Detection of 23 Immunogenic Pneumococcal Proteins Using Convalescent-Phase Serum

GREGOR ZYSK,^{1*} ROGER J. M. BONGAERTS,¹ ELKE TEN THOREN,¹ GESINA BETHE,¹
REGINE HAKENBECK,² AND HANS-PETER HEINZ¹

Department of Medical Microbiology and Virology, University of Duesseldorf, Duesseldorf,¹ and
Department of Microbiology, University of Kaiserslautern, Kaiserslautern,² Germany

Received 29 November 1999/Accepted 29 February 2000

A genomic expression library of *Streptococcus pneumoniae* was screened with a convalescent-phase serum for immunoreactive proteins. Six known and 17 unknown pneumococcal proteins were detected. Five of the known proteins were surface-located virulence factors, and eight of the unknown proteins were putative membrane proteins.

Streptococcus pneumoniae is the major cause of otitis media, pneumonia, and meningitis. Recent studies have presented new insights into the pathogenesis of pneumococcal infection (17). In addition to the antiphagocytic polysaccharide capsule, several proteins have been described as virulence factors for *S. pneumoniae*, most of which are located on the pneumococcal cell surface and contribute to colonization, adherence, and invasion during infection of various animal models (12, 13). We constructed a genomic expression library of an *S. pneumoniae* strain (strain 3.B, serotype 1; collection of the Department of Medical Microbiology and Virology, University of Duesseldorf) which was isolated from a patient suffering from disseminated pneumococcal infection. Here we describe the screening of this library with a serum which was taken from the patient during the convalescent phase, 26 days after diagnosis.

A number of pneumococcal proteins were highly reactive with immunoglobulin G (IgG) antibodies of the patient serum. Immunoblot analysis of pneumococcal cell lysates showed immunodominant bands from 60 up to 130 kDa (data not shown). To identify these immunoreactive proteins, an expression library of the pneumococcal genome was created in *Escherichia coli*. Pneumococcal DNA was digested partially by *Sau3A*, and fragments of 500 to 1,600 bp were ligated with the expression vector pET15b (Novagen) and used for transformation of *E. coli*. About 61,000 recombinants were screened for expression of proteins which were reactive with IgG antibodies of the convalescent-phase serum by colony immunoblot analysis. Seventy-eight recombinant proteins were detected. Western blot analysis of whole-cell lysates confirmed expression of immunoreactive proteins in 56 clones, from which plasmid DNA was isolated and the pneumococcal DNA was sequenced. The DNA sequences were used for a similarity search. The nucleotide sequence of the identified gene was completed by the sequence data obtained from the unfinished pneumococcal genome (available on the website of The Institute for Genomic Research [TIGR] [<http://www.tigr.org>]).

Identification of known pneumococcal proteins. In total, 20 clones contained pneumococcal DNA of previously identified genes. They contained eight different DNA inserts from six genes (Fig. 1). Five of the deduced proteins were surface lo-

cated and one (DnaK) was presumably of cytoplasmic origin. Parts of the pneumococcal surface protein A (PspA) were expressed in three recombinants. PspA, a choline-binding protein, has been characterized as a lactoferrin-binding protein and was shown to be protective in immunization studies (6, 10). The protein is composed of a variable N-terminal domain followed by an 82-amino-acid proline-rich region and a C-terminal choline-binding domain composed of 10 highly conserved 20-amino-acid repeats (18). All three clones contained DNA coding for the proline-rich domain of PspA. The clone SP-52 expressed the N-terminal part of the protein, including the first 58 amino acids of the proline-rich domain. SP-11 and SP-30 encoded the C-terminal third of the proline-rich domain and parts of the flanking choline-binding domain. IgG antibodies of the convalescent-phase serum showed high reactivity with the overexpressed protein of clone SP-52 and showed low binding with the recombinant proteins SP-11 and SP-30. This is in agreement with previous results suggesting that the proline-rich domain has an important function in immunogenicity (3). Three recombinants contained DNA coding for the C-terminal part of the proline-rich and repetitive domain of the *S. pneumoniae* surface protein A (SpsA), another choline-binding protein. It has been shown that SpsA binds the secretory component of IgA and plays a role in pneumococcal adhesion to eukaryotic cells (7, 15). Another seven recombinants contained the 5' part of the neuraminidase A (*nanA*) gene. NanA has previously been identified as a surface protein on the basis of its C-terminal LPXTG motif. It has been assumed that the enzyme cleaves the terminal sialic acids on cell surface glycolipids, thereby exposing the carbohydrate ligand by which *S. pneumoniae* attaches to host cells (for review see reference 11). The recombinant SP-14 expressed the central part of β -N-acetylhexosaminidase (StrH), an enzyme predicted to be anchored in the bacterial membrane by its C-terminal LPXTG motif. The protein is one of six pneumococcal glycosidases. It has been assumed that StrH plays a role in pathogenesis since GlcNAc β 1-linked residues are common components of several cell surface molecules on host tissue cells (4). Five clones expressed parts of the AliB protein, a membrane-bound lipoprotein and part of a multicomponent oligopeptide transporter. The protein has been suggested to play a role in signal transduction and genetic competence (1, 2). The pneumococcal DNA cloned into SP-50 coded for the central region of AliB, and the four recombinants SP-18, -59, -67, and -68 expressed the N-terminal region of the protein. The N-terminal region of

* 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: 49-211-8112496. E-mail: zysk@uni-duesseldorf.de.

TABLE 1. Immunogenic proteins of putative or unknown function^a

Class	SP strain(s)	Length of open reading frame (bp)	PSORT analysis finding	Protein exhibiting similarity	% Amino acid identity (% similarity)	Species	Immunoblotting result (denaturing/native) ^d
Carbohydrate metabolism	70 3, 77 21, 74 22, 23, 35, 55	759 6,702 3,735 4,200	Membrane Membrane Membrane Membrane	Triosephosphate isomerase β-Galactosidase Amylopullulanase Endo-β-N-acetylglucosaminidase	66 (78) 23 (40) 52 (67) 32 (52)	<i>Lactobacillus delbrueckii</i> <i>Thermotogaobacter ethanolicus</i> <i>Bacillus</i> sp. <i>Arthrobaacter protophormiae</i>	x/+ +++/ +++/ +++/ +++
Other metabolic functions	44, 51 54 9 33, 38	468 1,341 804 1,344	Not clear Cytoplasm Not clear Cytoplasm	Lumazine synthase (riboflavin synthase beta subunit) Exodeoxyribonuclease VII large subunit Putative cyclophilin-related prolyl- <i>cis</i> <i>trans</i> isomerase Probable RNA helicase	68 (82) 32 (52) 40 (54) 48 (69)	<i>Actinobacillus pleuropneumoniae</i> <i>Bacillus subtilis</i> <i>Schizosaccharomyces pombe</i> <i>Bacillus subtilis</i>	+/x +/x +/+ x/+++
Transporter	12 69 75	2,166 1,365 1,350	Membrane Cytoplasm Membrane	ABC transporter ABC transporter Sodium-dependent transporter	33 (54) 66 (84) 47 (61)	<i>Bacillus subtilis</i> <i>Bacillus subtilis</i> <i>Bacillus subtilis</i>	+++/ x/+ +/+++
Other	6, 10, 13, 15, 32, 40, 43, 57 45 24, 41, 71, 34, 58 48 72 5, 26	6,435 — ^b 2,139 2,499 — 2,166	Membrane — Cytoplasm Membrane — Cytoplasm	Cell wall-associated serine proteinase Iron-regulated lipoprotein precursor Unknown protein Unknown protein Unknown protein	25 (39) 56 (73) ^c 39 (50) 34 (46)	<i>Lactococcus lactis</i> subsp. <i>cremoris</i> <i>Corynebacterium diptheriae</i> <i>Streptococcus agalactiae</i> <i>Streptococcus agalactiae</i> No homology No homology	+++/ +++/ +++/ +++/ +++/ x/+++

^a The DNA sequences obtained from the inserted fragments of immunoreactive recombinants were used to search the TIGR database of unfinished genomes. The deduced amino acid sequences of the entire genes were then used for identification of protein homologs. The intensity of IgG binding to the recombinant proteins on immunoblots as observed under standard conditions is indicated.
^b The entire open reading frame could not be obtained from the unfinished sequence data (TIGR website). Therefore, PSORT analysis was not done.
^c Homology was calculated based on the first 143 amino acids, which were taken from the unfinished sequence data (TIGR website).
^d If multiple immunoreactive protein bands were visible the main band was scored. x, no reactivity; +, and +++, weak, medium, and strong reactivities, respectively.

FIG. 1. Immunogenic pneumococcal proteins of known function. The parts of the proteins which were encoded by the cloned DNA inserts of the expression library are hatched. The intensity of IgG binding to the expressed proteins on the Western blot using whole-cell lysates of *E. coli* recombinants is indicated. x, no reactivity; +, ++, and +++, weak, medium, and strong reactivities, respectively; N, native conditions; D, denaturing conditions.

the heat shock protein DnaK (hsp70 family) was identified in SP-39. DnaK is one of three pneumococcal heat shock proteins. Immunoblot analysis with mouse antipneumococcal sera has revealed that DnaK is a major immunogen (5, 9). According to the PSORT analysis of the deduced amino acid sequence, DnaK appears to be a cytoplasmic protein, in contrast to the other proteins described above.

Identification of unknown pneumococcal proteins. Twenty distinct inserts corresponding to 17 genes not previously described were identified in 36 recombinant plasmids, all of which expressed IgG-reactive proteins. The entire nucleotide sequence and the corresponding amino acid sequence for 15 of the 17 detected unknown pneumococcal genes could be deduced by comparison with the preliminary sequence data of the pneumococcal genome (TIGR website). PSORT analysis of the proteins was used for the prediction of their cellular localization. In addition, these amino acid sequences were subjected to a similarity search through the National Center for Biotechnology Information database, using the BLAST programs (Table 1). Four of these proteins showed similarity with enzymes involved in carbohydrate metabolism, all of which appeared to be membrane bound. Three proteins, which were homologs of β -galactosidase, endo- β -N-acetylglucosaminidase, and amylopullulanase, showed strong IgG binding using the convalescent-phase serum. Four deduced proteins were also homologs of metabolic enzymes and probably located in the cytoplasm. Another three were homologs to ATP-binding cassette (ABC) transporters, whereby the two membrane-localized proteins were more immunogenic than the cytoplasmic SP-69 homolog. Among the remaining proteins, a homolog to a cell wall-associated serine proteinase was detected. Also, a putative RNA helicase (a member of the DEAD protein family) revealed strong binding to IgG under non-denaturing conditions. Remarkably, this protein showed 61% amino acid similarity to the autoaggregation protein AggH of *Lactobacillus reuteri* (14). In addition, one deduced protein was highly related to DtxR, an iron-dependent regulatory DNA binding protein of *Corynebacterium diphtheriae* (16). Finally, four proteins were identified in this screen which showed no homology to other known bacterial proteins.

In summary, we detected by the described library screening 13 pneumococcal proteins which have features of mem-

brane proteins; 8 were from a group of 17 unknown proteins, and 5 were from a group of 6 known proteins. PSORT analysis indicated only five of the unknown proteins and DnaK to be of cytoplasmic origin. Nevertheless, cytoplasmic proteins had been immunogenic during infection or preceding colonization as well. Because cytoplasmic proteins are released into the pneumococcal surroundings during autolysis of the bacteria, these proteins, besides surface-associated and secreted pneumococcal proteins, can play an important role in the pathogenesis of infection and can cause seroconversion in humans. This has been demonstrated for the cytosolic pneumococcal virulence factor pneumolysin (8). This study identified major pneumococcal virulence factors, indicating that the experimental approach described herein reveals proteins important during the pathogenesis of pneumococcal infection. Further detailed study of the unknown immunogenic pneumococcal proteins detected here will answer the question of whether some of these proteins contribute to the virulence of *S. pneumoniae*.

We acknowledge Karl Köhrer and Sibylle Scheuring, BMFZ, University of Duesseldorf, for DNA sequencing. We thank Wilfried Schwiipert and Gabriele Zysk for excellent technical assistance.

REFERENCES

- Alloing, G., C. Granadel, D. A. Morrison, and J. P. Claverys. 1996. Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. *Mol. Microbiol.* **21**:471-478.
- Alloing, G., P. de Philip, and J. P. Claverys. 1994. Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of the gram-positive *Streptococcus pneumoniae*. *J. Mol. Biol.* **241**:44-58.
- Brady, L. J., D. G. Cvitkovitch, C. M. Geric, M. N. Addison, J. C. Joyce, P. J. Crowley, and A. S. Bleiweis. 1998. Deletion of the central proline-rich repeat domain results in altered antigenicity and lack of surface expression of the *Streptococcus mutans* P1 adhesin molecule. *Infect. Immun.* **66**:4274-4282.
- Clarke, V. A., N. Platt, and T. D. Butters. 1995. Cloning and expression of the β -N-acetylglucosaminidase gene from *Streptococcus pneumoniae*. *J. Biol. Chem.* **270**:8805-8814.
- Hamel, J., D. Martin, and B. B. Brodeur. 1997. Heat shock response of *Streptococcus pneumoniae*: identification of immunoreactive stress proteins. *Microb. Pathog.* **23**:11-21.
- Hammerschmidt, S., G. Bethe, P. H. Remane, and G. S. Chhatwal. 1999. Identification of pneumococcal surface protein A as a lactoferrin-binding protein of *Streptococcus pneumoniae*. *Infect. Immun.* **67**:1683-1687.
- Hammerschmidt, S., S. R. Talay, P. Brandtzaeg, and G. S. Chhatwal. 1997. SpsA, a novel pneumococcal surface protein with specific binding to secre-

- tory immunoglobulin A and secretory component. *Mol. Microbiol.* **25**:1113–1124.
8. **Kanclerski, K., S. Blomquist, M. Granström, and R. Möllby.** 1988. Serum antibodies to pneumolysin in patients with pneumonia. *J. Clin. Microbiol.* **26**: 96–100.
 9. **Kim, S. W., I. H. Choi, S. N. Kim, Y. H. Kim, S. N. Pyo, and D. K. Rhee.** 1998. Molecular cloning, expression, and characterization of *dnaK* in *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* **161**:217–224.
 10. **McDaniel, L. S., J. S. Sheffield, P. Delucchi, and D. E. Briles.** 1991. PspA, a surface protein of *Streptococcus pneumoniae*, is capable of eliciting protection against pneumococci of more than one capsular type. *Infect. Immun.* **59**:222–228.
 11. **Navarre, W. W., and O. Schneewind.** 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* **63**:174–229.
 12. **Paton, J. C.** 1998. Novel pneumococcal surface proteins: role in virulence and vaccine potential. *Trends Microbiol.* **6**:85–87.
 13. **Paton, J. C., A. M. Berry, and R. A. Lock.** 1997. Molecular analysis of putative pneumococcal virulence proteins. *Microb. Drug Resist.* **3**:1–10.
 14. **Roos, S., S. Lindgren, and H. Jonsson.** 1999. Autoaggregation of *Lactobacillus reuteri* is mediated by a putative DEAD-box helicase. *Mol. Microbiol.* **32**:427–436.
 15. **Rosenow, C., P. Ryan, J. N. Weiser, S. Johnson, P. Fontan, A. Ortqvist, and H. R. Masure.** 1997. Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol. Microbiol.* **25**:819–829.
 16. **Schmitt, M. P., and R. K. Holmes.** 1994. Cloning, sequence, and footprint analysis of two promoter/operators from *Corynebacterium diphtheriae* that are regulated by the diphtheria toxin repressor (DtxR) and iron. *J. Bacteriol.* **176**:1141–1149.
 17. **Tuomanen, E. I.** 1997. The biology of pneumococcal infection. *Pediatr. Res.* **42**:253–258.
 18. **Yother, J., and D. E. Briles.** 1992. Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis. *J. Bacteriol.* **174**:601–609.

Editor: S. H. E. Kaufmann