INFECTION AND IMMUNITY, Jan. 1995, p. 238–247
0019-9567/95/$04.00+0

Vol. 63, No. 1

Revised Sequence of the *Porphyromonas gingivalis* PrtT Cysteine Protease/Hemagglutinin Gene: Homology with Streptococcal Pyrogenic Exotoxin B/Streptococcal Proteinase

THERESA E. MADDEN,[1]* VIRGINIA L. CLARK,[1] AND HOWARD K. KURAMITSU[2]

Departments of Dental Research and Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642,[1] and Department of Oral Biology, State University of New York at Buffalo, Buffalo, New York 14214[2]

Received 21 July 1994/Returned for modification 30 August 1994/Accepted 20 October 1994

The *prtT* gene from *Porphyromonas gingivalis* ATCC 53977 was previously isolated from an *Escherichia coli* clone possessing trypsinlike protease activity upstream of a region encoding hemagglutinin activity (J. Otogoto and H. Kuramitsu, Infect. Immun. 61;117–123, 1993). Subsequent molecular analysis of this gene has revealed that the PrtT protein is larger than originally reported, encompassing the hemagglutination region. Results of primer extension experiments indicate that the translation start site was originally misidentified. An alternate open reading frame of nearly 2.7 kb, which encodes a protein in the size range of 96 to 99 kDa, was identified. In vitro transcription-translation experiments confirm this size, and Northern (RNA) blot experiments indicate that the protease is translated from a 3.3-kb mRNA. Searching the EMBL protein database revealed that the amino acid sequence of the revised PrtT is similar to sequences of two related proteins from *Streptococcus pyogenes*. PrtT is 31% identical and 73% similar over 401 amino acids to streptococcal pyrogenic exotoxin B. In addition, it is 36% identical and 74% similar over 244 amino acids with streptococcal proteinase, which is closely related to streptococcal pyrogenic exotoxin B. The similarity is particularly high at the putative active site of streptococcal proteinase, which is similar to the active sites of the family of cysteine proteases. Thus, we conclude that PrtT is a 96- to 99-kDa cysteine protease and hemagglutinin with significant similarity to streptococcal enzymes.

Porphyromonas gingivalis has been shown to be a primary etiologic agent in periodontitis (24, 28, 54). Two potentially important virulence factors of *P. gingivalis* are the production of proteases and agglutination of erythrocytes (hemagglutination) (37). Proteases are able to degrade serum proteins involved in immunity which normally eradicate invading microorganisms (50, 60). *P. gingivalis* proteases also degrade connective tissue proteins directly and indirectly by activating latent human fibroblast procollagenase and neutrophil interstitial procollagenases (59).

Hemagglutination is thought to reflect the ability of a periodontal microorganism to adhere to and colonize oral tissues (55, 56), and in *P. gingivalis*, it is mediated by fimbrial and nonfimbrial molecules (6, 39, 68). Some of the hemagglutinins possess proteolytic activity, and others are found complexed with proteases (22, 41, 45), suggesting that hemagglutination may be involved in processes other than colonization.

We report the further characterization of the *P. gingivalis* trypsinlike protease gene, *prtT*, which was previously identified downstream of collagenase (*prtC*) and superoxide dismutase (*sod*) genes on a 5.9-kb *P. gingivalis* DNA fragment. Hemagglutination activity expressed from this fragment was localized to the 3′ end of the clone and believed to be encoded by a separate gene downstream of *prtT* (42). However, we demonstrated that the *prtT* gene is larger than originally reported and that it encompasses the region encoding hemagglutination.

PrtT is translated from a different open reading frame than was originally reported. The deduced amino acid sequence of the 96- to 99-kDa protein translated from the revised reading frame has significant homology to two streptococcal proteins. We present the corrected nucleotide sequence of *prtT* and accompanying sequence analysis.

MATERIALS AND METHODS

Bacterial strains, plasmid vectors, and culture conditions. *Escherichia coli* HB101 and JM109 were used as host strains for pUC18 and pUC19. JM109 was used as the host strain for pBluescript KS+. *E. coli* clones were grown on Luria agar plates. Supplements included ampicillin (100 μg/ml), X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 50 mg/ml), and 50 mM IPTG (isopropyl-β-D-thiogalactopyranoside; Sigma Chemical Co., St. Louis, Mo.). *P. gingivalis* ATCC 53977, originally designated A7A1-28, was isolated from the subgingival dental plaque of a patient with periodontitis and non-insulin-dependent diabetes from the Pima Indian Reservation in Arizona (35). Strain ATCC 53977 (A7A1-28) was grown in PY broth (23) supplemented (per liter) with 5.0 mg of hemin, 1.0 mg of vitamin K, 0.42 g of $NaHCO_3$, and 1.0 g of soluble starch and on prereduced paromomycin-vancomycin laked blood agar (Carr Scarborough Microbiologicals, Decatur, Ga.). Cultures were incubated at 37°C in an anaerobic chamber (Coy Laboratory Products, Ann Arbor, Mich.) under an atmosphere of 5% CO_2–10% H_2–85% N_2. Purity of the cultures was assessed by Gram stain and phase-contrast microscopy.

Preparation of DNA. *P. gingivalis* chromosomal DNA was prepared by using 1% cetyltrimethylammonium bromide in 0.7 M NaCl and a CsCl density gradient (1). Plasmid DNA was isolated by using Wizard or Magic Minipreps (Promega, Madison, Wis.) or a CsCl density gradient.

DNA sequencing. Double-stranded sequencing was carried out on both DNA strands (Fig. 1), using the dideoxy-chain termination method (48). Sequenase version 2.0 DNA polymerase, [^{35}S]dATP, and Sequenase kit reagents were used as recommended by the manufacturer (United States Biochemical Corp., Cleveland, Ohio). Sequencing gels (8% polyacrylamide) were run by using a model S2 sequencing gel electrophoresis system (Bethesda Research Laboratory [BRL], Gaithersburg, Md.). Cloning vectors pUC18, pUC19, and pBluescript KS+ (Stratagene, La Jolla, Calif.) were used for subcloning. Compressions were resolved by the use of 7-deaza-dGTP. Automated sequencing was performed directly on the *Hae*III inverse PCR (IPCR) product, using a model 370A DNA sequencer (Applied Biosystems, Foster City, Calif.). The 2.9-kb *Tha*I IPCR product was cloned into the *Sma*I site of pBluescript KS+. Oligonucleotide primers were synthesized with a model 380A DNA synthesizer (Applied Biosystems).

* Corresponding author. Present address: VA Medical Center, Research Section (151C), Portland, OR 97201. Phone: (503) 273-5110. Fax: (503) 273-5351.

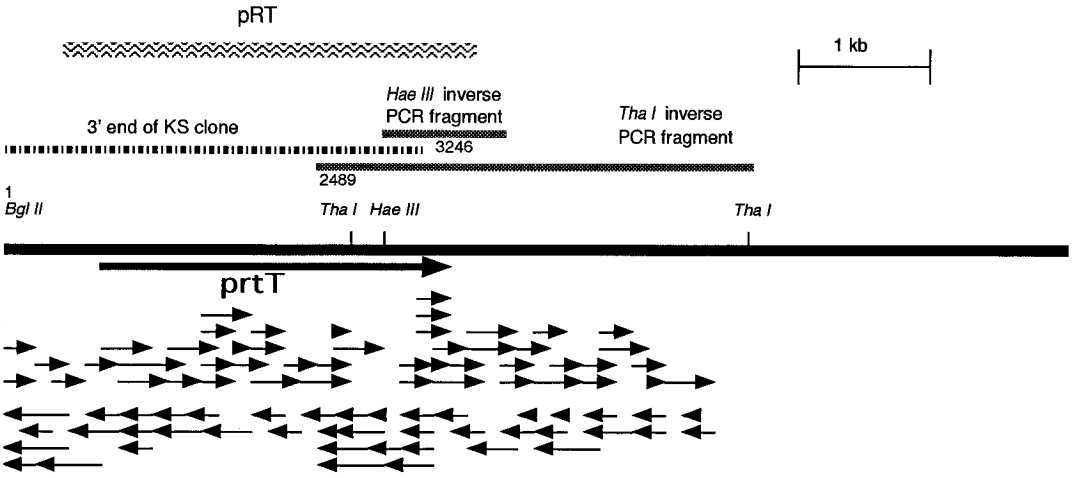


FIG. 1. Genetic map of the *prtT* locus and sequencing and cloning strategies. Nucleotide 1 occurs at the *Bgl*II site, as in the previous report (42). Sequence derived from the original pS1 clone and KS subclone extends to position 3246, and downstream sequence was derived from *Hae*III and *Tha*I IPCR fragments. Individual arrows represent the approximate region sequenced by each reaction. A clone (pRT) containing the entire PCR-amplified *prtT* open reading frame is described in the text.

IPCR. Since the entire *prtT* open reading frame was not contained on the original pS1 clone, IPCR was used to isolate contiguous downstream chromosomal DNA fragments. Two samples of 1.2 μg of strain 53977 chromosomal DNA were digested with the restriction endonucleases *Hae*III and *Tha*I (BRL) at 37 and 60°C, respectively. The fragments were precipitated, resuspended in a 350-μl ligation reaction mixture, and incubated with 5 μl of T4 ligase (BRL) at 15°C overnight; 10 μl of this product was used for each PCR as previously described. Primer pairs (5′GTAATAGTCCCCAAACAAAT3′ and 5′TATGTC CATGTAGCCATTCC3′) successfully amplified an 800-bp fragment from the *Hae*III religation product and a 2.9-kb fragment from the *Tha*I religation product.

PCR. Following identification of the *prtT* open reading frame, a subclone (pRT) which contained the entire *prtT* open reading frame was constructed (Fig. 1). PCR was used to amplify a 3.7-kb DNA molecule from *P. gingivalis* ATCC 53977 chromosomal DNA. Twenty-five cycles were carried out, using a denaturing temperature of 95°C, annealing temperature of 55°C, and extension temperature of 72°C, in a model 60 Coy Tempcycler (Coy Laboratory Products, Ann Arbor, Mich.). Similar conditions were used for amplification of IPCR products. Approximately 0.5 μg of DNA template was used in each reaction with buffer, deoxynucleoside triphosphates (dNTPs), and *Taq*I DNA polymerase (GeneAmp; Perkin-Elmer, Norwalk, Conn.) according to the manufacturer's recommendations.

Isolation of *prtT* subclones. Plasmid pS1, a 5.9-kb *P. gingivalis* clone with collagenase, benzoyl-L-arginine-*p*-nitroanilide (BAPNA) hydrolase, superoxide dismutase, and hemagglutination activities, was previously isolated (27). Subclones were constructed by *Kpn*I-*Cla*I double digestion of pS1 and ligation into the similarly digested pUC and pBluescript KS+ vectors. The pRT clone containing the entire *prtT* open reading frame was prepared by PCR with primers 5′GGGGTACCCTCAATCCTATCAAGCC3′ (*Kpn*I site plus bases 195 to 213) and 5′GCTCTAGAATTCAAGAGAGCTATGAAGGC3′ (*Xba*I site plus bases 3760 to 3777) (see Fig. 2 for nucleotide position numbers). The PCR fragment was digested with *Kpn*I and *Xba*I and ligated into pUC18. Restriction digestion with *Sca*I, *Pst*I, and *Hin*dIII as well as partial DNA sequencing confirmed that pRT contained the *prtT* gene. Three sequencing reactions (primers 5′GGGC-CGATTGCTATATAGCTTTG3′, 5′TGAGGAGTAGCAAACGGG3′, and 5′GGGGATGTGCTGCAAGGCG3′) which spanned bp 190 to 259 and 3673 to 3818 were completed to confirm that the clone contained the entire *prtT* open reading frame in the desired orientation.

Sequence analysis. DNA and amino acid sequence, amino acid composition, and Pustell DNA matrix analyses were performed, in part, with MacVector sequence analysis software (International Biotechnologies, New Haven, Conn.). Free energy of each potential stem-loop RNA structure was calculated by the method of Tinoco et al. (62). Protein homology was searched in the GenBank/EMBL sequence data library and the SwissProt and PIR protein sequence data libraries. FindPattern analysis was performed on a 12-nucleotide repeat occurring in the downstream region of the *prtT* gene (Fig. 2).

In vitro transcription-translation. Each reaction was performed with 2.5 μg of CsCl-prepared plasmid DNA, using a prokaryotic DNA-directed translation kit (Amersham, Arlington Heights, Ill.). Samples were incubated for 45 min at 37°C with [^{35}S]methionine as recommended by the manufacturer. Reaction products along with ^{14}C-labeled methylated molecular weight standard proteins (Amersham) were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. The gels were fixed for 30 min in isopropanol–H_2O–acetic acid (25:65:10) followed by 15 min in Amplify fluorographic reagent (Amersham) and dried under vacuum. Dried gels were scanned for digital imaging by using a Phosphore screen and PhosphorImager together with ImageQuant version 3.2 software (Molecular Dynamics, Sunnyvale, Calif.) and printed with a LaserJet 4 printer (Hewlett-Packard, Cupertino, Calif.).

Preparation of RNA. A single-step method was used for isolation of RNA (11). Three-day cultures of *P. gingivalis* ATCC 53977 were processed with TRI (total RNA isolation) reagent (Molecular Research Center, Inc., Cincinnati, Ohio) or total RNA isolation reagent (BRL) according to the manufacturer's directions, with the following modifications. The pellet from a 500-ml culture was resuspended in 7 ml of TRI reagent and incubated 60 min at room temperature with 50 μg of RNase-free DNase I, 100 mM $MgCl_2$, 10 mM dithiothreitol, and 10 μl of RNAsin at 37°C. The reaction was stopped by incubation at 60°C in 2.5 ml of 50 mM EDTA–1.5 M sodium acetate (pH 7.0)–1% SDS. Phenol-chloroform extraction and ethanol precipitation followed. Purified RNA samples were stored at −70°C.

Primer extension. Primer 5′GTGGCGCGAGATTTTTTTTCG3′ was end labeled with [γ-^{32}P]ATP (New England Nuclear, Boston, Mass.); 10^5 cpm was added to 1.5 μg of *P. gingivalis* RNA–10 μM dithiothreitol–10 mM dNTPs–200 U of Superscript H$^-$ reverse transcriptase enzyme (BRL) and incubated at 42°C for 30 min. The Sequenase reaction was run simultaneously on an 8% polyacrylamide gel. Dried gels were scanned for digital imaging, using a PhosphorImager and ImageQuant version 3.2 software as described above.

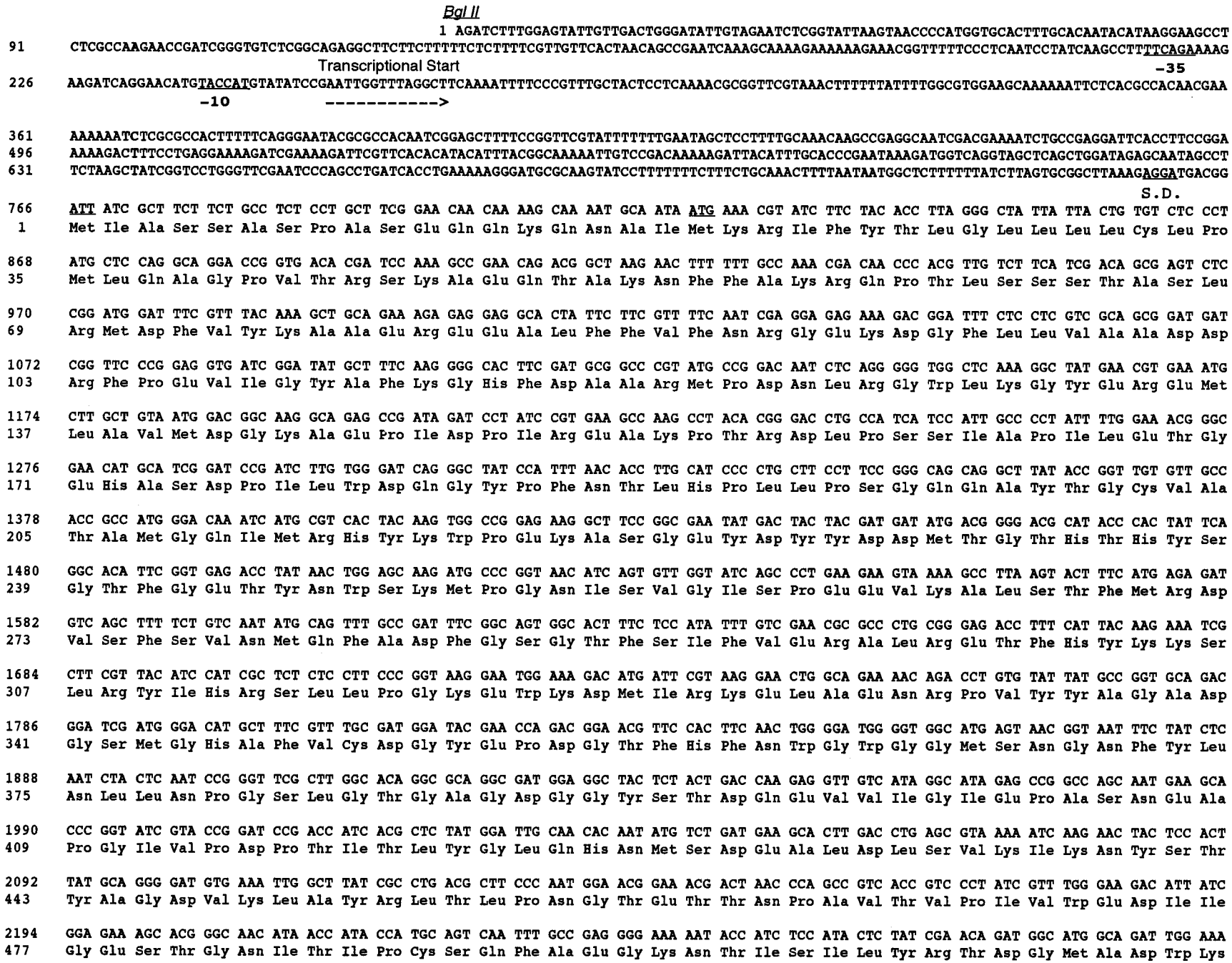
Northern (RNA) blotting. RNA was electrophoretically separated in 1.2% agarose–formaldehyde, and the gel was washed in 10× SSC (1.5 M NaCl, 0.15 M sodium citrate) to remove formaldehyde. Overnight capillary transfer of RNA onto a Hybond N filter (Amersham) was carried out. The *Kpn*I-*Hin*dIII restriction fragment of pRT containing the entire *prtT* gene was used as a probe. Probe preparation and hybridization were carried out by using ECL (enhanced chemiluminescence) direct nucleic acid labeling and detection system (Amersham) according to the manufacturer's directions. The prepared membranes were exposed to Hyperfilm ECL (Amersham).

Nucleotide sequence accession number. The *P. gingivalis prtT* gene sequence deposited in the GenBank database (accession number M83096) has been updated.

RESULTS

Identification of a transcriptional start site and potential promoter upstream of *prtT*. The results of primer extension analysis (Fig. 3) demonstrated that transcription begins at the adenine at nucleotide position 256 (Fig. 2). This rules out the possibility that translation of *prtT* originates at nucleotide 239, as previously hypothesized (42). Putative −35 (TTCAGA) and −10 (TACCAT) promoter sequences with homology to the *E. coli* consensus sequence were identified upstream of the transcription start site (Fig. 2).

Nucleotide sequence of the *prtT* gene. The nucleotide sequence downstream of the *prtT* gene, including DNA frag-




```
                                                    Bgl II
                                                    1 AGATCTTTGGAGTATTGTTGACTGGGATATTGTAGAATCTCGGTATTAAGTAACCCCATGGTGCACTTTGCACAATACATAAGGAAGCCT
91      CTCGCCAAGAACCGATCGGGTGTCTCGGCAGAGGCTTCTTCTTTTTCTCTTTTCGTTGTTCACTAACAGCCGAATCAAAGCAAAAGAAAAAAGAAACGGTTTTTCCCTCAATCCTATCAAGCCTTTTCAGAAAAG
                                                    Transcriptional Start                                                                  -35
226     AAGATCAGGAACATGTACCATGTATATCCGAATTGGTTTAGGCTTCAAAATTTTCCCGTTTGCTACTCCTCAAAACGCGGTTCGTAAACTTTTTTATTTTGGCGTGGAAGCAAAAAATTCTCACGCCACAACGAA
                                -10             ---------->

361     AAAAAATCTCGCGCCACTTTTTCAGGGAATACGCGCCACAATCGGAGCTTTTCCGGTTCGTATTTTTTTGAATAGCTCCTTTTGCAAACAAGCCGAGGCAATCGACGAAAATCTGCCGAGGATTCACCTTCCGGA
496     AAAAGACTTTCCTGAGGAAAAGATCGAAAAGATTCGTTCACACATACATTTACGGCAAAAATTGTCCGACAAAAAGATTACATTTGCACCCGAATAAAGATGGTCAGGTAGCTCAGCTGGATAGAGCAATAGCCT
631     TCTAAGCTATCGGTCCTGGGTTCGAATCCCAGCCTGATCACCTGAAAAAGGGATGCGCAAGTATCCTTTTTTTCTTTCTGCAAACTTTTAATAATGGCTCTTTTTTATCTTAGTGCGGCTTAAAGAGGATGACGG
                                                                                                                                        S.D.
766     ATT ATC GCT TCT TCT GCC TCT CCT GCT TCG GAA CAA CAA AAG CAA AAT GCA ATA ATG AAA CGT ATC TTC TAC ACC TTA GGG CTA TTA TTA CTG TGT CTC CCT
  1     Met Ile Ala Ser Ser Ala Ser Pro Ala Ser Glu Gln Gln Lys Gln Asn Ala Ile Met Lys Arg Ile Phe Tyr Thr Leu Gly Leu Leu Leu Leu Cys Leu Pro

868     ATG CTC CAG GCA GGA CCG GTG ACA CGA TCC AAA GCC GAA CAG ACG GCT AAG AAC TTT TTT GCC AAA CGA CAA CCC ACG TTG TCT TCA TCG ACA GCG AGT CTC
 35     Met Leu Gln Ala Gly Pro Val Thr Arg Ser Lys Ala Glu Gln Thr Ala Lys Asn Phe Phe Ala Lys Arg Gln Pro Thr Leu Ser Ser Ser Thr Ala Ser Leu

970     CGG ATG GAT TTC GTT TAC AAA GCT GCA GAA AGA GAG GAG GCA CTA TTC TTC GTT TTC AAT CGA GGA GAG AAA GAC GGA TTT CTC CTC GTC GCA GCG GAT GAT
 69     Arg Met Asp Phe Val Tyr Lys Ala Ala Glu Arg Glu Glu Ala Leu Phe Phe Val Phe Asn Arg Gly Glu Lys Asp Gly Phe Leu Leu Val Ala Ala Asp Asp

1072    CGG TTC CCG GAG GTG ATC GGA TAT GCT TTC AAG GGG CAC TTC GAT GCG GCC CGT ATG CCG GAC AAT CTC AGG GGG TGG CTC AAA GGC TAT GAA CGT GAA ATG
 103    Arg Phe Pro Glu Val Ile Gly Tyr Ala Phe Lys Gly His Phe Asp Ala Ala Arg Met Pro Asp Asn Leu Arg Gly Trp Leu Lys Gly Tyr Glu Arg Glu Met

1174    CTT GCT GTA ATG GAC GGC AAG GCA GAG CCG ATA GAT CCT ATC CGT GAA GCC AAG CCT ACA CGG GAC CTG CCA TCA TCC ATT GCC CCT ATT TTG GAA ACG GGC
 137    Leu Ala Val Met Asp Gly Lys Ala Glu Pro Ile Asp Pro Ile Arg Glu Ala Lys Pro Thr Arg Asp Leu Pro Ser Ser Ile Ala Pro Ile Leu Glu Thr Gly

1276    GAA CAT GCA TCG GAT CCG ATC TTG TGG GAT CAG GGC TAT CCA TTT AAC ACC TTG CAT CCC CTG CTT CCT TCC GGG CAG CAG GCT TAT ACC GGT TGT GTT GCC
 171    Glu His Ala Ser Asp Pro Ile Leu Trp Asp Gln Gly Tyr Pro Phe Asn Thr Leu His Pro Leu Leu Pro Ser Gly Gln Gln Ala Tyr Thr Gly Cys Val Ala

1378    ACC GCC ATG GGA CAA ATC ATG CGT CAC TAC AAG TGG CCG GAG AAG GCT TCC GGC GAA TAT GAC TAC TAC GAT GAT ATG ACG GGG ACG CAT ACC CAC TAT TCA
 205    Thr Ala Met Gly Gln Ile Met Arg His Tyr Lys Trp Pro Glu Lys Ala Ser Gly Glu Tyr Asp Tyr Tyr Asp Asp Met Thr Gly Thr His Thr His Tyr Ser

1480    GGC ACA TTC GGT GAG ACC TAT AAC TGG AGC AAG ATG CCC GGT AAC ATC AGT GTT GGT ATC AGC CCT GAA GAA GTA AAA GCC TTA AGT ACT TTC ATG AGA GAT
 239    Gly Thr Phe Gly Glu Thr Tyr Asn Trp Ser Lys Met Pro Gly Asn Ile Ser Val Gly Ile Ser Pro Glu Glu Val Lys Ala Leu Ser Thr Phe Met Arg Asp

1582    GTC AGC TTT TCT GTC AAT ATG CAG TTT GCC GAT TTC GGC AGT GGC ACT TTC TCC ATA TTT GTC GAA CGC GCC CTG CGG GAG ACC TTT CAT TAC AAG AAA TCG
 273    Val Ser Phe Ser Val Asn Met Gln Phe Ala Asp Phe Gly Ser Gly Thr Phe Ser Ile Phe Val Glu Arg Ala Leu Arg Glu Thr Phe His Tyr Lys Lys Ser

1684    CTT CGT TAC ATC CAT CGC TCT CTC CTT CCC GGT AAG GAA TGG AAA GAC ATG ATT CGT AAG GAA CTG GCA GAA AAC AGA CCT GTG TAT TAT GCC GGT GCA GAC
 307    Leu Arg Tyr Ile His Arg Ser Leu Leu Pro Gly Lys Glu Trp Lys Asp Met Ile Arg Lys Glu Leu Ala Glu Asn Arg Pro Val Tyr Tyr Ala Gly Ala Asp

1786    GGA TCG ATG GGA CAT GCT TTC GTT TGC GAT GGA TAC GAA CCA GAC GGA ACG TTC CAC TTC AAC TGG GGA TGG GGT GGC ATG AGT AAC GGT AAT TTC TAT CTC
 341    Gly Ser Met Gly His Ala Phe Val Cys Asp Gly Tyr Glu Pro Asp Gly Thr Phe His Phe Asn Trp Gly Trp Gly Gly Met Ser Asn Gly Asn Phe Tyr Leu

1888    AAT CTA CTC AAT CCG GGT TCG CTT GGC ACA GGC GCA GGC GAT GGA GGC TAC TCT ACT GAC CAA GAG GTT GTC ATA GGC ATA GAG CCG GCC AGC AAT GAA GCA
 375    Asn Leu Leu Asn Pro Gly Ser Leu Gly Thr Gly Ala Gly Asp Gly Gly Tyr Ser Thr Asp Gln Glu Val Val Ile Gly Ile Glu Pro Ala Ser Asn Glu Ala

1990    CCC GGT ATC GTA CCG GAT CCG ACC ATC ACG CTC TAT GGA TTG CAA CAC AAT ATG TCT GAT GAA GCA CTT GAC CTG AGC GTA AAA ATC AAG AAC TAC TCC ACT
 409    Pro Gly Ile Val Pro Asp Pro Thr Ile Thr Leu Tyr Gly Leu Gln His Asn Met Ser Asp Glu Ala Leu Asp Leu Ser Val Lys Ile Lys Asn Tyr Ser Thr

2092    TAT GCA GGG GAT GTG AAA TTG GCT TAT CGC CTG ACG CTT CCC AAT GGA ACG GAA ACG ACT AAC CCA GCC GTC ACC GTC CCT ATC GTT TGG GAA GAC ATT ATC
 443    Tyr Ala Gly Asp Val Lys Leu Ala Tyr Arg Leu Thr Leu Pro Asn Gly Thr Glu Thr Thr Asn Pro Ala Val Thr Val Pro Ile Val Trp Glu Asp Ile Ile

2194    GGA GAA AGC ACG GGC AAC ATA ACC ATA CCA TGC AGT CAA TTT GCC GAG GGG AAA AAT ACC ATC TCC ATA CTC TAT CGA ACA GAT GGC ATG GCA GAT TGG AAA
 477    Gly Glu Ser Thr Gly Asn Ile Thr Ile Pro Cys Ser Gln Phe Ala Glu Gly Lys Asn Thr Ile Ser Ile Leu Tyr Arg Thr Asp Gly Met Ala Asp Trp Lys
```

```
2296   GAG CTT AAG CAC ATA CTT ATG GGG CTT GTC AAT AAG ATA GAA GTG ACT ATG CCT GCC GGA GAT GTA GCT TAC TCG GTT GCC GAT GCC AGG ATC GTT CTC AAA
511    Glu Leu Lys His Ile Leu Met Gly Leu Val Asn Lys Ile Glu Val Thr Met Pro Ala Gly Asp Val Ala Tyr Ser Val Ala Asp Ala Arg Ile Val Leu Lys

2398   GAC GGT TCT CTT TCG CAC AAC CTG AAA GCA TAT TCC GAC TGC AAA CTC AGT GCC ACA GTC TAC AAC CCG GGT ACG GAA GAA TTT CGT AGT CGC GTC ACC TTT
545    Asp Gly Ser Leu Ser His Asn Leu Lys Ala Tyr Ser Asp Cys Lys Leu Ser Ala Thr Val Tyr Asn Pro Gly Thr Glu Glu Phe Arg Ser Arg Val Thr Phe

2450   GCC CTC CGC AAT ACA GAG GGA CGG CTC TAT TTC CTT GGC AGA CAT TTA GTA GAA TTA CAC CCG GGA GAT GAA GAC GGC GAA AAA GTT TCG CTG ACA ATA ACA
579    Ala Leu Arg Asn Thr Glu Gly Arg Leu Tyr Phe Leu Gly Arg His Leu Val Glu Leu His Pro Gly Asp Glu Asp Gly Glu Lys Val Ser Leu Thr Ile Thr

2602   GGG CTC AAG GCT CGT GCA GGA CAA TAC ATG CTT GTC TGT ACG GGC GAT ATG GAA TTG CTT ATG GAA GAT GCC TCA TGG ATA GAG TTG GCT TCC ATA GAA GTA
613    Gly Leu Lys Ala Arg Ala Gly Gln Tyr Met Leu Val Cys Thr Gly Asp Met Glu Leu Leu Met Glu Asp Ala Ser Trp Ile Glu Leu Ala Ser Ile Glu Val

2704   GCA GAG CAT ACG AGC ACC CAT TCA TCG TTA CTG GTG GCC TCC AAC CCA CAG ATC GAT CTT CTC ACA GTT CAT CGG GCC AAT CCC GAG ACA TTG CCG ACT TTC
647    Ala Glu His Thr Ser Thr His Ser Ser Leu Leu Val Ala Ser Asn Pro Gln Ile Asp Leu Leu Thr Val His Arg Ala Asn Pro Glu Thr Leu Pro Thr Phe

2806   AGC ATT ACA AAT GAA GGT GGT GCT ACT TTC TCC GGG AAA ATC GAA ATA GTG GCT ATA AAG GCT TTC TCG GAA ACT TTC TTC CAA GCG AAA GAA GAA CAC ATG
681    Ser Ile Thr Asn Glu Gly Gly Ala Thr Phe Ser Gly Lys Ile Glu Ile Val Ala Ile Lys Ala Phe Ser Glu Thr Phe Phe Gln Ala Lys Glu Glu His Met

2908   AGT CTC GCC CAA GGG GAA ACC AAA GTA TTG TCT CCG GAG CTG ACT GCG AAC TCT TCT CTC TAT ACA AAT GCC GAA CTC TTT CCC GAT GGC ATC TAT TAC ATT
715    Ser Leu Ala Gln Gly Glu Thr Lys Val Leu Ser Pro Glu Leu Thr Ala Asn Ser Ser Leu Tyr Thr Asn Ala Glu Leu Phe Pro Asp Gly Ile Tyr Tyr Ile

3004   GTC ATC AGA GAG CAG GGA TTT TGG GAT CCG ATC GAT TTG TTT GGG GAC TAT TAC TAT CGT ATC CGT CTC ATT ACG GAT CTA TCC TCT TCG GAC ATC GCT GGT
749    Val Ile Arg Glu Gln Gly Phe Trp Asp Pro Ile Asp Leu Phe Gly Asp Tyr Tyr Tyr Arg Ile Arg Leu Ile Thr Asp Leu Ser Ser Ser Asp Ile Ala Gly

3112   AAG GAT GTT TCT ACT ATA GTA CTT TAT CCC AAT CCT GCT CAC GAC TAT GTC CAT GTA GCC ATT CCT CCC ACA TAT GCG GGC AGC ACA CTT CGT TTG TTC GAT
783    Lys Asp Val Ser Thr Ile Val Leu Tyr Pro Asn Pro Ala His Asp Tyr Val His Val Ala Ile Pro Pro Thr Tyr Ala Gly Ser Thr Leu Arg Leu Phe Asp

3214   ATT CAA GGG CGA ATG CAG CTC TCC ACG AAG ATC CGA ATC TGC CGA TAT GCG TCT CGA CGT CGA ACG TCT TCC GAA GGG CAC CTA TAT CGT TGT GGT AGA AGA
817    Ile Gln Gly Arg Met Gln Leu Ser Thr Lys Ile Arg Ile Cys Arg Tyr Ala Ser Arg Arg Arg Thr Ser Ser Glu Gly His Leu Tyr Arg Cys Gly Arg Arg

3316   CAT GGT AGG GAA GCT CTT TAT TCG CTG ATT CAG TCT GTC AAC AGC AAA ACC GAC AAG GGC AAT GCC GGA TGG GAG ATC CCC TCA GGG CAT TGC CCT TGT ACC
851    His Gly Arg Glu Ala Leu Tyr Ser Leu Ile Gln Ser Val Asn Ser Lys Thr Asp Lys Gly Asn Ala Gly Trp Glu Ile Pro Ser Gly His Cys Pro Cys Thr

3418   GTT TGT TAG GATATACCAAAGAAAGGAGATAAAGAGTTCGGTTTGCTTCGTAATAATAGATGATTTTCCTTGTCATGCCGCTTTATGTCATTGATTATACCTAAATACACGAAATAATC
885    Val Cys End

3537
GGCTGACACACAATATAATGAATGAATAAAAGCACCTGTATATTGAGCAAAGGTCTCTATAAGAAAACAAAAGCTATATAGCTTTCGGTCGATTCCTACATAGGATTGAGCTGATTACTATATAGCTTT
                                                                              ------------>    stem-loop #1 =    -16.4      <------------
                                                                                                                             --------->
3666   CGTTTGATTGCTATATAGGATTGGGCCGATTGCTATATAGCTTTGTTTACACTCCTATATAGTCATCCCCTAAAAAGTCCCTTTTAGGGAGATATTGTCAGCAAACAGTATCTCACAGATTG

       #2 = -12.4<---------          --------> #4 = -7.4  <--------     -------->-13.4 <--------
                                                                              #5
                ---------->       #3 = -12.1            <----------

3788   TTTTGATAAAGTCCAAAAGAGAGCCTTCATAGCTCTCTTGAATTCATAGGCTGCTGCCGCAAGAAGTATGTTCACGGTGTCCACTCTCATACCTTTGTATAAGTTGTAACCTAAGCGATAACCTTCTCTTTTTC
       -----> -0.7 <-----
              #6
3922   TTAGATGATTTTCCTTGTCGTGCCATAGCTTTGTGTCATTGATCATACCATAAACACACGGAATAATCGGCTGATACGCAAATAATAAATGAATAAAAGCACCTATGTATCG
4034   AGGGCTTTTTCATGTGCAATTCC
```

FIG. 2. Nucleotide and deduced amino acid sequences of the *P. gingivalis prtT* gene. Transcriptional start site, ribosome binding site (SD), and potential promoter are indicated. Potential translational start codons are underlined at nucleotide positions 766 and 820. Downstream of *prtT* are seven regions of dyad symmetry (arrows), three direct repeats of 12 nucleotides (underlined), and a pair of long direct repeats (boldface). A decanucleotide sequence also found downstream of papain is underlined. Free energies are indicated for the six potential stem-loop termination structures. Nucleotide 1 at the *Bgl*II site is identical to that reported previously (42).

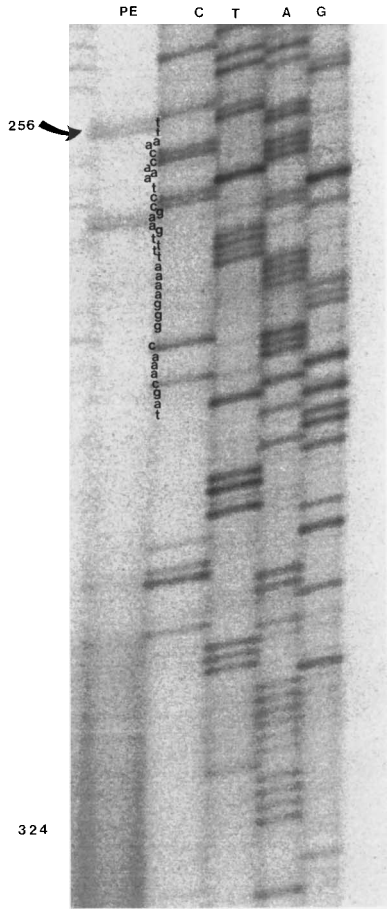


FIG. 3. Primer extension (PE) indicates the start of transcription at nucleotide 256.

ments encompassing the 3′ end of the gene, was found to differ from the published sequence. Therefore, the entire gene was resequenced. Nine single-nucleotide corrections were made between the *Bgl*II site at position 1 and the *Bam*HI site at position 1280. In addition, the previously published sequence between the *Bam*HI sites at 1280 and 1740 was incorrect. Multiple overlapping sequence reactions involving several different subclones and PCR products yielded identical sequence results (Fig. 1). Hence, the corrected sequence is presented here with confidence. In the original cloned fragment (plasmid pS1), no stop codon for the *prtT* gene was identified. Therefore, sequence downstream of the cloned fragment was obtained by IPCR. The stop codon TAG was then located at position 3426. The G+C content of the *prtT* gene is 46.3%, which corresponds to that expected for *P. gingivalis* (52).

The *prtT* gene product and identification of potential start of translational. Potential start codons for the *prtT* open reading frame were identified at positions 820 (ATG) and 766 (ATT). The former has no identifiable ribosome binding site. It does, however, encode a characteristic signal sequence, 5′MKRIF YTLGLLLLCLPMLQAGPV3′ (65). The predicted size of the gene product translated from position 820 is 96.4 kDa.

On the other hand, translation may start at position 766, utilizing the alternate start codon ATT. Seven nucleotides upstream is the sequence AGAGGA, a possible Shine-Dalgarno sequence (20). However, the 5′ amino amino acid sequence (MIASSASPASEQQKQNAIM) does not appear to be a signal

TABLE 1. Amino acid composition of PrtT[a]

Amino acid	No. (%) of residues
Nonpolar	
Ala	70 (7.9)
Val	44 (5.1)
Leu	77 (8.9)
Ile	52 (5.9)
Pro	45 (5.1)
Met	25 (2.8)
Phe	37 (4.2)
Trp	12 (1.4)
Polar	
Gly	74 (8.5)
Ser	69 (7.8)
Thr	60 (6.9)
Cys	11 (1.3)
Tyr	41 (4.7)
Asn	33 (3.7)
Gln	22 (2.5)
Acidic	
Asp	48 (5.5)
Glu	57 (6.6)
Basic	
Lys	40 (4.5)
Arg	44 (5.1)
His	24 (2.8)

[a] Based on start of translation at nucleotide position 766.

sequence. If translation begins at this site, the primary translation product is 98.3 kDa. The amino acid composition of the *prtT* gene product translated from position 766 is shown in Table 1 and has a deduced pI of 5.48.

Proteins obtained from in vitro transcription-translation experiments support our conclusion about the size of the *prtT* gene product. The in vitro product made from the pRT plasmid template containing the entire *prtT* gene is in the size range of 96 to 99 kDa (Fig. 4b, lane 4). The in vitro product made from the KS plasmid template (Fig. 4b, lane 5, and 4a, lane 2) is approximately 108 kDa. Since KS does not contain a stop codon for the *prtT* open reading frame, the in vitro transcription-translation reads through to the first in-frame stop codon within the pUC vector, located 469 nucleotides past the 3′ end of the insert. The predicted molecular size of this chimeric open reading frame is 107.7 kDa.

Size of the *prtT* transcript. Results of Northern blot analysis indicate that the transcript in *P. gingivalis* is approximately 3.3 kb (Fig. 5). This finding supports our hypothesis that transcription begins at position 256 and continues through the stop codon at position 3424. Transcription most likely ends within the downstream region of potential secondary structure described below.

Homology results. The PrtT deduced amino acid sequence was found to have 30.7% identity and 72.6% similarity over 401 amino acids with streptococcal pyrogenic exotoxin B (SPE B) and 36.1% identity and 74% similarity over 244 amino acids with streptococcal proteinase (Fig. 6). Amino acid identity of PrtT was particularly high in the putative active site of the streptococcal proteinase. When aligned with the active center of other cysteine proteases, striking conservation was found at the reactive cysteine, histidine, and glutamine residues (Fig. 7). Significant homology of PrtT with these other proteases was not found beyond amino acids of the active site cleft.

Potential secondary structures. Downstream from the stop codon for the *prtT* gene are six regions of overlapping dyad symmetry, which may be rho-dependent transcription termina-

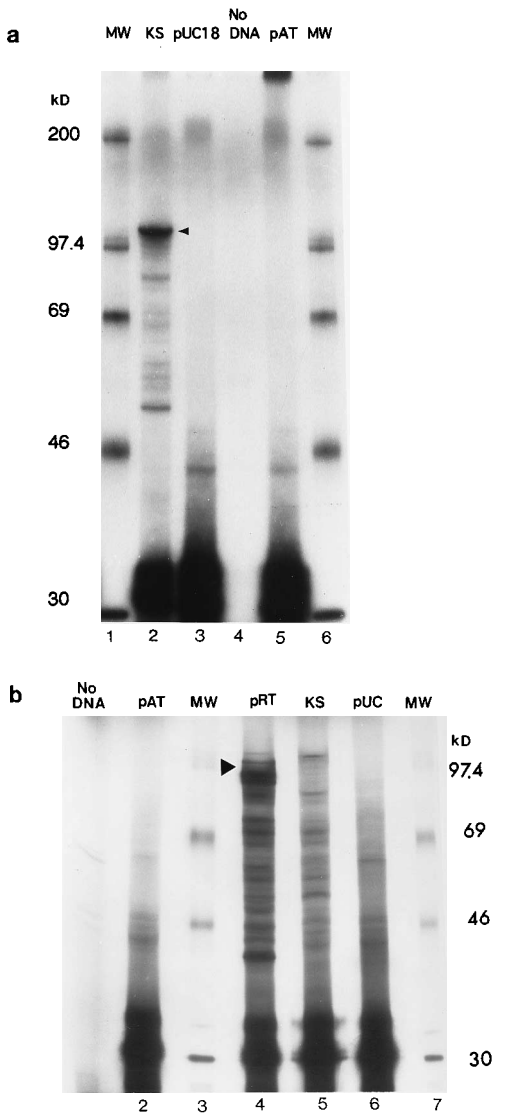


FIG. 4. (a) In vitro transcription-translation. Lanes: 1 and 6, molecular weight (MW) standards; 2 to 5, 108-kDa product obtained by using the KS template (lane 2) and control templates (lanes 3 to 5) (pAT is an Amersham kit control template). (b) In vitro transcription-translation. The product is approximately 96 to 99 kDa with use of the pRT template (lane 4); other lanes are as in panel a.

tors (33, 44). The calculated free energies (62) of the potential RNA hairpin loops formed by these sequences are −16.4, −12.4, −12.1, −7.4, −13.4, and −0.7 kcal (1 kcal = 4.184 kJ), respectively, indicating that loops 1, 2, 3, 5, and possibly 4 form in vivo. Overlaps within the stem regions, however, would limit the number of hairpins to three at any given time (hairpins 1, 3, and 5 or 2, 4, and 5). Within this same region is a 12-nucleotide direct repeat, CTATATAGCTTT, which occurs three times within the overlapping stems of the potential RNA stem-loop secondary structure (underlined in Fig. 2). Results of a FindPattern analysis of this 12-nucleotide sequence revealed no significant similarity to other prokaryotic sequences. However, if this were a binding site for a regulatory protein, formation of hairpins 1, 2, 4, and possibly 3 would be blocked. This region is flanked by one pair of long direct repeats. Between the stop codon and stem loop 1 are 107 nucleotides nearly identical (92%) to 110 nucleotides repeated after dyad

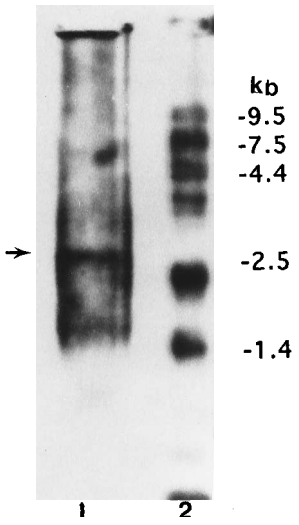


FIG. 5. Northern blot probed with a 3-kb fragment containing the entire *prtT* open reading frame. The transcript is approximately 3.3 kb in size (lane 1). Molecular weight markers are in lane 2.

symmetry 6 (Fig. 2, in boldface). Within the second repeat is the decanucleotide AAATAATAAA at positions 4009 to 4018, which also appears downstream of the papain gene but has no known function (13).

DISCUSSION

The gene *prtT*, shown previously to have trypsinlike protease and hemagglutinin activities, is nearly 2.7 kb and encodes a protein in the size range of 96 to 99 kDa. Its amino acid sequence is similar to those of *Streptococcus pyogenes* proteinase and SPE B, two closely related proteins. The sequence similarity spans the entirety of the streptococcal proteins, with the greatest conservation being at the putative active-site amino acids. Eleven cysteine residues are found in the deduced amino acid sequence of PrtT (Table 1), making possible the formation of extensive intramolecular disulfide bridges under nonreducing conditions. If the enzymatic mechanism of PrtT resembles that of the other cysteine proteinases, Cys-202 and His-345 are likely reactive amino acids. This conclusion is based on similarity of two regions of PrtT to the identified or putative active sites of papain, actinidin, cathepsins, streptococcal proteinase, and other cysteine proteases (Fig. 7) (7, 61).

Like PrtT, streptococcal proteinase, is similar at the amino acid level to other cysteine proteinases solely across their active sites. The putative active cysteine is the sole cysteine residue found in streptococcal proteinase. Thus, intramolecular disulfide bridges do not form. The mature protein is derived from a 40,314-Da, 398-amino-acid precursor. Cleavage of the 27-amino-acid signal sequence is followed by successive proteolysis to yield a stable 27.6-kDa protein (21). Streptococcal proteinase precursor and SPE B are thought to be identical proteins or variants of the same protein. Comparison of the presently deposited SPE B and streptococcal proteinase precursor sequences shows differences at 16 amino acid residues (Fig. 6) (21). Both are encoded by a single gene, *speB* (9, 17, 21). By insertional inactivation of *speB*, wild-type extracellular protease activity of strain NZ131 was significantly reduced (9).

Similarity between PrtT and SPE B and streptococcal proteinase suggests that PrtT may have significant biologic activity. SPE B is one of three pyrogenic exotoxins produced by group

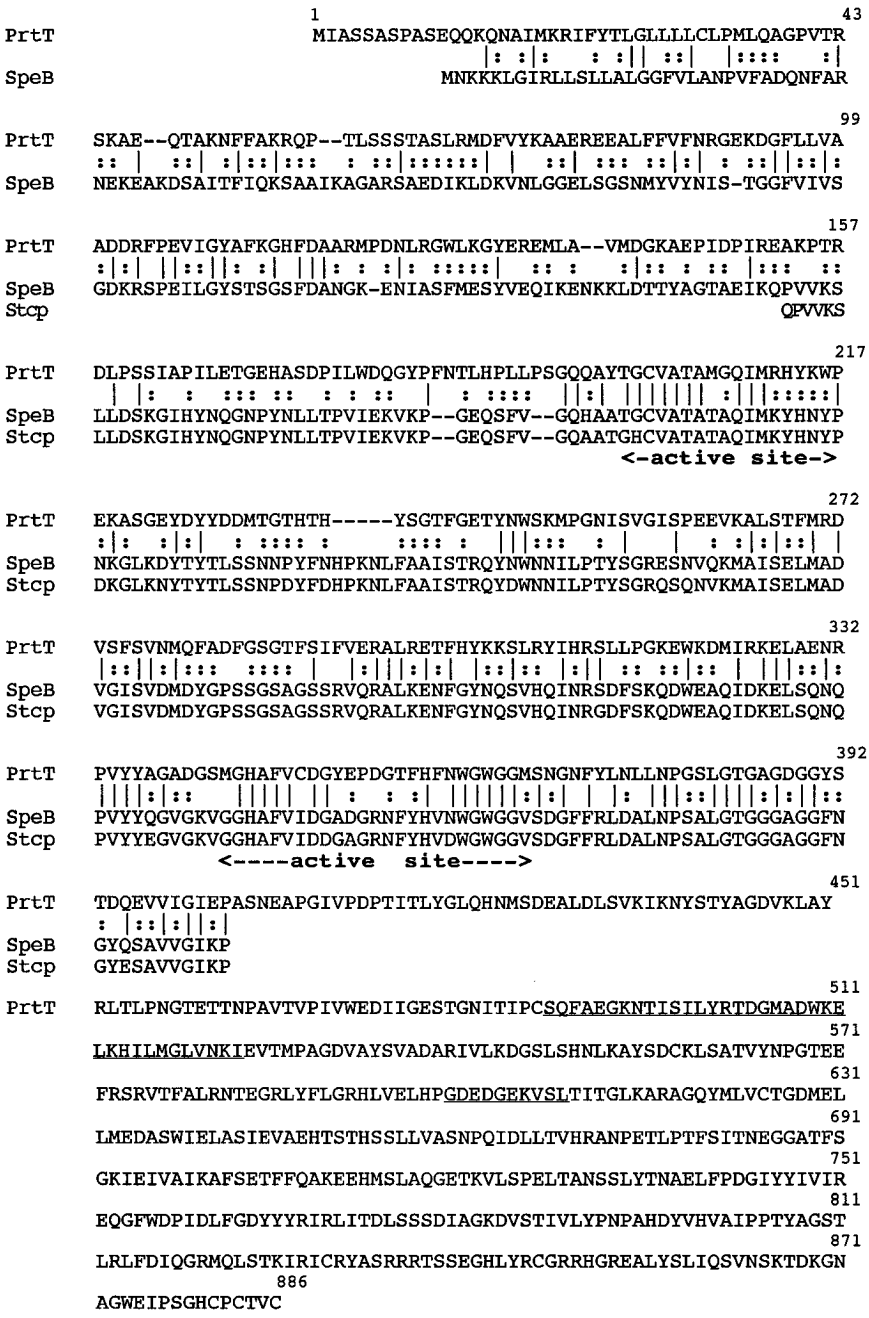


FIG. 6. Amino acid alignment of PrtT with SPE B and streptococcal proteinase (Stcp). Putative streptococcal active-site residues are indicated.

A beta-hemolytic streptococci (SPE A, B, and C). The SPE proteins are also known as scarlatinal, Dick, cardiac, erysipelas, and erythrogenic toxins (4, 32). The cause of scarlet fever and possibly toxic shock-like syndrome, they induce hypersensitive skin rash, fever, and T-lymphocyte mitogenesis. They enhance myocardial damage due to streptolysin O, potentiate endotoxic shock, and suppress immunoglobulin secretion (32). Preparations of SPE from strains NY5 and ATCC 19165 potently augment natural killer cell activity of human peripheral blood mononuclear cells (47). In the presence of splenic macrophages, SPE preparations augment glucose consumption and superoxide anion production as well as inhibit phagocytosis by rabbit macrophages (40).

Sequence similarity between enzymes in dissimilar bacterial genera may have arisen by one of several mechanisms. Certainly, divergent evolution of an enzyme possessed by a common ancestor of *Streptococcus* and *Porphyromonas* species is a possible explanation of the origin of *prtT* in *P. gingivalis*. A second mechanism would be one of genetic exchange. Close proximity of the two bacteria at some point in their life cycles would enhance such opportunities for genetic exchange. Although beta-hemolytic streptococci are cultured from the periodontal lesions of 33% of patients with advanced periodontitis, isolates did not include group A (*S. pyogenes*) (18). However, group A streptococci and *P. gingivalis* (classified as *Bacteroides melaninogenicus* at that time) have been commonly isolated

QQAYTG**C**VATAMGQIM	MG**H**AFVCDGYEPDGTFHFNWGWGGMSNGNF	*P. gingivalis prtT*
QHAATG**C**VATATAQIM	GG**H**AFVIDGADGRNFYHVNWGWGGVSDGFF	SpeB (21)
QAATGH**C**VATATAQIM	GG**H**AFVIDDGAGRNFYHVDWGWGGVSDGFF	streptococcal proteinase (61)
QGSCGS**C**YAFSAVVTI	VD**H**AVAAVGYNPGYILIKNSWGTGWGE	papain (13)
QGSCGS**C**WAFSAVATV	VD**H**AVTAVGYGKSGGKGYILIKNSWGTAWD	Papaya cysteine proteinase omega (38; 67)
QGYCES**C**WAFSTVATV	VD**H**AVTAVGYGKSGGKGYILIKNSWGPGWG	Papaya cysteine proteinase IV (46)
QGECGG**C**WAFSAIATV	VD**H**AIVIVGYGTEGGVDYWIVKNSWDTTWG	kiwi actinidin (7)
QGSCGS**C**WAFGAVEAM	GG**H**AIRILGWGIENGVPYWLVANSWNVDWG	cathepsin B (7)
QGACGS**C**WTFSTTGAL	VN**H**AVLAVGYGEQNGLLYWIVKNSWGSNWG	cathepsin H (26)
QNPCGA**C**WAFGAIATV	LN**H**AVTAIGYNKAEFGDGSGKKARWG	stem bromelain (7)
QGIAGD**C**YMLAALSAI	AY**H**AYTVLGYTVSNGAYYLIIRNPWGVTEP	*P. gingivalis tpr* (5)
QGALGD**C**WLLAAIASL	KG**H**AVSVTGAKQVNYRGQVVSLIRMRNPWG	human calpain(m large subunit (58)
QGALGD**C**WLLAVVASI	GS**H**AYSLTGVYPVNYRGRTQWLMRLRNPWG	Shistosoma mansomi calpain (12)
QGQCGS**C**WAFSTI	LD**H**GVLLVGYNDNSNPPYWIVKNSW	*Trypanosoma brucei* cysteine protease (16)
QGACGS**C**WAFSTIATV	LD**H**AVTAVGYGTSDGKNYIIIKNSWGPNWG	chymopain B (67)
QGGCGS**C**WAFSGVAAT	NY**H**AVNIVGYSNAQGVDYWIVRNSWDTNVG	*Dermatophagoides pteronyssinus* cysteine protease (12)

FIG. 7. Comparison of active-site amino acid residues of various cysteine proteinases. Reactive cysteine and histidines are in boldface, as is contributory glutamine (7, 29, 34).

together from acutely infected tonsils (8). Both *Streptococcus* and *Bacteroides* species contain conjugal transposons, and *P. gingivalis* has been successfully transformed with elements from *Bacteroides* species (49, 53, 64). In addition, there does not appear to be a barrier to conjugal transfer between gram-negative and gram-positive bacteria. The *Streptococcus tetM* gene freely infects *Bacteroides* hosts, conferring high-level resistance to tetracycline and minocycline (14, 30, 57).

In determining the *prtT* start of translation, two possible sites were identified. The codon at nucleotide position 766 is ATT, is preceded by a ribosome binding site, but does not encode a signal sequence. In *P. gingivalis*, alternate (not ATG) start codons have been identified for the methylase (*pgiM*) and the fimbrial subunit (*fimA*) genes (2, 19). It is not known whether their use affects the rate of translation of these proteins. The ATG at position 820 encodes a signal sequence but is not preceded by a classic ribosome binding site (Fig. 2). A protein initiated at position 766 would be 96.4 kDa, and a protein initiated from position 820 would be 98.3 kDa. A size difference this small could not be resolved from the in vitro transcription-translation experiments carried out in this work.

The transcriptional start at position 256 is preceded by the putative promoter sequences TTCAGA (−35) and TACCAT (−10) (Fig. 2). The consensus prokaryotic Pribnow box is −10 $T_{89}A_{89}T_{50}A_{65}A_{65}T_{100}$, with subscripts indicating percent agreement of each nucleotide position with known promoters. Two cytosine residues of the putative *prtT* promoter occur at positions 3 and 4 (underlined in Fig. 2), which are the bases most likely to diverge from the consensus (33). Therefore, the 3.3-kb transcript (Fig. 5) begins at position 256 and includes 0.5 kb of 5′ untranslated message, 2.6 kb of translated message, and approximately 0.14 kb of 3′ untranslated message.

If translation begins at position 766, the amino terminus of PrtT does not have a signal sequence. It is not known if PrtT is secreted in vivo from *P. gingivalis* or if secretion requires a signal peptide. While extracellular proteases from gram-positive bacteria depend on signal peptides for secretion, proteases from gram-negative bacteria utilize a variety of mechanisms to reach the extracellular environment. Two metalloproteases which do so without signal peptides are the products of *prtB* and *prtC* of *Erwinia chrysanthemi* (66).

We show that PrtT is translated as a 96- to 99-kDa protein. Previous purification of recombinant chimeric PrtT expressed in *E. coli* yielded an active protein with an approximate molecular size of 51 to 53 kDa (42). As previously stated, *prtT* in the KS clone lacks the 3′ end of the gene, and so the recombinant protein lacks its carboxyl terminus. In addition to the possibility that the signal peptide is removed from the protein, it apparently undergoes additional size reduction by processing in *E. coli*. Precedence for size reduction of bacterial cysteine proteases is seen with streptococcal proteinase as well as clostripain from *Clostridium histolyticum*. We have not yet determined the nature of this processing, nor do we know the size of the mature PrtT protein when expressed in *P. gingivalis*. Although not all gram-negative proteases require signal peptides for secretion, finding one in this work suggests that PrtT may use this mechanism to reach the extracellular environment (66).

PrtT appears to be unique among a number of *P. gingivalis* trypsinlike, cysteine proteases being characterized at the molecular level. Porphypain is a large, cell surface cysteine protease from strain W12 with arginyl and lysyl endopeptidase activities. The porphypain open reading frame is 4.6 kb, encoding 1,830 amino acids, and has no homology to *prtT* (31). It is, however, related to a hemagglutinin gene, *hagA*. The significance of this relatedness is not clear; however, similar binding epitopes may be used by the hemagglutinins and the proteases of *P. gingivalis* (45).

Currently, the biochemical and molecular characterization of similarly acting arginine- and lysine-specific cysteine proteases is under way in several laboratories. It remains unclear whether these proteins are related to each other. Gingipain is a calcium-stabilized cysteine protease purified from strain H66, which specifically cleaves peptides at the carboxyl-terminal side of arginine residues (43). The amino-terminal sequence is not homologous with any region of *prtT* (10). The gene for gingipain appears to have a much larger open reading frame than would be predicted from the 50-kDa molecular size of the purified enzyme (63), suggesting posttranslational modification of the enzyme. Further studies have indicated that there are two forms of gingipain with different molecular weights and activities (43). Two of these subunits have been termed arg-gingipain and lys-gingipain because they cleave specifically at the carboxyl-terminal side of arginine and lysine residues, respectively. Arg-gingipain is 50 kDa, and lys-gingipain is 60 kDa, and both form noncovalent bonds with a 44-kDa protein and with 30-, 27-, and 17-kDa proteins thought to be subunits of the 44-kDa protein. These associating proteins are believed to be

responsible for the hemagglutination of high-molecular-weight gingipain complexes (43).

A distinct 64-kDa thiol protease has been cloned and sequenced from strain W83. The *tpr* gene was isolated by screening recombinant *E. coli* clones for general protease activity on agar plates containing skim milk. It also has activity against bovine serum albumin, azocoll, denatured collagen, casein, and fibrinogen but lacks activity against BAPNA and native collagen (5). The *tpr* gene encodes a 482-amino-acid protein which has homology across a 24-amino-acid region with eukaryotic and prokaryotic cysteine proteases, encompassing their active sites (Fig. 7). The deduced amino acid sequence of *tpr* does not have other homology with that of *prtT*.

A 97- to 110-kDa protease, PrtH, which cleaves human complement component C3 has recently been cloned from *P. gingivalis* W83. By direct comparison of the *prtT* sequence, we have found that this protease is expressed from a different gene (36). Mutant strain HF18, produced by allelic exchange, has diminished proteolytic activity against C3 and is noninvasive and avirulent in the mouse model. HF18 accumulates activated C3b opsonins on its surface to a level 12-fold higher than that of the wild-type strain (49). This finding suggests that *prtH* encodes a protease that may enable *P. gingivalis* to evade complement-mediated killing during the immune response.

Others have characterized an enzyme from *P. gingivalis* ATCC 33277 and named it lys-gingivain (51). It is a 68- to 70-kDa thiol lysyl-amidase with potent activity against high-molecular-weight kininogen and fibrinogen. This dual activity against kininogen and fibrinogen mimics the host enzymes tryptase, plasmin, factor Xla, and neutrophil elastase.

Three additional *P. gingivalis* proteases are under study, but sequence information is not yet available for comparison with known genes. A 150-kDa thiol gelatinase/hemagglutinin which stimulates collagen degradation by rat oral mucosal keratinocytes by activation of tissue matrix metalloproteinases has been purified (15). A 45-kDa trypsinlike enzyme increases interstitial collagenase production 10-fold by cultured human gingival fibroblasts but with no effect on tissue inhibitor of metalloproteinase (25). A 55-kDa arginine-specific cysteine protease with a pH optimum of 8.0 has been purified from the culture supernatant of *P. gingivalis* 381 and is able to degrade collagens, C3, α1 and β chains of fibrinogen, fibronectin, α1-antitrypsin, α2-macroglobulin, apotransferrin, albumin, and the β chain of insulin (3).

Further molecular characterization of this group of enzymes will add clarity in understanding the relationship among multiple *P. gingivalis* proteases and hemagglutinins beyond that achieved solely from biochemical purification and characterization. In summary, *prtT* encodes a cysteine protease with hemagglutinin activity. Its sequence similarity with streptococcal proteinase and the active site of the family of cysteine proteases supports experimental findings that it is a cysteine protease. Because of its similarity with SPE B, the possibility exists that it also has toxin activity. However, additional experiments with the purified enzyme are necessary to confirm this hypothesis.

ACKNOWLEDGMENTS

We thank T. Karunakaran for performing the Northern blot experiments.

This work was supported by Public Health Service grants 5R01 DE08512 and DE08293 from the National Institute for Dental Research. T. Madden was the recipient of Dentist-Scientist Award 5K16 DE00159 and National Research Science Award DE0 7061-16 from the National Institute for Dental Research.

REFERENCES

1. **Ausubel, F., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1994. Current protocols in molecular biology, p. 2.4.5. Greene and Wiley-Interscience, New York.
2. **Banas, J. A., J. J. Ferretti, and A. Progulske-Fox.** 1991. Identification and sequence analysis of a methylase gene in *Porphyromonas gingivalis*. Nucleic Acids Res. **15:**4189–4192.
3. **Bedi, G. S., and T. Williams.** 1994. Purification and characterization of a collagen-degrading protease from *Porphyromonas gingivalis*. J. Biol. Chem. **269:**599–606.
4. **Betley, M., D. Borst, and L. Regassa.** 1992. Staphylococcal enterotoxins, toxic shock syndrome toxin and streptococcal pyrogenic exotoxins: a comparative study of their molecular biology. p. 2–37. *In* B. Fleischer (ed.), Biological significance of superantigens. Karger, Basel.
5. **Bourgeau, G., H. Lapointe, P. Peloquin, and D. Mayrand.** 1992. Cloning, expression and sequencing of a protease gene (*tpr*) from *Porphyromonas gingivalis* W83 in *Escherichia coli*. Infect. Immun. **60:**3186–3192.
6. **Boyd, J., and B. McBride.** 1984. Fractionation of hemagglutinating and bacterial binding adhesins of *Bacteroides gingivalis*. Infect. Immun. **45:**403–409.
7. **Brockelhurst, K., F. Willenbrock, and E. Salih.** 1987. Cysteine proteinases, p. 39–158. *In* A. Neuberger and K. Brockelhurst (ed.), Hydrolytic enzymes. Elsevier, Amsterdam.
8. **Brook, I., and A. E. Gober.** 1983. *Bacteroides melaninogenicus*. Its recovery from tonsils of children with acute tonsillitis. Arch. Otolaryngol. **109:**818–820.
9. **Chaussee, M. S., D. Gerlach, C. E. Yu, and J. J. Ferretti.** 1993. Inactivation of the streptococcal erythrogenic toxin B gene (*speB*) in *Streptococcus pyogenes*. Infect. Immun. **61:**3719–3723.
10. **Chen, Z., P. Potempa, A. Polanowski, M. Wikstrom, and J. Travis.** 1992. Purification and characterization of a 50-kDa cysteine proteinase (Gingipain) from *Porphyromonas gingivalis*. J. Biol. Chem. **267:**18896–18901.
11. **Chomczynski, P., and N. Sacchi.** 1987. Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
12. **Chua, K. Y., G. A. Stewart, W. R. Thomas, R. J. Simpson, R. J. Dilworth, T. M. Plozza, and K. J. Turner.** 1988. Sequence analysis of cDNA coding for a major mouse dust mite allergen. J. Exp. Med. **167:**175–182.
13. **Cohen, L. W., V. M. Coghlan, and L. C. Dihel.** 1986. Cloning and sequencing of papain-encoding cDNA. Gene **48:**219–227.
14. **de Barbeyrac, B., B. Dutilh, C. Quentin, H. Renaudin, and C. Debear.** 1991. Susceptibility of *Bacteroides ureolyticus* to anti-microbial agents and identification of a tetracycline resistance determinant related to *tetM*. J. Antimicrob. Chemother. **27:**721–731.
15. **DeCarlo, A. A., L. J. Windsor, M. K. Bodden, W. G. I. Moore, G. J. Harber, B. Birkedal-Hansen, and H. Birkedal-Hansen.** 1994. Activation and molecular processing of matrix metalloproteinases by *Porphyromonas gingivalis* thiol proteinase. J. Dent. Res. **73:**314.
16. **Eakin, A. E., J. Bouvier, J. A. Sakanari, C. S. Craik, and J. H. McKerrow.** 1990. Amplification and sequencing of genomic DNA fragments encoding cysteine proteinases from protozoan parasites. Mol. Biochem. Parasitol. **39:**1–8.
17. **Ferretti, J. J. (Oklahoma Health Science Center).** 1994. Personal communication.
18. **Flynn, M., and J. Slots.** 1993. Beta-hemolytic streptococci in advanced periodontitis. Oral Microbiol. Immunol. **8:**295–297.
19. **Fujiwara, T., S. Morishima, I. Takahashi, and S. Hamada.** 1993. Molecular cloning and sequencing of the fimbrilin gene of *Porphyromonas gingivalis* strains and characterization of recombinant proteins. Biochem. Biophys. Res. Commun. **197:**241–247.
20. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35:**365–403.
21. **Hauser, A., and P. Schlievert.** 1990. Nucleotide sequence of the streptococcal pyrogenic exotoxin type B gene and relationship between the toxin and the streptococcal proteinase precursor. J. Bacteriol. **172:**4536–4542.
22. **Hayashi, H., A. Nagata, D. Hinode, M. Sato, and R. Nakamura.** 1992. Survey of a receptor protein in human erythrocytes for hemagglutinin of *Porphyromonas gingivalis*. Oral Microbiol. Immunol. **7:**204–211.
23. **Holdeman, L., E. Cato, and W. Moore.** 1977. Anaerobe laboratory manual. Virginia Polytechic Institute, Blacksburg.
24. **Holt, S. C., J. Ebersole, J. Felton, M. Brunsvold, and K. S. Kornman.** 1988. Implantation of *Bacteroides gingivalis* in nonhuman primates initiates progression of periodontitis. Science **239:**55–57.
25. **Howard, E. W., U. Kucich, G. W. Shen, W. Abrams, and J. Rosenbloom.** 1994. Purification and characterization of a *Porphyromonas gingivalis* trypsin-like protein. J. Dent. Res. **73:**314.
26. **Ishidoh, K., S. Imajoh, Y. Emori, D. Ohno, H. Kavaski, Y. Minami, E. Kominani, N. Katunuma, and K. Suzuki.** 1987. Molecular cloning and sequencing of cDNA for rat cathepsin H. FEBS Lett. **226:**33–37.
27. **Kato, T., N. Takahashi, and H. Kuramitsu.** 1992. Sequence analysis and characterization of the *Porphyromonas gingivalis prtC* gene, which expresses

a novel collagenase activity. J. Bacteriol. **174:**3889–3895.
28. **Klausen, R., N. Ramamurthy, L. Golub, C. Sfintescu, and R. Genco.** 1992. Periodontopathic potential of two strains of *Porphyromonas gingivalis* in gnotobiotic rats. Arch. Oral Biol. **37:**813–819.
29. **Kortt, A., and T. Liu.** 1973. On the mechanism of action of streptococcal proteinase. Biochem. **12:**320–327.
30. **LeBouguenec, C., G. deCespedes, and T. Horaud.** 1990. Presence of chromosomal elements resembling the composite structure Tn*3701* in streptococci. J. Bacteriol. **172:**727–734.
31. **Lantz, M. (University of Pittsburgh).** 1994. Personal communication.
32. **Lee, P., and P. Schlievert.** 1989. Quantification of toxicity of group A streptococcal pyrogenic exotoxins in an animal model of toxic shock syndrome-like illness. J. Clin. Microbiol. **27:**1890–1892.
33. **Lewin, B.** 1987. Genes III, p. 250. John Wiley & Sons, New York.
34. **Liu, T., and S. Elliott.** 1971. Streptococcal proteinase, p. 609–647. *In* P. Boyer (ed.), Hydrolysis. Academic Press, New York.
35. **Loos, B., D. Mayrand, R. J. Genco, and D. P. Dickenson.** 1990. Genetic heterogeneity of *Porphyromonas* (*Bacteroides*) *gingivalis* by genomic DNA fingerprinting. J. Dent. Res. **69:**1488–1493.
36. **Fletcher, H. M., H. A. Schenkein, and F. Macrina.** 1994. Cloning and characterization of a new protease gene (*prtH*) from *Porphyromonas gingivalis*. Infect. Immun. **62:**4279–4286.
37. **Mayrand, D., and S. Holt.** 1988. Biology of asaccharolytic black-pigmented *Bacteroides* species. Microbiol. Rev. **52:**134–152.
38. **McKee, R. A., S. Adams, J. A. Matthews, C. J. Smith, and S. H. Smith.** 1986. Molecular cloning of two cysteine proteinases from paw-paw (*Carica papaya*). J. Biochem. **237:**105–110.
39. **Mouton, C., N. Eidhin, M. Deslauriers, and L. Lamy.** 1991. The hemagglutinating adhesin HA-Ag2 of *Bacteroides gingivalis* is distinct from fimbrilin. Oral Microbiol. Immunol. **6:**6–11.
40. **Murai, T., Y. Ogawa, and H. Kawasaki.** 1990. Effect of streptococcal pyrogenic exotoxin on rabbit macrophage functions in vitro: mediation by splenic lymphocytes. FEMS Microbiol. Lett. **69:**201–204.
41. **Nishikata, M., and F. Yoshimura.** 1991. Characterization of *Porphyromonas* (*Bacteroides*) *gingivalis* hemagglutinin as a protease. Biochem. Biophys. Res. Commun. **178:**336–342.
42. **Otogoto, J., and H. Kuramitsu.** 1993. Isolation and characterization of the *Porphyromonas gingivalis prtT* gene, coding for protease activity. Infect. Immun. **61:**117–123.
43. **Pike, R., W. McGraw, J. Potempa, and J. Travis.** 1994. Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. J. Biol. Chem. **269:**406–411.
44. **Platt, T., and D. G. Bear.** 1983. Role of RNA polymerase, rho factor, and ribosomes in transcription termination, p. 123–161. *In* J. Beckwith, J. Davies, and J. Gallant (ed.), Gene function in prokaryotes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
45. **Progulske-Fox, A., V. Rao, N. Han, G. Lepine, J. Whitlock, and M. Lantz.** 1993. Molecular characterization of hemagglutinin genes of periodontopathic bacteria. J. Periodontal Res. **28:**473–474.
46. **Ritonja, A., D. J. Buttle, N. D. Rawlings, V. Turk, and A. J. Barrett.** 1989. Papaya proteinase IV amino acid sequence. FEBS Lett. **258:**109–112.
47. **Sacks, L., F. Petermann, C. Lapham, P. John, M. Yuille, and R. Tomar.** 1991. A streptococcal erythrogenic toxin preparation augments natural killer activity of peripheral blood mononuclear cells. J. Infect. Dis. **164:**522–526.
48. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
49. **Schenkein, H., H. Fletcher, C. Berry, and F. Macrina.** 1994. Restoration of C3-fragment accumulation on a protease-deficient mutant of *Porphyromonas gingivalis* W83. J. Dent. Res. **73:**189.
50. **Schenkein, H. A.** 1988. The effect of periodontal proteolytic *Bacteroides* species on proteins of the human complement system. J. Periodontal Res. **23:** 187–192.
51. **Scott, C., E. J. Whitaker, B. F. Hammond, and R. W. Colman.** 1993. Purification and characterization of a potent 70-kDa thiol lysyl-proteinase (Lys-gingivain) from *Porphyromonas gingivalis* that cleaves kininogens and fibrinogen. J. Biol. Chem. **268:**7935–7942.
52. **Shah, H., and M. Collins.** 1988. Proposal for reclassification of *Bacteroides asaccharolyticus*, *Bacteroides gingivalis*, and *Bacteroides endodontalis* in a new genus, *Porphyromonas*. Int. J. Sys. Bacteriol. **38:**128–131.
53. **Simpson, W., B. M. Odusany, M. O. Lassiter, and C. A. Genco.** 1994. Characterization of the hemin uptake defect in a Tn4351 of *Porphyromonas gingivalis*. J. Dent. Res. **73:**349.
54. **Slots, J.** 1982. Importance of black-pigmented *Bacteroides* in human periodontal disease. p. 27–45. *In* R. Genco and S. Mergenhagen (ed.), Host-parasite interactions in periodontal diseases. American Society for Microbiology, Washington, D.C.
55. **Slots, J., and R. Genco.** 1984. Black-pigmented *Bacteroides* species, *Capnocytophaga* species, and *Actinobacillus actinomycetemcomitans* in human periodontal disease: virulence factors in colonization, survival, and tissue destruction. J. Dent. Res. **63:**412–421.
56. **Slots, J., and R. 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.
57. **Smith, C. J., M. B. Rogers, and M. L. McKee.** 1992. Heterologous gene expression in *Bacteroides fragilis*. Plasmid **27:**141–154.
58. **Sorimachi, H., and K. Suzuki.** 1992. Sequence comparison among muscle-specific calpain, p94, and calpain subunits. Biochim. Biophys. Acta **1160:**55–62.
59. **Sorsa, T., T. Ingman, K. Suomalainen, M. Haapasalo, Y. T. Konttinen, O. Lindy, H. Saari, and V. Uitto.** 1992. Identification of proteases from periodontopathogenic bacteria as activators of latent human neutrophil and fibroblast-type interstitial collagenases. Infect. Immun. **60:**4491–4495.
60. **Sunqvist, G., J. Carlsson, B. Herrmann, and A. Tarnvik.** 1985. Degradation of human immunoglobulins G and M and complement factors C3 and C5 by black-pigmented *Bacteroides*. J. Med. Microbiol. **19:**85–94.
61. **Tai, J., A. Kortt, T. Liu, and S. Elliot.** 1976. Primary structure of streptococcal proteinase. J. Biol. Chem. **251:**1955–1959.
62. **Tinoco, I., P. Borer, B. Dengler, M. Levine, O. Uhlenbeck, D. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246:**40–41.
63. **Travis, J. (University of Georgia).** 1993. Personal communication.
64. **Valentine, P. J., N. B. Shoemaker, and A. A. Salyers.** 1988. Mobilization of *Bacteroides* plasmids by *Bacteroides* conjugal elements. J. Bacteriol. **170:** 1319–1324.
65. **von Heijne, G.** 1987. Sequence analysis in molecular biology. Academic Press, San Diego, Calif.
66. **Wandersman, C.** 1989. Secretion, processing and activation of bacterial extracellular proteases. Mol. Microbiol. **3:**1825–1831.
67. **Watson, D. C., M. Yaguchi, and K. R. Lynn.** 1990. The amino acid sequence of chymopain from *Carica papaya*. Biochem. J. **266:**75–81.
68. **Yoshimura, F., K. Takahashi, Y. Nodosdaka, and T. Suzuki.** 1984. Purification and characterization of a novel type of fimbriae from the oral anaerobe *Bacteroides gingivalis*. J. Bacteriol. **160:**949–957.