5546 Nucleic Acids Research, Vol. 18, No. 18

Nucleotide sequence and deduced amino acid sequence of a *cryIA(c)* gene variant from *Bacillus thuringiensis*

Freddy Dardenne, Jef Seurinck, Bart Lambert and Marnix Peferoen
Plant Genetic Systems, J. Plateaustraat 22, 9000 Gent, Belgium

Submitted August 17, 1990

EMBL accession no. X54159

From a *Bacillus thuringiensis* isolate, highly active against several lepidopteran larvae (e.g. *Heliothis virescens*), a crystal protein gene was isolated and a fragment containing the toxin encoding region was sequenced. The gene fragment was aligned with other *Bacillus thuringiensis* crystal protein genes, and was found to be highly homologous to the *cryIA(c)* gene (1). Compared to the *cryIA(c)* gene, there are 10 point mutations resulting in 3 amino acid changes. In addition, there is a deletion of a triplet, resulting in the removal of an asparagine.

REFERENCES

1. Adang,M.J., Staver,M.J., Rocheleau,T.A., Leighton,J., Barker,R.F. and Thompson,D.V. (1985) *Gene* **36**, 289–300.
2. Höfte,H. and Whiteley,H.R. (1989) *Microbiol. Rev.* **53**, 242–255.
3. Wong,H.C., Schnepf,H.E. and Whiteley,H. (1983) *J. Biol. Chem.* **258**, 1960–1967.

```
   1 TAAGTATTTCTACATTACCACAAATTCTCAATTTGTATATGTAAAATAGGAAAAGTGGATTTTATATATAAGTAT
  76 AAAAAGTAATAAGACTTTAAAATAAGTTAACGGAATACAAACCCTTAATGCATTGGTTAAACATTGTAAAGTCTA
 151 AAGCATGGATAATGGGCGAGAAGTAAGTAGATTGTTAACACCCTGGGTCAAAAATTGATATTTAGTAAAATTAGT
 226 TGCACTTTGTGCATTTTTTCATAAGATGAGTCATATGTTTTAAATTGTAGTAATGAAAAACAGTATTATATCATA
 301 ATGAATTGGTATCTTAATAAAAGAGATGGAGGTAACTTATGGATAACAATCCGAACATCAATGAATGCATTCCTT
                                •••••     M  D  N  N  P  N  I  N  E  C  I  P  Y
 376 ATAATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGAAAGAATAGAAACTGGTTACACCCCAATCGATA
      N  C  L  S  N  P  E  V  E  V  L  G  G  E  R  I  E  T  G  Y  T  P  I  D  I
 451 TTTCCTTGTCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTGCTGGATTTGTGTTAGGACTAGTTGATA
      S  L  S  L  T  Q  F  L  L  S  E  F  V  P  G  A  G  F  V  L  G  L  V  D  I
 526 TAATATGGGGAATTTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATTGAACAGTTAATTAACCAAAGAA
      I  W  G  I  F  G  P  S  Q  W  D  F  L  V  Q  I  E  Q  L  A  I  N  Q  R  I
 601 TAGAAGAATTCGCTAGGAACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTATCAAATTTACGCAGAAT
      E  E  F  A  R  N  Q  A  I  S  R  L  E  G  L  S  N  L  Y  Q  I  Y  A  E  S
 676 CTTTTAGAGAGTGGGAAGCAGATCCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAATTCAATGACATGA
      F  R  E  W  E  A  D  P  T  N  P  A  L  R  E  E  M  R  I  Q  F  N  D  M  N
 751 ACAGTGCCCTTACAACCGCTATTCCTCTTTTGGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTATATGTTC
      S  A  L  T  T  A  I  P  L  L  A  V  Q  N  Y  Q  V  P  L  L  S  V  Y  V  Q
 826 AAGCTGCAAATTTACATTTATCAGTTTTGAGAGATGTTTCAGTGTTTGGACAAAGGTGGGGATTTGATGCCGCGA
      A  A  N  L  H  L  S  V  L  R  D  V  S  V  F  G  Q  R  W  G  F  D  A  A  T
 901 CTATCAATAGTCGTTATAATGATTTAACTAGGCTTATTGGCAACTATACAGATTATGCTGTACGCTGGTACAATA
      I  N  S  R  Y  N  D  L  T  R  L  I  G  N  Y  T  D  Y  A  V  R  W  Y  N  T
 976 CGGGATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGTAAGGTATAATCAATTTAGAAGAGAATTAACAC
      G  L  E  R  V  W  G  P  D  S  R  D  W  V  R  Y  N  Q  F  R  R  E  L  T  L
1051 TAACTGTATTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAAGATATCCAATTCGAACAGTTTCCCAAT
      T  V  L  D  I  V  A  L  F  P  N  Y  D  S  R  R  Y  P  I  R  T  V  S  Q  L
1126 TAACAAGAGAAATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTTCGAGGCTCGGCTCAGGGCATAG
      T  R  E  I  Y  T  N  P  V  L  E  N  F  D  G  S  F  R  G  S  A  Q  G  I  E
1201 AAAGAAGTATTAGGAGTCCACATTTGATGGATATACTTAACAGTATAACCATCTATACGGATGCTCATAGGGGTT
      R  S  I  R  S  P  H  L  M  D  I  L  N  S  I  T  I  Y  T  D  A  H  R  G  Y
1276 ATTATTATTGGTCAGGGCATCAAATAATGGCTTCTCCTGTCGGTTTTTCGGGGCCAGAATTCACGTTTCCGCTAT
      Y  Y  W  S  G  H  Q  I  M  A  S  P  V  G  F  S  G  P  E  F  T  F  P  L  Y
1351 ATGGAACCATGGGAAATGCAGCTCCACAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGAACATTAT
      G  T  M  G  N  A  A  P  Q  Q  R  I  V  A  Q  L  G  Q  G  V  Y  R  T  L  S
1426 CCTCTACTTTTTATAGAAGACCTTTTAATATAGGGATAAATAATCAACAACTATCTGTTCTTGACGGGACAGAAT
      S  T  F  Y  R  R  P  F  N  I  G  I  N  N  Q  Q  L  S  V  L  D  G  T  E  F
1501 TTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTATACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAA
      A  Y  G  T  S  S  N  L  P  S  A  V  Y  R  K  S  G  T  V  D  S  L  D  E  I
1576 TACCACCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAGTCATCGATTAAGCCATGTTTCAATGTTTCGTT
      P  P  Q  N  N  N  V  P  P  R  Q  G  F  S  H  R  L  S  H  V  S  M  F  R  S
1651 CAGGCTCTAGTAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCTGAATTTAATA
      G  S  S<>S  S  V  S  I  I  R  A  P  M  F  S  W  I  H  R  S  A  E  F  N  N
1726 ATATAATTGCATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAACTTTCTTTTTAATGGTTCTGTAATTT
      I  I  A  S  D  S  I  T  Q  I  P  A  V  K  G  N  F  L  F  N  G  S  V  I  S
1801 CAGGACCAGGATTTACTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAATAACATTCAGAATAGAGGGTATA
      G  P  G  F  T  G  G  D  L  V  R  L  N  S  S  G  N  N  I  Q  N  R  G  Y  I
1876 TTGAAGTTCCAATTCACTTCCCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATGCTTCTGTAACCCCGA
      E  V  P  I  H  F  P  S  T  S  T  R  Y  R  V  R  V  R  Y  A  S  V  T  P  I
1951 TTCACCTCAACGTTAATTGGGGTAATTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACGTCATTAGATA
      H  L  N  V  N  W  G  N  S  S  I  F  S  N  T  V  P  A  T  A  T  S  L  D  N
2026 ATCTACAATCAAGTGATTTTGGTTATTTTGAAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGTAGGTG
      L  Q  S  S  D  F  G  Y  F  E  S  A  N  A  F  T  S  S  L  G  N  I  V  G  V
2101 TTAGAAATTTTAGTGGGACTGCAGGAGTGATAATAGACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG
      R  N  F  S  G  T  A  G  V  I  I  D  R  F  E  F  I  P  V  T* A  T  L  E  A
2176 CTGAATATAATCTGGAA
      E  Y  N  L  E
```

Figure 1. Nucleotide sequence and deduced amino acid sequence of the toxin encoding DNA fragment of a *cryIA(c)* gene variant. Point mutations and amino acid changes are underlined, and the amino acid deletion is marked (< >). The amino acid delineating the C-terminus of the putative minimum toxic fragment (2) is indicated (*). Putative promoter sequences (3) are shadowed and the putative ribosome binding site is marked (°).