J. Microbiol. Biotechnol. (2000), **10**(1). 86–90

JOURNAL OF
MICROBIOLOGY
AND
BIOTECHNOLOGY

Cloning, Sequencing, and Characterization of Enterotoxin Pathogenicity Islet from *Bacteroides fragilis* 419

RHIE, GI-EUN, GYUNG TAE CHUNG, YONG-JIN LEE, WON KEUN SUNG, AND HEE-BOK OH*

Laboratory of Bacterial Toxins, National Institute of Health, 5 Nokbun-Dong, Eunpyung-Gu, Seoul 122-701, Korea

Received: June 30, 1999

Abstract We have earlier reported on the cloning and identification of *bft-k* from an enterotoxigenic strain of *Bacteroides fragilis* 419, which was isolated from the blood of a Korean patient who suffered from systemic infections [4, 5]. The *bft-k* gene encodes a 397-amino-acids metalloprotease enterotoxin, and the protein has been identified as a new isoform of *B. fragilis* enterotoxins (BFTs), which are cytopathic to intestinal epithelial cells to induce fluid secretion and tissue damage in ligated intestinal loops [4, 6, 18, 20]. This report describes the cloning and sequencing of the enterotoxin pathogenicity islet of *B. fragilis* 419 which contains the *bft-k* gene. The cloned enterotoxin pathogenicity islet was found to have 6.045 bp in length and to contain 12-bp direct repeats near its end. In the pathogenicity islet, in addition to the BFT-K, two putative open reading frames (ORFs) were identified; (1) the *t-3* gene encoding a 396-amino-acids protein of a putative metalloprotease; (2) the third gene encoding an ORF of a 59-amino-acids protein, whose function has not yet been characterized. The expression of the *t-3* gene in *B. fragilis* 419 was verified by western blot analysis.

Key words: *Bacteroides fragilis* 419, enterotoxin pathogenicity islet, BFT-K, T-3, nucleotide sequence

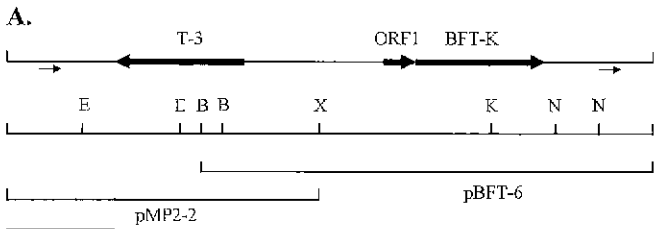
Bacteroides fragilis is a gram-negative nonsporulating anaerobic rod that inhabits the large bowel of humans and animals, and constitutes about 1 to 2% of the normal human colonic microflora [15]. Several strains of *B. fragilis* have been implicated as a cause of diarrhea in some animals such as lambs, calves, and humans [3, 17, 21], and the strains are termed enterotoxigenic *B. fragilis* (ETBF). These strains are reported to stimulate secretory response in a lamb ligated intestinal segment [23], and this secretory response has been attributed to a 20-kDa enterotoxin protein which exhibits biological activities such as a loss of cell-cell attachments, rounding, and swelling in the human epithelial cell line HT29/C_1 *in vitro* [6, 16, 18, 20, 23]. Enterotoxins have been shown to act through the cleavage of the zonula adherens protein, E-cadherin [4, 24].

Recently, two genes designated as *bft-1* and *bft-2*, encoding two BFT isoforms, were cloned from lamb-isolated *B. fragilis* VPI 13784 and piglet-isolated *B. fragilis* 86-5443-2-2, respectively [7, 12]. Sequence analyses of the cloned genes revealed a zinc-binding consensus motif (HEXXHXXGXXH) characteristic of metalloproteases and metzincins [1, 7, 12, 22], and purified BFT has been shown to hydrolyze G (monomeric) actin, gelatin, and azocoll *in vitro* [14].

In recent years, evidence has been accumulating to indicate that virulence genes of pathogenetic bacteria are often clustered within definable genetic elements termed pathogenicity islands [8, 9, 10]. The *bft-1* gene of *B. fragilis* VPI 13784 has also been reported to be present within a small genetic element termed the fragilysin (*bft-1*) pathogenicity islet [13]. Besides BFT-1, two ORFs have been identified in the pathogenicity islet; T-3 and a third ORF. A sequence analysis of the ORFs in the pathogenicity islet has revealed that the *bft-1* and *t-3* genes encode metalloprotease proteins with similar sizes as well as structural features, both of which contain a zinc-binding motif and methionine-turn region [13]. A third open reading frame located immediately upstream of the *bft-1* gene appears to encode a small protein that shows some sequence identity with cobra cytotoxins [12, 13]. However, besides BFT-1, the involvement of other gene products in pathogenicity has not been studied yet.

Previously, we have identified 34 ETBF strains from Korean patients by using cell culture assays, colony blot hybridization, and PCR [4]. Eleven of them expressed a new isoform of BFT, which was more closely related to BFT-2 than BFT-1, and exhibited the same biological activity as the other two isoforms [4]. The gene for the protein was cloned and sequenced from *B. fragilis* 419, which was isolated from the blood of a patient suffering from systemic

*Corresponding author
Phone: 82-2-380-1465; Fax: 82-2-380-1487;
E-mail: hbotox@chollian.net



B.

Fig. 1. The enterotoxin pathogenicity islet of *B. fragilis* 419.

(A) Schematic diagram of the enterotoxin pathogenicity islet. pBFT-6. and pMP2-2. The ORFs are represented by thick arrows. The small arrows at the ends represent direct repeats The restriction sites are *Bgl*II (B), *Eco*RI (E), *Kpn*I (K), *Nde*I (N), and *Xmn*I (X) (B) Nucleotide and deduced amino acid sequences of the enterotoxin pathogenicity islet. The GenBank accession number is AF103902. The stop codons are indicated by *asterisks* The putative TATA boxes and -35 sequences are indicated by *underlines*. Potential ribosome binding sites and direct repeats are indicated with *bold letters*. The ends of the pathogenicity islet and primers. BFTA and BFTB. used for the PCR are denoted by the arrows above the nucleotide sequence.

infections [4]. This gene was referred as *bft-k* [4], and attempts were made determine the genetic structure of the enterotoxin pathogenicity islet of *B. fragilis* 419. This report describes the molecular characterization of the *bft-k*-containing enterotoxin pathogenicity islet of the strain.

Cloning of the BFT Pathogenicity Islet in *B. fragilis* 419

The plasmid pBFT-6, which contained the *bft-k* gene of *B. fragilis* 419, was cloned and sequenced as previously reported [4]. Since the pBFT-6 did not contain the full length of the BFT pathogenicity islet, another screening process was carried out to obtain a complete sequence. For this screening, two oligonucleotide primers, whose sequences were derived from the published *t-3* sequences, were synthesized [13]. The forward BFTB and reverse BFTA primer sequences were 5'-TTTTTGTATGCATTCCA-AC-3' and 3'-TTCTTCTAGCAGTCGTGTAC-5', respectively (Fig. 1). Using these two primers, a putative T-3 gene fragment of 1,136 bp was amplified from the *B. fragilis* 419 genomic DNA by PCR using *Taq* DNA polymerase [19], and subcloned into pTOPO (Invitrogen). Sequencing of the inserted fragment revealed that it contained the expected T-3 nucleotide sequence. Next, a λZAP library of the *B. fragilis* 419 genomic DNA was screened using the 1,136 bp PCR fragment as a probe. Several positive plaques were isolated, and the λZAP Express™ vectors were excised *in vivo* according to the manufacturer's protocol (Stratagene). The excised plasmids from the positive clones were digested with *Eco*RI, and subjected to Southern blot with the same probe. Among the several positive clones, one of the excised plasmids, pMP2-2, which had a 4.0 kb insert, was hybridized with the probe and sequenced in both directions. The pMP2-2 contained the missing part of the BFT pathogenicity islet that was not included in the pBFT-6. In addition, approximately 1 kb of the pMP2-2 insert perfectly overlapped with the pBFT-6 insert (Fig. 1).

A sequence analysis of the cloned pathogenicity islet revealed that it consisted of 6,045 bp and contained three putative ORFs: BFT-K, T-3, and ORF1. The 419 pathogenicity islet also contained nearly perfect direct repeats of 12 bp close to its ends (Fig. 1). The ends of the pathogenicity islet were defined by comparing the nucleotide sequences around the direct repeats (Fig. 1) with those of the previously reported pathogenicity islet of VPI 13784 [13]. The nucleotide sequence of the 419 pathogenicity islet was deposited in the GenBank Database under accession number AF103902.

Sequence Analysis of the ORFs in the BFT Pathogenicity Islet of *B. fragilis* 419

As recently reported, the isolated BFT pathogenicity islet contains a complete open reading frame (BFT-K) of 1,191 bp, which encodes a protein product of 397 residues with a predicted molecular weight of 44,396 with pI of 5.13 (Fig. 1) [4]. A comparison of the alignment of the predicted amino acid sequence of the *bft-k* gene with those of strains VPI 13784 and 86-5443-2-2 revealed that the BFT-K shared 93% and 95% identity with the BFT-1 and BFT-2, respectively [4]. The alignment also suggested that the BFT-K was synthesized as a preproprotein where the initial 18 amino acids comprised a signal peptide, followed by a 193-residue 'pro' region and an active mature region of 186 residues containing the zinc-binding signature motif (HELGHILGANH) and Met-turn [4]. Similar to the BFT-1 and BFT-2, the purified BFT-K exhibited cytotoxicity on HT29/C1 cells, and cleavage of E-cadherin on HT29/C1 by BFT-K was observed [4].

Approximately 4 kb upstream of the start codon was sequenced. Although the transcription start site was not determined, a putative TATA-like box and -35 sequence were found between bases -66 and -71, and -95 and -100 from the start codon, respectively (Fig. 1).

Interestingly, ETBF strains that contain BFT-K have only been identified in the intestinal and extraintestinal isolates of Japanese and Koreans subjects, but not in isolates of Americans [4]. This fact raises intriguing questions regarding the molecular evolution of ETBF strains, and further studies are to clarify the global distribution of ETBFs and their toxin subtypes.

Another open reading frame (T-3) of 1,188 bp locating 1,687 bp upstream of the *bft-k* gene encodes the predicted 396-amino-acids protein (Fig. 1). The predicted protein has a calculated molecular weight of 44,405 Da and a pI of 5.12. The deduced amino acid sequence of the *t-3* gene showed some sequence homology to many metalloproteases including BFTs. The predicted amino acid sequences of the *t-3* and *bft-k* genes contain similar features, such as a zinc-binding motif, Met-turn, and hydrophobic signal peptide sequence in their N-terminus [25]. The alignment of their amino acid sequences showed 26% sequence identity and 57% similarity when conservative substitutions were considered (Fig. 2).

The nucleotide and deduced amino acid sequences of the *t-3* genes from strains 419 and VPI 13784 [13] shared over 99% identity. Among the proteins encoded by the genes in the BFT pathogenicity islet, the T-3 protein appeared to be more conserved than the other gene products in the islet. This may suggest that the T-3 protein plays a more important role in the pathogenicity of ETBF than BFT. However, so far, no studies on the toxic properties of the T-3 protein have been reported.

To verify the expression of the *t-3* gene in strain 419, a recombinant mature form of T-3 protein was produced, and an immunoblot assay was performed using polyclonal antibodies raised in mice against the recombinant protein. Since the N-terminus of T-3 contains an 18-amino-acids hydrophobic signal sequence, and an Arg213-Ala214 processing site is located in the middle of the protein in a

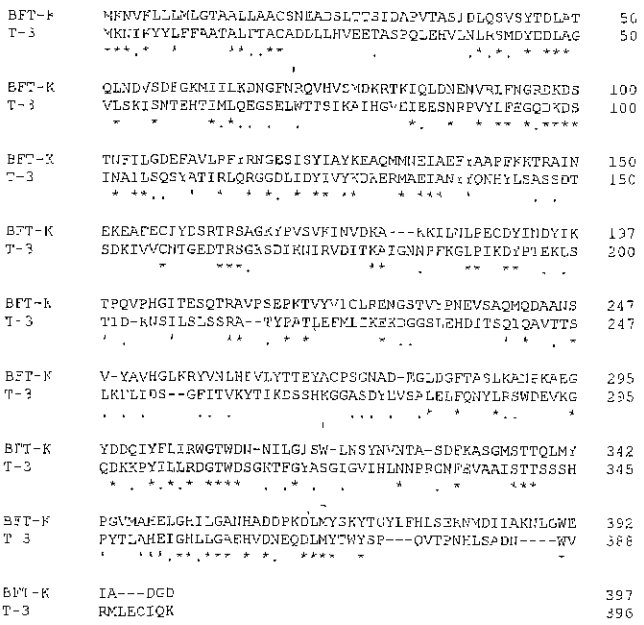


Fig. 2. Alignment of the predicted amino acid sequences of the *t-3* and *bft-k* genes.
The *asterisks* indicate identical amino acid residues in the sequence and *dots* indicate conservative substitutions.

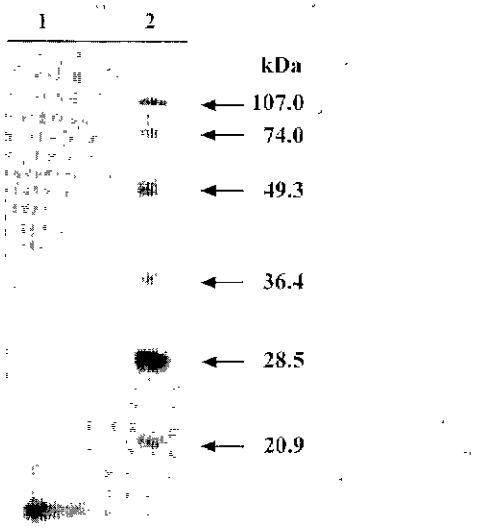


Fig. 3. Western blot analysis of the T-3 protein from *B. fragilis* 419.
Phosphorylase *b* (107 kDa), bovine serum albumin (74 kDa), ovalbumin (49.3 kDa), carbonic anhydrase (36.4 kDa), soybean trypsin inhibitor (28.5 kDa), and lysozyme (20.9 kDa) were used as molecular standards (lane 2). Lane 1 is the concentrated cell culture supernatant fraction of *B. fragilis* 419. Polyclonal antibodies raised in mice against the mature form of the recombinant protein of T-3 produced in *E. coli* were used.

position similar to the one found in BFT proteins, it can be speculated that T-3 is processed and released from the cytoplasm through two processing steps, as found in BFT. If processed at these sites, the released mature T-3 would be a 183-residue protein with a molecular weight of 20.4 kDa. Thus, the culture supernatant fluid of *B. fragilis* 419 cells grown in a brain heart infusion broth to an early stationary phase was collected and concentrated by precipitation with trichloroacetic acid at a final concentration of 10% [2]. The precipitated proteins were dissolved in 0.1 N NaOH and dialyzed in 50 mM Tris-HCl (pH 7.5) at 4°C. Next, an SDS-PAGE (15%) was performed to separate the proteins, and immunoblotting was conducted according to the standard method [2]. In the concentrated cell culture supernatant fraction, one band of 16.0 kDa in size was detected (Fig. 3, lane 1). The polyclonal antibodies did not detect the purified BFT-K in a control immunoblotting experiment, eliminating the possibility of the antibodies crossreacting with BFT-K (data not shown). Accordingly, these results suggest that T-3 is expressed, processed, and secreted from the cytoplasm in *B. fragilis* 419. In contrast, the recombinant T-3 did not show any biological activity on HT29/C_1 *in vitro* (Rhie *et al.*, unpublished result).

A putative Shine-Dalgano (SD) sequence was found -2 bp from the start codon, and the putative TATA box and -35 sequences were found between bases -48 and -53, and -78 and -83 from the start codon, respectively (Fig. 1).

A third ORF comprising 177 bp was found 93 bp upstream of the *bft-k* gene, and encoded a protein composed of potential 59 amino acids with a molecular weight of 6,955 Da in a different open reading frame of BFT-K (Fig. 1). ORF1 has a putative SD sequence (5'-AGG-3') located 9 nucleotide upstream of the start codon. A putative TATA box and potential -35 sequences were found at -63 and -91 nucleotides upstream of the start codon, respectively (Fig. 1).

The ORF1 identified in strain 419 showed a 73% homology to the unknown ORF of the fragilysin pathogenicity islet in VPI 13784 [12], although the ORF1 in strain 419 showed a pre-termination of the protein due to a frame-shift. The unknown ORF in VPI 13784 has been reported to have 88 amino acids, 23% overall homology with the cytotoxin CTXIIb from the cobra *NAJA mossambica mossambica*, and a 44% homology with a final 40% of the cobra toxin [12]. However, the ORF1 in strain 419 did not show any significant homology with these cytotoxins. The expression and involvement of this protein with bacterial pathogenicity remains unanswered.

The flanking DNA sequence of the strain 419 BFT pathogenicity islet almost perfectly aligned with that of the fragilysin pathogenicity islet of VPI 13784 [13], suggesting that the strain 419 BFT pathogenicity islet was integrated into a specific site in the chromosome as a fragilysin pathogenicity islet. Furthermore, Southern blot assay with the genomic DNA digested with several restriction endonucleases, including *Eco*RI, *Bgl*II, *Xho*I, and *Hin*dIII, detected only one hybridizing band (data not shown) when the PCR fragment, which contained a part of

the *bft-k* gene, was used as a probe. These results indicate that the BFT pathogenicity islet was integrated only once into a specific site in the chromosome. The G+C content of the entire pathogenicity islet in strain 419 was 35%, while that of the total DNA of *B. fragilis* was 41 to 44% [11].

In summary, the BFT pathogenicity islet was cloned and sequenced from *B. fragilis* 419 isolated from the blood of a systemically infected Korean patient. The structure and order of the genes in the pathogenicity islet were very similar to those reported in VPI 13784 [13]. Furthermore, three ORFs were identified in the pathogenicity islet: BFT-K, T-3, and ORF1. The expression of the *t-3* gene was verified by Western blot analysis. Further experiments are now in progress to determine whether the gene products of the BFT pathogenicity islet, in addition to BFT-K, are involved in the pathogenicity of the bacteria.

Acknowledgments

This work was supported by a research grant (1998) to the National Institute of Health (Korea) from the Korean Ministry of Health and Welfare.

References

1. Bode, W., F. Gomis-Ruth, and W. L Stockler. 1993. Astacins, serralysins, snake venom and matrix metalloproteases exhibit identical zinc-binding environment (HEXXHXXGXXH and met-turn) and topologies and should be grouped into a common family, the metzincins. *FEBS Lett.* **331:** 134–140.
2. Bollag, D. M. and S. J. Edelstein. 1991. *Protein Methods*, pp 71–94. Wiley-Liss, Inc. New York, U.S.A.
3. Border, M., B. D. Firehammer, D. S. Shoop, and L. L. Myers. 1985. Isolation of *Bacteroides fragilis* from feces of diarrheic calves and lambs. *J. Clin. Microbiol.* **21:** 472–473.
4. Chung, G. T., A. A. Franco, S. Wu, R. Cheng, G. Rhie, H. B. Oh, and C. L. Sears. 1999 Identification of a third metalloprotease toxin gene in extraintestinal isolates of *Bacteroides fragilis*. *Infect. Immun.* **67:** 4945–4949.
5. Chung, G T., W. K. Seong, K. S. Shin, K. W. Lee, and H. B. Oh. 1997. The detection of enterotoxin gene from *Bacteroides fragilis* isolates in Korea by polymerase chain reaction. *J. Kor. Soc. Microbiol.* **32:** 521–528.
6. Donelli, G., A. Fabbri, and C. Fiorentini. 1996. *Bacteroides fragilis* enterotoxin induces cytoskeletal changes and surface blebbing in HT-29 cells. *Infect. Immun.* **64:** 113–119.
7. Franco, A. A., L. M. Mundy, M. Trucksis, S. Wu, J. B. Kaper, and C. L. Sears. 1997. Cloning and characterization of the *Bacteroides fragilis* metalloprotease toxin gene. *Infect. Immun.* **65:** 1007–1013.
8. Groisman, E. A. and H. Ochman. 1996. Pathogenicity islands: Bacterial evolution in quantum leaps. *Cell* **87:** 791–794.
9. Groisman, E. A. and H. Ochman. 1997. How *Salmonella* became a pathogen. *Trends Microbiol.* **5:** 343–349.
10. Hacker, J., G. Blum-Oehler, I. Muhldorfer, and H. Tschape. 1997 Pathogenicity islands of virulence bacteria: Structure, function and impact on microbial evolution. *Mol. Microbiol.* **23:** 1089–1097.
11. Johnson, J. L. 1978. Taxonomy of the *Bacteroides*. I. Deoxyribo nucleic acid homologies among *Bacteroides fragilis* and other saccharolytic *Bacteroides* species. *Int. J. Syst. Bacteriol.* **28:** 245–256.
12. Kling, J. K., R. L. Wright, J. S. Moncrief, and T. D. Wilkins. 1997. Cloning and characterization of the gene for the metalloprotease enterotoxin of *Bacteroides fragilis*. *FEMS Microbiol. Lett.* **146:** 279–284.
13. Moncrief, J. S., A. J. Duncan, R. Rhonda, L. Wright, L. A. Barroso, and T. D. Wilkins. 1998. Molecular characterization of the fragilysin pathogenicity islet of enterotoxigenic *Bacteroides fragilis*. *Infect. Immun.* **66:** 1735–1739.
14. Moncrief, J. S., R. Obiso Jr., L. A. Barroso, J. J. Kling, R. L. Wright, R. L. Van Tassel, D. M. Lyerly, and T. D. Wilkins. 1995. The enterotoxin of *Bacteroides fragilis* is a metalloprotease. *Infect. Immun.* **63:** 175–181.
15. Moore, W. E. C., E. P. Cato, and L. V. Holdeman. 1978. Some current concepts in intestinal bacteriology. *Am. J. Clin. Nutr.* **31:** S33–S42.
16. Mundy, L. M. and C. L. Sears. 1996. Detection of toxin production by *Bacteroides fragilis*. *Clin. Infect. Dis.* **23:** 269–276.
17. Myers, L. L., B. D. Firehammer, D. S. Shoop, and M. M. Border. 1984. *Bacteroides fragilis*: A possible cause of acute diarrheal disease in newborn lambs. *Infect. Immun.* **44:** 241–244.
18. Obiso Jr., R. J., D. M. Lyerly, R. L. Van Tassel, and T. D. Wilkins. 1995. Proteolytic activity of the *Bacteroides fragilis* enterotoxin causes fluid secretion and intestinal damage *in vivo*. *Infect. Immun.* **63:** 3820–3826.
19 Park Y., J. M. Kim, H. Choi, S. K. Kim, and Y. Kim. 1998. Construction of two metal-ion binding sites to improve the 3-5 exonuclease activity of *Taq* DNA polymerase. *J. Microbiol. Biotechnol.* **8:** 471–477.
20. Saidi, R. F. and C. L. Sears. 1996. *Bacteroides fragilis* toxin rapidly intoxicates human intestinal epithelial cells (HT29/C_1) *in vitro*. *Infect. Immun.* **64:** 5029–5034.
21. Sears, C. L., L. L. Myers, A. Lazenby, and R. L. Van Tassel. 1995. Enterotoxigenic *Bacteroides fragilis*. *Clin. Infect. Dis* **20** (Suppl. 2)**:** S142–S148.
22. Stockler, W. L., F. Grams, U. Bauman, P. Reiner, F. X. Gomis-Ruth, D. B. McKay, and W. Bode. 1995. The metzincins - topological and sequential relations between the astacins, akamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc peptidases. *Protein Sci.* **4:** 823–840.
23. Van Tassel, R. L., D. M. Lyerly, and T. D. Wilkins. 1992. Purification and characterization of an enterotoxin from *Bacteroides fragilis*. *Infect. Immun.* **60:** 1343–1350.
24. Wu, S., K. Lim, J. Huang, R. F. Saidi, and C. L. Sears. 1999. *Bacteroides fragilis* enterotoxin cleaves the zonular adherens protein E-cadherin. *Proc. Natl Acad. Sci. USA* **95:** 14979–14984.
25. Wu, H. C. and M. Tokunaga. 1986. Biogenesis of lipoproteins in bacteria. *Curr Top. Microbiol. Immunol.* **125:** 127–157.