

二次回归正交组合设计与综合相关系数法对耐热纤维素酶基因工程菌发酵条件的优化与分析

杜瑞卿¹, 吕文平^{2,*}, 王 丽¹

(1. 南阳师范学院生命科学与技术学院, 河南 南阳 473061;

2. 江南大学 食品科学与技术国家重点实验室, 江苏 无锡 214122)

摘 要: 目的: 研究耐热纤维素酶基因工程菌的规模化生产。方法: 采用二次回归正交组合设计试验对内切纤维素酶重组菌的发酵条件进行优化, 利用综合相关系数分析法分析因素影响程度以及相对重要性。测定菌体生长的生物量、内切纤维素酶活性。结果: 影响因素与纤维素酶活性间的二次回归方程为: $Y_2 = 45.858 - 1.653Z_1 - 1.258Z_2Z_3 - 2.012Z_1^2 - 1.499Z_2^2 - 1.415Z_3^2 - 1.719Z_4^2$, 方差分析具有显著性。内切纤维素酶重组菌产酶的最佳发酵条件为: 添加 15.9 g/L 乳糖、10 g/L 麸皮、15 g/L 硫酸铵、5 g/L 玉米浆、1.0 g/L 碳酸钙; 理论产酶量最好为 46.5357 U/mL; 在诱导时间约为 4 h, 实际实验酶产量达 70.78 U/mL。综合相关系数分析表明: 4 个影响因素的重要性依次为: 诱导剂 X_1 (乳糖) > 氮源 X_3 (硫酸铵和玉米浆) > 钙质量浓度 X_4 (碳酸钙) > 碳源 X_2 (麸皮), 相互间通过交互对纤维素酶活性的间接影响都非常大, 直接影响作用较小。结论: 通过使用综合相关系数分析与二次回归正交组合分析, 两者结合使用有互补性, 能够较好地指导耐热纤维素酶基因工程菌发酵条件的优化。

关键词: 耐热纤维素酶; 响应面; 综合相关系数; 优化

Optimization of Fermentation Medium for Thermostable Cellulase Production by Engineered Bacteria Using Quadratic Regression Orthogonal Composite Design Coupled with Integrated Correlation Coefficient Analysis

DU Rui-qing¹, LÜ Wen-ping^{2,*}, WANG Li¹

(1. School of Life Science and Technology, Nanyang Normal University, Nanyang 473061, China;

2. State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China)

Abstract : Objective: To provide reference for large-scale production of thermostable cellulase by engineered bacteria fermentation. Methods: The optimal fermentation medium of recombinant *Escherichia coli* obtained from *Clostridium thermocellum* for the production of thermostable cellulase were investigated using quadratic regression orthogonal composite design coupled with integrated correlation coefficient analysis, where four factors and their relative importance were analyzed. Results: The quadratic regression equation between four experimental factors and cellulase activity was presented as follows: $Y_2 = 45.858 - 1.653Z_1 - 1.258Z_2Z_3 - 2.012Z_1^2 - 1.499Z_2^2 - 1.415Z_3^2 - 1.719Z_4^2$. Variance analysis exhibited a significant difference. The optimal fermentation medium of the recombinant strain for the production of thermostable cellulase consisted of 15.9 g/L lactose, 10 g/L wheat bran, 15 g/L ammonium sulfate, 5 g/L corn syrup and 1.0 g/L calcium carbonate. The theoretical cellulase activity was 46.19 U/mL after 4 h of induction time. However, the actual cellulase activity was 70.78 U/mL. Integrated correlation coefficient analysis showed the importance of four factors according to the following order of inducer X_1 (lactose) > nitrogen source X_3 (ammonium sulfate and corn syrup) > calcium concentration X_4 (calcium carbonate) > carbon source X_2 (wheat bran). The indirect effects from cross-interaction on cellulase activity were all larger than the direct effects. Conclusion: The combinatorial application of integrated correlation coefficient and quadratic regression orthogonal composite design exhibits a complementary effect, which benefits the optimization of fermentation medium of recombinant *Escherichia coli* obtained from *Clostridium thermocellum* for the production of thermostable cellulase.

收稿日期: 2009-05-10

基金项目: 国家自然科学基金项目(30972120); 南阳师范学院生物化学与分子生物学重点学科科研资助项目

作者简介: 杜瑞卿(1968—), 男, 副教授, 硕士, 主要从事实验设计优化和生物统计学研究。E-mail: duruiqing8@163.com

* 通信作者: 吕文平(1968—), 男, 副教授, 博士, 主要从事食品营养与功能因子研究。E-mail: wplu@jiangnan.edu.cn

Key words: *Clostridium thermocellum* cellulase; quadratic regression orthogonal composite design; integrated correlation coefficient; optimization

中图分类号: S182

文献标识码: A

文章编号: 1002-6630(2010)03-0160-05

世界上每年生成1500~2000亿吨植物性有机物,其中约一半为纤维素物质,如何将纤维素物质降解转化为糖、乙醇、甲烷等容易利用的小分子有机物是当前国际上的重大课题之一。目前,对纤维素的降解利用主要采用酸碱化处理等化学手段,以及汽爆及蒸气加热等物理手段^[1]。生物法由于对设备要求低,分解后的产物易回收利用,对环境污染较小等特点而备受重视^[2-3]。

纤维素酶是一类将纤维素降解为葡萄糖的多组分酶系的总称,是一个复杂酶系,根据其功能的不同可分为3大类:作用于纤维素非结晶区的内切葡聚糖酶(又称CMCase),作用于纤维素无定型区切割糖苷键的外切葡聚糖酶,以及作用于纤维二糖的葡萄糖苷酶。它们协同作用,分解纤维素产生寡糖和纤维二糖,最终水解为葡萄糖。现已广泛应用在纺织、能源、食品、饲料等工业。提高纤维素酶的热稳定性可提高反应温度,降低反应体系的黏度,提高底物浓度,加快降解速度,所以研究耐热纤维素具有重要的实用价值和广泛的应用前景^[4-7]。

实验在前期构建了嗜热梭菌内切纤维素酶基因工程菌,该重组内切纤维素酶的最适温度为60℃,在70℃处理2min,酶的活性基本不受影响;酶的活性较高,发酵液中羧甲基纤维素酶活性为33.2U/mL。但该工程菌以进口的原料配制的LB培养基生长良好,而且诱导剂IPTG价格昂贵,大大增加了工业化生产成本。近年来,国内学者对大肠杆菌基因工程菌培养的诱导条件、培养基的优化、装瓶量、发酵温度、初始pH值和溶解氧浓度等进行了研究。笔者在上述研究文献的基础上,结合内切纤维素酶的特点以及规模生产的实际问题,选取4个因素:以乳糖为诱导剂、麸皮为碳源、硫酸铵和玉米浆为氮源、碳酸钙的钙离子质量浓度,进行四因素五水平二次回归正交组合设计分析,寻找最佳优化条件,再利用综合相关系数分析法,揭示各因素的影响程度以及因素间相互交互影响程度,以期耐热纤维素酶基因工程菌的规模化生产提供参考。

1 材料与方法

1.1 材料

1.1.1 菌种

Escherichia coli ECE12, 含内切纤维素酶重组质粒PET-30a, 由本实验室保藏^[8]。

1.1.2 试剂

CMC-Na 美国Sigma公司;蛋白胨、酵母膏和牛肉膏 上海化学试剂公司;玉米芯粉、麸皮和甘蔗渣均在本地市场购买;实验中所用其他试剂均为分析纯。

1.1.3 培养基

种子培养基为LB培养基(含卡那霉素50mg/L, 蛋白胨10g/L, 酵母提取物5g/L, NaCl 10g/L, pH7.0);发酵培养基:以麸皮、硫酸铵、玉米浆、碳酸钙为基础, 乳糖作为诱导剂。

1.2 方法

1.2.1 培养方法

将活化菌种接于5mL LB培养基中, 37℃振荡培养12h, 然后吸取300μL菌悬液至装有30mL发酵培养基的250mL三角瓶中, 在37℃, 200r/min条件下培养12h至对数生长期, 再加入乳糖诱导一定时间后, 测定其生物量和酶活性。同样发酵条件重复5次, 取其平均值作为实验数据。

1.2.2 菌体生长的生物量测定

准确称取10mL发酵液, 4000r/min离心20min, 用蒸馏水洗涤离心2次, 85℃烘干至恒重。

1.2.3 内切纤维素酶活性测定

1.2.3.1 制作标准葡萄糖曲线

取标准葡萄糖溶液(10mg/mL)2.5mL, 定容至100mL, 取其溶液2.5mL, 再加入DNS溶液2.5mL, 混匀后煮沸5min, 冷却加水10.0mL摇匀, 540nm波长处测定光密度(OD)值, 并用葡萄糖质量浓度(mg/mL)与其OD值作图, 即为标准曲线。

1.2.3.2 纤维素酶活力测定

取2只25mL具塞试管, 分别加入1% CMC-Na磷酸缓冲溶液(pH8)2.0mL, 0.5mL发酵破碎上清液, 60℃保温30min; 取出后加入2.5mL DNS, 煮沸5min; 冷却后加水10mL摇匀; 540nm波长处测定OD值, 以磷酸缓冲液代替菌悬液作为空白。从标准曲线上查出相应的葡萄糖含量折算成酶活单位(U/mL)^[9]。纤维素酶的单位定义为: 在上述条件下, 每分钟分解底物生成1μg葡萄糖所需要的酶量为1个酶活力单位(U)。

1.2.3.3 工程菌的遗传稳定性

将阳性工程菌株在LB液体培养基中培养12h, 稀释后分别涂布含卡那霉素(50mg/L)和不含卡那霉素的LB平板上培养24h, 用灭菌牙签挑取100个单菌落点种于LB平板上, 置37℃培养, 每隔24h取其菌落边缘部分

点种于下一个LB平板上。计数生长的菌落数,并据此计算生长阳性单菌落百分率。

1.2.4 实验方案及分析方法

1.2.4.1 二次回归正交组合设计

试验因素及水平编码见表1。

表1 试验因素、因素水平及水平编码

Table 1 Coded factors and levels in quadratic regression orthogonal composite design

水平	因素				g/L
	X ₁ 诱导剂	X ₂ 碳源	X ₃ 氮源	X ₄ 钙质量浓度	
上星号臂 γ	34	16	30/5	1.7	
上水平1	30	14	25/5	1.5	
零水平0	20	10	15/5	1.0	
下水平-1	10	6	10/10	0.5	
下星号臂 $-\gamma$	6	4	2/10	0.3	
Δ	10	4	2	0.5	

注: $\Delta_j = (X_j - X_{j0})/\gamma$, $\gamma = 1.414$; 诱导剂为乳糖; 碳源为麸皮; 氮源为玉米浆和硫酸铵; 钙质量浓度为碳酸钙质量浓度。

对因素 X_j 的各个水平进行线性变换,得到水平的编码 $z_j = \frac{x_j - x_{j0}}{\Delta_j}$,对4个因素进行四元二次回归正交组合设计试验^[10-12]。

表2 四元二次(1/2实施)回归正交组合设计试验方案及结果

Table 2 Arrangement and results of quadratic regression orthogonal composite design

试验号	因素				酶活性/(U/mL)	
	Z ₁ 诱导剂	Z ₂ 碳源	Z ₃ 氮源	Z ₄ 钙质量浓度	Y ₁ 实际值	Y ₂ 预测值
1	1	1	1	1	36.7	36.3
2	1	1	-1	-1	39.2	38.8
3	1	-1	1	-1	39.2	38.8
4	1	-1	-1	1	36.7	36.3
5	-1	1	1	-1	39.6	39.6
6	-1	1	-1	1	42.1	42.1
7	-1	-1	1	1	42.1	42.1
8	-1	-1	-1	-1	39.6	39.6
9	γ	0	0	0	38.6	39.5
10	$-\gamma$	0	0	0	44.4	44.2
11	0	γ	0	0	42.4	42.9
12	0	$-\gamma$	0	0	42.7	42.9
13	0	0	γ	0	42.7	43.0
14	0	0	$-\gamma$	0	42.7	43.0
15	0	0	0	γ	40.2	42.4
16	0	0	0	$-\gamma$	44.0	42.4
17	0	0	0	0	46.5	45.9
18	0	0	0	0	46.5	45.9

1.2.4.2 综合相关系数分析法

设有变量组 X_1, X_2, \dots, X_n 与变量组 Y 。每个变量有 m 个观测值,实验中 $n=4$, $m=18$ 。

1) 主成分的求解与合成^[13]

可求解出变量组 X_1, X_2, \dots, X_n 的主成分 F_1, F_2, \dots, F_n , 特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$, 方差贡献率(B):

$$B = \left(\frac{\lambda_1}{\sum_{i=1}^n \lambda_i}, \frac{\lambda_2}{\sum_{i=1}^n \lambda_i}, \dots, \frac{\lambda_n}{\sum_{i=1}^n \lambda_i} \right) = (b_1, b_2, \dots, b_n) \quad (1)$$

$$F = F_1 b_1 + F_2 b_2 + \dots + F_n b_n = C_1 X_1' + C_2 X_2' + \dots + C_n X_n' \quad (2)$$

F 被称为变量组 X_1, X_2, \dots, X_n 的合成主成分, X_1', X_2', \dots, X_n' 为 X_1, X_2, \dots, X_n 标准化处理后的数值。其中: $C_i = b(a_{i1} + a_{i2} + \dots + a_{in})$

C_i 反映了 X_i 在 X_1, X_2, \dots, X_n 中的重要性,称为影响系数。令变量 Y 影响系数矩阵 $T = (T_i)$ 。

2) 综合相关系数的求解及其分解

设变量组 X_1, X_2, \dots, X_n 与变量 Y 直接相关系数矩阵为 r ,令 $C = \text{diag}(C_1, C_2, \dots, C_n)$, $T = (T_i)$, F 与变量 Y 的综合相关系数为 R 可分解为:

$$R = TrC = \sum_{i=1}^n \sum_{j=1}^n Tr_{ij} C_j \quad (4)$$

其中 $Tr_{ij} C_j$ 表示变量 Y 与变量 X_j 的综合相关系数,它包含了直接相关系数 r_{ij} ,也包含了变量 Y 与变量 X_j 各自的影响系数 T_i 和 C_j ,称为变量综合相关系数。

Y 与变量组 X_1, X_2, \dots, X_n 的综合相关系数:

$$R(1) = \sum_{j=1}^n Tr_{ij} C_j \quad (5)$$

以上计算,采用统计分析软件SPSS15.1和工程计算软件MATLAB2007编程处理。

2 结果与分析

2.1 内切纤维素酶发酵条件优化

在确定诱导时机为 OD_{540nm} 值达0.6、诱导时间为4h、初始pH值为7、菌株生长阶段温度37℃、诱导阶段32℃,接种量为1.5%的条件下,对影响基因工程菌内切纤维素酶表达的几个主要因子(诱导剂(乳糖)剂量、碳源(麸皮)浓度、氮源(硫酸铵和玉米浆)浓度、钙离子(碳酸钙)质量浓度,采用四因素五水平二次回归正交组合设计试验,结果见表2。

利用MATLAB7.0软件,对表2数据进行二次多元回归拟合,通过逐步回归法,得到酶活性 Y_2 与四因素 Z_1, Z_2, Z_3, Z_4 之间的二次回归方程:

$$Y_2 = 45.858 - 1.653Z_1 - 1.258Z_2Z_3 - 2.012Z_1^2 - 1.499Z_2^2 - 1.415Z_3^2 - 1.719Z_4^2 \quad (6)$$

回归相关系数0.964, $F=24.42$, $P < 0.0001$,说明回归方程极其显著。

利用回归方程(6),进行预测,预测值见表2。从预测值与实测值得对比来看,也说明回归模型与实际情况拟合的很好。对回归方程(6)求导,解逆矩阵分析表明:

$Z_1 = -0.41$, $Z_2 = 0$, $Z_3 = 0$, $Z_4 = 0$ 时, $Y_2 = 46.5357$ 。

依据 $z_j = \frac{x_j - x_{j0}}{\Delta_j}$,可得对应的变量 X 的值: $X_1 =$

15.9, $X_2 = 10$, $X_3 = 15/5$, $X_4 = 1.0$ 时, Y_2 取得最大预测值 46.5357。在回归方程(6)中, 各项的系数均极其显著。

2.2 验证性实验

在最优化条件下(诱导剂 $X_1=15.9\text{g/L}$, 碳源 $X_2=10\text{g/L}$, 氮源 $X_3=15/5\text{g/L}$, 钙质量浓度 $X_4=1.0\text{g/L}$), 进行纤维素酶的产酶动态实验。内切纤维素酶基因工程菌的生长和产酶是不同步的, 在未经乳糖诱导时, 菌体生物量大量增加, 但产酶水平是非常低的; 当加入乳糖时, 菌体迅速开始产酶, 诱导时间约为 4h, 产酶量达 70.78U/mL。

2.3 综合相关系数分析法结果

根据表 1 和表 2 的数据, 实验 4 个因素的影响系数 C 为 -0.9566、-0.8454、-0.6672、0.7441。酶活性 Y 的影响系数: $T = 1$ 。酶活性 Y 与 4 个因素间的直接相关系数为 -0.406、-0.051、-0.138、0.066。由表 3 可知, 酶活性 Y 与 4 个因素间的综合相关系数为 $R_1(1)=0.710$, $P < 0.05$, 表示变量 Y 与变量组 X_1, X_2, \dots, X_n 的综合相关系数 $T_{ij}C_j$ 、变量 Y 与变量组 X_1, X_2, \dots, X_n 的综合相关系数 $R_1(1) = \sum_{j=1}^n T_{ij}C_j$ 。

表 3 酶活性和 4 个因素之间的综合相关系数
Table 3 Integrated correlation coefficients between cellulose activity and four experimental factors

项目	$T_{ij}C_j$				$R_1(1)$
	X_1	X_2	X_3	X_4	
综合相关系数	0.388	0.043	0.230	0.049	0.710*

注: *. $P < 0.05$ 。

由 4 个因素的影响系数 C 可知, 可以对酶活性 Y 起间接作用的大小依次为: 诱导剂 X_1 、碳源 X_2 、钙浓度 X_4 和氮源 X_3 。

由 Y 与 4 个因素的直接相关系数 r 可知, 可以对酶活性 Y 起直接作用的大小依次为: 诱导剂 X_1 、氮源 X_3 、钙浓度 X_4 和碳源 X_2 。

由表 3 可以看出, 4 个因素对酶活性 Y 的综合影响大小为: 诱导剂 $X_1 >$ 氮源 $X_3 >$ 钙浓度 $X_4 >$ 碳源 X_2 。

2.4 质粒稳定性和酶活稳定性分析

将基因工程菌用连续点种法 37℃ 继代培养, 以考察所含纤维素酶基因的质粒在宿主中的稳定性。结果显示, 经连续转接 16 次, 工程菌质粒的稳定性为 98%。此外, 将培养后的细胞重悬在 PBS 溶液中, 于 4℃ 放置 35d, 细胞仍保持 90% 以上的酶活性, 说明该基因工程菌具有较高的质粒稳定性和酶活稳定性。

3 结 论

通过以上分析, 可以得出以下基本结论:

3.1 二次回归正交组合试验分析表明, 发酵培养基对提

高内切纤维素酶活性的最佳条件为: 诱导剂 $X_1=15.9\text{g/L}$, 碳源 $X_2=10\text{g/L}$, 氮源 $X_3=15/5\text{g/L}$, 钙质量浓度 $X_4=1.0\text{g/L}$, 理论预测酶活性最值为 46.5357U/mL。在理论最优条件下, 验证性实验表明, 菌体迅速开始产酶, 诱导时间约为 4h, 产酶量达 70.78U/mL。理论与实际相符。

3.2 二次回归方程为: $Y_2=45.858 - 1.653Z_1 - 1.258Z_2 - 2.012Z_3 - 1.499Z_4 - 1.415Z_5 - 1.719Z_6$ 。方差分析也表明回归方程显著。依据回归方程进行预测, 预测结果与实际非常相似。也说明二次回归正交组合实验应用是合理成功的。

3.3 综合相关系数分析表明, 4 个影响因素的重要性依次为: 诱导剂 X_1 (乳糖) > 氮源 X_3 (硫酸铵和玉米浆) > 钙质量浓度 X_4 (碳酸钙) > 碳源 X_2 (麸皮)。相互间通过交互对纤维素酶活性的间接影响都非常大, 直接相关作用较小。综合相关系数分析与二次回归正交组合分析的结果基本一致, 而且各有侧重, 两者结合使用有互补性。

3.4 通过连续 16 代培养, 纤维素酶基因工程菌质粒较稳定, 便于工业化生产。

4 讨 论

嗜热梭菌是一种高温厌氧细菌。它的纤维素酶和半纤维素酶构成了一个多酶复合体^[14], 纤维素酶主要由内切 β -1,4-葡聚糖酶(C1 酶), 外切 β -1,4-葡聚糖酶(Cx 酶)和葡萄糖苷酶 3 种酶组成。其中, 内切 β -1,4-葡聚糖酶(EC3.2.1.4) 是分解纤维素最主要的酶。嗜热梭菌产生的纤维素酶耐热性好, 但嗜热梭菌又是一种严格的厌氧菌, 其纤维素酶的分泌必须在高度厌氧的条件下进行^[15], 酶产量相对较低, 培养基要求高, 产酶周期长等缺点, 不适合在工业中大规模发酵生产。但将来源于嗜热梭菌的内切纤维素酶基因在 *E.coli* BL21 中表达, 内切纤维素酶基因工程菌的最适酶活性为 33.2U/mL, 是出发菌的 2.8 倍, 酶的耐热性与原出发产酶菌株相似, 培养条件为好氧, 培养基成分简单, 克服了出发菌株的不足。

参考文献:

- [1] BELTRAME, CARNITI P, VISEGLIO A, et al. Fractionation and bioconversion of steam-exploded wheat straw[J]. *Bioresour Technol*, 1992, 39: 165-171.
- [2] 吴丹, 邓泽元, 范亚平, 等. 一株纤维素分解菌的分离、鉴定及产酶条件优化[J]. *食品科学*, 2008, 29(6): 219-221.
- [3] 罗新颖, 欧阳嘉, 许娟, 等. 耐热纤维素酶产生菌的筛选、鉴定及产酶条件优化[J]. *食品与生物技术学报*, 2007, 26(1): 85-88.
- [4] ANNISON G, CHOCT M. Anti nutritive activities of cereal non starch polysaccharides in broiler diets and strategies for minimizing their effects [J]. *Journal of Poultry Science*, 1991, 47(3): 232-242.
- [5] HAKI G D, RAKSHIT S K. Developments in industrially important thermostable enzymes: a review[J]. *Bioresource Technology*, 2003, 89 (1): 17-34.

- [6] 李淑彬, 陆广欣, 林如妹, 等. 嗜热菌: 工业用酶的新来源[J]. 中国生物工程杂志, 2003, 23(7): 67-71.
- [7] 陈燕勤, 毛培宏, 曾宪贤. 细菌纤维素酶结构和功能的研究进展[J]. 化学与生物工程, 2004(6): 4-6.
- [8] 吕文平, 康秀英, 乐国伟, 等. 嗜热梭菌内切纤维素酶的表达及其酶特性研究[J]. 江苏农业学报, 2009, 25(1): 185-187.
- [9] 高培基. 纤维素酶活力测定方法研究进展[J]. 工业微生物, 1985(6): 5-8.
- [10] 李云雁, 胡传荣. 试验设计与数据处理[M]. 北京: 化学工业出版社, 2005: 131; 141; 226.
- [11] 王允祥, 吕凤霞, 陆兆新. 林伞发酵培养基的响应曲面法优化研究[J]. 南京农业大学学报, 2004, 27(3): 89-94.
- [12] 冯培勇, 钟旭生, 杨立红, 等. 利用响应面法优化茶薪菇产纤维素酶的发酵条件[J]. 食品科学, 2009, 30(7): 162-165.
- [13] 李加纳. 数量遗传学概论[M]. 重庆: 西南师范大学出版社, 1995: 199-200.
- [14] DAE-HYUK K, NAM S H, KYUNG M P. Over-production of *Phytolacca* insulars protein in batch and fed-batch of *Escherichia coli* [J]. *Process Biochemistry*, 2001, 36(6): 537-542.
- [15] IRVIN J E, TEAFIHER R M. Cloning and expression of a Bactericides succinogenes mixed linkage β -glucanase gene in *Escherichia coli*[J]. *Apply Environment Microbiology*, 1998, 54: 2672-2676.