A unique sequence of the *Bordetella pertussis* toxin operon

S.M.Loosmore*, J.D.Cunningham, W.R.Bradley, F.L.Yao, G.A.Dekaban and M.H.Klein

Molecular Genetics Group, Connaught Centre for Biotechnology Research, 1755 Steeles Avenue W., Willowdale, Ontario M2R 3T4, Canada

Submitted September 14, 1989

EMBL accession no. X16347

The entire pertussis toxin operon of the Connaught vaccine production strain 10536 has been cloned and sequenced. The portion of the operon coding for the B oligomer subunits is identical to published sequences for four different strains (1) (2), while four unique bases have been identified in the promoter and the S1 structural gene (figure 1). The presence of three different bases in the S1 gene results in two unique amino acids, GLU34 and VAL198. It is of interest to note that three of the four nucleotide differences (first, second, and fourth) are also found in *B.parapertussis*, *B.bronchiseptica*, and *B.pertussis* 18323 (1).

```

10536 GAAATCGTCGCGCTCGCCCTGGTTCGCGGTCAATGGCCCCAAGGAAACGACCCCAAGATAATCGTCTGCTCAACCGCCACATCAACGAGGCGCTGCAGT 100
10536 CCAAGGCGGTCGTCGAGGCGCTTTGCCGCCAAGGCGCCACGCGGTTCATGCCACGCGGATCAGACCGCGCGGTTTCATGCGAGACGAGATCCAGCGCTG 200
10536 GGC CGCGTGTGTCGCGGAAACCGGCGCAAGCTGAAGTAGCAGCGCCCTCCAAACGCGCCATCCCGGTCCGCGCGGCAACCATCCCGCATACGCTTTGG 300
10536 CAACCGCCACGCGTATGCGTGCAGATTGTCGTCAAAAACCTCGATTCTTCGTACATCCCGCTACTGCAATCCAACGCGCATGAACGCTCCTTCGG 400
BP165 C
3779 C
10536 CGCAAAGTCGCGCGATGGTACCGGTCAACGTCGCGGACCGTGTGACCCCCCTGCATGGTGTGATCCGTAATAAGGCCACCATCAAAACGACAGGGGAA 500
10536 GACGGGATCGGTTGCACTCGGGCAATTGCCAAAACGCAAGAACAGGCTGGCTGACGTGGCTGGCGATTCTTGCCGTACGCGCGCGGTGACTTCGCGG 600
10536 CATGGGCGGACGATCCTCCCGCCACCGTATACCGCTATGACTCCCGCGCGCGGAGGACGTTTTCCAGAACGGATTTCAGCGGTGGGGAAACACGACAA 700
10536 TGTGCTCGAACAATCTGACCGGAGCTTCTGCCAGGTGCGCAGCAGCAACAGCGCTTTGCTCTCCACCGACGAGCGCGGCGCTATACCGAGGTCTATCTC 800
BP165 C
3779 C
10536 GAACATCGCATGCAGGAACGCGTCGAGGCCGAACGCGCGGACGGGACCGGCACTTCATCGGCTACATCTACGAAGTCGCGCGGACAACAATTTCT 900
10536 ACGGCGCGCGCAGCTCGTACTTCGAATACGTGCACACTTATGGCGACAATGCCGGCGGTATCCTCGCGCGGCGGCTGGCCACCTACGAGCGGAATATCT 1000
10536 GGCACACCGGCGCATTTCCGCCGAAAAACATCGCAGGTTAACGCGGCTATATCAACCGGCATCACCGCGGAGCAGCAGCAACGAGGATTCCAAACGCT 1100
10536 CGTACGTGAGCAGCAGACTCGCGCCAAATCCCAACCCCTACACATCGCGAAGGTCCGTAGCGTGCATGTCGCGCACATTGGTGGCGCATGGCGCGGTGG 1200
BP165 A
3779 ** A
10536 TGGCGCGTTGCATGGCGCGGAGGCGGAAAGCTCCGAGGCCATGGCAGCTGGTCCGAACGCGCGGCGAGGCGATGGTTCTCGTGTACTACGAAAGCAT 1300
BP165 A
3779 A*
10536 CGCGTATTCGTTCTAG * indicates a deletion _ is the leader start of S1

```

*To whom correspondence should be addressed

REFERENCES

- (1) Arico, B. et al (1987) *Molec. Micro.* 1:301
- (2) Lochter, C. and Keith, J.M. (1986) *Science* 232:1258