

虚拟筛选辅助新药发现的研究进展

刘艾林, 杜冠华*

(中国医学科学院、北京协和医学院药物研究所, 北京 100050)

摘要: 在新药发现过程中, 虚拟筛选的应用可以富集活性化合物, 降低筛选成本, 提高药物筛选的可行性, 因此已成为新药发现的重要方法。虚拟筛选与生物活性筛选的结合, 可以优势互补, 有效地促进新药的发现。本文介绍了非类药化合物排除、假阳性化合物排除、药效团搜索、分子对接计算以及分子相似性分析等几种方法在药物发现中的应用及其发展趋势, 以期更好地应用虚拟筛选方法, 促进新药的快速发现。

关键词: 虚拟筛选; 非类药化合物; 药效团; 分子对接; 分子相似性

中图分类号: R913

文献标识码: A

文章编号: 0513-4870 (2009) 06-0566-05

Research progress of virtual screening aided drug discovery

LIU Ai-lin, DU Guan-hua*

(Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China)

Abstract: In the process of new drug discovery, the application of virtual screening can enrich active compounds, reduce the cost of drug screening, and increase the feasibility of drug screening. Therefore virtual screening technology has become an important approach for new drug discovery. As virtual screening and bioactivity screening possess different advantages, their combination can effectively promote new drug discovery. In the present paper, the application and the trend of removal of non-drug compounds, removal of false positive compounds, pharmacophore searching, molecular docking, and molecular similarity in the process of drug discovery are introduced in order to obtain more benefit from virtual screening strategy for new drug discovery.

Key words: virtual screening; non-drug compound; pharmacophore; molecular docking; molecular similarity

药物筛选是药物发现的最初阶段和关键步骤, 其目的是发现新药。由于采用动物模型进行药物筛选的成本高、效率低、速度慢、样品需要量大等特点, 近年来已逐步从以体内实验筛选为主, 转变为以体外实验筛选为主特别是以高通量筛选 (high throughput screening, HTS) 为主的新药发现模式^[1]。高通量筛选采用自动化操作系统执行实验过程, 以灵敏快速的检测方法和手段采集生物活性数据, 每天可以完成数万甚至数十万个样品的活性检测, 从而极大地提高了药物筛选的速度。

高通量筛选技术的应用, 使得待筛样品的数量成为药物筛选的重要瓶颈, 因此新的化学合成技术特别是组合化学合成技术得到快速发展和广泛应用。据理论推算, 一个组合化学工作者一年可合成 $1 \times 10^4 \sim 1 \times 10^5$ 个化合物, 若将所有的组合化学库加在一起, 目前的化合物数量可达 1×10^9 甚至更多^[2], 大量化合物的出现虽然满足了高通量药物筛选的需要, 但是对于如此海量化合物的活性评价, 如果采取高通量筛选的方法, 其阳性率很低, 一般不到万分之一。在这种情况下, 仅仅依赖于高通量筛选技术发现先导化合物, 其成本高且成功率低, 将造成大量不必要的浪费。因此, 目前开展药物筛选, 还需要引入其他新理论、新方法和新技术, 以提高新药发现的几率, 降低新药发现的成本。

收稿日期: 2008-12-15.

基金项目: 国家科技基础条件平台工作项目 (2005DKA32400); 国家自然科学基金资助项目 (30171148).

*通讯作者 Tel / Fax: 86-10-63165184, E-mail: dugh@imm.ac.cn

药物虚拟筛选是基于药物设计理论, 借助计算机技术和专业应用软件, 从大量化合物中挑选出一些有苗头的化合物, 进行实验活性评价, 其阳性率一般在 5%~20%, 远远高于高通量筛选的阳性率^[3], 但虚拟筛选的计算结果最终还需要实验结果来验证。因此, 虚拟筛选技术与实验筛选技术优势互补, 如果将两者结合起来, 将有利于新药的快速发现。

本文根据虚拟筛选所采取的方法策略, 分别介绍非类药化合物排除、假阳性化合物排除、药效团搜索、分子对接计算以及分子相似性分析等几种方法在药物发现中的应用及其发展趋势。

1 非类药化合物排除

尽管目前可用于药物筛选的化合物是海量的, 但绝大部分并不产生药效活性。因此, 在药物发现的早期, 就应该排除这些无活性化合物, 以富集活性化合物, 降低筛选成本。在药物筛选的化合物样品准备阶段, 非类药化合物排除法常为首要采取的方法, 它是根据化合物类药性 (drug-likeness) 的特点^[4], 排除化合物数据库中违背化合物类药性特征的化合物。这种方法简单易行, 在化合物数据库管理系统中即可完成。这种排除法可以排除非特异性的无活性化合物, 适用于大多数高通量药物筛选前的化合物样品准备。但也有例外, 如抗肿瘤药物的筛选不排除金属有机化合物, 麻醉用药的筛选不排除小分子量的化合物等。

2 假阳性化合物排除

从理论上讲, 具有化学反应活性的化合物本不应该收集到用于新药发现的化合物样品库中, 但实际上化学合成工作者已经为新药发现提供了化学反应中间产物, 如醛、环氧化物、卤代烷以及分子整体为共轭体系的化合物等。这些化合物易于与生物大分子发生化学反应, 在基于受体、酶或细胞检测实验中总是表现为阳性, 而实际上为假阳性 (false positive)^[4, 5], 从而干扰药物研发的进程, 因此应尽早予以排除。

这些表现为假阳性的化合物, 其特征基团一般在水解条件下易于分解, 可以与蛋白质及生物亲核试剂 (如谷胱甘肽、二巯苏糖醇等) 产生化学反应, 在血清中稳定性很差。虽然从化合物数据库中排除所有可疑的假阳性化合物并不现实, 但对于从事药物筛选的工作者来说, 具备识别假阳性化合物的基础知识和判断能力, 能够在先导物发现的早期剔除假阳性化合物是至关重要的。

化合物是否为假阳性取决于化合物与生物大分子之间产生的相互作用是源于生物学反应还是化学反应。如果产生生物学反应的化合物是我们寻找的目标, 那么产生化学反应的化合物则被认为是假阳性化合物, 可以通过这种方法予以排除。但特殊情况下, 如果与生物大分子发生某种化学反应的化合物才是我们筛选与发现的目标, 这种排除方法并不适用。

3 药效团搜索

随着蛋白质结构快速鉴定技术的发展和运用, 越来越多的药物靶点的晶体结构得到解析, 但目前仍有许多药物靶点的结构是未知的, 特别是分布于细胞膜上的药物靶点^[6]。当药物靶点的结构未知时, 对于这些靶点的药物筛选, 可以应用基于药效团的虚拟筛选 (pharmacophore-based virtual screening), 在已知活性化合物结构基础上, 建立药效团模型, 从大量化合物数据库中搜索符合药效团模型的化合物, 从而富集活性化合物, 为药物筛选提供优质的待筛化合物。当蛋白质的结构已知时, 药效团搜索法常常与其他虚拟筛选方法联合应用, 使虚拟筛选的预测能力得到显著提高。

目前, 常用的药效团数据库搜索软件有 Tripos 公司的 Discotech 和 Unity 模块^[7], Accelrys 公司的 Pharmacophore 模块, Schrodinger 公司的 Phase 模块^[6-8]等。在此基础上, 又出现了基于柔性药效团的虚拟筛选^[9], 基于平行药效团的虚拟筛选^[10], 药效团比对搜索^[11]等。Yamazaki 等^[11]应用两个中心的药效团描述参数和宏观描述参数, 进行分类和回归分析, 使虚拟筛选的计算结果具有很高的预测能力, 且获得的先导候选物具有结构多样性。

4 分子对接计算

当药物靶点结构或同源蛋白结构已知时, 可以基于药物靶点或通过同源模建获得蛋白结构, 应用分子对接软件, 从化合物数据库中, 挑选空间上和化学性质上均与药物靶点活性位点相契合的化合物。另外, 当药物靶点具有多个亚型时, 需根据治疗疾病的特点, 选择对药物靶点特定亚型契合度高, 对其他亚型契合度低的化合物, 以提高化合物的选择性, 减少化合物的毒副作用。目前, 常用的分子对接软件有 Dock^[12]、Autodock^[13]、FlexX^[14]、ICM^[15]、Gold^[16]、Ligand-Fit^[17]以及 Glide^[18]等。当参与对接的分子数量过大时, 需要采用大规模分子对接的方法策略^[19, 20]。蛋白质结构的柔性化及溶剂化问题一直是制约分子对接计算结果的重要因素, 最近有关大分子柔性化

和溶剂化的软件开发及其在 HCV 聚合酶抑制剂虚拟筛选中的应用结果表明, 其预测结果优于其他软件^[21], 具有良好的开发和应用前景。

5 分子相似性分析

分子相似性分析是基于分子相似性, 将一个或多个与蛋白结合的化合物结构作为数据库搜寻的条件, 从化合物数据库中提取符合相似性标准的化合物, 用于生物活性评价。药效团搜索法仅利用了配体的部分结构特点, 对化合物数据库进行搜索, 而分子相似性分析则利用了分子整体结构特点, 与对照化合物进行比较, 从化合物数据库中搜索符合条件的化合物^[22]。为了有效地开展基于相似性的虚拟筛选, 分子相似标准应当把可以与同一蛋白密切结合的所有分子视为相似分子。

目前, 基于相似性的方法作为实用工具已被广泛应用于药物设计过程中。由于组合化学技术的应用, 化合物的合成速度大大提高, 因而产生大量不同结构的化合物, 在此基础上, 最常应用的虚拟筛选方法就是基于分子相似性的方法。另外, 这种方法还用于确定待筛化合物结构类型, 相似性配体分类以后, 构建生物活性与亚结构或功能团出现与否的 SAR 模型, 这些模型用于决定化合物参与实验筛选的优先权。常用的分子相似性虚拟筛选软件有 Cerberus^[23]、FlexS^[24]、GASP^[25]、MIMIC^[26]以及 GRID^[27]等。近年来, 分子相似性分析应用较多的还有骨架迁跃(scaffold hopping)^[28, 29]和形状识别(shape discrimination)^[30]。Talevi 等^[31]基于分子相似性分析, 提出了可以识别抗痉挛药物和非抗痉挛药物的新计算模型, 该计算模型将有助于抗痉挛新药的大规模虚拟筛选。

6 虚拟筛选的应用及优势

在药物筛选过程中, 应用虚拟筛选技术方法指导药物筛选, 首先需要确定药物筛选的疾病相关药物靶点, 在此基础上设计虚拟筛选策略。化合物数据库的处理需利用非类药化合物排除法和假阳性化合物排除法, 对化合物数据库进行初步过滤, 然后应用基于配体的虚拟筛选方法和(或)基于受体的虚拟筛选方法, 对化合物数据库作进一步过滤, 从中挑选部分有苗头的化合物进行生物活性评价。虚拟筛选完成以后, 库中的化合物分为两个部分: 一部分为被排除的化合物, 另一部分为待筛化合物。待筛化合物经实验筛选后, 一般情况下可以发现一些活性化合物, 将这些化合物的结构特点, 反馈到化合物数据库, 可以用于评价虚拟筛选的方法策略(图 1)。对于特定药

物靶点的药物筛选, 通过虚拟筛选和实验验证的多次反复, 可以不断完善两者结合的方法策略, 为基于其他靶点的药物筛选提供借鉴。

近年来, 虚拟筛选和高通量筛选等多种新药发现技术在新药发现中的整合应用, 出现了许多成功的实例, 如 HIV Nef 的蛋白-蛋白相互作用的抑制作用研究^[32], 抗肿瘤药物研究^[33], 抗流感病毒神经氨酸酶抑制剂的发现^[34], 抗锥虫 cruzitans 唾液酸酶抑制剂的发现^[35], 钾离子通道阻断剂的发现^[36], 蛋白激酶 CK2 抑制剂的发现^[37], 新型 falcipain-2 (FP-2) 抑制剂的发现^[38], G 蛋白偶联受体拮抗剂的发现^[39]等。

本实验室应用高通量筛选技术和虚拟筛选技术对样品库中的 4 855 个化合物同时进行了乙酰胆碱酯酶的抑制活性评价和虚拟筛选评价。实验结果发现 34 个活性化合物, 其 $IC_{50} < 1 \times 10^{-4} \text{ mol} \cdot \text{L}^{-1}$, 阳性率为 0.7%。应用多种虚拟筛选方法所得到的待测化合物, 其阳性率均有所提高, 阳性率最高可达 6% 以上。经虚拟筛选, 待筛化合物的数量减少了 10 余倍, 但其中活性化合物的阳性率比高通量筛选普筛的阳性率却提高了 7 倍多^[8]。虚拟筛选与实验筛选结果的比较表明, 虚拟筛选可以大大提高活性化合物的富集度, 减少待筛化合物的数量, 从而降低筛选成本, 提高活性化合物发现的几率和新药发现的速度。

随着高通量筛选等药物发现新技术的普及应用, 以及虚拟筛选计算软件的日趋成熟和广泛应用, 通

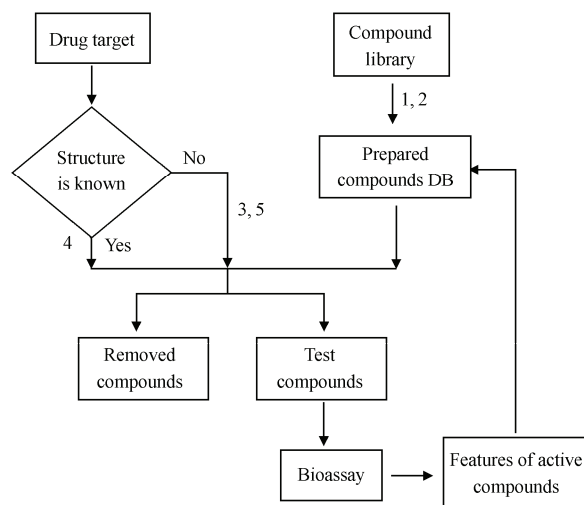


Figure 1 The application of the combination of virtual screening and bioactivity screening. 1: Removal of non-drug compounds; 2: Removal of false positive compounds; 3: Pharmacophore searching; 4: Molecular docking; 5: Molecular similarity

过药理学和计算化学等多学科研究者的通力协作,必将为新药发现注入新的活力,并有力地推动新药的快速发现。

References

- [1] Du GH. High Throughput Screening (高通量药物筛选) [M]. Beijing: Chemical Industry Publishers, 2002.
- [2] Langer T, Hoffmann RD. Virtual screening: an effective tool for lead structure discovery? [J]. *Curr Pharm Des*, 2001, 7: 509–527.
- [3] Waszkowycz B. Towards improving compound selection in structure-based virtual screening [J]. *Drug Discov Today*, 2008, 13: 219–226.
- [4] Liu AL, Du GH. Studies on the predict methods of drug-like properties of compounds [J]. *Chin Pharm J (中国药学杂志)*, 2003, 38: 644–648.
- [5] Lipinski CA, Lombardo F, Dominy BW, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery [J]. *Adv Drug Deliv Rev*, 1997, 23: 3–26.
- [6] Rishton GM. Reactive compounds and *in vitro* false positives in HTS [J]. *Drug Discov Today*, 1997, 2: 382–384.
- [7] Sun H. Pharmacophore-based virtual screening [J]. *Curr Med Chem*, 2008, 15: 1018–1024.
- [8] Liu AL, Du GH. Method research on virtual screening of acetylcholinesterase inhibitors [J]. *Comput Appl Chem (计算机与应用化学)*, 2003, 20: 547–550.
- [9] von Korff M, Freyss J, Sander T. Flexophore, a new versatile 3D pharmacophore descriptor that considers molecular flexibility [J]. *J Chem Inf Model*, 2008, 48: 797–810.
- [10] Steindl TM, Schuster D, Wolber G, et al. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening [J]. *J Comput Aided Mol Des*, 2006, 20: 703–715.
- [11] Yamazaki K, Kusunose N, Fujita K, et al. Identification of phosphodiesterase-1 and 5 dual inhibitors by a ligand-based virtual screening optimized for lead evolution [J]. *Bioorg Med Chem Lett*, 2006, 16: 1371–1379.
- [12] Rizzi A, Fioni A. Virtual screening using PLS discriminant analysis and ROC curve approach: an application study on PDE4 inhibitors [J]. *J Chem Inf Model*, 2008, 48: 1686–1692.
- [13] Shoichet BK, Stroud RM, Santi DV, et al. Structure-based discovery of inhibitors of thymidylate synthase [J]. *Science*, 1993, 259: 1445–1450.
- [14] Shrestha AR, Ali HI, Ashida N, et al. Antitumor studies. Part 5: Synthesis, antitumor activity, and molecular docking study of 5-(monosubstituted amino)-2-deoxy-2-phenyl-5-deazaflavins [J]. *Bioorg Med Chem*, 2008, 16: 9161–9170.
- [15] Price S, Bordogna W, Bull RJ, et al. Identification and optimization of a series of substituted 5-(1H-pyrazol-3-yl)-thiophene-2-hydroxamic acids as potent histone deacetylase (HDAC) inhibitors [J]. *Bioorg Med Chem Lett*, 2007, 17: 370–375.
- [16] Cavasotto CN, Ortiz MA, Abagyan RA, et al. In silico identification of novel EGFR inhibitors with antiproliferative activity against cancer cells [J]. *Bioorg Med Chem Lett*, 2006, 16: 1969–1974.
- [17] Kellenberger E, Springael JY, Parmentier M, et al. Identification of nonpeptide CCR5 receptor agonists by structure-based virtual screening [J]. *J Med Chem*, 2007, 50: 1294–1303.
- [18] Sharma P, Ghoshal N. Exploration of a binding mode of benzothiazol-2-yl acetonitrile pyrimidine core based derivatives as potent c-Jun N-terminal kinase-3 inhibitors and 3D-QSAR analyses [J]. *J Chem Inf Model*, 2006, 46: 1763–1774.
- [19] McInnes C. Improved lead-finding for kinase targets using high-throughput docking [J]. *Curr Opin Drug Discov Dev*, 2006, 9: 339–347.
- [20] Zhang S, Kumar K, Jiang X, et al. DOVIS: an implementation for high-throughput virtual screening using AutoDock [J]. *BMC Bioinformatics*, 2008, 9: 126.
- [21] Corbeil CR, Englebienne P, Yannopoulos CG, et al. Docking ligands into flexible and solvated macromolecules. 2. Development and application of fitted 1.5 to the virtual screening of potential HCV polymerase inhibitors [J]. *J Chem Inf Model*, 2008, 48: 902–909.
- [22] Lengauer T, Lemmen C, Rarey M, et al. Novel technologies for virtual screening [J]. *Drug Discov Today*, 2004, 9: 27.
- [23] Lushington GH, Guo JX, Wang JL, et al. Whither combine? New opportunities for receptor-based QSAR [J]. *Curr Med Chem*, 2007, 14: 1863–1877.
- [24] Feher M. Consensus scoring for protein-ligand interactions [J]. *Drug Discov Today*, 2006, 11: 421–428.
- [25] Jain AN. Scoring functions for protein-ligand docking [J]. *Curr Protein Pept Sci*, 2006, 7: 407–420.
- [26] Tirado-Rives J, Jorgensen WL. Contribution of conformer focusing to the uncertainty in predicting free energies for protein-ligand binding [J]. *J Med Chem*, 2006, 49: 5880–5884.
- [27] Rockey WM, Elcock AH. Structure selection for protein kinase docking and virtual screening: homology models or crystal structures? [J]. *Curr Protein Pept Sci*, 2006, 7: 437–457.
- [28] Cheeseright TJ, Mackey MD, Melville JL, et al. FieldScreen: virtual screening using molecular fields. Application to the DUD Data Set [J]. *J Chem Inf Model*, 2008, 48: 2108–2117.
- [29] Wale N, Karypis G, Watson IA. Method for effective virtual

- screening and scaffold-hopping in chemical compounds [J]. Comput Syst Bioinformatics Conf, 2007, 6: 403–414.
- [30] Kortagere S, Krasowski MD, Ekins S. The importance of discerning shape in molecular pharmacology [J]. Trends Pharmacol Sci, 2009, 30:138–147.
- [31] Talevi A, Prieto JJ, Bruno-Blanch LE, et al. New similarity-based algorithm and its application to classification of anticonvulsant compounds [J]. J Enzyme Inhib Med Chem, 2007, 22: 253–265.
- [32] Guido RV, Oliva G, Andricopulo AD. Virtual screening and its integration with modern drug design technologies [J]. Curr Med Chem, 2008, 15: 37–46.
- [33] Geromichalos GD. Importance of molecular computer modeling in anticancer drug development [J]. J BUON, 2007, Suppl 1: S101–118.
- [34] Liu AL, Cao HP, Du GH. Drug screening for influenza neuraminidase inhibitors [J]. Sci Sin C (中国科学 C 辑), 2005, 48: 1–5.
- [35] Neres J, Brewer ML, Ratier L, et al. Discovery of novel inhibitors of Trypanosoma cruzitrans-sialidase from *in silico* screening [J]. Bioorg Med Chem Lett, 2009, 19: 589–596.
- [36] Liu H, Gao ZB, Yao Z, et al. Discovering potassium channel blockers from synthetic compound database by using structure-based virtual screening in conjunction with electrophysiological assay [J]. J Med Chem, 2007, 50: 83–93.
- [37] Golub AG, Yakovenko OY, Bdzhola VG, et al. Evaluation of 3-carboxy-4(1H)-quinolones as inhibitors of human protein kinase CK2 [J]. J Med Chem, 2006, 49: 6443–6450.
- [38] Zhu J, Chen T, Liu J, et al. 2-(3,4-Dihydro-4-oxothieno[2,3-d]pyrimidin- 2-ylthio) acetamides as a new class of falcipain-2 inhibitors. 3. Design, synthesis and biological evaluation [J]. Molecules, 2009, 14: 785–797.
- [39] Engel S, Skoumbourdis AP, Childress J, et al. A virtual screen for diverse ligands: discovery of selective G protein-coupled receptor antagonists [J]. J Am Chem Soc, 2008, 130: 5115–5123.

《药学报》荣获 2009 年度中国科协精品科技期刊示范项目 B 类资助

5 月 8 日, 中国科协公布了获得 2009 年度中国科协精品科技期刊示范项目和英文版期刊国际推广项目的期刊名单。全国共有 145 个期刊被列为精品科技期刊示范项目, 其中 A 类项目 5 个, 每项资助 25 万元/年, B 类项目 40 个, 每项资助 15 万元/年, 非资助项目 100 个。《药学报》在连续两年获得中国科协精品科技期刊工程项目 C 类资助, 一年获得 B 类资助的基础上, 今年又获得了 B 类资助。

中国科协组织实施此项目, 旨在进一步促进科技期刊深化改革、创新发展, 增强科技期刊的核心竞争力, 努力提高科技期刊学术水平和出版质量, 推进科技期刊网络化与数字化建设, 大力培育精品科技期刊。《药学报》将再接再厉, 不懈努力, 争创精品期刊、一流期刊, 为我国药学事业的发展做贡献。

本刊编辑部