Nucleotide Sequence of the Lepidoptera-toxic Protein Gene of *Bacillus thuringiensis* subsp. *dendrolimus* T84A1

Masashi Ogo, Shoichi Yamada, Yasuo Kobayashi, Jiro Shibata and Yasunori Nagamatsu

Faculty of Applied Biological Science, Hiroshima University, Higashi-Hiroshima 724, Japan

Received September 26, 1990

Abstract A 3882 nucleotides sequence including the complete crystal protein gene of *Bacillus thuringiensis* (B. t.) *dendrolimus* T84A1 was determined by the dideoxy-chain termination method. It was revealed that the gene of *B. t. dendrolimus* is highly homologous to that of *B. t. sotto* (Shibano et al., 1986). According to the suggested nomenclature, this gene is classified into *cry IA(a)* (Höfte and Whiteley, 1989). The open reading frame encoded a 133487.71 Da protein consisting of 1180 amino acid residues. The hydrophobicity of the predicted crystal protein was also analyzed. When the crystal protein gene was expressed in *Escherichia coli* MV184 using the vector pUC118, the lower cultivation temperature (25°C) yield larger accumulation of the crystal protein than that at ordinary temperature (37°C). Based on the sequence determined in this study, the functional structure of the crystal protein was discussed.

INTRODUCTION

*Bacillus thuringiensis* is a gram-positive bacterium, and is well known to produce an insecticidal protein crystal during sporulation (δ-endotoxin). Some δ-endotoxins have been widely used to control certain species of insects.

Crystal proteins from a number of *B. t.* strains can be classified into four Cry protein classes and one Cyt (cytolytic) protein, based on the host spectra of insecticidal activities and the primary structures. The four Cry classes are (I) Lepidoptera-specific, (II) Lepidoptera-and Diptera-specific, (III) Coleoptera-specific, (IV) Diptera-specific, and the genes coding these proteins are called *cry* genes. And the remaining cytolytic protein gene is designated *cytA*.

From its lepidoptera-specific spectrum, the crystal protein produced by *B. t. dendrolimus* may belong to CryI. CryI proteins dissolve in the larval lumen and release smaller toxic polypeptide fragments. A trypsin-resistant polypeptide (Fragment T, app. M. W. 60–kDa) from *B. t. dendrolimus* crystal protein is the minimum unit and its N-terminal sequence was determined previously (Nagamatsu et al., 1984). Several recent studies demonstrated the correlation between the insecticidal specificity and the interaction of the toxic fragments with the receptors in the brush border membrane of insect midgut (Hoffmann et al., 1988, Rie et al., 1989).

It is important to know the full primary structure of the crystal protein, because it may
help to understand which part of the crystal protein interacts with the insect larval midgut, and how their biochemical mechanisms are. To date, several cryI genes from different B. t. strains were cloned and sequenced, but not B. t. dendrolimus. In this study, we present its complete nucleotide sequence and the deduced amino acid sequence of the crystal protein. Temperature influence on expression of the crystal protein in E. coli MV1184 was also examined.

**MATERIALS AND METHODS**

**Hosts and Plasmids.**

*Escherichia coli* MV1184 [ara, Δ(lac-pro), strA, thi, (φ80Δ1acZΔM15), Δ(srl-recA)306::Tn10(tet'), F' :traD36, proAB, lacIΔZ Δm15] was used as a host with plasmids, pUC118 and pUC119. Plasmid pH206 was used as the source of full length of the crystal protein gene (SHIBATA et al., 1989).

**Nucleotide sequencing.**

Plasmid pH206 containing the 5 kb PsI fragment of *B. t. dendrolimus* was digested with some restriction enzymes and the generated fragments were inserted into pUC118 or pUC119 vectors to create subclones for sequencing. Some of them were further digested with exonuclease III and Mung bean nuclease to generate overlapping deletion mutants (HENIKOFF, 1984). The template single-stranded DNAs were prepared by infecting with helper phage M13K07 (VIEIRA and MESSING, 1987), and sequenced by the dideoxy chain-termination method (SANGER et al., 1977).

**Analysis of the open reading frame and the transcriptional terminator of the crystal protein gene, and the hydropathy profile of the crystal protein.**

Searching for the ORF (open reading frame) and the transcriptional terminator and calculation of the average hydropathy indexes (KYTE and DOOLITTLE, 1982) were carried out by using the computer program (GENETYX: SDC software development inc.) at Center for Gene Science Hiroshima University.

**Construction of pPP03.**

The 5 kb PsI fragment containing the crystal protein gene on pH206 was ligated with pUC118 and introduced into *E. coli* MV1184.

**Analysis of the crystal protein produced in E. coli.**

*E. coli* MV1184 cells carrying the pPP03 were grown in 100 ml of L-broth at 37°C or 25°C for 48 hr. Cell concentrations of *E. coli* MV1184 in a medium were measured by the absorbance at 600 nm.

The cells were harvested by centrifugation at 7000 × g for 10 min at 4°C, resuspended with 20 ml PBS, and sonicated at 30 W for 10 min in an ice-bath. After centrifugation at 15000 × g and 4°C for 90 min, the crystal protein in the precipitate was purified by selective solubilization in sodium carbonate buffer (Na⁺; 0.1 M), pH 10.2, including 10 mM dithiothreitol. Samples prepared as described above were subjected to SDS (sodium dodecyl sulfate)-7.5% PAGE (polyacrylamide gel electrophoresis) analysis (LAEMMLI et al., 1970). Protein bands were visualized by Coomasis Brilliant Blue staining. The crystal protein was immunodetected after electrotransfer to a polyvinylidene difluoride membrane as described previously (SHIBATA et al., 1989).
Transformation, media and other techniques.

Transformation of *E. coli* MV1184 was performed as described by Chung et al., (1989). Media and other techniques used in this study were described by Sambrook et al., (1989).

**RESULTS**

Shibata et al., (1989) isolated a 9.91 kb *PstI* fragment including the crystal protein gene from the 51.1 kb plasmid DNA of *Bacillus thuringiensis* *dendrolimus* T84A1. The entire *cry* gene resided on the 5.0 kb *HpaI*-*PstI* segment within the DNA fragment and it was the only *cry* gene in this *B. t.* strain. The 5.0 kb DNA fragment was attached with a *PstI* linker and inserted into vector pBR322 to generate plasmid pH206.

Figure 1 shows the restriction enzyme map of the DNA fragment and the sequencing strategy. By analyses of 48 partial sequences, a 3882-base continuous sequence from the *HpaI* site to the second *KpnI* site was determined.

In this sequence, two open reading frames (ORF) were found together with the two identical potential ribosome binding regions (Fig. 2). The shorter one (ORF2, 6 bases) is of course too short to encode the crystal protein. The calculated molecular weight (133,487
Fig. 3. Nucleotide sequence and the deduced amino acid sequence of the crystal protein.
Bacillus thuringiensis Crystal Protein Gene

970 980 990 1000 1010 1020
ACCCGATTTAGAAAATTTTTGATGTTAGTTTCTGTAATGGATGCTAGAAGATAGAAGAG
ProValLeuGluAsnPheAspGlySerPheArgGlyMetAlaGlnArgIleGluGlnAsn

1030 1040 1050 1060 1070 1080
ATATAGCAGCAACACACATCTTTCTAGTATCTAGGGATAGATCAGCAGCTCAGCCGAT
IleArgGlnProHisLeuMetAspIleLeuAsnArgIleThrIleTyrThrAspValHis

1090 1100 1110 1120 1130 1140
ATAGAGGCTTAAATTATGGTCAGGGCATCAAAATAACAGCTTTCTCCTGGATAGGGTTTCG
ArgGlyPheAsnTyrTrpSerGlyHisGlnIleThrAlaSerProValGlyPheSerGly

1150 1160 1170 1180 1190 1200
GCCAGAATTCCGATTTACCTTTTATTTGGAATGCGGGGAATGAGCTCCACCCGACTTG
ProGluPheAlaPheProLeuPheGlyAsnAlaGlyAsnAlaAlaProValGluVal

1210 1220 1230 1240 1250 1260
TCTCATTAACCTGTTGGGATTTTTAGACACTACCTCCTTCTATAGAAGAATA
SerLeuThrGlyLeuGlyIlePheArgThrLeuSerSerProLeuTyrArgIleIle

1270 1280 1290 1300 1310 1320
TACTTGATTGACCACAAATAACAGGAACGTTGCTGAGAGATTGTTTCTCTGTT
LeeGlySerGlyProAsnAsnGlnGluLeuPheValLeuAspGlyThrGluPheSerPhe

1330 1340 1350 1360 1370 1380
TTGCCCTTCAACGCAACACCTGCTCTCCACTATACTATAGACAAAGGTGACGAGC
AlaSerLeuThrThrAsnProSerAlaSerThrIleTyrArgGlnGlyArgTyrValAspSer

1390 1400 1410 1420 1430 1440
CAGCATGATAATACGTTGCAAGATAATAGTGATCCACCTGATGCGGGATTTGAT
LeuAspValIleProProGlnAspSerValSerProProArgAlaGlyPheSerHisArg

1450 1460 1470 1480 1490 1500
GATTTGCTGATGGTTTACCAATGGCTGACGCAAGCAGCTGAGAGTTTACCTGAGG
LeeSerHisValThrMetLeuSerGlnGluAlaAlaValAlaValThrThrLeuPro

1510 1520 1530 1540 1550 1560
CAACGTTTTCTGGCAGCATCGATGCTGAATTTTAAATAATAATATATTCTTCCTCG
ThrPheSerTrpGlnHisArgSerAlaGluPheAsnIleIleProSerSerGlnIle

1570 1580 1590 1600 1610 1620
TTCAGCAAATACCTTTACTCTACCATCTGCTGTAAGAGTATTCTCTCTGCTTTAAAG
ThrGlnIleProLeuLysSerThrAsnLeuGlySerGlyGlySerValValLysGly

1630 1640 1650 1660 1670 1680
GACCAGTTTACAGGAGGAGATTTCTTCGCAAGAATTTCCACGCTGGCAGATTTCCACT
ProGlyPheThrGlyGlyIleLeuArgArgThrSerProGlyGlnIleSerThrLeu

1690 1700 1710 1720 1730 1740
TAAGATGAAATTACGACCATTACCATCAAGGATTACGCTGGAAGAGTTCCCTCG
ArgAsnAsnIleThrAlaProLeuSerGlnArgTyrArgValArgIleArgTyrAsa

1750 1760 1770 1780 1790 1800
CTACTACAAATTTCTCAATATCATACATACCATCAATTGAGGAGACCTATTTAATCA
ThrThrAsnLeuGlnPheHisThrSerIleAspGlyArgProIleAsnGlnGlyAsnPhe

1810 1820 1830 1840 1850 1860
TTCCGACACATGAGTGTGGAGGATTTTACGCTCGGAAGCTTTAGGACTGAGGTT
SerAlaThrMetSerSerGlySerAsnLeuGlnSerGlySerPheArgThrValGlyPhe

1870 1880 1890 1900 1910 1920
TTACTACCTGGTTAACTCTTTCAATGAGGATCTGAGTATTTACGTTAAGTGGCTCAG
ThrThrProPheAsnPheSerAsnGlySerSerValPheThrLeuSerAlaHisValPhe

Fig. 3. Continued.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TCAATT</td>
<td>CAGCAAT</td>
<td>GGAAGT</td>
<td>TTTATAT</td>
<td>TAGATCG</td>
<td>ATGTAAT</td>
</tr>
<tr>
<td>AsnSerGlyAsnGluValTyrIleAspArgIleGluPheValProAlaGluValThrPhe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>2000</td>
<td>2010</td>
<td>2020</td>
<td>2030</td>
<td>2040</td>
</tr>
<tr>
<td>TTGAGG</td>
<td>CCAATAT</td>
<td>TAGATTT</td>
<td>GAAGAAC</td>
<td>CAAAAG</td>
<td>GCGGTGA</td>
</tr>
<tr>
<td>GluAlaGluTyrAspGluArgAlaGlnLysAlaValAsnGluLeuPheThrSerSer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2050</td>
<td>2060</td>
<td>2070</td>
<td>2080</td>
<td>2090</td>
<td>2100</td>
</tr>
<tr>
<td>CCAATCAATCAGGTGTTTAAAACAGATGTCGACGATTATCAATGACTCTAACAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AsnGlnIleGlyLeuLysThrAspValThrAspTyrHisIIeAspGlnValSerAsnLeu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2110</td>
<td>2120</td>
<td>2130</td>
<td>2140</td>
<td>2150</td>
<td>2160</td>
</tr>
<tr>
<td>TAGTTGATGTGTATATCGATGAAATTTGCTGGATGAAAACAAAGAGTTGTCGAGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ValGluCysLeuSerAspGluPheCysLeuAspGluLysGlnGluValSerGlyVal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2170</td>
<td>2180</td>
<td>2190</td>
<td>2200</td>
<td>2210</td>
<td>2220</td>
</tr>
<tr>
<td>TCAACCAATGCGAAGGCACCATAGTAGTGGAGGCGGAATTTACCTCAATCAGACAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LysHisAlaLysArgLeuSerAspGluArgAsnLeuLysGluAspProAsnGlyArgGly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2230</td>
<td>2240</td>
<td>2250</td>
<td>2260</td>
<td>2270</td>
<td>2280</td>
</tr>
<tr>
<td>GGTCAATTAAACACTAGAACGCTGCGTGGAGAGAGTTGCGGATTTATGCAATCAGAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IleAsnArgGlnLeuAspArgGlyTrpArgGlySerThrAspIleThrIleGlnGlyGly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2290</td>
<td>2300</td>
<td>2310</td>
<td>2320</td>
<td>2330</td>
<td>2340</td>
</tr>
<tr>
<td>GGGATGACGTGATTTCAAAAGAAGATGGATTTGCTAACCTGTTGATGAGTCGATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AspAspValPheLysGluAsnValThrValLeuHisLeuGluThrPheAspAsnGlyPro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2360</td>
<td>2370</td>
<td>2380</td>
<td>2390</td>
<td>2400</td>
<td></td>
</tr>
<tr>
<td>CAGCTATTTATATCAAAATAGATGAGTGGCAGAAATTTAAGCCTATACCCGTTTATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ThrTyrLeuValGlnLysIleAspGluSerLysLeuLysAlaThrThrArgTyrGlnLeu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2410</td>
<td>2420</td>
<td>2430</td>
<td>2440</td>
<td>2450</td>
<td>2460</td>
</tr>
<tr>
<td>TAAGGGTGATATCGAGATGACGACTTTAGAAATCTATTTAATTGCTACTAAGGAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ArgGlyTyrIleGluAspSerGlnAspLeuGluIleTyrLeuIleArgTyrAsnAlaLys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2470</td>
<td>2480</td>
<td>2490</td>
<td>2500</td>
<td>2510</td>
<td>2520</td>
</tr>
<tr>
<td>AACATTAAACAGTAAATGCGGCAATGGTTCTCTATGGCCGCTTCAGCCCCAGAAGTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HisGluThrValAsnValProGlyThrGlySerLeuTrpProLeuSerAlaGlnSerPro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2530</td>
<td>2540</td>
<td>2550</td>
<td>2560</td>
<td>2570</td>
<td>2580</td>
</tr>
<tr>
<td>CAACTCAGGAAAGTCTGGAGGCGGACCTGATGGCCTCCATCGATGATCTGACT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IleGlyLeuAspGluAspArgCysAlaProHisLeuGluTrpAspProAsnLeuAsp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2590</td>
<td>2600</td>
<td>2610</td>
<td>2620</td>
<td>2630</td>
<td>2640</td>
</tr>
<tr>
<td>TAGATTCTGTGTCAGGGATGAGAAAGAATGTTGCCCATCATTTCCATACCTTTCTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AspCysSerGluArgSerGluLysCysAlaHisSerHisHisPheSerLeuAsp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2650</td>
<td>2660</td>
<td>2670</td>
<td>2680</td>
<td>2690</td>
<td>2700</td>
</tr>
<tr>
<td>ACATTGATGGCTGATCACACTTAAATGAGGACTTTAGTGTATGGTGATATCAGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IleAspValGluCysThrAspLeuAspGluAspLeuGluValPheLysIle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2710</td>
<td>2720</td>
<td>2730</td>
<td>2740</td>
<td>2750</td>
<td>2760</td>
</tr>
<tr>
<td>TTAAGGGCCAGTGGCCTGCAAGACTTAAATGAGGACTTTAGTGTATGGTGATATCAGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LysThrGlnAsnGluHisAlaArgLeuGluAspGluLeuPheGluLysProLeu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2770</td>
<td>2780</td>
<td>2790</td>
<td>2800</td>
<td>2810</td>
<td>2820</td>
</tr>
<tr>
<td>TAGTTAGGAGAACAGCAGCTCCTGCTGGAAAGAAGGAGGAAAGAAATGAGGAGACAAAAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ValGluGluAlaLeuAlaArgValLysArgAlaGluLysTrpArgAspLysArgGlu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2830</td>
<td>2840</td>
<td>2850</td>
<td>2860</td>
<td>2870</td>
<td>2880</td>
</tr>
<tr>
<td>AAAATTGGGAAAGGAAAACAATTTTATTAAAAAGACGAAAAAGAATCTTGAGACGTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LysLeuGluTrpAsnIleValTyrLysGluAlaLysGluSerValAspAlaLeu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Continued.
Fig. 3. Continued.
Da) of the ORF1 protein is well consistent with the apparent molecular weight 140,000 estimated on SDS-PAGE of the crystal protein. Among the several stem loop structures (data not shown), the largest (20 bases) one locating behind the two ORFs was revealed to be a potential transcriptional terminator because of the highest stability and its location. The minimum free energy (ΔG) for the structure formation was calculated as -25.1 kcal/mol.

Detailed nucleotide sequence and the deduced amino acid sequence were shown in Fig. 3. ORF1 (#153-3692) encodes 1180 amino acid residues. The ribosome binding sequence is located six bases upstream from the initiation codon. The N-terminal amino acid sequence of the crystal protein and the several partial sequences within the toxic trypsin-resistant fragment (Fragment) were also determined (NAGAMATSU et al., 1984, and our unpublished data). The sequences underlined were identical with those deduced from the nucleotide sequences, giving the evidence that ORF1 is the crystal protein gene. The N-terminal and the C-terminal residues of Fragment T correspond to #29 and #618 amino acid residues of the crystal protein, respectively. The minimum toxic unit resides on the N-terminal half of the crystal protein. Compared with the N-terminal half and the remaining C-terminal half, most of cysteine and lysine residues distributed in the latter region.

From the modes of action of several well-studied protein toxins, it was expected that Fragment T could be consisted of several functional domains. A hydrophyt profile of the crystal protein was created based on the deduced amino acid sequence (Fig. 4). The N-ter-

![Graph showing hydrophobic index](image-url)

**Fig. 4.** Hydrophyt profile of the crystal protein.

The location of the minimum toxic region (Fragment T) and the polypeptides generated by subtilisin cleavage were shown by the boxes. Trypsin and subtilisin cleaving sites are given by arrows.
minal half of the crystal protein is more hydrophobic as than the C-terminal half. Trypsin cleavage sites, the N- and C-terminals of Fragment T, are located in boundary regions between the hydrophilic regions and the hydrophobic regions. Further digestion of the low specific protease, subtilisin gave two (TS–35, TS–26) or three (TS–30, TS–5, TS–26) polypeptides. The results from their N-terminal and C-terminal sequence analysis and the sequence determined in this study, these polypeptides was positioned as illustrated in Fig. 4. These subtilisin cleavage sites were also positioned in the boundary regions between the hydrophilic regions and the hydrophobic regions. These generated polypeptides associate each other, making a nicked protein missing the insecticidal activity. Highly hydrophobic region can be seen around #50 amino acid residue.

The expression of the crystal protein gene in *E. coli* was demonstrated, but the productivity was quite low (Shibata, et al., 1989). Using the recombinant plasmid pPP03 (vector pUC118–host *E. coli* MV1184 system), influence of the cultivated temperature on expression of the crystal protein gene was examined. Figure 5 shows the SDS-PAGE patterns of proteins in the soluble and insoluble fractions. In the insoluble fraction (lane 5) from *E. coli*

![SDS-PAGE analysis of the crystal protein produced in *E. coli* cultivated at 25°C and 37°C. *E. coli* carrying the recombinant plasmid, pPP03 (pUC118 containing the crystal protein gene), was cultivated at 37°C (lanes 1, 4 and 7) and 25°C (lanes 2, 5 and 8) for 48 hr. As a control, *E. coli* carrying vector pUC118, was also cultivated (lanes 3, 6 and 9). The cell suspension was disrupted by sonication and fractionated into the soluble (lanes 1, 2, 3) and the insoluble (lanes 4, 5, 6). Each of the insoluble fractions were extracted with the buffer, pH 10.2, containing 10 mM dithiothreitol (lanes 7, 8, 9).]
carrying pPP03 cultivated at 25°C, a protein band of the same molecular size as that of authentic crystal protein (lane C), was clearly detected. This band was not seen in the fraction from E. coli carrying the vector (control, lane 6). The amount of the protein band from the culture at 37°C (lane 4) was quite low comparing with the case of 25°C (lane 5). The bulk of the crystal protein produced was recovered in the insoluble fraction, indicating its accumulation as inclusion bodies. The crystal protein in the insoluble fraction was selectively extracted with sodium carbonate buffer (Na⁺; 0.1 M), pH 10.2, containing 10 mM dithiothreitol (lane 8).

Figure 6 shows growth curves of E. coli cultivated at 25°C and 37°C. The growth rate at 25°C was lower during the log phase than that at 37°C, but cell concentrations in the stationary phase were on a similar level at those two temperatures. Influence of cultivation temperature on the crystal protein production was again examined by immunoblot analysis with anti-Fragment T antibody. Expression of the crystal protein gene in this E. coli system was confirmed, and accumulation of the crystal protein was higher at the lower temperature.

DISCUSSION

In this study, the crystal protein structural gene of Bacillus thuringiensis dendrolimus
was found to be consisted of 3540 nucleotides encoding 1180 amino acid residues.

Comparison of the sequence for the *dendrolimus* strain with other sequences for several *B. t.* strains revealed its high homology with *cryIA(a)* genes. Between the *kurstaki* HD-1 Dipel gene (Schnepf *et al.*, 1985) and the *dendrolimus* gene, there are 57 nucleotides differences; 19 amino acid residues are substituted and 4 residues are deleted in the *kurstaki* protein. Within the toxin region (the N-terminal half of the crystal protein), only 3 amino acids changes were seen. Comparing with the *sotto* gene (Shibano *et al.*, 1986), only one nucleotide difference at position 2795, C for *sotto*, G for *dendrolimus* was found. Structural differences were revealed, in 5′-region (Shibata *et al.*, 1989). Subspecies *sotto* is categorized into the same serotype as subsp. *dendrolimus*.

Yamada (1990) showed two transcriptional start sites of the *B. t. dendrolimus* crystal protein gene, Bt I (#83 base in early sporulation), and Bt II (#69 base in mid sporulation). Results for the *kurstaki* HD-1 Dipel gene (Wong *et al.*, 1983) were coincident with the case of *dendrolimus*, and the promoter regions of both strains were well conserved.

A very short open reading frame (ORF2) with the identical ribosome binding sequence was found in the 3′-terminal regions of the crystal protein gene (Fig. 2). Influence of the second ribosome binding sequence on the crystal protein production remains to be studied.

The hydropathy and the distribution of certain amino acid revealed the differences in protein structural features between the N- and C-terminal halves. The C-terminal half containing most of cysteine residues may be involved in crystal formation. Bietlot *et al.*, (1990) suggested that the thiol groups of the crystal protein are exposed on the surface of the molecule, and form symmetrical interchain disulfide bridges. Comparing with the C-terminal half, several hydrophobic peaks were clearly seen in the toxic N-terminal half, suggesting that the existence of several structural domains. Subtilisin digestion of the toxic fragment gave three polypeptide fragments, generating no toxic protein. One of the hydrophobic peaks located between the two subtilisin cleavage sites, and this part corresponding to TS-5 domain (Fig. 4). Comparison of many crystal protein sequences revealed two hypervariable regions in the minimum toxic unit (Geiser *et al.*, 1986). One of them located from #286-457 corresponds to the boundary regions of three domains with TS-5 as the central figure. The other one #466-618 corresponds to the C-terminal remaining part of TS-26 region, and most of TS-30 region are well conserved. The combination of the primary sequence analysis and the domain searching with limited proteolysis would be a great help to study the relation between the function and the structure. Subtilisin-cleaved nicked protein missed the insecticidal activity, suggesting some synergistic action of these domains. Binding of this toxin to the receptor, a specificity-determining step, might be exerted by the interface region or the C-terminal part of TS-26 polypeptide. In addition, highly hydrophobic region is observed around #50 amino acid residue. Role of the potential membrane-spanning sequence is very interested from the point of view that this protein is a membrane-acting toxin.

The production of the crystal protein in *E. coli* was higher at 25°C than that at 37°C in contrast to the growth rates at these temperatures. As the explanation of the preferential accumulation at the lower temperature, the following three factors could be considered; ① the gene may has a temperature-dependent regulation system, ② the temperature difference
would affect the protein synthesis rate, the degradation rate of the crystal protein might change at different temperatures. *B. t. dendrolimus* incubated at temperatures above 30°C decrease the crystal productivity. So temperature-dependent expression of the crystal protein may be regulated under a similar mechanism in both organisms.

**ACKNOWLEDGEMENT**

We would like to thank all staff members of Center for Gene Science Hiroshima University for their kind offer of the facilities.

**REFERENCES**


Bacillus thuringiensis Crystal Protein Gene


Bacillus thuringiensis dendrolimus T84A1
の殺虫性蛋白質遺伝子の塩基配列の決定

尾崎 正志・山田 彰一・小林 泰久
柴田 治郎・永松 康徳

広島大学生物産学部、東広島市 724

Bacillus thuringiensis dendrolimus T84A1 の殺虫性蛋白質遺伝子を含む 3882 塩基の配列を決定した。
本菌の結晶蛋白質は 1180 個のアミノ酸（分子量133, 487 Da）から成ることが推定された。
蛋白質の研究と今回決定した配列の結果から、最小の殺虫活性断片である Fragment T が結晶蛋白質の
N-末端側の半分に位置することがわかった。結晶蛋白質のハイドロバシー分析から N-末端側の半分の領域
は疎水的で C-末端側は親水的であることが明らかになった。その C-末端側に大部分のシテイン、リジン
残基が偏って存在していることから、C-末端側は結晶の形成に関与していることが予測された。
サテラインによる Fragment T の限定水解とハイドロパシーの解析によって 3 つのドメインから成る
毒性発現領域の構築が明確になった。
また、大腸菌における殺虫性結晶蛋白質の生産に与える培養温度の影響を検討したところ、通常の培養温
度である37℃に比べてより低い25℃の方が顕著な蓄積を示すことが明らかになった。