Cloning, Nucleotide Sequence, and Hybridization Studies of the Type IIb Heat-Labile Enterotoxin Gene of *Escherichia coli*

CAROL L. PICKETT,† EDDA M. TWIDDY, CHRISTOPHER COKER, AND RANDALL K. HOLMES*

Department of Microbiology, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814

Received 22 March 1988/Accepted 3 June 1989

Type IIb heat-labile enterotoxin (LT-IIb) is produced by *Escherichia coli* 41. Restriction fragments of total cell DNA from strain 41 were cloned into a cosmid vector, and one cosmid clone that encoded LT-IIb was identified. The genes for LT-IIb were subcloned into a variety of plasmids, expressed in minicells, sequenced, and compared with the structural genes for other members of the *Vibrio cholerae-E. coli* enterotoxin family. The A subunits of these toxins all have similar ADP-ribosyltransferase activity. The A genes of LT-IIa and LT-IIb exhibited 71% DNA sequence homology with each other and 55 to 57% homology with the A genes of cholera toxin (CT) and the type I enterotoxins of *E. coli* (LT-I and LTp-I). The A subunits of the heat-labile enterotoxins also have limited homology with other ADP-ribosylating toxins, including pertussis toxin, diphtheria toxin, and *Pseudomonas aeruginosa* exotoxin A. The B subunits of LT-IIa and LT-IIb differ from each other and from type I enterotoxins in their carbohydrate-binding specificities. The B genes of LT-IIa and LT-IIb were 66% homologous, but neither had significant homology with the B genes of CT, LT-I, and LTp-I. The A subunit genes for the type I and type II enterotoxins represent distinct branches of an evolutionary tree, and the divergence between the A subunit genes of LT-IIa and LT-IIb is greater than that between CT and LT-I. In contrast, it has not yet been possible to demonstrate an evolutionary relationship between the B subunits of type I and type II heat-labile enterotoxins. Hybridization studies with DNA from independently isolated LT-II-producing strains of *E. coli* also suggested that additional variants of LT-II exist.

More than a decade ago, the heat-labile enterotoxins of *Escherichia coli* and *Vibrio cholerae* were recognized to be a family of related protein toxins with similarities in structure, mode of action, and immunochemistry (12, 13). They are oligomeric proteins composed of A and B polypeptides, and they activate adenylate cyclase in susceptible eucaryotic target cells. The A1 fragment, derived by proteolytic cleavage and reduction of the A polypeptide, catalyzes ADP-ribosylation of the Gs regulatory subunit of adenylate cyclase, resulting in stimulation of the cyclase activity (15, 16, 31, 46). The B subunit mediates binding of toxin to the plasma membrane, and ganglioside GMI is a functional receptor for toxin binding (46).

Recently, new heat-labile enterotoxins were discovered (17, 19, 21), and the *V. cholerae-E. coli* enterotoxin family was divided into two distinct antigenic groups (36). Cholera toxin (CT) (12) and the type I *E. coli* heat-labile enterotoxins (LT-I), including the antigenic variants LT-I and LTp-I (22), belong to serotype I, and antisera to any one of them will neutralize the other type I toxins. In contrast, type II *E. coli* enterotoxins (LT-II) are not neutralized by antisera against type I toxins, but they are neutralized by antisera to the prototype LT-II (19, 21, 36). Two antigenic variants of type II enterotoxin, designated LT-IIa and LT-IIb, were characterized (19, 21). Both consist of A and B subunits, which are similar in size to the subunits of CT and LT-I. The ADP-ribosyltransferase activity of these LT-II toxins is similar to that of CT and LT-I (8; P. P. Chang, S.-C. Tsai, R. Adamik, B. C. Kunz, J. Moss, E. M. Twiddy, and R. K. Holmes, submitted for publication), but the ganglioside-binding specificities of LT-IIa and LT-IIb are different from those of CT and LT-I and from each other (14, 19, 21).

We recently reported the nucleotide sequence of the LT-IIa genes (37). The A gene of LT-IIa exhibited 55 to 57% sequence homology with the A genes of CT and LT-I (29, 37, 49). Most of this homology was derived from the region of the A gene that encodes the enzymatically active A1 fragment (15, 16, 37). The B gene of LT-IIa showed no apparent sequence homology to the B genes of CT and LT-I (29, 37, 49). We cloned and sequenced the genes for LT-IIb from *E. coli* 41, a strain isolated from food in Brazil (18). In addition, we performed hybridization experiments using probes for the A and B genes of LT-IIb to study their relatedness with the toxin structural genes in other LT-II-producing strains of *E. coli*.

**MATERIALS AND METHODS**

Bacterial strains, bacteriophage, and plasmids, *E. coli* 41, the LT-IIb-producing reference strain, is described by Guth et al. (18). The other LT-II-producing strains isolated from food or humans in Brazil, 34/4, IAL190, 538, TR442/2, and 0-4, were also characterized by Guth et al. (18). *E. coli* SA53, which produces prototype LT-IIa, is described by Green et al. (17), Pickett et al. (36), and Holmes et al. (21). Four other LT-II-producing *E. coli* strains, SA76, SA100, SA31, and SA35, all isolated from animals in Thailand, are described by Pickett et al. (37). The *E. coli* minicell-producing strain P678-54, is described by Adler et al. (1). *E. coli* JM103 was from D. Ward (30), and the phage M13mp18 and M13mp19 were purchased from New England BioLabs, Inc. (Beverly, Mass.) (33). The cosmid pH79 is described by Hohn and Collins (20). The plasmids pBR322, pBR328, and pUC18 have been described previously (5, 43, 50). The plasmid pCP706, which contains the intact LT-IIa genes, is described by Pickett et al. (36).

**Media**. Bacterial strains tested for LT-IIb toxin production were grown in glucose-sucrose medium (17), and assays for toxin were performed with cultured mouse Y1 adrenal cells.
as described by Pickett et al. (36). For all other experiments, bacteria were grown in L medium (42).

DNA procedures. Total bacterial DNA was isolated as described by Silhavy et al. (42). Sau3A-generated DNA fragments, to be inserted into the cosmid vector, were isolated from 5 to 20% sucrose gradients (36). Plasmid DNA was isolated and purified by the Ish-Horowicz modification (26) of the alkaline extraction procedure of Birnboim and Doly (4). Restriction endonucleases, T4 DNA ligase, and lambda packaging kits were used according to supplier specifications (Bethesda Research Laboratories, Inc., Gaithersburg, Md., and Boehringer Mannheim Biochemicals, Indianapolis, Ind.). Transformations were performed with CaCl₂-treated cells as described by Cohen et al. (9). Agarose gel electrophoresis was performed by standard procedures (26), and specific DNA fragments were isolated from agarose gels with DEAE-cellulose paper as described previously (11, 37).

Minicell experiments. Minicell experiments were performed by the method of Meagher et al. (27) as described by Pickett et al. (36).

Hybridizations. DNA was transferred from agarose gels to nitrocellulose paper by the method of Southern (44). The blots were then used in hybridization experiments exactly as described by Maniatis et al. (26) except that the blots were washed at 55°C to allow approximately 30% base pair mismatch (28).

Nucleotide sequence analysis. Restriction fragments carrying portions of the LT-I lb genes were cloned into replicative-form M13 vectors (39). Sequencing of both strands of the LT-I lb genes was performed using [γ-³²P]dATP (3) in the dideoxy-chain termination method of Sanger et al. (40). Synthetic oligonucleotides were made on a DNA synthesizer (model 380A; Applied Biosystems, Inc., Foster City, Calif.) and used unpurified in the same manner as the universal primer. MicroGenie sequence software (Beckman Instruments, Inc., Palo Alto, Calif.) was used to analyze and compare the sequence data. Local hydropathy values were calculated by the method of Hopp and Woods (23) with a hexapeptide window, using the MicroGenie sequence software.

RESULTS

Cloning of LT-I lb toxin genes. DNA was isolated from E. coli 41 and partially digested with Sau3A, and fragments of 30 to 50 kilobase pairs (kb) were isolated from sucrose gradients. These fragments were ligated to BamHI-cut, alkaline phosphatase-treated pH79 DNA. The ligated DNA was packaged into lambda phage particles which were transduced into E. coli HB101. Ampicillin-resistant transductants were screened for toxin production in the cultured mouse Y1 adrenal cell assay. One toxin-positive transductant was found from the screening of 335 transductants. This toxin-producing clone contained a plasmid, named pCP4127, with a 45-kb insert. When pCP4127 was cut with various restriction endonucleases, a single EcoRV fragment of about 2.4 kilobase pairs (kb) hybridized with a DNA probe (36) carrying the LT-I lb genes. EcoRV fragments from pCP4127 were therefore cloned into the EcoRV site of pBR328, and a resultant toxin-positive subclone was found to carry the 2.4-kb EcoRV fragment. A restriction map of the EcoRV insert in this subclone, pCP4185, is shown in Fig. 1.

Toxin assays of E. coli 41, HB101(pCP4127), and HB101(pCP4185) indicated that the strain carrying the cosmid clone produced no more toxin than the parental strain. E. coli HB101(pCP4185), however, produced about 100 times more toxin than strain 41. Transcription of the LT-I lb operon of pCP4185 was opposite in orientation to transcription from the promoter for the tetracycline resistance gene in pBR328; therefore, the tetracycline promoter was not responsible for the increased expression of the LT-I lb genes in pCP4185.

Several subclones of pCP4185 were generated by standard procedures. These were tested for toxin activity, and the results indicated that the LT-I lb genes were located between the BglII and Hpal sites (Fig. 1). Minicell analysis of pCP4185 and selected subclones suggested that the PsI site lies within the gene encoding the A polypeptide and that the Nhel site lies within the gene encoding the B polypeptide (Fig. 2).

Nucleotide sequence of LT-I lb. Three overlapping DNA fragments derived from pCP4185, the 1.0-kb XhoI-Psil fragment, the 1.3-kb BglII-Nhel fragment, and the 1.1-kb HindIII-Psil fragment, were cloned into M13mp18 and M13mp19. The strategy used to sequence the LT-I lb genes with these clones is shown in Fig. 1. The nucleotide sequence of the LT-I lb genes is shown in Fig. 3 and compared with data reported previously for the LT-I la and LT-I b genes.

The overall nucleotide sequence homology of the A genes of LT-I lb and LT-I a was 71%. The overall homology of the LT-I lb and LT-I a A gene sequences was 55%; slightly less than the 57% overall homology between the LT-I la and LT-I a A gene sequences (37). The proposed leader peptide for the A gene of LT-I lb is two amino acids longer than the leader peptide for the A subunit of LT-I la, LT-I, or CT. The leader sequences of the LT-I a, LT-I lb, and LT-I b A polypeptides showed considerable variation, but the proposed Ala-Asn cleavage site is conserved. The calculated molecular weight for the mature A polypeptide of LT-I lb is 27,199. Treatment of LT-I lb with trypsin cleaves the mature A polypeptide into the disulfide-linked fragments A1 and A2 (19), similar in size to the A1 and A2 fragments of LT-I a, LT-I, and CT. The two cysteine residues at positions 185 and 197, which flank the proposed cleavage site for the formation of A1 and A2, are also conserved in all these toxins. A possible trypsin cleavage site for the formation of the A1 and A2 fragments exists following the lysine at position 190. The
The nucleotide sequence homology between the B genes of LT-IIb and LT-IIa was 66% (Fig. 3). The B gene of LT-IIb, like that of LT-IIa (37), showed no significant sequence homology to the B gene of LT-I or CT. The amino acid sequences of LT-IIb and LT-IIa were 58% homologous (67% if conservative changes were included). The leader peptides for the B subunits of LT-IIa and LT-IIb are 23 amino acids long, which is 2 amino acids longer than the B polypeptide leaders of CT and LTh-I (49). These cleavage sites were confirmed by amino-terminal sequencing of purified B subunits of LT-IIa and LT-IIb (R. K. Holmes and E. M. Twiddy, unpublished data), and the cleavage site for LT-IIa is one of several possible sites deduced from the sequence for the B gene of LT-IIa (35, 37). The two cysteine residues at positions 10 and 81 in the B polypeptides of LT-IIb and LT-IIa are conserved. CT, LTh-I, and Ltp-I also have two cysteines, at positions 9 and 86. The positions of the two tryptophan residues at positions 47 and 92 in the B polypeptides of LT-IIb and LT-IIa are also conserved, but these are not at the same location as the sole tryptophan residue at position 88 in the CT, LTh-I, and Ltp-I B polypeptides (29, 49).

Hydropathy plots of the A and B subunits of LT-IIb were generated from the predicted amino acid sequences. The A subunit was very much like that of the A subunit of LT-IIa, except for some slight differences in the A2 regions (Fig. 4). The B subunit plots of LT-IIa and LT-IIb were quite similar despite having only 58% amino acid homology (Fig. 4).

Hybridization studies. Two nonoverlapping probes from LT-IIb gene sequences were used to examine the relatedness of LT-IIb to other LT-II toxin-coding sequences. The A gene probe was the 0.3-kb PstI-AccI fragment which contains only A gene sequences (Fig. 1). The B gene probe was the 0.5-kb HindIII-NheI fragment which contains the A2 region and most of the B gene (Fig. 1). These two fragments were used as probes in Southern hybridization experiments with digests of total cell DNA from several LT-II-producing strains (Fig. 5). The A gene probe hybridized under conditions of moderate stringency to a single restriction fragment in each of these strains. The B gene probe only hybridized to DNA from strain 41, the parental LT-IIb-producing strain. These results, together with the hybridization results reported by Pickett et al. (37), suggest that considerable heterogeneity exists among the B genes from independently isolated LT-II-producing strains.

DISCUSSION

The LT-IIb-producing strain E. coli 41 was isolated from cooked meat in Brazil. Initial experiments showed that LT-IIb, like LT-IIa, was encoded on chromosomal genes and that LT-IIa gene probes hybridized with less efficiency to DNA from strain 41 than to DNA from other LT-II-producing strains in our collection (18). LT-IIb was purified and found to have several properties that distinguished it from LT-IIa (19). It had only partial antigenic identity with LT-IIa, and its isoelectric point of approximately 5.4 was lower than the value of 6.8 for LT-IIa. LT-IIb had lower specific toxicity against Y1 adrenal cells and was activated to a greater extent by treatment with trypsin; the toxic dose of LT-IIb was 94 pg before treatment and 3 pg after treatment with trypsin, versus 0.5 pg before or after trypsin treatment for LT-IIa. LT-IIa and LT-IIb also differed from each other and from CT or LTh-I with respect to their specificity to bind to the oligosaccharide moieties of particular gangliosides (14, 19).
The LT-IIb genes from *E. coli* 41 were cloned and sequenced. The LT-IIb genes were organized in a manner similar to that seen for LT-IIa, LTh-I, and Ltp-I. The A and B genes of both LT-IIa and LT-IIb overlap by 11 base pairs, and possible Shine-Dalgarno sequences are present just upstream of the A and B genes of both LT-IIa and LT-IIb (37, 41). The A1 regions of the LT-IIa and LT-IIb genes are 79% homologous, which is not surprising considering that both toxins have similar enzymatic activities.

We compared specific regions of the LT-IIa and LT-IIb A1 fragments with the catalytic regions of other ADP-ribosylating toxins. Recent work has shown that diphtheria toxin and *Pseudomonas aeruginosa* exotoxin A (ETA) have significant homology with respect to the amino acid residues that form the NAD-binding cleft in the crystal structure for ETA (7). Furthermore, a histidine, a glutamate, and a tryptophan residue are conserved at nearly identical distances from each other in CT, diphtheria toxin, ETA, and the S1 subunit of pertussis toxin (S1), and LTh-I contains an aspartate residue instead of a glutamate residue at the corresponding position (47). In diphtheria toxin and ETA, these glutamate residues (Glu-148 and Glu-553, respectively) are required for ADP-ribosylating activity and are believed to be part of the active site (10, 45). Substitution of Asp for Glu at these positions in diphtheria toxin and ETA by site-directed mutagenesis lowered ADP-ribosylating activity.

---

**FIG. 3.** Comparison of nucleotide sequences of the LT-IIb, LT-IIa, and LTh-I genes. The nucleotides are numbered above the sequences, with the adenylate residue of the initiating ATG codon for the A gene of LT-IIb indicated as 1. Translation for the A and B genes of LT-IIb is proposed to start at positions 1 and 782 and end at positions 792 and 1150, respectively. Amino acid residues are numbered below the sequence, with the first amino acid of each mature subunit polypeptide designated as 1, and the last amino acid of the leader sequence of each subunit designated as −1. The sequence of the A gene is compared with the sequence of the LT-IIa A gene and the LTh-I A gene (37, 49). Nucleotides and amino acids of LT-IIa that differ from those in LT-IIb are shown immediately above and below the LT-IIb sequences, respectively. Nucleotides and amino acids in LTh-I that differ from those in LT-IIb are shown on the second line above and below the LT-IIb sequences, respectively. The B gene of LT-IIb is compared with the sequence of the LT-IIa B gene in a similar manner. The LTh-I B gene sequence is not included since no significant homology could be found between the B genes of LTh-I and LT-IIa or LT-IIb. The proposed leader polypeptide for the A subunit of LT-IIb is two amino acids longer than the leader polypeptides of LT-IIa and LTh-I; the locations of the two absent codons, indicated by dashed lines at positions −16 and −15, were assigned to maximize amino acid homology. Dashed lines also indicate two other regions where the lengths of the LT-IIa, LT-IIb, and LTh-I A genes do not agree. Probable Shine-Dalgarno (41) sequences are indicated by solid lines above the nucleotide sequence at positions −10 and 772. Probable promoter sequences (labeled −10 and −35) are indicated by solid lines below the nucleotide sequence upstream of the A gene (38). The ATG corresponding to the first codon of the B gene is overlined and underlined in the A gene sequence.

**FIG. 4.** Hydropathy plots of the A and B polypeptides of LT-IIa and LT-IIb. Amino acid position 1 corresponds to the first amino acid of the mature subunit polypeptide. Positive values indicate hydrophobicity; negative values indicate hydrophilicity.
of the mutant toxin to less than 1% of the wild-type level (10, 45).

The corresponding His and Trp residues are conserved in LT-IIa and LT-IIb at positions 42 and 172, respectively, but the Glu residue is not conserved. The amino acid substituted for Glu at position 168 was Asn in LT-IIa and Gly in LT-IIb. Neither of these is a conservative change, and it is unlikely that position 168 is an active site residue in LT-IIa and LT-IIb. There are, however, conserved Gln and Trp residues in LT-IIa, LT-IIb, and LTH-I at positions 174 and 177, respectively, in Fig. 3, and this Glu residue is a possible active site residue. The Arg residue at position 144 in Fig. 3 is conserved in the type I and type II enterotoxins and is the site for self-ADP-ribosylation in the type I enterotoxins. Substitution of Gly for Arg at the corresponding position 146 in LTP-I eliminated the acceptor site for self-ADP-ribosylation but did not reduce the activity of the toxin in the vascular permeability assay (34). Hence, this Arg does not appear to be an active site residue.

Homologies exist between the N-terminal regions of the S1 subunit of pertussis toxin and the type I heat-labile enterotoxins and may be part of the NAD-binding sites of these toxins (25, 32). These regions are also conserved in LT-IIa and LT-IIb and correspond to residues 4 to 11 and 58 to 65 in Fig. 3. Comparison of the sequences for the type I and type II enterotoxins reveals that the residues at positions 10 and 62 in Fig. 3 do not need to be conserved to retain ADP-ribosylating activity. Residue 10 is Thr in the type II enterotoxins and Pro in the type I enterotoxins. Residue 62 is Thr in LT-IIb, Val in LT-IIa, and Leu or Ile in the type I enterotoxins. Site-directed mutations affecting Arg-9, Asp-11, and Arg-13 in the S1 subunit of pertussis toxin caused decreased ADP-ribosylating activity (2, 6). The homologous residues are conserved in the heat-labile enterotoxins and correspond with residues 5, 7, and 9 in Fig. 3.

The A2 regions of the heat-labile enterotoxins were much less homologous than their A1 subunits. The conservation of the two cysteine residues at positions 185 and 195 in Fig. 3 and the trypsin-susceptible site between them may reflect the necessity to cleave the A subunit into the A1 and A2 fragments to achieve full activation of the toxin. Whether the differences in the A2 regions of LT-IIa, LT-IIb, and LT-I reflect the requirement for the A subunits to associate with significantly different B subunits to form their respective holotoxins is unknown.

Neither the LT-IIa or LT-IIb B genes showed significant nucleotide sequence homology to the LT-I or CT B genes. The B genes of LT-IIa and LT-IIb were themselves only 66% homologous. The nucleotide substitutions between the B genes of LT-IIa and LT-IIb were widely and quite evenly distributed throughout the genes. The nonconservative amino acid substitutions, however, appeared to be clustered near the two cysteine residues at positions 10 and 81. The hydropathy plots of the LT-IIa and LT-IIb B subunits also showed their greatest differences in these areas (roughly between amino acid positions 5 and 20 and between amino acid positions 75 and 90). Whether the substitutions within these regions are responsible for the different ganglioside affinities of LT-IIa and LT-IIb is not yet known.

The G+C content of the A gene of LT-IIb is 38%, close to the 40% G+C content of LT-IIa and the same as the G+C content of the LTH-I A gene (48). These values are low compared with the average G+C content of 50% for E. coli sequences. The frequency of usage of optimal codons in E. coli for the A genes of LT-IIb and LTH-I is 0.38, and for the A gene of LT-IIa it is 0.39 (24, 37, 48). These frequencies are also low compared with those of other E. coli genes and may indicate that the genes for LT-I and LT-II originated in some bacterial species other than E. coli (24, 26). Comparisons of the sequence homologies among the A genes of LT-IIa, LT-IIb, and LT-I, and CT suggest that the LT-I and LT-II enterotoxin groups diverged from a common ancestor before the divergence of CT and LT-I. Indeed, since the LT-IIa and LT-IIb A genes have only 71% sequence homology compared with the 78% sequence homology between the A genes of CT and LT-I, the divergence of the LT-IIa and LT-IIb genes may also have occurred prior to the divergence of the CT and LT-I sequences.

The B genes of LT-IIa and LT-IIb did not exhibit significant sequence homology to the B genes of LT-I and CT. The LT-IIa and LT-IIb B genes, in common with the A genes, have a lower G+C content (39 and 41%, respectively) than typical E. coli sequences. The frequency of usage of optimal codons in E. coli for the B genes of LT-IIa and LT-IIb is 0.54 and 0.58, respectively, which is higher than the 0.44 value for the LTH-I B gene (37, 48). These values suggest that the LT-II B genes have evolved a codon usage more favorable than that of the A genes for expression in E. coli.

The hybridization experiments with the LT-IIb A and B gene probes confirm and extend the results obtained with
probes derived from LT-IIa genes (17, 36, 37). The LT-IIB A gene probe hybridized to one restriction fragment in each LT-II-producing strain examined, although the efficiency of hybridization varied greatly from strain to strain (Fig. 5). The LT-IIB B gene probe, however, hybridized only to DNA obtained from strain 41, the parent LT-IIB-producing strain. This result suggests that there is at least 30% base pair mismatch between the LT-IIB B gene and the B genes of the other LT-II-producing strains tested. We reported earlier that a B gene probe derived from the LT-IIB B gene also failed to hybridize with DNA from LT-II-producing strains other than the parental strain (37). Apparently, the B genes for type II toxins from independently isolated LT-II-producing strains of E. coli exhibit considerable nucleotide sequence heterogeneity. Whether these differences in the B gene sequences will reflect additional differences in toxin receptor specificities is unknown.

ACKNOWLEDGMENTS

We thank M. Jackson and M. Flora for advice with the sequencing and R. McNally for assistance with the toxin assays.

The work reported here was supported in part by Public Health Service grant SR22 AI14107 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED


