JOURNAL OF BACTERIOLOGY, Sept. 1990, p. 5490–5493
0021-9193/90/095490-04$02.00/0

Vol. 172, No. 9

The Heat-Stable Toxin I Gene from *Escherichia coli* 18D

WALTER S. DALLAS

Department of Molecular Genetics and Microbiology, The Wellcome Research Laboratories, Burroughs Wellcome Co., Research Triangle Park, North Carolina 27709

Received 20 February 1990/Accepted 14 June 1990

The heat-stable toxin I gene in the human *Escherichia coli* isolate 18D is the *estA1* allele. The gene is not part of a composite transposon, but inspection of the flanking DNA sequences suggests that it was at one time part of a transposon. The hypothetical transposon originated from an event other than the occurrence that formed Tn*1681*.

Enterotoxigenic *Escherichia coli* that cause diarrheal disease in some young animals and in humans of all ages express adherence factors and toxins (5, 6, 19, 23). Different *E. coli* enterotoxins were first identified by heat sensitivity. The more heat-susceptible toxin was described as the heat-labile toxin, or LT, and the more heat-stable toxin was called ST (22). The toxin genes have always been found on plasmids (8, 24). Since the initial description of an ST, this toxin class has now been shown to consist of two different groups of peptides. The two groups are distinguishable by biological assays, DNA sequences of the genes, and amino acid sequences of the toxins and are designated STI and STII (1, 15, 18, 27, 30). Both classes of ST genes can exist as parts of different composite transposons that are fully proficient in moving from one locus to a site on another replicon (14, 26). The two ST gene classes are designated *estA* (STI) and *estB* (STII). The nucleotide sequences of four *estA* alleles have been determined; the base sequences may vary as much as 30%, and the predicted amino acid sequences may vary by 38% (7, 16, 27, 29).

Chan and Giannella were the first to report the amino acid sequence for STI (which was later revised) and showed the peptide to be 18 amino acids in length with a remarkable cysteine content of six residues (2, 30). The toxin was purified from *E. coli* 18D, which had been isolated in Kentucky from an infant with diarrhea. The amino acid sequence of the peptide matched the carboxy-terminal portion of a 72-amino-acid protein predicted from the nucleotide sequence of an STI gene (*estA1*) from a calf isolate (27). These results suggested that STI is made as a precursor with 54 amino acids at the amino-terminal end that must be processed to release the mature toxin. Results from in vitro coupled transcription-translation experiments have been consistent with the existence of an STI precursor (13). The STI biological activity has been shown to reside in a chemically synthesized octodecameric peptide, thereby verifying the 18-amino-acid peptide as the toxin moiety (12). Although features of several STI alleles have been published, no genetic information about the STI gene from strain 18D (from which STI was first purified) has been reported. The STI gene from this strain is described in this report.

The first experiments were designed to identify the replicon carrying the ST gene. Plasmids were isolated from *E. coli* 18D (provided by R. Giannella) by alkaline lysis followed by density gradient centrifugation (3). Examination of the plasmids by electrophoresis in 0.7% agarose gels (15-cm vertical gels electrophoresed at 120 V for 4 h in Tris acetate buffer) indicated the presence of at least two large plasmids. A size estimate for each plasmid was made by using two plasmids (157 and 64 kilobase pairs [kb]) whose sizes had been assessed by reconstruction from overlapping cosmid clones (data not shown). The larger plasmid was estimated to be 145 kb, and the smaller plasmid was estimated to be 100 kb. The STI gene was shown to be associated with the larger plasmid in a Southern transfer hybridization experiment with a mixed-site oligonucleotide probe (Fig. 1) (9, 28). The enterotoxin plasmid was designated pST18.

To clone the STI gene, plasmids from strain 18D were first cut with restriction enzymes that have single sites within the *lacZ* portion of the cloning vector pKAN1 (a derivative of pBS+ in which the ampicillin resistance gene had been replaced with a kanamycin resistance gene) (21). DNA fragments were separated on 1% agarose gels, and Southern transfers to nitrocellulose were made. Hybridization with the mixed-site oligonucleotide probe revealed the size of the DNA fragment containing the STI gene for each digest (data not shown). *Kpn*I was the enzyme chosen for STI gene isolation, since cutting the plasmids with this enzyme yielded a modest number of DNA fragments (about 15) and since the fragment (6.5 kb) that hybridized to the ST gene probe was smaller than 10 kb. pKAN1 and the plasmids from strain 18D were cut with *Kpn*I, ligated (10 μg of pKAN1 and 50 μg of target DNA per ml), and transformed into DH5α (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) (3). Transformants with a high probability of inserts were recognized as white colonies on Luria agar plates with 50 μg of kanamycin per ml and 50 μg of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside per ml and were screened by colony hybridization (31, 32). A hybridization-positive transformant was identified. The recombinant plasmid, pST181, was shown to be composed of pKAN1 and a 6.5-kb *Kpn*I DNA fragment. Restriction enzyme mapping, Southern hybridizations, and DNA sequencing (see below) placed the STI gene 1,400 base pairs (bp) from one end of the cloned fragment (Fig. 2). (The location of a unique *Mlu*I site in IS*1* made the placement of *estA1* unambiguous [see below].)

The *estA1* allele can be part of a composite transposon (Tn*1681*) in which the toxin gene is flanked by inverted repeats of the insertion sequence IS*1* (26). The nucleotide sequence of IS*1* has been determined, and the element is 768 bp (17). The arrangement of four restriction enzyme recognition sites (*Aat*II, *Fsp*I, *Mlu*I, and *Pst*I), each present once in IS*1*, defines the location of the element. The positions of these recognition sequences were determined in pST181, and a single grouping of the sites consistent with the pres-

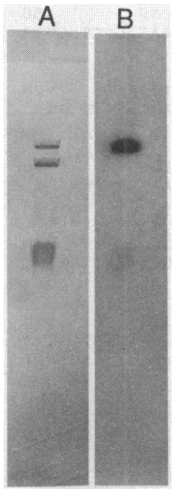

FIG. 1. Agarose gel (A) and Southern transfer hybridization autoradiogram (B) of plasmids from strain 18D. The plasmids were hybridized to a mixed-site STI oligonucleotide probe (GCTGT GAA[T,C]T[T,G]TGTTGTAATCC) (9). The amorphous band in the middle of the gel is fragmented plasmid and chromosome.

ence of IS*1* was found. The grouping of sites was within 300 bp of the STI gene (Fig. 2).

By using the published sequence of *estA1* (from Tn*1681*), oligonucleotide sequencing primers were made, and the sequence for both strands of the gene and flanking DNA was determined by the chain termination method (20). The STI gene in strain 18D was found to be the *estA1* allele. The absolute agreement in nucleotide sequences between STI genes isolated from both a human and an animal strain is in conspicuous contrast to the dissemination of the heat-labile toxin gene, in which a human isolate allele can be distinguished from its porcine counterpart (32). On the 5′ side of the STI gene, a complete IS*1* sequence was identified 294 bp from the initiation codon proposed by So and McCarthy (27). In Tn*1681*, 218 bp separate the two loci (Fig. 3). The contiguous 218 bp of 5′ flanking sequences were the same for the two gene isolates. IS*1* in pST181 was opposite in orientation when compared to the 5′ IS*1* element in Tn*1681*. In addition, the IS*1* element in pST181 was found to be bordered on the opposite side by another insertion sequence, IS*30*, starting at bp 537 and extending at least to bp 460 (the extent of IS*30* sequence present was not determined) (4). In Tn*1681*, another copy of IS*1* begins 88 bp from the termination codon of *estA1*. The DNA sequence 3′ to the termination codon of *estA1* in pST181 was extended 154 bp, but no sequences homologous to IS*1* or IS*30* were found. The contiguous 88 bp at the 3′ end of *estA1* were identical in the two gene isolates (Fig. 3).

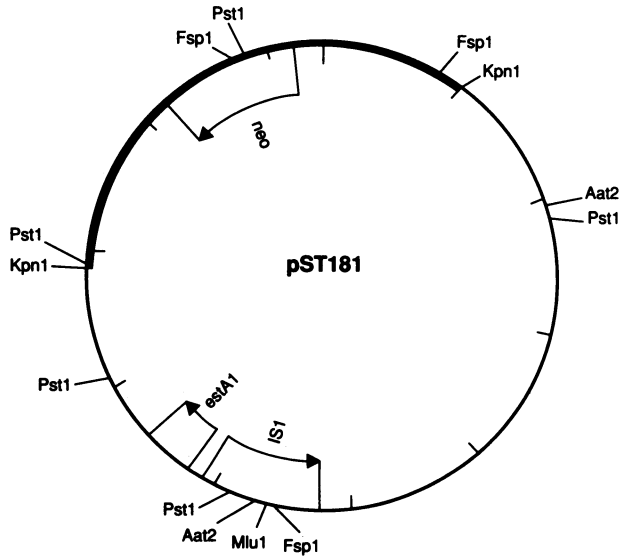

FIG. 2. Schematic diagram of pST181. The heavy line denotes the vector portion of the plasmid (3,368 bp). Arrows show direction of gene transcription (*estA1*, *neo*) or the numbering convention for IS*1* (base furthest from arrowhead is nucleotide 1). The STI allele on pST181 was determined to be *estA1* by DNA sequencing. Each tick on the inside of the circle represents 950 bp.

So et al. showed that *estA1* could be part of a transposon, which they called Tn*1681* (26). The ability of the STI gene to transpose explained its location on a variety of replicons (25). It should be noted that the exceptionally high AT content (70%) of the STI gene suggests that it was probably imported into *E. coli* (which has an AT content of 50%) from another bacterium. Several features of the nucleotide sequences bordering *estA1* from strain 18D suggest that the toxin gene was at one time part of a composite transposon different from Tn*1681*. First, the bias for AT nucleotides is also a feature of the bases 5′ to both *estA1* copies up to but not including the IS*1* elements. IS*1* has an AT content of 47% (17). The additional 75 bp before the start of the IS*1* sequence in pST181 also have a high AT content (68%). On the termination codon end of *estA1* from pST181, the same 88 bp are present, as in Tn*1681*, and this sequence has an AT content of 70.4%. The sequence of an additional 66 bp has been determined. For the next 12 bases, an AT nucleotide bias is apparent but the AT tendency soon disappears as the sequence is extended. It is conceivable that the flanking AT-rich sequences surrounding *estA1* were part of the original DNA that was transferred into *E. coli*. The STI gene on pST18 has additional AT-rich flanking sequences compared with those sequences bordering *estA1* in Tn*1681*. Second, the orientation of IS*1* in pST181 is opposite to the 5′ IS*1* element in Tn*1681*. Finally, when IS*1* transposes as a 768-bp element, a short region at the site of insertion is duplicated (10). The nucleotides surrounding IS*1* in pST181 are not duplicated.

Iida, Meyer, and Arber have demonstrated in a stepwise manner the formation of an IS*1* composite transposon and have presented examples of decay products from these mobile structures (11). With consideration of this work, a possible model for the origin of *estA1* on pST18 is presented. A DNA segment that included *estA1* was transferred into *E. coli* from an organism that had DNA with a high AT content. The ST gene was captured by IS*1* elements in at least two separate events. One event formed Tn*1681*. The second capture event included more of the original DNA flanking *estA1*. The composite transposon from the second event eventually became situated, through transposition, in an IS*30* element on pST18. The transposon then decayed by a deletion event which removed one IS*1* sequence (on the 3′ side of *estA1*) along with part of the IS*30* sequence.

The speculation on the origin of *estA1* on pST18 presented here is interesting in two respects. First, it is another demonstration of the genetic flexibility for bacterial virulence determinants. These genes have been found on bacteriophages and plasmids as well as on the chromosome. Tn*1681* is an example of a transposon that includes a toxin

```
                                    10         20         30         40         50
IS30------IS1-------ATTGATGTTTTACAGACATAAAAGTATCAATTTATATAAGAGCCTAGTCA

       60         70         80         90        100        110        120
TCCTACCATCCTGCACTGGAAAATACGAACAAGAAAAGGATAAAACCATCACTTAATATGATATTGATAG

      130        140        150        160        170        180        190
TAAAAAAAACCAGATAGCCAGACAATTAACATATATAACACTATGAAAATAATATATAAAAAGCGAGTGT

      200        210        220        230        240        250        260
ACCTCGACATATAACATGATGCAACTCACAAAAAAAAAATAAAAAAATTGCAAAATCCGTTTAACTAATCT

      270        280        290        300        310        320        330
CAAATATCCGTGAAACAACATGACGGGAGGTAACATGAAAAAGCTAATGTTGGCAATTTTTATTTCTGTA
                                    MetLysLysLeuMetLeuAlaIlePheIleSerVal

      340        350        360        370        380        390        400
TTATCTTTCCCCTCTTTTAGTCAGTCAACTGAATCACTTGACTCTTCAAAAGAGAAAATTACATTAGAGA
LeuSerPheProSerPheSerGlnSerThrGluSerLeuAspSerSerLysGluLysIleThrLeuGluT

      410        420        430        440        450        460        470
CTAAAAAGTGTGATGTTGTAAAAAACAACAGTGAAAAAAAATCAGAAAATATGAACAACACATTTTACTG
hrLysLysCysAspValValLysAsnAsnSerGluLysLysSerGluAsnMetAsnAsnThrPheTyrCy

      480        490        500        510        520        530        540
CTGTGAACTTTGTTGTAATCCTGCCTGTGCTGGATGTTATTAAAAAGCATAGAGGGAATCTTTATTTTGA
sCysGluLeuCysCysAsnProAlaCysAlaGlyCysTyr

      550        560        570        580        590        600        610
TTCCCTCTATATTATATGTATTGGTGCTGAGAAAATCATAATAAAAAACAACGCTTTTATATTATATGTA

      620        630        640        650        660
AGTGCCTCATCTGAACCGTCTGGTCTGTTTCCTCCGGCTCCACAAAAATAATGTCCA
```

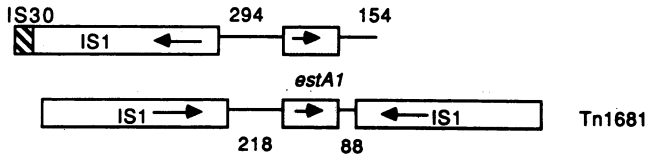

FIG. 3. Nucleotide sequence of *estA1* and flanking regions from pST18. IS*30* sequences are not shown but extend from at least bp 460 to bp 537 (4). IS*1* sequences are not shown but extend from bp 768 to bp 1 (17). The relative position of the 72-residue precursor form of STI is shown below the nucleotide sequence (27). Arrows indicate the position of IS*1* in Tn*1681*. Below the sequence is a schematic diagram depicting the arrangement of sequences flanking *estA1* for pST18 and Tn*1681*. The numbers in the schematic are base pairs and represent distances between boxed regions.

gene (26). Second, it is probable that the STI gene did not originate in *E. coli* but was recruited from another source at least two separate times. This suggests that the genesis of a particular bacterial pathogen through the acquisition of genes that are determinants of virulence need not be a one-time event, but, rather, appropriate circumstances for pathogen formation may occur more than once and pathogen formation is probably an ongoing process.

LITERATURE CITED

1. **Burgess, M. N., R. J. Bywater, C. M. Cowley, N. A. Mullan, and P. M. Newsome.** 1978. Biological evaluation of a methanol-soluble, heat-stable *Escherichia coli* enterotoxin in infant mice, pigs, rabbits, and calves. Infect. Immun. **21**:526–531.
2. **Chan, S., and R. A. Giannella.** 1981. Amino acid sequence of heat-stable enterotoxin produced by *Escherichia coli* pathogenic for man. J. Biol. Chem. **256**:7744–7746.
3. **Dallas, W. S.** 1983. Conformity between heat-labile toxin genes from human and porcine enterotoxigenic *Escherichia coli*. Infect. Immun. **40**:647–652.
4. **Dalrymple, B., P. Caspers, and W. Arber.** 1984. Nucleotide sequence of the prokaryotic mobile genetic element IS30. EMBO J. **3**:2145–2149.
5. **Evans, D. G., R. P. Silver, D. J. Evans, Jr., D. G. Chase, and S. L. Gorbach.** 1975. Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. Infect. Immun. **12**:656–667.
6. **Gorbach, S. L., and C. M. Khurana.** 1972. Toxigenic *Escherichia coli*: a cause of infantile diarrhea in Chicago. N. Engl. J. Med. **287**:791–795.
7. **Guzman-Verduzco, L.-M., and Y. M. Kupersztoch.** 1989. Rectification of two *Escherichia coli* heat-stable enterotoxin allele sequences and lack of biological effect of changing the carboxy-terminal tyrosine to histidine. Infect. Immun. **57**:645–648.
8. **Gyles, C., M. So, and S. Falkow.** 1974. The enterotoxin plasmids in *Escherichia coli*. J. Infect. Dis. **130**:40–49.
9. **Hill, W. E., W. L. Payne, G. Zon, and S. L. Moseley.** 1985. Synthetic oligodeoxyribonucleotide probes for detecting heat-stable enterotoxin-producing *Escherichia coli* by DNA colony hybridization. Appl. Environ. Microbiol. **50**:1187–1191.
10. **Iida, S., R. Hiestand-Nauer, and W. Arber.** 1985. Transposable element IS1 intrinsically generates target duplications of variable length. Proc. Natl. Acad. Sci. USA **82**:839–843.
11. **Iida, S., J. Meyer, and W. Arber.** 1980. Genesis and natural history of IS-mediated transposons. Cold Spring Harbor Symp. Quant. Biol. **45**:27–43.
12. **Klipstein, F. A., R. F. Engert, and R. A. Houghten.** 1983. Properties of synthetically produced *Escherichia coli* heat-stable enterotoxin. Infect. Immun. **39**:117–121.
13. **Lathe, R., P. Hirth, M. DeWilde, N. Harford, and J.-P. Lecocq.** 1980. Cell-free synthesis of enterotoxin of *E. coli* from a cloned

gene. Nature (London) **284**:473–474.
14. **Lee, C.-H., S.-T. Hu, P. J. Swiatek, S. L. Moseley, S. D. Allen, and M. So.** 1985. Isolation of a novel transposon which carries the *Escherichia coli* enterotoxin STII gene. J. Bacteriol. **162:** 615–620.
15. **Lee, C. H., S. L. Moseley, H. W. Moon, S. C. Whipp, C. L. Gyles, and M. So.** 1983. Characterization of the gene encoding heat-stable toxin II and preliminary molecular epidemiological studies of enterotoxigenic *Escherichia coli* heat-stable toxin II producers. Infect. Immun. **42**:264–268.
16. **Moseley, S. L., J. W. Hardy, M. I. Huq, P. Echeverria, and S. Falkow.** 1983. Isolation and nucleotide sequence determination of a gene encoding a heat-stable enterotoxin of *Escherichia coli*. Infect. Immun. **39**:1167–1174.
17. **Ohtsubo, H., and E. Ohtsubo.** 1978. Nucleotide sequence of an insertion element, IS1. Proc. Natl. Acad. Sci. USA **75**:615–619.
18. **Picken, R. N., A. J. Mazaitis, W. K. Maas, M. Rey, and H. Heyneker.** 1983. Nucleotide sequence of the gene for heat-stable enterotoxin II of *Escherichia coli*. Infect. Immun. **42**:269–275.
19. **Sack, R. B., S. L. Gorbach, J. G. Banwell, B. Jacobs, B. D. Chatterjee, and R. C. Mitra.** 1971. Enterotoxigenic *Escherichia coli* isolated from patients with severe cholera-like disease. J. Infect. Dis. **123**:378–385.
20. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
21. **Short, J. M., J. M. Fernandez, J. A. Sorge, and W. D. Huse.** 1988. Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. Nucleic Acids Res. **16**:7583–7600.
22. **Smith, H. W., and C. L. Gyles.** 1970. The relationship between two apparently different enterotoxins produced by enteropathogenic strains of *Escherichia coli* of porcine origin. J. Med. Microbiol. **3**:387–401.
23. **Smith, H. W., and S. Halls.** 1967. Studies on *Escherichia coli* enterotoxin. J. Pathol. Bacteriol. **93**:531–543.
24. **Smith, H. W., and S. Halls.** 1968. The transmissible nature of the genetic factor in *Escherichia coli* that controls enterotoxin production. J. Gen. Microbiol. **52**:319–334.
25. **So, M., H. W. Boyer, M. Betlach, and S. Falkow.** 1976. Molecular cloning of an *Escherichia coli* plasmid determinant that encodes for the production of heat-stable enterotoxin. J. Bacteriol. **128**:463–472.
26. **So, M., F. Heffron, and B. J. McCarthy.** 1979. The *E. coli* gene encoding heat stable toxin is a bacterial transposon flanked by inverted repeats of IS1. Nature (London) **277**:453–456.
27. **So, M., and B. J. McCarthy.** 1980. Nucleotide sequence of bacterial transposon Tn1681 encoding a heat-stable (ST) toxin and its identification in enterotoxigenic *Escherichia coli* strains. Proc. Natl. Acad. Sci. USA **77**:4011–4015.
28. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
29. **Stieglitz, H., L. Cervantes, R. Robledo, R. Fonseca, L. Covarrubias, F. Bolivar, and Y. M. Kupersztoch.** 1988. Cloning, sequencing, and expression in ficoll-generated minicells of an *Escherichia coli* heat-stable enterotoxin gene. Plasmid **20**:42–53.
30. **Thompson, M. R., and R. A. Giannella.** 1985. Revised amino acid sequence for a heat-stable enterotoxin produced by an *Escherichia coli* strain (18D) that is pathogenic for humans. Infect. Immun. **47**:834–836.
31. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
32. **Vinal, A. C., and W. S. Dallas.** 1987. Partition of heat-labile-enterotoxin genes between human and animal *Escherichia coli* isolates. Infect. Immun. **55**:1329–1331.