

文章编号: 1006-6144(2009) 04-0473-05

分子对接方法的应用与发展

段爱霞, 陈 晶, 刘宏德, 刘秀辉, 卢小泉*

(西北师范大学化学化工学院, 甘肃兰州 730070)

摘 要: 分子对接方法加快了药物开发周期, 具有快速、准确度高等优点。本文详述了分子对接方法的基本原理, 及分子对接空间和能量的匹配要求和优化时的各种方法。综述了该方法在药物设计、药理分析和探测生命体系等方面的应用。

关键词: 分子对接; 蛋白质; 抑制剂; 药物设计

中图分类号: O657 **文献标识码:** A

1 引言

分子对接方法(Molecular Docking Method)已成为药物设计方法中比较成熟的直接药物设计方法。

已知结构的受体(靶蛋白或活性位点)和配体出发, 通过化学计量学方法模拟分子的几何结构和分子间作用力来进行分子间相互作用识别并预测受体-配体复合物结构的方法称为分子对接。准确预测蛋白质复合物的结构和结合位点在药物设计中显得非常重要^[1]。

分子对接方法最初是科学家在分子水平上研究生物体系中的化学问题时提出的。早在 1894 年, E. Fisher 在他的论文里用“锁和钥匙”的比喻描述酶与底物的专一性结合, 称之为识别。同时, Langley 提出的受体学说认为大多数药物与细胞膜上或细胞内某些特定分子结合, 才能发生效应, 这些特定的分子被称为受体。所以, 受体学说中的受体与分子对接中的受体概念是基本相同的, 为分子对接方法奠定了理论基础。高通量虚拟筛选技术(并行化的分子对接方法)筛选量大, 每天可完成对包含数十万, 甚至数百万个化合物的数据库的测试。通常, 虚拟筛选比经验筛选更加高效率。分子对接技术已成为与高通量筛选互为补充的寻找先导化合物的方法。

2 分子对接的原理

2.1 分子对接的一般原理

分子对接是将已知三维结构数据库中的分子逐一放在靶标分子的活性位点处。通过不断优化受体化合物的位置、构象、分子内部可旋转键的二面角和受体的氨基酸残基侧链和骨架, 寻找受体小分子化合物与靶标大分子作用的最佳构象, 并预测其结合模式、亲和力和通过打分函数挑选出接近天然构象的与受体亲和力最佳的配体的一种理论模拟分子间作用的方法。

2.2 分子对接的互补性

影响复合物分子稳定性的主要因素是疏水作用和键合力大小。Audie 等^[2]建立了一种新的经验自由能公式, 该方法能够预测蛋白质-蛋白质键合力大小, 准确率达到 89%。而影响键合力的因素有作用位点空间位的互补、静电相互作用和氢键等, 且溶解熵对稳定受体-配体复合物起着重要的作用^[3], 所以分子对接的过程主要包括分子间的空间互补和电学性质互补。空间互补是分子间发生相互作用的基础, 能量互补是分子间保持稳定结合的基础。

收稿日期: 2008-05-28 修回日期: 2008-07-23

基金项目: 国家自然科学基金(No. 20775060, 32335030); 教育部第四届高校青年教师教学科研奖励计划(TRAPOYT)

* 通讯联系人: 卢小泉, 男, 教授, 博士生导师, 研究方向: 电化学及化学计量学。

2.2.1 几何互补 目前,最常见的研究受体大分子结合部位三维结构的方法是 X-单晶衍射(蛋白质晶体学),其次是二维核磁共振和多维核磁研究液相结构,电镜三维重组、电子衍射、中子衍射和各种频谱学方法研究生物高分子的空间结构。尽管受体大分子的三维结构能够通过以上方法得到,但当不同的配体靠近受体时,往往会因为分子的柔性而导致受体分子的结构发生变化。Richards^[4,5] 小组运用模糊模拟识别算法解决了上述问题并且提高了搜索速度。

2.2.2 能量匹配 配体与受体能量匹配包括静电作用、氢键作用、范德华相互作用和疏水作用等。其中静电作用、氢键作用和范德华力控制药物-受体结合。疏水作用是药物-受体结合的驱动力。Liu 等^[6] 运用 MOLCAD 程序对六缩氨酸和 α 淀粉酶分子表面分析,揭示静电作用(包括氢键)和范德华力是它们的主要亲和力。为了与实际试验相吻合,他们固定缩氨酸,用色谱法试验进一步证明:缩氨酸和 α 淀粉酶的主要作用力是静电作用(包括氢键)和范德华力。

2.3 分子对接方法分类

各种分子对接方法对对接体系均有一定的简化,根据简化的程度和方式,可以将分子对接方法大致分为三类:

2.3.1 刚体对接 在对接过程中,研究体系(主体与客体)的构象不发生变化。适合考察结构比较大的体系,如蛋白质和蛋白质间以及蛋白质和核酸等大分子间的对接。Stoddard 等^[3] 用二元对接方法成功地对接麦芽糖和蛋白质进行了对接。对接时,他们把配体和受体主链结构当作刚性处理。

2.3.2 半柔性对接 在对接过程中,研究体系尤其指配体的构象允许在一定的范围内变化,但大分子是刚性的,构象不变化。适合处理大分子和小分子间的对接,因为受体大分子构象的变化相对底物分子而言变化不大。比如蛋白质分子与配体小分子的作用^[7-10]。进行半柔性对接时,一般使用快速的对接软件 FlexX,因为它^[11] 允许配体是柔性,而受体是刚性的。

2.3.3 柔性对接 在对接过程中,底物和受体的构象是允许发生变化的。一般用于精确考虑分子间的识别情况。分子的柔性主要来自于可旋转键的旋转。这种变化包括三个平动自由度、三个转动自由度以及底物分子的部分二面角的变化, Mangoni 等^[12] 用柔性的配体对柔性的受体进行了对接。

3 分子对接的研究进展

计算机和计算科学的迅速发展使得药物设计工作者能够运用计算机处理大量数据,大大推动了分子对接技术的发展。1982 年,加利福尼亚大学旧金山分校 Kuntz 教授的研究小组开发了 DOCK(分子对接软件和结果数据库)软件,该软件通过计算机预先筛选小分子数据库,运用打分函数评价受体小分子与配体的匹配程度,筛选出能高度和受体匹配的小分子,从而扩展了分子对接方法的应用,增强了它的实用性,并使其进入计算机量化处理的全新阶段。此时的分子对接方法已经与我们现在研究的分子对接方法原理基本相同。

研究发现许多蛋白质与不同的受体分子接触时,其侧链甚至是主链要发生移动,同类的蛋白质会发生相似的移动。分子对接方法也在不断地发展,从 Kuntz 研究小组开发的 DOCK 1.0 软件(基于空间匹配的刚性模型)发展到现在的 DOCK 4.0(基于空间匹配和能量匹配的柔性模型),模型的优化使分子对接方法更能准确地模拟分子间的相互作用。Koshland^[13] 提出了诱导契合(Induced Fit Theory)的概念:当底物与受体相互结合时,受体将采取一个最佳的构象同底物达到最佳结合,同时底物也会改变原有的构象以更好地同受体结合,这个过程也被称为分子对接的构象重组织。并且,他们^[14] 用沙丁格糊精,环状糊精和 β -环糊精与红苕淀粉酶的反应验证了诱导契合理论。分子对接计算过程中,对接分子的柔性就是在此基础上通过计算机模型发展起来的。已经有很多文章报道:当配体接近 RNA 受体时, RNA 结构发生了变化^[15-19]。

考虑了分子柔性的分子对接方法所涉及的搜索空间非常巨大,需要耗费大量的时间、财力和物力,并且对计算环境也有较高的要求。近年来,柔性对接和对接速率成为当今研究的热点。Richards^[20] 用互联网把世界上两百多个国家的 350 万台个人计算机连接起来,其能够在几天之内从数十亿的药分子库里筛选出目标蛋白质的配体。这个筛选系统大约是世界上最快超型计算机的三倍。筛选出的数百个小分子合成并测试得 10% 的小分子有活性,它们非常符合工厂的标准。郭权等^[21] 利用网格技术对分子对接过程

进行了优化,有效利用了网格结点资源,并降低了对接时间。

将优秀对接方法和软件应用于药物分子对接中,有效解决了上述提到的问题。为推动分子对接方法的发展,欧洲生物信息学中心举办了 CAPRI 蛋白质-蛋白质复合物结构预测竞赛。2001 年至今已成功举行了 9 届国际对接竞赛,一些拥有大量成功预测范例可用来评定或合理化配体-受体对接模式的软件和算法相继出现,有力促进了分子对接方法的发展。分子对接的常用软件有: DOCK(<http://dock.compbio.ucsf.edu/>)(注意: DOCK 软件只能在 Unix 和 Linux 系统上运行)、AUTODOCK(<http://www.scripps.edu/mb/olson/doc/autodock/>),SYBYL(http://tripos.com/index.php?family=Modules,SimplePage,discovery_info),Program SievGene(http://www.jbic.or.jp/activity/st_pr_pj/mypresto/index_mypr.html),Arguslab(<http://www.arguslab.com/>)等人们经常运用的软件。为了尽可能模拟实际分子间相互识别的过程,分子对接方法不断被改进,新的软件层出不穷,例如:Rosetta++(http://depts.washington.edu/ventures/UW_Technology/Express_Licenses/Rosetta),该软件不但考虑了蛋白质-蛋白质对接时生物大分子侧链、环区(loop)、结构域(domain)等的柔性,还考虑了蛋白质主链的柔性。eHiTS(<http://www.simbiosys.ca>)软件能够快速、准确地对柔性配体进行对接。该软件运用 3 种打分函数组合的方法提高了打分的稳健性,避免了分子对接中的假阴性问题,即没有使研究人员失去发现好的药物先导化合物的机会。Zheng 等^[22,23]用 Multiple Copy Simultaneous Search (MCSS) 方法和柔性对接技术成功地构造了 Bcl-2 抑制剂的三维药效团模型。Feature Point Pharmacophores (FEPOPS)^[24]比早期的基于形状和基于电荷的算法表现出更好的优越性。

空间、能量优化方面常采用模拟退火^[25]、遗传算法、片断生长等方法来寻找受体和配体最佳的结合位置。格点能量计算方法、人工神经网络、分子动力学^[26]、蒙特卡洛、禁忌搜索^[27]、局部搜索、拉马克遗传算法(LGA)等各种方法结合进行优化,其中 LGA 将遗传算法与局部搜索方法相结合。多目标优化问题的粒子群优化算法(Multiobjective Particle Swarm Optimization)^[28],Metadynamics^[29]算法,该方法不但能够找到最佳的空间位置、预测对接的键合力,还可以探究从对接溶液到对接空穴的整个过程。Gervasio^[30]用该方法对四组分子在溶液里进行了柔性对接,取得了较好的结果,Gervasio^[31]还指出对接时受体-配体相互识别的机理对药物设计起着重要的作用。包括配合物与底物的结合位点^[3](如:DNA 的大沟、小沟或碱基位置的特异性结合等)、生物活性构象^[32]和键合模式的确定。

分子对接方法的键合模式有插入、沟结合和静电作用等模式。在确定小分子化合物与 DNA 作用模式方面,流体力学方法常用来验证小分子化合物是否以插入方式与 DNA 作用。当含有平面芳香环的小分子与 DNA 碱基对以插入方式结合时,会导致 DNA 长度的增加,使其溶液粘度增大。当以部分插入方式与碱基对结合时,常产生 DNA 双链的扭结,使溶液粘度下降。当化合物以其它方式与 DNA 键合时,DNA 溶液粘度变化不大。电泳法亦可验证插入方式,在拓扑异构酶存在下,超螺旋环状 DNA 与具有平面芳香环化合物作用时,导致超螺旋解旋,不同解旋度的环状 DNA 在凝胶电泳中电泳速度不一样。解旋度越大,泳动速度越小。而以沟、静电结合的化合物,一般不会导致超螺旋的解旋(该方法与粘度法结合使用会更好)。此外还有热力学方法、电化学方法等,不过这些方法在研究小分子化合物与 DNA 键合机理方面不常用。配合物的形状结构与底物(DNA)的匹配程度决定了配合物与 DNA 的结合模式。

分子对接方法中结合位点的确立非常的重要。Collins^[33]等第一次运用多尺度算法成功地确定了蛋白质表面的结合位点^[28]并对分子进行了柔性对接,大大推进了分子对接的发展。Stoddard 等^[3]用二元对接方法对蛋白质复合物的结构进行了预测并得到了它们的结合位点。利用二维 NOESY 谱可判断出配合物与 DNA 键合的具体位点,它是更深层次上探测小分子化合物与 DNA 键合机理所采用的一种技术^[33-36]。

目前,发展的重点包括打分函数^[37]的完善、受体活性位点和对接时分子间力场的确定等。分子模型柔性的考虑不够充分,不能成功地处理复合物形成中较大的构象变化,如蛋白质分子域间类似铰链的运动。

4 分子对接的用途

分子对接方法在药物研究领域发挥了重要作用,为先导化合物的发现和优化提供了有效的工具。研

究小分子探针与细胞内生物大分子的相互作用, 确认小分子在生物体内的作用靶点, 为新药开发寻找新的突破口; 以结构生物学为基础, 对正常生理过程中及与肝癌、肝炎等疾病相关的重要蛋白质的结构和功能进行系统的研究与分析, 得到蛋白质药靶的三维结构, 进行药物与靶标蛋白相互作用的动力学模拟研究; 同时充分利用现有药物分子或中草药有效成分, 经修饰和优化, 设计具有更高活性的先导化合物。分子对接时, 用比较好的软件成功率高达 70% ~ 80%^[38, 39]。

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。药物研究过程一般是先利用化学计量学方法找到在分子和细胞水平活性较好的化合物, 然后进行实验室动物实验, 即小分子要发展成药物必须满足: 具有良好的吸收、分布、代谢和排泄性质以及较低的毒性。在这个过程中, 人们认识到越早淘汰那些性质较差的化合物, 就越能更有效地减少人力、物力和药物的开发周期。因此, 发展、利用各种分子对接方法进行新药的发现将成为一个研究热点。

其次, 因为小分子化合物能够与靶标生物大分子作用并抑制或激活蛋白质的生物功能。许多生物学家利用有机小分子作为有效的工具来探测生命体系^[40]。以分子对接为前提, 帮助人们了解 DNA 损伤和修复的机理, 合成出对某些与基因突变有关疾病的治疗药物。

此外, Lin 等^[41]用分子对接方法研究短小芽孢杆菌木聚糖酶与底物木聚糖的作用。因为木聚糖主要存在于农业产物中, 如: 秸秆, 玉米棒子, 所以木聚糖酶可以降解这些农业废物。

5 展望

分子对接方法的优势在于各种化合物数据库中的分子均是已知化合物, 且相当大一部份可以通过购买得到或根据已知的合成路线合成, 可以较快地进行后续的药理测试。目前有许多的商用数据库, 如可用化合物数据库(ACD)、剑桥晶体结构数据库(CSD)、世界药物索引(WDL)、药用化合物数据库(CMC)以及可用化合物搜索数据库(ACDSC)等, 为科研工作者提供了诸多方便。而且它可以模拟药物-受体相互作用, 阐明药物的作用机理, 增加药物准确性、灵敏性、特殊性、预测性, 这为药物开发提供了一个比较好的工具。近年来, 计算机技术的发展、靶酶晶体结构数据和算法数量的快速增长和商用小分子数据库的不断更新, 使得分子对接在药物设计中取得了巨大成功。

参考文献:

- [1] Fahmy A, Wagner G. J. Am. Chem. Soc. [J], 2002, **124**(7): 1241.
- [2] Audie J, Scarlata S. Biophysical Chemistry [J], 2007, **129**: 198.
- [3] Stoddard B L, Koshland D E. Nature [J], 1992, **358**(27): 774.
- [4] Robinson D D, Lyne P D, Richards W G. J. Chem. Inf. Comput. Sci. [J], 2000, **40**: 503.
- [5] Robinson D D, Lyne P D, Richards W G. J. Chem. Inf. Comput. Sci. [J], 1999, **39**: 594.
- [6] Liu F F, Wang T, Dong X Y, Sun Y. Journal of Chromatography A [J], 2007, **1146**: 41.
- [7] Sun Y, Ewing T J A, Skillman A G, Kuntz I D. J. Comput.-Aided Mol. Des. [J], 1998, **12**: 597.
- [8] Maurer M C, Trosset J Y, Lester C C, DiBella E E, Scheraga H A. Prot. Struct. Funct. Genet. [J], 1999, **34**: 29.
- [9] Oshiro C M, Kuntz I D, Dixon J S. J. Comput.-Aided Mol. Des. [J], 1995, **9**(2): 113.
- [10] Wang J, Kollman P A, Kuntz I D. Prot. Struct. Funct. Genet. [J], 1999, **36**: 1.
- [11] Glick M, Robinson D D, Grant G H, Richards W G. J. Am. Chem. Soc. [J], 2002, **124**(10): 2337.
- [12] Mangoni M, Roccatano D, Nola A D. Prot. Struct. Funct. Genet. [J], 1999, **35**(2): 153.
- [13] Koshland D E. Proc. Natl. Acad. Sci. [J], 1958, **44**(2): 98.
- [14] Thoma J A, Koshland D E. J. Am. Chem. Soc. [J], 1960, **82**: 3329.
- [15] Ruff M, Krishnaswamy S, Boegl M, Poterszman A, Mitschler A, Podjarny A, Rees B, Thierry J C, Moras D. Science [J], 1991, **252**: 1682.
- [16] Allain F H T, Gubser C C, Howe P W A, Nagai K, Neuhaus D, Varani G. Nature [J], 1996, **380**(18): 646.
- [17] Peterson R D, Feigon J. J. Mol. Biol. [J], 1996, **264**: 863.
- [14] Dieckmann T, Suzuki E, Nakamura G K, Feigon J. RNA [J], 1996, **2**: 628.
- [19] Jiang F, Kumar R A, Jones R A, Patel D J. Nature [J], 1996, **382**: 183.
- [20] Richards W G. Journal of Molecular Graphics and Modelling [J], 2007, **26**: 596.

- [21] GUO Quan(郭 权), WANG Xi cheng(王希诚), LI Churr lian(李纯莲). Journal of Computer Research and Development (计算机研究与发展)[J], 2004, **41**(12): 2054.
- [22] Miranker A, Karplus M. Prot. Struct. Funct. Genet. [J], 1991, **11**: 29.
- [23] Zheng C H, Zhou Y J, Zhu J, Ji H T, Chen J, Li Y W, Sheng C Q, Lu J G, Jiang J H, Tang H, Song Y L. Bioorganic & Medicinal Chemistry[J], 2007, **15**: 6407.
- [24] Nettles J H, Jenkins J L, Williams C, Clark A M, Bender A, Deng Z, Davies J W, Glick M. Journal of Molecular Graphics and Modelling[J], 2007, **26**: 622.
- [25] Goodsell D S, Olson A J. Prot. Struct. Funct. Genet. [J], 1990, **8**(3): 195.
- [26] Hansson T, Oostenbrink C, Gunsteren W F. Curr. Opin. Struct. Biol. [J], 2002, **12**(2): 190.
- [27] Cvijovic D, Klinowski J. Science[J], 1995, **267**: 664.
- [28] Janson S, Merkle D, Middendorf M. Applied Soft Computing[J], 2007, **8**: 666.
- [29] Laio A, Parrinello M. Proc Natl. Acad. Sci. USA[J], 2002, **99**(20): 12562.
- [30] Gervasio F L, Laio A, Parrinello M. J. Am. Chem. Soc. [J], 2005, **127**: 2600.
- [31] Gane P G, Dean P M. Curr. Opin. Struct. Biol. [J], 2000, **10**: 401.
- [32] Glick M, Grant G H, Richards W G. J. Med. Chem. [J], 2002, **45**(21): 4639.
- [33] Collins J G, Shields T P, Barton J K. J. Am. Chem. Soc. [J], 1994, **116**: 9840.
- [34] Greguric I, Aldrich Wright J R, Collins J G. J. Am. Chem. Soc. [J], 1997, **119**(15): 3621.
- [35] Collins J G, Sleeman A D, Aldrich Wright J R, Greguric I, Hambley T W. Inorg. Chem. [J], 1998, **37**(13): 3133.
- [36] Eriksson M, Leijon M, Hiort C, Norden B, Graslund A. J. Am. Chem. Soc. [J], 1992, **114**: 4933.
- [37] Perola E. Prot. Struct. Funct. Bioi. [J], 2006, **64**(2): 422.
- [38] Taylor P D, Jewsbury P J, Essex J W. J. Comput.-Aided Mol. Des. [J], 2002, **16**: 151.
- [39] Congreve M, Murray C W, Blundell T L. Drug Discovery Today[J], 2005, **10**(13): 895.
- [40] Stockwell B R, Nature[J], 2004, **432**: 846.
- [41] LIN Jir xia(林锦霞), ZHANG Liao yuan(张燎原), ZHANG Guang ya(张光亚), FANG Bai shan(方柏山). Chin. J. Biotech(生物工程学报)[J], 2007, **23**(4): 715.

Applications and Developments of Molecular Docking Method

DUAN Ai xia, CHEN Jing, LIU Hong-de, LIU Xiur hui, LU Xiaσ quan^{*}
(*College of Chemistry and Chemical Engineering, Northwest
Normal University, Lanz hou 730070*)

Abstract: In this article, we introduce the principle of molecular docking method and its requirements of geometry and energy in detail. Some ways of optimization are illustrated. The applications in drug design, pharmacological research and exploring life system are included. Many specific advantages of it have been demonstrated, including convenience, rapidity and high accuracy. 41 papers are cited.

Keywords: Molecular docking; Protein; Inhibitor; Drug design