

Nucleotide sequence of the *murD* gene encoding the UDP-MurNAc-L-Ala-D-Glu synthetase of *Escherichia coli*

Dominique Mengin-Lecreux and Jean van Heijenoort

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

Submitted December 6, 1989

EMBL accession no. X17609

The *murD* gene coding for the UDP-MurNAc-L-Ala-D-Glu synthetase of *Escherichia coli* has been recently identified at 2-min on the chromosome, within a large cluster of genes from *phbB* to *envA* that code for proteins involved in peptidoglycan synthesis or cell division (1). Its nucleotide sequence has now been determined within the left part of a 2.6 kilobase-pair EcoRI-SalI chromosomal fragment (1) originating from pLC 26-6 of the Clarke and Carbon library. The coding region contains 1281 nucleotides which are translated into 427 amino acids corresponding to a protein with a molecular weight of 45893. This latter value is in agreement with that of the purified synthetase (2) or that of the protein expressed by the maxicell technique (1). The synthetase activity requires ATP hydrolysis

and, as previously observed with the *murF* gene product (3), a domain A characteristic of ATP binding proteins is found between amino acids 100 and 110. *murD* is preceded by the X region (1) coding for an unknown protein and is immediately followed by the recently sequenced *ftsW* gene (4).

REFERENCES

1. Mengin-Lecreux, D. *et al.* (1989) *J. Bacteriol.* **171**, 6126-6134.
2. Michaud, C. *et al.* (1987) *Eur. J. Biochem.* **166**, 631-637.
3. Parquet, C. *et al.* (1989) *Nucl. Acids Res.* **17**, 5379.
4. Ikeda, M. *et al.* (1989) *J. Bacteriol.* **171**,

```

-226 GAATTCCTGCTGGTGATTATGGGGGGCGTGTTCGTTGGTAGAAAACGCTTTCTGTGCATCCTGCGAGGTGCGGCTCCTTTAAACTGCGC
-142 GGACAACGTATTTTCCGCATGGCACCGATTTCATCACCACATGAACGAAAGGCTGGCCGAACCGCGCTCATTGTGGTTTCT
- 58 GGATTATTTCCGCTGATGCTGGTTCTGATTGGTCTGGCAACGCTGAAGGTACGTTAATCATGGCTGATTATCAGGGTAAAAATGT
                                     M A D Y Q G K N V
 26 CGTCATTATCGGCCCTGGCCCTCACCGGGCTTTCTGCGTGGACTTTTTCTCGCTGCCGGTGTGACGCGCGCGGTTATGGATAC
   V I I G L G L T G L S C V D F F L A A G V T P R V M D T
110 GCGTATGACACCGCTGGCCTGGATAAATTACCGAAGCCGTAGAACGCCACACGGGCAGTCTGAATGATGAATGGCTGATGGC
   R M T P P G L D K L P E A V E R H T G S L N D E W L M A
194 GGCAGATCTGATTGTGCGCAGTGTGCGGCTGATGCCGGAATCGAAATCGTTGGCGATATCGAGCTGTTCTGTGCGGAAGCACA
   A D L I V A S A A A D A G I E I V G D I E L F C R E A Q
278 AGCACCGATTGTGGCGATTACCGGTTCTAACGGCAAAAGCACGGTCACCACGCTAGTGGGTGAAATGGCGAAAGCGGGGGGT
   A P I V A I T G S N G K S T V T T L V G E M A K A A G V
362 TAACGTTGGTGGTGGCAATATTGGCCTGCCCTGCGTGTGCTACTGATGATGAGTGTGAAGTGTGACTGTGCTGGAAGTGTG
   N V G V G G N I G L P A L M L L D D E C E L Y V L E A L S
446 GAGCTTCCAGCTGGAACCACCTCCAGCTTACAGGCGGTAACGGCGACCATTCTGAACGCTGACTGAAGATCATATGGATCGCTA
   S F Q L E T T S S L Q A V T A T I L N V T E D H M D R Y
530 TCCGTTTGGTTTACAACAGTATCGTGCAGCAAAACTGCGCATTTACGAAACGCGAAAGTTTGGCTGGTTAATGCTGATGATGC
   P F G L Q Q Y R A A K L R I Y E N A K V C V V N A D D A
614 CTTAACAATGCCGATTTCGCGGTGCGGATGAACGCTGCGTCAGCTTTGGCGTCAACATGGGTGACTATCACCTGAATCATCAGCA
   L T M P I R G A D E R C V S F G V N M G G D Y H L N H Q Q
698 GGGCGAAACCTGGCTGCGGTTAAAGGCGAGAAAGTGC TGAATGTGAAAGAGATGAAACTTTCCGGGCAGCATAACTACACCAA
   G E T W L R V K G E K V L N V K E M K L S G Q H N Y T N
782 TGCGCTGGCGCGCTGGCGCTGGCAGATGCTGCAGGGTTACCGCGTGCCAGCAGCCTGAAAGCGTTAACCACATTCCTGGTCT
   A L A R V A L A D A A G L P R A S S L K A L T T F T G L
866 GCCGCATCGCTTTGAAGTTGTGCTGGAGCATAACGGCGTACGTTGGATTAACGATTCGAAAGCGACCAACGTCGGCAGTACGGA
   P H R F E V V L E H N G V R W I N D S K A T N V G S T E
950 AGCGGCGTGAATGGCCTGCACGTAGACGGCACACTGCAATTTGTTGCTGGGTGGCGATGGTAAATCGGGGACTTTAGCCCACT
   A A L N G L H V D G T L H L L L L G G D G K S A D F L S P L
1034 GGGCGTTACCTGAATGGCGATAACGTACGCTGTATTGTTTCCGGTCTGACGGCGCGCAGCTGGCGGCGCTACGCCGGAAGT
   A R Y L N G D N V R L Y C F G R D G A Q L A A L R P E V
1118 GGCAGAAAACCGAACTATGGAACAGGCGATGCGCTTGCTGGCTCCGCGTGTTCAGCCGGCGATATGGTTCTGCTCCTCCC
   A E Q T E T M E Q A M R L L A P R V Q P G D M V F L S P
1202 AGCCTGTGCCAGCCTTGATCAGTTCAAGAATTTGAACAACGAGGCAATGAGTTTGCCCGTCTGGCGAAGGAGTTAGGTTGATG
   A C A S L D Q F K N F E Q R G N E F A R L A K E L G *
1286 CGTTTATCTCTCCCTGCGCTGAAAATGCCGCGCTGCCAGGATTCAGTATCCTGGTCTGGATCTCCACGGCGCTAAAGGGCTGG
    
```