JOURNAL OF BACTERIOLOGY, Apr. 1992, p. 2391–2393
0021-9193/92/072391-03$02.00/0

Vol. 174, No. 7

Revised Nucleotide Sequence of the *gltP* Gene, Which Encodes the Proton-Glutamate-Aspartate Transport Protein of *Escherichia coli* K-12

B. TOLNER, B. POOLMAN, B. WALLACE,† AND W. N. KONINGS*

Department of Microbiology, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received 16 October 1991/Accepted 22 January 1992

The gene encoding the proton-glutamate carrier (GltP) of *Escherichia coli* K-12 was sequenced, and the primary structure of the protein was analyzed. The nucleotide sequence was found to differ in several aspects from the previously published sequence (B. Wallace, Y. Yang, J. Hong, and D. Lum, J. Bacteriol. 172:3214–3220, 1990). The corrected open reading frame encodes a protein of 437 (instead of 395) amino acids. Hydropathy analysis predicts 12 membrane-spanning α-helical regions. The complementary strand does contain an open reading frame possibly encoding a highly hydrophilic polypeptide of 272 amino acids.

In *Escherichia coli*, three L-glutamate transport systems have been identified: (i) a binding-protein-dependent, sodium-independent glutamate-aspartate system (inhibited by cysteate); (ii) a binding-protein-independent, sodium-independent glutamate-aspartate system (inhibited by β-hydroxyaspartate and cysteate), designated GltP; and (iii) a binding-protein-independent, sodium-dependent glutamate-specific system (inhibited by α-methylglutamate), designated GltS (6, 9, 11). Genes encoding the last two transport proteins have been cloned, and their nucleotide sequences have been reported (2, 3, 8, 15).

So far, two types of L-glutamate transport mechanisms have been identified in thermophilic bacteria. In *Bacillus stearothermophilus*, L-glutamate (or L-aspartate) transport proceeds via a sodium-proton symport mechanism with a 1:1:1 stoichiometry (4, 7). In *Clostridium fervidus*, an electrogenic sodium-glutamate symport mechanism with a stoichiometry of 2 has been identified (12). To further characterize the molecular properties of the sodium–proton–L-glutamate symport system of *B. stearothermophilus*, the gene encoding the transport protein was cloned and sequenced (14). Comparison of the primary structure of the sodium–proton–L-glutamate symport system (GltT) of *B. stearothermophilus* with the published sequence of the proton-glutamate carrier (GltP) of *E. coli* K-12 (15) revealed regions of homology, but other regions differed completely. By translating the nucleotide sequence of *gltP* in different reading frames and by comparing the translated sequences with that of GltT, it became apparent that the sequence divergence between GltP and GltT was likely to be caused by sequencing errors, i.e., base substitutions, deletions, and insertions, in the L-glutamate transport gene (*gltP*) of *E. coli*. To test this explanation, the *Mlu*I-*Nsi*I fragment of pBW8, containing the gene encoding the proton-glutamate carrier protein of *E. coli* K-12 as described by Wallace et al. (15), was resequenced. The *Mlu*I-*Nsi*I fragment of pBW8 was removed as an *Eco*RI-*Pst*I fragment and ligated into the corresponding sites of pUC18. The nucleotide sequences of both strands of the *Eco*RI-*Pst*I fragment or subclones of this fragment in pUC18 or M13mp18/19 (*Rsa*I, *Sau*3A, *Hpa*I, and *Hin*cII fragments) were determined (Fig. 1) by using the dideoxy-chain termination method (10). A T7 sequencing kit (Pharmacia) was used in sequencing either single- or double-stranded DNA, using T7 forward and pUC/M13 reverse synthetic oligonucleotide primers. The programs Micro Genie (release 5.0) and PC-Gene (release 6.26) were used for computer-assisted sequence analysis. The sequence of the *Mlu*I-*Nsi*I fragment is shown in Fig. 2.

A comparison of the sequence (Fig. 2) with the reported sequence of *gltP* (15) revealed 10 differences (indicated by roman numerals in Fig. 2): (i) at position 447, a G should be

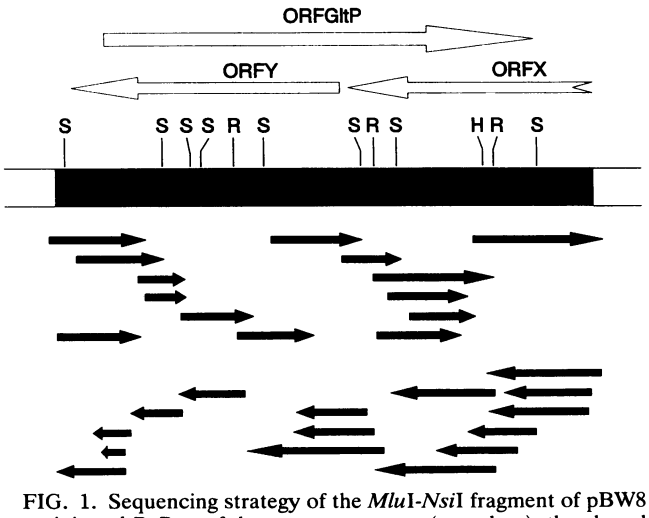

FIG. 1. Sequencing strategy of the *Mlu*I-*Nsi*I fragment of pBW8 containing *gltP*. Part of the vector sequence (open box), the cloned fragment (black box), the position and direction of transcription of *gltP* (large arrow above the sequence), and the positions and directions of transcription of putative ORFX and ORFY genes (small arrows above the sequence) are shown. The starts and extent of the data derived from individual sequencing reactions are shown by arrows below the sequence. S, *Sau*3A; R, *Rsa*I; H, *Hin*dIII.

* Corresponding author.
† Present address: School of Microbiology, The University of New South Wales, Kensington, New South Wales, 2033 Australia.

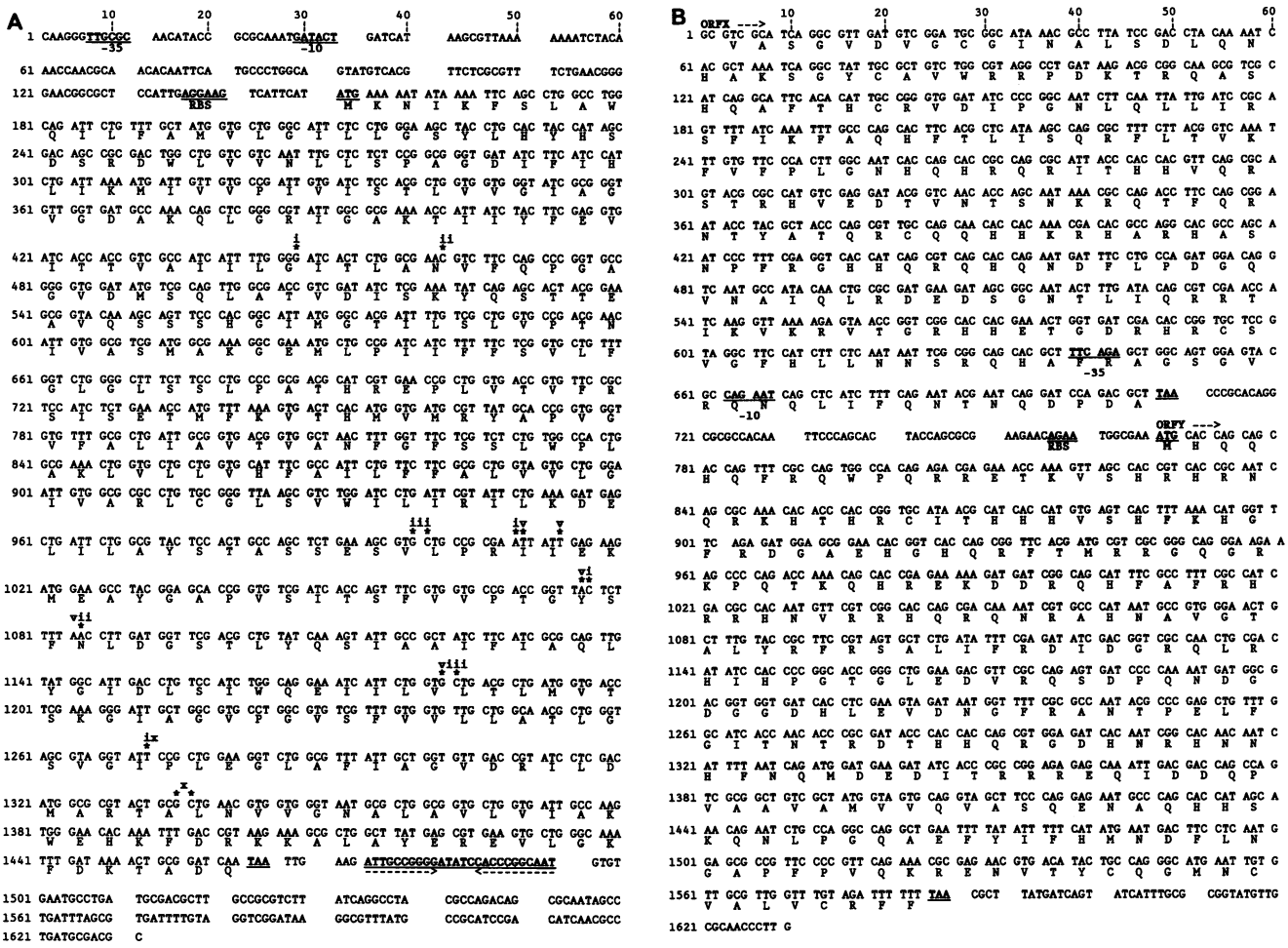

FIG. 2. Nucleotide sequence and flanking regions of the *gltP* gene encoding the proton-glutamate transport system of *E. coli* K-12 (A) and of the reverse strand on which ORFX and ORFY are located (B). The putative promotor regions (−35 and −10), the possible ribosome-binding site (RBS), and the terminator sequences (<-- and -->) are underlined. The translated amino acid sequences of GltP, ORFX, and ORFY are shown below the DNA sequence. The positions of corrected sequence elements are indicated by asterisks and roman numerals above the sequence.

inserted; (ii) at position 462, a C should be deleted; (iii) the CG at positions 999 to 1000 should be GC; (iv) the CA at positions 1009 to 1010 should be AT; (v) at position 1014, a T should be inserted; (vi) the AACC at positions 1076 to 1077 should be AC; (vii) at position 1085, an A should be inserted; (viii) the CG at positions 1182 to 1183 should be GC; (ix) at position 1272, a T should be inserted, and (x) the CG at positions 1335 to 1336 should be GC. The amino acid sequence deduced from the revised nucleotide sequence differs greatly from the published sequence because of the frameshifts (15). On the basis of the ribosome-binding site at positions 137 to 142 and the amino acid sequence homology between GltP and GltT (14), the translation initiation codon was assigned to the ATG at positions 151 to 153. Putative −35 and −10 promotor regions can be found upstream of *gltP*. The first stop codon (TAA at positions 1462 to 1464) is followed immediately by a putative rho-independent transcription terminator sequence with a ΔG° of −27.8 kcal (1 cal = 4.184 J)/mol according to the calculations of Tinoco et al. (13). The revised primary structure of GltP contains 437 (instead of 395 [15]) amino acids with a molecular mass of 47.2 kDa.

The carrier protein contains 66.3% nonpolar and 32.7% polar residues, an amino acid composition typical for membrane proteins (1). Of the 437 residues, 30 (6.9%) are basic and 28 (6.4%) are acidic, resulting in a basic protein with an excess of two positive charges at neutral pH. Hydropathy analyses by the method of Eisenberg et al. (5) predicts 12 putative membrane-spanning α-helical regions (Fig. 3).

Interestingly, the complementary strand contains two open reading frames, ORFX and ORFY. The start codon of ORFX is located upstream of the sequenced region, and the stop codon is located at position 708. ORFY is located between positions 768 and 1584 and could encode a highly hydrophilic polypeptide of 272 amino acids. ORFX and ORFY were not homologous to any of the proteins present in the SWISS-PROT protein sequence data bank (version 1.40).

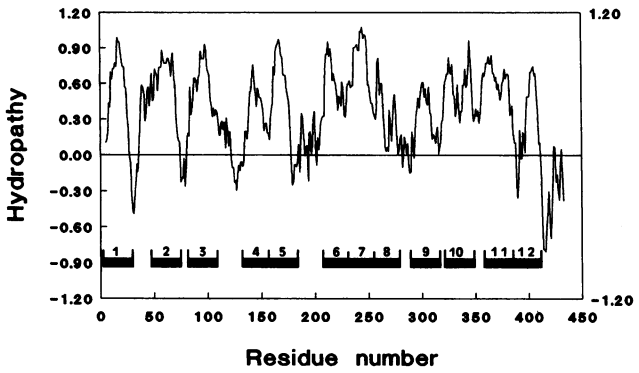

FIG. 3. Hydropathy profile of the amino acid sequence of the proton-glutamate of *E. coli* K-12. The hydropathy profile was calculated according to the method of Eisenberg et al. (5), with a window of 21 amino acids. Vertical bars and numbers indicate the positions of the 12 putative membrane-spanning segments.

Nucleotide sequence accession number. The GenBank accession number for the sequence of the *gltP* gene is M84805.

REFERENCES

1. **Büchel, D. E., B. Gronenborn, and B. Müller-Hill.** 1980. Sequence of the lactose permease gene. Nature (London) **283**:541–545.
2. **Deguchi, Y., I. Yamato, and Y. Anraku.** 1989. Molecular cloning of *gltS* and *gltP*, which encode glutamate carriers of *Escherichia coli* B. J. Bacteriol. **171**:1314–1319.
3. **Deguchi, Y., I. Yamato, and Y. Anraku.** 1990. Nucleotide sequence of *gltS*, the Na^+/glutamate symport carrier gene of *Escherichia coli* B. J. Biol. Chem. **265**:21704–21708.
4. **de Vrij, W., R. A. Bulthuis, P. R. van Iwaarden, and W. N. Konings.** 1989. Mechanism of L-glutamate transport in membrane vesicles from *Bacillus stearothermophilus*. J. Bacteriol. **171**:1118–1125.
5. **Eisenberg, D., E. Schwarz, M. Komarony, and R. Wall.** 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. **179**:125–142.
6. **Halpern, Y. S., H. Barash, S. Dover, and K. Druck.** 1973. Sodium and potassium requirements for active transport of glutamate by *Escherichia coli* K-12. J. Bacteriol. **114**:53–58.
7. **Heyne, R. I. R., W. de Vrij, W. Crielaard, and W. N. Konings.** 1991. Sodium ion-dependent amino acid transport in membrane vesicles of *Bacillus stearothermophilus*. J. Bacteriol. **173**:791–800.
8. **Kalman, M., D. R. Gentry, and M. Cashel.** 1991. Characterization of the *Escherichia coli* K12 *gltS* glutamate permease gene. Mol. Gen. Genet. **225**:379–386.
9. **Miner, K. M., and L. Frank.** 1974. Sodium-stimulated glutamate transport in osmotically shocked cells and membrane vesicles of *Escherichia coli*. J. Bacteriol. **117**:1093–1098.
10. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
11. **Schellenberg, G. D., and C. E. Furlong.** 1977. Resolution of the multiplicity of the glutamate and aspartate transport systems of *Escherichia coli*. J. Biol. Chem. **252**:9055–9064.
12. **Speelmans et al.** Unpublished data.
13. **Tinoco, I., P. N. Borer, B. Dengler, M. K. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) **246**:40–41.
14. **Tolner, B., B. Poolman, and W. N. Konings.** Unpublished data.
15. **Wallace, B., Y. Yang, J. Hong, and D. Lum.** 1990. Cloning and sequencing of a gene encoding a glutamate and aspartate carrier of *Escherichia coli* K-12. J. Bacteriol. **172**:3214–3220.