

基于中药数据库的 HIV抑制剂的筛选

高维娜^{1,2}, 李云¹, 张瑞¹, 高慧¹, 徐为人², 李爱秀³,
杜奇石¹, 张欣¹, 魏冬青^{1*}

(1. 天津师范大学 生物信息与药物开发研究所, 天津 300074; 2. 天津药物研究院, 天津 300193;
3. 武警医学院, 天津 300162)

摘要: 目的 用现代化学信息学手段和中药化学数据库寻找新的抗 HIV 药物。方法 利用现有的抗 HIV 的药物作模版, 进行中药分子数据库搜索, 然后采用分子对接方法得到 HIV-1 蛋白酶受体与亚叶酸的合理复合物结构, 用分子动力学方法对对接结果分别进行无水存在 200 ps 和有水存在 50 ps 的模拟。结果 根据数据库搜索和对接, 认为亚叶酸可以作为药物开发的起点。通过分析分子动力学数据, 可以了解配体各个区域与受体相互作用的细节和变化规律。结论 本工作得到的 HIV-1 蛋白酶与中药分子抑制剂的结合模式信息将有助于设计和改造出效果更好的抗 HIV-1 蛋白酶抑制剂。

关键词: 蛋白酶抑制剂; 中药分子数据库搜索; 分子对接; 分子动力学; 亚叶酸

中图分类号: R916 文献标识码: A 文章编号: 0513 - 4870(2006)03 - 0241 - 06

Screening of new HIV inhibitors based on the database of traditional Chinese medicine

GAO Wei-na^{1,2}, LI Yun¹, ZHANG Rui¹, GAO Hui¹, XU Wei-ren², LI Ai-xiu³,
DU Qi-shi¹, ZHANG Xin¹, WEI Dong-qing^{1*}

(1. Institute of Bioinformatics and Drug Discovery, Tianjin Normal University, Tianjin 300074, China; 2. Tianjin Institute of Drug Discovery, Tianjin 300019, China; 3. Medical College of Chinese People's Armed Police, Tianjin 300162, China)

Abstract: **Aim** To report the preliminary result of the HIV inhibitor screening based on cheminformatics tools and the traditional Chinese medicine database. **Methods** Database search was carried out with saquinavir molecule as a template, further screening was made with docking. Detailed studies using molecular dynamics simulation of 50 ps and 200 ps were made with respect to a potential leading compound, leucovorin. **Results** The leucovorin molecule distinguished from other molecules as a potential drug candidate and is subject to extensive studies. The bonding profile and energy were calculated with MD simulations. **Conclusion** Our results could be very helpful when we modify leucovorin or design new inhibitors against HIV.

Key words: HIV protease; traditional Chinese medicines database search; docking; molecular dynamics; leucovorin

艾滋病^[1]是 20 世纪危害人类健康和生命最严重的疾病之一, 其病原体为人免疫缺陷病毒 (human

immunodeficiency virus, HIV)。HIV 蛋白酶归属于天冬氨酸基蛋白酶, 由两条含 99 个氨基酸的多肽链形成的 C₂ 对称的均二聚体, 每一亚基上天冬氨酸残基形成了酶的活性中心。HIV 蛋白酶 (HIV PR) 对该病毒复制周期正常运转和病毒毒粒成熟至关重要, 是病毒复制必需的酶, 是抗 HIV 药物的重要靶点^[2]。

收稿日期: 2005-06-08.

基金项目: 天津市科委基础科学重点项目 (033801911) 和计算专项基金资助项目。

* 通讯作者 Tel / Fax: 86 - 22 - 23540187,
E-mail: dongqing@ce mm. concordia. ca

至今已有许多抗 HIV 的药物。药物设计、筛选的一般做法是以现有的药物作模板进行数据库搜索,即基于药效团的搜索或相似性搜索。但是,现有的小分子化学数据库已被广泛利用,而中药、天然产物有效成分的提取进展非常迅速。中药化学数据库已经建立起来^[3,4]。如何利用数据库筛选新的活性高、毒性低的化合物是中药现代化的一个重要环节。本文报道利用现代化学信息学手段和中药化学数据库寻找新的抗 HIV 药物的初步结果。

方法

中药库相似性搜索 本研究所用的数据库是中国科学院工程研究所分子设计课题组开发的中药化学数据库 (traditional Chinese medicine database, TCMD 2003.1)^[4]。搜索的方法是相似性搜索。首先计算数据库里的所有分子的指纹 (fingerprint),将其存入存放中药分子的数据库里。指纹是由特征组成的,比如:

$U = \{ \text{是否芳族-化合物, 有无环, 卤素, C, N, O, S, P} \}$

有几种方法计算指纹,比如, MDL MACCS 钥匙 (keys)^[5,6], 类型图形距离 (typed graph distance, TGD) 和类型图形三角 (typed graph triangle, TGT)。然后计算模板分子的指纹并与数据库里分子的指纹比较。为此需要定义一个量度,一般是采用 Tanimoto 系数^[5],它是介于 0 和 1 之间的一个数值,0 和 1 分别代表最不相似和最相似。Tanimoto 系数的定义可用以下公式表示:

$$\#AB / (\#A + \#B - \#AB)$$

#AB, #A 和 #B 分别代表 A, B 共同的特征数、A 的特征数和 B 的特征数。将模板分子的指纹与数据库里的中药化学分子的指纹逐个比较,相似性超过 70% 的分子被标记并存放于结果数据库里,即得到了搜索结果。沙奎那韦 (saquinavir, invirase, SQV) 是较早上市的抗 HIV 药物,可作为数据库搜索的模板。TCMD 数据库有 9 000 多个分子,根据以上所描述的搜索方法,选用 MDL MACCS 钥匙 (keys)^[6] 方法计算指纹,得到了约 50 个与沙奎那韦相似的分子 (图 1)。下一步是以对接 (docking) 做进一步筛选。

分子对接 (docking) 从晶体蛋白库中选择有蛋白酶-底物复合物的晶体结构,进行分子叠合选出最有代表性的蛋白酶结构 (1A30) 作为受体蛋白分子。

分子对接方法就是两个或多个分子之间通过几

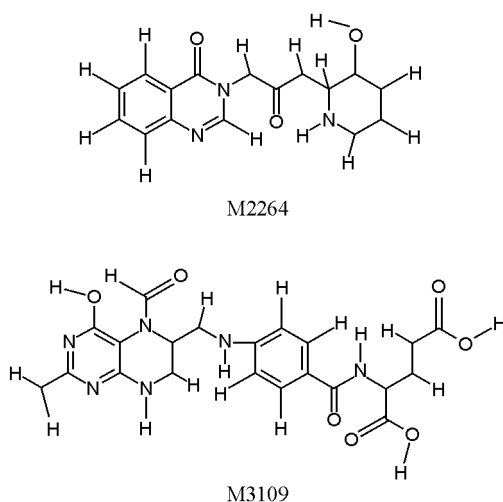


Figure 1 Typical molecules similar to saquinavir according to similarity search

何匹配和能量匹配而相互识别的过程,是药物设计的重要手段。在考虑药物分子与靶酶相互结合时,首先需要两个分子充分接近,采取合适的取向,使两者在必要的部位相互契合,发生相互作用,继而通过适当的构象调整,得到一个稳定的复合物构象。通过分子对接确定复合物中两个分子正确的相对位置和取向,研究两个分子的构象特别是底物构象在形成复合物过程中的变化,是确定药物作用机制、设计新药的基础。

药物分子与受体结合时,应遵循以下互补匹配规则: (1) 几何形状互补匹配; (2) 静电相互作用互补匹配 (正电荷对应负电荷); (3) 氢键相互作用互补匹配 (氢键供体对应氢键受体); (4) 疏水相互作用互补匹配 (疏水区对应疏水区)。

Flex X 是一种典型的基于片断的、快速、准确的柔性对接 (flexible docking) 算法,在对接时考虑了配体分子的许多可能的构象。该方法首先在配体分子中选取一个核心部分,并将其对接到受体的活性部位,然后再通过树搜寻方法连接其余的片断。它采用 Bohm 的结合自由能函数来评价分子间结合自由能,可以对化合物数据库进行虚拟筛选,还可以运行在多个 CPU 的 SG1 机器上,加快了筛选的速度。

本文采用 Sybyl 7.0 软件包的 FlexX/Run mutiligand 模块和 Amber 99 力场,取复合物中原配体周围 0.8 nm 范围为活性位点 (active-site) 进行分子对接,采用经验结合自由能函数作为打分函数来评价对接结果。

分子动力学模拟 分子动力学方法是根据分子的势能函数,得到作用在每个原子上的力,利用牛顿

第二定律求解运动方程,得到原子在势能面上的运动轨迹,从而达到构象搜索的目的。

为了考察水对模拟体系的影响,本实验分别在有水分子和无水分子存在的情况下进行分子动力学模拟。从分子对接结果中取配体中药分子亚叶酸与受体 1A30 的复合物,应用 Hyper Chem 7.0 软件包,选用 PM3 方法对复合物计算电荷,加入水盒子 (4.5 nm × 4.5 nm × 5.0 nm),使活性位点部位充满水分子。选中活性位点包含的残基结构、小分子配体及水分子进行局部优化(离小分子 1 nm 之外的受体残基保持刚性),选择部分结构进行局部优化后,作为初构象开始分子动力学模拟,温度为 310 K,步长 0.000 5 ps,每 50 步保存一次数据。平衡计算一般为 5 ~ 10 ps,统计平均的计算需要 50 ~ 200 ps。不加水进行的分子动力学模拟,除去加水一步,方法同上。

结果与讨论

1 分子对接

小分子配体与 1A30 的初步对接结果表明,亚叶酸分子的匹配最佳。因此,系统研究了亚叶酸分子与 1A30 的相互作用。系统搜索亚叶酸分子的构象,共得到了 2 000 多个构象,全部与 1A30 进行对接,结果表明,不同构象的对接结果不同,所以分子构象的选择对于分子对接具有一定的影响。

本文只报道一个对接筛选得到的药物分子并对对接模拟结果进行详细分析。这个分子是亚叶酸(leucovorin),图 2 显示其立体结构,它是从乌药中提取。此分子与钙结合形成亚叶酸钙,主要用作叶酸拮抗剂(如甲氨蝶呤、乙胺嘧啶或甲氧苄啶等)的

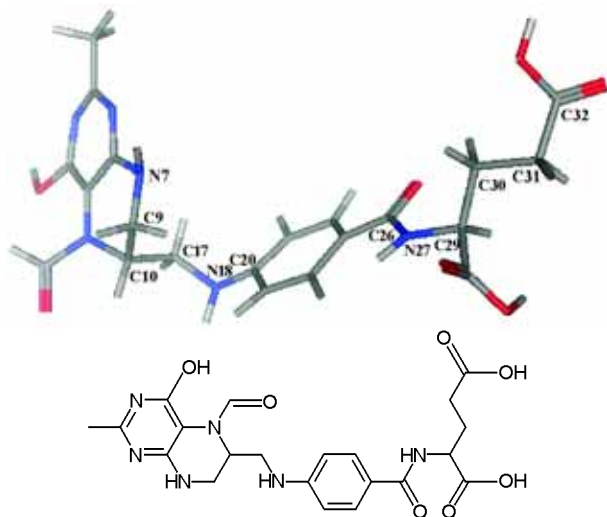


Figure 2 The structure of leucovorin

解毒剂,与氟尿嘧啶联合应用时,用于治疗晚期结肠癌、直肠癌等。

将筛选的分子与模板分子作柔性重叠^[7-9],证明该药物分子是一个好的抑制剂候选物。柔性对接时基于药效团方法(pharmacophore elucidation),其基本思想是若两个配体的生物特性结合的较好,二者会有好的重叠。柔性对接是基于分子的药效团重叠原理进行的。作者采用分子的体积、芳环组、氢键给体和受体为药效团,并定义其几率密度为一个高斯函数,通过优化相似函数 $-KT\log F + U$ 而得到分子的最佳重叠^[7,8]。U 是势能函数, $F = F_{\text{aro}} + F_{\text{don}} + F_{\text{acc}} + F_{\text{vol}}$,其中 F_{aro} , F_{don} , F_{acc} 和 F_{vol} 分别是芳环组、氢键给体、氢键受体以及体积的几率函数。

图 3 显示的是亚叶酸分子与其模板沙奎那韦的柔性重叠结果。药效团芳环组(aromatic groups)重叠得很好。由于亚叶酸分子中间部位单键较多,其转动势能垒很低,分子在活性位点较为灵活,这样可以保证其他药效团与受体较好的结合。除芳环组外,最主要的药效团是氢键给体和受体。氢键的形成可以作为对接好坏的表示之一。从图 4 可以看出,有 3 个氢键形成,表明亚叶酸分子与受体结合得很好。

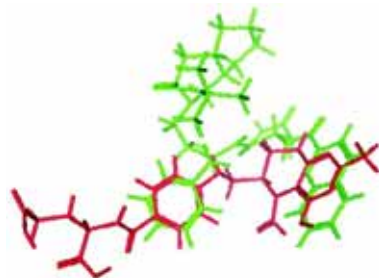


Figure 3 The flexible alignment of leucovorin and saquinavir, where red (darker) and green (lighter) colors represents leucovorin and saquinavir, respectively

为了进一步理解此亚叶酸分子和 1A30 蛋白质之间的作用机制,为药物改造、设计提供思路,用分子动力学方法^[10]研究了它和 1A30 蛋白质之间的相互作用,探测受体活性中心的势能面轮廓。从对接结果中选出得分最佳的构象和受体的复合结构(图 2,其对接得分为 -49.35)作为进行分子动力学研究的初始构象。

2 分子动力学模拟

2.1 系统的描述 在体系活性位置无水(200 ps)的情况下,势能和温度在 10 ps 后趋于平衡。对稳定 MD 数据进行统计,分析亚叶酸分子与蛋白相互作用。体系活性位置含水(50 ps)的情况下,势能和

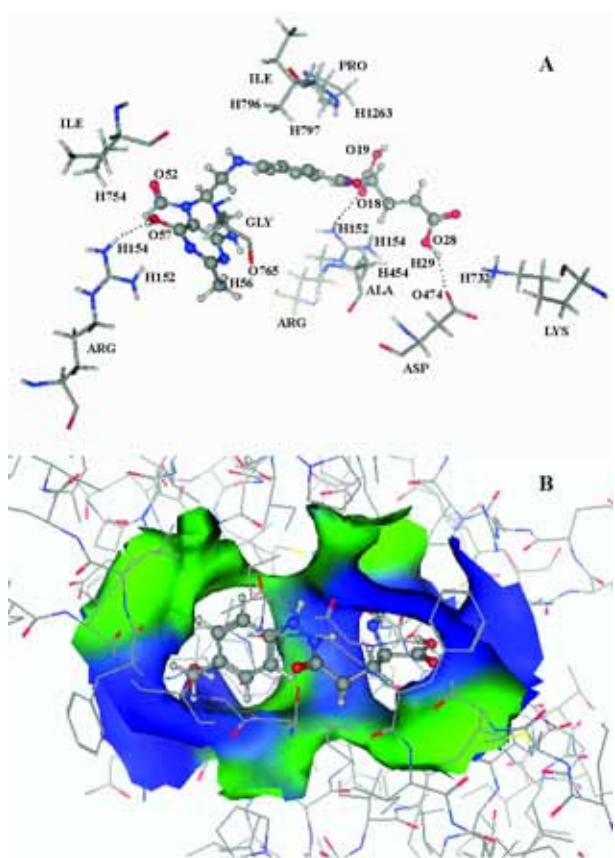


Figure 4 The close view of binding of leucovorin with the residues of the receptor around it. In A, the dotted lines represent the hydrogen bond, the leucovorin is in the ball-stick representation, the residues of the receptor is in the stick representation, oxygen, nitrogen, carbon and hydrogens are colored in red, blue, gray and white, respectively. In B, the molecular surface rendered with hydrophobicity near the ligand is displayed where the blue and green colors refer to hydrophobic and hydrophilic areas, respectively. It is quite clear that the ligand is well situated in the pocket and the hydrophobic features of ligand and receptor are well matched

温度在 5 ps 后趋于平衡。

2.2 配体的构象变化 选取亚叶酸分子的 5 个二面角, 定义为二面角 $N_7-C_9-C_{10}-C_{17}$, $C_{10}-C_{17}-N_{18}-C_{20}$, $C_{26}-N_{27}-C_{29}-C_{30}$, $C_{29}-C_{30}-C_{31}-C_{32}$ 和 $C_{30}-C_{31}-C_{32}-O_{33}$, 观察这 5 个二面角在动力学过程中的变化 (图 5)。原子的编号如图 2 所示。

数据表明二面角 $N_7-C_9-C_{10}-C_{17}$ 在动力学过程中的变化主要介于 $60^\circ \sim 90^\circ$ 之间, 变化范围较小, 它与整个杂环的转动有关, 阻力较大, 该二面角的变化幅度较小。二面角 $C_{10}-C_{17}-N_{18}-C_{20}$ 主要介于 $-150^\circ \sim -50^\circ$ 之间, 其变化幅度较大, 原因是 4 个原子都是单键连接, 转动能垒较低, 它的两边都是芳环。由于这个二面角柔性较大, 分子可以较好的与受体对接。后面的 3 个二面角都是由单键相连的原子组成。二面角 $C_{26}-N_{27}-C_{29}-C_{30}$ 主要介于 $70^\circ \sim 110^\circ$ 之间。二面角 $C_{29}-C_{30}-C_{31}-C_{32}$ 在 180 度左右小幅变化, 表明 4 个原子几乎停留在 1 个平面之上。二面角 $C_{30}-C_{31}-C_{32}-O_{33}$ 角度变化范围极大, 几乎没有固定值, 二面角自由变化, 主要是由于 C-O 键的自由转动。作者注意到二面角 $C_{29}-C_{30}-C_{31}-C_{32}$ 和 $C_{30}-C_{31}-C_{32}-O_{33}$ 变化产生不同的构型的变化趋势与势能的变化趋势基本一致, 并且它们的变化与强氢键的形成有关, 所以在搜索能量最低构象时应着重考虑这两个二面角。

2.3 配体与受体之间的相互作用 众所周知, 配体和受体是在不断运动的。分子动力学可以模拟其随时间的变化, 并计算其各种性质的统计平均值以了解药物分子与配体结合的情况, 为评估它成为先导化合物的可能性, 也为修饰该化合物提供依据。图 6 所示的是配体亚叶酸分子的一些原子与周围受体原子的相关径向分布函数, 它代表了一个原子周围找到另一个在特定位置上的原子的几率。原子的编

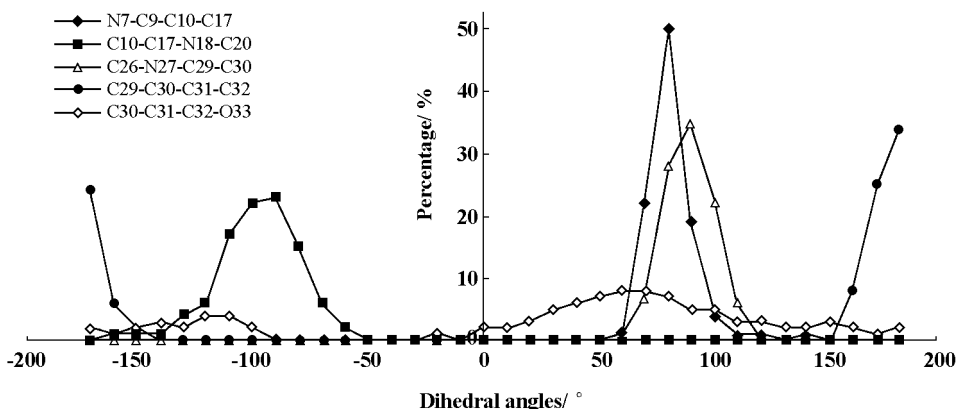


Figure 5 The statistical distribution of the important dihedral angles in arbitrary unit (the angle is in degrees), \diamond , \square , \triangle , \bullet and \diamond represent dihedral angles $N_7-C_9-C_{10}-C_{17}$, $C_{10}-C_{17}-N_{18}-C_{20}$, $C_{26}-N_{27}-C_{29}-C_{30}$, $C_{29}-C_{30}-C_{31}-C_{32}$ and $C_{30}-C_{31}-C_{32}-O_{33}$, respectively

号如图 4 所示。O18 与 H152 和 H154 都可以形成氢键。二者都是氮连接的氢,由图所示,它们形成氢键的几率几乎相同。而 H1263 与 O19 之间只是范德华作用,比氢键弱许多,所以其峰较宽。

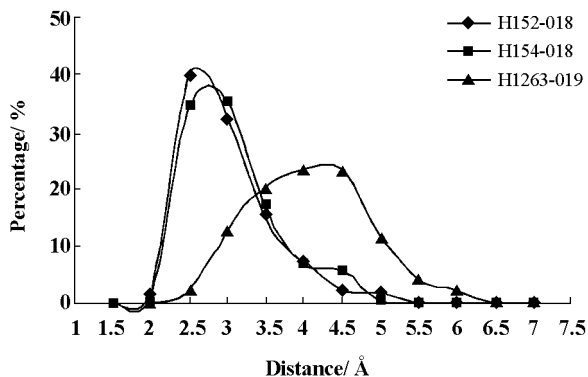


Figure 6 The radial distribution functions between atoms of ligand and those of the residues of the receptor, \diamond , \square , and \triangle represent the radial distribution functions between H152-O18, H154-O18 and H1263-O19, respectively

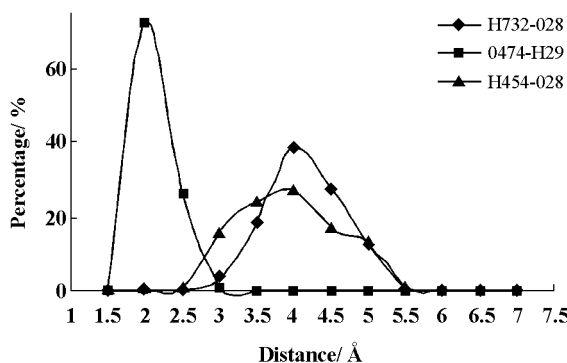


Figure 7 The radial distribution function between atoms of ligand and those of residues of the receptor, \diamond , \square and \triangle represent the radial distribution functions between H732-O28, H29-O474 and H454-O28, respectively

从图 7 可以看出, O474 与 H29 形成了很强的氢键,另外两对之间的相互作用较弱,活动范围较大,峰较宽。受体蛋白酶与亚叶酸分子可形成 3 个氢键,分别为受体蛋白残基 ARG8(左)的 154 号氢原子与亚叶酸分子的 57 号氧原子, ASP30 的 474 号氧原子与亚叶酸分子的 29 号氢原子, ARG8(中间)的 152 或 154 号氢原子与亚叶酸分子的 18 号氧原子。这些氢键受热运动的影响随时间变化,较弱氢键键长的变化幅度较大,但它对于配体分子在靶点的位置起重要作用。

当模拟达到平衡后,收集统计模拟数据,分析小分子配体与受体相互作用及配体各基团与受体残基作用情况。表 1 是无水分子和有水分子存在时的模拟结果,其中 VDW 和 Elec 分别表示范德华作用和静电作用能。

Table 1 The energetics of ligand in complex with the receptor ($\text{kcal} \cdot \text{mol}^{-1}$)

	VDW	Elec	Total binding energy
Without water	- 56.607	- 14.430	- 71.037
With water	- 63.315	- 28.509	- 91.824

VDW: Van Der Waals interaction; Elec: Electrostatic interactions

由无水的能量数据可以看出,配体亚叶酸分子与受体 HIV 蛋白酶 (1a30) 总结合自由能较低,配体与受体之间发生的范德华作用和静电作用同等重要。从有水的能量数据可以看出,加水后的模拟体系比无水参与的模拟体系的范德华作用及静电作用均有所增强,但静电作用增强幅度较大,总结合自由能因此下降,基团活性有所提高。

结论

通过分子对接方法得到的艾滋病蛋白酶与亚叶酸分子最佳构象的复合结构,并进行分子动力学模拟,可以了解配体功能团与受体残基结合的细节和变化规律。本文得到的 HIV-1 蛋白酶与亚叶酸分子的结构模式信息将有助于设计和改造出效果更好的抗 HIV-1 蛋白酶的抑制剂。下一步的工作是亚叶酸的活性实验研究并根据实验结果提出修饰亚叶酸分子的方案。

References

- [1] Debouck C. The HIV-1 protease as a therapeutic target for AIDS [J]. AIDS Res Hum Retroviruses, 1992, 8: 153 - 164.
- [2] Weber IT. Comparison of the crystal structure and intersubunit interaction of human immunodeficiency and rous sarcoma virus proteases [J]. Biol Chem, 1990, 265: 10492 - 10495.
- [3] Qiao XB, Hou TJ, Zhang W, et al. A 3D structure database of components from Chinese traditional medicinal herbs [J]. Chem Inf Comp Sci, 2002, 42: 481 - 495.
- [4] Zhou JJ, Xie GR, Yan XJ, et al. Traditional Chinese Medicines: Molecular Structures, Natural Sources and Applications (中药原植物化学成分手册) [M]. Beijing: Chemical Industry Press, 2005.

- [5] Sheridan RP, Miller MD, Underwood DJ, et al. Chemical similarity using geometric atom pair descriptors [J]. Chem Inf Comp Sci, 1996, 36: 128 - 136.
- [6] Brown RD, Martin YC. Use of structure-activity data to compare structure-based clustering methods and descriptors for use in compound selection [J]. Chem Inf Comp Sci, 1996, 36: 572 - 584.
- [7] Kearsley S. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap [J]. Tetrahedron Comput Methodol, 1990, 3: 615 - 633.
- [8] Martin YC, Bures MG, Danaher EA, et al. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists [J]. Comput Aided Mol, 1993, 7: 83 - 102.
- [9] Miller MD, Sheridan RP, Kearsley SK, et al. A program for rapidly producing pharmacophorically relevant molecular superpositions [J]. Med Chem, 1999, 42: 1505 - 1514.
- [10] Chen KX. Computer Aided Drug Design——principle, method and application (计算机辅助药物设计——原理、方法及应用) [M]. Shanghai: Shanghai Science and Technology Publishers, 2001: 270.