The complete nucleotide sequence of tetanus toxin

Neil F. Fairweather and Val A. Lyness

Bacterial Genetics Group, Department of Molecular Biology, Wellcome Biotechnology Ltd.,
Beckenham, Kent BR3 3BS, UK

Submitted 5 September 1986

We have previously isolated and sequenced a portion of the structural gene for tetanus toxin from *C. tetani* CN3911 (1). Overlapping clones encoding the entire gene have now been isolated and the full sequence is reported below. Our designation of the start of translation is based on amino terminal analysis (2) and partial amino acid sequence of the toxin (3). The methionine residue before the proline is presumably cleaved after synthesis. The DNA sequence contains one open reading frame encoding a protein of molecular weight 150 491, in agreement with other measurements of the size of the toxin (4).

```

      .           .           .           .           .           .           .           .           .           .
      32           62           92
GCTTCATATGAGTCGTCAAGCTGTATATAAAATAAGBTTTATGACATTAAGAAAAATAGAACCTATAGTAATAAATTAATTAATATATA

      .           .           .           .           .           .           .           .           .           .
      122          152          182
GTTTTTATAATTTAATTATGAATAATATTCTTAAGATAAAAAGTAATTTTAAAAATTTAAATTTTCAGTTTACAAAAATAACCTGAT

      .           .           .           .           .           .           .           .           .           .
      212          242          272
TATGTTATATGTAATTGTAAAAACATATAAAAAATCAGAAAAATTTAGBAGGTATATTATTAAATGGATTAAATAATAATTTTTTAATTT

      .           .           .           .           .           .           .           .           .           .
      302          332          362
ACTTTTGAATTAATAATATTAAATGTTTATTTAATTAGGAGATGATACGTATGCCAATAACCATAAATAATTTTAGATATAGTGATCCT
MetProIleThrIleAsnAsnPheArgTyrSerAspPro

      .           .           .           .           .           .           .           .           .           .
      392          422          452
GTTAATAATGATACAATTATTATGATGGAGCCACCATCTGTAAAGGCTCTAGATATCTATTATAAGGCTTTCAAAATAACAGATCGTATT
ValAsnAsnAspThrIleIleMetMetGluProProTyrCysLysGlyLeuAspIleTyrTyrLysAlaPheLysIleThrAspArgIle

      .           .           .           .           .           .           .           .           .           .
      482          512          542
TGGATAGTGCCGGAAGGTATGAATTTGGGACAAAACCTGAAGATTTTAACCCACCATCTTATTAAATAGAGGTGCATCTGAGTATTAC
TrpIleValProGluArgTyrGluPheGlyThrLysProGluAspPheAsnProProSerSerLeuIleGluGlyAlaSerGluTyrTyr

      .           .           .           .           .           .           .           .           .           .
      572          602          632
GATCCAAATTATTTAAGBACTGATTTCTGATAAGATAGATTTTACAAACCATGTTAAAGTCTTTAACAAGAAATAAAAACAATGTAGCA
AspProAsnTyrLeuArgThrAspSerAspLysAspArgPheLeuGlnThrMetValLysLeuPheAsnArgIleLysAsnAsnValAla

      .           .           .           .           .           .           .           .           .           .
      662          692          722
GGTGAAGCCTTATTAGATAAGATAATAATGCCATACCTTACCTTGGAAATTCATATTCCTTACTAGACAAGTTTGATACAAACTCTAAT
GlyGluAlaLeuLeuAspLysIleIleAsnAlaIleProTyrLeuGlyAsnSerTyrSerLeuLeuAspLysPheAspThrAsnSerAsn

      .           .           .           .           .           .           .           .           .           .
      752          782          812
TCAGTATCTTTTAATTTATAGACAAGACCCAGTGGAGCAACTACAAATCAGCAATGCTGACAAATTTAATAATATTGGACCTGGG
SerValSerPheAsnLeuLeuGluGlnAspProSerGlyAlaThrThrLysSerAlaMetLeuThrAsnLeuIleIlePheGlyProGly

```


2372 2402 2432
 TTAGAATATATTCCAGAAATTACTTTACCAAGTAATTGCAGCTTTATCTATAGCAGAAAGTAGCACACAAAAGAAAAGATAATAAAAAACA
 Leu61uTyrIlePro61uIleThrLeuProValIleAlaAlaLeuSerIleAla61uSerSerThr61nLys61uLysIleIleLysThr

2462 2492 2522
 ATAGATAACTTTTTAGAAAAAGATATGAAAAATGGATTGAAGTATATAAAGTAGTAAAGCAGAAATGGTTAGGCACAGTTAATACGCAAA
 IleAspAsnPheLeu61uLysArgTyr61uLysTrpIle61uValTyrLysLeuValLysAlaLysTrpLeu61uThrValAsnThr61n

2552 2582 2612
 TTCCAAAAAGAAAGTTATCAAATGTATAGATCTTTAGAAATATCAAGTAGATGCAATAAAAAAATAATAGACTATGAATATAAAATATAT
 Phe61nLysArgSerTyr61nMetTyrArgSerLeu61uTyr61nValAspAlaIleLysLysIleIleAspTyr61uTyrLysIleTyr

2642 2672 2702
 TCAGGACCTGATAAGBAACAAATGCCGACGAAATTAATAATCTGAAAAACAACTTGAAGAAAAAGGCTAATAAGCAATGATAAACATA
 Ser61yProAspLys61u61nIleAlaAsp61uIleAsnAsnLeuLysAsnLysLeu61u61uLysAlaAsnLysAlaMetIleAsnIle

2732 2762 2792
 AATATATTTATGAGGAAAGTTCTAGATCATTTTGTATTAATCAAATGATTAAACGAGCTAAAAAGCAGTTATTAGAGTTTGATACCTCAA
 AsnIlePheMetArg61uSerSerArgSerPheLeuValAsn61nMetIleAsn61uAlaLysLys61nLeuLeu61uPheAspThr61n

2822 2852 2882
 AGCAAAATATTTTAATGCAATATATAAGCAAAATCTAAATTTATAGGTATACTGAACTAAAAAAATAGAAATCAAAAAATAACAAA
 SerLysAsnIleLeuMet61nTyrIleLysAlaAsnSerLysPheIle61yIleThr61uLeuLysLysLeu61uSerLysIleAsnLys

2912 2942 2972
 GTTTTTCAACACCAATTCATTTCTTATCTAAAAATCTGGATTGTTGGGTTGATATGAGGAAGATAGATGTTATATTAAGAAAG
 ValPheSerThrProIleProPheSerTyrSerLysAsnLeuAspCysTrpValAspAsn61u61uAspIleAspValIleLeuLysLys

3002 3032 3062
 AGTCAATTTTAAATTTAGATATTAATAATGATATTATATCAGATATATCTGGGTTTAATTCATCTGTAATAACATATCCAGATGCTCAA
 SerThrIleLeuAsnLeuAspIleAsnAsnAspIleIleSerAspIleSer61yPheAsnSerSerValIleThrTyrProAspAla61n

3092 3122 3152
 TTGGTCCCGGAATAAATGGCAAGCAATACATTTAGTAAACAAATGAATCTTCTGAAGTTATAGTGCATAAAGCTATGGATATTGAATAT
 LeuValPro61yIleAsn61yLysAlaIleHisLeuValAsnAsn61uSerSer61uValIleValHisLysAlaMetAspIle61uTyr

3182 3212 3242
 AATGATATGTTTAATAATTTTACCGTTAGCTTTTGGTTGAGGGTTCTTAAAGTATCTGCTAGTCATTAGAACAAATATGGCACAATGAG
 AsnAspMetPheAsnAsnPheThrValSerPheTrpLeuArgValProLysValSerAlaSerHisLeu61u61nTyr61yThrAsn61u

3272 3302 3332
 TATTCAATAATTAGCTCTATGAAAAACATAGTCTATCAATAGGATCTGGTTGGAGTGTATCACCTTAAGGTAATAACTTAATATGGACT
 TyrSerIleIleSerSerMetLysLysHisSerLeuSerIle61ySer61yTrpSerValSerLeuLys61yAsnAsnLeuIleTrpThr

3362 3392 3422
 TTAAAGATTCCGCGGGAGAAGTTAGACAAAATAACTTTTGGGATTTACCTGATAAATTTAATGCTTATTAGCAAAATAAATGGGTTTTT
 LeuLysAspSerAla61y61uValArg61nIleThrPheArgAspLeuProAspLysPheAsnAlaTyrLeuAlaAsnLysTrpValPhe

3452 3482 3512
 ATAATACTACTAATGATAGATTATCTCTGCTAATTTGTATATAAATGGAGTACTTATGGGAAGTGCAGAAATTAAGTGGTTATGGAGCT
 IleThrIleThrAsnAspArgLeuSerSerAlaAsnLeuTyrIleAsn61yValLeuMet61ySerAla61uIleThr61yLeu61yAla

3542 3572 3602
 ATTAGAGAGGATAATAATAACATTAAAGTATAGATAGTGAATAATAATAATCAATACGTTTCTATTGATAAATTTAGGATATTTTGC
 IleArg61uAspAsnAsnIleThrLeuLysLeuAspArgCysAsnAsnAsnAsn61nTyrValSerIleAspLysPheArgIlePheCys

3632 3662 3692
 AAAGCATTAAATCCAAAAGAGATTGAAAAATTATACACAAGTTATTTATCTATAACCTTTTAAAGABACTTCTGGGGAACCCCTTTACGA
 LysAlaLeuAsnProLys61uIle61uLysLeuTyrThrSerTyrLeuSerIleThrPheLeuArgAspPheTrp61yAsnProLeuArg

3722 3752 3782
 TATGATACAGAAATATTATTAAACCAAGTACCTTCTAGTTCTAAGATGTTCAATTGAAAAATATAACAGATTATATGATTTTACCAAT
 TyrAspThr61uTyrTyrLeuIleProValAlaSerSerSerLysAspVal61nLeuLysAsnIleThrAspTyrMetTyrLeuThrAsn

3812 3842 3872
 GCBCCATCGTACTAACGGAATTTGAATATATATATAGAAGGTTATATAATGGACTGAAAAATTTATTATAAAAAAGATATACACCTAAT
 AlaProSerTyrThrAsn61yLysLeuAsnIleTyrTyrArgArgLeuTyrAsn61yLeuLysPheIleIleLysArgTyrThrProAsn

```

      3902      3932      3962
AATGAAATAGATTCTTTTGTAAATCAGGTGATTTTATTAAATTATATGTATCATATAACAATAATGAGCATTGTAGGTTATCCGAAA
AsnGluIleAspSerPheValLysSerGlyAspPheIleLysLeuTyrValSerTyrAsnAsnAsnGluHisIleValGlyTyrProLys

      3992      4022      4052
GATGAAATGCTTTAATAATCTTGTATGAAATCTAAGATAGGTTATAATGCCCCAGGTATCCCTCTTTATAAAAAATGGAGCAGTA
AspGlyAsnAlaPheAsnAsnLeuAspArgIleLeuArgValGlyTyrAsnAlaProGlyIleProLeuTyrLysLysMetGluAlaVal

      4082      4112      4142
AAATTGCGTATTTAAAAACCTATTCTGTACAACCTAAATTATATGATGATAAAATGTCATCTTTAGSACTAGTAGGTACCCATAATGGT
LysLeuArgAspLeuLysThrTyrSerValGlnLeuLysLeuTyrAspAspLysAsnAlaSerLeuGlyLeuValGlyThrHisAsnGly

      4172      4202      4232
CAATAGGCAACGATCCAAATAGGGATATATTAATTGCAAGCAACTGGTACTTTAATCATTAAAGATAAAATTTAGGATGTGATTGG
GlnIleGlyAsnAspProAsnArgAspIleLeuIleAlaSerAsnTrpTyrPheAsnHisLeuLysAspLysIleLeuGlyCysAspTrp

      4262      4292      4322
TACTTTGTACCTACAGATGAAGGATGGACAATGATTAAACAGATTGATATGTTTCATGATTACTCTATATAAAAAATTAAATAATATAAC
TyrPheValProThrAspGluGlyTrpThrAsnAsp***

AATCTAGCTATATTATT

```

ACKNOWLEDGEMENTS

We thank Hugh Spence for synthesis of oligonucleotides and H. Niemann for discussion of results.

REFERENCES

1. Fairweather, N.F., Lyness, V.A., Pickard, D.J., Allen, G. and Thomson, R.O. (1986) *J.Bacteriol.* **165**, 21-27.
2. Neubauer, V and Helting, T.B. (1979) *Biochem.Biophys.Res.Comm.* **86**, 635-642.
3. Robinson, J.P. and Nash, J.H. (1982) *Molec.Cell.Biochem.* **48**, 33-44.
4. Bizzini, B. (1979) *Microbiol.Rev.* **43**, 224-240.