JOURNAL OF BACTERIOLOGY, Oct. 1982, p. 506–509
0021-9193/82/100506-04$02.00/0

Vol. 152, No. 1

Sequence Analysis of the Heat-Labile Enterotoxin Subunit B Gene Originating in Human Enterotoxigenic *Escherichia coli*

TATSUO YAMAMOTO,[1]* TAKA-AKI TAMURA,[2] MASARU RYOJI,[3] AKIRA KAJI,[3] TAKESHI YOKOTA,[1] AND TOSHIYA TAKANO[2]

Department of Bacteriology, School of Medicine, Juntendo University, 2-1-1 Hongo, Bunkyo-ku, Tokyo, Japan[1]; Department of Microbiology, School of Medicine, Keio University, Shinjuku-ku, Tokyo[2]; and Department of Microbiology, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104[3]

Received 26 March 1982/Accepted 11 June 1982

In this study, we determined the amino-terminal coding sequence, covering the signal peptide and the amino-terminus of the mature peptide, of the heat-labile enterotoxin subunit B (LT-B) gene originating in human enterotoxigenic *Escherichia coli*. Neither the signal sequence nor the amino-terminal sequence of the mature LT-B was identical to those sequences from porcine enterotoxigenic *E. coli*, but there was an extensive homology.

Enterotoxigenic *Escherichia coli* produce enterotoxins, such as heat-labile toxin (LT) and heat-stable toxin (ST), that cause diarrheal diseases in humans and food animals (13). LT and *Vibrio cholerae* enterotoxin (cholera toxin) are similar toxins. They share similar modes of action (5) and similar subunit structures (4, 15), have common antigenic determinants (1, 18), and manifest extensive homology of amino acid sequence (2, 14).

The LT operon is polycistronic, having at least two genes for LT subunits A and B (LT-A and LT-B); the latter is distal to the former (3, 20). Modification of this LT operon by in vitro recombination techniques allowed us to construct a series of recombinant plasmids producing toxoids of LT (LT-B and LT-A*), which we hope are useful as vaccines (20). These techniques also allowed us to study the transcomplementation of the LT genes (18, 20) and the secretion mechanism of the LT subunits through bacterial membrane(s) (19).

Although LT operons from human and porcine enterotoxigenic *E. coli* resemble each other in many aspects (3, 20), recent studies demonstrated considerable heterogeneity between the LT gene products of human and porcine enterotoxigenic *E. coli*. (i) LT-B produced from a human enterotoxigenic *E. coli* is apparently larger in molecular weight than is the cholera toxin subunit B (18, 20). In contrast, LT-B from porcine enterotoxigenic *E. coli* is similar in molecular weight to cholera toxin subunit B (11). (ii) There exist immunologically nonidentical determinant(s) between LTs from human and porcine enterotoxigenic *E. coli* (7). Because LT genes might have been widely disseminated among bacterial replicons as transposons (18), we have been interested in the evolutionary relationships of the LT genes from human and porcine enterotoxigenic *E. coli*.

In previous experiments, the LT operon (*toxA* and *toxB*) originating in human enterotoxigenic *E. coli* strain H10407 was cloned into pBR322 to construct an LT-producing recombinant plasmid pJYL2299 (17). By using pJYL2299, a peptide with a molecular weight of 11,800 to 12,000 was identified as a major *toxB* product in *E. coli* minicells (18, 20). This 11,800- to 12,000-molecular-weight LT-B from pJYL2299, which was labeled with ^{14}C-amino acid mixture in *E. coli* minicells, was purified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, eluted from the gel, and subjected to the Edman degradation (Fig. 1). Five cycles of Edman degradation revealed that the amino acid sequence at the amino-terminus of the

$$\text{LT-B is } \overset{1}{\text{A}}\text{la-}\overset{2}{\text{P}}\text{ro-}\overset{3}{\text{G}}\text{ln-}(\overset{4}{-})\text{-}\overset{5}{\text{I}}\text{le(or leu)----;}$$

the fourth amino acid residue remained ambiguous. This sequence is similar to that

$$(\overset{1}{\text{A}}\text{la-}\overset{2}{\text{P}}\text{ro-}\overset{3}{\text{G}}\text{ln-}\overset{4}{\text{T}}\text{hr-}\overset{5}{\text{I}}\text{le----)}$$

of the LT-B of porcine enterotoxigenic *E. coli* (1).

The structure of the LT operon (approximately 1 kilobase in length) originating in human enterotoxigenic *E. coli* is summarized in Fig. 2. Of this LT operon, the *Eco*RI-*Hin*fI region (166 nucleotides) encoding part of the *toxB* gene was sequenced by the chemical method of Maxam and Gilbert (10), and the first 102 nucleotides and the corresponding amino acid sequence

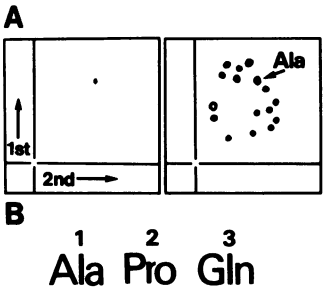

FIG. 1. Amino-terminal analysis of the LT subunit B by the Edman degradation. (A) The LT-B was synthesized in *E. coli* minicells carrying pJYL2299 in the presence of ^{14}C-labeled amino acid mixture as described (18) except the labeling conditions; in this experiment, 19 ^{14}C-labeled amino acids (L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, glycine, L-histidine, L-isoleucine, L-leucine, L-lysine, L-phenylalanine, L-proline, L-serine, L-threonine, L-tyrosine, L-valine, L-methionine, L-tryptophan, L-glutamine, and L-asparagine) at approximately 0.6 μCi per ml each were used. The minicells were lysed, and the lysates were subjected to 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (18). The peptide bands were detected by autoradiography, and the peptide with a molecular weight of 11,800 to 12,000 was eluted from the gel. Subsequent procedures were carried out as previously described (12; M. Ryoji and A. Kaji, Anal. Biochem., in press). Phenylthiohydantoin amino acids obtained from each cycle of the Edman degradation were separated by two-dimensional chromatography on a polyamide sheet together with phenylthiohydantoin amino acid standards as internal markers. The ^{14}C-labeled amino-termini were detected by autoradiography (7- to 30-day exposure) as shown at the left of this figure. The markers were detected under the UV lamp as shown at the right of the figure. The figure shows the result of the 1st cycle of the Edman degradation. (B) The amino-termini from the first three cycles of the Edman degradation.

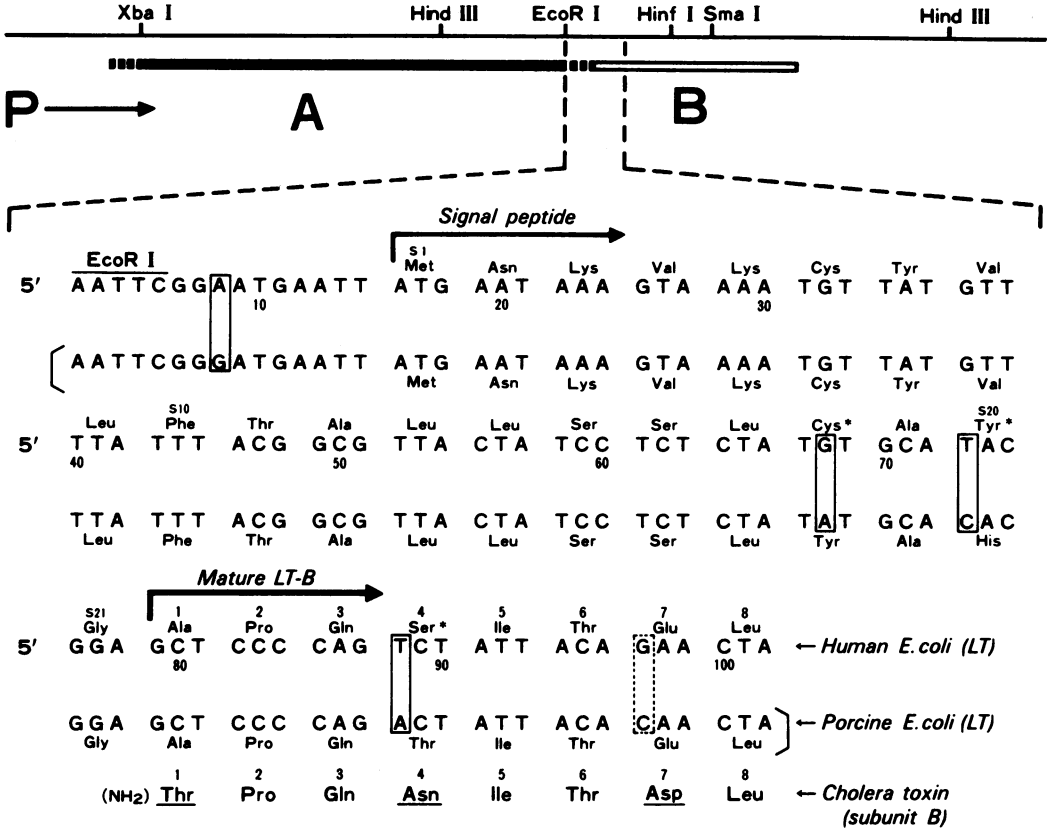

FIG. 2. Analysis of the amino-terminal sequence of the LT-subunit B gene. The top line shows the LT operon (*toxA,B*) map (18, 20). P, A, and B respectively represent the LT promoter, *toxA* (filled bar), which encodes LT subunit A, and *toxB* (open bar), which encodes LT subunit B. The location of the carboxy-terminus of the *toxA* is described (18, 20). The arrow indicates the direction of transcription. The *Eco*RI-*Hin*fI region (166 nucleotides) was sequenced by the method of Maxam and Gilbert (10) as described previously (16), and the first 102 nucleotides are shown in this figure (lower part). The nucleotide sequence (and amino acid sequence) of the corresponding region originating in porcine enterotoxigenic *E. coli* is from Dallas and Falkow (2). Base changes between human and porcine enterotoxigenic *E. coli* are marked with a box. A base change, marked with a dotted box, remains ambiguous (see the text). Asterisks represent amino acid residues that differ from those of porcine enterotoxigenic *E. coli*. The amino-terminal amino acid sequence of cholera toxin subunit B is from Klapper et al. (8) and Kurosky et al. (9). The underlined amino acid residues of cholera toxin subunit B differ from those of the LT-B sequences.

translated from the nucleotide sequence are shown in Fig. 2. The initiation codon for the *toxB* product (precursor, pre-LT-B) is located 15 nucleotides from the site cut by *Eco*RI. A processing site within the precursor (pre-LT-B) to convert into the mature LT-B was determined to be 78 nucleotides away from the *Eco*RI site; the signal sequence of the *toxB* is 63 nucleotides in length. This conclusion is based on the amino-terminal amino acid sequence obtained with the LT-B in Fig. 1. The signal peptide (21 amino acid residues) has a calculated molecular weight of 2,308 and is enriched in hydrophobic amino acid residues (~70%); the hydrophobicity is required to pass through membranes (6). In a previous communication, we showed that even the LT-B alone, which was produced from the cloned *toxB*, can be secreted into the culture supernatant through the bacterial membrane(s) (19) and that the 11,800- to 12,000-molecular-weight peptide (the mature LT-B) is a major product of the *toxB* gene in *E. coli* minicells (18, 20). This signal peptide probably facilitates the secretion of the LT-B through the bacterial inner membrane.

When this signal sequence of the *toxB* of human enterotoxigenic *E. coli* is compared with that of porcine enterotoxigenic *E. coli*, two base changes are found. Consequently, two amino acids are replaced. This comparison shows 97% homology at the nucleotide level and 90% homology at the amino acid level. An extensive homology was also found in the amino-terminal sequence encoding the mature LT-B. So far, only one base change was identified at nucleotide 88 (Fig. 2). This results in one amino acid replacement. It should be noted that this residue, the fourth amino acid from the amino-terminus of the mature LT-B, differs from that of porcine enterotoxigenic *E. coli* and from that of cholera toxin subunit B (Fig. 2). Another nucleotide change is located at nucleotide 97. However, this does not appear to change the amino acid assigned, Glu, to the corresponding codon, as has been reported (2). No changes were seen in the additional 64 nucleotides up to the *Hin*fI endonuclease site, between human and porcine enterotoxigenic *E. coli*. The sequenced amino-terminal region of the mature LT-B totals 88 nucleotides, which corresponds to approximately 30% of the entire mature LT-B-encoding sequence.

Based on the product analysis of the cloned *tox* regions, we have pointed out the possibility that the termination codon for the LT-A (*toxA* product) may be located downstream (right side in Fig. 2) of the *Eco*RI site (18, 20). If so, the termination codon for LT-A would be at one of three possible sites: nucleotides 21 to 23 (TAA; ocher), nucleotides 17 to 19 (TGA, opal), or nucleotides 10 to 12 (TGA, opal). The first two possibilities imply that the two LT genes *toxA* and *toxB* overlap.

We thank Y. Takeda for encouragement, T. Arai for helpful suggestions, and R. Nozawa for reading the manuscript.

This work was supported by the Scientific Research Fund of the Ministry of Education, Science and Culture of Japan.

LITERATURE CITED

1. **Clements, J. D., R. J. Yancey, and R. A. Finkelstein.** 1980. Properties of homogeneous heat-labile enterotoxin from *Escherichia coli*. Infect. Immun. **29**:91–97.
2. **Dallas, W. S., and S. Falkow.** 1980. Amino acid sequence homology between cholera toxin and *Escherichia coli* heat-labile toxin. Nature (London) **288**:499–501.
3. **Dallas, W. S., D. M. Gill, and S. Falkow.** 1979. Cistrons encoding *Escherichia coli* heat-labile toxin. J. Bacteriol. **139**:850–858.
4. **Gill, D. M., J. D. Clements, D. C. Robertson, and R. A. Finkelstein.** 1981. Subunit number and arrangement in *Escherichia coli* heat-labile exterotoxin. Infect. Immun. **33**:677–682.
5. **Gill, D. M., and S. H. Richardson.** 1980. Adenosine diphosphate-ribosylation of adenylate cyclase catalyzed by heat-labile enterotoxin of *Escherichia coli*: comparison with cholera toxin. J. Infect. Dis. **141**:64–70.
6. **Heijne, G. V.** 1980. Trans-membrane translocation of proteins. A detailed physico-chemical analysis. Eur. J. Biochem. **103**:431–438.
7. **Honda, T., T. Tsuji, Y. Takeda, and T. Miwatani.** 1981. Immunological nonidentity of heat-labile enterotoxins from human and porcine enterotoxigenic *Escherichia coli*. Infect. Immun. **34**:337–340.
8. **Klapper, D. G., R. A. Finkelstein, and J. D. Capra.** 1976. Subunit structure and N-terminal amino acid sequence of the three chains of cholera enterotoxin. Immunochemistry **13**:605–611.
9. **Kurosky, A., D. E. Markel, J. W. Peterson, and W. M. Fitch.** 1977. Primary structure of cholera toxin β-chain: a glycoprotein hormone analog? Science **195**:299–301.
10. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
11. **Palva, E. T., T. R. Hirst, S. J. S. Hardy, J. Holmgren, and L. Randall.** 1981. Synthesis of a precursor to the B subunit of heat-labile enterotoxin in *Escherichia coli*. J. Bacteriol. **146**:325–330.
12. **Ryoji, M., R. Berland, and A. Kaji.** 1981. Reinitiation of translation from the triplet next to the amber termination codon in the absence of ribosome-releasing factor. Proc. Natl. Acad. Sci. U.S.A. **78**:5973–5977.
13. **Smith, H. W., and C. L. Gyles.** 1970. The relationships between two apparently different enterotoxins produced by enteropathogenic strains of *Escherichia coli* of porcine origin. J. Med. Microbiol. **3**:387–401.
14. **Spicer, E. K., W. M. Kavanaugh, W. S. Dallas, S. Falkow, W. H. Konigsberg, and D. E. Schafer.** 1981. Sequence homologies between A subunit of *Escherichia coli* and *Vibrio cholerae* enterotoxins. Proc. Natl. Acad. Sci. U.S.A. **78**:50–54.
15. **Takeda, Y., T. Honda, S. Taga, and T. Miwatani.** 1981. In vitro formation of hybrid toxins between subunits of *Escherichia coli* heat-labile enterotoxin and those of cholera enterotoxin. Infect. Immun. **34**:341–346.
16. **Tamura, T., M. Noda, and T. Takano.** 1981. Structure of the baboon endogenous virus genome: nucleotide sequences of the long terminal repeat. Nucleic Acids Res. **9**:6615–6626.
17. **Yamamoto, T., and T. Yokota.** 1980. Cloning of deoxyribonucleic acid regions encoding a heat-labile and heat-

stable enterotoxin originating from an enterotoxigenic *Escherichia coli* strain of human origin. J. Bacteriol. **143**:652–660.

18. **Yamamoto, T., and T. Yokota.** 1981. *Escherichia coli* heat-labile enterotoxin genes are flanked by repeated deoxyribonucleic acid sequences. J. Bacteriol. **145**:850–860.
19. **Yamamoto, T., and T. Yokota.** 1982. Release of heat-labile enterotoxin subunits in *Escherichia coli*. J. Bacteriol. **150**:1482–1484.
20. **Yamamoto, T., T. Yokota, and A. Kaji.** 1981. Molecular organization of heat-labile enterotoxin genes originating in *Escherichia coli* of human origin and construction of heat-labile toxoid-producing strains. J. Bacteriol. **148**:983–987.