INFECTION AND IMMUNITY, July 1996, p. 2387–2390 Vol. 64, No. 7
0019-9567/96/$04.00+0

Conservation of the C5a Peptidase Genes in Group A and B Streptococci

ILONA CHMOURYGUINA,[1] ALEXANDER SUVOROV,[1] PATRICIA FERRIERI,[2] AND P. PATRICK CLEARY[3]*

Institute of Experimental Medicine, St. Petersburg, Russia,[1] and Departments of Laboratory Medicine and Pathology and Pediatrics[2] and Microbiology,[3] University of Minnesota, Minneapolis, Minnesota 55455

Received 22 January 1996/Returned for modification 21 February 1996/Accepted 10 April 1996

The chromosome of group B streptococci (GBS) contains a gene which is related to the C5a peptidase gene (*scpA*) of group A streptococci (GAS). *scpA* encodes a surface-associated peptidase (group A streptococcal C5a peptidase [SCPA]) which specifically cleaves C5a, a major chemoattractant generated in serum by activation of complement. The entire *scpA*-like gene (*scpB*) was cloned from a GBS strain and sequenced. The gene encodes an open reading frame of 3,450 bp, which corresponds to a deduced protein (SCPB) of 1,150 amino acids with a molecular weight of 126,237 Da. Nucleotide and deduced amino acid sequences of SCPB were found to be highly homologous to those of SCPAs from GAS. Unexpectedly, *scpA*12 is more similar to *scpB* than to another GAS gene, *scpA*49. The sequence 5′ of the open reading frame, including transcription start and a termination site in the signal sequence, is also similar to that of *scpA*, although less conserved than the coding sequences. The near identity of GBS and GAS peptidases is consistent with horizontal transmission of the *scp* gene between these species. Recombinant SCPB was expressed in *Escherichia coli* by using the expression vector plasmid pGEX-4T-1 and was shown to be identical in size to the enzyme extracted from the parental GBS strain 78-471.

Group B streptococcus (GBS) is an important human pathogen which is responsible for a broad range of human diseases. This microorganism is known as a major cause of sepsis and meningitis in newborns and infants. In spite of its medical significance, its mechanisms of pathogenicity are poorly understood. Major GBS virulence factors include polysaccharide capsules and several surface-bound proteins, the C-protein complex. The functional role of these proteins and the means by which they interact with the host immune system are still poorly defined. Hill and collaborators discovered a C5a-ase enzyme on the surface of some strains of GBS and suggested that it could contribute to virulence of the organism (9). They showed that C5a-ase cleaves the human C5a chemotaxin at residue His-67 in the neutrophil binding site, a position identical to where the group A streptococcus (GAS) enzyme acts (2, 7). The GAS enzyme is also surface bound and has been shown to retard the influx of neutrophils and enhance clearance of streptococci from sites of streptococcal infection in mice (4, 10, 15). We previously reported that the restriction map of a gene (*scpB*) in the GBS genome is similar to that of *scpA* in GAS (5). In the present work we cloned and sequenced the entire *scpB* gene, and we defined its relationship to the C5a-ase, which was originally described by Hill et al. (9) and further characterized by Bohnsack and colleagues (2, 3), and its relationship to group A streptococcal C5a peptidase (SCPA) (4).

MATERIALS AND METHODS

Bacterial strains and plasmids. Group B streptococcal strain 78-471 (type IIα+ β), used for this study, was a clinical isolate from the collection of P. Ferrieri. *Escherichia coli* DH5α was used as the recipient for transformation in all cloning experiments. Two plasmid vectors were used for cloning: pUC18 and pGEX-4T-1 (Pharmacia, Uppsala, Sweden). Streptococcal strains were grown in Todd-Hewitt broth with 1% yeast extract (THY) or on THY with 1% agar containing sheep blood. *E. coli* strains were grown in Luria broth or Luria agar. *E. coli* transformants were selected on Luria agar with 100 μg of ampicillin per ml.

DNA techniques. Restriction enzymes and T4 DNA ligase were purchased from Pharmacia Inc. PCRs were carried out in PCR buffer containing 0.2 mM deoxynucleoside triphosphate, 2.5 mM $MgCl_2$, 50 mmol of each primer, and 2.5 U of *Taq* polymerase, obtained from Promega (Madison, Wis.). Plasmid DNAs were purified by using Wizard kits purchased from Promega. DNA and hybridizations were detected by DNA probes labeled with the Boehringer Mannheim Genius DNA labeling and detection kit. PCR conditions were 1 min at 94°C, 1 min at 45 to 55°C, and 1 to 4 min at 72°C for 30 cycles. DNA primers were synthesized by the Microchemical Facility at the University of Minnesota and by VNII Genetica, Moscow, Russia. Primers used for PCR and sequencing were ForW (bp 835 to 853), Z119bio (bp 4350 to 4328), scpForprom (bp 542 to 563), PFRevscp (bp 1142 to 1130), L1 (bp 2146 to 2162), L2 (bp 2776 to 2760), L3 (bp 2865 to 2883), L4 (bp 3042 to 3028), L5 (bp 3165 to 3180), L6 (bp 3753 to 3768), N1 (bp 1243 to 1260), and N2 (bp 2388 to 2361). The nucleotide positions of primers correspond to those of *scpA* (3). DNA sequencing reactions used Sequenase kits (U.S. Biochemicals, Cleveland, Ohio). Two different approaches were used for sequencing. Most of the sequence was determined by using GBS chromosomal DNA as a template for amplification of *scpB* with a set of PCR primers in which one was biotinylated. The PCR product was denatured, the biotinylated single strand was purified with Dynabeads (Dynal Inc., Oslo, Norway), and then the strand was used as a template in sequencing reactions. DNA sequencing reactions were carried out with DNA sequencing kits (U.S. Biomedical, Cleveland, Ohio). A portion of the sequence was determined from double-stranded sequencing reactions of *E. coli* plasmids containing fragments of the *scpB* gene (5, 23). The region between the primers scpFORprom and PFRevscp, which contains 300 bp of sequence 5′ of *scpB* and the first 300 bp of that gene, was cloned into pUC18 and subjected to double-stranded sequencing. The region between the N1 and N2 was sequenced by the Microchemical Facility Laboratory at the University of Minnesota Health Sciences Center with the DNA Sequenser 370A from Applied Biosystems, Forster, Calif.

Protein studies. A glutathione transferase-*scpB* gene fusion was constructed by ligation of the open reading frame (ORF) into the expression vector pGex-4T-1, supplied by Pharmacia Biotech Inc. The entire *scpB* gene sequence was amplified by using the primers FoRW and Z119, which contain restriction sites for *Bam*HI and *Eco*RI. The amplified product was purified by agarose gel electrophoresis and digested with *Bam*HI and *Eco*RI before it was ligated to pGEX-4T-1 vector DNA that was cut with the same enzymes. *E. coli* DH5α was transformed with ligated DNA by electroporation. Recombinant clones, carrying *scpB*, were identified by hybridization to the same *scpB* PCR fragment labeled with digoxigenin.

* Corresponding author. Mailing address: Department of Microbiology, Box 196 UMHC, University of Minnesota, Minneapolis, MN 55455. Phone: (612) 624-3932. Electronic mail address: cleary@lenti.med.umn.edu.

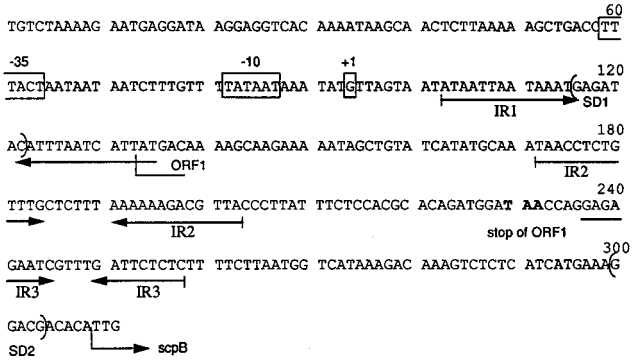

FIG. 1. Properties of the leader and promoter sequence of *scpB*. SD, Shine-Dalgarno site.

Protein from *E. coli* strains, harboring recombinant plasmids, was prepared according to the protocol provided by Pharmacia Inc. Isolation of SCPB from GBS was described previously (3). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot (immunoblot) analyses were performed on continuous 7% polyacrylamide gels as described by Laemmli (12). Proteins were transferred to nitrocellulose membranes as previously described (4). SCP antigens on blots were detected with SCPA12-specific hyperimmune rabbit serum, followed by goat anti-rabbit immunoglobulin G conjugated to alkaline phosphatase (Sigma, St. Louis, Mo.) (6).

Nucleotide sequence accession numbers. The *scpB* sequence was deposited into GenBank under accession number UL56908.

RESULTS

Our earlier study suggested that GBS harbors a peptidase gene which is similar to that of GAS (5, 23). The putative *scpB* gene was cloned and sequenced to further assess this similarity and to determine the relationship of the SCPB protein to the C5a-ase described by Bohnsack et al. (2, 3). DNA fragments were amplified by using primers that correspond to the *scpA* sequence (4). Chromosomal DNA extracted from strain 78-471 served as a template for PCR to produce a single-stranded biotinylated template for sequencing reactions. All primers, which were based on the *scpA* sequence, annealed sufficiently to amplify targeted sequences. Nearly 50% of the sequence was verified by sequencing of the opposite DNA strand. The sequence was read from at least three independent reactions. The longest ORF encodes a deduced protein of 1,150 amino acids, starts with a TTG codon, and terminates at a TAG stop codon which is followed by two additional stop codons within a stretch of 18 bp. This ORF is designated *scpB*. Sequence comparisons showed a high degree of similarity with C5a peptidase genes from GAS, 98% similarity to *scpA*12 from serotype M12 streptococci (4) and 97% similarity to *scpA*49 from a serotype M49 strain (16). A major difference between these DNA sequences was found in the 3′ portion of the *scpB* gene. *scpB* contained a 51-bp deletion, starting from nucleotide 3424. This region corresponds to the end of the first C repeat and the beginning of the second C repeat of *scpA* (4). As expected, the *Hin*dIII restriction site at nucleotide 1596 in *scpA* which was shown to be absent from *scpB* by restriction analysis (5) is lacking in the *scpB* sequence.

Promoter region. The 312 bp preceding the *scpB* coding sequence is also highly similar to that preceding *scpA* in the GAS chromosome (Fig. 1). The experimentally confirmed transcription start of *scpA*, a G at position 94 bp, and predicted −10 and −35 consensus sequences are conserved (18). In addition, three sets of inverted repeats—IR1, IR2, and IR3—were observed. IR1 and IR3 are conserved in both *scpA*12 and *scpA*49 sequences. IR3 terminates transcription within the *scpA*12 leader sequence and accounts for down regulation of that gene relative to *emm* genes (1, 18). The first IR contains a second putative ribosomal binding site and ATG codon which start a small ORF. This ORF is confined to the leader sequence and codes for a peptide of 32 amino acids. The role of this putative peptide in regulation of *scpB* expression is under investigation. Alignment of promoter regions of *scpB* and *scpA* revealed that the *scpB* promoter region has diverged more from the *scpA*49 than from the *scpA*12 promoters. Interestingly, the *scpA*12 promoter is more similar to *scpB* than it is to *scpA*49.

Amino acid comparison. Analysis of the deduced amino acid sequence of SCPB also demonstrated a high level of homology with SCPA (Fig. 2). The signal peptide sequence is nearly identical to that of SCPA; therefore, we presume that SCPB protein is processed in a manner similar to SCPA by cleavage between alanine 31 and arginine 32. The reactive serine 509, histidine 193, and aspartic acid 130, known to be a critical part of the reactive center of serine proteases, are conserved (4, 22). Alignment of the SCPB amino acid sequence with the two known SCPA sequences show significant similarity. SCPB differs from the SCPA12 protein by only 23 residues, in addition to a 17-amino-acid deletion. SCPA12 and SCPB have a common 3-amino-acid insertion near their carboxyl end and contain many common amino acid substitutions relative to SCPA49. A phylogenetic tree built by progressive alignment (8) demonstrated that SCPA12 is phylogenetically closer to SCPB than to SCPA49.

Characterization of recombinant SCPB. *scpB* was subcloned and expressed in order to compare it with C5a-ase produced by GBS cells. The ORF of *scpB* was cloned into the thrombin recognition site carried by the high-expression C-level vector pGEX-4T-1. A DNA fragment of 3.5 kb was PCR amplified by using the primers ForW and Z119. ForW anneals to *scpA* 2 bp before the ribosome binding site; therefore, the fusion includes the signal sequence. These primers carry the restriction sites for *Bam*HI and *Eco*RI, respectively. By utilizing these restriction sites, amplified DNA was cloned into the vector which was previously digested with the same restriction enzymes. *E. coli* transformants with the desired 3.5-kb insert were identified and analyzed for expression of the fusion protein. This purified protein was analyzed by Western blots which were probed with serum directed against the SCPA12 protein (Fig. 3). The serum reacted equally well with recombinant SCPB, native streptococcal SCPB, and SCPA12. SCPA and glutathione transferase-free SCPB appear to be approximately the same size, between 120 and 140 kDa. SDS-PAGE was previously shown to overestimate the size of streptococcal cell wall proteins. The aberrant migration of these proteins has been attributed to the proline-rich cell wall anchor region, near the carboxyl-terminal ends (4, 21). Although SCPB is smaller than SCPA by 17 amino acids, the cleaved fusion protein is expected to be larger because it retains the signal sequence of 31 amino acids. Native SCPB which was extracted from GBS cells with mutanolysin lacks the signal sequence and, therefore, migrates farther during SDS-PAGE.

DISCUSSION

Early events in the pathogenesis of GBS disease are poorly understood. The means by which streptococci colonize the vaginal mucosa without profound inflammation has not been investigated. Expression of capsule and/or the C-protein complex on the surface correlates with virulence, and antibody to either capsular polysaccharide or components of the C-protein complex is protective in animal studies (13, 14). Hill and col-

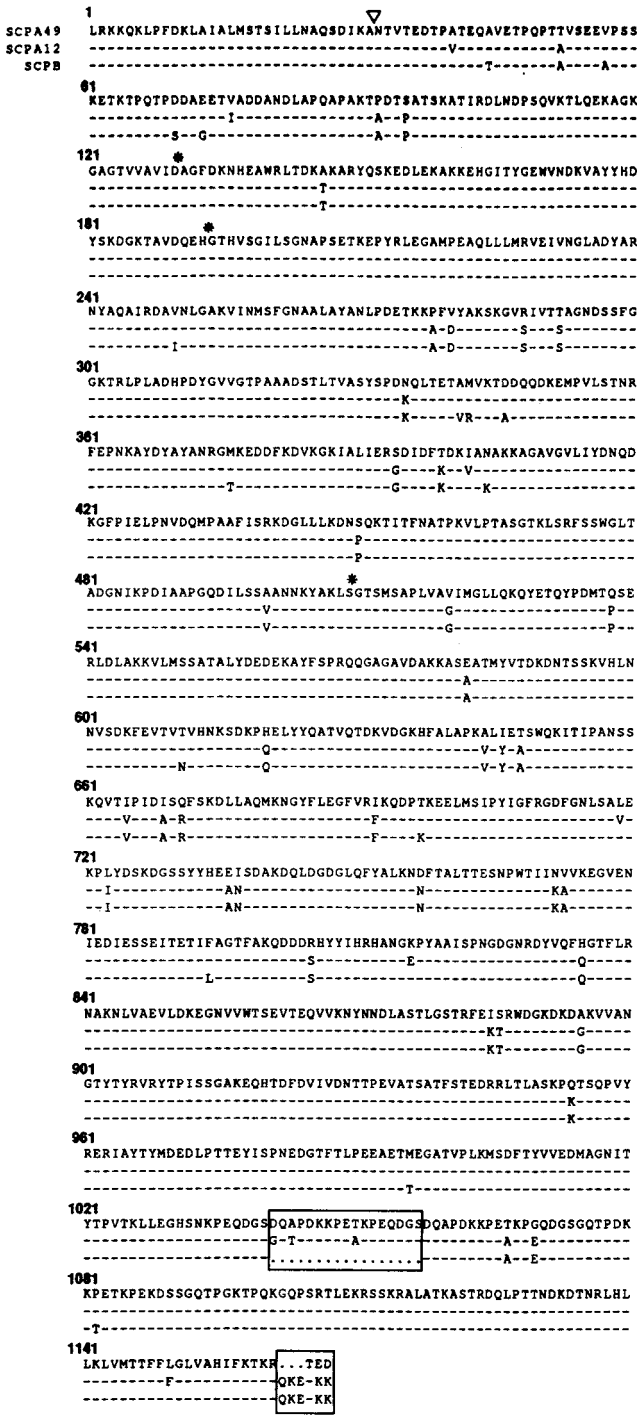

FIG. 2. Alignment of protein sequences of *scpB*, *scpA*12, and *scpA*49. Amino acids predicted to be in the enzyme's active site are marked by asterisks. Deletions in *scpB* are boxed. ▽, predicted cleavage point of the signal peptidase.

laborators discovered that GBS produces a protease which is functionally similar to the GAS C5a peptidase and suggested that it may contribute to the pathogenesis of GBS disease (2, 3, 9). The C5a-ase they described has an apparent M_r of 120,000 and hydrolyzes C5a at His-67 in the neutrophil binding site (2). We previously showed that the GBS chromosome contains a sequence which is similar to the *scpA* gene of GAS. Our experiments did not, however, address the relationship between the C5a-ase and the *scpB* gene product (5, 23). Purified C5a-ase does not bind to anti-SCPA which was prepared in rabbits immunized with a native group A streptococcal SCPA49 protein extracted from M49 streptococci (9). In contrast, rabbit serum which was directed against recombinant SCPA12, encoded by the *scpA* gene from serotype M12 streptococci, binds C5a-ase protein which was extracted from GBS (unpublished data). Experiments described here showed that recombinant SCPB protein and a high-molecular-weight protein extracted from GBS react with antiserum directed against anti-SCPA12 (Fig. 3). As predicted by our previous investigations (5, 23), the DNA and amino acid sequences that correspond to SCPB are highly homologous to those of SCPA. Sequence alignments may explain the failure of C5a-ase to bind antiserum directed at SCPA49 protein (9). Alignments revealed that SCPB and SCPA12 proteins are significantly more similar to each other than they are to SCPA49 protein. This argument assumes that C5a-ase and SCPA are the same protein. This conclusion, however, remains open to question without knowledge of the C5a-ase amino acid sequence. Moreover, SCPA (SCPB) is able to cleave mouse C5a (10). In contrast, mouse C5a was resistant to C5a-ase (1). Although it seems unlikely that GBS would produce two peptidases of the same size that cleave human C5a at the same site, the identity of SCPB and C5a-ase is still in doubt.

The SCPB sequence, like SCPA, contains amino acids which are predicted to contribute to charge transfer in the active center of the enzyme and characteristic of bacterial serine proteases (5, 22). The C-terminal half of SCPB contains sequences known to function in anchorage to the cell wall. The LPTTND motif which bonds to peptidoglycan is conserved at position 1213 (21).

The only significant difference between *scpB* and *scpA* is a deletion in the region encoding the C-terminus portion of the molecule. This 51-bp deletion located in the C-repeat region encodes 17 amino acids. Interestingly, this sequence is highly related to peptide sequences which contribute to proteolytic instability of eukaryotic proteins (19). Perhaps loss of this sequence stabilizes *scpB* on the surface of GBS, which in turn favors colonization in the protease-rich milieu of the colon and vagina.

The yield of extractable C5a peptidase from GBS is low relative to that from GAS (3). The near identity of the protein sequences and conservation of sequences around amino acids predicted to be associated with the catalytic center of the enzyme suggest that they should have equivalent specific ac-

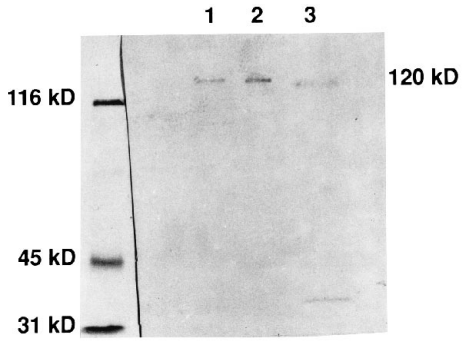

FIG. 3. Western blot analysis of the recombinant SCPB protein. Lane 1, purified SCPA12; lane 2, affinity-purified recombinant SCPB; lane 3, crude mutanolysin extract of GBS strain 78-471. Molecular mass markers are on the left.

tivity. It is not known whether *scpB* and *scpA* are equally transcribed and translated. The sequence preceding the *scpB* coding region, although less conserved than that of the structural gene, resembles the sequence 5′ of *scpA*. It is important to point out that the promoter region of *scpB* is more similar to the corresponding region of *scpA*12 than to that preceding *scpA*49. The location of the ATG codon in the stem of the first inverted repeat and the existence of a short ORF is analogous to the translational attenuator which regulates expression of erythromycin methylases (11). We hypothesize that translation of the short ORF, located in the *scpA* and *scpB* leader sequences, influences formation of the stem and loop formed by IR3 and thereby alters readthrough into the *scpB* structural gene. Deletion of IR3 alone increases the level of *scpA* transcription (18).

The first 200 bp 5′ of *scpA* in GAS contains the sequence upon which the *mga* gene product (formally known as VirR) acts to initiate transcription. A highly conserved sequence 5′ of the predicted promoter, beginning at position 426, was postulated to be an Mga binding site (17). Although this sequence is conserved in the GBS chromosome, there is no evidence that expression of *scpB* or other GBS surface protein genes is controlled by an Mga-like transcriptional activator.

The overall similarity in the *scpB* and *scpA* gene sequences (97%) is consistent with horizontal transmission of DNA between these two species of streptococci. Alternatively, group A and B streptococci evolved from a common ancestor which had a peptidase gene. Because 30 clinical strains representing serotypes III, III/R, II, Ia/c, NT/c, and NT/c/R1 were shown to carry a copy of *scpB* (reference 23 and unpublished data), transmission of the gene between these species is presumed to have preceded the evolution of serotypes in GBS. Group G streptococci also harbor a copy of the *scpA* gene, and one strain carries a nearly identical copy of the *emm12* gene (6). Together these observations suggest either that beta-hemolytic streptococci evolved from a common ancestor or that horizontal gene transfer is a relatively common event in nature.

Recent experiments conclusively showed that SCPA is essential for full virulence of GAS. Strains with a plasmid insertion or deletion internal to *scpA* are more rapidly cleared and fail to disseminate systemically in contrast to wild-type cultures in a connective-tissue air sac model in mice (10). By analogy, we propose that SCPB contributes to the virulence of GBS. We speculate that SCPB retards the inflammatory response during the first few hours of contact with the vaginal or oral mucosa. It may be less important in colonization and subsequent invasion of premature infants and compromised neonates. Experiments with newborn rats are now in progress to directly test the role of SCPB in virulence.

ACKNOWLEDGMENTS

We thank Tim Rast for technical assistance.

The project was supported by grant AI20016-11 (P.P.C.) and AI13926 (P.F.) from the National Institute of Allergy and Infectious Diseases.

Editor: V. A. Fischetti

REFERENCES

1. **Bohnsack, J. F., J.-K. Chang, and H. R. Hill.** 1993. Restricted ability of group B streptococcal C5a-ase to inactivate C5a prepared from different animal species. Infect. Immun. **61:**1421–1426.
2. **Bohnsack, J. F., K. W. Mollison, A. M. Buko, J. C. Ashworth, and R. H. Hill.** 1991. Group B streptococci inactivate C5a by enzymatic cleavage at the carboxy terminus. Biochem. J. **273:**635–640.
3. **Bohnsack, J. F., X. Zhou, P. A. Williams, P. P. Cleary, C. J. Parker, and H. R. Hill.** 1991. Purification of the proteinase from group B streptococci that inactivates human C5a. Biochim. Biophys. Acta **1079:**222–228.
4. **Chen, C. C., and P. P. Cleary.** 1989. Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus pyogenes. J. Biol. Chem. **265:**3161–3167.
5. **Cleary, P. P., J. Handley, A. N. Suvorov, A. Podbielski, and P. Ferrieri.** 1992. Similarity between the group B and A streptococcal C5a peptidase genes. Infect. Immun. **60:**4239–4244.
6. **Cleary, P. P., J. Peterson, C. Chen, and C. Nelson.** 1991. Virulent human strains of group G streptococci express a C5a peptidase enzyme similar to that produced by group A streptococci. Infect. Immun. **59:**2305–2310.
7. **Cleary, P. P., U. Prahbu, J. B. Dale, D. E. Wexler, and J. Handley.** 1992. Streptococcal C5a peptidase is a highly specific endopeptidase. Infect. Immun. **60:**5219–5223.
8. **Feng, D., and R. F. Doolittle.** 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. **25:**351–360.
9. **Hill, H. R., J. F. Bohnsack, E. Z. Morris, N. H. Augustine, C. J. Parker, P. P. Cleary, and J. T. Wu.** 1988. Group B streptococci inhibit the chemotactic activity of the fifth component of complement. J. Immunol. **141:**3551–3556.
10. **Ji, Y., L. McLandsborough, A. Kondagunta, and P. P. Cleary.** 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. Infect. Immun. **64:**503–510.
11. **Kamimiya, S., and B. Weisblum.** 1988. Translational attenuation control of *ermSF*, an inducible resistance determinant encoding rRNA *N*-methylase from *Streptomyces fradiae*. J. Bacteriol. **170:**1800–1811.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
13. **Martin, T. R., J. T. Ruzinski, C. E. Rubens, E. Chi, and C. B. Wilson.** 1992. The effect of type-specific polysaccharide capsule on the clearance of group B streptococci from the lung of infant and adult rats. J. Infect. Dis. **165:**306–314.
14. **Michel, J. L., L. C. Madoff, D. E. Kling, D. L. Kasper, and F. M. Ausubel.** 1991. Cloned alpha and beta C-protein antigens of group B streptococci elicit protective immunity. Infect. Immun. **59:**2023–2028.
15. **O'Connor, S. P., and P. Cleary.** 1987. In vivo Streptococcus pyogenes C5a peptidase activity: analysis using transposon- and nitrosoguanidine-induced mutants. J. Infect. Dis. **156:**495–504.
16. **Podbielski, A., A. Flosdorff, and J. Weber-Heynemann.** 1995. The group A streptococcal *virR49* gene controls expression of four structural *vir* regulon genes. Infect. Immun. **63:**9–20.
17. **Podbielski, A., J. A. Peterson, and P. P. Cleary.** 1992. Surface protein-CAT reporter fusions demonstrate differential gene expression in the vir regulon of Streptococcus pyogenes. Mol. Microbiol. **6:**2263–2265.
18. **Pritchard, K. H., and P. Cleary.** 1996. Expression of Streptococcus pyogenes scpA is down-regulated by an inverted repeat in the leader mRNA. Mol. Gen. Genet. **250:**207–213.
19. **Roger, S., R. Wells, and M. Rechsteiner.** 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science **234:**364–369.
20. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
21. **Schneewind, O., A. Fowler, and K. F. Faull.** 1995. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. Science **268:**103–106.
22. **Siezen, R. J., W. M. deVos, J. A. Leunissen, and B. W. Dijkstra.** 1991. Homology modeling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. Protein Eng. **4:**719–737.
23. **Suvorov, A. N., P. P. Cleary, and P. Ferrieri.** 1991. C5a peptidase gene from group B streptococci, p.230–232. *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.