

文章编号:1672-058X(2011)05-0547-04

医学结构基因组学在研制新药中的作用

姜怀春¹, 李 宏²

(1. 重庆工商大学 学术期刊社, 重庆 400067; 2. 重庆教育学院 生命科学与化学系, 重庆 400067)

摘 要:从药靶的选择、蛋白质片段为基础的新药开发、以寡聚酶为新药开发的靶、工业化的结构和蛋白质数据库以及专门研究传染病的结构基因组学工程等 5 个方面综述了医学结构基因组学在研制新药中的作用。

关键词:医学结构基因组学; 药靶; 蛋白质片段; 蛋白质数据库

中图分类号:Q139

文献标志码:A

自从结构基因组学出现以后,注册的新药数量远远没有达到预期,用高通量生物化学方法筛选新药失败的理由还不清楚,但是这反映出用于药物筛选的分子化学结构库里的分子结构信息十分有限,缺乏用于药靶的结构信息。事实上,以结构为基础的新药开发在新药设计中不断地起着重要作用,在已经确认的一些新药中,他们都完全或者部分来源于用 X-射线结晶衍射和核磁共振对蛋白质靶的分析的结构信息以及对蛋白质-配体复合物的结构分析信息。这说明蛋白质结构信息已经帮助人们开发出新药,包括以蛋白质的天然结构信息、蛋白质-配体复合物的结构信息、蛋白质-抑制子(inhibitor)的结构信息,这些信息可以增加药物对靶的亲合性(affinity)或者在保持靶的亲合性时,改善药物的性质^[1]。

随着人类基因组和一些病原体基因组的完成,以及在蛋白质分子结构测定方法的不断改进,出现了结构基因组学,结构基因组学的目标是解析尽可能多的来自一个有机体或多个有机体的有用的蛋白质结构。在过去 10a 中,全世界有 20 个结构基因组学研究所在从事蛋白质的分子结构研究^[2]。在以后的 5a 中,成千上万新的大蛋白质的结构将被阐明,这些结构信息包括许多与配体结合的蛋白质结构信息,这将奠定以结构为基础的新药研制的基础,也将帮助设计病原体蛋白质的化学治疗试剂(chemotherapeutic agent),Voorhis 等认为医学结构基因组学就是结构生物学与化学和生物学的交叉学科,也就是医学上确认的药靶的结构怎样作为设计抑制子的起点和新药开发的起点,并建议医药工业应该通过及时地使传染病微生物的药靶结构公开来完善公共的结构基因组学对分子结构的测定,通过公共-私人的联合来提供最大的协作,进而确认药靶、确认结构测定、用结构来设计新药^[1]。

1 药靶的选择

医学结构基因组学的先决条件是那些被测定了结构的蛋白质必须被很好地确定为药靶,有制药能力的(drugability)这个术语常常用于松散地描述任何给定的靶是怎样成为后选药(drug candidate)而进行新药开发的。对于那些传染病微生物,定义有制药能力的关键因素是靶蛋白(target protein)必须是微生物生存所必须的蛋白,这可以用基因敲除(gene knockout)和 RNA 干扰(RNA interference)这些技术来确认,有时,还要用化学生物学方法来补充,而且,要估计这些实验的意义常常会非常困难,因为寄主和病原体的相互作用常常是非常复杂的,例如,Wright 等花费了巨大的努力来开发细菌脂肪酸合成导路的靶蛋白的拮抗药^[3],有效的

药物类似分子已经被开发出来,这些分子能在胞外能有效地关闭细菌的复制。这些化合物分子在后来的动物实验中证明是无效的,然而,因为脂肪酸在脊椎动物中含量非常丰富,因此,即使在自己的脂肪酸合成通路被关闭的情况下,细菌可以用寄主的脂肪酸生存和生长。所以,为了改进结构基因组学中靶的选择,必须与化学和生物学家合作,研制出在适当条件下能引起病原体死亡的化合物^[4]。

如果一种药物的靶蛋白结构已经弄清楚了,医学结构基因组学可以提供快速和高效的方法来获得配体-蛋白质结构,就是应用高通量 X-射线衍射和/或核磁共振方法来弄清分子结构。相反,当一个有细胞活力的化合物的靶还不知道时,医学结构基因组学研究可以提供纯化的蛋白质用于潜在的药靶筛选,可以用象热稳定性这样的生物物理方法来筛选靶蛋白^[5]。原生动植物病原体医学结构基因组学研究小组(The Medicinal Structural Genomics of Protozoan Pathogens)已经开始用这种方法筛选抗疟疾化合物^[1]。目的是了解合成的小分子对靶蛋白的生物学作用。以后将有更多医学结构基因组学研究中心与化学生物学研究小组合作来收集表现型定义的化合物(phenotype-defined compounds),结果将是多个单位协同来进行靶的确认,用蛋白质结构来设计药物^[1]。

2 蛋白质片段为基础的新药开发

以蛋白质片段为基础的新药开发引起医药工业的巨大兴趣,以片段为基础的新药开发包括用“三个规则”建立的蛋白质大分子靶结构库筛选化合物,这些大分子的靶结构是用 X-射线衍射、核磁共振(NMR)、表面等离子体共振(surface Plasmon resonance)、分热变性(differential thermal denaturation)、荧光极化(fluorescence polarization)和其他技术测定的^[6]。“三个规则”包括分子量小于 300 道尔顿(dalton)、分子含有小于或等于 3 个旋转键(rotatable bonds)、分子含有小于或等于 3 个氢键供体/受体(hydrogen bond donors/acceptors)、另外,Clog P 小于 3 (Clog P 是计算的辛醇/水分配系数的对数)^[1]。如果这些片段更可能象药物的话,这些化合物通常包括片段或者是叫做可用药物的建筑模型(building block),这些片段成为新药开发的引导(lead)^[7]。生命片段库(The Fragments of Life library, FOL)收集了大约 1 400 个分析了结构的小分子,这些小分子是从细胞环境、代谢物、自然产物、和他们的衍生物或同分异构体得到的,这个库也收集了一系列的二芳基小分子来模拟蛋白质二级结构^[1]。因此,这些片段集对于酶的活性位点靶和含有变构小分子结合位点的更复杂的蛋白质表面靶以及蛋白质-蛋白质接触表面靶是非常重要的^[8]。

3 以寡聚酶为新药开发的靶

在核糖体和核孔复合物中分布有简单的蛋白质二聚体和及其复杂的多聚体,这些蛋白质的聚集形成了微生物生命活动的基础,参与的大分子之间常常有很多接触点。蛋白质-蛋白质接触表面并不是必然伴随有自由能(free energy)的产生,但是,小分子已经表现出能阻止关键的蛋白质-蛋白质相互作用^[1]。这个发现使 Wells 等用结构为基础的方法开发小分子抗生素,这些抗生素通过结合到多蛋白质内的蛋白质亚基之间的表面来调节蛋白质活性^[9],细菌的无机焦磷酸酶就是一个典型的例子,因为焦磷酸酶以六聚体的状态存在,在把无机焦磷酸转化为磷酸的过程中,需要构象变化^[1]。而且,所有的细菌无机焦磷酸酶以同源六聚体(homohexamer)的状态才起作用,真核生物的细胞质和线粒体的无机焦磷酸酶以同源二聚体的形式起作用^[10]。因此,与细菌的无机焦磷酸酶相比,真核无机焦磷酸酶有不同的寡聚体相互作用表面^[1]。Voorhis 等认为可以用寡聚体状态作为药靶,而不是以高度保守的活性位点作为药靶来抑制细菌焦磷酸酶的活力^[1]。一种相似的办法已经被用于鉴定胆色素原合成酶(porphobilinogen synthase)活性的种专一性调控代谢物(species-specific modulator)^[11]。西雅图传染病结构基因组学研究中心(Seattle Structural Genomics Center for Infectious Disease, SSGCID)已经解析了来自病原微生物 *Burkholderia pseudomallei* 的无机焦磷酸酶的高分辨率 X-射线晶体结构,FOL 正在从鉴定的几个片段中筛选药靶,这几个片段能够在同源六聚体的两个三体(trimer)的分子表面专一性地结合到多寡聚体的口袋中(multiple oligomerization pocket)^[1],然而,对于无机焦磷酸酶活性的种专一性抑制,这些片段仍然需要确认,但是,他们提供了开发新抗生素的潜在起点^[1]。

4 工业化的结构和蛋白质数据库

结构基因组学工程不断地增加解析了结构的蛋白质数量,并且这些信息是可以随便从蛋白质数据库(Protein Data Bank, PDB)得到的,得到这些信息不需要任何成本,也没有版权的限制。大多数出版商要求作者在发表文章时,把蛋白质数据储存在蛋白质数据库,所以,全世界的科学家测定的蛋白质结构在全世界得到了很好的传播^[12]。相比较而言,药学工业分析的重要病原体的蛋白质-配体复合物的结构数据还没有运用到更广阔的科学队伍。Voorhis 等建议美国国家健康研究所(the United States National Institute of Health)与其他的国家或国际研究机构一起把自己数据库里昂贵的结构数据转移到 PDB^[11]。

5 专门研究传染病的结构基因组学工程

结构基因组学的出现还不到 10 年,结构基因组学利用了传统科学研究的事实,那就是有相似的氨基酸序列的蛋白质具有相似的结构,因此,如果有人想要得到一个蛋白质家族的结构信息,这个家族中的任何一个蛋白质都可以提供,应用高通量结构方法(high throughput structural method)来获得蛋白质家族成员的结构信息,允许研究人员过滤和停止研究那些不能得到结构信息的蛋白质,而是,继续研究那些容易得到结构信息的蛋白质。这种方法加大了结构基因组学研究的成本,但是,却增加了这个蛋白质家族的一些成员的结构被测定的可能性,测定蛋白质家族成员的结构不同于前面提到的专一地确认一个蛋白质的药靶^[1,13]。有效地利用结构基因组学方法研究传染病需要对常规的方法进行一点点修饰,这个方法不是从大量的微生物中寻找最适合的代表蛋白质(representative protein),而是从所研究的病原体中选择一些有潜在药靶的蛋白质。虽然这些蛋白质可能不是先前确认的药靶,但是这种方法可以产生后选蛋白(candidate protein)的许多结构,可以进一步研究和检验来确认他们是潜在的药靶。即使蛋白质靶(protein target)不能满足工业上药物开发的标准,但是这些结构扩大了人们对选择的病原体的知识的了解。另外,大量的蛋白质表达载体和大量的纯化的蛋白质可以为研究人员提供很好的研究材料^[13]。

利用结构基因组学来研究潜在药靶的一个可能的问题是研究人员遇到对一些引导化合物(lead compound)过敏的微生物时可能终止研究,在小分子和蛋白质之间的相互作用的一点变化可能导致亲和性(affinity)的大量变化。过去 10 年中,许多大规模研究引起传染病微生物的研究已经开展,美国国家过敏和传染病研究所(National Institute for Allergy and Infectious Disease, NIAID)最近已经开始研究从所选择的病原体中的蛋白质,西亚图传染病结构基因组学研究中心和美国传染病结构基因组学研究中心(Center for Structural Genomics of Infectious Disease, CSGID)也在大规模从事引起传染病的微生物中蛋白质的结构基因组学研究。一个国际联合研究小组对结核病的研究就很成功,现在已经在 PDB 储存了 607 个相关蛋白质的结构^[13]。

6 展 望

随着结构基因组学研究的不断深入,特别是对蛋白质家族的研究,将有越来越多的蛋白质结构被阐明,越来越多的引起传染病的微生物的蛋白质结构被解析,也就有越来越多的药靶结构被测定和使用,越来越多的对付传染病的新药就会被开发出来,蛋白质家族的研究还有可能让人类从基因组顺序来了解药靶,设计新药,因此,结构基因组学研究将给人类带来巨大的福音。

参考文献:

- [1] VAN V W, HOL W, MYLER P, et al. The role of medical structural genomics in discovering new drugs for infectious disease [J]. PLOS Computational Biology, 2009, 10(5): 1371-1377
- [2] HAQUIN S, OEUILLET E, PAJON A, et al. Data management in structural genomics: An overview[J]. Methods Mol Biol,

- 2008, 426(1): 49-79
- [3] WRIGHT H, REYNOLDS K. Antibacterial targets in fatty acid biosynthesis [J]. *Curr Opin Microbiol*, 2007, 10(2): 447-453
- [4] HOON S, SMITH A, WALLACE I. An integrated platform of genomic assays reveals small-molecule bioactivities [J]. *Nat Chem Biol*, 2008, 4(2): 498-506
- [5] ERICSSON U, HALLBERG B, DETITTA G, et al. Thermofluor-based high-throughput stability optimization of proteins for structural studies[J]. *Analytical Biochemistry*, 2006, 357: 289-298
- [6] CONGREVE M, CHESSARI G, TISI D, et al, Recent developments in fragment-based drug discovery[J]. *J Med Chem*, 2008, 51: 3661-3689
- [7] BOSCH J, ROBIEN M, MEHLIN C, et al, Using fragment cocktail crystallography to assist inhibitor design of *Trypanosoma brucei* nucleoside 2-deoxyribosyltransferase[J]. *J Med Chem*, 2006, 49: 5939-5946
- [8] DAVIES D, MAMAT B, MAGNUSSON O, et al. , Discovery of leukotriene A4 hydrolase inhibitors using metabolomics biased fragment crystallography [J]. *J Med Chem*, 2009, 52: 4694-4715
- [9] WELLS J, MCCLENDON C, Reaching for high-hanging fruit in drug discovery at protein-protein interfaces[J]. *Nature*, 2007, 450(2): 1001-1009
- [10] SIVULA T, SALMINEN A, PARFENYE V, et al. Evolutionary aspects of inorganic pyrophosphatase[J]. *FEBS Lett*, 1999, 454 (1): 75-80
- [11] LAWRENCE S, RAMIREZ U, TANG L, et al. Shape shifting leads to small-molecule allosteric drug discovery [J]. *Chem Biol*, 2008, 15: 586-596
- [12] EDWARDS A, BOUNTRA C, KERR D, et al. Open access chemical and clinical probes to support drug discovery, *Nat Chem Biol*, 2009(5): 436-440
- [13] ANDERSON W. Structural genomics and drug discovery for infectious disease [J]. *Infect Disord Drug Targets*, 2009, 9 (5): 507-517

The Role of Medical Structural Genomics in New Drug Discovery

JIANG Huai-chun¹ , LI Hong²

- (1. Academic Periodical Office, Journal of Chongqing Technology and Business University, Chongqing 400067;
2. Department of Life Science and Chemistry, Chongqing Education College, Chongqing 400067, China)

Abstract: The role of medical structural genomics in new drug discovery is reviewed from such five aspects as drug target selection, protein fragment-based new drug development, target based on oligomeric enzymes, industry-generated structures and Protein Data Bank and structural genomics projects focused on infectious diseases.

Key words: medical structural genomics; drug target; protein fragment; Protein Data Bank

责任编辑:田 静