Volume 16 Number 14 1988 Nucleic Acids Research

Nucleotide sequence of a *Bacillus circulans* xylanase gene

Robert C.A.Yang*, C. Roger MacKenzie and Saran A.Narang

Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario, K1A 0R6, Canada
Submitted June 14, 1988 Accession no.X07723

A 4.3 kb EcoRI fragment of Bacillus circulans DNA was inserted into pUC19 and cloned in Escherichia coli. The insert of the resulting chimeric plasmid, pBCX-549R3, was subcloned and one subclone, pBCX-600, was subjected to dideoxy-sequencing from both strands (1). The sequence has an open-reading frame for a protein of 213 amino acid residues. Partial N-terminal peptide sequence analysis revealed that the protein consists of a 28-residue signal peptide and a 185-residue mature enzyme. A comparison of its primary structure with that of other xylanases showed more than 50% homology (2,3). The 5'-flanking sequence contains a modified promotor sequence (boxed) and two pairs of tandem repeats (underlined) resembling some animal virus enhancer elements (4). The 3'-flanking sequence contains a pair of inverted repeats (arrows) which have been proposed as transcription termination signals (5).

```
GAATTCGGTGATGATGGATTCGATGCGCTTTAGTTCTGAGGTGATCACGTTGAAGTAAAGGGCGTAGTCTCCCTCGACACTCCCTTTCAGAAGCTGAATG     100
AAGCCTTTAAGAGCAGTCATCGGATTTCTGATTTCATGAGCAATACCGGCTGCGAGTTCGCCGACAACATGAAGGGTATCTGATTTCGCAGGCGCTCCTC     200
CAATTCTTTTCGCTCTGTTACGTCTTTAAAGATAGCCAAGTTCATATTTTGAATAATATTCCGTTTAAATGAGAATTCGTGGTATTATACTGAAGGGGAC     300
GATCAAAAGCTTTGGCGTTAGTAATTAAAAATGTTTTAAATGTATACGAGTGCTGCCTCAAAGTTGGAAAAAATATTATAGGAGGTAACAT ATG TTT      397
AAG TTT AAA AAG AAT TTC TTA GTT GGA TTA TCG GCA GCT TTA ATG AGT ATT AGC TTG TTT TCG GCA ACC GCC TCT          472
GCA GCT AGC ACA GAC TAC TGG CAA AAT TGG ACT GAT GGG GGC GGT ATA GTA AAC GCT GTC AAT GGG TCT GGC GGG          547
AAT TAC AGT GTT AAT TGG TCT AAT ACC GGA AAT TTT GTT GTT GGT AAA GGT TGG ACT ACA GGT TCG CCA TTT AGG          622
ACG ATA AAC TAT AAT GCC GGA GTT TGG GCG CCG AAT GGC AAT GGA TAT TTA ACT TTA TAT GGT TGG ACG AGA TCA          697
CCT CTC ATA GAA TAT TAT GTA GTG GAT TCA TGG GGT ACT TAT AGA CCT ACT GGA ACG TAT AAA GGT ACT GTA AAA          772
AGT GAT GGG GGT ACA TAT GAC ATA TAT ACA ACT ACA CGT TAT AAC GCA CCT TCC ATT GAT GGC GAT CGC ACT ACT          847
TTT ACG CAG TAC TGG AGT GTT CGC CAG TCG AAG AGA CCA ACT GGA AGC AAC GCT ACA ATC ACT TTC ACG AAT CAT          922
GTG AAC GCA TGG AAG AGC CAT GGA ATG AAT CTG GGC AGT AAT TGG GCT TAC CAA GTC ATG GCG ACA GAA GGA TAT          997
CAA AGT AGT GGA AGT TCT AAC GTA ACA GTG TGG TAA CAGATCATCCTTAATCAGGGGTAGCTAACGGGCTGCTGATCGTTCCTTGAGA      1085
AGTTTTATAATCAATGATTATTAAAATCGTTAGTAAGGGTTAAAGGTTGTTTTCTACTAGGTGAACGGCCTTGCAATTGCTGGAGGTAGGGTATTCTCCA      1185
TCTGGTTTTATAACTTTTCCTATAGGTTAATAGAATGGTATTTAAATGAGAATGCTACAATTTTTTCTAGTCAGCGCTTGCTCACAACAGACACCTTTAC      1285
ATAACTCTTCTTTATCAAACATAAGCCTTATTCAAAATAAAAATATCTAGTAGTTGACCTGCAG                                          1349
```

Acknowledgement. We thank Dr. Makoto Yaguchi for help with the amino acid sequence analysis.

*To whom correspondence should be addressed

References.

1. Sanger, F. et al. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
2. Fukusaki, E. et al. (1984) FEBS Lett. 171, 197-201.
3. Paice, M. et al. (1986) Arch Microbiol. 144, 201-206.
4. Yang, R.C.A. et al. (1979) Nucleic Acids Res. 7, 651-668.
5. Platt, T. (1981) Cell 24, 10-23.