The complete nucleotide sequence of a 130 kDa mosquito-larvicidal delta-endotoxin gene of *Bacillus thuringiensis* var. israelensis

Sumalee Tungpradubkul, Chatri Settasatien and Sakol Panyim*

Center for Molecular Genetics, Genetic Engineering and Biochemistry Department, Faculty of Science, Mahidol University, Rama VI Rd, Bangkok 10400, Thailand

A 130 kDa mosquito-larvicidal delta-endotoxin gene of *B. thuringiensis* var. israelensis (B.t.i.) was identified as previously described (1). The complete nucleotide sequence of the gene is presented below. It encodes a protein of 1136 amino acids with a calculated molecular weight of 127, 863 Da. The protein shows little amino acid homology to a 130 kDa endotoxin of *B. thuringiensis* subsp. kurstaki (2), but its C-terminal region (under-lined) is almost identical to another 130 kDa delta-endotoxin of B.t.i. (3).

```
5'aattgttcataggaatccgtatcaatttttcaaggaatatgtatttgcacttttggtcttttaaatcgtatgaattcaaaatagttta 90

  tatcaatctttgttacaccagaaaaagattgtatccaatgtgaatatgggaggaataaatATGAATTCAGGCTATCCGTTAGCGAATGAC 180
                                                               M  N  S  G  Y  P  L  A  N  D
TTACAAGGGTCAATGAAAAACACGAACTATAAAGATTGGCTAGCCATGTGTGAAAATAACCAACAGTATGGCGTTAATCCAGCTGCGATT 270
L  Q  G  S  M  K  N  T  N  Y  K  D  W  L  A  M  C  E  N  N  Q  Q  Y  G  V  N  P  A  A  I
AATTCTTCTTCAGTTAGTACCGCTTTAAAAGTAGCTGGAGCTATCCTTAAATTTGTAAACCCACCTGCAGGTACTGTCTTAACCGTACTT 360
N  S  S  S  V  S  T  A  L  K  V  A  G  A  I  L  K  F  V  N  P  P  A  G  T  V  L  T  V  L
AGCGCCGGTGCTTCCTATTCTTTGGCCGACTAATACTCCAACGCCTGAAAGAGTTTGGAATGATTTCATGACCAATACAGGGAATCTTATT 450
S  A  V  L  P  I  L  W  P  T  N  T  P  T  P  E  R  V  W  N  D  F  M  T  N  T  G  N  L  I
GATCAAACTGTAACAGCTTATGTACGAACAGATGCAAATGCAAAAATGACGGTTGTGAAAGATTATTTAGATCAATATACAACTAAATTT 540
D  Q  T  V  T  A  Y  V  R  T  D  A  N  A  K  M  T  V  V  K  D  Y  L  D  Q  Y  T  T  K  F
AACACTTGGAAAAGAGAGCCTAATAACCAGTCCTATAGAACAGCAGTAATAACTCAATTTAACTTAACCAGTGCCAAACTTCGAGAGACC 630
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
GCAGTTTATTTTAGCAACTTAGTAGGTTATGAATTATTGTTATTACCAATATACGCACAAGTAGCAAATTTCAATTTACTTTTAATAAGA 720
A  V  Y  F  S  N  L  V  G  Y  E  L  L  L  L  P  I  Y  A  Q  V  A  N  F  N  L  L  L  I  R
GATGGCCTCATAAATGCACAAGAATGGTCTTTAGCACGTAGTCGTGGTGACCAACTATATAACACTATGGTGCAGTACACTAAAGAATAT 810
D  G  L  I  N  A  Q  E  W  S  L  A  R  S  R  G  D  Q  L  Y  N  T  M  V  Q  Y  T  K  E  Y
ATTGCACATAGCCATTACATGGTATAATAAAGGTTTAGATGTACTTAGAAATAAATCTAATGGACAATGGATTACGTTTAATGATTATAAA 900
I  A  H  S  I  T  W  Y  N  K  G  L  D  V  L  R  N  K  S  N  G  Q  W  I  T  F  N  D  Y  K
AGAGAGATGACTATTCAAGTATTAGATATACTCGCTCTTTTTGCCAGTTATGATCCACGTCGATACCCTGCGGACAAAATAGATAATACG 990
R  E  M  T  I  Q  V  L  D  I  L  A  L  F  A  S  Y  D  P  R  R  Y  P  A  D  K  I  D  N  T
AAACTATCAAAAACAGAATTTACAAGAGAGATTTATACAGCTTTAGTAGAATCTCCTTCTAGTAAATCTATAGCAGCACTGGAGGCAGCA 1080
K  L  S  K  T  E  F  T  R  E  I  Y  T  A  L  V  E  S  P  S  S  K  S  I  A  A  L  E  A  A
CTTACACGAGATGTTCATTTATTCACTTGGCTAAAGAGAGTAGATTTCTGGACCAATACTATATATCAAGATTTAAGATTTTTATCTGCC 1170
L  T  R  D  V  H  L  F  T  W  L  K  R  V  D  F  W  T  N  T  I  Y  Q  D  L  R  F  L  S  A
AATAAAATTGGGTTTTCATATACAAATTCTTCTGCAATGCAAGAAAGTGGAATTTATGGAAGTTCTGGTTTTGGTTCAAATCTTACTCAT 1260
N  K  I  G  F  S  Y  T  N  S  S  A  M  Q  E  S  G  I  Y  G  S  S  G  F  G  S  N  L  T  H
CAAATTCAACTTAATTCTAATGTTTATAAAACTTCTATCACAGATACTAGCTCCCCCTCTAATCGAGTTACAAAAATGGGATTTCTACAAA 1350
Q  I  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
ATTGATGGTACTCTTGCCTCTTATAATTCAAATATAAACACCAACTCCTGAAGGTTTAAGGACCACATTTTTTGGATTTTCAACAAATGAG 1440
I  D  G  T  L  A  S  Y  N  S  N  I  T  P  T  P  E  G  L  R  T  T  F  F  G  F  S  T  N  E
AACACACCTAATCAACCAACTGTAAATGATTATACGCATATTTTAAGCTATATAAAAACTGATGTTATAGATTATAACAGTAACAGGGTT 1530
N  T  P  N  Q  P  T  V  N  D  Y  T  H  I  L  S  Y  I  K  T  D  V  I  D  Y  N  S  N  R  V
TCATTTGCTTGGACACATAAGATTGTTGACCCTAATAATCAAATATACACAGATGCTATCACACAAGTTCCGGCCCGTAAAATCTAACTTC 1620
S  F  A  W  T  H  K  I  V  D  P  N  N  Q  I  Y  T  D  A  I  T  Q  V  P  A  V  K  S  N  F
TTGAATGCAACAGCTAAAGTAATCAAGGGACCTGGTCATACAGGGGGGGGATCTAGTTGCTCTTACAAGCAATGGTACTCTATCAGGCAGA 1710
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
```

```
ATGGAGATTCAATGTAAAACAAGTATTTTTAATGATCCTACAAGAAGTTACGGATTACGCATACGTTATGCTGCAAATAGTCCAATTGTA 1800
 M  E  I  Q  C  K  T  S  I  F  N  D  P  T  R  S  Y  G  L  R  I  R  Y  A  A  N  S  P  I  V
TTGAATGTATCATATGTATTACAAGGAGTTTCTAGAGGAACAACGATTAGTACAGAATCTACGTTTTCAAGACCTAATAATATAATACCT 1890
 L  N  V  S  Y  V  L  Q  G  V  S  R  G  T  T  I  S  T  E  S  T  F  S  R  P  N  N  I  I  P
ACAGATTTAAAATATGAAGAGTTTAGATACAAAGATCCTTTTGATGCAATTGTACCGATGAGATTATCTTCTAATCAACTGATAACTATA 1980
 T  D  L  K  Y  E  E  F  R  Y  K  D  P  F  D  A  I  V  P  M  R  L  S  S  N  Q  L  I  T  I
GCTATTCAACCATTAAACATGACTTCAAATAATCAAGTGATTATTGACAGAATCGAAATTATTCCAATCACTCAATCTGTATTAGATGAG 2070
 A  I  Q  P  L  N  M  T  S  N  N  Q  V  I  I  D  R  I  E  I  I  P  I  T  Q  S  V  L  D  E
ACAGAGAACCAAAATTTAGAATCAGAACGAGAAGTTGTGAATGCACTGTTTACAAATGACGCGAAAGATGCATTAAACATTGGAACGACA 2160
 T  E  N  Q  N  L  E  S  E  R  E  V  V  N  A  L  F  T  N  D  A  K  D  A  L  N  I  G  T  I
GATTATGACATAGATCAAGCCGCAAATCTTGTGGAATGTATTTCTGAAGAATTATATCCAAAAGAAAAAATGCTGTTATTAGATGAAGTT 2250
 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
AAAAATGCGAAACAACTTAGTCAATCTCGAAATGTACTTCAAAACGGGGATTTTGAATCGGCTACGCTTGGTTGGACAACAAGTGATAAT 2340
 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
ATCACAATTCAAGAAGATGATCCTATTTTTAAAGGGCATTACCTTCATATGTCTGGGGCGGAGAGACATTGATGGTACGATATTCCGACC 2430
 I  T  I  Q  E  D  D  P  I  F  K  G  H  Y  L  H  M  S  G  A  R  D  I  D  G  T  I  F  P  T
TATATATTCCAAAAAATTGATGAATCAAAATTAAAACCGTATACACGTTACCTAGTAAGGGGATTGTAGGAAGTAGTAAAGATGTAGAA 2520
 Y  I  F  Q  K  I  D  E  S  K  L  K  P  Y  T  R  Y  L  V  R  G  F  V  G  S  S  K  D  V  E
CTAGTCGGTTTCACGCTATGGGGAAGAAATTGATGCCATCATGAATGTTCCAGCTGATTTAAACTATCTGTATCCTTCTACCTTTGATTGT 2610
 L  V  V  S  R  Y  G  E  E  I  D  A  I  M  N  V  P  A  D  L  N  Y  L  Y  P  S  T  F  D  C
GAAGGGTCTAATCGTTGTGAGACGTCCGCTGTGCCGGCTAACATTGGGAACACTTCTGATATGTTGTATTCATGCCAATATGATACAGGG 2700
 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
AAAAAGCATGTCGTATGTCAGGATTCCCATCAATTTAGTTTCACTATTGATACAGGGGCATTAGATACAAATGAAAATATAGGGGTTTGG 2790
 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
GTCATGTTTAAAATATCTTCTCCAGATGGATACGCATCATTAGATAATTTAGAAGTAATTGAAGAAGGGCCAATAGATGGGGAAGCACTG 2880
 V  M  F  K  I  S  S  P  D  G  Y  A  S  L  D  N  L  E  V  I  E  E  G  P  I  D  G  E  A  L
TCACGCGTGAAACACATGGAGAAGAAATGGAACGATCAAATGGAAGCAAAACGTTCGGAAACACAACAAGCATATGATGTAGCGAAACAA 2970
 S  R  V  K  H  M  E  K  K  W  N  D  Q  M  E  A  K  R  S  E  T  Q  Q  A  Y  D  V  A  K  Q
GCCATTGATGCTTTATTCACAAATGTACAAGATGAGGCTTTACAGTTTGATACGACACTCGCTCAAATTCAGTACGCTGAGTATTTGGTA 3060
 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
CAATCGATTCCATATGTGTACAATGATTGGTTGTCAGATGTTCCAGGTATGAATTATGATATCTATGTAGAGTTGGATGCACGAGTGGCA 3150
 Q  S  I  P  Y  V  Y  N  D  W  L  S  D  V  P  G  M  N  Y  D  I  Y  V  E  L  D  A  R  V  A
CAAGCCGCGTTATTTGTATGATACAAGAAATATTATTAAAAATGGTGATTTTACACAAGGGGTAATGGGGTGGCATGTAACTGGAAATGCA 3240
 Q  A  R  Y  L  Y  D  T  R  N  I  I  K  N  G  D  F  T  Q  G  V  M  G  W  H  V  T  G  N  A
GACGTACAACAAATAGATGGTGTTTCTGTATTGGTTCTATCTAATTGGAGTGCTGGCGTATCTCAAAATGTCCATCTCCAACATAATCAT 3330
 D  V  Q  Q  I  D  G  V  S  V  L  V  L  S  N  W  S  A  G  V  S  Q  N  V  H  L  Q  H  N  H
GGGTATGTCTTACGTGTTTATTGCCAAAAAAGAAGGACCTGGAAATGGGTATGTCACGCTTATGGATTGTGAGGAGAATCAAGAAAAATTG 3420
 G  Y  V  L  R  V  I  A  K  K  E  G  P  G  N  G  Y  V  T  L  M  D  C  E  E  N  Q  E  K  L
ACGTTTACGTCTTGTGAAGAAGGATATATTACGAAGACAGTAGATGTATTCCCAGATACAGATCGTGTACGAATTGAGATAGGCGAAACC 3510
 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
GAAGGTTCGTTTTATATCGAAAGCATTGAATTAATTTGCATGAACGAGTGAttaataaaaaataactaaagctttaaaaaccatggagaa 3600
 E  G  S  F  Y  I  E  S  I  E  L  I  C  M  N  E  *
agttttctccatggtttttaatttctgcatttattaattctggtacaaaaaatatatagaaaacataaaaaatagatatctaga3'
```

*To whom correspondence should be addressed

REFERENCES
(1) Angsuthanasombat, C. et al. (1987) Mol. Gen. Genet. 208, 384–389. (2) Geiser, M. et al. (1986) Gene 48, 109–118. (3) Ward, E.S. and Ellar, D. (1987) Nucl. Acids Res. 15, 1975. This work was supported by the National Center for Genetic Engineering and Biotechnology, Ministry of Science, Technology and Energy, Bangkok and USAID grant No. 936-5542. We thank S. Yasothornsrikul and M. Pao-intara for assistance.