

## 苏云金杆菌毒性蛋白基因5'端的 PCR 合成、克隆及序列测定

郑 琦<sup>1</sup> 刘中柱<sup>1</sup> 陈章良<sup>2</sup>

(<sup>1</sup>福建省农科院生物技术研究中心, 福州 350003;

<sup>2</sup>北京大学蛋白质工程和植物基因工程国家重点实验室, 北京 100871)

**摘 要** 以苏云金杆菌戈尔斯德亚种 (*Bacillus thuringiensis* var. *kurstaki* HD-1) 的总 DNA 为模板, 用多聚酶链反应 (PCR) 的方法, 成功合成了该菌毒性蛋白基因 5' 端片段, 并克隆到大肠杆菌 Bluescript 质粒载体上, 获得了重组质粒 pQC20。核酸限制酶谱分析表明, pQC20 的重组片段属于 Kronstad 的毒性蛋白基因限制酶片段长度多型性分类系统中 5.3 Kb 类型基因。对该基因片段的全序列分析结果证明, 在以 PCR 合成的 1959 个核苷酸中与国外已发表的 5.3 Kb 类型基因的相应片段仅有一个核苷酸差异。该基因片段具有表达有生物活性的毒性蛋白所必需的全部核苷酸序列。

**关键词:** 苏云金杆菌毒性蛋白; PCR; 基因克隆; DNA 序列测定

### PCR Synthesizing, Cloning and Sequencing of 5' End of BT Toxin Gene

Zheng Qi<sup>1</sup>, Liu Zhongzhu<sup>1</sup> and Chen Zhangliang<sup>2</sup>

(<sup>1</sup>Institute of Biotechnology, Fujian Academy of Agri. Sci., Fuzhou 350003;

<sup>2</sup>National Lab. of Protein Engineering and Plant Genetic Engineering, Beijing University, Beijing 100871)

**Abstract:** By Polymerase Chain Reaction (PCR), the 5' end of *Bacillus thuringiensis* var. *kurstaki* HD-1 toxin gene has been synthesized, using the total DNA of the bacterium as template. The recombinant plasmid pQC20 has been obtained, cloning the PCR fragment into the Bluescript plasmid vector in *E. coli*. According to the restriction map, the insert of pQC20 belongs to the 5.3 Kb type gene in the Kronstad's classification system for restriction fragment of length polytype of toxin protein gene. The sequencing result showed that, the synthesized 1959 bp nucleotides are almost the same as the published sequence of 5.3 Kb type toxin gene with only one-nucleotide difference. This 5' end fragment contains the complete sequence for encoding the activated toxin.

**Key Words:** BT toxin; PCR; Cloning; DNA sequencing

苏云金杆菌在其芽孢形成期能产生对许多昆虫具有特异性毒杀作用的伴孢晶体蛋白<sup>[13]</sup>。随着植物基因工程技术的发展, 将编码该杀虫毒性蛋白的基因导入植物细胞, 培育自身抗虫的基因工程植株, 已在许多实验室取得成功<sup>[2,4,5,12]</sup>。对苏云金杆菌毒性蛋白基因的深入研究表

明,仅为全长毒素基因一半的5'端片段已足以表达有活性的毒性蛋白,且该3'端缺失的基因在转基因植物中的表达效果优于完整的毒素基因<sup>[2,8,10]</sup>。因此,在克隆获得毒素基因之后,有必要对其进行3'端缺失,以获得能在植物细胞中有效表达的5'端毒素基因片段。本文报道通过设计适当的寡核苷酸引物,以PCR合成的方法,直接克隆具有完整起始和终止密码的苏云金杆菌毒素基因5'端的工作。国内已有实验室报道了克隆苏云金杆菌毒性蛋白基因的工作,但未见发表过其DNA序列。因此,我们对所获得的5'端基因片段进行了全序列测定,以确定其与国外报道的同类型基因的核苷酸序列异同,并确证由PCR合成的毒性蛋白基因的可靠性。

## 1 材料与方法

**1.1 菌株和质粒** 苏云金杆菌戈尔斯德亚种HD-1 (*Bacillus thuringiensis* var. *kurstaki* HD-1) 购自中国菌种保藏中心。用作PCR合成模板的DNA按文献[1]方法提取。基因克隆所用质粒载体为Bluescript,为序列测定进行亚克隆所用质粒为Bluescript和pGEM,均按文献[7]方法提纯。受体大肠杆菌菌株为DH-5。克隆与亚克隆的具体方法步骤除以下特别提到的外,均按文献[7]方法进行。

**1.2 生化试剂和仪器** 各种核酸限制性内切酶、T<sub>4</sub>核苷酸激酶、T<sub>4</sub>核苷酸连接酶等均购自华美生物工程公司。PCR试剂盒和序列测定试剂盒Sequenase Version 2.0系Promega公司产品。 $\alpha$ -<sup>32</sup>P-dATP购自NEN公司。90A-PCR仪由中国科学院遗传所技术开发室制造。

**1.3 寡核苷酸引物** 寡核苷酸引物的设计以S. Kondo等(1987)<sup>[11]</sup>发表的苏云金杆菌戈尔斯德亚种HD-1毒性蛋白基因序列为依据。第一引物为5.3Kb基因的起始18个核苷酸序列,即:5'-ATGGATAACAATCCGAAC-3';第二引物选择5.3Kb基因第1942到1959的18个核苷酸的互补序列,并在其5'端增设2个终止密码,即5'-TTATCACTCAACTAAATTGGTATC-3'。寡核苷酸引物由北京大学生命研究中心DNA合成仪合成。

**1.4 PCR合成** PCR合成依文献[9]所述方法进行。反应体系中各成份按试剂盒说明加入。PCR反应条件如下:93℃,0:30;45℃,1:30;65℃,2:30;持续30个循环后继续在65℃保持5分钟。

**1.5 PCR产物的克隆** PCR产物以电泳分离,切下含目的DNA条带的琼脂糖凝胶,通过冻融法回收目的DNA。为提高克隆效率,按文献[3]方法对回收的DNA先用蛋白酶K处理,再以核苷酸激酶在5'端加上磷酸基,按常规方法克隆到Bluescript的EcoR V位点。

**1.6 DNA序列测定** 毒性蛋白V基因5'端片段以核酸限制酶切成6个小片段,分别亚克隆到pBS或pGEM质粒载体的相应位点上,分别以KS、SK、T<sub>7</sub>或SP<sub>6</sub>为测序引物,按Sequenase Version 2.0试剂盒所述双脱氧终止法进行DNA序列测定。

## 2 结果和讨论

**2.1 毒性蛋白基因5'端的PCR合成和克隆** 按“材料和方法”所述的条件、方法,由PCR合成获得了大小约1.9Kb的DNA片段。电泳未见其他的DNA条带,电泳照片见图1。这表明,由两个寡核苷酸引物引导合成的1.9Kb片段系PCR的特异性产物,其大小符合预定的毒性蛋白基因5'端长度。据Crowe等(1991)<sup>[3]</sup>报道,PCR合成片段的克隆成功率很低,原因可能是经PCR合成的DNA片段末端附着残余的Taq多聚酶,阻碍了DNA连接酶发挥作用。他们用蛋白

酶K对PCR合成产物进行预处理后,大大提高了该片段的克隆效率。我们也参照此经验进行蛋白酶K处理,却未能明显提高克隆效率,其部份原因可能是我们设计的寡核苷酸引物没有可利用的限制酶识别位点,只能采用效率较低的平末端连接所致。用煮沸法快速提质粒筛选,从大约一千多个白色菌落中获得一个含有PCR合成片段的重组菌株。对该重组菌的质粒pQC20进行了核酸限制酶谱分析,结果如图2所示。对照Kronstad (1986)<sup>[6]</sup>对苏云金杆菌戈尔斯德亚种毒性蛋白基因的Hind III酶切片长度多型性的分类结果,我们克隆获得的基因属于5.3 Kb类型。由于所设计的寡核苷酸引物是选择5.3 Kb和4.5 Kb两类基因的同源序列,理论上PCR产物中应包含两种类型基因,可能是因为克隆效率低,尚未克隆到4.5 Kb类型基因。

**2.2 苏云金杆菌毒性蛋白基因5'端的DNA序列测定** 根据核酸限制酶谱分析的结果,我们将1.9 Kb的毒性蛋白基因5'端切割为6个小片段,分别亚克隆到pBS或pGEM质粒载体的相应酶切克隆位点,并分别以质粒载体上的KS、SK、T<sub>7</sub>或SP<sub>6</sub> Primer为测序引物进行DNA序列测定。测定结果与S. Kondo (1987)<sup>[11]</sup>等发表的5.3 Kb基因的相应序列进行比较,结果见图3。在1959个核苷酸中,仅在第99位有一个核苷酸差异,由C取代了T,这一差异恰位于密码子的第3个核苷酸,并未导致所编码氨基酸的改变。测序结果表明,PCR合成可以有效地达到约2 Kb长度,而仍能保持严格的复制,至少在本实验中,Taq多聚酶的复制误差远低于人们的预料。测序结果还表明,该基因片段包含了编码有活性的毒性蛋白所需要的全部核苷酸序列,寡核苷酸引物中设置的起始密码和终止密码被忠实地保留。

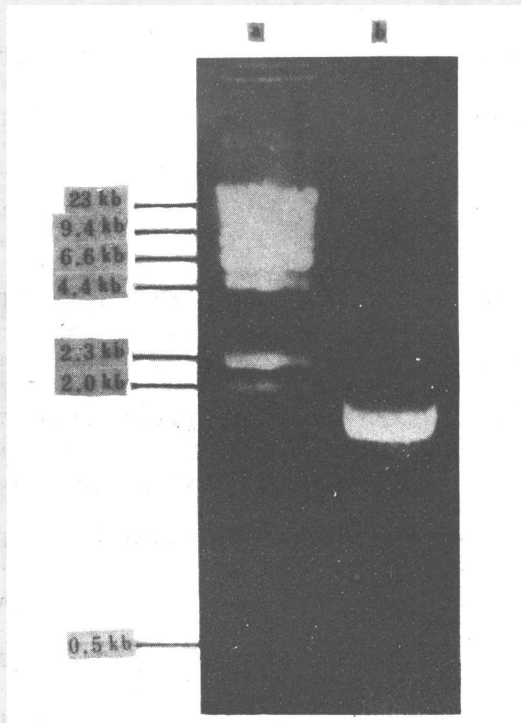


图1 PCR产物的琼脂糖凝胶电泳照片

Fig. 1 Photograph of agarose gel electrophoresis for PCR product

- a)  $\lambda$  Hind III 分子量标准,  $\lambda$  Hind III marker.  
b) PCR 产物, PCR product.

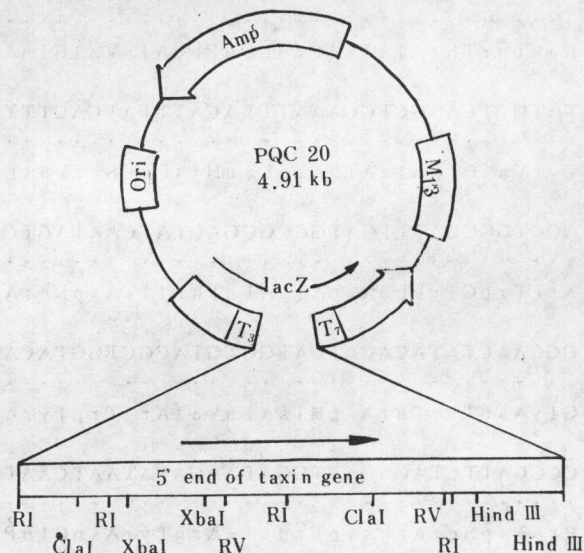


图2 重组质粒pQC20限制酶酶切图谱

Fig. 2 Restriction map of recombinant plasmid pQC20

ATGGATAACAATCCGAACATCAATGAATGCATTCTTTATAATTGTTTAAGTAACCCTGAA ..... Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu Ser Asn Pro Glu	60
GTAGAAGTATTAGGTGGAGAAAGAATAGAACTGGTTACACCCCAATCGATATTTCTTG .....T..... Val Glu Val Leu Gly Gly Glu Arg Ile Gly Thr Gly Tyr Thr Pro Ile Asp Ile Ser Leu	120
TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTGCTGGATTTGTGGGAGGACTA ..... Ser Leu Thr Gln Phe Leu Leu Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu	180
GTTGATATAATATGGGGAATTTTGGTCCCTCTCAATGGGACGCATTTCTTGACAAATT ..... Val Asp Ile Ile Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile	240
GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGAACCAAGCCATTTCTAGATTA ..... Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile Ser Arg Leu	300
GAAGGACTAAGCAATCTTTATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT ..... Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu Ser Phe Arg Glu Trp Glu Ala Asp	360
CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAATTCAATGACATGAACAGTGCC ..... Pro Thr Asn Pro Ala Leu Arg Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala	420
CTFACAACCGCTATTCTCTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA ..... Leu Thr Thr Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val	480
TATGTTCAAGCTGCAAAATTTACATTTATCAGTTTTGAGAGATGTTTCAGTGTTTGGACAA ..... Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser Val Phe Gly Gln	540
AGGTGGGGATTTGATGCCGCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT ..... Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg Tyr Asn Asp Leu Thr Arg Leu Ile	600
GGCAACTATACAGATCATGCTGTACGCTGGTACAATACGGGATTAGAGCGTGTATGGGGA ..... Gly Asn Tyr Thr Asp His Ala Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly	660
CCGGATTCTAGAGATTGGATAAGATATAATCAATTTAGAAAGAGAATTAACACTAACTGTA ..... Pro Asp Ser Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val	720
TTAGATATCGTTTCTCTATTTCCGAACATGATAGTAGAACGTATCCAATTCGAACAGTT ..... Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro Ile Arg Thr Val	780

TCCCAATTAACAAGAGAAATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT	840
.....	
SerGlnLeuThrArgGluIleTyrThrAsnProValLeuGluAsnPheAspGlySerPhe	
CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTCCACATTTGATGGATATACTT	900
.....	
ArgGlySerAlaGlnGlyIleGluGlySerIleArgSerProHisLeuMetAspIleLeu	
AACAGTATAACCATCTATACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA	960
.....	
AsnSerIleThrIleTyrThrAspAlaHisArgGlyGluTyrTyrTrpSerGlyHisGln	
ATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATTCACTTTTCCGCTATATGGAAC	1020
.....	
IleMetAlaSerProValGlyPheSerGlyProGluPheThrPheProLeuTyrGlyThr	
ATGGGAAATGCAGCTCCACAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
.....	
MetGlyAsnAlaAlaProGlnGlnArgIleValAlaGlnLeuGlyGlnGlyValTyrArg	
ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAGGGATAAATAATCAACAATA	1140
.....	
ThrLeuSerSerThrLeuTyrArgArgProPheAsnIleGlyIleAsnAsnGlnGlnLeu	
TCTGTTCTTGACGGGACAGAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
.....	
SerValLeuAspGlyThrGluPheAlaTyrGlyThrSerSerAsnLeuProSerAlaVal	
TACAGAAAAAGCGGAACGGTAGATTTCGCTGGATGAAATACCGCCACAGAATAACAACGTG	1260
.....	
TyrArgLysSerGlyThrValAspSerLeuAspGluIleProProGlnAsnAsnAsnVal	
CCACCTAGGCAAGGATTTACTCATCGATTAAGCCATGTTTCAATGTTTCGTTACGGCTTT	1320
.....	
ProProArgGlnGlyPheSerHisArgLeuSerHisValSerMetPheArgSerGlyPhe	
AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCT	1380
.....	
SerAsnSerSerValSerIleIleArgAlaProMetPheSerTrpIleHisArgSerAla	
GAATTTAATAATATAATTCCTTCATCACAAATTACACAAATACCTTTAACAAAATCTACT	1440
.....	
GluPheAsnAsnIleIleProSerSerGlnIleThrGlnIleProLeuThrLysSerThr	
AATCTTGGCTCTGGAACCTCTGTTCGTTAAAGGACCAGGATTTACAGGAGGAGATATTCTT	1500
.....	
AsnLeuGlySerGlyThrSerValValLysGlyProGlyPheThrGlyGlyAspIleLeu	
CGAAGAACTTCACCTGGCCAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTATCA	1560
.....	
ArgArgThrSerProGlyGlnIleSerThrLeuArgValAsnIleThrAlaProLeuSer	

```

CAAAGATATCGGGTAAGAATTTCGCTACGCTTCTACCACAAATTTACAATTCCATACATCA 1620
.....
GlnArgTyrArgValArgIleArgTyrAlaSerThrThrAsnLeuGlnPheHisThrSer

ATTGACGGAAGACCTATTAATCAGGGGAATTTTTTCAGCAACTATGAGTAGTGGGAGTAAT 1680
.....
IleAspGlyArgProIleAsnGlnGlyAsnPheSerAlaThrMetSerSerGlySerAsn

TTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTCCGTTTAACTTTTCAAATGGA 1740
.....
LeuGlnSerGlySerPheArgThrValGlyPheThrThrProPheAsnPheSerAsnGly

TCAAGTGTATTTACGTTAAGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAGAT 1800
.....
SerSerValPheThrLeuSerAlaHisValPheAsnSerGlyAsnGluValTyrIleAsp

CGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAGAATATGATTTAGAAAGAGCA 1860
.....
ArgIleGluPheValProAlaGluValThrPheGluAlaGluTyrAspLeuGluArgAla

CAAAAGGCGGTGAATGAGCTGTTTACTTCTTCCAATCAAATCGGGTTAAAAACAGATGTG 1920
.....
GlnLysAlaValAsnGluLeuPheThrSerSerAsnGlnIleGlyLeuLysThrAspVal

ACGGATTATCATATTGATCAAGTATCCAATTTAGTTGAGTGATAA 1965
.....
ThrAspTyrHisIleAspGlnValSerAsnLeuValGlu

```

图3 Bt 毒性蛋白基因5'端核苷酸序列及推导出的氨基酸序列

Fig. 3 Nucleotide sequence of Bt toxin gene 5' end and the deduced amino acid sequence

第一行: PCR 合成的毒性蛋白基因5'端核苷酸序列; 第二行: S. Kondo 发表的5.3 Kb 类型毒性蛋白基因部份核苷酸序列, ... 表示相同核苷酸; 第三行: 由核苷酸序列推导出的氨基酸序列

1st line: Nucleotide sequence of PCR synthesized Bt toxin gene 5' end; 2nd line: Partial nucleotide sequence of 5.3 Kb type Bt toxin gene published by S. Kondo, ... indicate the identical nucleotide; 3rd line: The deduced amino acid sequence from nucleotide sequence

从上实验结果表明, 我们通过 PCR 合成的手段, 成功克隆获得了苏云金杆菌戈尔斯德亚种 HD-1 5.3 Kb 类型毒性蛋白基因的5'端。此项工作证明, PCR 方法可用于扩增长达2 Kb 的 DNA 序列, Taq 多聚酶的复制错误率不高, 扩增的片段完全可用作基因克隆的材料。此外, PCR 技术在基因克隆工作上的应用, 使我们能够借助国内外已有的研究结果, 在较短时间内, 以较少的花费直接克隆我们需要的目的基因。我们克隆获得的毒性蛋白基因5'端片段具备完整的起始和终止序列, 无需再行3'端缺失的工作, 就可直接连接到植物表达启动子35S 下游, 这为下一步进行植物基因转化的工作提供了便利条件。这方面的工作目前正在进行中。

## 参 考 文 献

- [1] Alik Honigman, Galit Nedjar-Pazerini et al. 1986. Cloning and expression of the lepidopteran toxin produced by *Bacillus thuringiensis* var. *thuringiensis* in *Escherichia coli*. *Gene*. 42:69—77
- [2] Barton K A, H R Whiteley et al. 1987. *Bacillus thuringiensis* -endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to Lepidopteran insects. *Plant Physiol.* 85:1103—1109
- [3] Crowe J S, H J Cooper et al. 1991. Improved cloning efficiency of PCR products after proteinase K digestion. *Nucleic Acids Research*. 19 (1) :184
- [4] Fischhoff D A, K S Bowdish et al. 1987. Insect tolerant transgenic tomato plants. *Bio/Technology*. 5:807—813
- [5] Frederick, J Perlak et al. 1990. Insect resistant cotton plants. *Bio/Technology*. 8:939—943
- [6] Kronstad J W and Whiteley H R. 1986. Three classes of homologous *Bacillus thuringiensis* crystal-protein genes. *Gene*. 43:29—40
- [7] Maniatis T, E F Fritsch et al. 1982. *Molecular cloning-laboratory manual*. Cold Spring Harbor, New York
- [8] Michael J A, Michael J S et al. 1985. Characterized full-length and truncated plasmid clones of *Bacillus thuringiensis* subsp. *kurstaki* HD-73 and their toxicity to *Manduca sexta*. *Gene*. 36:289-300
- [9] Sambrook J, E F Fritsch et al. 1989. *Molecular Cloning-laboratory manual* (2nd ed) . Cold Spring Harbor, New York
- [10] Schnepf H E and H R Whiteley. 1985. Delineation of a toxinencoding segment of a *Bacillus thuringiensis* crystal protein gene. *J. Biol. Chem.* 260:6273—6280
- [11] Shin-ichi Kondo, Nobuya Tamura et al. 1987. Cloning and nucleotide sequencing of two insecticidal -endotoxin genes from *Bacillus thuringiensis* var. *kurstaki* HD-1 DNA. *Agri. Biol. Chem.* 51:455—463
- [12] Vacck M, A Reynaerts et al. 1987. Transgenic plants protected from insect attack. *Nature*. 328:33—37
- [13] Whiteley H R and Schnepf H E. 1986. The molecular biology of parasporal crystal body formation in *Bacillus thuringiensis*. *Annu. Rev. Microbiol.* 40:549—576