JOURNAL OF BACTERIOLOGY, Apr. 1988, p. 1658–1665
0021-9193/88/041658-08$02.00/0

Vol. 170, No. 4

Molecular Cloning and Sequencing of the Gene Encoding the Fimbrial Subunit Protein of *Bacteroides gingivalis*

DOUGLAS P. DICKINSON,†* MICHAEL A. KUBINIEC, FUMINOBU YOSHIMURA,‡ AND ROBERT J. GENCO†

Department of Oral Biology, State University of New York at Buffalo, Buffalo, New York 14214

Received 2 November 1987/Accepted 5 January 1988

The gene encoding the fimbrial subunit protein of *Bacteroides gingivalis* 381, fimbrilin, has been cloned and sequenced. The gene was present as a single copy on the bacterial chromosome, and the codon usage in the gene conformed closely to that expected for an abundant protein. The predicted size of the mature protein was 35,924 daltons, and the secretory form may have had a 10-amino-acid, hydrophilic leader sequence similar to the leader sequences of the MePhe fimbriae family. The protein sequence had no marked similarity to known fimbrial sequences, and no homologous sequences could be found in other black-pigmented *Bacteroides* species, suggesting that fimbrilin represents a class of fimbrial subunit protein of limited distribution.

Black-pigmented members of the anaerobic gram-negative genus *Bacteroides* are frequent colonists of the human oral cavity (6, 25, 26). In particular, *Bacteroides gingivalis* has been shown to comprise a significant proportion of the cultivable flora in established periodontal lesions (37, 46) and has been implicated in the pathogenesis of periodontal disease (35).

It is widely accepted that one of the first stages in colonization by bacteria is adherence (1, 2). Bacterial fimbriae (pili) have been shown to mediate specific cell adhesion to surfaces (1–3, 8), and their presence has been demonstrated on several black-pigmented *Bacteroides* species, including *B. gingivalis* (14, 30, 35, 36). However, the adhesive function of *B. gingivalis* fimbriae has yet to be established. Some reports have correlated the hemagglutinating activity of *B. gingivalis* (29, 30, 34, 36) with fimbriation (29, 36), indicating a potential role in binding to oral epithelial cells. Other studies have found no correlation between fimbriae and hemagglutination in this organism (5, 44). In one of these studies, fimbriae were purified from *B. gingivalis* 381 by shearing, and hemagglutinating activity did not purify with the fimbriae (44). These different results could have been due to copurification of the hemagglutination factor with fimbriae in certain preparations or they could have been caused by the presence of different types of fimbriae on the different *B. gingivalis* strains used in these studies.

The fimbrial subunit protein of *B. gingivalis* 381, fimbrilin, is approximately 43 kilodaltons, as determined by gel electrophoresis (44). The N-terminal sequence of the subunit protein appears to be unrelated to the N termini of fimbriae from other species, including the nonpigmented *Bacteroides nodosus* (45). No data are available on the relationships, if any, between *B. gingivalis* fimbriae and those of the other black-pigmented *Bacteroides* species.

We have sought to isolate a recombinant clone carrying the fimbrilin gene from *B. gingivalis* 381. We had two specific aims: first, to provide basic information on the structure of the subunit protein, which might provide further clues to its function and to any relationship with other fimbriae; and second, to provide a DNA hybridization probe which would allow us to investigate the presence of related genes in other black-pigmented *Bacteroides* species and the expression of the fimbrilin gene in fimbriate strains of *B. gingivalis*.

(Part of this work appeared as an abstract at the 1986 International Association of Dental Research annual meeting [D. P. Dickinson, M. Kubiniec, F. Yoshimura, and R. J. Genco, J. Dent. Res. **65**:24, 1986].)

MATERIALS AND METHODS

Strains and media. *B. gingivalis* 381 (kindly provided by S. Socransky) was maintained on blood agar plates (40 g of tryptic soy agar per liter, 1 g of yeast extract per liter, and 5% defibrinated sheep blood, supplemented with 1 mg of vitamin K_1 per liter). Cells were transferred to fresh plates every 4 to 7 days. Growth was at 37°C in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 85% nitrogen, 10% hydrogen, and 5% carbon dioxide. For liquid cultures, 10 ml of 0.5× brain heart infusion medium (per liter: 20 g of brain heart infusion, 0.75 g of cysteine, 5 mg of hemin, and 1 mg of vitamin K_1) was inoculated with a single colony and grown anaerobically at 37°C for 1 to 2 days. This preculture was then used to inoculate larger volumes of 0.5× brain heart infusion medium for growth under the same conditions. Colony purity was assessed by Gram stain, phase-contrast microscopy, and anaerobic and aerobic culture on blood agar plates. *Escherichia coli* JM83 [F^- *ara* Δ(*lac-pro*) *strA thi* φ80d *lacZ* Δ*M15*) was used as the host strain for pUC13 plasmid subclones (40). *E. coli* DH-1 (F^- *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1*) was used as the host strain for cosmid cloning (13) with the cosmid vector pJB8 (17). Cells were maintained as frozen glycerol stocks, and each strain colony was purified before use by plating on Luria agar plates (21). Routine liquid cultures were prepared by single-colony inoculation of a 5-ml Luria broth overnight culture, which was then used to inoculate (1:100) larger volumes of medium. Growth was aerobic at 37°C with shaking (250 rpm). The selective medium contained 50 μg of ampicillin per ml.

Isolation of high-molecular-weight genomic DNA. DNA was isolated from *B. gingivalis* 381 and *E. coli* JM83 and DH-1 by a modification of the method of Meyer et al. (24).

* Corresponding author.

† Present address: Department of Biological Chemistry, Dental Branch, University of Texas Health Science Center at Houston, 6516 John Freeman Avenue, P.O. Box 20068, Houston, TX 77025.

‡ Present address: School of Dentistry, Hokkaido University, Kita-ku, Sapporo 060, Japan.

Briefly, cells in the late logarithmic phase of growth were lysed in 1.6% sodium dodecyl sulfate, 50 mM NaCl, 50 mM Tris hydrochloride, and 50 mM EDTA (pH 7.5). RNA and protein were digested by successive treatment with RNase A (1 μg/ml) and pronase (0.5 mg/ml), and high-molecular-weight DNA was purified by repeated phenol extraction followed by ammonium acetate-isopropanol precipitation (21). The yields were ca. 10 to 30 mg of DNA per liter of culture, and the DNA was >100 kilobases (kb), as judged by gel electrophoresis on 0.3% agarose gels.

Construction and testing of oligonucleotide probes. Several 17- to 20-residue mixed-base probes were synthesized to correspond to the N-terminal coding sequence of fimbrilin by using phosphoramidite chemistry (4) on an Applied Biosystems model 380A DNA synthesizer. Each probe was desalted by gel exclusion chromatography on a Sephadex G-50 column (18 by 0.7 cm) eluted with 10 mM triethylammonium bicarbonate (pH 7.0) and lyophilized. The probes were then 5′ end-labeled with ^{32}P in a 20-μl reaction containing 300 pmol of oligonucleotide, 70 mM Tris hydrochloride (pH 7.6), 10 mM $MgCl_2$, 5 mM dithiothreitol, 0.1 mM spermidine, 150 μCi of [γ-^{32}P]rATP (7,000 Ci/mmol), and 10 U of T4 polynucleotide kinase (Bethesda Research Laboratories, Inc.). The reaction was incubated at 37°C for 1 h, and the labeled probes were used without further purification. To test oligonucleotides as probes for detecting the fimbrilin gene, 5-μg samples of *B. gingivalis* 381 and *E. coli* genomic DNA and 1 μg of both pJB8 and pUC13 vector DNAs were digested with *Bam*HI, electrophoresed on 1% agarose gels, and transferred to nylon membranes (Zetabind; AMF Cuno) (38). The baked filters were rinsed at 65°C for 30 min in 3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0)–0.1% sodium dodecyl sulfate and then prehybridized with 6× SSC–1× Denhardt solution (21)–100 μg of sheared denatured salmon sperm DNA per ml–0.05% sodium pyrophosphate at 1 ml/cm^2 of the blot. Hybridization with the probes was performed in the same volume of fresh prehybridization mix containing 2 × 10^6 to 3 × 10^6 cpm/ml of probe. Hybridization was continued overnight at 37°C, after which the blots were rinsed four times at room temperature with 500 ml of 6× SSC–0.05% sodium pyrophosphate. To establish probe specificities and wash stringencies, the filters were further washed at increasing temperatures. Each successive wash was performed for 15 min with 500 ml of 6× SSC–0.05% sodium pyrophosphate at 37, 47, 50, and 52°C. Between each temperature increment, the filter was removed, blotted dry, and autoradiographed overnight at −70°C with an intensifying screen.

Cosmid cloning and screening of the cosmid library. A 1,200-μl reaction mixture containing 250 μg of *B. gingivalis* genomic DNA, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris hydrochloride (pH 7.5), and 100 μg of bovine serum albumin per ml was prewarmed to 37°C for 20 min. Then 2.5 U of *Sau*3AI (New England BioLabs, Inc.) was added, and at intervals of 10, 25, and 40 min, 400-μl samples were withdrawn and added to 20 μl of 0.5 M EDTA (pH 8.0). This protocol yielded *Sau*3AI fragments in the range of 20 to 50 kb, as shown by electrophoresis of samples of each fraction in the cold at 3 V/cm on 0.3% agarose gels with TAE buffer (21) containing 0.5 μg of ethidium bromide per ml. Bacteriophage λ DNA and suitable restriction fragments of λ DNA were used as size markers. DNA fragments in the pooled samples were separated on a 30-ml 5 to 25% sucrose gradient containing 10 mM Tris hydrochloride and 1 mM EDTA (pH 7.5) by centrifugation for 14 h at 22,500 rpm and 15°C in a Beckman SW-28 rotor. Fractions (1.2 ml) were collected and analyzed on 0.3% agarose gels as described above. DNA fragments in the 30- to 50-kb range were recovered by precipitation with ammonium acetate-isopropanol from appropriate fractions. Cosmids were constructed essentially as described by Grosveld et al. (12). The cosmid vector pJB8 (17) was linearized with *Bam*HI and treated with phosphatase. The DNA fragments were ligated into the vector, and the ligation mix was packaged into bacteriophage λ heads with a commercial packaging kit (Promega-Biotec). The packaged heads were then used to infect *E. coli* DH-1 as described previously (21). Cosmids were replicated onto nitrocellulose disks and lysed in situ by sequential treatment with sodium dodecyl sulfate and NaOH (21). After being baked, the filters were hybridized with oligonucleotide BGFIM2 (see Results) essentially as described above.

Plasmid cloning. To clone *B. gingivalis* genomic DNA fragments into pUC13 (40), 100 μg of *B. gingivalis* 381 genomic DNA was digested to completion with *Sac*I restriction endonuclease as recommended by the manufacturer (New England BioLabs). This enzyme was chosen because it produced the smallest fragment (2.5 kb) that carried a sequence complementary to BGFIM2 (see Results) and that was compatible with the vector polylinker cloning sites. The restricted DNA was electrophoresed on a 1% agarose preparative gel, the region of the gel containing fragments of 2 to 3 kb was excised, and the DNA (about 1 μg) was recovered by electroelution (model 1750 sample concentrator; ISCO). The pUC13 vector was linearized with *Sac*I and treated with calf intestinal alkaline phosphatase (21). The *Sac*I genomic DNA fragments were ligated into the vector, and the ligation mix was used to transform JM83 cells rendered competent by calcium chloride treatment (21). 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside was used to identify recombinants as white colonies (40).

Screening of *Sac*I clones. Normally, recombinants carrying cloned sequences are identified by replication onto filters, followed by colony lysis and probe hybridization. Positive colonies are then purified, retested by filter colony screening, and characterized by restriction endonuclease analysis and Southern blotting. However, since the electroeluted genomic DNA sample used to construct the pUC13 library contained only 5 to 10 *Sac*I fragments (not shown), 1 of which contained a sequence complementary to our oligonucleotide probe, the probability of finding a positive clone in the first 22 recombinants examined at random (assuming equal cloning efficiencies) was >90%. We therefore employed a rapid, direct Southern blot analysis of isolated plasmid DNA to initially test individual recombinants and simultaneously provide direct information on the clone inserts. White colonies were picked off the primary plate and inoculated both onto a selective agar plate (to provide a stock) and into 1 ml of selective Luria broth. After overnight growth, plasmid DNA from 22 broth cultures was isolated by the rapid alkaline lysis method of Ish-Horowicz and Burke (17) and digested with *Sac*I for 1 h, and the fragments were separated on a single 0.7% agarose gel containing 0.5 μg of ethidium bromide per ml. The DNA was transferred to a nylon membrane (38) for 2 h. Transfer of most of the DNA was verified by examination of the gel and filter under UV. The filter was baked for 1 h at 80°C in vacuo and then hybridized with the oligonucleotide probe essentially as described above, except that the hybridization time was reduced to 5 h, the filter was washed directly at 52°C, and X-ray film was exposed overnight with an intensifying screen.

Sequencing. A putative fimbrilin clone (pUC13Bg12.1; see

Results) was characterized by restriction mapping (21) and Southern blotting (38) with oligonucleotide probe BGFIM2. Subclones were then constructed in either pUC13 or pUC18 by using appropriate restriction fragments. These subclones were then linearized by restriction endonuclease digestion at an appropriate site in the vector polylinker region. For 5'-to-3' sequencing, the linearized DNA was treated with alkaline phosphatase (21) and separated from residual RNA by agarose gel electrophoresis and electroelution. The DNA was end labeled essentially as described above, except that a 50-μl reaction mixture with 200 μCi of [γ-^{32}P]rATP was used. Free label was removed by precipitating the DNA with 2.5 mM spermidine–10% dimethyl sulfoxide at −70°C for 10 min, thawing it on ice for 1 h, and recovering the DNA by centrifugation. The labeled DNA was washed three times with cold 70% ethanol and dried. For 3'-to-5' sequencing, the DNA was labeled with the appropriate deoxynucleotide triphosphate by using a Klenow fill-in reaction (21). Labeled DNAs were then restricted with a second enzyme, and the fragments were purified by gel electrophoresis and electroelution. The fragments were sequenced by the method of Maxam and Gilbert (23). A search was conducted of the Protein Sequence Data Base of the Protein Identification Resource (release no. 7.0, 1986; National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.) for protein sequences similar to the derived fimbrilin sequence.

Southern blot analysis of black-pigmented *Bacteroides* species by using the fimbrilin clone. The 2.5-kb *Sac*I insert in pUC13Bg12.1 was labeled by the random primed method of Feinberg and Vogelstein (7) to a specific activity of $>3 \times 10^8$ cpm/μg with [α-^{32}P]dCTP. High-molecular-weight genomic DNA was isolated from the American Type Culture Collection type strains of black-pigmented *Bacteroides* species, essentially as described above for *B. gingivalis*. The strains used were *B. endodontalis* ATCC 35406, *B. asaccharolyticus* ATCC 25260 and ATCC 27067, *B. gingivalis* ATCC 33277, *B. intermedius* ATCC 25611 and ATCC 25261, *B. corporis*, ATCC 33547, *B. melaninogenicus* ATCC 25845, *B. denticola* ATCC 33185, *B. loescheii* ATCC 15930, and *B. levii* ATCC 29147. After restriction of 1 μg of DNA from each strain with *Pst*I and electrophoresis on a 0.7% agarose gel, the DNA fragments were transferred to nylon membrane as described above. The baked filters were rinsed at room temperature in 6× SSC and then incubated at 65°C for 1 h in 200 ml of 10× Denhardt solution–4× SSC–0.1% sodium dodecyl sulfate–0.1% sodium pyrophosphate per filter. The filters were then prehybridized for 4 h at 60°C in 25 ml of 5× SSC–5× Denhardt solution–0.1% sodium dodecyl sulfate–10% dextran sulfate–100-μg/ml denatured salmon sperm DNA per filter; then 2×10^6 cpm of denatured probe per ml was added. After overnight incubation at 60°C, the filters were washed several times at 60°C. Each wash was performed for 45 min with 200 ml of 5× SSC–0.1% sodium dodecyl sulfate. The dried filters were then exposed to X-ray film with an intensifying screen at −70°C for various times.

RESULTS AND DISCUSSION

Oligonucleotide probes for detecting the fimbrilin gene. The availability of the N-terminal peptide sequence of *B. gingivalis* 381 (44) allowed us to design oligonucleotide probes which would hybridize to the corresponding region of the genome. Synthesized probes were radiolabeled and tested by Southern blot hybridization analysis of *B. gingivalis* genomic DNA. A 20-mer probe designated BGFIM2 (5'-GCNAARTTRACNGTNATGGT-3'OH) yielded a single band of ca. 6.5 kb with *B. gingivalis* genomic DNA, even at the most stringent temperature tested, and yielded no detectable signal with an equivalent amount of *Bam*HI-digested *E. coli* genomic DNA or with a >100-fold molar excess of vector DNA, even at the lower stringencies tested (data not shown). In contrast, the other probes we tested gave either weak signals, cross-hybridization with control DNAs, or both. We therefore used BGFIM2 as a hybridization probe to screen for recombinants carrying fimbrilin sequences.

Cosmid library construction and screening. We initially sought to construct a cosmid library containing *B. gingivalis* genomic DNA fragments, since a clone carrying the fimbrilin subunit gene might also carry accessory genes involved in fimbriation, as has been found in other cases (20, 28, 31; for a review, see reference 8). At least 94% of the clones carried an appropriate insert, as shown by *Bam*HI digestion of small-scale cosmid preparations (17) and gel electrophoresis (not shown). Despite screening over 2,000 clones (for a probability of >99.9% that the library contained the fimbrilin gene, assuming all fragments cloned with equal efficiency [21]), we failed to detect a clone that gave a significant hybridization signal, even after chloramphenicol amplification (21).

Plasmid cloning. One possible explanation for our failure to detect clones in the cosmid library which carried fimbrilin sequences was a deleterious effect on *E. coli* of sequences closely linked to the fimbrilin gene. Cloning of substantially smaller DNA fragments could obviate this potential difficulty. Preliminary Southern blot data obtained by using BGFIM2 as probe (not shown) indicated that *Sac*I produced a ca. 2.5-kb genomic DNA fragment that hybridized with the probe. Other enzymes tested that would give appropriately sized fragments and that would clone efficiently into the polylinker region of the pUC13 vector all gave larger fragments. Therefore, *Sac*I fragments of *B. gingivalis* 381 genomic DNA in the range of 2 to 3 kb were cloned into pUC13. We estimate that there were 5 to 10 fragments in this size range, judging by the number and intensity of bands seen on ethidium bromide-stained agarose gels. Of the first 22 clones screened by direct Southern blot analysis, 1 carried a 2.5-kb *Sac*I fragment that strongly hybridized with the BGFIM2 probe. This clone was designated pUC13 Bg12.1.

Sequence analysis of pUC13Bg12.1. A map of selected restriction sites for clone pUC13Bg12.1 and the sequencing strategy used are shown in Fig. 1. To facilitate sequencing by the method of Maxam and Gilbert (23), subclones were constructed in pUC13. During subcloning, fragments derived from the region between the 3' end of the fimbrilin gene and the downstream *Sac*I site were inserted into the polylinker of pUC13 in the orientation opposite to that in the parent clone. Such constructs were highly deleterious to *E. coli*. Clones containing these inserts could be distinguished on plates as flat, transparent gray colonies which microscopic examination showed to contain many dead cells and cell wall ghosts. The gene fragments caused no adverse effect when cloned in the orientation opposite to that of the *lacZ* gene in pUC18. The apparently lethal effect of this region of the *B. gingivalis* genome on *E. coli* may explain our failure to obtain a cosmid clone.

The predicted sequence of the fimbrilin gene is shown in Fig. 2. The region comprising the major open reading frame that includes the BGFIM2 sequence (see below) was sequenced on both strands, and the majority of each strand

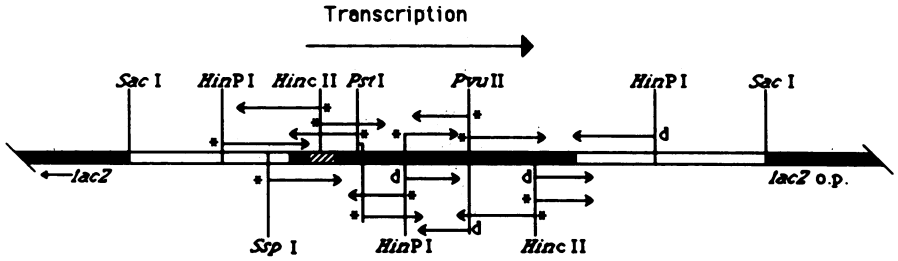

FIG. 1. Sequencing strategy used to determine the nucleotide sequence of pUC13Bg12.1. With the exception of the sequence derived from the *Ssp*I site, which was determined by directly labeling *Ssp*I-digested pUC13Bg12.1, the sequences were determined from fragments subcloned into pUC13 or pUC18 by using appropriate restriction sites in the polylinkers. Symbols: *, 5′ end labeled; d, 3′ end labeled; ▬, vector DNA (o., *lac* operator; p., *lac* promoter); ▭, region encoding the fimbrilin protein; ▨, N-terminal region complementary to the oligonucleotide probe. The presumptive direction of transcription in *B. gingivalis*, based on subsequent nucleotide sequencing of the clone, is indicated by the large arrow. Small arrows indicate direction of sequencing.

was covered in both directions. All cloning sites were sequenced through, and no ambiguities remained.

pUC13Bg12.1 and complete fimbrilin gene correspondence. The region of pUC13Bg12.1 that hybridized to the BGFIM2 oligonucleotide probe was localized by Southern blot analysis to the sequence bounded by the *Ssp*I site and the adjacent *Pst*I site. The sequence of this region includes a span of 20 nucleotides that corresponds exactly to one of the isomers in the mixed oligonucleotide probe BGFIM2 and that contains a *Hinc*II site (Fig. 1 and 2). As further confirmation of the uniqueness of this probe, Southern blot analysis of *Hinc*II-digested pUC13Bg12.1 failed to give a signal due to cleavage within the region recognized by the probe, since it gave fragments with regions of homology too short to form stable hybrids (not shown). This 20-base sequence defined an open reading frame which extended from base 113 to base 1277.

Beginning at base 266 is a stretch of 24 codons that codes precisely for the known N-terminal sequence of mature *B. gingivalis* 381 fimbrilin (45). The open reading frame extended downstream from the N-terminal alanine codon for a total of 337 codons, until it was terminated by TAA. Assuming no further processing of the polypeptide, this would produce a mature fimbrilin polypeptide of 35,924 daltons, or somewhat smaller than the estimate of 43,000 daltons derived from sodium dodecyl sulfate gel electrophoresis (44). One possibility was that the clone had undergone a deletion of a portion of the coding sequence. This would require the loss of approximately 175 bases during cloning while maintaining an open reading frame (unless the loss occurred at the C-terminal region). To test this possibility, we first performed detailed restriction site mapping of the *B. gingivalis* fimbrilin gene based on information obtained from the clone sequence. When this experiment was designed, small fragments were chosen to maximize the probability of detecting differences. We found no discrepancy between the clone and the genomic DNA it represented (data not shown), which strongly suggests that a deletion had not occurred during cloning. Since the coding region was sequenced on both strands by using overlapping fragments, we had no indications of a sequencing error. To rule out the possibility that we had erred in the assignment of the termination codon, the three C-terminal amino acids were determined by the carboxypeptidase method and shown to be in agreement with our sequence data (F. Yoshimura, unpublished observations). A 20% discrepancy in size could give rise to differences between the predicted amino acid composition and that found by direct amino acid analysis of the protein. There was very close agreement between these two approaches, again suggesting that a deletion or an error had not occurred (Table 1). The simplest explanation for the observed size difference is that the fimbrilin polypeptide migrated aberrantly on sodium dodecyl sulfate gels compared with globular protein size markers. Such discrepancies are not unusual; for example, chromogranin A has an apparent molecular weight of 75,000 daltons but a true size of 48,000 daltons as determined by DNA sequencing (15).

Codon usage. Codon usage has been proposed as a factor in determining the efficiency of translation of messages (10, 11). The codon usage for the fimbrilin gene is shown in Table 2. There was a marked bias toward the use of U and against the use of C in the third position of quartet codons (U:C = 59:32), consistent with favoring codon-anticodon complexes of intermediate energy (10). For individual codons, the usage was comparable to that of highly expressed genes in other bacteria. Cognate tRNAs for codons AUA, CGG, AGA, AGG, CUA, CGA, GGA, and GGG are rare in *E. coli* or interact only weakly, and the use of these codons is avoided in highly expressed proteins (10, 11). With the exception of GGA and GGG, which were used only once, the remaining codons were not used at all in the fimbrilin gene. Other codons used rarely in highly expressed bacterial genes are UCA, UCG, ACG, CUC, CUU, CCU, and CCC (10). Of these, UCG, ACG, and CCC were used rarely or not at all in the fimbrilin gene. Although the leucine codons CUU and CUC accounted for nearly half of the leucine codons, CUG was clearly the preferred codon, as has been noted for other bacterial genes (10). Exceptions in the fimbrilin gene to the general pattern of codon usage in efficiently expressed bacterial genes were the preferential use of UCA instead of UCC, use of CCU instead of CCG, and extensive use of GGA. These exceptions may indicate that the cognate tRNA species are more abundant in *B. gingivalis* than in other bacterial species.

Shine-Dalgarno, initiation codon, and leader sequences. Secretory proteins are generally synthesized with an oligopeptide sequence at the N terminus called the leader or signal peptide. This is usually 15 to 20 amino acids long and contains a highly hydrophobic domain with the general exception of at least 1 positively charged residue close to the methionyl initiation residue (16, 41, 42). The open reading frame extends upstream from the known N-terminal alanine codon for a total of 50 codons and is closed by a TAA termination codon at base 113. Since the initiation tRNA of *E. coli* can recognise XUG (including reinitiation codons [18, 27]) and the initiation codon in bacteria is bounded by certain sequence characteristics (39), including a Shine-Dalgarno sequence of the form AGGAGGU (9, 33, 39), we searched

FIG. 2. Nucleotide sequence of the region of pUC13Bg12.1 spanning the fimbrilin gene. The sequence is numbered from base 1 at the 5'-most *Hin*PI site (numbers below the sequence), and the conceptual amino acid sequence is numbered from residue 1 at the putative GTG initiation codon located at base 236 (numbers above the sequence). The known N-terminal amino acid sequence begins at residue 11. The putative leader sequence, the nucleotide sequence complementary to BGFIM2, and restriction sites used for sequencing are underlined and labeled.

TABLE 1. Comparison of the amino acid compositions of fimbrilin from *B. gingivalis* 381 derived by direct analysis and deduced from the clone pUC13 Bg 12.1

Amino acid	Amt[a]	
	Measured[b]	Deduced[c]
Asx	43	44
Thr	27	29
Ser	11	12
Glx	33	32
Pro	13	13
Gly	27	26
Ala	43	43
1/2-Cys-CM[d] Cys	3	3
Val	23	25
Met	5	6
Ile	17	17
Leu	24	24
Tyr	19	19
Phe	7	7
His	5	5
Lys	22	23
Arg	5	5
Trp	2[e]	4

[a] Molar ratio relative to 43 mol of alanine.
[b] Hydrolyzed with 6 N HCl at 110°C for 24 h after reduction, followed by *S*-carboxymethylation of the sample for cysteine and cystine.
[c] Derived from the sequence shown in Fig. 2.
[d] CM, Carboxymethylated.
[e] Derived from an analysis of methanesulfonic acid hydrolysis.

this upstream region for sequences which met these criteria, indicating potential translation initiation sites. There were four sites with homology to the 3′ end of *B. fragilis* 16S rRNA (Table 3) (43), but only two of these sites were within an appropriate distance of a potential initiation codon, i.e., ATG at base 173 and GTG at base 236. The leader sequence resulting from initiation at the ATG would be 31 residues long and hydrophilic (average hydropathy index [19], −0.66) and would contain 6 charged residues, with 4 acidic and 2 basic. This putative leader sequence therefore would not resemble any known format for secretory signal peptides. The leader generated from the GTG would be 10 residues long and hydrophilic (average hydropathy index, −0.78) and would contain two positive and no negative charges. This leader would be remarkably similar to those of the MePhe group of fimbrial subunit proteins, which are also found in *B. nodosus* (22) (Table 4). Residues 4 and 5 in the putative *B. gingivalis* leader (Lys-Thr) would be conserved, and residues 9 to 11 would be conservative substitutions. Further, to increase alignment between the MePhe subunit protein leader sequences, a gap had to be introduced in the same position as that required to align the *B. gingivalis* leader, and residue 2 in mature *B. gingivalis* fimbrilin, Phe, aligned with the N-terminal Phe of the MePhe fimbriae. However, no significant homology was found between the rest of the fimbrilin gene sequence and the MePhe subunit genes when they were compared by the homology matrix method of Pustell and Kafatos (32). No homology to other published sequences was found at the protein level in a search of the

TABLE 3. Analysis of potential Shine-Dalgarno sites and initiation codons in fimbrilin clone pUC13 Bg 12.1

Base no.	Shine-Dalgarno sequence	Initiation codon	Stormo et al. rules[a]				
			i	ii	iii	iv	v
123	ACGAGAT						
128		ATG	G	2	A	C	None
149		TTG	T	1	G	T	20
152		TTG	T	1	A	C	23
154	GGGACTT						
173		ATG	G	1	A	G	12
195	ACGAGGC						
214	AGAAGGT						
236		GTG	A	1	A	A	15
242		TTG	G	3	G	G	21

[a] With the first base of the initiation codon designated zero, the Stormo et al. (39) rules are as follows: (i) not G at −3, (ii) less than two Gs in the −1 to −7 region, (iii) A at +5, (iv) A at +10, and (v) a Shine-Dalgarno sequence within an appropriate distance of the initiation codon. The 3′-terminal sequence of *B. fragilis* 16S rRNA is 3′OH-UCUUUCCUCCAC-5′ (43).

TABLE 2. Codon usage in the *B. gingivalis* fimbrilin gene

Codon[a]	Amino acid	No.[b]	%[c]	Codon	Amino acid	No.	%	Codon	Amino acid	No.	%	Codon	Amino acid	No.	%
TTT	Phe	2	0.6	TCT	Ser	3	0.9	TAT	Tyr	9	2.7	TGT	Cys	2	0.6
TTC	Phe	5	1.5	TCC	Ser	0	0.0	TAC	Tyr	10	3.0	TGC	Cys	1	0.3
TTA	Leu	2	0.6	TCA	Ser	3	0.9	TAA	CT	1	0.3	TGA	CT	0	0.0
TTG	Leu	10	3.0	TCG	Ser	0	0.0	TAG	CT	0	0.0	TGG	Trp	4	1.2
CTT	Leu	3	0.9	CCT	Pro	9	2.7	CAT	His	2	0.6	CGT	Arg	4	1.2
CTC	Leu	3	0.9	CCC	Pro	1	0.3	CAC	His	2	0.6	CGC	Arg	1	0.3
CTA[c]	Leu	0	0.0	CCA	Pro	2	0.6	CAA	Gln	4	1.2	CGA	Arg	0	0.0
CTG	Leu	6	1.8	CCG	Pro	1	0.3	CAG	Gln	7	2.1	CGG	Arg	0	0.0
ATT	Ile	10	3.0	ACT	Thr	8	2.4	AAT	Asn	18	5.2	AGT	Ser	2	0.6
ATC	Ile	7	2.1	ACC	Thr	5	1.5	AAC	Asn	13	3.9	AGC	Ser	4	1.2
ATA	Ile	0	0.0	ACA	Thr	12	3.6	AAA	Lys	8	2.4	AGA	Arg	0	0.0
ATG	Met	6	1.8	ACG	Thr	4	1.2	AAG	Lys	15	4.5	AGG	Arg	0	0.0
GTT	Val	10	3.0	GCT	Ala	14	4.2	GAT	Asp	7	2.1	GGT	Gly	8	2.4
GTC	Val	2	0.6	GCC	Ala	12	3.6	GAC	Asp	6	1.8	GGC	Gly	7	2.1
GTA	Val	9	2.7	GCA	Ala	15	4.5	GAA	Glu	12	3.6	GGA	Gly	9	2.7
GTG	Val	4	1.2	GCG	Ala	2	0.6	GAG	Glu	9	2.7	GGG	Gly	2	0.6

[a] Underlined codons represent rate tRNA species in *E. coli* (11).
[b] Total number of residues specified by a codon.
[c] Percentage of total codons (338, including termination codon) represented by a given codon.

TABLE 4. Alignment of the putative leader sequence of the fimbrilin protein with those of the MePhe fimbrial protein family[a]

Species	Leader sequence[b]
Methylococcus bovis β	MN - - - AQKG / FT L ....
Neisseria gonorrhoeae	MNT - - LQKG / FT L ....
Bacteroides nodosus	MKS - - LQKG / FT L ....
Pseudomonas aeruginosa	MK - - - AQKG / FT L ....
	1 10
Bacteroides gingivalis[c]	MVLKTSNSNR / AFG V ....

[a] MePhe fimbrial-protein-family data are from Marrs et al. (22). Gaps have been introduced to enhance the alignment.
[b] /, Cleavage site.
[c] Residue numbers are given above the line.

Protein Identification Resource of the National Biomedical Research Foundation.

Southern blot analysis of black-pigmented *Bacteroides* species. To test the possibility that homologous genes might exist in other black-pigmented *Bacteroides* species, we performed Southern blot analysis (38) of *Pst*I-digested genomic DNA isolated from the American Type Culture Collection type strains with pUC13Bg12.1 as the probe (Fig. 3). *B. gingivalis* ATCC 33277 gave two strongly hybridizing fragments, as would be expected for a single-copy gene with two closely spaced internal *Pst*I sites. Eight other independently isolated *B. gingivalis* strains tested showed a hybridization pattern consistent with a single-copy gene, and strain 381 gave a hybridization pattern identical to ATCC 33277 (not shown). None of the other species tested gave a significant hybridization signal, even under the very-low-stringency wash conditions used. Prolonged exposure of a film resulted in the appearance of numerous weak bands in all lanes, including those containing *B. gingivalis* DNA, but their relative intensity was about 1/100 that obtained for the fimbrilin gene.

We conclude that we have isolated a recombinant clone carrying a sequence representing the gene encoding the *B. gingivalis* fimbrial subunit protein fimbrilin. Since we have failed to detect plasmids in this organism using a number of approaches (D. P. Dickinson, P. S. Rivoli, J. M. Brewster, and R. J. Genco, J. Dent. Res. **66**:223, abstr. no. 933, 1986), this gene appears to be present as a single copy in the chromosome of this organism. No similar sequences appear to be present in other black-pigmented *Bacteroides* species, and we could find no clear homology with other published fimbrial sequences, suggesting that *B. gingivalis* fimbriae are assembled from a previously undescribed type of subunit. Preliminary experiments have shown that pUC13Bg12.1 detects a ca. 1,450-base transcript in total RNA isolated from cultures of *B. gingivalis* (data not shown). Therefore, this probe will be valuable in future studies of environmental modulation of fimbriation in this organism.

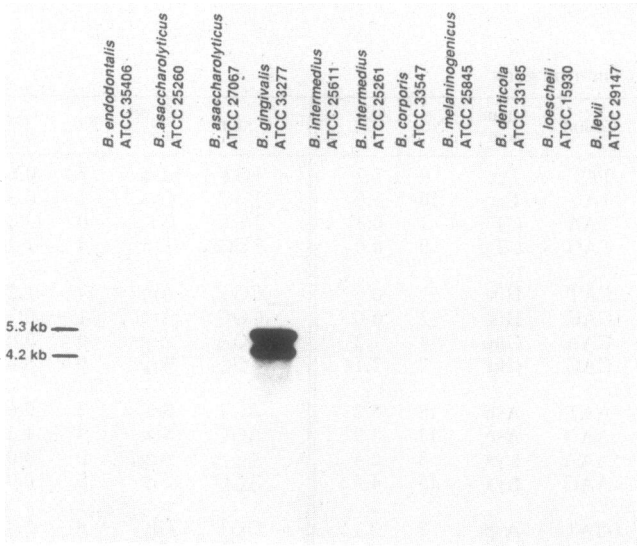

FIG. 3. Southern blot analysis of *Pst*I-digested genomic DNA from American Type Culture Collection strains of black-pigmented *Bacteroides* species isolated by using the pUC13Bg12.1 probe. The blot shown is a 6-h exposure with a screen. The weak upper band in the *B. gingivalis* lane is a trace amount of a partial digestion product. Molecular sizes (in kilobases) are indicated on the left.

ACKNOWLEDGMENTS

This work was supported in part by Public Health Service grant DE04898 from the National Institutes of Health. M.A.K. was supported in part by Public Health Service grant DE07106 from the National Institutes of Health and by a Student Research Grant from the American College of Dentists. F.Y. was supported by grant-in-aid 61570874 for scientific research from the Ministry of Education, Science and Culture of Japan.

We thank L. A. Tabak for the use of his laboratory facilities; J. Zambon, H. Reynolds, and B. Loos for providing bacterial cultures; P. Dressel for photography; and Karen Willis for providing excellent technical assistance.

LITERATURE CITED

1. **Abraham, S. N., and E. H. Beachey.** 1985. Host defenses against adhesion of bacteria to mucosal surfaces. Adv. Host Def. Mech. **4**:63–88.
2. **Beachey, E. H.** 1980. Bacterial adherence. Chapman and Hall, Ltd., London.
3. **Beachey, E. H.** 1981. Bacterial adherence: adhesin-receptor interactions mediating the attachment of bacteria to mucosal surfaces. J. Infect. Dis. **143**:325–345.
4. **Beaucage, S. L., and M. H. Caruthers.** 1981. Deoxynucleoside phosphoramides—a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. **22**:1859–1862.
5. **Boyd, J., and B. C. McBride.** 1984. Fractionation of hemagglutinating and bacterial binding adhesins of *Bacteroides gingivalis*. Infect. Immun. **45**:403–409.
6. **Dzink, J. L., A. C. R. Tanner, A. D. Haffajee, and S. S. Socransky.** 1985. Gram negative species associated with active destructive periodontal lesions. J. Clin. Periodontol. **12**:648–659.
7. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
8. **Gaastra, W., and F. K. de Graaf.** 1982. Host-specific fimbrial adhesins of noninvasive enterotoxigenic *Escherichia coli* strains. Microbiol. Rev. **46**:129–161.
9. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translation initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
10. **Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier.** 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9**:r43–r74.
11. **Grossjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimum codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199–209.
12. **Grosveld, F. G., H.-H. M. Dahl, E. deBoer, and R. A. Flavell.** 1981. Isolation of β-globin-related genes from a human cosmid library. Gene **13**:227–237.
13. **Hanahan, D.** 1983. Studies on the transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
14. **Handley, P. S., and L. S. Tipler.** 1986. An electron microscope

survey of the surface structures and hydrophobicity of oral and non-oral species of the bacterial genus *Bacteroides*. Arch. Oral Biol. **31**:325–335.

15. **Iacangelo, A., H.-U. Affolter, L. E. Eiden, E. Herbert, and M. Grimes.** 1986. Bovine chromogranin A sequence and distribution of its messenger RNA in endocrine tissues. Nature (London) **323**:82–86.
16. **Inouye, M., and S. Halegoua.** 1980. Secretion and membrane localisation of proteins in *Escherichia coli*. Crit. Rev. Biochem. **7**:339–371.
17. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2998.
18. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**: 1–45.
19. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
20. **Lindberg, F. P., B. Lund, and S. Normark.** 1984. Genes of pyelonephritogenic *E. coli* required for digalactoside specific agglutination of human cells. EMBO J. **3**:1167–1173.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Marrs, C. F., G. Schoolnik, J. M. Koomey, J. Hardy, J. Rothbard, and S. Falkow.** 1985. Cloning and sequencing of a *Moraxella bovis* pilin gene. J. Bacteriol. **163**:132–139.
23. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
24. **Meyer, T. F., N. Mlawer, and M. So.** 1982. Pilus expression in *Neisseria gonorrhoeae* involves chromosomal rearrangement. Cell **30**:45–52.
25. **Moore, L. V. H., W. E. C. Moore, E. P. Cato, R. M. Smibert, J. A. Burmeister, A. M. Best, and R. R. Ranney.** 1987. Bacteriology of human gingivitis. J. Dent. Res. **66**:989–995.
26. **Moore, W. E. C., L. V. Holdeman, R. M. Smibert, I. J. Good, J. A. Burmeister, K. G. Palcanis, and R. R. Ranney.** 1982. Bacteriology of experimental gingivitis in young adult humans. Infect. Immun. **38**:651–667.
27. **Napoli, C., L. Gold, and B. S. Singer.** 1981. Translational reinitiation in the rIIB cistron of bacteriophage T4. J. Mol. Biol. **149**:433–449.
28. **Norgren, M., S. Normark, D. Lark, P. O'Hanley, G. Schoolnik, S. Falkow, C. Svanborg-Eden, M. Baga, and B. E. Uhlin.** 1984. Mutations in *E. coli* cistrons affecting adhesion to human cells do not abolish Pap pili fiber formation. EMBO J. **3**:1159–1165.
29. **Okuda, K., J. Slots, and R. J. Genco.** 1981. *Bacteroides gingivalis*, *Bacteroides asaccharolyticus* and *Bacteroides melaninogenicus* subspecies: cell surface morphology and adherence to erythrocytes and human buccal epithelial cells. Curr. Microbiol. **6**:7–12.
30. **Okuda, K., and I. Takazoe.** 1974. Hemagglutinating activity of *Bacteroides melaninogenicus*. Arch. Oral Biol. **19**:415–416.
31. **Orndorff, P. E., and S. Falkow.** 1984. Organization and expression of genes responsible for type 1 piliation in *Escherichia coli*. J. Bacteriol. **159**:736–744.
32. **Pustell, J., and F. C. Kafatos.** 1982. A high speed, high capacity homology matrix; zooming through SV40 and polyoma. Nucleic Acids Res. **10**:4765–4782.
33. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
34. **Slots, J., and R. J. Genco.** 1979. Direct hemagglutination technique for differentiating *Bacteroides asaccharolyticus* oral strains from nonoral strains. J. Clin. Microbiol. **10**:371–373.
35. **Slots, J., and R. J. Genco.** 1984. Black-pigmented *Bacteroides* species, *Capnocytophaga* species and *Actinobacillus actinomycetemcomitans* in human periodontal disease: virulence factors in colonisation, survival and tissue destruction. J. Dent. Res. **63**:412–421.
36. **Slots, J., and R. J. Gibbons.** 1978. Attachment of *Bacteroides melaninogenicus* subsp. *asaccharolyticus* to oral surfaces and its possible role in colonization of the mouth and of periodontal pockets. Infect. Immun. **19**:254–264.
37. **Socransky, S. S., A. C. R. Tanner, A. D. Haffajee, J. D. Hillman, and J. M. Goodson.** 1981. Present status of studies on the microbial etiology of periodontal diseases, p. 1–12. *In* R. J. Genco and S. E. Mergenhagen (ed.), Host-parasite interactions in periodontal diseases. American Society for Microbiology, Washington, D.C.
38. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
39. **Stormo, G. D., T. D. Schnieder, and L. M. Gold.** 1982. Characterisation of translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2971–2996.
40. **Viera, 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.
41. **Vlasuk, G. P., S. Inouye, H. Ito, K. Itakura, and M. Inouye.** 1983. Effects of the complete removal of basic amino acid residues from the signal peptide on secretion of lipoprotein in *Escherichia coli*. J. Biol. Chem. **258**:7141–7148.
42. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
43. **Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese.** 1985. Natural relationship between Bacteroides and Flavobacteria. J. Bacteriol. **164**:230–236.
44. **Yoshimura, F., K. Takahashi, Y. Nodosaka, and T. Suzuki.** 1984. Purification and characterization of a novel type of fimbriae from the oral anaerobe *Bacteroides gingivalis*. J. Bacteriol. **160**: 949–957.
45. **Yoshimura, F., T. Takasawa, M. Yoneyama, T. Yamaguchi, H. Shiokawa, and T. Suzuki.** 1985. Fimbriae from the oral anaerobe *Bacteroides gingivalis*: physical, chemical, and immunological properties. J. Bacteriol. **163**:730–734.
46. **Zambon, J. J., H. S. Reynolds, and J. Slots.** 1981. Black-pigmented *Bacteroides* spp. in the human oral cavity. Infect. Immun. **32**:198–203.