

Nucleotide sequence of the cell-envelope *murG* gene of *Escherichia coli*

Dominique Mengin-Lecreux, Laurent Texier and Jean van Heijenoort

URA 1131 du Centre National de la Recherche Scientifique, Biochimie Moléculaire et Cellulaire, Bâtiment 432, Université Paris-Sud, 91405 Orsay, France

Submitted March 26, 1990

EMBL accession no. X52540

The *murG* gene of *Escherichia coli* earlier identified at 2-min on the chromosome (1) has now been precisely located within a large cluster of genes from *pbpB* to *envA* that codes for proteins involved in peptidoglycan synthesis or cell division (2, 3). Its nucleotide sequence has been determined within the left part of a 1.7 kilobase-pair *SalI*-*KpnI* chromosomal fragment (2) originating from pLC 26-6 of the Clarke and Carbon collection. Complementation of the *murG* mutant strain GS58 (1) with various constructed plasmids shows that the *murG* gene begins between the *MluI* and *SmaI* sites at positions 41 and 141 respectively. Two potential ATG initiation codons are found in this area at positions 82 and 106, the first one overlapping by 4 bases the end of the preceding *ftsW* gene (4). Both are preceded by a consensus sequence of ribosome binding site: GAGGT (at 72) and AAGGA (at 92). The coding region may thus contain

1041 or 1065 nucleotides which are translated into a protein of 347 or 355 amino acids, corresponding to a molecular weight of 36914 or 37771. This finding is in agreement with the expression of a 38000 molecular weight labeled protein from plasmids carrying the *SalI*-*KpnI* insert previously observed in maxicell experiments (2). The figure shows only the 37771 molecular weight protein. The final proof will be provided by determining the N-terminal amino acid sequence of the purified *murG* encoded protein.

REFERENCES

1. Salmond, G.P.C. *et al.* (1980) *J. Bacteriol.* **144**, 438-440.
2. Mengin-Lecreux, D. *et al.* (1989) *J. Bacteriol.* **171**, 6126-6134.
3. Ishino, F. *et al.* (1989) *J. Bacteriol.* **171**, 5523-5530.
4. Ikeda, M. *et al.* (1989) *J. Bacteriol.* **171**, 6375-6378.

SalI

```

GTCGACAGCCATCATGATGCTGTTGCGTATTGATTATGAAACGCGTCTGGAGAAAGCGCAGGCGTTTGTACGAGGTTACAGATGAGTGGTCAA 93
                                     M S G Q
GGAAAGCGATTAATGGTGTATGGCAGGCGGAACCGGTGGACATGTATTCCCGGGACTGGCGGTTGCGCACCATCTAATGGCTCAGGGTTGGCAA 186
G K R L M V M A G G T G G H V F P G L A V A H H L M A Q G W Q
GTTGCGTGGCTGGGGACTGCCGACCGTATGGAAGCGGACTTGTGCAAAACATGGCATCGAAATTGATTTTCATTCGTATCTCTGGTCTGCGT 279
V R W L G T A D R M E A D L V P K H G I E I D F I R I S G L R
GGAAAAGGTATAAAGCACTGATAGTGCCTGCGTATCTTCAACGCTGGCGTCAGGCGCGGGCGATTATGAAAGCGTACAAAACCTGAC 372
G K G I K A L I A A P L R I F N A W R Q A R A I M K A Y K P D
GTGGTGCTCGGTATGGGAGGCTACGTGTGTCAGGTCCAGGTGGTCTGGCCGCGTGGTTCGTTAGGCATTCCGGTTGTACTTCATGAACAAAACGGT 465
V V L G M G G Y V S G P G G L A A W S L G I P V V L H E Q N G
ATTGCGGGCTTAACCAATAAATGGCTGGCGAAGATTGCCACCAAAAGTGTATGCAGGCGTTTCCAGGTGCTTCCCTAATGCGGAAGTAGTGGGT 558
I A G L T N K W L A K I A T K V M Q A F P G A F P N A E V V G
AACCCGGTGCATACCGATGTGTTGGCGCTGCCGTTGCCGAGCAACGTTTGGCTGGACGTGAAGGTCGGTTCGTGTGCTGGTAGTGGGTGGT 651
N P V R T D V L A L P L P Q Q R L A G R E G P V R V L V V G G
TCTCAGGGCGCACGATCTTAAACAGACAATGCCGACGTTGCTGCGAAACTGGGTGATTCAGTCACTATCTGGCATCAGAGCGGCAAAAGGT 744
S Q G A R I L N Q T M P Q V A A K L G D S V T I W H Q S G K G
TCGCAACAATCCGTTGAACAGGCGTATGCCGAAGCGGGGCAACCGCAGCATAAAGTGACGGAATTTATTGATGATATGGCGGCGGCGTATGCG 837
S Q Q S V E Q A Y A E A G Q P Q H K V T E F I D D M A A A Y A
TGGGCGGATGTCGTCGTTTGGCGCTCCGGTGCCTAACGGTGAAGTGAATCGCCGCGGAGGAGTACCGGCGTTTGTGTCGCTTCAACAT 930
W A D V V V C R S G A L T V S E I A A A G L P A L F V P F Q H
AAAGACCGCCAGCAATACTGGAATGCGCTACCGCTGGAATAAAGCGGGCGCAGCAAAATTCGAGCAGCCACAGCTTAGCGTGGATGCTGTC 1023
K D R Q Q Y W N A L P L E K A G A A K I I E Q P Q L S V D A V
GCCAACCCCTGGCGGTTGGTGGCGAGAAACCTTATTAACCATGGCAGAACCGCGCGCTGCATCCGATCCGACCCGAGCGAGT 1116
A N T L A G W S R E T L L T M A E R A R A A S I P D A T E R V
GCAAATGAAGTGAGCCGGTGGTCCCGGGCGTAATTGTAGCGATGCCTTTTGCATCGTATGAATTTAAGAAGTTAATGGCGTAAAGAATGAATA 1209
A N E V S R V A R A *
CACAAACAATTGGCAAACTGCGTCCATCGTGCCCGAAATGCGTTCGCGTTCGGCACATACATTTTGTGGCATTGGTGGTGGCGGATGGGCG 1302
GTATTGCCGAAGTTCTGGCCAATGAAGGTTATCAGATCAGTGGTTCCGATTTAGCGCCAAATCCGGTACGCGAGCAGTTAATGAATCTGGGTG 1395
CGACGATTTTCAACCATCGCCCGGAAAACGTACGTGATGCCAGCGTGGTTCGTTTCCAGCGCATTTCTGCGGATAACCCGGAATTTG 1488
TCGCGGCTCATGAAGCGGCTATTCCGGTGTATCCGTCGTCGCGAAATGCTGGCTGAGTTAATGCGTTTTTCGTCATGGCATCGCCATTGCCGGA 1581
CGCACGGCAAAACGACAAACCCGCGATGGTTCCAGCATCTACGCAGAAGCGGGGCTCGACCAACCTTCGTTAACGGCGGGCTGGTAAAAAG 1674
CGGCGGGGTTTATGCGCGTTTGGGGCATGGTCCGTACC 1713

```

KpnI