

Antigenicity, Expression, and Molecular Characterization of Surface-Located Pullulanase of *Streptococcus pneumoniae*

ROGER J. M. BONGAERTS,† HANS-PETER HEINZ, ULRICH HADDING, AND GREGOR ZYSK*

Department of Medical Microbiology and Virology, University of Duesseldorf, Duesseldorf, Germany

Received 24 April 2000/Returned for modification 5 June 2000/Accepted 20 September 2000

A putative pullulanase-encoding gene from *Streptococcus pneumoniae* was identified by screening a genomic expression library with human convalescent-phase serum. The 3,864-bp gene encoded a 143-kDa protein. Surface location and pullulanase activity of the protein, designated SpuA, was demonstrated. SpuA was present in all investigated pneumococcal isolates of different serotypes. The *spuA* 5' end was highly conserved among clinical isolates except for a 75-bp region. The properties of SpuA reported here indicate that this novel immunogenic surface protein might have potential as a vaccine target.

Streptococcus pneumoniae is an important infectious agent of humans who are at the extremes of age or who are immunocompromised. It is responsible for invasive diseases like pneumonia, sepsis, and meningitis, as well as less severe but highly prevalent infections such as sinusitis and otitis media. Although capsular polysaccharides from *S. pneumoniae* have long been considered effective immunogens against pneumococcal infection, pneumococcal proteins are thought to form the basis of novel vaccines. To identify novel pneumococcal proteins that are immunogenic in humans we screened a genomic DNA expression library of *S. pneumoniae* using human convalescent-phase serum. Twenty-three clones encoding pneumococcal polypeptides were identified which had been immunogenic to the host during invasive pneumococcal disease or prior colonization (10). The sequences of the inserts of clones SP-21 and SP-74 were found to be highly homologous to a putative gene from the pneumococcal genome (contig 3836, nucleotides [nt] 140,383 to 136,542) published by The Institute for Genomic Research (<http://www.tigr.org>). The sequence of the complete gene from *S. pneumoniae* strain 3.B (clinical isolate, serotype 1) was analyzed and deposited in GenBank (of the National Center for Biotechnology Information) (see below). The gene was homologous to the 3' region of the *Bacillus* sp. gene encoding alkaline amylopullulanase (*apuA*), encoding the site with pullulanase activity (5). Therefore, the pneumococcal gene (3,864 nt) was designated *spuA* (for *S. pneumoniae* pullulanase A). Southern blot analysis demonstrated that *spuA* is a single-copy gene (data not shown), and comparison with the pneumococcal genome (<http://www.tigr.org>) indicates that the mRNA is monocistronic. *spuA* encoded a protein of 1,287 amino acid residues (Fig. 1), with a predicted molecular mass of 143,317 Da. The calculated molecular mass correlates with the molecular mass of the main protein band determined by Western blot analysis (~140 kDa) using a specific polyclonal rabbit serum raised against the N-terminal fragment (amino acids [aa] 118 to 542) of SpuA. In addition, Western blot analysis demonstrated the continuous presence of SpuA during all growth phases (Fig. 2).

For prediction of protein localization and motif identifica-

tion, the PSORT WWW server (<http://psort.nibb.ac.jp>), the Motif WWW server (<http://motif.genome.ad.jp>), and the Pfam search engine at the Sanger Institute (<http://www.sanger.ac.uk/Software/Pfam>) (2, 9) were used. Analysis of the *spuA* deduced amino acid sequence revealed, at amino acids 670 to 674, a YNWGY motif that is common to pullulanases (1). In addition, four regions (I, II, III, and IV), that are conserved in amylolytic enzymes and form a catalytic domain (aa 717 to 723, aa 781 to 789, aa 812 to 817, and aa 897 to 903) with the putative catalytic triad Asp₇₈₅-Glu₈₁₄-Asp₉₀₂ (Fig. 1) were identified. Moreover, tryptophan residue W₈₁₆ and tyrosine residue Y₈₉₇, in regions III and IV, respectively, are specific for pullulanases.

Functional analysis was performed by electrophoresis of pneumococcal extract in native 6% polyacrylamide gels at 4°C. For analysis of pullulanase activity, gels containing 0.15% Red-pullulan (Gamma Chemie, Munich, Germany) were incubated in buffer A (0.1 M NaAc, 1 mM dithiothreitol [pH 6.0]) for 3 h at 37°C. For analysis of α -amylase activity, gels containing 1% soluble starch were stained with I₂ and Lugol's iodine solution as described by Brooker and McCarthy (3) and Satoh et al. (7), respectively. Functional analysis confirmed the results from the deduced amino acid analysis and showed that a protein with pullulanase activity corresponded to a protein recognized by anti-SpuA (Fig. 3). Immunoblotting of the supernatant revealed a very weak but visible band as well. As we did

TABLE 1. Grouping of investigated pneumococcal strains^a

Serotype	No. of strains with <i>spuA</i> of sequence type:		
	I	II	III
1	4	2	1
2		1	
4			1
6A		1	
6B	2		
7F		3	
9V			1
14	3		
18C	1		
18F		1	
19F		1	1
23F		1	

^a Grouping is based upon the 5' region of *spuA* showing variation within a 75-bp region. The strains are divided into three sequence types based upon the presence of all or the absence of 21 or 66 nucleotides.

* Corresponding author. Mailing address: Department of Medical Microbiology and Virology, University of Duesseldorf, Geb. 22.21/U1, Universitaetsstrasse 1, D-40225 Duesseldorf, Germany. Phone and fax: 49-211-811 24 96. E-mail: zysk@uni-duesseldorf.de.

† Present address: Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, United Kingdom.

1 MRKTPSHTEK KMVYSIRSLK **NG** **GSV**LIGA SLVLLAMATP TISSDESTPT TNEPNNRNT**II** TLAQPLTDTA AGSGKNE**SDI** SSPGNANA**SL** EKTEKPA**TE** 100
A **B**
101 **PTTPA**ASPAD **PAPQTG**DRS **SEPTT**STSPV **TTETK**AEEPI EDNYFRIHVK **KLPEENK**DAQ **GLNTW**DDEVER **PSENW**PNGAL **SFKDAK**KDDY **GYLDV**KLKG 200
201 **EQAKKIS**FLI **NNAG**KNL**IG** **DKSVE**KLVPK MNEAWL**DQDY** **KVFSY**EPQPA **GTVRV**NYIRT **DGNYD**KKSLW **YWGDK**VPN**SS** **AQWP**DGTDFT **ATGKY**GRYID 300
FN type III
301 **IPLNEA**AREF **GFLLLD**ESKQ **GDDV**KIRKEN **YKFTD**LKNHS **QIFL**KDDDES **IYTN**PYYVHD **IRMTG**AQHV**G** **TSSIE**SSFST **LVGAK**KEDIL **KHSNI**TNHLG 400
401 **NKVTIT**DVAI **DEAGK**KV**YS** **GDFS**DTKHPY **TVSYN**S**DQFT** **TRKTS**WRLKDE **TYSYD**GK**LGA** **DLKEE**GK**QVD** **LTLWS**PSADK **VSVVV**YDKND **PDKVV**GTVAL 500
501 **EKGERG**TW**Q** **TLDST**NK**LGI** **TDFTG**YYY**QY** **QIERO**GK**TVL** **ALDPY**AKSLA **AWNSD**DAKID **DAHKV**A**AAF** **VDPAK**LGP**QD** **LTYGK**IHN**FK** **TREDA**VIYEA 600
601 **HVRDFT**SDPA **IAKDL**TK**PF**G **TFEAF**IEKLD **YLKDL**GV**THI** **QLLPV**LSY**YF** **VNELN**KNEHL **SDYAS**S**SNY** **NWGYD**PQ**NYF** **SLTGM**YSS**DP** **KNPEK**RIAEF 700
701 **KNLINE**IH**KR** **GMGAI**I**DVVY** **NHAK**V**DIFE** **DLEPN**Y**YHF**M **DADGT**P**RTSF** **GGRLG**TT**HH** **MTKRL**L**VDSI** **KYLVD**TY**KVD** **GFRF**DM**MGD**H **DAASIE**EAYK 800
801 **AARALN**PNLI **MLGEG**WR**TYA** **GDENM**PT**KAA** **DQDM**K**H**TD**T** **VAVFS**DDIRN **NLKS**GYP**NEG** **QPAFI**T**G**G**KR** **DVNTI**F**KNLI** **AQPTN**FE**ADS** **PGDVI**Q**YTAA** 900
901 **HDNI**L**LF**DII **AQSIK**K**DPSK** **AENYA**E**I**HRR **LRLGN**LM**VL**T **AQGT**PF**I**HS**G** **QEYGR**TK**QFR** **NPAYR**TP**VAE** **DKVPN**K**S**HL**L** **RDKDG**NP**FDY** **PYFIH**DSY**S** 1000
1001 **SDAVN**K**FDWT** **KATDG**K**AYPE** **NVKS**RD**Y**M**K**G **LIALR**Q**STDA** **FRLK**SL**QDIK** **DRVHL**IT**VP**G **QNGVE**KED**VV** **IGYQI**TAP**NG** **DIYAV**F**V**NAD **EKARE**FN**LGT** 1100
1101 **AF AHL**R**NAEV** **LADEN**Q**AGSV** **GIANP**K**GLEW** **TEKGL**KL**NAL** **TATV**LR**V**S**Q**N **GH**SE**STAE**E **KPDST**PK**PE** **HQNEA**SH**PAH** **QDPA**E**ARP**D **STKPD**AK**Q**AD 1200
1201 **AENK**PS**Q**ATA **DSQA**E**Q**PA**QE** **AQASS**V**K**EA**V** **RKES**V**ENS**SK **ENIS**A**T**PD**RQ** **AELP**NT**G**IK**N** **ENKLL**F**AGIS** **LLALL**GL**GFL** **LKNK**KEN 1287

FIG. 1. The deduced amino acid sequence of *spuA* (GenBank accession no. AF217414). The N-terminal part of SpuA encoded by inserts of clones SP-21 and SP-74 is depicted in bold. The four conserved regions, designated I, II, III, and IV, which form the catalytic domain are boxed. The amino acids marked with asterisks in boxes II, III, and IV form the putative catalytic triad Asp₇₈₅-Glu₈₁₄-Asp₉₀₂. A YNWGY sequence motif is marked with a gray box. An LPXTG motif, typical for cell wall anchoring in gram-positive cocci, is overlined. A boxed fibronectin type III repeat signature is named FN type III, and two 18-amino-acid repeats are underlined. Possible O-glycosylated amino acid residues are indicated in white on a gray background. The 7 boxed amino acid residues annotated with A are absent in *spuA* sequence type III. The 22 boxed amino acid residues annotated with B are absent in *spuA* sequence type II.

not solubilize the peptidoglycan, the amount of soluble SpuA is probably primarily due to autolysis of pneumococci. No corresponding α-amylase activity could be detected in the pullulanase-containing extracts (data not shown). However, it cannot be completely ruled out that SpuA contains such activity as, e.g., properties of amyolytic enzymes vary and can be linked to environmental circumstances.

The deduced amino acid sequence of *spuA* revealed the

presence of a fibronectin type III repeat signature, two 18-amino-acid repeats, and several potential O glycosylation sites, whose functions in SpuA are presently unknown (Fig. 1). Although no clear signal sequence was present, the LPXTG motif at the carboxy terminus (Fig. 1) represents a typical sequence of gram-positive bacteria, indicating that the protein is attached to the cell wall (4, 6, 8). The LPXTG motif was preceded by a typical proline-rich segment, thought to span the cell wall, and was followed by a hydrophobic transmembrane domain with several positively charged amino acids at its tail (Fig. 1). Immunocytometric analysis and immunofluorescence microscopy with anti-SpuA confirmed the surface localization of SpuA (data not shown). Considering the cellular localization of SpuA, SpuA might play a role in colonization and may solubilize mucus or affect exposition of host glycoconjugates

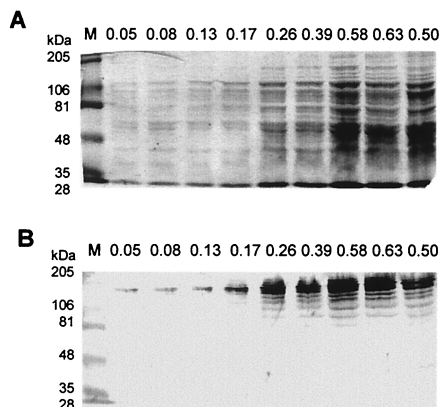


FIG. 2. Expression of SpuA during different growth phases. (A) Polyacrylamide gel electrophoresis of whole-cell extracts from *S. pneumoniae* strain 3.B harvested from the same culture as the optical density at 621 nm increased. The gel was Coomassie brilliant blue stained. (B) Western blot analysis of whole-cell extracts (the same as those used for polyacrylamide gel electrophoresis) with anti-SpuA. Lanes M, molecular mass standard sizes shown at left; all other lanes are labeled with optical density values of the culture.

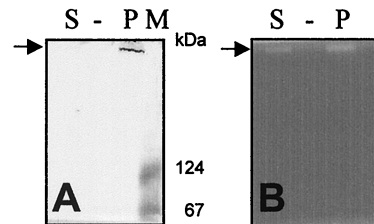


FIG. 3. SpuA functional analysis. The same polyacrylamide gel in which the *S. pneumoniae* strain 3.B native whole-cell extract (P) and the culture supernatant (S) were separated was used for Western blot analysis (A) and detection of pullulanase activity (B). The protein band recognized by anti-SpuA and possessing pullulanase activity is indicated by arrows. Lane M, molecular mass standard (sizes shown at right).

```

210                                                                                               309
I AGCTGGCTCTGGTAAGAACGAAAGTGATATTTCTTACCTGGAAATGCAAACGCTTCCCTAGAGAAAACAGAAGAAAAACCTGCAACAGAGCCAACT
II AGCTGGCTCTGGTAAGAACGAAAGTGATATTTCTTACCTAGAAATGCAAACGCTTCCCTAGAGAAAACAGAAGAAAAACCTGCAACA-----
III AGCTGGCTCTGGTAAGAACGAAAGTGATATTTCTTACCTGGAAATGCAAACGCTTCCCTAGAGAAAACAGAAGAAAAACCTGC-----
310                                                                                               409
I CCTGTGCAAGCCCAGCCGATCCAGCACCACAACTGGACAAGATCGTTCAGTGAGCCAACTACTTCTACTAGTCCAGTAACAACCTGAAACTAAGGCAG
II -----GAGCCAACTACTTCTACTAGCCAGTAACAACCTGAAACTAAGGCAG
III --TGTGCAAGCCCAGCCGATCCAGCACCACAACTGGACAAGATCGTTCAGTGAGCCAACTACTTCTACTAGTCCAGTAACAACCTGAAACTAAGGCAG

```

FIG. 4. Nucleotide sequence comparison of the region from nucleotide 210 to nucleotide 409 of *spuA*. Three sequence types (types I, II, and III) could be distinguished in the 5' end of the *spuA* gene (nucleotides 72 to 834) from 25 pneumococcal strains representing 12 distinct serotypes. The nucleotides that differed between the three types are depicted in bold. Short sequence repeats are underlined. The absence of nucleotides is depicted by a minus.

for the pneumococcus by its α -1,6-glucanohydrolase activity. On the other hand, SpuA might contribute to virulence or survival of pneumococci in the host by scavenging potential carbon sources.

Western blot analysis with native extracts from 41 pneumococcal strains of 17 distinct serotypes with anti-SpuA showed that SpuA was recognized in all extracts, indicating that SpuA is serologically highly conserved. Western blot analysis using native extracts of *Escherichia coli*, *Klebsiella oxytoca*, *Staphylococcus aureus*, *Moraxella catarrhalis*, *Bacillus subtilis*, *Enterococcus faecalis*, and *Streptococcus bovis* showed no immunoreactive band (data not shown).

DNA sequence analysis of the 5' end of the *spuA* genes of 25 distinct pneumococcal strains representing 12 serotypes demonstrated limited variation. However, one 75-bp stretch differed among pneumococcal strains. Based upon the presence of all nucleotides and absence of 21 or 66 nucleotides the strains could be divided into three sequence types (Table 1 and Fig. 4). The differences include the missense of 7 or 22 amino acids, with an overlap of 2 amino acids (Fig. 1). The pneumococcal strains of serotypes 1 ($n = 7$) and 19F ($n = 2$) are distributed among the three and among two *spuA* sequence types, respectively. In contrast, the strains of serotypes 6B ($n = 2$), 7F ($n = 3$), and 14 ($n = 3$) each group within one sequence type. In addition, the 5' *spuA* gene fragments of the serotype 4 pneumococcal strain and the serotype 4 type strain used for genome sequencing (<http://www.tigr.org>) were identical. Furthermore, the 5' *spuA* sequence of strain D39 and its unencapsulated derivative, R6x, were identical.

The four regions forming the catalytic domain of amylolytic enzymes, like SpuA, are strongly conserved and therefore highly homologous between microorganisms and mammals. In that respect this region of SpuA is not suitable for vaccine development. In contrast, the N-terminal half of SpuA shows interesting characteristics, as this region is immunogenic, located at the surface of the pneumococcus, and species specific. Furthermore, the 5' region of *spuA* is highly conserved, except for a 75-bp region, and SpuA is present in all investigated pneumococcal strains. Inactivation of SpuA and virulence studies will enable us to understand its role in the pathogenicity of the pneumococcus.

Nucleotide sequence accession number. The sequence of the complete *spuA* gene for *S. pneumoniae* strain 3.B was deposited in GenBank under accession number AF217414.

We thank S. Scheuring and K. Köhrer, BMFZ, Düsseldorf, for carrying out DNA sequencing. R. R. Reinert, National Reference Center for Streptococci, Aachen, Germany, is acknowledged for serotyping of pneumococcal clinical isolates. W. Schwippert and V. Werner are kindly acknowledged for expert technical assistance.

REFERENCES

1. Albertson, G. D., R. H. McHale, M. D. Gibbs, and P. L. Bergquist. 1997. Cloning and sequence of a type I pullulanase from an extremely thermophilic anaerobic bacterium, *Caldicellulosiruptor saccharolyticus*. *Biochim. Biophys. Acta* **1354**:35–39.
2. Bateman, A., E. Birney, R. Durbin, S. R. Eddy, R. D. Finn, and E. L. Sonnhammer. 1999. Pfam 3.1: 1313 multiple alignments match the majority of proteins. *Nucleic Acids Res.* **27**:260–262.
3. Brooker, J. D., and J. M. McCarthy. 1997. Gene knockout of the intracellular amylase gene by homologous recombination in *Streptococcus bovis*. *Curr. Microbiol.* **35**:133–138.
4. Fischetti, V. A., V. Pancholi, and O. Schneewind. 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. *Mol. Microbiol.* **4**:1603–1605.
5. Hatada, Y., K. Igarashi, K. Ozaki, K. Ara, J. Hitomi, T. Kobayashi, S. Kawai, T. Watabe, and S. Ito. 1996. Amino acid sequence and molecular structure of an alkaline amylopullulanase from *Bacillus* that hydrolyzes α -1,4 and α -1,6 linkages in polysaccharides at different active sites. *J. Biol. Chem.* **271**:24075–24083.
6. Pancholi, V., and V. A. Fischetti. 1989. Identification of an endogenous membrane anchor-cleaving enzyme for group A streptococcal M protein. Its implication for the attachment of surface proteins in gram-positive bacteria. *J. Exp. Med.* **170**:2119–2133.
7. Satoh, E., Y. Ito, Y. Sasaki, and T. Sasaki. 1997. Application of the extracellular α -amylase gene from *Streptococcus bovis* 148 to construction of a secretion vector for yogurt starter strains. *Appl. Environ. Microbiol.* **63**:4593–4596.
8. 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.
9. Sonnhammer, E. L. L., S. R. Eddy, E. Birney, A. Bateman, and R. Durbin. 1998. Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* **26**:320–322.
10. Zysk, G., R. J. M. Bongaerts, E. ten Thoren, G. Bethe, R. Hakenbeck, and H.-P. Heinz. 2000. Detection of 23 immunogenic pneumococcal proteins using convalescent-phase serum. *Infect. Immun.* **68**:3740–3743.