

doi:10.3969/j.issn.1672-5565.2015.04.07

结合分子相似性、药效团和分子对接 筛选新的 HIV-1 蛋白酶抑制剂

汪 滢,唐国荣,刘澍楠*

(中国人民解放军第476医院,福州 350002)

摘要:结合分子相似性、药效团和分子对接建立兼顾计算效率和预测准确度的 HIV-1 蛋白酶抑制剂筛选方法。首先通过对现有 HIV-1 蛋白酶抑制剂分子进行相似性分析,选取代表性的 HIV-1 蛋白酶抑制剂作为模板分子,构建和优化药效团模型,并从1万个化合物中优先筛选出500个化合物。而后采用分子对接方法进一步考察化合物与 HIV-1 蛋白酶结合情况,得到4个新的活性候选化合物,并进行其结合自由能计算和抗突变性分析。结果表明新候选化合物 ST025723 和 HIV-1 蛋白酶表现出较好的相互作用和抗突变性,具有深入研究的价值,同时也证明分子相似性、药效团和分子对接相结合能够快速有效地发现新颖活性候选化合物。

关键词:HIV-1 蛋白酶抑制剂;分子相似性;药效团;分子对接

中图分类号:R91 **文献标志码:**A **文章编号:**1672-5565(2015)04-244-07

Discovery of new HIV-1 protease inhibitors by integrating molecular similarity, pharmacophore and docking methods

WANG Ying, TANG Guorong, LIU Shunan*

(Department of Pharmacy, The 476th Hospital of PLA, Fuzhou 350002, China)

Abstract: An efficient and accurate method was established for HIV-1 protease inhibitor screening by integrating molecular similarity, pharmacophore and docking methods. First, the similarity of known HIV-1 protease inhibitors was analyzed, and representative inhibitors were chosen as templates to build pharmacophore model for database screening. Five hundred compounds were picked out from 10 000 compounds by pharmacophore screening, which were further submitted to docking experiment. Four new candidate compounds were obtained, and their binding free energies and antimutagenic abilities were further evaluated. The candidate compound ST025723 showed good interactions with HIV-1 protease and antimutagenic ability, and deserved to be further studied. Therefore, the combination of molecular similarity, pharmacophore, and docking could be an efficient strategy to discover novel active candidates.

Keywords: HIV-1 protease inhibitor; Molecular similarity; Pharmacophore; Docking

获得性免疫缺陷综合征(AIDS)是由人类免疫缺陷病毒(HIV)感染引起的传染病,对公共健康构成严重威胁^[1]。HIV蛋白前体由结构基因 *gag*、*pol* 和 *env* 编码而成^[2]。HIV-1蛋白酶是 HIV 病毒粒子组装、成熟过程的关键酶,其主要功能是特异性地裂解 *gag* 和 *gag-pol* 多聚蛋白前体形成具有活性的病毒结构蛋白和酶^[3,4]。HIV-1蛋白酶以二聚体形式

发挥其活性,干扰二聚体形成或破坏已形成的二聚体能够有效地抑制其活性^[5]。因此,HIV-1蛋白酶可作为抗艾滋病治疗的重要靶点,其抑制剂也成为目前抗 HIV 药物复合疗法的重要组成部分^[6,7]。近二十年,FDA 批准了9个 HIV-1蛋白酶抑制剂药物,其中包括最新的 Darunavir。但是 HIV 病毒变异率高,耐药病毒株不断出现,而且耐药毒株的耐药性

收稿日期:2015-10-14;修回日期:2015-11-17.

作者简介:汪滢,女,药师,研究方向:化合物数据库的构建与虚拟筛选;E-mail:yingwang912@163.com;

* 通信作者:刘澍楠,男,主任药师,研究方向:创新药物研究与药理学信息学;E-mail:liusnfj@gmail.com.

越来越强^[8,9]。Darunavir 是至今研发出的活性最高的 HIV-1 蛋白酶抑制剂,也是唯一一个被美国 FDA 批准作为治疗耐药型 HIV-1 的蛋白酶抑制剂。但是, Darunavir 具有肝脏毒性,特别是对伴有乙型和丙型肝炎病毒感染的 HIV 患者^[10]。另外, Darunavir 还可诱发严重的过敏反应^[11]。因此,开发新型的、安全性高的、抗耐药性的 HIV-1 蛋白酶抑制剂具有重要意义。

目前,虚拟筛选包括分子对接、药效团、三维定量构效关系以及机器学习,已经成功应用于先导化合物的发现^[12-15]。虽然单独的分子对接以及单个分子为模板的药效团模型已被用于 HIV-1 蛋白酶抑制剂的发现^[16,17],但这并不能满足海量化合物库筛选和新型活性骨架捕获的需要。DataWarrior 是全新的集分子描述符计算和数据分析于一体的化学信息学软件,能够利用自带的分子描述符对小分子化合物进行快速的相似性分析和聚类^[18]。LigandScout 是基于 3D-药效团模型进行精确虚拟筛选的集成平台。支持基于配体或基于结构的药效团建模,还包含基于药效团的叠合,共有药效团特征的创建,以及自动生成 ROC 曲线以便进行性能评估。其筛选速度快,特别适合于大化合物库的筛选^[19]。

因此,本研究采用分子相似性,药效团和分子对接相结合的策略。首先,对已知的 HIV-1 蛋白酶抑制剂进行相似性分析,选出代表性化合物作为模板建立药效团筛选模型,根据 Darunavir 和 HIV-1 蛋白酶间的相互作用指导药效团模型的特征优化,并利用已报道的 HIV-1 蛋白酶抑制剂作为训练集对模型筛选能力进行评价。最后,我们将优化好的模型应用于化合物库筛选,对打分较高的化合物进行分子对接,观察化合物与 HIV-1 蛋白酶的结合情况,选取活性候选化合物进行结合自由能计算和抗突变性分析。

1 材料与方法

1.1 软件和材料

采用软件 DataWarrior (开源软件), LigandScout 3.12 (Inte: Ligand), PyMOL (开源软件), Gold 软件 (CCDC Software Ltd), MOE (Chemical Computing Group Inc.) 和 Discovery Studio Visualization (Accelrys)。蛋白质晶体结构下载于 PDB (Protein Data Bank), 收集文献已报道的 HIV-1 蛋白酶抑制剂组成训练集 (Training Set)^[20-23], 类药性伪活性化合物库 (Drug-Like Ligand Decoys Set) 来自 Schrödinger, 化合物筛选库采用 TimTec 的多样化合

物筛选库 (Diversity-screening-set)。

1.2 药效团模型的建立、优化及验证

化合物包括 Training Set、Decoys Set 和 Screening Set 均采用 LigandScout 中的 OMEGA 生成 ldb 格式的化合物构象数据库, 单个化合物最多构象限制为 300。Darunavir 的 HIV-1 蛋白晶体复合物 (PDB ID: 4LL3) 从 PDB 获得, 并利用 PyMOL 对 Darunavir 和 HIV-1 蛋白酶的相互作用模式进行分析。采用 DataWarrior 对 141 个已知的 HIV-1 蛋白酶抑制剂进行相似性分析, 选取 7 个代表性的化合物进行药效团模型构建, 并根据 Darunavir 和 HIV-1 蛋白酶的相互作用分析指导药效团特征的优化。在模型建立和优化过程中, 我们利用训练集和类药性伪活性化合物库对模型进行评价, 考察模型对已知 HIV-1 蛋白酶抑制剂的富集能力。

1.3 基于药效团模型的化合物筛选和分子对接

将优化好的药效团模型作为提问结构, 应用 LigandScout 分子模拟软件包对 TimTec 公司的多样化化合物筛选库 (含 10 000 个化合物) 进行试筛, 得到与所建药效团模型配备较高的前 500 个化合物。再利用 Gold 软件对这些化合物进行分子对接, 以 HIV-1 蛋白酶晶体结构 (PDB ID: 4LL3) 作为受体筛选模型, 定义复合物中 Darunavir 结合区域作为活性位点, 综合考虑受体-配体之间的极性和非极性相互作用评价对接结果。最后选取打分高于或相当于 Darunavir 的化合物作为活性候选物。

1.4 结合自由能计算和抗突变性分析

现有的分子对接软件很少或很难考虑受体大分子的柔性, 采用 MOE 软件包中的 MMFF94 力场进行能量优化可以同时考虑受体和配体的柔性, 得到更为合理的结合自由能^[24]。在 HIV-1 蛋白酶抑制剂研发过程中, 药物学家已经发现了至少 50 种 HIV 蛋白酶抗性突变, Darunavir 的抗耐药性源于它能与蛋白酶活性中心的氨基酸骨架形成广泛而强的氢键作用, 而且大部分作用残基并不位于突变位点, 能有效抵抗蛋白酶局部突变所致的亲和力下降^[9]。因此分析活性候选物的氢键形成位点和蛋白酶突变位点, 有助于了解化合物的抗突变性。

2 结果与讨论

2.1 蛋白晶体复合物及相互作用分析

从蛋白数据库 (PDB) 下载 Darunavir 的 HIV-1 蛋白酶晶体结构 (PDB ID: 4LL3)。通过 Discovery Studio Visualization 对结构进行分析, 可以看到晶体复合物中 HIV-1 蛋白酶由 AB 两条链共同构成了抑

制剂 Darunavir 的结合口袋(见图 1a)。采用 PyMOL 软件对结合口袋进行相互作用分析,发现 Darunavir 与活性口袋匹配较好,能够同 HIV-1 蛋白酶上的 ASP-25, GLY-27, ASP-29 以及 ASP30 氨基酸残基产生强的氢键相互作用(见图 1b),另外 Darunavir 还

可与 HIV 蛋白酶的 Leu23, Gly49, Ile50, Pro81, Val82 以及 Ile84 等残基有较强的范德华相互作用,所有这些特性对于它的高活性和抗突变型蛋白酶的能力来说都是至关重要的。

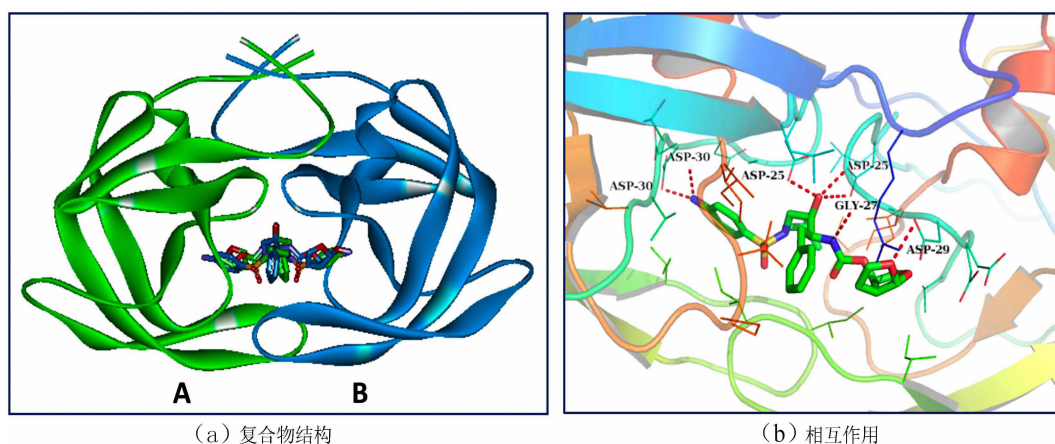


图1 Darunavir 和 HIV-1 蛋白酶的晶体复合物结构及其相互作用

Fig.1 The crystal structure of Darunavir bound to HIV-1 protease and their interactions

2.2 药效团的构建、验证和优化

利用 DataWarrior 对已知的 HIV-1 蛋白酶抑制剂进行相似性分析(见图 2a),发现其大致可分为 7 类,每类的中心分子被选作模板分子。利用 LigandScout 软件构建药效团模型并进行优化和验证,优化后的药效团模型如图 2b 所示,其包括 3 个芳香环,4 个氢键受体和 2 个氢键供体。采用训练集(由 141 个已知的 HIV-1 蛋白酶抑制剂组成)和

类药性伪活性化合物库(999 个来自 Schrödinger 的类药分子构成)对模型进行评价,考察药效团模型将活性分子和伪活性分子(Decoys)分开的能力。通过计算 ROC 曲线下面积(Area Under Curves, AUC)和 1%, 5%, 10%, 100% 四个水平的富集因子(Enrichment Factors, EFs)作为模型验证的基准。如图 3 所示,最后优化好的模型 AUC 为 0.80,四个水平的 EFs 分别为:8.1, 7.8, 5.2 和 2.8。

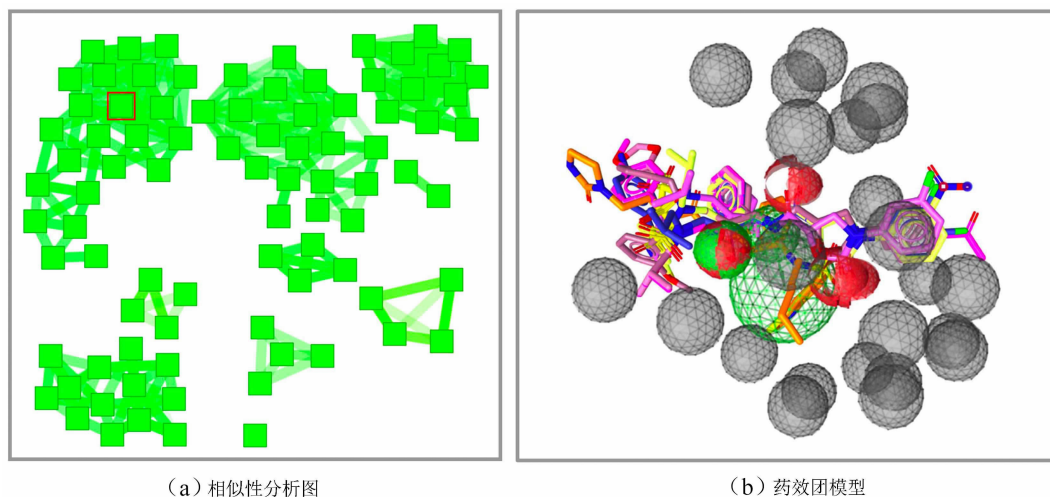


图2 HIV-1 蛋白酶抑制剂的相似性分析和以代表性分子为模板构建的药效团模型

Fig.2 The similarity analysis of known HIV-1 protease inhibitors using DataWarrior and the pharmacophore derived from seven representative HIV-1 protease inhibitors using LigandScout

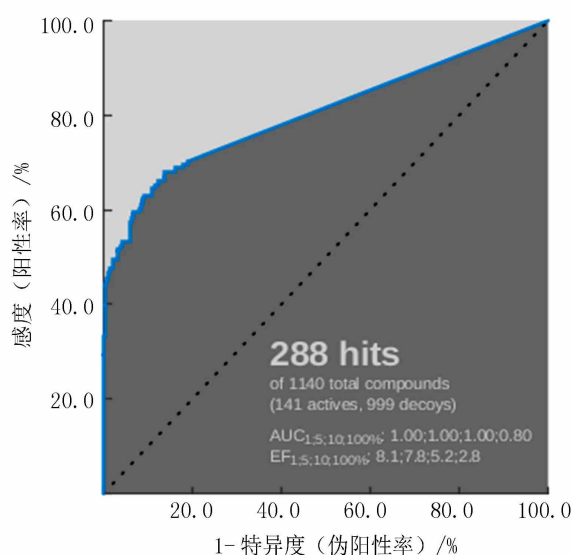


图3 药效团模型筛选能力的评价 (AUC=0.80)

Fig.3 The screening performance of the optimized pharmacophore model (AUC=0.80)

2.3 药效团虚拟筛选和分子对接

首先利用优化好的药效团模型对 TimTec 的 Diversity-screening-set (10 000 个化合物) 进行筛选,

保留打分较高的前 500 个化合物,然后利用 Gold 软件进行分子对接。为了检验对接方法的可靠性,选择将原有配体 Darunavir 重新对接到 HIV-1 蛋白酶受体中,然后比较配体对接后预测构象与晶体构象的差异。结果预测构象可以重现原晶体中的构象,表明分子对接参数能够有效用于化合物的进一步筛选。最后从分子对接结果中选取打分高于或相当于 Darunavir 的 4 个候选化合物做进一步分析(见表 1)。如图 4 所示,对获得的 4 个候选化合物与 HIV-1 蛋白酶间的相互作用进行分析,发现候选化合物 ST070845 可与 HIV-1 蛋白酶的 ARG-8, ASP-25, ASP-30 以及 GLY-48 形成多个氢键,而且分子中的苯并咪唑母核可以很好地嵌入疏水口袋。化合物 ST088084 中的三氮唑母核能够同时和蛋白酶 A、B 两条链的 ILE50 残基形成强的氢键作用,另外其末端氨基亦可与蛋白酶中的 GLY-48 残基形成氢键。化合物 ST025723 与蛋白酶的 ASP-25 和 GLY-27 残基产生氢键相互作用,这两个残基也是 Darunavir 的作用残基。化合物 ST025679 的酰胺结构可与 ILE-50 形成氢键,同时其甲氧基也可与 ASP-30 残基有氢键作用。

表 1 化合物的分子对接打分,结合自由能和氢键作用

Table 1 Docking scores, binding free energies and hydrogen bond interactions

化合物	分子对接打分	结合自由能(kcal/mol)	氢键作用残基
Darunavir	35.97	-121.75	ASP-25, GLY-27, ASP-29, ASP-30
ST070845	38.36	-42.35	ARG-8, ASP-25, ASP-30, GLY-48
ST088084	37.14	-62.37	GLY-48, ILE50
ST025723	37.06	-67.84	ASP-25, GLY-27
ST025679	36.74	-106.45	ILE-50, ASP-30

2.4 结合自由能计算、抗突变性分析和结构新颖性评价

分子对接可以快速地对配体小分子进行一定空间的构象搜索,以保证在合理的时间内从大量数据库中搜索到类药的化合物。但是,现有的分子对接软件很少或很难考虑受体大分子的柔性,难以在对接时适时调整受体的构象。对受体-配体复合物进行能量优化和结合自由能计算可以同时考虑受体和配体的柔性,受体活性中心的构象可以随着配体构象的改变而改变,真正做到“诱导契合”,得到更为合理的结合自由能^[25]。通过结合自由能计算,发现这些化合物的结合自由能要高于 Darunavir,但是化合物 ST025679 的结合自由能非常接近于 Darunavir,另外化合物 ST088084 和 ST025723 也显示出较低的结合自由能,而化合物 ST070845 虽然对接打分较高,但也呈现出较高的结合自由能。耐药性是艾滋

病治疗过程中的一个重要问题,主要原因是 HIV-1 蛋白酶序列中的某些氨基酸发生了抗性突变,按照它们在蛋白酶结构中的位置,又可以分为活性位点突变和非活性位点突变,其中活性位点突变占了较大比重。根据 HIV 蛋白酶常见突变位点和 4 个新活性候选化合物的作用残基,其中化合物 ST088084 和 ST025679 可能会因 I50V 突变导致其活性下降^[26,27]。最后,采用 FCFP_4 和 MACCS 两种分子指纹对化合物 ST025723 与 FDA 批准上市的 9 个 HIV-1 蛋白酶抑制药物进行相似性分析,发现 ST025723 与药物 Atazanavir 最为相似,其相似系数分别为 0.34 和 0.45,说明 ST025723 结构新颖,不同于已有的 HIV-1 蛋白酶抑制药物。因此,综合考虑结合自由能、抗突变性和结构新颖性,化合物 ST025723 具有更潜在的深入研究价值。

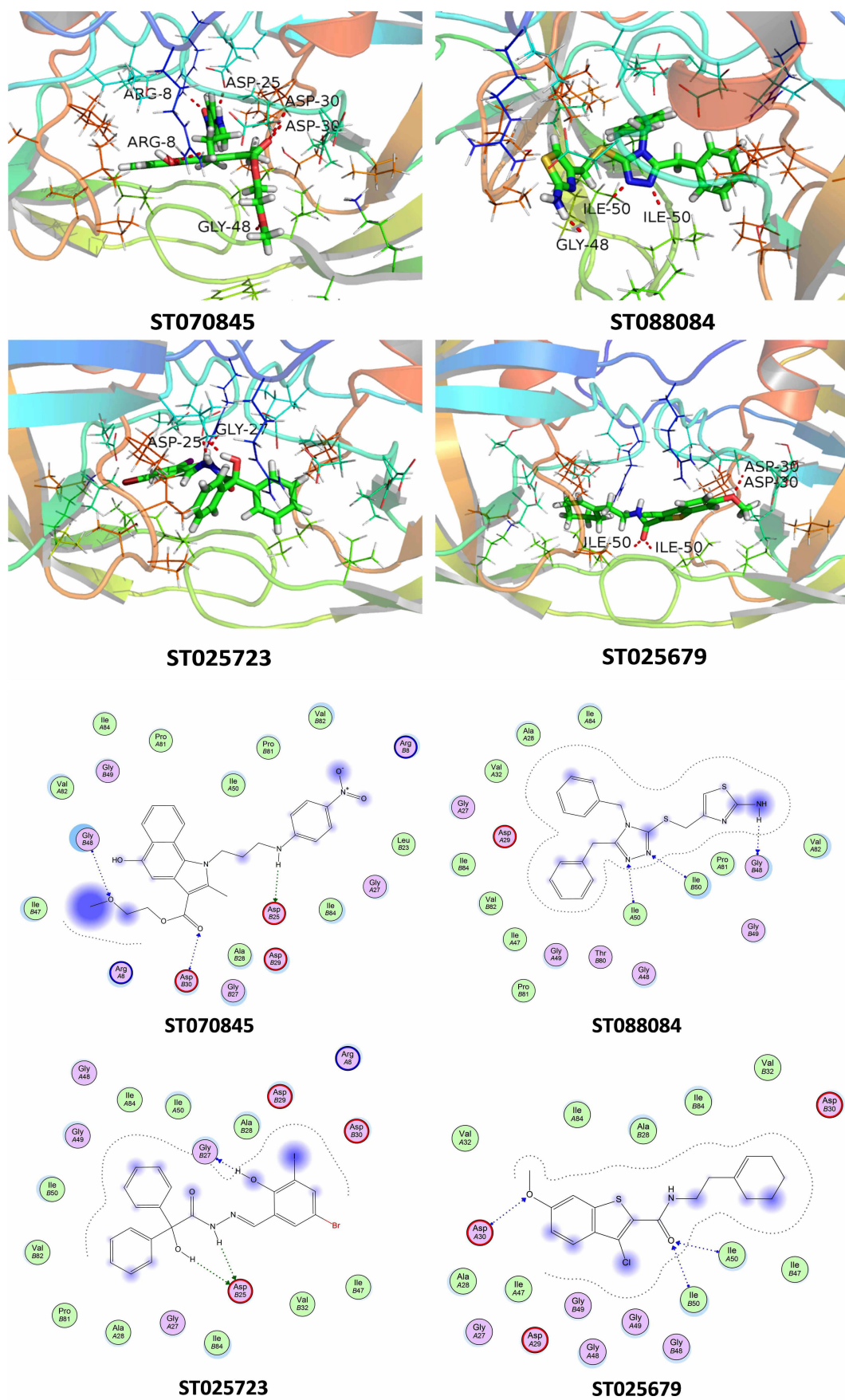


图4 HIV-1蛋白酶抑制剂候选化合物和HIV-1蛋白酶的相互作用模式
Fig.4 The interactions between candidate inhibitors and HIV-1 protease

3 结 论

HIV-1 蛋白酶是 HIV 复制过程中的关键酶,将其作为药物靶点并开发其抑制剂具有广阔的前景。本研究采用分子相似性、药效团和分子对接相结合的虚拟筛选策略,使各自的优点得到了很好的融合。由于药效团筛选速度快,适合于对大化合物库进行快速筛选。而后采用分子对接,对药效团筛选结果进行进一步的筛选,而且还能对获得的候选化合物进行相互作用模式分析,最后采用结合自由能计算,同时考虑受体和配体的柔性,使活性评价更为合理和准确。此外,通过化合物结合位点和 HIV-1 蛋白酶常见突变位点进行抗突变性分析,进一步从筛选到的活性候选化合物中挑选出具有抗耐药性的活性化合物。该策略确保了筛选结果的质量,达到在更短时间内更高效地发现抗耐药活性候选化合物的目的。根据对 TimTec 多样化化合物库筛选结果,4 个活性最好的候选化合物,虽然结构和现有 HIV-1 蛋白酶抑制剂 Darunavir 存在较大差异,但其作用模式却类似,这也充分展现了组合药效团模型在捕获新颖活性骨架方面的优势。

参考文献

- [1] MOORE C, CIARALD E. Quality of care and service expansion for HIV care and treatment[J]. *Current HIV/AIDS Reports*, 2015, 12(2): 223-230.
- [2] MCGETTIGAN J, NAPER K, ORENSTEIN J, et al. Functional human immunodeficiency virus type 1 (HIV-1) Gag-Pol or HIV-1 Gag-Pol and env expressed from a single rhabdovirus-based vaccine vector genome[J]. *Journal of Virology*, 2003, 77(20): 10889-10899.
- [3] ADAMSON C. Protease-mediated maturation of HIV: inhibitors of protease and the maturation process[J]. *Molecular Biology International*, 2012, 604261: 1-13.
- [4] 袁啸尘, 钮冰, 尹京苑. 基于二层特征筛选的 HIV-1 蛋白酶特异位点预测[J]. *生物信息学*, 2012, 10(4): 274-279.
CHEN Yuanxiao, NIU Bing, YIN Jingyuan. HIV-1 protease cleavage site prediction based on two stage feature selection method[J]. *Chinese Journal of Bioinformatics*, 2012, 10(4): 274-279.
- [5] HAYASHI H, TAKAMUNE N, NIRASAWA T, et al. Dimerization of HIV-1 protease occurs through two steps relating to the mechanism of protease dimerization inhibition by darunavir[J]. *Proceedings of the National Academy of Sciences*, 2014, 111(33): 12234-12239.
- [6] DEBOUCK C. The HIV-1 protease as a therapeutic target for AIDS[J]. *AIDS Research and Human Retroviruses*, 1992, 8(2): 153-164.
- [7] POTEPA M, LEE S, WOLFENDEN R, et al. The triple threat of HIV-1 protease inhibitors[J]. *Current Topics in Microbiology and Immunology*, 2015, 389: 203-241.
- [8] HOU L, WANG H, DUAN S, et al. The prevalence and determinants of drug-resistance-associated mutations in the HIV-1-infected MSM population of Henan Province in China[J]. *Archives of Virology*, 2015, 160(8): 2051-2061.
- [9] 郝格非, 杨光富. 新型非拟肽磺胺类 HIV 蛋白酶抑制剂 Darunavir 的研究进展[J]. *有机化学*, 2008, 28(9): 1545-1552.
HAO Gefei, YANG Guangfu. Progress in the new generation nonpeptide sulfonamide anti-HIV protease inhibitor darunavir[J]. *Chinese Journal of Organic Chemistry*, 2008, 28(9): 1545-1552.
- [10] VISPO E. Warning on hepatotoxicity of darunavir[J]. *AIDS Reviews*, 2008, 10(1): 62-63.
- [11] LIN K, CHENG C, YANG C, et al. Skin rash related to once-daily boosted darunavir-containing antiretroviral therapy in HIV-infected Taiwanese: incidence and associated factor[J]. *Journal of Infection and Chemotherapy*, 2014, 20(8): 465-470.
- [12] LAVECCHIA A. Machine-learning approaches in drug discovery: methods and applications[J]. *Drug Discovery Today*, 2015, 20(3): 318-331.
- [13] VUORINEN A, SCHUSTER D. Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling[J]. *Methods*, 2014, 71: 113-134.
- [14] 宋新蕊, 李达, 陈洁, 等. 计算机辅助药物筛选平台及应用[J]. *生物信息学*, 2014, 12(4): 300-304.
SONG Xinrui, LI Da, CHEN Jie, et al. Computer aided drug screening platform and its application[J]. *Chinese Journal of Bioinformatics*, 2014, 12(4): 300-304.
- [15] LAGARDE N, ZAGURY J, MONTES M. Benchmarking data sets for the evaluation of virtual ligand screening methods: review and perspectives[J]. *Journal of Chemical Information and Modeling*, 2015, 55(7): 1297-1307.
- [16] 谷万港, 张丽, 张旋. 基于结构的 HIV-1 蛋白酶抑制剂的虚拟筛选[J]. *昆明医科大学学报*, 2013, 34(8): 19-22.

- GU Wangang, ZHANG Li, ZHANG Xuan. Discovery of protease inhibitors of HIV-1 through structure based virtual screening [J]. Journal of Kunming Medical University, 2013, 34(8): 19-22.
- [17] 缪有盼, 李爱秀, 刘涛, 等. 基于“底物包膜”假说筛选新型 HIV-1 蛋白酶抑制剂 [J]. 化学学报, 2009, 67(10): 1098-1102.
- MIAO Youpan, LI Aixiu, LIU Tao, et al. Screening novel HIV-1 protease inhibitors based on the substrate envelope hypothesis [J]. Acta Chimica Sinica, 2009, 67(10): 1098-1102.
- [18] SANDER T, FREYSS J, VON KORFF M, et al. DataWarrior: an open-source program for chemistry aware data visualization and analysis [J]. Journal of Chemical Information and Modeling, 2015, 55(2): 460-473.
- [19] WOLBER G, LANGER T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters [J]. Journal of Chemical Information and Modeling, 2005, 45(1): 160-169.
- [20] WEI Y, LI J, CHEN Z, et al. Multistage virtual screening and identification of novel HIV-1 protease inhibitors by integrating SVM, shape, pharmacophore and docking methods [J]. European Journal of Medicinal Chemistry, 2015, 101: 409-418.
- [21] PARAI M, HUGGINS D, CAO H, et al. Design, synthesis, and biological and structural evaluations of novel HIV-1 protease inhibitors to combat drug resistance [J]. Journal of Medicinal Chemistry, 2012, 55(14): 6328-6341.
- [22] ISLAM M, PILLAY T. Exploration of the structural requirements of HIV-protease inhibitors using pharmacophore, virtual screening and molecular docking approaches for lead identification [J]. Journal of Molecular Graphics and Modeling, 2015, 56: 20-30.
- [23] ALI A, REDDY G, NALAM M, et al. Structure-based design, synthesis and structure-activity relationship studies of HIV-1 protease inhibitors incorporating phenyloxazolidinones [J]. Journal of Medicinal Chemistry, 2010, 53(21): 7699-7708.
- [24] DEETH R, FEY N, WILLIAMS-HUBBARD B. DommiMOE: an implementation of ligand field molecular mechanics in the molecular operating environment [J]. Journal of Computational Chemistry, 2005, 26(2): 123-130.
- [25] 曹冉, 李伟, 孙汉资, 等. 计算化学方法在基于受体结构的药物分子设计中的基础理论及应用 [J]. 药学学报, 2013, 48(7): 1041-1052.
- CAO Ran, LI Wei, SUN Hanzi, et al. Computational chemistry in structure-based drug design [J]. Acta Pharmaceutica Sinica, 2013, 48(7): 1041-1052.
- [26] PARTALEDIS J, YAMAGUCHI K, TISDALE M, et al. In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethyl amino sulfonamide inhibitors of HIV-1 aspartyl protease [J]. Journal of Virology, 1995, 69(9): 5228-5235.
- [27] MITTAL S, BANDARANAYAKE R, KING N, et al. Structural and thermodynamic basis of amprenavir/darunavir and atazanavir resistance in HIV-1 protease with mutations at residue 50 [J]. Journal of Virology, 2013, 87(8): 4176-4184.