Bacillus sphaericus strain 2362: identification and nucleotide sequence of the 41.9kDa toxin gene

Colin Berry and John Hindley

University of Bristol, Department of Biochemistry, Medical School, University Walk, Bristol BS8 1TD, UK
Submitted June 17, 1987 Accession no. Y00378

Oligonucleotide probes designed on the basis of the known N-terminal 40 amino -acid sequence of the B.sphaericus 2362 larvicidal toxin (1) were used to identify an EcoRI-HindIII fragment containing the entire coding sequence. Sequence analysis showed an ORF of 1110 nucleotides corresponding to a 41.9-kDa protein, in agreement with an estimated 43-kDa by gel electrophoresis (1). Features of note are (i) an additional four amino-acids at the N-terminus which are not found in the purified toxin, (ii) homology of the tetrapeptide with that deduced for the N-terminus of B.thuringiensis var. israelensis and morrisoni (2), i.e. MRNL and MENL respectively, (iii) Cys replacing a reported Ser at residue 31 (no.27 in ref.1), (iv) extensive homology in the upstream control regions to B.t. subsp. kurstaki and israelensis (overlined). The putative Shine-Dalgarno sequence is boxed and a downstream inverted repeat underlined.

```
ACCTATAACTAATCCACTTACGCTAACAAAACATACAATTATTCGATGTGAAAATAGTTACGATGGACACATATTTAAAACACCTTTAATCTTTAAAAATGGTGAAGTTATTGTAAAAAC
                                                                                                                        120
GAATGAAGAATTAATACCTAAAATTAACCAGTGATACTTTAACTTCAAATATTCATTACCATGTTATTTAAAATAGTAGATAGATGAAATAAATAGTATATATTAAGACAACAACTTAAT
                                                                                                                        240
                                                                                                  M  R  N  L  D  F  I  D  S  F  I  P
TTTGACACATAAGAATAATTTTTAAATGTATAAATAGTATTTAGAGTGTTATTGCAATATATTTTTGAAAGGGAGCTAAAAGACATGAGAAATTTGGATTTTATTGATTCTTTTATACC
                                                                                                                        360
 T  E  G  K  Y  I  R  V  M  D  F  Y  N  S  E  Y  P  F  C  I  H  A  P  S  A  P  N  G  D  I  M  T  E  I  C  S  R  E  N  N
CACAGAAGGAAAGTACATTCGCGTTATGGATTTTTATAATAGCGAGTATCCTTTCTGTATACATGCACCCTCAGCCCCTAATGGGGATATCATGACAGAAATCTGTAGCAGAGAAAATAA
                                                                                                                        480
 Q  Y  F  I  F  F  P  T  D  D  G  R  V  I  I  A  N  R  H  N  G  S  V  F  T  G  E  A  T  S  V  V  S  D  I  Y  T  G  S  P
TCAATATTTTATTTTTTTTCCTACTGATGATGGTCGAGTAATTATTGCAAATAGGCATAATGGGTCCGTTTTTACCGGAGAAGCCACAAGTGTAGTATCAGATATCTATACTGGTAGCCC
                                                                                                                        600
 L  Q  F  F  R  E  V  K  R  T  M  A  T  Y  Y  L  A  I  Q  N  P  E  S  A  T  D  V  R  A  L  E  P  H  S  H  E  L  P  S  R
ATTACAGTTTTTTAGAGAGGTCAAAAGAACTATGGCAACTTATTATTTAGCGATACAAAATCCTGAATCCGCAACAGATGTGAGAGCTCTAGAACCGCATTCCCATGAGCTGCCATCTCG
                                                                                                                        720
 L  Y  Y  T  N  N  I  E  N  N  S  N  I  L  I  S  N  K  E  Q  I  Y  L  T  L  P  S  L  P  E  N  E  Q  Y  P  K  T  P  V  L
CCTTTATTACACTAACAATATTGAAAATAATAGCAACATATTAATTTCTAATAAGGAACAAATATATTTAACCTTGCCTTCACTTCCAGAAAACGAGCAATACCCTAAAACTCCAGTATT
                                                                                                                        840
 S  G  I  D  D  I  G  P  N  Q  S  E  K  S  I  I  G  S  T  L  I  P  C  I  M  V  S  D  F  I  S  L  G  E  R  M  K  T  T  P
AAGCGGTATCGATGATATAGGACCTAATCAATCAGAGAAATCAATAATAGGAAGTACTCTTATCCCATGTATAATGGTTTCGGATTTTATTAGTTTGGGGGAGAGAATGAAAACCACTCC
                                                                                                                        960
 Y  Y  Y  V  K  H  T  Q  Y  W  Q  S  M  W  S  A  L  F  P  P  G  S  K  E  T  K  T  E  K  S  G  I  T  D  T  S  Q  I  S  M
ATATTATTATGTAAAGCACACTCAATATTGGCAAAGCATGTGGTCCGCGCTCTTTCCACCCGGCTCTAAAGAGACAAAAACTGAGAAATCAGGTATCACTGACACTTCTCAAATAAGTAT
                                                                                                                       1080
 T  D  G  I  N  V  S  I  G  A  D  F  G  L  R  F  G  N  K  T  F  G  I  K  G  G  F  T  Y  D  T  K  T  Q  I  T  N  T  S  Q
GACTGACGGGATTAATGTTTCAATCGGAGCAGATTTCGGATTAAGGTTTGGAAATAAAACGTTTGGAATTAAGGGGGGGTTCACCTATGATACAAAGACTCAAATAACTAATACCTCCCA
                                                                                                                       1200
 L  L  I  E  T  T  Y  T  R  E  Y  T  N  T  E  N  F  P  V  R  Y  T  G  Y  V  L  A  S  E  F  T  L  H  R  S  D  G  T  Q  V
ATTGTTAATAGAAACAACTTATACTAGAGAATACACAAATACAGAAAATTTTCCTGTTAGATATACAGGCTATGTTTTAGCGTCAGAATTTACTTTACATCGTAGTGATGGAACTCAGGT
                                                                                                                       1320
 N  T  I  P  W  V  A  L  N  D  N  Y  T  T  I  A  R  Y  P  H  F  A  S  E  P  L  L  G  N  T  K  I  I  T  D  D  Q  N  *
TAATACGATCCCATGGGTTGCTTTAAACGATAACTATACAACAATAGCAAGATATCCACATTTTGCAAGTGAACCTTTACTAGGAAATACAAAGATTATTACAGATGATCAAAACTAAAT
                                                                                                                       1440
TTAAACAATATTCTTGAACTAATAGATGTTAAATAGAACAATTAATAACAATTTAAGTACTTTTGGATTATAGTGAAGGGACCTATAAGCATAGCTTTTAGGTCCCTTTTAAGTTGCTTT
                                                                                                                       1560
TTTTCGTTTTTAGAATAGTATAGATAGGCTACACTACACTAAGTTGGACAGATAAAATAAGGGGTTGTAAACTTAGACTATTAAAAAAGGGAGAGTGCTACTATGACACGTCAACATCGA
                                                                                                                       1680
```

REFERENCES. (1) Baumann, P., Unterman, B.M., Baumann, L., Broadwell, A.H., Abbene, S.J. and Bowditch, R.D. (1985). J. Bact. 163: 738-747.
(2) Earp, D.J. and Ellar, D.J. (1987). Nucleic Acids Res. 15: 3619.