

Porphyromonas gingivalis Genes Isolated by Screening for Epithelial Cell Attachment

M. J. DUNCAN,^{1*} S. A. EMORY,¹ AND E. C. ALMIRA²

Department of Molecular Genetics, Forsyth Dental Center, Boston, Massachusetts 02115,¹ and Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, Florida 32611²

Received 9 February 1996/Returned for modification 16 April 1996/Accepted 13 June 1996

***Porphyromonas gingivalis* is associated with chronic and severe periodontitis in adults. *P. gingivalis* and the other periodontal pathogens colonize and interact with gingival epithelial cells, but the genes and molecular mechanisms involved are unknown. To dissect the first steps in these interactions, a *P. gingivalis* expression library was screened for clones which bound human oral epithelial cells. Insert DNA from the recombinant clones did not contain homology to the *P. gingivalis* *fimA* gene, encoding fimbrillin, the subunit protein of fimbriae, but showed various degrees of homology to certain cysteine protease-hemagglutinin genes. The DNA sequence of one insert revealed three putative open reading frames which appeared to be in an operon. The relationship between *P. gingivalis* attachment to epithelial cells and the activities identified by the screen is discussed.**

The diseases collectively known as periodontitis are bacterial infections which begin with inflammation of the periodontium and progress to loss of periodontal attachment. Untreated infections can lead to alveolar bone resorption and may result in tooth loss. The change from health to disease is associated with the displacement of a predominantly gram-positive bacterial flora in the gingival sulcus by gram-negative anaerobic bacteria. A small number of gram-negative species have been consistently linked with specific forms of periodontal disease; *Porphyromonas gingivalis* is associated with chronic and severe disease in adults. The organism colonizes the surface of gingival epithelial cells, but to infiltrate and compromise the underlying tissue, it must cross the epithelial cell barrier. The mechanisms used to penetrate tissue are unknown, although a possible pathway was suggested by previous demonstrations that *P. gingivalis* attached to and invaded oral epithelial cells (9, 18, 29, 33, 34). It has been demonstrated that *P. gingivalis* fimbriae mediate attachment to saliva-coated hydroxyapatite (20, 21). In addition, monoclonal antibodies raised against purified fimbriae of *P. gingivalis* 381 inhibit adhesion to buccal epithelial cells (16). The gene encoding fimbrillin (*fimA*), the protein subunit of fimbriae, was cloned and sequenced (8), and two independently isolated *fimA* insertion mutants show a loss of ability to adhere not only to saliva-coated hydroxyapatite (25) but also to human epithelial cells (12). However, in the latter case, it is not known whether fimbrillin itself is the adhesin, fimbriae form the supporting structure for associated adhesins, or fimbria-mediated adhesion is the only mechanism. We wish to dissect the first steps in the interaction between *P. gingivalis* and epithelial cells. In this report, we describe a screening procedure for the isolation of clones containing *P. gingivalis* genes involved in epithelial cell attachment and preliminary characterization of the clones. Insert DNA from one of the clones was analyzed in detail.

MATERIALS AND METHODS

Chemicals, media, bacterial strains, and plasmids. Chemicals and antibiotics were obtained from Sigma Chemical Co., St. Louis, Mo. *Escherichia coli* was

grown aerobically at 37°C in Luria-Bertani broth (LB) or on plates of the same medium containing 1.5% agar. As required, ampicillin was added to broth or plates at 50 µg/ml (LBamp). *P. gingivalis* ATCC 33277 was grown as described previously (9). Plasmid pUC13Bgl2.1 containing the *P. gingivalis* *fimA* gene was obtained from D. Dickinson, University of Texas, Houston. Clones containing probes for hemagglutinin genes *hagA*, *-B*, *-C*, and *-D* were kindly provided by A. Progulsk-Fox, University of Florida, Gainesville.

Recombinant DNA methods. Unless otherwise stated, standard procedures for DNA manipulation were as described by Sambrook et al. (32). Enzyme digests were carried out as recommended by the suppliers. Chromosomal DNA was extracted from *P. gingivalis* ATCC 33277 by the lysis procedure described by Kado and Liu (17). DNA was precipitated from the lysate with ethanol, phenol-chloroform extracted, reprecipitated with ethanol, and dissolved in Tris-EDTA. DNA fragments from restriction enzyme digests were excised from 0.7% separating gels, and purified using GeneClean II (Bio101, La Jolla, Calif.).

Construction of a *P. gingivalis* ATCC 33277 expression library. Chromosomal DNA was partially digested with *Sau3AI* and fractionated by agarose gel electrophoresis. The portion of the gel containing 5- to 10-kb DNA fragments was excised and purified as described above. Phagemid vector pBlueScriptII SK (pBS; Stratagene, La Jolla, Calif.) was digested to completion with *Bam*HI, treated with phosphatase, and gel purified. Chromosomal fragments and vector DNA were mixed in a 2:1 ratio and ligated overnight with T4 DNA ligase. The ligation mix was used to transform calcium-competent cells of *E. coli* NM522 (Stratagene), and ampicillin-resistant transformant clones were selected by overnight growth at 37°C on LBamp plates. The transformation frequency was 6×10^5 transformants per µg of vector DNA. More than 80% of the transformants contained inserts as calculated from the number of white transformant clones on LBamp containing isopropyl-β-D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal).

Southern hybridization. DNA fragments were transferred to either Gene Screen hybridization transfer membranes (NEN Research Products, Boston, Mass.) or nitrocellulose. The enhanced chemiluminescence (ECL) gene detection system (Amersham PLC, Little Chalfont, England) was used without modification of the protocol supplied by the manufacturer. For chemiluminescent signal detection, filters were exposed to Hyperfilm-ECL (Amersham).

Screening the *P. gingivalis* library for clones which bound epithelial cells. *E. coli* recombinant clones containing the *P. gingivalis* library were replicated to nitrocellulose filters and lysed with lysozyme and sodium dodecyl sulfate (9), since it was not anticipated that expressed proteins would be transported to the *E. coli* surface. Protein from the clones was UV cross-linked to the filters (Stratalinker; Stratagene). The blocking of filters and incubation with and detection of epithelial cells were as described previously (9).

DNA sequencing. Plasmid DNA for sequencing was prepared by the modified alkali lysis/polyethylene glycol precipitation procedure for use with the *Taq* DyePrimer (no. 401386) and *Taq* DyeDeoxyTerminator (no. 401388) cycle-sequencing protocols developed by Applied Biosystems (Perkin-Elmer Corp., Foster City, Calif.). The labelled extension products were analyzed on an Applied Biosystems model 373A DNA sequencer (Perkin-Elmer Corp.). Oligonucleotide primers were synthesized at the DNA Synthesis Core Laboratory, University of Florida, Gainesville. Nucleotide sequences were aligned and assembled with programs in the Sequencher software package (Gene Codes Corp., Ann Arbor, Mich.).

* Corresponding author. Phone: (617) 262-5200. Fax: (617) 262-4021.

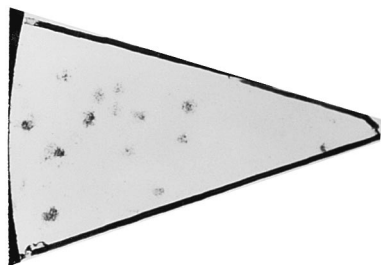


FIG. 1. Nitrocellulose filter replicate of purified *E. coli* recombinant clone isolated in the screen. Dark aggregates are alkaline phosphatase-stained KB epithelial cells attached to single colonies.

DNA and protein analysis. DNA and protein sequences were analyzed with DNAsis and Prosis (Hitachi) software. Hydropathy analysis was performed by the method of Hopp and Wood (14). Homology searches of databases were carried out with the BLAST algorithm (2).

Nucleotide sequence accession number. The DNA sequences in this study have been deposited with GenBank under accession number U60208.

RESULTS

Quality of the *P. gingivalis* library. The quality of the *P. gingivalis* library was assessed by determining the frequency of *fimA* DNA. One thousand eighty *E. coli* clones were replicated to nitrocellulose filters and treated by standard procedures to release and denature DNA, which was cross-linked to the filters with UV light. A *SacI* fragment containing the *fimA* gene was isolated from plasmid pUC13Bgl2.1 (8) and used as a probe. Of the 1,080 clones tested, 4 (1 in 270) hybridized to the *fimA* probe (data not shown), demonstrating that this DNA was represented with good frequency in the library.

Isolation of clones expressing *P. gingivalis* epithelial cell binding proteins. We used a modified version of a qualitative filter-binding assay (9, 22) to screen for *E. coli* clones expressing *P. gingivalis* genes for epithelial cell-binding proteins. Clones which bound epithelial cells were purified from the master plate and retested. After screening approximately 150,000 *E. coli* clones from two independent libraries, we identified 15 which bound KB epithelial cells. Clones from the first library were designated by numbers, and those from the second were designated by letters. Figure 1 shows dark-stained aggregates of epithelial cell binding to single colonies of one of the purified clones. Of the 15 clones, 2 failed to grow after subculture on LBamp plates and several others segregated blue colonies during growth on LBamp plates containing IPTG and X-Gal, suggesting that potentially toxic protein products were expressed. Intact cells of the *E. coli* clones did not attach to epithelial cell monolayers at levels significantly higher than those of the host strain NM522 containing pBluescript. This was not unexpected, since no attempts were made to direct expressed gene products to the bacterial cell surface.

Southern hybridization analysis of clones. Plasmid DNA from the clones was digested with *PvuII*, fractionated by agarose gel electrophoresis, and transferred to membranes for Southern hybridization. All plasmids contained 5- to 9-kb inserts displaying different *PvuII* restriction patterns (Fig. 2A), an indication that they were derived from independent recombinant clones. All the inserts hybridized to whole genomic DNA from *P. gingivalis* (data not shown). None of the inserts hybridized to the *SacI* fragment probe containing the *fimA* gene, although this probe hybridized to control *P. gingivalis* genomic DNA and the *fimA* clone, which were run on the same gel (data not shown), and clones carrying this gene were detected in the library (see above).

Because adhesins often possess hemagglutinating activity (35, 36), the same blot was probed with fragments of the hemagglutinin genes *hagA*, *-B*, *-C*, and *-D* isolated from *P. gingivalis* 381 (23, 24, 31). Figure 2B and C show the *hagA* and *-D* fragments (solid lines) used as probes and the hybridizing fragments from the inserts. The *hagA* and *-D* genes contain homology not only to each other (23) but also to cysteine protease genes (3). Cross-hybridization was also observed between the internal fragments of *hagA* and pBluescript. No homology to probes from *hagB* or *-C* could be detected.

By restriction fragment length polymorphism (RFLP) analysis, the clones fell into two classes. As shown in Fig. 3, the inserts of one class, represented by clones C and E, hybridized to single fragments of *PstI*-digested chromosomal DNA from strains ATCC 33277 and 381. The other class of inserts, represented by clones D and F, hybridized to multiple genomic fragments, indicating that they contained either repetitive sequences or multicopy genes. Strain ATCC 33277 possesses two arg-gingipain genes (26), and repetitive sequences are found both within the cysteine protease genes themselves and at flanking chromosomal sites (reviewed in reference 3).

Restriction map and RFLP analysis of p24, *hagA*, and *hagD*.

A restriction map was derived for clone p24, which contained the smallest insert and was a representative of clones which hybridized to single *PstI* genomic fragments. The map of the p24 insert was different from those of *hagA* and *-D* (Fig. 4A) and indicated that it contained a unique DNA fragment. RFLP analysis was used to define the genomic fragment of the p24 insert and to determine whether *P. gingivalis* ATCC 33277 and 381 both contained the gene(s) on this fragment, in addition to *hagA* and *-D*. Genomic DNA (5 μ g) was digested with *PstI*, fractionated on an agarose gel, and transferred to nitrocellulose. The blot was hybridized with (i) the 1.3-kb *KpnI* repeat fragments from *hagA* (Fig. 2B); (ii) the 1.9-kb internal *PstI* fragment from *hagD* (Fig. 2C), which also has homology to the *hagA* *KpnI* fragments; and (iii) the 2.5-kb *PstI* fragment of p24 (Fig. 4A). The results are shown in Fig. 4B. First, all the probes hybridized to identically sized *PstI* fragments in both ATCC 33277 and 381 genomic digests, establishing that the chromosomes of the two strains are similar with respect to these three loci. Second, the *hagA* and *-D* probes hybridized to the same genomic fragments. The 1.9-kb *PstI* probe from *hagD* hybridized to the 1.9-kb *PstI* chromosomal fragment, as expected, and also to a 13.6-kb fragment which was presumed to contain the chromosomal copy of *hagA*, since the gene does not have an internal *PstI* site. The 1.3-kb *KpnI* fragment from *hagA* hybridized to the same *PstI* genomic fragments as the *hagD* probe, as predicted by their homology (data not shown). The 2.9-kb *PstI* fragment from p24 again hybridized to the *PstI* fragments of *hagA* and *-D* and also to a new 4.2-kb band present in both ATCC 33277 and 381. The p24 insert has a single *PstI* site, and the probe was generated from this site and the *PstI* site in the pBluescript vector. Therefore, the 4.2-kb genomic fragment was generated from this internal *PstI* site and another in the flanking genome sequence. These results confirm that the p24 insert contains a different gene(s); however, under the conditions used, the *hagA* and *-D* probes did not hybridize to the p24 genomic fragment. The interpretation of this result is that the homology between the genes was too weak for detectable quantities of the *hagA* and *-D* probes to hybridize with a single genomic copy of the p24 gene, but when p24 DNA was in excess as the probe, detectable quantities hybridized to the genomic copies of *hagA* and *-D*.

DNA sequence analysis. The DNA sequence of the p24 insert (Fig. 5) confirmed the restriction map. Also shown in Fig. 5 are the deduced amino acid sequences of the putative

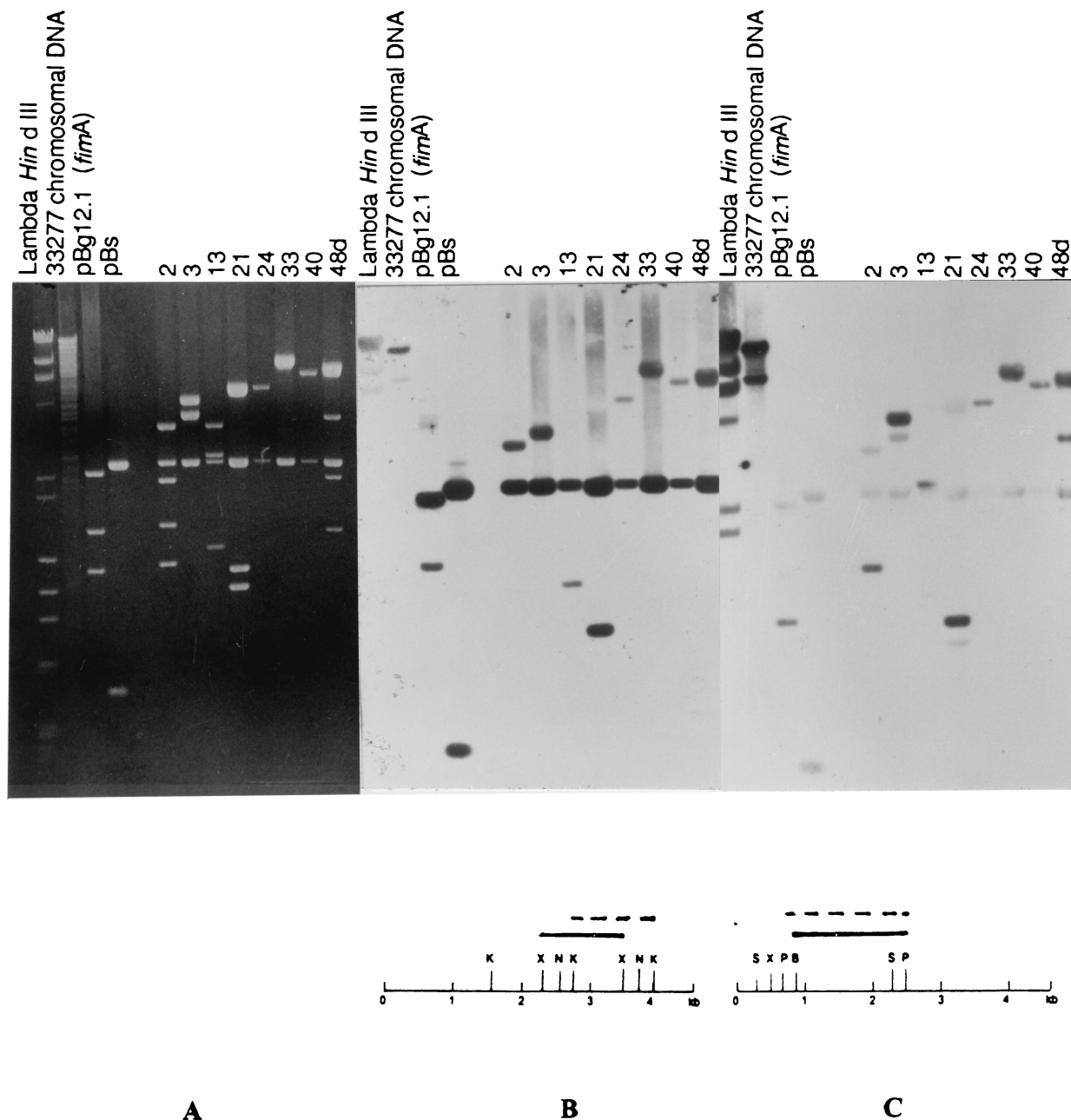
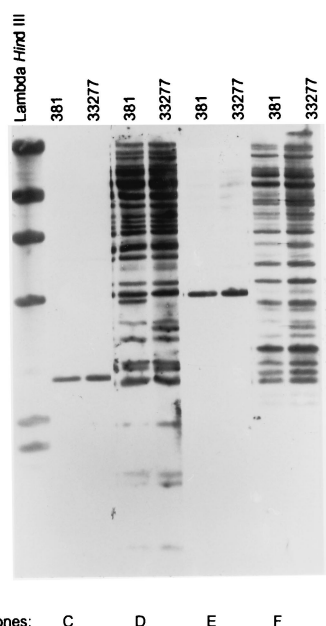


FIG. 2. Hybridization of insert DNA with *P. gingivalis* *hagA* and *-D* probes. (A) Ethidium bromide-stained *Pvu*II digests of recombinant (lanes 2 to 48d) and control plasmids. (B) Southern hybridization with *Xmn*I-digested *hagA* probe (solid overline). (C) Hybridization with *Bam*HI-*Pst*I-digested *hagD* probe (solid overline). P, *Pst*I; B, *Bam*HI; N, *Nco*I; S, *Ssp*I; K, *Kpn*I; X, *Xmn*I. Broken overlines are probes used in RFLP analyses depicted in Fig. 4.

proteins encoded by the three most significant open reading frames (ORFs), which were 36.8, 21, and 80 kDa. Upstream from the 124-ATG translational start codon of the ORF encoding the 36.8-kDa protein were candidate -35 and -10 *E. coli* promoter consensus sequences, together with a putative ribosome-binding site (37). Similar consensus promoter sequences were not found upstream (5') of the ORFs encoding the 21- and 80-kDa proteins, although possible ribosome-binding sites were identified. Terminator regions (4) were not detected in the 3' regions of the ORFs encoding the putative 36.8- and 21-kDa proteins. These observations suggest that the p24-encoding insert may contain a sequence of genes in an

operon. In the filter assay, p24 was able to bind epithelial cells without IPTG induction, indicating that binding entities may be expressed from a *P. gingivalis* promoter. Weak DNA homology (72% over 57 bp) was found between the p24 gene and *hagA* (3). However, in the p24 gene, the homology was localized to the noncoding region 5' of the ORF encoding the 80-kDa protein, while in *hagA*, the homology was within a sequence which was repeated several times (3), and a similar region is present in *hagD*. Thus, the p24 probe could detect the multicopy sequences in *hagA*, but under the hybridization conditions used, the *hagA* probe could not detect the single-copy sequence in the p24 clone. With the BLASTN (2) search al-



Probed with clones: C D E F

FIG. 3. RFLP analysis showing two classes of clones. Chromosomal DNA (5 μ g) from strains 381 and ATCC 33277 was digested with *Pst*I, subjected to electrophoresis, and transferred to nitrocellulose. Inserts from clones C, D, E, and F (from the second library) were used as probes.

gorithm, no significant homology between the p24-encoding DNA sequence and sequences deposited in the databases was observed.

Protein sequence analysis. Plots of antigenic sites (14) of the 36.8- and 80-kDa ORFs revealed N-terminal sequences containing cores of hydrophobic amino acids, and potential signal peptidase cleavage sites (37) were identified predicting signal sequences of 23 amino acids for both putative proteins (Fig. 6). The hydrophobicity profile of the 21-kDa protein was consistent with that of a cytoplasmic membrane protein, and six membrane-spanning domains in its ORF were predicted by the TopPred II program (6). BLASTP homology searches gave highest scores for positive identity between large stretches of the 36.8-kDa protein ORF and a transmembrane protein encoded by the *whiB* region of *Streptomyces coelicolor* (7). Lower similarities were observed between three regions of the 38.6-kDa protein sequence and that of a hypothetical protein encoded by the *E. coli rfb* region involved in O-specific lipopolysaccharide synthesis (40); the *rfbC* gene of *Myxococcus xanthus*, also involved in O-antigen biosynthesis (11); and an ORF in a multicistronic locus of *Sphingomonas* strain S88, which contains genes required for polysaccharide capsule biosynthesis (39). No significantly high scores were obtained for the other two ORFs.

DISCUSSION

Periodontal pathogens possess a variety of virulence factors, but relatively little is known about their interactions with gingival epithelial cells and tissue at the molecular level. After the initial contact between the bacterium and the epithelial cell, stable attachments involving specific bacterial ligands and epithelial cell receptors are formed. Bacterial surface structures such as pili, fimbriae, and outer membrane proteins with adhesion functions are known in other systems to mediate these specific attachments. This study describes a strategy to identify *P. gingivalis* genes which participate in this interaction. A qual-

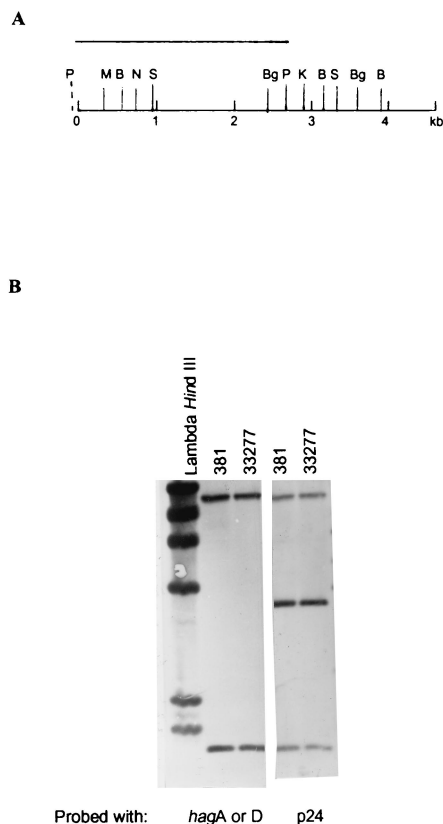


FIG. 4. Restriction map of p24 and RFLP analysis. (A) Restriction map showing the *Pst*I probe (solid overline) used in RFLP analysis. M, *Mlu*I; Bg, *Bgl*II; other enzymes are as described in the legend to Fig. 2. (B) Strain 381 and ATCC 33277 DNA (5 μ g) was digested with *Pst*I for RFLP analysis. The *hagA* and *hagD* probes are depicted as broken overlines in Fig. 2B and C.

itative filter-binding assay (9, 22) was modified and used to screen a *P. gingivalis* library for recombinant clones expressing *P. gingivalis* genes involved in attachment. The screen depends on the expression in *E. coli* of a *P. gingivalis* gene or the part of the gene which contains a binding motif for an epithelial cell receptor. Since the clones were lysed before being probed with epithelial cells, the screen does not depend on the localization of the expressed protein on the *E. coli* outer surface.

Surprisingly, none of the clones hybridized to the *P. gingivalis fimA* gene, even though we knew that this DNA was well represented, since it was used as a probe to assess library quality. A possible explanation for this result is that *fimA* is expressed poorly, if at all, in *E. coli* from the vector used in this study (41), although there are reports of detectable levels of recombinant protein expressed from similar vectors (10). Alternatively, *fimA* may have been expressed in a biologically inactive form and thus escaped detection in the screen. It is not known whether fimbriin itself is the adhesive protein which binds to the epithelial cell surface or forms the supporting structure for an associated adhesin, as in other gram-negative bacteria (15). The protein heterogeneity of *P. gingivalis* fimbrial preparations (19, 41) suggests that they are composite structures. There is immunological evidence that the hemagglutinating adhesin HA-Ag2 is distributed along the length of fimbriae in strain ATCC 33277 (5). Thus, the fimbriin itself may not participate directly in the ligand-receptor interaction. Lastly, it is likely that *P. gingivalis* does not depend solely on

-35

1 CTGACGCCGGGGGATCGTCCGAATGCTTTGGAGCCACCGCCGCTGAAAGCTCCCGG 60

-10 RBS

61 ATTGCATAATACCCGACACCATGACACCCGAGTCATCATCTCACTGGACCGGAGCT 120

121 AAGATGCTGGAGAGTTCCTCCCGCTCTTATAGCACATACACCGGACAAATCGGCTGT 180
 M L E E F L P S L I A H T P R Q S A R

181 CTTATCGTAGCGGATAATGGCTCCGACGACGACTCCGTTGAATTTCTGCGTTCCGATAT 240
 L I V A D N G S T D D S V E F L R C Q Y

241 CCACAGTGGAACTGATCCTTTCCCGGAGAATACGGCTTTCCGGAAGGCTACAACCGT 300
 P Q V E L I L Y F P E N Y G R F A E G Y N R

301 GCTATCGCACAGCCGAATGCGACTGCGTGGTACTGCTCAATAGCGATGTGGAAGTCTCC 360
 A I A Q T E C D C V V L L N S D V E V S

361 GAAGGATGGCTGGATGCTCCGCTGCGATGCTCCAGAGCGGAAGCGATGTTCCGCGTCTC 420
 F G Y D R W F A L G L Q E R S D V A A V

421 CAACCCAAAATCCGCGCTCTCGAGACAGCAGCGCTTCAGTACCGCGGCTGCTCCGCGC 480
 Q P K I R A L R D K H A F E Y A G A A G

481 GGCTATGATGATGCTGGGGTTATCCGCTCTCCGCTGACGATCTTCGATACGGTAGAA 540
 G Y V D R V A L R L G F P C R G R I F E D T V E

541 ACTGATACCGCCCAATAGATATAGCAGAGTATTCTGGGCTCCGCGTGCAGCCCTT 600
 T D T G Q Y D D I A E V F W A S G A A L

601 TTCATCCGACAGGATGATACGGCTAGGAGGATGATCCCGCTTTTTCGTCAT 660
 F I R R Q V Y M A V G G L D P R F F A H

661 CAGGAGAAATAGACTCTGTTGGCGACTCCATGCGCTGGCTGGCTATAGCCGTAGCA 720
 Q E E I D L C W R L H S R G W R I A V A

721 CCAAGTCTACTGTTTCCATTTGGGGGAGGTCCCTGATAGAAAATCCGCGCAAA 780
 P S S T V F H L G G S L D M K N P R K

781 GTGATCTCAACTCCGAAACATCTGCTGATGCTGTACAAGATTTCCGCGAAACAGCC 840
 V Y L N F R N N L L M L Y K N L P E T A

841 CTCACCGCCCAATGCGCATCCGCGCTACTGGATTCCTCTCCGCGATGTTTTCGTTG 900
 L K P T M R I R A L L D F L S A M V F L

901 CTCACCGCCCAATGCGCATATAGCGGCTGTTGCTGTCAGCTCCGCGACTTCCTTCGT 960
 L T G K F R H M Q A V V R A R R D F L R

961 ATGCGACCGCAATTCCTCCGCTCCGCGCTCCAGAACTCTCTCCGCGCTATCCCTGCA 1020
 M R P D F S A R S K N L L L G S I P A

1021 CCGATGAAGCCTTATCCATGCTCTCAATTTTATATCCGCGCGGTAAAATATCTCT 1080

299 P M K P Y S I V F Q F Y I R R R K I F S 319

1081 CGCCTCCCTAATCTTCCTCCCTCTCTGTAATAACTTTTCGTTTATAATTTGGTTTATA 1140
 R L P *

1141 GTAACATAATTATATTTGCATCGGAAGCGGTAAGAAATGCTTCCGCGAGACATTTTCAA 1200

1201 ACGGAATACGAACCACTAACGGATAGCAAAAAAATGGAAGAAAACAGCTTACACCCCAA 1260

RBS

1261 GAGACGTACAGATGATCGAAGATGCTCGAACAGACAGCAACGACTGATCCGGGGA 1320
 M I E K M L E Q T R K R L I R G

1321 GGAGGCATCCCGAGCTGATAGGGATACGCTCACTTTGGCAGAGTCTGCTGATCCTT 1380
 A G I P S L I W G Y V T F A T S L L I L

1381 TTCGTTTACCCCATATCGGTTATCGAGCAATTAATCTGTGGATGCTGATCCCGATAGTA 1440
 F V Y P H I G Y R A N Y L W M L I P I V

1441 GGAGGGAGCTGACATATTTGTAATAGGAAGACAGAAAGACATGCAAGCTACA 1500
 G S L T I I C N R K R Q K E A H A R T

1501 CAGATACGCCATTTATAGACAGGATGATACCATCGGCTGAATGTACGGCATG 1560
 Q I D R F I D T T W I T I G L N V T A L

1561 AGCATACTGGCTATCGCTCCCGTGGCTATCTCCGCTGTACTCATCTCATAGGA 1620
 S I L A Y R F P L A I L P L V L I L I G

1621 ATCGCTACGGCATACCGGATCTCGCACAGGTAACTGCTCAATAACAGCTCATA 1680
 I A T A I T G F S H K V T L L K Y S S I

1681 TTCGGCATACTGTAGGCTATATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 F G I L V G Y M L L V V P M N G K L M V

1741 CTGATCTTCGGCTTACTCTCTCTGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 L I F G L T F F L M H C V P G H Y L C Y

1801 CTCGAACGAAAATTTCCGCGAGTGTAGACTCTCGACCCGCTGCTCATGCTCCGAAT 1860
 L E R K I L R D A *

1861 GAGGTTGGCCATCATGCTCCCTCATGATGTTGGAGGAGCCGACTTCCTCTACCTCAA 1920

1921 GGAGGTGACGGGGCTACTTCGGCAACATCAGTGTACAACTGGACAACTGAGCACAGC 1980

1981 CGGCTACATCGAAATAGAAAAGGCTACAAAGGTAACGCCCGCTACCACTGTAGAGC 2040

2041 CACCGATCGCGCGTGGAGCCTTCCTCCGACATTCGAAGCACTCAAGTCTACCTGCC 2100

RBS

2101 GACAGACTCTACTACTAATCGAAAAACAAAAAGGCATGTCACATCAGGGATTCCGA 2160

2161 CAATGGTGAATCCCGTTTGAACCAATCATTAAAGTAAATGAAGCGTTTTAAATCGGAA 2220
 M K R L I S R

2221 CCCTTATCAGGCTTCTCCTCATCTTTGCATCCCTATGCTCTCGCTGCTCAGCAGACA 2280
 S T L I R L L L I L C I P H S L A A Q R Q

2281 TAGCCGTAAGTCTCAGCAGACCAAGACCGGCTATCCGTTTGGCTAATGCTCATCCGA 2340
 28 I A G T V T D T K D R P I V L A N V Y P

2341 TCAATAGCTTCGAGGATGCTGCTACCGATAGCTTGGACGGTTCGTTTACGACGAGCA 2400
 48 I N S F D G M A T D S L G R F R F T T S

2401 GCACCTTCGGGATGCTCTCGTGGTTTCGATATCAATTATCCTCCGATACGGTTACCA 2460
 68 S T F P M R L V V S H I N Y H S D T L T

2461 TTCGAAAGGAGATCCGATGGAAGCAATCCGATCCGCTGCGGAGCAGCATGACTACA 2520
 88 I R K G D P M E N I R I R L R G D L S M

2521 AGCTCCCGAAGTCTGCTGCTGCTCTCCGCTTTCCGGTAGGAGCAAGCAACAAGA 2580
 108 K L S E V V V S V S A F G V G G K Q N K

2581 CGGTGCTCAATCCTTGGATGCTACAGCAATCAAGTGGTAAAGGAGTCTCTCGATG 2640
 128 T V L N P M D V Y T N P S G N G D L S M

2641 CGCTGCTCAGACACCGCGCTGAGGATGCGGAGATGCGAGGCTTTTTCGTTCCGG 2700
 148 A L R Q T P G L Q D V G D R E G F F V R

2701 GAGGAGCCTCTTGGGAAACCGCGCTGACGATGAAGCAATAGAGTGAAGCGCTTTTTCG 2760
 168 G A S W E T A V T I E G I R V K R F F

2761 GCAAGAAGATTCGACGCTCTGCTGCTCACGTTTTCGAGACAGGAATGTCAGCGCC 2820
 188 G K N R F D A P A R S R F E T G M F S G

2821 TTTCCGCTCCACCGGAGGCTTGGGCAACGAGCGGCTGCGCTGAGCGGACTACTCC 2880
 208 L S L S T G G Y G A T E G G A L S G L L

2881 GGTGGATGGGGGTAAGTCTCCCTGCTCCGCTGGGATAGGAATATCCCGCTTTTTCG 2940
 228 R L R L A G K S P S S V G I G I S P L F

2941 TAAAGGTTGGGGTGGGCTCTTCCAGGTCGAATGTTTCTATATAGCAAAATGCT 3000
 248 V N G G G G H L S R S N R F Y I E Q N A

3001 CAGTGAAGGATCCGCTATATCCGCTGCTGCTCAAGCGGAGTACAACATGCCCGGTA 3060
 268 S V S D A A Y I R L L L K P E Y K L P G

3061 CCAACCGCTCTATGCTCAATGACAGGAGATGGAACCGGACTTCGCAAGTAGA 3120
 288 T N R S Y A Y N A R T I W N P I T S Q D E

3121 TGAAGGGCTTTTCTCTGCTACAGCCTTCTCTCTCCGCTTGGCGATACCCCTCTC 3180
 308 V K G L F L F V H D L S S S A L R I P S

3181 CTGCAACTGCATATATGCTTATCGGGAGCAACCACTATGATTCGGATGCGCTGTT 3240
 328 P A T A Y M L Y A G R N T Y G F G M A V

3241 GGAGGCATGACTTCGAAGGAGCAAAACGACTCTACCCCTGCGCCGCTATAGCAAG 3300
 348 W R H D F E G G K T T S T L S A G Y S E

3301 ACCAACCGGCTGACCGCCAGCTTGAATCGAGCGGACATAGAGAGGACTTTCCGCA 3360
 368 D H N R L H R Q L E I E P D I E S T F R

3361 CGTAGAAGCGGATGCAATATCCGCTTCCGCTTCTGATTCGGATTAAAGCTTGGAGGC 3420
 388 T V E R D A N I R L R F D S R I K A W R

3421 TCTCTACGGATCCGATATACATATAGCAAGCAAGTCTTCTGTTGGGAGGATAAAT 3480
 408 L S Y G S D Y T Y S E A K S S V R E D K

3481 CACGAGCTTACCCACTGCGGGAACACTTGTGGAGCCTATCGCAAGCCCTCTCC 3540
 428 S S A L P T L R E H L V A A Y A E A L F

3541 CTCTGCGAATGGATTCGCTACAGCGGATGCGCGCGAATATCCGGACTTACCG 3600
 448 P L S N R I S A T A G L R A E Y S G L T

3601 AGTGGCTATGCTCTACCCCTTTTGTGAGCACTTACAAGCTCTCTCTCCGAGCCGA 3660
 468 E S A I V L P R L S A T Y K L S P S S R

3661 TGACCTTGGATGAGGAGCTATGCCGCCCCGCGACTACTCTCCCATCGCAATG 3720
 488 I T L D A G G Y A A P G D Y Y L S H R I

3721 TGCCGAAAAGCGGCAACTCCGACGATACAACTGACCTATGAATGGCGGCTGCG 3780
 508 V P Q K R E H S Q Q Y N L T Y E W R P S

3781 GTACAAAGTCTGCTTCCAGCGCTATGCAAAAGATATCCCGGCTGACCACTGG 3840
 528 R T Q V L R P Q A Y D K E Y S R L T T L

3841 CTCGCGATGGCAGCCTCCCACTCCGCAAGGATATGCAAGGGGCTGACTTCTTCT 3900
 548 A P D G T A S N L G K G Y A R G L D P F

3901 GGAAGCATCCGAGTATCAGATCTTTCGAGCAITGGTCTCTCAATTCCTATACCGATG 3960
 568 W K A S G L I R S F E H W F S Y S Y T D

3961 CTCGAGGAGTATCTTTTCCCGGATCAGGACGACCGGATTCGTCGCAAGCAGCA 4020
 588 A R R Q Y L F S P D Q E R P D F V A K H

4021 CTCTTCCCGCTGCAAGTATGGTGTCTCTCTACGCTCACTATTCAAGTGTGCA 4080
 608 T L S T V L K Y W C A P I S S L F N V S

4081 TGCTTGGCTCCGGCATGACCTACCAAGTCCCAACATAGCCTCCCGGCTATTGGA 4140
 628 M S W R S G M T Y H D P N I A S P A Y L

4141 ACGCATCGTACCGGCTAATTCAGCATGAGTGCCTGTACAATACCAATCAAGTACA 4200
 648 N A S L P A N F S M S A S Y N Y P F K Y

4201 AGAAGCCGGGCTACTTGTCTCAGCGTGCACAACTGTTCAATTCGATCCCACT 4260

```

668 K K A G G V L V F S V H N L F N S D P T
4261 ACGGCTATCGCTTCGGGGGCAAAACCCATCGGAGCGGTATCCATCGGTAGAATCTC 4320
688 Y G Y R F G G Q N P S E A Y I H R L E S
4321 TACCCCTTACCGACAGTTTTATATGGTAGTCTGTTCATCAACTTCGGAATAGACCCGG 4380
708 L P L T D S F I W *
4381 ACAAGAATCATGAACCCGATATAAAAATCAAATCNAGATGAAAACCTTCAAAGAATC 4440
4441 GCCCTGCTCTCGTCGGAGGCTTTGCGAGGCTGTGTGCAACATCCGCACAAGGGGCGGCT 4500
4501 TATGCCGAAGTGATGAATCGAAAAGTAGCTGCTCTGGACAGTGTGCCACCGACGGAATAT 4560
4561 GCCACACTGGCTGCGGACTTTCCCGGATAACAGCCGTGGAAGGATCC 4608
    
```

FIG. 5. DNA sequence analysis of clone p24. Putative promoter sequences and ribosome-binding sites (RBS) are overlined. Nucleotides are numbered at both ends of each line; amino acids are numbered at one end only.

fimbria-mediated binding to epithelial cells and that other surface molecules will participate in this function.

All the clones isolated in the screen showed various levels of DNA homology to two known *P. gingivalis* hemagglutinin genes, *hagA* and *-D*, which contain homology not only to each other (23) but also to a cysteine protease gene (3, 31). Cell-bound *P. gingivalis* hemagglutinins are associated with cysteine protease activity (27), and evidence that these enzymes are also able to interact with epithelial cells comes from separate work

on their effect on *P. gingivalis* attachment (8a). In the present study, one class of cloned inserts hybridized to multiple *Pst*I genomic fragments from ATCC 33277 and 381, indicating that they contain or are located close to multicopy sequences in the genome. Numerous reports demonstrate the association of repetitive sequences with cysteine proteases. Genetic analysis established that *P. gingivalis* ATCC 33277 contained two arg-gingipain genes, *rgpA* and *-B* (26). Sequence analysis of the arg-gingipain cysteine protease genes from several strains established that they contain an active site and multiple repeat sequences designated adhesin domains (3, 28, 30). Using the *P. gingivalis* W50 arg-gingipain homolog as a probe, Aduse-Opoku et al. (1) showed that the ATCC 33277 and 381 genomes contained multiple cross-hybridizing fragments. More recently, a repeat region within the porphyain coding sequence was shown to hybridize to multiple genomic fragments, and furthermore, the gene was located upstream from a sequence related to the *P. gingivalis* insertion element *IS1126* (3). Cloned inserts which hybridized to multiple genomic fragments are being analyzed to define and characterize the repeat regions and to determine the extent of their homologies to the genes encoding cysteine proteases.

Clone p24, by contrast, was a representative of the unique, single-genomic-copy insert class. Its restriction map was not

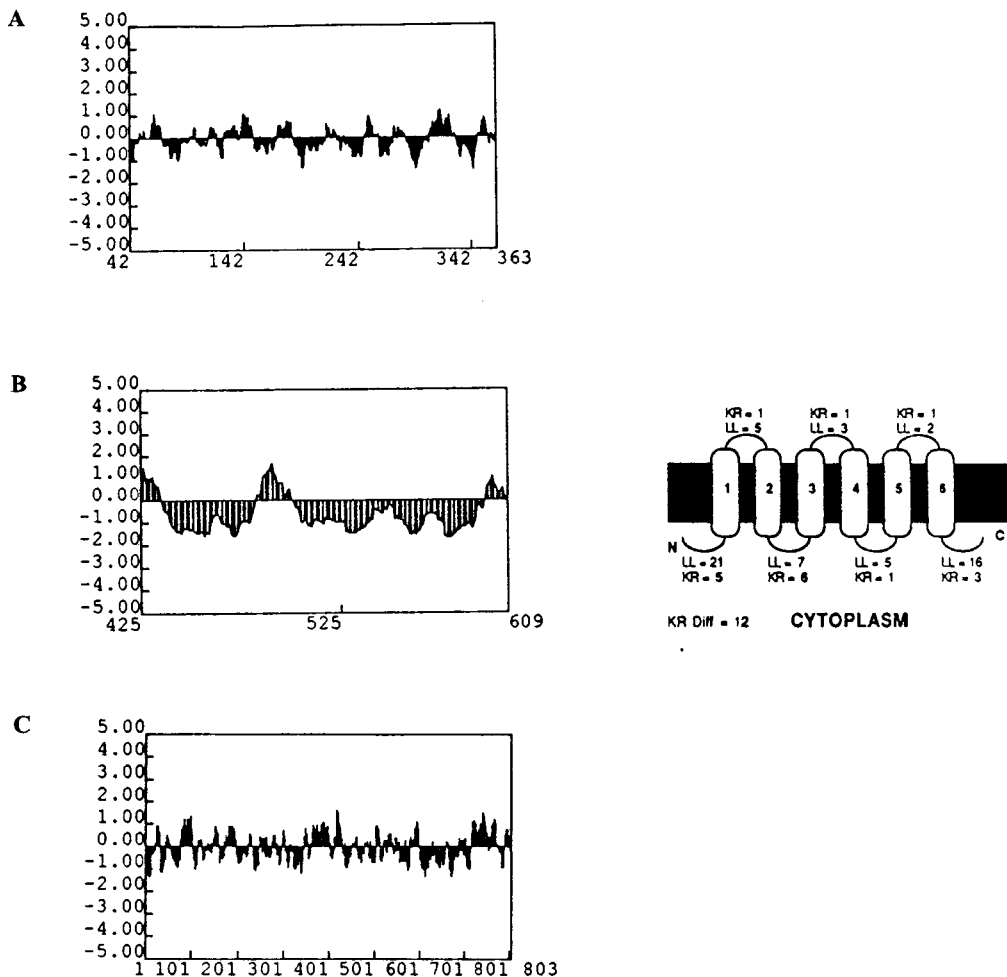


FIG. 6. Hopp and Wood plots of antigenic sites showing the hydrophathy ORFs contained in the clone p24 DNA. (A) The ORF encoding the 36.8-kDa protein. (B) The ORF encoding the 21-kDa protein, including the cytoplasmic membrane domain plot. (C) The ORF encoding the 80-kDa protein.

obviously similar to that of *hagA* or *-D*, and RFLP analysis showed that it contained a different DNA fragment. A small region of DNA homology between a p24-noncoding region and repeat 2 of the *hagA* and porphypain-coding sequence (3) was responsible for the cross-hybridization observed. The DNA insert of clone p24 contained sequences for three potential ORFs coding for proteins with molecular masses of 36.8, 21, and 80 kDa. At this time, there are no known bona fide *P. gingivalis* promoter and ribosome-binding sequences. However, a sequence with *E. coli* promoter characteristics was identified upstream from the ORF encoding the 36.8-kDa protein. Thus, one possibility is that the three ORFs are part of an operon. In anticipation of potential virulence being clustered and possibly coregulated within an operon, the library was made to contain relatively large DNA inserts. Hydrophathy plots of the deduced amino acid sequences indicated that the putative 36.8- and 80-kDa proteins contained N-terminal signal peptides, suggesting they are exported out of the cytoplasm. Only one ORF (that encoding the 36.8-kDa protein) showed significant homology to those encoding other proteins and lower but perhaps more meaningful homology to those encoding proteins from other gram-negative species which are involved in oligosaccharide synthesis. These proteins are transcribed from operons which in the case of *M. xanthus* and *Sphingomonas* strain S88 also encode ABC transporters (11, 39). Interestingly, the hydrophobicity profile of the next ORF (that encoding the 21-kDa protein) predicted a protein with six cytoplasmic membrane-spanning domains, reminiscent of the integral membrane domains of transporters (13). ATP-binding motifs (38) were not found within the sequence encoding the 21-kDa protein, although they may be contained in related ORFs flanking the p24 insert. The total size of the *P. gingivalis* transcript containing the p24-encoding ORFs is being determined by Northern (RNA) analysis.

The screening procedure was expected to identify a spectrum of *P. gingivalis* genes whose products could potentially interact with host cells and proteins, among which a subset might be involved in colonization and pathogenesis. As yet, only a small number of *P. gingivalis* activities have been identified at the gene level; therefore, the putative proteins encoded in clone p24 are unique, and their role in epithelial cell attachment remains to be determined.

ACKNOWLEDGMENT

This work was supported in part by a grant from Warner-Lambert Co., Morris Plains, N.J., to M.J.D.

REFERENCES

- Aduse-Opuku, J., J. Muir, J. M. Slaney, M. Rangarajan, and M. A. Curtis. 1995. Characterization, genetic analysis, and expression of a protease antigen (PrpRI) of *Porphyromonas gingivalis* W50. *Infect. Immun.* **63**:4744-4754.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Barkocy-Gallagher, G. A., N. Han, J. M. Patti, J. Whitlock, A. Progulsk-Fox, and M. S. Lantz. 1996. Analysis of the *prtP* gene encoding porphypain, a cysteine protease of *Porphyromonas gingivalis*. *J. Bacteriol.* **178**:2734-2741.
- Brendel, V., and E. N. Trifonov. 1984. A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res.* **12**:4411-4427.
- Chandad, F., and C. Mouton. 1995. Antigenic, structural, and functional relationships between fimbriae and the hemagglutinating adhesin HA-Ag2 of *Porphyromonas gingivalis*. *Infect. Immun.* **63**:4755-4763.
- Claros, M. G., and G. von Heijne. 1994. TopPred II: an improved software for membrane protein structure predictions. *Comput. Applic. Biosci.* **10**:685-686.
- Davis, N. K., and K. F. Chater. 1992. The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol. Gen. Genet.* **232**:351-358.
- Dickinson, D. P., M. K. Kubinec, F. Yoshimura, and R. J. Genco. 1988. Molecular cloning and sequencing of the gene encoding the fimbrial subunit of *Bacteroides gingivalis*. *J. Bacteriol.* **170**:1658-1665.
- Duncan, M. J. Unpublished data.
- Duncan, M. J., S. Nakao, Z. Skobe, and H. Xie. 1993. Interactions of *Porphyromonas gingivalis* with epithelial cells. *Infect. Immun.* **61**:2260-2265.
- Fujiwara, T., S. Morishima, I. Takahashi, and S. Hamada. 1993. Molecular cloning and sequencing of the fimbriin gene of *Porphyromonas gingivalis* strains and characterization of recombinant proteins. *Biochem. Biophys. Res. Commun.* **197**:241-247.
- Guo, D., M. G. Bowden, R. Pershad, and H. S. Kaplan. 1996. The *Myxococcus xanthus rfbABC* operon encodes an ATP-binding cassette transporter homolog required for O-antigen biosynthesis and multicellular development. *J. Bacteriol.* **178**:1631-1639.
- Hamada, N., K. Watanabe, C. Sasakawa, M. Yoshikawa, F. Yoshimura, and T. Umemoto. 1994. Construction and characterization of a *fimA* mutant of *Porphyromonas gingivalis*. *Infect. Immun.* **62**:1696-1704.
- Higgins, C. F. 1992. ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**:67-113.
- Hopp, T. P., and K. R. Wood. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **78**:3824-3828.
- Hultgren, S. J., S. Normark, and S. N. Abraham. 1991. Chaperone-assisted assembly and molecular architecture of adhesive pili. *Annu. Rev. Microbiol.* **45**:383-415.
- Isogai, H., E. Isogai, F. Yoshimura, T. Suzuki, W. Kagota, and K. Takano. 1988. Specific inhibition of an oral strain of *Bacteroides gingivalis* 381 to epithelial cells by monoclonal antibodies against the bacterial fimbriae. *Arch. Oral Biol.* **33**:479-485.
- Kado, C. I., and S.-T. Liu. 1981. Rapid procedure for detection of large and small plasmids. *J. Bacteriol.* **145**:1365-1373.
- Lamont, R. J., D. Oda, R. E. Persson, and G. R. Persson. 1992. Interaction of *Porphyromonas gingivalis* with gingival epithelial cells maintained in culture. *Oral Microbiol. Immunol.* **7**:364-367.
- Lamont, R. J., C. A. Bevan, S. Gil, R. E. Persson, and B. Rosan. 1993. Involvement of *Porphyromonas gingivalis* fimbriae in adherence to *Streptococcus gordonii*. *Oral Microbiol. Immunol.* **8**:272-276.
- Lee, J.-Y., H. T. Sojar, G. S. Bedi, and R. J. Genco. 1992. Synthetic peptides analogous to the fimbriin sequence inhibit adherence of *Porphyromonas gingivalis*. *Infect. Immun.* **60**:1662-1670.
- Lee, J.-Y., H. T. Sojar, G. S. Bedi, and R. J. Genco. 1993. Active domains of fimbriin involved in adherence of *Porphyromonas gingivalis*. *J. Periodontal Res.* **28**:470-472.
- Leong, J., L. Moitoso de Vargas, and R. R. Isberg. 1992. Binding of cultured mammalian cells to immobilized bacteria. *Infect. Immun.* **61**:2260-2265.
- Lepine, G., and A. Progulsk-Fox. 1993. Isogenic mutations in two hemagglutinin genes of *Porphyromonas gingivalis*. *J. Dent. Res.* **72**:118. (Abstract.)
- Lepine, G., and A. Progulsk-Fox. 1993. Molecular biology, p. 293-319. In H. N. Shah (ed.), *Biology of the species Porphyromonas gingivalis*. CRC Press, Inc., Boca Raton, Fla.
- Malek, R., J. G. Fisher, A. Caleca, M. Stinson, C. J. Van Oss, J.-Y. Lee, M.-I. Cho, R. J. Genco, R. T. Evans, and D. W. Dyer. 1994. Inactivation of the *Porphyromonas gingivalis fimA* gene blocks periodontal damage in gnotobiotic rats. *J. Bacteriol.* **176**:1052-1059.
- Nakayama, K., T. Kadowaki, K. Okamoto, and K. Yamamoto. 1995. Construction and characterization of arginine-specific cysteine protease (arg-gingipain)-deficient mutants of *Porphyromonas gingivalis*. Evidence for significant contribution of arg-gingipain to virulence. *J. Biol. Chem.* **270**:23619-23626.
- Nishikata, M., and F. Yoshimura. 1991. Characterization of *Porphyromonas (Bacteroides) gingivalis* hemagglutinin as a protease. *Biochem. Biophys. Res. Commun.* **178**:336-342.
- Okamoto, K., Y. Misumi, T. Kadowaki, M. Yoneda, K. Yamamoto, and Y. Ikehara. 1995. Structural characterization of argingipain, a novel arginine-specific cysteine proteinase as a major periodontal pathogenic factor from *Porphyromonas gingivalis*. *Arch. Biochem. Biophys.* **316**:917-925.
- Papapanou, P. N., J. Sandros, K. Lindberg, M. J. Duncan, R. Niederman, and U. Nanmark. 1994. *Porphyromonas gingivalis* may multiply and advance within stratified human junctional epithelium *in vitro*. *J. Periodontal Res.* **29**:374-375.
- Pavloff, N., J. Potempa, R. N. Pike, V. Prochazka, M. M. Kiefer, J. Travis, and P. J. Barr. 1995. Molecular cloning and structural characterization of the arg-gingipain proteinase of *Porphyromonas gingivalis*. Biosynthesis as a proteinase-adhesion polyprotein. *J. Biol. Chem.* **270**:1007-1010.
- Progulsk-Fox, A., V. Rao, N. Han, G. Lepine, J. Witlock, and M. Lantz. 1993. Molecular characterization of hemagglutinin genes of periodontopathic bacteria. *J. Periodontal Res.* **28**:473-474.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sandros, J., P. N. Papapanou, and G. Dahlen. 1993. *Porphyromonas gingivalis* invades oral epithelial cells *in vitro*. *J. Periodontal Res.* **28**:219-226.
- Sandros, J., P. N. Papapanou, U. Nanmark, and G. Dahlen. 1994. *Porphy-*

- romonas gingivalis* invades human pocket epithelium *in vitro*. J. Periodontal Res. **29**:62-69.
35. **Tuomanen, E., and A. Weiss.** 1985. Characterization of two adhesins of *Bordetella pertussis* for human ciliated respiratory-epithelial cells. J. Infect. Dis. **152**:118-125.
 36. **Vaisanen, V., J. Elo, L. G. Tallgren, A. Siitonen, P. H. Makela, C. Svanborg-Eden, G. Kallenius, S. B. Svensson, H. Hultberg, and T. Korhonen.** 1981. Mannose-resistant hemagglutination and P antigen recognition are characteristic of *Escherichia coli* causing primary pyelonephritis. Lancet **ii**:1366-1369.
 37. **Von Heijne, G.** 1987. Sequence analysis in molecular biology. Academic Press, Inc., New York.
 38. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gray.** 1982. Distantly related sequences in α - and β -subunits of ATP synthase, myosin kinase and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **8**:945-951.
 39. **Yamazaki, M., L. Thorne, M. Mikolajczak, R. W. Armentrout, and T. J. Pollack.** 1996. Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. J. Bacteriol. **178**:2676-2687.
 40. **Yao, Z., and M. A. Valvano.** 1994. Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotypes Y and 4a. J. Bacteriol. **176**:4133-4143.
 41. **Yoshimura, F., Y. Takahashi, E. Hibi, T. Takasawa, H. Kato, and D. P. Dickinson.** 1993. Proteins with molecular masses of 50 and 80 kilodaltons encoded by genes downstream from the fimbriin gene (*fimA*) are components associated with fimbriae in the oral anaerobe *Porphyromonas gingivalis*. Infect. Immun. **61**:5181-5189.

Editor: P. J. Sansonetti