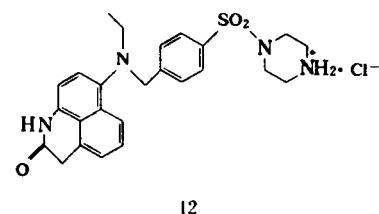
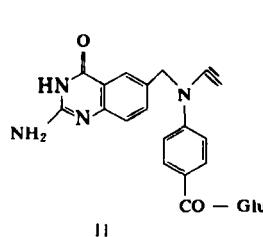
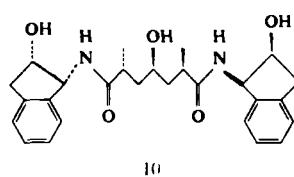
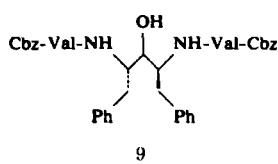
3.1.1 HIV-1蛋白酶抑制剂的设计 HIV-1是导致AIDS病的主要原因,一旦病毒部分地侵入正常细胞,则随逆转录酶的启动诱发产生若干蛋白质和酶,它们能干扰宿主细胞的正常机理,发生病变。研究表明,HIV-1蛋白酶由99个氨基酸残基组成,该酶抑制

剂必须具有渗透细胞膜的能力。研究发现^[11],若干ACE抑制剂和肾素抑制剂对HIV-1蛋白酶具有较强的拮抗活性,这表明该酶与ACE、肾素具有部分相似的活性域。通过对HIV-1蛋白酶与天门冬氨酸蛋白酶的结构进行比较,并借助X-衍射波谱的结果,人们获得了高精度(1.8 Å)的HIV-1蛋白酶的三维结构,并建立起该酶的结构模型。随后,以该酶模型为靶,将已知的ACEI类抗高血压药物与之契合,通过结构修饰不断提高吻合程度。根据“浸透”理论,用水分子模拟包围该复合物,利用分子力学计算对整个集合体构象进行优化,借助计算机图形学分析判断酶与抑制剂的作用方式。结果发现,两者通过六个氢键和一个疏水域相互作用。随后,设计几种新结构分子,使之与模型蛋白酶受点吻合,再将这些分子的图形重迭,找出药效构象。

同时,运用CADD建立的HIV-1蛋白酶的脂质双分子层细胞膜模型,借助分子动力学及LogP值判断化合物是否具有渗透膜的能力。这在基于受体结构的药物设计中是必不可少的。并藉助药代动力学预测化合物的生物利用度提供的可行性建议,最终设计出结构近乎对称的HIV-1蛋白酶抑制剂(9,10)。

3.1.2 胸腺嘧啶核苷酸合成酶抑制剂的设计 在

5,10-乙撑四氢叶酸存在的条件下,胸腺嘧啶核苷酸合成酶(TS)能催化脱氧尿苷酸甲基化转变成胸腺苷酸。抑制该酶可阻断关键的生化代谢途径,达到抗肿瘤目的。由于E. coli TS与人的TS具有较高同源性(46%),而且结合位点的同源性达75%,因此,在人的TS结构未知时可由E. coli的TS代替,根据E. coli的TS结构进行药物设计。文献报道^[11],某个三元复合物(即E. coli TS-5-氟尿苷酸-化合物11)的三维结构已知,该酶通过氢键、疏水键和静电作用与化合物11相互作用。在该复合物模型的基础上,运用计算机图形学、探针法及等势面计算法,设计了两种该酶抑制剂(12,13),结构分析表明,这两类化合物通过一个疏水区和两个氢键与酶相互作用,即苯环与酶的Ile 79形成疏水作用,化合物12中的内酰胺NH和CO分别与酶的Asp169和Ala263形成氢键(距离分别为2.9 Å和2.8 Å);类似地,化合物13中的N₃和N₁分别与Asp169和Ala263形成氢键(2.9 Å和2.8 Å)。接着,以此为先导,不断进行结构修饰,得到两种新型TS酶抑制剂(14,15)。药理实验结果显示,化合物14和15对小鼠L1210的抗肿瘤活性分别比化合物12和13提高16倍和10倍。



3.1.3 抗多重耐药性药物的设计 多重耐药性蛋白质(MDRP)的结构特征是人们非常感兴趣的研究领域^[12]。这种被称为P-170的大分子糖蛋白存在于癌细胞膜上,也存在于其它正常的上皮组织中,如肝、肾、肠、胎盘、肾上腺体、子宫等。该糖蛋白由mdrl基因编码,分子量为141 KDa,含有两个纵列的重复部分,由N端疏水域(具渗透细胞膜作用)和C端亲水域(具ATP结合位点)组成(图6)。而且,P-170与大多数细菌的转运蛋白具有大量相同的氨基酸片段,即片段同源性。这些转运蛋白包括大肠杆菌的 α -溶血素输出泵和革兰氏阴性菌的“震颤敏感”营养吸收蛋白等。同时发现,多重耐药性蛋白质的底物和抑制剂与P-170结合在同一位置上,即疏水性结合位点。可见,当mdrl基因被启动时,所表达的产物能活跃地将药物泵出细胞外,导致耐药性出现。因此,P-170是一种ATP-水解转膜泵(ATP-hydrolysing transmembrane pump),能增强细胞对药物的耐受性。而使mdrl基因表达受阻,可望找到减少细胞对药物耐受性的新药。

另外发现,在小鼠和人的细胞中,随着谷胱甘肽

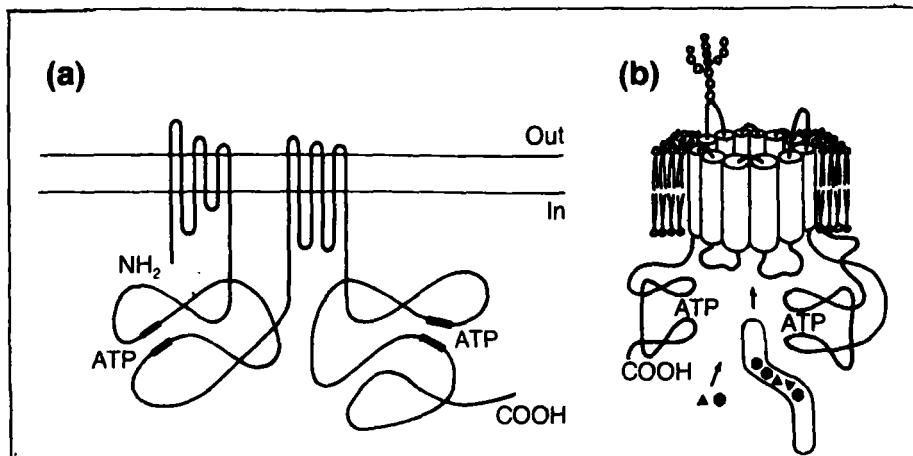


Fig 6. The multi drug resistance pump^[12]

(a) Hypothetical transmembrane folding and ATP binding sites based on primary sequence. (b) Possible three-dimensional arrangement showing glycosylation site and two possible pathways for drug exit, either free or on a hypothetical carrier protein.

S-转移酶(GST)浓度的升高,多重耐药性蛋白质含量也升高;对烷化剂有耐受性的肿瘤细胞中,谷胱甘肽(GSSG)的浓度明显升高,类似的情况也出现在耐阿霉素的人卵巢癌细胞中。由此可推测,GST与多重耐药性具有一定的相关性,多重耐药性蛋白质可能是一种与谷胱甘肽相连的疏水性药物泵。因此,多重耐药性蛋白质抑制剂的药效构象可能包括芳香环、羰基以及 α 、 β -不饱和双键等。

3.2 以核酸为受体的药物设计

随着分子生物学的不断发展,对DNA的双螺旋结构的认识越来越清楚。运用扫描显微镜技术^[13](Scanning Force Microscopy)已探明其三维结构,这为基于受体结构药物设计提供了明确的靶受体。

3.2.1 抗肿瘤化合物的设计 关于肿瘤的癌变机理,人们普遍认为是由于基因在突变导致基因表达失调和细胞无限增殖所引起的。因此,可将癌基因作为药物设计的靶受体,利用反义技术^[14](antisense technology)达到抑制细胞增殖的目的。反义技术是指

根据所调控的RNA序列产生与之互补的RNA序列的技术,该互补的RNA序列即为反义RNA。目前,正从分子生物学的角度来研究癌基因所转录表达的反义RNA,通过逆转录病毒为载体将它引入癌细胞中,能明显地抑制癌细胞增殖。例如,通过痘苗病毒为载体,将ras癌基因的反义RNA引入患肝癌的鼠体中,可抑制细胞增殖,鼠存活率可达30%~50%。可见,以此为突破口,合成短核苷酸(即反义RNA),可望成为新型抗肿瘤药物。

以已知的抗肿瘤药物为先导,运用CADD技术设计新的抗癌药物也正在开展。现已知道诺霉素和阿霉素等抗癌药是以嵌入或塞进的方式与DNA分子相互作用(图7),分子的体积不能太大,以直径为120 Å的球体为最佳。以此为依据设计的新化合物将有利于化合物插入DNA分子中,破坏DNA结构,干扰其基因表达过程,达到抗肿瘤目的。

3.2.2 抗真菌药物的设计 真菌的细胞、代谢物以及基因表达产物都是导致疾病的原因之一,然而在

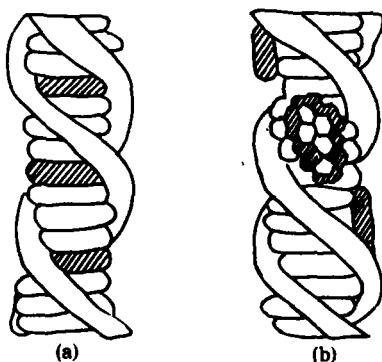


Fig. 7. Binding between DNA and antitumor drug^[15]
(a) intercalation site (b) stacking

哺乳类宿主内没有其相应的副本。因此,寻找新型靶受体对于设计新型抗真菌药物十分重要。目前临幊上使用的抗真菌药物有一类含吡咯环的化合物,如酮康唑、依康唑和氟康唑等。其作用机理是通过与细胞色素 P-450 形成复合物,从而抑制真菌中的依赖于 P-450 的 14α-甲基化酶,达到广谱抗真菌活性。

开发新型抗真菌药物的关键在于寻找能抗衡致病因子(virulence factors)的药物^[16]。真菌在感染宿主时,通过甘露糖蛋白将真菌细胞连接在宿主的上皮细胞上,同时分泌天冬氨酸蛋白酶。研究发现,天冬氨酸蛋白酶的参与能增强真菌的致病能力。可见,这种分泌的天冬氨酸蛋白酶(Secretory Aspartic Proteinase, SAP)可能是一种致病因子。由此推測可运用分子生物学方法找出编码该蛋白酶的基因组,以此为靶设计化合物干扰基因表达,达到抗真菌目的。

4 展望

综上所述,已知配基的结构或活性信息、靶受体及相关靶受体的结构,以及受体配基复合物的结构,对发展 CADD 是很重要的。目前,基于受体结构药物设计的研究领域已拓展到上述多领域的研究。在这些领域中普遍存在药物的靶受体,如病毒的穿衣-脱衣过程(cating-uncoating)、酵母的 mating factor、寄生虫的特异酶以及 DNA 的小沟处的 AT 碱基片段^[17]等。另外, G-蛋白耦联受体^[18]也可作为基于受体结构药物设计的靶受体,例如,以 β_2 -肾上腺能受体^[19]为靶设计一类新型 β_2 -阻断剂,可作为抗高血压药物。同时,抗原也可作为基于受体结构药物设计的靶受体^[20,21],例如,以肿瘤抗原为靶,其单克隆抗体直接与该抗原结合,激活体内免疫系统,可达到杀死癌细胞的目的。然而,由 CADD 获得的大部分化合物不

可能全部成为药物。只有不断提高 CADD 水平,不断总结经验,才能提高药物设计的命中率。

参考文献

- Cohen NC, Blaney JM, Christine Humbert, et al. Molecular modeling software and methods for medicinal chemistry. *J Med Chem*, 1990;33(3):883
- Hibert MF, Gittos MW, Middlemiss DN, et al. Graphics computer-aided receptor mapping as a predictive tool for drug design: development of potent, selective and stereospecific ligands for the 5-HT_{1A} receptor. *J Med Chem*, 1988;31:1087
- Frank Blaney. Molecular modeling in the pharmaceutical industry. *Chem & Ind*, 1990;23(12):791
- Appelt K, Bacquet RJ, Bartlett CA, et al. Design of enzyme inhibitors using iterative protein crystallographic analysis. *J Med Chem*, 1991;34(7):1925
- Dixon JS. Computer-aided drug design: getting the best results. *TIBTECH*, 1992;10:357
- Yoshihiko N, Akiko I. Automatic creation of drug candidate structures based on receptor structure. Starting point for artificial lead generation. *Tetrahedron*, 1991;47(43):8985
- Paul S, Alan L, Yuan X. Predicting protein secondary structure using neural net and statistical methods. *J Mol Biol*, 1992;225:363
- 孙 喻,甘 强,韦 钦. 基于多层网络的构效关系分析. 药学进展,1993;17(1):10
- Irwin DK. Structure-based strategies for drug design and discovery. *Science*, 1992;257(11):1078
- Shoichet BK, Bodian DL, Kuwitz ID. Molecular docking using shape descriptors. *J Comput Chem*, 1992;13(3):380
- James P. Computer-assisted drug design. part 1. conditions in the 1980s. *Med Res Rev*, 1991;11(6):641
- West IC. What determines the substrate specificity of the multi-drug resistance pump. *TIBS*, 1990;15:42
- Eehansern F, Adrian M, Heggeler-Brdier BH, et al. Zmapping of DNA by scanning Force Microscopy. *J Struct Biol*, 1992;108:69
- Bejarano ER, Lichtenstein CP. Prospects for engineering virus resistance in plants with antisense RNA. *TIBTECH*, 1992;10(11):383
- 刘次全. 量子生物学导论. 北京:科学出版社,1989;174
- Fuite MF. Antifungal drug development: the identification of new targets. *TIBTECH*, 1992;10:235
- Brown DG, Sanderson MR, Garman E, et al. Crystal structure of a Berenil-d (CGCAAATTTGCG) complex. *J Mol Biol*, 1992;226:481
- Linden ME, Gilman AG. G Proteins. *Scientific American*, 1992;8:36
- Huss KM, Lybrand TP. Three-dimensional structure for the β_2 -adrenergic receptor protein based on computer modeling studies. *J Mol Biol*, 1992;225:859
- Lobuglio AF, Salen MN. Advances in mono clonal antibody therapy for cancer. *Am J Med Sci*, 1992;304(3):214
- Walls PH, Sternberg MJ. New algorithm to model protein-protein recognition based on surface complementarity. *J Mol Biol*, 1992;228:277