Rapid Paper

Cloning and Nucleotide Sequences
of the Two 130 kDa Insecticidal
Protein Genes of *Bacillus*
thuringiensis var.
*israelensis*Kikuo SEN, Gouichi HONDA, Naoto KOYAMA,
Masato NISHIDA, Akio NEKI, Hiroshi SAKAI,
Michio HIMENO* and Tohru KOMANOLaboratory of Biochemistry,
Department of Agricultural Chemistry,
Kyoto University, Kyoto 606, Japan

*Department of Polymer Engineering,

Faculty of Technology,

Tokyo University of Agriculture and Technology,
Tokyo 184, Japan

Received November 10, 1987

Two types of 130 kDa insecticidal protein genes isolated from large plasmid DNAs of *Bacillus thuringiensis* var. *israelensis* HD522 were cloned in an *Escherichia coli* plasmid vector. Analysis of the nucleotide sequences revealed that the two genes encoded a 127,500-dalton protein (ISRH3) consisting of 1,135 amino acids and a 134,400-dalton protein (ISRH4) consisting of 1,180 amino acids. The two insecticidal proteins were identical in a region of the C-terminal 467 amino acids.

Strains of *Bacillus thuringiensis* var. *israelensis* (BTI) are gram-positive bacteria which produce parasporal crystal proteins known as δ -endotoxin at the sporulation stage. The crystal proteins are highly toxic to mosquito and black fly larvae,¹⁾ and solubilized crystal proteins are hemolytic.²⁾ The crystals are composed of four major protein components of 28 kDa, 38 kDa, 65 kDa, and 130 kDa in their relative molecular masses.^{3,4)} BTI carries seven plasmids,⁵⁾ and it is reported that the 70 Md plasmid, pBTI-6, has the insecticidal protein gene(s).^{5,6)} Several investigators have cloned insecticidal protein genes from BTI, established their DNA sequences, and tested for activities of the cloned gene products.^{7~15)} The 130 kDa protein is insecti-

cidal but not hemolytic.^{11,16)} Bourguoin *et al.* cloned two types of 130 kDa insecticidal protein gene.¹³⁾ The Type I gene has been cloned^{10,14)} and partially sequenced in the 5'-flanking and 5'-terminal regions.¹⁴⁾ Recently Ward and Ellar have reported the base sequence of the coding stretch of the Type II gene.¹⁵⁾

In this paper, we report cloning of both types of 130 kDa insecticidal protein genes. The Type I and Type II genes correspond to ISRH3 and ISRH4 genes respectively, in this paper. And we show complete DNA sequences of the two genes and compare them with other insecticidal protein genes.

MATERIALS AND METHODS

Bacterial strains and plasmids. *B. thuringiensis* var. *israelensis* strain HD522 and strain 4Q1 have been described in our previous paper.⁵⁾ *E. coli* strains HB101¹⁷⁾ and plasmid vector pBR322¹⁸⁾ have been previously described.

Preparation of DNA. BTI HD522 plasmid DNAs were prepared by the alkaline method¹⁹⁾ except for the following modifications: bacterial cells were treated with lysozyme at 30°C for 2 hr. Plasmid DNAs were purified by centrifugation on a CsCl-ethidium bromide density gradient. Large plasmid DNAs (>63 Md) were fractionated by centrifugation on sucrose density gradients. Plasmid DNAs from *E. coli* were prepared by the alkaline method.¹⁹⁾ If it was needed, the plasmid DNAs were purified by centrifugation on CsCl-ethidium bromide density gradients.

Cloning and screening of recombinants. One μ g of large plasmid DNAs from BTI HD522 was digested with 10 units of *Hind*III (Takara Shuzo Co., Ltd.) in medium-salt buffer (10 mM Tris-HCl (pH 7.5) containing 50 mM NaCl, 10 mM MgCl₂, and 7 mM 2-mercaptoethanol) at 37°C for 2 hr. The plasmid DNA digests were inserted into the *Hind*III site of pBR322. Recombinant plasmids were introduced into *E. coli* HB101 cells by the calcium method.¹⁷⁾ Colonies which were ampicillin resistant but tetracycline sensitive were stored as a gene bank. When they were screened for desired recombinant plasmids, colonies on a replica of the bank were treated with chloroform vapor and lifted onto Biodyn A nylon membrane filters. The filters were blocked with 1% bovine serum albumin (BSA) in PBS (0.125 M NaCl-0.02 M potassium phosphate (pH 7.2)) by incubation overnight at 4°C and were soaked in PBS containing 0.5% BSA and

0.2~1 µg/ml horseradish peroxidase (HRPO)-immunoglobulin G (IgG) conjugate at 4°C overnight. The filters were washed with PBS and soaked in PBS containing 0.5 mg/ml 3,3'-diaminobenzidine and 0.05% H₂O₂. Colored colonies, which reacted with HRPO-IgG conjugate, were stored for further analysis. An IgG against *BTI* 4Q1 total crystal proteins was kindly prepared by Sumitomo Chemical Co., Ltd., Japan, and was bound to HRPO by the maleimide method.²⁰⁾

DNA sequencing analysis. DNA sequences were analyzed by the dideoxy method of Messing²¹⁾ and the deletion method of Perron *et al.*¹⁷⁾ The sequences obtained were analyzed by the DNASIS Program from Hitachi Software Engineering Co., Ltd., Japan.

RESULTS AND DISCUSSION

Isolation and characterization of clones

The large plasmid DNA preparation from *BTI* HD522 contained almost equal amounts of 63 Md plasmid (pBTI-5) and 70 Md plasmid (pBTI-6) and a smaller amount of 82 Md plasmid (pBTI-7). From about 1,000 recombinants, 7 positive clones were obtained. One of them harbored a recombinant plasmid, pBGH3, which had a 7.8 kb *Hind*III insert at the *Hind*III site of pBR322 (Fig. 1). Another harbored a recombinant plasmids pBGH4 which had a 9.4 kb *Hind*III insert at the *Hind*III site of pBR322 (Fig. 1). Comparison of their restriction maps with those of the

other recombinant plasmids and an enzyme-linked immunosorbent assay of the subclones (data not shown) revealed that insecticidal protein genes were on a 5.2 kb *Hind*III fragment from pBGH3 and on a 5.0 kb *Hind*III fragment from pBGH4 (Fig. 1).

Comparison of nucleotide sequences of the two 130 kDa insecticidal protein genes

Nucleotide sequences of the 4.2 kb *Cl*aI fragment on pBGH3 and 5.3 kb *Hind*III-*Cl*aI fragment on pBGH4 were analyzed, and a long open reading frame was found in each of the two (Fig. 1). One open reading frame, ISRH3, found in pBGH3 consisted of 3408 nucleotides, encoding a 127,500-dalton protein (Fig. 2). The other open reading frame, ISRH4, found in pBGH4 consisted of 3543 nucleotides, encoding a 134,400-dalton protein (Fig. 2). In the 5'-flanking region was found a 3'-terminal part (621 bases in length) of an unknown open reading frame. Values of the calculated molecular masses of the two predicted proteins were consistent with those found by immunoblotting analysis (data not shown). A Shine-Dalgarno sequence was found 7 nucleotides upstream of the initiation codon in each of the two genes (Fig. 2). Nucleotide sequences were highly conserved in the 5'-flanking region (79/100 nucleotides).

Fig. 1. Restriction Map of Cloned DNA Fragments from the Large Plasmids of *BTI* HD522.

With pBGH3 and pBGH4, pBR322 and *E. coli* HB101 were used as a vector and host bacterium, respectively. Thick arrows indicate the direction of transcription and the extent of open reading frames ISRH3 and ISRH4. The area of the solid box in the thick arrow indicates a stretch where ISRH3 and ISRH4 are identical in their nucleotide and amino acid sequences. The area of the hatched box in the thick arrow indicates a stretch where lower homology was found. H, *Hind*III; C, *Cl*aI; E, *Eco*RI; X, *Xba*I; P, *Pst*I.

(a) ISRH3

5' CGATTTGAATTTTCTGAATATCGAACAAATATATTATTTGGATGCTTGAATACCACTAACGATATGTATGGAAAATTTTGAAGTGAAAAATATGGTCAAAATAAAATGGAATAATT 120
 ATATTGGTACAGAAATATGATTGGGATTAGTACTCTATAATATAGAAAGGAATGTTTGGTTTTGTATATAGTTGAAAAAGATTCTGTAAATTTCCAGAGACGTATGTGTAGAT 240
 TGAGTATTGGAAACATATCGTTAATTTTATATATGAATATGAAATATACAAGGTCTAGATAAGAAATGTTTCATAGGAATCCGTATCAATTTTTTCAAGGAATATGTATTTGCA 360
 CTTTGGTCTTTTTAAATCGTATGAATTCAAAATAGTTTATATCAATCTTTGTTACACAGAAAAAGATTGTATCCAATGTGAATATGGAGGAATAAAATGAATTCAGGCTATCCGTT 480
 SD M N S G Y P L
 AGCGAATGACTTACAGGGTCAATGAAAAACACGAACATAAAGATTGGCTAGCCATGTGTGAAATAACCAACAGTATGGCGTTAATCCAGCTCCGATTAATCTCTCTCAUATAGTAC 600
 A N D L Q G S M K N T N Y K D W L A M C E N N Q Q Y Q V N P A A I N S S S V S T
 CGCTTTAAAAAGTACGCTGGAGCTATCCTTAAATTTGTAAACCCAGCTGCGAGTACTGTCTTAACCGTACTTACGCGGCTCTCTCTATTTTGGCCGACTAATCTCCACGCCCTGAAAG 720
 A L K V A G A I L K F V N P P A G T V L T V L S A V L P I L W P T N T P T P E R
 AGTTTGGAAATGATTTCATGACCAATACAGGGAATCTTATGTATCAAACTGTAAACAGCTTATGTACGAACAGATGCAAAATGCAAAATGACGGTTGTGAAGATATTATAGATCAATATAC 840
 V W N D F R H T N T G N L I D Q T V T A Y V R T D A N A K H T V V K D Y L D Q Y T
 AACTAAATTAACACTTGGAAAAAGAGGCTTAATAACCACTCTATAGAACAGCAATAACTCAATTAACCTAACCAAGTCCGAGAGACCGCAAGTTATTTAGCAACTT 960
 T K F N T W K R E P N N Q S Y R T A V I T Q F N L T S A K L R E T A V Y F S N L
 AGTAGGTTGAATATTGTTTATACCAATATACGCACAAGTAGCAAAATTCAAATTTACTTTTAAAGAGATGGCCATATAATGCACAAAGATGGTCTTTAGCATGTCTGCTGGTGACCA 1080
 V O Y E L L L L P I Y A V A N F N L L L I R D G L I N A Q E W S L A C A G D W T
 ACTATTAACACTATGCTGCTACATTAAGAAATATATTCACACATAGCATACATGGTATAATAAAGGTTTAGATGTACTTAGAATAAATCTAATGGCAATGGATACGTTTAAATGA 1200
 L Y N T M V L I R F S A N K I G F S Y T N S S A M Q E S G I Y G S G F C S N L
 TTATAAAGAGAGTACATTTCAAGTATTAGATATACCTGCTCTTTTGGCAGTTATGATCCACGTGATACCTGGGACAAAATAGATAATCAAACTATCAAAAGCAATTTAC 1320
 Y K R E M T I Q V L D I L A L F A S Y D P R R L P A D K I D N T K L S K T E F T
 AAGAGAGATTATACAGCTTTAGTAGAATCTCTCTAGTAAATCTATAGCAGCACTGGAGGCAAGCTTACAGGAGATGTTCAATTTACTTGGCTAAAGAGAGTATTTTTCGGAC 1440
 R E S I Y T A L V E S P S I A A L E T R D V H L F T W K L R V Y F T
 CAATACTATATCAAGATTAAAGATTTTATCGCCATAAAATTTGGGTTTCATATCAAAATCTCTGCAATGCAAGAAAGTGGAAATTTGGAAGTTTGGTTTGGTTCAAACTC 1560
 N T I Y Q D L R F L S A N K I G F S Y T N S S A M Q E S G I Y G S G F C S N L
 TACTCATCAATCAACTTAATCTAATGTTTATAAACTCTATACAGCACTAGCTGCCCCCTTAATCGAGTTACAAAAATGGATTCTCAAAATATGATGGTACTTTGCTCTTA 1680
 T H Q I T A Q L N S N V Y K T S I T D T S S P S N R V T K M D F Y K I D G D T L A S Y
 TAATTAATATCAACCACTCTGAGGTTTAAAGGACACATTTTGGATTTTCAACAAATGAGAACACACCTAATCAACCACTGTAAGATGATTAACCATTTTAACTATAT 1800
 N S N I T P T P E Q L R T T F F G F S T N E N T P N Q P A C T V N D Y T H I L S Y I
 AAAAACTGATTTATAGATTATACAGTAAACAGGTTTCATTTGCTTGGACACATAGATTTGTCACCTTAATCAAAATATACACAGATGCTATCACAAAGTCCGGCCGTAATATC 1920
 K T D N I D Y N S P R I V F A W T H K I V D P N N Q I Y T D A I T I F V P A V K S
 TAACTCTTGAATGCAACAGCTAAAGTAAATCAAGGACCTGTCATACAGGCGGAGATCTAGTCTCTTACAAGCAATGGTACTCTATACAGGCAAGATGAGATCAATGTAAAAACAG 2040
 N F L N A T A K V I K G P G H T G G D L V A L T S N G T L S G R M E I Q C K T S
 TATTTTAAATGCTCTACAAGAGTTACGGATTACGCATATGCTGCAAAATGCTCAATTTGATGATGATATGATTTACAAGGAGTTTCTAGAGGAACACAGTATAGTAC 2160
 I F N D P T R S Y G L R I R Y A A N S P I V L N V S Y V L Q G G V S R G T T I S T
 ADAATACGCTTTCAAGACCTAATAATATACCTACAGATTAAAAATGAGAGGTTTATGATACAAAGATGCTTTGATGCAATTTGACCAATGATGATGATGATGATGATGATGAT 2280
 E S T F S R P N I I P T D L K Y E F R Y K D P F D A I Y V P H L S N N G I
 AACTATAGCTATCAACCAATCAACGACTTCAATATCAAGTATTATGACAGAACTGAAATTTTCAACTCACTCAATCTGATATAGATGAGACAGAGCAAAATTTAGATCT 2400
 T I A I Q P L N M T S N N Q I I D R I E I I P I T Q S V L D T E N Q N L E S
 AGAACGAGAAATTTGTAATGCACTGTTTCAAAATGACGCAAGATGCAATTAACATTTGGAACGACAGATTATGACATAGATCAAGCGCAAACTTTGTAAGATGATTTTCTGAAGATT 2520
 E R E V A A A A L F T N D A K D A L N I G T T D Y D I D Q A A N L V E C I S E E L
 ATATCAAAAAGAAAATGCTTTTATAGATGATTAATAAGTACGAACTTGTCAATCTGCAAAATGCTCAAAACGCGGATTTGAATCGCTACCGTGGTGGACACAAAG 2640
 Y P K E K H L L L L D E V K N A K Q L S Q S R N V L Q N G D F E S A T L G W T T S
 TGATAATATCAACATTCAGAAAGATGATCTATTTTAAAGGCGATTACCTTCATATGCTGCGGCGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
 D N I T I Q E D D P N I F K G H Y L H M S G A R D I D G T I F P T Y I F Q P A V K S
 ATCAAAATTAACCGCTATACAGCTTACCTAGTAAAGGATTTGTAGAAAGTATGAGAACTAGTGGTTTACGCTATGCGGAGAAATGATGCCATCATGAATTTCCAGC 2880
 S K L K P Y T R Y L V R G F V G S S K D V E L V V S R Y G E E I D A I H N V P A
 TGATTTAACTATCTGATCTCTTACCTTTGATTTGAAAGGCTTAATGTTGTGAGAGCTGCGCTGTCGCGCTAACATTGGGAACACTTGTGATATGTTGATTTATGCAATATGA 3000
 D L N Y L Y P S T F D C E G S N R C E T S A V P A N I G N T S D M L Y S C Q Y D
 TACAGGAAAAAGCATGCTGATGTCAGGATTTCCCATCAATTTAGTTTCTACTATTGATACAGGGGCTTATGATACAAATGAAATATAGGGGTTTGGGTCATGTTTAAATATCTTCTCC 3120
 T G K K H V V C Q D S H Q F S F T I D T G A L D T N E N I G V W M F K I S S P
 AGATAGATCGCATTAAGATATTTAGAAATTAATGAAGAGGCGCAATAGATGGAAGAGCACTGTACGCGTGAAACACATGGAGAGAAATGGAACGATCAATGGAGGAGCAAAAG 3240
 D O Y A L D N L E I E E G P I D G E A L S R V K H M E K K W Q M E A K R
 TTCGGAACACACAAAGCATATGATGATGCAAGAACAGCCATTGATGCTTTATTCACAAATGACAAAGTACAGGCTTACAGTTTATGATGACGACCTGCTCAAAATCAGTACGCTGAGTA 3360
 S E T Q Q A Y D V A K Q A I D A L F T N V Q D E A L Q F D T T L A Q I Q Y A E Y
 TTTGGTACAACTGATTTCCATATGTTGAT 3480
 L V Q S I P Y V Y N D W L S D V P Q M N Y D I Y V E L D A R V A Q A R Y L Y D T
 AAGAAATATTTAAAAATGGTGATTTTACCAAGGGTAAATGGGGTGGCATGTAATGGAATGACAGAGTACAACAAATAGATGGTGTCTGTATTTGGTTCTATCAATTTGAGATGC 3600
 R N I I K N G N F I T G V M Q W H V T G N A D V Q Q I D G V S L V L S N W S A
 TGGCGATGCTCAAAATGCTCACTCCCAATATCATGGTATGCTTACCTGTTTATGCAAAAAAGAGGACCTGGAATGGGATGTCACGCTTATGATGTTGGAGAGATCAAGA 3720
 G V S Y V H L Q H N H O Y V L R V I A K K E G P G N G Y V T L H D C E E N Q E
 AAAATTGAGCTTTACGCTTTGTGAAGAAGGATATATTACGAAGACAGTATGATGTTTCCAGATACAGATGCTGTACGAATGAGATGAGGCAACCGAAGGTTGTTTATATCGAAAG 3840
 K L T F T S C E E G Y I T K T V D V F P D T D R V R I E I G E T E G S F Y I E S
 CATTGAATTAATTTGCATGAACGAGTGAATTAATAAAAAATCACTAAAGCTTTAAAAACCATGGGAGAAAGTTTCTCCATGGTTTTAATTTCTGCATTTTAAATTTCTGGTACAAAAAT 3960
 I E L I C H N E
 ATATAGAAACATAAAAAATAGATATCTAGAGGACATAAAATTTATCAAAATATCAATTTTCTAGTATAGAACGTTTATCCAATAAATATTCACCATTTAAACTATCCAACTACAA 4080
 TTGTGATCCAGTTTATGATGATTAATAGATGAAGTATTTAGGATTTGAAATAGATTTTATCAACCAAAATAGGGGAAAGAAATAGCAGGGGAAATCAACAATCG 3'

(b) ISRH4

5' AAGCTTTCGTGAT // ATATTAAATACGAATCTTTAAATGCATAGGGGAAATTTGATAGCTTGGAGTGTAAAGAAAGAGAGATTACTACATAGGAGAGGATAAATGCATAAGG 1080
S F R D

GTCAAAGATACCAAGAAATCTTGTAAGAAGAGTTTCTTATTTTGTGTGAGGTTTGTACAAACCATATGAAATACAAAGTTTAAAAATTTAACGCAAAAAATTTTAAAAATGTATAA 1200

ATTACGAATAATTATATTTGGTACATAATATGTATATAGGATTTGTAAAAGTAATCAATATTTTCAATTTTGTACTTTTAAACCATATCAATTTAAAAATGATTTAAAAATATCTTG 1320

TTGCTACAGAAAAAGAGTGTATCTAAATTTGAGTATGGGAGGACAAATATCAATCTTATCAAAAATAAAGTAATGAAACATTAATGCTTCACAAAAAAATTTAAATATCTCTAAT 1440
M N P Y K N K N E Y E T L N S A N K I L N R S

SD

AATTATACAAGATATCCAATAGAAAAATGTCACAAACAAATTTACAAAGTACAAATATAAAGATTTGGCTCAATATGTGTCAACAGAAATCAGCAGTATGTGGAGATTTGAAACTTTT 1560
N Y T R Y P I E N S P K Q L L Q S T N Y K D W L N M C Q Q N Q Q Y G G D F E T F

ATTGATAGTGTGAACCTGACCTATACTATTGTAGTTGGGACCTGACTGCTTTGGGGTTTCAACACCCCTTAGGACTTGCTTTAAATAGGTTTGGTACATTAATACCAAGTTCTT 1680
I D S G E L S A Y T I V V G T V L T G F G F T T P L G L A L I G F G T L I P V L

TTTCCAGCCCAAGACCAATCTAACACATGGAGTACTTTATACACAACTAAAAATATTATAAAAAAGAAATAGCATCAACATATATAAGTAATGCTATAAAATTTTAAACAGGTGG 1800
F P A Q D Q S N T W S D F I T Q T K N I I K K E I A S T Y I S N A N K I L N R S

TTTAATGTTATCAGCACTATCATATAACCTTTAAACATGGGAGAAATATCCAAACCCACAAATACTCAGGATGTAAAGACACAAATCCAGCTAGTTCATTACCATTTCCTAAAAATGTC 1920
F N V I S T Y H N H L K T W E N N P N P Q N T Q D V R T Q I Q L V H Y H F Q N V

ATTCAGAGCTTTGTAACCTCTTCTCCTCTAATCTAGTGATTGGGATTAATAACATAGTATTAATCTAGTATGACAGCAGGAACTTACATCTGACTGTTATTAATCAAGGC 2040
I P E L V N S C P P N P S D C D Y Y N I L V L S S Y A Q A A N L H L T V L N Q A

GTCAAATTTGAAGCGTATTTAAAAACAAATCGCAATCTGATTATTTAGAGCCTTTGCCAACGCAATGATTATTTCCAGTATTTGCAATAGGATTTAGAGATTACACTAATTTGT 2160
V K F E A Y L K N N R P Q F D Y L P T A I D Y Y P V L T K A I E D Y T N Y C

GTAACACTTTAAAAAGGATTAATTTAAACGACCGCTGATGATTAATCTTGGGAAATATAAAGTGAACACATCAATCGTATGGAACAAAAATGACTACTGCTGCTGATTA 2280
V T T Y K K G L N L I K T T P D S N L D G N I N W N T Y N T K M T T A V L

GATCTTGTGCACTCTTCTCAATATATAGTATGATTAATCAATAGGTGCTCAATCTGAAGTACTGAGAAATTTATCAGGTAATTAATCTGGAAGAAAGCCCTATAAATATTAT 2400
D L V A L F P N Y D V G K Y P I G V Q S E L T R E I V Q V L N F E E S P Y K Y Y

GACTTTCAATCAAGAGGATTTCACTACACGTAGACCGGATTTATTTACTTGGCTGATTTCTTGAATTTTATGAAAAGCGGAACTTCTCAATATTTTTCACGAGCCATTAT 2520
D F Q Y Q E D S L T R R P H L F T W L D S L N F Y E K A Q T T P N N F F T S H Y

AAATGTTTCACTACACCTTGATTAATATATCCCAAAATCTACTGCTTTTGGAAATCAACAATGATAAATTTAAATCTTTGGTTGGGCAACAAATATTTATATTTTATTA 2640
N M F H Y T L D N I K K S S V F G N H N V T D K L K S L G L K L N I Y I F L L

AATGTCATAAGCTTAGATAAATAATCTAATGATTATATAAATATAGTAAATGGATTTTTTATACTAATGGTACTAGACTTTTGGGAAGAAAGAACTTACAGCAGGATCTGGGCAA 2760
N V I S L D K K Y L N D Y N N I S K M D F F I T N G T R L L E K E L T T A G S G Q

ATAACTTATGATGATAAATAAATTTTGGGTTTACCAATCTTAAACGAGAGAGAACTCAAGGAAACCTTACCTTTTCCAAACATATGATTAATCTTATTTATCTATTAT 2880
I T Y D V N K N I F G L P I L K R R E N Q G N P T L F P T Y D N Y S H I L S F I

AAAAGCTTAGTATCCCTGCAACATAAATACTCAAGTGTATACGTTTGGCTGGACACACTCTAGTGTGATCTAAAAATACAAATTTATACACATTTAATACCCAAATCCAGCTGTA 3000
K S L S I P A T Y K Y T F A W T H S S V D P K N T I Y T H L T T G I P A V

AAAGCGAATTTCACTGGGAGTCTCTCAAGGTTGTTCAAGGACCTGTCATACAGGAGGGGATTAATGATTTCAAGATCACTTTCAAAATACATGTCACCACTCAATTTTCAACAA 3120
K A N S L G T A S K V Y Q P G H T G G D L I D F K D H F K I T C Q H S N F Q Q

TCGTATTTATAAGAAATCTGATGCTTCAAAATGGAAGCCCAATCTCAAGCTGTTATAAATCTTAGTATCCAGGGGTAGCAGCAATGGGTATGGCACTCAACCCCAATTTTCTGGT 3240
S Y F I R I R Y A S N O S A N T R A V I N L S I P G V A E L G M A L N P T F S G

ACAGATTATACGAATTTAAATATAAGATTTTTCAGTACTAGAAATTTTCTCAAGAGGTGAAATTTGCTCCAAATCAAAATCATCTCTCTGTTTAACTGCTGGATGTATATCAAACT 3360
T D Y T N L K Y K D F Q Y L E F S N E V K F A P N Q N I S L V F N R S D V Y T A

ACACAGTACTTATGATAAATGAAATTTTTCGCAATTTACTCGTTCTATAAGAGAGGATAGAGAGAAACAAATTTAGAAACAGTACAAACAAATTAATACATATTTTATGCAATCTCT 3480
T T V L I D K I E L P I T R S I R E D D R E K Q K L E T V C Q I I N T F Y A N P

ATAAAAAACCTTTCAACATCAGAACTTACAGATTATGACATAGATCAAGCCCAATCTTGTGGAATGATTTCTGAAGAAATATATCCAAAAAGAAAAATGCTGTTATAGATGAAGTT 3600
I K N T L Q S E L T D Y D I D Q A A N L V E C I S E E L Y P K E K M L L L D E V

AAAAAGGAAACCACTTACTGATGCTTCAAAATGACTTCAAAAGGGGATTTTGAATCGGCTACGCTTGGTTGGACAAAGTGAATAATCAACAATCAAGAAAGATGCTCTATTTT 3720
K N A K Q L S Q S R N V L Q N G D F E S A T L G W T T S D N I T I Q E D D P I F

AAAGGCAATACCTTCATATGCTGGGGCGAGACATATGATGATGATGATTTCCGACCTATATATCCAAAAATTTGATGAATCAAAATTTAAACCGGTATACACGTTACCTAGTAAGG 3840
K G H Y L H M S G A R D I D G T I F P T Y I F Q K I D S E K L P Y T R Y L V R

CGATTGTAGGAAGTATGATGAGATGAGATTTTCAAGCTTGGGAGAAATGATGCCATCAATGATGCTTCAAGTATTTAACTATCTGATCTCTTCACTTCTGATTTGATTTG 3960
G F V G S S K D V E L V V S R Y G E E I D A I M N V P A D L N Y L P C T S T F D C

GAAGGCTAATCGTTGTGAGACGTCGCTGTGCGGCTAACATTTGGGAACACTTCTGATATGTTGATTCATGCCAATATGATACAGGAAAAAGCATGTCGATATGTCAGGATTCCTAAT 4080
E G S N R C E T S A V P A N I G N T S D M L Y S C Q Y D T G K K H V V C Q D S H

CAATTTAGTTTCACTATTGATACAGGGGCAATAGATACAAATGAAATATAGGGGTTTGGGTCATGTTTAAAAATCTTCTCCAGATGGATACGCATATAGATAATTTAGAAGTAAT 4200
T F S F T I D T G A L D T N E I I G V W V H F K I S S P D G Y A S L D N L E V I

GAAGAAGGCCAATAGATGGGAGCAGCTGACCGGTGAACACATGGAAGAAATGGAACGATCAATGGAAGCAAAACGTTTGGGAAACACAAAGCATATGATGATAGGCAAAACAA 4320
E E G P I D G E L A G S R N V L Q N G D F E S A T L G W T T S D N I T I Q E D D P I F

GCCATTGATGCTTTATTCACAAATGACAAAGATAGGCGTTTACAGTTTGTATACGACACTCGCTCAAAATTCAGTACGCTGAGTATTTGGTACAAATCGATTCCATATGTCACAAATGAT 4440
A I D A L F T N V Q D E A L Q F D T T L A Q I Q Y A E Y L V Q S I P Y V Y N D W

TTGTCAAGTGTCCAGGTATGAATATGATATCTATGATGAGTTGGATGACGAGTGGCACAAGCGGCTATTGTTGATGATACAGAAATATTTAAAAATGATGATTTTACACAAGGG 4560
L S D V P G M N Y D I Y V E L D A R V A Q A R Y L Y D T R N I I K N G D F T Q G

GTAATGGGGTGGCATGTAAGTGAATGCAAGCATACAAATAGATGTTGTTCTGATTTGGTTCTATCTAATTTGGAAGTGTGCTGGCGTATCTCAAAATGTCCTTCCAAACATATCAT 4680
V M G W H V T G N A D I D G V S V L V L S N W S A G V S L V F N R S D V Y T A

GGGTATGTTAGCTGTTTATGCAAAAGAGAAAGACCTGGAAATGGGATGTCAGCTTATGAGTGTGAGGAGAAATGAGGAAATTTGAGCTTTACGCTTCTGGAAGAGGATATAT 4800
G Y V L R V I A K K E G P G N G Y V T L M D C E N Q E K L T F T S C T G E E G Y I

ACGAAGACAGTATGATGATTTCCAGATACAGATCTGTATGCAATTTGAGATAGGCGAAACGGAAGGTTGCTTTATATCGAAAGCATTTGAATTAATTTGATGAGCAGGATTAATAAAAA 4920
T K T V D V F P D T D R V R I E I G E T E G S F Y I E S I E L I C M N E

AAAACTAAAGCTTTAAAAACCATGGAGAAAGTTTCTCCATGGTTTAAATTTTCTGCAATTTATTAATTTGCTGTTACAAAAATATATAGAAAAACATAAAAAATAGATATCTAGAGGACAT 5040
AAAAATTTATACAAATATCAATTTTCTAGTATGAGACGTTTATCCAAATTAATATTCACCATTTAACTATCCAAACTACAATTTGTATCCAAAGTTATGATGATTAATAGATGAACTGA 5160
TTTAGGATTTGAAATAGATTTTCTATCAACAAAAATGAGGAAAGAAATATGACAGGAAATCAACAATCG 3'

Fig. 2. Nucleotide Sequences and Deduced Amino Acids for the Genes of ISRH3 (a) and ISRH4 (b).

Both sequences are shown in the 5'-to-3' direction. SD indicates a Shine-Dalgarno sequence and X represents an unknown protein. Underlined bases indicate the locations where amino acids are replaced in the insecticidal protein reported by Ward and Ellar.¹⁵ The tract of 963 bases (from T₁₄ to G₉₇₆) upstream of the ISRH4 gene was omitted.

Nucleotide sequences in the 3'-terminal regions (about 40% of the structural gene) of ISRH3 and ISRH4 were identical. Those in the 3'-flanking regions (318 nucleotides) were also identical. Twenty-two nucleotides downstream of the termination codon there was an inverted repeat which may function as a termination signal (Fig. 2). In the 5'-terminal or N-terminal region (about 60% of the structural gene), homology between the two genes was relatively low with respect to the nucleotide sequences and the predicted amino acid sequences.

Comparison among various insecticidal protein genes from B. thuringiensis var. israelensis

Angsuthanasommbat *et al.* reported a nucleotide sequence of 5'-flanking and 5'-terminal regions of a 130 kDa protein gene.¹⁴⁾ The sequence is identical to that of the ISRH3 gene except an additional A at nucleotide position 863 in our sequence. The nucleotide sequence in the ISRH4 structural gene was identical to that of a 130 kDa protein gene which Ward and Ellar have recently reported,¹⁵⁾ except that the C at 2284, C at 4523, C at 4693, and T at 4749 were replaced by G, T, G and G, respectively (Fig. 2). These base changes resulted in four amino acid replacements, from Leu to Val, from Thr to Ile, from Arg to Gly, and from Cys to Trp, whose location are indicated by underlines in Fig. 2. The divergence in the amino acid sequence may come from using different bacterial strains to isolate the insecticidal protein genes.

Thorne *et al.*¹²⁾ reported a 58 kDa protein gene (ORF1) and an unknown protein gene (ORF2) which was 69 nucleotides downstream of ORF1. We could recognize homology in amino acid sequences between the two reported protein (ORF1 and 2) and ISRH3, and that between the two reported proteins and ISRH4.

We analyzed homology between the 28 kDa protein^{8,9)} and ISRH3, and that between 28 kDa protein and ISRH4, but there was no homology. This evidence indicates that *BTI* can produce at least two different types of

insecticidal proteins. We have confirmed synthesis of the two insecticidal 130 kDa proteins under the direction of cloned ISRH3 and ISRH4 genes in *E. coli* cells. They are insecticidal against mosquito larvae (the results will be reported elsewhere).

We suppose that there are two major groups of insecticidal protein genes. One is the 28 kDa protein gene, whose products are insecticidal and hemolytic.^{7,13)} The other is a group of the 58 kDa and two types of the 130 kDa protein genes, whose products are insecticidal but not hemolytic.^{11,13,14)} It seems that only the 28 kDa protein is evolutionally different from the other insecticidal protein of *BTI*. This may relate to a difference of toxicity between the 28 kDa protein and the other.

Acknowledgments. We wish to thank Dr. H. Ohkawa, (Sumitomo Chemical Co., Ltd., Japan) for preparing antibodies, and Dr. T. Yamamoto (Cotton Insects Res., ARS, USDA, U.S.A.) and D. M. Ellis (*Bacillus* Genetic Stock Center, Ohio State Univ., U.S.A.) for their kind supplies of *B. thuringiensis* strains.

This investigation was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- 1) L. J. Golderg and J. Margalit, *Mosquito News*, **37**, 355 (1977).
- 2) W. E. Thomas and D. J. Ellar, *J. Cell Sci.*, **60**, 181 (1983).
- 3) M. A. Pfannenstiel, E. J. Ross, V. C. Kramer and K. W. Nickerson, *FEMS Microbiol. Lett.*, **21**, 39 (1984).
- 4) S. G. Lee, W. Eckblad and L. A. Bulla, Jr., *Biochem. Biophys. Res. Commun.*, **126**, 953 (1985).
- 5) M. Himeno, M. Ikeda, K. Sen, N. Koyama, T. Komano, H. Yamamoto and I. Nakayama, *Agric. Biol. Chem.*, **49**, 573 (1985).
- 6) J. M. Gonzalez, Jr. and B. C. Carlton, *Plasmid*, **11**, 28 (1984).
- 7) E. S. Ward, D. J. Ellar and J. A. Todd, *FEBS Lett.*, **175**, 377 (1984).
- 8) E. S. Ward and D. J. Ellar, *J. Mol. Biol.*, **191**, 1 (1986).
- 9) C. Waalwijk, A. M. Dulleman, M. E. S. van Workum and B. Visser, *Nucl. Acids Res.*, **13**, 8207 (1985).
- 10) V. Sekar and B. C. Carlton, *Gene*, **33**, 151 (1985).
- 11) V. Sekar, *Biochem. Biophys. Res. Commun.*, **137**, 748 (1986).

- 12) L. Thorne, F. Garduno, T. Thompson, D. Decker, M. Zounes, M. Wild, A. M. Walfeld and T. J. Pollock, *J. Bacteriol.*, **166**, 801 (1986).
 - 13) C. Bourgouin, A. Klier and G. Rapoport, *Mol. Gen. Genet.*, **205**, 390 (1986).
 - 14) C. Angsuthanasombat, W. Chungjatupornchai, S. Kertbundit, P. Luxananil, C. Settasatian, P. Wilairat and S. Panyim, *Mol. Gen. Genet.*, **208**, 384 (1987).
 - 15) E. S. Ward and D. J. Ellar, *Nucl. Acids Res.*, **15**, 7195 (1987).
 - 16) B. Visser, M. van Workum, A. Dulleman and C. Waalwijk, *FEMS Microbiol. Lett.*, **30**, 211 (1986).
 - 17) C. Y.-Perron, J. Vieira and J. Messing, *Gene*, **33**, 103 (1985).
 - 18) K. W. C. Peden, *Gen*, **22**, 277 (1983).
 - 19) H. C. Birnboim and J. Doly, *Nucl. Acids Res.*, **7**, 1513 (1979).
 - 20) M. Imagawa, S. Yoshitake, Y. Hamaguchi, E. Ishikawa, Y. Niitsu, I. Urushizaki, R. Kanazawa, S. Tachibana, N. Nakazawa and H. Ogawa, *J. Appl. Biochem.*, **4**, 41 (1982).
 - 21) J. Messing, *Methods Enzymol.*, **103**, 20 (1983).
-