

近平滑假丝酵母 (R)-专一性羰基还原酶基因的克隆与表达*

许娜 王海燕 聂尧 徐岩** 肖荣

(江南大学生物工程学院教育部工业生物技术重点实验室 无锡 24036)

摘要: 根据纯化得到的 (R)-专一性羰基还原酶 (rCR) 蛋白质测序结果推导出的核苷酸序列设计引物, 以筛选得到的近平滑假丝酵母 (*Candida parapsilosis*) CCTCC M203011 基因组为模板, 通过 PCR 扩增目的片段, 克隆后测序。核苷酸序列测定结果表明 *rcr* 基因全长 1011bp, 共编码 336 个氨基酸, 分子量为 35.9 kD。将序列递交 NCBI 比对, 与醇脱氢酶超家族成员序列同源性达 99%。在大肠杆菌 (*Escherichia coli*) JM109 中表达 *rcr* 基因, 重组菌可还原 β -羟基苯乙酮得到 (R)-苯乙二醇, 光学纯度为 100% e. e., 摩尔产率为 80.4%, 在反应体系中无需外加辅酶再生系统即可完成转化。

关键词: 羰基还原酶, 表达, 不对称还原, (R)-苯乙二醇

中图分类号: Q78 **文献标识码:** A **文章编号:** 0253-2654 (2006) 04-0112-07

Cloning and Expression of Gene Encoding (R)-specific Carbonyl Reductase from *Candida parapsilosis* CCTCC M203011 in *Escherichia coli**

XU Na WANG Hai-Yan NIE Yao XU Yan** XIAO Rong

(Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology
Southern Yangtze University, Wuxi 214036)

Abstract: The gene which encodes (R)-specific carbonyl reductase (rCR) from *Candida parapsilosis* CCTCC M203011 was cloned. The cloned sequence includes an open reading frame (ORF) consisting of 1,011bp, encoding a protein of 336 amino acids, with a molecular weight of 35.9 kD. The nucleotide sequence showed 99% similarity to those of the other members of the alcohol dehydrogenase superfamily. A recombinant *Escherichia coli* JM109 strain harboring the expression plasmid, produced (R)-1-phenyl-1, 2-ethanediol (100% e. e., 80.4% yield) from β -hydroxyacetophenone without any additive to regenerate NAD^+ from NADH.

Key words: Carbonyl reductase, Expression, Asymmetric reduction, (R)-1-phenyl-1, 2-ethanediol

手性化合物在人们生活中具有重要作用, 由于两个对映体在药理、毒理及功能作用等各方面均有所不同, 制备光学纯的手性模块化合物在医药、农业、材料, 以及环保等领域都具有广泛的价值。光学纯苯乙二醇不仅是液晶材料中不可缺少的重要手性添加剂, 而且已成为制备具有光学活性的医药、农药和功能材料的重要中间体^[1-3]。

生物法合成光学纯度手性化合物已成为近年来科学研究的热点, 利用酶不对称还原潜手性酮是得到手性醇的有效途径。Kataoka 等人将羰基还原酶基因与葡萄糖脱氢酶

* 国家重点基础研究发展规划 (973 计划) (No. 2003CB716008)
长江学者和创新团队发展计划 (No. IRT0532)

** 通讯作者 Tel: 0510-5864735, Fax: 0510-5864112, E-mail: yxu@sytu.edu.cn

收稿日期: 2005-11-07, 修回日期: 2006-03-07

基因协同表达来生产 (S) 或 (R) 型 4-氯-3 羟基-乙基丁酯, 产物光学纯度和产率均可达 90%, 在反应体系中利用葡萄糖脱氢酶来完成辅酶再生循环^[4]。Weckbecker 等人利用嘧啶核苷转氢酶、醇脱氢酶和甲酸脱氢酶共表达来生产 (R)-苯乙醇产率达 66%, 反应体系中利用嘧啶核苷转氢酶完成辅酶再生循环^[5]。Kataoka 等人克隆单一的聚酮还原酶, 在反应体系中外加辅酶和葡萄糖脱氢酶用以还原泛解酮酸内酯生产 (D)-泛酸内酯, 其产物光学纯度和产率也可达 90%^[6]。

本实验室通过筛选得到一株可用于制备光学纯苯乙二醇的菌株近平滑假丝酵母 CCTCC M203011。对反应过程中的酶进行纯化得到二个酶, 一个是 (R)-专一性羰基还原酶, 可将 β -羟基苯乙酮不对称还原为 (R)-苯乙二醇; 另一个是 (S)-专一性羰基还原酶, 可将 β -羟基苯乙酮不对称还原为 (S)-苯乙二醇。目前已对二个纯酶蛋白进行了测序 (相关内容将另文报道), 本文对 (R)-专一性羰基还原酶基因 *rcr* 进行了克隆与表达, 并对其全基因序列进行了分析。利用 *E. coli* JM109 表达 *rcr* 基因, 在反应体系中无需外加辅酶或其他酶即可还原 β -羟基苯乙酮得到光学纯产物 (R)-苯乙二醇。

1 材料与方法

1.1 菌株和质粒

近平滑假丝酵母 (*Candida parapsilosis*) CCTCC M203011, 大肠杆菌表达质粒 pTrc 99 A, 宿主菌 *E. coli* XL-1, *E. coli* JM109 为本实验室保存, pMD18-T Vector 购于 Takara Biotech 公司, pT-rCR, pTrc-rCR 为本实验构建。

1.2 培养基

2 × YT 培养基: 蛋白胨 20g/L, 酵母提取物 10g/L, 氯化钠 10g/L, pH7.2; 蓝白斑筛选平板: 蛋白胨 10g/L, 酵母提取物 5g/L, 氯化钠 10g/L, 琼脂粉 15g/L, X-Gal 4mg/mL, IPTG 2.4mg/mL, pH7.2。以上各种培养基使用前加氨苄青霉素至终浓度为 50 mg/L。

1.3 各种工具酶及其它试剂

蜗牛酶购于北京拜尔迪公司, 蛋白酶 K 购于 MERCK 公司, dNTP 购自上海生工公司, 限制性内切酶、T4 DNA Ligase、*Taq* DNA Polymerase、DNA Marker DL2000、X-Gal、IPTG 购于 Takara 宝生物公司, 3S Spin Agarose Gel DNA Purification Kit 购于上海申能博彩生物科技有限公司, β -羟基苯乙酮为本实验室合成, 其余试剂均为国产分析纯。

1.4 基因组 DNA 的提取

取 0.1g 菌体 TE 洗涤两次; β -巯基乙醇重悬, 室温振荡 30min; SCE 洗涤, 5% 蜗牛酶 37℃ 摇床孵育 3h; SCE 洗涤两次, 1% SDS, 100 μ g/mL 蛋白酶 K, 37℃ 水浴 30min; 1mL 酚-氯仿抽提两次取上清, 加入 2 倍体积 100% 乙醇沉淀 DNA; 75% 乙醇洗涤, 晾干, TE 溶解 DNA, -20℃ 保存。

1.5 rCR 蛋白质序列测定

LC/MS/MS 系统进行测序, 上海基康生物公司完成。

1.6 *rcr* 基因的 PCR 扩增条件

94℃ 预变性 3min; 94℃ 1min、45℃ 1min、72℃ 1min, 30 个循环; 72℃ 延伸

10min。取 5 μ L 扩增产物用 1% 琼脂糖凝胶电泳检测。

1.7 PCR 产物的纯化、连接、转化

PCR 扩增产物的纯化、连接按照 pMD18-T Vector 使用说明书进行, 连接产物转化 *E. coli* XL-1 感受态细胞, 蓝白斑筛选阳性克隆。

1.8 阳性克隆酶切鉴定

用碱裂解法^[7]进行质粒的小量制备, 得到的 pT-rCR 质粒做酶切鉴定。

1.9 克隆基因的序列测定

测序由上海申能博彩生物科技有限公司完成, 采用 M13-R 和 M13-F 通用引物由克隆基因的两侧进行测序得到克隆基因序列。

1.10 表达载体的构建

按文献[7]方法进行。

1.11 表达产物 SDS-PAGE 分析

按文献[8]方法进行。

1.12 重组蛋白酶活测定

2 \times YT 培养基培养菌体 8h, IPTG 诱导培养 16h, 离心发酵液将得到的湿菌体在缓冲液中混匀, 超声波破碎细胞, 离心取上清 Vallee&Hoch 法测定粗酶酶活^[9]。

1.13 重组菌不对称还原反应及其检测

取 0.1g 湿菌体在 1mL 0.5% β -羟基苯乙酮溶液中混匀, 30 $^{\circ}$ C 摇床反应 48h; 离心取上清反应液, 2mL 乙酸乙酯萃取; 萃取液经 HPLC 分析, 得到产物 (R)-苯乙二醇光学纯度 e. e 值及产率。

2 结果与分析

2.1 rCR 蛋白质序列测定

LC/MS/MS 法对酶蛋白进行测序分析得到 3 段肽段序列如下: NDLPVHKPK-AGQLLLK; VACVGPNGCGGCKYCR; LKELPEYIEK。将 3 段肽段序列与 PBD 上已发表的蛋白序列进行比对分析, 发现三段序列均与文献[10]报道的 CpSADH 酶的序列相匹配, 可知 *rcr* 基因与 *CpSADH* 基因具有很高的同源性, 因此可以 *CpSADH* 基因序列为依

据来设计引物, 经 PCR 扩增得到 *rcr* 全基因。primer1: 5'-aa gtcgac atg tca att cca tca a-3'; primer2: 5'-aa ctcgag cta tgg att aaa aac aac-3'; 下划线序列为引入的限制性内切酶酶切位点: *Sal* I、*Xho* I。

2.2 基因的 PCR 扩增、回收、连接和转化

PCR 反应结果见图 1, 扩增出一条约 1kb 的特异性条带, 与预期目标一致。回收 PCR 扩增产物, 经与 pMD18-T Vector 通过 TA 互补连接后转化 *E. coli* XL-1 感受态细胞, 蓝白斑筛选得到阳性克隆。

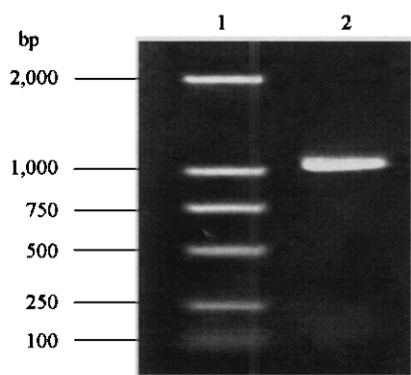


图 1 PCR 产物电泳结果

1 DL2000 DNA markers, 2 PCR fragment

2.3 重组质粒鉴定

小量制备所获得的重组质粒 pT-rCR，利用 primer1 和 T Vector 上都具有的 *Sal*I 位点酶切后，琼脂糖凝胶电泳得到长度为 1kb 的目的基因片段，如图 2 所示，可知已成功构建测序载体 pT-rCR。

2.4 *rcr* 全基因序列的测定结果

rcr 全基因核苷酸编码序列见图 3，其中下划线部分分别是上下游引物的对应序列，由 ATG 开始至 TAG 的基因全长 1,011 bp，将基因序列递交 NCBI 比对，其与 *CpSADH* 基因的同源性达 99%。基因序列翻译得到的多肽具有 336 个氨基酸残基，计算得其分子量为 35.9 kD。该基因已提交到 GenBank 数据库，数据库登记号为：DQ295067。

AAGTCGACATGTCAATTCCATCAAGCCAGTACGGATTGTTATCAATAAGCAATCAGGACTTAAGTTGA
GAAATGATTGCGCTGCCACAAGCCCAAGCGGGTCAATTGTTGTTGAAAGTTGATGCTGTGGATTG
TGTCATTCTGATTACATGTCATTTACGAAGGGTTGGATTGTGGTGATAATTATGTCATGGGACATGAAA
TTGCTGGAAGTGTGCTGCTGTGGGTGATGATGTCATTAATAAGGTTGGTGATCGTGTGCGCTGTG
TCGGACCCAATGGATGTGGTGGGTGCAAGTATGTCGTGGTGCCATTGACAATGTATGTAAAAACGCAT
TTGGTGATTGGTTCGGATTGGGGTACGATGGTGGGTATCAACAGTACTGTTGGTTACTAGACCAAGTA
ACTTGCTCTGATCCCAGATAACGTATCTGCAGACGTGGCTGCGGCTTCAACTGATGCTGTATTGACAC
CATATCACGCAATCAAGATGGCTCAAGTGTACCAACTTCGAATATCTTGCTTATTGGTGCTGGTGGAT
TGGGTGGAAATGCAATTCAAGTTGCCAAGGCATTGGTGCGAAAGTTACTGTTTGGACAAAAAAG
GAGGCTCGTGACCAAGCAAAAGAAAGTTGGGTGCTGATGCAGTTTATGAAACATTGCCAGAATCCATTTT
TCCTGGCTCTTTTTCAGCATGTTTGTATTGTTTTCAGTGCAAGCTACATTTGATGTATGTCAAAAGTAT
GTTGAACCAAAAGGGTGTAAATTATGCCCGTGGGACTCGGTGCTCCTAATTATCGTTTAATTTGGGAGAT
TTGGCATTGAGAGAAATTCGAATCTTGGGTAGTTTTGGGGAACACTAATGATTGGATGATGTTTGG
AAATTGGTTAGTGAAGGTAAAGTTAAACCCGTTGTGAGAAGTGCCAAATTGAAGGAATTGCCAGAGTA
TATTGAAAAATTGAGAAACAATGCTTATGAAGGTAGAGTTGTTTTTAATCCATAGCTCGAGTTAATCTC
TAGAGGATCCCCGGT

图 3 *rcr* 全基因序列的测定结果

2.5 *rcr* 基因结构域与系统进化分析

利用该基因的核苷酸编码序列，递交 NCBI (<http://www.ncbi.nlm.nih.gov/>) 通过对 GenBank's non-redundant Database (Nr Database) 进行 BLASTX ($E < 1e-27$) 搜索，在其标准下有 19 个同源基因返回。利用 clustalw (<http://www.ebi.ac.uk/clustalw/>) 程序对 rCR 酶的同源基因间进行多序列联配 (图略)，并进一步利用 InterProScan (<http://www.ebi.ac.uk/InterProScan/>) 对其编码区序列进行功能域鉴定。结果显示，rCR 酶蛋白具有 3 个功能域，分别为 Adh_ zinc 锌离子结合域，NAD⁺_ Binding 辅酶结合域以及 GroEs_ like 折叠腔域，见图 4。将 *rcr* 基因的核苷酸编码序列翻译为氨基酸序列，运行 Phylip 3.57 软件包下的 neighbor-joining 程序构建系统进化树，见图 5，阴影部分为 *Candida parapsilosis* CCTCC M203011，与其亲缘关系最近的是 *Candida orthopsilosis*

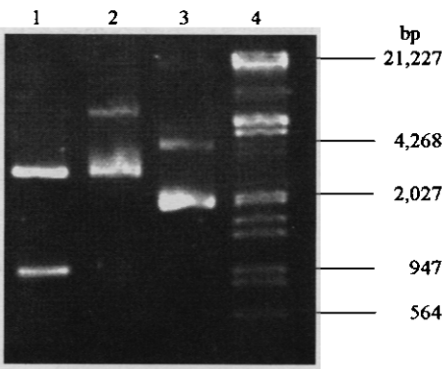


图 2 重组质粒 pT-rCR 酶切电泳结果
1 pT-rCR + *Sal*I, 2 pT-rCR plasmid, 3 pMD18-T plasmid, 4 λ DNA/*Eco*RI + *Hind*III

和 *Candida metapsilosis*, 亲缘关系最远的是 *Zymomonas mobilis*。

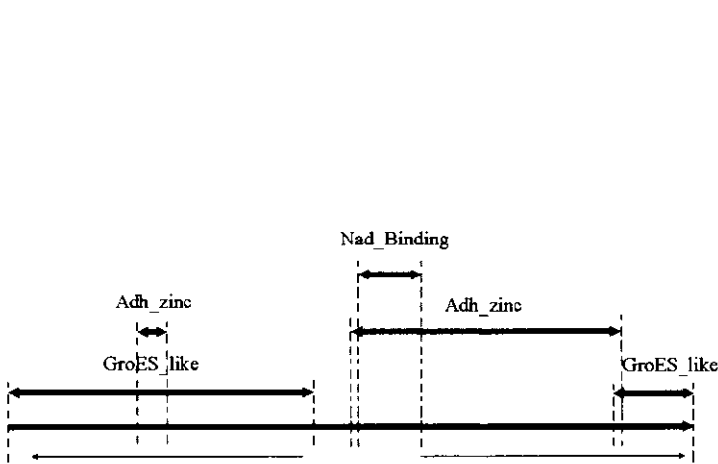


图 4 *rcr* 基因的功能域分布

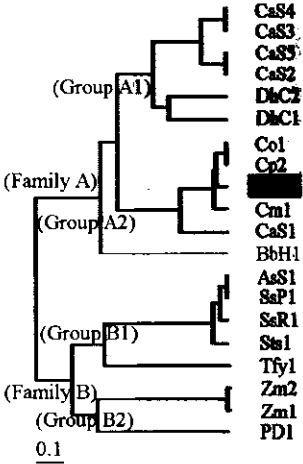


图 5 *rcr* 基因的系统进化树

2.6 pTrec-rCR 表达载体的构建及酶切检测

pTrec-rCR 表达载体的构建过程见图 6。利用 primer1 和 pTrec 99 A 载体上都具有的

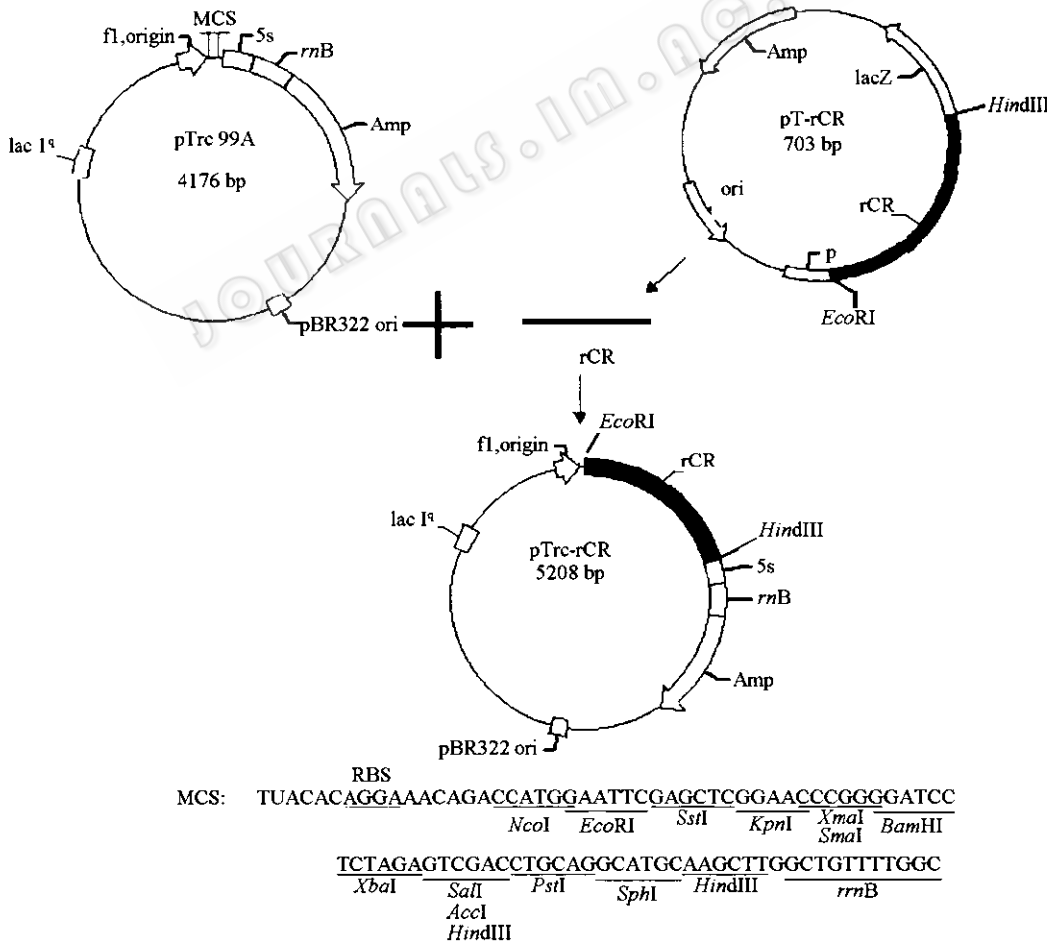


图 6 pTrec-rCR 表达载体构建图

SalI 酶切位点对 pTrc-rCR 表达载体进行酶切检测, 由图 7 可见切下 1kb 大小的片段, 证明 *rcr* 基因已插入至 pTrc 99 A 载体中, 成功构建了 pTrc-rCR 表达载体。将构建成功的 pTrc-rCR 质粒转入 *E. coli* JM109 感受态细胞, 得到重组菌待用。

2.7 表达产物 SDS-PAGE 分析

将重组菌细胞破碎液与 5 × 样品缓冲液混合煮沸 5min, 离心取上清 15μL 作 SDS—PAGE 分析, 如图 8 所示, 与不含重组质粒的 *E. coli* JM109 细胞破碎液相比较, 经 IPTG 诱导后的重组菌细胞破碎液中出现了一个新的条带, 可知 *rcr* 基因已在 *E. coli* JM109 中得到了表达。

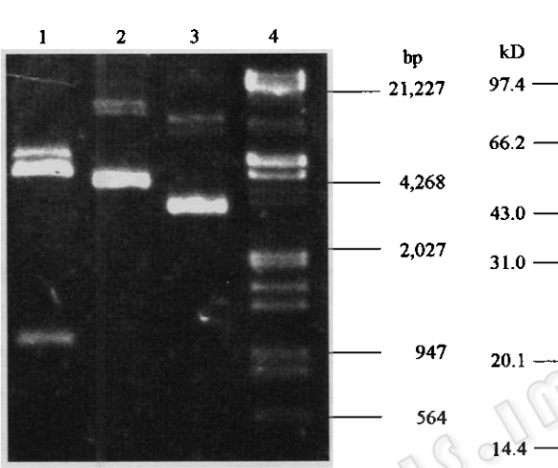


图 7 pTrc - rCR 表达载体酶切电泳图

1 pTrc-rCR + *SalI*, 2 pTrc-rCR plasmid, 3 pTrc 99 A plasmid, 4 λDNA/*EcoRI* + *HindIII*

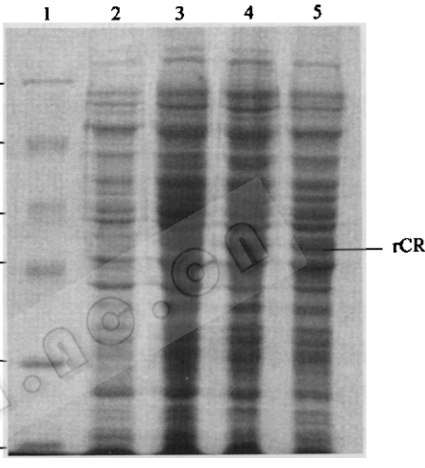


图 8 rCR 重组蛋白 SDS - PAGE 结果

1 Protein molecular weight markers, 2 *E. coli* JM109 (pTrc 99 A), 3 ~ 5 *E. coli* JM109 (pTrc-rCR) after IPTG induction

2.8 rCR 重组蛋白酶活测定

以 β-羟基苯乙酮为底物测定重组蛋白酶活, 结果见表 1。不含 *rcr* 基因的菌株其酶活很低, 而含有重组蛋白的菌株酶活明显增高, 说明 *rcr* 基因在重组菌中已得到表达。而经 1mmol/L IPTG 诱导后的重组蛋白酶活高于未诱导的酶活, 说明重组蛋白可在 IPTG 作用下得到更好的表达。

表 1 重组蛋白酶活测定

底物	空载体蛋白比活 (U/mg)	未诱导重组蛋白比活 (U/mg)	诱导重组蛋白比活 (U/mg)
β-羟基苯乙酮	0.04	0.129	0.223

2.9 重组菌不对称还原反应

以 β-羟基苯乙酮为底物做重组菌全细胞转化试验, 结果见表 2。结果分析, *E. coli* JM109 空载体菌株自身也存在一些与 rCR 重组蛋白有类似作用的酶类, 但其 (R)-苯乙二醇产率与 e. e. 值远低于作者克隆得到的 rCR 重组蛋白。后又以 *E. coli* JM109 为模板加入 primer1 和 primer2 做菌液 PCR 并未能克隆出基因条带, 可见 *E. coli* JM109 空载体菌株自带的酶与 rCR 重组蛋白并不相同, 而只是一些功能类似的酶。重组菌转化β-羟

基苯乙酮生产 (R)-苯乙二醇, 产物光学纯度达 100% e. e, 摩尔产率达 80.4%。

表 2 重组菌全细胞转化结果

菌株	摩尔产率 (%)	R-e. e. (%)
<i>E. coli</i> JM109 (pTrc 99 A)	18.5	91.5
<i>E. coli</i> JM109 (pTrc-rCR)	80.4	100

3 讨论

以酶蛋白测序结果为依据克隆得到了 *rcr* 基因全序列, 经 NCBI 比对以及对其结构域分析, 其具有典型的 Adh₂ zinc 和 NAD⁺ Binding 结合域。纯化得到的 rCR 纯酶性质显示, 它是一个含锌金属酶, 酶活依赖于辅酶 NAD⁺, 此结果与 *rcr* 基因的结构域分析相符合, 进一步确认了克隆的 *rcr* 基因的正确性。*rcr* 基因全序列的获得, 为深入研究 rCR 酶蛋白性质, 催化机理, 基因的结构与功能, 以及基因改造等后期工作打下了良好的基础。

由近平滑假丝酵母 CCTCC M203011 中克隆得到了一个 (R)-专一性羧基还原酶基因, 在大肠杆菌中进行表达, 重组菌可将 β -羟基苯乙酮不对称还原为 (R)-苯乙二醇, 产物光学纯度达 100% e. e, 摩尔产率达 80.4%。文献报道其他研究者克隆表达羧基还原酶基因, 通过重组菌不对称还原反应得到手性化合物, 其反应或者是在反应液中外加辅酶, 或者是协同表达另一个基因来完成辅酶再生循环, 无论何种方法都为手性化合物的生产增加了复杂性也提高了反应成本。作者所克隆表达的重组大肠杆菌在反应过程中可通过细胞自身的循环系统调节辅酶的再生循环, 无须外加辅酶或其它辅酶再生系统, 是一个高效便利的手性醇合成系统, 为进一步深入研究反应过程提供了良好的基础, 为工业化生产 (R)-苯乙二醇做了准备, 也为利用基因工程手段得到更高效的立体专一性酶奠定了基础。

参考文献

[1] Schmid A, Dordick J S, Hauer B, *et al.* Nature, 2001, **409** (1): 258 ~ 268.
[2] Bevinakatti H S, Banerji A A. J Org Chem, 1991, **56** (18): 5372 ~ 5375.
[3] Shinobu O, Yutaka K, Yasushi N. Biosci Biotech Biochem, 1992, **56** (8): 1216 ~ 1220.
[4] Kataoka M, Yamamoto K, Kawabata H, *et al.* Appl Microbiol Biotechnol, 1999, **51**: 486 ~ 490.
[5] Weckbecker A, Hummel W. Biotechnology Letters, 2004, **26**: 1739 ~ 1744.
[6] Kataoka M, Delacruz-Hidalgo A R G, Akond M A, *et al.* Appl Microbiol. Biotechnol, 2004, **64**: 359 ~ 366.
[7] 萨姆布鲁斯 J. 弗里奇 E F, 曼尼阿蒂斯 T. 分子克隆实验指南 (第二版). 北京: 科学出版社, 1993.
[8] 汪家政, 范 明. 蛋白质技术手册. 北京: 科学出版社, 2004.
[9] Stellmach B, 钱嘉渊 (译). 酶的测定方法. 北京: 中国轻工业出版社, 1992.
[10] Yamamoto H, Kawada N, Matsuyama A, *et al.* Biosci Biotechnol Biochem, 1999, **63** (6): 1051 ~ 1055.