

文章编号: 1004-0374(2011)09-0869-06

催化元件和途径的人工设计与组装

韩云宾, 黄琛, 冯雁*

(上海交通大学生命科学技术学院微生物代谢国家重点实验室, 上海 200240)

摘要: 催化元件以及由多个催化元件组成的合成途径的设计与组装为人工合成体系的建立奠定了基础, 是合成生物学的重要研究内容。除从自然生物中挖掘大量的天然酶和途径可供人工合成体系使用外, 将计算生物学、蛋白质工程以及组合生物合成等技术相结合, 理性地、有目的地进行催化元件和途径的人工设计与组装, 将提供新功能酶以及新物质合成途径。介绍了催化元件和合成途径人工设计与组装的研究策略和最新进展。

关键词: 催化元件; 合成途径; 合成生物学; 从头设计

中图分类号: O621.33 **文献标志码:** A

Design of the enzymes and assembly of the novel synthetic pathways

HAN Yun-Bin, HUANG Chen, FENG Yan*

(State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Design of enzymes, as well as assembly of synthetic pathways, which are composed of multiple enzymes, are the basis for constructing artificial synthetic system, therefore they are important issues in synthetic biology. The researchers are demanded not only to explore the reservoirs of the natural enzymes but also design the new functional biocatalysts and assemble the novel synthetic pathways by combining the multidisciplinary theories and techniques, such as computational biology, protein engineering, combinatorial biosynthesis and so on. The research strategies and recent progress in artificial design and assembly of enzymes and synthetic pathways are reviewed here.

Key words: enzymes; synthetic pathways; synthetic biology; *de novo* design

自从 2000 年在 *Nature* 杂志上报道合成转录调控震荡网络及构建基因拨动开关的研究以来^[1-2], 合成生物学研究在全球范围内得到关注。在合成新生命领域, 美国著名科学家 Venter 所领导的研究团队创造了一个由化学合成的基因组控制的人造支原体细胞^[3], 为人造生命提供了可能。在对现有生物体系设计和改造的应用领域中, 已建立了多种制备重要药物、精细化工产品以及生物燃料等的人工合成体系, 其中加州大学伯克利分校 Keasling 教授研究组所研究的抗疟疾药物青蒿素前体合成系统即将实现商业化^[4]。

合成生物学研究的基础是对生物元件进行有序组装, 以实现复杂的、新的生物功能。其中催化元件, 即功能酶分子, 是合成生物学研究中重要的基本生

物元件之一; 由多个催化元件构成的催化途径, 可以为细胞提供能量、传递及处理各种信号、合成或降解其他生物分子等。虽然在人工合成途径组装中可以跨物种的应用天然酶基因资源来重构特定的代谢途径, 但是, 由于现有的酶的功能还不能直接满足合成生物学的一些特殊要求, 如在异源宿主中表达量低、生理条件不适应、上下游酶活性不匹配等, 都使得人工设计的合成途径无法实现预想功能。因

收稿日期: 2011-07-18; 修回日期: 2011-07-20

基金项目: 国家重点基础研究发展计划(“973”项目)(2011CBA00800); 国家自然科学基金项目(30970632)

*通信作者: E-mail: yfeng2009@sjtu.edu.cn

此通过对催化元件的人工设计和组装来提供具有新功能的元件成为人工合成途径构建的基础。由于许多有重要应用价值的化合物是天然催化途径所不能产生的或者产生途径目前还不清楚,例如美国能源部列举的来自于生物质的12种重要化学品中有一半的天然合成途径是已知的^[5]。因此,对新功能元件的挖掘及人工设计和组装合成途径将有望快速获得重要产品。本文介绍了催化元件和途径的人工设计与组装的研究策略和最新进展。

1 催化元件的人工设计与组装

催化元件的人工设计和组装主要包括以下三点:(1)人工预测自然界中未发现的天然酶和已经发现的天然酶的新活性;(2)从头设计具有新功能的酶;(3)对天然酶和从头设计的人工酶进行功能优化。

1.1 新酶功能的探测

自然界是获得催化元件的一个重要来源,目前已经表征的酶仅占自然界中酶的少部分,还有大量的天然酶没有被发现或者表征。随着测序成本的降低,蛋白质序列信息急剧增加,通过生物信息学分析方法对这些未表征功能的蛋白质进行功能注释可以为人工途径设计提供新的催化元件。最常用的方法是通过序列相似性比对推断可能的功能^[6],但由于蛋白质序列存在同源异途进化现象,在选择性的压力下获得不同的功能。因此,仅依靠基于序列比对并不能准确地预测蛋白质功能,结合蛋白质的特异位点信息以及结构数据信息可以增加功能预测的准确性^[7]。由于目前完成全基因组测序的生物只是少部分,当我们想从一些没有测序的生物中获得它可能存在的某种酶时,可以从该物种的已经完成全基因组测序的近源物种出发,设计探针引物,把可能存在的酶钓取出来。Keasling 研究组采用这种方法获得了在酵母细胞中合成青蒿酸的关键酶——来源于黄花蒿的细胞色素 P450 单加氧酶(CYP71AV1)。他们通过搜索紫菀目的两种植物向日葵和莴苣的表达序列标签数据库,找到了一些细胞色素 P450 单加氧酶的表达标签,然后根据这些标签设计引物成功地从黄花蒿中得到了 CYP71AV1^[8]。

虽然很多天然酶都已被表征,但是我们获得的只是通过生化实验验证的酶对少数几种底物催化能力的相关知识。目前通过虚拟筛选小分子数据库,可以鉴定一个已知结构酶家族的可能底物,具体过程是通过分子对接技术将这些小分子对接到活性位

点,并基于它们作用的能量情况给予评分。例如在探索来源于喜温海洋杆菌(*Thermotoga maritima*)酰胺水解酶超家族中未表征的酶 Tm0936 的底物特异性时,加州大学圣弗朗西斯科分校 Hermann 等^[9]首先从 KEGG 数据库中寻找该超家族催化反应的可能底物,然后将这些化合物转换成它们各自的高能反应中间体形式,最后进行分子对接。在筛选了 4 000 多个潜在的底物后,分子对接结果表明 Tm0936 可能催化硫代腺苷-L-高半胱氨酸脱氨基生成 S-次黄嘌呤苷-L-高半胱氨酸,并经生化实验得到验证。这项研究表明,采用基于结构的分子对接技术可以用于预测未表征酶的功能。

1.2 从头设计产生新酶

从头设计新酶是指针对一些天然酶无法催化的重要化学反应,通过人工设计催化位点和蛋白质骨架的方法获得能够催化这些反应的新酶。这种方法可以为人工合成生物体系提供新的催化元件。华盛顿大学的 David Baker 研究组在从头设计新酶这一领域取得了许多重要进展。

Baker 研究组采用下列步骤设计新功能酶:第一步是计算化学反应中涉及的所有中间体和过渡态的结构;第二步是采用量子力学计算和一般化学原理相结合的方法设计由蛋白质功能基团组成(氢键供体和受体、阳离子和阴离子、芳香环等)的活性位点,并且以有利于催化的方式将这些功能基团定位于反应中间体和过渡态的周围;第三步是设计一个能在折叠后形成的口袋中包含这些活性位点的蛋白质序列。有两种方法可以实现第三个步骤:第一种方法是从头蛋白质设计,他们设计出由不同数目的二级结构元件构成的、非常稳定的铁氧还蛋白和 Rossmann 折叠,并通过实验进行了验证;第二种方法是通过搜索具有结合口袋的已知结构的蛋白质,寻找能够与活性位点几何兼容的蛋白质主链位置。为了实现这个目的,他们开发了一个叫做 Rosetta-Match 的程序,该程序使用几何散列方法识别能够与活性位点兼容的主链位置^[10]。

Baker 研究组已经成功地应用上述方法获得许多新功能酶,并用于催化天然酶无法催化的化学反应。例如应用传统的酸-碱催化机制设计得到能够催化 Kemp 消除反应的酶, Kemp 反应是苯异恶唑开环生成邻羟基苯甲腈的反应,反应机制是双分子消去反应^[11]。该新功能酶通过天冬氨酸、谷氨酸的侧链羧基作为碱基,或者组氨酸-天冬氨酸、组氨酸-谷氨酸二联体中的咪唑基作为碱基催化底物碳原子

的去质子化, 在设计的 59 个蛋白质中, 有 8 个能够催化 Kemp 消除反应, 其中两个活力最高的蛋白质, 一个活性位点是谷氨酸, 另一个是组氨酸-天冬氨酸二联体^[12]。另一个成功的例子是获得了能够催化 4-羟基-4-(6-甲氧基-2-萘基)-2-丁酮发生 C-C 键断裂生成 6-甲氧基-2-萘醛和丙酮的逆-醛缩酶 (Retroaldolase)。该酶利用赖氨酸作为活性位点与底物上的酮基形成一个希夫碱作为电子陷阱用于促进键的断裂, 在设计的 72 个蛋白质中, 有 32 个具有活力^[13]。在成功地应用 Rosetta 计算设计方法得到的两个催化键裂解的、自然界中不存在的新酶后, 随后他们设计了一个能够催化键形成的新酶, 该酶能够催化立体选择性的 Diels-Alder 反应。Diels-Alder 反应是有机合成的基础, 由共轭双烯与含有烯键或炔键的化合物相互作用反应生成六元环, 形成两个碳碳键和四个新的手性中心, 天然酶无法催化这个双分子反应。他们通过前线轨道理论设计了该酶的活性位点, 用谷氨酰胺或者天冬酰胺的羰基氧作为氢键受体来提高最高占据轨道的能量和稳定过渡态的正电荷积累, 用来自于丝氨酸、苏氨酸或者酪氨酸的羟基作为氢键供体来降低未占用的最低分子轨道的能量和稳定过渡态中的负电荷积累。在设计的 84 个蛋白质中, 有 50 个蛋白质是可溶表达的, 其中有 2 个蛋白质能够催化 Diels-Alder 反应^[14]。

从头设计产生新酶的成功是酶研究领域的重要进展, 不仅可以加深我们对酶催化机制的认识^[15-18], 而且能为代谢途径的人工合成提供新功能的催化元件。虽然目前获得的新酶活力较低, 但可以通过定向进化等方法有效提高酶活力, 例如通过定向进化使催化 Kemp 消除的酶活力提高了 200 倍以上^[12]。

1.3 催化元件的功能优化

在构建体内或者体外人工代谢途径时, 现有的催化元件往往不能满足特殊要求, 例如一些酶不能在异源宿主中可溶性表达, 或者在体内具有副反应, 热稳定性差, 不耐酸碱, 活力低等。目前可以利用多种方法对催化元件进行功能优化, 例如密码子优化、酶的模块组装、定向进化等。

合成途径的构建经常需要组装一系列来自于不同生物来源的催化元件。因为密码子偏好性, 很多酶在异源宿主中不能有功能的表达, 特别是对于在亲缘关系比较远的宿主中表达时更是如此。通过一些密码子分析工具, 例如 Gene Designer^[19], 可以很好地解决表达水平上的密码子偏好性问题。植物基因在大肠杆菌中很难进行有活力的表达^[20],

Keasling 课题组成功表达了一个来源于黄花蒿中的经过密码子优化过的紫穗槐-4,11-二烯 (amorpho-4,11-diene) 合成酶, 催化法呢基焦磷酸发生自身环化生成紫穗槐-4,11-二烯, 用于生产抗疟疾药物青蒿素的前体^[21]; 塔夫茨大学 Pfeifer 研究组和麻省理工学院 Stephanopoulos 研究组合作在大肠杆菌中运用密码子优化、改变 N-末端转膜序列, 以及来自于紫杉中的细胞色素 P450 还原酶融合表达等方式, 在大肠杆菌中成功表达出了具有活性的植物细胞色素 P450 单加氧酶, 用于合成治疗癌症的重要药物紫杉醇的前体紫杉二烯-5a-醇 (taxadien-5a-ol)^[22]。

由于天然酶在序列、结构和功能等不同层次中均具有一定程度的模块性, 通过酶的模块组装技术可以对催化元件的功能进行优化, 特别是在扩宽酶的底物特异性方面。拥有 B 型构象 (两个 Rossmann 折叠区域通过柔性连接组成的一种拓扑结构) 的糖基转移酶在结构上具有较强的模块性, 其 N-末端及 C-末端结构域分别负责结合糖受体底物与糖供体底物。汉城国立大学 Park 等^[23] 将具有不同糖受体底物与糖供体底物特异性的糖基转移酶卡那霉素糖基转移酶 (kanamycin GT) 和万古霉素糖基转移酶 (vancomycin GT) 进行了重组, 获得了底物选择性明显拓宽的嵌合体。

当我们不清楚一些催化元件的结构及作用机理时, 可以通过定向进化技术对它们进行功能优化, 包括提高热稳定性^[24] 和耐酸性^[25], 提高底物的化学、区域和对映体选择性^[26], 删除不需要的生物化学活性 (例如副反应)^[27], 这在构建人工代谢途径中得到了多方面的应用。瓦格宁根大学 Machielsen 等^[28] 对来自于嗜热细菌火球菌属 (*Pyrococcus furiosus*) 的醇脱氢酶采用定向进化的方法, 提高了该酶在低温时的催化活力, 其中一个突变体在 30℃ 时催化 2,5-己二酮还原成 2,5-己二醇的活力是野生型的 10 倍; 加州理工大学的 Arnold 研究组通过使用定向进化技术改变了来自于巨大芽孢杆菌 (*Bacillus megaterium*) 的细胞色素 P450 BM3 的区域选择性, 使其能够对正烷烃的末端进行羟基化, 例如可以羟基化乙烷生成乙醇, 具有使用石油化学品作为原料生产运输燃料的潜力^[29]; 加州大学洛杉矶分校 Liao 研究组对来源于闪烁古球菌 (*Archaeoglobus fulgidus*) 的牻牛儿基牻牛儿基二磷酸合酶编码基因进行定向进化, 从超过 2 000 个突变体的突变库中筛选获得了 8 个活力增加的突变体, 提高了在大肠

杆菌中产番茄红素这一人工合成途径的效率, 其中一个突变体使番茄红素产量增加一倍^[30]。

2 催化途径的人工设计与组装

2.1 计算机设计合成途径

催化途径人工设计的第一步是获取相关催化元件的信息, 一些综合性蛋白质和代谢数据库可以提供许多催化元件的相关信息, 例如 BRENDA^[31]、KEGG^[32]、Metacyc^[33] 和 Swiss-Prot^[34]。近期, 还专门开发了一些用于代谢途径设计的数据库, 例如 BNICE(Biochemical Network Integrated Computational Explorer), 它可以用于发现两个化合物之间可能存在的代谢路径^[35], 在芳香族氨基酸生物合成的研究中使用 BNICE 发现在分支酸与苯丙氨酸、酪氨酸和色氨酸之间有超过 40 万种理论上的生物化学途径^[36]; 麻省理工学院 Prather 研究组开发了一个包括与酶催化相关的 600 多种保守结构的数据库, 叫做 ReBit(Retro-Biosynthesis 工具, <http://www.retro-biosynthesis.com>), 输入一个分子或者功能基团结构, 输出所有能够与这个结构发生反应或者产生这个结构的相关酶类^[37]; 明尼苏达大学开发的生物催化/生物降解数据库 (UM-BBD) 利用一系列广义的反应规则逐步设计代谢途径, 特别适用于设计外源物质的降解途径^[38]。

运用前面所述的数据库工具对于一个给定的化合物可以获得大量可能的代谢途径, 但关键在于如何选出最可行的代谢途径, 即采用什么方法来评价一种途径是可行的。一种方法是参考自然界中已经存在的一些代谢途径来过滤除去不太可能实现的途径, UM-BBD 采用的就是这种方法来选择一些最有可能实现的代谢途径。另一种策略是计算代谢途径的热力学情况, 通过计算每个步骤中的 Gibbs 能量变化给予评分^[39]。此外, 还有其他一些方法用于评价代谢途径的可行性, 例如 DESHARKY 方法^[40]通过对代谢负担进行定量, 以此为度量来评价一个催化途径的可行性。随着这些数据库的不断完善, 可以为催化途径的人工设计和组装提供强有力的支持。

2.2 体外催化途径的组装

目前体内代谢途径组装取得了一些重大进展, 除了 Keasling 研究组在酵母中构建了一条非天然的代谢途径合成青蒿酸这一成功例子外, 最近加州大学伯克利分校 Chang 研究组在大肠杆菌中构建了一条高效产丁醇的代谢途径, 也有望应用于规模生

产^[41]。然而, 在实际操作过程中并不是所有的人工设计途径都能在生物体内表现出预期的效果, 原因大多有以下几点: 基因回路的表现难以预测、生物体系的复杂性难以把握、基因元件的兼容性不佳或者目标产物与宿主的相互抑制等。针对在细菌或酵母体内进行代谢途径组装的一些缺点, 范德比尔特大学 Forster 和哈佛医学院 Church 共同提出了在活细胞体外进行代谢途径组装的设想^[42]。相比于调控一个活细胞的代谢, 体外的无细胞合成体系显得更具灵活性及可操作性。无细胞合成体系是指通过组合一系列酶及辅酶来构建复杂的生化代谢网络获得目标产物^[43]。体外代谢途径的组装不仅能在一定程度上克服体内途径组装的弱点, 而且还拥有产量高、纯度高以及反应速度快等优点, 目前已经成为合成高价值化合物和生物能源的一种新选择^[44]。

在体内进行合成途径构建时, 数以千计的蛋白质都被用于微生物本身的复制和代谢, 只有非常少量的蛋白质起到了合成目标产物的作用, 但是如果利用体外体系, 所有引入体系的酶或辅酶都能够相对直接地用于目标产物的生成, 从而在效率上更优^[45]。设计一条体外的多酶代谢途径一般需要经过五个步骤^[46]: (1) 代谢途径的拆解及组合; (2) 酶的选择; (3) 酶的优化; (4) 目标产物试生产; (5) 体系的优化。

最近体外催化途径组装在生物产氢这一研究领域取得了较大的突破。美国橡树岭国家实验室 Woodward 等^[47]首次利用葡萄糖-6-磷酸作为反应起点, 在体外成功组装了一条生物产氢的催化途径, 最终每摩尔葡萄糖-6-磷酸经过一系列反应得到了 11.6 摩尔氢气。然而, 葡萄糖-6-磷酸作为反应起始点存在以下几个弱势: 葡萄糖-6-磷酸比较昂贵; 游离的磷酸作为最终产物会结合 Mg^{2+} , 从而影响部分酶的活力^[48]; 磷酸积累造成的 pH 值变化过大, 不利于体系的平衡^[49]。弗吉尼亚理工大学 Zhang 等^[50]为了解决这些问题, 将葡聚糖作为反应起始点, 在原有途径基础上引入了从葡聚糖到葡萄糖-6-磷酸的两步反应, 由于葡聚糖价格更低 (约 0.15 \$/kg)、更易获得, 因此新途径能有效降低生产成本 (2 \$/kg), 可以预见, 通过使用更为廉价的淀粉等固体多糖可以进一步降低成本, 为生物产氢工业化奠定基础。

利用聚酮合酶的模块性进行体外组装成产聚酮类抗生素成为近年来的研究热点, 一方面通过对来源于不同模块中的催化功能域进行组装可以合成

许多新的、非天然的化合物, 为微生物药物的筛选提供丰富的化合物资源; 另一方面在体外组装避免了基因工程菌在合成抗生素时产生的自我抑制情况^[51]。斯坦福大学 Khosla 等^[52] 针对 I 型聚酮合酶中最简单的 6-脱氧红霉素内酯 B 合成酶进行了模块分解和再组装, 将第三个重复模块中的酮还原酶用其他模块中的相同部分进行替换并测试其活力, 发现重复模块中的每个部分都有其特殊性, 并不能进行简单替代。加州大学圣地亚哥分校 Moore 研究组首次在体外构建了 II 型聚酮类化合物的完整合成途径, 该体系利用苯甲酸及丙二酰 CoA 作为反应底物, 在消耗 ATP 及 NADPH 的状态下得到了天然抗生素——肠球菌素^[53]。加州大学洛杉矶分校 Ma 等^[54] 在体外利用真菌中的 LovB 聚酮合酶构建了降血脂药物洛伐他丁的合成途径, 他们首先尝试将 LovB 聚酮合酶以及 LovC 烯醇还原酶在酿酒酵母中表达纯化后, 以丙二酰辅酶 A 作为底物合成洛伐他丁, 但是由于 LovC 的错误修饰而无法得到目标产物; 随后用 LovC 的同功酶 MlcG 代替它, 成功地合成了洛伐他丁。在聚酮类化合物碳骨架延伸的过程中需要丙二酰辅酶 A 等昂贵的底物, 德州大学奥斯汀分校 Hughes 等^[55] 为降低成本, 在体外利用来源于天蓝色链霉菌 (*Streptomyces coelicolor*) 中的丙二酰辅酶 A 生成酶 MatB 合成了丙二酰辅酶 A, 该方法能够满足体外酶法合成聚酮类化合物时对丙二酰辅酶 A 的需求。

3 结语与展望

合成生物学作为一门新兴交叉学科, 无论在加深我们对于生命本质的认识, 还是在生物工程应用领域, 都已经取得重要进展。由于现有的天然催化元件和途径无法满足合成生物学的一些特殊要求, 催化元件和途径的人工设计和组装成为必要。在获得满足标准化、通用性要求的催化元件后, 进一步应用这些元件来设计和组装完整的催化途径, 有望合成具有重要应用价值的化合物。目前已经开发了许多方法辅助催化元件及途径的设计, 成功地实现了催化途径在体内及体外的组装。虽然催化元件和途径的研究近年来取得了快速进展, 但是仍有许多问题需要解决, 例如, 迄今为止, 采用从头设计得到的新功能酶非常少, 并不能满足研究需要, 对酶功能快速优化和组装的技术以及从头设计的研究需要进一步发展。综上所述, 随着催化元件和途径人工设计与组装研究的不断深入, 必将极大推动合成

生物学这一领域的发展。

[参 考 文 献]

- [1] Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*, 2000, 403(6767): 335-8
- [2] Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 2000, 403: 339-42
- [3] Gibson DG, Glass JI, Venter JC, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 2010, 329(5987): 52-6
- [4] Kwok R. Five hard truths for synthetic biology. *Nature*, 2010, 463(7279): 288-90
- [5] Werpy T, Petersen G. Top value added chemicals from biomass volume 1-Results of screening for potential candidates from sugars and synthesis gas [R]. US Department of Energy, 2004
- [6] Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *J Mol Biol*, 1990, 215: 403-10
- [7] Weeks AM, Chang MCY. Constructing *de novo* biosynthetic pathways for chemical synthesis inside living cells. *Biochemistry*, 2011, 50: 5404-18
- [8] Ro DK, Paradise EM, Ouellet M, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, 440(7086): 940-3
- [9] Hermann JC, Marti-Arbona R, Fedorov AA, et al. Structure-based activity prediction for an enzyme of unknown function. *Nature*, 2007, 448: 775-9
- [10] Zanghellini A, Jiang L, Wollacott AM, et al. New algorithms and an *in silico* benchmark for computational enzyme design. *Protein Sci*, 2006, 15(12): 2785-94
- [11] Casey ML, Kemp D, Paul KG, et al. Physical organic chemistry of benzisoxazoles. I. Mechanism of the base-catalyzed decomposition of benzisoxazoles. *J Org Chem*, 1973, 38 (13): 2294-301
- [12] Röthlisberger D, Khersonsky O, Wollacott AM, et al. Kemp elimination catalysts by computational enzyme design. *Nature*, 2008, 453(7192): 190-5
- [13] Jiang L, Althoff EA, Clemente FR, et al. *De novo* computational design of retro-aldol enzymes. *Science*, 2008, 319(5868): 1387-91
- [14] Siegel JB, Zanghellini A, Lovick HM, et al. Computational design of an enzyme catalyst for a stereo selective bimolecular Diels-Alder reaction. *Science*, 2010, 329(5989): 309-13
- [15] Baltzer L, Nilsson J. Emerging principles of *de novo* catalyst design. *Curr Opin Biotechnol*, 2001, 12(4): 355-60
- [16] Hill RB, Raleigh DP, Lombardi A, et al. *De novo* design of helical bundles as models for understanding protein folding and function. *Acc Chem Res*, 2000, 33(11): 745-54
- [17] Moffet DA, Hecht MH. *De novo* proteins from combinatorial libraries. *Chem Rev*, 2001, 101: 3191-203
- [18] Kaplan J, DeGrado WF. *De novo* design of catalytic proteins. *Proc Natl Acad Sci USA*, 2004, 101(32): 11566-70
- [19] Villalobos A, Ness JE, Gustafsson C, et al. Gene designer:

- a synthetic biology tool for construction artificial DNA segments. *BMC Bioinformatics*, 2006, 7: 285
- [20] Martin VJ, Yoshikuni Y, Keasling JD. The *in vivo* synthesis of plant sesquiterpenes by *Escherichia coli*. *Biotechnol Bioeng*, 2001, 75: 497-503
- [21] Martin VJ, Pitera DJ, Withers ST, et al. Engineering a mevalonate pathway in *Escherichia coli* for the production of terpenoids. *Nat Biotechnol*, 2003, 21: 796-802
- [22] Ajikumar PK, Xiao WH, Tyo KEJ, et al. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science*, 2010, 330(6000): 70-4
- [23] Park SH, Park HY, Sohng JK, et al. Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnol Bioeng*, 2009, 102(4): 988-94
- [24] Shi C, Lu X, Ma C, et al. Enhancing the thermostability of a novel β -agarase AgaB through directed evolution. *Appl Biochem Biotechnol*, 2008, 151: 51-9
- [25] Liu YH, Lu FP, Li Y, et al. Acid stabilization of *Bacillus licheniformis* α amylase through introduction of mutations. *Appl Microbiol Biotechnol*, 2008, 80: 795-803
- [26] Asako H, Shimizu M, Itoh N. Engineering of NADPH-dependent aldo-keto reductase from *Penicillium citrinum* by directed evolution to improve thermostability and enantioselectivity. *Appl Microbiol Biotechnol*, 2008, 80: 805-12
- [27] Kelly RM, Leemhuis H, Rozeboom HJ, et al. Elimination of competing hydrolysis and coupling side reactions of a cyclodextrin glucanotransferase by directed evolution. *Biochem J*, 2008, 413: 517-25
- [28] Machielsen R, Leferink NGH, Hendriks A, et al. Laboratory evolution of *Pyrococcus furiosus* alcohol dehydrogenase to improve the production of (2S,5S)-hexanediol at moderate temperatures. *Extremophiles*, 2008, 12: 587-94
- [29] Meinhold P, Peters MW, Hartwick A, et al. Engineering cytochrome P450 BM3 for terminal alkane hydroxylation. *Adv Synth Catal*, 2006, 348: 763-72
- [30] Wang CW, Oh MK, Liao JC. Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production. *Biotechnol Progr*, 2000, 16(6): 922-6
- [31] Schomburg I, Chang A, Ebeling C, et al. BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res*, 2004, 32(suppl 1): 431-3
- [32] Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res*, 2006, 34(suppl 1): 354-7
- [33] Caspi R, Foerster H, Fulcher, et al. MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res*, 2006, 34: 511-6
- [34] Wu CH, Apweiler R, Bairoch A, et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res*, 2006, 34: 187-91
- [35] Li C, Henry CS, Jankowski MD, et al. Computational discovery of biochemical routes to specialty chemicals. *Chem Eng Sci*, 2004, 59(22-23): 5051-60
- [36] Hatzimanikatis V, Li C, Ionita JA, et al. Exploring the diversity of complex metabolic networks. *Bioinformatics*, 2005, 21(8): 1603-9
- [37] Martin CH, Nielsen DR, Solomon KV, et al. Synthetic metabolism: engineering biology at the protein and pathway scales. *Chem Biol*, 2009, 16(3): 277-86
- [38] Ellis LBM, Roe D, Wackett LP. The University of Minnesota Biocatalysis/Biodegradation Database: the first decade. *Nucleic Acids Res*, 2006, 34: 517-21
- [39] Jankowski MD, Henry CS, Broadbelt LJ, et al. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys J*, 2008, 95:1487-99
- [40] Rodrigo G, Carrera J, Prather KJ, et al. DESHARKY: Automatic design of metabolic pathways for optimal cell growth. *Bioinformatics*, 2008, 24: 2554-6
- [41] Bond-Watts BB, Bellerose RJ, Chang MCY. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat Chem Biol*, 2011, 7: 222-7
- [42] Forster AC, Church GM. Synthetic biology projects *in vitro*. *Genome Res*, 2007, 17(1):1-6
- [43] Zhang YH. Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. *Biotechnol Bioeng*, 2010, 105(4): 663-77
- [44] Hirao I, Ohtsuki T, Fujiwara T, et al. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat Biotechnol*, 2002, 20(2):177-82
- [45] Schultheisz HL, Szymczyna BR, Scott LG, et al. Pathway engineered enzymatic *de novo* purine nucleotide synthesis. *ACS Chem Biol*, 2008, 3(8): 499-511
- [46] Zhang YH, Sun JB, Zhong JJ. Biofuel production by *in vitro* synthetic enzymatic pathway biotransformation. *Curr Opin Biotechnol*, 2010, 21(5): 663-9
- [47] Woodward J, Orr M, Cordray K, et al. Biotechnology: Enzymatic production of biohydrogen. *Nature*, 2000, 405(6790): 1014-5
- [48] Jewett MC, Swartz JR. Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng*, 2004, 86(1): 19-26
- [49] Wang Y, Zhang YH. Cell-free protein synthesis energized by slowly-metabolized maltodextrin. *BMC Biotechnol*, 2009, 9: 58
- [50] Zhang YH, Evans BR, Mielenz JR, et al. High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PLoS One*, 2007, 2(5): e456
- [51] Gao X, Wang P. Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl Microbiol Biotechnol*, 2010, 88(6): 1233-42
- [52] Khosla C, Kapur S, Cane D. Revisiting the modularity of modular polyketide synthases. *Curr Opin Chem Biol*, 2009, 13(2): 135-43
- [53] Cheng Q, Xiang L, Izumikawa M, et al. Enzymatic total synthesis of enterocin polyketides. *Nat Chem Biol*, 2007, 3(9): 557-8
- [54] Ma SM, Li JW, Choi JW, et al. Complete reconstitution of a highly reducing iterative polyketide synthase. *Science*, 2009, 326(5952): 589-92
- [55] Hughes AJ, Clay AK. Enzymatic extender unit generation for *in vitro* polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem Biol*, 2011, 18(2): 165-76