INFECTION AND IMMUNITY, Oct. 1998, p. 4748–4754 Vol. 66, No. 10
0019-9567/98/$04.00+0

Comparison of the PspA Sequence from *Streptococcus pneumoniae* EF5668 to the Previously Identified PspA Sequence from Strain Rx1 and Ability of PspA from EF5668 To Elicit Protection against Pneumococci of Different Capsular Types

LARRY S. MCDANIEL,[1,2,3]* D. OLGA MCDANIEL,[2] SUSAN K. HOLLINGSHEAD,[3] AND DAVID E. BRILES[3,4]

Departments of Microbiology[1] and Surgery,[2] The University of Mississippi Medical Center, Jackson, Mississippi 39216, and Departments of Microbiology,[3] Comparative Medicine, and Pediatrics,[4] The University of Alabama at Birmingham, Birmingham, Alabama 35294

Received 20 April 1998/Returned for modification 2 June 1998/Accepted 24 July 1998

PspA (pneumococcal surface protein A) is a serologically varied virulence factor of *Streptococcus pneumoniae*. In mice, PspA has been shown to elicit an antibody response that protects against fatal challenge with encapsulated *S. pneumoniae*, and the protection-eliciting residues have been mapped to the α-helical N-terminal half of the protein. To date, a published DNA sequence for *pspA* is available only for *S. pneumoniae* Rx1, a laboratory strain. PspA/EF5668 (EF5668 indicates the strain of origin of the PspA) is serologically distinct from PspA/Rx1. Sequencing of the gene encoding PspA/EF5668 revealed 71% identity with that of PspA/Rx1. The greatest amount of divergence between the two proteins was seen in their α-helical portions, which are surface exposed and probably under selective pressure to diversify serologically. In spite of the diversity within the α-helical regions of PspAs, we have observed that recombinant PspA (rPspA)/EF5668, like rPspA/Rx1, can elicit cross-protection against pneumococci of different capsular and PspA serological types.

PspA (pneumococcal surface protein A) plays a role in the ability of a pneumococcus to cause disease (17) and is present on all pneumococcal isolates. We have previously demonstrated that this protein is antigenically variant among different pneumococcal isolates (8). Despite this variation, we have observed that immunization with a limited number of variant PspAs can elicit cross-protection against a diverse number of pneumococcal isolates (15, 23). The cross-protection results indicate that while there is variation among PspAs, there must also be conserved PspA epitopes. Understanding the basis of the variation and conservation among PspAs is important in determining the mechanism of cross-protection elicited by PspAs.

PspA is attached to the surface to a pneumococcus by binding to choline in the pneumococcal lipoteichoic acids (27). While this attachment mechanism is novel when compared to that of most other gram-positive surface proteins, a number of other pneumococcal surface proteins have also been observed to bind choline (2, 10, 20).

To date the only complete nucleotide sequence of a *pspA* gene has been determined for *pspA*/Rx1 (Rx1 indicates the strain of origin). The deduced amino acid sequence of *pspA*/Rx1 reveals four distinct domains. The N-terminal half of the protein has a sequence expected to be an α-helical coiled-coil structure similar to those of many surface proteins of other gram-positive bacteria (6, 25). The α-helix of PspA/Rx1 is followed by a region of 81 amino acids, of which 23 are proline. Within the proline-rich domain, the prolines are clustered in Pro-Ala-Pro-Ala-Pro consensus repeats at either end of a 31-amino-acid nonproline region. This proline-rich domain is followed by 10 20-amino-acid repeats that interact with choline to anchor PspA to the pneumococcal surface. These repeats constitute the choline-binding domain (CBD) (25, 27). The CBD is followed by a slightly hydrophobic tail of 17 amino acids. The absence of these last 17 amino acids apparently does not affect the attachment of PspA to a cell surface (25, 27). Previous studies of PspA using antibodies to α-helical-region epitopes and DNA probes for α-helical-region sequences have indicated that α-helical regions are especially varied (15, 22).

Here we report a comparison of the complete nucleotide sequence of a second *pspA* with that of *pspA*/Rx1. The *pspA*/EF5668 gene encoded a PspA that had a greater molecular size than that of PspA/Rx1. Our studies demonstrated that the greater size of PspA/EF5668, compared to that of PspA/Rx1, was due to additional nucleotides in the α-helical coding region of the gene. This finding further confirmed the variability of the α-helical regions of PspAs. Additionally, we observed that in spite of differences in the sequences encoding PspA/EF5668 and PspA/Rx1, both molecules were able to elicit cross-protection against several different pneumococcal isolates.

MATERIALS AND METHODS

Bacterial strains, plasmids, and transformation. Strains and plasmids used in this study are listed in Table 1. Pneumococcal and *Escherichia coli* strains were grown and stored as previously described (13, 17). *E. coli* was transformed by the method of Hanahan (11).

Hybridoma antibodies and immunoblot procedure. The anti-PspA hybridoma cell lines that secrete the monoclonal antibodies (MAb) Xi126 and XiR278 and the epitopes recognized by these MAb have been described previously (8, 12–14). Immunoblot analysis with MAb to detect PspA was carried out as previously described (15).

Amplification of *pspA*. PCR was carried out on genomic DNA isolated from *Streptococcus pneumoniae* EF5668 as previously described (13) with the oligonucleotide primers LSM2 (3′) and LSM112 (5′), whose designs (22) are based on the sequence of *pspA*/Rx1 (25). LSM2 and LSM112 started at nucleotides 1990 and 47 of the *pspA*/Rx1 sequence. LSM2 and LSM112 contain *Sal*I and *Bam*HI restriction sites at their 5′ ends.

* Corresponding author. Mailing address: Department of Microbiology, The University of Mississippi Medical Center, 2500 North State St., Jackson, MS 39216. Phone: (601) 984-6880. Fax: (601) 984-1708. E-mail: LSMCD@fiona.umsmed.edu.

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)	Reference or source
Strains		
S. pneumoniae		
WU2	Capsular type 3, PspA type 1	5
D39	Capsular type 2, PspA type 25	1
R36A	Nonencapsulated mutant of D39, PspA type 25	1
Rx1	Nonencapsulated variant of R36A, PspA type 25	21
A66	Capsular type 3, PspA type 13	1
BG7322	Capsular type 6, PspA type 24	16
EF5668	Capsular type 4, PspA type 12	16
E. coli DH5α	Host strain for plasmids	11
Plasmids		
pJY4163	Em^r	26
pKSD2106	pJY4163::*Bam*HI-*Sal*I PCR product from EF5668, carries DNA that encodes PspA/EF5668	This study

DNA sequencing and analysis. Both strands of DNA encoding PspA/EF5668 were sequenced. The data were obtained by direct sequencing of plasmid pKSD2106, which contains the entire *pspA* gene from *S. pneumoniae* EF5668. Sequencing was done either with Sequenase (U.S. Biochemicals) or on an ABI 377 automated DNA sequencer (Perkin-Elmer, Foster City, Calif.). Sequencing primers were prepared as needed to facilitate sequencing of the cloned pneumococcal DNAs. In a few cases data were confirmed by sequencing of PCR-amplified fragments from the cloned pneumococcal DNAs. Sequence analyses were performed with the programs of the University of Wisconsin's Genetics Computer Group (GCG), MacVector 5.0 (Oxford Molecular), Sequencer 3.0 (GeneCodes, Inc.), and GeneJockey 1.5 (Biosoft, Cambridge, United Kingdom). The Matcher program was used to determine what portions of the sequence matched the 7-amino-acid motif characteristic of coiled-coil proteins (9). To provide direct comparison between the potential structural characteristics of PspA/EF5668 and PspA/Rx1 sequences, we analyzed both sequences using the Matcher program.

Purification of recombinant PspA/EF5668. *E. coli* KSD2106 was grown to mid-log phase as determined by optical density in 500 ml of Luria-Bertani medium. The cells were centrifuged and osmotically shocked (18) to release the periplasmic contents. NaCl was added to the solution to a final concentration of 0.25 M. This solution was passed over a choline-Sepharose column preequilibrated with 50 mM Tris acetate buffer (pH 6.9) containing 0.25 M NaCl (TAB). The column was subsequently washed with 10 bed volumes of TAB. The column was eluted with TAB containing 2% choline chloride, and 1-ml-volume fractions were collected. The presence of PspA/EF5668 was detected in the individual fractions by dot spotting 1 μl of 1/3 serial dilutions of each fraction onto nitrocellulose. The presence of PspA/EF5668 on the membranes was detected by anti-PspA MAb XiR278 followed by alkaline phosphatase-conjugated anti-mouse immunoglobulin. Those fractions containing recombinant PspA/EF5668 were pooled and further analyzed with silver stain (Silver Stain Plus; Bio-Rad, Hercules, Calif.) following sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Immunization and challenge. Immunization studies used CBA/N mice (Jackson Laboratory, Bar Harbor, Maine), which carry the X-linked immunodeficiency trait and fail to respond to polysaccharide antigens, making them very susceptible to pneumococcal challenge (3). Mice were injected subcutaneously with approximately 5 μg of isolated recombinant PspA (rPspA)/EF5668 in 0.2 ml of Freund's complete adjuvant. The mice were boosted at 2 weeks with an additional 5 μg of rPspA/EF5668 in incomplete Freund's adjuvant. Control mice were injected with adjuvant and a comparable volume of a comparable column fraction from an *E. coli* strain that did not express PspA. Approximately 7 days later, the mice were challenged intravenously with a minimum of 100 times the 50% lethal dose of the indicated virulent encapsulated pneumococcal isolate.

Passive protection experiments (15) were performed to examine the protective capacity of sera from some of the mice immunized with PspA/EF5668. CBA/N mice were injected intraperitoneally with 0.1 ml of a 1/40 dilution of pooled mouse sera from immunized or nonimmune mice 1 h prior to intravenous challenge with *S. pneumoniae* A66 or BG7322.

Nucleotide sequence accession number. The nucleotide sequence of *pspA*/EF5668 has been deposited in GenBank under the accession no. U89711.

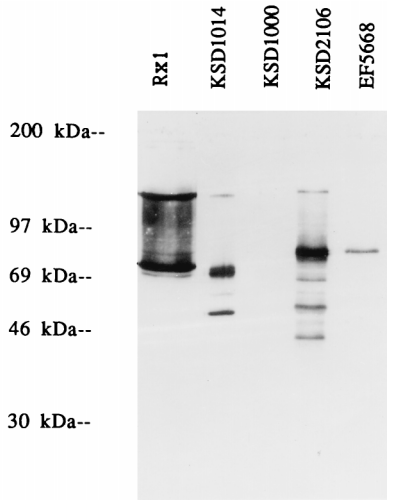

FIG. 1. Immunoblot of PspA detected with the anti-PspA MAb XiR278. The lanes labeled Rx1 and EF5668 contain choline-eluted PspAs from those pneumococcal strains. The Rx1 lane was purposely overloaded to demonstrate the complex banding pattern that can be seen with PspA. The lanes labeled KSD1014 and KSD2106 contain rPspA purified from osmotically shocked *E. coli* expressing PspA/Rx1 and PspA/EF5668, respectively. Lane KSD1000 contains a similar *E. coli* preparation from a strain whose vector plasmid contains no insert.

RESULTS

Cloning and expression of *pspA*/EF5668. LSM112, the 5′ primer, contained an additional *Bam*HI site at the 5′ end, while LSM2, the 3′ primer, contained an additional *Sal*I site at the 5′ end. These sites were used to clone the PCR product amplified from EF5668 genomic DNA into the *Bam*HI/*Sal*I site of pJY4163 (26). The resulting plasmid was designated pKSD2106 and carried a gene that expressed a 92-kDa protein that reacted with XiR278, an anti-PspA MAb (Fig. 1). The apparent molecular mass of PspA/EF5668 was approximately 10 kDa more than that observed for PspA/Rx1. Both rPspA/Rx1 and rPspA/EF5668 had several lower-molecular-weight bands than was observed for material derived from the strains from which the genes encoding PspA were cloned. Following the purification of the rPspA, the lower-molecular-weight bands were reduced in intensity or not detected at all by Western blotting (data not shown).

Sequence analysis of *pspA*/EF5668. The complete nucleotide sequence of *pspA*/EF5668 was determined. The open reading frame of *pspA*/EF5668 includes 102 more nucleotides than were found for the open reading frame of *pspA*/Rx1. The nucleotides within the open reading frame were sufficient to encode a mature PspA with an expected molecular mass of 69,871.98 Da.

Amino acid sequence analysis. The deduced amino acid sequence of PspA/EF5668 is shown aligned with the previously published sequence of PspA/Rx1 (Fig. 2). The percentages of identity of the different regions of PspA from the two pneumococcal strains are shown in Table 2. The overall identity between the two molecules was 71%. Two blocks of 11 and 34 amino acids were present in the α-helical region of PspA/EF5668 which were absent from PspA/Rx1. The larger of these two blocks fell within a region of PspA/Rx1 to which protective MAb had previously been mapped (13). Despite this divergence, analysis of amino acids 1 to 352 of the mature PspA/EF5668, like the corresponding region of PspA/Rx1, predicts the formation of a coiled-coil α-helical structure (Fig. 3).

The number of amino acids of the proline-rich domain of

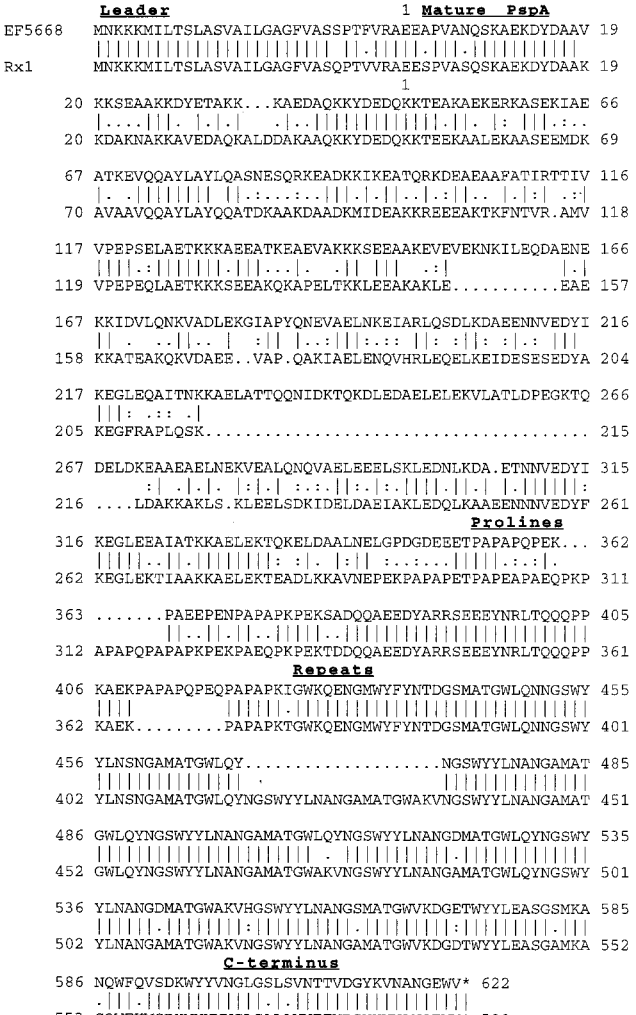

FIG. 2. Alignment of the amino acid sequence of PspA/EF5668 with that of PspA/Rx1. Sequences are shown in standard one-letter amino acid code. The vertical lines indicate identity. Double and single dots between sequences indicate strong and weak similarity, respectively, based on Pam250 matrix alignment with the GCG program. Dots within a sequence indicate gaps in the alignment. The different domains of PspA are indicated above the sequences in boldface type. The words Leader, Mature PspA, Prolines, Repeats, and C-terminus each begin at the first codon that marks their respective domains.

PspA/EF5668 was 1 less than that found in PspA/Rx1 (Fig. 2). The most conserved region within the proline-rich domain was the 27 nonproline amino acids between the two proline-containing sequences. The greatest differences in the PspA/EF5668 and PspA/Rx1 proline-rich domains were in the two blocks of proline-containing amino acids within each domain. Compared to PspA/EF5668, PspA/Rx1 contains five extra prolines (10 extra amino acids) at the N-terminal end of the proline-rich domain and four fewer prolines (9 amino acids) at the C-terminal end of the domain. Additionally, the CBD of PspA/EF5668 contains nine repeat regions, one less than the CBD of PspA/Rx1.

Based on analysis with the Matcher program, the α-helical domain of PspA/EF5668 was largely predictive of a coiled-coil conformation from amino acids 10 through 344, but like the PspA/Rx1 sequence, it contained several breaks in the coiled-coil motif. In the previous analysis of the α-helical domain of PspA/Rx1, it was observed that the frequencies of the different amino acids at each position (a, b, c, d, e, f, or g) of the coiled-coil motif were generally consistent with those of known coiled-coil sequences (Fig. 4).

In a few cases, however, the relative frequencies of particular amino acids at some positions appeared to deviate from those of other coiled-coil sequences. Through the comparison of the PspA/EF5668 sequence with that of PspA/Rx1 and composite data for other known coiled-coil sequences, it has been possible to determine which of these deviations may be particular to PspAs in general and which ones may simply indicate random variation between proteins.

We observed four instances where the frequencies of amino acids in the coiled-coil motif of these two PspAs were highly deviant from those of the known non-PspA sequences (Fig. 4). At each of the seven sequence positions (a, b, c, d, e, f, and g) in the coiled-coil motif, there were fewer arginines than expected. At position f in other coiled-coil sequences, 13% of amino acids were arginines, and at position f in the two PspAs, no arginines were observed, contrary to what was expected ($P = 0.0003$). Instead of using the positively charged arginines at positions b, c, e, f, and g, PspAs used lysines almost exclusively. At position f, 43 and 37% lysines were observed in the two PspAs, compared to 11% in non-PspA coiled-coil proteins ($P < 0.0001$). The numbers of lysines in positions c and e were also somewhat higher than expected. The other significant difference was at position d, where PspAs had on average 42% alanines in their α-helical domains but where non-PspA molecules had 22% alanines in their coiled-coil regions ($P \leq 0.0098$).

Since 140 similar comparisons were made, it was expected that <0.014, <0.043, and <1.37 of the comparisons would be significant at P values of ≤0.0001, ≤0.0003, and ≤0.0098, respectively. We observed one significant comparison at a P value of ≤0.0001, two at ≤0.0003, and three at ≤0.0098. Thus, it is likely that at least two of the observations referred to above may represent real differences between PspA sequences and those of other coiled-coil proteins.

As in most coiled-coil proteins, alanine and leucine were the predominant amino acids seen at position d. It was of note, however, that in the first half of each α-helical domain the most common position d amino acid was alanine but that in the second half leucine was the most common. That this was true even in regions of the PspAs that were not highly identical at other amino acid positions suggests a functional necessity for the differences in the position d compositions of the two halves of the α-helical domain.

The breaks in the coiled-coil motif prevent contiguous motif regions from being larger than 42 amino acids. The PspA/EF5668 sequence contains seven major coiled-coiled blocks of 26 to 42 amino acids and six smaller blocks of 10 to 18 amino acids that can be aligned to fit a coiled-coil motif. These blocks

TABLE 2. Percentages of identity of amino acid residues of different domains of PspA/EF5668 compared to those of PspA/Rx1

Amino acid residues[a]	Region of PspA	% Identity
−31 to −1	Leader	94
1 to 345	α-Helix	54
346 to 524	Proline-rich	84
426 to 605	Repeat (CBD)	94
606 to 622	C terminus	100
1 to 622	Mature PspA	71

[a] The numbering is based on the numbering of the amino acid sequence of PspA/EF5668 as aligned with that of PspA/Rx1 in Fig. 2.

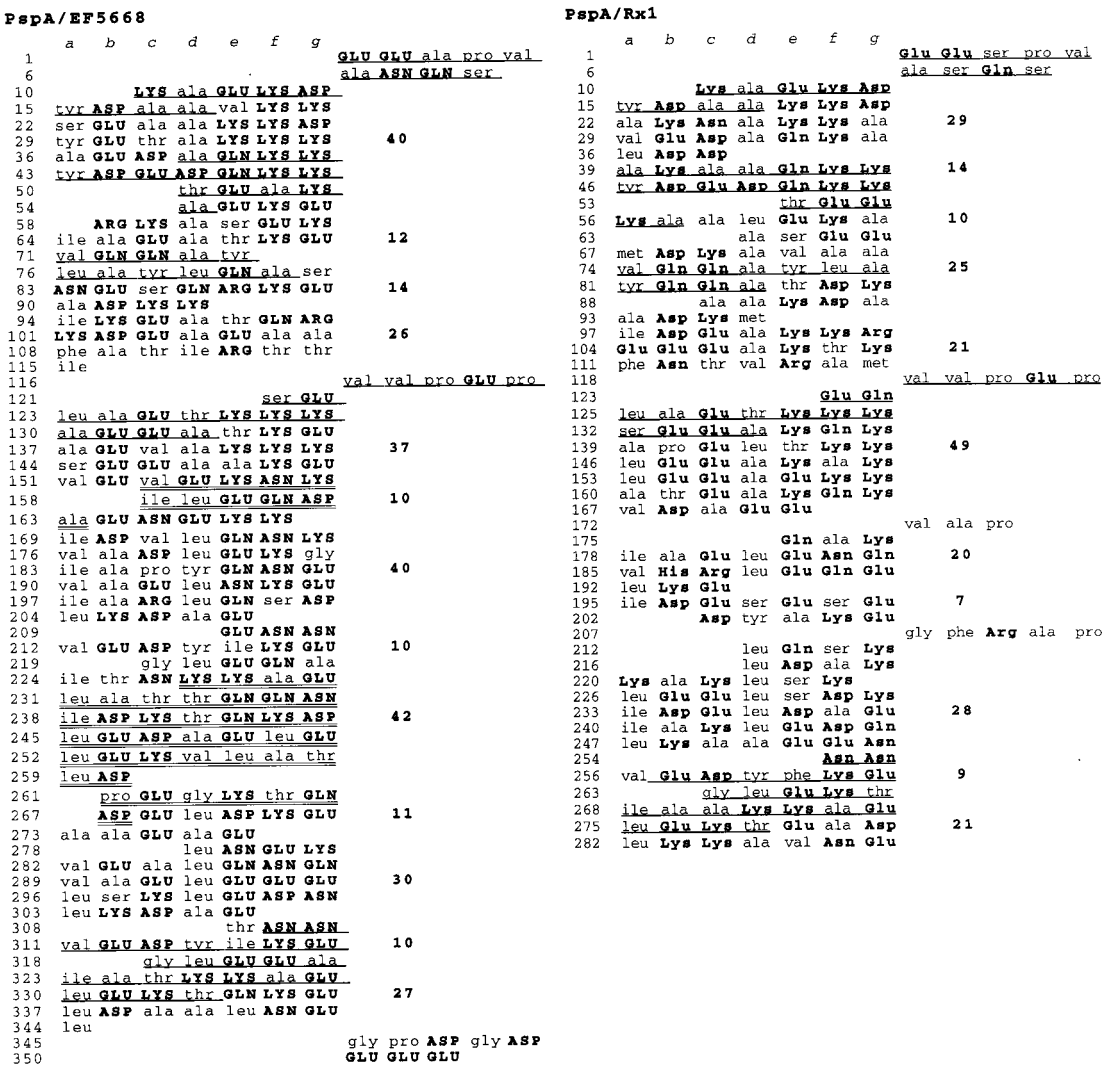

FIG. 3. Amino acid sequences of the N-terminal 352 and 288 amino acids of PspA/EF5668 and PspA/Rx1, respectively, arranged to illustrate the coiled-coil motif. In this motif the amino acids in positions a and d are generally nonhydrophilic, whereas amino acids at positions b, c, e, f, and g are generally hydrophilic and frequently charged. The most hydrophilic amino acids are indicated in boldface type. Amino acids compatible with hydrophilic environments are indicated in boldface, and amino acids compatible with nonsurface exposure are indicated in lightface. Regions of high identity between PspA/EF5668 and PspA/Rx1 are indicated by a single underline. Regions present in the PspA/EF5668 sequence that are absent from the PspA/Rx1 sequence, based on GCG Best Fit alignment (Fig. 3), are indicated by a double underline. Numbers to the left of each column denote the number of the first amino acid in each column. Numbers to the right of each column indicate the numbers of amino acids in contiguous blocks of the uninterrupted coiled-coil sequence motif.

of coiled-coil motif are separated by 11 frameshifts in the coiled-coil motif and a 5-amino-acid region (116 to 120) that could not be aligned into a coiled-coil motif.

Within the α-helical domain there are five regions of high conservation between PspA/EF5668 and PspA/Rx1. Based on the numbering of the PspA/EF5668 sequence, these are amino acids 1 to 18, 39 to 54, 71 to 81, 116 to 133, and 309 to 332. One of these conserved sequences includes the N-terminal non-α-helical–non-coiled-coil region, and the others each include significant breaks in the coiled-coil motif (Fig. 3). For PspA/Rx1 it has been shown that amino acids 192 to 260 contain highly cross-reactive protection-eliciting epitopes (13). This region corresponds to amino acids 202 to 315 of PspA/EF5668. The last conserved α-helical sequence, amino acids 309 to 332, overlaps slightly this region. The first two highly conserved α-helical sequences fall within the first 115 amino acids of PspA/Rx1, where one of five known protection-eliciting epitopes exists (4, 13). The 116- to 134-amino-acid sequence is in the portion of PspA with which no MAb are known to react.

Region of alternating charge. In the analysis of the PspA/Rx1 sequence it was observed that for 12 consecutive turns of the α-helix (amino acids 125 to 166), the charge alternated between negative and positive. The positive charges were associated with lysines at many of the e, f, and g positions, and the negative charges were associated with glutamic acids at many of the b and c positions (25) (Fig. 4). The sequence of the coiled-coil region of PspA/EF5668 reveals a similar alternating charge motif for the 10 turns of the α-helix between amino acids 123 and 157.

In the PspA/EF5668 sequence an alternation of charge with consecutive turns of the α-helix was also observed for the 10 turns of the α-helix from amino acids 15 to 49. This time, however, about half of the negatively charged amino acids in positions b and c were aspartic acid. In light of this finding, a

		a			*b*			*c*			*d*			*e*			*f*			*g*		
		5668	Rx1	other	5668	Rx1	other	5668	Rx1	other	5668	Rx1	other	5668	Rx1	other	5668	Rx1	other	5668	Rx1	other
−	**Glu**	0	3	**1**	34	24	**21**	27	37	**19**	4	3	**6**	29	28	**32**	8	10	**15**	36	23	**20**
+	**Lys**	2	6	**8**	7	15	**15**	16	18	**12**	6	3	**1**	18	25	**9**	43	37	**11**	23	35	**15**
−	**Asp**	0	0	**0**	20	27	**13**	16	11	**13**	2	3	**1**	1	5	**4**	2	10	**10**	11	8	**8**
+	**Arg**	0	0	**6**	2	0	**6**	2	3	**8**	0	0	**1**	4	3	**6**	0	0	**13**	2	3	**10**
+	**His**	0	0	**1**	0	3	**3**	0	0	**2**	0	0	**1**	0	0	**1**	0	0	**3**	0	0	**1**
	Asn	2	0	**4**	0	3	**4**	4	3	**5**	0	0	**1**	4	0	**6**	14	7	**5**	9	5	**3**
	Gln	0	0	**1**	2	6	**9**	0	5	**8**	2	0	**4**	20	13	**14**	8	7	**6**	4	8	**13**
	ala	12	14	**10**	27	18	**12**	9	18	**8**	38	46	**22**	2	3	**4**	12	20	**11**	6	15	**9**
	thr	0	0	**1**	2	3	**4**	9	3	**4**	10	5	**2**	8	8	**5**	4	2	**4**	4	3	**4**
	ser	5	3	**2**	2	0	**4**	2	0	**8**	0	3	**2**	2	8	**5**	4	5	**8**	2	0	**5**
	thy	7	9	**4**	0	0	**1**	2	0	**0**	6	5	**6**	2	8	**0**	0	0	**0**	0	0	**0**
	pro	0	0	**0**	2	3	**0**	2	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**
	cys	0	0	**1**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**1**	0	0	**0**
	gly	0	0	**1**	0	0	**2**	4	3	**4**	2	0	**1**	0	0	**1**	0	0	**4**	2	0	**1**
	trp	0	0	**0**	0	0	**0**	0	0	**0**	0	0	**1**	0	0	**0**	0	0	**0**	0	0	**0**
	leu	29	26	**32**	0	0	**2**	0	0	**4**	26	28	**35**	4	0	**6**	2	2	**4**	0	0	**6**
	Ile	21	17	**13**	0	0	**1**	2	0	**2**	2	0	**6**	4	0	**2**	0	0	**2**	0	0	**2**
	val	19	17	**9**	0	0	**2**	4	0	**2**	2	3	**6**	2	5	**2**	0	0	**3**	0	0	**3**
	phe	2	3	**2**	0	0	**0**	0	0	**1**	0	0	**2**	0	3	**0**	0	0	**0**	0	0	**0**
	met	0	3	**5**	0	0	**1**	0	0	**1**	0	3	**2**	0	0	**1**	0	0	**1**	0	3	**0**

FIG. 4. Comparison of the PspA/EF5668 coiled-coil structure with that of PspA/Rx1 and pooled data (7) from other known coiled-coil proteins. The charges of the amino acid side chains are listed to the right of amino acids Glu, Lys, Asp, Arg, and His. Those listed in boldface type include the most hydrophilic amino acids and include all charged amino acids. Those listed in lowercase letters and underlined are the most hydrophobic amino acids. Amino acid positions in the coiled-coil motif (a through g) are listed at the top of the figure. The observed amino acid frequencies for the pooled PspA data (PspA/EF5668 and PspA/Rx1) differed from those of the coiled-coiled regions of other proteins in several cases (indicated by shading) as follows: for the lysine at position e, $P = 0.036$; for the lysine at position f, $P < 0.0001$; for the arginine at position f, $P = 0.0003$; and for the alanine at position d, $P = 0.0098$.

reexamination of the PspA/Rx1 sequence revealed a similar (but slightly less distinct) alternating charge motif for the seven consecutive turns from amino acids 15 to 38. Similar repetitive alternation of charge in consecutive turns of an α-helix has not been reported for other coiled-coil sequences.

Cross-protection with PspA/EF5668. To assess the ability of PspA/EF5668 to elicit protective immune responses against pneumococci, mice were immunized with recombinant PspA/EF5668 or a comparably prepared column fraction from *E. coli* that contained the vector with no pneumococcal insert. Different groups of mice immunized with each antigen preparation were challenged by intravenous injection with one of five pneumococcal isolates representing four capsular types and five PspA serotypes. In all cases mice immunized with PspA/EF5668 showed significant survival compared to that of the controls. With two of the five strains, all immunized mice survived (Table 3).

The protective capacity of antibody-containing serum was tested with sera from additional groups of mice immunized in a similar fashion with recombinant PspA/EF5668. Mice were passively immunized with a 1/40 dilution of pooled mouse serum and challenged intravenously with pneumococci. When challenged with the capsular type 3 strain A66, three of three mice receiving immune sera from PspA/EF5668-immunized mice survived while all three control mice died. Similar experiments conducted with the capsular type 6 strain BG7322 resulted in the survival of all three passively immunized mice and one of three control mice. From an analysis of pooled data from experiments with immune and control mice challenged with both capsular types, it was observed that survival was statistically more common for mice receiving immune serum ($P < 0.008$).

DISCUSSION

PspA, either by itself or in combination with additional pneumococcal proteins, offers the possibility of a protein-based pneumococcal vaccine. Such a vaccine could be important in targeting groups at risk for pneumococcal infection. Although a capsular conjugate vaccine will be effective in affording specific protection, there are possible limitations to this approach. If PspA is to serve as a component of a pneumococcal vaccine, it is important to define the antigenic diversity that is seen in PspAs. One approach to this problem is to obtain additional

TABLE 3. Cross-protection of CBA/N mice immunized with recombinant PspA/EF5668[a]

Challenge strain			Median no. of days to death of mice		No. of mice that lived:no. of mice that died		*P* value[b] for PspA$^+$ mice vs PspA$^-$ mice
Name	Capsular serotype	PspA serotype	PspA$^+$	PspA$^-$	PspA$^+$	PspA$^-$	
WU2	3	1	>10	2	8:0	0:6	<0.0003
A66	3	13	>10	2.5	7:1	0:6	<0.008
BG7322	6	24	>10	6	5:1	1:5	<0.003
D39	2	2	>10	3	3:2	0:5	<0.075
EF5668	4	4	>10	2	5:0	1:4	<0.002

[a] Mice were immunized with two injections of 5 μg of rPspA/EF5668 or a comparable volume of material prepared from JY4163, which contains the vector but no *pspA* insert.

[b] *P* values were calculated for day-of-death data obtained with mice separately infected with each challenge strain by using the Wilcoxon two-sample rank sum test to compare time to death values of immunized mice with those of mice given adjuvant and diluent but no antigen.

nucleotide sequence information of the genes encoding PspAs from many different pneumococcal isolates.

Previous studies have shown that all protective MAb recognize the α-helical regions of PspAs, and most were found to recognize epitopes in the C-terminal 100 amino acids of the α-helical region (13). Moreover, immunization with fragments of PspA containing portions of the α-helical region has been shown to elicit protection (13, 19). Present data indicate that immunization with the C-terminal end of the α-helical region shows similar cross-protection to immunization with the intact molecule (13). Earlier studies using fragments of *pspA*/Rx1 as DNA probes found the greatest degree of conservation among PspAs to be in their CBDs and proline-rich domains. The α-helical domains were found to be the most diverse (16). This was confirmed in studies using specific oligonucleotides as hybridization probes (22). Studies of the expression of epitopes detected by MAb in different PspAs have also demonstrated high serological variability of the α-helical regions (8, 24).

The present sequence provides the first insight into the diversity and conservation within PspAs at the sequence level. The encoded sequences displayed almost complete conservation in their leader regions, proline-rich regions, CDBs, and 17-amino-acid C termini. The only significant diversity was observed in the α-helical domains, and even then there were some regions that were much more conserved than others. It was of particular interest that the region from amino acids 254 to 278 (numbered based on the numbering of PspA/Rx1) was the largest highly conserved portion in the α-helical region and that it partially overlapped the region (amino acids 192 to 260) that has been identified as particularly cross-protective (13).

The most conserved regions of the α-helix are all associated with major frameshifts or gaps in the coiled-coil motif. This finding may suggest that these regions are particularly important in the conformation of the protein. Alternatively, these regions may be unimportant in the elicitation of protection and thus not be subject to evolutionary selection for antigenic variation.

Although there were considerable differences in the α-helical regions of the two PspAs, they both conformed to a largely coiled-coil motif. Both had similar amino acids at the several positions of the motif that differed significantly from amino acids of non-PspA coiled-coil proteins. One striking aspect of both sequences was the relatively high usage of lysine rather than of other positively charged amino acids such as arginine. It was also of note that the PspAs made more frequent use of alanine at nearly all positions than was observed in other coiled-coil proteins. We previously proposed that the long lysine side chains allow PspA to interact more effectively with the negative charges of most capsular polysaccharides.

Another aspect of both the α-helical sequences that appeared to be relatively unique to PspA was that there were two regions in each PspA sequence where consecutive turns of the α-helix alternated in charge. It seems likely that these unusual regions and the high frequencies of alanines and lysines in PspA may be important to the still unknown function of PspA.

In addition to being more varied than any other part of *pspA*, the portion of the gene encoding the α-helical region is also more varied than the noncoding DNA 5′ to the leader. This finding suggests that the greater diversity of the α-helical region than of the rest of the molecule is because of selection favoring diversity rather than a lack of selection favoring function. As we have suggested in the past, it is likely that the diversity of this region may have been the result of selection favoring antigenic variability that helped pneumococci escape responses to prior pneumococcal infections. If this is the case, then it would be a strong argument for anti-PspA-mediated protection against pneumococci in humans, since humans are thought to be the primary reservoir of this pathogen.

ACKNOWLEDGMENTS

We are grateful to Edwin Swiatlo for suggestions and for critically reading the manuscript.

This study was supported by NIH grants AI27201 (L.S.M.) and AI21548 (D.E.B. and S.K.H.). Support was also provided by a Biomedical Research Grant (L.S.M.) from the University of Mississippi Medical Center.

REFERENCES

1. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. **79:**137–158.
2. **Briese, T., and R. Hakenbeck.** 1983. Interaction between choline and the N-acetyl-L-alanine amidase of *Streptococcus pneumoniae* p. 173–178. *In* R. Hakenbeck, J. V. Holtjeandand, and H. Labischinski (ed.), The target of penicillin. The murein sacculus of bacterial cell walls, architecture, and growth. Walter de Gruyter, Berlin, Germany.
3. **Briles, D. E., J. L. Claflin, K. Schroer, and C. Forman.** 1981. Mouse IgG3 antibodies are highly protective against infection with *Streptococcus pneumoniae*. Nature **294:**88–90.
4. **Briles, D. E., J. D. King, M. A. Gray, L. S. McDaniel, E. Swiatlo, and K. A. Benton.** 1996. PspA, a protection-eliciting pneumococcal protein: immunogenicity of isolated native PspA in mice. Vaccine **14:**858–867.
5. **Briles, D. E., M. Nahm, K. Schroer, J. Davie, P. Baker, J. Kearney, and R. Barletta.** 1981. Antiphosphocholine antibodies found in normal mouse serum are protective against intravenous infection with type 3 *Streptococcus pneumoniae*. J. Exp. Med. **153:**694–705.
6. **Briles, D. E., R. C. Tart, E. Swiatlo, J. P. Dillard, P. Smith, K. A. Benton, B. A. Ralph, A. Brooks-Walter, M. J. Crain, S. K. Hollingshead, and L. S. McDaniel.** Pneumococcal diversity: considerations for new vaccine strategies with emphasis on pneumococcal surface protein A (PspA). Clin. Microbiol. Rev., in press.
7. **Cohen, C., and D. A. D. Parry.** 1990. α-Helical coiled coils and bundles: how to design an α-helical protein. Genetics **7:**1–15.
8. **Crain, M. J., W. D. Waltman II, J. S. Turner, J. Yother, D. E. Talkington, L. M. McDaniel, B. M. Gray, and D. E. Briles.** 1990. Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. Infect. Immun. **58:**3293–3299.
9. **Fischetti, V. A., G. M. Landau, J. P. Schmidt, and P. Sellers.** 1993. Identifying periodic occurrences of a template with applications to protein structure. Inf. Process. Lett. **45:**11–18.
10. **Garcia, P., and J. L. Garcia.** 1989. Purification and characterization of the autolytic glycosidase of *Streptococcus pneumoniae*. Biochem. Biophys. Res. Commun. **158:**251–256.
11. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
12. **Kearney, J. F., A. Radbuch, B. Liesegang, and K. Rajewsky.** 1979. A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. J. Immunol. **123:**1548–1550.
13. **McDaniel, L. S., B. A. Ralph, D. O. McDaniel, and D. E. Briles.** 1994. Localization of protection-eliciting epitopes on PspA of *Streptococcus pneumoniae* between amino acid residues 192 and 260. Microb. Pathog. **17:**323–337.
14. **McDaniel, L. S., G. Scott, J. F. Kearney, and D. E. Briles.** 1984. Monoclonal antibodies against protease sensitive pneumococcal antigens can protect mice from fatal infection with *Streptococcus pneumoniae*. J. Exp. Med. **160:**386–397.
15. **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.
16. **McDaniel, L. S., J. S. Sheffield, E. Swiatlo, J. Yother, M. J. Crain, and D. E. Briles.** 1992. Molecular localization of variable and conserved regions of *pspA*, and identification of additional *pspA* homologous sequences in *Streptococcus pneumoniae*. Microb. Pathog. **13:**261–269.
17. **McDaniel, L. S., J. Yother, M. Vijayakumar, L. McGarry, W. R. Guild, and D. E. Briles.** 1987. Use of insertional inactivation to facilitate studies of biological properties of pneumococcal surface protein A (PspA). J. Exp. Med. **165:**381–394.
18. **Osborn, M. J., and J. Munson.** 1974. Separation of the inner (cytoplasmic) and outer membranes of gram negative bacteria. Methods Enzymol. **31:**642–653.
19. **Ralph, B. A., D. E. Briles, and L. S. McDaniel.** 1994. Cross-reactive protec-

tion eliciting epitopes of pneumococcal surface protein A. Ann. N. Y. Acad. Sci. **730:**361–363.
20. **Sanchez-Pulles, J. M., J. M. Sanz, J. L. Garcia, and E. Garcia.** 1990. Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases. Gene **89:**69–75.
21. **Shoemaker, N. B., and W. R. Guild.** 1974. Destruction of low efficiency markers is a slow process occurring at a heteroduplex stage of transformation. Mol. Gen. Genet. **128:**283–290.
22. **Swiatlo, E., A. Brooks-Walter, D. E. Briles, and L. S. McDaniel.** 1997. Oligonucleotides identify conserved and variable regions of *pspA* and *pspA*-like sequences of *Streptococcus pneumoniae*. Gene **188:**279–294.
23. **Tart, R. C., L. S. McDaniel, B. A. Ralph, and D. E. Briles.** 1996. Truncated Streptococcus pneumoniae PspA molecules elicit cross-protective immunity against pneumococcal challenge in mice. J. Infect. Dis. **173:**380–386.
24. **Waltman, W. D., II, L. S. McDaniel, B. Andersson, L. Bland, B. M. Gray, C. Svanborg-Eden, and D. E. Briles.** 1988. Protein serotyping of *Streptococcus pneumoniae* based on reactivity to six monoclonal antibodies. Microb. Pathog. **5:**159–167.
25. **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.
26. **Yother, J., G. L. Handsome, and D. E. Briles.** 1992. Truncated forms of PspA that are secreted from *Streptococcus pneumoniae* and their use in functional studies and cloning of the *pspA* gene. J. Bacteriol. **174:**610–618.
27. **Yother, J., and J. M. White.** 1994. Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. J. Bacteriol. **176:**2976–2985.

Editor: V. A. Fischetti