

# 植物挥发性有机物数据库的建立及应用

张彦博 钟国华 胡美英 柳晓磊 耿鹏 陈光毅

华南农业大学天然农药与化学生物学教育部重点实验室, 广州 510642

**摘要** 通过收集已有植物挥发性有机物信息,利用 MDL ISIS Base 化学信息管理系统,构建了植物挥发性有机物数据库;同时,运用 Discovery Studio 模拟结合板块,对数据库进行了虚拟筛选。该数据库包含 2 410 个挥发性有机物的分子式、结构式、分子质量、CAS 号等基本参数;以斜纹夜蛾普通气味结合蛋白 II (SlitGOBP II) 为靶标,应用此数据库对其进行虚拟筛选并得到 6 个备选活性化合物,其中泼尼松(prednisone)和 SlitGOBP II 能够结合成稳定的复合物。植物挥发性有机物数据库的构建,以及 Discovery Studio 的分子模拟对接技术的应用,可为今后新型农药的研制提供理论依据。

**关键词** 植物挥发性有机物;数据库;ISIS BASE;分子对接;Discovery Studio

**中图分类号** TQ 453.3 **文献标识码** A **文章编号** 1000-2421(2011)05-0604-05

目前已鉴定的植物挥发性有机物(plant volatile organic compounds, PVOs)达 1 700 余种,主要为烃类、烯烃类、醇类和脂肪酸衍生物等类的化合物。随着检测技术的提高,植物挥发性有机物种类还在不断增加<sup>[1-3]</sup>。植物挥发性有机物不仅能够吸引传粉者、促进种子传播,而且还可以通过抑制其他植物的生长发育来提高自身的竞争能力<sup>[4-5]</sup>。就植食者取食后诱导植物产生的挥发性有机物而言,国内外学者已对其类型、传导途径、释放机制等进行了综述<sup>[2,6-10]</sup>。植食性昆虫在寻找寄主阶段,大都通过嗅觉感受器接收并识别寄主植物的挥发性有机物,而这些有机物会诱导植食性昆虫产生寄主定向行为或逃避行为;其次,植食性昆虫的种群增殖受植物挥发性有机物的明显影响,不但影响植食性昆虫的交配、产卵等繁殖行为,还影响其取食行为<sup>[11]</sup>。因此,收集整理已有的植物挥发性有机物对其进行系统分析,成为新药研发的重要环节。

但是,传统的生测方法,每测试一个化合物都需耗费大量的人力、物力和财力,如果使用这种传统的方法对数千个挥发性有机物进行逐一测试,其难度可想而知。随着计算机技术的不断发展,计算机辅助药物设计<sup>[12]</sup>已经成为新药研发的一种重要手段,以数据库为基础工具,运用分子对接、药效团等技术

在虚拟药物筛选中应用广泛<sup>[13-16]</sup>。分子对接起源于 Fisher 于 1894 年提出的“锁与钥匙”学说,是把配体分子放在受体活性位点的位置,按照几何互补、能量互补以及化学环境互补的原则来实时评价配体与受体相互作用的好坏,并找到两个分子之间最佳结合模式的过程。目前,分子对接技术已广泛应用于基于受体药物设计的数据库搜寻以及虚拟组合库设计和筛选的研究中<sup>[17-19]</sup>。

为了通过分子模拟的方法研究植物挥发性有机物与昆虫气味相关蛋白之间的关系,笔者收集了已有报道的植物挥发性有机物,构建了植物挥发性有机物数据库(plant volatile organic compound database, PVOCD),并以斜纹夜蛾(*Spodoptera litura*)的普通气味结合蛋白 II (general odorant binding protein II, GOBP II)(简称 SlitGOBP II)为靶标蛋白,通过 Discovery Studio 软件模拟与 PVOCD 中的化合物分子对接,对此数据库进行了筛选,旨在验证其可靠性和该方法的可行性,为研发新一代植物性杀虫剂提供理论依据。

## 1 材料与方法

### 1.1 化合物信息来源

构建数据库的化合物信息来自国内外文献报道

收稿日期: 2010-12-21

基金项目: 国家自然科学基金项目(31071713)和全国优秀博士学位论文作者专项资金项目(2004061)

张彦博, 博士研究生。研究方向: 天然源农药。E-mail: fiend\_wing\_bo@yahoo.com.cn

通讯作者: 胡美英, 教授。研究方向: 昆虫毒理。E-mail: humy@scau.edu.cn

的各种 POVCs,同时在化学文摘数据库中利用相关关键词搜索 PVOCs 相关的信息。

## 1.2 应用软件

在构建 PVOCD 时使用 ISIS BASE 数据库软件(MDL 公司),在 PVOCD 的应用中使用 Discovery Studio 2.0 分子模拟软件(Accelry 公司)。

## 1.3 PVOCD 构建的方法

将搜集得到的化合物信息使用 MDL ISIS BASE 数据库软件构建 PVOCD,一级目录包括序号、名称、结构、分子式、分子质量和 CAS(chemical abstracts service)号,二级目录包括氢受体、氢供体、可旋转键数、脂水分配系数(clogP)和 5 规则(rule of 5)符合数。

## 1.4 PVOCD 应用的步骤

使用 Discovery Studio 2.0 软件完成以下的模拟试验。首先,使用 Discovery Studio 软件的 Prepare Ligands 将植物挥发性有机物库中导出的每个小分子进行原子类型和化学键归属,将 2D 结构转变成 3D 结构并进行结构优化,组成 3D 小分子数据库。使用 Build Homology Models 同源模建 SlitGOBP II 的三维结构,并使用 Verify Score 和拉式图(Ramachandran Plot)对其模型进行可靠性分析。使用 Binding Site 搜索 SlitGOBP II 编码蛋白可能的活性位点,位点分析过程采用 CHARMm 力场。使用 Ligandfit 刚性对接程序对植物挥发性有机物库进行初筛,所得结果采用 Score Ligand Pose 程序中的 5 种不同打分函数进行评价,并采用 Consensus Score 程序对其进行打分排列,根据一致性打分结果选择 Consensus Score $\geq 4$  的化合物。初筛所

得化合物使用 CDOCKER 对接程序进行复筛,对接结果中的能量评价项:—COOCKER\_Interaction\_Energy 与初筛中的 5 项打分函数结果共 6 项函数再使用 Consensus Score 进行二次评价,根据一致性打分结果选择 Consensus Score $\geq 5$  的化合物。为了进一步确定 SlitGOBP II 编码蛋白与植物挥发性有机物对接时哪些氨基酸起到关键作用,本研究采用 Discovery Studio 软件中 Caculate Interaction Energy 程序对植物挥发性有机物与活性位点每个氨基酸的结合自由能计算。

# 2 结果与分析

## 2.1 PVOCD 的构建

通过查阅大量文献,收集了 PVOCs 近 3 000 个,去除重复的以及找不到 2D 结构的化合物,最后共收集到 PVOCs 共 2 410 个。搜集得到的化合物信息主要包括以下内容:PVOCs 的名称、结构、分子式、分子量和 CAS 登记号;已确定结构的 PVOCs 二维化学结构式及利用分子力学方法优化得到的三维结构式;PVOCs 的分子性质,包括氢受体、氢供体、可旋转键数、脂水分配系数和 5 规则符合数。植物挥发性有机物数据库基本架构和基本界面如图 1 所示。

PVOCD 涵盖的化合物类别及其数量:烃类 299,酮类 214,烯炔类 347,酸类 135,醇类 378,酯类 600,酚类 71,胺类 87,醚类 45,醛类 99,其他 135。其中烯炔类、醇类和酯类的 3 类化合物之和占整个数据库的 55%。可见这 3 类化合物在挥发性化合物中所占比例较大,也是研究中比较常见的。

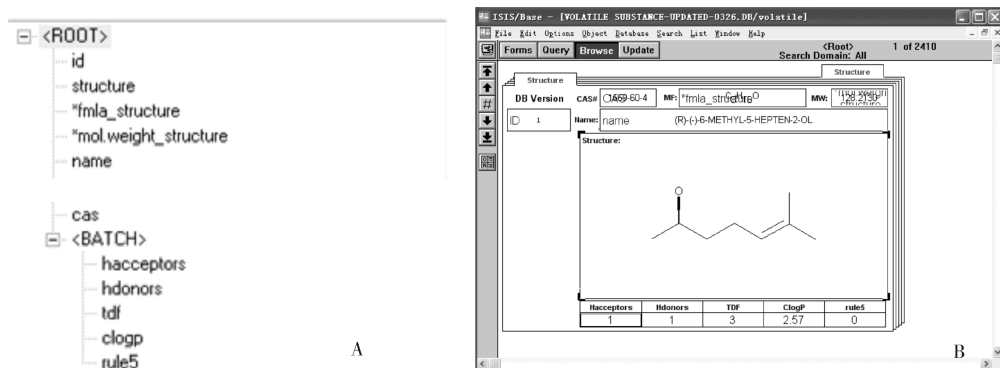


图 1 植物挥发性有机物数据库基本架构(A)和基本界面(B)

Fig.1 The basic framework(A) and basic interface(B) of PVOCD

## 2.2 PVOCD 的应用

为验证所构建的数据库是否能针对靶标快速筛

出相应的化合物,笔者选择斜纹夜蛾的普通气味结合蛋白 SlitGOBP II 作为靶标进行测试。首先,使用

Build Homology Models 程序模拟构建 SlitGOBP II 的三维结构,得到 10 种备选结构,使用 Verify Score 对模型的可靠性进行评价,结果如图 2 所示,除 N 端和 C 端(肽链的两端的得分允许小于 0)外,绝大部分氨基酸残基的得分均大于 0,这说明几乎所有的氨基酸残基都位于合理位置;使用拉式图对模型的可靠性进行评价,结果如图 3 所示,模建蛋白位于蓝色和紫色区域内的氨基酸残基结构>97%,因此模建蛋白的氨基酸残基结构也是比较合理的。由上述的 Verify Score 结果和拉式图分析可知,同源模建的 SlitGOBP 的三维结构是可靠的,可以用于下一步的分子对接(图 4)。

使用 Ligandfit 刚性对接程序对植物挥发性有机物库进行初筛,此对接程序在靶标蛋白的疏水口袋中产生刚性的格点(图 5),将配体化合物逐一与靶点进行匹配,共有 1 107 个化合物对接成功,其余

1 303 个化合物对接失败。所得结果采用 Score Ligand Pose 程序中的 5 种不同打分函数,并采用 Consensus Score 程序对其进行打分排列,此程序可选出在 5 种打分结果中排名均在前 10 的化合物,根据一致性打分结果选择 Consensus Score $\geq 4$  的化合物共 429 个。初筛所得化合物使用 CDOCKER 对接程序进行复筛,此程序是基于 CHARMm 的柔性对接程序,采用 soft-core potentials 和 optional grid representation 将配体分子与受体活性位点进行对接。首先采用动力学的方法随机搜索小分子构象,随后采用模拟退火的方法将各个构象在受体活性位点区域进行优化,从而使对接结果更加准确。对接结果中的能量评价项—CDOCKER\_Interaction\_Energy 与初筛中的 5 项打分函数结果共 6 项函数再使用 Consensus Score 进行二次评价,根据一致性打分结果选择 Consensus Score $\geq 5$  的化合

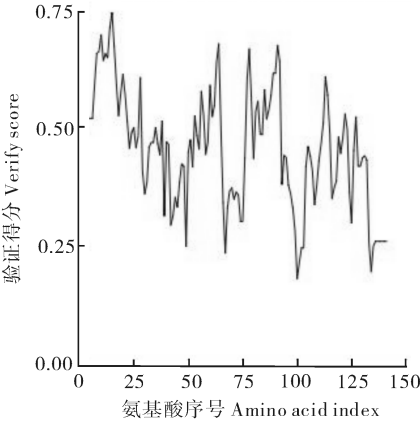


图 2 SlitGOBP II 三维模型的 Verify Score 结果  
Fig.2 The Verify Score of SlitGOBP II

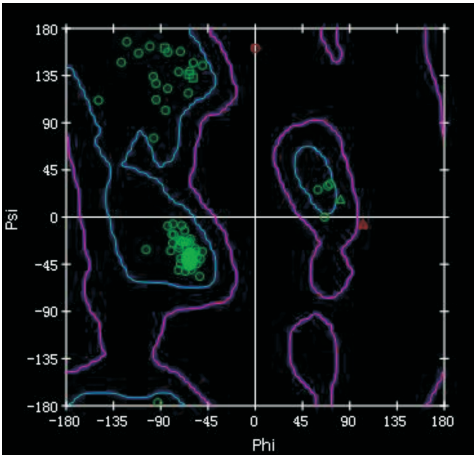


图 3 SlitGOBP II 三维模型的拉式图分析  
Fig.3 The Ramachran Plot of SlitGOBP II

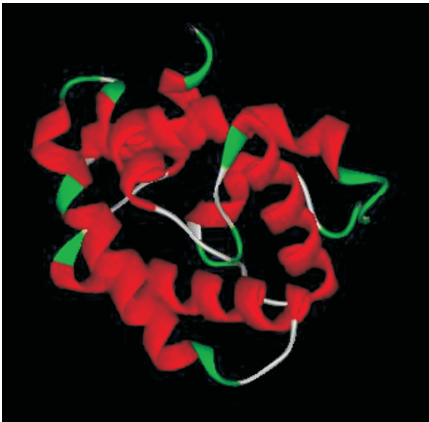


图 4 SlitGOBP II 的三维结构模型  
Fig.4 The 3D model of SlitGOBP II

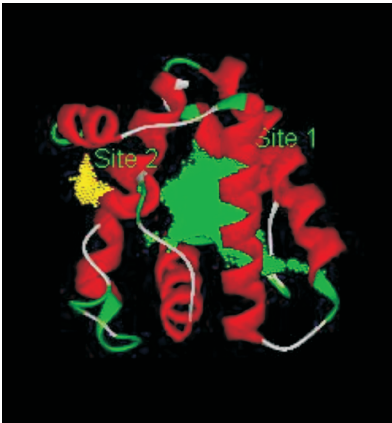


图 5 SlitGOBP II 三维模型的结合位点预测  
Fig.5 The binding site prediction of SlitGOBP II

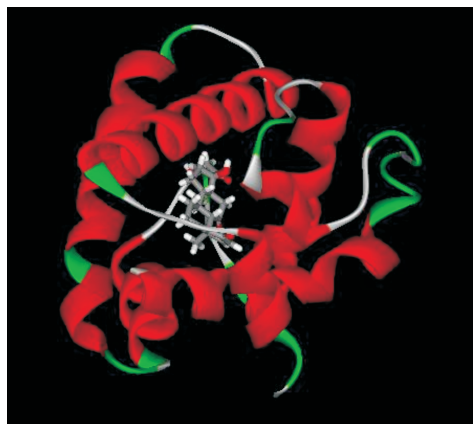


图6 Prednisone 与 SlitGOBP II 的结合模式图

Fig. 6 The docked mode of prednisone and SlitGOBP II

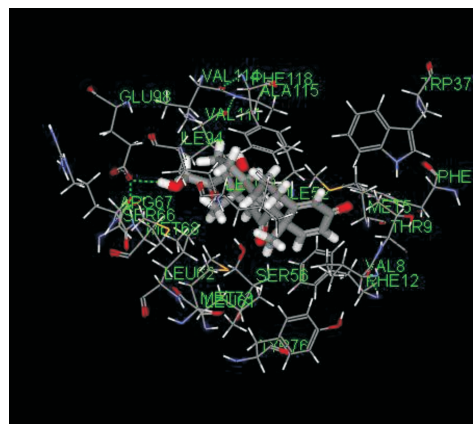


图7 Prednisone 与 SlitGOBP II 活性位点的结合模式图(绿色虚线表示氢键)

Fig. 7 The binding mode of prednisone with the binding site of SlitGOBP II

(green broken lines represent hydrogen bonds)

物共 137 个。在 6 种打分函数结果最好的前 10 种化合物中, prednisone 和 sucrose 同时出现在 3 项打分函数排名前 10 的化合物中, pregnanolone、strychnine、1H, 1H, 2H, 2H-perfluoro-1-decanol 和 lactose 同时出现在 2 项打分函数排名前 10 的化合物中。使用 Calculate Interaction Energy 进一步研究 prednisone、sucrose、lactose、pregnanolone、1H, 1H, 2H, 2H-perfluoro-1-decanol 和 strychnine 6 种化合物与 SlitGOBP II 的结合发现 prednisone 与 SlitGOBP II 的结合作用最强(图 6), prednisone 位于 SlitGOBP II 的  $\alpha$  螺旋所围成的疏水孔穴中, 而且 prednisone 的 1 个氢原子与 SlitGOBP II 的 98 位谷氨酸的氧原子形成 1 个氢键在结合中起重要作用(图 7), 因此推测 prednisone 和 SlitGOBP II 能够结合成稳定的复合物。

### 3 讨论

灵敏的嗅觉对于昆虫的正常生存和环境的适应具有重要作用。自从 Viehovey 于 1918 年首次报道水汽蒸馏棉叶而得到的挥发性油能够吸引棉铃象甲以来, PVOCs 在昆虫行为调控和生物防治中的作用越来越受到重视, 日益显示出其在生态系统中的作用。尤其在今后, 因过分使用有机化学农药而造成的害虫抗药性、害虫再猖獗和环境污染等问题, 迫切需要寻求以化学防治以外的有效方法控制害虫, 利用 PVOCs 作为引诱剂可以提高杀虫剂的效能, 或作为忌避剂干扰害虫的寄主识别与寄主定位、交配和寻找产卵场所等行为。由于蛋白质的三维结构在很大程度上决定了其功能, 根据此原理, 通过研究化合物与靶标蛋白在结构上的契合来寻找靶标蛋白的活性化合物, 这种方法比随机性的通过生测的方法寻找活性化合物更具合理性和针对性, 因此这也成为通过计算机模拟其结构来寻找活性化合物的理论依据。笔者使用计算机辅助药物设计的方法, 以昆虫气味相关的某特定蛋白为靶标, 通过基于分子对接的虚拟筛选方法就可以去除掉大部分不合适的化合物, 而留下一小部分在结构契合层面上分析有可能对靶标蛋白有活性的化合物, 再通过生测方法确认对昆虫行为有影响的化合物。此方法可以大幅度缩减发现活性化合物的研究成本及其所需的时间, 这对于加速研发新一代的昆虫引诱剂和忌避剂具有重要意义。

本试验首次对 PVOCs 进行了收集、归类 and 整理, 并构建了 PVOCs 的数据库, 这可为未来系统研究 PVOCs 的特点和作用提供参考, 同时对昆虫引诱剂和忌避剂的研发也有一定的推动作用。

### 参考文献

- [1] PICHERSKY E, GERSHENZON J. The formation and function of plant volatiles: perfumes for pollinator attraction and defense[J]. Current Opinion in Plant Biology, 2002(5): 237-243.
- [2] DUDAREVA N, NEGRE F, NAGEGOWDA D A, et al. Plant volatiles: recent advances and future perspectives[J]. Critical Reviews in Plant Sciences, 2006, 25: 417-440.
- [3] DUDAREVA N, NEGRE F. Practical applications of research into the regulation of plant volatile emission[J]. Current Opinion in Plant Biology, 2005(8): 113-118.
- [4] GOFF S A, KLEE H J. Plant volatile compounds: sensory cues for health and nutritional value? [J]. Science, 2006, 311: 815-



- 819.
- [5] SINGH H P, BATISH D R, KAUR S, et al. Comparative phytotoxicity of four monoterpenes against *Cassia occidentalis* [J]. *Annals of Applied Biology*, 2002, 141: 111-116.
- [6] LOU Y G, CHENG J A. Herbivore-induced plant volatiles: primary characteristics, ecological functions and its release mechanisms [J]. *Acta Ecologica Sinica*, 2001, 20 (6): 1097-1106.
- [7] KONG C H, HU F. Advance in the research on chemical communication between plants [J]. *Acta Phytoecologica Sinica*, 2003, 27: 561-566.
- [8] GUI L Y, LIU S S, CHEN Z M. Plant resistance to insects induced by application of exogenous jasmonic acid and methyl jasmonate [J]. *Acta Entomologica Sinica*, 2004, 47 (4): 507-514.
- [9] YANG W W, LI Z Z, AN Y, et al. Plant volatile organic compounds (VOCs): a review [J]. *Chinese Journal of Ecology*, 2008, 27(8): 1386-1392.
- [10] CAI X M, SUN X L, DONG W X, et al. Herbivore-induced plant volatiles: from induction to ecological functions [J]. *Acta Ecologica Sinica*, 2008, 28(8): 3969-3980.
- [11] 卢伟, 侯茂林, 文吉辉, 等. 植物挥发性次生物质对植食性昆虫的影响 [J]. *植物保护*, 2003, 33(3): 7-11.
- [12] GARALAND R M. Computer-aided drug design [J]. *Annual Review of Pharmacology and Toxicology*, 1987, 27: 193-213.
- [13] MESCH S, MOSER D, STRASSER D S, et al. Low molecular weight antagonists of the myelin-associated glycoprotein: synthesis, docking, and biological evaluation [J]. *Journal of Medicinal Chemistry*, 2010, 53(4): 1597-1615.
- [14] LARSEN C A, BISSON W H, DASHWOOD R H. Tea catechins inhibit hepatocyte growth factor receptor (MET Kinase) activity in human colon cancer cells: kinetic and molecular docking studies [J]. *Journal of Medicinal Chemistry*, 2009, 52 (21): 6543-6545.
- [15] EKONOMIUK D, SU X C, OZAWA K, et al. Flaviviral protease inhibitors identified by fragment-based library docking into a structure generated by molecular dynamics [J]. *Journal of Medicinal Chemistry*, 2009, 52(15): 4860-4868.
- [16] GROLLA A A, POSESTA V, CHION M G, et al. Synthesis, biological evaluation, and molecular docking of ugi products containing a zinc-chelating moiety as novel inhibitors of histone deacetylases [J]. *Journal of Medicinal Chemistry*, 2009, 52(9): 2776-2785.
- [17] WU J S, PENG Y H, WU J M, et al. Discovery of non-glycoside sodium-dependent glucose co-transporter 2 (SGLT2) inhibitors by ligand-based virtual screening [J]. *Journal of Medicinal Chemistry*, 2010, 53(24): 8770-8774.
- [18] MOHAMMAD A K, RONLD A H, AMAL K, et al. Discovery of novel GSK-3 $\beta$  inhibitors with potent in vitro and in vivo activities and excellent brain permeability using combined ligand- and structure-based virtual screening [J]. *Journal of Medicinal Chemistry*, 2010, 53(24): 8534-8545.
- [19] WEI D G, ZHENG H, SU N F, et al. Binding energy landscape analysis helps to discriminate true hits from high-scoring decoys in virtual screening [J]. *Journal of Chemical Information and Modeling*, 2010, 50(10): 1855-1864.

## Development and application of database of plant volatile organic compounds

ZHANG Yan-bo ZHONG Guo-hua HU Mei-ying LIU Xiao-lei GENG Peng CHEN Guang-yi

*Key Laboratory of Natural Pesticide and Chemical Biology, Ministry of Education,*

*South China Agricultural University, Guangzhou 510642, China*

**Abstract** The study aimed to build the plant volatile organic compound database (PVOCD), investigate the relationship between PVOCs and insect behavior and provide theoretic basis for the development of new pesticides. The PVOCD was constructed on the basis of MDL ISIS Base scientific information management system and its information was collected from papers published at home and abroad. A virtual screening was carried out on the PVOCD by Discovery Studio. The PVOCD contained 2 410 compounds, including molecular formula, structure, molecular weight, CAS registry number and so on. 6 candidate compounds for general odorant binding protein II of *Spodoptera litura* (SlitGOBP II) were screened out from PVOCD by molecular docking, and prednisone was the best one binding with SlitGOBP II. The construction of PVOC and the application of molecular docking could provide theoretic basis for the development of new pesticide.

**Key words** plant volatile organic compounds; database; ISIS BASE; molecular docking; Discovery Studio

(责任编辑:陈红叶)