Volume 15 Number 9 1987 Nucleic Acids Research

Sequence of an active fragment of *B. polymyxa* beta amylase

Craig Rhodes*, Jeanne Strasser+ and Felix Friedberg[1]

Genex Corporation, Gaithersburg, MD 20877 and [1]Department of Biochemistry, Howard University, Washington, DC 20059, USA

Submitted February 24, 1987 Accession no. Y00150

We described the isolation of the gene coding for beta-amylase which causes synthesis of an enzymatically active protein from B. polymyxa (ATCC 8523) by molecular cloning in B. subtilis (1). Here we report the sequence of the insert which exhibits no translational stop codon because the insert is not long enough. However, this insert induces production of an enzymatically active protein (1). A Shine-Dalgarno sequence (underlined) occurs immediately upstream from a signal sequence (italics) which contains 38 amino acids. A large direct repeat structure within the gene extends from base 1462 to base 2103 (underlined). At 2427 the insert terminates and the sequence continues into the plasmid itself.

1. Friedberg, F., et al., (1986) J. Bacteriol. 165, 819-824.

```
AAATTCTTCTTTTTATGCAACCGCTTGCTCTATACTATTTTTGTAGACGTTTACAAAATGTTGTCGAATTAGACTATGCTTTTGGGGAGG   90
ATGATCGGTTTGACCTTGTATCGAAGTCTATGGAAAAAAGGGTGTATGCTTCTACTAAGCCTTGTGCTGTCTGTGACCGCCTTCATTGGC  180
MetIleGlyLeuThrLeuTyrArgSerLeuTrpLysLysGlyCysMetLeuLeuLeuSerLeuValLeuSerLeuThrAlaPheIleGly
TCACCCTCTAATACGGCGAGCGCCGCGGTTGCAGATGATTTTCAGGCTTCCGTTATGGGACCACTGGCGAAAATCAACGACTGGGGTTCC  270
SerProSerAsnThrAlaSerAlaAlaValAlaAspAspPheGlnAlaSerValMetGlyProLeuAlaLysIleAsnAspTrpGlySer
TTTAAAAAGCAACTGCAAACGTTGAAAAGCAATGGCGTGTATGCCATCACTACCGACGTCTGGTGGGGCTATGTGGAGAGCGCAGGAGAT  360
PheLysLysGlnLeuGlnThrLeuLysSerAsnGlyValTyrAlaIleThrThrAspValTrpTrpGlyTyrValGluSerAlaGlyAsp
AACCAGTTCGATTGGAGCTACTACAAAACCTATGCAGATGCCGTGAAAGAAGCGGGACTGAAGTGGGTTCCTATTATCTCTACGCATAAG  450
AsnGlnPheAspTrpSerTyrTyrLysThrTyrAlaAspAlaValLysGluAlaGlyLeuLysTrpValProIleIleSerThrHisLys
TGTGGAGGAAATGTGGGAGATGATTGCAACATCCCACTGCCAAGCTGGCTGTCAAGCCAAGGGTAGTGCAGATGAAATGGCAGTTCAAGGAC  540
CysGlyGlyAsnValGlyAspAspCysAsnIleProLeuProSerTrpLeuSerSerLysGlySerAlaAspGluMetGlnPheLysAsp
GAAAGCGGCTATGCCAATAACGAAGCCCTTTCCCCATTGTGGAGCGGGACTGGCAAACAATATGACGAGTTGTATGCCTCATTCGCTCAA  630
GluSerGlyTyrAlaAsnAsnGluAlaLeuSerProLeuTrpSerGlyThrGlyLysGlnTyrAspGluLeuTyrAlaSerPheAlaGln
AACTTTGCCGGGTATAAAAGTATCATTCCCAAAATTTATTTAAGCGGCGGTCCTTCCGGGGAATTGCGCTATCCGTCCTACTACCCGGCA  720
AsnPheAlaGlyTyrLysSerIleIleProLysIleTyrLeuSerGlyGlyProSerGlyGluLeuArgTyrProSerTyrTyrProAla
GCGGGCTGGAGTTACCCGGGACGAGGCAAGTTCCAAGCCTATACTGAAACAGCCAAAAAGTCATTCCGCACAGCTATGAATGACAAGTAC  810
AlaGlyTrpSerTyrProGlyArgGlyLysPheGlnAlaTyrThrGluThrAlaLysLysSerPheArgThrAlaMetAsnAspLysTyr
GGATCATTGGACAAGATTAATGCTGCTTGGGGCACCAAGCTTACCAGCCTTAGCCAAATCAATCCACCGACCGACGGGGACGGATTTTAC  900
GlySerLeuAspLysIleAsnAlaAlaTrpGlyThrLysLeuThrSerLeuSerGlnIleAsnProProThrAspGlyAspGlyPheTyr
ACGAACGGCGGGTATAATTCCGCGTATGGAAAGGATTTTCTGTCGTGGTACCAAAGTGTGCTTGAGAAGCATCTAGGTGTCATCGGAGCA  990
ThrAsnGlyGlyTyrAsnSerAlaTyrGlyLysAspPheLeuSerTrpTyrGlnSerValLeuGluLysHisLeuGlyValIleGlyAla
GCTGCCCACAAGAACTTCGATTCTGTGTTTGGTGTCCGGATTGGGGCCAAAATTTCCGGTCTGCACTGGCAGATGAACAATCCTGCTATG 1080
AlaAlaHisLysAsnPheAspSerValPheGlyValArgIleGlyAlaLysIleSerGlyLeuHisTrpGlnMetAsnAsnProAlaMet
CCGCACAGCACAGAACAAGCAGGTGGATATTATGATTACAACCGTCTCATTCAGAAATTCAAGGATGCCGATCTAGACTTGACATTTACT 1170
ProHisSerThrGluGlnAlaGlyGlyTyrTyrAspTyrAsnArgLeuIleGlnLysPheLysAspAlaAspLeuAspLeuThrPheThr
TGCCTGGAGATGAGTGACAGTGGCACAGCACCGAATTATTCGTTGCCGTCCACCCTGGTTGATACCGTCTCCTCCATCGCGAATGCCAAG 1260
CysLeuGluMetSerAspSerGlyThrAlaProAsnTyrSerLeuProSerThrLeuValAspThrValSerSerIleAlaAsnAlaLys
GGTGTTCGCCTGAATGGTGAAAATGCACTCCCGACAGGCGGCAGCCGGCTTCCAAAAAATCGAGGAGAAAATCACGAAGTTCGGTTATCAT 1350
GlyValArgLeuAsnGlyGluAsnAlaLeuProThrGlyGlySerGlyPheGlnLysIleGluGluLysIleThrLysPheGlyTyrHis
GGGTTTACGTTATTGCGCATCAGCAATCTCGTCAACAATGACGGCTCACCAACCGGTGAGTTGAGTGGATTCAAACAATATATCATCAGC 1440
GlyPheThrLeuLeuArgIleSerAsnLeuValAsnAsnAspGlySerProThrGlyGluLeuSerGlyPheLysGlnTyrIleIleSer
AAAGCCAAGCCTGACAATAATGGCGGTACAGGCAACAAAGTAACGATATACTACAAAAAAAGGCTTCAACTCGCCTTACATTCACTACCGC 1530
LysAlaLysProAspAsnAsnGlyGlyThrGlyAsnLysValThrIleTyrTyrLysLysGlyPheAsnSerProTyrIleHisTyrArg
CCGGCTGGCGGAAGCTGGACTGCTGCTGCTGGTGTGAAAATGCAGGCTGCCGAAATTAGCGGCTATGCTAAAATTACCGTCGATATCGGT 1620
ProAlaGlyGlySerTrpThrAlaAlaProGlyValLysMetGlnAlaAlaGluIleSerGlyTyrAlaLysIleThrValAspIleGly
TCTGCCTCGCAATTGGAAGCCGCCTTAACGACGGCAATAACAACTGGGACAGCAACAACACTAAAAACTACCTTTTCAGCACTGGAACT 1710
SerAlaSerGlnLeuGluAlaAlaPheAsnAspGlyAsnAsnAsnTrpAspSerAsnAsnThrLysAsnTyrLeuPheSerThrGlyThr
TCTACCTACACGCCCGGTAACAGCGGCAATGCAGGAACGATTACATCAGGCGCTCCAGCGGGCACAAATCCCGGGGATGGCGGCGGTACA 1800
SerThrTyrThrProGlyAsnSerGlyAsnAlaGlyThrIleThrSerGlyAlaProAlaGlyThrAsnProGlyAspGlyGlyGlyThr
ACTAACAAAGTAACGGTATACTACAAAAAAGGCTTCAACTCACCTTACATTCACTACCGTCCGGCTGGCGGTAGCTGGACTGCTGCACCT 1890
ThrAsnLysValThrValTyrTyrLysLysGlyPheAsnSerProTyrIleHisTyrArgProAlaGlyGlySerTrpThrAlaAlaPro
GGTGTGAAAATGCAGGATGCCGAAATTAGCGGCTATGCTAAAATTACCGTCGATATCGGTTCTGCCTCCCAATTGGAAGCCGCCTTTAAC 1980
GlyValLysMetGlnAspAlaGluIleSerGlyTyrAlaLysIleThrValAspIleGlySerAlaSerGlnLeuGluAlaAlaPheAsn
GACGGCAATAACAACTGGGACAGCAACAACACTAAAAACTACCTTTTCAGCACTGGAACTTCTACCTACGCAACCAACCTCAACGGAGCC 2070
AspGlyAsnAsnAsnTrpAspSerAsnAsnThrLysAsnTyrLeuPheSerThrGlyThrSerThrTyrThrProGlySerAsnGlyAla
GCAGGAACTATCCGAACAGGGACACCTTCAGGGAGTGTCCTCAGCGTAGTGACTTCCACCTACGCAACCAACCTCAACGGAGTTGAAGTAACAGCT 2160
AlaGlyThrIleArgThrGlyThrProSerGlySerValLeuSerValValThrSerThrTyrAlaThrAsnLeuAsnGluValAlaGly
CCTATCGAAACAGAAAAACTTTCAGGTGTATCCTTAAACGTAAGCACTTCTACCTACGCACCCAACTCTAACCGGAGTTGAAGTAACAGCT 2250
ProIleGlnThrGluLysLeuSerGlyValSerLeuAsnValSerThrSerThrTyrAlaProAsnSerAsnGlyValGluValThrAla
CAAACGGAGGCGGCTTCGGGCGTCTTCAGCAGCTATGGATCTGGGTACTCTCTCCAAGGCCACCTCCCTCAGCACAGACTGGAGCAAACAG 2340
GlnThrGluAlaProSerGlyValPheSerProMetAspLeuGlyThrLeuSerLysProThrSerLeuSerThrAspTrpSerLysGln
AGCATTTATTTTATCATGACAGATCGTTTCAACAACGGTGATGCATCCAATGATAACTACGGTGGCTTCAATTCCAACAACAGTGATGCA 2430
SerIleTyrPheIleMetThrAspArgPheAsnAsnGlyAspProSerAsnAspAsnTyrGlyGlyPheAsnSerAsnAsnSerAspPro
GTTGCTCAAAAAAATCTCGGTCAGATGTTACTAGCAACTCATTTACAAGAACAGCATCTTTCCTCGTTTTCTTGTACCTGTTTTTTGTG 2520
ValAlaGlnLysAsnLeuGlyGlnMetLeuLeuAlaThrHisLeuGlnGluGlnHisLeuSerSerPhePheLeuTyrLeuPhePheVal
ATTCAATAA
IleGlnTer                                                                                   2529
```

Present addresses: *Laboratory of Development Biology and Anomalies, National Institute of Dental Research, Bethesda, MD 20205 and +Chem Gen Corporation, Rockville, MD 20850, USA

 © IRL Press Limited, Oxford, England.