INFECTION AND IMMUNITY, Oct. 1996, p. 3957–3966 Vol. 64, No. 10
0019-9567/96/$04.00+0

Characterization of the *Streptococcus pneumoniae* Immunoglobulin A1 Protease Gene (*iga*) and Its Translation Product

KNUD POULSEN,[1]* JESPER REINHOLDT,[2] AND MOGENS KILIAN[1]

Departments of Medical Microbiology and Immunology[1] and Oral Biology,[2] University of Aarhus, Aarhus, Denmark

Received 4 March 1996/Returned for modification 5 June 1996/Accepted 3 July 1996

Bacterial immunoglobulin A1 (IgA1) proteases constitute a very heterogenous group of extracellular endopeptidases which specifically cleave human IgA1 in the hinge region. Here we report that the IgA1 protease gene, *iga*, of *Streptococcus pneumoniae* is homologous to that of *Streptococcus sanguis*. By using the *S. sanguis iga* gene as hybridization probe, the corresponding gene from a clinical isolate of *S. pneumoniae* was isolated in an *Escherichia coli* lambda phage library. A lysate of *E. coli* infected with hybridization-positive recombinant phages possessed IgA1-cleaving activity. The complete sequence of the *S. pneumoniae iga* gene was determined. An open reading frame with a strongly biased codon usage and having the potential of encoding a protein of 1,927 amino acids with a molecular mass of 215,023 Da was preceded by a potential −10 promoter sequence and a putative Shine-Dalgarno sequence. A putative signal peptide was found in the N-terminal end of the protein. The amino acid sequence similarity to the *S. sanguis* IgA1 protease indicated that the pneumococcal IgA1 protease is a Zn-metalloproteinase. The primary structures of the two streptococcal IgA1 proteases were quite different in the N-terminal parts, and both proteins contained repeat structures in this region. Using a novel assay for IgA1 protease activity upon sodium dodecyl sulfate-polyacrylamide gel electrophoresis, we demonstrated that the secreted IgA1 protease was present in several different molecular forms ranging in size from approximately 135 to 220 kDa. In addition, interstrain differences in the sizes of the pneumococcal IgA1 proteases were detected. Southern blot analyses suggested that the *S. pneumoniae iga* gene is highly heterogenous within the species.

Immunoglobulin A1 (IgA1) is the predominant immunoglobulin isotype involved in the protection of mucosal membranes of the upper respiratory tract in humans (12). Although generally resistant to traditional proteases, IgA1, including its secretory form, is susceptible to IgA1 proteases, which are extracellular, bacterial enzymes that specifically cleave human IgA1 in the hinge region at one of several postproline peptide bonds absent in the IgA2 molecule (reviewed in references 14, 26, and 27). The resulting Fab fragments have retained antigen binding capacity, whereas Fc-mediated effector functions are eliminated. A number of human pathogens, including *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Ureaplasma urealyticum*, which infect and invade mucosal membranes, are known to have IgA1-cleaving activity. In addition, species of *Prevotella* and *Capnocytophaga*, which are suspected periodontal pathogens, as well as three streptococcal species that are commensals of the upper respiratory tract, *Streptococcus sanguis*, *Streptococcus mitis* biovar 1, and *Streptococcus oralis*, secrete IgA1 protease. All streptococcal IgA1 proteases cleave the same Pro-Thr bond at positions 227 and 228 in the hinge region of human IgA1, which is the only known substrate.

Human IgA1-cleaving activity has evolved among bacteria through convergent evolution in at least three independent lines, as the enzymes belong to three different classes of proteinases. Cloning and characterization of *iga* genes encoding IgA1 protease have confirmed that the proteases of *Haemophilus* and *Neisseria* spp. are genetically related serine proteases (3, 20, 28, 29), whereas that of *S. sanguis* is a Zn-metalloproteinase (8). The pneumococcal IgA1 protease has also been classified as a metalloproteinase, as it is inhibited by EDTA (13), and the enzymes from some strains show an antigenic relationship to that of *S. sanguis* (19), although DNA:DNA hybridization studies indicated that the two IgA1 proteases are not closely related at the gene level (7). On the basis of inhibition studies, the IgA1 protease of *Prevotella melaninogenica* has been classified as a cysteine proteinase (25).

Considerable diversity among the IgA1 proteases from *S. pneumoniae* has been demonstrated with enzyme-neutralizing antibodies raised in rabbits against selected IgA1 proteases (18). By this method several serologically distinct versions, termed inhibition types, of pneumococcal IgA1 proteases were identified. A similar serological heterogeneity of the IgA1 protease has been observed among *H. influenzae* strains (16) and appears to serve immune escape purposes by allowing a succession of clones of the same species to colonize one host (21).

Because of the specificity for human IgA1, no relevant animal model is available to test the biological function of the IgA1 proteases. However, several lines of indirect evidence suggest that IgA1 protease activity enables bacteria to evade the mucosal immune barrier and is an important colonization factor. A hypothetical model suggesting an essential role of the IgA1 proteases in certain invasive infectious diseases has been proposed (15).

The emerging problems with antibiotic resistance among *S. pneumoniae* strains (2) have raised an urgent need for a better understanding of the pathogenesis of this important human pathogen. To elucidate the phylogeny, characteristics, and biological significance of the pneumococcal IgA1 protease, we have cloned and sequenced the corresponding *iga* gene from a clinical isolate of *S. pneumoniae* and examined its translation

* Corresponding author. Mailing address: Department of Medical Microbiology and Immunology, The Bartholin Building, University of Aarhus, DK-8000 Aarhus C, Denmark. Phone: 45 89 42 17 36. Fax: 45 86 19 61 28.

TABLE 1. Characteristics of streptococcal strains used

Bacterial strain	Serotype or biovar	IgA1 protease activity[a]
S. pneumoniae		
PK76	Capsular serotype 3	+ (3)
PK80	Capsular serotype 7F	+ (8)
PK81	Capsular serotype 14	+ (9)
PK85	Capsular serotype 19F	+ (15)
PK88	Capsular serotype 23F	+ (4)
PK89	Capsular serotype 14	+ (9)
PK90	Capsular serotype 6B	+ (4)
PK91	Capsular serotype 6B	+ (6)
PK92	Capsular serotype 14	+ (10)
PK93	Capsular serotype 18C	+ (11)
S. sanguis SK1 (ATCC 10556)		+
S. oralis SK2 (ATCC 10557)		+
S. mitis		
SK141	Biovar 1	+
SK149	Biovar 2	−

[a] Numbers in parentheses indicate the inhibition type detected by enzyme neutralization assays (18).

product. The gene sequence is compared with the published sequence of the *iga* gene of *S. sanguis*.

MATERIALS AND METHODS

Bacterial strains and cloning vectors. The streptococcal strains used in this study are listed in Table 1. The streptococci were grown at 37°C in an atmosphere of air plus 5% CO_2 on blood agar and in 2× YT medium (33). The latter medium is devoid of high-molecular-weight proteins. Bacteriophage λL47.1 (17) was used as a *Bam*HI substitution vector, and recombinant phages were plated on *Escherichia coli* K802 (37) as described previously (33). *E. coli* XL1-Blue (Stratagene, La Jolla, Calif.) was used as a host for M13mp18 and M13mp19 derivatives (38) as well as for recombinant forms of pBluescriptII SK and KS (Stratagene). *E. coli* XL1-Blue derivatives were grown in 2× YT medium supplemented with antibiotics when appropriate.

DNA isolation, Southern blotting, and hybridizations. Whole-cell DNA was isolated as described previously (10). Briefly, the cells were washed in Na_3-citrate buffer, treated with lysozyme, lysed with sodium dodecyl sulfate (SDS), digested with proteinase K, and extracted with phenol-chloroform. Nucleic acids were eventually precipitated with ethanol. The DNA was dissolved in TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 7.4), and the quality and quantity were examined by agarose gel electrophoresis. For Southern blot analysis, approximately 2 μg of whole-cell DNA was digested with *Eco*RI as recommended (Boehringer Mannheim, Mannheim, Germany). After treatment with RNase (DNase free; Boehringer Mannheim), the DNA fragments were electrophoresed in a 1% agarose gel and subsequently transferred and fixed onto Nytran nylon membranes (Schleicher & Schuell, Keene, N.H.). Hybridizations were performed as described previously (33) except that the membranes were soaked in 1% Triton X-100 prior to prehybridization and 0.05% sodium pyrophosphate was included in all hybridization and washing buffers. Highly stringent hybridizations were performed at 65°C, and the final posthybridization wash was at 65°C in 0.5× SET (1× SET is 0.15 M NaCl, 0.5 mM EDTA, and 20 mM Tris-HCl; pH 7.0) with 0.1% SDS and 0.05% sodium pyrophosphate. Hybridizations at low stringency included prehybridization at 65°C and hybridization at temperatures declining from 65 to 45°C over a period of 8 h followed by posthybridization washes in 4× SET–0.1% SDS–0.05% sodium pyrophosphate at 45°C. For rehybridization, the filter was soaked in boiling 0.1% SDS for 10 min to remove previously hybridizing probe.

Two hybridization probes representing the *S. sanguis iga* gene were used. The two DNA fragments consisted of genomic DNA from *S. sanguis* ATCC 10556 amplified by PCR and cloned into plasmid PCR by using the TA-Cloning kit (R&D Systems Europe Ltd., Abingdon, United Kingdom). The two pairs of PCR primers were chosen according to the published nucleotide sequence of the *S. sanguis iga* gene and represented bp 571 to 3395 and bp 3369 to 5702, respectively (8). The hybridization probe representing the *S. pneumoniae iga* gene consisted of fragments of the recombinant phage λPK81*iga*-1, isolated in this study, subcloned into plasmids as described below. Vector sequences were removed from the DNA probes by digestion with appropriate restriction enzymes followed by agarose gel electrophoresis and elution with the Gene Clean Kit (BIO 101, Vista, Calif.). The DNA fragments used as probes were labelled with [α-^{32}P]dATP by nick translation (33).

Construction and screening of the *S. pneumoniae* PK81 genomic library. A *Sau*3A partial digest of whole-cell DNA from *S. pneumoniae* PK81 was fractionated by agarose gel electrophoresis. Fragments in the size range from 12 to 20 kb were extracted from the gel by electroelution and used to prepare a genomic library by using λL47.1 as a *Bam*HI substitution vector. Positive plaques among the recombinant phages packaged in vitro as described previously (33) with Gigapack II Packaging Extracts (Stratagene) and plated on *E. coli* K802 were identified by in situ hybridization with two DNA fragments labelled with [α-^{32}P]dATP by nick translation as probes. The low-stringency conditions used in the hybridization were as described above. Positive plaques were purified by replating on *E. coli* K802, and DNA was isolated from 20-ml phage cultures (24).

PCR. For amplification of genomic DNA sequences, we used 27-mer oligonucleotides as primers and *Taq* DNA polymerase (Life Technologies) as recommended by the supplier. The thermocycling program consisted of denaturation for 5 min at 94°C, 30 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 2 min followed by extension at 72°C for 7 min.

DNA sequencing. Restriction fragments of λPK81*iga*-1 were ligated into the vectors M13mp18/19 and pBluescriptII (SK/KS) by using appropriate restriction sites and transformed into *E. coli* XL1-Blue. DNAs of individual clones were sequenced from single-stranded M13 DNA or from double-stranded plasmid DNA with an Applied Biosystems model 373A DNA sequencer, using the *Taq* DyeDeoxy-Terminator Cycle Sequencing Kit (Perkin-Elmer, Foster City, Calif.). The sequence of the *iga* gene was determined by using the universal sequencing primers as well as synthetic oligonucleotides designed on the basis of the previous sequence. The DNA sequence was determined for both strands of the *iga* gene. Computer analysis of the resulting sequence was performed with the Genetics Computer Group (University of Wisconsin, Madison) package.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were separated on 1-mm-thick SDS-polyacrylamide gradient gels (4 to 10% polyacrylamide; 7 by 8 cm) as described previously (1) and either visualized by staining with Coomassie brilliant blue or renatured for subsequent analysis of IgA1-cleaving activity (see below).

IgA1 protease assays. IgA1-cleaving activity was detected by immunoelectrophoresis or enzyme-linked immunosorbent assay (ELISA) analysis of human IgA1 myeloma protein solutions incubated overnight at 37°C with the sample as previously described (23, 30). The titer of IgA1 protease activity was determined as the reciprocal of the dilution that caused cleavage of 50% of substrate IgA1 as measured by ELISA (30). For detection of IgA1 protease activity after SDS-PAGE, inhibition of SDS was overcome and proteins were renatured by incubation for 2 h at room temperature in 50 mM Tris-HCl (pH 7.4)–0.154 M NaCl (TS) containing 1% Triton X-100 and 0.5% bovine serum albumin (32). Subsequently, excess fluid was drained off and the gel was laid on a polyvinylidene difluoride membrane (pore size, 0.45 μm; Millipore, Bedford, Mass.) onto which human myeloma IgA1 had been immobilized in advance by the following procedure: the membrane was (i) wetted in alcohol, washed in water, and then incubated for 1 h with rabbit antibody directed against mouse immunoglobulin (code Z0259; DAKO, Glostrup, Denmark) diluted 1:500 in 15 ml of TS, (ii) washed (four times for 5 min each in TS containing 0.15% Tween 20 [TTS]) and incubated for 2 h with a murine monoclonal antibody recognizing the Fc_α fragment of human IgA1 (code M0728; DAKO) diluted 1:200 in TTS, and (iii) washed in TTS and incubated for 2 h in purified human myeloma IgA1(κ) (30) at 50 μg/ml in TTS and finally washed. In principle, this procedure was adopted from the ELISA for titrating IgA1 protease activity (30). The proteins in the gel were allowed to react with IgA1 on the membrane for 20 to 60 min at 37°C in a humid chamber. After the membrane was washed, cleavage of IgA1 was visualized by incubation with peroxidase-conjugated rabbit antibody to human κ light chains (code P0129; DAKO) diluted 1:1,000 in TTS, followed by washing first in TTS then in 50 mM Na-acetate (pH 5.5) and development with a chromogenic substrate of 0.004% 3-amino-9-ethyl carbazole and 0.015% H_2O_2 in 0.05 M Na-acetate (pH 5.5) (4). In this assay IgA1 protease activity in the gel is reflected as a loss of light chains (as part of Fab fragments) and detected on the membrane as a lack of staining on a red background. To allow for estimation of molecular masses, gel lanes containing molecular mass standards were stained with Coomassie brilliant blue and then equilibrated in the buffer used for renaturation of proteins in the gel. Theoretically, decreased staining of the membrane could also be due to proteolysis by IgA1 protease-unrelated proteins in the gel of either of the two antibodies that connect the human IgA1 molecules to the membrane. To assay for this possibility, an SDS-PAGE gel of a crude preparation of IgA1 protease from *S. pneumoniae* PK81 was renatured as described above and incubated for 60 min in direct contact with a membrane sequentially incubated with rabbit anti-mouse immunoglobulin and the murine monoclonal antibody recognizing the Fc_α fragment as described above. After being removed from the gel and washed, the membrane was incubated with the human IgA1 myeloma protein and then with the conjugated rabbit anti-kappa light-chain antibody and finally washed and developed as described above. No areas of the filter with decreased staining were observed. Thus, the lack of reaction observed when IgA1 protease activity was assayed after SDS-PAGE of crude preparations of proteins from *S. pneumoniae* PK81 does represent specific cleavage of human IgA1.

Preparation of IgA1 protease from *S. pneumoniae*. The proteins in 1 liter of culture supernatant were precipitated by adding ammonium sulfate to 60% saturation. After centrifugation, the pellet was dissolved in 5 ml of 0.05 M Tris-HCl (pH 8.2) and loaded onto a column (1.6 by 50 cm) of Sepharose 12

(prep grade; Pharmacia, Uppsala, Sweden). Upon elution with the same buffer, fractions were assayed for IgA1 protease activity, and the peak fractions were pooled and subjected to ion-exchange chromatography on a Mono Q HR 5/5 column (Pharmacia) which was eluted with a linear gradient of 0 to 0.5 M NaCl in 0.05 M Tris-HCl (pH 8.2). The titer of IgA1 protease activity in each of the resulting 0.5-ml fractions was determined as described above.

Secreted proteins of eight pneumococcal strains were concentrated approximately 50-fold by size exclusion centrifugation of the culture supernatant with Centriprep concentrators (Amicon, Beverly, Mass.) and used for IgA1 protease assay. The titers of the resulting IgA1 protease preparations were approximately 250.

Nucleotide sequence accession number. The nucleotide sequence of the *iga* gene of *S. pneumoniae* PK81 reported here has been deposited in the EMBL nucleotide sequence database under accession number X94909.

RESULTS

Cloning and isolation of the pneumococcal *iga* gene. The IgA1 proteases of *S. sanguis* and *S. pneumoniae* are both metalloproteinases which cleave the same peptide bond in human IgA1. Despite these common characteristics, Gilbert et al. found that the cloned IgA1 protease gene, *iga*, of *S. sanguis* had no detectable homology to the *S. pneumoniae iga* gene as revealed by lack of hybridization to pneumococcal DNA, even under low-stringency conditions (7). We reexamined this by using low stringency hybridization conditions including a sequential decrease in hybridization temperature, which presumably favors specific annealing. In Southern blot analyses of *Eco*RI-digested whole-cell DNA of *S. pneumoniae* PK81, we used as probes two DNA fragments representing each half of the *S. sanguis iga* gene prepared by PCR. Each probe recognized one and the same *Eco*RI fragment of 11 kb, indicating that the hybridizations were specific for the pneumococcal *iga* gene (data not shown).

A gene library of partially *Sau*3A-digested whole-cell DNA of *S. pneumoniae* PK81 was constructed in *E. coli* K802 by using the vector λL47.1. The two fragments of the *S. sanguis iga* gene were used together as DNA probes to screen this genomic library under low-stringency conditions. Three positive plaques were isolated and selected for further analysis. Liquid *E. coli* lysates generated by two of these recombinant phages, λPK81*iga*-1 and λPK81*iga*-5, were found to possess IgA1-cleaving activity as demonstrated by immunoelectrophoresis of human IgA1 incubated with the supernatants. This confirms the cloning of the IgA1 protease gene. The lysate generated by the third phage, λPK81*iga*-3, had no detectable IgA1-cleaving activity. DNAs from the three clones were characterized by mapping sites for the restriction enzymes *Eco*RI, *Hin*dIII, and *Bam*HI. As shown in Fig. 1, the restriction maps of the three phage inserts overlap, suggesting that the *iga* gene in *S. pneumoniae* PK81 is a single-copy gene. The localization and orientation of the *iga* gene within the recombinant phages were determined by Southern blot analysis of *Eco*RI-, *Hin*dIII-, and *Bam*HI-restricted phage DNA with the individual *S. sanguis iga* gene fragments as probes and hybridization conditions previously described. The lack of IgA1-cleaving activity in the lysate generated by λPK81*iga*-3 is presumably caused by the lack of the 3′ part of the *iga* gene in this phage.

DNA sequence analysis. Appropriate DNA restriction fragments of λPK81*iga*-1 were subcloned into phage M13 or plasmid pBluescriptII, and both strands of the region shown in Fig. 1 were sequenced. The resulting 6,478-nucleotide sequence (Fig. 2) contains a large open reading frame (nucleotides 259 to 6,039) encoding a deduced polypeptide of 1,927 amino acids with an M_r of 215,023 and a pI of 5.61. The open reading frame is preceded by a typical ribosome binding site (Shine-Dalgarno sequence) (34). Just upstream, the sequence TAAATT is in an appropriate position and contains the four best conserved nucleotides of the consensus −10 promoter region (6, 9), whereas a −35 promoter sequence is not evident. An inverted repeat structure downstream of the open reading frame may constitute a rho-independent transcription terminator. The codon usage of the open reading frame is similar to that reported for other pneumococcal genes (22) as well as to that of the *S. sanguis iga* gene (8). The open reading frame has a GC content of 37.0%, which is close to the 38.5% estimated for the *S. pneumoniae* genome by chemical analysis (31). The low GC content is caused by a codon usage biased towards codons with adenine and thymine in the second and third nucleotide positions, with such codons representing 65.4 and 73.1%, respectively, whereas the frequency of codons with adenine or thymine in the first position is 50.4%.

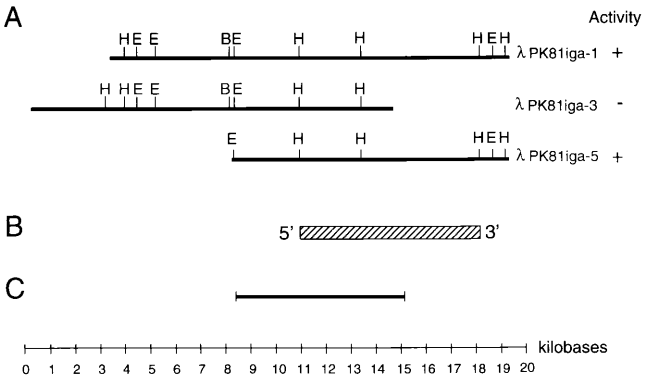
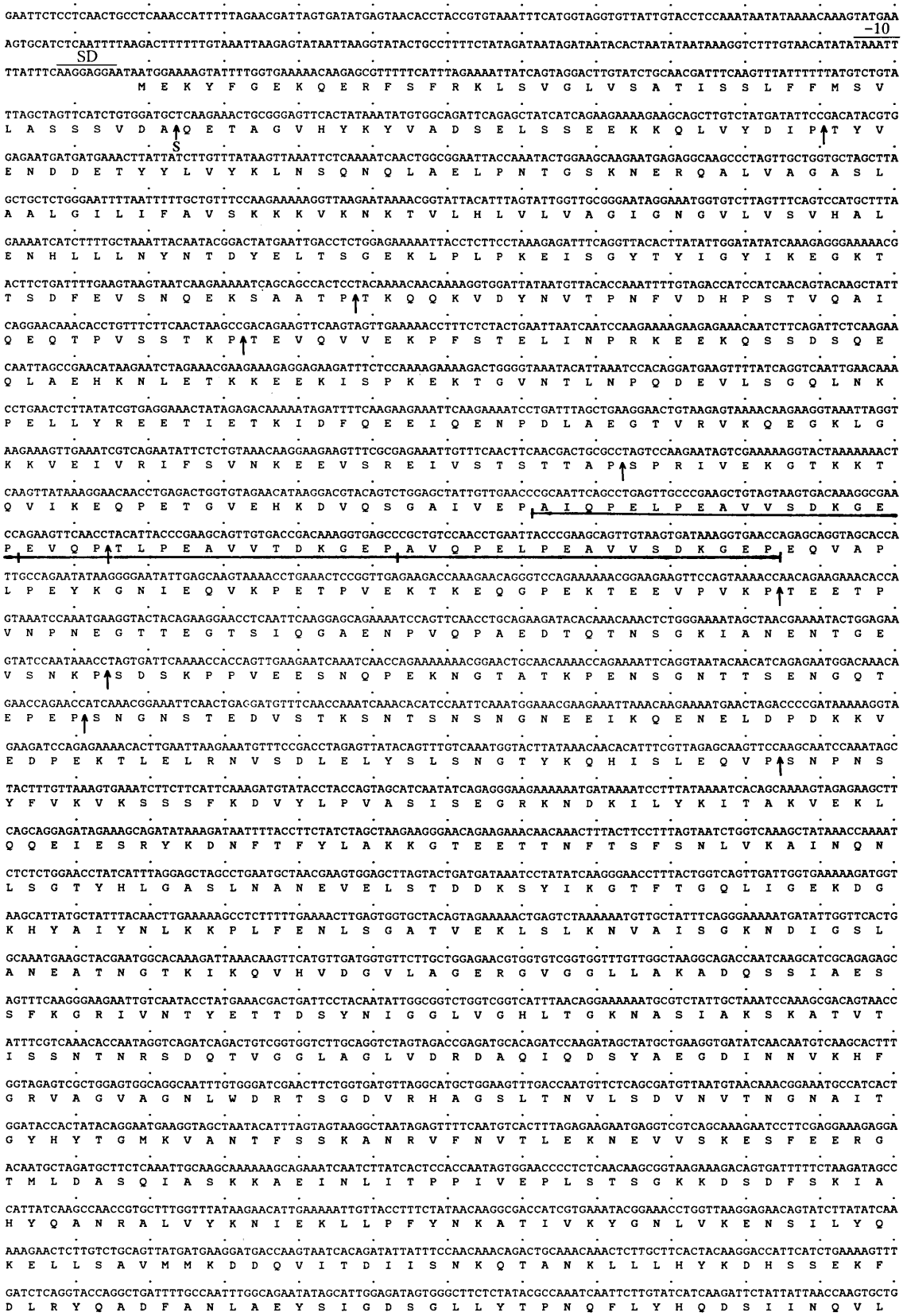


FIG. 1. Structure of the *S. pneumoniae* PK81 *iga* gene. (A) Restriction maps of the three recombinant lambda phage inserts analyzed with homology to the *S. sanguis iga* gene probes. Symbols: E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; + or −, presence or absence, respectively, of IgA1-cleaving activity in culture supernatants of *E. coli* infected with the individual recombinant phages. (B) Restriction fragments hybridizing to the two *iga* gene probes from *S. sanguis*. The deduced orientation of the *iga* gene is indicated by 5′ and 3′. (C) The region of the pneumococcal *iga* gene sequenced in this study.

Homology between the cloned pneumococcal *iga* gene and other streptococcal *iga* genes. Alignment of the nucleotide sequence of the *S. pneumoniae* PK81 *iga* gene and the published *S. sanguis iga* gene sequence (8) revealed that the genes are approximately 70% homologous. The previously observed failure of the two *iga* genes to cross-hybridize (7) is conceivably due to the very few long stretches (i.e., more than 20 nucleotides) of identical sequence in the two genes (results not shown). In the regions flanking the *iga* gene we detected no significant homology between the two species.

The *iga* genes of an additional nine *S. pneumoniae* strains were examined by using the *iga* gene of strain PK81 (from the *Msp*I site around position 1728 to the *Hin*fI site around position 6159 in Fig. 2) as a probe in Southern blot analysis of *Eco*RI-restricted whole-cell DNA (Fig. 3). In genomic DNAs from 7 of the 10 pneumococcal strains analyzed, the *iga* gene probe hybridized with a single fragment. In DNAs of the remaining three strains, additional fragments were recognized by the probe, presumably because of the presence of an *Eco*RI site within the *iga* gene combined with incomplete digestion of the genomic DNAs of two strains (PK88 and PK85). These results are consistent with the *iga* gene being a single-copy gene. The sizes of *Eco*RI restriction fragments hybridizing to the PK81 *iga* gene probe varied among the strains, except for two isolates (PK81 and PK89) which were both of capsular serotype 14. These two strains presumably represent the same clonal type, as they have previously been assigned to closely related multilocus enzyme electrophoretic types and produce the same antigenic version of the IgA1 protease (18). The



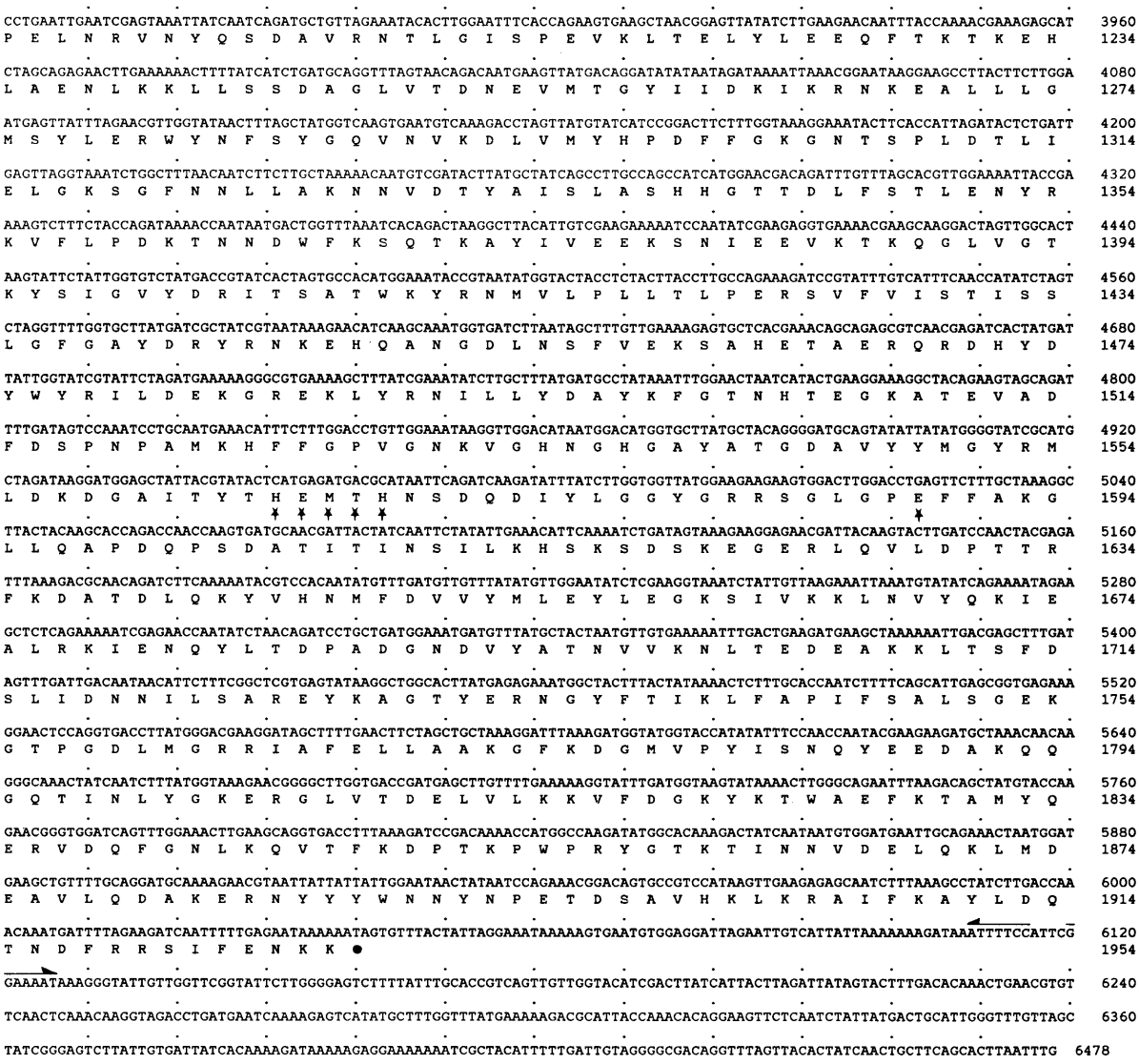


FIG. 2. Nucleotide sequence of the *iga* gene from *S. pneumoniae* PK81. The deduced amino acid sequence of the large open reading frame is shown below the gene sequence. The stop codon is indicated by a dot. The proposed −10 promoter sequence and the putative Shine-Dalgarno sequence (SD) are overlined. The inverted repeat structure, which is a possible rho-independent transcriptional terminator, is indicated by divergent arrows. The proposed cleavage site for the N-terminal signal peptide is indicated by an arrow marked S. Vertical arrows indicate suggested autoproteolytic sites, and the zinc-binding motif is shown by stars below the sequence. The repeat structure is underlined.

differences in intensity of hybridization among the pneumococcal strains after low-stringency washing (Fig. 3A) can be explained by differences in the amount of DNA applied to the gel combined with incomplete transfer of large DNA fragments.

Differences in response to posthybridization washing stringency were used to assay for the degree of sequence homology to the *S. pneumoniae* PK81 *iga* gene probe. As shown in Fig. 3, all 10 pneumococcal *iga* genes tested had a high degree of homology to the probe, as they hybridized even under high-stringency washing conditions. However, the decline in the hybridization signal as a result of more-stringent washing was more pronounced for some strains (PK92, PK91, and PK80), suggesting that there is a difference in homology to the probe among the pneumococcal *iga* genes (Fig. 3B).

The Southern blot experiment also included strains representing each of the other three streptococcal species known to produce IgA1 protease activity, i.e., *S. mitis* biovar 1, *S. oralis*, and *S. sanguis*. All three strains contained genomic DNA sequences with homology to the pneumococcal *iga* gene probe, whereas a strain of *S. mitis* biovar 2, which is devoid of IgA1-cleaving activity, did not (Fig. 3). The results of the hybridization experiments including different conditions of stringency indicate that the *iga* gene of *S. pneumoniae* PK81 is highly homologous to those of the *S. oralis* and *S. mitis* biovar 1 strains analyzed and more distantly related to the *S. sanguis* ATCC 10556 *iga* gene.

The same Southern blot filter was used for hybridization with the two *S. sanguis iga* gene fragments described above. In genomic DNAs from the 10 *S. pneumoniae* strains as well as the *S. mitis* biovar 1 strain, this *iga* gene probe hybridized to *Eco*RI fragments of the same size as those recognized by the PK81 *iga* gene probe only when low-stringency washing conditions were used, whereas the hybridization was washed off by

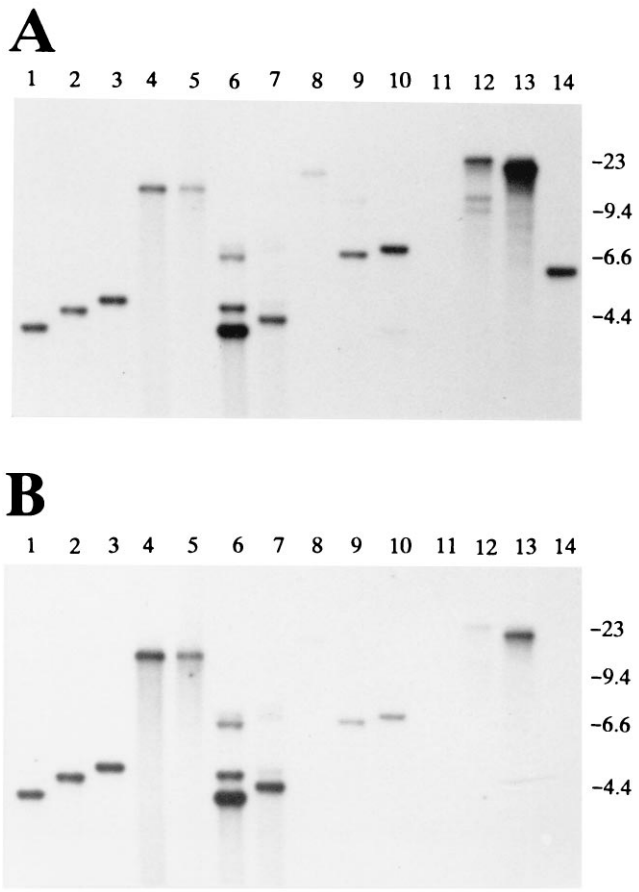


FIG. 3. Southern hybridization analysis. Whole-cell DNA was digested with *Eco*RI, gel electrophoresed, transferred to a nylon membrane, and hybridized with the *S. pneumoniae* PK81 *iga* gene probe. Lanes: 1 to 10, *S. pneumoniae* strains PK90, PK76, PK93, PK89, PK81, PK88, PK85, PK92, PK91, and PK80, respectively; 11, *S. mitis* SK149 (biovar 2); 12, *S. mitis* SK141 (biovar 1); 13, *S. oralis* SK2; 14, *S. sanguis* SK1. Molecular size markers (in kilobases) are indicated to the right. (A) Hybridization after low-stringency wash. (B) Hybridization after high-stringency wash.

high-stringency conditions. In contrast, the *S. sanguis iga* gene probe hybridized to the single *Eco*RI fragment of the *S. oralis* genomic DNA even under high-stringency conditions, although the intensity was highly reduced compared with that under the low-stringency washing conditions (data not shown).

Homology between the *S. pneumoniae* and *S. sanguis* IgA1 proteases. The deduced sequence of the translation product of the *S. pneumoniae* PK81 IgA1 protease gene is 64% identical and 76% similar to that of the published *S. sanguis* IgA1 protease when aligned as shown in Fig. 4. The homology is unevenly distributed along the protein, with a pronounced diversity including several insertions and deletions in the N-terminal third. In contrast, only a few short gaps were introduced in the sequences to obtain the alignment in the middle and C-terminal parts of the proteins, except for the extreme C terminus, which differed significantly. Notably, the zinc-binding sequence is conserved in the two proteins (around amino acid 1565 in Fig. 4), indicating that they have a common catalytic mechanism. Like most extracellular proteins of streptococci, the two IgA1 proteases contain no cysteine residues.

At the N terminus of the open reading frame shown in Fig. 2, a basic N-terminal region followed by a stretch of hydrophobic residues suggests the presence of a signal sequence with a likely cleavage site at Ala-Gln in positions 42 and 43 (35). Presumably, the *S. sanguis* IgA1 pre-protease has a similar signal peptide, since, as noted previously (8), the same structural features are found in the translation product deduced from the *S. sanguis iga* gene, assuming that the translation initiates at the methionine denoted position 1 in Fig. 4.

In the *S. sanguis* ATCC 10556 IgA1 protease a highly uniform sequence of 20 amino acids is tandemly repeated 10 times (amino acid positions 349 to 546 in the *S. sanguis* IgA1 protease sequence shown in Fig. 4) (8). In that area the two streptococcal IgA1 proteases are quite different, and in the *S. pneumoniae* protein a repeat structure of 17 amino acids is present in three tandemly arranged copies (Fig. 2).

Crude preparation and analysis of the *S. pneumoniae* IgA1 protease. The IgA1 protease of *S. pneumoniae* PK81 was prepared by sequential gel filtration and anion-exchange chromatography of culture supernatant proteins concentrated by ammonium sulfate precipitation. The elution profile upon gel filtration of the enzyme measured as IgA1-cleaving activity was reasonably consistent with the M_r predicted from the sequence data. The two procedures of chromatography combined resulted in a protease preparation with highly elevated activity. The maximum titer of fractions resulting from anion-exchange chromatography was 5,267, compared with 10 in the culture supernatant. However, the IgA1-cleaving activity was eluted as a broad peak at concentrations of NaCl ranging from 0.29 to 0.37 M, suggesting electrochemical heterogeneity of the IgA1 protease molecules.

Direct evidence for the presence of a diverse population of IgA1 protease molecules in the eluent fractions was obtained by SDS-PAGE and subsequent application of a novel assay for detection of IgA1 protease activity. With this technique, the IgA1-cleaving activity could be ascribed to several distinct molecular species with M_rs spanning the range from 135,000 to 220,000 (Fig. 5). Proteins with estimated molecular masses of 135, 150, and 220 kDa accounted for the major part of the IgA1-cleaving activity. Comparison of the gel stained for total protein (Fig. 5A) with the localization of IgA1-cleaving activity (Fig. 5B and C) indicated that IgA1 protease was only a minor constituent of the proteins in each of the eluent fractions.

An attempt to perform N-terminal amino acid sequencing was unsuccessful because of the small amounts of proteins in the bands in combination with their high molecular masses and contamination with other proteins with similar mobilities.

The assay for detection of IgA1-cleaving activity after SDS-PAGE was applied to the supernatants of an additional seven strains of *S. pneumoniae*. In all strains, active IgA1 protease was present in several molecular forms (Fig. 6). Thus, the presence of different sizes of the enzyme appears to be a general feature in the species. Although the activity of a high-molecular-weight form ($M_r \geq 200{,}000$) was predominant in all strains, the activity profiles suggested that the spectra of IgA1 protease molecules may differ among *S. pneumoniae* strains.

DISCUSSION

For a number of bacterial pathogens that colonize mucosal surfaces, IgA1 protease production represents a putative virulence factor. Notably, all of the three leading causes of bacterial meningitis secrete IgA1 protease. In the present study we have characterized an *S. pneumoniae* IgA1 protease gene, *iga*, and its translation product.

In contrast to the case for previous studies (7), we were able to detect specific hybridization of the *S. sanguis iga* gene with pneumococcal DNA by using a low-stringency hybridization protocol. This cross-hybridization was used to isolate the IgA1

FIG. 4. Amino acid sequence homology between the IgA1 proteases of *S. pneumoniae* PK81 (upper sequence) (this study) and *S. sanguis* ATCC 10556 (lower sequence) (8). The amino acid sequences were deduced from the *iga* gene sequences. Identical residues are indicated by broken bars, and functionally equivalent and similar residues are indicated by double and single dots, respectively. Gaps introduced in the sequences to obtain maximal homology are shown by dots. An arrow between positions 42 and 43 indicates the proposed cleavage site for the N-terminal signal peptide. The zinc-binding motif is shown by stars.

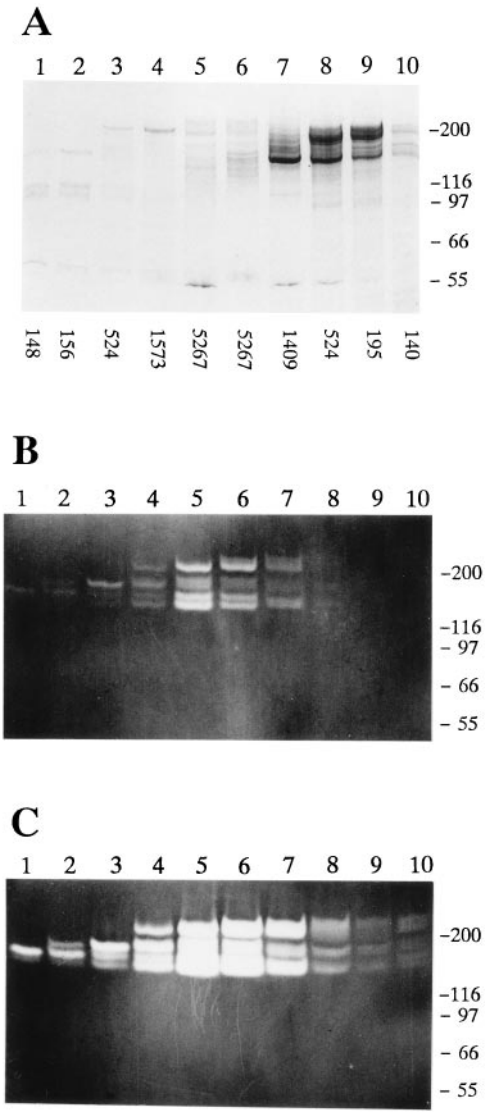


FIG. 5. Molecular diversity of the IgA1 protease from *S. pneumoniae* PK81. (A) SDS-PAGE followed by Coomassie blue staining of 10 consecutive eluent fractions that constituted the peak of IgA1 protease activity upon ion-exchange chromatography. The titer of activity is indicated below each lane, and the mobilities of molecular mass markers (in kilodaltons) are shown to the right. From each 0.5-ml fraction 10 μl was loaded onto the gel. (B and C) IgA1 protease activity assays of a gel identical to the one in panel A. The proteins in the gel were renatured and placed for 20 min (B) or 60 min (C) on a filter coated with human IgA1 as described in Materials and Methods. IgA1-cleaving activity was subsequently visualized on the filter by the lack of reaction with anti-light-chain antibodies.

protease gene in an *E. coli* phage library of genomic DNA from *S. pneumoniae* PK81. The nucleotide sequence of the *iga* gene revealed a large open reading frame surrounded by regulatory sequence elements typical of streptococcal genes (6). The presence of a typical −10 promoter sequence suggests that the pneumococcal *iga* gene is not part of a larger operon as has been suggested for the *S. sanguis* IgA1 protease gene (8). A biased codon usage implies that the GC content of the gene is low and very similar to that of the pneumococcal genome, suggesting that it is an ancient gene in the species, although it may have been acquired by horizontal gene transfer from related streptococcal species.

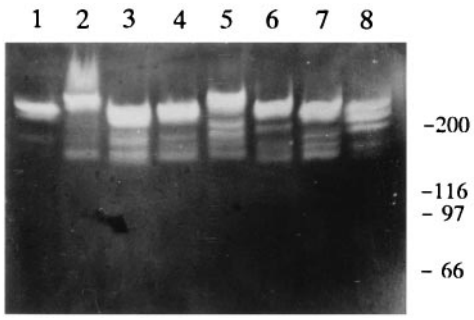


FIG. 6. IgA1-cleaving activities of different *S. pneumoniae* strains. Proteins in the culture supernatant concentrated by size exclusion centrifugation were separated according to size by SDS-PAGE, renatured in the gel, and allowed to react with human IgA1 bound to a filter for 60 min, and IgA1-cleaving activity was subsequently visualized on the filter by the lack of reaction with anti-light-chain antibodies as described in Materials and Methods. Lane 1, strain PK91; lane 2, strain PK85; lane 3, strain PK80; lane 4, strain PK81; lane 5, strain PK93; lane 6, strain PK90; lane 7, strain PK89; lane 8, strain PK88. Molecular mass markers (in kilodaltons) are indicated to the right.

The amino acid sequence of the *S. pneumoniae* IgA1 protease deduced from the *iga* gene sequence is homologous to that of the *S. sanguis* IgA1 protease. The *S. sanguis* protease is known to be a Zn-metalloproteinase (8), and the sequence similarities indicate that the pneumococcal IgA1 protease has the same catalytic mechanism, as the zinc-binding motif is conserved in the two enzymes. This is in accordance with the previous finding, based on inhibition studies, that the enzyme is a metalloproteinase (13). The primary structures of the two streptococcal IgA1 proteases show no similarity to those of the functionally equivalent proteins from *Haemophilus* and *Neisseria* spp., which are serine-type proteases.

Amino acid residues important for enzymatic specificity and activity are thought to be conserved in the IgA1 protease. Therefore, the high degree of homology in the middle and C-terminal parts of the two IgA1 proteases suggests that these parts of the protein are important to enzymatic function, whereas the significant diversity in the N-terminal third suggests that this part, with the possible exception of small conserved stretches like the proposed signal peptide, is not essential. The latter idea is supported by the observation that when expressed in *E. coli*, the *S. sanguis* IgA1 protease lacks the 209 N-terminal amino acids as compared with the translation product proposed in Fig. 4 but has retained enzymatic activity (8).

Considerable serological diversity among pneumococcal IgA1 proteases has been observed (18). One possible mechanism to create diversity is the presence of repeat sequences which may be subject to variation in number and sequence as a result of frequent unequal homologous recombination followed by mutations. Repeat structures are frequent features of surface-exposed and secreted proteins in streptococci. A 20-amino-acid sequence is tandemly repeated 10 times in the N-terminal part of the *S. sanguis* IgA1 protease (8), whereas in the *S. pneumoniae* protease analyzed, a quite different 17-amino-acid repeat structure in the same area was observed and was in only three copies. The biological significance of these repeat structures and their contribution to some of the size variations observed among strains are uncertain. Characterization and comparison of additional streptococcal IgA1 proteases may contribute to localize the regions involved in serological variation as well as those essential for enzymatic activity.

Using a novel assay for demonstration of IgA1-cleaving activity, we were able to associate IgA1 protease activity with

distinct protein bands in SDS-PAGE. When applied to *S. pneumoniae* PK81, this analysis revealed that in the culture supernatant the IgA1 protease was present in several different molecular sizes. This intrastrain variation in sizes of proteins capable of cleaving human IgA1 is a general feature of *S. pneumoniae*, as similar variation was observed for seven other strains analyzed. The translation product deduced from the *iga* gene sequence analyzed has a molecular mass of 215,023 Da, and when cleavage of the proposed signal peptide is taken into account, the secreted IgA1 protease is estimated to be 210,578 Da. Although proteins with such molecular masses with IgA1-cleaving activity were detected, additional proteins with several different, lower masses were found to cleave human IgA1. This indicates that the protease is posttranslationally processed, as several lines of evidence exclude multiple copies of the *iga* gene. The presence of a precursor with an additional polypeptide segment located between the signal peptide and the mature enzyme is characteristic of other extracellular proteases of gram-positive bacteria (36). In that context, it is noteworthy that the N-terminal part of the deduced amino acid sequence of the IgA1 protease is rich in proline residues and contains a number of Pro-Thr and Pro-Ser sites resembling the Pro-Thr peptide bond cleaved in substrate IgA1 by streptococcal IgA1 proteases (Fig. 2). Cleavage at the nine sites indicated in Fig. 2 would result in IgA1 proteases with masses in the range of 207,294 to 141,654 Da, in agreement with the number and sizes of proteins with IgA1-cleaving activity observed upon SDS-PAGE (Fig. 5). However, further studies are needed to clarify whether the pneumococcal IgA1 protease is actually subject to autoproteolytic processing, in analogy with the serine-type IgA1 proteases, which are autoproteolytically cleaved during the secretion process (28, 29). Apparently, the lower-molecular-mass versions of the IgA1 protease are formed when high concentrations of the protease are present for a prolonged time as they are formed during the purification process, whereas in the freshly concentrated culture supernatant the high-molecular-mass form predominates. It should be noted that other, yet unknown factors such as pH and ion strength may influence formation of the low-molecular-mass forms.

A total of four streptococcal species, *S. sanguis*, *S. pneumoniae*, *S. oralis*, and *S. mitis* biovar 1, are known to produce IgA1 protease activity. Our hybridization experiments revealed that all of these proteases are related at the gene level. Apparently, the degree of homology varies among the species, with the *S. sanguis* IgA1 protease being most distantly related to the others. This order of relatedness is in accordance with the phylogeny of streptococci as revealed by 16S rRNA sequences (11). The hybridization analyses, in addition, revealed a pronounced diversity both in *Eco*RI restriction fragment length patterns and in degrees of sequence homology of the *iga* genes among different *S. pneumoniae* isolates, suggesting that this is an extremely variable gene within the species.

One consequence of the pattern of homology between the *S. sanguis* and *S. pneumoniae* IgA1 protease genes analyzed is that the two *iga* genes hybridize only under extremely low-stringency conditions. As specific annealing is a prerequisite for homologous recombination, this may suggest that horizontal transfer of *iga* gene sequences between the two species is not efficient. Horizontal gene transfer among streptococci is known to occur in vivo (5), and the divergent evolution of the *iga* gene sequences might be an adaptation of either of the two species to avoid sharing epitopes of the other IgA1 protease.

In summary, the present study has revealed that the *S. pneumoniae* and *S. sanguis* IgA1 proteases are homologous enzymes and that the genetic relationships of *iga* genes among different streptococcal species appear to be in accordance with the assumed phylogeny of the genus. An interesting aspect is the detection of proteins of different sizes with IgA1-cleaving activity within and between pneumococcal strains. A number of open questions remain, particularly concerning the molecular nature of the different forms of the *S. pneumoniae* IgA1 protease and the mechanism of antigenic diversity. The cloning and characterization of the pneumococcal *iga* gene provide a means to further study the IgA1 protease protein, e.g., by genetic modifications for functional and serological analyses. The results may help to elucidate the role of the IgA1 proteases in the process of infection.

ACKNOWLEDGMENTS

This work was supported by Danish Medical Research Council grant no. 12-1615.

We are grateful for the technical assistance of Ella Brandt, and we thank L. Sottrup-Jensen for his attempt to perform N-terminal amino acid sequence analysis.

REFERENCES

1. **Ahl, T., and J. Reinholdt.** 1991. Detection of immunoglobulin A1 protease-induced Fab_{α} fragments on dental plaque bacteria. Infect. Immun. **59:**563–569.
2. **Appelbaum, P. C.** 1992. Antimicrobial resistance in *Streptococcus pneumoniae*: an overview. Clin. Infect. Dis. **15:**77–83.
3. **Bachovchin, W. W., A. G. Plaut, G. R. Flentke, M. Lynch, and C. A. Kettner.** 1990. Inhibition of IgA1 proteinases from *Neisseria gonorrhoeae* and *Haemophilus influenzae* by peptide prolyl boronic acids. J. Biol. Chem. **265:**3738–3743.
4. **Bjerrum, O. J., and N. H. H. Heegaard.** 1988. Handbook of immunoblotting of proteins. CRC Press, Inc., Boca Raton, Fla.
5. **Dowson, C. G., A. Hutchison, N. Woodford, A. P. Johnson, R. C. George, and B. G. Spratt.** 1990. Penicillin-resistant viridans streptococci have obtained altered penicillin-binding protein genes from penicillin-resistant strains of *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **87:**5858–5862.
6. **Ferretti, J. J., and R. Curtiss III.** 1987. Compilation of nucleotide sequences that signal the interaction of transcription and translation in streptococci, p. 293. *In* J. J. Ferretti and R. Curtiss III (ed.), Streptococcal genetics. American Society for Microbiology, Washington, D.C.
7. **Gilbert, J. V., A. G. Plaut, Y. Fishman, and A. Wright.** 1988. Cloning of the gene encoding streptococcal immunoglobulin A protease and its expression in *Escherichia coli*. Infect. Immun. **56:**1961–1966.
8. **Gilbert, J. V., A. G. Plaut, and A. Wright.** 1991. Analysis of the immunoglobulin A protease gene of *Streptococcus sanguis*. Infect. Immun. **59:**7–17.
9. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
10. **Hohwy, J., and M. Kilian.** 1995. Clonal diversity of the Streptococcus mitis biovar 1 population in the human oral cavity and pharynx. Oral Microbiol. Immunol. **10:**19–25.
11. **Kawamura, Y., X.-G. Hou, F. Sultana, H. Miura, and T. Ezaki.** 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. Int. J. Syst. Bacteriol. **45:**406–408.
12. **Kett, K., P. Brandtzaeg, J. Radl, and J. J. Haaijman.** 1986. Different subclass distribution of IgA-producing cells in human lymphoid organs and various secretory tissues. J. Immunol. **136:**3631–3635.
13. **Kilian, M., J. Mestecky, R. Kulhavy, M. Tomana, and W. T. Butler.** 1980. IgA1 proteases from *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Neisseria meningitidis*, and *Streptococcus sanguis*: comparative immunochemical studies. J. Immunol. **124:**2596–2600.
14. **Kilian, M., and J. Reinholdt.** 1986. Interference with IgA defence mechanisms by extracellular bacterial enzymes, p. 173–208. *In* C. S. F. Easmon and J. Jeljaszewics (ed.), Medical microbiology, vol. 5. Academic Press, Inc., London.
15. **Kilian, M., and J. Reinholdt.** 1987. A hypothetical model for the development of invasive infection due to IgA1 protease-producing bacteria, p. 1261–1269. *In* J. Mestecky, J. R. McGhee, J. Bienenstock, and P. L. Ogra (ed.), Recent advances in mucosal immunology, part B. Plenum Publishing Corp., New York.
16. **Kilian, M., B. Thomsen, T. E. Petersen, and H. Bleeg.** 1983. Molecular biology of *Haemophilus influenzae* IgA1 proteases. Mol. Immunol. **20:**1051–1058.
17. **Loenen, W. A. M., and W. J. Brammar.** 1980. A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. Gene **10:**249–259.
18. **Lomholt, H.** 1995. Evidence of recombination and an antigenically diverse

immunoglobulin A1 protease among strains of *Streptococcus pneumoniae*. Infect. Immun. **63:**4238–4243.
19. **Lomholt, H., and M. Kilian.** 1994. Antigenic relationships among immunoglobulin A1 proteases from *Haemophilus*, *Neisseria*, and *Streptococcus* species. Infect. Immun. **62:**3178–3183.
20. **Lomholt, H., K. Poulsen, and M. Kilian.** 1995. Comparative characterization of the *iga* gene encoding IgA1 protease in *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Haemophilus influenzae*. Mol. Microbiol. **15:**495–506.
21. **Lomholt, H., L. van Alphen, and M. Kilian.** 1993. Antigenic variation of immunoglobulin A1 proteases among sequential isolates of *Haemophilus influenzae* from healthy children and patients with chronic obstructive pulmonary disease. Infect. Immun. **61:**4575–4581.
22. **Martin, B., and J.-P. Claverys.** 1991. Codon usage patterns for *Streptococcus pneumoniae* and *Escherichia coli*, p. 295–296. *In* G. M. Dunny, P. P. Cleary, and L. L. McKay (ed.), Genetics and molecular biology of streptococci, lactococci, and enterococci. American Society for Microbiology, Washington, D.C.
23. **Mestecky, J., and M. Kilian.** 1985. Immunoglobulin A. Methods Enzymol. **116:**37–75.
24. **Mikkelsen, B. M., M. E. Clark, G. Christiansen, O. M. Klintebaek, J. T. Nielsen, K. K. Thomsen, and J. P. Hjorth.** 1985. The structure of two distinct pancreatic amylase genes in mouse strain YBR. Biochem. Genet. **23:**511–524.
25. **Mortensen, S. B., and M. Kilian.** 1984. Purification and characterization of an immunoglobulin A1 protease from *Bacteroides melaninogenicus*. Infect. Immun. **45:**550–557.
26. **Mulks, M. H.** 1985. Microbial IgA proteases, p. 81–104. *In* I. A. Holder (ed.), Bacterial enzymes and virulence. CRC Press, Boca Raton, Fla.
27. **Plaut, A. G.** 1983. The IgA1 proteases of pathogenic bacteria. Annu. Rev. Microbiol. **37:**603–622.
28. **Pohlner, J., R. Halter, K. Beyreuther, and T. F. Meyer.** 1987. Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature (London) **325:**458–462.
29. **Poulsen, K., J. Brandt, J. P. Hjorth, H. C. Thøgersen, and M. Kilian.** 1989. Cloning and sequencing of the immunoglobulin A1 protease gene (*iga*) of *Haemophilus influenzae* serotype b. Infect. Immun. **57:**3097–3105.
30. **Reinholdt, J.** 1996. A method for titration of inhibiting antibodies to bacterial immunoglobulin A1 proteases in human serum and secretions. J. Immunol. Methods **191:**39–48.
31. **Rotta, J.** 1986. Pyogenic hemolytic streptococci, p. 1047–1054. *In* P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 2. Williams & Wilkins, Baltimore.
32. **Russell, R. R. B.** 1979. Use of Triton X-100 to overcome the inhibition of fructosyltransferase by SDS. Anal. Chem. **97:**173–175.
33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
35. **Von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
36. **Wandersman, C.** 1989. Secretion, processing and activation of bacterial extracellular proteases. Mol. Microbiol. **3:**1825–1831.
37. **Wood, W. B.** 1966. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affected the restriction and modification of DNA. J. Mol. Biol. **16:**118–133.
38. **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: V. A. Fischetti