JOURNAL OF BACTERIOLOGY, May 1987, p. 2301–2306
0021-9193/87/052301-06$02.00/0

Vol. 169, No. 5

Entire Nucleotide Sequence of the Pullulanase Gene of *Klebsiella aerogenes* W70

NOBUHIRO KATSURAGI, NOBORU TAKIZAWA,† AND YOSHIKATSU MUROOKA*

Department of Fermentation Technology, Faculty of Engineering, Hiroshima University, Saijo, Higashi-Hiroshima 724, Japan

Received 12 December 1986/Accepted 11 February 1987

We determined the entire nucleotide sequence of the *Klebsiella aerogenes* W70 pullulanase gene (*pulA*) contained on a 4.2-kilobase-pair fragment of plasmid pPB174. The amino acid composition deduced from an open reading frame of 3,288 base pairs agreed closely with that determined for the intracellular pullalanase. The precursor enzyme consisted of 1,096 amino acid residues and contained a hydrophobic N-terminal signal peptide and the consensus sequence for the bacterial prelipoprotein signal peptide cleavage site.

Pullulanase (EC 3.2.1.41, pullulan 6-glucanohydrolase), a starch-debranching enzyme, hydrolyzes (1-6)-α-glucosidic linkages in pullulan and starch to form maltotriose. This enzyme is produced by the gram-negative bacterium *Klebsiella aerogenes* and secreted from the cells into the culture broth (2). We recently cloned the pullulanase gene (*pulA*) of *K. aerogenes* W70 (19), which is located within a 4.2-kilobase-pair segment of the chromosomal DNA on plasmid pPB174. *K. aerogenes* cells carrying plasmid pPB174 overproduce both extracellular and intracellular pullulanase when the cells are induced by maltose. The two enzymes are immunologically identical to those produced by *K. aerogenes* W70 (19). We sequenced the entire pullulanase gene (*pulA*) and showed that the precursor enzyme contained N-terminal signal peptide. The precursor and mature proteins deduced from the DNA sequence showed a high degree of correspondence to the total amino acid compositions of the intracellular and extracellular enzymes.

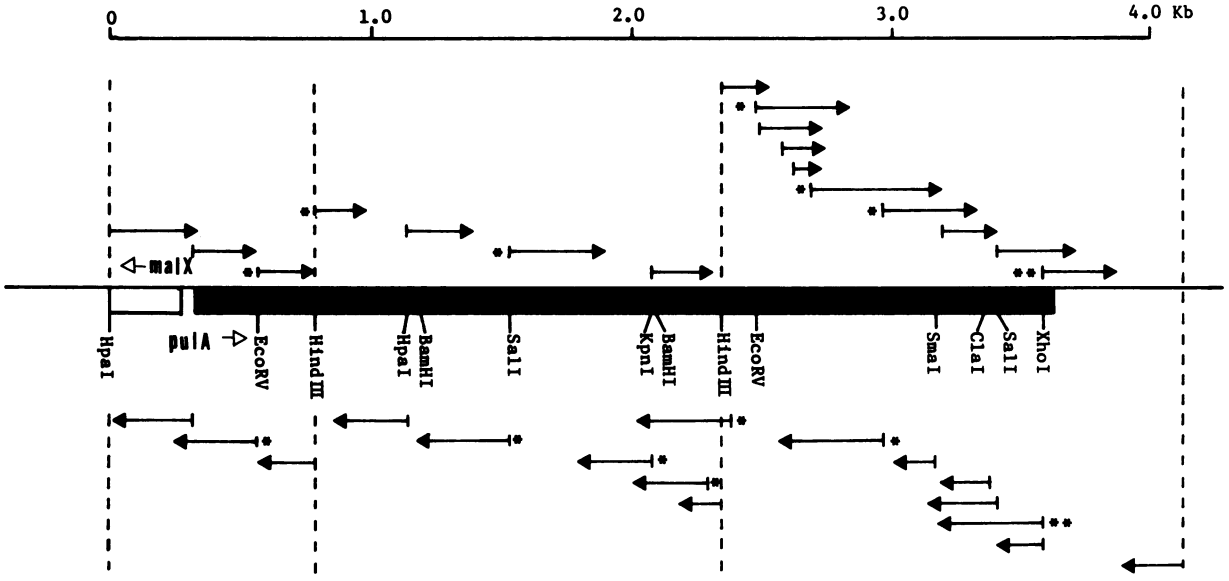

FIG. 1. Sequencing strategy for the 4.2-kilobase-pair fragment of pPB17 containing the *pulA* gene. The restriction sites shown within the pullulanase-coding sequence were used to subclone DNA fragments in the M13 vectors. The closed and open boxes indicate the *pulA* and *malX* genes, respectively. The open arrows show the transcriptional directions of the two genes. The closed arrows show the lengths and directions of the sequences determined. The fragments were reanalyzed by the modification method with methoxyamine and bisulfite (*) and with deoxy-7-deazaguanosine triphosphate (**).

LB and PMA media (18) were used for routine cultivation of *Escherichia coli* and *K. aerogenes* cells. We prepared intracellular and extracellular pullulanases from *K. aerogenes* W70 (pMP1) (19) and purified each to a single protein as described by Eisele et al. (4), except that DEAE-Toyopearl 650S and Toyopearl HW55S (Toyosoda Co. Ltd., Tokyo, Japan) were used instead of DEAE-cellulose and Sephadex G-200, respectively. The amino acid composition of the enzyme was determined by using a Hitachi model 835 amino acid analyzer. Restriction endonucleases, T4 DNA

* Corresponding author.
† Present address: Okayama University of Science, Ridai, Okayama 700, Japan.

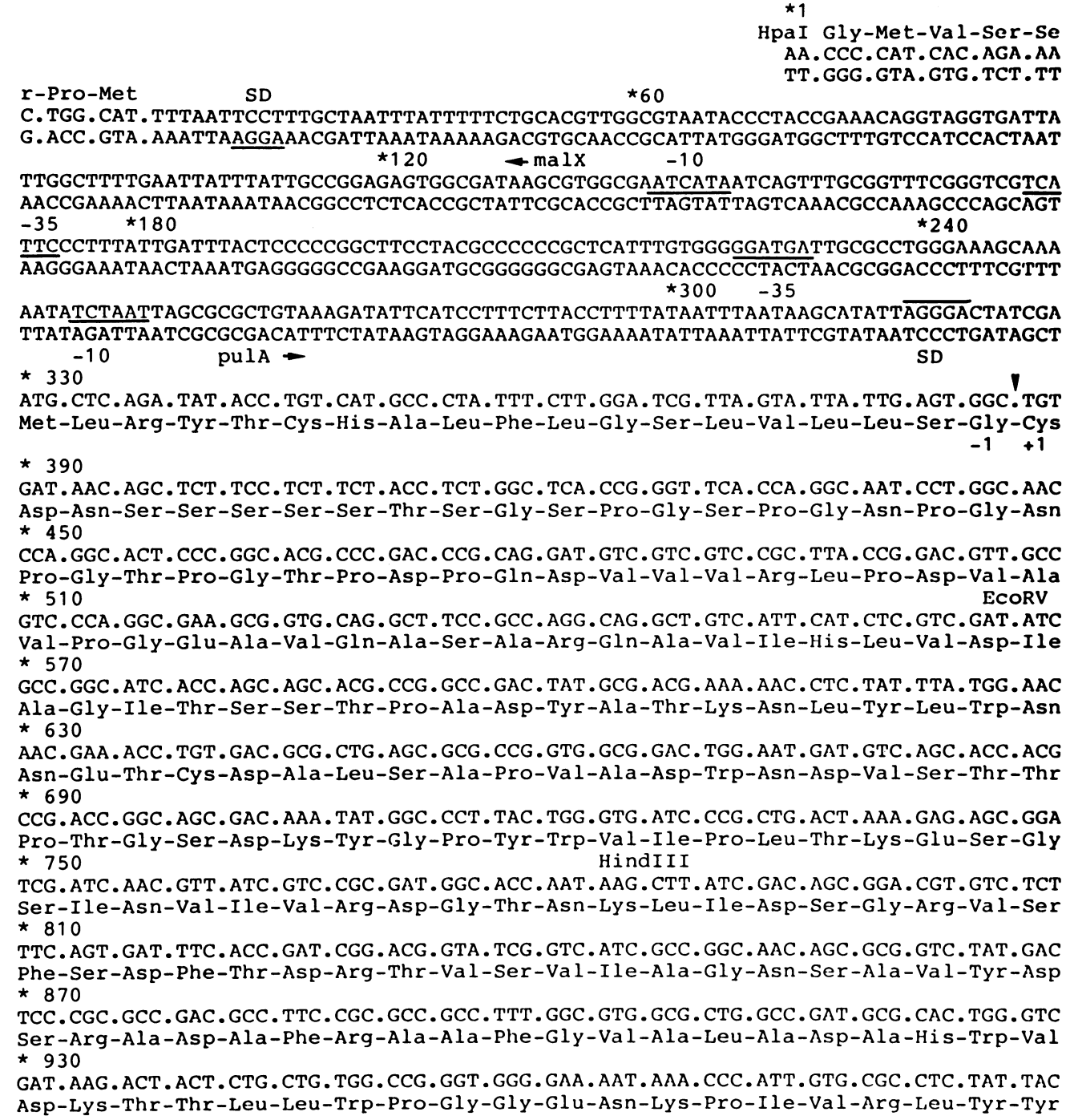

FIG. 2. Nucleotide sequence of the DNA fragment containing the pullulanase gene. The presumptive −35, −10, Shine-Dalgarno ribosome-binding site (SD), +1 processing site, and the stem-loop structure are indicated beneath the sequence. The restriction endonuclease sites are shown above the sequence. The initiation signals for the *malX* gene are as described by Chapon and Raibaud (3).

ligase, alkaline phosphatase, BAL 31 nuclease, and Klenow fragment (obtained from Takara Shuzo Co. Ltd., Kyoto, Japan, or Toyobo Co. Ltd., Osaka, Japan) were used to subclone the fragments of pPB174 containing the *pulA* gene (19). The dideoxy chain terminator sequence method (16) was used to determine the DNA sequence. The Takara sequence kit (Takara Shuzo) was used for the reactions. [α-^{32}P]dCTP (>400 Ci/mmol) was obtained from Amersham Corp., Arlington Heights, Ill.

The appropriate fragments (Fig. 1) were ligated to replicative-form DNA of the M13mp18 or mp19 cloning vectors and used to transfect *E. coli* JM105 (21). Bacteriophages from colorless plaques were used to infect 2.0-ml cultures of JM105, and double-stranded replicative-form DNA was isolated from the infected cells. Minilysate DNA (8) was then examined to select the correct inserts and transfected into strain JM105. Single-stranded DNA was then purified from bacteriophage liberated by these transfected cells (21). The

```
* 990
AGC.CAC.AGC.AGT.AAG.GTG.GCC.GCC.GAC.AGT.AAC.GGC.GAA.TTT.AGC.GAT.AAA.TAT.GTC.AAG
Ser-His-Ser-Ser-Lys-Val-Ala-Ala-Asp-Ser-Asn-Gly-Glu-Phe-Ser-Asp-Lys-Tyr-Val-Lys
*1050
CTG.ACC.CCC.ACC.ACC.GTC.AAC.CAG.CAG.GTA.AGC.ATG.CGC.TTC.CCG.CAT.CTC.GCC.AGC.TAT
Leu-Thr-Pro-Thr-Thr-Val-Asn-Gln-Gln-Val-Ser-Met-Arg-Phe-Pro-His-Leu-Ala-Ser-Tyr
*1110                          HpaI
CCA.GCC.TTT.AAA.CTG.CCG.GAT.GAT.GTT.AAC.GTC.GAT.GAA.TTG.CTG.CAG.GGC.GAC.GAC.GGT
Pro-Ala-Phe-Lys-Leu-Pro-Asp-Asp-Val-Asn-Val-Asp-Glu-Leu-Leu-Gln-Gly-Asp-Asp-Gly
*1170                    BamHI
GGG.ATC.GCG.GAA.AGC.GAT.GGG.ATC.CTG.AGC.CTC.AGC.CAC.CCA.GGT.GCA.GAC.CGC.CGG.CGT
Gly-Ile-Ala-Glu-Ser-Asp-Gly-Ile-Leu-Ser-Leu-Ser-His-Pro-Gly-Ala-Asp-Arg-Arg-Arg
*1230
GCT.GGA.CGA.TAC.CTA.TGC.CGC.CGC.GCC.GAG.GCG.CTG.AGC.TAT.GGC.GCC.CAG.CTA.ACC.GAT
Ala-Gly-Arg-Tyr-Leu-Cys-Arg-Arg-Ala-Glu-Ala-Leu-Ser-Tyr-Gly-Ala-Gln-Leu-Thr-Asp
*1290
AGC.GGC.GTG.ACC.TTC.CGC.GTC.TGG.GCG.CCC.ACG.GCG.CAG.CAG.GTC.GAG.CTG.GTG.ATC.TAT
Ser-Gly-Val-Thr-Phe-Arg-Val-Trp-Ala-Pro-Thr-Ala-Gln-Gln-Val-Glu-Leu-Val-Ile-Tyr
*1350
AGC.GCG.GAC.AAG.AAA.GTG.ATA.GCC.AGT.CAC.CCG.ATG.ACC.CGC.GAC.AGC.GCC.TCC.GGC.GCC
Ser-Ala-Asp-Lys-Lys-Val-Ile-Ala-Ser-His-Pro-Met-Thr-Arg-Asp-Ser-Ala-Ser-Gly-Ala
*1410
TGG.TCC.TGG.CAG.GGG.GGA.AGC.GAC.CTG.AAG.GGC.GCC.TTC.TAC.CGC.TAC.GCG.ATG.ACG.GTC
Trp-Ser-Trp-Gln-Gly-Gly-Ser-Asp-Leu-Lys-Gly-Ala-Phe-Tyr-Arg-Tyr-Ala-Met-Thr-Val
*1470
TAC.CAC.CCG.CAG.TCG.CGT.AAA.GTC.GAG.CAG.TAC.GAA.GTG.ACC.GAT.CCC.TAC.GCC.CAC.AGT
Tyr-His-Pro-Gln-Ser-Arg-Lys-Val-Glu-Gln-Tyr-Glu-Val-Thr-Asp-Pro-Tyr-Ala-His-Ser
*1530 SalI
TTG.TCG.ACC.AAC.TCG.GAG.TAC.AGC.CAG.GTG.GTC.GAT.CTC.AAC.GAC.AGC.GCG.CTG.AAG.CCG
Leu-Ser-Thr-Asn-Ser-Glu-Tyr-Ser-Gln-Val-Val-Asp-Leu-Asn-Asp-Ser-Ala-Leu-Lys-Pro
*1590
GAA.GGC.TGG.GAC.GGG.CTG.ACG.ATG.CCG.CAC.GCG.CAG.AAA.ACC.AAA.GCC.GAT.CTG.GCG.AAA
Glu-Gly-Trp-Asp-Gly-Leu-Thr-Met-Pro-His-Ala-Gln-Lys-Thr-Lys-Ala-Asp-Leu-Ala-Lys
*1650
ATG.ACG.ATT.CAC.GAA.TCG.CAT.ATT.CGC.GAT.CTC.TCC.GCC.TGG.GAT.CAA.ACC.GTG.CCT.GCG
Met-Thr-Ile-His-Glu-Ser-His-Ile-Arg-Asp-Leu-Ser-Ala-Trp-Asp-Gln-Thr-Val-Pro-Ala
*1710
GAG.CTG.CGC.GGT.AAG.TAT.CTG.GCG.CTC.ACC.GCC.CAG.GAA.AGC.AAT.ATG.GTC.CAG.CAT.CTG
Glu-Leu-Arg-Gly-Lys-Tyr-Leu-Ala-Leu-Thr-Ala-Gln-Glu-Ser-Asn-Met-Val-Gln-His-Leu
*1770
AAA.CAG.CTG.TCG.GCC.TCG.GGC.GTG.ACC.CAT.ATT.GAG.CTG.CTG.CCG.GTC.TTC.GAT.CTG.GCG
Lys-Gln-Leu-Ser-Ala-Ser-Gly-Val-Thr-His-Ile-Glu-Leu-Leu-Pro-Val-Phe-Asp-Leu-Ala
*1830
ACG.GTC.AAT.GAG.TTC.AGC.GAC.AAG.GTC.GCC.GAT.ATT.CAG.CAG.CCG.TTC.AGC.CGC.CTG.TGC
Thr-Val-Asn-Glu-Phe-Ser-Asp-Lys-Val-Ala-Asp-Ile-Gln-Gln-Pro-Phe-Ser-Arg-Leu-Cys
*1890
GAG.GTC.AAT.AGC.GCG.GTG.AAG.AGC.AGC.GAG.TTT.GCG.GGC.TAT.TGC.GAC.AGC.GGT.TCG.ACG
Glu-Val-Asn-Ser-Ala-Val-Lys-Ser-Ser-Glu-Phe-Ala-Gly-Tyr-Cys-Asp-Ser-Gly-Ser-Thr
*1950
GTT.GAA.GAG.GTG.CTG.ACT.CAG.CTG.AAG.CAG.AAC.GAC.AGC.AAG.GAT.AAC.CCG.CAG.GTG.CAG
Val-Glu-Glu-Val-Leu-Thr-Gln-Leu-Lys-Gln-Asn-Asp-Ser-Lys-Asp-Asn-Pro-Gln-Val-Gln
```

Figure continued on following page

labeled fragments were separated by polyacrylamide gel electrophoresis under denaturing conditions. To eliminate the G-C band compression, several fragments shown in Fig. 1 were reanalyzed by the modification method with methoxyamine and bisulfite (1) and with deoxy-7-deazaguanosine triphosphate (9). Computerassisted sequence analysis was carried out with GENETYX programs (SDC Software Development Co. Ltd., Tokyo, Japan).

A 4,130-base-pair fragment of pPB174 DNA containing the *pulA* gene is shown in Fig. 2. We found an open reading frame of 3,288 bases, with the ATG initiation codon at position 330 and the TGA termination codon at position 3618. A possible Shine-Dalgarno ribosome-binding site (17) with the sequence AGGGA was located 8 bases upstream of the initiation codon. Potential −35 and −10 promoter regions were detected at positions 230 (GGATGA) and 260 (TCTAAT). The *Xho*I site, which was suggested by deletion analysis (19) to be near the 3′ end of the *pulA* gene in a region essential for pullulanase activity, was 36 bases upstream from the stop codon TGA. Furthermore, 150 bases down-

```
*2010
GCG.TTG.AAT.ACG.CTG.GTG.GCG.CAG.ACC.GAC.TCC.TAT.AAC.TGG.GGC.TAC.GAT.CCG.TTC.CAC
Ala-Leu-Asn-Thr-Leu-Val-Ala-Gln-Thr-Asp-Ser-Tyr-Asn-Trp-Gly-Tyr-Asp-Pro-Phe-His
*2070    KpnI          BamHI
TAC.ACG.GTA.CCG.GAA.GGA.TCC.TAC.GCC.ACC.GAT.CCG.GAA.GGC.ACA.GCG.CGT.ATT.AAA.GAG
Tyr-Thr-Val-Pro-Glu-Gly-Ser-Tyr-Ala-Thr-Asp-Pro-Glu-Gly-Thr-Ala-Arg-Ile-Lys-Glu
*2130
TTC.CGC.ACC.ATG.ATT.CAG.GCG.ATC.AAG.CAG.GAT.CTG.GGA.ATG.AAC.GTC.ATT.ATG.GAC.GTG
Phe-Arg-Thr-Met-Ile-Gln-Ala-Ile-Lys-Gln-Asp-Leu-Gly-Met-Asn-Val-Ile-Met-Asp-Val
*2190
GTG.TAC.AAC.CAC.ACC.AAC.GCC.GCC.GGC.CCG.ACC.GAT.CGC.ACC.TCG.GTA.CTG.GAT.AAG.ATC
Val-Tyr-Asn-His-Thr-Asn-Ala-Ala-Gly-Pro-Thr-Asp-Arg-Thr-Ser-Val-Leu-Asp-Lys-Ile
*2250
GTC.CCC.TGG.TAC.TAC.CAG.CGT.CTG.AAC.GAA.ACC.ACC.GGC.AGC.GTG.GAA.TCG.GCT.ACC.TGT
Val-Pro-Trp-Tyr-Tyr-Gln-Arg-Leu-Asn-Glu-Thr-Thr-Gly-Ser-Val-Glu-Ser-Ala-Thr-Cys
*2310                                        HindIII
TGC.TCC.GAC.TCG.GCG.CCA.GAG.CAC.CGG.ATG.TTC.GCC.AAG.CTT.ATC.GCC.GAT.TCA.CTG.GCG
Cys-Ser-Asp-Ser-Ala-Pro-Glu-His-Arg-Met-Phe-Ala-Lys-Leu-Ile-Ala-Asp-Ser-Leu-Ala
*2370
GTA.TGG.ACC.ACC.GAT.TAT.AAG.ATC.GAT.GGC.TTC.CGC.TTC.GAC.CTG.ATG.GGC.TAC.CAC.CCG
Val-Trp-Thr-Thr-Asp-Tyr-Lys-Ile-Asp-Gly-Phe-Arg-Phe-Asp-Leu-Met-Gly-Tyr-His-Pro
*2430                                                    EcoRV
AAA.GCG.CAG.ATC.CTC.TCG.GCC.TGG.GAA.CGC.ATT.AAA.GCG.CTG.AAC.CCG.GAT.ATC.TAC.TTC
Lys-Ala-Gln-Ile-Leu-Ser-Ala-Trp-Glu-Arg-Ile-Lys-Ala-Leu-Asn-Pro-Asp-Ile-Tyr-Phe
*2490
TTT.GGC.GAA.GGC.TGG.GAT.TCC.AAC.CAG.AGC.GAT.CGC.TTT.GAA.ATT.GCC.TCG.CAA.ATC.AAT
Phe-Gly-Glu-Gly-Trp-Asp-Ser-Asn-Gln-Ser-Asp-Arg-Phe-Glu-Ile-Ala-Ser-Gln-Ile-Asn
*2550
CTC.AAA.GGC.ACC.GGG.ATC.GGC.ACG.TTC.TCC.GAT.CGT.CTG.CGC.GAC.GCC.GTG.CGC.GGC.GGC
Leu-Lys-Gly-Thr-Gly-Ile-Gly-Thr-Phe-Ser-Asp-Arg-Leu-Arg-Asp-Ala-Val-Arg-Gly-Gly
*2610
GGG.CCG.TTC.GAC.TCC.GGT.GAC.GCA.TTA.CGC.CAG.AAC.CAG.GGC.GTG.GGC.AGC.GGG.GCT.GGC
Gly-Pro-Phe-Asp-Ser-Gly-Asp-Ala-Leu-Arg-Gln-Asn-Gln-Gly-Val-Gly-Ser-Gly-Ala-Gly
*2670
GTT.CTG.CCG.AAT.GAG.CTG.ACC.ACC.CTG.AGC.GAC.GAT.CAG.GCG.CGT.CAC.CTC.GCC.GAT.CTG
Val-Leu-Pro-Asn-Glu-Leu-Thr-Thr-Leu-Ser-Asp-Asp-Gln-Ala-Arg-His-Leu-Ala-Asp-Leu
*2730
ACC.CGT.CTC.GGC.ATG.GCC.GGT.AAC.CTT.GCG.GAC.TTT.GTG.CTG.ATC.GAC.AAA.GAC.GGC.GCG
Thr-Arg-Leu-Gly-Met-Ala-Gly-Asn-Leu-Ala-Asp-Phe-Val-Leu-Ile-Asp-Lys-Asp-Gly-Ala
*2790
GTG.AAG.AGA.GGC.AGC.GAG.ATT.GAT.TAT.AAC.GGC.GCG.CCA.GGC.GGC.TAT.GCG.GCT.GAT.CCG
Val-Lys-Arg-Gly-Ser-Glu-Ile-Asp-Tyr-Asn-Gly-Ala-Pro-Gly-Gly-Tyr-Ala-Ala-Asp-Pro
*2850
ACG.GAA.GTC.GTG.AAC.TAT.GTG.TCA.AAA.CAC.GAT.AAC.CAG.ACG.CTG.TGG.GAC.ATG.ATC.AGC
Thr-Glu-Val-Val-Asn-Tyr-Val-Ser-Lys-His-Asp-Asn-Gln-Thr-Leu-Trp-Asp-Met-Ile-Ser
*2910
TAT.AAA.GCC.GCT.CAG.GAG.GCG.GAT.CTC.GAT.ACC.CGC.GTC.CGG.ATG.CAG.GCG.GTG.TCG.CTG
Tyr-Lys-Ala-Ala-Gln-Glu-Ala-Asp-Leu-Asp-Thr-Arg-Val-Arg-Met-Gln-Ala-Val-Ser-Leu
*2970
GCG.ACG.GTG.ATG.CTC.GGC.CAG.GGG.ATC.GCC.TTT.GAC.CAG.CAG.GGC.TCG.GAG.CTG.CTG.CGC
Ala-Thr-Val-Met-Leu-Gly-Gln-Gly-Ile-Ala-Phe-Asp-Gln-Gln-Gly-Ser-Glu-Leu-Leu-Arg
```

stream from the C-terminal codon of pullulanase was a palindromic sequence (ΔG, −20.9 kcal/mol [ca. −87.4 kJ/mol]) that might act as a transcription terminator (15). We also detected the promoter signals of *malX* at positions 174 (GAATGA) and 146 (TATGAT) in the opposite orientation to *pulA*, as shown in *Klebsiella pneumoniae* (3). The function of the *malX* gene is unknown, but it is controlled by *malT* (3).

The amino acid sequence deduced from this reading frame showed that a 19-amino-acid signal peptide could precede the mature pullulanase sequence, with a glycine residue at the potential cleavage site where processing of the precursor would occur to yield the mature pullulanase of *K. pneumoniae* (3, 14). This sequence was followed by a cysteine residue, which was modified by palmitate and probably becomes the amino terminus of the mature protein (14). We also found that both intracellular and extracellular pullulanases of *K. aerogenes* were labeled with [^{3}H]palmitate (Y. Murooka, N. Takizawa, and N. Katsuragi, Proc. 14th Int. Congr. Microbiol., p. 298, 1986). The sequence

```
*3030
TCT.AAA.TCC.TTT.ACC.CGC.GAT.TCG.TAT.GAT.TCC.GGC.GAC.TGG.TTT.AAC.CGC.GTG.GAC.TAC
Ser-Lys-Ser-Phe-Thr-Arg-Asp-Ser-Tyr-Asp-Ser-Gly-Asp-Trp-Phe-Asn-Arg-Val-Asp-Tyr
*3090
TCC.CTG.CAG.GAC.AAC.AAC.TAC.AAC.GTC.GGT.ATG.CCG.CGC.AGC.AGC.GAT.GAT.GGC.AGC.AAT
Ser-Leu-Gln-Asp-Asn-Asn-Tyr-Asn-Val-Gly-Met-Pro-Arg-Ser-Ser-Asp-Asp-Gly-Ser-Asn
*3150                SmaI
TAT.GAC.ATT.ATC.GCC.CGG.GTG.AAA.GAC.GCG.GTG.GCT.ACT.CCG.GGT.GAA.ACG.GAG.CTC.AAG
Tyr-Asp-Ile-Ile-Ala-Arg-Val-Lys-Asp-Ala-Val-Ala-Thr-Pro-Gly-Glu-Thr-Glu-Leu-Lys
*3210
CAG.ATG.ACC.GCG.TTT.TAT.CAG.GAG.CTG.ACC.GCG.CTG.CGT.AAA.TCG.TCT.CCG.CTG.TTT.ACC
Gln-Met-Thr-Ala-Phe-Tyr-Gln-Glu-Leu-Thr-Ala-Leu-Arg-Lys-Ser-Ser-Pro-Leu-Phe-Thr
*3270
CTC.GGC.GAC.GGC.GCG.ACG.GTG.ATG.AAG.CGC.GTG.GAC.TTC.CGC.AAT.ACT.GGC.GCC.GAT.CAG
Leu-Gly-Asp-Gly-Ala-Thr-Val-Met-Lys-Arg-Val-Asp-Phe-Arg-Asn-Thr-Gly-Ala-Asp-Gln
*3330                        ClaI
CAG.ACG.GGT.CTG.CTG.GTG.ATG.ACC.ATC.GAT.GAC.GGG.ATG.CAG.GCT.GGC.CGC.CAG.TCT.GGA
Gln-Thr-Gly-Leu-Leu-Val-Met-Thr-Ile-Asp-Asp-Gly-Met-Gln-Ala-Gly-Arg-Gln-Ser-Gly
*3390      SalI
CAG.CCG.TGT.CGA.CGG.CAT.CGT.GGT.GGC.GAT.CAA.CGC.CGC.GCC.GGA.AAG.CCG.GAC.GCT.GCA
Gln-Pro-Cys-Arg-Arg-His-Arg-Gly-Gly-Asp-Gln-Arg-Arg-Ala-Gly-Lys-Pro-Asp-Ala-Ala
*3450
GGA.CTT.CGC.CGG.CAC.ATC.GCT.CCA.GCT.GAG.CGC.TAT.TCA.GCA.GGC.GGC.GGG.CGA.CCG.GTC
Gly-Leu-Arg-Arg-His-Ile-Ala-Pro-Ala-Glu-Arg-Tyr-Ser-Ala-Gly-Gly-Gly-Arg-Pro-Val
*3510
GCT.GGC.GAG.CGC.GTG.CAG.GTT.GCC.GCT.GAC.GGT.TCG.GTC.ACG.CTG.CCG.GCC.TGG.TCG.GTT
Ala-Gly-Glu-Arg-Val-Gln-Val-Ala-Ala-Asp-Gly-Ser-Val-Thr-Leu-Pro-Ala-Trp-Ser-Val
*3570    XhoI
GCC.GTT.CTC.GAG.TTG.CCG.CAG.GCG.AGT.CGC.AGG.GCG.CTG.GCC.TGC.CGG.TGA.GCAGTAAATAA
Ala-Val-Leu-Glu-Leu-Pro-Gln-Ala-Ser-Arg-Arg-Ala-Leu-Ala-Cys-Arg-***
                                                          *3690
CATAACGTAGGGCGCCCGTAAGCGTCTGGCTGATGACAAACCTGTTTCGTTATTGTGCTCGCCTTTCTGTCGGGTGGCG
                                                  *3750
GCTTTCGCTTACCCGGCCTACGAAAGATTGCAATTTCAAACAGTTATGCAGTTATCTGTAGGCCCGCGCAAGTAGCGCG
                          *3810
CCGGCCACTTCCTGAAGATTGCACTTTGTCAGCAGTCTGGGGCATCCGTAAGGGCGCCTTCTCGTTATCGTGAGGGAAC
 *3870                                                              *3930
AGGAGACGGTGGATGTCGGATAAGTGCATGAAACAGTATGCAGTATCTGTAGCCGCCAACGGTACGCCGCCGGGACTTC
                                                      *3990
CTGAAGATTGCACTTTGTCAGCAGTCTGGGGCATCCGTAAGGGCGCCTTCTCGTTATCGTGAGGGAACAGGAGACGGAT
                                  *4050
TGAGTATGATTGTTCGTGATGACAGTGAAAAGCATCCGCCTCTGGTCGCATCAGCCGCCGTCCTCGCTGGATGATCGCC
          *4110
GGTGAGCCGCGGTCGGATGGTGCAGATC
```

upstream from the cysteine residue (Leu-Leu-Ser-Gly) is the same as the consensus sequence for the bacterial prelipoprotein cleavage site (6, 10–12, 14).

The extracellular pullulanase of *K. aerogenes* was previously shown by Eisele et al. (4) to be a single polypeptide chain molecule with a molecular weight of about 143,000, whereas Ohba and Ueda (13, 20) estimated the molecular weights of extracellular and intracellular enzymes of the same organism to be about 70,000 and 80,000, respectively. The translated sequence shows that the precursor pullulanase contained 1,096 amino acid residues, giving a molecular weight of 119,334, and that the predicted mature protein contains 1,077 amino acids, giving a molecular weight of 117,258. Thus, we conclude that pullulanase is a monomeric protein.

The amino acid compositions of the precursor and mature proteins deduced from the DNA sequence were compared with our amino acid analysis data and with data reported earlier (4). The values deduced from the DNA sequence for each residue agree closely with those determined for the analyzed intracellular and extracellular enzymes. The preferential codon usage for the *pulA* gene is generally similar to that demonstrated for several genes of *E. coli* (5; data not shown). The hydropathy plot (7) showed that several regions in the precursor form of the pullulanase contained a relatively high concentration of hydrophobic amino acids. The first region was in the signal peptide at the extreme N-terminal end of the molecule. The other hydrophobic regions were dispersed throughout the mature pullulanase molecule (data not shown).

The transcription regulatory signals for the *pulA* gene are identical to those reported by Chapon and Raibaud (3), who sequenced the promoter regions of *pulA* and *malX* of *K. pneumoniae* ATCC 15050 and demonstrated the transcriptional startpoints. We found about 75% homology between our DNA sequence and the 5′ end of the *K. pneumoniae*

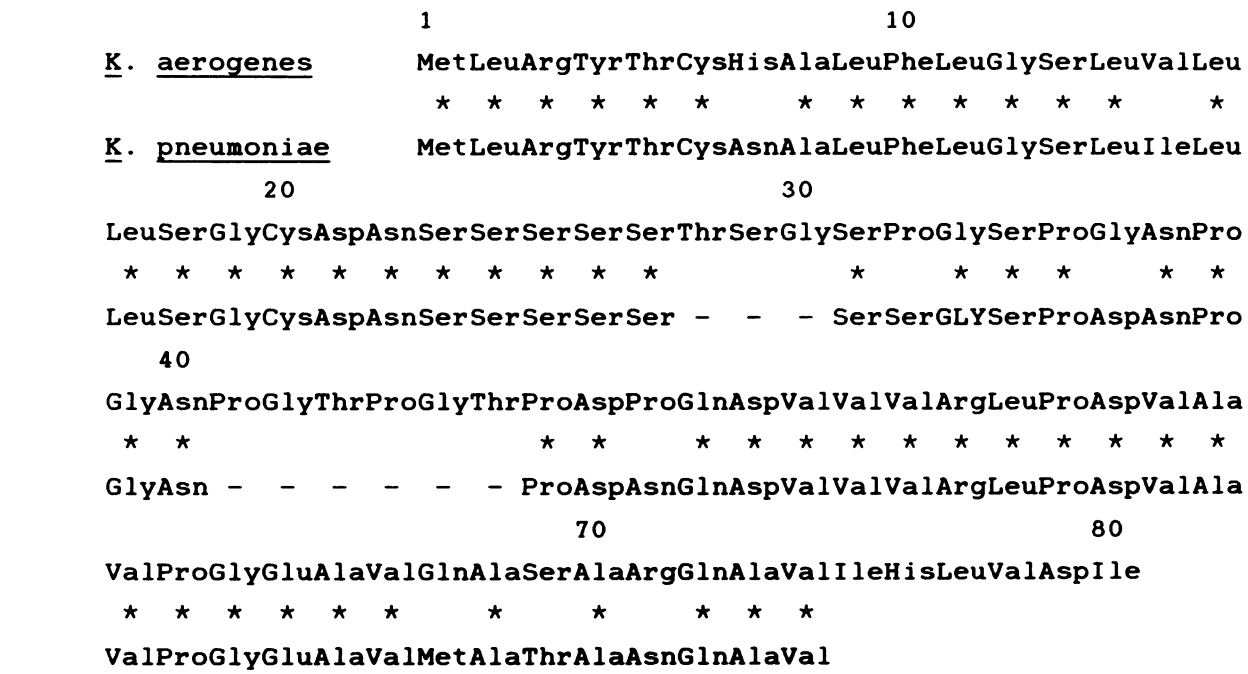

FIG. 3. Comparison of the N-terminal sequences of the *pulA* gene products from *K. aerogenes* W70 and *K. pneumoniae* ATCC 15050 (3).

DNA sequence (674 base pairs) (3). Although 25 of 27 amino acid residues of the N-terminal hydrophobic regions of *K. aerogenes* and *K. pneumoniae* are identical, parts of the N-terminal region of predicted mature proteins of *K. pneumoniae* were deleted (Fig. 3). Thus, at least the deleted parts of the N-terminal region of the mature enzyme of *K. aerogenes* were not necessary for the enzyme activity.

We thank Y.-M. Hong of Otsuka Pharmaceutical Co. Ltd., Tokushima, Japan, for the amino acid analysis and M. Ikemi for helpful discussions.

This work was supported by grant 60560118 to Y.M. from the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

1. **Ambartsumyan, N. S., and A. M. Mazo.** 1980. Elimination of the secondary structure effect of gel sequencing of nucleic acids. FEBS Lett. **114**:265–268.
2. **Bender, H., and K. Wallenfels.** 1961. Untersuchungen an Pullulan. II. Spezifisher Abbau durch ein Bacterielles Enzym. Biochem. Z. **334**:79–95.
3. **Chapon, C., and O. Raibaud.** 1985. Structure of two divergent promoters located in front of the gene encoding pullulanase in *Klebsiella pneumoniae* and positively regulated by the *malT* product. J. Bacteriol. **164**:639–645.
4. **Eisele, B., I. R. Rasched, and K. Wallenfels.** 1972. Molecular characterization of pullulanase from *Aerobacter aerogenes*. Eur. J. Biochem. **26**:62–67.
5. **Grosjean, H., and W. Fiers.** 1962. Preferential codon usage in procaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199–209.
6. **Huang, Y., G. Ching, and M. Inouye.** 1983. Comparison of the lipoprotein genes among *Enterobacteriaceae*: DNA sequence of *Morganella morganii* lipoprotein gene and its expression in *Escherichia coli*. J. Biol. Chem. **258**:8139–8145.
7. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
8. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 366–367. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
10. **Mizushima, S.** 1984. Post-translational modification and processing of outer membrane prolipoproteins in *Escherichia coli*. Mol. Cell. Biochem. **60**:5–15.
11. **Neugebauer, K., R. Spengel, and H. Schaller.** 1981. Penicillinase from *Bacillus licheniformis*: nucleotide sequence of the gene and implications for the biosynthesis of a secretory protein in a gram-positive bacterium. Nucleic Acids Res. **9**:2577–2588.
12. **Nielson, J. B. K., M. P. Caulfield, and J. O. Lampen.** 1981. Lipoprotein nature of *Bacillus licheniformis* membrane penicillinase. Proc. Natl. Acad. Sci. USA **78**:3511–3515.
13. **Ohba, R., and S. Ueda.** 1973. Purification, crystallization, and some properties of extracellular pullulanase from *Aerobacter aerogenes*. Agric. Biol. Chem. **36**:2381–2391.
14. **Pugsley, A. P., C. Chapon, and M. Schwartz.** 1986. Extracellular pullulanase of *Klebsiella pneumoniae* is a lipoprotein. J. Bacteriol. **166**:1083–1088.
15. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
16. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
17. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity of nonsense triplets and ribosomal binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
18. **Takizawa, N., and Y. Murooka.** 1984. Intergeneric transfer of the pullulanase gene between *Klebsiella aerogenes* and *Escherichia coli* by in vivo genetic manipulation. Agric. Biol. Chem. **48**:1451–1458.
19. **Takizawa, N., and Y. Murooka.** 1985. Cloning of the pullulanase gene and overproduction of pullulanase in *Escherichia coli* and *Kelbsiella aerogenes*. Appl. Environ. Microbiol. **49**:294–298.
20. **Ueda, S., and R. Ohba.** 1972. Purification, crystallization, and some properties of extracellular pullulanase from *Aerobacter aerogenes*. Agric. Biol. Chem. **36**:2381–2391.
21. **Yanisch, C., J. Viera, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.