J. Biochem. 120, 398-406 (1996)

Cloning and Sequencing of the Gene Encoding a Novel Lysine-Specific Cysteine Proteinase (Lys-Gingipain) in *Porphyromonas gingivalis*: Structural Relationship with the Arginine-Specific Cysteine Proteinase (Arg-Gingipain)[1]

Kuniaki Okamoto,*[,2] Tomoko Kadowaki,* Koji Nakayama,† and Kenji Yamamoto*[,3]

*Departments of *Pharmacology and †Microbiology, Kyushu University Faculty of Dentistry, Higashi-ku, Fukuoka 812-82*

Received for publication, April 15, 1996

Lys-gingipain (KGP), so termed due to its peptide cleavage specificity for lysine residues, is a cysteine proteinase produced by the Gram-negative anaerobic bacterium *Porphyromonas gingivalis*. Mixed oligonucleotide primers designed from the NH_2-terminal sequence of the purified enzyme were used to clone the KGP-encoding gene (*kgp*) from the organism. The nucleotide sequence of *kgp* had a 5,169-bp open reading frame encoding 1,723 amino acids with a calculated molecular mass of 218 kDa. As the extracellular mature enzyme had an apparent molecular mass of 51 kDa in gels, the precursor of KGP was found to comprise at least four domains, the signal peptide, the NH_2-terminal prodomain, the mature proteinase domain, and the COOH-terminal hemagglutinin domain, and to be proteolytically processed during its transport. Importantly, the COOH-terminal region contained three direct repeats of two different amino acid sequences, LKWD(or E)AP and YTYTVYRDGTKI, and the subdomains located between the two repeats exhibited strong similarity to those of Arg-gingipain (RGP), another major cysteine proteinase produced by the organism and having cleavage specificity for arginine residues, although the arrangement of the subdomains was not necessarily identical in the two enzymes. Since the KGP activity was greatly decreased in RGP-deficient mutants and since the most probable site of the propeptide cleavage was present in the homologous sequence highly susceptible to proteolysis by RGP, the precursor of KGP is likely to be processed by RGP to form the mature enzyme.

Key words: Arg-gingipain, Lys-gingipain, lysine-specific cysteine proteinase, *Porphyromonas gingivalis*, precursor structure.

The destruction of periodontal tissues, including alveolar bone, is one of the most important pathological conditions of progressive periodontal disease. Although the actual mechanism of periodontal tissue breakdown during disease progression remains to be clarified, proteolytic enzymes derived from both host cells and oral microorganisms are thought to be key factors in the pathogenicity of the disease. *Porphyromonas gingivalis*, a Gram-negative anaerobic bacterium frequently isolated from subgingival lesions of periodontitis patients, is thought to be one of the most important pathogens of the disease (*1-7*) and to produce a variety of proteases in both cell-associated and secretory forms (*6, 8-15*). Despite a great deal of work, little is known about the pathogenic capability of individual proteases and the mechanism by which they function, although some of the enzymes have been suggested to relate to the development of disease (*16-22*). We have recently identified the unique arginine-specific cysteine proteinase Arg-gingipain (RGP), formerly argingipain, in the culture supernatant of *P. gingivalis* 381 strain (*23*). Based on biochemical and structural characterization (*23, 24*), the proteolytic activity of RGP was shown to be closely associated with the virulence of the organism. To clarify to what extent RGP contributes to the virulence of *P. gingivalis*, we have constructed RGP-deficient mutants by disruption of the RGP gene with suicide plasmid systems (*25*). In the course of this study, we found two separate RGP-encoding genes (*rgpA* and *rgpB*) located on the chromosome of this organism (*25*). By analysis of the *rgpA rgpB* double mutant, RGP was shown to contribute as a major periodontal pathogenic factor to the virulence of *P. gingivalis*.

Several studies have shown that *P. gingivalis* produces distinct cysteine proteinases that have peptide cleavage

[1] This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. The nucleotide sequence reported in this paper has been submitted to the GenBank™/EMBL Data Bank with Accession No. D83258.

[2] Okamoto, K., Research Fellow of the Japan Society for the Promotion of Science.

[3] To whom correspondence should be addressed. Tel: +81-92-641-1151 (Ext. 4261), Fax: +81-92-632-1226

Abbreviations: PAGE, polyacrylamide gel electrophoresis; KGP, Lys-gingipain; RGP, Arg-gingipain; PCR, polymerase chain reaction; *P. gingivalis, Porphyromonas gingivalis*; SDS, sodium dodecyl sulfate; SSC, standard saline citrate; Boc, *t*-butyloxycarbonyl; MCA, 4-methyl-7-coumaryl-amide.

specificity for lysine residues (*13, 22, 26, 27*). Although some of these enzymes have been suggested to participate in host tissue damage (*13, 22*), it is unclear whether these lysine-specific cysteine proteinases are identical entities and to what extent they contribute to the virulence of *P. gingivalis*. To gain an insight into the pathological potential of such lysine-specific cysteine proteinases, we considered it important to examine their genetic basis.

In this paper we report: (i) the nucleotide sequence of the gene for a lysine-specific cysteine proteinase from *P. gingivalis* 381 and the deduced amino acid sequence; (ii) the structural characteristics of its precursor; (iii) its comparison with RGP. Since, described later, the deduced amino acid sequence of the enzyme contained sequences identical to the NH_2-terminal sequences of Lys-gingipain fragments (60-, 44-, 30-, and 27-kDa polypeptides) reported by Pike *et al.* (*26*), the present proteinase is hereafter referred to as Lys-gingipain (KGP).

EXPERIMENTAL PROCEDURES

Culture of P. gingivalis and Extraction of Total DNA—*P. gingivalis* 381 strain was grown in brain heart infusion broth (37 mg/ml) (Difco Chemical) containing yeast extract (5 mg/ml), hemin (5 mg/ml), and vitamin K (1.0 mg/ml) at 37°C as previously described (*24*). After cultivation in an anaerobic chamber under 90% N_2, 5% H_2, and 5% CO_2, the cells were pelleted by centrifugation and the pellet was suspended in TESS solution (30 mM Tris-HCl, pH 7.5, 5 mM Na_2EDTA, 50 mM NaCl, 25% sucrose). After addition of 10% SDS (final concentration, 0.5%) and proteinase K (10 mg/ml), the solution was incubated for 2 h at 50°C. The suspension was mixed with $CHCl_3$/phenol and centrifuged at 14,000 × *g* for 5 min. The supernatant was centrifuged again under the same conditions and dialyzed against TE (10 mM Tris-HCl, pH 7.5, 1 mM Na_2EDTA) for 2 d. Two volumes of absolute ethanol were added and mixed gently with the solution. The precipitated DNA was washed in 70% ethanol and suspended in TE. After addition of RNase A (10 mg/ml), the suspension was incubated for 1 h at 37°C. The DNA was precipitated by ethanol and resuspended in TE.

Purification and NH_2-Terminal Sequence Analysis of KGP—The procedures for purification of KGP from the culture supernatant of *P. gingivalis* are detailed elsewhere. Briefly, the culture supernatant was separated from *P. gingivalis* cells by centrifugation. Solid ammonium sulfate was added to the supernatant to give a 65% saturation. The precipitated proteins were collected by centrifugation and resuspended in 10 mM sodium phosphate buffer, pH 7.0. After dialysis against the same buffer, the enzyme in the dialyzate was subjected to chromatography on DEAE-Sephacel and CM-Toyopearl 650S, isoelectric focusing in an LKB 8101 column, gel filtration on a TSKgel G2000SW, and chromatography on arginine-Sepharose, lysine-Sepharose and Mono Q, in that order. The purified enzyme was electrophoresed on SDS gels by the method of Laemmli (*28*), then transferred onto polyvinylidene difluoride membranes and stained with Coomassie Blue R-250. The stained band was excised and the adsorbed protein was applied to an automatic protein/peptide sequencer (Applied Biosystems Model 470A).

Oligonucleotides—Two degenerate synthetic oligonucleotide primers for polymerase chain reaction (PCR) amplification of the KGP gene fragment were designed based on the amino acid sequences of the NH_2-terminal part of the purified enzyme (from 231 to 236 and from 248 to 253). The primers PR1 = 5′-TANACNGANCANGGNGAN-3′ and PR2 = 5′-NTTNAANTTNGCNCCNGC-3′ (N denotes complete degeneracy), which correspond to the opposite ends of the peptide sequence, were synthesized.

PCR Amplification of the KGP-Encoding DNA Fragment Gene—PCR amplification of the KGP-encoding DNA fragment was carried out using the two degenerate oligonucleotide primers [(20 mM Tris-HCl, pH 8.0/100 mM KCl/0.1 mM EDTA/1 mM dithiothreitol/0.5% Tween 20/0.5% Nonidet P-40/50% glycerol/200 mM dATP/200 mM dCTP/200 mM dGTP/200 mM dTTP/2.5 units of Amplitaq) (Perkin Elmer-Cetus)]. The following PCR thermal cycle was used: step 1, 92°C for 1 min; step 2, 50°C for 1 min; step 3, 70°C for 1 min. This was repeated 40 times. PCR products were purified by $CHCl_3$/phenol treatment and cloned into the plasmid vector pT7Blue(R) (Novagen). Four clones containing the inserts were sequenced by the dideoxynucleotide chain termination method (*29*) using a Sequenase DNA-sequencing kit (U.S. Biochemical, Cleveland, OH). All clones were found, as anticipated, to encode NH_2-terminal amino acid sequences. As all clones were identical in the nucleotide sequence, a 69-base pair (bp) insert DNA fragment from these clones was used as a probe for further screenings.

Southern Blot Analysis—Total DNA from *P. gingivalis* was digested for 1 h at 37°C with either *Bam*HI, *Bgl*II, *Eco*RI, *Sac*I, *Sph*I, *Eco*RV, *Nru*I, *Pvu*II, or *Sca*I. Each digest was subjected to electrophoresis in 0.8% agarose, and blotted to a nitrocellulose filter (Schleicher & Schnell). The filter was washed for 5 min in 2× standard saline citrate (SSC, 0.15 M NaCl/0.015 M sodium citrate). After being baked for 2 h at 80°C, the filters were prehybridized for 3 h at 65°C in a hybridization solution (6×SSC containing 0.1% SDS and 10% Denhardt). The filter was washed once in the hybridization solution, then immersed in the fresh hybridization solution containing a ^{32}P-labeled 69-bp probe which was generated with a MEGALABEL labeling kit (Takara Shuzo) and incubated for 20 h at 65°C. The filter was washed twice in 2×SSC containing 0.1% SDS, dried and subjected to autoradiography at −80°C overnight. For another Southern blot analysis, total DNA from *P. gingivalis* was digested for 1 h at 37°C with either *Bam*HI, *Eco*RI, *Hin*dIII, *Bam*HI-*Eco*RI, or *Bam*HI-*Hin*dIII, and Southern blotting was performed by the same method as above. A synthetic oligonucleotide (probe I: 5′-CCACAGGGTGGACAGCATTGGATGC-3′) obtained from Greiner (Tokyo) was labeled with fluorescein-dUTP (Amersham International plc, Little Chalfont, UK), and used for hybridization with ECL 3′-oligolabeling and detection systems (Amersham).

Library Construction and Screening—The *Eco*RI and *Eco*RV-digested DNA was separated by electrophoresis on a 0.8% agarose gel. The fractions containing the 2.7- and 5.5-kbp DNA fragments hybridized with the ^{32}P-labeled 69-bp DNA probe were transferred to DEAE cellulose filters (Wattman DE81). The blotted DNAs were eluted with 0.1 M NaCl and then in *Eco*RV-digests subjected to ligation with *Eco*RI adapters (Promega) and phosphorylation with kinases. The DNA fragments were further

purified by 0.8% agarose gel electrophoresis and ligated into the *Eco*RI site of λZAPII (*30*). The ligated DNAs were packaged into bacteriophages by using a Gigapack GoldII packaging kit (Stratagene, San Diego), followed by plating onto dishes with Sure (Stratagene) bacterial hosts (*31*). The phage solution was used as the *P. gingivalis* genomic DNA library. Two hundred thousand clones of the *P. gingivalis* genomic DNA library were screened by plaque hybridization using the ^{32}P-labeled 69-bp DNA probe. The DNA inserts of isolated positive clones were subcloned into pBluescript SK$^-$ plasmid vectors by automatic excision process (*30*). The DNA inserts were characterized by restriction endonuclease mapping. Finally, single DNA clones with 2.7- and 5.5-kbp DNA inserts were obtained (pNKI and pNKV, respectively). For further sequencing of the gene 3′-region, the *Bam*HI-*Hin*dIII digested DNA was separated by electrophoresis on a 0.8% agarose gel. The fractions containing the 5.0-kbp DNA fragment hybridized with the labeled 25-bp DNA probe were ligased into the *Bam*HI-*Hin*dIII site of pUC118. This solution was used as the *P. gingivalis* genomic DNA library. The DNA library was transformed into *Escherichia coli* DH5α and the plasmids with the *Bam*HI-*Hin*dIII fragment inserted were collected. The DNA inserts were characterized by restriction endonuclease mapping and finally a single clone with a 5.0-kbp DNA insert was obtained (pNKBH).

DNA Sequencing—The DNA clone obtained by autoexcision was digested with *Eco*RI and subcloned into pUC118. Then three kinds of plasmids were digested with appropriate restriction enzyme and subcloned again. Both strands of all regions were sequenced by the dideoxynucleotide chain termination method (*29*) using an AutoRead Sequencing Kit (Pharmacia Biotech) and a Thermo Sequenase core sequencing reagent kit (Amersham). The sequencing strategy is illustrated in Fig. 2.

Computer Analysis of Genomic DNA and Protein—Nucleotide and protein sequences were analyzed by using the Swiss Prot System. Hydropathy analysis was carried out in accordance with Kyte and Doolittle (*32*).

Determinations—The KGP activity was measured using two synthetic substrates (final conc. 10 μM), *t*-butyloxycarbonyl (Boc)-Val-Leu-Lys-Lys-4-methyl-7-coumarylamide (MCA) and Boc-Glu-Lys-MCA, in 20 mM sodium phosphate buffer, pH 7.5, containing 5 mM cysteine in a total volume of 1.0 ml. After incubation at 40°C for 10 min, the reaction was terminated by adding 1.0 ml of 10 mM iodoacetic acid, pH 5.0, and the released 7-amino 4-methylcoumarin was measured at 460 nm (excitation at 380 nm). One unit of enzyme activity was defined as the amount of enzyme required to release 1 μmol of 7-amino-4-methylcoumarin/ml under these conditions.

RESULTS

Analysis of the NH_2-Terminal Amino Acid Sequences—When the final preparation of KGP was incubated in the SDS-solubilizing buffer in the presence of 100 μM leupeptin and 100 μM *p*-toluenesulfonyl-L-lysine chloromethyl ketone at 37°C for 30 min or 100°C for 5 min followed by SDS-PAGE under reducing conditions, it showed a single protein band with an apparent molecular mass of 51 kDa (not shown). The NH_2-terminal amino acid sequence of the purified KGP was analyzed by automatic Edman degradation of the sample electroblotted on a polyvinylidene difluoride membrane. The first 25 amino acid residues of the NH_2-terminus of the purified enzyme were directly determined as follows: D^{229}VYTDHGDLYNTPVRMLVV-AGAKFK^{253} (the numberings refer to the sequence of KGP precursor predicted from its nucleotide sequence). This amino acid sequence was identical with the NH_2-terminal sequence of the 60-kDa Lys-gingipain described by Pike *et al.* (*26*).

PCR Amplification and Isolation of Lys-Gingipain Gene—To clone the KGP-encoding gene from *P. gingivalis*, two degenerate oligonucleotide primers (PR1 and PR2) which correspond to parts of the NH_2-terminal amino acid sequence of the purified enzyme were designed. Using these primers, a DNA fragment contained in the KGP-encoding gene was amplified from total *P. gingivalis* DNA. After 40 cycles of amplification, PCR products were cloned into the pT7Blue(R) plasmid. A 69-bp fragment containing the sequence corresponding to the NH_2-terminal amino acid sequence of the enzyme was obtained and used for initial screening of the total *P. gingivalis* DNA.

As shown in Fig. 1, Southern blot analysis revealed that the radiolabeled 69-bp fragment and the synthesized 25-bp oligonucleotides labeled by fluorescein-11-dUTP hybridized with each blot of the total *P. gingivalis* DNA digested with *Bam*HI, *Bgl*II, *Eco*RI (Fig. 1A), *Sac*I, *Sph*I, *Eco*RV, *Nru*I, *Pvu*II, *Sca*I (Fig. 1B), *Bam*HI, *Eco*RI, *Hin*dIII, *Bam*HI-*Eco*RI, and *Bam*HI-*Hin*dIII (Fig. 1C) to give single major bands. Of these restriction enzymes, *Eco*RI, *Eco*RV,

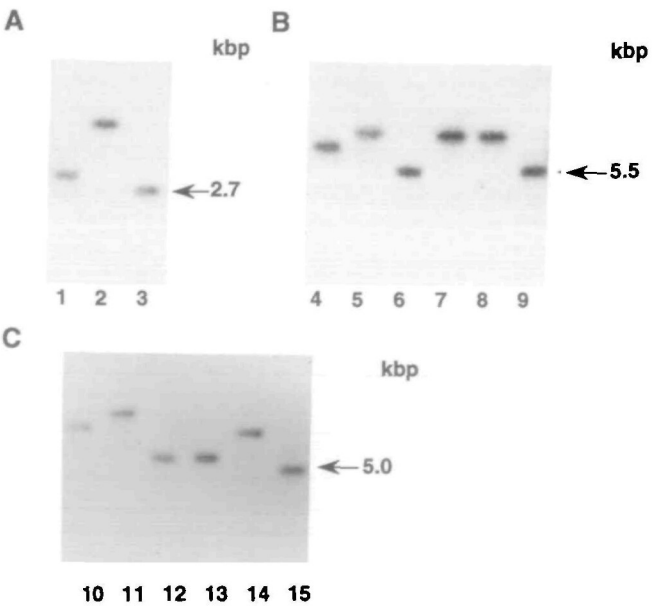

Fig. 1. Southern blot analysis of the genomic DNA of *P. gingivalis* probed with the KGP-specific DNA fragment and oligonucleotide. The *P. gingivalis* DNA was digested with *Bam*HI (lane 1), *Bgl*II (lane 2), *Eco*RI (lane 3), *Sac*I (lane 4), *Sph*I (lane 5), *Eco*RV (lane 6), *Nru*I (lane 7), *Pvu*II (lane 8), *Sca*I (lane 9), *Bam*HI (lane 10), *Eco*RI (lane 11), *Hin*dIII (lane 12), *Kpn*I (lane 13), *Bam*HI-*Eco*RI (lane 14), and *Bam*HI-*Hin*dIII (lane 15). The digests were separated by agarose gel electrophoresis and transferred to a nitrocellulose filter, followed by hybridization with the ^{32}P-labeled 69-bp fragment encoding a portion of KGP (A and B) and the fluorescein-11-dUTP-labeled synthesized 25-bp oligonucleotide, 5′-CCACAGGGTGGACAGCATTGGATGC-3′ (C).

and *Bam*HI-*Hin*dIII were selected and used for construction of the partial *P. gingivalis* library. Two hundred thousand clones of the *P. gingivalis* DNA library from *Eco*RI and *Eco*RV, and 36 colonies of the *P. gingivalis* DNA

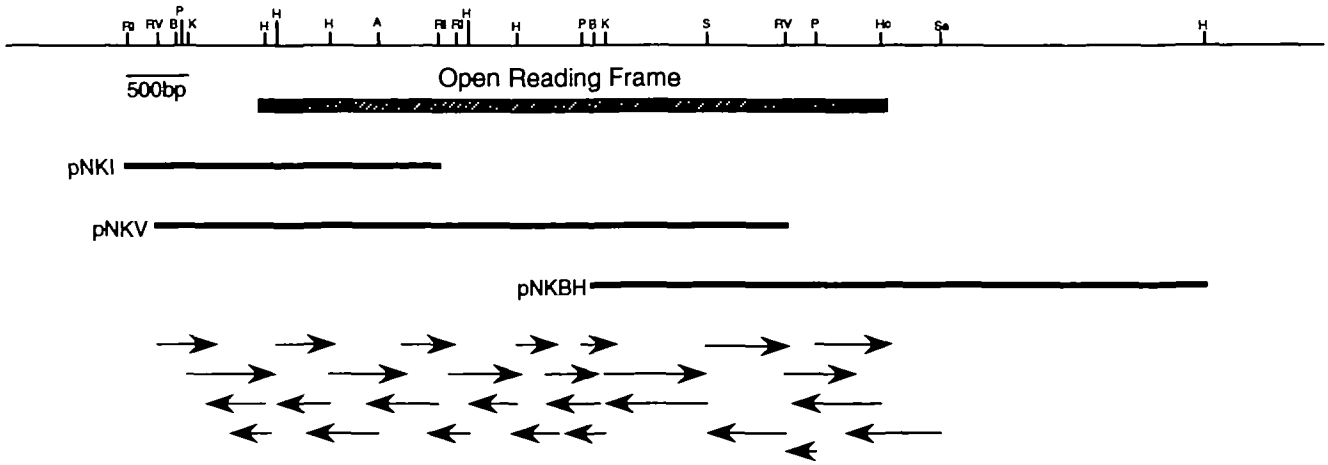

Fig. 2. **Restriction enzyme map and nucleotide sequencing strategy for the KGP-encoding gene (*kgp*).** Representative restriction sites are shown on the top. The sequencing strategy is summarized on the bottom. The hatched bar represents the predicted open reading frame. The horizontal arrows indicate the direction and extent of sequences determined. pNKI, pNKV, and pNKBH are single DNA clones with 2.7-, 5.5-, and 5.0-kbp DNA inserts, respectively. A, *Acc*I; B, *Bam*HI; H, *Hin*dIII; Hc, *Hin*cII; K, *Kpn*I; P, *Pst*I; RI, *Eco*RI; RV, *Eco*RV; S, *Sma*I.

A

```
TTTTTGTTTTTTTAACCCGGCCGTGGTTCTCTGAATCACGACCATAAATTGTTTTAAAGTATGAGGAAATTATTATTGCTGATCGCGGCGTCCCTTTTGGGAGTTGGTCTTTACGCCCAA  120
                                                    *** M  R  K  L  L  L  L  I  A  A  S  L  L  G  V  G  L  Y  A  Q   20

AAGCGCCAAGATTAAGCTTGTGCTCCGACTACTCGAACGACATGTACGAACAATAGCTTCAAGCAGTTCGATGCAAGCTTTTCGTTCAATGAAGTCGAGCTGACAAAGGTGGAGACCAAA  240
S  A  K  I  K  L  D  A  P  T  T  R  T  T  C  T  N  N  S  F  K  Q  F  D  A  S  F  S  F  N  E  V  E  L  T  K  V  E  T  K   60

GGTGGTACTTTCGCCTCAGTGTCAATTCCGGGTGCATTCCCGACCGGTGAGGTTGGTTCTCCCGAAGTGCCAGCAGTTAGGAAGTTGATTGCTGTGCCTGTCGGAGCCACACCTGTTGTT  360
G  G  T  F  A  S  V  S  I  P  G  A  F  P  T  G  E  V  G  S  P  E  V  P  A  V  R  K  L  I  A  V  P  V  G  A  T  P  V  V  100

CGCGTGAAAAGTTTTACCGAGCAAGTTTACTCTCTGAACCAATACGGTTCCGAAAAACTCATGCCACATCAACCCTCTATGAGCAAGAGTGATGATCCCGAAAAGGTTCCCTTCGTTTAC  480
R  V  K  S  F  T  E  Q  V  Y  S  L  N  Q  Y  G  S  E  K  L  M  P  H  Q  P  S  M  S  K  S  D  D  P  E  K  V  P  F  V  Y  140

AATGCTGCTGCTTATGCACGCAAAGGTTTTGTCGGACAAGAACTGACCCAAGTAGAAATGTTGGGGACAATGCGTGGTGTTCGCATTGCAGCTCTTACCATTAATCCTGTTCAGTATGAT  600
N  A  A  A  Y  A  R  K  G  F  V  G  Q  E  L  T  Q  V  E  M  L  G  T  M  R  G  V  R  I  A  A  L  T  I  N  P  V  Q  Y  D  180

GTGGTTGCAAACCAATTGAAGGTTAGAAACAACATCGAAATTGAAGTAAGCTTTCAAGGAGCTGATGAAGTAGCTACACAACGTTTGTATGATGCTTCTTTTAGCCCTTATTTCGAAACA  720
V  V  A  N  Q  L  K  V  R  N  N  I  E  I  E  V  S  F  Q  G  A  D  E  V  A  T  Q  R  L  Y  D  A  S  F  S  P  Y  F  E  T  220

GCTTATAAACAGCTCTTCAATAGAGATGTTTATACAGATCATGGCGACTTGTATAATACGCCGGTTCGTATGCTTGTTGTTGCAGGTGCAAAATTCAAAGAAGCTCTCAAGCCTTGGCTC  840
A  Y  K  Q  L  F  N  R  D  V  Y  T  D  H  G  D  L  Y  N  T  P  V  R  M  L  V  V  A  G  A  K  F  K  E  A  L  K  P  W  L  260

ACTTGGAAGGCTCAAAAGGGCTTCTATCTGGATGTGCATTACACAGACGAAGCTGAAGTAGGAACGACAAACGCCTCTATCAAGGCATTTATTCACAAGAAATACAATGATGGATTGGCA  960
T  W  K  A  Q  K  G  F  Y  L  D  V  H  Y  T  D  E  A  E  V  G  T  T  N  A  S  I  K  A  F  I  H  K  K  Y  N  D  G  L  A  300

GCTAGTGCTGCTCCGGTCTTCTTGGCTTTGGTTGGTGACACTGACGTTATTAGCGGAGAAAAAGGAAAGAAAACAAAAAAAGTTACCGACTTGTATTACAGTGCAGTCGATGGCGACTAT  1080
A  S  A  A  P  V  F  L  A  L  V  G  D  T  D  V  I  S  G  E  K  G  K  K  T  K  K  V  T  D  L  Y  Y  S  A  V  D  G  D  Y  340

TTCCCTGAAATGTATACTTTCCGTATGTCTGCTTCTTCCCCAGAAGAACTGACGAACATCATTGATAAGGTGTTGATGTATGAAAAGGCTACCATGCCGGATAAGAGCTATTTGGAAAAG  1200
F  P  E  M  Y  T  F  R  M  S  A  S  S  P  E  E  L  T  N  I  I  D  K  V  L  M  Y  E  K  A  T  M  P  D  K  S  Y  L  E  K  380

GCCCTCTTGATTGCCGGTGCTGACTCCTACTGGAATCCTAAGATAGGCCAGCAAACCATCAAATATGCTGTACAGTATTACTACAATCAAGATCATGGCTATACAGATGTGTACAGTTAC  1320
A  L  L  I  A  G  A  D  S  Y  W  N  P  K  I  G  Q  Q  T  I  K  Y  A  V  Q  Y  Y  Y  N  Q  D  H  G  Y  T  D  V  Y  S  Y  420

CCTAAAGCTCCTTATACAGGCTGCTATAGTCACTTGAATACCGGTGTCGGCTTTGCCAACTATACAGCGCATGGATCTGAGACATCATGGGCAGATCCGTCGCTGACCGCCACTCAAGTG  1440
P  K  A  P  Y  T  G  C  Y  S  H  L  N  T  G  V  G  F  A  N  Y  T  A  H  G  S  E  T  S  W  A  D  P  S  L  T  A  T  Q  V  460

AAAGCACTCACAAATAAGGACAAATACTTCTTAGCTATTGGGAACTGCTGTGTTACAGCTCAATTCGATTATCCACAGCCTTGCTTTGGAGAGGTAATGACTCGTGTCAAGGAGAAAGGT  1560
K  A  L  T  N  K  D  K  Y  F  L  A  I  G  N  C  C  V  T  A  Q  F  D  Y  P  Q  P  C  F  G  E  V  M  T  R  V  K  E  K  G  500

GCTTATGCCTATATCGGTTCATCTCCGAATTCTTATTGGGGCGAGGACTACTATTGGAGTGTCGGTGCTAATGCCGTATTTGGTGTTCAGCCTACTTTTGAAGGTACGTCTATGGGTTCT  1680
A  Y  A  Y  I  G  S  S  P  N  S  Y  W  G  E  D  Y  Y  W  S  V  G  A  N  A  V  F  G  V  Q  P  T  F  E  G  T  S  M  G  S  540

TATGATGCTACATTCTTTGGAAGATTCGTACAACACAGTGAATTCTATTATGTGGGCAGGTAATCTTGCCGCTACTCATGCTGGAAATATCGGCAATATTACCCATATCGGTGCTCATTAC  1800
Y  D  A  T  F  L  E  D  S  Y  N  T  V  N  S  I  M  W  A  G  N  L  A  A  T  H  A  G  N  I  G  N  I  T  H  I  G  A  H  Y  580

TATTGGGAAGCTTATCATGTCCTTGGCGATGGTTCGGTTATGCCTTATCGTGCAATGCCTAAGACCAATACTTATACGCTTCCTGCTTCTCTGCCTCAGAATCAGGCTTCTTATAGCATT  1920
Y  W  E  A  Y  H  V  L  G  D  G  S  V  M  P  Y  R  A  M  P  K  T  N  T  Y  T  L  P  A  S  L  P  Q  N  Q  A  S  Y  S  I  620

CAGGCTTCTGCCGGTTCTTACGTAGCTATTTCTAAAGATGGAGTTTTGTATGGAACAGGTGTTGCTAATGCCAGCGGTGTTGCGACTGTGAATATGACTAAGCAGATTACGGAAAATGGT  2040
Q  A  S  A  G  S  Y  V  A  I  S  K  D  G  V  L  Y  G  T  G  V  A  N  A  S  G  V  A  T  V  N  M  T  K  Q  I  T  E  N  G  660

AATTATGATGTAGTTATCACTCGCTCTAATTATCTTCCTGTGATCAAGCAAATTCAGGCAGGAGAGCCTAGCCCCTACCAGCCTGTTTCCAACTTGACTGCTACAACGCAGGGTCAGAAA  2160
N  Y  D  V  V  I  T  R  S  N  Y  L  P  V  I  K  Q  I  Q  A  G  E  P  S  P  Y  Q  P  V  S  N  L  T  A  T  T  Q  G  Q  K  700

GTAACGCTCAAGTGGGATGCCCCGAGCGCAAAGAAGGCAGAAGCTTCCCGT
V  T  L  K  W  D  A  P  S  A  K  K  A  E  A  S  R
```

Fig. 3 (Continued on next page)

library from *Bam*HI-*Hin*dIII were screened and single positive clones of 2.7, 5.5, and 5.0 kbp were isolated (Fig. 2).

DNA Sequencing and Structure Analysis—Figure 2 shows the restriction map and sequencing strategy of the isolated DNA. The nucleotide sequence of the DNA and the deduced amino acid sequence are shown in Fig. 3. The regions covering 6.4 kbp were sequenced. The nucleotide sequence (6,366 nucleotides) includes the complete coding region and parts of the 5′- and 3′-noncoding regions. The open reading frame consisting of 5,169 nucleotides was found to encode 1,723 amino acid residues with a calculated mass of 218 kDa. The peptide sequence determined for the purified enzyme is identified in the predicted sequence (Fig. 3), confirming that the isolated DNA is the gene for KGP.

The deduced amino acid sequence suggests that the first 22 amino acid residues containing a hydrophobic amino acid cluster may represent a signal peptide. The hydropathy of the NH_2-terminus of the precursor is high enough for it to function as a signal sequence translocating the protein across the membrane (not shown). The next 206-residue sequence is considered to be an NH_2-terminal propeptide,

B

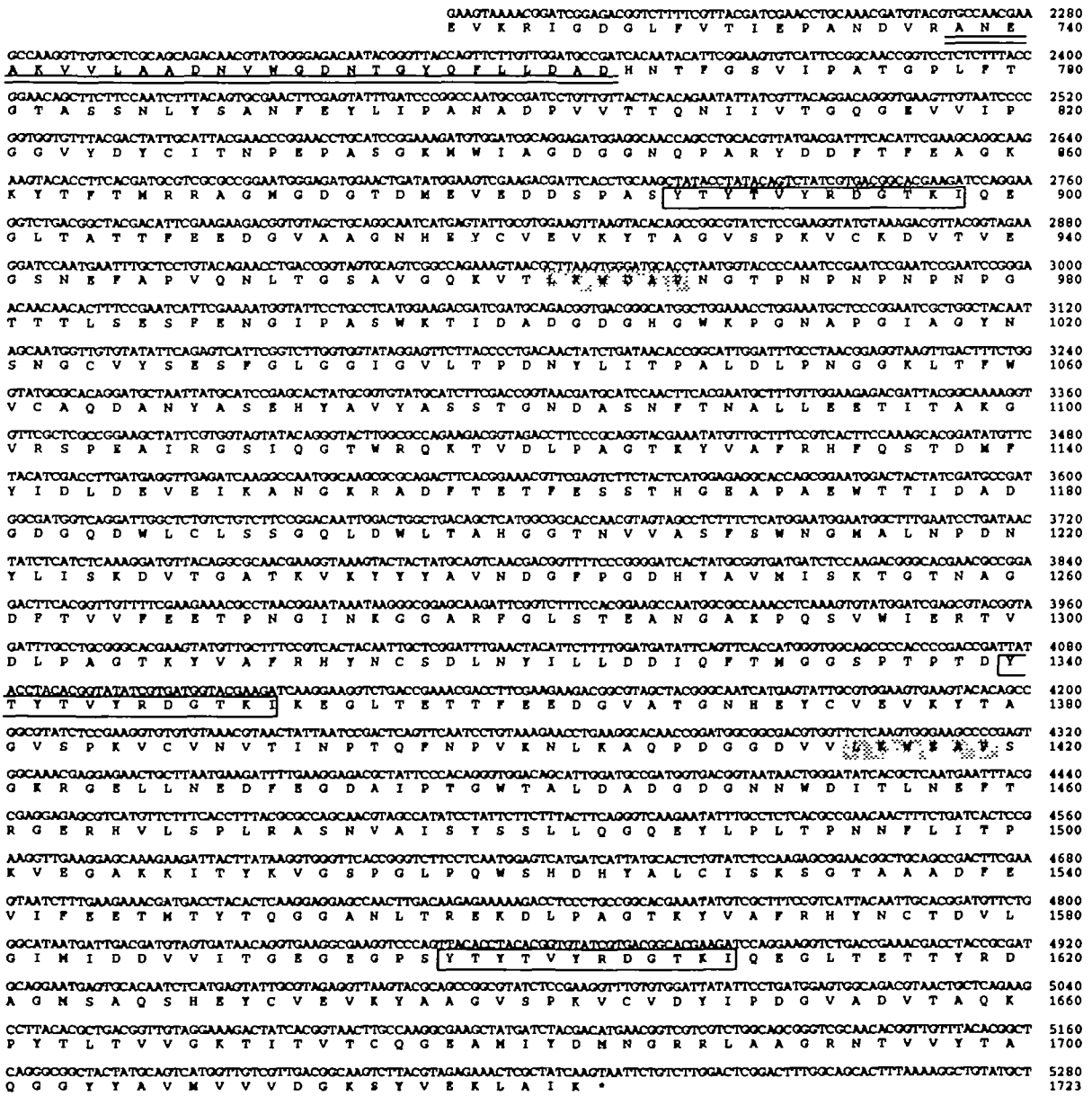

Fig. 3. **Nucleotide and deduced amino acid sequences of the *kgp* gene.** Nucleotides and predicted amino acids are numbered on the right. Amino acids are shown in the single letter code. Panel A shows the region including the signal peptide, the NH_2-terminal prosequence and the proteinase domain, and panel B represents the COOH-terminal hemagglutinin region. Amino acid sequence determined by Edman degradation of the purified enzyme is indicated by underlining. Shadows and boxes indicate the sequence repeats of LKWD(or E)AP and YTYTVYRDGTKI, respectively. The translational initiation and termination codons (TAA) are indicated by triple asterisks.

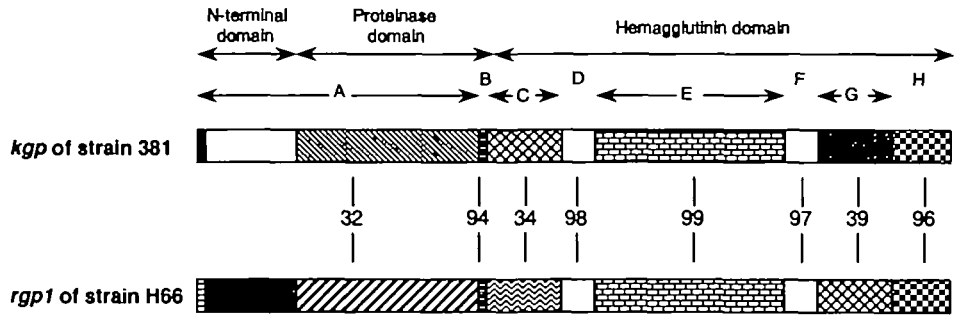

Fig. 4. **Structural similarity between the *kgp* gene of strain 381 and the *rgp1* gene of strain H66.** Nucleotide sequences of the *kgp* and *rgp1* genes were divided into eight portions based on the similarity in sequences and lengths between the two genes (A, the nucleotide number 60-2100; B, 2101-2166; C, 2167-2718; D, 2719-2943; E, 2944-4077; F, 4078-4299; G, 4300-4851; H, 4852-5229). The A region includes the signal peptide, the NH_2-terminal prosequence, and most of the proteinase domain. The A and G regions of the *kgp* gene were different from the corresponding regions of the *rgp1* gene, whereas more than 94% nucleotide sequence identity was observed in the regions B, D, E, F, H, and 33 bp after the terminal codons between the two genes. The amino acid sequences of the C regions of *kgp* and *rgp1* exhibited moderate similarity (42% identity).

TABLE I. **Similarity of the D and F regions of *kgp* and *rgp1*.** The upper right and lower left parts show the identity (%) in nucleotide and amino acid sequences, respectively. The identity was determined by using Gene Works software (IntelliGenetics).

	kgp(D)	*kgp*(F)	*rgp1*(D)	*rgp1*(F)
kgp(D)	—	76	98	76
kgp(F)	72	—	74	97
rgp1(D)	96	75	—	73
rgp1(F)	71	97	74	—

TABLE II. **Similarity of the C and G regions of *kgp* and *rgp1*.** The upper right and lower left parts show the identity (%) in nucleotide and amino acid sequences, respectively. The identity was determined by using Gene Works software (IntelliGenetics).

	kgp(C)	*kgp*(G)	*rgp1*(C)	*rgp1*(G)
kgp(C)	—	48	34	99
kgp(G)	19	—	34	39
rgp1(C)	42	15	—	24
rgp1(G)	98	20	42	—

because the sequence is followed by 25 amino acid residues identical with the NH_2-terminal sequence of the purified enzyme. The amino acid sequence of the mature enzyme was found to start with the 229th Asp residue and to end with the putative site of the 717th Arg residue. The remaining COOH-terminal portion is thought to be the COOH-terminal prosequence, which contains the hemagglutinin-related sequence starting with the 738th Ala residue, identical to the NH_2-terminal sequence of *P. gingivalis* hemagglutinin (*26, 33, 34*). Interestingly, the COOH-terminal region of KGP precursor contains three direct repeats of two different amino acid sequences, LKWD(or E)AP and YTYTVYRDGTKI. Thus, it is most likely that the precursor of KGP consists of at least four domains: the signal peptide, the NH_2-terminal prosequence, the proteinase domain, and the COOH-terminal hemagglutinin domain.

A comparison of the amino acid sequence of KGP with sequences of other proteins in the GenEMBL Sequence Data Library and the SwissProt Protein Sequence Data Library revealed that the sequences of the proteinase domain and the NH_2-terminal prosequence had no significant similarity with any other protein sequences, whereas the sequence of the COOH-terminal region exhibited significant similarity to those of RGP (*24, 35*) and PrpRI (*34*). It is of special interest to note that the three direct repeats of two different amino acid sequences, LKWD(or E)AP and YTYTVYRDGTKI, observed in the COOH-terminal region of KGP precursor were present in the COOH-terminal region of RGP-1 (*35*). To compare the nucleotide sequences of the *kgp* and *rgp1* genes and their deduced amino acid sequences precisely, *kgp* was divided into eight regions (Fig. 4) based on the similarity to *rgp1*. The A region of *kgp*, which was composed of the signal peptide, the NH_2-terminal domain, and most of the proteinase domain region, did not resemble that of *rgp1*, despite the similarity in the length of both regions. On the other hand, there was close similarity (more than 94% identity) in nucleotide sequence between the other regions (B, D, E, F, H, and 33 bp after the terminal codon) of *kgp* and *rgp1*. In addition, the D region was similar to the F region in both nucleotide and amino acid sequences (Table I). The nucleotide sequences of the C regions of *kgp* and *rgp1* were not similar to each other, but exhibited moderate similarity in amino acid sequence (42% identity), while the nucleotide and amino acid sequences of the G regions of the two genes were not similar (Table II). Surprisingly, the C region of *kgp* shared extremely high similarity (99% identity) in nucleotide sequence with the G region of *rgp1*.

TABLE III. **Proteolytic activity in the culture supernatants from single and double knock-out mutants.** KDP110 and KDP111 are single knock-out mutants. KDP112 is a double knock-out mutant.

	Proteolytic activity (%)	
	Boc-Val-Leu-Lys-MCA	Boc-Glu-Lys-Lys-MCA
ATCC33277	100	100
KDP110	21.3	30.2
KDP111	29.0	39.0
KDP112	19.1	24.4

Proteolytic Activity of KGP in RGP-Deficient Mutants— To examine the effect of disruption of the *rgp* gene on the translocation and activity of KGP, we constructed *rgp* knock-out mutants by using suicide plasmid systems which contain an internal DNA region of the *rgp* gene, as previously described (*25*), and measured the KGP activity in the supernatants of the respective mutants by using two synthetic substrates. As shown in Table III, in either *rgpA*- or *rgpB*-deficient single knock-out mutants (KDP110 and KDP111, respectively), the KGP activity in each supernatant was greatly decreased to 30-40% of the supernatant of the wild-type ATCC33277 strain. The KGP activity was more efficiently reduced in the *rgpA rgpB* double mutant (KDP112) than in the single mutants. These results indi-

cate that RGP is involved in the processing and translocation of KGP precursor in the cell.

DISCUSSION

The potential role of the proteolytic activity produced by *P. gingivalis* as a virulence factor has led to much work on the isolation and biological roles of the enzymes involved. Of these, the lysine-specific cysteine proteinases and the arginine-specific cysteine proteinase RGP are considered to be major etiologic enzymes of *P. gingivalis* because of their abilities to degrade various physiologically important proteins (*13, 22*). However, much of the data on these proteinases with regard to molecular mass, association with hemagglutinin activity, substrate specificity, and sensitivity to various protease inhibitors is ambiguous. Recently, Potempa *et al.* (*36*) reported that the single enzyme KGP is responsible for the trypsin-like activity with lysine-X specificity associated with this organism. Although partial NH_2-terminal amino acid sequences of its fragments have been presented (*26*), little is known about the gene for KGP. In this paper, we present for the first time the complete sequence of the cloned gene for the lysine-specific cysteine proteinase KGP from *P. gingivalis*. The initial translation product is a large precursor, which is composed of four functional regions (the signal peptide, the NH_2-terminal prosequence, the mature proteinase domain, and the COOH-terminal hemagglutinin domain). Since the secreted proteinase is devoid of the signal peptide, the NH_2-terminal prosequence, and the COOH-terminal hemagglutinin region, the initial precursor is found to be completely processed before or just after secretion from the cell into the culture medium. The signal peptide is thought to be essential for the inner membrane transport of the organism. The NH_2-terminal prosequence is assumed to stabilize the proprotein structure during transport. Since the COOH-terminal region contains amphipathic β-sheet structures interspersed by hairpin turns and loops, as revealed by the method of Chou and Fasman (*37*), and since such characteristic structures have also been shown in the COOH-terminal prosequence of other extracellular proteases from Gram-negative bacteria such as *Neisseria gonorrhoeae* (*38, 39*), *Serratia marcescens* (*40*), and *Thermus aquaticus* (*41*), and suggested to contribute to secretion through the outer membrane, this region may be important in translocating the proenzyme across the outer membrane, besides its association with hemagglutination. At present the precise active site cysteine residue of KGP remains to be established. However, considering that the structural feature of each domain of KGP is basically similar to that of the counterpart of RGP and that the reactive cysteine residue of RGP is assigned to Cys^{471} (numbering according to Ref. *24*) based on the structural similarity to the active site region of clostripain (*24*), or more directly by labeling the active site cysteine residue with $N\alpha$-acetyllysine chloromethyl ketone (*42*), the putative active cysteine residue of KGP may be Cys^{477} or Cys^{476}.

Although both the NH_2-terminal prepropeptide and the COOH-terminal propeptide were eliminated from the initial product during its transport to form the mature enzyme, the enzyme(s) responsible for this processing are not precisely known. However, we have recently reported that RGP is a processing proteinase responsible for the cell surface proteins, such as fimbrilin and the 75-kDa protein (*43*), and RGP itself (*24*). Taken with this finding, the observation that the proteolytic activity of KGP was markedly decreased in the culture media of the RGP-deficient mutants strongly suggests that RGP is closely related to the processing and secretion of KGP.

Another striking structural feature of KGP is found in the COOH-terminal region, in which there are three direct repeats of the two different sequences LKWD(or E)AP and YTYTVYRDGTKI. These two sequences are also found in the COOH-terminal region of the *rgp* genes (*rgpA*, formerly *agp*, of strain 381 and *rgp1* of strain H66) (*24, 35*). Although the *rgpA* gene of 381 (*24*) is apparently homologous to the *rgp1* gene of H66 (*35*), there is a significant size difference in the COOH-terminal region between the two genes. The *rgp1* gene of H66 has the same three direct repeats of the two different sequences as the present *kgp* gene, whereas the *rgpA* gene of 381 is devoid of two of the three repeats (Fig. 5). Several genes, including *prtH* of W83 (*44*), *rgpA* of 381, *rgp1* of H66, *prtR* of W50 (*33*), and *prpR1* of W50 (*34*), for arginine-specific cysteine proteinases associated with the organism have been described so far. Thus, the reason for the size difference in the COOH-terminal region between *rgpA* of 381 and *rgp1* of H66 is not known at this time. It may be due to strain differences or cloning artifacts.

Comparison of the nucleotide sequences of *kgp* and *rgp1* revealed that the nucleotide sequence ranging from the B region to the site located 33 bp downstream of the terminal codon, except for the C and G regions, is almost identical between *kgp* and *rgp1*. Recombinational rearrangement such as transposition or gene conversion may have occurred in this nucleotide region between *kgp* and *rgp1*. At least two

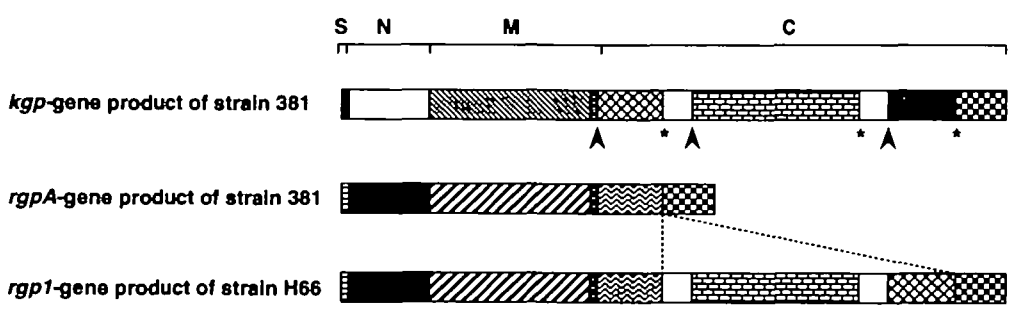

Fig. 5. Comparison of the preproprotein structure of KGP encoded by *kgp* gene of strain 381 and the precursor structures of RGP encoded by *rgpA* of strain 381 and *rgp1* of strain H66. The proposed domain structures of KGP of 381 were compared with those of the gene products of *rgpA* of 381 and *rgp1* of H66. S, signal peptide; N, NH_2-terminal prosequence; M, mature proteinase domain; C, COOH-terminal prosequence. The arrowheads and asterisks indicate the repeated sequences of LKWD(or E)-AP and YTYTVYRDGTKI, respectively. A part of the COOH-terminal prosequence of KGP as well as the *rgp1* gene product is missing in the COOH-terminal prosequence of the *rgpA* gene product, which is marked by dashed lines.

other DNA regions on the *P. gingivalis* chromosome share homology with this region (*25*). Therefore, these DNA regions may also have taken part in this recombinational event. In addition, it is possible that these homologous DNA has been supplied from the chromosomal DNA of other *P. gingivalis* cells (horizontal gene transfer). The C region of *kgp* has very close similarity in nucleotide sequence with the G region of *rgp1*. This result strongly suggests that recombinational shuffling has taken place between these regions. We found that each of the two direct repeats was located at each end of the C and G regions. These nucleotide repeats may have assisted in this rearrangement.

Although we cannot determine whether these recombinational events have been generated through intrachromosomal rearrangement or interchromosomal rearrangement by horizontal gene transfer, we tend to consider that gene-conversion type intrachromosomal rearrangement accounts for these events since there has been no report concerning natural transformation in *P. gingivalis* and since gene-conversion type recombination in *P. gingivalis* has been observed in a previous study (*45*). Gene conversion with respect to bacterial pathogenesis has been found in the pilin of *Neisseria gonorrhoeae* and the variable major protein of *Borrelia hermsii* (*46, 47*). Antigenic variation of the proteins of these bacteria, which is caused by gene conversion, provides them with the advantage of protection from immunological challenge by their hosts. Therefore, it is of particular interest to determine whether the recombinational rearrangement of the *kgp* and *rgp* genes encoding major cysteine proteinases of *P. gingivalis* contributes to its virulence.

REFERENCES

1. Mayrand, D. and Holt, S.C. (1988) Biology of asaccharolytic black-pigmented *Bacteroides* species. *Microbiol. Rev.* **52**, 134-152
2. Moore, W.E.C., Holdeman, L.V., Smibert, R.M., Hash, D.E., Burmeister, J.A., and Ranney, R.R. (1982) Bacteriology of severe periodontitis in young adult humans. *Infect. Immun.* **38**, 1137-1148
3. Zambon, J.J., Reynolds, H.S., and Slots, J. (1981) Black-pigmented *Bacteroides* spp. in the human oral cavity. *Infect. Immun.* **32**, 198-203
4. Marsh, P.D., McKee, A.S., McDermid, A.S., and Dowsett, A.B. (1989) Ultrastructure and enzyme activities of a virulent and an avirulent variant of *Bacteroides gingivalis* W50. *FEMS Microbiol. Lett.* **50**, 181-185
5. Grenier, D. and Mayrand, D. (1987) Selected characteristics of pathogenic and nonpathogenic strains of *Bacteroides gingivalis*. *J. Clin. Microbiol.* **25**, 738-740
6. Birkedal-Hansen, H., Taylor, R.E., Zambon, J.J., Barwa, P.K., and Neiders, M.E. (1988) Characterization of collagenolytic activity from strains of *Bacteroides gingivalis*. *J. Periodont. Res.* **23**, 258-264
7. Sundqvist, G., Carlsson, J., and Hänström, L. (1987) Collagenolytic activity of black-pigmented *Bacteroides* species. *J. Periodont. Res.* **22**, 300-306
8. Otsuka, M., Endo, J., Hinode, D., Nagata, A., Maehara, R., Sato, M., and Nakamura, R. (1987) Isolation and characterization of protease from culture supernatant of *Bacteroides gingivalis*. *J. Periodont. Res.* **22**, 491-498
9. Sorsa, T., Uitto, V.-J., Suomalainen, K., Turto, H., and Lindy, S. (1987) A trypsin-like protease from *Bacteroides gingivalis*: Partial purification and characterization. *J. Periodont. Res.* **22**, 375-380
10. Smalley, J.W., Birss, A.J., and Shuttleworth, C.A. (1988) The degradation of type I collagen and human plasma fibronectin by the trypsin-like enzyme and extracellular membrane vesicles of *Bacteroides gingivalis* W50. *Arch. Oral Biol.* **33**, 323-329
11. Grenier, D. (1992) Inactivation of human serum bactericidal activity by a trypsinlike protease isolated from *Porphyromonas gingivalis*. *Infect. Immun.* **60**, 1854-1857
12. Chen, Z., Potempa, J., Polanowski, A., Renvert, S., Wikström, M., and Travis, J. (1991) Stimulation of proteinase and amidase activities in *Porphyromonas* (*Bacteroides*) *gingivalis* by amino acids and dipeptides. *Infect. Immun.* **59**, 2846-2850
13. Scott, C.F., Whitaker, E.J., Hammond, B.F., and Colman, R.W. (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
14. Curtis, M.A., Ramakrishnan, M., and Slaney, M. (1993) Characterization of the trypsin-like enzymes of *Porphyromonas gingivalis* W83 using a radiolabelled active-site-directed inhibitor. *J. Gen. Microbiol.* **139**, 949-955
15. Chen, Z., Potempa, J., Polanowski, A., Wikstrom, M., and Travis, J. (1992) Purification and characterization of a 50-kDa cysteine proteinase (gingipain) from *Porphyromonas gingivalis*. *J. Biol. Chem.* **267**, 18896-18901
16. Sundqvist, G., Carlsson, J., Herrmann, B.F., and Tärnvik, A. (1985) Degradation of human immunoglobulins G and M and complement factors C3 and C5 by black-pigmented *Bacteroides*. *J. Med. Microbiol.* **19**, 85-94
17. Sundqvist, G.K., Carlsson, J., Herrmann, B.F., Höfling, J.F., and Väätäinen, A. (1984) Degradation in vivo of the C3 protein of guinea-pig complement by a pathogenic strain of *Bacteroides gingivalis*. *Scand. J. Dent. Res.* **92**, 14-24
18. Wingrove, J.A., DiScipio, R.G., Chen, Z., Potempa, J., Travis, T., and Hugli, T.E. (1992) Activation of complement components C3 and C5 by a cysteine proteinase (gingipain-1) from *Porphyromonas* (*Bacteroides*) *gingivalis*. *J. Biol. Chem.* **267**, 18902-18907
19. Nilsson, T., Carlsson, J., and Sundqvist, G. (1985) Inactivation of key factors of the plasma proteinase cascade systems by *Bacteroides gingivalis*. *Infect. Immun.* **50**, 467-471
20. Lantz, M., Rowland, R.W., Switalski, L.M., and Höök, M. (1986) Interactions of *Bacteroides gingivalis* with fibrinogen. *Infect. Immun.* **54**, 654-658
21. Carlsson, J., Herrmann, B.F., Höfling, J.F., and Sundqvist, G.K. (1984) Degradation of the human proteinase inhibitors alpha-1-antitrypsin and alpha-2-macroglobulin by *Bacteroides gingivalis*. *Infect. Immun.* **43**, 644-648
22. Fujimura, S., Shibata, Y., and Nakamura, T. (1992) Comparative studies of three proteases of *Porphyromonas gingivalis*. *Oral Microbiol. Immunol.* **7**, 212-217
23. Kadowaki, T., Yoneda, M., Okamoto, K., Maeda, K., and Yamamoto, K. (1994) Purification and characterization of a novel arginine-specific cysteine proteinase (argingipain) involved in the pathogenesis of periodontal disease from the culture supernatant of *Porphyromonas gingivalis*. *J. Biol. Chem.* **269**, 21371-21378
24. Okamoto, K., Misumi, Y., Kadowaki, T., Yoneda, M., Yamamoto, K., and Ikehara, Y. (1994) 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
25. Nakayama, K., Kadowaki, T., Okamoto, K., and Yamamoto, K. (1995) Construction and characterization of arginine-specific cysteine proteinase (Arg-gingipain)-deficient mutants of *Porphyromonas gingivalis*. *J. Biol. Chem.* **270**, 23619-23626
26. Pike, R., McGraw, W., Potempa, J., and Travis, J. (1994) Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. *J. Biol. Chem.* **269**, 406-411
27. Ciborowski, P., Nishikata, M., Allen, R.D., and Lantz, M.S. (1994) Purification and characterization of two forms of a high-molecular-weight cysteine proteinase (porphypain) from *Porphyromonas gingivalis*. *J. Bacteriol.* **176**, 4549-4557
28. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
29. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequenc-

ing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
30. Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D. (1988) λZAP: A bacteriophage λ expression vector with *in vivo* excision properties. *Nucleic Acids Res.* **16**, 7583-7600
31. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, pp. 368-369, Cold Spring Harbor Laboratory, Cold Spring Harbor
32. Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132
33. Kirszbaum, L., Sotiropoulos, C., Jackson, C., Cleal, S., Slakeski, N., and Reynolds, E.C. (1995) Complete nucleotide sequence of a gene *prtR* of *Porphyromonas gingivalis* W50 encoding a 132 kDa protein contains an arginine-specific thiol endopeptidase domain and a haemagglutinin domain. *Biochem. Biophys. Res. Commun.* **207**, 424-431
34. Aduse-Opoku, J., Muir, J., Slaney, J.M., Rangarajan, M., and Curtis, M.A. (1995) Characterization, genetic analysis, and expression of a protease antigen (PrpRI) of *Porphyromonas gingivalis* W50. *Infect. Immun.* **63**, 4744-4754
35. Pavloff, N., Potempa, J., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J., and Barr, P.J. (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
36. Potempa, J., Pike, R., and Travis, J. (1995) The multiple forms of trypsin-like activity present in various strains of *Porphyromonas gingivalis* are due to the presence of either Arg-gingipain or Lys-gingipain. *Infect. Immun.* **63**, 1176-1182
37. Chou, P.Y. and Fasman, G.D. (1978) Prediction of the secondary structure of proteins from their amino acid sequence in *Advances in Enzymology* (Meister, A., ed.) Vol. 47, pp. 45-148, Wiley, New York/Chichester/Brisbane/Toronto
38. Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T.F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325**, 458-462
39. Klauser, T., Krämer, J., Otzelberger, K., Pohlner, J., and Meyer, T.F. (1993) Characterization of the *Neisseria* Iga_β-core. The essential unit for outer membrane targeting and extracellular protein secretion. *J. Mol. Biol.* **234**, 579-593
40. Yanagida, N., Uozumi, T., and Beppu, T. (1986) Specific excretion of *Serratia marcescens* protease through the outer membrane of *Escherichia coli*. *J. Bacteriol.* **166**, 937-944
41. Terada, I., Kwon, S.-T., Miyata, Y., Matsuzawa, H., and Ohta, T. (1990) Unique precursor structure of an extracellular protease, aqualysin I, with NH_2- and COOH-terminal pro-sequences and its processing in *Escherichia coli*. *J. Biol. Chem.* **265**, 6576-6581
42. Nishikata, M. and Yoshimura, F. (1995) Active site structure of a hemagglutinating protease from *Porphyromonas gingivalis*: Similarity to clostripain. *Biochem. Mol. Biol. Int.* **37**, 547-553
43. Nakayama, K., Yoshimura, F., Kadowaki, T., and Yamamoto, K. (1996) Involvement of arginine-specific cysteine proteinase (Arg-gingipain) in fimbriation of *Porphyromonas gingivalis*. *J. Bacteriol.* **178**, 2818-2824
44. Fletcher, H.M., Schenkein, H.A., and Macrina, F.L. (1994) Cloning and characterization of a new protease gene (*prtH*) from *Porphyromonas gingivalis*. *Infect. Immun.* **62**, 4279-4286
45. Nakayama, K. (1994) Rapid viability loss on exposure to air in a superoxide dismutase-deficient mutant of *Porphyromonas gingivalis*. *J. Bacteriol.* **176**, 1939-1943
46. Seifert, H.S. and So, M. (1988) Genetic mechanisms of bacterial variation. *Microbiol. Rev.* **52**, 327-336
47. Zang, Q.Y., DeRyckere, D., Lauer, P., and Koomey, M. (1992) Gene conversion in *Neisseria gonorrhoeae*: Evidence for its role in pilus antigenic variation. *Proc. Natl. Acad. Sci. USA* **89**, 5366-5370