INFECTION AND IMMUNITY, Dec. 1995, p. 4744–4754 Vol. 63, No. 12
0019-9567/95/$04.00+0

Characterization, Genetic Analysis, and Expression of a Protease Antigen (PrpRI) of *Porphyromonas gingivalis* W50

JOSEPH ADUSE-OPOKU,* JUSTINE MUIR, JENNIFER M. SLANEY, MINNIE RANGARAJAN, AND MICHAEL A. CURTIS

MRC Molecular Pathogenesis Group, Department of Oral Microbiology, London Hospital Medical College, London E1 2AA, England

Received 19 May 1995/Returned for modification 4 August 1995/Accepted 20 September 1995

Previous studies of the serum immunoglobulin G antibody response of periodontal patients have demonstrated significant reactivity to a cell surface or extracellular arginine-specific protease of *Porphyromonas gingivalis* which migrates as an ~50-kDa band on sodium dodecyl sulfate-polyacrylamide gels. In the present report, two forms of the enzyme (ArgI and ArgIA) with this electrophoretic behavior were isolated. ArgI is a heterodimer of α and β subunits, and ArgIA is a monomer composed of the catalytically active α component alone. The gene encoding ArgI (*prpR1* encoding protease polyprotein ArgI) was cloned from *Sau3*AI digests of *P. gingivalis* W50 DNA into pUC18. Sequence analysis demonstrated that the α and β components are contiguous on the initial translation product and are flanked by large N- and C-terminal extensions. *prpR1* is 97.5% identical to the *rgp-1* gene from *P. gingivalis* H66. *prpR1* expression in *Escherichia coli* demonstrated the presence of an internal transcription-translation initiation site which could permit independent expression of different regions of the polyprotein. Immunochemical analysis of *P. gingivalis* mid-logarithmic-phase cultures suggested that the processing of PrpRI may be closely coupled to its synthesis, with only the final stages taking place at the cell surface. Southern hybridization studies demonstrated that the *prpR1* gene is widely distributed in other *P. gingivalis* strains and that a second homologous locus to the α component and at least two other homologous loci to the β component are present on the *P. gingivalis* chromosome. These data indicate that the ArgI protease of *P. gingivalis* is a member of a family of sequence-related gene products which may share both functional and antigenic properties.

The periodontal diseases, a group of chronic inflammatory conditions of the tooth-supporting tissues, are initiated by components of the microbial biofilm which develops on the hard tooth surface adjacent to the soft tissues of the periodontium. In chronic adult periodontal disease, the most prevalent of these conditions and probably the most common chronic infectious disease in humans, the biofilm comprises an exceedingly complex assortment of different bacterial species, many of which have been implicated in the inflammatory disease process (22). Of these, the anaerobic, gram-negative rod *Porphyromonas gingivalis* is considered to be particularly relevant to the etiology of the disease because of its low frequency of isolation from healthy mouths and sites (24), its association with progressive disease (40), and the demonstration of high serum immunoglobulin G (IgG) antibody levels to this bacterium in individuals with a history of destructive disease (11).

The specific immune response to this organism and other cocolonizing bacteria is considered to play a significant and pivotal role in the defense of the periodontal tissues. The orchestration of the phagocytic cell and complement arms of the host defense, the neutralization of extracellular bacterial products, and the prevention of antigen entry into the tissues are all highly dependent on the presence of a fully functional specific antibody, which is therefore likely to comprise an essential component of the controlling influence of the subgingival microbiota. Conversely, however, overzealous complement activation, phagocytic cell recruitment and degranulation, and immune complex formation, which may be considered to be the hallmarks of a chronically inflamed, gingival crevice (5), are also potential outcomes of this IgG-mediated defensive strategy. Hence, effective defense can be viewed as a delicate balance between the competing forces of the microbial insult and the host response to this challenge. Shifts in that equilibrium, toward either a more uncontrolled microbial challenge on the one hand or a more extreme, deregulated inflammatory response on the other, would be predicted to have deleterious consequences on the integrity of the surrounding tissues.

In order to examine aspects of this dynamic, we have been pursuing the characterization of the structure and function of the major surface antigens of *P. gingivalis* to serve as model antigens in the study of host-parasite interactions in periodontal diseases. Initially, the IgG antibody response to soluble and outer membrane preparations of this organism was determined in sera from a case-control study of adult periodontal disease, and a number of discriminatory surface antigens were identified (7). One of these antigens, with an M_r of ~50,000, is readily released from the cell surfaces of all *P. gingivalis* strains examined to date but is absent from type and reference strains of other related oral and intestinal anaerobic bacteria (10) and thus satisfies some of the criteria needed for a model antigen.

Preliminary assignment of the function of this protein was determined via radiolabelling of cell sonicates and culture supernatants of *P. gingivalis* with an irreversible protease inhibitor, ^{125}I-tyrosyl-alanyl-lysyl-arginyl chloromethyl ketone (^{125}I-YAKR), which binds to and inactivates serine and cysteine proteases with specificity for arginine in the P1 position of target peptide bonds (6). These experiments demonstrated that a significant proportion of the total cell-bound and extracellular Arg-X (where X is any amino acid) protease activity of this organism is associated with a protein with electrophoretic

* Corresponding author. Mailing address: MRC Molecular Pathogenesis Group, Department of Oral Microbiology, The London Hospital Medical College, 32 Newark St., London E1 2AA, United Kingdom. Phone: (0171) 377 0444. Fax: (0171) 247 0657. Electronic mail address: jaduse@lhmc.ac.uk.

behavior identical to that of the ~50-kDa surface antigen. This observation has prompted us to examine the structural and biochemical characteristics of the extracellular Arg-X protease activity in *P. gingivalis* (1, 34).

The aims of this study were to determine the primary structure, genomic organization, and distribution of the protease antigen of *P. gingivalis* and to examine expression of the cloned gene in *E. coli.*

MATERIALS AND METHODS

Bacteria, culture conditions, and plasmids. The following bacteria were used in this study: *P. gingivalis* W50, W50Be1, W50Br1, W83, LB13D-3, A7436, 381, 33277, and 11834; *Porphyromonas asaccharolytica* ATCC 25260 (all from sources described previously, (25); and *E. coli* XL-1 Blue (Stratagene). *Porphyromonas* species were maintained on blood agar plates or grown in brain heart infusion broth supplemented with hemin in an atmosphere of N_2-H_2-CO_2 (80:10:10) at 37°C. *E. coli* strains were grown in Luria broth (LB) (tryptone, 1%; yeast extract, 0.5%; NaCl, 0.5%), and the medium for XL-1 Blue was supplemented with 20 μg of tetracycline ml^{-1}. Ampicillin was added to a final concentration of 50 μg ml^{-1} for plasmid selection. Where appropriate, protease substrates were added to the LB agar to a final concentration of 0.4% for azoalbumin and azocasein or 1% for skimmed milk. The plasmid pUC18 (46) was used throughout. Pure plasmid DNA was prepared by ion-exchange chromatography on Qiagen columns (Qiagen Ltd.), and minipreps were prepared as described by Clark-Curtiss and Curtiss (4). *P. gingivalis* W50 was used for all genomic cloning.

Purification of Arg-X proteases of *P. gingivalis* W50. Enzymes with specificity for Arg-X peptide bonds were purified from 6-day-culture supernatants of *P. gingivalis* W50 by a combination of ammonium sulfate precipitation, gel filtration, affinity chromatography on arginine-agarose, and ion exchange with L-benzoyl-arginyl *para*-nitroanilide (BApNA) to monitor activity (33a, 34).

N-terminal amino acid sequences of the resulting proteins were generated via automated Edman degradation by using an Applied Biosystems model 377A gas-phase automated sequencer. Internal protein sequences of the enzymes were obtained following cyanogen bromide digestion, separation of the resulting fragments by reverse-phase chromatography, and automated N-terminal sequencing and by sequencing the autolytic degradation products.

Antiserum Rb3158 was raised against *P. gingivalis* W50 culture supernatant proteins eluted from arginine-Sepharose as described previously (6), followed by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). A homogeneous band with an M_r of ~50,000 was excised from the gel, finely chopped, and used as the immunogen. Two immunizations of 150 μg of protein were given 3 weeks apart to hybrid lop-earred Dutch rabbits, and blood was withdrawn and separated for antiserum after an additional 2 weeks.

The collection of periodontal case and control sera was described in a previous report (7).

SDS-PAGE. SDS-PAGE was performed according to the method of Laemmli (19) on 6 or 10% separating gels. Western blotting (immunoblotting) was carried out in a bicarbonate transfer buffer (3 mM Na_2CO_3, 10 mM $NaHCO_3$, [pH 9.9], 20% [vol/vol] methanol) at a constant current of 400 mA for 2.5 h. After blotting, the nitrocellulose membranes were blocked in 5% bovine serum albumin and incubated overnight in primary antibody (1:1,000). Antibody binding was detected by using horseradish peroxidase-conjugated anti-species immunoglobulins and diaminobenzidine (0.05%). For N-terminal microsequencing of purified *P. gingivalis* enzymes and recombinant products, proteins were electroblotted onto polyvinylidene difluoride Immobilon-P (Millipore) membranes, and the membranes were briefly stained with Coomassie blue R for identification of the protein bands.

Preparation of bacterial chromosomal DNA. The majority of the molecular biology techniques were performed as described by Sambrook et al. (37). Stationary-phase bacterial cultures (20 ml) were harvested by centrifugation and resuspended in 10 ml of a solution containing 10 mM EDTA–50 mM Tris-HCl buffer (pH 8.0) (TE) and 100 μg of RNase A ml^{-1}. After 10 min at room temperature, 50 mg of lysozyme-HCl was added, and the mixture was incubated at 37°C for a further 30 min. Proteinase K (1 mg) was added, and incubation was continued for a further 30 min. SDS was added to a final concentration of 2%, and after gentle mixing, the solution was incubated for another 10 min. DNA was further purified by repeated phenol-chloroform extraction followed by precipitation with twice the volume of absolute ethanol in the presence of 0.5 M ammonium acetate. It was then spooled onto a glass rod, washed extensively in 70% ethanol, and dissolved in 3 to 5 ml of TE. The yield and purity of the DNA were assessed following agarose gel electrophoresis and restriction enzyme digestion.

Construction of recombinant plasmids. *P. gingivalis* W50 chromosomal DNA was partially digested with the restriction endonuclease *Sau3*AI. The fragments were fractionated by preparative agarose gel electrophoresis. DNA within the size range of 0.5 to 12 kb was excised from the gel, electroeluted into TE buffer, and purified by extraction with the WIZARD clean-up kit (Promega). *P. gingivalis* DNA fragments were ligated to *Bam*HI-restricted pUC18 with T4 DNA ligase (Pharmacia). The ligation mixture was used to transform competent *E. coli* XL-1 Blue, and transformants were selected on LB agar plates supplemented with ampicillin and tetracycline. Colony lifts were performed on either nitrocellulose (for immunoscreening) or Hybond N^+ membranes (for DNA hybridization). The nitrocellulose membranes were placed on LB agar with isopropyl-β-D-thiogalactopyranoside (IPTG), and growth was continued for a further 4 h prior to cell lysis and screening with Rb3158.

Screening of genomic library. *E. coli* XL-1 Blue transformants were screened by using antiserum Rb3158, preadsorbed with an *E. coli* extract, at a final dilution of 1:2,000. The second antibody, swine anti-rabbit IgG conjugated to horseradish peroxidase (Dako), was used at a dilution of 1:500. 3,3′-Diaminobenzidine (0.05%) was used as the developing substrate.

Restriction mapping and localization of gene. Subclones of one of the immunoreactive clones were constructed by removing various insert DNA fragments from the recombinant plasmid with restriction enzymes. DNA fragments were excised out of agarose gels, electroeluted in TE, further purified with WIZARD (Promega), and recircularized with T4 DNA ligase (sometimes, Ready To Go DNA ligase [Pharmacia] was used). The ligation mixture was then used to transform competent *E. coli* XL-1 Blue. Recombinant protein expression in ampicillin- and tetracycline-resistant transformants was then examined via Western blotting with antiserum Rb3158 preadsorbed with *E. coli* XL-1 Blue (pUC18).

DNA hybridization. Cells immobilized on Hybond N^+ (Amersham) were lysed, and the DNA was fixed with 0.4 M NaOH in accordance with the manufacturer's instructions. After being rinsed, the membranes were prehybridized for 1 h at 65°C in Rapid Hyb (Amersham) prior to hybridization for 4 to 18 h in the same solution with heat-denatured ^{32}P-labelled probes. The membranes were then washed thoroughly under stringent conditions with 0.2× standard saline citrate containing 0.1% SDS at 65°C. Southern hybridization was performed in a similar manner, except DNA was electrophoresed in agarose gels, and, following depurination in 0.25 M HCl, was transferred under alkaline conditions (0.4 M NaOH) onto Hybond N^+ by vacuum blotting (Vacu-aid; Hybaid). DNA was labelled by the random primer method by using the T7-Quickprime or Ready To Go labelling system (Pharmacia) and [α-^{32}P]dCTP (3,000 mCi/mmol; Amersham). Hybridization with inosine (I) containing oligonucleotides, ADU21, (5′-AA[A,G]TA[T,C]GAIGGIGA[C,T]ATIAA[A,G]GA[C,T]TT[C,T]GTIGA[C,T]TGGAA-3′), designed to the N terminus of ArgIα, was performed at 42°C as described by Monaco et al. (26). DNA was end labelled with T4-polynucleotide kinase and [α-^{32}P]dATP. After completion of the reaction, the unincorporated nucleotides were removed by passage over NAP-5 columns (Pharmacia). Autoradiography was performed by using Kodak X-Omat-LS films at −70°C, with intensifying screens, for 1 to 18 h.

DNA sequencing. Double-stranded nested deletion derivatives of pJM2 were generated by using a kit from Pharmacia and *Sst*I (residual multiple cloning site [MCS] in pJM2) and *Nco*I. Plasmids with progressive deletion of the insert were prepared on a small scale, and the extent of insert deletion was assessed by restriction with *Hin*dIII and *Eco*RI; both enzymes restrict only at the residual MCS of pJM2. For sequencing, plasmid DNA was prepared by ion-exchange chromatography. *Taq* cycle sequencing was performed on pJM2 derivatives by using PRISM (Applied Biosystems) and both M13 universal forward and reverse primers, and the products were analyzed on an automated DNA sequencer (model 373A; Applied Biosystems). Oligonucleotide primers were used to complete the sequence and were synthesized on either a Gene Assembler 4 Plus (Pharmacia) or a 392 DNA/RNA synthesizer (Applied Biosystems) followed by purification over NAP-10 columns (Pharmacia). DNA sequence data were analyzed by using the University of Wisconsin's Genetics Computer Group software (9) in conjunction with the Staden DNA and protein analysis softwares (National Institute for Medical Research, London, England).

RESULTS

Purification of extracellular Arg-X proteases of *P. gingivalis* W50. Separation of the culture supernatant of *P. gingivalis* W50 on gel filtration, arginine-agarose, and ion-exchange chromatography yielded two distinct forms of protease with specificity for Arg-X peptide bonds (ArgI and ArgIA) which migrate as 50-kDa proteins on reducing SDS-PAGE gels (Fig. 1A). Separation of the two forms was achieved largely through their behavior on arginine-agarose at pH 6 in 25 mM MOPS [3-(*N*-morpholino)propanesulfonic acid] buffer containing 200 mM NaCl and 10 mM $CaCl_2$ Under these conditions, ArgI remained tightly bound to the resin, while ArgIA was not retained. A third form which migrated as a diffuse band in the 70- to 80-kDa region (ArgIB) on similar gels was also obtained during the course of these separations (data not shown). Together these three forms comprise >90% of the total extracellular Arg-X protease activity of *P. gingivalis* W50. N-terminal sequence analysis revealed that all three forms possess the

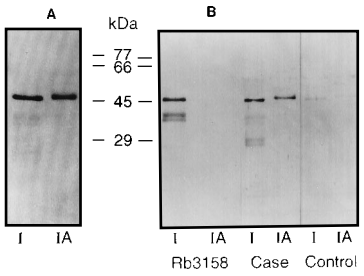


FIG. 1. Comparison of two forms of Arg-X proteases from *P. gingivalis*. (A) SDS-PAGE gel of 2 μg of ArgI (I) and ArgIA (IA), developed with Coomassie blue; (B). Western blot of ArgI and ArgIA probed with Rb3158 and pooled periodontal patients' (case) and control serum samples. The molecular sizes of the marker proteins are shown between the two panels.

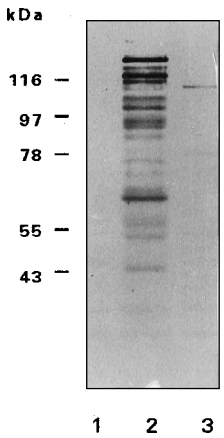


FIG. 2. Western blot analysis of whole-cell lysates from 15 μl of stationary-phase cultures of clones isolated by immunoscreening with Rb3158 antiserum. Lanes: 1, UC18 control; 2, JM2; 3, JM7. The molecular sizes of the marker proteins are indicated on the left.

same N terminal sequence (YTPVEEK. . .), but while ArgIA and ArgIB are monomeric, ArgI is a heterodimeric species which contains a second comigrating polypeptide on SDS-PAGE gels with an N-terminal sequence of SGQAE. . . (34). The comigrating polypeptide was extremely susceptible to autolytic degradation, which in some preparations led to the formation of numerous lower-molecular-weight species on SDS-PAGE gels, all of which had the N-terminal sequence SGQAE. . .. The ArgI preparation shown in Fig. 1A was largely intact as determined by Coomassie blue-staining and SDS-PAGE, but Western blotting with hyperimmune sera did demonstrate the presence of the lower-molecular-weight breakdown products of this comigrating species.

On the basis of these observations we have adopted the following nomenclature to describe these proteases. ArgI has an α and β subunit structure in which the α component (N-terminal sequence of YTPVEEK. . .) bears the enzyme active site. ArgIA represents the α component alone. ArgIB is a highly posttranslationally modified form of the α component, which may explain its aberrant behavior on SDS-PAGE gels and has a considerably reduced solubility compared with that of the other two forms. A more complete description of the purification strategy and the physical and kinetic differences between these enzymes will be presented elsewhere. The internal protein sequences derived from cyanogen bromide digestion and from a peptide with a molecular weight of 14,000 derived from autolysis of ArgI are described later.

In order to verify that ArgI and (or) ArgIA represented the discriminatory antigen that we had previously described in sonicates of *P. gingivalis*, their immunoreactivities with a pool of sera from periodontal patients, which had previously been shown to contain high levels of IgG antibody reactive with *P. gingivalis* whole cells, were examined. Both forms were strongly recognized by Western blots with these sera compared with the reaction with pooled control human sera at the same dilution (Fig. 1B). We were therefore able to conclude that at least the α component of ArgI is recognized. In addition however, ongoing studies have now demonstrated that a significant immune response to the β chain of the ArgI heterodimer is also evident in sera from periodontal patients (data not shown).

Rabbit antiserum (Rb3158) prepared to an ~50-kDa band excised from preparative SDS-PAGE gels of proteins bound to arginine-Sepharose was employed as a preliminary screening reagent in the cloning of the ArgI gene. While Rb3158 is a high-titer antiserum for the detection of ArgI, subsequent immunoblots demonstrated that it reacts only poorly with ArgIA (Fig. 1B), indicating that the β chain had been the dominant immunogen in the immunization protocol for this animal. As we describe in the next section, Southern hybridization experiments using oligonucleotides to the N terminus of ArgIα revealed that both the α and β components of ArgI were on the same *P. gingivalis* DNA fragment, and hence an initial screening for transformants expressing either component would be a satisfactory strategy for cloning the relevant gene. Rb3158 antiserum was used to screen a pUC18 library in *E. coli* XL-1 Blue containing the *P. gingivalis* W50 DNA fragments.

Cloning of gene encoding ArgI. Three immunoreactive clones (JM2, JM7, and JM11) were isolated. Western blots of these clones with Rb3158 antiserum are shown in Fig. 2 (also, see Fig. 7). These were also screened with a ^{32}P-labelled oligonucleotide (with inosine residues at positions of base ambiguity) to the N terminus of the α subunit of ArgI (ADU21). JM2 was the only clone which hybridized with ADU21 (data not shown). Subsequent restriction mapping and nucleotide sequence analysis demonstrated that pJM11 contained an internal fragment of the pJM2 insert and that pJM7 was a related but not an identical plasmid. None of the three clones expressed protease activity in assays using either protein or chromogenic substrates. The lack of activity was not due to inhibition by *E. coli* products since sonicates of XL-1 Blue have no effect on the BApNA activity of purified ArgI from *P. gingivalis*.

The induction of β-galactosidase by the addition of 1 mM IPTG to mid-logarithmic-phase JM2 broth cultures for an hour did not alter the level of immunoreactive recombinant protein expression. Furthermore, subsequent DNA sequencing of the recombinant plasmid verified that the direction of transcription of the cloned *P. gingivalis* DNA was opposite that of the *lac* promoter of the vector. Thus, the gene is most likely expressed from its own promoter or promoter-like sequence that is efficiently recognized by the *E. coli* transcription-translation system. Examination of the recombinant proteins by Western blotting with antiserum Rb3158 showed that JM2 whole-cell proteins gave many positive bands, the largest molecular size of which was ~160 kDa (Fig. 2 and 7). A recombinant protein of this size was also detectable via Coomassie brilliant blue R staining of JM2 cellular proteins on SDS-PAGE gels (data not shown); this detected protein is indicative of a strong and efficient *P. gingivalis* promoter.

Automated N-terminal sequence analysis of the maximum-size recombinant protein of JM2 failed to generate any sequence. However, the sequence MNEIL. . . was obtained from

an M_r-120,000 immunoreactive protein band. The sequence NEIL. . . was also obtained from a cyanogen bromide digestion product of ArgI and was later found in the deduced amino acid sequence of the ArgI gene (Fig. 3). The numerous, immunoreactive, lower-molecular-weight proteins produced by JM2 may have been generated by extensive proteolysis of the large, >160-kDa protein or as a result of either aborted synthesis or the presence of alternative transcriptional-translational start sites along the pJM2 insert or its corresponding mRNA. These possibilities are discussed in more detail below.

Since the maximum molecular mass of the recombinant protein suggested that almost the entire pJM2 insert must represent a single open reading frame (ORF), sequence analysis was performed on the whole insert.

DNA sequencing and analysis of gene encoding ArgI. The nucleotide sequence of pJM2 was obtained by alignments of a minimum of two independent sequences, which, in some cases, were obtained by primer walking on both strands of a particular region. Difficult areas required alignments of a maximum of five separate DNA sequences.

Analysis of the 5,205-bp DNA sequence (Fig. 3) revealed an incomplete ORF of 4,577 bp starting at nucleotide (nt) 628, with a coding potential for 1,526 amino acids and a calculated M_r of 166,574, which is in close agreement with the maximum size of the recombinant protein. A putative ribosome binding site, 5′-AAGGAGGC-3′, which shares 87.5% identity with the *E. coli* consensus sequence (39), is shown at nt 322 to 330, 312 bp upstream of the putative methionine start codon at nt 628. The complete sequence has an overall G-C base composition of 46.4 mol%. This is in agreement with previously reported values from *Porphyromonas* DNA melting studies of 46 to 48 mol% (38). However, the putative noncoding region, nt 1 to 627, has a relatively low G-C content (41.0 mol%) compared with that of the remainder of the insert (47.2%), analogous to the base usage in *E. coli* and *Salmonella typhimurium* where the G-C composition of the coding regions reflects the overall G-C content of the organism (33). Imperfect nucleotide repeat sequences (85.8% identity) are found at nt 3220 to 3381 and nt 4591 to 4752, and within these is a 27-bp sequence with 100% identity beginning at nt 3334 and nt 4705, respectively.

Experimentally generated peptide sequences from ArgI of *P. gingivalis* accurately mapped within the deduced sequence and thereby confirmed the identity of the cloned and expressed gene. The N terminus of the α component of ArgI maps to amino acid position Y-228. Two other internal protein sequences, derived via N-terminal sequence analysis of cyanogen bromide digests of purified, monomeric *P. gingivalis* ArgIA (CNBr1 & 2) are located at N-522 to T-536 and L-560 to V-573. The N terminus of the β component of ArgI is also located within the deduced amino acid sequence of pJM2 at amino acid position S-720, and the N-terminal sequence of an ~14-kDa peptide fragment derived from ArgI autodigestion was also found immediately following an arginine residue at A-1139 to T-1158. On the basis of a molecular weight of the α component of approximately 52,000, we have introduced a putative processing site at R-719 (i.e., immediately preceding the N terminus of the β component). This corresponds to a calculated M_r of 53,900 for the α component, which is therefore flanked by two processing sites, R-227 to Y-228 at the N terminus and R-719 to S-720 at the C terminus. Since both α and β components migrate to the same position on SDS-PAGE gels, we introduced a further putative processing site at R-1262 to F-1263 to give a calculated M_r of 58,100 for the β component.

The α component is preceded by a polypeptide with a deduced M_r of 25,600. A Kyte and Doolittle hydropathy plot of the deduced protein shows a hydrophobic N terminus which has a low surface probability and exhibits features similar to those of signal peptides of *E. coli* (35). The N-terminal sequence of 22 amino acids has two charged lysines in the first 6 residues and is followed by a hydrophobic core of 16 residues with a potential signal peptidase site at A-23 to Q-24 (Fig. 3), which would satisfy the criteria of von Heijne (42).

Carboxy-terminal to the β component on the initial translation product is a region that we have termed the γ domain (Fig. 4 and 5). The γ domain on the ORF of pJM2 was incomplete and comprised 264 amino acids (M_r of 29,046). Extension cloning of the pJM2 insert, described in a later section, demonstrated that the γ region corresponds to a protein with a size of approximately 42 kDa.

At the primary sequence level, five peptide repeats, mostly in the β domain, containing 8 to 54 amino acids labelled c1/c2, d1/d2, e1/e2, f1/f2, and g1/g2 with identities of 100, 90.7, 80, 87.5, and 100%, respectively, are found in the deduced sequence; the organization of these repeat sequences and some of the amino acid signatures within them resemble bacterial proteins that bind to mammalian extracellular matrix proteins (44).

On the basis of the size and organization of the deduced polypeptide from pJM2, we have given the name PrpRI (protease polyprotein ArgI) to the unprocessed product of the gene encoding ArgI (*prpR1*).

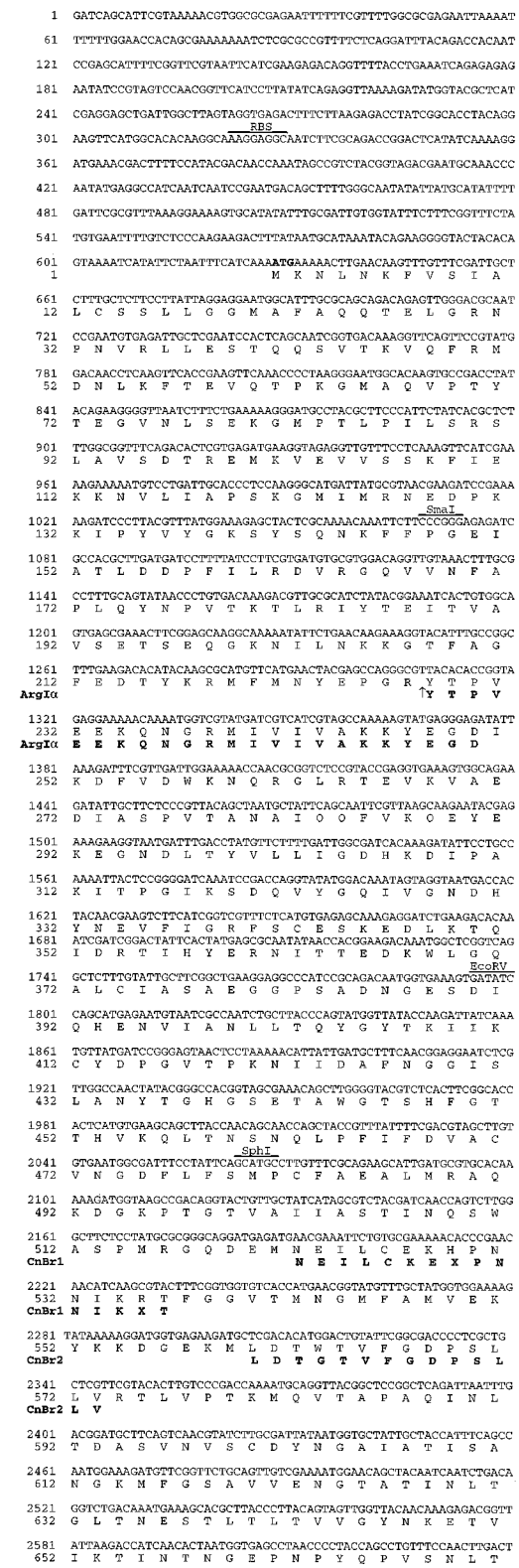
Database searches. A search of the protein and nucleic acid databases (National Center for Biotechnology Information, Bethesda, Md.) with the sequences in Fig. 3 revealed several sequences from *P. gingivalis* with significant identities to PrpRI. *P. gingivalis* W50 PrpRI is almost identical (97.5% identity) to the product of the *arg-gingipain-1* protease gene (accession no. U15282 [30]) from *P. gingivalis* H66. A notable difference between the two gene products is the presence of a hexamer of Pro-Asn $(PN)_6$ in PrpRI, which is found as a pentamer in Arg-gingipain-1. Other subtle differences are mostly conservative amino acid substitutions. In contrast, the identity of *prtH* (accession no. L27483 [14]) encoding a complement-degrading enzyme from *P. gingivalis* W83 is confined to the 3′ end of *prpR1* (nt 3386 to 5205). At K-1082 to D-1495, PrpRI is 94% identical to PrtH (K-32 to D-446).

In addition, at three positions (K-246 to N-259, G-439 to L-463, and T-671 to L-718) the α component shares 60 to 64% similarity to the p60 protein of *Listeria seeligeri* and *Listeria monocytogenes*, a major extracellular protein of these pathogens with a suggested role in the invasion of phagocytic cells (18). The deduced sequence of the β component of ArgI shows similarity at two positions (D-748 to V-767 [55%] and N-947 to G-1006 [56%]) to the circumsporozoite protein of parasitic *Plasmodium falciparum* (8) and to a hemagglutinin homolog of *Mycoplasma gallisepticum* (accession no. M83178 [23]) at position N-947 to G-963 (76%).

An additional feature of PrpRI which was revealed by examination of the Prosite dictionary of the protein motif database is the presence of an adenylation signature, EXKX(N) GXRA (41), at E-1131 to A-1140 (Fig. 3). This sequence is typical of ATP-dependent ligases of eukaryotes, archaebacteria, viruses, and bacteriophages, with the lysine residue being the contact site for protein-adenylate intermediate formation.

Confirmation of genomic organization of *prpR1* by Southern hybridization. In order to confirm the genomic organization of *prpR1* and to determine its distribution in other *P. gingivalis* strains, Southern hybridizations were performed with probes specific to the α and β regions.

The nucleotide sequence of *prpR1* (Fig. 3) indicated that a 3.2-kb *Sma*I fragment of *P. gingivalis* W50 genomic DNA en-



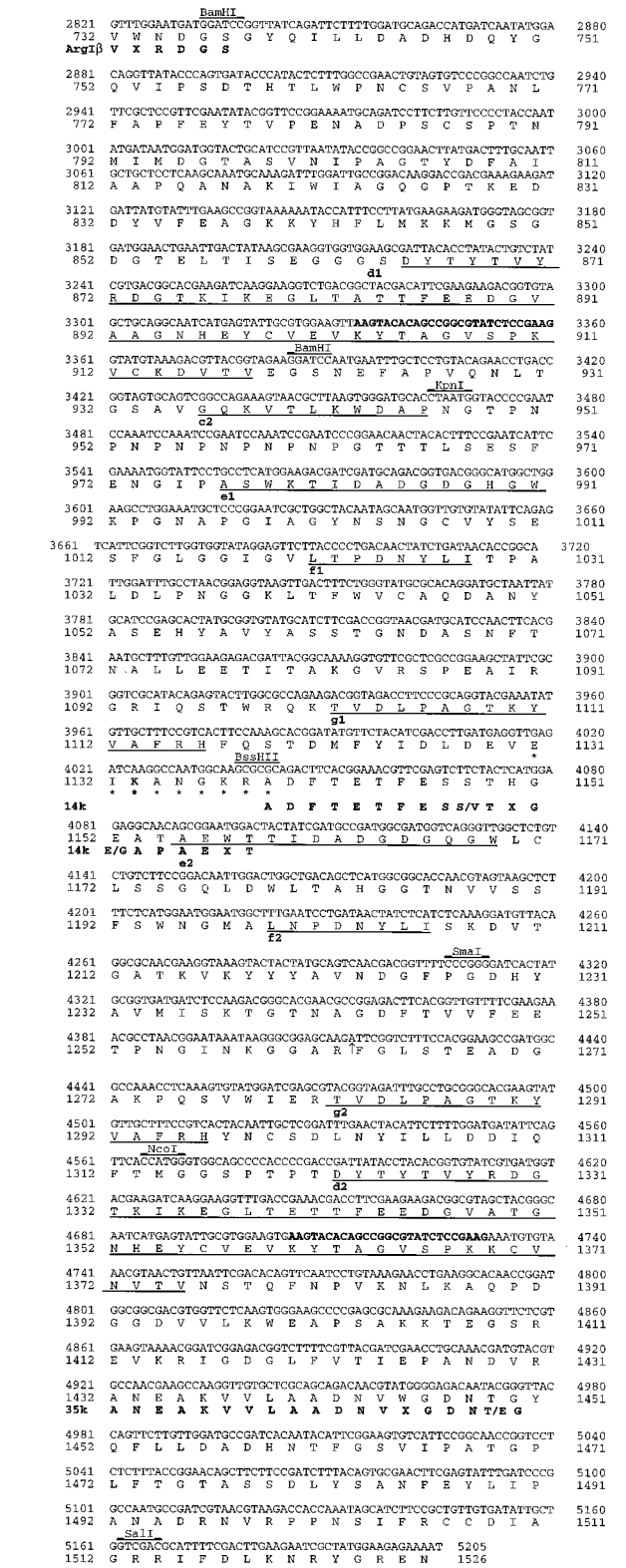


FIG. 3. Nucleotide sequence of *prpR1* of *P. gingivalis* W50. The putative ribosome binding site (RBS) is shown above the sequence, and the methionine ATG start codon is in bold. Restriction enzyme sites are also shown above the sequence. Experimentally determined protein sequences are shown under the deduced protein sequence and labelled as follows (in bold): ArgIα and ArgIβ (N termini of the two polypeptide chains), CNBr1 and CNBr2 (cyanogen bromide-generated peptides of ArgIA), and 14k (an internal peptide from ArgI with a molecular mass 14 kDa). A unique 27-bp sequence which is repeated at two positions is in bold. Peptide repeats c1/c2, d1/d2, e1/e2, f1/f2, and g1/g2 are underlined. Amino acid residues corresponding to an adenylation motif are indicated with asterisks, and the presumed catalytic lysine is in boldface. Arrows mark the positions of the putative processing sites of PrpRI. The sequence will appear in the EMBL, GenBank, and DDBJ nucleotide sequence databases under accession no. X82680.

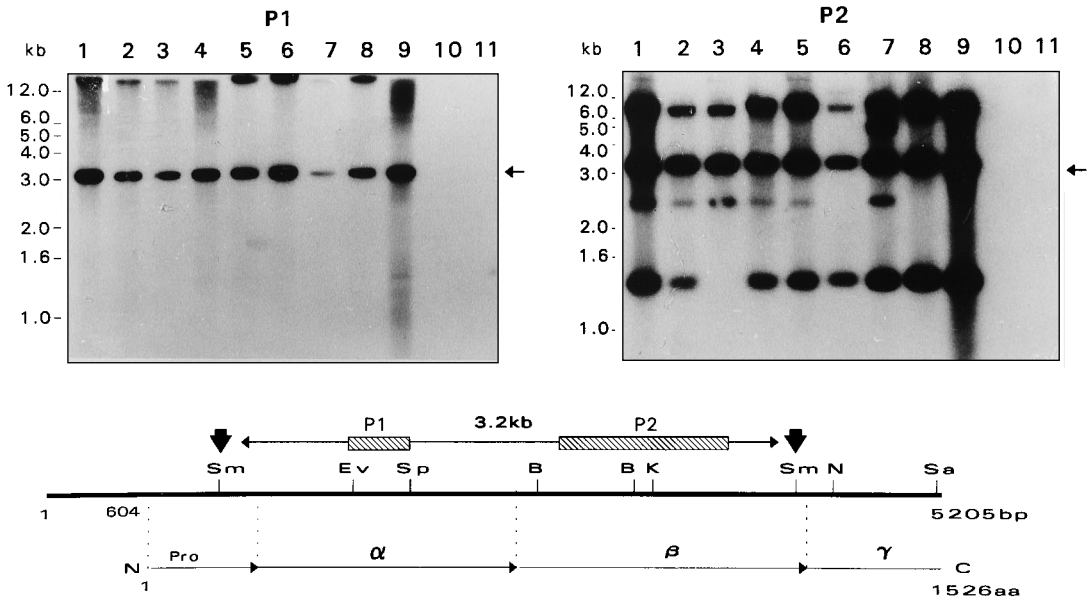


FIG. 4. Southern blot analysis of *prpR1* in *P. gingivalis* strains. DNA was completely restricted with *Sma*I and probed with DNA contained within the coding regions for either the α component (P1-*Eco*RV-*Sph*I) or β component (P2-pJM11) of ArgI. Lanes: 1, *P. gingivalis* W50; 2, W50Be1; 3, W50Br1; 4, W83; 5, LB13 D-3; 6, A7436; 7, 381; 8, 33277; 9, 11834; 10, *P. asaccharolytica* ATCC 25260; 11, *E. coli* XL-1 Blue. The horizontal arrows mark the position of a 3.2-kb *Sma*I fragment which is common to all *P. gingivalis* strains. In the schematic diagram below the blots, the pJM2 insert is shown (bold line), with selected restriction enzyme sites above and arrowheads indicating the positions of the *Sma*I sites on the insert. The probes used in the Southern hybridizations, P1 and P2, are shown as hatched boxes above the restriction map. The direction of transcription and domain organization of the translation product are shown on the thin line (N→C). Sm, *Sma*I; Ev, *Eco*RV; Sp, *Sph*I; B, *Bam*HI; K, *Kpn*I; N, *Nco*I; Sa, *Sal*I; aa, amino acids.

codes both the α and β components of ArgI. *Sma*I-digested DNA from nine different *P. gingivalis* strains (Fig. 4) was probed with two distinct DNA fragments derived from pJM2. A probe to the α domain (an *Eco*RV-*Sph*I fragment) hybridized to the expected 3.2-kb *Sma*I fragment from all *P. gingivalis* strains (Fig. 4). In addition however, a second hybridizing band, greater than 12 kb, was also detected in all strains. The presence of this homologous locus to the α region of *prpR1* was confirmed by performing double digestion of the genomic DNA of *P. gingivalis* W50 with combinations of restriction endonucleases which would generate a hybridizing fragment of known size from *prpR1* (Fig. 5). In all cases an additional locus was detected. Therefore, at least a portion of the sequence in the *Eco*RV-*Sph*I fragment of the α region is present at two separate loci on the *P. gingivalis* genome.

Complex hybridization profiles were obtained by using a probe to the β region (p2; pJM11 insert, nt 2977 to 4019 [Fig. 3 and 4]). In addition to the expected 3.2-kb *Sma*I fragment, three bands varying in size from 1.5 to 6 kb were detected in most strains. These data suggest that the β region of PrpRI may represent a domain, possibly involved in adhesion, in at least three *P. gingivalis* gene products.

Extension cloning of the complete *prpR1* gene. Sequence analysis (Fig. 3) showed that the ORF on pJM2 was incomplete, and the following strategy was used to extend the 3′ end. *P. gingivalis* W50 genomic DNA was restricted to completion

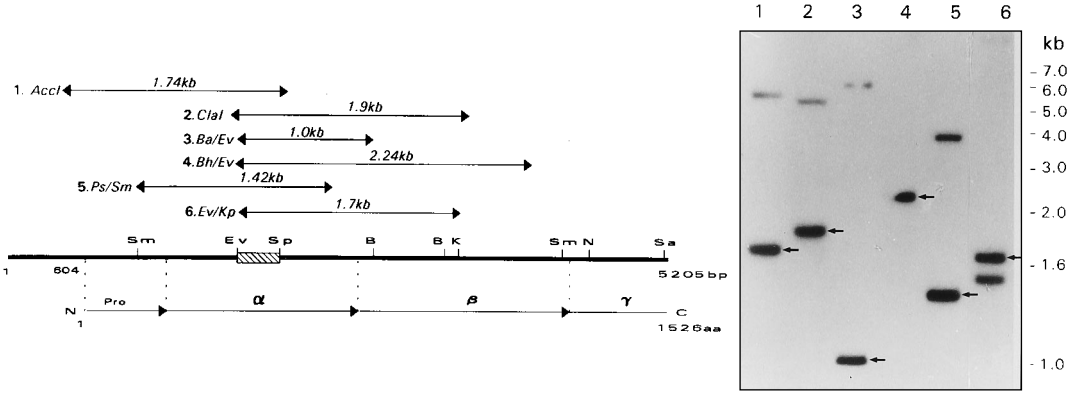


FIG. 5. Southern blot analysis of *P. gingivalis* W50 DNA. *P. gingivalis* DNA was restricted with various endonucleases, either singly or in combination, and then probed with the P1 DNA fragment described for Fig. 4. Thin lines (numbered 1 to 6) above the restriction map denote the expected fragments and are indicated by arrows in the corresponding lanes of the blot on the right. The numbers on the lines indicate the fragment sizes. Lanes: 1, *Acc*I; 2, *Cla*I; 3, *Bam*HI and *Eco*RV (Ba/Ev); 4, *Bss*HII and *Eco*RV (Bh/Ev); 5, *Pst*I and *Sma*I (Ps/Sm); 6, *Eco*RV and *Kpn*I (Ev/Kp). aa, amino acids.

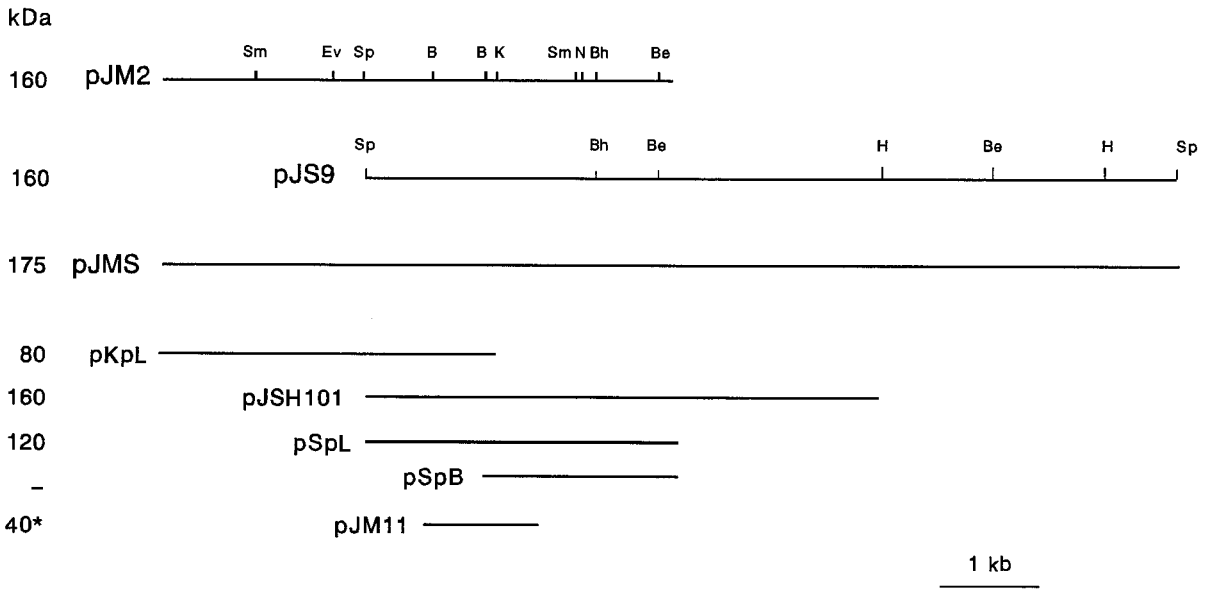


FIG. 6. Physical map and extension of pJM2. The original pJM2 insert with unique restriction enzyme sites is shown at the top. Southern hybridization analysis of total *Sph*I-digested *P. gingivalis* W50 DNA and subsequent targetted cloning of the band with the labelled pJM11 insert as a probe gave three distinct clones. Restriction mapping, Southern hybridization profiles, and DNA sequencing of the ends of the inserts of the plasmids confirmed that only one plasmid, pJS9, contained the extension sequences to pJM2. Subclones of either pJMS or pJM2 were constructed as described in the text, and their relative positions are shown. All inserts, with the exception of that of pJM11, are in the same orientation relative to that of pUC18, and the *Eco*RI site of the residual MCS is to the right of each insert. The sizes of the expressed proteins in the clones are indicated on the left. An asterisk indicates protein expression from the *lac* promoter of pUC18. B, *Bam*HI; Be, *Bst*EII; Bh, *Bss*HII; E, *Eco*RV; H, *Hin*dIII; K, *Kpn*I; N, *Nco*I; Sa, *Sal*I; Sm, *Sma*I; Sp, *Sph*I.

by using *Sph*I and probed on Southern blots with labelled pJM11 insert. This generated hybridizing DNA fragments in the size range of 8 to 12 kb (data not shown). These fragments were cloned into pUC18, and three clones containing inserts with approximate sizes of 10, 11, and 8.3 kb were obtained. Restriction endonuclease mapping and signature Southern hybridization profiles indicated that the plasmids represented different genomic loci. DNA sequencing with M13 forward and reverse primers demonstrated that the insert of one plasmid, pJS9, contained terminal sequences which overlapped with the pJM2 insert. Thus, pJS9 contained the 3′ extension of the *prpR1* gene.

A complete *prpR1* gene was constructed by employing the unique *Bss*HII site in both pJM2 and pJS9 and the *Eco*RI site in the residual MCS (on the righthand side of the inserts [Fig. 6]) of the recombinant plasmids. Since both inserts are in the same relative orientation, the 1.0-kb *Bss*HII-*Eco*RI fragment in pJM2 was replaced with a 6.2-kb *Bss*HII-*Eco*RI fragment from pJS9 to construct pJMS. Recombinant protein expression from JMS, JM2, and other related clones was then compared by Western blotting (Fig. 7).

Recombinant protein expression and demonstration of an internal transcription-translation site. With the exception of pJM11, all the inserts shown in Fig. 6 are in the same orientation in pUC18 with the vector *lac* promoter positioned at the 3′ end of pJM2 (Fig. 3). Hence, recombinant protein production by these clones is initiated by sequences within the cloned *P. gingivalis* DNA. The largest recombinant protein immunoreactive with Rb3158 antiserum produced by JMS was shown by Western blot analysis to be in the range of 170 to 180 kDa compared with the 160-kDa-maximum-size product of JM2 (Fig. 7, lanes 1 and 3). These data demonstrate that the extension cloning procedure described above, which extended the 3′ end of the pJM2 insert by approximately 5 kb, generated a *prpR1* ORF in pJMS which was less than 1 kb larger than the original pJM2 ORF.

As in JM2, no protease activity was demonstrated in JMS and a large number of lower-molecular-weight immunoreactive proteins were present. Treatment of the transformants with protease inhibitors prior to sonication and electrophoresis failed to reduce the number of lower-molecular-weight forms. This suggested that at least some of these proteins may have arisen through internal initiation of transcription-translation

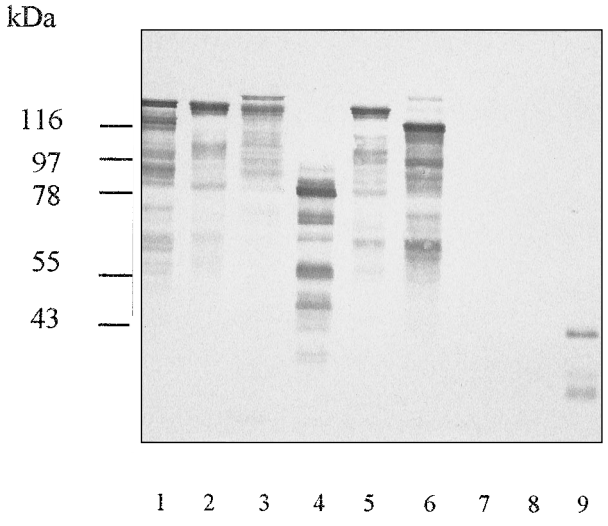


FIG. 7. Western blot analysis of JMS derivatives. Proteins from whole-cell pellets of clones indicated in Fig. 6 were resolved by SDS-PAGE, blotted onto nitrocellulose membranes, and probed with Rb3158 antiserum. Note that although expression of recombinant protein in JM11 is under vector control, there is leaky synthesis without IPTG. Lanes: 1, JM2; 2, JS9; 3, JMS; 4, KpL; 5, JSH101; 6, SpL; 7, SpB; 8, UC18; 9, JM11. The molecular sizes of the marker proteins are indicated on the right.

within the cloned *P. gingivalis* DNA or corresponding mRNA. This possibility was examined by subcloning.

Truncation of the 3′ end of the pJM2 insert by 1.9 kb via *Kpn*I digestion generated pKpL. As expected, KpL produced a truncated maximum-size product with a size of ~80 kDa (Fig. 7, lane 4). To determine whether protein expression could be initiated from elsewhere on the protease gene in *E. coli*, pSpL was generated by deletion of a 2.2-kb *Sph*I(MCS)-*Sph*I(insert) fragment from the 5′ end of the pJM2 sequence. The deleted fragment thus included the controlling elements of pJM2, pJMS, and pKpL expression. Similarly, JS9 and JSH101, the plasmids of which share the same 5′ end as that of pSpL but contain longer insert sizes, were also examined. The clones containing these constructs expressed recombinant proteins. In SpL the maximum size of the major immunoreactive product was approximately 120 kDa (Fig. 7, lane 6), while JS9 and JSH101 produced identical maximum-size proteins with a size of approximately 160 kDa (Fig. 7, lanes 2 and 5) and identical SDS-PAGE banding profiles. (The pJSH101 insert in the opposite orientation also expressed a protein with the same size and the same banding pattern [data not shown]). Removal of a 1.3-kb *Sph*I-*Bam*HI fragment from the 5′ end of pSpL generated pSpB, which did not express any immunoreactive recombinant protein (Fig. 7, lane 7) and thus localized the internal initiation site to between nt 2050 and 3387 (Fig. 3).

Supportive evidence for the presence of an internal transcription-translation initiation site was obtained by examination of the coding probability of the nucleotide sequence of pJM2 with the TESTCODE algorithm of Fickett (13). This sequence shows sharp periodic dips at the extremes of the ORF as expected and also within the ORF in the region of nt 2000 (data not shown), which is consistent with the subcloning expression studies.

These data demonstrate the presence of at least one alternative *E. coli*-compatible transcription-translation initiation site within *prpR1* which may contribute to the number of recombinant proteins with different molecular weights observed on SDS-PAGE gels and Western blots. However, the failure to express any active recombinant protease and the large number of lower-molecular-weight recombinant protein forms in *E. coli* also indicate that *prpR1* is inappropriately processed in this host.

Determination of precursor production in *P. gingivalis*. The finding that the coding regions for the α and β components of ArgI are contiguous on the same gene led us to attempt to demonstrate the intact precursor of ArgI in *P. gingivalis*. Initially, cultures (50 ml) in mid-logarithmic, early stationary, and late stationary phases were rapidly chilled on ice. The cells were harvested by centrifugation at 14,000 × *g* for 5 min at 4°C and quickly resuspended in ice-cold 10% formic acid to arrest any further proteolysis. Following sonication, the soluble and membrane-associated cellular proteins were analyzed by SDS-PAGE and Western blotting with Rb3158 antiserum. In all cases, regardless of the time of sampling, the maximum-size immunoreactive protein which was detected corresponded to the β component of the ArgI dimer, suggesting that the processing of the precursor may be closely coupled to its synthesis (data not shown).

Since the putative processing sites in the deduced sequence of the initial translation product (R-227 to Y-228, R-719 to S-720, and R-1262 to F-1263) are consistent with an autolytic processing mechanism of PrpRI, we next examined the effect of the addition of leupeptin, an irreversible inhibitor of ArgI, on PrpRI processing in mid-logarithmic-phase *P. gingivalis* cultures. The results are shown in Fig. 8. The addition of leupeptin (final concentration, 0.7 mM) in early logarithmic phase

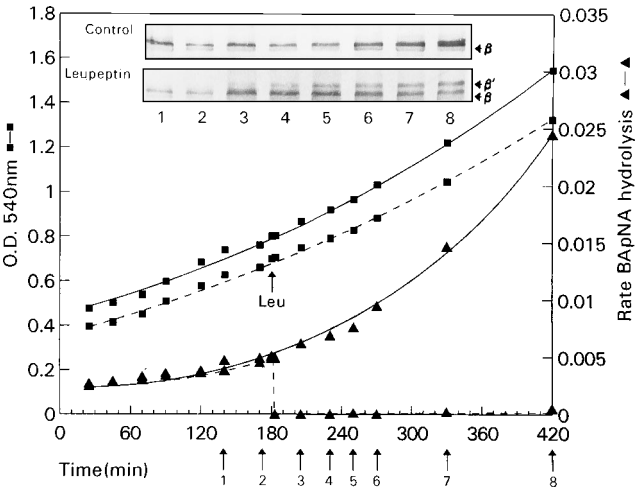


FIG. 8. PrpRI expression during growth of *P. gingivalis* W50 and effect of leupeptin addition. The solid lines represent the control culture (no leupeptin), and the dotted lines indicate the test culture (leupeptin added at 180 min). (Insert) Appearance of ArgI β component on Western blots of formic acid-treated whole cells from control and leupeptin-treated cultures at the sampling points (numbered 1 to 8 on the lower axis). The optical density (O.D.) of the cultures (■) was measured at 540 nm, and the total Arg-X enzyme activity (△) was expressed as the rate of BapNA hydrolysis per 20 μl of whole-cell cultures.

abolished all extracellular activity against BApNA for the duration of the experiment but had no effect on the mean generation time of the cells compared with that of the control. Western immunoblots of the formic acid-treated cells (insert in Fig. 8) developed with Rb3158 antiserum failed to demonstrate the presence of an intact PrpRI precursor but did show that following the addition of leupeptin there was a time-dependent accumulation of a novel immunoreactive protein which was approximately 5 kDa greater in mass than the β component (labelled β′ in Fig. 8). These data suggest that the initial processing of the PrpRI precursor may occur intracellularly and hence be inaccessible to the exogenously added inhibitor. The final stages of processing may, however, occur at the cell surface and include the proteolytic removal of a small ~5-kDa fragment from the β component.

DISCUSSION

Characterization of *prpR1*. This report describes the cloning, sequence analysis, and expression of the protease antigen of *P. gingivalis*, PrpRI, which we had previously identified by examination of the specific immune responses of individuals with a history of destructive periodontal disease. In the absence of an active recombinant enzyme, confirmation of the identity of the gene was based on the following criteria. First, the experimentally determined N-terminal protein sequences of the purified monomeric and dimeric forms of the enzyme from *P. gingivalis* as well as the internal protein sequences of the monomeric form were found in the deduced sequence of the *prpR1* ORF. Secondly, the recombinant protein expressed by the cloned DNA was shown to contain protein sequences which were present in both the purified *P. gingivalis* enzymes and the deduced translation product of the gene. In addition, the recombinant protein carried epitopes which were recognized by IgG antibody in periodontal patients' sera, as would be predicted for this discriminatory protease antigen (data not shown).

The complete gene has the coding potential for an approximately 175-kDa protein with a domain structure which resembles that of some other large extracellular proteases from other

microbial species. The initial translation product is a large precursor, which we have divided into four contiguous regions (pro, α, β, and γ [Fig. 4 and 5]). The α component of PrpRI is preceded by a polypeptide with a deduced M_r of 25,600 in the pro region. Bacterial proteases are commonly synthesized as inactive precursors with an additional polypeptide, the propeptide, which is absent from the secreted protein. Activation of the precursor is achieved either via autolysis or through the action of a second enzyme as part of the secretion process (43). The propeptide may have an essential role in guiding the folding of the protein into the correct conformation and also in the inhibition of the mature enzyme by tight binding at the active site, as has been established for subtilisin E of *Bacillus subtilis* (28). The presence of a signal peptide-like sequence at the N terminus of PrpRI implies that the secretion process utilizes components of the general secretory pathway to guide the protein through the cytoplasmic membrane.

Downstream from the propeptide on the initial precursor is the α region, which is flanked by two putative processing sites: R-Y at the amino terminus to give the sequence YTPV. . ., the N-terminal sequence of the monomeric ArgIα, and R-S at the carboxy terminus, which simultaneously generates the N-terminal sequence of the β component of the ArgI dimer, SGQE. . .. Both α and β monomers appear to be identical in size on reducing SDS-PAGE gels. This suggests that there is another putative processing site at R-1262 to F-1263 which maintains the peptide bond specificity of processing cleavages and would generate a β component with a size of ~58 kDa.

Relationship of PrpRI to adhesins. The β component of the ArgI dimer contains a number of features typical of adhesins and binding proteins. For example, the regions of PrpRI with similarity to the circumsporozoite protein of *Plasmodium falciparum* sporozoites, which is involved in attachment of the parasite to hepatocytes (8), and to the hemagglutinin of *Mycoplasma gallisepticum* (23) are all within the β component. In addition, the β region contains the dipeptide repeat $(PN)_6$, and polypeptides which contain repetitive short proline-rich sequences such as $(XP)_n$ or $(XPY)_n$ are frequently involved in binding processes (for a review, see reference 45). Thus, the $(AP)_n$ motif in myosin light-chain kinase binds to actin. A repeating XP sequence in TonB of *Serratia marcescens* (15) is involved in the interaction of this protein with the outer membrane ferrichrome receptor protein, FhuA (2), and a proline containing tripeptide repeat sequence in the M protein of *Streptococcus equi* binds to the peptidoglycan layer of the cell wall. Furthermore, Lantz et al. (20) have suggested that a thiol-activated cell surface protease of *P. gingivalis* W12 could be associated with the cellular binding of fibronectin. It may be significant, therefore, that the β component of PrpRI contains hydrophilic repeat sequences characteristic of bacterial proteins which bind to the mammalian extracellular matrix.

These findings support earlier observations from a number of laboratories that the hemagglutination properties of *P. gingivalis* are closely related to proteolytic enzyme function (for example, see reference 31). In physiological terms, the molecular association of an adhesin-like β component with the catalytically active α component in the ArgI heterodimer may represent a device by which the bacterium targets the action of the protease component to either host macromolecules or cells. This would be analogous to bacterial toxins with intracellular targets in which the catalytic domain is activated and released from other domains on the molecule with translocation or receptor-binding functions, following interaction with the appropriate cell (16).

Processing aspects of PrpRI. This polyprotein configuration of PrpRI has parallels with the IgA proteases of pathogenic species of *Neisseria* and *Haemophilus* and the serine protease of *Serratia marcescens* which are secreted via a signal peptide-dependent autotransporter pathway (36). In these cases, the extracellular enzymes are synthesized as large precursors in which the domain that will become the mature protease is flanked by a typical secretion signal at the amino terminus and a large helper domain at the carboxy terminus. The C-terminal peptide of the polyprotein (the helper peptide) contains the determinants necessary to direct the protein across the outer membrane where it remains anchored via the same C-terminal peptide. Extracellular enzyme release is then achieved by autocatalytic proteolysis, leaving the helper peptide embedded in the outer membrane (21).

Although we were unable to demonstrate the presence of the intact ArgI precursor in *P. gingivalis* during mid-logarithmic-phase growth, the formation of a slightly higher-molecular-weight form of the β component was revealed in these cells when leupeptin was added to the extracellular medium. Hence, while initial processing may be accomplished intracellularly, probably closely coupled to synthesis of the full-length precursor, the final stages may occur at the surface. However, an additional aspect to the potential mechanism of formation of the final products of the *prpR1* gene was observed. During the recombinant protein expression studies of *E. coli*, an internal transcription-translation start site on *prpR1* was detected within the coding region. If this site is also operative in *P. gingivalis*, it may permit independent expression of different regions of the polyprotein and provide an alternative explanation for the failure to detect a large precursor in *P. gingivalis*. The presence of an internal transcription-translation site has recently been demonstrated in *Peptostreptococcus* sp. strain 312; the gene for the immunoglobulin-binding protein L contains three sites for internal initiation of translation, resulting in three active proteins (27).

Nucleotide and protein database searches demonstrated that the *prpR1* is homologous to the *arg-gingipain-1* protease gene of *P. gingivalis* H66. However, the biochemical characterization of the products of this gene in culture supernatants of strain H66 (3, 31) has led to significant differences in the interpretation of how the initial translation product is processed (30). Our data and those of Chen et al. (3) are in complete agreement that an ~50-kDa monomeric form is produced (i.e., ArgIA and gingipain). The discrepancy lies in the nature of the nonmonomeric form. While our analyses of *P. gingivalis* W50 culture supernatants revealed a heterodimer (ArgI) composed of the α and β components of the PrpRI, the studies of Pike et al. (31) suggest a more complex subunit structure in which the ~50-kDa catalytic subunit is associated with four other polypeptides (with M_rs of 44,000, 27,000, 17,000, and 15,000) in nonequimolar amounts, derived from the C terminus of the initial translation product. The resulting complex, high-molecular-weight gingipain, has a final molecular weight of 95,000. The relationship of ArgIB to the glycosylated forms of gingipain described by Potempa et al. (32) is still uncertain. Despite these inconsistencies, we are in general agreement that these three forms represent the majority of the total extracellular Arg-X protease activity of these two strains.

Homologs of *prpR1*. The sequences of several other protease genes of *P. gingivalis* demonstrate significant regional homology to *prpR1* (14, 17, 29). The apparent multiplicity of related genes was confirmed in the present work during the examination of the genomic organization of *prpR1* by Southern hybridization, which identified several homologous loci to different regions of *prpR1* elsewhere on the *P. gingivalis* genome. Significantly, these included a second locus which hybridized to a

DNA fragment from the coding region for the catalytic component of ArgI.

Hybridization studies using a probe to the β region (i.e., adhesin domain) of the *prpR1* demonstrated the presence of at least two other homologous loci. It is therefore possible that non-protease-related functions of the polyprotein are also shared by other gene products of *P. gingivalis*. Indeed, a clone (JM7) containing one of these loci was obtained during the initial immunoscreening with ArgI antiserum, and sequence analysis of this clone has revealed a region of 1,419 bp with 95% identity within the coding region for the C terminus of the polyprotein. Such loci may have arisen through homologous recombination events involving genome rearrangements partly facilitated by the presence of the short tandem repeats (12) found within *prpR1*.

In summary, the protease antigen of *P. gingivalis* is synthesized as part of a large polyprotein precursor which is processed to a number of different forms of the extracellular product. Different regions of the polyprotein are likely to form domains of other gene products of this organism, and these domains may be functionally and antigenically related to ArgI. The development of appropriate expression systems to produce active recombinant ArgI protease, the determination of the function of antigenically related gene products, and the assessment of the role of the host's immune response in the control of these activities will form the next phase of these investigations.

ACKNOWLEDGMENTS

We thank T. O'Keeffe and P. A. Pemberton for assistance with the DNA and protein sequencing, respectively, and K. A. Young for assistance with the microbiology.

This study was supported by the Medical Research Council (grant no. PG 9318173).

REFERENCES

1. **Aduse-Opoku, J., J. Muir, J. M. Slaney, and M. A. Curtis.** 1994. Cloning and characterisation of a protease gene of *Porphyromonas gingivalis* W50. J. Dent. Res. **73:**247.
2. **Brewer, S., M. Tolley, I. P. Trayer, G. C. Barr, C. J. Dorman, K. Hannavy, C. F. Higgins, J. S. Evans, B. A. Levine, and M. R. Wormald.** 1990. Structure and function of X-prodipeptide repeats in the TonB proteins of *Salmonella typhimurium* and *Escherichia coli*. J. Mol. Biol. **216:**883–895.
3. **Chen, X., J. Potempa, A. Polanowski, M. Wikstrom, and J. Travis.** 1992. Purification and characterisation of a 50-kDa cysteine proteinase (gingipain) from *Porphyromonas gingivalis*. J. Biol. Chem. **267:**18896–18901.
4. **Clark-Curtiss, J. E., and R. Curtiss III.** 1983. Analysis of recombinant DNA using *Escherichia coli* minicells. Methods Enzymol. **101:**347–362.
5. **Curtis, M. A., I. R. Gillett, G. S. Griffiths, M. F. J. Maiden, J. A. C. Sterne, D. T. Wilson, J. M. A. Wilton, and N. W. Johnson.** 1989. Detection of high-risk groups and individuals for periodontal diseases: laboratory markers from analysis of gingival crevicular fluid. J. Clin. Periodontol. **16:**1–11.
6. **Curtis, M. A., M. Ramakrishnan, and J. M. Slaney.** 1993. Characterisation of the trypsin-like enzymes of *Porphyromonas gingivalis* W83 using a radiolabelled active-site-directed inhibitor. J. Gen. Microbiol. **139:**949–955.
7. **Curtis, M. A., J. M. Slaney, R. J. Carman, and N. W. Johnson.** 1991. Identification of the major surface protein antigens of *Porphyromonas gingivalis*. Oral Microbiol. Immunol. **6:**321–326.
8. **Dame, J. B., J. L. William, T. F. McCutchan, J. L. Weber, R. A. Wirtz, W. T. Hockmeyer, W. L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, G. S. Sanders, E. P. Reddy, C. L. Diggs, and L. H. Miller.** 1984. Structure of the gene encoding the immunodominant surface-antigen on the sporozoite of the human malarial parasite *Plasmodium falciparum*. Science **225:**593–599.
9. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
10. **Duncan, A. J., R. J. Carman, F. H. Harper, G. S. Griffiths, and M. A. Curtis.** 1992. *Porphyromonas gingivalis*: presence of a species-specific antigen which is discriminatory in chronic inflammatory adult periodontal disease. Microb. Ecol. Health Dis. **5:**15–20.
11. **Ebersole, J. L., M. A. Taubman, D. J. Smith, and D. E. Frey.** 1986. Human immune responses to microorganisms: patterns of systemic antibody levels to *Bacteroides* spp. Infect. Immun. **51:**507–513.
12. **Ehrlich, S. D., H. Bierne, E. d'Alençon, D. Villet, P. Noirot, and B. Michel.** 1993. Mechanisms of illegitimate recombination. Gene **135:**161–166.
13. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10:**5303–5318.
14. **Fletcher, H. M., H. A. Schenkein, and F. L. Macrina.** 1994. Cloning and characterization of a new protease gene (*prtH*) from *Porphyromonas gingivalis*. Infect. Immun. **62:**4279–4286.
15. **Gaisser, S., and V. Braun.** 1991. The *tonB* gene of *Serratia marcescens*: sequence, activity and partial complementation of *Escherichia coli tonB* mutants. Mol. Microbiol. **5:**2777–2787.
16. **Gordon, V. M., and S. H. Leppla.** 1994. Proteolytic activation of bacterial toxins: role of bacterial and host cell proteases. Infect. Immun. **62:**333–340.
17. **Kirszbaum, L., C. Sotiropoulos, C. Jackson, S. Cleal, N. Slakeski, and E. Reynolds.** 1995. Complete nucleotide sequence of a gene *prtR* of Porphyromonas gingivalis W50 encoding a 132 kDa protein that contains an arginine-specific thiol endopeptidase domain and a haemagglutinin domain. Biochem. Biophys. Res. Commun. **207:**424–431.
18. **Kuhn, M., and W. Goebel.** 1989. Identification of an extracellular protein of *Listeria monocytogenes* possibly involved in the intracellular uptake by mammalian cells. Infect. Immun. **57:**55–61.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
20. **Lantz, M. S., R. D. Allen, L. W. Duck, J. L. Blume, L. M. Switaski, and M. Hook.** 1991. Identification of *Porphyromonas gingivalis* components that mediate its interactions with fibronectin. J. Bacteriol. **173:**4263–4270.
21. **Lory, S.** 1992. Determinants of extracellular protein secretion in gram-negative bacteria. J. Bacteriol. **174:**3423–3428.
22. **Maiden, M. J. F., R. J. Carman, M. A. Curtis, I. R. Gillett, G. S. Griffiths, J. A. C. Sterne, J. M. A. Wilton, and N. W. Johnson.** 1990. Detection of high-risk groups and individuals for periodontal diseases: laboratory markers based on the microbiological analysis of subgingival plaque. J. Clin. Periodontol. **17:**1–13.
23. **Markham, P. F., M. D. Glew, K. G. Whithear, and I. D. Walker.** 1993. Molecular cloning of a member of the gene family that encodes pMGA, a hemagglutinin of *Mycoplasma gallisepticum*. Infect. Immun. **61:**903–909.
24. **Marsh, P.** 1991. Do bacterial markers exist in subgingival plaque for predicting periodontal disease susceptibility?, p. 365–388. *In* N. W. Johnson (ed.), Risk markers for oral diseases: periodontal diseases, vol. 3. Cambridge University Press, Cambridge.
25. **Millar, D. J., E. E. Scott, J. M. Slaney, S. U., P. Benjamin, and M. A. Curtis.** 1993. Production and characterisation of monoclonal antibodies to the principal sonicate antigens of *Porphyromonas gingivalis* W50. FEMS Immunol. Med. Microbiol. **7:**211–222.
26. **Monaco, L., J. J. Murtagh, K. B. Newman, S.-C. Tsai, J. Moss, and M. Vaughan.** 1990. Selective amplification of an mRNA and related pseudogene for a human ADP-ribosylation factor, a guanine nucleotide-dependent protein activator of cholera toxin. Proc. Natl. Acad. Sci. USA **87:**2206–2210.
27. **Murphy, J. P., C. J. Duggleby, M. A. Atkinson, A. R. Trowern, T. Atkinson, and C. R. Goward.** 1994. The functional units of a peptostreptococcal protein L. Mol. Microbiol. **12:**911–920.
28. **Ohta, Y., H. Hojo, S. Aimoto, T. Kobayashi, X. Zhu, F. Jordan, and M. Inouye.** 1991. Pro-peptide as an intermolecular chaperone: renaturation of denatured subtilisin E with a synthetic pro-peptide. Mol. Microbiol. **5:**1507–1510.
29. **Okamoto, K., Y. Misumi, T. Kadowaki, M. Yoneda, K. Yamamoto, and Y. Ikehara.** 1995. Structural characterisation of argingipain, a novel arginine-specific cysteine proteinase as a major periodontal pathogenic factor from *Porphyromonas gingivalis*. Arch. Biochem. Biophys. **316:**917–925.
30. **Pavloff, N., J. Potempa, R. N. Pike, V. Prochazka, M. C. Kiefer, J. Travis, and P. J. Barr.** 1995. Molecular cloning and structural characterisation of the Arg-gingipain proteinase of *Porphyromonas gingivalis*. Biosynthesis as a proteinase-adhesin polyprotein. J. Biol. Chem. **270:**1007–1010.
31. **Pike, R., W. McGraw, J. Potempa, and J. Travis.** 1994. Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. Isolation, characterisation, and evidence for the existence of complexes with hemagglutinin. J. Biol. Chem. **269:**406–411.
32. **Potempa, J., R. Pike, and J. Travis.** 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.
33. **Raha, M., M. Kihara, I. Kawagishi, and R. M. Macnab.** 1993. Organisation of the *Escherichia coli* and *Salmonella typhimurium* chromosomes between flagellar regions IIIa and IIIb, including a large non-coding region. J. Gen. Microbiol. **139:**1401–1407.

33a. **Rangarajan, M., et al.** Unpublished data.

34. **Rangarajan, M., S. Smith, and M. A. Curtis.** 1994. Proteases specific for arg-X and lys-X peptide bonds from *Porphyromonas gingivalis* W50. J. Dent. Res. **73:**247.
35. **Rusch, S. L., and D. A. Kendall.** 1994. Transport of an export-defective protein by a highly hydrophobic signal peptide. J. Biol. Chem. **269:**1243–1248.
36. **Salmond, G. P. C., and P. J. Reeves.** 1993. Membrane traffic wardens and protein secretion in Gram-negative bacteria. Trends Biochem. Sci. **18:**7–12.

37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Shah, H. N., and M. D. Collins.** 1988. Proposal for reclassification of *Bacteroides asaccharolyticus*, *Bacteroides gingivalis*, and *Bacteroides endodontalis* in a new genus, *Porphyromonas*. Int. J. Syst. Bacteriol. **38:**128–131.
39. **Shine, J., and L. Dalgarno.** 1974. The 3 prime-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosomal binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
40. **Slots, J., L. Bragd, M. Wikström, and G. Dahlén.** 1986. The occurrence of *Actinobacillus actinomycetemcomitans*, *Bacteroides gingivalis* and *Bacteroides intermedius* in destructive periodontal disease in adults. J. Periodontol. **13:** 576–577.
41. **Tomkinson, A. E., N. F. Totty, M. Ginsburg, and T. Lindahl.** 1991. Location of the active site for enzyme-adenylate formation in DNA ligases. Proc. Natl. Acad. Sci. USA **88:**400–404.
42. **von Heijne, G.** 1985. Structural and thermodynamic aspects of the transfer of proteins into and across membranes. Curr. Top. Membr. Transp. **24:**151–179.
43. **Wandersman, C.** 1989. Secretion, processing and activation of bacterial extracellular proteases. Mol. Microbiol. **3:**1825–1831.
44. **Westerlund, B., and T. K. Korhonen.** 1993. Bacterial proteins binding to mammalian extracellular matrix. Mol. Microbiol. **9:**687–694.
45. **Williamson, M. P.** 1994. The structure and function of proline-rich regions in proteins. Biochem. J. **297:**249–260.
46. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.

Editor: A. O'Brien