

Reconstruction of *Erwinia carotovora* subsp. *atroseptica* SCRI1043 Metabolic Network and Its Application in Screening Potential Targets

Jun Wang, Cheng Wang, Dexin Kong, Lingling Chen*

College of Life Science and Technology, Huazhong Agricultural University, Wuhan
Email: *llchen@mail.hzau.edu.cn

Received: Feb. 10th, 2012; revised: Feb. 25th, 2012; accepted: Mar. 2nd, 2012

Abstract: *Erwinia carotovora* subsp. *atroseptica* SCRI1043 (*Eca* SCRI1043) is a widespread phytopathogen that causes blackleg and soft rot disease in potatoes. In this paper, we reconstructed the metabolic network of *Eca* SCRI1043 based on its genomic information. Through the topology and flux balance analysis, hub nodes of the network were selected. After that TTD database was used to screen those hubs and find out the candidate targets. Undecaprenyl pyrophosphate synthetase (Upps) was chosen to do homology modeling and virtual screening by using the commercialize compounds database provided by specs company. Finally, 73 compounds were screened manually in the top scoring 400 compounds.

Keywords: *Erwinia carotovora* subsp. *atroseptica* SCRI1043; Metabolic Networks; Flux-Balance Analysis; Homology Modeling; Virtual Screening

胡萝卜软腐欧文氏菌代谢网络重构及其在靶标筛选中的应用

王君, 王成, 孔德信, 陈玲玲*

华中农业大学生命科学技术学院, 武汉
Email: *llchen@mail.hzau.edu.cn

收稿日期: 2012年2月10日; 修回日期: 2012年2月25日; 录用日期: 2012年3月2日

摘要: 欧文氏杆菌(*Erwinia*)是一类重要的农作物致病细菌, 侵染宿主范围广, 在世界范围内造成了严重的经济损失, 其中 *Erwinia carotovora* subsp. *atroseptica* SCRI1043 (*Eca* SCRI1043) 可导致马铃薯感染黑胫软腐病, 危害极大。本文构建了 *Eca* SCRI1043 基因组的代谢网络, 并进行拓扑结构和流平衡分析, 以此筛选出网络的中心节点。若中心节点所代表的酶收录在 TTD 数据库中, 则此酶可作为农用杀菌剂的候选靶标。随后选取靶标十一异戊二烯焦磷酸酶(Upps), 采用 specs 公司的商品化合物数据库作为小分子数据库, 对靶标进行高通量虚拟筛选, 对综合打分结果最好的前 400 个化合物进行目筛后得到 73 种先导化合物, 这 73 种先导化合物可以进一步进行后续生测实验以确定其杀菌活性。

关键词: 胡萝卜软腐欧文氏菌; 代谢网络重构; 流平衡分析; 同源模建; 虚拟筛选

1. 引言

细菌性软腐病作为一种分布极广的植物病害, 在

全球造成了严重的经济损失, 其中胡萝卜软腐欧文氏菌(*Erwinia carotovora* subsp. *atroseptica* SCRI1043, *Eca* SCRI1043)是其主要病原菌, 可危害胡萝卜、马铃薯、大白菜等农作物, 并引起储存期蔬菜的腐烂^[1]。

*通讯作者。

Eca SCRI1043 是在土壤中传播的植物致病菌,可借助昆虫、地下害虫或灌溉水传播,从根茎部伤口或地上叶片气孔及水孔侵入,主要侵害地下部肉质根,病组织通常表现为软腐与湿腐,如果侵染维管束组织,能阻碍水分输送而导致地上茎叶出现萎蔫等症状^[2]。对于这种土传病害,目前尚无很好的治疗手段,急缺高效可靠的杀菌剂。

随着基因组及生物化学数据的大量累积,通过基因组学方法构建物种的代谢网络,可从系统角度出发,筛选对维持物种生存所必须的酶或核酸作为靶标,提高药物开发的效率^[3]。较农用杀菌剂而言,药物的研发流程相对成熟,积累了大量的实验方法和靶标数据,借鉴药物研发的策略来研究农用杀菌剂也是一种捷径,其中基于靶标的药物设计也必将成为农药杀菌剂研发的重要方向。

2. 材料与方法

2.1. 代谢网络的重构

全基因组尺度代谢网络的重构包括重构、模拟、评估及修正。其中模型的重构最为关键,它包括基因注释、代谢反应及反应列表的建立、代谢网络的建立、计量学矩阵的建立以及代谢途径的优化。近几年,国际上对代谢网络重构的研究如火如荼,出现了多种代谢网络重构工具,如 GEM System^[4]、metaSHARK^[5]、MODEL SEED^[6]等。其中 MODEL SEED 依托于比较基因组数据库 SEED 及其高精度基因组注释工具 RAST,可生成包含有完整的 Gene-Protein-Reaction (GPR)反应列表的初始代谢草图。不过目前的网络构建工具都只能生成初始草图,还需要对 GPR 反应列表进行进一步的验证和修正,而这一阶段仍然只能通过手工完成,所以说成功构建一个高精度的代谢网络是一个劳动密集的手工产物。Christopher S Henry 利用 MODEL SEED 构建了 130 个细菌的代谢网络,计算这 130 个网络的准确性得知:草图的准确性平均为 66%,优化后的网络的准确性为 87%^[6]。考虑到 SEED 数据库中含有 *Eca* 的基因组注释信息,故本文采用 MODEL SEED 生成胡萝卜软腐欧文氏菌(SCRI104)的代谢草图。

2.1.1. 反应方向的修正

代谢草图中的反应信息源于 KEGG,而 KEGG 作

为网络重构的主要数据库,其反应方向存在很多错误。Ma 和 Zeng 对 KEGG 数据库的所有反应信息进行核查,并根据一定的生理学规则,提出了确立反应方向的 11 条规则^[7]。由于 *Eca* 与大肠杆菌属于肠杆菌科 (Enterobacteriaceae) 的不同属,亲源关系很近,本文即根据 MA 和 Zeng 提出的方法,并同时参考大肠杆菌等模式生物的反应方向,结合 BioCYC 等数据库来确定胡萝卜软腐欧文氏菌的反应方向。若现有的信息和方法都无法最终确定反应方向,则默认该反应是可逆的。

2.1.2. 通用代谢物的剔除

通用代谢物一般是用于电子转移以及一些功能基团的转移,如磷酸基团、氨基基团、一碳单元、甲基基团等。除了上述物质外,我们认为 H₂O、NH₃、O₂、CO₂、Phosphate 也是通用代谢物^[8]。通用代谢物参与众多反应,导致这些节点的连接度偏高,从而使通用代谢物成为网络的中心节点,而无法挖掘出真正有用的数据。考虑到当通用代谢物在反应中充当原始反应物时不能去除,故需手工对涉及通用代谢物的反应进行审查。

2.1.3. 网络的可视化

目前绝大多数代谢网络都是用节点代表代谢物,节点之间的连线代表代谢物之间的反应,而反应方向则由箭头表示。如 Ma 等人对可逆反应 R00006:2C00022 ↔ C00900 + C00011 表示为 C00022 ↔ C00900 和 C00022 ↔ C00011,其中 C00022、C00900、C00011 为均 KEGG 编号的化合物^[7]。考虑到本文构建代谢网络的主要目的是筛选合适的杀菌剂靶标,故在进行代谢网络可视化时,将催化反应的酶也用节点来表示,以便在后续拓扑筛选时,便于提取靶标信息。例如反应 A + B → C 是由酶 EC 催化的,则可表示为 A → EC, B → EC, EC → C。若一个反应是由多个酶催化的,则将反应拆分为多个,其中每个子反应有单一酶所控制。利用这种方式,可将整个菌种的代谢反应整理为 Excel 表格,便于利用 cytoscape 实现网络的可视化。

2.2. 拓扑分析

生物网络作为复杂网络,可以运用图论理论进行分析。其中连接度和中介性这两个参数可以反应节点

在网络中的影响力和重要性^[9], 故在对胡萝卜软腐欧文氏菌的代谢网络进行拓扑结构分析时, 认为最佳靶标应同时具有较高的连接度和中介性。

2.3. TTD 数据库筛选

TTD 数据库收集了成功的和正在研究的靶标以及靶标对应的疾病、药物、序列等信息^[10]。对拓扑分析筛选到的靶标, 将其蛋白序列与 TTD 数据库进行序列比对, 进一步筛选被 TTD 数据库收录的靶标。

2.4. 流平衡分析

Flux Balance Analysis (FBA) 是一种分析代谢网络的数学工具, 不需要代谢物浓度或者酶动力学信息便可预测生物必需的代谢反应。本文中采用 Crobra Toolbox v2.0^[11], 对通过拓扑分析和 TTD 数据库筛选的靶标, 计算代谢网络的流平衡。

2.5. 靶标的同源建模及虚拟筛选

SYBYL-X 是 Tripos 公司研发的一个相当全面的药物与分子设计专业工具, 提供了综合的分子模拟工具, 同源建模采用的是 SYBYL-X 1.2 中的 Orchestrar 模块。虚拟筛选则先用 Dock 6.4 进行初筛^[12], 随后用 Autodock 4.2 进行精筛^[13], 小分子数据库采用 specs 公司提供的商品化合物数据库, 以保证化合物的可购买性。

3. 结果与讨论

3.1. 代谢网络的分析

Eca 的代谢网络用上述“反应物 - 酶 - 生成物”的表示方式整理为 cytoscape 可读的 Excel 格式, 利用 cytoHubba 插件舍弃孤立反应后含有 1416 个节点, 2041 条边(见图 1)。为验证本文构建的代谢网络在结构上符合代谢网络的基本规律, 故对网络的节点度分布和平均路径长度进行了分析。

目前的研究表明, 代谢网络具有典型的“小世界”和“无标度”特征, 也就是说代谢网络具有较小的网络平均路径长度, 且节点的度符合幂律分布。分析 *Eca* SCRI1043 的代谢网络, 结果显示, 采用“反应物 - 酶 - 生成物”的方式表示的代谢网络也符合“小世界”及“无标度”的规律。

3.1.1. 小世界效应

Jeong 等人对 43 种不同物种的代谢网络进行研究后发现无论是古菌、原核还是真核, 其代谢网络的平均路长都约为 3.2, 表明任何两种代谢物之间平均只需 3 步即可完成转化^[14]。但他们的研究并未去除通用代谢物, 故平均步长过低, 与实际生化反应不一致。Ma 等人去除通用代谢物后重构了 80 个物种的代谢网络, 研究表明原核的平均路径长度为 7.73(不包含寄生虫), 且平均路径长度随着网络节点的增加逐渐增大^[7]。本文构建的胡萝卜软腐欧文氏菌的代谢网络的节点包含了反应的信息, 故节点数目较大为 1416, 依据 Ma 等人的研究可知, 本文构建的代谢网络的平均路长为 9.26 符合平均路径长度随着网络节点的增加逐渐增大的结论, 很好地反映了代谢网络的结构信息。几种典型的模式物种的代谢网络的平均步长也很好的验证了这点(见表 1)。

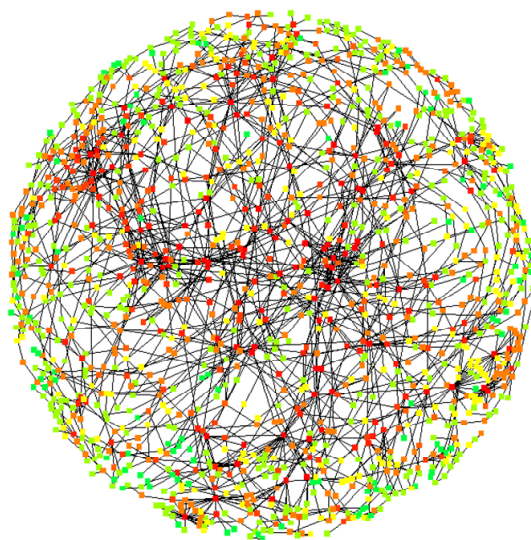


Figure 1. Metabolic network of *Eca* SCRI1043
图 1. *Eca* SCRI1043 的代谢网络

Table 1. Average distance and diameter of multi-bacteria
表 1. 模式物种及 *Eca* SCRI1043 的代谢网络的平均路径长度

物种	平均路径长度	网络直径
<i>B. subtilis</i>	8.48	23
<i>E. coli</i>	8.20	23
<i>H. influenzae</i>	8.35	27
<i>H. pylori</i>	7.91	24
<i>S. typhimurium</i>	8.22	24
<i>Eca</i> SCRI1043	9.26	28

从表 1 中可以看出, 胡萝卜软腐欧文氏菌代谢网络的平均路径长度为 9.26, 直径为 28, 仍在代谢网络规定的对数尺度范围内。考虑到 *Eca* 代谢网络节点的数目, 本文构建的代谢网络符合小世界的基本规律。

平均路径长度和网络的直径是评价网络节点间信息交流速度的重要依据。这表明, 胡萝卜软腐欧文氏菌代谢网络的大多数代谢物只需通过 10 步之内的反应就可以完成相互转化, 从而可以保证代谢物浓度的变化能迅速的传到整个网络。整个网络的直径为 28, 说明网络中相距最远的两个代谢物也只需 28 步就可完成相互转换。这两个指标意味着代谢物和酶的浓度的变化信息能够很快的传递给整个网络, 使得整个菌种能够迅速的对外界环境的变化作出反应。

3.1.2. 无标度网络

代谢网络节点度的概率密度函数为 $p(k) = ak^{-r}$, 其中 a 与 k 为常数, 幂指数 r 约为 2.2。符合这一条件即为无标度网络。无标度网络的关键节点控制着整个网络的结构, 使得网络对随机改变(如突变、环境的变化等)有很强的鲁棒性^[15]。Ma 和 Zeng 的研究表明所有代谢网络的中心节点相似^[7], 它们控制着网络的整体结构。鉴于本文研究目的的不同, 所构建的代谢网络中节点除了表示代谢物外, 还包含了反应信息, 以便筛选靶标信息。对 *Eca* SCRI1043 代谢网络的节点度分布取对数后进行统计分析(节点度参数见附录 1), 节点度分布符合公式: $p(k) = 1324.778 k^{-2.233}$ 。*Eca* SCRI1043 代谢网络节点度的幂指数为 2.233, 符合无标度网络的幂律分布(见图 2)。

从图中可以看出, 连接度高的节点数目较少, 绝大多数节点的连接度都较低, 也就是说, 少数连接度高的节点连接着为数众多的节点, 这些高连接度的节点控制着整个网络的结构, 对维持网络的稳定性起着重要的作用。

3.2. 靶标筛选

在对胡萝卜软腐欧文氏菌进行靶标筛选时, 首先筛选网络中节点的度和中介性都大于均值的节点, 得到 111 个候选酶。之后对这 111 个酶进行 TTD 数据库筛选, 得到 44 个候选靶标。随后用 FBA 验证这 44 个候选靶标对胡萝卜软腐欧文氏菌的生存是否重要, 得到 6 个对菌种维持生命所必须的靶标(见表 2)。

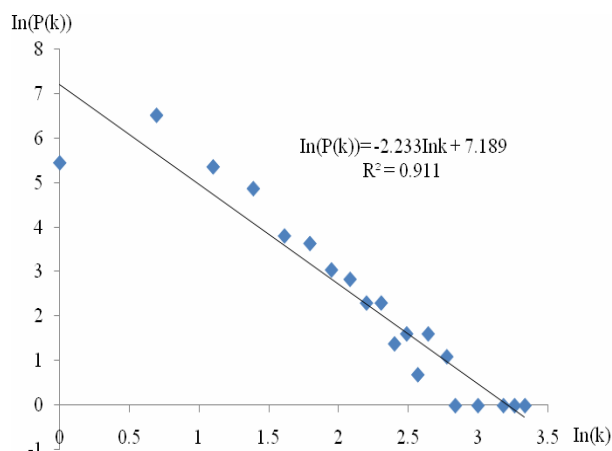


Figure 2. Degree distribution of the metabolic network of *Eca* SCR I1043

图 2. *Eca* SCRI1043 代谢网络的度分布

Table 2. Information of potential targets

表 2. 潜在靶标信息

酶号	名称	TTD 中靶标状态
2.3.1.41	Beta-ketoacyl-ACP synthase	Research target
2.4.1.227	Glycosyltransferase MurG	Research target
2.5.1.31	Undecaprenyl pyrophosphate synthetase	Research target
2.5.1.9	Riboflavin synthase	Research target
2.7.1.33	Pantothenate kinase	Research target
3.1.3.1	Alkaline phosphatase	Successful target

潜在靶标十一异戊二烯焦磷酸酶(酶号 2.5.1.31, 简称 Upps)在 PDB 数据库中有相似性很高的同源蛋白的晶体结构, 在宿主植物中的没有序列相似性高的同源蛋白, 因此选取 Upps 作为潜在靶标, 进行下一步的虚拟筛选工作。

3.3. 靶标 Upps 的同源模建

十一异戊二烯焦磷酸酶 Upps 是由 253 个氨基酸残基组成的单亚基蛋白质。它催化反应的产物 UPp 是细菌细胞壁主要成份肽聚糖生物合成中的重要载体, 功能是运载 N-乙酰胞壁酸五肽到细胞膜与 N-乙酰葡萄糖胺结合。作为一个新的潜在的抗生素靶标, 它的催化机制和结构构像得到了深入的研究^[16-19], 也有关于基于结构的抑制剂的筛选的报道^[18]。

在同源模建中, 模板的选择至关重要, 对结果的准确性起着决定性的影响。本文从 PDB 数据库中选

取了 *Escherichia coli* 的 Upps 结构 1X06 作为胡萝卜软腐欧文氏菌 Upps 的模板，二者是高度同源蛋白，序列全同性(Identity)为 80.3%，1X06 的晶体结构具有较高的分辨率(1.9 Å)，包含 3 个配体小分子，分别为镁离子，反应底物 IPP 以及 FPP 结构类似物 FSPP^[19]，故 1X06 是 *Eca* SCRI1043 的 Upps 进行同源建模的一个很好模板。

利用 SYBYL-X1.2 的同源建模模块 ORCHESTRAR 对 *Eca* SCRI1043 的 Upps 进行同源建模，建模的结果采用 Procheck 进行评价，拉氏图结果表明建模蛋白的主链结构合理：主链二面角 φ 、 ψ 落于中心区域和“可接受区域”的残基占 100%，绝大多数 φ - ψ 角均在正常范围内，没有落在不允许区域内的残基(图 3)。

建模蛋白的结构显示 Upps 是一个主要由 α 螺旋和 β 折叠组成的单体蛋白质分子(图 4)。从图 4 中可以看到 Upps 的活性口袋是一个狭长的疏水通道，由 4 个 β 折叠(黄色的条带)和 2 个 α 螺旋(红色的螺旋)组成了一个桶装空间结构。活性口袋中结合的配体小分子从上到下依次是镁离子，FSPP 和 IPP。

Asp25 是 Upps 的反应催化位点^[19]，Upps 与底物类似物 FSPP 的相互作用图(图 5)显示 Asp25 与 FSPP 作用密切，此外可以看出 FSPP 的活性部位是焦磷酸基团，而 Upps 的活性空腔中 Arg29, His42 和 Arg76 也是焦磷酸基团作用的位点。

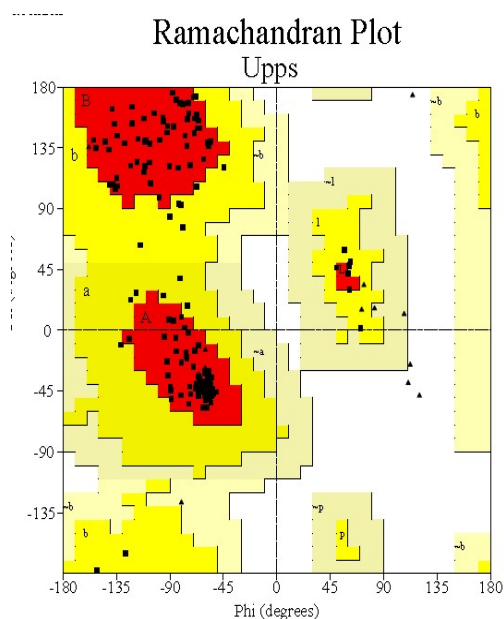


Figure 3. Ramachandran map of Upps in *Eca* SCRI1043
图 3. *Eca* SCRI1043 中 Upps 同源建模的 Ramachandran 图

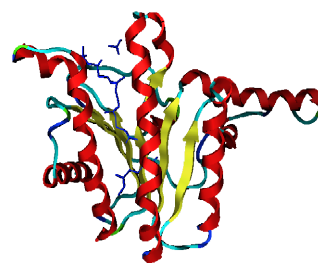


Figure 4. The homology modeling of Upps in *Eca* SCRI1043
图 4. *Eca* SCRI1043 中 Upps 的同源建模结构图

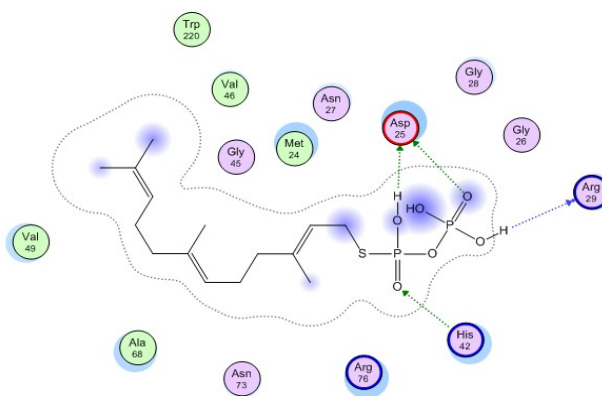


Figure 5. Interactions between Upps and substrate analogue FSPP
图 5. Upps 与底物类似物 FSPP 的相互作用图

3.4. 对 specs 数据库的虚拟筛选

使用 Dock 6.4 以同源建模的 *Eca* SCRI1043 Upps 为受体蛋白质，对 specs 数据库中的 10 万多个化合物进行初筛，采用 autoMS 产生分子表面，再对活性口袋部位分子表面的位点进行聚类，并以配体 FSPP 周围 4 Å 半径范围内的氨基酸残基来确定口袋部位的球集。然后用 GRID 获取口袋部位的网格点数据，并对初筛结构进行打分(Dock 6.4 中的对接参数见附录 2)，取打分前 1600 的化合物分子用 Autodock 4.2 进行精筛(Autodock 4.2 中的对接参数见附录 3)，分子对接的网格是以配体 FSPP 为中心，长宽高分别为 38 Å，20 Å，52 Å，网格间距 0.375 Å，包含了底物结合的全部空间，取打分前 400 位的化合物进行目筛，在目筛时没有考虑化合物和蛋白的结合能，筛选标准如下：1) 小分子占据和 FPP 尽量相同的结合位置，去掉大部分露在活性空间外的分子；2) 考虑小分子和活性空间的亲水和疏水空间及小分子和受体之间的氢键结合情况。最终得到 73 个潜在活性化合物。在后续工作中可进一步对这些化合物进行生测实验，鉴定其抑菌活性。

4. 结论

本文重构了胡萝卜软腐欧文氏菌的代谢网络, 通过拓扑结构分析并结合 TTD 数据库和 FBA 成功筛选到 6 个 *Eca* SCRI1043 维持正常生命活动所必须的酶, 为研制杀菌剂提供了潜在的作用靶标。并对其中的靶酶 Upps 进行了同源建模和虚拟筛选, 并获得了 73 个潜在的活性杀菌剂(见附录 4)。本文的研究表明利用代谢网络和靶标筛选的策略研发农药杀菌剂具有可行性, 下一步的工作将对筛选到的活性物质进行相应的抑菌试验, 以验证其活性。

5. 致谢

本文由国家自然科学基金(31071659)资助。感谢张红雨教授, 戢风琴博士和马彬广副教授的悉心指导。

参考文献 (References)

- [1] 远方, 屈淑平, 崔崇士. 一株新的胡萝卜软腐欧文氏菌的分离和鉴定[J]. 微生物学报, 2004, 44(2): 136-140.
- [2] 雷玉明, 张建朝, 邢会琴. 几种杀菌剂对胡萝卜软腐欧文氏菌的毒力测定[J]. 长江大学学报(自然版), 2010, 7(3): 3-5.
- [3] B. Elhanan, K. Martin, W. Marcus, et al. Large-scale reconstruction and phylogenetic analysis of metabolic environments. Proceeding of the National Academy of Science USA, 2008, 105(38): 14482-14487.
- [4] K. Arakawa, Y. Yamada, K. Shinoda, et al. GEM System: Automatic prototyping of cell-wide metabolic pathway models from genomes. BMC Bioinformatics, 2006, 7: 168.
- [5] J. W. Pinney, M. W. Shirley, G. A. McConkey, et al. metaSHARK: Software for automated metabolic network prediction from DNA sequence and its application to the genomes of *Plasmodium falciparum* and *Eimeria tenella*. Nucleic Acids Research, 2005, 33(4): 1399-1409.
- [6] C. S. Henry, M. DeJongh, A. A. Best, et al. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nature Biotechnology, 2010, 28(9): 977-984.
- [7] H. Ma, A. P. Zeng. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics, 2003, 19(2): 270-277.
- [8] M. Huss, P. Holme. Currency and commodity metabolites: Their identification and relation to the modularity of metabolic networks. IET Systems Biology, 2007, 1(5): 280-285.
- [9] 王卓, 陈琦, 刘雷. 代谢网络进化过程中拓扑结构与功能之间的关联[J]. 科学通报, 2009, 54(5): 776-782.
- [10] F. Zhu, Z. Shi, C. Qin, et al. Therapeutic target database update 2012: A resource for facilitating target-oriented drug discovery. Nucleic Acids Research, 2012, 40(D1): D1128-1136.
- [11] J. Schellenberger, R. Que, R. M. Fleming, et al. Quantitative prediction of cellular metabolism with constraint-based models: The COBRA toolbox v2.0. Nature Protocols, 2011, 6(9): 1290-1307.
- [12] P. T. Lang, S. R. Brozell, S. Mukherjee, et al. DOCK 6: Combining techniques to model RNA—small molecule Complexes. RNA, 2009, 15(6): 1219-1230.
- [13] G. M. Morris, R. Huey, W. Lindstrom, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry, 2009, 30(16): 2785-2791.
- [14] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabási. The large-scale organization of metabolic networks. Nature, 2000, 407(6804): 651-654.
- [15] S. H. Strogatz. Exploring complex networks. Nature, 2001, 410(6825): 268-276.
- [16] Y.-P. Lu, H.-G. Liu, K.-H. Teng, et al. Mechanism of cis-prenyl-transferase reaction probed by substrate analogues. Biochemical and Biophysical Research Communications, 2010, 400(4): 758-762.
- [17] W. Sinko, C. Oliveira, S. Williams, et al. Applying molecular dynamics simulations to identify rarely sampled ligand-bound conformational states of undecaprenyl pyrophosphate synthase, an antibacterial target. Chemical Biology & Drug Design, 2011, 77(6): 412-420.
- [18] C. J. Kuo, R. T. Guo, I. L. Lu, et al. Structure-based inhibitors exhibit differential activities against *Helicobacter pylori* and *Escherichia coli* undecaprenyl pyrophosphate synthases. Journal of Biomedicine and Biotechnology, 2008: Article ID 841312.
- [19] R. T. Guo, T. P. Ko, A. P. Chen, et al. Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: Roles of the metal ion and conserved residues in catalysis. The Journal of Biological Chemistry, 2005, 280(21): 20762-20774.

附录 1: *Eca* 代谢网络的连接度统计数据

Degree	Number of nodes
1	232
2	674
3	212
4	130
5	45
6	38
7	21
8	17
9	10
10	10
11	4
12	5
13	2
14	5
15	0
16	3
17	1
18	0
19	3
20	1
21	0
22	0
23	0
24	1
25	0
26	1
27	0
28	1

附录 2: Dock 6.4 中的对接参数

ligand_atom_file	specs.mol2
limit_max_ligands	no
skip_molecule	no
read_mol_solvation	no
calculate_rmsd	no
use_database_filter	no
orient_ligand	yes
automated_matching	yes
receptor_site_file	selected_spheres.sph
max_orientations	500
critical_points	no
chemical_matching	no
use_ligand_spheres	no
use_internal_energy	yes
internal_energy_rep_exp	12
flexible_ligand	yes
min_anchor_size	40

pruning_use_clustering	yes
pruning_max_orients	100
pruning_clustering_cutoff	100
pruning_conformer_score_cutoff	25.0
use_clash_overlap	no
write_growth_tree	no
bump_filter	no
score_molecules	yes
contact_score_primary	no
contact_score_secondary	no
grid_score_primary	yes
grid_score_secondary	no
grid_score_rep_rad_scale	1
grid_score_vdw_scale	1
grid_score_es_scale	1
grid_score_grid_prefix	grid
dock3.5_score_secondary	no
continuous_score_secondary	no
gbsa_zou_score_secondary	no
gbsa_hawkins_score_secondary	no
amber_score_secondary	no
minimize_ligand	yes
minimize_anchor	yes
minimize_flexible_growth	yes
use_advanced_simplex_parameters	no
simplex_max_cycles	1
simplex_score_converge	0.1
simplex_cycle_converge	1.0
simplex_trans_step	1.0
simplex_rot_step	0.1
simplex_tors_step	10.0
simplex_anchor_max_iterations	500
simplex_grow_max_iterations	500
simplex_grow_tors_premin_iterations	0
simplex_random_seed	0
simplex_restraint_min	no
atom_model	all
vdw_defn_file	vdw_AMBER_p arm99.defn
flex_defn_file	flex.defn
flex_drive_file	flex_drive.tbl
ligand_outfile_prefix	dockout
write_orientations	no
num_scored Conformers	1
rank_ligands	yes
max_ranked_ligands	1600

附录 3: Autodock 4.2 中的对接参数

```

DPF> autodock_parameter_version 4.2 # used by autodock to validate parameter set
DPF> outlev 1 # diagnostic output level
DPF> intelec # calculate internal electrostatics
DPF> seed pid time # seeds for random generator
DPF> ligand_types HD OA # atoms types in ligand
DPF> fld receptor.maps.fld # grid_data_file
DPF> map receptor.HD.map # atom-specific affinity map
DPF> map receptor.OA.map # atom-specific affinity map
DPF> elecmap receptor.e.map # electrostatics map
DPF> desolvmap receptor.d.map # desolvation map
DPF> move ligand.pdbqt # small molecule
DPF> about 0.378 0.6623 0.648 # small molecule center
DPF> tran0 random # initial coordinates/A or random
DPF> axisangle0 random # initial orientation
DPF> dihe0 random # initial dihedrals (relative) or random
DPF> tstep 2.0 # translation step/A
DPF> qstep 50.0 # quaternion step/deg
DPF> dstep 50.0 # torsion step/deg
DPF> torsdof 0 # torsional degrees of freedom
DPF> rmstol 2.0 # cluster_tolerance/A
DPF> extnrg 1000.0 # external grid energy
DPF> e0max 0.0 10000 # max initial energy; max number of retries
DPF> ga_pop_size 150 # number of individuals in population
DPF> ga_num_evals 2500000 # maximum number of energy evaluations
DPF> ga_num_generations 27000 # maximum number of generations
DPF> ga_elitism 1 # number of top individuals to survive to next generation
DPF> ga_mutation_rate 0.02 # rate of gene mutation
DPF> ga_crossover_rate 0.8 # rate of crossover
DPF> ga_window_size 10 #
DPF> ga_cauchy_alpha 0.0 # Alpha parameter of Cauchy distribution
DPF> ga_cauchy_beta 1.0 # Beta parameter Cauchy distribution
DPF> set_ga # set the above parameters for GA or LGA
DPF> sw_max_its 300 # iterations of Solis & Wets local search
DPF> sw_max_succ 4 # consecutive successes before changing rho
DPF> sw_max_fail 4 # consecutive failures before changing rho
DPF> sw_rho 1.0 # size of local search space to sample
DPF> sw_lb_rho 0.01 # lower bound on rho
DPF> ls_search_freq 0.06 # probability of performing local search on individual
DPF> set_psw 1 # set the above pseudo-Solis & Wets parameters
DPF> unbound_model bound # state of unbound ligand
DPF> ga_run 10 # do this many hybrid GA-LS runs
DPF> analysis # perform a ranked cluster analysis

```


附录 4: 虚拟筛选获得的 73 个候选化合物

idnumber	formula	logp	logs	name
AG-401/37257007	C17H10F3N5O	4.41	-4.273	1-phenyl-5-{5-[4-(trifluoromethyl)phenyl]-4-isoxazolyl}-1H-tetrazole
AK-968/11369248	C17H18N2O5S	2.37	-3.799	N-{4-[(acetylamino)sulfonyl]phenyl}-4-ethoxybenzamide
AG-690/15433317	C13H11N7O3S2	2.63	-3.136	N-(5-{4-nitrophenyl}-1,3,4-thiadiazol-2-yl)-2-[(5-methyl-4H-1,2,4-triazol-3-yl)sulfanyl]acetamide
AG-690/15435795	C23H23NO5	4.43	-5.766	2-(4-methylphenyl)-2-oxoethyl 2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)hexanoate
AG-690/15436763	C16H12BrNO2S	4.22	-5.855	N-(2-bromophenyl)-2-naphthalenesulfonamide
AG-690/15436768	C20H23NO2S	4.57	-5.994	N-(2-adamantyl)-2-naphthalenesulfonamide
AG-690/15437728	C21H22N2O3S	4.21	-4.807	N-(2,4-dimethylphenyl)-2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)-4-(methylsulfonyl)butanamide
AG-690/15440295	C18H21ClN2O3S	3.88	-5.478	N-(4-chlorophenyl)-2-[4-isopropyl(methylsulfonyl)anilino]acetamide
AG-690/15442007	C18H20N4O2	3.76	-4.136	N-[2-(4-methoxyphenyl)-6-methyl-2H-1,2,3-benzotriazol-5-yl]-2-methylpropanamide
AG-690/40135635	C16H12N2O3S	3.76	-4.646	2-[(6-methyl-1,3-benzothiazol-2-yl)amino]carbonyl benzoic acid
AE-641/40790040	C20H17N3O3S	4.1	-5.408	N-(1,3-benzothiazol-2-yl)-2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)-3-methylbutanamide
AK-105/40832021	C17H14Cl2N2	5.41	-5.997	N-(2,4-dichlorobenzylidene)-N-[2-(1H-indol-3-yl)ethyl]amine
AG-670/40725019	C21H25N3O2S	6.16	-5.993	4-tert-butyl-N-(4-[[propionylamino]carbothioyl]amino)phenyl)benzamide
AH-487/40936238	C18H14BrFN2O4	3.17	-5.763	5-{2-bromo-4-[(2-fluorobenzyl)oxy]-5-methoxybenzylidene}-2,4-imidazolidinedione
AH-487/40935573	C23H18N2O3S2	5.61	-6.321	4-[[5-[(3-methyl-2-thienyl)methyl]ethoxy]-4-oxo-2-(phenylimino)-1,3-thiazolidin-3-yl]methyl]benzamide
AK-918/40909289	C23H21NO4	4.2	-4.508	4-methylphenyl 4-(3,5-dioxo-4-azatricyclo[5.2.1.0~2,6~]dec-4-yl)benzoate
AG-205/41004595	C18H15NO5S	1.57	-1.626	6-(ethylsulfonyl)-3-(2-oxo-2-phenylethylidene)-3,4-dihydro-2H-1,4-benzoxazin-2-one
AG-205/41006043	C12H11Cl2N5O2S	2.01	-3.261	2-[(4-amino-6-methyl-5-oxo-4,5-dihydro-1,2,4-triazin-3-yl)sulfanyl]N-(3,4-dichlorophenyl)acetamide
AG-205/41005081	C20H15BrF3N5O3S	3.82	-6.517	methyl 4-[[[(4-bromobenzylidene)amino]-5-(trifluoromethyl)-4H-1,2,4-triazol-3-yl]sulfonyl]benzoate
AG-670/40987986	C18H19NO4	3.85	-5.166	methyl 2-[(4-isopropoxybenzoyl)amino]benzoate
AG-670/41019316	C22H21NO3	4.72	-6.054	4-(benzyloxy)-N-(4-ethoxyphenyl)benzamide
AG-670/41011160	C18H16ClN3O3S	1.94	-4.954	2-[(2-chlorophenyl)imino]-4-hydroxy-2,5-dihydro-1,3-thiazol-5-yl)-N-(4-methoxyphenyl)acetamide
AG-205/10365003	C18H19N3O9S	2.41	-4.39	2-[[2-nitro-4-[(3-nitrophenyl)sulfonyl]anilino]ethoxy]ethyl acetate
AE-848/11421830	C17H19NO3S	4.86	-5.3	ethyl 5-ethyl-2-[(3-methylbenzoyl)amino]-3-thiophenecarboxylate
AG-690/11763413	C18H19N2O4	3.6	-2.837	2-[[4-(acetylamino)anilino]carbonyl]cyclohexanecarboxylic acid
AK-968/11840063	C18H16ClN3O2S	2.59	-4.18	2-(5-chloro-2-thienyl)-N-(4-morpholinyl)-4-quinolinecarboxamide
AK-968/11842268	C20H16N2O2	2.6	-4.272	N-(4-pyridinylmethyl)-9H-xanthene-9-carboxamide
AK-968/11842323	C19H17ClN2O2S	3.94	-5.42	2-(5-chloro-2-thienyl)-N-(tetrahydro-2-furanyl)ethyl)-4-quinolinecarboxamide
AK-968/11986087	C22H19N3O3S	5.95	-6.577	N-(4,6-dimethyl-2-pyridinyl)-2-(5-methyl-2-thienyl)-4-quinolinecarboxamide
AJ-030/12105040	C19H16Cl2N2O2	3.95	-6.044	4-[[1-(4-chlorobenzyl)amino]propylidene]-2-(4-chlorophenyl)-1,3-oxazol-5(4H)-one
AK-918/13399079	C18H11Cl2NO3	4.29	-5.336	2-(4-chlorophenyl)-2-oxoethyl 2-chloro-4-quinolinecarboxylate
AJ-916/14717070	C22H15ClN2O5	3.99	-5.024	ethyl 1-(4-chlorophenyl)-4-[(2-oxo-2H-chromen-3-yl)carbonyl]-1H-pyrazole-3-carboxylate
AA-516/25012196	C18H13ClN2O3S2	5.42	-6.592	2-[[[(3-chloro-1-benzothien-2-yl)carbonyl]amino]carbothioyl]amino]-5-methylbenzoic acid
AJ-916/37249013	C19H13N3O2S	4	-4.15	1-[4-(1,3-benzothiazol-2-ylcarbonyl)-1-phenyl-1H-pyrazol-3-yl]ethanone
AG-690/15436747	C14H16N2O3S	1.49	-3.066	N'-isobutyl-2-naphthalenesulfonohydrazide
AK-105/40836215	C18H25N3O3	3.69	-4.777	3-[[1-pyrrolidinylcarbonyl]amino]phenyl cyclohexylcarbamate
AF-399/40857327	C23H31N3O5	3.79	-5.805	N-[(4-benzyl-1-piperazinyl)carbothioyl]-1-adamantanecarboxamide
AH-487/40935978	C19H16ClN3O3S	5.32	-5.759	4-[[anilino]carbonyl]amino]-N-(4-chlorophenyl)benzenesulfonamide
AF-407/40919041	C14H14ClN5O4S	1.22	-2.664	3-(4-amino-3-[[2-(4-chloroanilino)-2-oxoethyl]sulfonyl]-5-oxo-4,5-dihydro-1,2,4-triazin-6-yl])propanamide
AG-670/41011459	C16H12BrN3O3S2	4.72	-6.64	methyl 5-[[[(3-bromobenzoyl)carbothioyl]amino]-4-cyano-3-methyl-2-thiophenecarboxylate
AK-968/11968312	C21H15BrN2O2	5.57	-6.601	N-[[2-(4-bromophenyl)-1,3-benzoxazol-5-yl]-2-phenylacetamide
AJ-916/13004022	C17H13BrN2O3S	4.57	-5.579	ethyl 1-(4-bromophenyl)-4-(2-thienylcarbonyl)-1H-pyrazole-3-carboxylate
AH-484/31178007	C24H18N4O2S	5.7	-7.228	3-oxo-3-phenyl-2-(phenylhydrazono)-N-(4-phenyl-1,3-thiazol-2-yl)propanamide
AG-205/41005649	C18H14N4O4	4.17	-5.774	5-{4-nitrophenyl}-2-furaldehyde N-phenylsemicarbazone
AK-105/40690081	C15H11N3O3	1.9	-2.772	2-[(1H-benzimidazol-1-ylamino)carbonyl]benzoic acid
AK-105/40837666	C18H25N3O3	3.76	-3.888	tert-butyl 2-tert-butyl-2-(1H-indol-2-ylcarbonyl)hydrazinecarboxylate
AG-670/40725010	C17H16ClN3O2S	5.08	-4.75	3-chloro-N-(3-[[propionylamino]carbothioyl]amino)phenyl)benzamide
AK-918/40711603	C18H18N2O4S	1.96	-2.785	2-[[2-[(4-methylphenyl)sulfonyl]ethyl]-1H-benzimidazol-1-yl]acetic acid
AG-670/40765640	C20H18ClN3O4S	3.74	-3.484	2-[[2-[(4-chlorophenyl)imino]-4-oxo-1,3-thiazolidin-5-yl]acetyl]amino]-4-methylphenyl acetate
AH-487/40785762	C20H18ClN3O3S	5.26	-6.331	4-[[anilino]carbonyl]amino]-N-(3-chloro-2-methylphenyl)benzenesulfonamide
AG-670/40910925	C18H19ClN2O2	5.28	-5.011	N-[3-(butylamino)phenyl]-2-chloro-4-methylbenzamide
AE-848/40939019	C17H14N4O3S2	2.58	-2.999	4-[[3-oxo-3-(2-thienyl)-1-propenyl]amino]-N-(2-pyrimidinyl)benzenesulfonamide
AG-205/41004875	C21H13FO4	4.38	-5.19	2-(4-fluoro-3-phenoxyphenyl)-3-hydroxy-4H-chromen-4-one
AG-205/41005840	C20H27N3O4	4.37	-6.386	N-1-N,N-4-dicyclohexyl-2-nitrotetraphthalamide
AG-205/41005473	C11H11Cl2N5OS	2.82	-2.963	2-[[4-amino-5-methyl-4H-1,2,4-triazol-3-yl]sulfonyl]-N-(3,4-dichlorophenyl)acetamide
AG-205/41005631	C20H27N3O2	1.77	-3.902	3-(1-piperidinyl)-1-[4-(1-piperidinyl)phenyl]-2,5-pyrrolidinedione
AG-205/41005649	C21H26N2O4S	3.02	-5.354	4-[[4-methyl-1-piperidinyl]sulfonyl]-N-(2-phenoxyethyl)benzamide
AG-205/41005732	C13H11NO6S	1.47	-2.925	2-hydroxy-5-[(4-hydroxyanilino)sulfonyl]benzoic acid
AH-487/41033649	C17H12N4O5	4.75	-5.635	N-(1,2,3-benzothiadiazol-5-yl)-N'-(2-naphthyl)urea
AH-487/41034388	C21H23N5O	2.91	-4.671	N-(2,3-dimethyl-6-quinoxalanyl)-4-phenyl-1-piperazinecarboxamide
AH-487/41034721	C18H18N4O2	4.17	-4.343	N-(2,3-dimethyl-6-quinoxalanyl)-N'-(3-methoxyphenyl)urea
AG-670/41065165	C18H13ClN4O2S	1.67	-4.722	2-[[2-[(2-chlorophenyl)imino]-4-hydroxy-2,5-dihydro-1,3-thiazol-5-yl]-N-(2-cyanophenyl)acetamide
AE-641/10061047	C24H28N2O3	5	-6.615	ethyl 2-[[4-tert-butylbenzoyl]amino]-3-(1H-indol-3-yl)propanoate
AK-968/11986066	C17H15NO3S	3.43	-4.807	4-methoxy-N-(2-naphthyl)benzenesulfonamide
AJ-916/13841080	C22H17N3O3	4.2	-5.324	4-(2-furoyl)-1-(4-methylphenyl)-N-phenyl-1H-pyrazole-3-carboxamide
AJ-916/12583126	C17H9Cl2NO3	4.72	-4.978	[3-(4-chlorobenzoyl)-4-isoxazolyl](4-chlorophenyl)methanone
AG-690/33057026	C23H17NO2S2	5.84	-6.685	3-benzyl-5-(3-phenoxybenzylidene)-2-thioxo-1,3-thiazolidin-4-one
AH-487/40936647	C17H17Cl3N2O3S2			N-{2-[(4-chlorophenyl)sulfonyl]ethyl}-2-[2,3-dichloro(methylsulfonyl)anilino]acetamide
AJ-916/13841085	C21H14ClN3O3			1-(4-chlorophenyl)-4-(2-furoyl)-N-phenyl-1H-pyrazole-3-carboxamide
AH-487/40936196	C20H18N2S2			N-phenyl-N'-[4-[(phenylsulfonyl)methyl]phenyl]thiourea
AJ-916/13841092	C22H17N3O2S			1-(4-methylphenyl)-N-phenyl-4-(2-thienylcarbonyl)-1H-pyrazole-3-carboxamide
AG-205/41006072	C17H14N4O3			4-[[3,4-dihydroxybenzylidene]amino]-3-methyl-6-phenyl-1,2,4-triazin-5(4H)-one
AG-205/41005325	C15H19N5O3S			N-{4-[(2-methyl-1-piperidinyl)sulfonyl]phenyl}-1H-1,2,4-triazole-3-carboxamide