

弗氏链霉菌丝氨酸蛋白酶基因的克隆及表达 Gene Cloning and Expression of Serine Protease SFP2 from *Streptomyces fradiae* var. k11

李江^{1,2,3}, 石鹏君², 张王照², 韩晓宇³, 徐玲玲³, 张会图², 姚斌^{2*}, 范云六¹
LI Jiang^{1,2,3}, SHI Peng-Jun², ZHANG Wang-Zhao², HAN Xiao-Yu³, XU Ling-Ling³, ZHANG Hui-Tu²,
YAO Bin^{2*} and FAN Yun-Liu¹

1. 中国农业科学院生物技术研究所, 北京 100081

2. 中国农业科学院饲料研究所, 北京 100081

3. 东华理工学院生物系, 江西抚州 344000

1. Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China

2. Feed Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China

3. Department of Biology, East China Institute of Technology, Fuzhou Jiangxi 344000, China

摘要 从一株具有极强的降解羽毛能力的弗氏链霉菌菌株(*Streptomyces fradiae* var. k11)中纯化得到了一种丝氨酸蛋白酶 SFP2。经蛋白测序,得到部分氨基酸序列,设计简并引物,PCR 扩增得到部分基因序列,通过构建基因文库,获得了包括信号肽序列在内的完整的基因 *sfp2* (EMBL 收录号 AJ784940)。开放阅读框全长 924bp,包括 114bp 的信号肽编码序列和 810bp 的酶原编码序列,其中成熟蛋白编码基因长 576bp,编码 191 个氨基酸,理论分子量为 19.112kD。酶原编码基因和成熟蛋白编码基因均在大肠杆菌和枯草芽孢杆菌中得到了表达,酶原编码基因表达产物具有正常的生物学活性,证明了克隆基因的生物学功能。

关键词 弗氏链霉菌,丝氨酸蛋白酶,基因克隆,表达

中图分类号 Q786 文献标识码 A 文章编号 1000-306X(2005)05-0782-07

Abstract Extracellular serine protease SFP2 from *Streptomyces fradiae* var. k11 with high feather-degrading activity was purified. The partial amino acid sequences of internal peptide of purified SFP2 were determined, and the partial gene encoding SFP2 was cloned by PCR using the degenerate primers designed according to the amino acid sequences. Complete *sfp2* gene was cloned by screening the genomic DNA library of *Streptomyces fradiae* var. k11. The Open Reading Frame of *sfp2* including pre-pro-enzyme is 924bp long (EMBL Accession number: AJ784940). The signal peptide sequence is as long as 114bp, the precursor sequence is 810bp and the mature enzyme is 576bp long, encoding 191 amino acid residues with the putative molecular weight of 19.112kD. In *E. coli* and *Bacillus subtilis*, the two sequences encoding SFP2 pro-enzyme and mature enzyme were both expressed successfully. The pro-enzyme expressed had normal biological function and its mature product had normal enzymatic activity.

Key words *Streptomyces fradiae* var. k11, serine protease, gene cloning, expression

Received: March 25, 2005; Accepted: May 25, 2005.

This work was supported by the grants from Chinese National Programs for High Technology Research and Development(863 X No. 2003AA214030) and Chinese Key International S&T Cooperation Projects(No. 2004DF100229).

* Corresponding author. Tel: 86-10-68975126; E-mail: yaobin@public3.bta.net.cn

国家高技术研究与发展计划(863计划)项目(No. 2003AA214030)和国防科技合作重点项目计划(No. 2004DF100229)资助。//journals.im.ac.cn

链霉菌(*Streptomyces*)属于原核生物放线菌,其中灰色链霉菌(*S. griseus*)、变铅青链霉菌(*S. lividans*)和弗氏链霉菌(*S. fradiae*)等链霉菌可产生多种蛋白酶^[1]。Mori-hara^[2]和 Nickerson^[3]等人报道了弗氏链霉菌能分泌多种蛋白酶,包括具有角蛋白酶活性的丝氨酸蛋白酶、一种胰蛋白酶等。1991年 Uma等从弗氏链霉菌中纯化了两种丝氨酸蛋白酶,其中一种具有角蛋白酶活性^[4]。1994年 Kitadokoro等纯化了弗氏链霉菌分泌的一种丝氨酸蛋白酶 Sfase-2,研究了其酶学性质,发现其有较强的降解角蛋白的能力,测定了其N端和C端氨基酸序列,并据此设计引物通过PCR扩增了其部分DNA片段,加上N端和C端氨基酸序列测定结果,推测出其成熟蛋白的氨基酸序列^[5]。2005年3月 Kitadokoro等又更新了他们的网上注册信息(Accession number P41140),对 Sfase-2成熟蛋白的174个氨基酸的二级结构进行了注释,但他们的工作只得到了成熟蛋白的部分基因序列(编码成熟蛋白的576bp中的492bp),蛋白酶的pro肽、信号肽编码序列及基因侧翼序列均未得到,而且未验证基因功能。本研究从一株具有很强的降解羽毛能力的弗氏链霉菌变种 *Streptomyces fradiae* var. k11 中纯化出一种丝氨酸蛋白酶 SFP2,进一步克隆了其完整的编码基因,并对其进行了表达,验证了基因功能。

1 材料和方法

1.1 实验材料

弗氏链霉菌 k11(*Streptomyces fradiae* var. k11), 枯草芽孢杆菌蛋白酶缺陷型菌株 DB104(*his nprE18 nprR2ΔaprA3*), 大肠杆菌(*Escherichia coli*) JM109、BL21(DE3), 质粒 pUC19、pHY300PLK、pET-22b(+) 均为本实验室保存; 枯草芽孢杆菌表达载体 pHY500A 为本实验室以质粒 pAMY40X(南开大学蒋如璋老师惠赠)^[6]和 pHY300PLK(TaKaRa公司产品)为基础构建得到(组成型启动子,四环素抗性标记); 实验所用的各种限制性酶、Klenow 大片段酶购自 TaKaRa 公司; T4 DNA 连接酶购自 Gibco 公司; Taq 酶为 TaKaRa 公司的 *TakaRa LA TaqTM* (with GC Buffer) 产品; 其他常规试剂均为国产分析纯试剂。

1.2 培养基及培养条件

1%脱脂牛奶平板: 1%脱脂奶粉, 1%琼脂溶于蒸馏水中, 自然 pH 值。0.3%羽毛培养基: 0.3%羽毛, 0.1% 尿素, 0.1% KH₂PO₄, 0.1% (NH₄)₂SO₄, 自然 pH 值。先将弗氏链霉菌 k11(*Streptomyces fradiae*

var. k11) 接种于高氏 1 号培养基, 使菌生长到对数期, 离心收集菌体, 高密度接种于 0.3% 羽毛培养基, 培养 36~40 h, 蛋白酶活性达到最高值时离心, 取上清, 用作蛋白酶纯化的粗酶液。蛋白酶活性测定按文献[7]方法略为改动(蛋白酶反应温度改为 50℃)。

1.3 蛋白酶的纯化

粗酶液经分子量截留为 5kD 的超滤膜包超滤浓缩并将缓冲溶液置换为 0.025 mol/L 的 Tris-HCl 缓冲溶液(pH9.0), 得到浓缩酶液。将浓缩酶液上 DEAE 阴离子层析柱, 用 0.025 mol/L 的 Tris-HCl 缓冲溶液(pH9.0)洗脱平衡柱子, 然后用相同缓冲溶液配制的 0~1mol/L 的 NaCl 线性梯度洗脱, 分部收集, 将检测有蛋白酶活性的收集样合并后超滤浓缩, 再经过分子筛 Superdex 75 HR 10/30 (Amersham Pharmacia Biotech 预装柱) 纯化, 用 0.020 mol/L 的 Tris-HCl 缓冲溶液(pH8.2)洗脱, 分部收集洗脱峰, 检测其蛋白酶活性, 得到纯化的 SFP2, 超滤浓缩后备用。取少量纯化样品进行 SDS-PAGE 分析。

凝胶电泳及考马斯亮蓝 R250 染色、银染主要根据 Maniatis 手册^[8]进行。酪蛋白活性染色按以下方法进行: 电泳结束后, 将胶取下, 置 50℃ 预热好的 1% 的酪蛋白溶液浸泡处理 20min, 然后进行常规的考马斯亮蓝 R250 染色、脱色处理。

1.4 蛋白酶的氨基酸序列测定

纯化得到的蛋白酶样品, 经 SDS-PAGE, 用考马斯亮蓝 R250 染色后送军事医学科学院国家生物医学分析中心进行胰蛋白酶酶解后的内肽测序。

1.5 蛋白酶基因部分序列的克隆

主要根据 Maniatis 手册^[8]进行。根据蛋白酶成熟蛋白氨基酸序列测序结果及链霉菌密码子使用的偏爱性^[9], 设计简并引物 P1 5' GAGGCCATCTACGC CGCSGG 3', P2 5' GTCTGCCCGACGTASCGGTT 3', 以弗氏链霉菌 k11(*Streptomyces fradiae* var. k11) 总 DNA 为模板, 进行 PCR 扩增。

PCR 反应参数: 98℃ 变性 3min, 94℃ 变性 1min, 60℃ 退火 1min, 72℃ 延伸 1min, 30 个循环后 72℃ 保温 10min。PCR 产物回收后克隆到 pGEM-T 载体上送博亚生物技术有限公司进行序列测定。

1.6 基因文库的构建和筛选

基因文库的构建采用半补齐法^[10]。Sau3A I 局部消化弗氏链霉菌 k11 染色体 DNA, 琼脂糖电泳回收 5~12kb 的 DNA 片段, 用 dATP、dGTP 在 Klenow 大片段酶作用下将末端 5'-GATC 3' 补上 2 个碱基, 但仍

留有 2 个碱基的 5' 突出(5'GA3')。Sal I 消化质粒 pUC19, 由 Klenow 大片段酶用 dCTP, dTTP 作末端半补齐。两种半补齐产物经 T4 DNA 连接酶连接, 转化 *E. coli* JM109 感受态细胞, 在含有 Amp 的 LB 平板上培养, 用蓝白斑法筛选含有外源基因的重组阳性克隆, 构建基因文库。将转化子平行转接二套同样的平板, 每个平板接种 100 个单菌落。用菌落 PCR 法筛选构建好的基因文库。将其中一套平板上的菌落以适量的 LB 培养液冲洗到 EP 管中, 每管中含有 100 个菌落组成的混合菌液, 先以 1000 个克隆的混合菌液(10 管)为模板进行 PCR 筛选, 结果为阳性的再以 100 个克隆(1 管)为单位进行筛选, 确定某管为阳性后, 从对应的另一套平板挑菌(每 10 个克隆挑到一管中)进行 PCR, 最后以单个克隆为模板进行 PCR 筛选, 筛出阳性克隆。所得阳性克隆送博亚生物技术有限公司测序。

1.7 丝氨酸蛋白酶 SFP2 的表达

大肠杆菌重组表达载体的构建及 SFP2 的诱导表达: 根据网上数据库分析及同源性比较确定 SFP2 的信号肽及 pro 肽。分别用 PCR 的方法获得除去信号肽编码序列的 SFP2 成熟蛋白基因和酶原基因, 利用引物上设计的 *Nco* I 和 *Hind* III 酶切位点克隆进表达载体 pET-22k(+) 中, 电击转化 BL21(DE3) 感受态细胞, 筛选出阳性转化子, 与对照[含质粒 pET-22k(+)]的 BL21(DE3) 同时接种于含 Amp(终浓度为 100 $\mu\text{g}/\text{mL}$) 和 IPTG(终浓度为 15 $\mu\text{mol}/\text{L}$) 的 1% 脱脂牛奶平板。同时接种于含 Amp(终浓度为 100 $\mu\text{g}/\text{mL}$) 的 LB 液体培养基中, 37 $^{\circ}\text{C}$ 振荡培养过夜, 次日按 1% 接种量转接, 37 $^{\circ}\text{C}$ 培养至 OD_{600} 为 0.6 ~ 0.8, 用终浓度为 15 $\mu\text{mol}/\text{L}$ 的 IPTG 于 28 $^{\circ}\text{C}$ 诱导表达 4h, 收集上清液, 检测蛋白酶活性。用 2 倍体积的丙酮沉淀上清液中的蛋白, 所得沉淀溶解于 0.025 mol/L 的 Tris-HCl 缓冲溶液(pH9.0)中, 进行 SDS-PAGE 分析。将 IPTG 诱导表达 4h 的 BL21-*sfp2pro*(含重组质粒 pET-22b-*sfp2pro*) 的上清液, 经超滤浓缩, 阴离子柱层析纯化后, SDS-PAGE 分析。

枯草芽孢杆菌重组表达载体的构建及 SFP2 的诱导表达: 利用引物上设计的 *Nhe* I 和 *Bam* HI 酶切位点, 将 SFP2 成熟蛋白基因和酶原基因分别克隆到组成型表达载体 pHY500A 中(图 1、图 2), 电击转化枯草芽孢杆菌蛋白酶缺陷型菌株 DB104 感受态细胞, 枯草芽孢杆菌转化方法参照 Gary 等方法略加改进^[11]。筛选出阳性转化子, 与对照(含质粒 pHY500A 的 DB104) 同时接种至含有四环素(终浓度

为 30 $\mu\text{g}/\text{mL}$) 的 1% 脱脂牛奶平板。同时接种于含四环素(终浓度为 30 $\mu\text{g}/\text{mL}$) 的 LB 液体培养基中, 37 $^{\circ}\text{C}$ 振荡培养过夜, 次日按 5% 接种量转接, 37 $^{\circ}\text{C}$ 培养, 分时取上清液检测其蛋白酶活性。

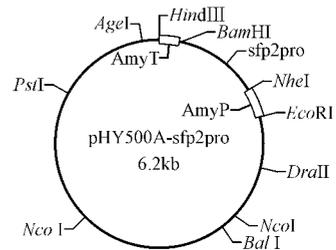


图 1 重组表达质粒 pHY500A-*sfp2pro* 的物理图谱

Fig. 1 Physical map of expression plasmid pHY500A-*sfp2pro*
AmyP: Fusion of α -amylase promoter with sequence encoding signal peptide, AmyT: Sequence encoding transcriptional termination signal of *Bacillus licheniformis* α -amylase.

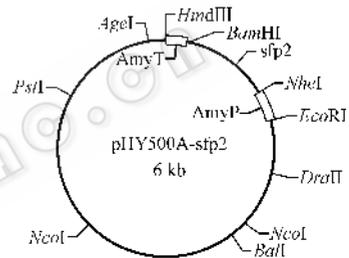


图 2 重组表达质粒 pHY500A-*sfp2* 的物理图谱

Fig. 2 Physical map of expression plasmid pHY500A-*sfp2*
AmyP: Fusion of α -amylase promoter with sequence encoding signal peptide, AmyT: Sequence encoding transcriptional termination signal of *Bacillus licheniformis* α -amylase.

2 结果与分析

2.1 丝氨酸蛋白酶 SFP2 的纯化

粗酶液经超滤浓缩、离子交换层析、凝胶层析纯化后, SDS-PAGE 结果(图 3)表明, 未经煮沸变性的纯化样品银染和酪蛋白活性染色均显示单一的条带。通过样品煮沸变性的 SDS-PAGE 测得其分子量约为 17.6 kDa(图 4)。

2.2 丝氨酸蛋白酶 SFP2 部分编码基因序列的克隆

经氨基酸序列测定, 纯化的蛋白酶成熟蛋白测得的三段内肽的氨基酸序列分别为: IAGGEAIYAAGGGR, AGTSFPGNDYGLIR, DITGAGNAYVQGTVQR。

根据氨基酸序列测定结果及链霉菌密码子使用偏爱性^[9]设计引物, PCR 扩增出约 300bp 的 DNA 片段, 回收 PCR 产物, 克隆到 pGEM-T 载体进行序列测定。结果表明(图 5), 根据这段 DNA 序列推测得到的氨基酸编码序列中含有另一段已测定的氨基酸序

列“AGTSFPGNDYGLIR”,说明 PCR 扩增所得到的片段是我们拟克隆的基因。

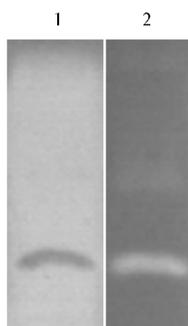


图3 丝氨酸蛋白酶 SFP2 纯化的 SDS-PAGE 分析
(样品未经煮沸变性)

Fig.3 SDS-PAGE analysis of purified SFP2
(Native purified SFP2)
1: stained with silver; 2: activity staining.

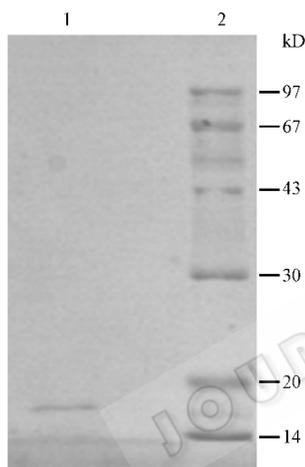


图4 丝氨酸蛋白酶 SFP2 纯化的 SDS-PAGE 分析
(样品经煮沸变性)

Fig. 4 SDS-PAGE analysis of purified SFP2
(SFP2 denaturalized under 100°C.)
1: purified SFP2; 2: standard protein molecular weight.

LAGGEATYAA GGGRCCSLGFN VRSSSGATYA LTAGHCTEIA STWYTNSSGQT
SLLGTRAGTS FPGNDYGLIRHSNASAADGR VVLYNGSYRDTITGAGNAYVVG QTVQR

图5 PCR 获得的 SFP2 部分基因序列推测的氨基酸序列
(方框内为氨基酸测序得到的内肽序列)

Fig.5 Deduced amino acid sequence from partial DNA
sequence obtained by PCR (The amino acids sequences
in pane represented sequences determined
by protein sequencing of SFP2)

2.3 基因文库的构建和筛选

一次连接反应的 $10\mu\text{L}$ 连接体系分 5 次转化 *E. coli* JM109 感受态细胞,共得到 16 800 个克隆,其中 99% 为白斑。随机挑取 10 个白斑提质粒鉴定,

阳性率为 90%,转入的 DNA 片段大小分布于 5 ~ 12kb 之间。所插入的外源基因片段平均大小约为 8kb 左右。以链霉菌的基因组大小为 9 000kb 计算,根据公式 $N = \ln(1 - p) / \ln(1 - f)$ 计算,达到 99% 的覆盖率,需 5 179 个克隆。所得到的克隆数大大超过此数,因此文库的构建是成功的。

公式中 N 为一个完全基因文库所应包含的重组 DNA 的转化子克隆数; p 为在重组体群体中出现目的基因序列的几率; f 为限制片段的平均大小与基因组 DNA 总量之比。

根据已知的 SFP2 部分 DNA 序列设计引物,从构建的基因文库中用菌落 PCR 法筛选了约 6 200 个白色重组子,经过 4 轮 PCR,从中筛选到两个重组子含有 SFP2 蛋白酶基因。测序结果及推测的氨基酸序列如图 6 所示。

分离得到的蛋白酶基因 *sfp2* (EMBL 收录号 AJ784940) 含一个全长 924bp 的开放阅读框,成熟蛋白基因 576bp, G + C 含量为 72%,编码 191 个氨基酸,理论分子量为 19.112kD。图 6 中粗斜体所示的 1-114bp 为信号肽编码序列,起始密码子为 GTG, 115-348bp 为 pro 肽编码序列,349-924bp 为成熟蛋白酶的编码序列。

2.4 丝氨酸蛋白酶基因 *sfp2* 的表达

2.4.1 在大肠杆菌中的表达:将丝氨酸蛋白酶 SFP2 成熟蛋白基因 *sfp2mat* 和酶原基因 *sfp2pro* 以正确的阅读框架克隆到表达载体 pET-22k(+) 上的信号肽编码序列 *pelB* 之后,得到重组质粒 pET-22b-*sfp2mat* 和 pET-22b-*sfp2pro*,受 T7 *lac* 启动子驱动表达。

在 1% 脱脂牛奶平板上,对照 BL21-pET-22b(+) (含质粒 pET-22k(+)) 和 BL21-*sfp2mat* (含重组质粒 pET-22b-*sfp2mat*) 不形成降解牛奶的透明圈,而 BL21-*sfp2pro* (含重组质粒 pET-22b-*sfp2pro*) 则降解牛奶形成明显的透明圈,表明其具有分泌蛋白酶活性(图 7)。

接种于 LB 液体培养基中的对照 BL21-pET-22b(+) 和 BL21-*sfp2mat* 经 IPTG 诱导后,其培养液上清中未检测出蛋白酶活性,而 BL21-*sfp2pro* 则检测到明显的蛋白酶酶活(22u/mL)。样品未经煮沸变性的 SDS-PAGE 检测结果表明(图 8),IPTG 诱导后的 BL21-*sfp2pro* 培养液上清浓缩液,考染和酪蛋白活性染色均显示一条特异性蛋白条带。对照 BL21-pET-22k(+) 则没有这一蛋白条带,这进一步表明 SFP2 蛋白酶原基因得到了正确表达,且表达产物具有正

常的生物学活性。

GCCCCACCCTTCCCCACCGGAGGATGCA
 1 *G*TGAACCTCAAGCGCTTCAACCCCGCGGGACTCGCGAGAGCGCGGGTGCACGCC
 ValAsnLeuLysArgPheThrProArgGlyGlyLeuAlaArgGlyAlaArgLeuThrAla
 61 *T*GCCCGCCCGCTGGTCAACGCCACCGCGCTTCCGCCCCGAGCGCGCGCGGAGACC
 ValAlaAlaAlaLeuValThrAlaThrAlaLeuAlaAlaProSerAlaGlyAlaGluThr
 121 GCCGACGCGACCCGGGCGAGCTGCGGAGCTGGCGCGCTCAGCGACGCCGTGCTCGAC
 AlaAspAlaThrArgAlaSerValAlaGluLeuAlaArgValSerAspAlaValLeuAsp
 181 GCCGACGTGCCGGCACCCGCTGGTACACCGACGCCGAGAGCGGCAAGTGGTGGTCAAC
 AlaAspValProGlyThrAlaTrpTyrThrAspAlaGluSerGlyLysLeuValValThr
 241 GCCGACGCCACCGTGTCCGCCCGGAGCTGGCCAGCTGAAGAAGCGCGGGCGACAAG
 AlaAspAlaThrValSerAlaAlaGluLeuAlaGlnLeuLysLysAlaAlaGlyAspLys
 301 GCCGGCCCGTGGAGATCAAGCGCACCCCGGGCACCTTCAACAAGCTCATCGCCGGCGGC
 AlaGlyAlaValGluIleLysArgThrProGlyThrPheAsnLysLeuIleAlaGlyGly
 361 *G*AGGCCATCTACGCGCGGGCGGGCGGCTTGGCTCCCTCGGCTCAACGTCFCGCGAGCAGC
 GluAlaIleTyrAlaAlaGlyGlyGlyArgCysSerLeuGlyPheAsnValArgSerSer
 421 *A*GCGGCGCGACCTACGCCCTGACGGCCGGCCATGCACGGAGATCGCCTCCACCTGGTAC
 SerGlyAlaThrTyrAlaLeuThrAlaGlyHisCysThrGluIleAlaSerThrTrpTyr
 481 *A*CGAACCCGCGCAGACCTCGCTGCTCGGCACCCGTCGCGGCAGCAGCTTCCCAGGCAAC
 ThrAsnSerGlyGlnThrSerLeuLeuGlyThrArgAlaGlyThrSerPheProGlyAsn
 541 *G*ACTACGGCCTGATCGCCACTCCAACGCGTCCGCGCGGAGCGCGCGTGTACCTGTAC
 AspTyrGlyLeuIleArgHisSerAsnAlaSerAlaAlaAspGlyArgValTyrLeuTyr

601 *A*ACGGCTCGTACCAGGACATCACCGCGCCGGCAACGCCCTACGTGGGCCAGACCGTCCAG
 AsnGlySerTyrArgAspIleThrGlyAlaGlyAsnAlaTyrValGlyGlnThrValGln
 661 *C*GCAGCGGCTCCACGACCCGGTCTGCACAGCGCCGCGTGCACGGGCCATCAACGCCACGGTC
 ArgSerGlySerThrThrGlyLeuHisSerGlyArgValThrGlyLeuAsnAlaThrVal
 721 *A*ACTACGGCGCGGCGACATCGTCTCCGGCTGATCCAGACCAACGCTGCGCGGAGCCC
 AsnTyrGlyGlyGlyAspIleValSerGlyLeuIleGlnThrAsnValCysAlaGluPro
 781 *G*CGCAGACGCGCGGCCCTGTTCGCGCGGCTCCACCGCCCTCGGCCGTACCTCCGGCGGC
 GlyAspSerGlyGlyAlaLeuPheAlaGlySerThrAlaLeuGlyLeuThrSerGlyGly
 841 *A*GCGGCAACTGCCGTACGGCGCGCACCCAGTCTTCCAGCCCGTCAACGAGGCCCTCAGC
 SerGlyAsnCysArgThrGlyGlyThrThrPhePheGlnProValThrGluAlaLeuSer
 901 *G*CGTACGGCGTCAAGCATCATCTGACCCCGCCACCGCCGGCATCCGCGCCGGTGACCCG
 AlaTyrGlyValSerIleIle *
 TCCGTGCGGGCCACCGGGCCGACCCGTGCCGGGAGGCCGCCCGCTCCCGCGGGGGTCC
 CCGCGCGGGCGGGCGGGCCCGCGCGCGGGCCCGCGGGCCAGGCGCGGGCTCCGTCC
 AGCGGGCGGGGGCGCCACACAGAGGGCCGTGCCCGTGTACGCGGGAGCGTCCGGCGCT
 CGGTGAAGCGGACGACGGGGCGCGGCTGTGGTCTGTCGCGCTCGTGAAGGCCACCGAGT
 CCGACTTGCTCATCCGGAGCTGGATCTCCCCGGTGTCCCGGCCCGCCAGGGTCCC
 CGCGAAGCCGACCTCCAGGTAGCGTCCGGCCCGCGACCCGGGGCGGACAGCGGGACGA
 CCGGAGCCCGCACCCCGCGCACCCGAGGGCGCGGTGGTGCACACCGCCCTGACCGAGG
 GGGCGCCCGTCCGGGTAAGTAGTAGCGGGCGGTACGGCCCGAGGTCGACGGCGG
 CGCGCCGGTGTGGCGATCCCGAGCTCGGGCGGATCGCGTGTTCGGTGAGGAGCCGT
 CGGTGGTGC

图 6 来源于 *Streptomyces fradiae* var. k11 的 *sfp2* 核苷酸序列及推导出的氨基酸序列 (斜体部分代表信号肽,下划线部分代表成熟蛋白,两者之间部分为 Pro 肽)

Fig.6 The DNA sequence of *sfp2* from *S. fradiae* var. k11 and the deduced amino acid sequence (The italic part represented signal peptide, the underline part represented the mature enzyme the part between the italic part and the underline part represented the pro-peptide)

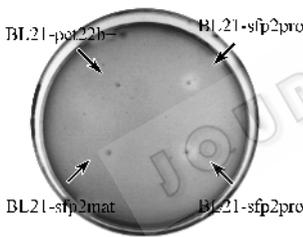


图 7 1%脱脂牛奶平板检测重组大肠杆菌蛋白酶酶活

Fig.7 Analysis of protease activity on the 1% deaged milk plate

BL21-pET-22k(+) : *E. coli* with plasmid pET-22k(+) BL21-*sfp2mat* : *E. coli* with plasmid pET-22b-*sfp2* BL21-*sfp2pro* : *E. coli* with plasmid pET-22b-*sfp2pro*.

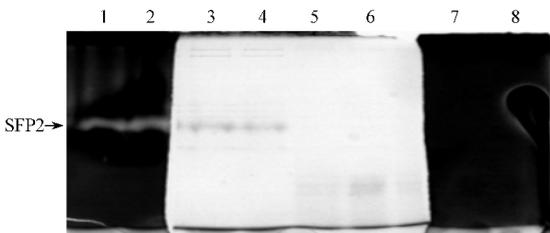


图 8 表达的 SFP2 的 SDS-PAGE 和活性染色

Fig.8 SDS-PAGE analysis and activity gel-staining of SFP2 expressed in *E. coli*

1~4 : medium supernatant of BL21-*sfp2pro* induced by IPTG ; 5~8 : medium supernatant of BL21-pET-22k(+) induced by IPTG.

将 IPTG 诱导表达 4h 的 BL21-*sfp2pro* 的上清液, 经超滤浓缩 阴离子柱层析纯化后, 进行样品煮沸变性的 SDS-PAGE 分析, 结果表明(图 9):表达的有蛋白酶活性的蛋白酶分子量大小约为 17.6kD, 明显小于蛋白酶原的理论分子量而与成熟蛋白酶分子量相当, 说明蛋白酶原的 pro 肽部分在蛋白酶成熟过程中被切除了。

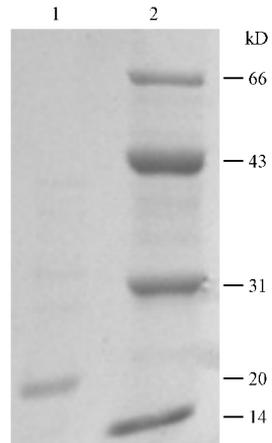


图 9 SFP2 酶原基因表达的 SDS-PAGE 分析

Fig.9 SDS-PAGE analysis of SFP2 pro-protein expressed in *E. coli*

1 : purified enzyme ; 2 : molecular mass standards.

SFP2 成熟蛋白编码基因和酶原编码基因以正确的阅读框架克隆到表达载体 pHY500A 的地衣芽孢杆菌 α 淀粉酶启动子和信号肽之后,得到重组质粒 pHY500A-*sfp2* 和 pHY500A-*sfp2pro*。

在 1% 脱脂牛奶平板上,对照 DB104pHY500A (含质粒 pHY500A) 不形成降解牛奶的透明圈, DB104*sfp2mat* (含重组质粒 pHY500A-*sfp2*) 仅形成微弱的透明圈,而 DB104*sfp2pro* (含重组质粒 pHY500A-*sfp2pro*) 则降解牛奶形成极明显的透明圈,表明其具有很强的分泌蛋白酶活性(图 10)。

接种于 LB 培养基中的对照 DB104pHY500A,其培养液上清未检测出蛋白酶活性, DB104*sfp2mat* 检测到极弱的蛋白酶活性,而 DB104*sfp2pro* 则检测到明显的蛋白酶酶活(7u/mL)。

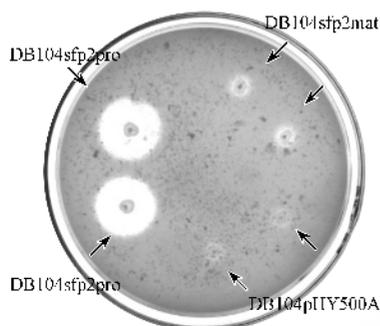


图 10 1% 脱脂牛奶平板检测重组菌的蛋白酶酶活

Fig. 10 Analysis of protease activity on the 1% deproteinized milk plate

DB104pHY500A : DB104 with plasmid pHY500A DB104*sfp2mat* : DB104 with plasmid pHY500A-*sfp2* DB104*sfp2pro* : DB104 with plasmid pHY500A-*sfp2pro*.

3 讨论

本研究中,由于样品中含有多种蛋白酶,其中仅丝氨酸蛋白酶就不只一种,故纯化过程中未计算酶的比活的提高倍数和得率。运用银染和蛋白酶活性染色均证明纯化所得样品为单一条带,说明所得样品纯度较高,可直接用于蛋白测序。

应用半补齐法建库,有效消除了载体的自身环化和插入片段自身彼此退火的能力,极大提高了连接产物中重组质粒的比例,使组成基因文库的总转化子数目大大减少,减轻了文库筛选的工作量。研究结果证明了此方法的正确性。

采用菌落 PCR 法筛选文库,简便,快捷,一个人一至二个工作日即可完成数千乃至上万个克隆的筛选,筛选的灵敏性和准确率也很好。与用核酸探针杂交筛选方法相比,既省时省力,又避免了放射性污

染,筛选结果可靠性高(我们用此法筛选到了 *Streptomyces fradiae* var. k11 中包括 *sfp2* 在内的 6 种蛋白酶基因的阳性克隆子,经 DNA 测序验证,全部是所要获得的克隆)。因此,在克隆到所筛基因的部分序列的情况下,采用此法筛选文库,是一种理想的选择。

蛋白酶的异源表达较困难,一方面由于蛋白酶的表达可能对受体菌的生长繁殖产生不利影响,另一方面受体菌自身蛋白酶对表达产物有影响,而且增加了表达蛋白酶检测的困难,这也是很多蛋白酶基因未能进一步验证其生物学功能的原因。本研究中,分别采用了革兰氏阳性菌枯草芽孢杆菌(蛋白酶缺陷型)和革兰氏阴性菌大肠杆菌为受体菌,成功地表达了 *sfp2* 基因,验证了所克隆基因的生物学功能。

研究发现,无论是在大肠杆菌还是在枯草芽孢杆菌中,表达成熟蛋白基因的蛋白酶酶活很低,而表达蛋白酶原基因的蛋白酶酶活则高得多,这可能是因为蛋白酶的 pro 肽对蛋白酶的正确折叠和运输起重要作用。Shih 等人在对角蛋白酶基因 *KerA* 进行表达时,也是采用表达其蛋白酶原基因的策略,并取得了成功^[12,13]。

克隆的 *sfp2* 其成熟酶的氨基酸序列与已报道的 Sfase-2 同源性高达 99.5% 基本上可认为是同一蛋白酶。以天青角蛋白(keratin azure)作为底物测定在大肠杆菌中表达的蛋白酶 SFP2 的部分酶学性质,结果表明,SFP2 的蛋白酶活性可被丝氨酸蛋白酶抑制剂 PMSF 所抑制,SFP2 是一种丝氨酸蛋白酶。SFP2 的最适 pH 为 9.5,最适反应温度为 60℃,具有较强的角蛋白酶活性,而且还原剂的存在,能极大地提高其角蛋白酶活性,表现出对角蛋白具有较强的底物专一性(详情另文发表)。SFP2 具有较高的角蛋白酶活性,是一极具应用潜力的蛋白酶,但 Kitadokoro 等工作只得到了成熟蛋白的部分基因序列(编码成熟蛋白的 576bp 中的 492bp)^[5],蛋白酶的 pro 肽、信号肽编码序列及基因侧翼序列均未得到,而且未验证基因功能,而蛋白酶的 pro 肽对蛋白酶的活性表达可能具有较重要的作用。我们从蛋白纯化出发,经蛋白测序,调出部分基因序列,通过构建基因文库,首次获得了包括信号肽、pro 肽编码序列在内的完整的基因,并且首次完成了此基因的表达,既验证了基因的功能,也为今后的开发利用打下了良好的基础。

有极强的降解羽毛的能力。角蛋白的降解机制极为复杂,目前的研究还难以得出一个清晰的结论。除蛋白酶以外,可能还涉及到脂肪酶、氧化还原酶类等,我们从此菌中已克隆得到包括 *sfp2* 在内的 6 种蛋白酶基因(其中四个已在 EMBL 注册,注册号为 AJ781828, AJ784159, AJ784940, AJ810091),其中 3 种与报道的蛋白酶基因同源性最高的均在 80% 以下。另外还克隆到一个脂肪酶编码基因。这些基因的克隆,为进一步研究它们在降解角蛋白过程中的时空表达顺序、协同作用机制等打下了基础。

REFERENCES(参考文献)

- [1] Chen K(陈 侃), Wang YQ(王以光). Protases from *Streptomyces* World Notes on Antibiotics(国外医药抗生素分册), 2002, **23** (2): 67-72
- [2] Morihara K, Oka T, Tsuzuki H. Multiple proteolytic enzymes of *Streptomyces fradiae*. Production, isolation, and preliminary characterization. *Biochim Biophys Acta*, 1967, **139**: 382-397
- [3] Nickerson WJ, Noval JJ, Robinson RS *et al.* Properties of the enzyme conjugate elaborated by *Streptomyces fradiae*. *Biochim Biophys Acta*, 1963, **77**: 73-86
- [4] Uma Sinha, Sarah AW, Pushkaraj JL. Two new extracellular serine proteases from *Streptomyces fradiae*. *Int J Biochem*, 1991, **23** (10): 979-984
- [5] Kitadokoro K, Tsuzuki H, Nakamura E *et al.* Purification, characterization, primary structure, crystallization and preliminary crystallographic study of a serine proteinase from *Streptomyces fradiae* ATCC 14544. *Eur J Biochem*, 1994, **220**(1): 55-61
- [6] Xin K(辛 丰), Jiang RZ(蒋如璋). Construction of a Secretion-expression vector of *Bacillus*. *Chinese Journal of Biotechnology*(生物工程学报), 1992, **8**(1): 40-47
- [7] Lab of biochemistry and microorganism, department of biology, sun yat-sen university(中山大学生物系生化微生物教研室). Introduction to Biochemical Technique(生化技术导论). People's Education Press(人民教育出版社), 1978
- [8] Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: A laboratory manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989
- [9] Frank Wright, Mervyn J Bibb. Codon usage in the G + C-rich *Streptomyces genome*. *Gene*, 1992, **113**(1): 55-65
- [10] Hong YQ(洪义国), SUN M(孙 谧). Cloning and sequencing of the marine low-temperature alkaline protease gene from *F. yellow sea* YS-9412-130. *Marine Fisher Research*(海洋水产研究) 2002, **21**(4): 46-53
- [11] Gary MD, Linda NL, Donald JL. Improved electroporation and cloning vector system for Gram-positive bacteria. *Applied and Environmental Microbiology*, 1991, **57**(4): 1194-1201
- [12] Lin X, Wong SL, Miller ES *et al.* Expression of the *Bacillus licheniformis* PWD-1 keratinase gene in *B. subtilis*. *Journal of Industrial Microbiology and Biotechnology*, 1997, **19**(2): 134-138
- [13] Wang JJ, Rojanatavorn K, Shih JC. Increased production of *Bacillus* keratinase by chromosomal integration of multiple copies of the *kerA* gene. *Biotechnol Bioeng*, 2004, **87**(4): 459-464