

分类号_____

密级_____

UDC_____

编号_____

博士后研究报告

霉酚酸微生物合成和 P450

单加氧酶与还原伴侣蛋白的相互作用机制研究

齐凤霞

工作完成日期 2013 年 12 月—2016 年 5 月

报告提交日期 2016 年 6 月

中国科学院青岛生物能源与过程研究所

二零一六年六月

霉酚酸微生物合成和
P450 单加氧酶与还原伴侣蛋白的相互作用机制研究

Microbial synthesis of mychophenolic acid and the study on the interactions
between P450 monooxygenases and redox parterners

博 士 后 姓 名 齐凤霞

流动站（一级学科）名称 中国科学院青岛生物能源与过程研究所
生物学流动站

专 业（二级学科）名称 生物化学与分子生物学

研究工作起始时间 2013 年 12 月 19 日

研究工作期满时间 2016 年 6 月 19 日

中国科学院青岛生物能源与过程研究所

2016 年 6 月

摘要

本报告围绕微生物生物合成霉酚酸（MPA）和“微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制”研究课题展开工作，本报告为为下一步的课题进展提供了平台。

内容包括以下三个方面：

1. 在酿酒酵母宿主中表达来自短密青霉的霉酚酸合成相关基因

a) 利用酿酒酵母同源重组系统组装大片段 DNA 的技术路线，证明了该技术在理性设计组装外源 DNA 和代谢工程改造酿酒酵母的可行性、高效性和可重复性；

b) 分子克隆了来自短密青霉的霉酚酸（MPA）合成相关基因 *mpaA*，*mpaDE* 和 *mpaC* 的。针对 *mpaC* 突变株开展了发酵实验，通过 HPLC 未能检测到代谢产物 5-MOA；三个蛋白均利用 SDS-PAGE 和 Western Blot 检测了蛋白表达情况。

c) 对 MPA 合成相关基因 *mpaA* 进行了密码子优化并蛋白纯化鉴定。

2. 国家优青项目“微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制”研究课题

a) 蓝细菌来源的 6 个还原伴侣 Fdx 的分子克隆和蛋白纯化。

b) 克隆表达来自天然色链霉菌的 18 个 P450 酶基因和 6 个铁氧化还原蛋白 Fdx、4 个铁氧化还原蛋白还原酶 FDR，对 Fdx6 密码子优化，并进行了蛋白纯化鉴定. P450 和 FDR 进行了光谱分析。

c) 分别采用黄体酮、睾酮、薯蓣皂苷元、丹参酮 IIA、麦新米星 IV(M-IV)和 YC17 等多种底物对这些 P450 和还原伴侣进行了酶催化反应筛选鉴（HPLC 检测）。首次发现了天蓝色链霉菌的还原伴侣蛋白用于催化支撑 PikC 催化 YC17 合成产物。并且证明在支撑 PikC 的效力层面，不同组合的还原伴侣可以达到不同的催化效率，合成不同的目的产物。为进一步获得对“P450 酶—Fdx—FDR”三蛋白系统相互作用和识别机制的规律性认识做出积累。

3. 其他：对序列未知的米曲霉表达载体 pTAex3 进行了全测序并获得了图谱等。

关键词：微生物合成；蛋白纯化；霉酚酸；P450；还原伴侣

ABSTRACT

The report is focused on the microbial biosynthesis of mycophenolic acid (MPA) and National Excellent Youth Project "The interaction mechanism of the microbial biosynthesized P450 oxygenase and reducing chaperones". This report provides a platform for further research of the subjects.

There are mainly three aspects:

1. Expression of genes involved in MPA biosynthesis from *Penicillium brevicompactum* using *Saccharomyces cerevisiae* as a host

A) A homologous recombination based DNA assembler technology was used for construction of large DNA fragments in *Saccharomyces cerevisiae*. This method is proved feasible, efficient and reproducible for the rational design of metabolic engineering pathways.

B) The *mpaA*, *mpaDE* and *mpaC* gene in for MPA synthesis in *Penicillium brevicompactum* strain NRRL864 were molecularly cloned. Through fermentation experiments, the *mpaC* mutant strain showed no production of the 5-MOA through HPLC/LC-MS detection. To detect protein expression of the three proteins, SDS-PAGE and Western blot were carried out.

C) *mpaA* was codon optimized and the protein was purified and verified.

2. The project "The interaction mechanism of the microbial biosynthesized P450 oxygenase and reducing chaperones".

A) Molecular cloning and protein purification of the 6 reduced chaperone Fdx from the cyanobacterium strain *Synechococcus elongatus* PCC7942.

B) Cloning and expression of the 18 P450 genes and 6 ferredoxins (Fdx), and 4 ferredoxin reductases (FDRs) from *Streptomyces coelicolor* A3(2). Fdx6 was codon optimized. The P450s were verified by CO-assay. The four FDR were analyzed using a spectral analysis.

C) Progesterone, testosterone, tanshinone IIA, diosgenin, mycinamicin IV (M-IV), and YC17 were used as substrates for the reaction by P450 and reducing partners (based on HPLC screening). It was discovered for the first time that the reductive chaperone protein of *Streptomyces sp.* are able to support PikC to catalyze YC17. And it is proved that the different combinations of the reducing partners have various catalytic efficiencies and can achieve different products. This helps to the understanding of the regularity of the interaction and recognition mechanism of the "Fdx - FdR - P450 system.

3. The expression vector in *Aspergillus oryzae*, pTAex3, was sequenced and the map was achieved.

Key Words: Microbial synthesis, protein purification, mycophenolic acid, P450, redox partners

目次

摘要.....	I
ABSTRACT.....	III
目次.....	V
第一章 引言.....	1
1 大肠杆菌遗传背景.....	1
1.1 大肠杆菌遗传系统.....	1
1.2 大肠杆菌常用菌株及质粒.....	1
2 酿酒酵母遗传背景.....	3
2.1 酿酒酵母常用的商业化表达系统.....	4
2.2 酿酒酵母常转化方法.....	4
3 霉酚酸合成研究进展.....	4
3.1 霉酚酸的应用价值.....	4
3.2 霉酚酸合成路线及现状.....	5
3.3 短密青霉霉酚酸合成路径相关基因.....	5
4 P450 与还原伴侣研究背景.....	6
4.1 P450 与还原伴侣及工业应用前景.....	6
4.2 聚球藻 <i>Synechococcus</i> sp. PCC7942 还原伴侣研究背景.....	9
4.3 天蓝色链霉菌 P450 和还原伴侣研究背景.....	10
第二章 实验方法与手段.....	13
1 实验原料、试剂及设备.....	13
1.1 主要实验仪器.....	13
1.2 实验试剂.....	13
1.3 实验菌株及菌株来源.....	14
2 大肠杆菌的培养.....	14
3 大肠杆菌 <i>E.coli</i> 感受态的制备与转化.....	15
4 酿酒酵母 <i>Saccharomyce scerevisiae</i> 的培养与电转化筛选.....	15
4.1 酿酒酵母培养方法.....	15
4.2 酿酒酵母电转化方法.....	15
4.3 酿酒酵母的基因组提取.....	16

4.4 酿酒酵母质粒提取	16
4.5 酿酒酵母菌落 PCR	17
4.6 酿酒酵母蛋白提取	17
5 DNA assembler 法在酿酒酵母系统中快速组装大片段 DNA	17
6 大肠杆菌和酿酒酵母转化子的基因鉴定方法	18
6.1 天蓝色链霉菌提取基因组总 DNA 及基因组 PCR	18
6.2 菌落 PCR 快速鉴定转化子基因型	18
7 聚合酶链式反应 PCR (Polymerase Chain Reaction)	19
8 大肠杆菌蛋白样品制备及聚丙烯酰胺凝胶电泳 (SDS-PAGE)	19
8.1 菌落 PCR 快速鉴定转化子基因型	19
8.2 大肠杆菌蛋白纯化	19
8.3 聚丙烯酰胺凝胶电泳 (SDS-PAGE)	20
9 蛋白质的免疫印迹 (Western blot)	20
10 酿酒酵母蛋白提取	21
11 CO 差光谱法测定 P450 单加氧酶的活性。	21
第三章 酿酒酵母中表达霉酚酸合成相关蛋白	23
1 材料与方法	23
1.1 酿酒酵母菌株的培养	23
1.2 菌株的构建过程	23
1.3 蛋白纯化与提取	25
1.4 聚合酶链式反应 PCR	25
1.5 质粒提取与鉴定	25
1.6 膜蛋白提取与鉴定	25
2 实验结果与分析	25
2.1 突变株的基因型鉴定	25
2.3 蛋白的表达纯化分析汇总	28
3 本章小结	30
第四章 蓝细菌还原伴侣的克隆与表达	33
1.1 菌株的培养	33
1.2 聚丙烯酰胺凝胶电泳 Western blot 检测目的基因的表达	33
1.3 聚合酶链式反应 PCR	33
1.4 蛋白纯化和 SDS-PAGE	33

2 实验结果与分析	33
2.1 基因检索分析	33
2.2 质粒和菌株的构建	39
2.2 突变株的基因型鉴定结果	39
2.3 蛋白纯化与鉴定	40
3 本章小结	41
第五章 天蓝色链霉菌还原伴侣基因克隆与蛋白纯化.....	43
1 材料与方法	44
1.1 大肠杆菌菌株的培养	44
1.2 菌株的构建	44
1.3 蓝细菌基因型的鉴定（菌落 PCR）	45
1.4 蛋白的诱导表达和纯化	45
1.5 CO 法测 P450	45
2 实验结果与分析	45
2.1 质粒和菌株的构建	45
2.2 天然色链霉菌 P450 及还原伴侣蛋白在大肠杆菌中的表达纯化与鉴定.....	47
2.3 P450 和 FDR 的光谱吸收.....	55
3 本章小结	57
第六章 天蓝色链霉菌还原伴侣和 P450 应用筛选	59
1 材料与方法	59
1.1 菌株的培养	59
1.2 菌株的构建	59
1.3 基因型鉴定	59
1.4 DNA 的重组.....	59
1.5 聚合酶链式反应 PCR	59
2 实验结果与分析	60
2.1 构建的质粒和菌株	60
2.2 酶反应底物的选择	60
2.3 HPLC 结果与分析	61
3 本章小结	66
第七章 结论.....	69
致 谢.....	71

参考文献.....	73
附 录.....	77
附录 1 符号表	77
附录 2 主要培养基	78
附录 3 本文涉及的质粒及菌株.....	82
附录 4 本文涉及的基因序列	87
结尾.....	117

第一章 引言

1 大肠杆菌遗传背景

1.1 大肠杆菌遗传系统

大肠埃希氏菌 (*Escherichia coli*) 通常被称为大肠杆菌, 是 Escherich 在 1885 年发现的, 故而得名。大肠杆菌是一种普通的原核生物, 属于革兰氏阴性细菌 (G-)。大肠杆菌表达系统遗传背景清楚, 目的基因表达水平高, 培养周期短, 抗污染能力强等特点, 是分子生物学研究和生物技术产业化发展进程中的重要工具。

野生大肠杆菌 (*E.coli*) 的基因组 DNA 中有 470 万个碱基对 (bp), 内含 4288 个基因。大肠杆菌基因组具有极高的可塑性, *E.coli* 基因组中包含有许多插入序列, 如 λ -噬菌体片段和一些其他特殊组分的片段, 这些插入的片段都是由基因的水平转移和基因重组而形成的。利用大肠杆菌基因组的高度可塑特性对其进行突变、插入或者缺失, 从而获得不同的基因型和表型, 这些不同基因型特性的菌株在基因工程的研究和生产中具有广泛的应用价值[1]。

大肠杆菌的主要基因型包括: 与基因重组相关的基因型 (如 *recA*、*recB* 和 *recC* 等)、与甲基化相关的基因型 (如 *dam*、*dcm*、*mcrA*、*mcrB* 和 *C*、*mrr* 和 *hsdM* 等)、与点突变相关的基因型 (如 *mutS*、*mutT*、*dut*、*ung* 和 *uvrB* 等)、与核酸内切酶相关的基因型 (如 *hsdR*、*hsdS* 和 *endA* 等)、与终止密码子相关的基因型 (如 *supE* 和 *supF*)、与抗药性相关的基因型 (*gyrA*、*rpsL* 和 *Tn5* 等) 及其他与能量代谢、氨基酸代谢、维生素代谢等相关的基因型[1]。

1.2 大肠杆菌常用菌株及质粒

1.2.1 大肠杆菌常用菌株

基因工程中经常使用的大肠杆菌几乎都来自于 K-12 菌株, 也使用由 B 株和 C 株来源的大肠杆菌。从分子生物学用途来分, 大肠杆菌可以分为质粒构建菌株和表达菌株。首先, 重组质粒的构建菌株一般选择遗传稳定、转化效率高和质粒产量高的菌株作为受体菌, 常用的有 *E.coli* DH5 α , *E.coli* JM 109, *E.coli* DH 10B, *E.coli* NovaBlue 等 *recA*⁻和 *endA*⁻型细胞。另一方面, 作为表达宿主菌必须具备遗传稳定、生长速度快、表达蛋白稳定等基本特点。具体操作过程中, 还应根据所使用的表达载体的特点, 以及目的基因密码子的组成等需求选择特定的表达宿主菌。以下是实验室常用的几种表达宿主:

1) BL21 (DE3): DE3 噬菌体溶源于 BL21 形成的带有染色体 T7 RNA 聚合酶基因大肠杆菌。IPTG 诱导的 *lacMV5* 启动子控制 T7 RNA 聚合酶基因表达 T7 RNA 聚合酶, 进而控制 T7 表达系统表达目的蛋白。

2) BL21 (DE3) 衍生系列: 在经典的 T7 表达系统 BL21 (DE3) 的基础上, Novagen 公司开发了一些特殊的表达宿主细胞。比如: *Origami* (DE3), *Origami B* (DE3) 和 *Rosetta-gami* (DE3) 菌株带有 *trxB* 和 *gor* 双突变。拥有 *trxB* 和 *gor* 突变的菌株比单具 *trxB* 突变的菌株更有可能促进二硫键的形成, 使蛋白可溶性更好, 活性更高。

3) Rosetta 系列: 是经过修饰, 专用于带有大肠杆菌稀有密码子的真核蛋白表达的菌株。经提高稀有 tRNA 水平, 可以提高一些真核基因表达效率。

4) BL21-Codon Plus 系列: 包括 BL21-CodonPlus (DE3)-RIPL, BL21-CodonPlus-RIL, BL21-CodonPlus (DE3) -RIL, BL21-CodonPlus -RP, BL21-CodonPlus (DE3) -RP 等。这些受体菌添加了大肠杆菌中编码精氨酸 (R), 亮氨酸 (L), 异亮氨酸 (I) 和脯氨酸 (P) 稀有密码子的 tRNA 基因, 更多用于表达一些真核生物的基因。其中 RIL 系列常用于 AT 含量高的基因, 而 RP 系列主要用于 GC 含量高的基因。

5) M15/SG13009: 自身表达 T5 RNA polymerase, 主要用于 pQE 系列载体的表达。

1.2.2 大肠杆菌常用质粒

按照用途可以分为克隆质粒和表达质粒。大肠杆菌常用的克隆质粒: pBR322, pUC18/19, pBlueScript-M13 等。T 载体是一种高效克隆 PCR 产物的专用质粒载体, 为线性化载体, 无需酶切可直接与具有 A 末端的 PCR 产物连接, 属于非定向克隆。常用商业化具有 T 末端的 T 载体, 如 pMD18-T (Takara 公司) 等。

表达载体是一类用于将特定基因引入特定宿主细胞的载体, 可以用于目的基因在宿主细胞中的蛋白表达和产物积累。表达载体具备调节序列、启动子、多克隆序列 (MCS)、终止序列和筛选标记和蛋白标签等多个元件。按照启动子类型可以分为可诱导表达载体和组成型表达载体。

1) pET 载体系列是大肠杆菌中最常用表达载体。pET 载体含有强启动子 T7 启动子, 通过诱导 T7 RNA 聚合酶的表达来启动 T7 启动子。T7 启动子启动效率极高, 在诱导短短几个小时之后就有超过 50% 的目的蛋白积累。pET 系列载体在不诱导的条件下可以达到几乎无声 (Silent)。通常采用两种方式诱导 pET 载体的表达。

第一, 在不含有 T7 RNA 聚合酶的宿主中转化目的载体, 通过感染 λ CE6 噬菌体从而引入 T7 RNA 聚合酶。此方法通常用于表达致死或高毒性产物时采用。

第二, 或者将目的质粒转化到基因组整合有 T7 RNA 聚合酶基因的宿主菌中。T7 RNA 聚合酶由 lacUV5 启动子启动。目的基因通过加入半乳糖或类似物 IPTG 启动表达, 目前有些商业公司也提供商业化的诱导培养基。pET 载体也可为基因组整合提供 T7 启动子及其调控序列元件。目前有几十种 pET 载体可供选择。

2 酿酒酵母遗传背景

酵母菌是单细胞真核微生物，在自然界分布很广，主要生长在偏酸性的环境中，必须以有机碳化物为碳源和能源物质。酵母菌是人类的第一种“家养微生物”，是目前人类直接食用量最大的一种微生物，其发酵产品大大改善和丰富了人类的生活。公元前 2000 多年，酵母就和人类的日常生活有着紧密的联系，人类就开始利用含酵母的“老酵”制作面包。并利用酵母制作发酵食品和酿酒。1680 年，列文虎克用显微镜从一滴啤酒中发现酵母细胞，不久，人类就开始有意识地利用酵母（啤酒酵母泥）发面，酵母的重要性逐渐引起工业界的注意。1876 年，法国人巴斯德关于空气中的氧能促进酵母繁殖理论的发表，为大规模通风培养生产酵母奠定了基础。20 世纪 20 年代起，酵母生产用原料扩大到使用糖蜜、木材水解液、亚硫酸纸浆废液和糖蜜酒精糟液等[2]。

酿酒酵母(*Saccharomyces cerevisiae*)是模式酵母菌之一，在分类学上属于真核生物界的真菌门、子囊菌亚门、半子囊菌纲、内孢霉目、酵母科的酵母属。又称面包酵母或者出芽酵母，酿酒酵母是与人类关系最广泛的一种酵母，不论在日常生活中，还是在现代分子和细胞生物学中，都具有重要意义。酿酒酵母为单细胞，形态有卵圆形、球形或椭圆形，细胞大小为 $(2.5\sim 10)\times(4.5\sim 21)\mu\text{m}$ ，酿酒酵母细胞的大小与菌龄、环境有关。细胞壁厚 $0.1\sim 0.3\mu\text{m}$ ，为三明治结构。菌落中等大小，扁平，光滑，湿润，折光，乳酪色到淡棕色，在沙保平板上有酒酿气味[3]。

酿酒酵母与同为真核生物的动物和植物细胞具有很多相同的结构。又容易培养。因此常被用作研究真核生物的模式生物。酵母的细胞有两种生活形态，单倍体和二倍体。单倍体的生活史较简单，通过有丝分裂繁殖。在环境压力较大时通常则死亡。二倍体细胞（酵母的优势形态）也通过简单的有丝分裂繁殖，但在外界条件不佳时能够进入减数分裂，生成一系列单倍体的孢子。

酿酒酵母的生长一般有 4 个时期：调整期，对数期，稳定期，衰亡期。将酿酒酵母以 1% 的接种量接种于培养基中。在 28°C ，250 rpm 下培养，定期测定酵母菌的 OD 值，得生长曲线。培养 0~6 h：吸光度值变化不大，酿酒酵母处于生长的调整期。培养 6~24 h：吸光度值迅速增加，酿酒酵母生长处在对数期。24h~：酿酒酵母的生长基本达到稳定状态，为稳定期。酵母菌的遗传物质由细胞核 DNA，线粒体 DNA，以及特殊的质粒 DNA 组成。

1996 年，酿酒酵母成为第一个完成基因组测序的真核生物，在基因组序列信息研究上取得了重大的进展，并且成为后基因组研究的主要模式材料[4]。为高等生物，以及人基因组的研究提供了很好的借鉴，并为深入认识酵母以及生命的进化提供了基本的信息。作为一种模式生物，酵母在许多研究方面都具有不可替代的优势，它不仅可作为一种很好的研究基因功能的宿主，而且极有可能成为一个强有力的反向遗传学研究的试验系统，用于确定与人类疾病相关基因的功能和抗癌、抗病毒药物的筛选。酿酒酵母含有 12156677 的碱基

对, 预测有 6275 个基因, 十六个染色体。其中只有大约 5800 个基因有功能。31% 的及酿酒酵母与人类基因组有同源性[5]。

2.1 酿酒酵母常用的商业化表达系统

酿酒酵母常用表达载体有 pYES2, pYES2NT, pYES2CT, pYES3, pYES6, pYCplac22-GFP, pAUR123, pRS303TEF, pRS304, RS305, RS306, Y13TEF, pY14TEF, pY15TEF, pY16TEF。酵母基因重组表达载体 pUG6, pSH47 等。酵母单杂载体有 pHISi, pLacZi, pHIS2, pGAD424 等。酵母双杂交系统质粒 pGADT7, pGBKT7; 对照质粒 pGBKT7-53, pGBKT7-lam, pGADT7-T, PCL1 等。本实验室存有的酿酒酵母质粒 pDR195, pYES2, pACT2, pSH47, YIP5-KanR, pYES2/NT A 等 (获赠于中科院能源所吕雪峰研究员)。

2.2 酿酒酵母常转化方法

酿酒酵母菌株有 INVSc1, YM4271, AH109, Y187, Y190 等。酿酒酵母常转化分为电转化法和化学转化法, 二者各有所长。其中, 电转化法操作简便, 转化效率高, 尤其适用于多 DNA 片段的高效转化。但电转化用的感受态细胞不可冻存, 需新鲜使用。化学法 (PEG/LiAc 法) 转化酵母转需要制备感受态细胞, 感受态可以在 -80°C 保存。LiAc 可使酵母细胞产生一种短暂的感受性状态, 此时它们能够摄取外源性 DNA。化学法转化依赖于鲑鱼精 DNA 的辅助质粒 DNA 提高转化效率。鲑鱼精担体 DNA 为短的线形单链 DNA, 在转化实验中主要是保护质粒免于被 DNA 酶降解; 另外还可能在酵母细胞摄取外源性环形质粒 DNA 中发挥协助作用。在每次使用前务必进行热变性, 使可能结合的双链 DNA 打开, 保证鲑鱼精担体 DNA 在转化实验体系中以单链形式存在。

3 霉酚酸合成研究进展

霉酚酸 (Mychophenolic Acid, MPA) 最早是在 1893 年由一位意大利药学家发现。合成霉酚酸的菌株被命名为灰绿青霉 (*Penicillium glaucum*), 又称为短密青霉 (*P. brevicompactum*)。霉酚酸具有抗菌活性, 1896 年分离得到了霉酚酸的晶体, 并证明霉酚酸可以用来对抗炭疽菌[6]。1912 年两名美国科学教将其命名为霉酚酸, 随后霉酚酸被证明是一种光谱抗菌药物, 具有抗真菌、抗细菌、抗肿瘤和抗牛皮癣活性[7]。

3.1 霉酚酸的应用价值

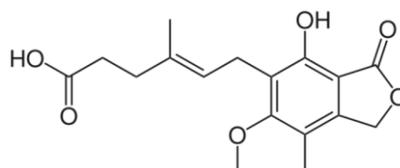


图 1-1 霉酚酸结构

经过近百年的发展，霉酚酸还具有免疫抑制活性，作为主要的免疫抑制剂已被国内外广泛应用于预防、治疗肝脏、肾脏及心脏移植引起的急性排异反应。能在体内通过抑制次黄（嘌呤核）苷酸脱氢酶（IMPDH）的作用阻断嘌呤核苷酸的合成，从而阻止 T 细胞，淋巴细胞的分裂及 B 细胞中抗体的形成[6,8]。此外，霉酚酸多种潜在医学应用价值，在治疗系统性红斑狼疮，硬皮病，寻常性天疱疮(pemphigus vulgaris)特发性血小板减少性紫癜(ITP)等领域具有很高的研究价值[9]。霉酚酸 1995 年被美国 FDA 批准用于肾脏移植。目前霉酚酸产品经济价值较高，Drugs.com 提供的售价大约为 600\$/g。霉酚酸是大型国际公司 CellCept (mycophenolate mofetil; Roche)和 Myfortic (mycophenolate sodium; Novartis)中的活性成分。

3.2 霉酚酸合成路线及现状

霉酚酸最早在 1893 年被发现，目前研究已经发现的可以产霉酚酸的生物有娄地青霉（*Penicillium roqueforti*）、雪白丝衣霉（*Byssosclamyces nivea*）和短密青霉（*Penicillium brevicompactum*）等。在青储饲料中也有发现霉酚酸的合成，这和青储饲料中的相关微生物有关。霉酚酸来自匍匐茎青霉菌（*Penicillium stoloniferum*）或短密青霉（*P. echinulatum*）[10]。霉酚酸酯在肝脏内代谢为活性组分霉酚酸。

3.3 短密青霉霉酚酸合成路径相关基因

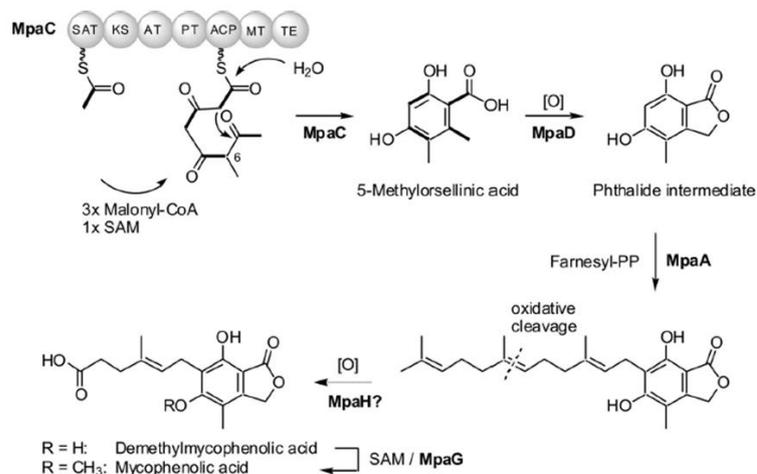


图 1-2 预测的短密青霉中霉酚酸合成路径[11]

霉酚酸历史悠久，但霉酚酸合成的分子基础仍未明朗。直到 2011 年以来，霉酚酸的生物合成分子机制才得到报道。在短密青霉中发现了一个与 MPA 合成相关的基因簇。该基因簇包括几个与 MPA 生物合成相关的基因：*mpaA*、*mpaB*、*mpaC*、*mpaDE*、*mpaG*。其中 *mpaA* 编码一个异戊烯基转移酶，*mpaB* 编码一个功能未知的蛋白，*mpaC* 编码一个聚酮合酶（PKS），*mpaDE* 编码一个含有 P450 域 和一个水解酶，*mpaF* 编码一个 与肌苷酸-5'单

磷酸脱氢酶 (IMPDH) 高度相似的蛋白, *mpaG* 编码一个邻-甲基转移酶, *mpaH* 编码一个氧化裂解酶[7,12,13]。

最新进展, 2016 年发表了一篇文章, 在娄地青霉 (*Penicillium roqueforti*) 中发现了一个与霉酚酸合成相关的基因簇。其组成与短密青霉中累相似[14]。

具体来讲, 在短密青霉中, *MpaC* 是一种聚酮合酶, 已经被证明参与催化 MPA 合成。在短密青霉中含有一个长达 25kb 的基因簇。通过敲除 *mpaC*, MPA 及其衍生物的合成终止。并预测了该基因簇中不同基因对 MPA 合成所发挥的作用。在所述报道中, *mpaC* 有 7,745 bp (包含内含子) [15]。本实验室对一株短密青霉的亚株 (*Penicillium brevicompactum* NRRL864) 进行了全基因组测序, 并预测了 MPA 合成的基因序列, 其中张伟等已经报道了 *mpaG* 的功能[15]。

4 P450与还原伴侣研究背景

4.1 P450 与还原伴侣及工业应用前景

细胞色素 P450 单加氧酶 (P450) 是一类亚铁血红素—硫醇盐(heme-thiolate)蛋白超家族。因其还原态与一氧化碳结合由于 P450 酶可在后在 450 nm 处的特征吸收峰而得名。P450 蛋白超家族是一类广泛存在的具有亚铁血红素活性中心单加氧酶超家族 (heme-thiolate enzymes), 能够催化包括碳-氢键选择性氧化在内的多达 20 余类反应, 可对海量化合物进行高度专一性地羟基化、环氧化、脱烷基等修饰性反应被誉为自然界的“万能催化剂”。P450 酶, 具有自然界最好的催化多能性, 在温和条件下高选择性地催化结构复杂有机化合物中惰性 C-H 键的氧化反应(如羟基化与环氧化), 因此, 细胞色素 P450 单加氧酶已成为近些年药物开发、生物催化及生物降解领域的研究热点[14,16]。

P450 接收来自 FDR 和 Fdx/flavoproteins 的电子。几乎所有细胞色素 P450 酶的催化过程都需要还原伴侣蛋白 (redox partner proteins) 的参与, 来源于 NAD(P)H 的 2 个电子经还原伴侣蛋白顺序传递到亚铁血红素反应中心实现 O₂ 分子的激活, 其中一个氧原子加入底物形成羟基、环氧基等基团(故而 P450 酶经常被称为单加氧酶), 另一个氧原子则参与形成一个水分子。根据还原力来源不同, P450 可以分为不同的类型[17] (图 1-3)。

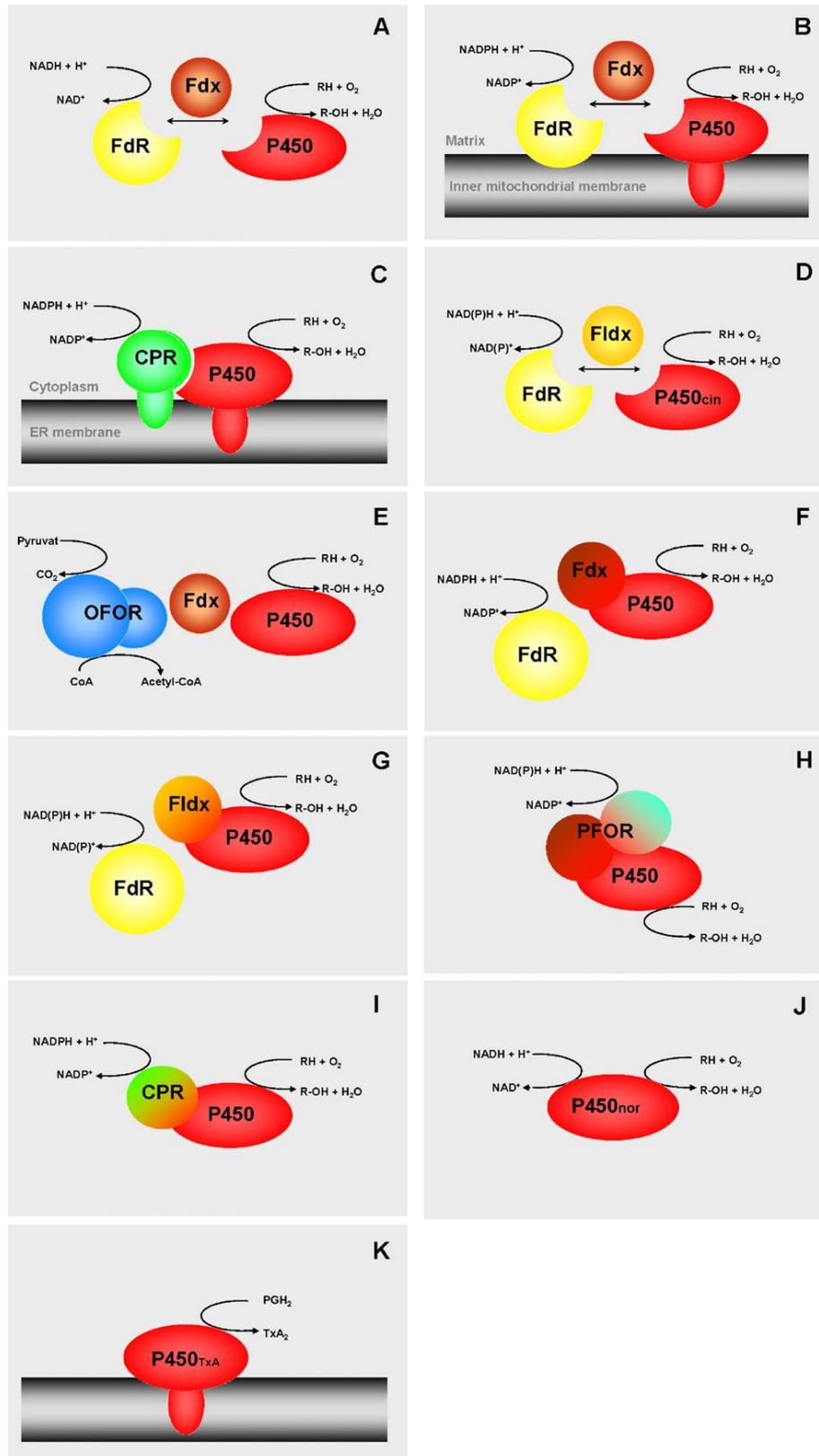


图 1-3 典型 P450 及催化反应[17]

(A) Class I, bacterial system; (B) class I, mitochondrial system; (C) class II microsomal system; (D) class III, bacterial system; example P450cin; (E) class IV, bacterial thermophilic system; (F) class V, bacterial [Fdx]-[P450] fusion system; (G) class VI, bacterial [Fldx]-[P450] fusion system; (H) class VII, bacterial [PFOR]-[P450] fusion system; (I) class VIII, bacterial [CPR]-[P450] fusion system; (J) class IX, soluble eukaryotic P450nor; (K) independent eukaryotic system, example P450TxA.

酶工程团队张伟等对来源于稀有放线菌淡灰红小单孢菌(*Micromonospora griseorubida*) 大环内酯类抗生素麦新米星 (mycinamicins) 生物合成途径中 P450 单加氧酶 MycG4 的研究结果首次对上述“常识”提出了挑战。张伟等发表在 JACS 上的数据发现, 在构建 MycG 与异源还原伴侣蛋白 RhFRED 的融合蛋白过程中, MycG 与 RhFRED 处于融合或分离状态时, MycG 以两种完全不同的方式对麦新米星结构进行修饰: 融合状态下 MycG 催化麦新米星 C14 位的羟基化和 C12/13 位的环氧化; 而在分离状态下, MycG 不仅能催化羟基化和环氧化反应, 还能针对不同底物催化全新的脱甲基反应, 产生 7 种新结构麦新米星衍生物。这一结果发表于 J. Am. Chem. Soc. 杂志[18]。不仅使人们重新认识还原蛋白在决定 P450 酶催化功能中的重要作用, 而且暗示了一种通过改变 P450 与还原伴侣作用方式来产生新结构“天然产物”的全新方法 (图 1-4)。

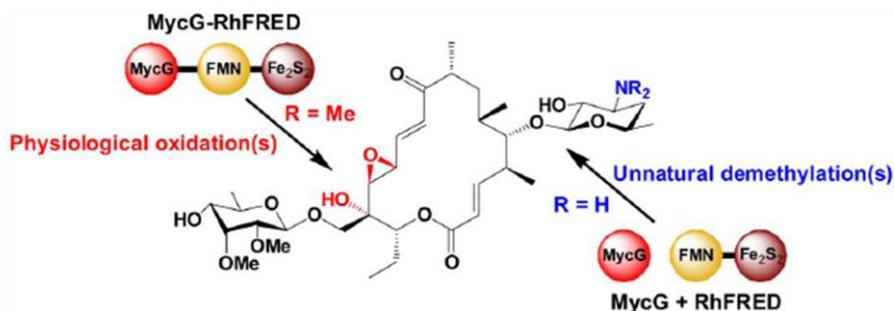


图 1-4 微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制

在微生物药物领域, 细胞色素 P450 酶的催化机制和工程化改造一直是国际研究热点, 但关于 P450 酶与还原伴侣的相互作用研究还比较欠缺, 尤其是一些 P450 酶还苦于不能寻获其最佳还原伴侣而无法进行详细研究和开发, 因此通用性高的还原伴侣选择及电子传递链优化策略的建立迫在眉睫。我们需要不断加强对参与植物和微生物药用活性天然产物生物合成中各种 P450 酶的基因资源挖掘、功能鉴定和酶工程研究, 按照合成生物学思路建立和充实 P450 催化元件库及还原伴侣蛋白库。在此基础上, 针对具有工业需求的目标化学反应开发最佳的细胞色素 P450 酶及其还原伴侣, 以优化的工程菌株作为载体, 开发新的工业 P450 酶和生物转化方案。细胞色素 P450 酶将在不久的将来更加广泛地应用于不同行业领域, 尤其是微生物药物创制领域[19]。

但是, 在少数情况下替代还原伴侣蛋白无法重建目标 P450 酶的催化活性。在多数情况下, 即使替代还原伴侣能够支撑目标 P450 酶的活性, 但支撑程度有限[20], 导致重建 P450 反应的效率低下, 这对于生物催化剂的工业化应用来说属于致命缺陷由于 P450 酶的低稳定性、低活力以及对还原伴侣蛋白和辅因子的依赖等因素限制了这一单加氧酶家族的工业化应用。因此研究和筛选可以高效支撑不同类型 P450 的还原伴侣系统, 筛选得到具有高效催化效率的催化体系具有重要的科研价值和应用价值 (图 1-5)。其中涉及到原核还原伴侣表达系统和真菌表达系统。本报告将分别对这两个平台加以构建及利用。

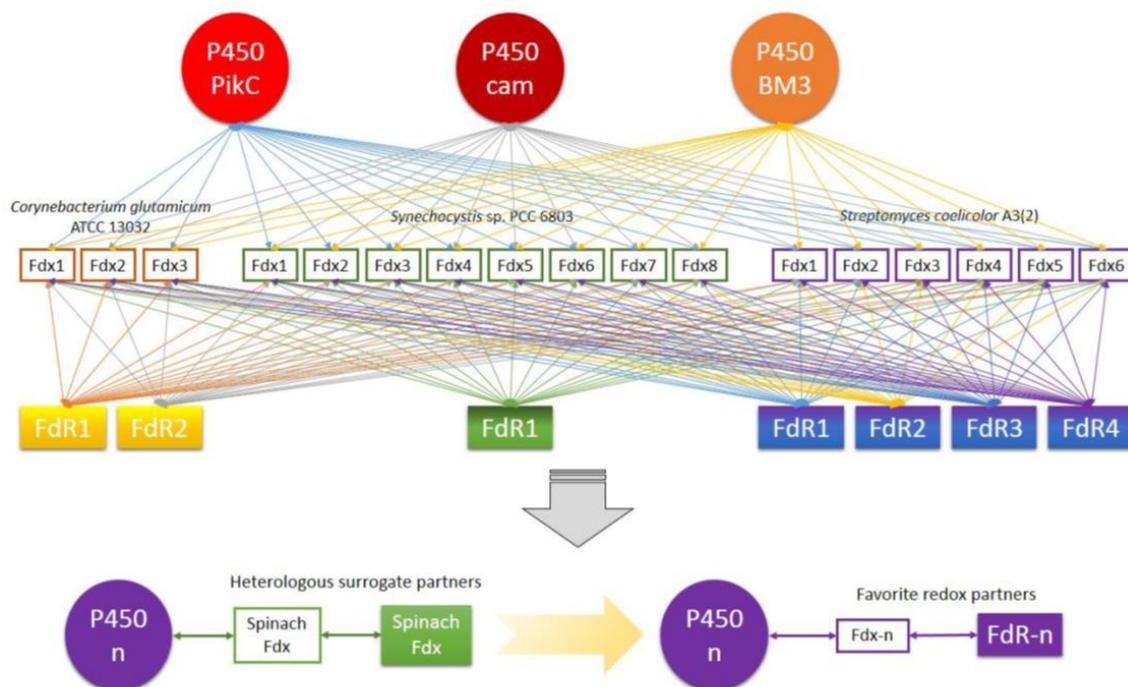


图 1-5 微生物生物合成 P450 单加氧酶与还原伴侣蛋白的研究策略

4.2 聚球藻 *Synechococcus* sp. PCC7942 还原伴侣研究背景

目前商业化的还原伴侣是来自菠菜的还原伴侣蛋白，其价格高，并且被酶制剂公司垄断生产。还原伴侣是广泛存在，蓝细菌是光合原核生物代表，和绿色光合生物具有很近的亲缘关系，同时还具备原核生物的特征，因此其还原伴侣的人工合成与效用具有很大的发展空间。蓝细菌的还原伴侣蛋白研究较为透彻，已有研究表明可以支撑多种反应进行。

本研究选取的蓝细菌聚球藻 (*Synechococcus* sp. PCC7942) 是一种单细胞嗜中温的淡水蓝细菌，其最适生长温度是 38℃，是一类专性光能自养原核微生物，它不能利用葡萄糖进行异养生长，仅能进行专性光合自养生长，是研究昼夜节律的重要模式原核生物。聚球藻 PCC7942 含有一个很小的基因组，大小为 2.7 Mb[21]，其内涵的两个质粒被全测序 (GenBank 序列号分别为 S89470 和 AF441790) [22]。聚球藻 PCC7942 是光合原核生物代表，与高等绿色生物具有极高同源性。

Fdx1 是最主要的铁氧还蛋白，其基因序列最早在 1986 年得到报道[23,24]，并被证明是不可缺少的 Fd[25]。ORF0338 (*petF*) 也有报道[26,27]。铁氧还蛋白-NADP 氧化还原酶 FNR 是一类可催化单电子载体和双电子载体 (携带 NADPH) 之间的还原力交换的黄素酶。其主要作用是在蓝细菌和叶片叶绿体中提供 NADPH 以供光合自养代谢。在根部质粒体中，FNR 可以催化相反的反应，即利用 NADHP 为氮吸收等异化作用提供电子。通常认为，在高等植物中有多个 FNR 的同工酶，但在蓝细菌中只有一个 FNR 基因 *petH*。然而科学家在 *Synechocystis* sp. strain PCC6803 发现了两个 FNR 同工酶。其中一个为质粒体 FNR (34 kDa)，

专为异化作用。另一个 FNR(46 kDa)则在 N 端还有末端结构可以使其与藻胆体发生关联。科学家认为小的 FNR 是 NADPH 氧化酶。而较大的 FNR 是一个 NADP 还原酶[28]。和集胞藻 PCC6803 不同,聚球藻 PCC7942 中只有一个 FNR 并且蛋白较小。体外实验已经证明,已有报道其 Fd1 和 FNR 支持 ADO 酶反应[29]。具有比原核表达比绿色植物提取更方便经济的优点。在替代商业化垄断产品方面具有一定的潜力。

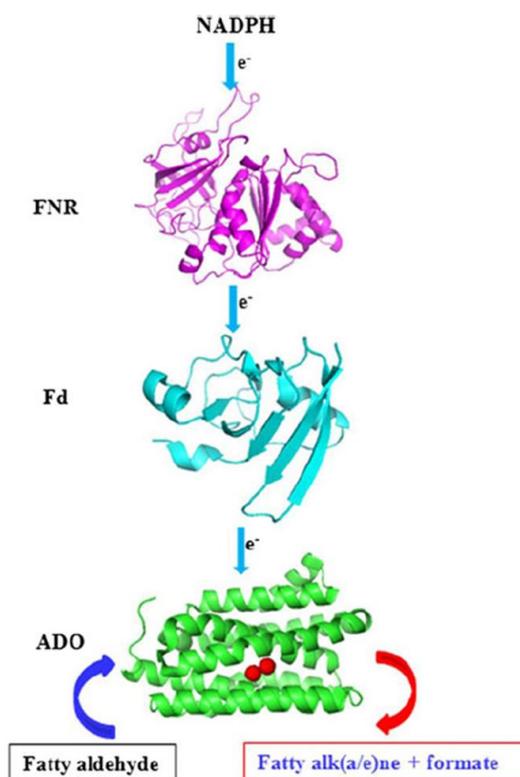


图 1-6 蓝藻 Fd1 和 FNR 支持 ADO 的电子传递示意图[29]

4.3 天蓝色链霉菌 P450 和还原伴侣研究背景

链霉菌属生产三分之二用于医药的天然抗生素以及共 9000 余种具生物活性物质,其次生代谢产物合成途径广受关注。天蓝色链霉菌是链霉菌属中的代表菌株,基因信息最全[30],是研究天然产物合成的模式生物(图 1-7)。

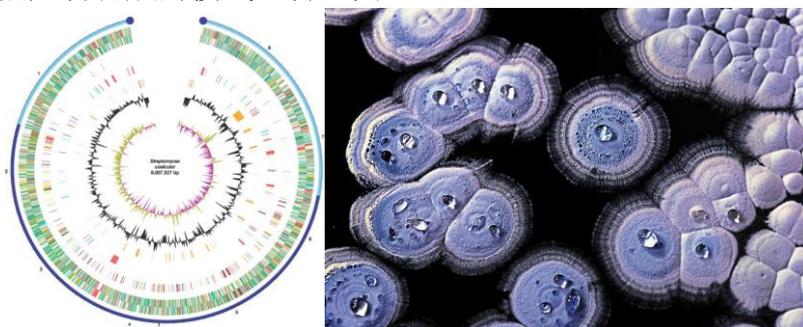


图 1-7 天蓝色链霉菌基因组(左)及菌落(右)

科学家们力图将它转化成更佳的药物开发工具。天蓝色链霉菌含有 8,667,507 bp,编码 7,825 预测基因,其中包括 20 余基因簇用来合成天然产物或未知化合物[31]。其中有 18 个

P450, 6 个 Fdx 和 4 个 FDR[32]。其中, 105D5 最匹配的还原伴侣是 FDR1 和 Fdx4[33]。105N1 的结构分析证明其参与 Zincophore, Coelibactin 的合成[34]。105N1 的晶体结构也得到了解析[35]。其他已经得到报道的 P450 还有 154A1[36], 同时 154C1 的晶体结构得到了解析[37]。

天蓝色链霉菌中含有 18 个 P450, 6 个 Fdx 和 4 个 FDR 基因, 其基因图谱信息见图 1-8 所示。

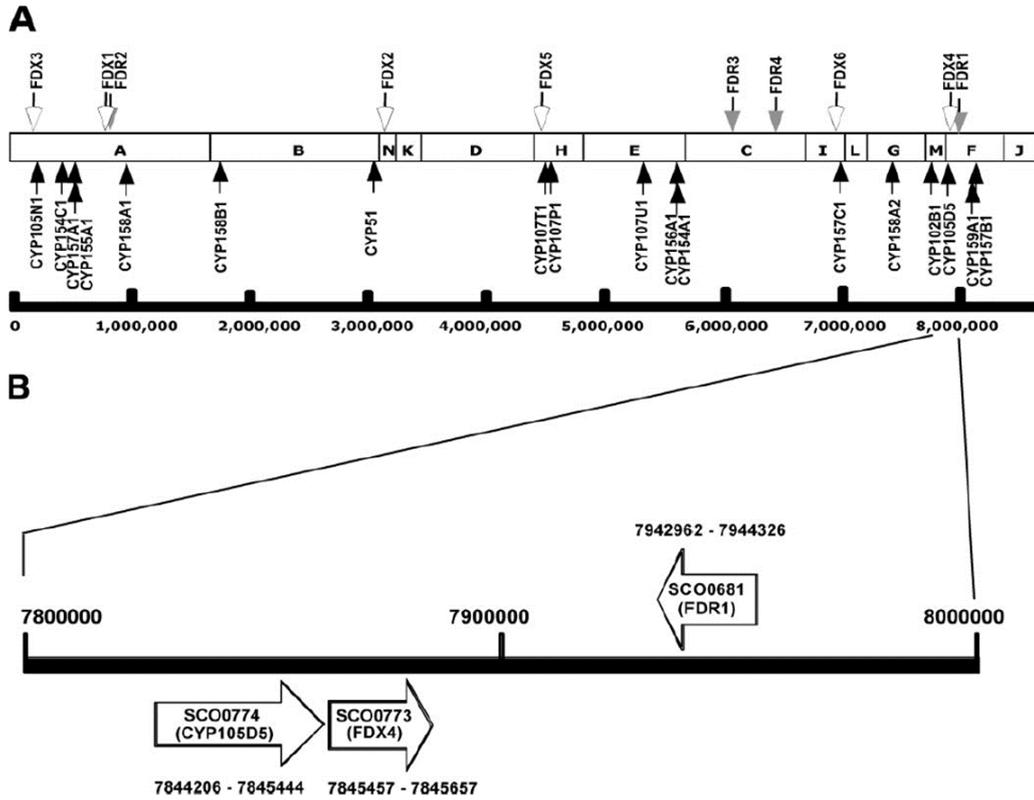


图 1-8 FDR, Fdx (FDX), 和 P450 (CYP)基因在天然色链霉菌基因组中的位置[33]

第二章 实验方法与手段

1 实验原料、试剂及设备

1.1 主要实验仪器

本实验中用到的主要仪器详见表 2-1。

表 2-1 本文使用的主要实验仪器

仪器	生产厂家
台式高速冷冻离心机	Eppendorf
蛋白电泳系统	Bio-RAD
Cary50Bio 紫外/可见分光光度计	Varian
GC-MS Agilent 7890-5795	Agilent
GC-MS Agilent 6890-5973	Agilent
Gel Doc XR 凝胶成像分析系统	Bio-RAD
30 °培养箱	上海志诚
Himac CR22G II 高速离心机	Hitachi
HVE-50 高压灭菌锅	日本 HIRAYAMA
LI-250 Light Meter	美国 LI-COR
PCR 仪	Bio-RAD
pH 计	美国丹佛
Power Pac Basic 型稳压稳流电泳仪	Bio-RAD
QHZ-98B 型光照摇床	太仓市华美生化仪器厂
Sigma-14 型台式高速离心机	Sigma
SW-CJ-2CD 标准型双人/单人净化工作台	苏净集团安泰公司
Tprofessional Thermocycler PCR 仪	Biometra
T6-新锐可见分光光度计	北京普析通用
雪花牌 XB-130 制冰机	宁波新芝生物科技股份有限公司
超声波破碎仪 VCX-500	美国 SONICS 公司
冰箱及 4℃冷藏柜	中国海尔
-80℃超低温冰箱	Thermo
恒温水浴器	上海精宏仪器厂
全温震荡摇床	上海精宏实验设备有限公司
移液器	Brand
小型离心机	Eppendorf
Thermo 大型离心机, 超速离心机	Thermo

1.2 实验试剂

本论文中使用的主要化学试剂以及主要分子生物学试剂详见表 2-2。

表 2-2 本文使用的主要实验试剂

Wisard® SV Gel and PCR Clean-up system	Promega
VB1	北京索莱宝科技有限公司
PVDF 膜	Roche (美国)
第一抗体 Mouse anti-His	北京索莱宝科技有限公司
第二抗体 AP-rabbit anti Mouse AgG	生工生物工程(上海)股份有限公司
IPTG	

BCIP-NBT	Amresco
DMSO	Sigma
dNTPs	Takara
DNA markers	Takara
酵母质粒提取试剂盒	Omega
Pfu DNA 聚合酶	MBI Fermentas
pMD18-T/simple	Takara 生物技术公司
SDS	北京索莱宝科技有限公司
Taq DNA 聚合酶	MBI Fermentas
Taq Master Mix	北京博迈德生物技术有限公司
5-ALA	Aladdin
Tris	生工生物工程（上海）股份有限公司
T4 DNA 连接酶	MBI Fermentas /Thermo
SC-Ura	北京泛基德分装
cOmplete ULTRA Tablets	Roche
质粒提取试剂盒	Omega/博迈德
琼脂糖	西班牙（北京索莱宝科技有限公司）
LB/LB 琼脂	博迈德、生工
限制性内切酶/Fast Digest	Takara/ MBI Fermentas/Thermo
抗生素	北京索莱宝科技有限公司
正己烷（色谱纯）	SK Chemicals 韩国
正十五烷醇	Sigma
正二十烷烃	Sigma
正十七酸	Sigma
其它试剂	国产分析纯
引物	上海桑尼生物科技有限公司
基因合成	金唯智
基因测序	上海桑尼，金唯智

1.3 实验菌株及菌株来源

本研究采用大肠杆菌作为克隆菌株，采用酿酒酵母作为表达系统。具体的菌株信息及来源详见表 2-3。

表 2-3 本文使用的宿主菌

Strains	Source
<i>E. coli</i> DH5 α	Takara company
<i>E. coli</i> BL21 (DE3)	Takara company
<i>Saccharomyce scerevisiae</i> INVSc1	Invitrogen
<i>Saccharomyce scerevisiae</i> BJ5464-NpgA	Gifted by Professor Tang Yi at UC Irvine

2 大肠杆菌的培养

大肠杆菌菌株于 LB 培养基中，置于 37°C 摇床培养（180 rpm）。如有需要添加氨苄抗生素（100 $\mu\text{g}/\text{mL}$ ），氯霉素（34 $\mu\text{g}/\text{mL}$ ），卡纳霉素（50 $\mu\text{g}/\text{mL}$ ）。酿酒酵母菌株于 YPD 培养基，30°C，250rpm 在摇床振荡培养。若在固体平板上培养，则采用 SC-Ura 培养基进行筛选尿嘧啶缺陷型的菌株。其他营养缺陷型培养基采用相应培养基即可。

3 大肠杆菌*E.coli*感受态的制备与转化

大肠杆菌感受态细胞的制备：

- (1) 用接种针挑取大肠杆菌 *E.coli* DH5⁻ (-80℃的甘油保存菌)，在 LB 平板培养基上分级划线，以能够出现单菌落为宜，将该平板倒置于 37℃ 的培养箱里培养过夜。
- (2) 从划线 LB 平板上挑取单菌落，接种到高压灭菌的 50 ml 新鲜液体 LB 培养基中（装在 250 ml 锥形瓶中），置于 37℃ 恒温摇床中（200 rpm）震荡培养。
- (3) 测定 OD₆₀₀ 值，当 OD₆₀₀ 达到 0.35-0.5 时，将培养物放置冰中停止培养，冰浴 10 min。
- (4) 将培养液置于预冷的离心管中，4000 rpm 离心 5 min，弃上清（尽量除尽上清）。
- (5) 在离心管中加入 5 mL 预冷的 0.1 M 的 CaCl₂ 溶液，轻轻弹动离心管使菌体悬浮（不要剧烈震荡），冰浴 10 min。
- (6) 4000 rpm 离心 5 min，弃上清（尽量除尽上清）。
- (7) 加入 1.6 mL 0.1 M 的 CaCl₂ 和 0.4 mL 80% 的甘油重悬细胞，50 μL/管分装感受态细胞，-80℃ 保存。

大肠杆菌的转化方法：

- (1) 取出一支 -80℃ 保存的感受态细胞，置于冰中融化 5 min。
- (2) 向融化的感受态细胞中加入 5-10 μL 连接产物或质粒，轻轻混匀，冰水浴中放置 30 min。
- (3) 42℃ 热激 90 秒，然后立即冰水浴 2 min。
- (4) 加入 900 μL LB 液体培养基，37℃、200 rpm 震荡复壮 1 h。
- (5) 5000 rpm 离心 3 min，弃上清，沉淀重悬于 100 μL LB 培养基中，取适量涂 LB（含一定浓度的相应抗生素）固体平板。
- (6) 将平板倒置于 37℃ 的培养箱里培养 12-16 h 后长出菌落。

4 酿酒酵母*Saccharomyce scerevisiae*的培养与电转化筛选

4.1 酿酒酵母培养方法

YPD 培养基 SC 培养基（适用于营养缺陷型）30℃ 培养箱和摇床 250rpm 培养。

4.2 酿酒酵母电转化方法

建立遗传转化系统是进行遗传操作和研究的前提条件之一。酿酒酵母可以通过电转化方法或者化学转化法吸收外源 DNA。酿酒酵母野生型和突变株在 30℃ YPD 培养基和 SC-Ura 中培养。用于提取和测试产物的菌株采用摇床 250 rpm 震荡培养，在 250 mL 三角瓶中装液量为 50 mL；具体步骤为：

1) 从-80℃取出菌种，在 YPD 培养基上划线培养两天。待单克隆长出后，挑取一个单克隆到 5mL YPD 液体中过夜摇菌（28-30℃，250rpm）。

2) 将 0.5mL 过夜培养的酿酒酵母培养液接种到 50mL YPD 培养基中，30℃，250 rpm 摇床培养 4-5 小时。当 OD600 达到 0.8-1.0 时，在 4℃离心力为 4000rpm 下离心 5min，去除上清液，将细胞沉淀用 50mL 冰置预冷的去离子水清洗两次，再用 20mL 预冷的 1M 山梨醇进行第二次清洗，最后定容于 150~250 μL 山梨醇中。

3) 在每个预冷的 0.2 cm 电转杯中加入 50 μL 酵母细胞和 4 μL DNA 混合液，置于冰上进行电转。电转条件：V=1.5 kV，C=25μF，R=200Ω，电击时间 T=4~5ms。

4) 迅速往电转杯中加入 1mL YPD 培养基，并将转化液转移到新的 1.5mL 离心管中，30℃摇床培养 1h。5000 rpm 离心 3min 收集细胞，用常温 1M 山梨醇清洗沉淀 2-3 次以去除培养基，最终定容于 1 mL 1M 山梨醇中。取 30-50 μL 酵母细胞液涂 SC-Ura 平板，30℃培养 2-4 天转化子可长出。

* 此方法也适用于转化质粒 DNA，应在电转化完成后往电转杯中迅速加入 1mL 1M 的预冷山梨醇。将转化液转移到 EP 管中，5min 后 5000rpm 离心 3min 收集细胞，取 30-50μL 酵母细胞液涂 SC-Ura 平板，30℃培养 2-4 天可长出转化子。

4.3 酿酒酵母的基因组提取

- 1) 取 1ml 过夜培养的酿酒酵母于 1.5mlEP 管中，3000rpm 离心 5min，弃上清
- 2) 加入 0.5ml TE 洗涤菌体，12000rpm 离心 1min，弃上清
- 3) 加入 200ul 石英砂悬液（10% w/v，研磨后的石英砂重悬于 TE 溶液）悬浮细胞，200ul 酚/氯仿/异戊醇(V/V 25:24:1)，高速漩涡振荡 5 min
- 4) 12000rpm 离心 5 分钟，将水相转移至干净 EP 管中，加入 1/5 体积的 5M 的 NaCl，2 倍体积的无水乙醇，-20℃沉降 1h
- 5) 4℃12000rpm 离心 10min，弃上清
- 6) 加 70%乙醇 500ul，12000rpm 离心 5 分钟弃上清
- 7) 12000rpm 再离心 1min，小心吸干残留液体，晾干 5min
- 8) 加入 20-50ul ddH₂O 溶解 DNA

4.4 酿酒酵母质粒提取

采用酵母质粒提取试剂盒。需要对酵母细胞进行石英砂磨碎处理。

质粒浓缩方法参照如下：

加入 1/5 V 的 5M NaCl 和两倍体积的无水乙醇，混匀后-20 度沉降 1 小时。12000rpm 离心 10min，沉淀用 70%乙醇清洗数次，在无菌台风干后，加入适量无菌水定容。

4.5 酿酒酵母菌落 PCR

收集酵母菌体，煮沸 10min。其余参照大肠杆菌菌落 PCR 方法进行。

4.6 酿酒酵母蛋白提取

1) NaOH 裂解法：以 OD600 = 0.2 起始培养至 OD600 = 0.8，收集 3 ml 菌液，13000 rpm 室温离心 2 min 后弃上清。加入 1ml 预冷的无菌水混合均匀，13000rpm 室温离心 2min 后弃上清。加入 100 μ l 无菌水，再加入 100 μ l 0.3M NaOH 混合均匀，室温静置 5min，离心去上清。加入 50 μ l 的 1X 上样缓冲液，95 度加热 5min。5000rpm 离心 5min，取 15 μ l 上样 SDS-PAGE。

2) 液氮研磨法：液氮研磨后直接样品处理，用来 SDS-PAGE。也可参照镍柱纯化蛋白方法进行大规模蛋白纯化分析。

5 DNA assembler 法在酿酒酵母系统中快速组装大片段 DNA

酿酒酵母 *Saccharomyce scerevisiae* 具备天然的重组系统，可以在重组片段 20-30 bp 重复片段的情况下高效同源重组。利用酿酒酵母天然的重组系统，可以将多个具有重叠区域的 assembler DNA 片段在酿酒酵母体内同源重组，组装成具有多个元件的表达质粒。此方法仅需要 PCR 准备 DNA 后电转化酿酒酵母即可完成，尤其适用于多个基因的串联组装、大片段基因和基因簇的快速组装[38]。

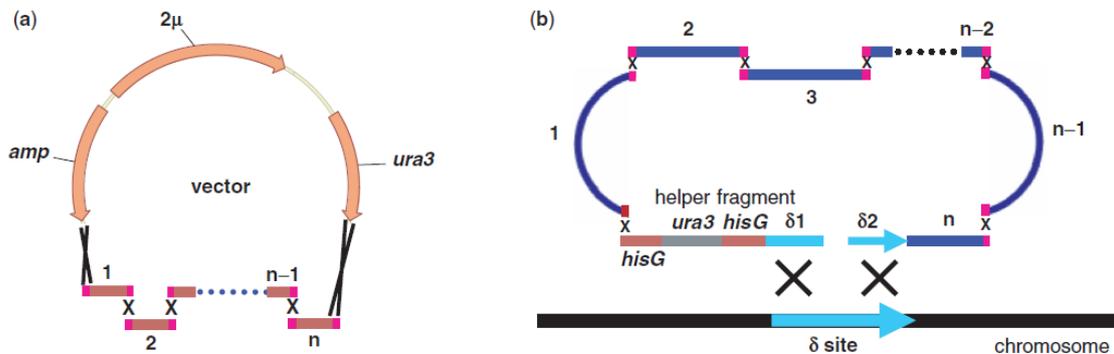


图 2-1 DNA assembler 法组装 PCR 片段

(a) 游离质粒法 (b) 整合质粒组装

具体方法为以下步骤：

1) Assembler 设计：如图中演示，将预期序列（表达载体+目的基因）切分为 3-5 个大小为 2-5kb 之间的 assembler 片段，根据 assembler 序列设计上下游引物，使各片段之间有 >50bp 的重叠区域以确保可以高效同源重组。设计通常可将筛选标记片段（如 Ura）切分到两个 assembler 中。

2) 采用 GXL 高保真酶 PCR 克隆各个 assembler 片段，一般来说 50 μ L 体系即可获得足够的 DNA。PCR 产物需经电泳鉴定大小，无误后，用 PCR 产物纯化试剂盒将各 assembler

片段浓缩到一个较小体积 (15-30 μ L)。纯化后的 assembler 混合液可直接用于转化酿酒酵母。

3) 操作注意事项:

a) 设计时应使各 assembler 之间有>50bp 的重叠区域以确保可以高效同源重组。

b) 感受态细胞制备过程中需采用 1M 山梨醇洗多次, 去除多余培养基和盐离子, 避免对电击时间产生影响。感受态细胞不宜在冰上放置太久, 应尽量马上制备马上转化。

c) DNA 的加入量及体积比对电转化效率影响较大, 应 50 μ L 酵母细胞加入<5 μ L DNA 混合液 (大约 5 μ g)。如 DNA 制备浓度不理想, 适当稀释转化体系后也可获得转化子, 但效率稍低。本实验室曾经尝试的比例有: 每 100 μ L 酵母细胞加入 20~30 μ L DNA 混合液 (大约 5 μ g) 也可成功。

d) 电转杯清洁处理: 一般先用去离子水冲洗、洗净, 然后再用无水乙醇冲洗、洗净; 放在超净台上吹风晾干后封口, 放入-20 $^{\circ}$ C 保存。

e) 转化后 1 mL 山梨醇的细胞重悬液不能全部涂在一块板子上, 一般 30-50 μ L 即可获得较多转化子。[38]

6 大肠杆菌和酿酒酵母转化子的基因鉴定方法

目前有两种鉴定蓝细菌转化子的方法, 分别为提取总基因组 DNA 后做 PCR 进行鉴定基因型, 和菌落 PCR 快速鉴定转化子的基因型。其中提取总基因组 DNA 的方法更为准确, 并且可以将总 DNA 保存在-20 $^{\circ}$ C 备用, 重复性好; 但同时也较为繁琐, 不适合鉴定样本数目庞大的转化子库。菌落 PCR 在大肠杆菌中是常用方法, 在蓝细菌也可以成功验证转化子基因型, 具有快速简洁的特点, 适合筛选大量样本库以及初步检测蓝细菌转化子基因型。

6.1 天蓝色链霉菌提取基因组总 DNA 及基因组 PCR

A. 链霉菌培养: 28 度, 250rpm 培养 2-4 天。

B. 链霉菌基因组提取: 根据 Practical Streptomyces Genetics, 采用 kirby-mix 方法提取链霉菌基因组。

C. 链霉菌基因组 PCR: 由于链霉菌基因组 GC 含量较高, 需在 PCR 反应体系中加入 5%-10% (v/v) 的 DMSO。基因组 PCR 时模板的量控制在 50ng/50 μ L 体系。

6.2 菌落 PCR 快速鉴定转化子基因型

6.2.1 大肠杆菌菌落 PCR

挑取少量大肠杆菌菌落细胞溶于 20 μ L 水, 取 1 μ L 作为模板加入 20 μ L 的 PCR 反应体系中; 大肠杆菌细胞经过 PCR 预热 95 $^{\circ}$ C 反应 5 min 即可裂解释放总 DNA。

6.2.2 酿酒酵母菌落 PCR

酿酒酵母细胞具有较厚夹膜层，需要沸水煮沸处理 10min。与常规的总 DNA 法 PCR 相比，具有一定的优势，可用于实验室快速检测聚球藻基因突变株的基因型。

7 聚合酶链式反应 PCR (Polymerase Chain Reaction)

PCR 反应参照 Takala 公司 GXL 酶使用说明，略有改动，反应体系如表 2-1。

表 2-4 PCR 反应体系

50 μ L PCR system	Volume
ddH ₂ O	27.5 μ L
10 \times Buffer	5 μ L
with (NH ₄) ₂ SO ₄ , w/o MgCl ₂	
MgCl ₂	4 μ L
dNTP (10 mM)	1 μ L
DNA template	μ L (1 pg ⁻¹ μ g)
GXL DNA polymerase (5U/ μ L)	0.5 μ L
Primer 1	0.5 μ L (100 pm)
Primer 2	0.5 μ L (100 pm)

8 大肠杆菌蛋白样品制备及聚丙烯酰胺凝胶电泳 (SDS-PAGE)

8.1 菌落 PCR 快速鉴定转化子基因型

200 mL 的培养液离心收集，用 10 mL 40 mmol/L Tris-Cl (pH 8.0) 重悬 细胞，加入终浓度 1 mmol/L 蛋白酶抑制剂 PMSF。在冰水浴中超声破碎细胞，每超声 30s 间隔 30s~1 min，有效超声时间为 5 min。随后，5 000 \times g 离心 30 min 去除细胞残渣。得到的蛋白样品经过 SDS-PAGE 和 Western blotting 检测。

8.2 大肠杆菌蛋白纯化

A. 蛋白的诱导表达

- 1) 将-80 $^{\circ}$ C 冻存的菌株取出，于相应抗体平板上划线；
- 2) 挑取单菌落，接种于 5 mL 相应抗性的 LB 液体中，37 $^{\circ}$ C 过夜培养；
- 3) 按照 1:100 的比例将菌株接种于添加终浓度为 1 mM 维生素 B1，5% 甘油和相应抗生素的 LB 培养基中，37 $^{\circ}$ C 培养 3-4 h，至其 OD₆₀₀ 为 0.6-1；取 0.5mL 培养液冻存(S1)
- 4) 加入 IPTG 至终浓度 0.2 mM，同时加入 5-ALA 至终浓度 0.5 mM，16 $^{\circ}$ C 诱导表达 20 小时。取 0.5mL 培养液冻存 (S2)。
- 5) 用 50 mL 离心管离心收集菌体，保存于-80 $^{\circ}$ C。

B. Ni 柱纯化蛋白

1) 将离心收集的菌体从-80 $^{\circ}$ C 取出，室温解冻后加入 40 mL Lysis buffer 裂解液重悬，超声波破碎 (离心管置于装满冰的烧杯中，超声波探头浸入液面下 2cm)，30% 功率，工作 5s，间歇 5s，总共时间为 30-min；

- 2) 离心管用天平配平。4℃ 10000 rpm 离心 30 min 两次，将上清转移到新的离心管。取出 100 μ L 上清 (S3) 和少许沉淀 (S4) 留做 SDS-PAGE。
- 3) 向上清液中加入 1 mL Ni-NTA 树脂，4℃ 摇床慢摇混合 1 h。
- 4) 取混匀后的溶液装柱，用 100 mL-200 mL Wash buffer 洗涤，并用 G250 染色液检测，直至流出的洗液中无蛋白组分；留存 0.5mL 流出液 (S5)。
- 5) 用 4 mL Elution buffer 对结合在 Ni-NTA 树脂上的目的蛋白进行洗脱收集；留存 0.5mL 流出液 (S6)。
- 6) 对收集的目标蛋白采用 Millipore 超滤离心管进行浓缩，浓缩至体积 1 mL 左右；
- 7) 采用 GE Healthcare 的 PD-10 脱盐柱对目标蛋白样品进行脱盐。脱盐前首先采用 5 倍体积 Desalting buffer 平衡 PD-10 脱盐柱柱，待柱子平衡后，将样品加到柱子上，待样品全部进入柱子内后，柱顶加入 Desalting buffer 洗脱。以 G250 检测变蓝开始分开收集多管。
- 8) 对收集后的脱盐目标蛋白进行混匀，分装，-80℃ 保存。

8.3 聚丙烯酰胺凝胶电泳 (SDS-PAGE)

聚丙烯酰胺凝胶电泳 (SDS-PAGE) 参考《分子克隆实验指南》Molecular Cloning 描述的方法，将离心好的蛋白样品与 5 \times SDS 凝胶上样缓冲液混合后，100℃ 煮沸 10 min。然后将煮沸后的样品经 12000 rpm 离心 1 min，取上清进行 SDS-PAGE 电泳。蛋白电泳参考《分子克隆实验指南》配制 12% SDS-PAGE 和 5% SDS-PAGE 的分离胶。浓缩胶的电泳条件是 50 V，60 min 至样品在浓缩胶上被压成一条细线；分离胶的电泳条件是 100 V，时间可根据实际电泳的情况和目的蛋白的大小灵活掌握。电泳完毕，将玻璃板分开，将胶剥离下来，以考马斯亮蓝染色液 (水/乙醇/乙酸 = 50: 40: 10, v/v, 0.1% 考马斯亮兰 R-250) 染色，将其放置在水平摇床上温和震荡数小时，之后将染色结束的胶用脱色液 (水/乙醇/乙酸 = 50: 40: 10, v/v) 脱色液脱色，将其放置在水平摇床上温和震荡数小时，至条带清晰时换掉脱色液，用 ddH₂O 漂洗，然后拍照存档。

9 蛋白质的免疫印迹 (Western blot)

Western blotting 采用 PVDF 膜 (Roche, 美国)。利用 6 \times Histidine-标签探针标记目的蛋白，最终使用碱性磷酸酶显色试剂盒 (Amresco, 美国) 进行检测。具体参考 Molecular Cloning 的方法，具体操作步骤如下：

- 1) 配置转膜液 (3.03 g/L Tris, 14.414 g/L 甘氨酸, 200 mL/L 甲醇)，将电极板和海绵泡在转膜液中 15 min；
- 2) 准备 PVDF 膜及滤纸：测量聚丙烯酰胺凝胶的大小 (一般为 6 \times 8 cm)，裁剪一张 PVDF 膜 (聚偏二氟乙烯，略比胶大)，同时裁剪 6 张滤纸 (略比胶小)；
- 3) 甲醇处理 PVDF 膜：先用甲醇处理 PVDF 膜 5 min 使其带上负电荷，待膜变得半透

明时，用纯水漂洗 5 min；

4) 转膜液平衡：将膜、滤纸和胶在转膜液中浸泡 15 min；

5) 铺板：按照负极板、海绵、三张滤纸、胶、膜、三张滤纸、海绵、正极板的顺序铺好，组成“三明治”结构，用板夹夹好。

6) 转移：将夹好后的“三明治”结构放入转膜仪，转膜时间根据目的蛋白分子量而定，大分子可适当长些，小分子时间短些，一般 30-70 KDa 的蛋白可选用稳压 100 V，1 h（将整个系统放在冰水里，电极板外用冰袋降温）；对于较小的目的蛋白，如 21 KDa 的 *E.coli* TesA' 一般 30 min 即可。

7) 封闭：待电转结束后，卸下电极板，通过预染 marker 或者丽春红染色的方法观察转膜效果。如果效果良好，将膜用水冲洗，用含有 5-10% 脱脂奶粉的 TBS (20 mM Tris-HCl, 150 mM NaCl) 溶液封闭 3 h，之后用 20 mL TBST (20 mM Tris-HCl; 150 mM NaCl 0.05%, v/v; Tween 20) 洗膜，摇床上温和振荡 5 min；

8) 一抗孵育洗涤：用 TBST (20 mM Tris-HCl, 150 mM NaCl, 0.05% (v/v) Tween 20) 按照 1: 1000 比例稀释鼠源一抗，将 PVDF 膜加入到一个塑料密封袋中，加入 10 ml TBST 和一抗，温和振荡 3 h。孵育结束后用 TBST 洗膜三次，摇床上温和振荡，每次 15 min；

9) 二抗孵育洗涤：将马抗鼠 IgG-AP 二抗稀释 2000 倍，将稀释后的二抗与膜继续孵育 1 h。在一个小的密封袋中加入 10 ml TBST 和二抗，放入膜，温和振荡孵育 1 h 后，用 TBST 洗膜三次，摇床上温和振荡，每次 15 min；

10) 显色：把膜放入密封袋中，加入 3~4 mL BCIP/NBT solution 碱性磷酸酶显色液避光显色，约 10-30 min 显出条带，将膜放入水中终止显色，拍照保存。

10 酿酒酵母蛋白提取

收集处于对数或平台生长期的酿酒酵母细胞 (OD₆₀₀ 在 左右)，离心去除上清液，加入 1 mL 50 mM pH 8.0 的 Tris-HCl 缓冲液，再加入适量的石英砂，放置冰中 5 min，然后将其在涡旋振荡器上剧烈震荡，震 20 s，停 15 s 并放置冰中冷却，震荡工作时间 10 min。吸取细胞破碎液转移至新的冷却的 EP 管中，4℃ 13000 rpm 离心 5 min，收集上清用于 Western blot 检测，用 Bradford 法测可溶的蛋白浓度，用 5 mg/mL 的 BSA 蛋白作为标准品绘制蛋白标准曲线，Western blot 检测时上样量在 20 μg 以上。

11 CO 差光谱法测定 P450 单加氧酶的活性。

本报告参照 Omura 等报道的方法进行 P450 蛋白浓度测定[39,40]，采用 CO 差异光谱对 P450 酶进行测定。CO 饱和的 P450 酶液采用瓦里安 UV/Vis 分光光度计在 400~500nm 进行全波长扫描，随后加入适量硫代硫酸钠还原后，再次进行全波长扫描。两次 P450 酶

液在 400~ 500 nm 的全波长扫描结果进行比较，并按照在 450~490nm 间的吸光度之差和摩尔吸光系数 ($91 \text{ L mmol}^{-1} \text{ cm}^{-1}$) 计算 P450 的含量。

其计算公式是： $[(A_{450} - A_{490})_{\text{observed}} - (A_{450} - A_{490})_{\text{baseline}}] / 0.091 = \text{nM P450 of per ml.}$

第三章 酿酒酵母中表达霉酚酸合成相关蛋白

本章主要通过酿酒酵母中表达来自短密青霉的霉酚酸合成相关基因，获得目的蛋白用于体外实验，同时检测酿酒酵母体内产物积累。

1 材料与方法

1.1 酿酒酵母菌株的培养

菌株在 30℃ 条件下进行培养，突变株 Syn-FQ20 和对照菌株 Syn-HB1567 在含有 10 mg/L 壮观素的 BG11 中培养。用于测定 β -半乳糖苷酶活性的藻株在装样量为 20 mL 的 50 mL 三角瓶中进行摇床培养，摇动速度为 140 rpm。用于 SDS-PAGE 及 Western Blotting 的藻株，在装样量为 300 mL 的 500 mL 三角瓶中进行通气培养。

1.2 菌株的构建过程

本章主要在酿酒酵母表达 MPA 三个相关基因 *mpaC*、*mpaDE*、*mpaA*（图 3-1）。其中 *mpaC* 较大，在大肠杆菌中不易表达，*mpaDE* 在大肠杆菌中表达失败，*mpaA* 在大肠杆菌中也未能表达。鉴于酿酒酵母是真核系统，比大肠杆菌具有一定的优点。

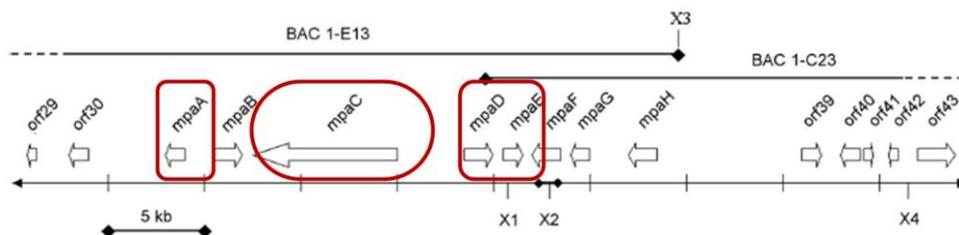


图 3-1 MPA 合成基因簇

1.2.1 *mpaC* 的克隆表达

本节 *mpaC* 的克隆及实验流程详见图 3-2，其中包括 *mpaC* 的克隆、蛋白表达纯化分析和突变株的发酵培养基产物的检测分析等。

因为 *mpaC* 序列较大，并且含有多种酶切位点，不适用于酶切连接方法插入到表达载体中。因此，本节采用一种新的方法，使 *mpaC* 与载体在酿酒酵母体内重组，并得到正确的表达质粒。

具体操作步骤参见第二章。

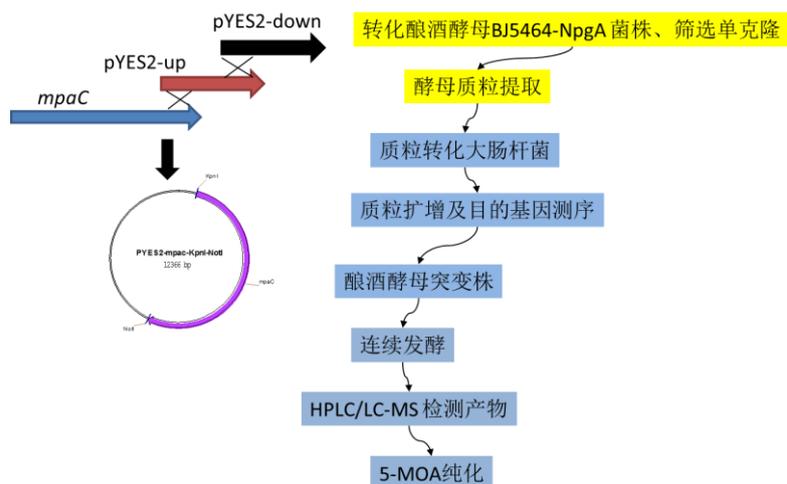


图 3-2 酿酒酵母表达 *mpaC* 的构建及实验步骤

首先，本章以天蓝色链霉菌的基因组 DNA 为模板，以表 3-3 中的引物 PCR 得到了 3 个 DNA 片段并电转化到酿酒酵母 BJ5464-*NpgA* 中。筛选克隆后经酵母菌落 PCR 筛选已经成功转化重组 DNA 片段的克隆并提取质粒。将质粒转化到大肠杆菌中进行扩增并送测序，所有序列经过测序确定无误。

表 3-3 *mpaC*-pYES2 的 PCR 引物

引物名称	5'→3'
Assembler 1: <i>mpaC</i> linker 克隆	
A1-sense:	ctctagatgcatgctcgagcggccgcctaaagctggacatgtgttctcaagaac
A1-anti-sense:	ctatagggaatattaagcttggtaccatgaattccacaaagggcaaccgaag
Assembler 2: <i>pyes2-Ura up</i> 克隆（以 <i>ura</i> 方向）含有 GAL promoter	
A2- sense :	cttcggttgccctttgtgaaattcatggtaccaagcttaattccctatag
A2-anti sense:	ctgggccaccacaccgtgtgcattc
Assembler 3: <i>pyes2-ura down</i> 克隆（以 <i>ura</i> 方向）	
A3- sense	cctatggagggcacagttaagccgct
A3-anti sense:	gttcttgagaacacatgtccagctttaggcggccgctcgagcatgcatctagag

1.2.2 *mpaDE* 的克隆及载体构建

同样，以天蓝色链霉菌的基因组 DNA 为模板，以表 3-4 中的引物 PCR 得到了 *mpaDE* 和 *mpaDE-NT* 片段并转化大肠杆菌中。筛选克隆后经菌落 PCR 筛选阳性克隆并提取质粒送测序，所有序列经过测序确定无误。*mpaDE* 在合成时有两组设计，分别为不带有标签和带有 *histag* 标签，记为 *mpaDE* 和 *mpaDE-NT*。设计引物见表 3-4。其中，*mpaDE-KpnI* 和 *mpaD/ER* 为 *mpaDE* 的合成引物；*mpaDEhis-R* 和 *mpaD/ER* 为 *mpaDE-NT* 的合成引物。*mpaDE* 和 *mpaDE-NT* 均采用酶切连接法插入到 pYES2 和 pYES2 NTA 的 KpnI 和 NotI 位点。重组质粒经过 PCR 和酶切双重验证，其序列经过测序确认，以上均无误后电转化到酿

酒酵母菌株 InvSC1 中。

表 3-4 *mpaDE* 的 PCR 引物

Primers	5'→3'
mpaDE-KpnI:	CGGGGTACCATGGAGTCTTTGTCGCTAACATGG
MPAD/ER:	ATTTGCGGCCGCTTACTTCTGTCCTTCTATGGC
mpaDEhis-R:	ATTTGCGGCCGCTTCTGTCCTTCTATGGCATT

1.2.3 *mpaA* 的密码子优化及质粒构建

本章以天蓝色链霉菌的基因组 DNA 为模板，以表 3-5 中的引物 PCR 得到了 *mpaA* 片段并电转化到片段并转化大肠杆菌中。筛选克隆后经菌落 PCR 筛选阳性克隆并提取质粒送测序，所有序列经过测序确定无误。

表 3-5 *mpaA* 的 PCR 引物

Primers	5'→3'
mpaA-KpnI:	cggggtacctATGACCAACGCAGTGGAGG 加入 T 防止移码
mpaA-NotI:	ATTTGCGGCCGCTcaAAGCTTAATGTACCCCTcaac 加入终止密码子

密码子优化的 *mpaA* 由金唯智合成，具体序列详见附表。

1.3 蛋白纯化与提取

参照第二章方法。

1.4 聚合酶链式反应 PCR

参照第二章方法。

1.5 质粒提取与鉴定

参照第二章方法。

1.6 膜蛋白提取与鉴定

采用碧云天膜蛋白提取试剂盒。

2 实验结果与分析

2.1 突变株的基因型鉴定

2.1.1 *mpaC* 质粒的鉴定

mpaC 克隆时以 cDNA 为模板，分别设计三组引物（表 3-3），PCR 得到三个具有 50bp 左右重叠片段的 DNA 产物 A1、A2、A3，纯化后共同转入酿酒酵母，在体内重组为一个完整的质粒。重组质粒经过 PCR 和酶切双重验证，其序列经过测序确认，以上均无误后电转化到酿酒酵母菌株 *Saccharomyces cerevisiae* BJ5464-*NpgA*[41]中。*Saccharomyces cerevisiae*

BJ5464-NpgA 菌株在基因组上整合了构巢曲霉的磷酸泛酰巯基乙胺基转移酶(PPTase)基因 *npgA*。具体流程详见图 3-2。

突变株中的外源基因表达元件均通过基因组 PCR 加以验证。所有的构建都过 PCR 及质粒酶切双重验证。所构建质粒按照第二章的方法转化宿主菌，得到突变株（表 3-4）。突变株的基因型经过 PCR 验证，如图 3-6 所示。

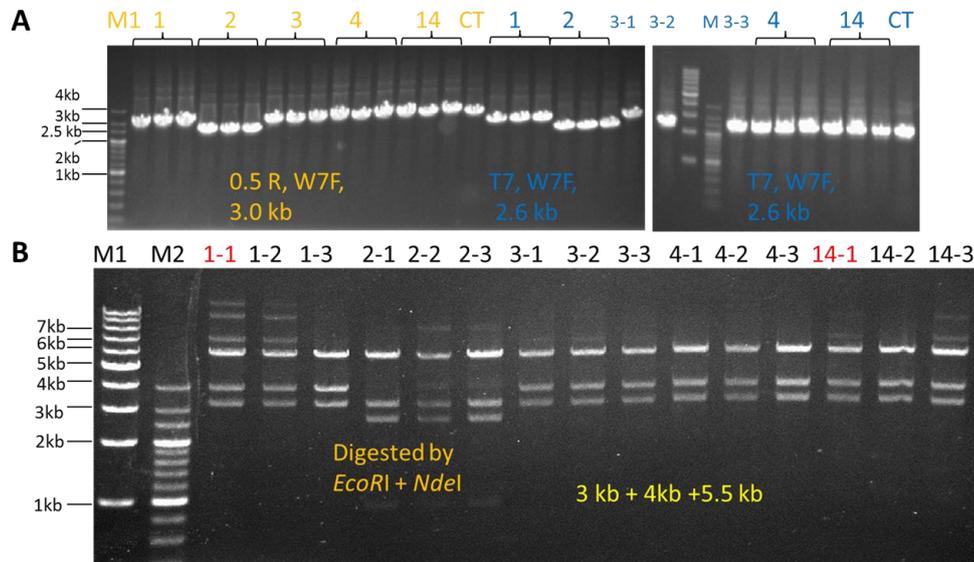


图 3-3 A. PCR 克隆 B. 酶切鉴定

PCR result for target genes in *E.coli* mutants. B. Enzymatic digestion of the constructed plasmids. W7F, 0.5 R, T7, W7F are primers used for colonial PCR.

2.1.2 *mpaC* 的内含子及外显子的鉴定

软件分析认为，本实验通过反转录得到 cDNA 再亚克隆到目的载体 pYES2，进行测序发现，通过测序确定了 *mpaC* 的正确序列，测序结果与原先软件预测的区域有所不同。*mpaC* 所有外显子序列最后确定详见附录。

2.1.3 *mpaC* 的产物检测

表达 *mpaC* 的酿酒酵母突变株中的产物检测结果见图 3-4。

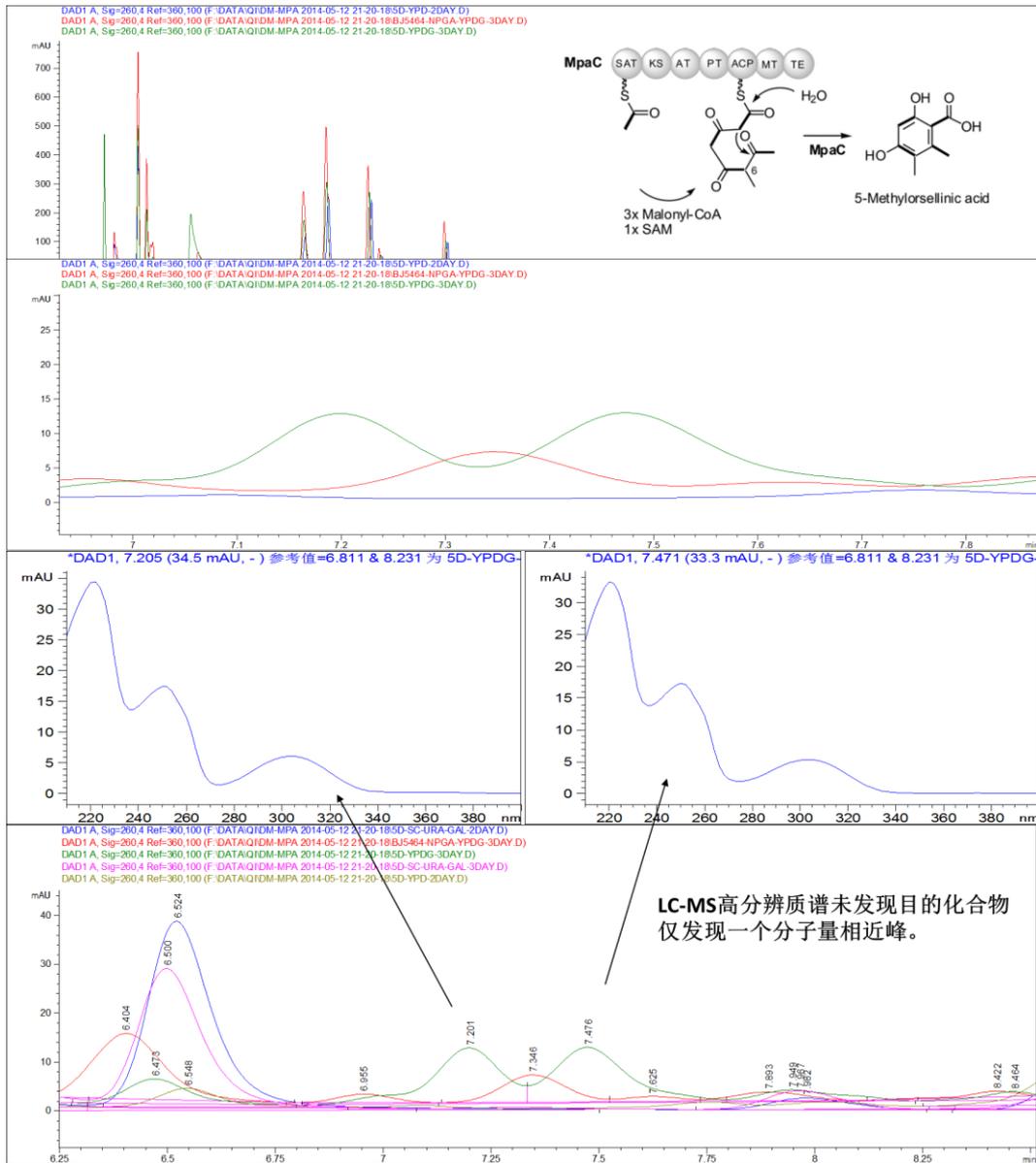


图 3-4 MpaC 催化的反应及 MpaC 催化产物 5-MOA 的检测

(A) the *lacZ* gene was inserted into the pFQ20 using *PrbcL* as the promoter. (B) Syn-LY2 was used as a negative control, harboring the Ω fragment in *slr0168* locus; Syn-HB1567 was used as the positive control applying the copper-induced promoter of *PpetE* and the *lacZ* reporter gene. Syn-FQ20 was the mutant strain that transformed by pFQ20.

2.3 蛋白的表达纯化分析汇总

2.2.2 MpaDE 的蛋白表达纯化分析及 Western Blot 检测

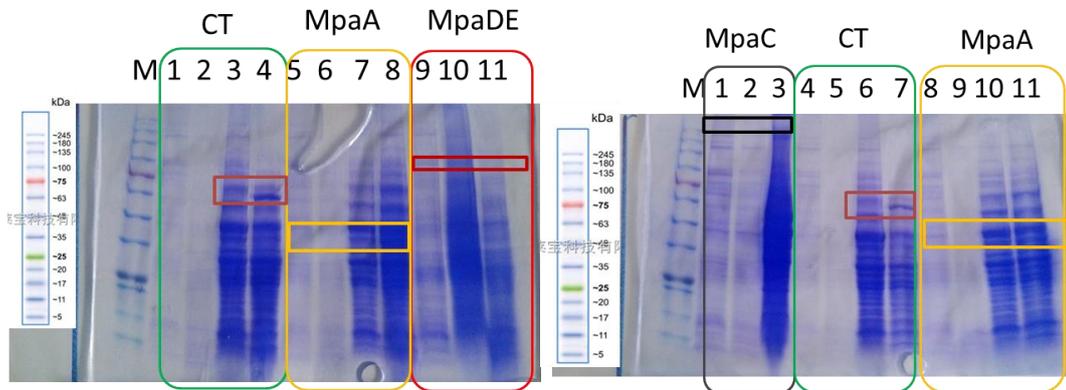


图 3-5 修饰后的启动子 SD-ATG 间隔序列对比

(左) CT pYES2-IMA 诱导前; 2. CT pYES2-IMA 诱导后; 3. CT pYES2-IGM1 诱导前; 4. CT pYES2-IGM1 诱导后; 5. mpaA-1 诱导前; 6. mpaA-1 诱导后; 7. mpaA-2 诱导前; 8. mpaA-2 诱导后; 9. mpaDE 诱导前; 10. mpaDE 诱导后; 11. mpaDE-YPD 培养诱导后。(右) 1. mpaC 诱导前; 2. mpaC 诱导后; 3. mpaC-YPD 诱导后; 4. CT pYES2-IGM1 诱导后; 4. CT pYES2-IMA 诱导前; 5. CT pYES2-IMA 诱导后; 6. CT pYES2-IGM1 诱导前; 7. CT pYES2-IGM1 诱导后; 8. mpaA-1 诱导前; 9. mpaA-1 诱导后; 10. mpaA-2 诱导前; 11. mpaA-2 诱导后;

SDS-PAGE 显示, MpaC 蛋白样品无法看到条带, 只有一条疑似条带, 但是蛋白大小不正确 (图 3-6)。

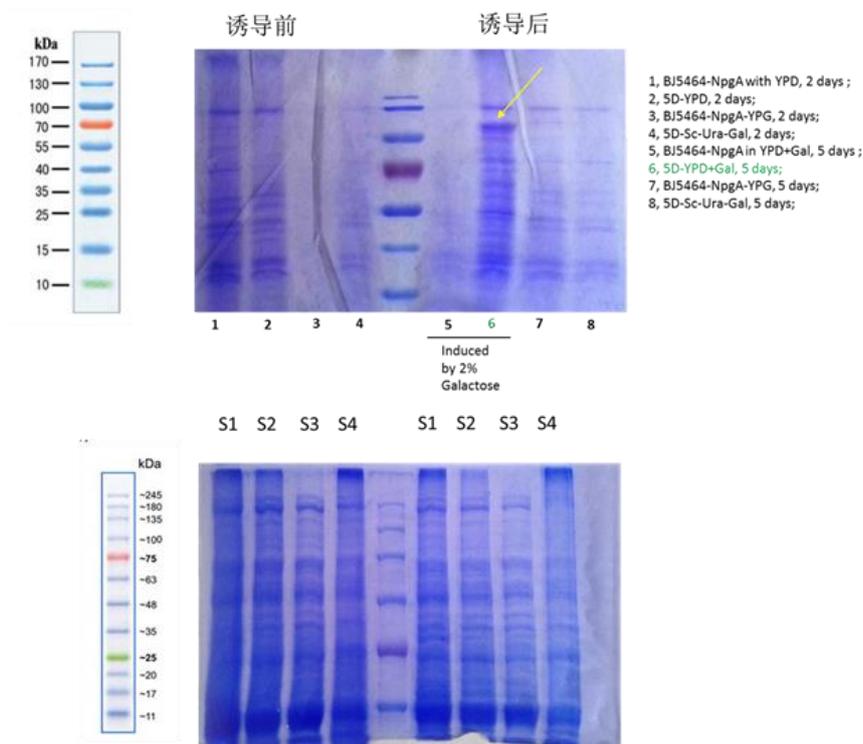


图 3-6 MpaC 的表达情况

同样, 本章对 100ml 表达 mpaDE-NT 基因的酿酒酵母蛋白进行纯化, 并进行了 SDS-

PAGE 分析，没有获得 mpaDE 蛋白的表达（图 3-7）。

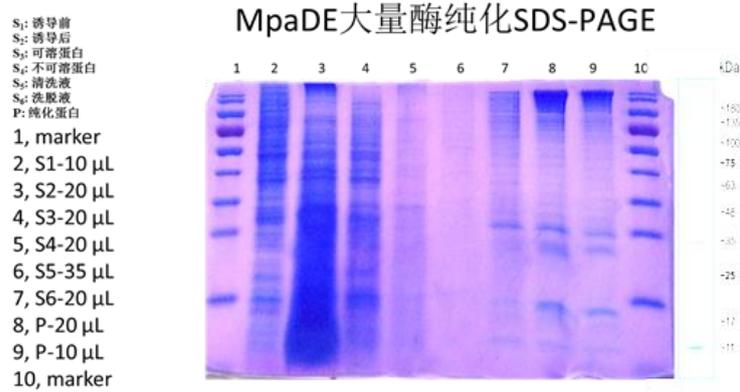


图 3-7 MpaDE 的大量表达纯化

mpaA 基因的酿酒酵母蛋白进行纯化及 SDS-PAGE 分析，没有 MpaA 蛋白的表达条带（图 3-8，图 3-9）。

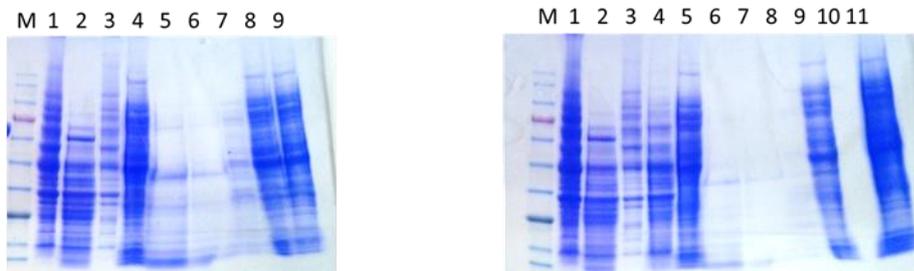


图 3-8 MpaA 的大量酶纯化

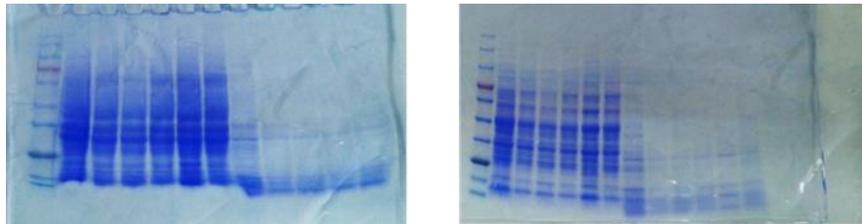


图 3-9 MpaA-Co 的表达情况

Western Blot 可以检测到 SDS-PAGE 胶图上肉眼不可见的目的蛋白条带。本章采用 histag 标签的对应抗体对 MpaDE 和 MpaA 进行了 Western Blot 检测分析（图 3-10，图 3-11）。

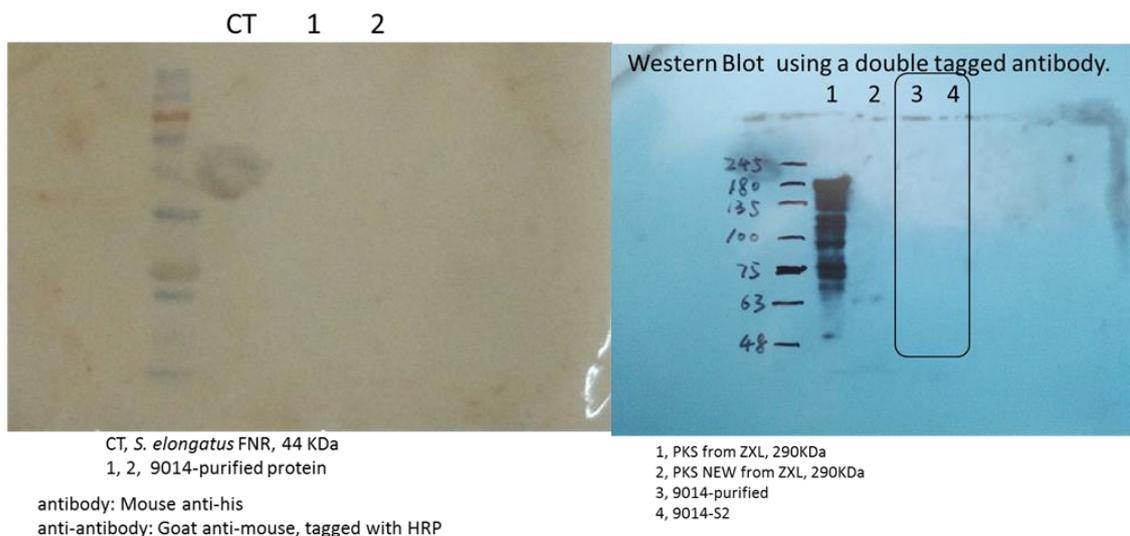


图 3-10 Western Blot 检测 MpaDE 的表达

(左) 利用 HRP 标记的抗体进行显色反应, (右) 利用两端标记的抗体进行检测

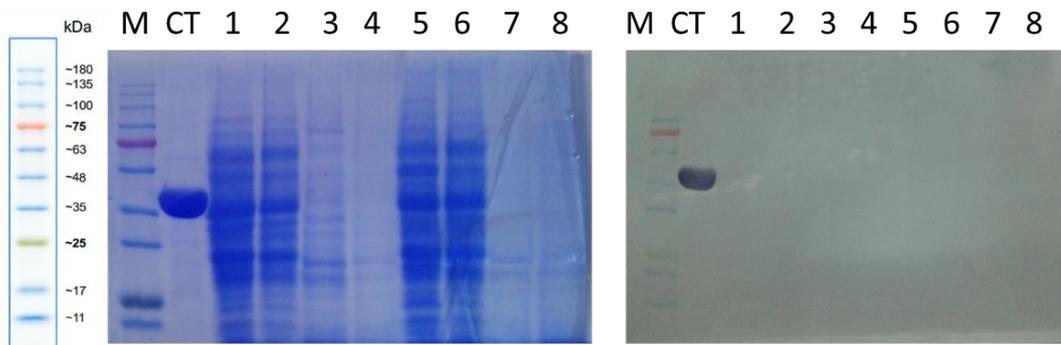


图 3-11 Western Blot 检测 MpaA 的表达

(左) SDS-PAGE 上样情况, (右) Western Blot 检测 MpaA 的表达; M: 蛋白 marker; CT: 具有 histag 标签的 FNR; 1: 克隆 1 总蛋白裂解液; 2: 克隆 1 总蛋白裂解液上清; 3: 克隆 1 膜提取试剂盒上清; 4: 克隆 1 膜提取试剂盒沉淀; 5: 克隆 2 总蛋白裂解液; 6: 克隆 2 总蛋白裂解液上清; 7: 克隆 2 膜提取试剂盒上清; 8: 克隆 2 膜提取试剂盒沉淀。

以上结果显示 MpaDE 和 MpaA 并没有表达。

3 本章小结

本章在酿酒酵母宿主中表达来自短密青霉的霉酚酸合成相关基因。首先成功打通了酿酒酵母同源重组系统组装大片段 DNA 的技术路线; 分别克隆了来自短密青霉的霉酚酸 (MPA) 合成相关基因 *mpaA*, *mpaDE* 和 *mpaC* 的。同时针对 *mpaC* 突变株开展了发酵实验, 通过 HPLC 未能检测到代谢产物 5-MOA; 三个蛋白均利用 SDS-PAGE 和 Western Blot 检测了蛋白表达情况。对 MPA 合成相关基因 *mpaA* 进行了密码子优化并蛋白纯化鉴定。但蛋白均未表达。

本节在酿酒酵母中构建表达 MPA 合成途径中三个关键酶, 均未能成功表达。分析可能是以下原因:

1) 单个基因是真核来源，在单细胞真核生物酿酒酵母中表达，具有较大的种属差异，本章均未对三个基因进行密码子优化，可能导致在大肠杆菌中蛋白表达受限。

2) 基因特殊性：首先，*mpaC* 大片段基因（7.5 kb）异源表达较难或不完整。其次，*mpaDE* 难度较大（可能需要其他元件），可溶性低，需要其他底物/辅因子等；再次，MpaA 是膜蛋白，具有多个跨膜区，酿酒酵母并不是优异的膜蛋白表达宿主。其 N 端序列对蛋白的折叠影响较大。

总之，酿酒酵母宿主在蛋白表达方面的短板，并非蛋白表达优良宿主，选择其他宿主如米曲霉可能更有利于来自于真菌短密青霉来源的蛋白的表达。

第四章 蓝细菌还原伴侣的克隆与表达

还原伴侣是 P450 重要支撑酶，不同来源的还原伴侣蛋白对 P450 具有不同的催化支撑效应。关于 P450 酶与还原伴侣的相互作用研究还比较欠缺，尤其是一些 P450 酶还苦于不能寻获其最佳还原伴侣而无法进行详细研究和开发，目前具体分子作用机制尚不明确。因此通用性高的还原伴侣选择及电子传递链优化策略的建立迫在眉睫。本章通过数据库搜索得到 7 个 Fdx 和 1 个 FNR。在大肠杆菌模式菌 *E.coli* BL21 (DE3) 中表达和纯化，是本报告中建立蛋白酶库工作之一。

1 材料与方法

1.1 菌株的培养

参照第二章材料与方法。

1.2 聚丙烯酰胺凝胶电泳 Western blot 检测目的基因的表达

方法参照第二章。

1.3 聚合酶链式反应 PCR

方法参照第二章。

1.4 蛋白纯化和 SDS-PAGE

参照第二章 Ni 柱纯化蛋白方法和 SDS-PAGE 方法。

2 实验结果与分析

2.1 基因检索分析

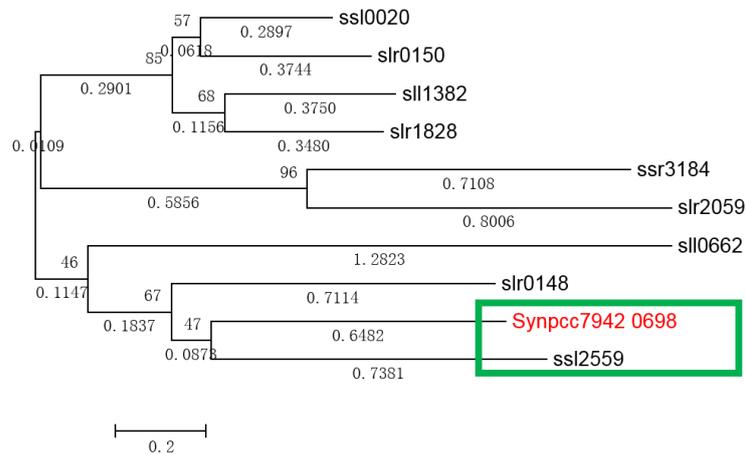
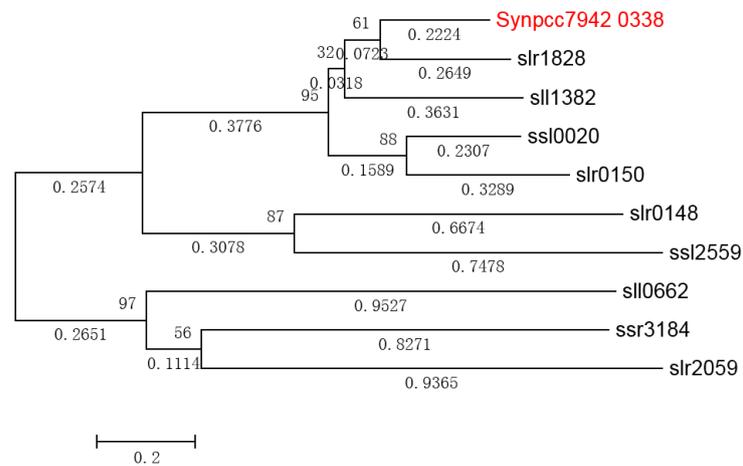
在总数据库中，含有 9 个 Fd (表 4-1) 和 1 个 FNR。

表 4-1 聚球藻 PCC7942 中的铁氧还蛋白信息汇总

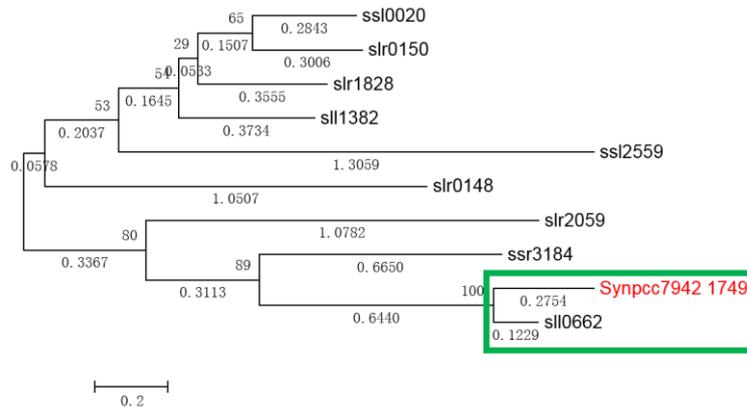
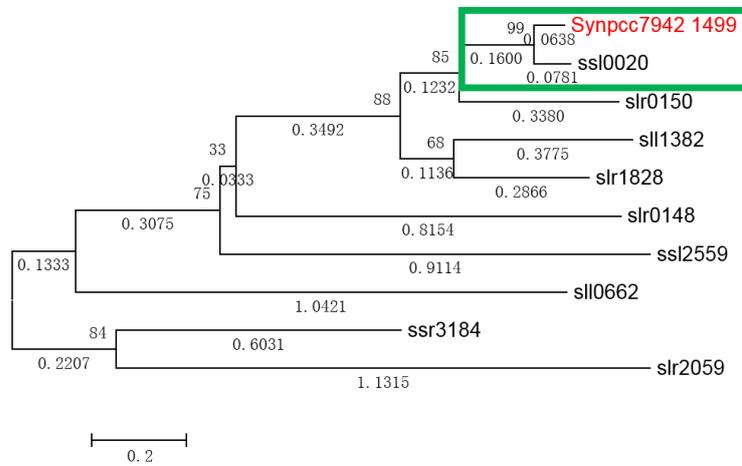
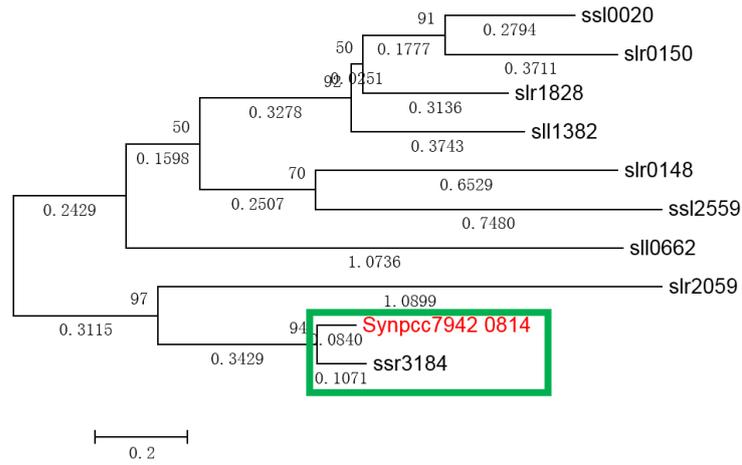
<i>Synechococcus</i> PCC7942	基因命名	Cyano base 注释	NCBI 注释
ORF0338	SYNPCC7942_0338	2Fe-2S ferredoxin like	Ferredoxin-2
ORF0698	petF5	2Fe-2S ferredoxin like	ferredoxin
ORF0814	petF1	4Fe-4S ferredoxins	ferredoxin
ORF0898	petF2	Thiordoxin like Fd	ferredoxin

ORF1499	petF3	2Fe-2S plant like Ferredoxin-1	Ferredoxin-1
ORF1749	petF4	[3Fe-4S] [4Fe-4S]	ferredoxin
ORF2581	SYNPCC7942_2581	2Fe-2S plant like	ferredoxin
ORF0624	SYNPCC7942_0624	4Fe-4S ferredoxins	4Fe-4S ferredoxin
ORF2129	SYNPCC7942_2129	4Fe-4S ferredoxins	ferredoxin

本章通过进化树建树，分别分析数据库中所有 7942Fd 与 6803 中 Fdx 蛋白的同源性 (图 4-1)。



第四章 蓝细菌还原伴侣的克隆与表达



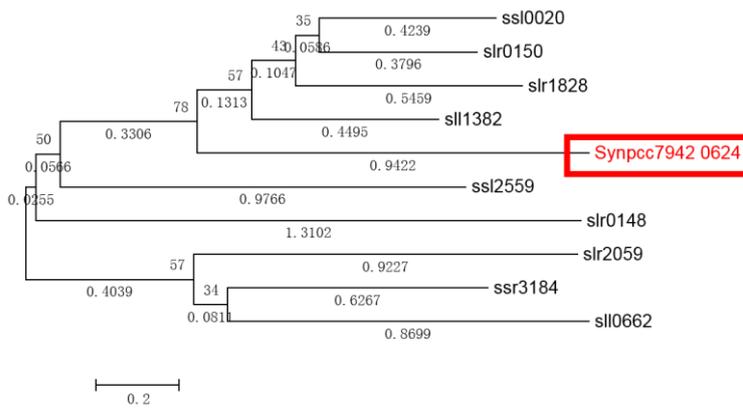
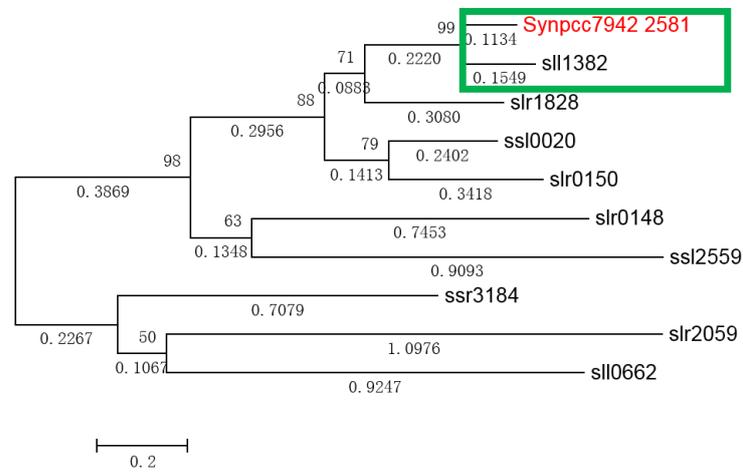
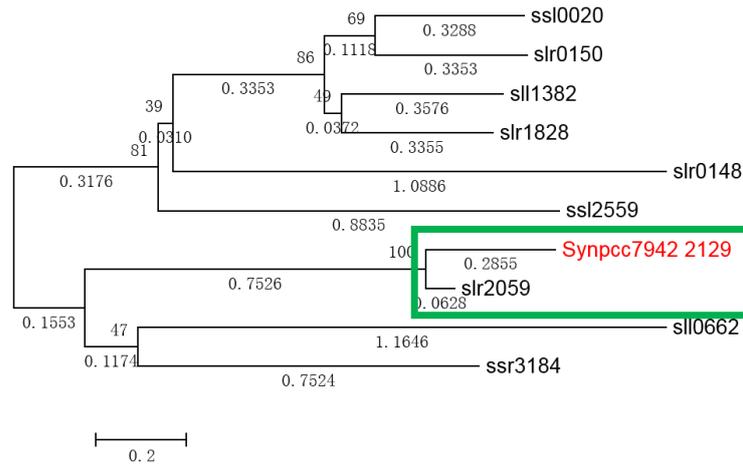




图 4-1 聚球藻 PCC7942 中的 Fd 与集胞藻 6803 中的 Fd 之间的进化树分析

本章以集胞藻 PCC6803 中的 9 个 Fd 为标杆分析了 72 株蓝藻中 Fd 的同源性。综合数据库数据，总共 9 个 Fdx。针对这 9 个 Fd 的生化分析中，只有 1499 有相关报道，其余均未报道；后面 2 个 Fd 只有基因组数据预测；与硫氧还蛋白毗邻，与已知亲缘关系较远（图 4-1）。

图 4-1，图 4-2 分析了聚球藻 PCC7942 中上下游毗邻基因信息，此分类结果与已发表文献结果相同[42]。

Cyanobacterial Species	Occurrence of Ferredoxin-Encoding Genes									
	Σ	[2Fe-2S]				Bacterial-type		[3Fe-4S] and [4Fe-4S]		
		<i>fed1</i>	<i>fed2</i>	<i>fed3</i>	<i>fed4</i>	<i>fed5</i>	<i>fed6</i>	<i>fed7</i>	<i>fed8</i>	<i>fed9</i>
<i>Synechococcus</i> sp. <i>WH8102</i>	4	+ ^b	+ ^c	+ ^d	-	-	-	+ ^g	+	-
<i>Synechocystis</i> sp. <i>PCC6803</i>	4	+	+ ^c	+	+ ^e	+ ^e	+ ^f	+ ^g	+ ⁱ	+
<i>Thermosynechococcus elongatus</i> <i>BP1</i>	4	+	+	+	-	+ ^e	-	+ ^g	+	+
<i>Synechococcus elongatus</i> <i>PCC6301</i>	3	+ ^b	+	+	-	-	-	+ ^g	+	+ ^j
<i>Synechococcus elongatus</i> <i>PCC7942</i>	3	+ ^b	+	+ ^d	-	-	-	+ ^g	+	+ ^j
<i>Synechococcus</i> sp. <i>CC9311</i>	6	2 ^b	+	+ ^d	-	-	-	+ ^g	+	-
<i>Synechococcus</i> sp. <i>CC9605</i>	6	+ ^b	+	+ ^d	-	-	-	2 ^g	+	-

图 4-1 聚球藻 PCC7942 与集胞藻 PCC6803 中 Fd 信息比较

霉酚酸微生物合成和 P450 单加氧酶与还原伴侣蛋白的相互作用机制研究

表 4-2 聚球藻 PCC7942 中 9 个 Fdx 的基因上下游信息及分类

329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	
hypothetical protein	ATP synthase subunit I	ATP synthase FOF1 subunit A	ATP synthase subunit C	ATP synthase subunit b	ATP synthase subunit B	ATP synthase subunit delta	ATP synthase subunit alpha	ATP synthase subunit gamma	Ferredoxin-2	hypothetical protein	hypothetical protein	abortive infection protein	hypothetical protein	hypothetical protein	photosystem II peptidase M23	
689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705
hypothetical protein	ATP-dependent protease adaptor ClpS	hypothetical protein	hypothetical protein	phosphoglycolate phosphatase	30S ribosomal protein S1	NrdR family transcriptional regulator	photosystem II center protein	photosystem II CP47 reaction center protein	ferredoxin	photosystem II reaction center protein	universal stress protein	hypothetical protein	hypothetical protein	DNA polymerase III	Vat family streptogramin A O-acetyltransferase	Cro/C1 family transcriptional regulator
805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821
hypothetical protein	16S rRNA (uracil(1498)-N(3))-methyltransferase	hypothetical protein	haloacid dehalogenase	hypothetical protein	hypothetical protein	sodium:proton antiporter	molecular chaperone DnaJ	ATP phosphoribosyltransferase regulatory	ferredoxin	ABC transporter ATP-binding protein	diguanylate cyclase	transcriptional repressor	phosphoribosylformylglycinamide synthase	phosphoribosylformylglycinamide synthase subunit PurQ	hypothetical protein	peptide ABC transporter substrate-binding protein
889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905
hypothetical protein	glutamate synthase	hypothetical protein	FO synthase subunit I	photosystem I reaction center subunit IX	shikimate kinase	hypothetical protein	septum site-determining protein MinD	cell division topological specificity factor	ferredoxin	hypothetical protein	glutamine synthetase	hydrolase	hypothetical protein	aconitate hydratase B	hypothetical protein	hypothetical protein
1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506
bicarbonate transport ATP-binding protein CmpC	bicarbonate transport ATP-binding protein CmpD	hypothetical protein	nucleoside triphosphate pyrophosphorylase	hypothetical protein	hypothetical protein	acetylglutamate kinase	membrane protein	hypothetical protein	Ferredoxin-1	ribosomal protein L11	phosphoglycerate dehydrogenase	hypothetical protein	RNA-binding protein	hypothetical protein	3-isopropylmalate dehydrogenase	hypothetical protein
1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	RS08930
	UDP-N-acetylmuramyl-L-alanine ligase	glyceraldehyde-3-phosphate dehydrogenase	NAD(P)H-quinone oxidoreductase subunit H	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	hypothetical protein	RNA polymerase sigma factor SigA2	aldo/keto reductase	hypothetical protein	ferredoxin	hypothetical protein	hypothetical protein	hypothetical protein	GNAT family N-acetyltransferase	no existed	sodium:proline symporter	hypothetical protein
RS13095	RS13100	RS13105	RS13110	2576	2577	2578	2579	2580	2581	2582	2583	2584	2585	2586	2587	2588
no annotation	no annotation	no annotation	16S ribosomal RNA	hypothetical protein	N-acetylmuramic acid 6-phosphate etherase	hypothetical protein	molecular chaperone DnaJ	chaperone protein dnaK3	ferredoxin	inositol monophosphatase	hypothetical protein	short-chain dehydrogenase/reductase	transcriptional regulator	hypothetical protein	hypothetical protein	hypothetical protein
615	616	617	618	619	620	621	622	623	624	625	626	RS03215	627	628	629	630
rhodanese-like domain-containing protein	HrcA family transcriptional regulator	hypothetical protein	adenosylhomocysteinase	alkaline phosphatase	membrane protein	hypothetical protein	ABC transporter ATP-binding protein	thioredoxin	4Fe-4S ferredoxin	hypothetical protein	dihydroxy-acid dehydratase	tRNA-Gln	hypothetical protein	polyamine aminopropyltransferase	hypothetical protein	hypothetical protein
2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136
"Ribonuclease III anti-sigma factor antagonist	carbamoyl-phosphate synthase small chain	anthranilate phosphoribosyltransferase	hypothetical protein	hypothetical protein	hypothetical protein	phycobilisome degradation protein NblA	thiol:disulfide interchange protein TxlA		ferredoxin	hypothetical protein	lytic transglycosylase	phosphoglucosamine mutase	hypothetical protein	N-acetylmuramoyl-L-alanine amidase	hypothetical protein	4-hydroxy-tetrahydrodipicolinate reductase
									Fed9 cluster J''				20			

注：九个 Fdx 分别是 0338,0698,0814,0898,1499,1749,2581,0624,2129；数字代表基因组中的位置及基因代码；根据 Fdx 基因上下游的信息，符合分类规范的有 Fdx1 属于 cluster b, Fdx9 属于 cluster J''。

2.2 质粒和菌株的构建

本章表达的聚球藻 PCC7942 的 7 个 Fd 和 1 个 FNR 信息如图 4-3 所示。

7942	基因命名	Cvano base	NCBI
0338	SYNPCC7942_0338	2Fe-2S ferredoxin like	Ferredoxin-2
0698	petF5	2Fe-2S ferredoxin like	ferredoxin
0814	petF1	4Fe-4S ferredoxins	ferredoxin
0898	petF2	Thiordoxin like Fd/2Fe-2S	ferredoxin
1499	petF3	2Fe-2S plant like	Ferredoxin-1
1749	petF4	Ferredoxin-1	ferredoxin
2581	SYNPCC7942_2581	[3Fe-4S] [4Fe-4S]	ferredoxin
2581	SYNPCC7942_2581	2Fe-2S plant like	ferredoxin

FDR Gene	gene description	Nt (bp)	Amino acids (aa)	MW (kDa)
FNR	NADPH dependent Ferredoxin reductase	1212	403	44

图 4-3 聚球藻 PCC7942 与集胞藻 PCC6803 中 Fd 信息比较

本章以聚球藻 PCC7942 的 DNA 为模板,以表 4-3 中的引物对 PCR 得到了 6 个 Fdx 片段并插入到 pET28b 的 NdeI 和 XhoI 位置。所有序列经过测序确定无误后转入到大肠杆菌 BL21 (DE3) 中进行表达。磁瓦,中科院青岛生物能源与过程研究所吕雪峰和李建军研究员[29]馈赠 pET28a-Fd1 和 pET28a-FNR 质粒,已经发表文章证明其有效性详见图 4-3。

表 4-3

Gene accession No.	Sense primer(NdeI)	Antisense primer(XhoI)
SYNPCC7942_0338	GGAATTC <u>CATATG</u> GCTACCTATCAA GTCGAAGTC	TCCGCTC <u>GAG</u> TCAGCCGGGCTGA CCAAAT
SYNPCC7942_0698	GGAATTC <u>CATATG</u> CCGAGCATTTCG TTTAT	TCCGCTC <u>GAG</u> TTAGGTCTCAGCT GTTGCGAT
SYNPCC7942_0814	GGAATTC <u>CATATG</u> GCGCACACCATT GTTAC	TCCGCTC <u>GAG</u> TTAGGGCGTTTGC TGCAACT
SYNPCC7942_0898	GGAATTC <u>CATATG</u> ACCCTCGCTGAA ACCCTG	TCCGCTC <u>GAG</u> CTATAGCGGATCG ACTGCTGGC
SYNPCC7942_1749	GGAATTC <u>CATATG</u> ACGACGCCGGA CCG	TCCGCTC <u>GAG</u> TTAAGTTCGTTCG CGTTGTTTC
SYNPCC7942_2581	GGAATTC <u>CATATG</u> AGCGACACTTAC ACCGTC	TCCGCTC <u>GAG</u> TTATTCCTCGTCCA AGGGCAG

2.2 突变株的基因型鉴定结果

本章所构建的表达质粒及突变株均经过基因组 PCR 验证其基因型 (图 4-4)。

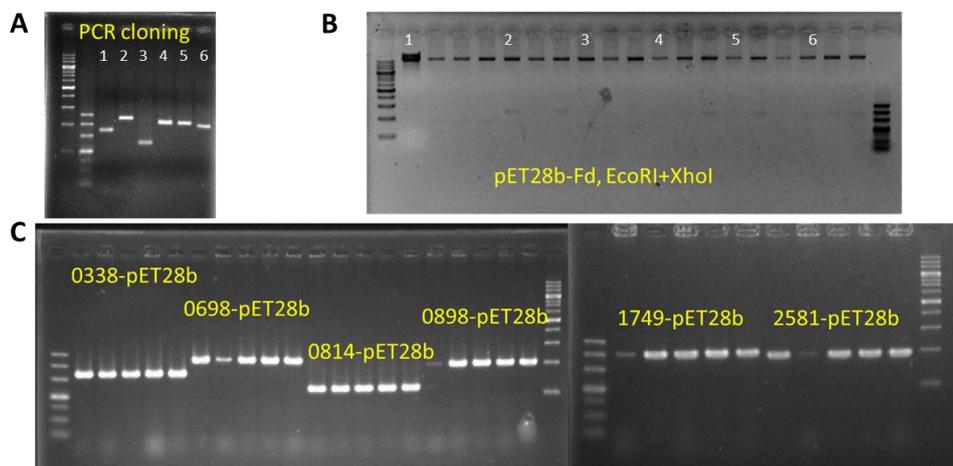


图 4-4 PCR 检测突变株的基因型

2.3 蛋白纯化与鉴定

2.3.1 SDS-PAGE 鉴定目的蛋白表达

图 4-5 展示的是突变菌株表达 FNR 和 Fdx 的情况。

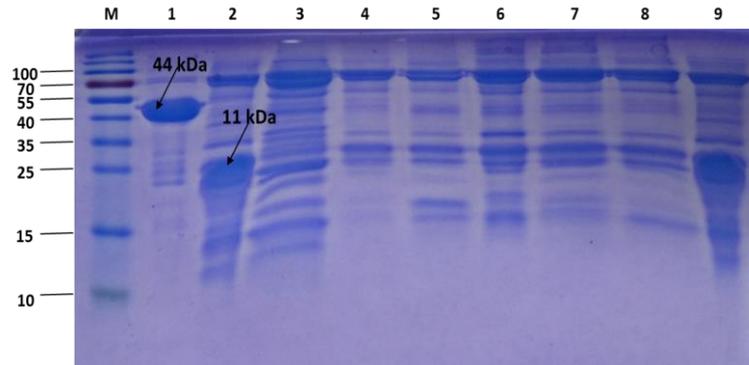


图 4-5 集胞藻突变株的生长曲线

1, FNR SYNPC7942_0978, NADPH dependent Ferredoxin reductase, 44 kDa; 2, petF3 (Fd) SYNPC7942_1499, ferredoxin (2Fe-2S), ferredoxin 1, 10.521 kDa; 3, SYNPC7942_0338 SYNPC7942_0338, ferredoxin (2Fe-2S), 11.132 kDa; 4, petF5, SYNPC7942_0698, ferredoxin 15.652 kDa; 5, petF1, SYNPC7942_0814, ferredoxin-like protein, 8.027 kDa; 6, petF2 SYNPC7942_0898, ferredoxin like protein, 14.98 kDa; 7, petF4 , SYNPC7942_1749 ferredoxin, 14.848; 8, SYNPC7942_2581, SYNPC7942_2581, ferredoxin (2Fe-2S), 13.786 kDa; 9, petF3 (Fd), SYNPC7942_1499, ferredoxin (2Fe-2S), ferredoxin 1, 10.521 kDa

以上结果显示, FNR 和 Fdx1 表达情况较好。未经过密码子优化的 Fdx2、Fdx3、Fdx4、Fdx5、Fdx6 在大肠杆菌中表达情况不好。

本章还表达菌株的蛋白组分进行了分析, 分别对诱导后蛋白沉淀和上清液以及全蛋白组分进行了 SDS-PAGE 分析 (图 4-6)。结果显示, 各个组分中并没有明显变化, 未能发现目的蛋白超量表达。

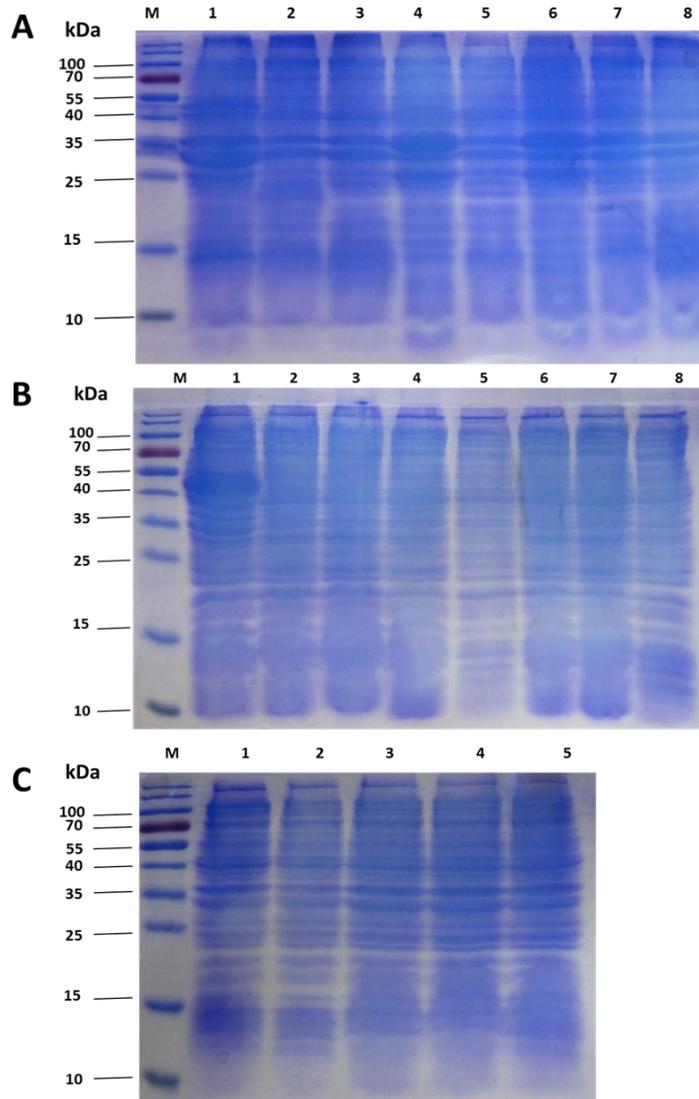


图 4-6 表达蓝细菌 Fdx 的蛋白组分分析

A: cell pellets; B: Soluble fractions; C: whole cell fractions; In A and B, 1, FNR, NADPH dependent Ferredoxin reductase, 44 kDa; 2, SYNPC7942_1499, ferredoxin (2Fe-2S), ferredoxin 1, 10.5213 kDa, SYNPC7942_0338, ferredoxin (2Fe-2S), 11.132 kDa; 4, SYNPC7942_0698, ferredoxin, 15.652; 5, SYNPC7942_0814, ferredoxin-like protein, 8.027 kDa; 6, SYNPC7942_0898, ferredoxin like protein, 14.98 kDa; 7, SYNPC7942_1749, ferredoxin, 14.848 kDa; 8, SYNPC7942_2581, ferredoxin (2Fe-2S), 13.786 kDa; In C: 1, SYNPC7942_0698, ferredoxin, 15.652 kDa; 2, SYNPC7942_0814, ferredoxin-like protein, 8.027; 3, SYNPC7942_0898, ferredoxin like protein, 14.98 kDa; 4, SYNPC7942_1749, ferredoxin, 14.848 kDa; 5, SYNPC7942_2581, ferredoxin (2Fe-2S), 13.786 kDa

3 本章小结

本章通承担了国家优青项目“微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制”研究课题中对蓝细菌来源的还原伴侣的分子克隆和蛋白纯化。首先通过数据库检索分析了 9 个 Fdx，同时选择其中 7 个 Fdx 和 1 个 FNR 进行克隆与表达。在未经密码子优化时，仅有 Fdx1 表达情况较好。本实验中其他 Fdx 表达量很低或不表达，推测可能与密

码子偏爱性有关，需要进一步密码子优化。本章为实验室进一步纯化和分析蓝细菌还原伴侣提供了指导，为下一步密码子优化等工作提供了信息。

第五章 天蓝色链霉菌还原伴侣基因克隆与蛋白纯化

本章按照合成生物学思路建立和充实 P450 催化元件库及还原伴侣蛋白库。本章在 *E.coli* 中表达来自天蓝色链霉菌的 18 个 P450 酶、6 个 Fdx 和 4 个 FDR 蛋白，并进行了镍柱纯化和光谱分析。通过尝试不同的大肠杆菌表达菌株提高蛋白表达量。

本章涉及的蓝色链霉菌的基因及蛋白信息如表 5-1 所示。

表 5-1

Genes (drnelson. utm.edu/Cytoc hromeP450.html)	Full name	DNA accession(Genbank/ Sanger center)	Protein_ID	Nt	Amino acids
Fdx1	ferredoxin	SCO7110/ SCO7110SC4B10.11	NP_631171.1	390	129
Fdx2	ferredoxin	SCO5135/ SCO5135SC9E12.20	NP_629284.1	321	106
Fdx3	ferredoxin	SCO7676/ SCO7676SC4C2.11	NP_631715.1	222	73
Fdx4	ferredoxin	SCO0773/ SCO0773SCF60.05c	NP_625075.1	201	66
Fdx5	ferredoxin	SCO3867/ SCO3867SCH18.04c	NP_628054.1	213	70
Fdx6	ferredoxin	SCO1649/ SCO1649SCI41.32	NP_625924.1	306	101
FDR1	ferredoxin/ferredoxin-NADP reductase	SCO0681/ SCO0681SCF15.02	NP_624989.1	1365	454
FDR2	ferredoxin reductase	SCO7117/ SCO7117SC4B10.18c	NP_631178.1	1263	420
FDR3	reductase	SCO2469/ SCO2469/ SC7A8.08c	NP_626711.1	1266	421
FDR4	oxidoreductase	SCO2106/ SC2C1A.02c	NP_626364.1	1224	407
CYP51	cytochrome P450	SCO5223/ SCO5223/ SC7E4.20	NP_629370	1386	461
CYP102B1	cytochrome P450	SCO0801/ SCO0801/ SCF43.12	NP_625102	1584	527
CYP105D5	cytochrome P450	SCO0774/ 3SCF60.06c	NP_625076	1239	412
CYP105N1	cytochrome P450	SCO7686/ SCO7686/ SC4C2.21	NP_631725	1236	411
CYP107P1	cytochrome P-450 hydroxylase	SCO3636/ SCO3636/ SCH10.14c	NP_627830	1236	411
CYP107T1	cytochrome P450 oxidoreductase	SCO3770/ SCO3770/ SCH63.17	NP_627960.1	1185	394
CYP107U1	cytochrome P450 hydroxylase	SCO3099/ SCO3099/ SCE41.08c	NP_627317.1	1302	433
CYP154A1	cytochrome P450	SCO2884/ SCO2884/ SCE6.21	NP_627112	1227	408
CYP154C1	cytochrome P450	SCO7417/ SC6D11.13c	NP_631466	1224	407
CYP155A1	cytochrome P450 (fragment), partial	SCO7444/ SC6D11.40c	NP_733733	1272	423
CYP156A1	cytochrome P450	SCO2883/ SCO2883/ SCE6.20	NP_627111	1233	410
CYP156B1	cytochrome P450	SCO6310/ SCO6310/ SCIF3.12	NP_630407	1344	447
CYP157A1	cytochrome P450-family protein	SCO7418/ SC6D11.14c	NP_631467	1242	413
CYP157B1	cytochrome P450	SCO0584/ SCO0584/ SCF55.08c	NP_624896	1263	420
CYP157C1	cytochrome P450	SCO1626/ SCO1626/ SCI41.09c	NP_625901	1497	498

		SCO6998/			
CYP158A1	cytochrome P450	SC8F11.24c	NP_631063	1224	407
CYP158A2	cytochrome P450	SCO12072/ SCG58.07	NP_625496	1215	404
CYP159A1	cytochrome P450	SCO0583/ SCF55.07	NP_624895	1224	407

1 材料与方法

1.1 大肠杆菌菌株的培养

野生型和突变株在 37°C, 1L 含有 5% 甘油的 LB 培养基中培养。如有需要, 向培养基中加入相应的抗生素 (卡那霉素 50 µg/mL, 氨苄霉素 100 µg/mL, 氯霉素 34 µg/mL 等), 多抗浓度各自减半。

1.2 菌株的构建

本章使用表 5-2 所列引物 PCR 克隆所涉及的 28 个目的基因, 并通过酶切连接克隆方法将目的基因连入 pET28b 表达质粒。将含有目的基因的质粒分别转化 BL21 (DE3) 并进行卡那霉素抗性筛选, 得到突变株 (附表 3)。将含有目的基因的质粒分别转 Codon plus RIPL 并进行卡那霉素、氯霉素、链霉素三抗性筛选, 得到突变株 (附表 3)。

表 5-2 PCR 引物

Genes	Sense Primer(NdeI)5'---3'	Anti-sense(EcoRI)5'---3'
Fdx1	GGAATTC <u>CATATG</u> ACTTACGTCATCGC	CCGGAATTC TCAGGAGGGGAACC
	ACAGC	ACAGAT
Fdx2	GGAATTC <u>CATATG</u> ACCTACGTCATCGC	CCGGAATTC TTACTGGTTCTGCGG
	GCAG	CGGCAG
Fdx3	GGAATTC <u>CATATG</u> ACCTTGGCAGGCCA	CCGGAATTC CTACCGAAGAGTGA
	GG	GCGCTCC
Fdx4	GGAATTC <u>CATATG</u> CACATCGGCATCGA	CCGGAATTC TCAGCCGACCCGCTC
	CAAGG	CGAG
Fdx5	GGAATTC <u>CATATG</u> AGGATCTCCGTCGA	CCGGAATTC TCAGCCCCGGACCCC
	CCCCG	TTCC
Fdx6	GGAATTC <u>CATATG</u> AGCGTGCAGCAGG	CCGGAATTC TCACTCTGAGTCCGG
	AGG	ACCGTAG
FDR1	GGAATTC <u>CATATG</u> CCCCGCCCTCTGC	CCGGAATTC TCAGGCGCCGCTCTC
	GGAATTC <u>CATATG</u> CCGCGTGCGAAGA	GC
FDR2	CG	CCGGAATTC TCACGCCTGGTCTCC
	GGAATTC <u>CATATG</u> GTCGACGCGGATCA	CGT
FDR3	G	CCGGAATTC CTATGCGACGAGGCT
	GGAATTC <u>CATATG</u> GTCGTGGTGGGCG	TTCG
FDR4	GGAATTC <u>CATATG</u> ACCGTCGAGTCCGT	CCGGAATTC CTACGCCGTGGCCGC
	CAAC	CGT
CYP51	GGAATTC <u>CATATG</u> GCCCAGACAGCGA	CCGGAATTC TCACCGCGCCACGGG
	GGGAAC	C
CYP102B1	TTAATTC <u>CATATG</u> ACGGACACCGACAC	CCGGAATTC TCAGTCACCCGCCCG
	GACGACGAACACCGACACGACG	GTG
CYP105D5	GGAATTC <u>CATATG</u> ACACCCCCGAATC	CCGGAATTC CTACCAGGCCACGGG
	CCCG	GAGTTCCAGC
CYP105N1	GGAATTC <u>CATATG</u> ACGGCTGCAACCG	CCGGAATTC TCACCAGGTCACCAT
	ATGG	CAGTTCTTCG
CYP107P1		CCGCTCGAG TCAGAGTTCCACGC
		GTAGCTC

CYP107T1	<u>GGAATTCATATG</u> GGGAGTGCTCCCCG TT	<u>CCGGAATTC</u> CAGCCGAGCCGCACCG G
CYP107U1	<u>GGAATTCATATG</u> ACCGGCAGCTCTTC CG	<u>CCGGAATTC</u> TCACTTTTGCCTCGG CGACG
CYP154A1	<u>GGAATTCATATG</u> GCGACCCAGCAGC C	<u>CCGGAATTC</u> TCAGCCGGCGTGACG CAG
CYP154C1	<u>GGAATTCATATG</u> ACGACCGGCACCG AAGAAG	<u>CCGGAATTC</u> TCAGGCCAGCCTGAC CGG
CYP155A1	<u>GGAATTCATATG</u> GCGGATCCTGGTGA CC	<u>CCGGAATTC</u> TCAGCCGCGGCCGGC C
CYP156A1	<u>GGAATTCATATG</u> ACGTTGCCCTCCAC CG	<u>CCGGAATTC</u> TTAGTGTGCGGCGGA TTGC
CYP156B1	<u>GGAATTCATATG</u> GACGCCACCACCCC CG	<u>CCGGAATTC</u> TCACCTCGTGAGCCA GTTGAG
CYP157A1	<u>GGAATTCATATG</u> AGCACCGACGCC AC	<u>CCGGAATTC</u> TCATGCTGACGTGCC TCCCAG
CYP157B1	<u>GGAATTCATATG</u> ACCGACATCGACCC GTCAC	<u>CCGGAATTC</u> TCAGGCCACGACGG GTGTGAAG
CYP157C1	<u>GGAATTCATATG</u> ACGCCTGAAAGCC ACTC	<u>CCGGAATTC</u> TCAGTAGCCGCGCCA CCAG
CYP158A1	<u>GGAATTCATATG</u> ACGCAGGAGACCA CCACG	<u>CCGGAATTC</u> CTACCAGGTGCAGGG CAGCG
CYP158A2	<u>GGAATTCATATG</u> ACTGAAGAAACGA TTTCCCAG	<u>CCGGAATTC</u> TCACCACGTCACCGG CAGG
CYP159A1	<u>GGAATTCATATG</u> TCCACCGCGCAGCA G	<u>CCGGAATTC</u> TCAGGCCGCGGGTGT GAAG

1.3 蓝细菌基因型的鉴定（菌落 PCR）

参照第二章 PCR 方法。

1.4 蛋白的诱导表达和纯化

蛋白的诱导表达纯化参照第二章提供的方法。对表达量较差的蛋白，浓缩直到看到颜色为止。

1.5 CO 法测 P450

方法参照第二章。

2 实验结果与分析

2.1 质粒和菌株的构建

本章采用表 5-1 的引物 PCR 合成了所有目的基因，包括 4 个 FDR、6 个 Fdx 和 18 个 P450。PCR 产物电泳结果见图 5-1。

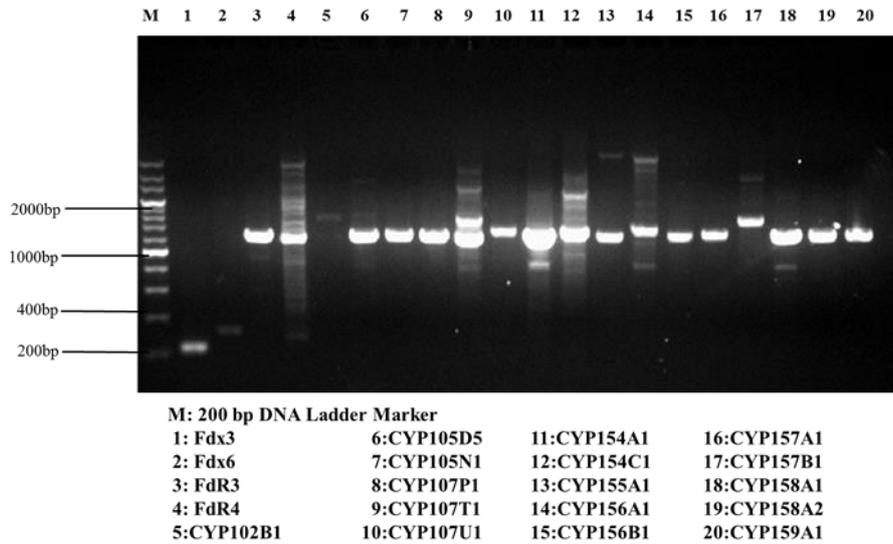
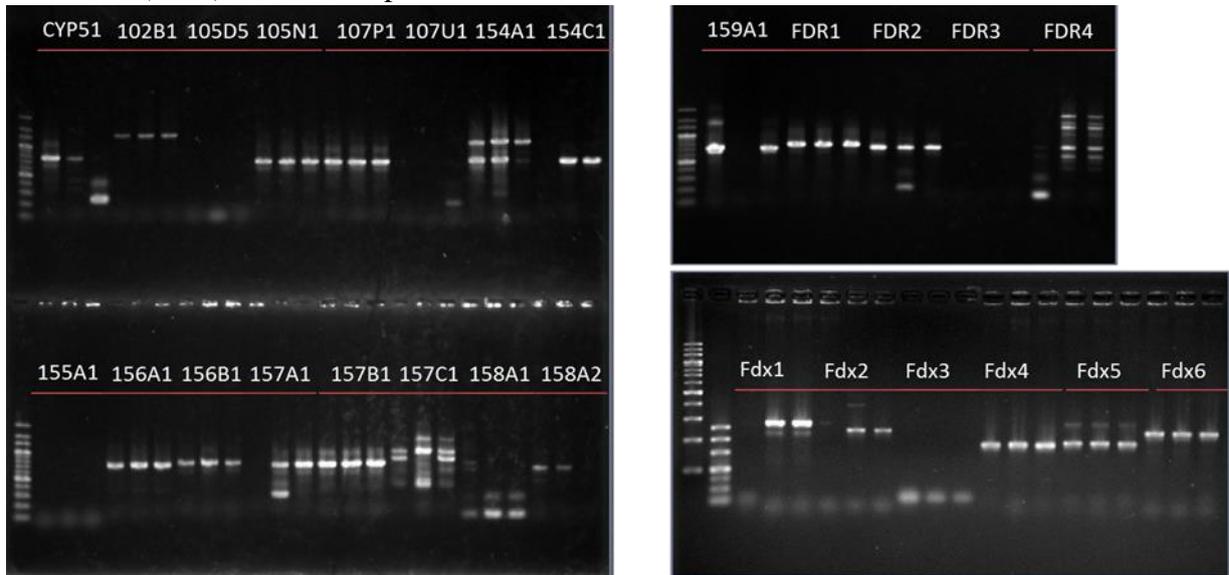


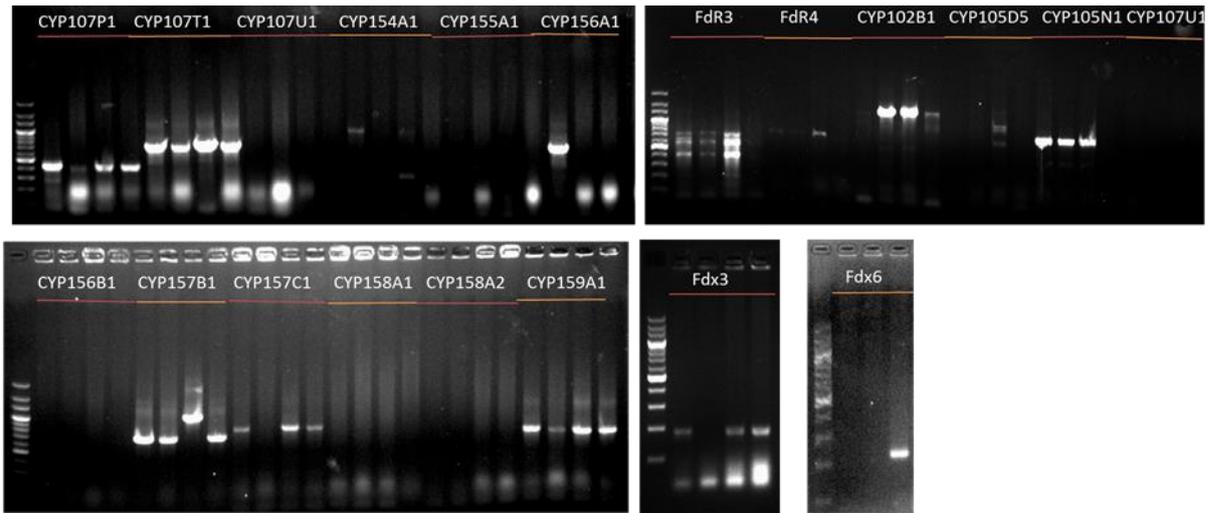
图 5-1 PCR 克隆目的基因

PCR 产物经酶切纯化后连接插入到 pET28b 的相应酶切位点，转化 *E.coli DH5α*，经过卡纳抗性筛选后得到转化子。所得克隆采用 T7 引物和目的基因的下游引物经过菌落 PCR 筛选得到正确的转化子（图 5-2，图 5-3）。挑取其中 4 个克隆进行液体培养后提取质粒，采用对应的酶进行双酶切验证无误后送测序。测序无误的克隆进行保存，同时质粒转化 *E.coli BL21(DE3)* 或者 codon plus RIPL 菌株，得到突变株，所有突变株信息详见附录 3。



Primers used: T7, antisense primer of each genes; Annealing temperature was 55 Celsius degree.

图 5-2 PCR 鉴定转化子基因型 (I)



Primers used: T7, antisense primer of each genes; Annealing temperature was 55 Celsius degree.

图 5-3 PCR 鉴定转化子基因型 (II)

2.2 天然色链霉菌 P450 及还原伴侣蛋白在大肠杆菌中的表达纯化与鉴定

所有突变株按照第二章所列方法进行培养和蛋白表达并进行 SDS-PAGE 检测分析。

2.2.1 天蓝色链霉菌 6 个 Fdx 的表达

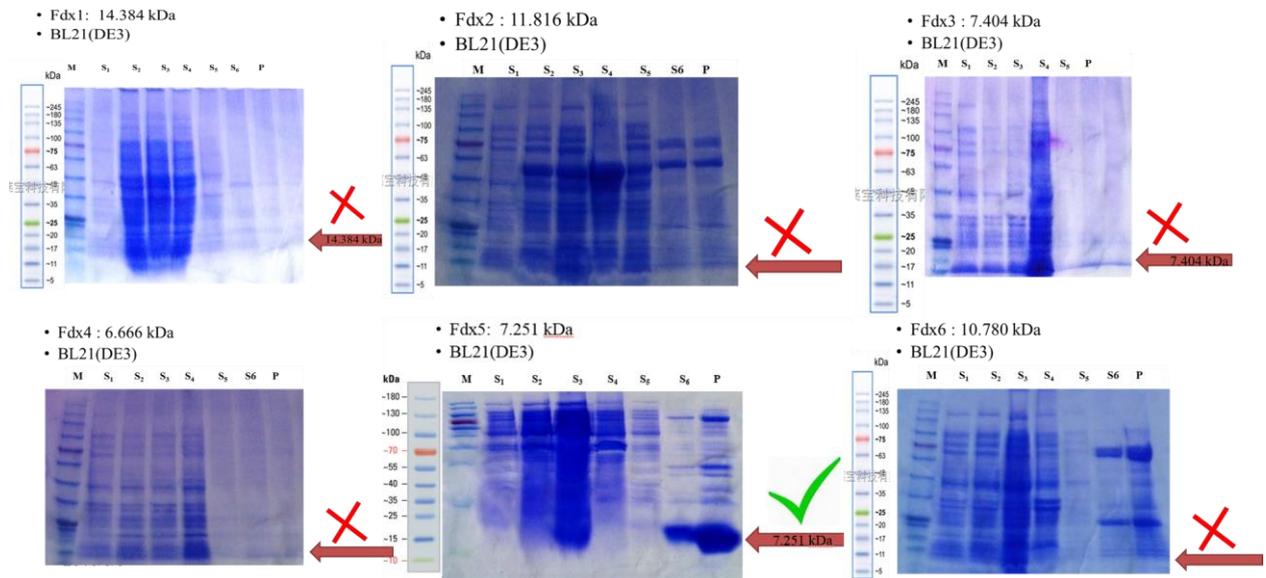


图 5-4 *Streptomyces coelicolor* A3(2)蛋白纯化 6 个 Fdx 转化 *E.coli* BL21(DE3)

以上结果显示, 在 BL21 (DE3) 菌株中, 未经过密码子优化的 Fdx 表达受到较大影响, 仅 Fdx5 表达较好。在尝试采用具有稀有密码子的表达菌株 RIPL 后 (图 5-5), 其余 Fdx 有效提高了蛋白的表达, 但 Fdx6 的表达量仍然较低。

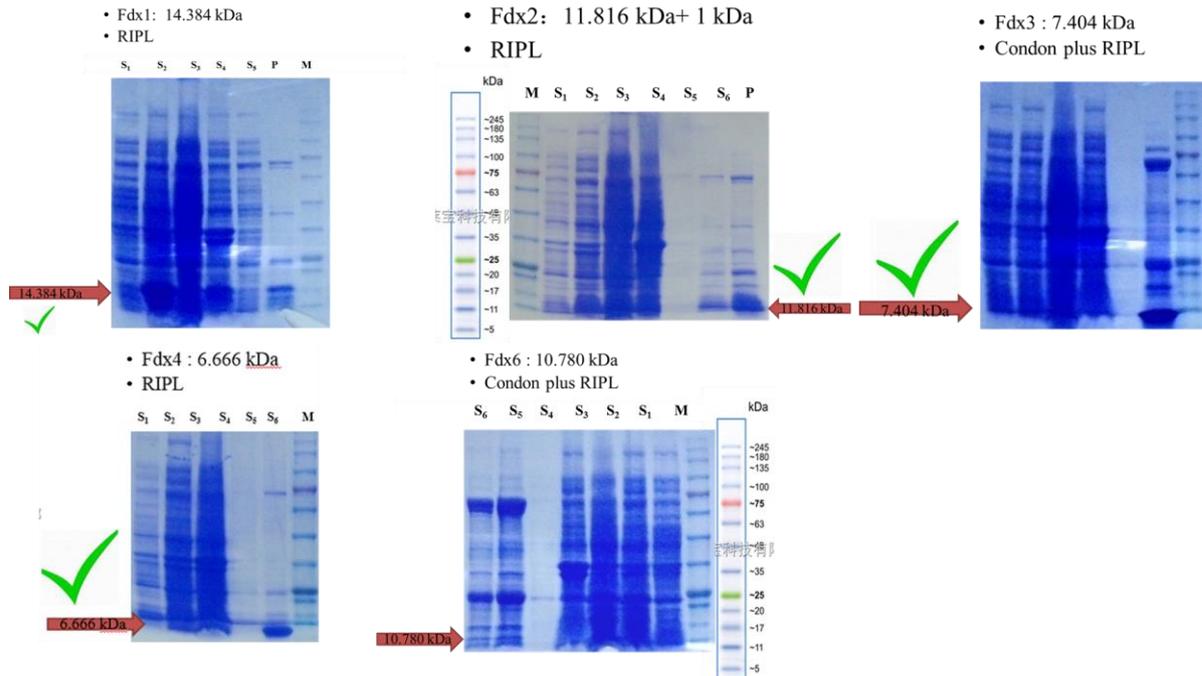


图 5-5 *Streptomyces coelicolor* A3(2)蛋白纯化 6 个 Fdx 转化 RIPL-codon plus strain

2.2.2 天蓝色链霉菌 4 个 FDR 的蛋白表达纯化

按照第二章所述方法进行菌株培养和蛋白诱导表达及镍柱纯化。四个 FDR 的蛋白胶见图 5-6。

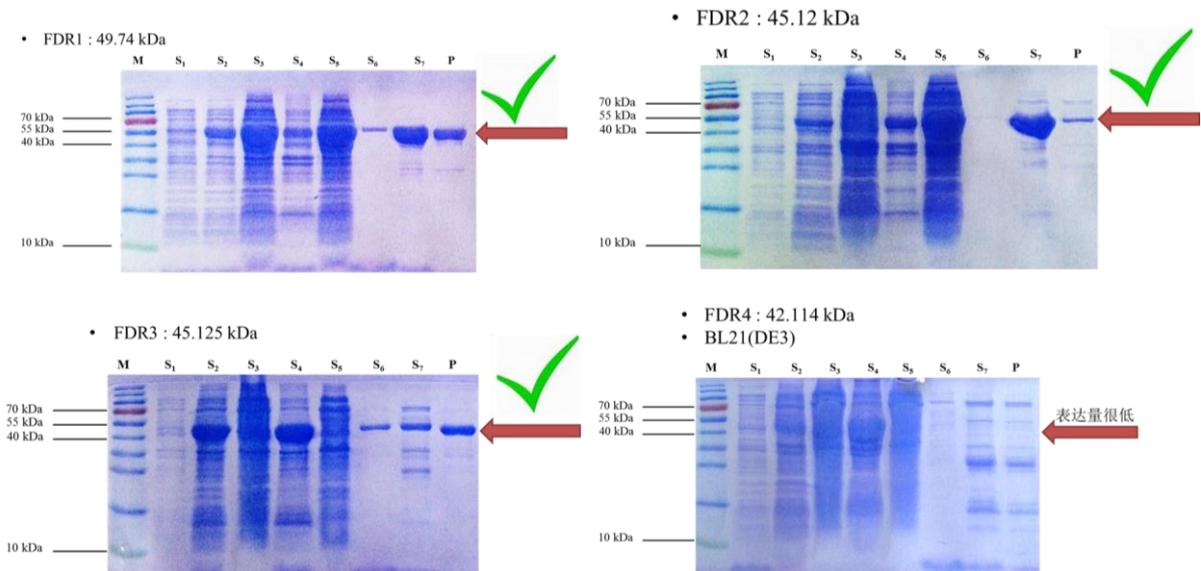


图 5-6 *E. coli* BL21(DE3)表达 4 个 FDR

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

以上结果显示, 在 BL21 (DE3) 菌株中, FDR1、FDR 2、FDR3 的表达量较好, 其中 FDR3 的表达量非常好, 蛋白溶液呈现明亮的荧光黄色。但是, FDR4 的表达量较低。本章在尝试采用具有稀有密码子的表达菌株 RIPL 后 (图 5-7), FDR4 的表达量仍然较低。

- FDR4 : 42.114 kDa
- Condon plus RIPL

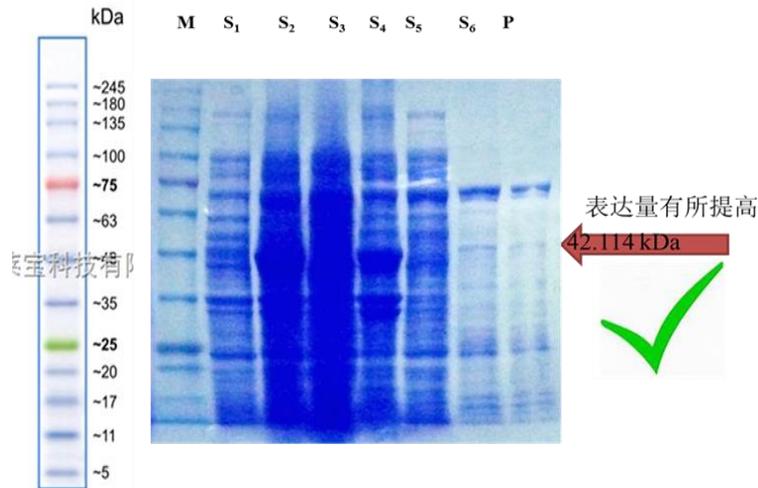


图 5-7 *codon plus RIPL* 表达 FDR4

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

最后, 本章将所有还原伴侣蛋白集中在一张胶图上。结果表明, FDR 蛋白表达尚可, Fdx5 表达较好, 其余 Fdx 在转化 RIPL 菌株后蛋白表达情况有所改善, 但 Fdx6 仍然表达量较差, 详见图 5-8。

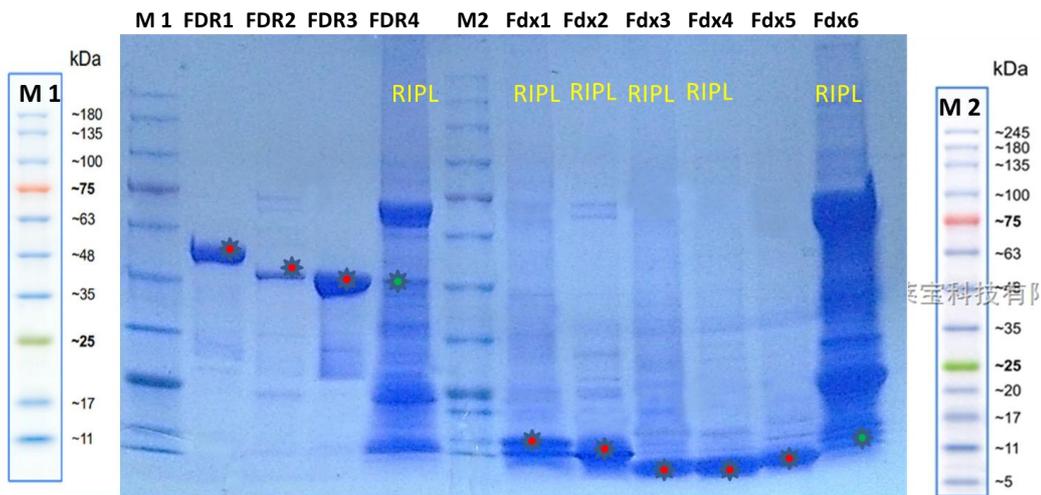


图 5-8 天蓝色链霉菌 FDR 和 Fdx 蛋白电泳
所有样品信息详见图内文字标注。

2.2.3 天蓝色链霉菌 18 个 P450 的表达

本章按照第二章的方法对 18 个 P450 进行了纯化, 首先尝试了 BL21(DE3)对蛋白进行表达纯化, 得到了若干表达较好的 P450 蛋白。其次, 对 BL21(DE3)表达不成功的 P450 重新转化 RIPL 菌株, 再次进行了蛋白表达筛选。

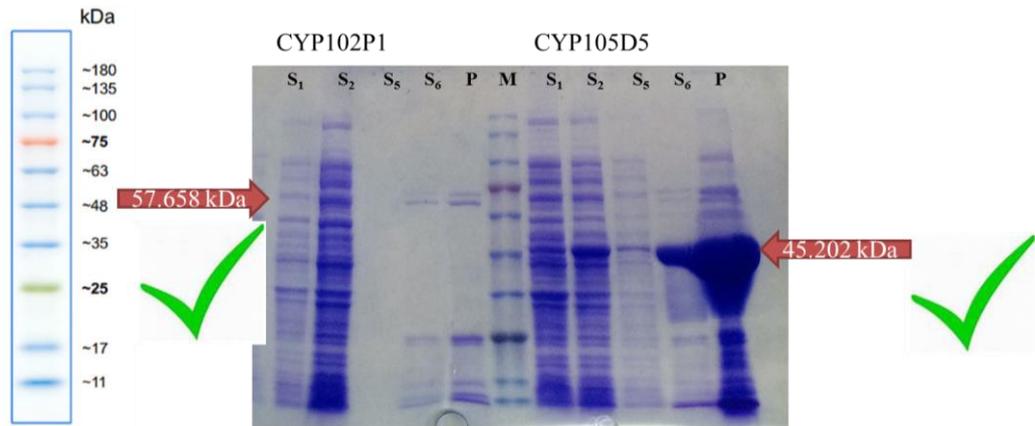


图 5-9 CYP102P1 和 CYP105D5 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

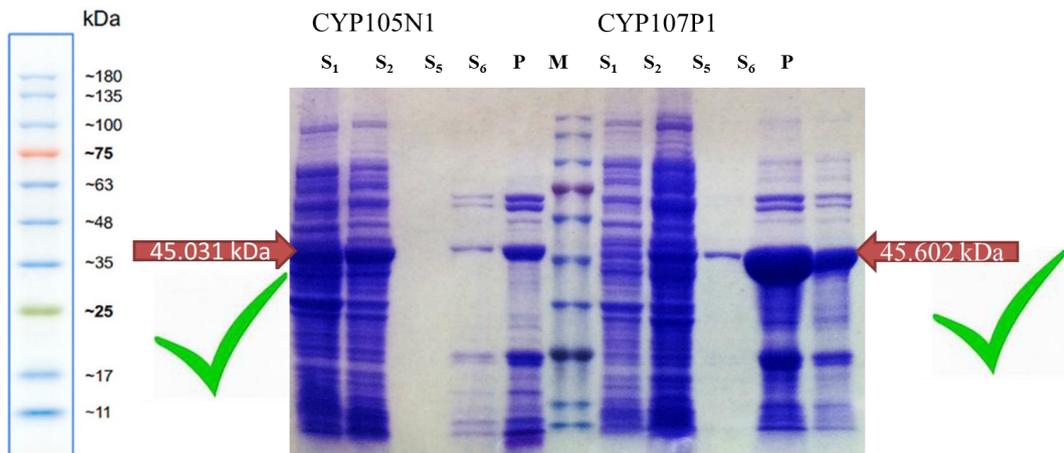


图 5-10 CYP105N1 和 CYP107P1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

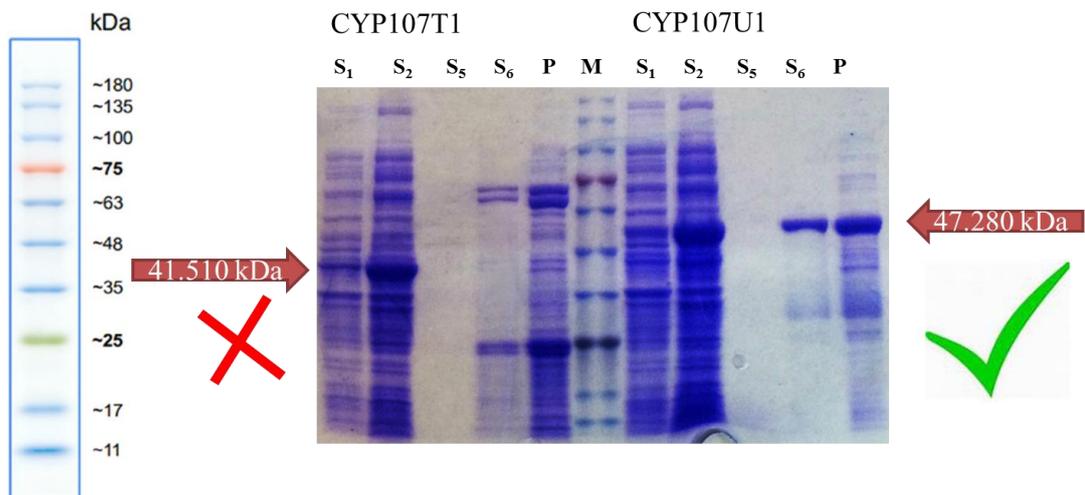


图 5-11 CYP107T1 和 CYP107U1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

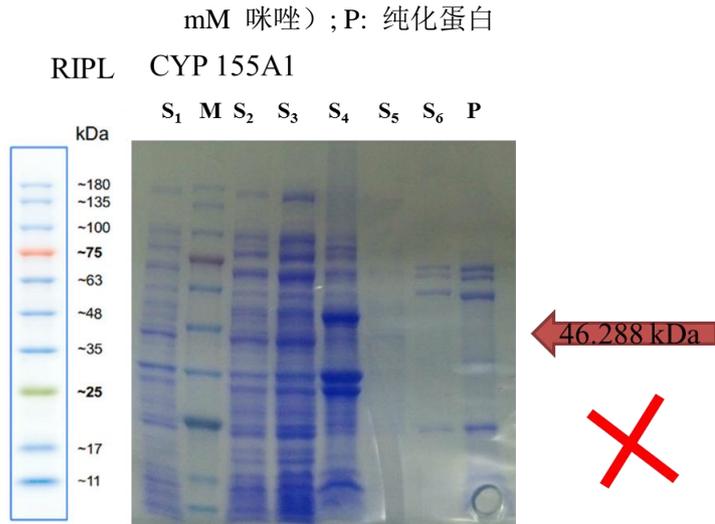


图 5-12 CYP155A1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

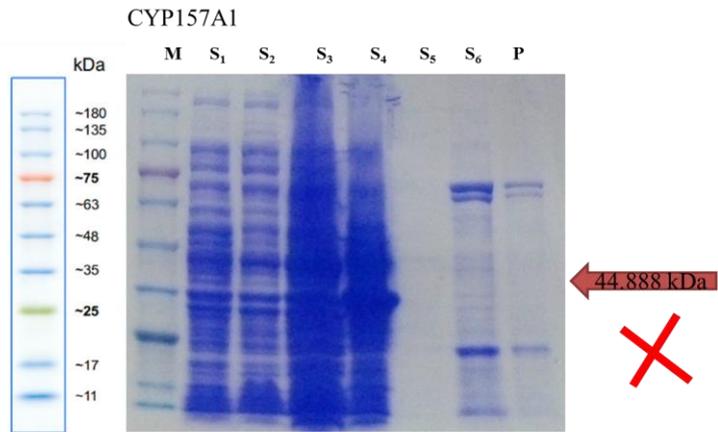


图 5-12 CYP157A1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

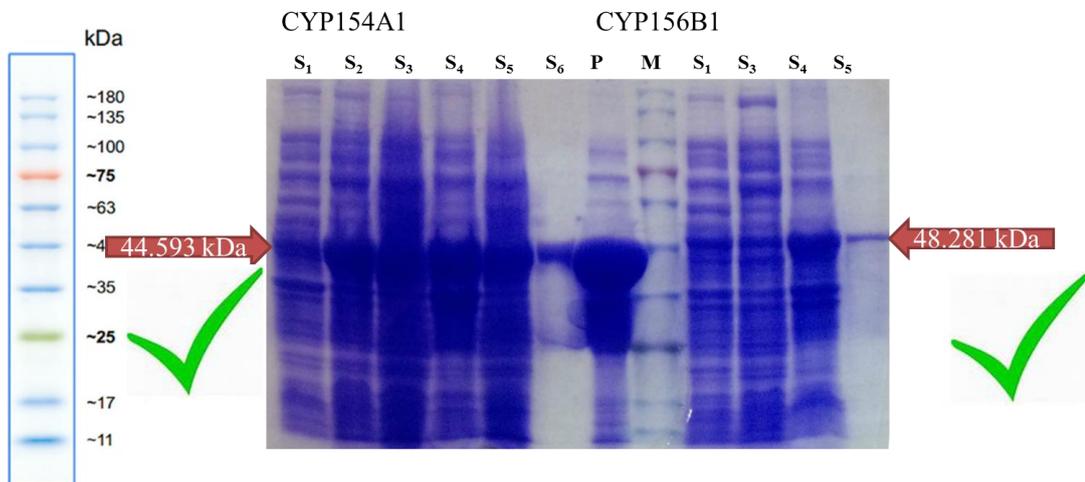


图 5-13 CYP154A1 和 CYP156B1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

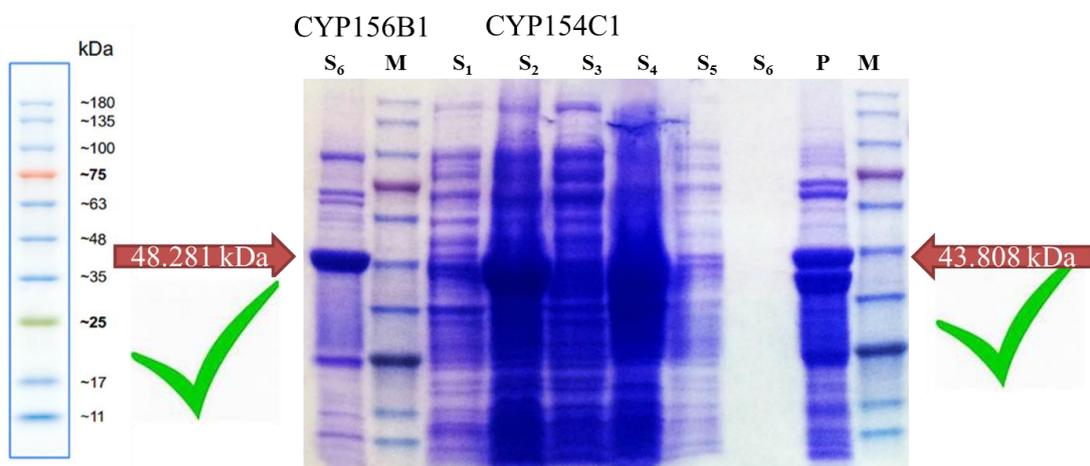


图 5-14 CYP156B1 和 CYP154C1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

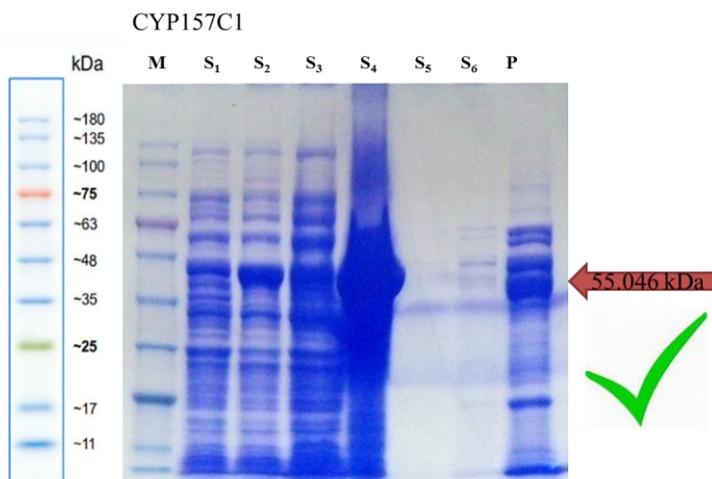


图 5-15 CYP157C1 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

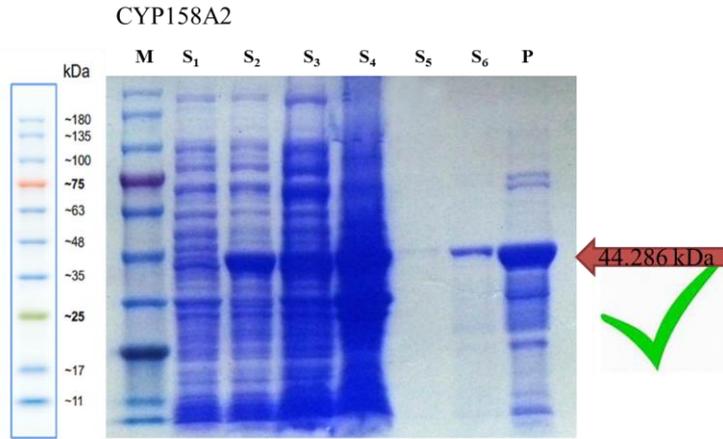


图 5-16 CYP158A2 的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

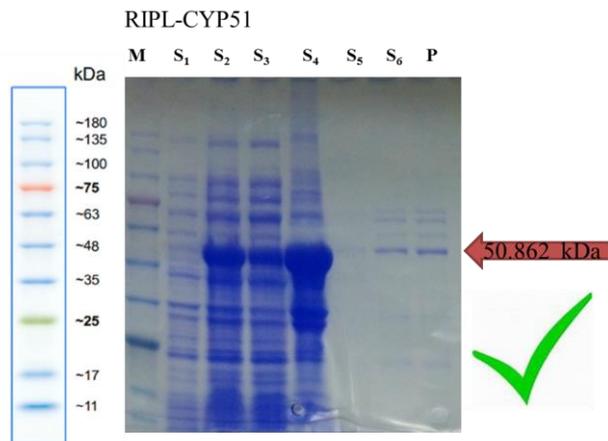


图 5-17 CYP51 在 RIPL 菌株中的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

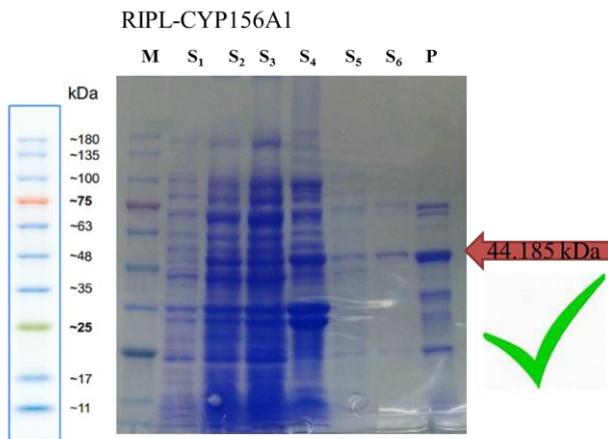


图 5-18 CYP156A1 在 RIPL 菌株中的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

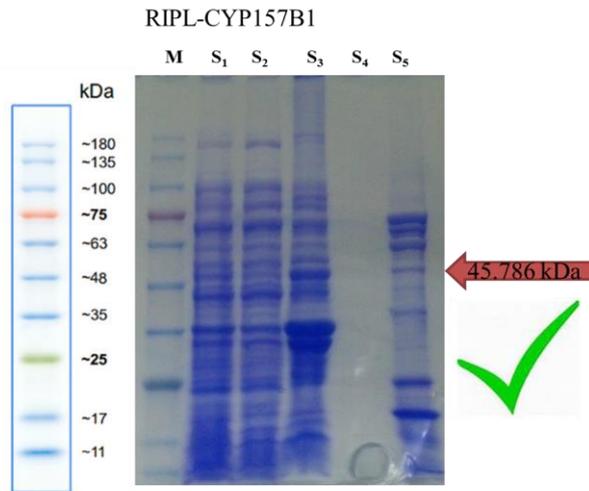


图 5-19 CYP157B1 在 RIPL 菌株中的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

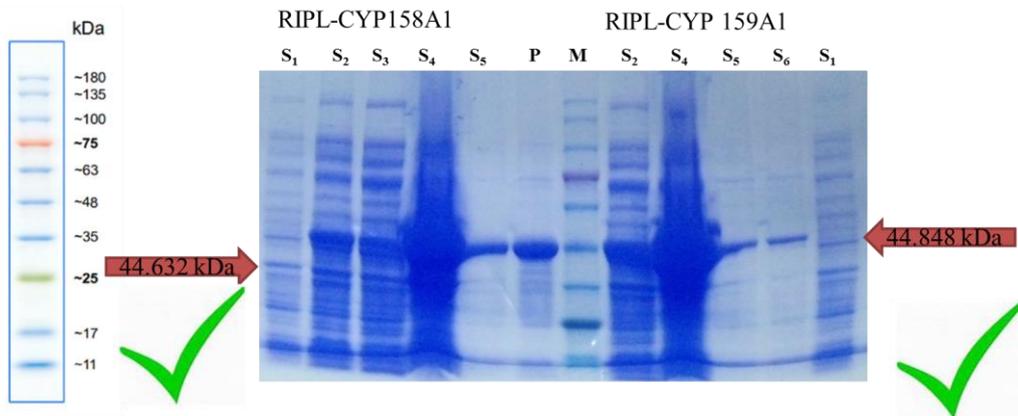


图 5-20 CYP158A1 和 CYP159A1 在 RIPL 菌株中的表达、纯化及 SDS-PAGE 分析

S1: 诱导前; S2: 诱导后; S3: 可溶蛋白; S4: 不可溶蛋白; S5: 清洗液 (20 mM 咪唑); S6: 洗脱液 (250 mM 咪唑); P: 纯化蛋白

以上结果表明, 未经密码子优化的来源于天然色链霉菌的 18 个 P450 酶在大肠杆菌中的表达情况有所不同。总体来讲, 表达情况较弱, 仅有 表达情况较好、表达量较大。其余表达量较低。部分 P450 在经过 RIPL 菌株优化表达后表达量得到提高, 仍有部分蛋白表达量很低 (图 5-21 和图 5-22)。

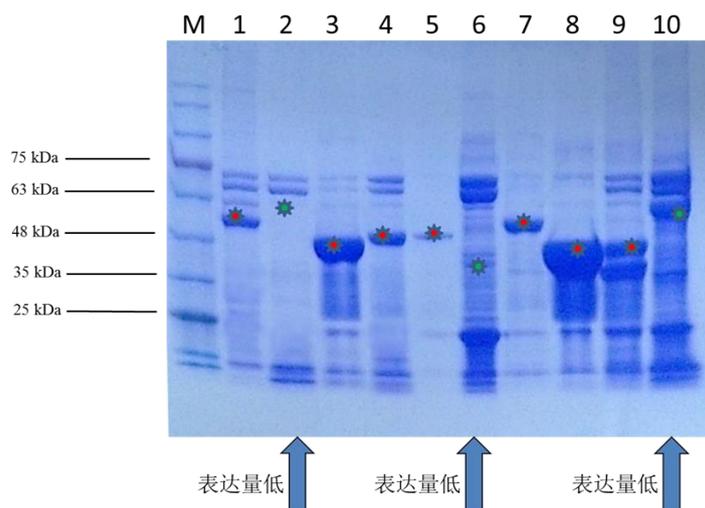


图 5-21 天蓝色链霉菌 18 个 P450 蛋白汇总 (I)

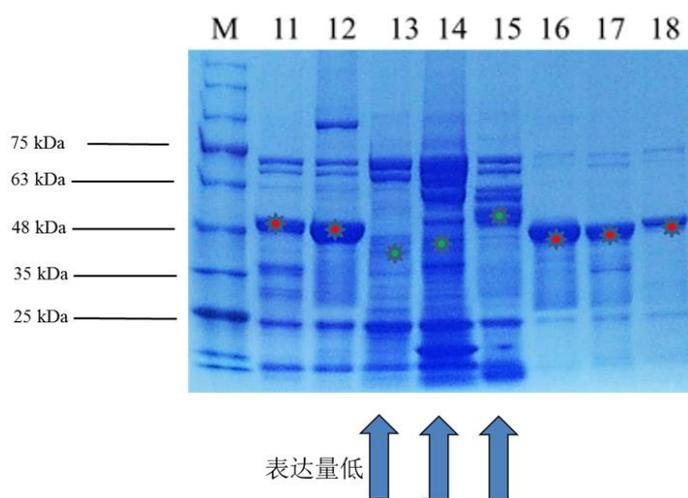


图 5-22 天蓝色链霉菌 18 个 P450 蛋白汇总 (II)

2.3 P450 和 FDR 的光谱吸收

2.3.1 P450 光谱吸收及浓度测定

按照第二章方法。图 5-23 为部分 P450 单加氧酶 CO 差光谱。

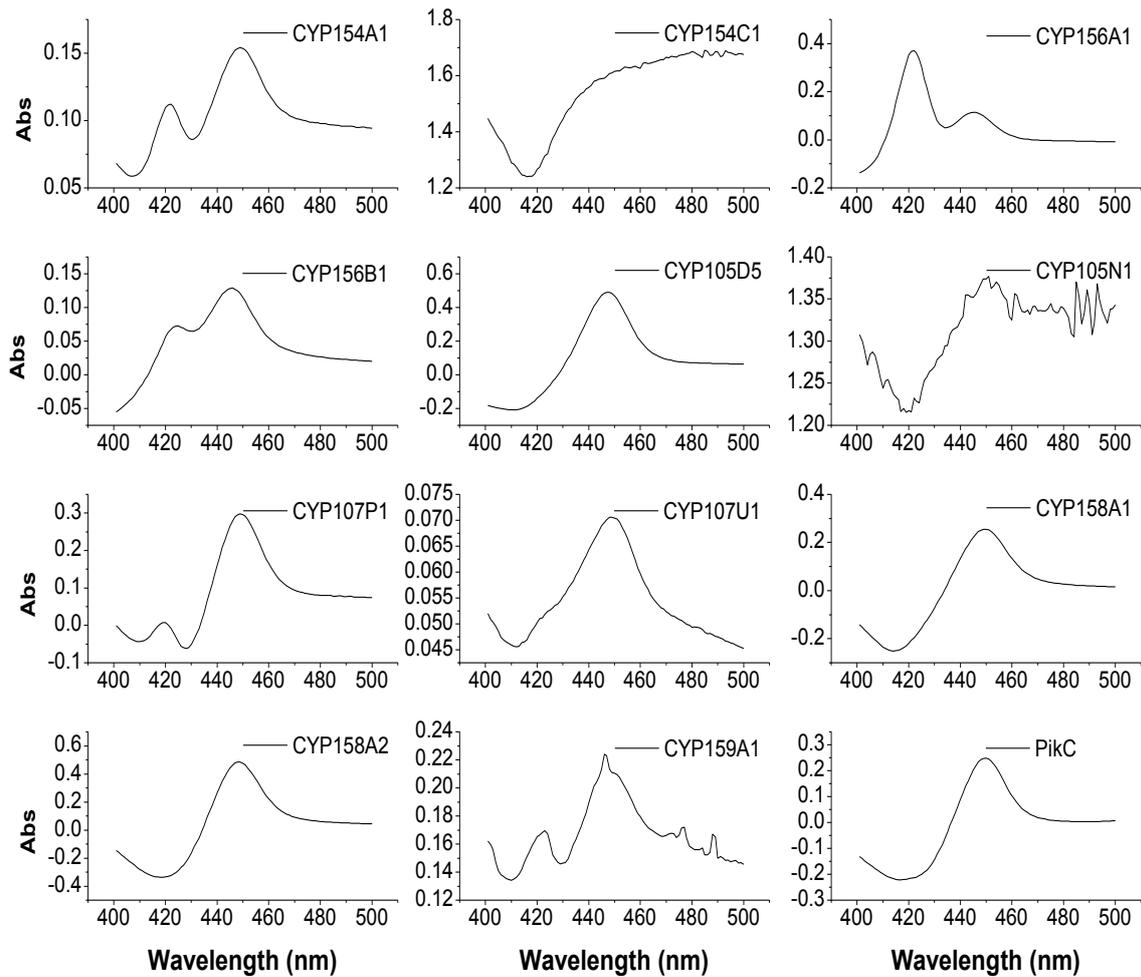


图 5-23 部分 P450 单加氧酶 CO 差光谱

PiKc 为对照蛋白。其余 P450 蛋白尚未进行光谱扫描。

2.3.2 4 个 FDR 的光谱吸收

本实验在 300-700nm 之间对 FDR 蛋白进行光谱扫描,得到的光谱吸收情况如图 5-4 所示。结果表明, FDR1 光吸收较低;其余 FDR 显示正确吸收光谱。此结果显示,虽然 FDR4 的表达量较低,但活性较好。FDR1 则光谱吸收较弱。

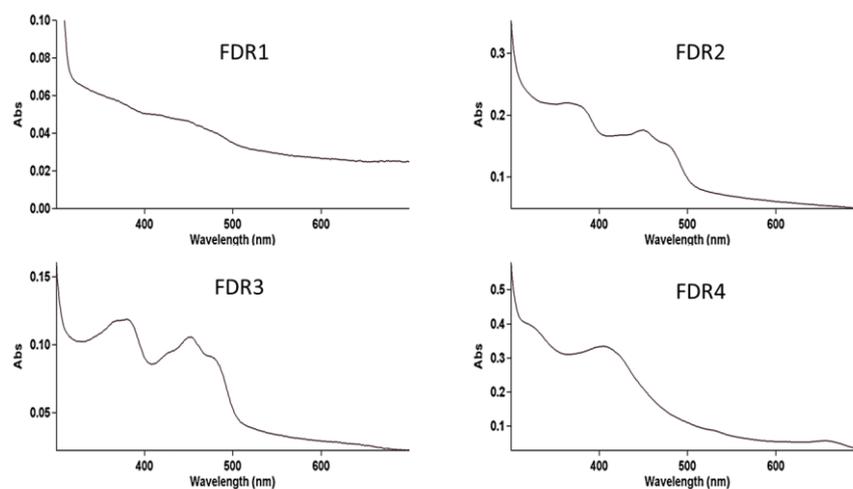


图 5-24 FDR 蛋白的紫外吸收光谱

3 本章小结

本章通过在大肠杆菌中表达来自天蓝色链霉菌的 18 个 P450 单加氧酶和 6 个 Fdx、4 个 FDR。成功构建了表达质粒，并获得了大肠杆菌突变株。本章对目的蛋白的表达纯化和 SDS-PAGE 进行了分析，同时对部分表达较差的蛋白转化 RIPL 菌株进行优化，对 Fdx6 进行了密码子优化，提高了表达量。为本实验室提供了多种还原伴侣和 P450 单加氧酶提供了平台。

第六章 天蓝色链霉菌还原伴侣和 P450 应用筛选

本章按照图 6-1 所示的策略对第五章纯化的 P450 酶、FDR 和 Fdx 进行了酶反应。

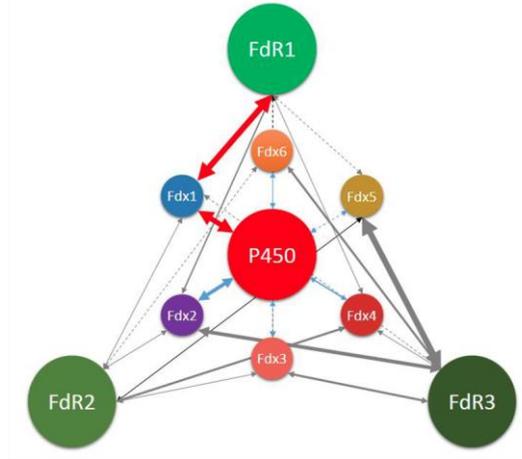


图 6-1 本章策略

1 材料与方法

1.1 菌株的培养

参照第二章方法。

1.2 菌株的构建

所构建表达质粒分别转化大肠杆菌 BL21(DE3)和 codon plus RIPL 菌株。BL21(DE3)用卡那霉素筛选转化子; codon plus RIPL 菌株用氯霉素、壮观霉素和卡那霉素筛选转化子,得到突变株(附录 3)。

1.3 基因型鉴定

参照第二章方法。

1.4 DNA 的重组

参照第二章方法。

1.5 聚合酶链式反应 PCR

参照第二章方法。

2 实验结果与分析

2.1 构建的质粒和菌株

本章构建的菌株和质粒信息见附表。

2.2 酶反应底物的选择

本章采用所纯化的 P450 和还原伴侣对四组底物进行催化筛选。第一组为黄体酮、睾酮；第二组为丹参酮 IIV 和薯蓣皂苷元；第三组为麦新米星的前体 M-IV；第四组为 YC17。这些化合物的结构见图 6-2 所示。

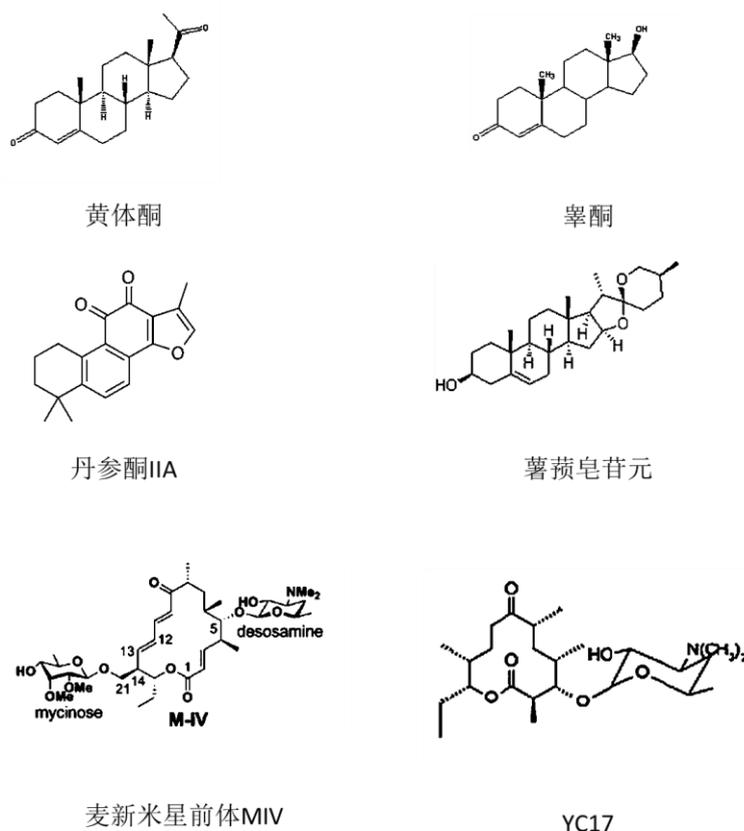


图 6-2 本章采用的酶反应底物

本章首先进行了混合酶的初步催化筛选，即采用天蓝色链霉菌的 P450 混合物和还原伴侣混合物对以上底物进行催化筛选，利用 HPLC 分析产物。此种方法简易但缺乏底物与 P450 之间的定向匹配，如利用软件进行分子模拟酶与底物之间的拟合等，这也是实验室目前的技术短板。

初步结果表明，天蓝色链霉菌的 P450 混合物和还原伴侣混合物对以上底物均不能催化。因此，本章在基础上设计了验证实验，以 YC17 为例，采用 YC17 为底物，采用原始 P450 酶 PickC 进行酶催化（图 6-3）；在选择还原伴侣时，则采用了来自蓝细菌的 Fdx 和 FNR 与天蓝色链霉菌的还原伴侣进行了交叉组合（表 6-1）。

2.3 HPLC 结果与分析

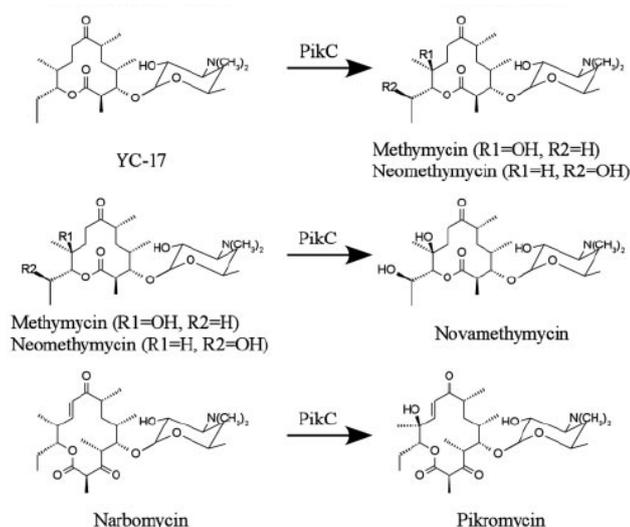


图 6-3 PikC 催化 Pikromycin 合成的步骤

采用的酶反应体系详见表 6-1。

表 6-1 采用 PikC 和不同还原伴侣组合进行催化 YC17 底物 (200 μ l 体系)

No	YC17 (20mM)	PikC	Glucose (1M)	NADPH (100mM)	GD H	S.c Fdx	S.c FDR	7942 -Fd	7942 -	other FNR
1	4	30	2	2	1		FDR1;5 μ l	5		to
2	4	30	2	2	1		FDR2;5 μ l	5		200u
3	4	30	2	2	1		FDR3; 5 μ l	5		1
4	4	30	2	2	1		FDR4; 5 μ l	5		
5	4	30	2	2	1		FDRmix;5*4 μ l	5		
6	4	30	2	2	1	Fdx1; 5 μ l				5
7	4	30	2	2	1	Fdx2; 5 μ l				5
8	4	30	2	2	1	Fdx3; 5 μ l				5
9	4	30	2	2	1	Fdx4; 5 μ l				5
10	4	30	2	2	1	Fdx5; 5 μ l				5
11	4	30	2	2	1	Fdx6; 5 μ l				5
12	4	30	2	2	1	Fdxmix;5*6 μ l				5
CT	4		2	2	2					

30 $^{\circ}$ 过夜 20h。优化后的反应条件为 30 $^{\circ}$ 过夜 20h 后，之后再 37 度反应 20min。加入等体积甲醇终止反应，1300rpm 离心 10min 取上清，HPLC 上样 100 μ l。

首先，本章采用 YC17+天蓝色链霉菌总 P450 + 6 Fdx-mix + 4 FDR-mix 进行反应时并

未合成产物。采用天蓝色链霉菌总 P450 和蓝藻还原伴侣 (Fd 和 FNR) 作为进行反应时并未合成产物 (图 6-4)。

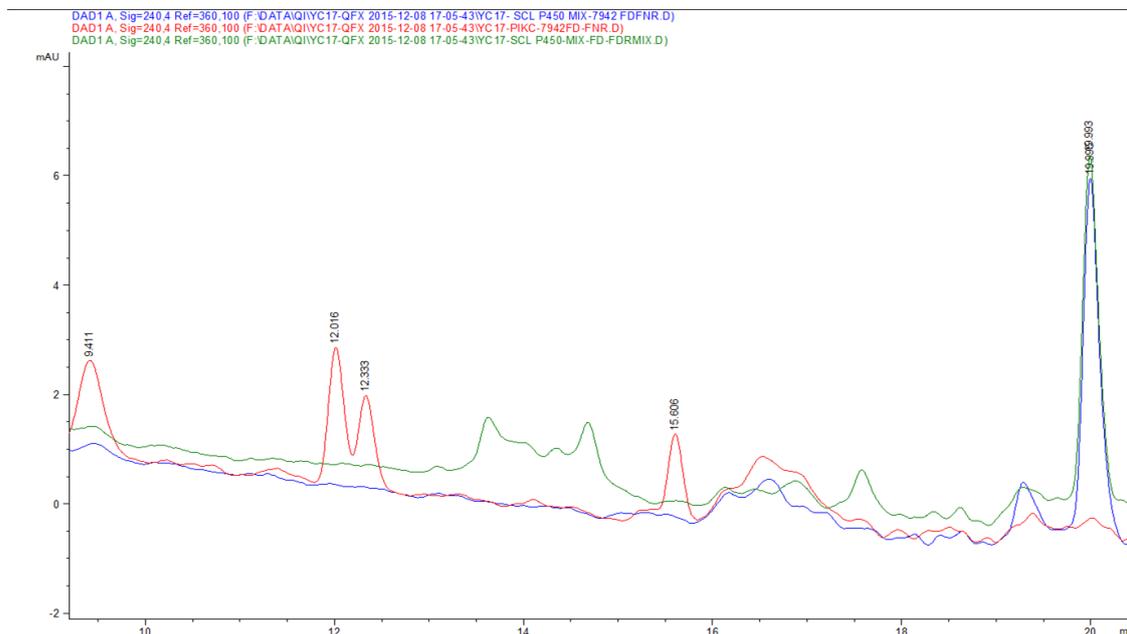


图 6-4 YC17+天蓝色链霉菌总 P450 + 6 Fdx-mix + 4 FDR-mix

红色为阳性对照；蓝色为天蓝色链霉菌总 P450 和蓝藻还原伴侣 (Fd 和 FNR)；绿色为 YC17+天蓝色链霉菌总 P450 + 6 Fdx-mix + 4 FDR-mix 进行酶反应。

本章还同时设计了若干反应，分析天蓝色链霉菌单个 P450 酶对 YC17 的催化反应，结果显示；YC17+天蓝色链霉菌单个 P450 +天蓝色链霉菌混合还原伴侣均无产物；只有 PikC 对照+蓝藻 FNR+Fd (图 6-5)。

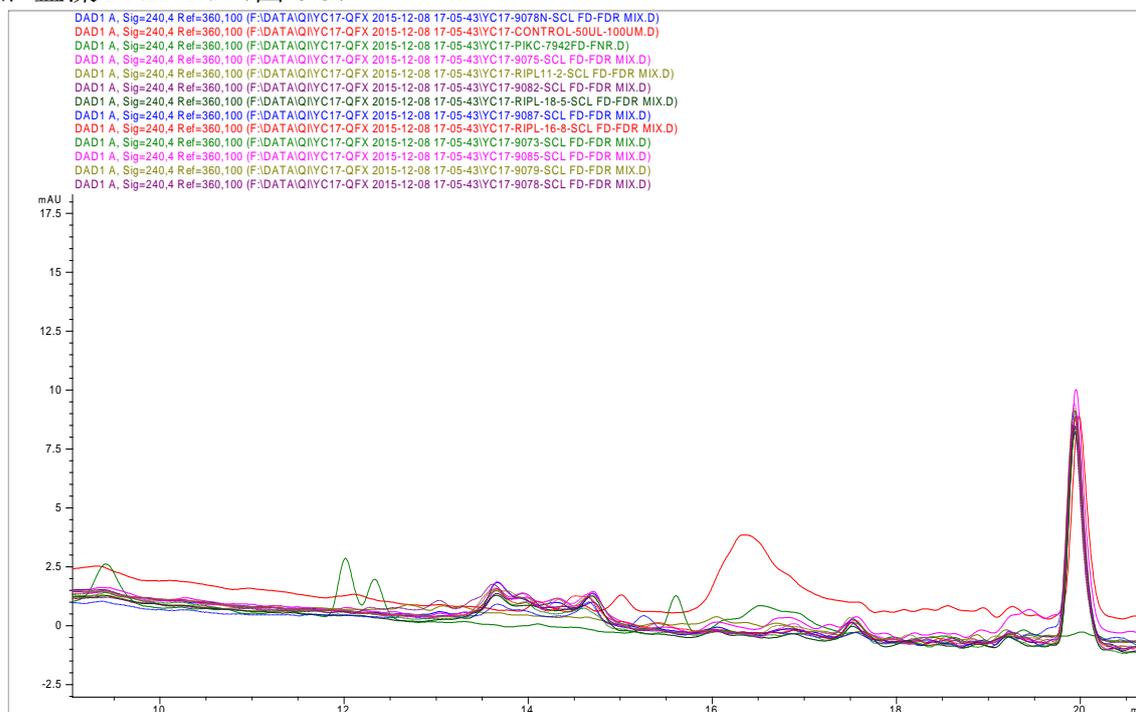


图 6-5 天蓝色链霉菌单个 P450 酶对 YC17 的催化反应

YC17+天蓝色链霉菌单个 P450 +天蓝色链霉菌混合还原伴侣均无产物；只有 YC17+PikC+蓝藻

FNR+Fd 可反应（绿色）。

图 6-5 和图 6-6 结果表明，本章所纯化的天蓝色链霉菌中的所有 P450 组合中没有催化 YC17 的功能酶。其中 CYP154C1 会在以后小节中单独分析。图 6-6 进行的反应是 CYP154C1 催化 YC17 的 HPLC 图谱，图谱显示并未有产物合成，此结果与已有报道不符合。其原因可能是 CYP154C1 酶无活性所致，因此本章对 CYP154C1 进行了光谱反应（图 6-7）发现 CYP154C1 在和 CO 结合后在 450nm 处并无特异吸收峰，即此酶没有活性。

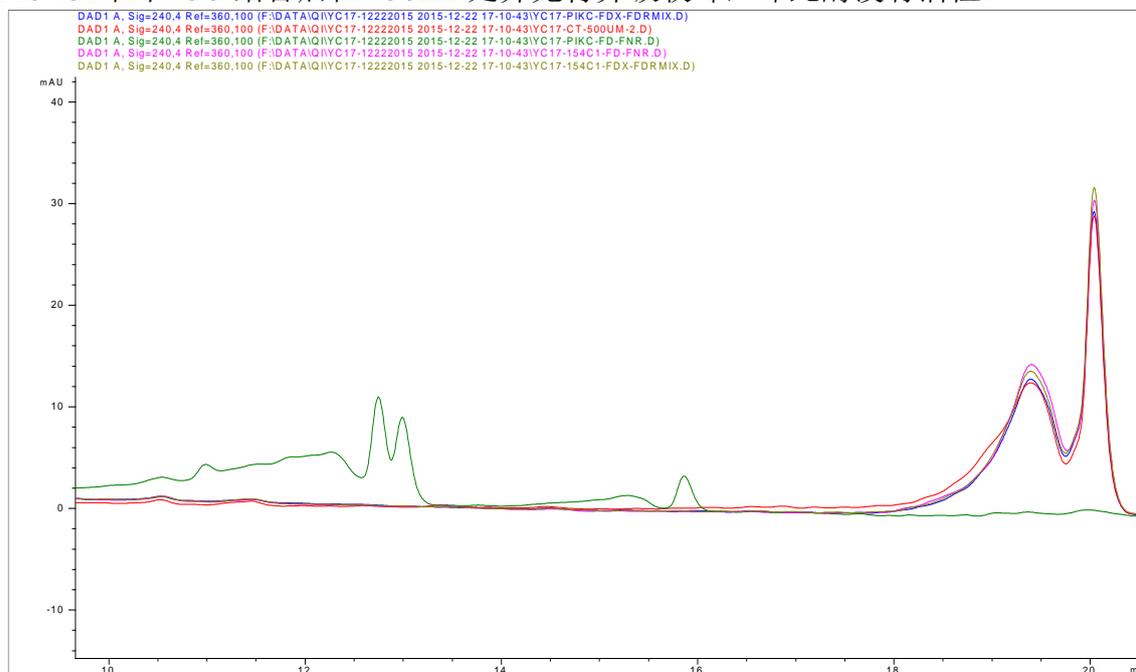


图 6-6 图 6-5 CYP154C1 对 YC17 的催化反应

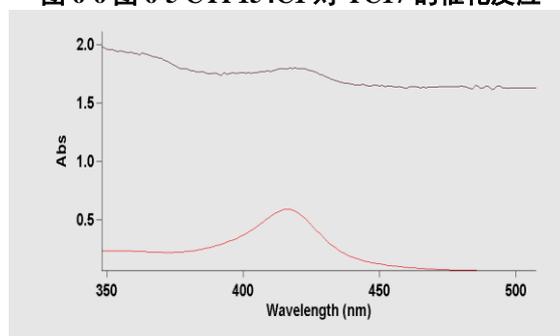


图 6-7 CYP154C1 的 CO 差异光谱吸收

FDR3 是天蓝色链霉菌中蛋白表达最强的 FDR。本节偶然发现作为对照组的 PikC+天蓝色链霉菌 FDR3+7942 Fdx1 可催化 YC17 合成产物，但只能见到到中间两个产物（图 6-8）。此结果首次表明，天蓝色链霉菌的还原伴侣蛋白也可以催化支撑的反应。此结果也揭示了，在支撑 PikC 的效力层面，通过采用不同组合（不同性质或不同来源）的还原伴侣，可以达到不同的催化效率，并合成不同的目的产物。其背后的机理，对于理解 P450 与还原伴侣之间的分子反应机制具有重要的作用。

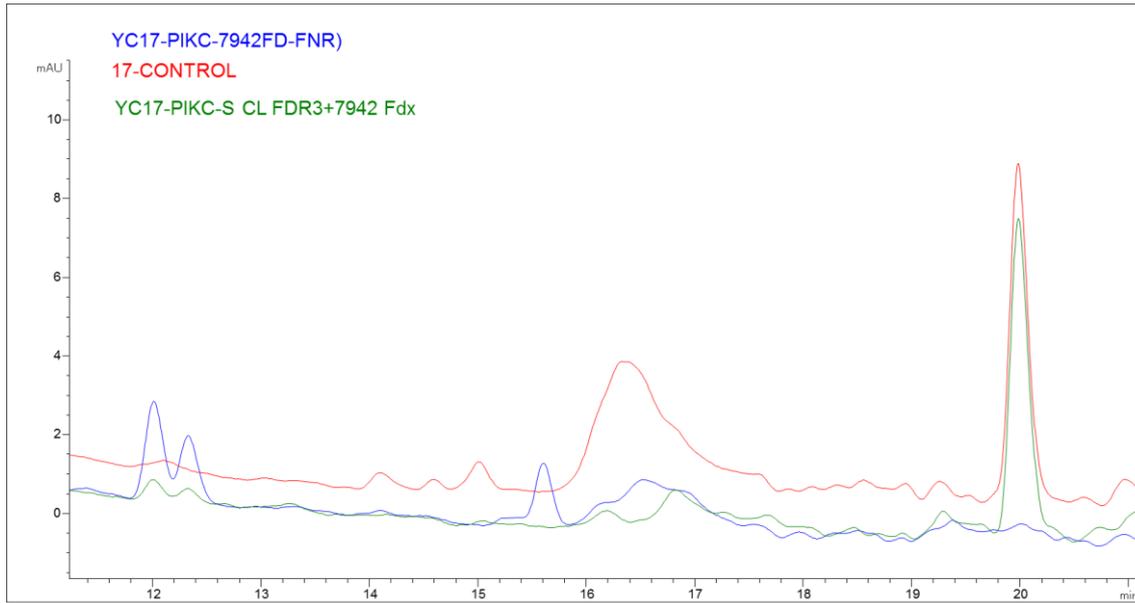


图 6-8 PikC+天蓝 FDR3+7942 Fdx1 可催化 YC17-FDR3 是天蓝色链霉菌中蛋白表达最强的 FDR;

本章在之后的实验中进一步发现证实，不同的还原伴侣可以催化 YC17 产生不同的产物类型。首先，本章采用蓝细菌聚球藻 PCC7942 的 Fdx1 和 FNR 催化 YC17 可以产生 4 个产物。本章继续进行了图 6-9 中的反应，发现采用天蓝色链霉菌四个 FDR 混合物与蓝细菌聚球藻 PCC7942 的 Fdx1 混搭可以支撑 PikC 催化 YC17 产生 2 个产物，但是产物量较低。相反，采用聚球藻 PCC7942 的 FNR 与天蓝色链霉菌的 6 个 Fdx 混合酶进行催化是则可以产生更高的产物，但同样，仍然只能催化产生其中两个产物。采用 PikC 与天蓝色链霉菌的 Fdx 和 FDR 则无法催化 YC17 产生产物。本实验经过重复实验确认。

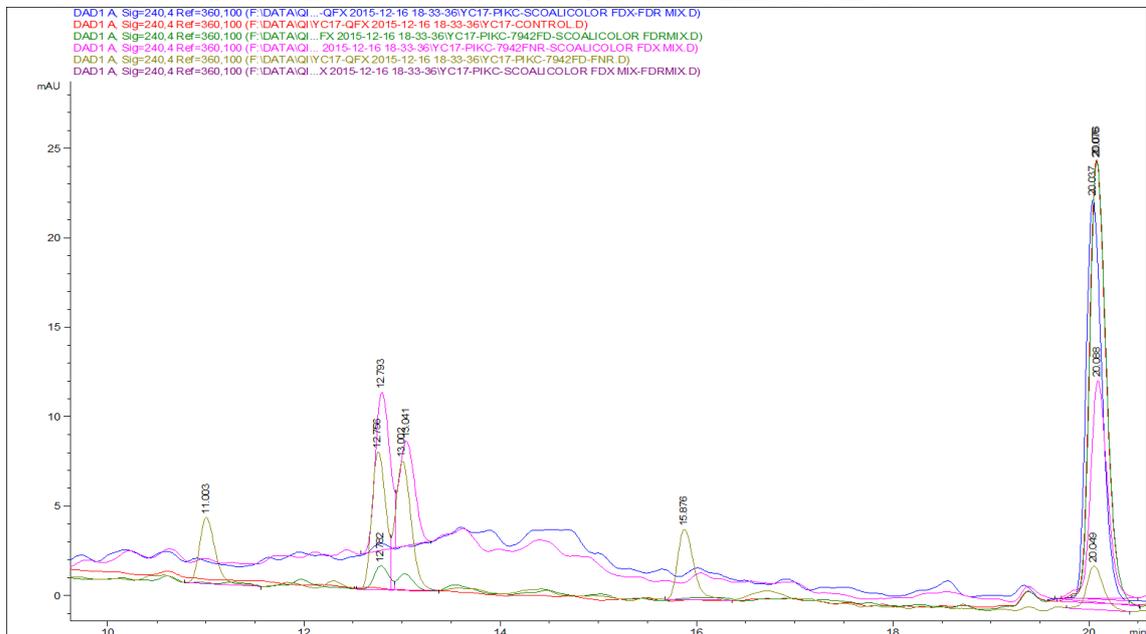


图 6-9 Fdx 和 FDR 的混搭

天蓝 FDR mix-7942 Fdx 可以产生 2 个产物；7942 FNR +天蓝 Fdx mix，产物量较高，但仍然只有 2 个产物。PikC 与天蓝色链霉菌的 Fdx 和 FDR 则无法催化 YC17 产生产物

在接下来的实验中，本章对单个天蓝色还原伴侣与聚球藻 PCC7942 的还原伴侣进行了

交叉组合分析 (图 6-10)。为了提高产物量, 本节特将反应条件进行了优化, 反应条件为 30°C 过夜 (大约 20h), 之后在 37°C 反应 30min 后终止反应。

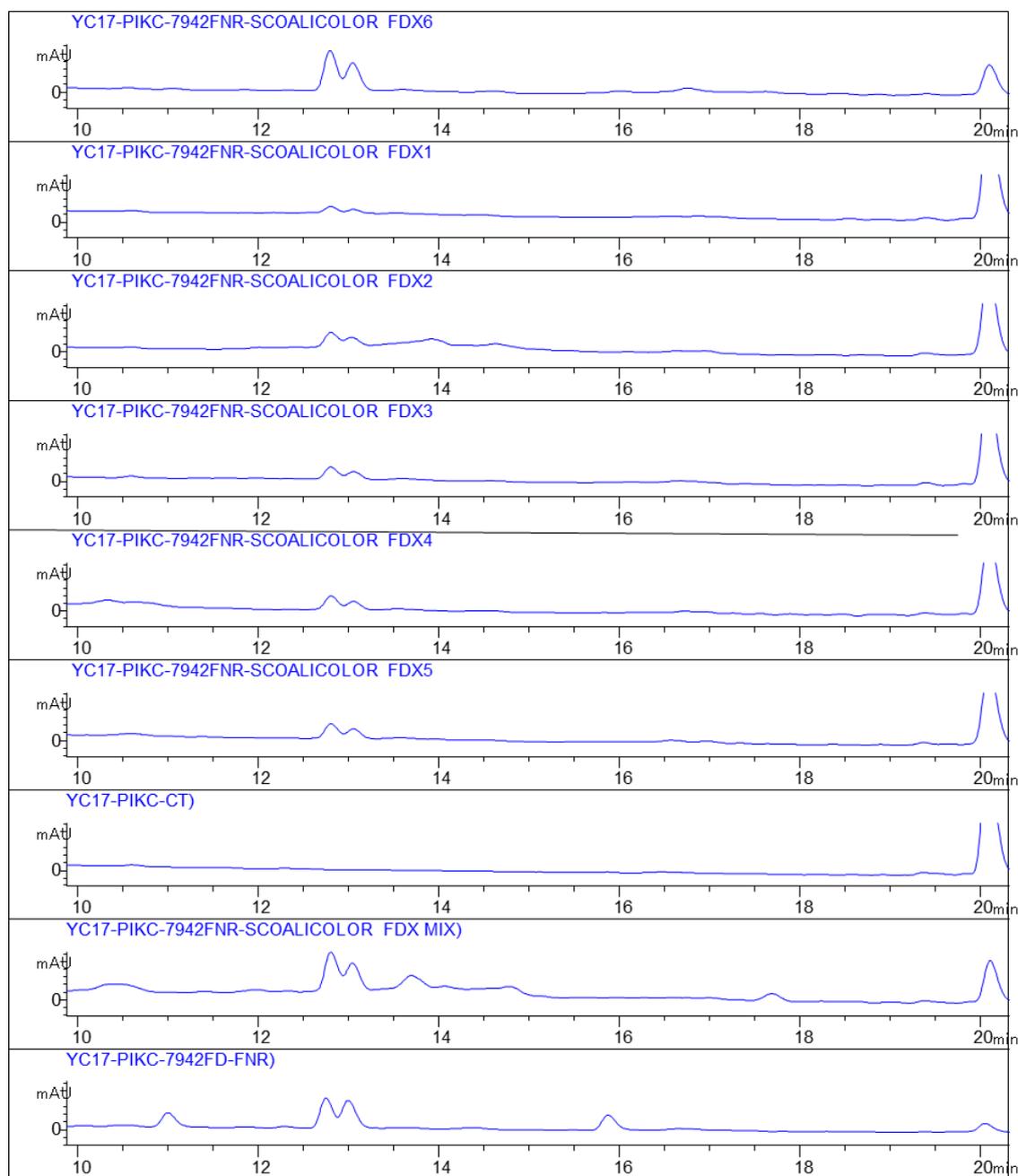


图 6-10 *PikC* 与蓝藻 FNR+单个天蓝 Fdx 互配均能催化 YC17

反应条件为 30°C 过夜 (then 37°C, 30min)

图 6-10 的结果表明, 天蓝色链霉菌单个 Fdx 和聚球藻 PCC7942 的 FNR 组合均可以催化反应, 但产物的量有所不同。同时, 6 个 Fdx 混合物与 FNR 反应所得的产物量与单个 Fdx 分别催化所得的产物量相当。

图 6-11 的结果表明, 天蓝色链霉菌单个 FDR 和聚球藻 PCC7942 的 Fdx1 组合均可以催化反应, 同样产物的量有所不同, FDR2 和 FDR3 的效率更高。然而, 6 个 Fdx 混合物与 FNR 反应所得的产物量与单个 Fdx 分别催化所得的产物量并不相当。结果显示, 混合后的

天蓝色链霉菌 FDR 比单个 FDR 产物高出许多，混合 FDR 可以高效的催化底物 YC17，其效率与蓝藻 Fd 和 FNR 组合相当。这里似乎可以揭示天蓝色链霉菌 FDR 在催化 P450 单加氧酶的反应时，相互之间存在某种协同和促进作用。

图 6-10 和图 6-11 比较可以发现，选用蓝细菌 FNR 匹配天蓝色链霉菌的 Fdx 获得了更多的产物。无论哪一种组合，只有蓝细菌 FNR 和 Fdx 的组合才能催化产生四种产物。更换其中一种还原伴侣后，只能积累两种产物（产量很高），但无法获得其余两种产物。图 6-10 表明，延长反应时间后，7942Fd+天蓝 FDRmix 的转化率显著提高，但产物仍以 2 个产物为主。蓝藻 FNR 在其余两个产物生成过程中起到了未知作用。图 6-11 则表明，蓝藻 Fdx 在其余两个产物生成过程中起到了未知作用。因此，PikC 在 YC17 催化中的限速步骤存在于后两个产物合成步骤。此结果揭示，蓝细菌聚球藻 PCC7942 的 Fdx1 和 FNR 在催化 PikC 反应时具有特别的机制。其具体作用原理尚不明确。

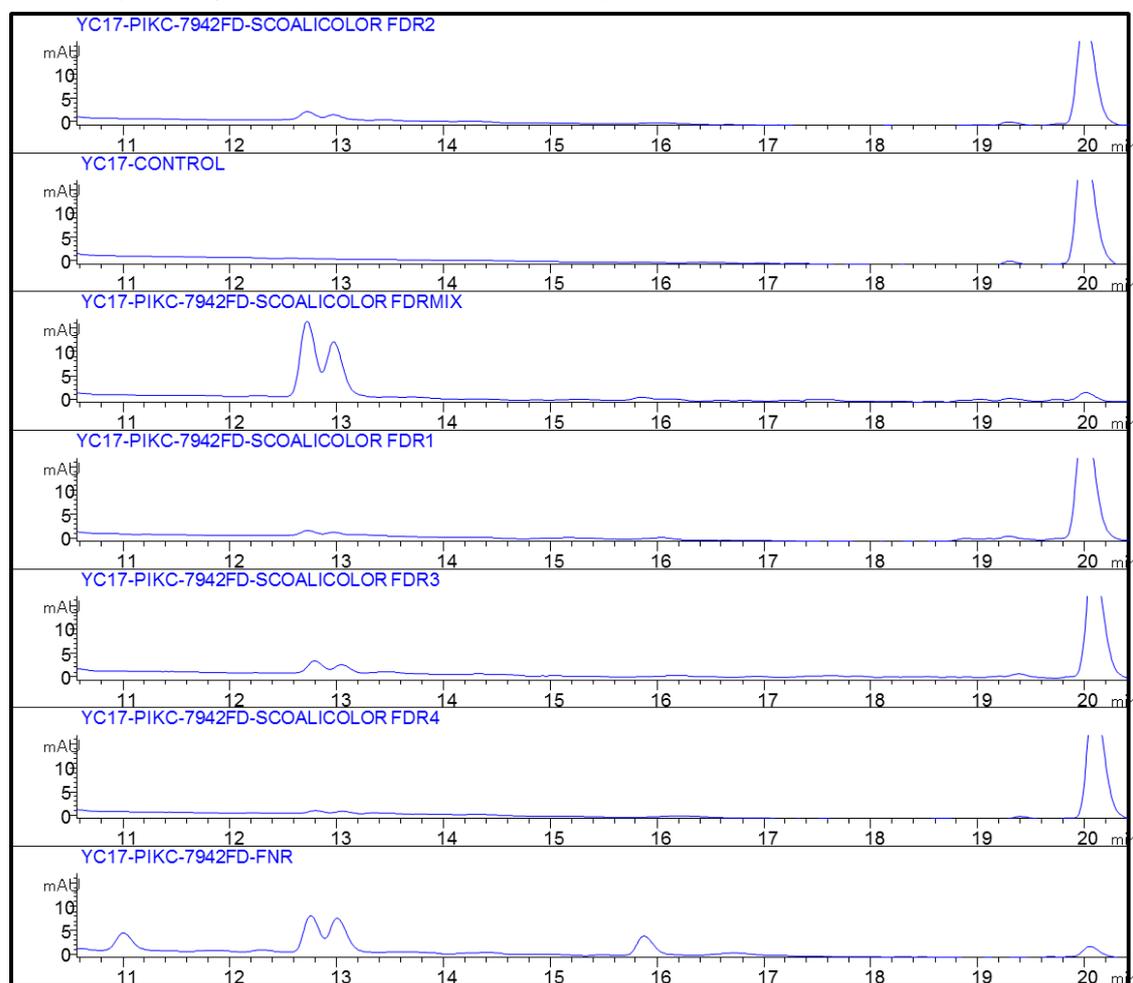


图 6-11 天蓝色链霉菌单个 FDR 和聚球藻 PCC7942 的 Fdx1
反应条件为 30°C 过夜 (then 37°C, 30min)

3 本章小结

本章主要完成了天蓝色链霉菌中 6 个 Fdx、4 个 FDR 和 18 个 P450 酶的蛋白表达优化

与鉴定。在尝试多种底物后，发现了天蓝色链霉菌还原伴侣支撑 **PikC** 催化 **YC17** 底物进行反应。具体结果汇总见表 6-1。本章仍有很多奥秘等待解释。

此结果首次表明，天蓝色链霉菌的还原伴侣蛋白催化支撑 **PikC** 催化 **YC17** 的反应。并且在支撑 **PikC** 的效力层面，通过采用不同组合（不同性质或不同来源）的还原伴侣，可以达到不同的催化效率，并合成不同的目的产物。同时，**PikC** 在 **YC17** 催化中的限速步骤存在于后两个产物合成步骤。蓝细菌聚球藻 **PCC7942** 的 **Fdx1** 和 **FNR** 在催化 **PikC** 反应时具有特别的作用机制。其具体作用机理，对于进一步理解 **P450** 与还原伴侣之间的分子反应机制具有重要的作用。

表 6-2 本章实验及结果汇总

Summary: gene cloning, expression and purification of the P450 and 10 redox partner proteins from <i>Streptomyces coelicolor</i> A3(2) (by December 2015)											
No	Genes (drnelson.utm.edu/CytochromeP450.html)	full name	DNA accession(Genbank Sanger center)	Protein_ID	Nt	Amino acids	MW(Da)	BL21 (DE3)	Codon plus RIPL	Enzyme Activity	UV spectra/CO difference spectrum
	Fdx1	ferredoxin	SC07110 SC07110SC4B10.11	NP_631171.1	390	129	14384	none	good epression	Yes	
	Fdx2	ferredoxin	SC05135 SC05135SC9E12.20	NP_629284.1	321	106	11816	none	good	Yes	
	Fdx3	ferredoxin	SC07676 SC07676SC4C2.11	NP_631715.1	222	73	7404	none	good	Yes	
	Fdx4	ferredoxin	SC00773 SC007733SCF60.05c	NP_625075.1	201	66	6666	none	good	Yes	
	Fdx5	ferredoxin	SC03867 SC03867SCH18.04c	NP_628054.1	213	70	7251	very good	strain freezed	Yes	
	Fdx6	ferredoxin	SC01649 SC01649SCI41.32	NP_625924.1	306	101	10780	None	not pure on PAGE	unknown	
	Fdx6-co	ferredoxin	chemically synthesized with codon optimization (co)	NP_625924.1	306	101	10780	very good	strain freezed	Yes	
	FDR1	ferredoxinferredoxin-NADP reductase	SC00681 SC00681SCF15.02	NP_624989.1	1365	454	49740	no color but very pure	strain freezed	Yes	correct
	FDR2	ferredoxin reductase	SC07117 SC07117SC4B10.18c	NP_631178.1	1263	420	45126	no color but very pure	strain freezed	Yes	correct
	FDR3	reductase	SC02469 SCTA8.08c	NP_626711.1	1266	421	45125	very good	strain freezed	Yes	correct
	FDR4	oxidoreductase	SC02106 SC2C1A.02c	NP_626364.1	1224	407	42114	very low	better expression, light yellow color, correct UV spectra	Yes	correct
1	CYP51	cytochrome P450	SC05223 SC7E4.20	NP_629370	1386	461	50962	None	better	unknown	
2	CYP102B1	cytochrome P450	SC00801 SCF43.12	NP_625102	1584	527	57658	Very low	strain freezed	unknown	
3	CYP105D5	cytochrome P450	SC00774 3SCF60.06c	NP_625076	1239	412	45202	OK	strain freezed	unknown	Yes
4	CYP105N1	cytochrome P450	SC07686 SC4C2.21	NP_631725	1236	411	45031	Very low	strain freezed	unknown	Yes
5	CYP107P1	cytochrome P-450 hydroxylase	SC03636 SCH10.14c	NP_627830	1236	411	45602	Good	strain freezed	unknown	Yes
6	CYP107T1	cytochrome P450 oxidoreductase	SC03770 SCH63.17	NP_627960.1	1185	394	41510	Very low	strain freezed	unknown	Yes
7	CYP107U1	cytochrome P450 hydroxylase	SC03099 SCE41.08c	NP_627317.1	1302	433	47280	Very low	strain freezed	unknown	Yes
8	CYP154A1	cytochrome P450	SC02884 SCE6.21	NP_627112	1227	408	44593	weak	strain freezed	unknown	Yes
9	CYP154C1	cytochrome P450	SC07417 SC6D11.13c	NP_631466	1224	407	43808	weak	strain freezed	No	No
10	CYP155A1	cytochrome P450 (fragment), partial	SC07444 SC6D11.40c	NP_733733	1272	423	46288	None	still low expression	unknown	
11	CYP156A1	cytochrome P450	SC02883 SCE6.20	NP_627111	1233	410	44185	None	much better, but not pure	unknown	Yes
12	CYP156B1	cytochrome P450	SC06310 SCIF3.12	NP_630407	1344	447	48281	OK	strain freezed	unknown	Yes
13	CYP157A1	cytochrome P450-family protein	SC07418 SC6D11.14c	NP_631467	1242	413	44888	low expression, most are inclusion bodies, hard to see on PAGE	strain freezed	unknown	
14	CYP157B1	cytochrome P450	SC00584 SCF55.08c	NP_624896	1263	420	45786	light yellow, very pure band on PAGE	light band on PAGE, not pure	unknown	
15	CYP157C1	cytochrome P450	SC01626 SCI41.09c	NP_625901	1497	498	55046	OK but not pure	strain freezed	unknown	Yes
16	CYP158A1	cytochrome P450	SC06998 SC8F11.24c	NP_631063	1224	407	44632	very very low	good	unknown	Yes
17	CYP158A2	cytochrome P450	SC012072 SCG58.07	NP_625496	1215	404	44286	OK	strain freezed	unknown	Yes
18	CYP159A1	cytochrome P450	SC00583 SCF55.07	NP_624895	1224	407	44848	none	good	unknown	Yes

注：本表格中颜色代表所纯化蛋白的颜色，颜色越深代表蛋白表达量越高。

第七章 结论

本人在博士后期间主要参与微生物生物合成霉酚酸 (MPA) 及国家优青项目“微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制”研究课题。

1. 在酿酒酵母中表达鉴定来自短密青霉的霉酚酸合成相关基因。

a) 利用酿酒酵母同源重组系统组装大片段 DNA 的技术路线, 证明了该技术在理性设计组装外源 DNA 和代谢工程改造酿酒酵母的可行性、高效性和可重复性;

b) 完成了短密青霉霉酚酸 (MPA) 合成相关基因 *mpaA*, *mpaDE* 和 *mpaC* 的分子克隆, 开展了发酵实验并 HPLC 检测代谢产物, 利用 SDS-PAGE 和 Western Blot 检测了蛋白表达情况;

c) 对 MPA 合成相关基因 *mpaA* 进行了密码子优化并蛋白纯化鉴定。

2. 国家优青项目“微生物生物合成 P450 单加氧酶与还原伴侣蛋白的相互作用机制”研究课题

a) 完成了蓝细菌来源的 6 个还原伴侣 Fdx 的分子克隆和蛋白纯化。

b) 克隆表达来自天然色链霉菌的 18 个 P450 酶基因和 6 个铁氧化还原蛋白 Fdx、4 个铁氧化还原蛋白还原酶 FDR, 并进行了蛋白纯化鉴定。对 Fdx6 密码子优化。c) 利用这些 P450 和还原伴侣进行了酶催化反应和 HPLC 检测。在尝试多种底物后, 发现了天蓝色链霉菌还原伴侣支撑 PikC 催化 YC17 底物进行反应。结果首次表明, 天蓝色链霉菌的还原伴侣蛋白催化支撑 PikC 催化 YC17 的反应, 并且在支撑 PikC 的效力层面, 通过采用不同组合 (不同性质或不同来源) 的还原伴侣, 可以达到不同的催化效率, 并影响目的产物类型。PikC 在 YC17 催化中的限速步骤存在于后两个产物合成步骤。为进一步获得对“P450 酶—Fdx—FDR”三蛋白系统相互作用和识别机制的规律性认识做出积累。

3. 其他工作:

a) 对短密青霉 *mpaC* 的预测序列进行纠正, 获得正确的内含子外含子区域注释。

b) 对序列未知的米曲霉表达载体 pTAex3 进行了文献调研, 设计引物全测序获得 pTAex3 的完整序列图谱, 为实验室米曲霉基因表达工作开展提供支撑等。

本论文仍有多处需要进一步开展实验进行改善。

致 谢

感谢国家博士后管理部门和青岛市政府提供博士后项目资助及补贴，同时感谢工作站和流动站对我的大力支持。

感谢我的两位合作导师对我的帮助。李盛英研究员知识渊博，为我提供了一流的科研环境和指导，他个人雷厉风行、思路犀利、学术严谨的风格让我印象深刻；青岛琅琊台集团股份有限公司董事长李悦明老师为我无私提供工作平台，为人温和、做事直爽的山东大汉形象在我心中特别高大；在此向他们二人表示衷心的感谢！

特别致谢马莉和张伟博士对酶纯化和 HPLC 及 LC-MS 工作给予的帮助。感谢徐卉芳、郑贤良、杜磊、齐飞飞、陈辉、阎金勇、刘琨、张兴旺、李众和刘奕、房博、孙悦、姚秋萍、宁琳琳、蒋媛媛等众位成员在实验和生活上给予的帮助！各位在我怀孕和哺乳期间充分发挥人性的光辉，生活和工作上对我尽可能的照顾，点点滴滴我和宝宝铭记于心！

衷心感谢我的博士导师吕雪峰研究员对我一如既往的支持，感谢提供酿酒酵母质粒和蓝藻 DNA；感谢已毕业的高政绪师兄无私分享酿酒酵母资料，同时感谢黄雪年博士提供酿酒酵母对照菌株。

感谢公共实验室白罚丽和王聪分别对 LC-MS 和 GC-MS 的支持，感谢人教处的苏华副处长以及阎星橙老师在出入站手续办理方面的帮助。

感谢我的爱人王喜成博士对我工作的支持和为家庭付出的辛勤劳动；感谢父母不遗余力帮我照顾孩子；感谢儿子王齐的到来。

齐凤霞

2016年6月

参考文献

1. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, et al. (1997) The complete genome sequence of *Escherichia coli* K-12. *science* 277: 1453-1462.
2. Ratledge C, Kristiansen B (2006) *Basic biotechnology*: Cambridge University Press.
3. Feldmann H (2011) *Yeast: molecular and cell biology*: John Wiley & Sons.
4. Botstein D, Chervitz SA, Cherry M (1997) Yeast as a model organism. *Science* 277: 1259-1260.
5. Goffeau A, Barrell BG, Bussey H, Davis R (1996) Life with 6000 genes. *Science* 274: 546.
6. Kitchin JES, Pomeranz MK, Pak G, Washenik K, Shupack JL (1997) Rediscovering mycophenolic acid: a review of its mechanism, side effects, and potential uses. *Journal of the American Academy of Dermatology* 37: 445-449.
7. Regueira TB, Kildegaard KR, Hansen BG, Mortensen UH, Hertweck C, et al. (2011) Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. *Applied and environmental microbiology* 77: 3035-3043.
8. Bentley R (2000) Mycophenolic acid: a one hundred year odyssey from antibiotic to immunosuppressant. *Chemical Reviews* 100: 3801-3826.
9. Mimouni D, Anhalt GJ, Cummins DL, Kouba DJ, Thorne JE, et al. (2003) Treatment of pemphigus vulgaris and pemphigus foliaceus with mycophenolate mofetil. *Archives of dermatology* 139: 739-742.
10. Anderson H, Bracewell J, Fraser A, Jones D, Robertson G, et al. (1988) 5-Hydroxymaltol and mycophenolic acid, secondary metabolites from *Penicillium echinulatum*. *Transactions of the British Mycological Society* 91: 649-651.
11. Wackett LP (2008) Microbial-based motor fuels: science and technology. *Microbial Biotechnology* 1: 211-225.
12. Hansen BG, Mnich E, Nielsen KF, Nielsen JB, Nielsen MT, et al. (2012) Involvement of a natural fusion of a cytochrome P450 and a hydrolase in mycophenolic acid biosynthesis. *Applied and environmental microbiology* 78: 4908-4913.
13. Zhang W, Cao S, Qiu L, Qi F, Li Z, et al. (2015) Functional Characterization of MpaG', the O - Methyltransferase Involved in the Biosynthesis of Mycophenolic Acid. *Chembiochem* 16: 565-569.
14. Del-Cid A, Gil-Durán C, Vaca I, Rojas-Aedo JF, García-Rico RO, et al. (2016) Identification and Functional Analysis of the Mycophenolic Acid Gene Cluster of *Penicillium roqueforti*. *PLoS one* 11: e0147047.
15. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, et al. (2011) Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Applied and environmental microbiology* 77: 3044-3051.
16. Urlacher VB, Girhard M (2012) Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in biotechnology* 30: 26-36.
17. Hannemann F, Bichet A, Ewen KM, Bernhardt R (2007) Cytochrome P450 systems—biological variations of electron transport chains. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1770: 330-344.
18. Zhang W, Liu Y, Yan J, Cao S, Bai F, et al. (2014) New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *Journal of the American Chemical Society* 136: 3640-3646.
19. 李众, 张伟, 李盛英 (2016) 细胞色素 P450 酶与微生物药物创制. *微生物学报* 56: 496-515.
20. Anzai Y, Li S, Chaulagain MR, Kinoshita K, Kato F, et al. (2008) Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics. *Chem Biol* 15: 950-959.
21. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, et al. (1996) Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence Determination of the Entire Genome and Assignment of Potential Protein-coding Regions. *DNA*

- Res 3: 109-136.
22. Holtman CK, Chen Y, Sandoval P, Gonzales A, Nalty MS, et al. (2005) High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA research* 12: 103-115.
 23. Van der Plas J, De Groot R, Woortman M, Weisbeek P, Van Arkel G (1986) Coding sequence of a ferredoxin gene from *Anacystis nidulans* R2 (*Synechococcus* PCC7942). *Nucleic acids research* 14: 7804.
 24. Reith ME, Laudenbach DE, Straus NA (1986) Isolation and nucleotide sequence analysis of the ferredoxin I gene from the cyanobacterium *Anacystis nidulans* R2. *Journal of bacteriology* 168: 1319-1324.
 25. Van Der Plas J, De Groot R, Woortman M, Cremers F, Borrias M, et al. (1988) Genes encoding ferredoxins from *Anabaena* sp. PCC 7937 and *Synechococcus* sp. PCC 7942: structure and regulation. *Photosynthesis research* 18: 179-204.
 26. Floß B, Igloi GL, Cassier-Chauvat C, Mühlhoff U (1997) Molecular characterization and overexpression of the *petF* gene from *Synechococcus elongatus*: evidence for a second site of electrostatic interaction between ferredoxin and the PS II subunit. *Photosynthesis research* 54: 63-71.
 27. Cozens AL, Walker JE (1988) Expression of a gene encoding a novel ferredoxin in the cyanobacterium *Synechococcus* 6301. *Biochemical journal* 252: 563-569.
 28. Thomas J-C, Ughy B, Lagoutte B, Ajlani G (2006) A second isoform of the ferredoxin: NADP oxidoreductase generated by an in-frame initiation of translation. *Proceedings of the National Academy of Sciences* 103: 18368-18373.
 29. Zhang J, Lu X, Li J-J (2013) Conversion of fatty aldehydes into alkenes by in vitro reconstituted cyanobacterial aldehyde-deformylating oxygenase with the cognate electron transfer system. *Biotechnology for biofuels* 6: 1.
 30. Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. *Journal of biotechnology* 124: 128-145.
 31. Bentley SD, Chater KF, Cerdeno-Tarraga A-M, Challis GL, Thomson N, et al. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* 417: 141-147.
 32. Lamb DC, Skaug T, Song H-L, Jackson CJ, Podust LM, et al. (2002) The cytochrome P450 complement (CYPome) of *Streptomyces coelicolor* A3 (2). *Journal of Biological Chemistry* 277: 24000-24005.
 33. Chun Y-J, Shimada T, Sanchez-Ponce R, Martin MV, Lei L, et al. (2007) Electron Transport Pathway for a *Streptomyces* Cytochrome P450 CYTOCHROME P450 105D5-CATALYZED FATTY ACID HYDROXYLATION IN *STREPTOMYCES COELICOLOR* A3 (2). *Journal of Biological Chemistry* 282: 17486-17500.
 34. Zhao B, Moody SC, Hider RC, Lei L, Kelly SL, et al. (2012) Structural analysis of cytochrome P450 105N1 involved in the biosynthesis of the zincophore, coelibactin. *International journal of molecular sciences* 13: 8500-8513.
 35. Lim Y-R, Hong M-K, Kim J-K, Doan TTN, Kim D-H, et al. (2012) Crystal structure of cytochrome P450 CYP105N1 from *Streptomyces coelicolor*, an oxidase in the coelibactin siderophore biosynthetic pathway. *Archives of biochemistry and biophysics* 528: 111-117.
 36. Podust LM, Bach H, Kim Y, Lamb DC, Arase M, et al. (2004) Comparison of the 1.85 Å structure of CYP154A1 from *Streptomyces coelicolor* A3 (2) with the closely related CYP154C1 and CYPs from antibiotic biosynthetic pathways. *Protein Science* 13: 255-268.
 37. Podust LM, Kim Y, Arase M, Neely BA, Beck BJ, et al. (2003) The 1.92-Å Structure of *Streptomyces coelicolor* A3 (2) CYP154C1 A NEW MONOOXYGENASE THAT FUNCTIONALIZES MACROLIDE RING SYSTEMS. *Journal of Biological Chemistry* 278: 12214-12221.
 38. Shao Z, Zhao H, Zhao H (2009) DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic acids research* 37: e16-e16.
 39. Guengerich FP, Martin MV, Sohl CD, Cheng Q (2009) Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nature Protocol* 4: 1245-1251.
 40. Omura T, Sato R (1964) The Carbon Monoxide-binding Pigment of Liver Microsomes. *Journal of Biological Chemistry*.

41. Ma SM, Li JW, Choi JW, Zhou H, Lee KK, et al. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* 326: 589-592.
42. Corinne Cassier-Chauvat FC (2014) Function and Regulation of Ferredoxins in the Cyanobacterium, *Synechocystis* PCC6803: Recent Advances. *Life* 4: 666-680.

附录

附录1 符号表

IPTG: Isopropyl β -D-1-Thiogalactopyranoside, 异丙基- β -D-硫代吡喃半乳糖苷

HPLC: High Performance Liquid Chromatography, 高效液相色谱

LC-MS: 液质联用色谱

GC-MS: 气质联用色谱

GC: 液相色谱

SDS-PAGE: 蛋白质变性凝胶电泳

Western Blot: 蛋白杂交

YPD: 培养基

M9: 基本培养基

SC(Synthetic complete): 酵母用合成全培养基

SC-Ura: 酵母用合成尿嘧啶缺陷型培养基

LB: 细菌用培养基

附录2 主要培养基

➤ LB 培养基

蛋白胨 10 g, 酵母粉 5 g, NaCl 10 g, 加水至 1 L, 高压灭菌。固体培养基加 1.5 g 琼脂。

➤ YPD (Yeast Extract Peptone Dextrose) 培养基

表 S-1 YPD 培养液

储备液	浓度
Yeast extract	1%
Peptone	2%
Glucose	2% (过滤除菌)
Agar	2%

➤ SC (Synthetic complete) 培养基

表 S-2 SC 培养液

储备液	浓度
Yeast nitrogen base	0.67%
Carbon source	2%
Agar	2%
Adenine	0.01%
Arginine	0.01%
Cysteine	0.01%
Leuine	0.01%
Threonine	0.01%
Tryptophan	0.01%
Uracil	0.01%
Aspartic acid	0.005%
Histidine	0.005%
Isoleucine	0.005%
Methionine	0.005%
Phenylalanine	0.005%
Proline	0.005%
Serine	0.005%
Tyrosine	0.005%
Valine	0.005%

注：氨基酸可混合起来溶解灭菌，配成 10×，-20℃ 保存

➤ 改良微量元素培养基 **M9 培养基**

表 S-3 SC 培养液

储备液	浓度
硫酸镁 MgSO ₄ ·7H ₂ O	25 g/100 mL 灭菌 4℃ 保存
氯化钙 CaCl ₂ ·2H ₂ O	1.1 g/100 mL 灭菌，4℃ 保存
硫酸铜 CuSO ₄ ·5H ₂ O	0.19 g/100 mL 灭菌，4℃ 保存
氯化锌 ZnCl ₂	0.131 g/100 mL 灭菌，4℃ 保存
钼酸钠 Na ₂ MoO ₄ ·2H ₂ O	0.2 g/100 mL 灭菌，4℃ 保存
氯化铁 FeCl ₃ ·6H ₂ O	2.7 g/100 mL 灭菌，4℃ 保存
硼酸 H ₃ BO ₃	0.05 g/100 mL 灭菌，4℃ 保存
Triton-X100	10 mL/100 mL 灭菌，4℃ 保存
维生素 B1	1.0 g/100 mL 过滤灭菌，4℃ 保存
Bis-Tris (pH=7.25)	20.9 g/100 mL 过滤灭菌，4℃ 保存
30% 葡萄糖	30 g/100 mL 灭菌，4℃ 保存

1. 按照上表分别配置 11 种母液；

2. 以配置 1 L 改良微量元素培养基培养基为例：首先将 15.1 g Na₂HPO₄ ·12H₂O, 3 g KH₂PO₄, 0.5 g NaCl, 2 g NH₄Cl 溶于 600 mL 超纯水中，灭菌待用；分别加入母液 1-6 各 1 mL 并混匀，然后再加入 10 mL 母液 8, 100 μL 母液 9, 200 mL 母液 10 及 100 mL 母液 11 混匀后即可分装使用。使用时，按照 1:1000 的比例加入 rare salt solution 及所需浓度的抗生素。

其他

DNA 琼脂糖凝胶电泳缓冲液 50×TAE Buffer:

1. 称取下列试剂置于 1L 烧杯中；

Tris	242g
EDTA	37.2g

2. 加入约 800ml 去离子水，充分溶解搅拌；

3. 加入 57.1ml 醋酸，充分搅拌；

4. 加去离子水定容至 1L，室温保存。

SDS-PAGE 电泳缓冲液 5×Tris-Glycine Buffer:

1. 称取下列试剂置于 1L 烧杯中；

Tris	15.1g
Glycine	94g
SDS	5g

2. 加入约 800ml 去离子水，搅拌溶解；

3. 加去离子水定容至 1L, 室温保存。

5×SDS-PAGE Loading Buffer

1. 将下列试剂置于 10ml 塑料离心管中;

1 M Tris (pH6.8)	1.25ml
SDS	0.5g
BPB	25mg
甘油	2.5ml

2. 加去离子水溶解后定容至 5ml, 加入 250ul β-巯基乙醇, 500μl/管分装, -20℃ 保存。

Western blot 转膜缓冲液:

1. 称取下列试剂置于 1L 烧杯中;

Glycine	2.9g
Tris	5.8g
SDS	0.37g

2. 向烧杯中加入 600ml 的去离子水, 搅拌溶解;

3. 加去离子水定容至 800ml, 加入 200ml 的甲醇, 室温保存。

Western blot 膜清洗液 TBST Buffer

1. 称取下列试剂置于 1L 烧杯中;

NaCl	8.8g
1 M Tris (pH 8.0)	20ml

2. 向烧杯中加入 800ml 的去离子水, 搅拌溶解;

3. 加入 0.5ml Tween 20 后充分混匀;

4. 加去离子水定容至 1L, 室温保存。

考马斯亮蓝 R-250 染色液

1. 称取 1g 考马斯亮蓝 R-250, 置于 1L 烧杯中;

2. 加入 500ml 去离子水, 搅拌溶解;

3. 加入 400ml 乙醇和 100ml 冰醋酸, 搅拌均匀;

4. 用滤纸除去颗粒物后室温保存。

考马斯亮蓝染色脱色液

1. 量取下列溶液, 置于 1L 烧杯中;

附录

去离子水	500ml
乙醇	400ml
冰醋酸	100ml

2. 充分混匀后室温保存。

附录3 本文涉及的质粒及菌株

表 S-4 本研究使用的质粒

Organism	Strain	Strain Name	Plasmid description	Medium	Resistance	Qifx#
<i>E.coli</i>	DH5 α		pDR195	LB	Amp	9001
<i>E.coli</i>	DH5 α		pYES2	LB	Amp	9002
<i>E.coli</i>	DH5 α		pACT2	LB	Amp	9003
<i>E.coli</i>	DH5 α		pSH47	LB	Amp	9004
<i>E.coli</i>	DH5 α		YIP5-KanR	LB	Amp	9005
<i>E.coli</i>	DH5 α		none	LB	none	9006
<i>E.coli</i>	DH5 α		pYES2/NT A	LB	Amp	9134
<i>E.coli</i>	DH5 α	FQ101	pMD19T- <i>mpaC</i>	LB	Amp	9007
<i>E.coli</i>	DH5 α	FQ102	pYES2- <i>mpaDE</i>	LB	Amp	9008
<i>E.coli</i>	DH5 α	FQ103	pYES2/NT A- <i>mpaDE</i>	LB	Amp	9009
			NT			
<i>S. cerevisiae</i>	BJ5464- <i>NpgA</i>	FQ104	genome integration of <i>NpgA</i>	YPD	none	9010
<i>S. cerevisiae</i>	INVSc1		none	YPD	none	9135
<i>S. cerevisiae</i>	BY4742		none	YPD	none	9136
<i>S. cerevisiae</i>	INVSc1	FQ105	pYES2/NT A	SC-Ura	none	9011
<i>S. cerevisiae</i>	BJ5464- <i>NpgA</i>	FQ106	pYES2- <i>mpaC</i>	SC-Ura	none	9012
<i>E.coli</i>	DH5 α	FQ107	pYES2- <i>mpaC</i>	LB	Amp	9013
<i>S. cerevisiae</i>	INVSc1	FQ108	pYES2/NT A- <i>mpaDE</i>	SC-Ura	none	9014
			NT			
<i>E.coli</i>	DH5 α	FQ109	pET28b-0338	LB	Km	9015
<i>E.coli</i>	DH5 α	FQ110	pET28b-0698	LB	Km	9016
<i>E.coli</i>	DH5 α	FQ111	pET28b-0814	LB	Km	9017
<i>E.coli</i>	DH5 α	FQ112	pET28b-0898	LB	Km	9018
<i>E.coli</i>	DH5 α	FQ113	pET28b-1749	LB	Km	9019

附录

<i>E.coli</i>	DH5 α	FQ114	pET28b-2581	LB	Km	9020
<i>E.coli</i>	BL21(DE3)	FQ115	pET28b-0338	LB	Km	9021
<i>E.coli</i>	BL21(DE3)	FQ116	pET28b-0698	LB	Km	9022
<i>E.coli</i>	BL21(DE3)	FQ117	pET28b-0814	LB	Km	9023
<i>E.coli</i>	BL21(DE3)	FQ118	pET28b-0898	LB	Km	9024
<i>E.coli</i>	BL21(DE3)	FQ119	pET28b-1749	LB	Km	9025
<i>E.coli</i>	BL21(DE3)	FQ120	pET28b-2581	LB	Km	9026
<i>S. cerevisiae</i>	BJ5464- <i>NpgA</i>	FQ121	pYES2- <i>mpaC</i>	SC- Ura	none	9027
<i>E.coli</i>	DH5 α	FQ122	pYES2- <i>mpaC</i>	LB	Amp	9028
<i>E.coli</i>	DH5 α	FQ123	pET28b-CYP51	LB	Km	9029
<i>E.coli</i>	DH5 α	FQ124	pET28b-CYP154C1	LB	Km	9030
<i>E.coli</i>	DH5 α	FQ125	pET28b-FDR1	LB	Km	9031
<i>E.coli</i>	DH5 α	FQ126	pET28b-FDR2	LB	Km	9032
<i>E.coli</i>	DH5 α	FQ127	pET28b-Fdx1	LB	Km	9033
<i>E.coli</i>	DH5 α	FQ128	pET28b-Fdx2	LB	Km	9034
<i>E.coli</i>	DH5 α	FQ129	pET28b-Fdx4	LB	Km	9035
<i>E.coli</i>	DH5 α	FQ130	pET28b-Fdx5	LB	Km	9036
<i>E.coli</i>	BL21(DE3)		pET28a-FNR	LB	Km	9037
<i>E.coli</i>	BL21(DE3)		pET28a-Fd	LB	Km	9038
<i>E.coli</i>	DH5 α	FQ131	pYES2- <i>mpaA</i>	LB	Amp	9039
<i>E.coli</i>	DH5 α	FQ132	pET28b-Fdx3	LB	Km	9040
<i>E.coli</i>	DH5 α	FQ133	pET28b-FDR3	LB	Km	9041
<i>E.coli</i>	DH5 α	FQ134	pET28b-FDR4	LB	Km	9042
<i>E.coli</i>	DH5 α	FQ135	pET28b-CYP102B1	LB	Km	9043
<i>E.coli</i>	DH5 α	FQ136	pET28b-CYP105D5	LB	Km	9044
<i>E.coli</i>	DH5 α	FQ137	pET28b-CYP105N1	LB	Km	9045
<i>E.coli</i>	DH5 α	FQ138	pET28b-CYP107P1	LB	Km	9046
<i>E.coli</i>	DH5 α	FQ139	pET28b-CYP156A1	LB	Km	9047
<i>E.coli</i>	DH5 α	FQ140	pET28b-CYP156B1	LB	Km	9048
<i>E.coli</i>	DH5 α	FQ141	pET28b-CYP157B1	LB	Km	9049
<i>E.coli</i>	DH5 α	FQ142	pET28b-CYP157C1	LB	Km	9050
<i>E.coli</i>	DH5 α	FQ143	pET28b-CYP158A1	LB	Km	9051
<i>E.coli</i>	DH5 α	FQ144	pET28b-CYP159A1	LB	Km	9052
<i>E.coli</i>	BL21(DE3)		none	LB	none	9053

<i>E.coli</i>	DH5 α	FQ145	pET28b-Fdx6	LB	Km	9054
<i>E.coli</i>	DH5 α	FQ146	pET28b-CYP107T1	LB	Km	9055
<i>E.coli</i>	DH5 α	FQ147	pET28b-CYP107U1	LB	Km	9056
<i>E.coli</i>	DH5 α	FQ148	pET28b-CYP154A1	LB	Km	9057
<i>E.coli</i>	DH5 α	FQ149	pET28b-CYP155A1	LB	Km	9058
<i>E.coli</i>	DH5 α	FQ150	pET28b-CYP157A1	LB	Km	9059
<i>E.coli</i>	DH5 α	FQ151	pET28b-CYP158A2	LB	Km	9060
<i>E.coli</i>	BL21(DE3)	FQ152	pET28b-Fdx1	LB	Km	9061
<i>E.coli</i>	BL21(DE3)	FQ153	pET28b-Fdx2	LB	Km	9062
<i>E.coli</i>	BL21(DE3)	FQ154	pET28b-Fdx3	LB	Km	9063
<i>E.coli</i>	BL21(DE3)	FQ155	pET28b-Fdx4	LB	Km	9064
<i>E.coli</i>	BL21(DE3)	FQ156	pET28b-Fdx5	LB	Km	9065
<i>E.coli</i>	BL21(DE3)	FQ157	pET28b-Fdx6	LB	Km	9066
<i>E.coli</i>	BL21(DE3)	FQ158	pET28b-FDR1	LB	Km	9067
<i>E.coli</i>	BL21(DE3)	FQ159	pET28b-FDR2	LB	Km	9068
<i>E.coli</i>	BL21(DE3)	FQ160	pET28b-FDR3	LB	Km	9069
<i>E.coli</i>	BL21(DE3)	FQ161	pET28b-FDR4	LB	Km	9070
<i>E.coli</i>	BL21(DE3)	FQ162	pET28b-CYP51	LB	Km	9071
<i>E.coli</i>	BL21(DE3)	FQ163	pET28b-CYP102B1	LB	Km	9072
<i>E.coli</i>	BL21(DE3)	FQ164	pET28b-CYP105D5	LB	Km	9073
<i>E.coli</i>	BL21(DE3)	FQ165	pET28b-CYP105N1	LB	Km	9074
<i>E.coli</i>	BL21(DE3)	FQ166	pET28b-CYP107P1	LB	Km	9075
<i>E.coli</i>	BL21(DE3)	FQ167	pET28b-CYP107T1	LB	Km	9076
<i>E.coli</i>	BL21(DE3)	FQ168	pET28b-CYP107U1	LB	Km	9077
<i>E.coli</i>	BL21(DE3)	FQ169	pET28b-CYP154A1	LB	Km	9078
<i>E.coli</i>	BL21(DE3)	FQ170	pET28b-CYP154C1	LB	Km	9079
<i>E.coli</i>	BL21(DE3)	FQ171	pET28b-CYP155A1	LB	Km	9080
<i>E.coli</i>	BL21(DE3)	FQ172	pET28b-CYP156A1	LB	Km	9081
<i>E.coli</i>	BL21(DE3)	FQ173	pET28b-CYP156B1	LB	Km	9082
<i>E.coli</i>	BL21(DE3)	FQ174	pET28b-CYP157A1	LB	Km	9083
<i>E.coli</i>	BL21(DE3)	FQ175	pET28b-CYP157B1	LB	Km	9084
<i>E.coli</i>	BL21(DE3)	FQ176	pET28b-CYP157C1	LB	Km	9085
<i>E.coli</i>	BL21(DE3)	FQ177	pET28b-CYP158A1	LB	Km	9086
<i>E.coli</i>	BL21(DE3)	FQ178	pET28b-CYP158A2	LB	Km	9087

附录

<i>E.coli</i>	BL21(DE3)	FQ179	pET28b-CYP159A1	LB	Km	9088
<i>S. cerevisiae</i>	INVSc1		pYES2-IGA	SC-	none	9089
				Ura		
<i>S. cerevisiae</i>	INVSc1		pYES2-IMG1	SC-	none	9090
				Ura		
<i>S. cerevisiae</i>	INVSc1	FQ180	pYES2-NTA-mpaA	SC-	none	9091
				Ura		
<i>S. cerevisiae</i>	INVSc1	FQ181	pYES2-NTA-mpaA	SC-	none	9092
				Ura		
<i>E.coli</i>	RIPL	FQ182	pET28b-Fdx1	LB	CmSmKm	9093
<i>E.coli</i>	RIPL	FQ183	pET28b-Fdx3	LB	CmSmKm	9094
<i>E.coli</i>	RIPL	FQ184	pET28b-Fdx6	LB	CmSmKm	9095
<i>E.coli</i>	RIPL	FQ185	pET28b-FDR4	LB	CmSmKm	9096
<i>E.coli</i>	RIPL	FQ186	pET28b-Fdx2	LB	CmSmKm	9097
<i>E.coli</i>	RIPL	FQ187	pET28b-CYP51	LB	CmSmKm	9098
<i>E.coli</i>	RIPL	FQ188	pET28b-CYP155A1	LB	CmSmKm	9099
<i>E.coli</i>	RIPL	FQ189	pET28b-CYP156A1	LB	CmSmKm	9100
<i>E.coli</i>	RIPL	FQ190	pET28b-CYP157A1	LB	CmSmKm	9101
<i>E.coli</i>	RIPL	FQ191	pET28b-CYP157B1	LB	CmSmKm	9102
<i>E.coli</i>	RIPL	FQ192	pET28b-CYP157C1	LB	CmSmKm	9103
<i>E.coli</i>	RIPL	FQ193	pET28b-CYP158A1	LB	CmSmKm	9104
<i>E.coli</i>	RIPL	FQ194	pET28b-CYP158A2	LB	CmSmKm	9105
<i>E.coli</i>	RIPL	FQ195	pET28b-CYP159A1	LB	CmSmKm	9106
<i>E.coli</i>	RIPL	FQ196	pET28b-Fdx2	LB	CmSmKm	9107
<i>E.coli</i>	RIPL	FQ197	pET28b-Fdx2	LB	CmSmKm	9108
<i>E.coli</i>	RIPL	FQ198	pET28b-Fdx4	LB	CmSmKm	9109
<i>E.coli</i>	Top10 F'	FQ199	pET28b-Fdx6-co	LB	Km	9110
<i>E.coli</i>	Top10 F'	FQ200	pYES2-mpaA-co- NThis	LB	Amp	9111
<i>E.coli</i>	BL21(DE3)	FQ201	pET28b-Fdx6-co	LB	Km	9112
<i>E.coli</i>	BL21(DE3)	FQ202	pET28b-pikC	LB	Km	9129
<i>E.coli</i>	BL21(DE3)	FQ203	pET30-BM3	LB	Km	9130
<i>E.coli</i>	BL21(DE3)	FQ204	pET?-BM3-F87A	LB	Km	9131
<i>E.coli</i>	BL21(DE3)	FQ205	pACYC-MycG	LB	Cm	9132
<i>E.coli</i>	RIPL		none	LB	CmSmKm	9133

<i>E.coli</i>	RIPL	FQ206	pET28b-FDR1	LB	CmSmKm	9113
<i>E.coli</i>	RIPL	FQ207	pET28b-FDR2	LB	CmSmKm	9114
<i>E.coli</i>	RIPL	FQ208	pET28b-FDR3	LB	CmSmKm	9115
<i>E.coli</i>	RIPL	FQ209	pET28b-CYP102B1	LB	CmSmKm	9116
<i>E.coli</i>	RIPL	FQ210	pET28b-CYP105D5	LB	CmSmKm	9117
<i>E.coli</i>	RIPL	FQ211	pET28b-CYP105N1	LB	CmSmKm	9118
<i>E.coli</i>	RIPL	FQ212	pET28b-CYP107P1	LB	CmSmKm	9119
<i>E.coli</i>	RIPL	FQ213	pET28b-CYP107T1	LB	CmSmKm	9120
<i>E.coli</i>	RIPL	FQ214	pET28b-CYP107U1	LB	CmSmKm	9121
<i>E.coli</i>	RIPL	FQ215	pET28b-CYP154A1	LB	CmSmKm	9122
<i>E.coli</i>	RIPL	FQ216	pET28b-CYP154C1	LB	CmSmKm	9123
<i>E.coli</i>	RIPL	FQ217	pET28b-CYP156B1	LB	CmSmKm	9124
<i>E.coli</i>	RIPL	FQ218	pET28b-Fdx5	LB	CmSmKm	9125
<i>E.coli</i>	RIPL	FQ219	pET28b-Fdx6-co	LB	CmSmKm	9126
<i>E.coli</i>	RIPL	FQ220	pET28b	LB	CmSmKm	9127
<i>E.coli</i>	BL21(DE3)	FQ221	pET28b	LB	Km	9128

Amp: 氨苄抗生素; Sp: 壮观霉素; Km: 卡那霉素; Cm: 氯霉素; Sm: 链霉素。

RIPL: CodonPlus RIPL

附录4 本文涉及的基因序列

> *mpaC*

ATGAATTTCCACAAAGGGCAACCGAAGGAAGATCTTCGGGTCCTCTTTGGACCACAATGCC
CGGACATCACAGATTCTATCACTCACATACGAGATGCCATATCCAAAGATCCCACAGGTTTG
GGATTTCTCACAAATATTCTCGACGAATTACCTCCTTATGGCCAACAATCGCCGGTGCCTG
GCCAGCTTTGAAGAATGTCGAAGGCGAGAATCAACTGCTTGCACTCGGCCGCCTTTTTGA
GCATGAGAGTGAAGTCAGGGTAGAAGCTTCAAACCTCATGATGACCCCAATAACTGTTATG
AGACACGTCGTTGATTTCTGGAATCTCCAAGATGTTGCCACACATCCAGCATTCCCTTCATC
TTCATTATCTGAGACAGAAATGCCAAGGATAGTTGATACTCAAGGATTCTGCGTGGGTCTGC
TTGCAGCAATTGCGGTGGCTTGCTCTCGAAACACACAGGAGTTCCAATATGTGGCGTCAAA
CGCCATACGCCTTTCTCTTTGCATTGGAGCGCTAGTGGACTTGGATGAGATTTTGTGTGGC
TCAACCACGTCACTGGCAGTCAGGTGGGAGTCCGTAGAGGACTTCAACCATCTCGAAAAA
ATCCTAAACAACAACACTGAAGGATACACATCATGCTACACAGACGTCAAAAGTGTCACCAT
CACTATCCCAATGACAGCGCAGAACGAGTGAAGCAGGAGATCCATGATCATGGGCTGCG
TACAAAGCAACTATCCCTCCGAGGGCGATTCCATCATGAGGCACACCGTGAGGGAATCCAA
CACATCATGAAGTTATGCATGAATGATTCTCGTTTCAAGCTTCTCGCAGTGATGCCTTATTA
ACTCCGCTTCGGTCAAGCCAGGGCGGGGAGATATTCCAACAAGAAGCATTACTACACACTG
TTGCACTGGACTCAATTCTTTGTGCCAAAGCAAACCTGGTACGACGTTGTCTCCGCTTTGAT
CAACAGCACAGAAATGACAGTTGATCAATCTCGCCTACTGTCTATTGGACCCGAGGAGTTC
GTTCCACGGTCTGCGAGGAGTAGGTGCGTGGCACGAAGAGAACTTGAGAGCTACGGAAT
GCAAGGATTTGCCAACGAAAGCCCACAGCCCTCCACGGCGTCCTTGTCAAACCTCTGTGCA
GACCTTCGACTCTCGTCCGCAGGCGGCTGAGGGCGTCGCCCATCGCCATCACCGGCATGG
CATGTCGGTATCCCAACGCAGACACACTCGCTCAGCTGTGGGAGCTACTTGAGTTAGGTC
GGTGCACAGTGAAGTCGCCCCCTGAGAGCCGATTCCATATGTCCGACTTGCAACGCGAGC
CCAAAGGCCCAATTCTGGGGTCACTTTCTGGAGCGCCCAGACGTCTTTGACCATCGCTTTTT
CAATATTTGGGCTCGGGAAGCCGAATCCATGGACCCCAGCAGCGCGTCCGCGCTGCAGG
TAGCATA CGAAGCCATGGAGTCTGCTGGATACTTAGGATGGCAGCCAAACAGGCTCTCACA
GGATATTGGGTGTTACGTCGGTGTGCGGATCTGAAGATTACACAGAGAATGTGGCATCAAGG
AACGCTAATGCGTTCTCGATCACGGGAACCCTGCAGTCGTTTATTGCAGGGCGGATCAGC
CATCACTTCGGCTGGTCAGGCCCTCAATCAGCTTGGACACAGCGTGCTCATCGGCGGCA
GTAGCTATTATCTGGCATGCAAAGCTCTGCAAACAACGATTGCAAGATCGCTTTGGCCG
GTGGTGTGAATGTTTTGACAAACCCAGAGTGTACCAGAACCTGAGTGCTGCGTCTTTCT
GTCACCAAGCGGAGCGTGCAAGCCATTTGACGCTTCCGCTGATGGATACTGCCGCGGAGA
GGGCGCCGGGTTATTCGTTCTGCGACCATTGCAGGATGCAATTGACAACGGAGATCCAATT

CTCGGTGTAATCGCTGGATCAGCTGTCAACCAGGGCTCCAACAATTCACCAATCACAGTCC
CCGATGCGGAAGCCCAACGGTCACTCTACAACAAGGCTATGTCGCTCGCTGGAGTTTCTC
CTGACGAAGTCACCTATGTGGAAGCCCATGGCACTGGCACCCAAGTCGGGGATCCCATTG
AGCTGGACAGCCTTCGAAGAACGTTTGGTGGCCCTCACCGCAGAAACAACCTTCACATCG
GATCGATCAAGGGAAATATCGGGCATACTGAGACATCCTCTGGGGCCGCGGGTCTGCTGA
AGACCATCCTTATGCTCCAACAACAACGCATCCCGAGACAGGCAAACCTTCAATCAGTTAAAT
CCCAAGGTCAAGTCTCTCACACCCGACCGGCTAGTGATCGCTTCAGAGTCGACTGAGTGG
GTGTGACCAAAACGAGTTGCCATGGTCAGCAACTACGGAGCATCGGGGAAGCAACGCGGC
GCTCATAGTGAAGGAGCACGCACCCATCGGCTCCGAGCAAAAATGGAACAGCCCCTGAGTA
TATCCAAAATGTACCAATCTTGGTCTCCGCTCGGTCTGAAGAATCCCTCCGGGCTTACTGC
GGTGCCCTTCGAGCGACTCTTTTGAGCCACCCACCATCTGAAACTCTGGTGCAGAAGTTG
GCATAAACCTCGCGATGAAGCAAAACCGGGACCTTCCACTCAATCTCACATTTTCAACTTC
TTCGGACACAACCTCGCTGGGCGCCCGACTGGAGGCCATTTCCACGGGCGCATCTGCCG
ATCTGATTCAGAAGCGGCCTTCCAATGAGCCCCAGTTGTTTTGTGTTTCGGTGGTCAGAA
TGGACTTACGGCAACCATTTTGAAGGAAGTTTTGACGCCTCCGCTCTTCTACGGACCCAC
CTCGAAGATTGCGAGGAAGTGGGCCGCACGCTTGGCCTTCTCAGTCTTTTCCCAACTATCT
TCAGCTCAGCCCCAATCACAACATCATCCATTTGCATTTCACTACTCTTCTATTAGTATG
CTTCTGCAAAGGCATGGCTGGACTCTGGACTTCGTGTCAGTCGCATTGTGGGACACAGCT
TTGGTCAGTTGACTGCGTTGTCTGTGGCGGGCAGCTTGAGCGTTCGCGATGGCATTACC
TAGTCACTGAGCGTGCTCGCCTCATTGAATCCAGCTGGGGCCAGAATCTGGGATCATGC
TCGCCGTTGAAGGGACGGAGATTGAGGTCCAACAGCTGCTAGACCAGACGGGTCATATCG
CTGACGTTGCCTGCTACAATGGGCCACGACAGCAAGTACTGGCAGGTACAGTGGAATCTAT
CGAGGCCATTGAGAACGCCGCTGCTAGAACCCCGTCTGCATCAAACCTTGCGGTTGACTCG
GCTGCAAAATTCCCATGCTTTCCATTCTCGACTTGTCGATAGCATTGTGCCCGGCATCATGG
AGGTGGCTGGGTCCCTTGTTTACCAGACACCGATCATTCCCATCGAAGCCTGCTCTGCAA
GTGGTACTGGTCGACCATCACAGCGGCTGAGATTGTGGAGCACAGCCGCATGCCTGTG
TATTTCCGACGTGCTGTTGAGCGAGTAGCTGAAAAGCTACAAGCCCCCGCGTCTGGCTG
GAAGCAGGATCCGCTTCACCCATCATAACCAATGGTACGCCGAGTGCTAGAGAGCTCTTCG
GTTGCGAACACATAACCAAAAATTGATTTGGGTGGTTCAAGCGGAGCCAAAATTTGGCAA
ACGTCCTAGCGCCCTCTGGGCACAGGGTGTACACGTTTCAGTTCTGGCCTTTTCGATCGGG
CCCAGCATGCAAGCTTCAAATGGATGAACCTTCCACCATATCAGTTTGCCAAAACAGCCA
CTGGGTTGACTTTGACCCTGCAGCGTTTTTCATCTGCTGGACCTTCGTCTGGAAAGCAGTCC
GCAGGACAAGAAGCAGGCCTTCTGTGTCAATTGAGCGAGAGCCCAGATGAGCGTCTCTAT
CATGTTAACATTCAAGATGCCCTCTACAGAGCATGCACCCAGGGTCACGCAGTTCTGAATC

AGACATTGTGTCCCGCCTCCATGTACATGGAGATGGTTTTGCGGGCAGCAGCCTCCATTT
CCCCACGGGCAACGCCTCTGAGCCAGCCATGTACATATCGAGGATCTCACCATCTCCTC
GCCCTTGGTATTGGACCCACAAGGAAAGGTGTTCTGAGACTCACCCGCGATGGAGCGG
GTCCTACTCGACCATGGCTCTTTTCGATCTTCAGCAGTGAGTCAAATGATCATAATTGGTC
CACGCTGAAGGTACCGTGTGTCTTACCAGGAGCGGTCCAGGGCTCTGGCGCGCTTCCA
GTCCATGGATCGACTGCTGGACTCGGCACGAAGCAAACCATTGAAGCGGACCCTGCTTC
AAACGGGTTGAAGGGCTCTACTGTCTATGCTGCTCTTGAATCAGTAACAACTATGGAGACT
ACTTCCGTGGCGTGAAGCAGGATTTGCAAACGGTCCGGGAAGCCAGTGGACTGGTGTCCA
TGATGCCATCAACCACCGAGACCAACTGCAACCCCATCTGCTAGACAACCTTCTCCAAGT
TGCTGGAATTCACGTCAACTGCCTTTCTGATCGCCAGTCAAGCGAGGTCTTTGTCTGCAAC
GCTATTGGGGAAACATTTGTTATCAACTCCTTGCTCAAGCAAGAGAATGGTGTCTCCCCCTC
GACATGGAAGGTGTACACCAGCTACGTCCGGCCGTCTAAGACTGAGATCGCGTGCACAT
CTATGTGATGGACTGCCAAACAGATACTCTCTGCGGCAATGATGGGTGTTTCGATTTACG
AGCGTCTCAATCCGCTCTCTCACTCGCGCGCTGGCCAAGCTGAATAACAATGTTCTAGAGA
CCGCTGAAGCTCAATCAGTTGTGAGGCCGCGATTCCGGCTGAGCAATCAGTTGTTACAG
CCACCCCTAGTGCACCGGCTGCCGATGGACATGCGGCTAACGATCTTGCCACTGTCCAGG
AAATGCTCTGTGAGCTCTTTGGAGTCAGTGTGCGCAGAGGTCTCTCCTTCAGTGTGCTGGT
AGACATTGGAGTGGACTCTCTCATGAGCACTGAAGTTCTTTCTGAAATCAAGAAGAGGTTT
CAAGTCGATATGTCGTATAACCACTTTGGTGGACATCCCAACATTCAAGGTCTGGTGGAAC
ATATCTTCCCCGGGCATTCTCACGCTGCCCCCTCACGGCCTGTTGTTGAAAAAGCGCCAGT
ACAGTCGGTGGCACCACAAGCTGTATCCCATGTACCGACACCGGCTAGCAATGGACCACC
ACTAGTCTCAGTGGCCCGCCAGTGTGTTGACACAACGCACGCAGCAGTATCTCACACCTCC
GATGCTCACTGGGCTGGCTTCTTCCACACCACTTACCCAAAGCAAATGTCACTTCTCACTG
CTTACATTCTCGAGGCATTCCGAGCCTTGGGCTCTTCGCTGGAGGCGAGTGAGCCAAACG
AAGTGCTCACCCCAATCGCAGTGCTTCCCCGTCACGAGCAGCTGAGAAAGCATCTTTACAA
AATCTTGGATTCTGTGCGGCTTGGTCCGCCAGATGTCCACTGGGGAGCTTGTGCGCACGAC
AACACCCATCCCGTTGTCGCGGTCGCATGACCTGCACACGCAGATAAGGGCGGAATATCC
TCCGTATGCCCTGGAGCATGATCTTCTTCAAATCACAGCCCCTCGGCTCGCTGACTGTCTC
ACCGGAAAAGCGGACGGAGTATCTCTGATTTTCCAGGATGCGAACACTCGGCGCTTGGTA
GGCGATGTGTACGCACAATCTCCAGTTTTCAAATCGGGTAACTTTATCTTGC GCGTTACCT
TTTGGATGTTGTTCAATCATTGGAAGCAGCCGGACTATCAAGATTCTTGAGATTGGTGCCG
GCACAGGAGGTACCACCAAGAATCTCCTCGAGAAATTGTCCACAATTCCTGGACTGTGAC
GCGTCTGGAGTACACATTCACAGACATCTCGCCATCACTTGTTGCAGCTGGACGGAAGAC
CTTCTCTAACTACAACCTTCATGCGATATGAACTCTGAATGTTGAAAATGACCCCCCTCAG
CTTTGAGTGGTCAATACGATATCGTGCTTTCGACAAATTGTGTCCATGCTACCCGCAACCTC

CAGGAATCTTGCACGAACATTTCGCAAGCTTCTGAGACCTGACGGAATTTTGTGTCTTGTGCG
AGCTGACTCGTGACATCTTCTGGTTGGATCTTGTCTTTGGTCTCCTGGAAGGCTGGTGGC
GATTCGAAGACGGACGTGAACACGCGCTGGCTACTGAGATGATGTGGGATCAGACTCTTC
GTCAATCTGGATTTGAGTGGGTTGACTGGACAAACAATGAGACCGTAGAGTCAAATGCTCT
GCGTGTCAATTGTCGCTTCGCCGACTGGAAACTCATCCGCCGCGACAATGTCTCCATCGAA
ACCCATAAAGATGGAAACTGTGGTTTTGGGGTGAGCGAGACAATTTGCAGCTCCGTGCAGA
CATCTACTATCCCGAGACCGTGGACACTACTCGGAAGCAGCGACCGATTGCCCTGATGATC
CACGGAGGTGGCCACGTATGCTATCTCGCAAGGACATCCGTCCAGCCCAGACTCAGACC
CTTCTGGATGCCGGTTTTCTGCCTGTTAGCATCGACTACCGCTTGTGCCCAGAAGTTTCAC
TAGCAGAAGGCCCAATGGCCGATGCCCGCGACGCTCTGTCTGGGTTCCGCCGCGTCTC
CCGAACATTCCTTGCTGCGCGCAGATATTCGACCCGACGGGAACCAGGTTGTTGCTATC
GGATGGTCTACTGGTGGTCATCTCGCCATGACACTACCTTTCACGGCCCCAGCTGCTGGTA
TTCCAGCTCCAGACGCTGTCTTGGCATTCTACTGCCCCGACCAATTACGAGGATCCATTCTG
GTCGAACCCAAACTTCCCCTTCGGACAGACAGTTCGCATCAAACGAGATGGAGTACGATGTT
TGGGAAGGATTGCAGAGCATGCCGATAGCCGGCTACAACCCTGCACTCAAGGAACGCCCT
CTCGGAGGCTGGATGTCAACGAGAGATCCGCGCAGTCGGATCGCACTGCACATGAACTG
GACTGGACAAACCTTGCTGTTCTGTTAAAGGCATGCACAATCAAGGGTAAACCCGAGAAA
TGCAGCCCCGATGACCTATCTCGTCCGACAGAGGAGGACATTCAGGCAGTCAGTCCCAAT
TATCAGATCCGCGTCGGTCGCTACAACACGCCGACATTTTGGATTCATGGAAACAAGTGATG
ACCTTGTTCTTGCTCAGACGGAGTCCACGCATGGTGCATTGACGGCCAGCGGAGTCG
AGGCGGAGCTTAGAGTGGTGCAGGAAGCTGCACACCTTTTCGACCTTACCCAGCCTCAC
ACGCTGGCCAGGAAGCAAAGGCCGAGTGGCAGAAGGATATGAGTTCTTGAGAACACATG
TCCAGCTTTAG

> *mpaDE*

ATGGAGTCTTTGTCGCTAACATGGATCACGGCTATTGCCGTGGTACTATATCTCGTCCAACG
CTATGTCAGGTCATATTGGCGGCTCAAGGATATAACAGGTCCGGTGTGGCAAAACTCACG
GATCTCCAACGTGTGTGGTGGGTGAAGACTGGCCGAGCACATGAGTTTCATCGTGACATG
CATGCCATGTATGGCCCTATCGTTCGATTTGGACCAAACATGGTTTCAGTCTCTGATCCTCG
TGTATACCAACCATCTACCCGAGTAGACCTGGATTCCCAGGGGTGACTTTTATCGCACGC
AAAAACCATATAACCCGGAATAAGGGCGCAATGCCCGCGGTCTTAAACTCAAGACGAGGA
TCTGCACAAGCAACTTCGAAGCCCGATCGCCTCGTTGTATTCCATGACAAATGTGGTGAGA
CTCGAGCCACTGGTCGATGAAACTCTTACAGTACTATCGAAACAACCTCGATGAGCGCTTCG
TGGGTACGAATGATAAGCCGTTTCGATCTCGGGGATTGGTTGCAATATTCGCCTTCGACTC
CATGGGCACATTGACCTTCTCCGAAGATACGGCTTTTTGGAGCAAGGTCGTGATATGCAT

GGAATCTTGCAGGAGATTTGGAATTCATGACCCGAGTCGCTGTGATGGGACAAATTCCGT
GGTTTGATGAGATCTGGAACAAGAATTCATTACTTTATTCAAGCGACCTACCGGGTTT
GGTGTGTTGAAGGTTGTCGATAACTTTATATCTCAACGAGTGTCCAGTCGTGAGAATGACGA
GAAAGCAGACGAAAAGGACATGCTTTCTCAATTCTTGAATATCCAAGCATCAAATCCTCATT
CTATCATGCCTTGGGCACCAAGGGCGTGGACCTTCTCTAACGTCATGGCTGGCTCAGACT
CTACTGCGAATGTCATGCGTACCATGATGTATAACCTTCTCGTTGACCGAGATACGTTGAAA
AGTCTGCGAGCCGAGCTTTTGAAGCGGAGAGCTCAAATGGCCTGTCTCGATCCTTGCCT
TCTTGGGATGGAGTGAGAAGCCTGCCTTATCTTGATGCATGCGTTCTTGAAGCTTTGCGTC
TGCATCCCCATTCTGTCTGCCCTTTGAACGAGTCGTCCCTGAAGGCGGTATCACAGTCTG
TGAGACATATCTTCCAGCGGGAACAGTTGTTGGAATTAGTCCCTACTTGGCCAACCGCGAT
AAACAAACATTTGGTGACGATGCAGATAAATGGCGACCTAGCAGATGGTTAGATTTGAGCC
GCGAAGATCGGGTCAAGCTTGAGAACAGTATTCTCACATTTGGCGCGGGACGTCGGACTT
GTTTAGGAAAGAACATCGCCATTCTTGAGATCAAGAAGCTCTTCCGATGTTGCTTTTGAAT
TATGAAATCGAAATTGTGAACCCTGAAAATTACCAGACCACAAATGCCTGGTTTTTCAGACA
GTGGGGCTTGACGCAGTGATTCGCAAACCTGCCAGCACCAGAACGAGATGATACCATTGA
GCAGAAAGCTTCCATCCCGCCTGCTTTGAATATTCCTCCTTCATCTTCGACAGTCGATGTGC
GAATTATTGACTCGGGAACACTACTTGATCTGAGACCTGATTTATTCTGGACACCTGATCTC
CCAGGGCTTCTGAAGGTCACAGCTCCAACCTTACTGCTTTTTGATCTCAAATGGCAGTCGGC
ATGTTTTGTTGACCTAGCCGTTTCGACAAGACTGGGAAAACCTCCCACCATCCATCGTGGC
CATGATCAAATCGCAAACGGTCATCCAAGAGCCACGCAACATCTCAGATGTTCTTACTCA
GACGAATCCTCCCTGGGGATCCGAAGCAAAGATATTGAAGCGATCATCTGGTTCGCATGCC
CATTTGACCACATTGGTGACCCATCAACTTTTCCCCATCTACCGAGCTAGTCGTCGGAC
CTGGCATCCGAGACACCCACTGGCCCGGGTTCCCAACCAACCCAGACGCAATAAACCTCA
ACACCGACATCCAAGGTCGCAATGTGCGTGAGATCTCCTTTGAAAAGACACAGAAGGGAG
CCACCAAGATCGGTTCCTTTGACGCCATGGACTATTTGGCGATGGGTCCTTTTATCTTCTA
GATGCTGCGGGTCATTCCGTGGGCCATATCGGTGCTCTTGCTCGCGTGACTACCTCTCCA
GACTCGTTTGTCTTCATGGGTGGTGACTCATGTCACCATGCCGGAGTGCTTCGACCCACAA
AATATCTTCCTTGTCCACTCGACTCTGGTGACACTTCACTTCCATGCAAATCCGACTCTGTT
TTCACGTTATCGCCTGCACTGCCAACTGATTACACTGCTGCTTTGAGGACAGTCGAGAATAT
TAAGGAGCTCGATGCCTGTGAGGATGTATTCGTCGTCCTTGCTCATGATGCTACCTTGAAA
GGGAAAGTCGACTTTTACCCTTCGAAAATCAATGATTGGAAGGCGAAAGAGTACGGCAAGA
AGACAAAATGGCTTTTTTACAAGGACATTGAGAATGCCATAGAAGGACAGAAGTAAA

> *mpaA*

ATGACCAACGCAGTGGAGGACTCTGGTCCTAGGGACCTGCTGTTCCCTACTGATCAGCACC
TCGCGGTTCAATCGATACATGCCTTATTACACGATGATGGCTGCTGTATGGGCCACTTTCAT

CGCCGGCGCTCTCAAATTACAACAGGATCCCGAAAGCCTTTCGATCGAGTTCATTTTGTAC
AAGGCCGGTCTCTGTTTCGTCCACTGTCTGCTGCTTTGCGGTGCTGGAAACACATGGAAC
GATCTTGTGGATCGCGATATTGACGCCAGAGTGGCACGCACAAAGATGCGGCCTTTGGCA
TCTGGAAAGGTCACTCTCACTGAGGCACTTCTTTGGATGACCGGTCAATATTTCTGTCTGT
CAAAATGCTGGATCTCATTTTGGATGGACGAAACATCTGGACCCTCATGCTGCCATTGACC
GCCAGCATCATGCTTTACCCCTACCTGAAGCGACCCATCTTCAGCAAGGTCTTCGTTTACC
CCCAATATATCCTGGGTCTTGCTATTGGTTACCCCGCCATCACGGGGTGGGCTTCGATCAC
CGGCAGCGAGGAGCCCCTTGGTGACATCATCAAGCATTGCATCCCAATCTGTCTTCTTGTC
TTCTTCTGGTGTGTGTACTTCAACACCGCATAACAGCCACCAAGATAGCGTCGATGACCGCA
AGATGAACATAAATTCTGCCTACGTCATTGCCGGTCAGCGCATTGACTGTTCTTGCCTTT
TTGAGTGTCTGCCACTTCTGACCATCCCATACATCATCTCTACGATCAACTCGCCGTGGTT
GTGGGTCTCCTGGATGGCTACCTGGACTGTTTCCATTGTTATGCAGATTGCTCAGTTGATT
CGCAGAAGCTTGAAAGTGGTGGTCGTATCCATTGGGACAACCTCCTGCTGGGATTGTGGA
CGATCGTTGCTTGCATGGTTGAAGTCGGATTGCAGAAGGTGGAATTCTGGAAGAATGTTGA
GGGGTACATTAAGCTTTGA

> *mpaA-codon optimization for yeast*

ATGACAAATGCTGTCGAGGACTCTGGTCCAAGAGATTTGTTGTTTTTATTGATTTCTACTTCT
AGATTAAATAGATATATGCCATACTACACTATGATGGCTGCTGTTTGGGCTACTTTCATTGCTG
GAGCTTTGAAGTTGCAACAGGACCCAGAGTCTTTGTCTATTGAATTTATTTTATATAAGGCTG
GTTTGTGTTTCGTCCACTGTTTGTGTTGTGCGGTGCTGGTAAACACTTGGAACGACTTGGT
TGATAGAGATATTGATGCTAGAGTCGCTAGGACAAAGATGAGGCCATTGGCTTCTGGTAAAG
GTCACTTTGACAGAGGCTTTGTTGTGGATGACAGGTCAATACTTCTTATCAGTTAAAATGTTA
GATTAATTTTAGACGGTAGAAATATTTGGACATTGATGTTGCCATTGACTGCTTCTATTATGT
TATATCCATATTTGAAAAGACCAATTTTCTCTAAAGTTTTTGTATCCACAGTACATTTTGGGT
TTGGCTATTGGTTACCCAGCTATACTGGTTGGGCTTCTATTACAGGTTCAGAAGAACCATT
GGGTGATATTATTAACATGATTGTATTCCAATTTGTTGTTAGTTTTCTTTTGGTGTGTTTACT
TTAATACTGCTTACTCTCATCAAGATTCTGTTGATGATAGAAAGATGAACATTAATTCTGCTTA
CGTTATAGCAGGTCAAAGAATTAGATTGTTTTTGGCTTTTTTGTCTGTCTTGCCATTGTTGAC
TATTCCATACATTATTTCTACAATTAACCTCACCTTGGTTGTGGGTCTCATGGATGGCTACTTG
GACTGTTTCAATTGTTATGCAAATAGCTCAATTTGATTCTCAAAAATTGGAATCTGGTGGTAG
AATTCATCTGGGATAACTTTTTGTTGGGTTTGTGGACAATTGTCGCTTGCATGGTCGAAGTCG
GTTTGCAAAAAGTTGAATTTTGAAGAATGTTGAAGGTTATATTAAATTGTAA

> *Synpcc7942_0338*

ATGGCTACCTATCAAGTCGAAGTCATCTATCAGGGTCAGTCCCAAACCTTCACGGCTGACT

CAGATCAAAGTGTTTTGGACTCAGCGCAAGCTGCGGGCGTTGATCTGCCGGCCTCTTGCT
TAACTGGAGTCTGTACAACCTGTGCCGCGCGGATTCTCAGCGGTGAAGTGGATCAGCCAG
ATGCCATGGGGGTAGGGCCCGAGCCTGCTAAGCAGGGCTACACCTTACTCTGTGTGGCCT
ATCCGCGATCGGACCTGAAGATCGAGACTCATAAGGAAGACGAACTCTACGCCCTGCAATT
TGGTCAGCCCGGCTGA

> *Synpcc7942_0698*

ATGCCGAGCATTTCGCTTTATCCGTGAAGACAAAGAGGTCTTTGCCGCTGATGGGGCCAAC
CTCCGCTTCAAAGCTGTAGAGAACCAAGTGGACCTCTACACCTTCGGCGGCAAGATGATG
AACTGCGGCGGCTACGGCCAATGCGGGACTTGTCATCGTCGAAATTGTTACAGGGCGCTGA
GAATCTCTCGCCCCGTACGTCTTTTGAGGAGCGCAAGCTCAAGCGCAAGCCTGATTCCTAC
CGTCTTGCCGCAAGCGACCGTCAATGGACCCGTGACGGTGTGACCAAACCCAACCCG
AAAGAGGCCCAACGCGGAGACTCTGATCGCTCAGGATTTAGCTCGGCCAATTCCGGTAACA
GCGCCGCTGCACTGCCACAACTGAGACTGAAGTGGCCGGCGATCCACCGTCGATCGC
AACAGCTGAGACCTAA

> *Synpcc7942_0814*

GTGGCGCACACCATTGTTACGAATACCTGCGAAGGGGTCGCTGACTGTGTGCGATGCCTGC
CCGGTTGCTTGCATCCAAGAAGGGCCGGGACGCAACCAAAGGGAACGACTTGGTACTG
GATTGACTTCAGCACCTGCATCGACTGCGGTATTTGCCTGCAAGTCTGTCCGGTTCGAAGGT
GCCATTCTGCCCGAAGAACGTCCCGAGTTGCAGCAAACGCCCTAA

> *Synpcc7942_0898*

ATGACCCTCGCTGAAACCCTGTCCGGTGCGGGTGCGGAGTCTCGGCCTCGATCAGATCGAT
CGCCACTTATTTCTCTGCGCCGACCAAATAAACCGCTTTGTTGCGATCGCGATCGCAGTC
TCGAAAGCTGGGAGTATCTGAAGCGCCGGCTGCGGGAGTTGGATCTCGATCGCCCCGAC
ACTGGCAAGCCTCTGGTGTTCGAACTAAGGCCAACTGTCTGCGCGTCTGTCAAGAAGGG
CCGATTCTTTGGTCTATCCCAGGGCATTGGTACGGTCGCGTCACCCCCGAAGCGATT
GAGCGGATTTGCGAGGAACATTTGCTAGGTGGTCAACCCGTGCAAGAGTTGATTTGTCATC
AACATGCGTTGCCAGCAGTCGATCCGCTATAG

> *Synpcc7942_0978*

ATGTTGAATGCGAGTGTGGCTGGCGGAGCAGCTACCACCACCTATGGCAACCGGCTCTTT
ATCTATGAAGTGATCGGTCTGCGCCAAGCCGAGGGCGAACCGTCCGACAGCTCAATCCGC
CGTAGTGGCAGCACCTTCTTCAAGGTGCCTTACAGCCGGATGAATCAAGAAATGCAACGGA
TTTTGCGCCTTGGCGGCAAATCGTTAGCATCCGGCCTGCGGAGGAAGCAGCCGCGAATA
ATGGTGCGGCTCCTCTACAGGCAGCAGCTGAAGAACCTGCTGCAGCACCAACCCCGCT
CCGGCTGCCAAAAACATTCAGCCGAAGACGTGCCTGTCAATATCTACCGGCCTAACAAAGC
CTTTCGTAGGCAAGGTGCTCTCGAACGAGCCCTTGGTTCAAGAAGGCGGGATTGGTGTG

TGCAGCACCTCACCTTCGATATTCGGAAGGCGATCTGCGCTACATCGAAGGTCAAAGTAT
CGGGATTATCCCGGATGGCACCGATGACAAAGGCAAGCCGCACAAGCTCCGTCTTTACTC
GATCGCATCCACTCGCCACGGCGACCACGTGGATGACAAAACCGTCTCGCTGTGCGTGC
GCCAGCTGCAGTACCAGAACGAAGCCGGCGAAACGATTAATGGCGTCTGCTCGACTTTCC
TCTGTGGTCTGAAGCCAGGCGATGACGTCAAGATCACCGGTCCTGTGGGCAAAGAAATGC
TCCTACCGGCGGACACAGACGCCAACGTGATCATGATGGGTACTGGCACCGGGATTGCTC
CGTTCCGAGCCTACCTATGGCGGATGTTAAAGACAACGAGCGAGCCATCAACAGCGAGTA
TCAATTCAACGGCAAGGCTTGGTTGATCTTCGGGATTCCGACGACCGCCAACATCCTCTAC
AAAGAGGAGCTGGAAGCGCTGCAGGCTCAGTATCCAGATAACTTCCGCCTGACCTACGCG
ATCAGCCGCGAGCAGAAAAATGAAGCGGGCGGCCGGATGTACATCCAAGACCGCGTCCG
TGAACATGCTGACGAGATCTGGAACCTACTCAAGGACGAAAAAACCCACGTCTATATCTGT
GGTTTTCGTGGCATGGAAGATGGGATCGATCAAGCCATGACCGTCGCAGCTGCCAAGGAA
GATGTGGTTTGGTCTGACTACCAACGCACCCTCAAGAAAGCGGGTCGTTGGCATGTTGAA
ACCTACTAG

> *Synpcc7942_1499*

ATGGCAACCTACAAGGTTACGCTCGTCAATGCTGCCGAAGGCTTGAACACCACGATCGAC
GTGGCTGACGATACTACATCTTGGACGCCGCTGAAGAGCAAGGCATTGACCTGCCTTACT
CCTGCCGTGCTGGTGCTTGCTCGACCTGTGCTGGCAAAGTCGTCTCTGGTACCGTCGACC
AATCGGATCAATCCTTCTTGGATGACGACCAAATTGCAGCAGGCTTTGTCTGACCTGCGT
CGCCTATCCGACCTCCGATGTGACGATCGAAACCCACAAAGAAGAAGACCTCTACTAA

> *Synpcc7942_1749*

ATGACGACGCCGGACCGTAGCGGGCTAGAACCAGAACTCGGAGGCAGTCTCCGTCATGG
TCAAGCCCAGCAGCGGCCTAGAGCCAGAGCTTGGGGGTGAGCTGCGTCAGAAATTAGTCT
GGGTAGACGAAGTGACCTGCATCGGCTGCCGCTACTGTTCCCACGTTGCCACCAATACCT
TCTACATCGAGCCGGACTACGGGCGATCGCGAGTGGTTCCGGCAGAATGGTGACCCTGAAG
AGCTCGTTCAAGAGGCGATTGATACCTGTCCCCTGGATTGCATCCACTGGGTCAATCCCAG
CGAACTCCGTCAACTAGAAGCTGAACGTCGCAATCAAGTCATCATGCCGCTGGGCTTTCCC
CAAGAGCGCTCGAAACAACGGCGACGAACTTAA

> *Synpcc7942_2581*

ATGAGCGACACTTACACCGTCCGCATTCGCGATCGCCGTACCGATGAGGAGTTTACGGTG
CAGGTTCCGCCCAGTCGCTACATCCTGCAAACCGCCGAAGAGCAAGGTTATGAGTTACCG
TTCTCCTGCCGCAATGGGGCTTGACAGCCTGTGCAGTGCAGTTCGCGGTTCTCGGCGGCGGAT
CGAGCAAACCGAGGCCATGGGACTCTCAGCTCCGCTACGCCAGCGCGGTTATGCCCTGC
TCTGCGTTAGCTATCCGCGATCGGATGTGATTGTGGAGACCCAAGACGAAGACGAAGTCTA

CATGCTGCAGTTTGGTCGCTATTTGGCCAAGGAAAAGTCAGCTTTGGCCTGCCCTTGGAC
GAGGAATAA

>CYP51

ATGACCGTCGAGTCCGTCAACCCCGAAACCCGCGCACCCGCGGCACCGGGAGCACCGG
AGCTGCGCGAGCCGCCCGTCGCGGGCGGTGGCGTTCCGCTGCTCGGCCACGGCTGGC
GGCTGGCCCGCGACCCGCTGGCCTTCATGTCCCAGCTGCGCGACCACGGCGACGTCGT
GCGCATCAAGCTCGGCCCGAAGACCGTCTACGCCGTCACGAACCCCGAGCTGACCGGCG
CCCTCGCCCTGAACCCCGACTACCACATCGCCGGTCCGCTGTGGGAGTCGCTGGAGGGC
CTGCTCGGCAAGGAGGGCGTGGCGACCGCCAACGGCCCGCTCCACCGCCGCCAGCGGC
GCACCATCCAGCCGGCGTTCCGGCTCGACGCCATCCCCGCTACGGGCCGATCATGGAG
GAGGAGGCGCACGCGCTACCGAGCGCTGGCAGCCGGGGAAGACCGTCGACGCCACCT
CGGAGTCCTTCCGGGTCGCCGTGCGCGTCGCGGCCCGCTGCCTGCTGCGCGGGCAGTA
CATGGACGAGCGGGCCGAGCGGCTCTGCGTCGCGCTCGCCACCGTCTTCCGGGGCATG
TACCGGCGGATGGTGGTCCCGCTCGGGCCGCTGTACCGGCTGCCGCTTCCGGCCAACCG
CCGATTCAACGACGCGTTGGCCGATCTGCACCTGCTGGTCGACGAGATCATCGCCGAGC
GCCGTGCATCCGGTCAAAGCCGGACGATTTGCTCACGGCATTGCTGGAGGCGAAGGAC
GACAATGGCGACCCGATCGGGGAACAGGAGATCCACGACCAGGTGGTCGCGATACTCAC
CCCGGGCAGTGAAACCATCGCGTCCACGATCATGTGGTTGCTCCAGGCACTTGCCGACCA
TCCGGAACATGCCGACCGCATAACGCGACGAAGTCGAAGCGGTCACCGGCCGGACGTCCCG
TGGCATTTCGAGGACGTCCGGAAGCTCAGGCACACCGGCAATGTCATCGTGGAGGCCATG
CGTTTTCGCTCCCGCCGTATGGGTATTGACCCGGCGCGCGGTGGCCGAGAGCGAACTCGG
TGGCTATCGCATTCCGGCCGGTGC GGACATCATCTACAGTCCGTACGCAATCCAACGCGAT
CCGAAGTCGTACGACGACAACCTGGAGTTCGATCCCGACCGGTGGCTACCGGAACGCGC
GGCGAATGTGCCGAAATACGCCATGAAGCCGTTTCAGCGCGGGCAAGCGGAAGTGCCCA
GTGACCACTTTTCGATGGCCAGCTCACGCTGATCACGGCCGCGCTCGCCACGAAGTACC
GCTTCGAACAGGTGGCGGGCTCGAACGACGCGGTCCGGGTCGGCATCACGCTGCGTCC
GCACGACCTGCTGGTCAGGCCCGTGGCGCGGTGA

>CYP102B1

ATGGCCCAGACAGCGAGGGAACCGGCCCGGGACGGACTGCCGAAGGGCTTCCGCAGCG
CCGAGCTGGGCTGGCCCGAACTGCACCGCATCCCGCACCCCCCGTACCGGCTGCCCTG
CTCGGAGACGTCGTGGGAGCGAGCAGGCGCACCCCGATGCAGGACTCGCTGCGGTACG
CGCGACGACTGGGGCCGATATTCCGGCGGGCGGGCCTTCGGCAAGGAGTTCGTGTTTCGTG
TGGGGTGCCGCCCTCGCCGCCGACCTGGCGGACGAGGCGCGGTTCCGCAAGCACGTGG
GCCTGGGGGTGGCCAACCTGCGTCCGGTGGCCGGGGACGGGCTGTTACGGCGTACAA
CCACGAGCCCAACTGGCAGCTCGCGCACGACGTGCTGGCGCCCGGCTTCAGCCGCGAG

GCCATGGCGGGGTACCACGTGATGATGCTGGACGTGGCCGCGCGGCTCACCGGCCACT
GGGACCTGGCCGAGGCGTCGGGCCGGGCGGTGGACGTGCCGGGCGACATGACCAAGC
TGACGCTGGAGACCATCGCGCGCACCGGTTTCGGCCACGACTTCGGCTCCTTCGAACGC
TCCCGCCTCCACCCCTTCGTCACCGCGATGGTGGGCACGCTCGGCTACGCGCAGCGCCT
CAACACCGTGCCCGCGCCGCTCGCCCCCTGGCTGCTGCGTGACGCGAGCCGCCGCAAC
GCCGCCGACATCGCCACCTCAACCGCACGGTCGACGACCTCGTCCGTGAACGCCGGGC
GAACGGCGGCACGGGCGGCGGAACCGGCAGCGGCAGCGGCAGCGGCACCTGCTCGA
CCGGATGCTGGAGACGGCCCATCCAGGACCGGTGAGCGGCTGTCGCCCCAGAACGTC
CGCCGGCAGGTCATCACCTTCCTGGTCGCCGGTCACGAGACCACCTCGGGCGCGCTCTC
CTTCGCCCTGCACTACCTCGCCAGCACCCCGACGTGCGGGCCCGGGCCCGCGCCGAG
GTGGACCGGGTGTGGGGCGACACCGAGGCACCCGGCTACGAGCAGGTGGCCAAACTGC
GCTACGTGCGCCGGGTGCTTGACGAGTCGCTGCGGCTGTGGCCGACCGCGCCGGGATT
CGCGCGCGAGGCCCGCGAGGACACGGTGCTCGGCGGTACGCATCCGATGCGCCGCGGG
GCCTGGGCGCTGGTGCTACGGGCATGCTGCACCGCGATCCCAGGTTGTGGGGCGCGG
ACGCCGAGCGGTTTCGACCCGGACCGCTTCGACGCGAAGGCCGTGCGGTCGCGCGCCCC
GCACACGTTCAAGCCGTTTCGGGACCGGGGCGCGGGCGTGATCGGGCGGCAGTTCGCG
CTGCACGAGGCCACGCTGGTCTTCGGGCTGCTGCTGCGCCGCTACGAGCTGCGGCCGG
AGCCCGGCTACCGGCTGCGGGTGACCGAGCGGCTGACGCTGATGCCGGAGGGGCTGCG
CCTGCACCTGGTCCGGCGGACCGCGGCCCGCGCCGGTCCCGCACGGCCGC
CCCGGGGGCGGCCGACGATGCCGGGGACACCGTCTCAGCGCCCGGCTGTCCAGTGAC
CGGGCGGGTGACTGA

>CYP105D5

ATGACGGACACCGACACGACGACGAACACCGACACGACGACGAACACCCATCCGGCAGC
CCCCGTCGCCTTCCCGCAGGACCGGACCTGTCCCTACCACCCGCCCGCCGCCTACGACC
CGCTGCGCGCCGCGCGTCCACTGGCCCGCATCACGCTCTTCGACGGCCGCCCGGCCTG
GCTGGTGACCGGGCACGCCGCCGCGCGCCGACTGCTGGCCGACCAGCGCCTGTCGACC
GACCGCACCCGCGACGGCTTCCCCGCCACGTCGGCGCGGCTGGCCGCCGTCCGCGAGC
GCAGGACCGCCCTGCTCGGCGTGGACGACCCCGAGCACCGCGCCAGCGCAGGATGGT
GCTCCCGGAGTTCACCCTGAAACGCGCCGGCGCACTGCGCCCAGCATCCGGCGGATC
GTCGGCGAACGGCTCGACGCGATGATCGCGCAGGGCCCGCCCGCCGACCTGGTGACGG
CCTTCGCGCTGCCCGTGCCCTCGATGGTGATCTGCGCGCTGCTCGGCGTCCCCTACGCC
GACCACGAGTTCTTCGAGGAACAGTCCCGCCGGTGTGTCGCGGGCCCCCTGCCCGCCG
ACACCAGGGACGCCCGCGACCGACTGGAGGCGTACCTGGGGGAGTTGATCGACCGCAA
GCGGCGGGCGCCCGGCGAGGGCCTGCTGGACGACCTGGTCCGGCGACAGGCGAGCGA

GGGCGCGACCGACCGCGAACAACCTGATCGCCTTCGCGGTGATCCTGCTGGTCGCGGGAC
ACGAGACCACCGCCAACATGATCTCGCTGGGCACCTACACCCTCCTGACCAACCCCGGGC
GGCTGGCCGAACCTGCGTGCGGACCCCGCGCTGCTGCCCGGCGCCGTGGAGGAGCTGAT
GCGCGTCCTGTGATCGCCGACGGGCTGCTGCGGATGGCCACCGAGGACATCGACGTG
GACGGGCAGACCATCCGGGCCGCGACGGGGTCTTCTCCACCTCCGTCATCAACCG
CGACGAGAGCGTCTACCCCGAACCGGACGCCCTCGACTGGCACCGCCCCGCCCGCCAC
CACGTGCGCTTCGGGTTCCGGCATCCACCAGTGCCTCGGTCAGAACCTGGCCCGTGCCGA
ACTGGAGATCGCCCTGGAGAGCCTCTTCGACCGGCTGCCACGCTGCGCCTGGCCGCTC
CGGCGGACGAGATCCCCTTCAAACCCGGCGACACGATCCAGGGGATGCTGGAACTCCCC
GTGGCCTGGTAG

>CYP105N1

ATGACACCCCCGAATCCCCGACGGCCTCCACACCCCCGGCGCCACCCCGCCCCGCG
ACTTCCCGATCCAGCGCGGCTGCCCTTCGCCGCGCCCGCGGAGTACGCCGCGCTGCG
CACCGACGACCCGGTCGCCCGCGTCACCCTGCCGACCCGGCGGGGAAGCCTGGGTGGTG
ACCCGCTACGACGACGTACGGGAACTGCTCTCCGACCCGCGCGTGAGCGCCGACATCCG
CCGCCCGGGCTTCCCCGCCCTCGGAGAGGGCGAGCAGGAGGCCGGGGCCAGGTTCCG
TCCCTTCATCCGCACCGACGCCCCGAGCACACCCGCTACCGGGCGCATGCTCCTGCCGG
CGTTCACCGTGCGCCGGGTCCGCGCGATGCGGCCCGCCGTACAGGCACGGGTGGACGA
GATCCTGGACGGCATGCTGGCCGCCGGCGGCCCGGTGGACCTCGTGAGCGCGTACGCC
AACGCCGTCTCCACGTCCGGTATCTGCGAACTGCTCGGCATCCCGCGGCACGACCTGGA
GTTCTTCCGGGACGTCACCCGGATCTCCGGCTCCCGCAACAGCACCGCCGAGCAGGTCT
CCGAGGCCCTCGGCGGCCTGTTCCGGCCTGCTCGGCGGGCTGGTCGCCGAGCGCCGGG
AGGAACCCCGCGACGACCTGATCTCGAAGCTCGTCACCGACCACCTGGTCCCCGGCAAC
GTGACGACGGAGCAGCTGCTGTCCACCCTCGGCATCACCATCAACGCCGGCCGCGAGAC
CACCACCAGCATGATCGCCCTGAGTACCCTGCTCCTCCTGGACCGTCCGGAGCTGCCGG
CGGAACTGCGCAAGGACCCCGACCTGATGCCCGCGGCCGTCGACGAACTCCTGCGCGT
GCTGTCCGTCGCCGACTCGATTCCGCTGCGCGTCGCCGCCGAGGACATCGAGCTGTCCG
GCCGACCGTCCCCGCCGACGACGGCGTCATCGCGCTCCTGGCCGGTGCCAACCACGA
CCCCGAGCAGTTCGACGACCCCGAGCGGGTCGACTTCCACCGCACGGACAACCACCAG
TGGCCTTCGGCTACGGAGTGCACCAGTGCCTCGGCCAGCACCTGGCGCGGCTGGAGCT
GGAGGTCGCCCTGGAGACACTCCTCCGCCGCGTCCCACGCTGCGCCTCGCGGGGGAG
CGGGACCAGGTCGTCGTCAGCACGACTCGGCCACCTTCGGCCTCGAAGAACTGATGGT
GACCTGGTGA

>CYP107P1

ATGACGGCTGCAACCGATGGTCCCATGTGTCCGGTCCCGCGTTCGACCCGTGGGACCC

CGCCTTCGTCGCCGACCCCTACCCCGCCTTCGCCGAGCTGCGGGCCCGGGGGCGCGTG
CTCTACTACGAGCCGAGCGACCAGTGGCTGGTCCCACACCACGCGGACGTCTCCGCGCT
GCTGCGGACCGCCGCCTCGGCCGGACCTACCAGCACCGTTTCACGCACGAGGACTTCG
GCCGCACGCCGCCCGCCGGAGCAGGAGCCCTTCCACACGCTCAACGACCACGGCAT
GCTCGACCTGGAGCCGCCGACCACACCCGATCCGGCGCCTGGTGTCAAGGCGTTCA
CGCCGCGCACCGTGGAGCGGCTGAAGCCGTACGTGCACGGCCTGGCGGACGACCTGGT
GGCCCGGCTGGTGGCGGCGGGCGGCGGCGATCTGCTCACCGACGTGGCCGAGCCGCT
GCCGGTCGCCGTGATCGCCGAGATGCTGGGCATCCCGGAGTCCGACCGGGCCCCGCTG
CGCCCCTGGTCGGCGGAGATCTGCGGGATGTACGAGCTGAACCCGTCCGAGGAGACGG
CGGCGAAGGCGGTCCGGGCCTCGCTCGACTTCTCGGACTACCTGCGCGCGCTGATCGCG
GCCCGCCGCAAGGAGCCCGGCGACGACCTGATCTCCGGTCTGATCGCGGCCACGACG
AGGACGACCGCCTCACCGAGCAGGAGATGATCTCGACCTGCGTCCTGCTCCTGAACGCC
GGTCACGAGGCCACGGTGAACGCCACCACCAACGGCTGGCTGGCGCTCTTCCGCCACCC
CGACCAGCTGGCCGCCCTGCGCGCCGACCACTCCCTCGTCCCGTCGGCCGTGGAGGAG
CTGATGCGCTACGACACCCCGCTCCAGCTTTTCGAACGTTGGGTCTGGACGAGATCGAG
ATCGACGGCACGACCCTCCCGCGCGGCGCGGAGGTGGCGATGCTGTTCCGGCTCGGCCA
ACCACGACCCGGCGGTCTTACCGACCCGGAACGCCTCGACCTACCCGCCGCGACAAC
CCGCACATCTCCTTCAGCGCCGGCATCCACTACTGCATCGGCGCCCCGCTGGCCCCGCATC
GAACTCGCGGCGTCCATGACGTCCCTGCTGAAGCGGGCGCCGGGCCTGCGCCTGGCGG
CGGAACCGGAACGCAGGCCGAACCTTCGTGATGCGCGGCCTGACGGAGCTACGCGTGAA
CTCTGA

>CYP107T1

ATGGGGAGTGCTCCCCGTTTAATTGTCTCTCTCACCGGAGGTCCCATGACCGCGCCCACC
TATGAGGAACTGGCCGCCCTGCGGGCCGTGGGCGCCGTGCACCGGGTGTTCGTGCCGG
GCAGCGGGGAGAGCCGGCTGGTGGTGACGCGCGACGCGGCACGGGCCGCGCTCACCG
ACCCGCGGCTGCGCAACGACATCCGGCACTCCGCTTCTGGGACAGCGACGGCGGGCA
CGCCATCGGCCACAACATGCTCCAGAGCGACCCGCCGAGCACACCCGTCTGCGCCGGC
TGGTCGCCGGACACTTACGCCGGGACGGACCGCCGCCCTGCGCCC GCGCGTCGAGCG
CATCGCGCACGACCTGCTGGACGCGCTGCCGCCGGCGGGCACGGCCGACCTGGTCCG
CGGTACGCGCTGCCGCTGCCGGTACGGTGATCTGCGAGCTGCTGGGCGTCCCGGAGA
GCGACCGCGGAACCTTCCACACCTGGTTCGAACGAGCTGGTCATGCCACTTCGCCCGAG
GCGGCCGGGTCCGGCGGCGACCGCGCTGACCGGCTACCTACCGAGCTGACCGACGCCA
AGCGCCGACCCCGGACGGCACGCTGCTCGGCGATCTGGTCGCCGCCGCGCCGACAGCGG
CGAACTACCCCGGCGAACTGCTCGGCATGGCCTTCTGATCCTGGTCCGCGGGGCACG

AGACGACCGTCAACCTGATCTCCGCCACCGTGACGGCCTGCTCACCCACCCCGGCCAG
CTCGCCCGCCTGCGCGCCGAGCCGGAGCTGACCGAGGCCCGCGTGAGGAGTCGCTGC
GGTACCACTCGCCCGTCCACGCCTCCGCGTTCCGTTTCGCGGCCGAGCCGTTGGAGCTG
GCGGGCACCGCGATCGCGGGCGGGCGACCCGGTGCTCGTCTCCCTCGCCGCCGCTCCC
GCGACCCGGCGCACTTCCCCGACCCCGACCGCTTCGACATCGGCCGCCGCCCGCGGGG
CCACCTCGGCTTCGGCCACGGCCCGCACCACTGCCTGGGCGCCCCGCTGGCCCGCGTC
GAGGCCGCCGTCGCCGTACGCCTGCTGCTCGACCGGCATCCCGCGCTCGCGCTCGCCG
CCGGCCCGGCGACGCTCACCTGGCGCACCAAGCACCTGCTGCGGGGGCTGGTGGAACT
GCCGGTGCGGCTCGGCTGA

>CYP107U1

ATGACCGGCAGCTCTTCCGCCCGGTCCCCGAGCTGTTAGCTGGGAGTTCGCGAGCGA
TCCGTACCCCGCCTACGCGTGGCTGCGCGAGCACGCCCCGTGCACCGCACCCGGTTGC
CCAGCGGAGTGGAGGCGTGGCTGGTCACCCGGTACGCCGACGCGAAACAGGCGCTCGC
CGACCCGCGGCTGTCCAAGAACCCGGGCCACCACGACGAGCCCGCGCACGCCAAGGGC
AAGACCGGCATCCCGGGGGAGCGCAAGGCCGAGCTGATGACCCATCTGCTCAATATCGA
CCCGCCCGACCACACCAGGCTGCGGCGGCTGGTCAGCAAGGCGTTCACACCGCGCCGC
GTGGCCGAGTTCGCGCCCCGGGTGCAGGAGCTGGCCGACGGTCTCATCGACCGGTTGC
CGGACACCGGCTCCGCCGACCTGATCCACGATTCGCCTTCCCGCTGCCCATCTACGCCA
TCTGCGACCTGCTCGGCGTCCCCGCGAGGACCAGGACGACTTCCGGGACTGGGCGGG
CATGATGATCCGTCACCAAGGGGGTCCGCGCGGGCGGCGTAGCCCGGTGCGTGAAGAAGA
TGCGCGGCTACCTCGCCGACCTCATCCACCGCAAGCGCGCCGCGCTGCCACCCGAGCCC
GCCCCGGCGAGGACCTCATCTCCGGCCTGATCCGCGCCTCCGACCACGGTGAGCACCT
CACCGAGAACGAGGCCGCGGCCATGGCCTTCATCCTGCTCTTCGAGGGTTCGAGACCA
CGGTCAACCTCGTGGGCAACGGCACGTACGCCCTGCTCACCCATCCCGAGCAGCGGGAG
CGGCTGCAGACCTCGCTGGCCGCCGGGGAGCGCGGCCTGCTGGAGACCGGCGTCGAG
GAACTCCTGCGGTACGACGGCCCGGTGGAGCTGGCCACCTGGCGCTTCGCCACCCGGC
CGCTGACCATCGGCGGCCAGGACGTCGCGGCCGGCGACCCGGTCCTCGTCGTGCTGGC
CGCCGCCGACCGGGACCCGGAGCGGTTACCGACCCGGACACCCTGGACCTCGCCCGC
CGCGACAGCCAGCACCTCGGGTACGGCCACGGCATCCACTACTGCCTCGGCGCCCCGCT
CGCCCGGCTGGAGGGGCAGACGGCGCTGGCCACGCTGCTGACCCGGCTGCCGGACCTC
CGCCTCGCCGCTGACCCGGCCGAACTCCGGTGGCGCGGGCGGCTCATCATGCGAGGTC
TGCGCACCTTGCCCGTCTCGTTACACCGCCCCGCGTCATCGGCCGGCAACGGTCCGTCG
CCGACGCAAAAGTGA

>CYP154A1

ATGGCGACCCAGCAGCCCGCCCTCGTCCTCGACCCACCGGCGCCGACCACCACACCGA

GCACCGCACCTGCGGGAGGGCGGCCCGGCCACCTGGGTGGACGTCCTCGGGGTGCA
GGCCTGGTCGGTCAGCGACCCCGTCTCTCAAGCAGCTGCTCACCAGCTCCGACGTCT
CCAAGGACGCCCCGGGCGCACTGGCCCCGCTTCGGGGAGGTCGTCGGCACCTGGCCGCT
GGCCCTGTGGGTGGCGGTGGAGAACATGTTACCGCGTACGGGCCAATCACCGCAAGC
TGCGCCGGCTGGTGGCGCCCGCCTTCAGCGCCCGTTCGCGTCGACGCGATGCGGCCGGC
CGTCGAGGCGATGGTGACCGGTCTCGTCGACCGGCTCGCCGAGCTTCCCGCCGGTGAG
CCGGTGGACCTGCGGCAGGAGCTGGCCTACCCGCTGCCATCGCGGTGATCGGTACCT
CATGGGCGTGCTCAGGACCGGCGCGACGGCTTCCGCGCCCTCGTGGACGGCGTCTTC
GACACCACCTGGACCAGGCCGAGGCCAGGCCAACACCGCGCGCCTGTACGAGGTCC
TCGACCAGCTCATCGCGGCCAAGCGCGCCACCCGGGCGACGACATGACCTCGCTGCTC
ATAGCCGCGCGGGACGACGAGGGGGACGGCGACCGGCTCTCCCCGAGGAGCTGCGC
GACACCCTGCTGCTGATGATCAGCGCCGGGTACGAGACCACCGTCAACGTCATCGACCAG
GCCGTGCACACCCTGCTGACCCGCCCGACCAGCTCGCGCTGGTCCGCAAGGGCGAGG
TCACCTGGGCGGACGTGGTGGAGGAGACGCTGCGCCACGAACCGGCGGTCAAGCACCT
GCCGCTGCGGTACGCGGTACCGACATCGCCCTGCCGGACGGGCGGACCATCGCCCGC
GGGGAGCCGATCCTCGCCTCGTACGCCGCCGCAACCGCCATCCGGACTGGCACGAGG
ACGCCGACACCTTCGACGCGACCCGCACCGTCAAGGAGCACCTGGCCTTCGGCCACGGC
GTCCACTTCTGCCTGGGCGCGCCGCTGGCCCGCATGGAGGTCACCCTCGCGCTGGAGA
GTCTCTTCGGCCGCTTCCCGGACCTCCGCTCGCCGATCCGGCCGAGGAGCTGCCGCC
GTGCCCTCCCTGATCAGCAACGGCCACCAGCGGCTCCCGTCTGCTGCACGCCGGCTG

A

>CYP154C1

ATGACGACCGGCACCGAAGAAGCCCGGATCCCGCTGGACCCGTTCTCACCAGCTGGA
CGGCGAGAGCGCTCGGCTGCGTGCGGCCGGACCGCTCGCCGCCGTGGAGCTGCCGGG
CGGCGTGCCCGTCTGGGCGGTACCCACCACGCGGAGGCCAAGGCCCTGCTCACGGAC
CCGCGGCTGGTCAAGGACATCAACGTGTGGGGCGCCTGGCGGCGCGGCGAGATCCCCG
CCGACTGGCCGCTGATCGGGCTGGCCAACCCGGGCGCTCGATGCTCACGGTGGACGG
CGCCGAGACCCGCCGGCTGCGCACCTGGTGGCGCAGGCGCTCACGGTGCGCCGGGT
GGAGCATATGCGGGGGCGGATACCGAGCTGACCGACCGCTGCTCGACGAACTGCC
GCGGACGGCGGCGTTCGTGGATCTCAAGGCGGCCTTCGCTATCCGCTGCCGATGTACGT
CGTCGCCGACCTGATGGGCATCGAGGAGGCGCGGCTGCCGCGGCTGAAGGTGCTGTT
GAGAAGTTCTTCGACGCAGACCCCGCGGAGGAGGTCGTGGCGACCCTCACCGAGCT
GGCCTCGATCATGACCGACACGGTCGCCGCGAAACGCGCCGCACCCGGGCGACGACCTG
ACGAGCGCGCTGATCCAGGCGTCCGAGAACGGCGATCACCTCACCGACGCGGAGATCGT

CTCGACCCTCCAGCTGATGGTGGCCGCCGGTCACGAGACGACGATCTCGCTCATCGTGA
ACGCCGTGGTCAACCTGTCCACCCATCCCAGCAGCGGGCGCTGGTGTCTCCGGCGAG
GCGGAGTGGTCCGCGGTGGTCGAGGAGACGCTGCGCTTCTCGACGCCACCTCGCACG
TCCTGATCCGGTTCGCGGCCGAGGACGTCCCGGTGCGCGACCGGGTGATAACCCGCCGG
GGACGCGCTCATCGTCTCGTACGGCGCGCTGGGCCGCGACGAGCGCGCCCACGGTCCG
ACGGCGGACCGGTTGACCTACCCGCACGTCCGGGAACCGGCACATCTCCTTCGGCCA
CGGCCCGCACGTGTGCCCGGTGCGGCCCTGTCCCGGATGGAGGCGGGCGTGGCGCTC
CCGGCCCTGTACGCCCGCTTCCCGCACCTGGACCTGGCGGTGCCGGCCGCGGAGCTGC
GCAACAAGCCGGTGGTCACGCAGAACGACCTGTTTGAAGTCCGGTTCAGGCTGGCCTGA
>**CYP155A1**

GTGGCGGATCCTGGTGACCACGTGTCGAGCCGGGTGGCCGCCGGGGCCGCACGGCC
GGAGTGGGGGGTGGCCACGTGGTGCAGGACGGACGGGTCTCCGCGCGGAAGGCCGAC
CGGGGCCGGGCCGGGCCGGCGTGCCCGGTGACCGGGCGGCGGACGGGACCTGGCG
GGTGCACGATTTCCCGTCCGCGCGGGCGCTGCTGCGCGGTCCCGGCACCGTACAGGCC
GGTCTCGGCATCGAGACGGTGGAGAAGCTCCCGCCGCGCGTCCGCCGGCCCGTGCTCT
ACCGGGACGGCCCCGAGCACCGGGAGCACCGCAGGCAGACCGCCCGCTACTTCACTCC
CCGCCGGGTGGACGAGCACTACCGCGAGCCGATGGTCCGGATCGCCGAGGAGCAGCTG
GCCGTTCTGCGCTCGGCCGGCGAGGCCCGCTCTCCGACCTCGCCTTCGGTCTGGCCGT
CGGCGTGGTGAGCGAGGTCGTCGGGCTGCGGTACAGCAGGCCCGGCATCCGCCGCAGG
CTGGAGCGGTTCTTCCCGGAGGAGTTCGGCGAGCCGGGACTGACCAGTGTCCGGGGTCT
GTA CTGGCTGGTGC GGCAGAACACCAACTGGCTGCGCATCCACCTGGCCGACGTGCGCC
CCGCCGTCCGCGCGCACCGCCGCCGCGAGCACGACGACCTGATCTCCACCTGATCGC
GGAGGGCTGCTCGGACGTCGAGATCCTCGGCGAGTGCCTGACCTTCGCCGCGGGCGGGC
ATGGTCACCACGCGCGAGTTCGTGTGCCTGGCCGCCTGGCACCTGTTCTCGGACGCCGA
GCTGCTCGGCCACTACCGGTGCGCGGACGAGACCGGGCGGCTGGCCGTGCTCCAGGAG
CTGCTGCGGCTGGAGCCCGTCATCGGGAGCCTGCGCAGGCGGGCGACCGGGCCGGTG
GAGCTGTCTGCGCGACGGGCCGGTGACGGTCCGCCCCGGCGAGTACGTGAGGTCC
ACCTGGACGACGCCAACGCGGACCCGAAGGCCGTGGGCGAGGAGCCGCTGCTCGTCCG
CCCGGAGCGGGCCGGGGCGGTGCGCGCGGGGCTCTCCTTCGGCGACGGGCCGCACCG
GTGCCCCGGGGCGCACATCGCCCTGCTGGAGACCGACGTGTTCTCAGCCGCCTGTTTCG
CCCTCGACGGCGTCCGGATGAGCGGTGGGCCGCGCGTCCGCTTCCAGGAGGCCATCGA
CGGCTACGAGATCCGCGACCTGACCGTGGCGCTCCCCGGGCCGGCCGCGGCTGA
>**CYP156A1**

ATGACGTTGCCCTCCACCGAGACCGCGCCGACGGGTGAACCGGGCCGGATCGCCCTGTA
CGCGCCGGAGTTCGCCGCCGACCCGCACGCCGCTACCGGAGCATGCGCCGCACCCAC

GGTCCGCTCGTCCCCGTGGACCTGGCGCCGGGGTCCCGGCGACCCTGGTGATCGGCT
ACTACCAGGCCCGCCGCATCCTGAACGACCCGCTGCGCTTCCCGGCCGACCCGCGGGC
CTGGGAGAAGCTGATCCCGGCGACCTGCCCGGTGCGGCCGATGATGGAGTGGCGGGCC
AACGCGCTGCGCTCGGGCGGCGCCGAGCACACCCGCTACCGGTCCGCGAACACCCACG
CCATCGACCAGGTGGACCAGCACGGGCTGCGCGCCCTCGTCGAGCAGGTGCGCTCCGA
CGCCATCGAGGGTTTCCGCACCCGCGGCTCGGCGGACCTGCTCACCCAGTACTCCTTCC
CCATCGCCTTCCGCGTGCTGAGCGCGCTGCTCGGCTGCCCGGACGAGATCGGGCAGCG
CATCGCCGACGGCATGGCGAAGATCTTCGACACCACCAACGCCGACCAGGGCAACCTGA
TCCTCGCGCAGGCCGTGTCCGACCTGGTGACCCTGCGCCGGACCCACCCCGGCGACGA
CATCACGTCCCGGCTGGCCCTGCACCCCGTCCGGCTGACCGACGAGGAGATGAGCCACC
AGCTCGTCACGCTCTACGGCGCCGGGATCGAACCCATGACCAACCTGATCAGCAACACGA
TCCTGAAGATCCTCACCGACGAGGAGTTCTCCGCGGACCTGCACGCCGGGCTCTCCACG
GTGCGCGACGCGCTCGACGCCGTCTCTACACCGACCCGCCGATGGCGAACTACTGCAT
CTCCTATCCCCGTACCCCGTCGACGTGGAGGGCGTGCTGCTGCCTGCCGACCAGCCGG
TGGTGATCTCGATGGCGGCCGCGAACAACGACCCGGCCCTCACCGAGGGCGTGCCCGC
CGGACAGCACGGCGGCAACCGCGCCCACCTGGCCTGGAGCACCGGTCCGCACACCTGC
CCCGCCCGCTCGCACGCCTACCTCATCGCCGAGACCGCGGTCACCCACCTCCTGGACGC
GCTCCCCGAGACGGACCTCGCCCGGCCCGCAGCCGAACTGGTGTGGCGCCCCGGCCCCG
TTCCACCGCGCCCTCGAATCCCTGCCCGTCACCTTCCCCGCCGCGCAATCCGCCGCACA
CTAA

>CYP156B1

ATGGACGCCACCACCCCGCGGTGCCCCCGCCGGATGCCCGCCCACGCCGACGCGC
GGATAACCGCTGTACGGGCCGGACTTCGCGGCCGACCCGCACGCCTACTACGACTACGCA
CGGTCATACGGCCCCTCCGCACCCGTCGAACTCGCCCCGGGTGTCGACGCCTCCCTCGT
CACCGACTACGCGACAGCACTGCGCCTGCTACAGGACAACGGCACCTTCCGAAAGGACG
CGCGCCGCTGGAAGGCCTTCAACGAGGGCCTGATCCCGGCGGACAGCCCCGTCGTCCC
GCTGCTCGCCTACCGCCCCAACGCCATGTTTCAGCGACGGCGCCGAACACCTGCGGCTGC
GCCAGGCGATCACCGACGCCATGGCCCGCATCGACACCGCCCGGCTGGCCCGCAGCAC
CGAACAGATCTCGGACTACCTGATCTCCAGTTCGGTTCCCGCGGCTCCGCCGACCTGAT
GGCCGACTACGCCAAGCAGCTCCCGCTGTTTCGTCTTCAACGAACTCTTCGGCTGCCCCGC
CGACATCGGCGACCGCATCCTGTTCCGGCATTTCGGCATGTTTCGACGGCGTCAACGCGGA
ACGGGCCGCCGAGGTCTTCTCGGTGCGGTGCGGTGAGCTGGTGGCTCTCAAGCGCAGC
CGCCCCGGCGAGGACGTGACCTCGTGGCTGATGCAGCACGAGACCAGGATGACCGACG
AGGAGATGGTCTATCAGCTCTCACTGATCCTGGGCGCGGGTGCCGATCCGCTGCGCAACC

TGATCGGCAACACCCTGCACCGCATCCTCATCCACGACGAGTACGCCCGCCAGGGCGGA
CTGATCGACGAGGCGATGGAGGACACCCTGTGGGAGAACCCCTCCGGTGCCCAATCTCGC
GCCCCACTATCCTGCCGCTGACGTGGAGTTCGCCGGGCAGAAAGTTCGAGGCGGGCGAGC
TGATCATGGTGAGTTTCGCCGCCGCGAACAAACAGCCCAGCCTGGCCGCCGCCCGGCAG
TCCGGCAGCAACCGTTCCACCTGGCCTGGAGCGCGGGCCCCGCACGCCTGCCCTCGA
AGGACCCCGCCCGCCAGATCACCATGGCCGCCGTCGAGAACCTCCTCAACCGGATCCCC
GACATCGCGGCCGCCGTGCCCGAGGACAGCCTGACGTGGCGCCCCGGGCCCTTACCC
GCGGCCTACCCGCGCTCCCGGCCCGCTTACCCCGCTCCCGGCCCCCGACCGGACGCC
TGCCCCTGCTCAGGCACCCGCAGGCGAGCGTGCCGAACAGTCCGGTACCGCGCGCAAG
GCCGCGGGCACCGGCCGTTGGAGCCAGTTCCTCAACTGGCTCACGAGGTGA

>CYP157A1

GTGAGCACCGACGCCACGACGTCCCGGGCGCCGTACCGCTCGGGCGGCCCCCGCTTCC
AGACCGATCCCGCCCTGCTGTACCGGCAGATGCGGCGCGAACACGGCGCGGTACGCC
GGTCGTGCTGGACGGCGACGTACCGGCCTGGCTGGTGCTCGGGTACCGCGAGCTGCAC
CAGGTGACCGGCGACCCGGTGCTGTTACGCCGGGACTCGGACCTGTGGAACCAGTGGG
AGAACATCCCGGACGACTGGCCGCTGCTGCCGATGATCGGCCGCAGGCAGCCGTCGATC
CTGTACACGGTCGGCGAGCGGCACCGCGAGCGCGCCCATGATCAGCGACGCGCTGG
AGGCCGTCGACCCGCACCTGCTGCGCGGTACGCCGAGCGGTTCCGCCGACGAACTCGT
CGACCGGCTCTGCGCGAAGGGCGAGGCCGACCTCGTCGGCGACTACCGGATGCTGCTG
CCGGTCCGGGTCCTGGCCCGTCTTTACGGGTTCCCGACGAGCAGGGTCCGGCACTGGT
CACCGCGCTCAACGACATGATCGACGGCAGGGAGCGGGCCCTGGCGGGCCAGACGCAC
CTCGGCACCTCCATGGCGCGGCTGCTGGCGGACCGGAAGGCGGGCGCCCGCCGACGACG
TCGCTCCCGGATGCTGGCCGACGAGAGCGGCTTACCGAGGAGGAGGTCGCCCAGGA
CCTGATGGTGATGATGGCCGCGGGCCACCAGCCGACGGCGGACTGGATCGGCAACTCGC
TGCGCCTGATGCTGACCGACGACCGCTTCGCCGCTCCCTCTTCGGCGGGCGCAACAGT
GTCGCCGAGGCCATGAACGAGGTGCTGTGGGAGGACACGCCACGCAGAACGTGGCCG
GCCGCTGGGCGCCCGGGACACCCAGCTCGGCGGACGCCGCATCCGGGCGGGCGATCT
GGTGCTGCTCGGGCTGCAGGGCGCCAACTCCGACCCGCAGGTCCGCACCGACGGCTCC
GCGCTACCGGCGGCAACAACGCGCACTTCTCCTTCGGCCACGGCGAGCACCGCTGCC
GTTCCCGCGCAGGAGGTCGCCGAGGTGATCGCGCGCACCGGCATCGAGGTCTGTGCTC
GACCGGCTGCCGGACATCGACCTGGCGGTGCCCGCCGGGTCCCTACCCGTCGCCCT
CGCCCTGGCTGCGGGGGCTGACCGAGTTGCCCGTGCGGTTACCCCGACCACTGCCCT
GGGAGGCACGTCAGCATGA

>CYP157B1

GTGACCGACATCGACCCGTCACCCACCCCGTCGCCGCACCGGGCTGCCCGTCCACCC

GGACGCCGTGCCGTTGGCCGACTGGAGTACCAGCAGACGCCGTCGGAGCTGTACCGA
GGTCTGCGGGCCGAACACGGCGCCGTCGCGCCCCTGCTCCTGGACGGCGGCATCCCCG
CCTGGCTGGTCTCGGCTACCCCGAGGTCAGCTATGTGACCAGTCACGACGAGCTGTTTCG
CCCGGGACTCGCGCCGCTGGAACCAGTGGGGCAGCATCCCGCCCGACTGGCCCCCTGCT
GCCCTACGTCGGCCACCAGCCGTCGGTCTGTTACCCGAGGGCGAGGAGCACCGGCGG
CGGGCCGGCGTCATCACCCAGGCACTGGCGGGCATCGACCAGTTCGAACTGGCCCCGGG
ACTGCAGGCACCTCGCCGACCGGCTCATCGCCGCCTTCGCCGGAAGCGGCCGGGCCGA
GCTGATGAGCGGCTACGCGCACCCCCTGCCGATGCGCGCCGCGTCCGCATGTGCGGG
ATGCCGCACACCGGCGTCGAGACCCGGCAACTCGTCGAGGACCTGCGCATCTCCCTGGA
CGCGGCCGAGGGAGACGACCCGGTCGCCGCCTACACGCGCGTGGGAGAGCGCATCCAC
CAACTGGTCCGCCACAAGCGCGAGCGCCCGGGCCCCGACGTCACCTCCCGCATGCTCAC
CCACCCGGCGGGACTGACCGACGAGGAGATCGTCCAGGACCTCATCTCCGTCATCGCCG
CCGCCAGCAGCCCACCGCCAACTGGATCGGCAACACGCTGCGCCTGCTGCTCACCGAC
GAACGCTTCGCCCTGAACGTCTCCGGCGGGCCGCCTCAGCGTGGGGGAGGCCCTCAACG
AGGTGCTGTGGCTGGACACACCCACGCAGAACTTCATCGGACGCTGGGCCGTCAACGAC
ACCCAGCTGGGCGGCCGGCACATCCGGGCCGGCGACTGCCTGGTGCTCGGCTTCGCCG
CGGCCAACACCGATCCGCAGCTGTGGCCCAGGCCACGTCGGCGCCGAGAACTCCGC
GCACCTGTCCTTCAGCAACGGCGAACACCGGTGCCCTACCCGGCACCCCTGCTCGCCG
ACGTCGTCGCCCGCACCGCGGTCGAGACCCTGCTGGAGCGCTCCCCGACCTCGTACTC
GCCCTCGAGCCCGGGGAGTTGACCTGGCGTCCGTCCATCTGGATGCGCGGCCTGACCG
CCCTGCCCGCGCTCTTCACACCCGTCGTGGCCTGA

>CYP157C1

GTGACGCCTGAAAGCCACTCCCCGACCGGAACGGGCGAACCCCTGCTCGAACCACCGCC
CGGCTGCCCCGCGCACGGCCTCGGGCCGGGCGGACTGCACCGGCTCCACGAGGCGGA
AGACCTGGAGGAGCTGTACGAGAAGCTGCGCGAGCAGCACGGGCCGGTGGCGCCCGCG
CTGCTCCACGACGACGTACCCATGTGGGTGGTCTCGGGCACGCCGAGAACCTGCACAT
GGTGAGCACCCCTCCCAGTTCTGCCGCGACAGCCGGATCTGGACCCCGCTCAACGAGG
GCATGGTCAAGCCCGACCATCCCCTCATGCCGCACATCGCCTGGCAGCCCATCTGCTCGC
ACGCCGAGGGCGACGAACACAAGCGGCTGCGCGGGCGCGGTACCAAGTGCCATGTCCGA
CCTCGACTACCGGGAGCTGCGCCGGCACATCAAGCGCTACACCCAGCGCGTCGTCAACC
GTTCTGCGAGGAGGGCCGCGCCGACCTGGTCAGCCAGTTCGCCGAGCACCTGCCGAT
GGGCGTGATGTGCCATCTGCTCGGCATGCCCGAGGAGTACAACGACCGGCTGGTGGAGG
CCGCCCGCGACACCCTCAAGGGCACCGAGACGGCGATCGCCAGCCACGCCTACGTCATG
GAGGCGCTGGGACGGCTCACCGCCACCCGTCGGGGCCGACCCGGCCGACGACATCGCGG

GCCGCCTGGTCACCCACCCGGCCGGGCTCACCGACGACGAGGTCCGCGAACACCTGCG
GGTGGTGTCTGCTGGCCGCCTACGAAGCGACCGTCAACCTGATCGGCAACGTGATGCGCG
TCGTGCTCACCGACCCGGGCTTCCGCGCCCAGCTCAGCGGCGGCCAGATGACGGTGCC
CCAGGCGGTTCGAGCAGTCCCTGTGGGACGAGCCGCCGTTACGACACCGTCTTCGCCTACT
TCGCCAAGCAGGAGACGGAACCTCGGCGGCCAGCGCATCCGGGCGGCGACGGCCTGCT
GCTCGGCATCGCGCCGGGCAACGTTCGACCCGCGCATCCGTCCCGACCTGGACGCCAGC
ATGATGGGCAACCGCGCCCACCTCGCCTTCGGCGGCGGCCCGCACGAATGCCCGGGGC
AGGACATCGGGCGGGCCATCGCCGACGCCGGCATCGACGCGCTGCTGATGCGGCTGCC
CGACCTCCAGCTCGACTGCGACGAGGACGACCTGCGCTGGCGGTTCGTCGATCGCCTCGC
GGCACCTGGTGGAACCTGCCGGTCCGTTTCGAGCCCAGGGCGCAGCAGGACATCATGCAG
CAGCCGAGTCACGCCCCGACCCCGGAACGGCACGCACCCTGGCACGTTCGGCCTGCCGA
AGCCGGAACGGCGGGCACAGCCGCCGCTGCCCGCCCAGCCGCCGCAGCCCGTGTCCG
TGACGGCCGCAGAGCCTCAGCAGGCACCGGGAGCGGGGCAGCCCCGCCCCGCGGTG
CCTGGCAGCGCTTCTGCTCTGGTGGCGCGGCTACTGA

>CYP158A1

ATGACGCAGGAGACCACCACGCTCACCGGCCAGTCGCCACCCCCGGTGCGGGACTGGC
CCGCCCTCGACCTGGACGGCCCGGAGTTCGACCCGGTCTCGCCGAGCTGATGCGCGA
GGGGCCGCTGACCCGCGTCCGGCTCCCGCACGGCGAGGGCTGGGCCTGGCTGGCGAC
GCGGTACGACGACGTGAAGGCGATACCAACGATCCGCGCTTCGGCCGGGCCGAGGTGA
CGAACGTCAGATCACGCGTCTCGCCCCGCACTTCAAGCCGCGGCCCGGCTCGCTCGCC
TTCGCCGACCAGCCCGACCACAACCGGCTGCGGGCGCGGGTCGCGGGCGCGTTACCCG
TCGGCGCCACGAAGCGGCTGCGGCCCGGGCGCAGGAGATCCTGGACGGTCTCGTGGA
CGGGATCCTCGCCGAGGGGGCCACCGGCCGACCTGGTTCGAGCGGGTCTGGAGCCCTTC
CCGATCGCCGTCGTCAGCGAGGTGATGGGCGTGCCGGCCGCCGACCGGGAGCGTGTGC
ACTCCTGGACCCGGCAGATCATCTCGACCTCCGGCGGTGCCGAGGCCGCCGAGCGGGC
CAAACGGGGCCTGTACGGGTGGATACCGAGACCGTGCGCGCCCGGGCCGGCAGCGAG
GGCGGCGACGTGTAICTGATGCTCGGGGCGCCGTGGGCGCGGGCGAGGTCGGCGAG
ACGGAGGCCGTCGGGCTGGCGGGGCGCTGCAGATCGGCGGCGAGGCCGTCACGCAC
AACGTCGGCCAGATGCTGTACCTGCTGCTACCCGCCGCGAGCTGATGGCCCGGATGCG
CGAGCGGCCCGGGGCGCGCGGCACGGCGCTGGACGAGCTGCTGCGCTGGATCTCGCA
CCGCACGTTCGGTGGGCCTCGCCCGCATCGCCCTGGAGGACGTCGAGGTGCACGGGACG
CGCATCGCCGCCGGTGTAGCCGGTGTACGTCTCGTACCTGGCCGCCAACCGCGACCCGGA
CGTCTTCCCCGACCCGGACCGCATCGACCTGGACCGCGACCCCAACCCGCACCTGGCGT
ACGGCAACGGGCACCACTTCTGCACCGGCGCCGTGCTCGCCCGGATGCAGACCGAGCT
GCTGGTCGACACGCTCCTGGAGCGGCTGCCGGGGTTCGGGCTCGCGGTCCCGGCCGAG

CAGGTGGCCTGGCGCCGCAAGACCATGATCAGGGGTCCGCGCACGCTGCCCTGCACCT
GGTAG

> **CYP158A2**

ATGACTGAAGAAACGATTTCCAGGCCGTGCCACCCGTCCGGGACTGGCCGGCCGTCGA
CCTTCCCGGCAGCGACTTCGACCCGGTGCTGACCGAGCTGATGCGCGAGGGTCCCGTCA
CCCGGATCTCGCTGCCCAACGGCGAGGGCTGGGCCTGGCTCGTGACCCGCCACGACGA
CGTCCGCCTGGTCACCAACGACCCCCGGTTCGGGCGCGAGGCCGTCATGGACCGCCAG
GTCACCCGGCTGGCCCCGCACTTCATCCCGGCGCGCGGGCGCGGTGGGCTTCCTGGACC
CGCCCGACCACACCCGGCTGCGCCGCTCGGTGGCCGCGGCCTTCACCGCGCGGGGGCGT
GGAGCGGGTGC GCGAGCGGTCCCGGGGCATGCTCGACGAGCTGGTCGACGCCATGCTG
AGGGCCGGTCCGCCCCGCCGACCTACCGAGGCGGTGCTGAGCCCGTTCCTCATCGCGG
TGATCTGCGAGCTGATGGGTGTGCCGGCCACCGACCGGCACTCCATGCACACCTGGACC
CAGCTGATCCTGTCTCTCGCACGGCGCCGAGGTCAGCGAGCGGGCCAAGAACGAGAT
GAACGCCTACTTCTCGGATCTCATCGGGCTCCGCTCCGACAGCGCGGGCGAGGACGTCA
CCTCGCTGCTGGGTGCCGCCGTGGGGCGGGACGAGATCACGCTGTCGGAGGCCGTCGG
GCTCGCGGTGCTGCTCCAGATCGGCGGGCGAGGCGGTACCAACAACAGCGGGCAGATG
TTCCACCTGCTGCTGAGCCGCCCGGAGCTGGCCGAACGCCTGCGCTCCGAGCCGGAGAT
CCGCCCCCGGGCCATCGACGAGCTGCTGCGCTGGATCCCGCACCCGCAACGCCGTGGGG
CTGTCCCGGATCGCCCTGGAGGACGTGGAGATCAAGGGGGTGC GGATCCGCGCGGGCG
ACGCCGTCTACGTCTCGTACCTGGCGGCCAACCGCGACCCGGAGGTGTTCCCCGACCCG
GACCGCATCGACTTCGAGCGCTCCCCAACCCGCACGTCTCCTTCGGCTTCGGCCCCGA
CTACTGTCCCGGCGGCATGCTGGCGCGGCTGGAGTCGGAGCTGCTCGTCGACGCGGTC
CTGGACCGCGTGCCGGGGCTGAAGCTCGCGGTGGCGCCGGAGGACGTGCCCTTCAAGA
AGGGTGCGCTGATCCGCGGGCCCCGAGGCCCTGCCGGTGACGTGGTGA

> **CYP159A1**

ATGTCCACCGCGCAGCAGGTCCCCGACATCCTCTCCCCGAGTTCGCGGGCAACCCGTA
CCCGGCCTACCGGACGATGCGGGACAGCGCTCCCTCATCCGGCACGAGGCCACCCAGA
GCTGGATCGTCTCGCGCTACGAGGACGTGGAACGCGTCTTCAAGGACCGCGCCGGTCAG
TTCACGACGGAGAACTACGACTGGCAGATCGAGCCCGTGCACGGCCGGACGATCCTCCA
GCTGAGCGGCCGGGAGCACGCGGTCCGGCGCGCCCTGGTCGCCCCGGCCTTCCGCGG
GGCCGACCTCCAGGAGCGGTTCTGCCGGTCATCGAGCGCAACTCCCGTGAGCTGATCG
ACGCGTTCGGCACACCGGCCGCGCCGATCTGGTCGCCGACTACGCGACCCGGTTCCTCC
GTCAACGTCATCGCGGACATGCTGGGCCTGGACAAGGCCGACCACGACCGCTTCCACGG
CTGGTACACGTCGGTCATCGCCTTCTCGGCAACCTCTCCGGCGACCAGGAGGTTCGCGG

CGGCCGGTGC GCGCACCCGGACGGAGTTCGCCGAGTACATGATCCCGGTCATCCGGGA
 GCGTCGTGAGAACCCGGGCGACGACCTGCTGTCCACGCTGTGCGCCGCGAGGTGGAC
 GGCGTCCGGATGAGCGACGAGGACATCAAGGCGTTCTGCAGTCTGCTGCTCGCGGGCGG
 GCGGCGAGACCACCGACAAGGCGATCGCCGGCATCTTCGCCAACCTGCTGGCCCATCCG
 GAGCAGTTGGCGGCCGTGCGCGAGGACCGGAGCCTGATCCCCCGCGCCTTCGCCGAGA
 CGCTGCGCTACACCCCGCCGGTCCACATGATCATGCGTCAGACGGGCGACCGACGTCACG
 CTCAGCGGCGGCACGATAACCGGCCGGTGCCACCGTCACCTGCCTGATAGGAGCGGCCAA
 CCGGGACGAGACGCGCTACCGCGACCCGGACCGCTTCGACATCATGCGCGACGACCTGA
 CCACGACGACCGCCTTCTCGGCCGCGCCGACCACCTCGCCTTCGCGCTCGGCCGGCA
 CTTCTGCGTCGGCGCCCTGCTGGCGAAGGCGGAGGTGAGATCGGGGTGCGTCAGCTC
 CTGGACGCGCTGCCCGGCCTGCGGACGGAGGACGGCTTCGAGGTGGTGGAGCGGGGC
 GTGTTACCCGGGGCCCGCAGTCGCTGCCGGTCCGCTTCACACCCGCGGCCTGA

>*S.coelicolor fdx1*

ATGACTTACGTCATCGCACAGCCCTGCGTCGACATCAAGGACCGGGCGTGTGTCACC
 GAATGCCCCGTGGACTGCATCTACGAGGGCGCGCGCACGCTCTACATCAACCCGGCG
 GAGTGCGTCGACTGCCATGCGTGCGAGCCCGTCTGCCCCGTGAGGGCGATCTTCCAC
 GAGGACGATCTGCCGCGGCACTGGGCGCACTACCTCGCCGTCAACGCGGAGTACTTC
 GACGAGGCGGCTTCCCCGTGCGCGCACGGCGTGACGCGGTGCGGCGACCACCCGGC
 GGTCGAGGCACTGCCCCCGCAGAGCGGGCCGCACAAGAAGGAGATCTCCTTCTTCA
 CCGTCCGCAAGGAGGAAGCGGCCGACGCGGATCTGTGGTTCCCTCTCTGA

> *S.coelicolor fdx2*

GTGACCTACGTCATCGCGCAGCCTTGTGTGCGACGTGAAGGACAAGGCGTGCATCGAG
 GAGTGCCCGGTGCGACTGCATCTACGAGGGCCAGCGGTCTTGTACATCCACCCGGAC
 GAATGCGTCGACTGTGGTGCCTGTGAGCCGGTCTGCCCCGTGAGGGCGATCTTCTAC
 GAGGACGACACTCCCGAGGAGTGGAAGGACTACTACAAGGCGAACGTCGAGTTCTT
 CGACGAGCTCGGCTCGCCCGGCGGCCAGCAAGCTGGGGCTGATCGAGCGCGACC
 ACCCTTCGTCGCCGCGCTGCCGCCGAGAACCAGTAA

> *S.coelicolor fdx3*

ATGACCTTGGCAGGCCAGGCGTTCGGATCCGGGAGAGTCGGCGCCGAGCGGGACCG
 GTGTGTCGGGGCGGGGCAGTGCGTCCTCGCCGCGCCGGGTGTCTTCGACCAGGACG
 AGGAGGACGGGCTCGTGCGGTCCTCGCCGAGAGGCCCTCCGCGGCCGAGTCGGAC
 GCCGTGCGCGCGGCCGTGCGGGCCTGCCCGTCTGGAGCGCTCACTCTTCGGTAG

> *S.coelicolor fdx4*

ATGCACATCGGCATCGACAAGGACACCTGCATCGGCGCGGGCCAGTGCGCCCTGACC
 GCCCCCGGCGTCTTCACCCAGGACGACGACGGCTACAGCACGCTCAGGCCCGGCCG

GGAGGACGGCGGCGGCAGCGCGCTGGTCCGGGAGGCGGCCCGCGCCTGCCCCGTCG
GCGCGATCACCGTCTCGGAGCGGGTCGGCTGA

> *S.coelicolor fdx5*

ATGAGGATCTCCGTTCGACCCCGAGCAGTGTACGGCTCCGGAGACTGCGTCCACCGG
GCCCCGTCCGTCTTACCCAGGTGGGCGGCCTCGGCGCCGTCATACCCGGCCGGGAG
CACGACGTCGACGCCCCGCGGGTGCGCGAGGCGGCCGAAGGGTGTCCGTCCGCCGC
GATCACCATCGCCCCGACGGAAGCGGAAGGGGTCCGGGGCTGA

> *S.coelicolor fdx6*

ATGAGCGTGCAGCAGGAGGCCGCTGTCGACGGCGAGGCTCTGGAGGTCTGGATCGA
CCAGGACCTGTGTACCGGCGACGGCATCTGCGCGCAGTACGCGCCCCGAGGTGTTTGA
GCTGGACATCGACGGTCTGGCCTACGTGAAGGGCGCGGACGACGAGCTTCTGCAGG
CCCCGGGCGCGACAACGCCCGTGCCGCTGACGCTTCTCACGGATGTGGTGGACTCGG
CGAAGGAGTGTCCGGGCGAGTGCATCCATGTACGTCGCGTTTCGGACAGGGCCGAG
ATCTACGGTCCGGACTCAGAGTGA

> *S.coelicolor fdr1*

ATGCCCCGCCCTCTGCGGGTAGCCATCGTTCGGATCCGGCCCCGGCCGGGATCTACGCCG
CCGACGCCCTGCTCAAGTCCGAAGTGGCCGCCGACCCCGGTGTTTCCATCGACATCT
TCGAGCGCATGCCCCGCCCGTTCGGCCTCATCCGGTACGGCGTCGCGCCCCGACCACC
CGCGGATCAAGGGCATCATCACGGCCCTCCACCAGGTGCTCGACAAGCCGCAGATCC
GCCTCTTCGGCAACGTGAACTACCCACCGACGTCAGCCTGGACGATCTGCGCGCCT
TCTACGACGGTGTGATCTTCGCCACCGGCGCCACGGCGGACCGGGACCTGTCCCTCC
CGGGCATCGACCTCGACGGCTCGTACGGCGCGGCCGACTTCGTTCGCTGGTACGACG
GCCACCCCGACTTCCCGCGCACCTGGCCGCTGGAGGCGGAGAAAGTCGCCGTCTC
GGTGTTCGGCAACGTCGCCCTGGACATCGCGCGCGTCTTCGCCAAGACGGCCGACGA
GCTGCTGCCGACCGAGATCCCGCCGAACGTCTACGAGGGCCTCAAGGCCAACAAGG
CGCTGGAGGTGCACGTCTTCGGCCGCCGCGGCCCGGCGCAGGCGAAGTTCAGCCCG
ATGGAGCTGCGGGAGCTGGACCACTCCCCAACATCGAGGTGATCGTCGACCCCGAG
GACATCGACTACGACGAGGGCTCGATCGCGACCCGGCGCGGCAACAAGCAGGCCGA
CATGGTCGCCAAGACCCTGGAGAACTGGGCGATCCGCGACGTCGGCGACCGGCCGC
ACAAGCTCTTCTGCACTTCTTCGAGTCGCCCCGCGGAGATCCTCGGGCAGGACGGCA
GGGTGACCGGCCTGCGCACCGAGCGCACGGAGCTGGACGGCACGGGCAACGTCAA
GGGCACCGGCGAGTTCAAGGACTGGGACGTCCAGGCGGTCTACCGGGCCGTCCGGCT
ACCTCTCCGACCAGCTGCCAAGCTGCCCTGGGACCTCGAGACGGGCACGGTCCCG
GACGCGGGCGGCCGGGTCGTCCAGGAGTCCGGCGAGCACCTCCAGTCGACGTACGT

CACCGGCTGGATCCGGCGCGGTCCGATCGGCCTGATCGGCCACACCAAGGGCGACGC
 CAACGAGACGGTGTCCAACCTGCTGGACGACTACGCGAACGGCCGTCTCCAGACGC
 CCTCTCCCCCGCTCCCGAGGCCGTGGACGCGTTCCTCGCCGAGCGGAACGTCCGCT
 TCACCACCTGGGACGGCTGGTACCGGCTCGACGCCGCGGAGAAGGCGCAGGGCGAA
 CCGCACGGGCGTGAGCGCGTGAAGTACGTGAGCGCGAGGACATGCTCCGCGAGAG
 CGGCGCCTGA

> *S.coelicolor fdr2*

ATGCCGCGTGCGAAGACGTTTCGTGATCGTCGGGGGCGGCCTGGCCGCCGGCAAGGC
 CGCGGAGGAACTGCGCGAGCACGGCCACGACGGGCCGCTTCTCGTGATCGGGGACG
 AGCGGGAACGACCGTACATCCGGCCGCCGCTGTCCAAGGGGTACCTGCTGGGCAAG
 GAGGACCGCGAGTCCATCCACGTGCACCCCGAGAGCTGGTACCGGGAGCACGACGT
 CGATCTGCTCCTCGGCACGAGCGTGACGTCCGTGACGCGCGTGGCCGGGCGGTGAC
 GCTGGACGACGGCCGTCGCGTGCCCTACGCCGGTCTGCTGCTGGCCACCGGTTCTC
 GCCGCGCCGCTGTCGGTGCCGGGCGCGGACCTGGAGGGCGTGCTGTACCTGCGGC
 GCGTGGGCGACAGCGAGCGCCTCAAGGAGGCGTTCACCGAAGGAGCCCGGATCGTG
 GTGGTTCGGCGGCGGCTGGATCGGGCTGGAGACGGCGGCGGCGGCCCGGGCGGCCGG
 CGCGGAGGTGACCGTGCTGGAGCGCGGTGAGCTGCCCTGCTGAAGGTCCTGGGCC
 GCGAGGCGGCCGAGGTCTTCGCCGGTCTGCACCGGGACCACGGTGTGGACCTGCGT
 CCCCATGCCCGGATCGAGGCCGTCACCGGCACCGGGGGCCGCGTCGACGGGGTCCG
 GCTCGCCGACGGCACCCACCTGCCCGCGGACGCCGTGGTTCGTGGGGGTGGGCATCA
 CGCCCAACGTCCGCCTGGCCGAGGAGGCGGGCCTCGACGTGCGCAACGGCATCGTG
 ACGGACGCCCCTGTCGGACCTCCGCCGCCGGGGTCCACGCCGCCGGTGACGTCGC
 CAACGCCTACCACCCCGGCTCGGCCGGCACCTGCGCGTGAGCACTGGGCCAACG
 CGCTGCACCAGCCCCGTACCGCCGCGCTGAGCATGCTCGGCCAGGACGCGGTGTACG
 ACCGGCTGCCGTACTTCTACACCGACCAGTACGACCTCGGCATGGAGTACACCGGGT
 ACGCCGAACCGGGCGGCTACGACCGCGTCGTCTTCCGCGGGTCGCGCGAGGAGCGG
 CGGTTCTGGCGTTCTGGATGTCCGGCGACCGGGTGTGGCGGGGATGAGCGTCAAC
 CTGTGGGACGTGATCGGGACGATCCGCGCCCTGATCGAGTCGGGCGCGGAGACGGA
 CGACGCCGCCCTGGCCGACCCCTCGGTCCCCTGGAGAGCCTGCTTCCCCCGCACGC
 GCGGCCGACGGGAGACCAGGCGTGA

> *S.coelicolor fdr3*

GTGGTTCGACGCGGATCAGACATTCGTCATCGTCGGAGGCGGCCTGGCGGGCGCGAA
 AGCGGCCGAGACGCTCCGCACGGAGGGCTTACCGGCCGGGTGATCCTCGTCTGCG
 ACGAACGCGACCACCCTACGAGCGCCCGCCGCTGTCCAAGGGCTACCTCCTGGGCA
 AGGAGGAGCGCGACAGCGTCTTCGTGCACGAGCCCGCCTGGTACGCCCGGCACGAC

ATCGAGCTGCACCTCGGCCAGACCGTCGTCGCGATCGACCGCGCCGCCAAGACCGTC
CACTACGGCGACGACGGCACCCACGTCAGCTACGACAAGCTGCTCATCGCGACCGGC
GCCGAGCCCCGCCGCCTGGACGTCCCCGGCACCGGCCTCGCGGGCGTCCACCACCT
GCGCCGCCTGGCGCACGCCGAGCGCCTCAAGGGCGTCCTCGCCACCCTCGGCCGGG
ACAACGGACACCTGGTGATCGCCGGCGCGGGCTGGATCGGCCTGGAGGTCGCGGCC
GCGGCCCGCGAGTACGGTGCGGAGGTCACCGTCATCGAGCCCCGCCCGACCCCGCT
GCACGGCGTCCTCGGTCCCGAGCTGGGCGCCGTCTTCGCCGAGCTGCACGAGTCGCG
CGGCGTCCGCTTCCGCTTCGGCGTGAAGCTGACCGAGATCGTCGGCCAGGACGGTGT
GGTGCTGGCCGCCCGCACCGACGACGGCGAGGAGCACCCCGCGCACGACGTGCTCG
CCGCGATCGGCGCCGCCCGCGCACCGCGCTCGCCCAGGCGGCCGGGTTGGAGATC
GCCGACCGCGCGCACGGCGGGCGGCATCGTCGTCGACGACCACCTGCGCACCTCCGA
CCCCGACATCTTCGCGGGCCGGCGACGTGGCCTCCTTCCACCACGCCCTCTTCGACAC
CAGCCTGCGCGTGGAGCACTGGGCCAACGCCCTGAACGGCGGTCCGGCCGCCGCC
GCGCGATGCTCGGCAGGGGCCTCGCCCACGACCGCGTGCCCTACTTCTTCACCGACC
AGTACGACCTGGGCATGGAGTACTCCGGCTGGGCGCCGGCCGGCTCGTACGACCAGG
TGGTGATCCGCGGGGACGCGGCGAAGCGCGAGTTCATCGCCTTCTGGGTGAAGGAG
GGCCGGGTGCTGGCCGGGATGAACGTCAACGTGTGGGACGTCACGGAGCCGATCCA
GCAGCTGATCCGCTCGAAGACCCGGGTGGACACGGAGGACCTGGCGAACCCGCACG
TATCCCTCGAAAGCCTCGTCGCATAG

> *S.coelicolor fdr4*

GTGGTTCGTGGTGGGCGCCGGGATGGCGGGCGTGCAGACCGCGGTTCGCCCTGCGGGA
ACAGGGCTTACCGGCCCGGTGACCCTGATCGGTGCGGAGCCGCACCAGCCGTACG
ACCGCCCCCGCTGTCCAAGGCCGTCTTGCTCGGCACCGCCGAGGGCTCCGCCTTCG
ACGTCGACTTCGAGGCCCTCGGCATCACCTGCGGCTGGGCTGCGAGGCCCTGGGGCG
TCCGCCCGCGGAGCACGTCTTGACACCTCCGAGGGGCCCCTGACCCTACGACGTCC
TCGTCTTGCGACCGGCGCCGAACCGGTGCGGCTCCCCGGCGCGGAGGGCGTGCC
GGCGTGACCTGCTGCGCACCTGGACGACGCGGAACGGTTGCGGCCGGTACTCGC
CCGGCAGCACGACGTTCGTGGTTCGTCGGCGCGGGCTGGATCGGCGCCGAGTTCGCCA
CCGCCGCGCGGAGGCGGGCTGCGCGGTGACCGTGGTGGAGGCCGCCGACCGGCCG
CTGGCCGACGCGCTGCCCGCCGAGGTCGCCGCGCCGATGACCGCCTGGTACGCCGAC
GCCGGCGCCGAACCTGCGCACCCGCGCGCGCGTGGAGCGCGTCGAGTCCGGCGGGCC
CGGTGAGTCCGGCCGGGTTCGTCTTCGACGACGGCACCCGGCTGGCCGCGGGCGCGG
TCGTTCGTCGGCATCGGCGCCCGCCCGGCCACCGGCTGGCTGGCCGGGTCCGGCATCG
CCCTCGGCGCGCACGGCGAGGTGCTGGCCGACGACCTCCTGCGCACCTCCGCGCCG

GACGTCCACGCGGTTCGGCGACTGCGCCTCCTTCCCCTCCGCCCCTACGGCGAGCGG
CTGCTGGTGCACCACTGGGACAACGCCCTCCAGGGCCCGCGCACGGTTCGCCGCCGA
CATCGTCCGGCACCCCGGGCGGCGAGACCCCGGGCGGCCTACGACCCGGTGCCCTACTT
CTGGTCCGAGCAGTTCGGCCGCTTCGTCCAGTACGCCGGTACCACACCGCCGCCGA
CACCACCCTGTGGCGCGGCGACCCCGCGAGCCCGGCGTGGACGGTCTGCTGGCTGC
GCGAGGGCCGCCTGGTTCGCCCTGCTCGCGGTGGGCGCCCCCGCGACCTGGGCGCAG
GGCAGGCGCCTGATCCAGGCGGGCACGCCGATGGACCCGGACCTGCTGGCCGACCC
GGCGAAGCCGTTGAAGGCGGCGACGGCGGCCACGGCGTAG

> pTAex3 全序列及图谱

CTCGTGCGCCTTCCC GGAGTGGCTCACGCGCCCGCCACCACACCCGAAGCAAGAGC
TCAACTCCTATTCATTTTAGCGACTCCAATCTTCAAGAGCAGAATGTGAACGAACCAC
TGCATCATCAGTCTAGAGGATCCCCATCATGGTGT TTTGATCATT TTAATTTTATATG
GCGGGTGGTGGGCAACTCGCTTGC GCGGGCAACTCGCTTACCGATTACGTTAGGGCT
GATATTTACGTAAAAATCGTCAAGGGATGCAAGACCAAAGTAGTAAAACCCCGGAGT
CAACAGCATCCAAGCCCAAGTCCTTCACGGAGAAACCCAGCGTCCACATCACGAG
CGAAGGACCACCTCTAGGCATCGGACGCACCATCCAATTAGAAGCAGCAAAGCGAA
ACAGCCCAAGAAAAAGGTCGGCCCGTCGGCCTTTTCTGCAACGCTGATCACGGGCA
GCGATCCAACCAACACCCTCCAGAGTGACTAGGGGCGGAAATTTAAAGGGATTAATT
TCCACTCAACCACAAATCACAGTCGTCCCCGGTATTGTCCTGCAGAATGCAATTTAAA
CTCTTCTGCGAATCGCTTGGATTCCCCGCCCTGGCCGTAGAGCTTAAAGTATGTCCC
TTGTCGATGCGATGTATCACAACATATAAATACTAGCAAGGGATGCCATGCTTGGAGG
ATAGCAACCGACAACATCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGC
AAGCTCCGAATTCGAGCTCGGTACCCGGGGATCTGTAGTAGCTCGTGAAGGGTGGAG
AGTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGAGTTTGAGTTTGATGTA
CAGTTGGAGTCGTTACTGCTGTCATCCCCTTATACTCTTCGATTGTTTTTCGAACCCTA
ACGCCAAGCAGCTAGTCTATTATAGGAAAGGATCCGGATTAATGTGTTTTTCATAACG
CGGTACTGTATGGTACTTCTGTATTATATCACCGAAGCTCATGTATCTTACATGTATATAT
TATACAGACACAACCTTGGTTACCCACCATGATGTTTCCTGCAGATAATCTCCTGACG
ATCAATCTTACCACAGGGATATGATGGCACCCAACCTGGCGCCTTCGCAACATCAAAT
TATTTGTGATATGATCAAGTGCGATCCATCACTTACTAATGCCAGATAGCTGAAGCTG
CTAACTGCAGCACACGCGCAATTCCTAGGATTCGGTCAAATCTCCGGCTATTCGGCAG
TAGCAAAGCCCCTCAAATAAAGGTGGACGCCACGAAGCATCTACCAATAATGCT
GGAGGCTCTTTGTGATCATCTTCTTGAAAAGCCTGATCTATACCTTGACGAAATGGCC
ATCTTTCTATGGGATGAGTTCCAAATATACGCAACTACATCTAGTATCAGGCGGGCTCT
GTCTTCTAAAGGTTGGTCCAAAAAGGCAGCTCGGCAGAAAGCAAAGGAACGGAATT

AATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCC
ACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAG
CTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAAACCTGTCTG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCGCTTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCC
GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGAC
GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC
AGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAG
CCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAG
GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCA
GATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCT
GACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA
AGGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTAT
ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCA
GCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTAC
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACG
CTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA
GAGTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA
AGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCCTC
CGATCGTTGTCAGAAGTAAGTTGGCCGCAAGTGTATCACTCATGGTTATGGCAGCACT
GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACT
CAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCCGGCGTC
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATATTGAAAA
CGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGT
AACCCACTCGTGACCCAACTGATCTTCAGCATCTTTACTTTCACCAGCGTTTCTGG

GTGAGCAAAAACAGGAAGGCCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGG
AAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGGGTTAT
TGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCC
GCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACA
TTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGGTTTCGGTGATG
ACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAG
CGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGT
CGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTAAGTGCACCATAAAA
ATTGTAACGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATT
TTTAACCAATAGGCCGAAATCGGCCAAAATCCCTTATAAATCAAAGAATAGCCCGAGA
TAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGTCCACTATTAAGAACGTGGACT
CCAACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCAT
CACCCAAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTA
AAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAG
GAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA
CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTACT
ATGGTTGCTTTGACGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATAC
CGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTA
AGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGGCCAGTGCC
AAGCTTGCATGCCTGCAGGTCGACTCTAGAGTCTGCTCTCCCAGGCTCAGCAGGCCT
GATTTGATATGGAGAATGACACTTCCATGAAGTAATAGACGGTCCAGTCTCCCTTCCA
GCCAGTCGACGGCGACGGAGCGCAAGCCCTCAATGTCGCCAGCAATCAAATCGGCTT
CGAAGTCAAGGTGATTCACATGGCCGTTGCGAATGCTAACCTTGATGGAAGGCAAGG
ATGCAGATCCAACAAGGGCATTTCGTATTCGGTAAGCGTACAATTACTTCAGACTG
GCCGGTTTTCGACCTCTTTCCCAATCCAGGTGACGAAACGCCCAATATTTCTCGCCGAT
TCACTTTTGACTCGATCAGCATCGAGGACAATATCCCCTTGTACCCTTGCCCTGACAC
CTTCGCTTGTGGCCGACACGACTGATAGGTTTGTCCGGTCTAAATATGGCTGCTTCCTT
GACGTATTGCTTTACTACCTCGGGAACAAAGAAGGCCAGAATGAGGATGACAATGAC
GGAACCGACAGCAAGGAGGCCGCATAATACCGTCCACCAAGCTTTATTTTCGCGGTTT
TTTGGGGTAGTCATCTAATGAAACAGACCCGGACGCAGCAGAGGAAGCCCGCGATGA
CTCTATAACCGTACGCCGATATATCATCATCGCGGCGATGGAGAAGTGGGGTTGAC
TCCGAAGACACTTCAAAGGAGCGACGCTGTTGATTTGTAGACGACGCTTGATAGGGA
GAAGCATTATTGTCGTGATGCTCGCCCAACAGAGGCCGACTCGCCTCATCCGTCATAA
CGAACGCTGTGTAAAGCGGAGTGGGGGGGAAAGTGTGGATTGTGGAGAGTATGCGA

TAGTGTTGAGGCTGATCAGACGGCGAATCGGGCCAGATATGACCAGTTTAGAGGCCT
CATTTGACTATAATTTACATAAATTAGATAAATAGAGATGAACGCATGCAATAATTGCA
GCAAATATTGATGAAGCGAGAGGTAGGACGATGAAGGACTGTGAGCAGTTCAAGGTA
TCAGCAGAGTCAAGGGCCTGATGCAATGGCGGTGATCCGTGATCAGCGAACGGAAG
GGGCGCTAACTCTGTTTCTTTACCAATGATCGGAAGCTCCTGCTGGCGGACTTATGAG
TCATTCACGAATCATTTCTCAGTTATTTGTGGATGCCCTCGTTCTGTCCACAATTTCTTT
CCGCCCCAAGTCTTTTAAGTTCTTTAACATCTATATTCTTGCACTTCCAATGGCATCCCT
TCGCTCCGTACTCAAGAGCCAGAGCTTGCACACACCCGTGCGATCCTACTCCTCGCA
AACCATGCCTCCCGCCTCACCTTTGCTCCCCGCCACTTCCTCTCCATTGCGGACCTC
TCGCCCTCCGAGTTCGCAACCCTTGTTTCGCAATGCCTCCTCACACAAACGGGCTATCA
AGTCGGGGTCAATGCCCAAACTTGCAGGGATCACTCCTTGGGAAAACGTGGCCA
TGATCTTCAGCAAACGAAGCACGAGGACAAGGGTATCTACAGAAGGGGGCCGTTGTG
CAGATGGGAGGTCATCCGATGTTCTTGGGCAAGGATGATATCCAACCTAGGTGTCAACG
AGTCCCTATACGACACCTCCGTTGTCATTTTCGTCCATGGTATCCTGCATTGTAGCCCGT
GTCGGTAAACATGCAGAGGTCGCAGATCTGGCGAAGCACTCTTCGGTTCCAGTCATC
AATGCTTTGTGTGACTCTTTCCACCCTCTCCAAGCCGTGGCCGATTTCCAGACCATCT
ATGAAGCATTACCCCCAAGGCGCACACCTTTCAAGTCTAGGGTTGGAAGGATTGA
AGATCGCTTGGGTGGGTGACGCCAACAAACGTCCTGTTTCGATATGGCCATTGCTGCTAC
AAAATGGGTGTCGACATTGCTGTCGCTACTCCAAGGGGTACGAAATCCCTCCTCA
CATGCTGGAGCTCATCAAGTCTGCTGGAGAGGGTGTCTCGAAACCAGGAAAGCTTCT
GCAAACCAATATTCCCGAAGAAGCGGTCAAGGACGCCGATATTCTGGTCCACAGACAC
CTGGGTCTCTATGGGCCAAGAGGAAGAGAAGGCTCAGAGGCTGAAGGAGTTTGATG
GTTTCCAAATCACTGCTGAACTCGCCAAGCGAGGAGGAGCTAAGGAGGGCTGGAAG
TTCATGCACTGTCTCCCGCGACACCCTGAGGAGGTCAGCGACGAGGTTTTCTACAGC
AACCGGTCACTTGTCTTCCCTGAGGCTGAGAACCGGTTATGGGCTGCGATTTCCGCCT
TGGAGGGTTTCGTTGTCAATAAGGGAAAAATTGAATAAATATAACCAGGCTTCCATTT
AAATATATAGAGGCTGGCGTTATCAAACCTGATGAGTTGACGGGTATGAGATCATTCCG
TCCCTAAATATATTTACTCCGATCACGTAAAAGCCTGTTAGTAGAAGCATTTTCCCAAT
TATCCTGACCAATTCTTTCTAGCATATATCAAATAACTAATTGACATGTTCTCTCGCTTC
CTTATATTAICTCAGAGTATTGGAAATGGGGCAAATCGCACCCCGGTGACTTTTACATGT
CACGAATGCGGAGTCGTCCTAGCCAAGGTAGATCCAGGCCTAACACACCCCAACCCT
CGACTCTCCTCATTCCATACTCTTGACCTCTATCCAGCACATTCTTCTGAGGTTTCGCA
ATGGCTGTAGGTCGACTTTACTACATCCACAATCCTGCTGATTCGCTAACAAAGCTGA
ACAGTCCTCGGCATCACAAGCTCGGCCCCCGCGCTCTTATCCATCCTTAAACCGCATC

TCACTCGCGCCTTTAACTCCCCATTACCCATTGACGACGACGACAATGAGTCGGACA
AAGACCCACAGGATTATTTACCCCGCGAAATGAGGCTATTGACAACCCCACTAGAA
CTTCCTACTTGTCTAGCTATTCGGTCCCAGGGACACCCGGCGTCCTCT

结尾

作者简介:

1985年2月出生于山东省淄博市沂源县。

2003年9月—2007年6月, 河南农业大学生命科学院, 学士。

2011年9月—2012年9月, University of California, San Diego, Susuan S Golden 实验室(Golden Lab), 博士联合培养。

2007年9月—2013年6月, 中国科学院青岛生物能源与过程研究所, 博士。

2013年12月—2016年6月, 青岛琅琊台集团股份有限公司工作站、中国科学院青岛生物能源与过程研究所博士后流动站, 博士后。

博士后期间发表的学术论文、专著、重要科研成果:

主持项目:

基因工程改造里氏木霉合成乙醇的研究, 青岛市博士后基础研究项目资助, 5万元

学术论文:

- [1] Zhang Wei, Cao Shaona, Qiu Li, **Qi Fengxia**, Li Zhong, Yang Ying & Li Shengying. (2015). Functional Characterization of MpaG', the O - Methyltransferase Involved in the Biosynthesis of Mycophenolic Acid. *Chembiochem*, 16 (4) , 565-569.
- [2] Zhang Wei, Sun Jingran, Li Zhong, **Qi Fengxia**, Li Shengying. Reconstitution of Biosynthetic Machinery for Immunosuppressant Mycophenolic Acid in *Aspergillus oryzae*. (In preparation)

申请专利:

- [1] 支撑 P450 脂肪酸脱羧酶活性的还原伴侣组合及其应用, 李盛英, 房博, **齐凤霞**, 张伟, 杜磊; 申请号: 201610151612.X

博士生期间发表的学术论文、专著、重要科研成果:

学术论文:

- [1] Yao, L.[#], **Qi, F.[#]**, Tan, X., & Lu, X. (2014) . Improved production of fatty alcohols in cyanobacteria by metabolic engineering. *Biotechnology for biofuels*, 7 (1) , 94. (#Co-first author)
- [2] **Qi, F.**, Yao, L., Tan, X., & Lu, X. (2013) . Construction, characterization and application of molecular tools for metabolic engineering of *Synechocystis sp.* *Biotechnology letters*, 35 (10) , 1655-1661.
- [3] **Qi, F.**, Tan, X., & Lu, X. (2013) . Construction and evaluation of efficient gene expression platforms in *Synechocystis sp.* PCC6803. *Chinese journal of biotechnology*, 29 (9) :

1332-1342

- [4] Tan, X., Yao, L., Gao, Q., Wang, W., **Qi, F.**, & Lu, X. (2011). Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metabolic engineering*, 13 (2), 169-176.

申请和授权专利:

- [1] 吕雪峰, **齐凤霞**, 谈晓明, 罗全, 一种集胞藻 PCC6803 表达外源基因的方法, 2010.1.15, 中国, ZL201010034404.4
- [2] 吕雪峰, 谈晓明, 姚伦, 高倩倩, **齐凤霞**. 用于合成脂肪醇的构建体, 载体, 蓝细菌, 以及在蓝细菌中生产脂肪醇的方法, 申请日期: 2010.06, 中国, 申请号: 201010213758.5
- [3] Xuefeng Lu, Xiaoming Tan, **Fengxia Qi**, Quan Luo, Lun Yao, Qianqian Gao “Constructs, vectors and cyanobacteria for the synthesis of fatty alcohols, and methods for producing fatty alcohols in cyanobacteria.” Applicants: SHELL INTERNATIONALE RESEARCH MAATSCHAPPIJ B.V. International Patent, Pub. No.: WO/2011/086189. Application No.: PCT/EP2011/050555
- [4] Xuefeng Lu, Xiaoming Tan, Lun Yao, Qianqian Gao, **Fengxia Qi**. “Constructs, vectors and cyanobacteria for the synthesis of fatty alcohols, and methods for producing fatty alcohols in cyanobacteria.” US Patent, SHELL OIL COMPANY (Houston, TX, US) Application No.: US20120029248A1, 2011