CLONING, EXPRESSION, PURIFICATION AND IN SILICO ANALYSIS OF N-TERMINAL DOMAIN OF PNEUMOCOCCAL SURFACE PROTEIN A FROM ALL SIX CLADES

1, 2, *Naeem Khan, 2Raies Ahmad Qadri and 1Devinder Sehgal

1Molecular Immunology Laboratory, National Institute of Immunology, New Delhi, India
2Department of Biotechnology, University of Kashmir, Srinagar, India

INTRODUCTION

Streptococcus pneumoniae commonly known as pneumococcus, a Gram positive and major respiratory mucosal bacterial pathogen, is one of the leading cause of pneumonia, bacteremia, meningitis, and otitis media (Bridy-Pappas et al. 2005). S. pneumoniae is responsible for mortality of both children and elderly worldwide. In both developed and developing countries, S. pneumoniae places a large burden on the medical community, as approximately one million cases occurred each year, majority of deaths occur in children below 5 years of age (Bogaert et al. 2004, Grijalva and Edwards 2006). With the ever increasing degree of antibiotic resistance acquired by pneumococcal strains, vaccination represents the most effective strategy for preventing pneumococcal infection. All virulent strains of S. pneumoniae are encapsulated and loss of capsule leads to avirulence, indicating that the polysaccharide capsule is the major virulence factor of pneumococci(Preston and Dockrell 2008). Humoral response is the primary protective mechanism against the pneumococcal infection. Pneumococcal vaccines that are currently available are capsular polysaccharide based and designed to cover the serotypes most frequently associated with invasive pneumococcal disease. The 23-valent pneumococcal polysaccharide vaccine was developed primarily based on epidemiological scenario of developed countries and has limited coverage of serotypes that cause diseases in children in developing countries. The second 7-valent pneumococcal conjugate vaccine that contains inactivated diphtheria toxin as a carrier protein is more effective in all age groups and immunization with it prevents nasopharyngeal colonization of strains include in the vaccine. Recently, a 10-valent and 13-valent conjugate vaccines has been approved (Bermal et al. 2009). To overcome the drawbacks of current pneumococcal vaccines, global efforts are on to identify proteins which are well conserved across all serotypes as potential vaccine candidates for a protein based pneumococcal vaccine (Adamou et al. 2001). Protein virulence factors present on the pneumococcal surface are potential vaccine candidates. Several pneumococcal proteins have been implicated in the virulence and pathogenesis which includes PspA, PspC, PsaA, neuraminidase, PiaA, PlyA, and LytA (Kadioglou et al. 2008, Tai 2006). In addition, many more pneumococcal surface proteins are being considered for vaccine formulation includes PspA, PcsB, and StkP (Griefing et al. 2008). PspA is a polymorphic cell surface choline binding protein, with variable molecular size ranging from 67 to 99 kDa (Holingshead et al. 2006). PspA, a virulence factor, is present in all strains of pneumococci. It appears to protect against the host complement system by blocking recruitment of C3b deposited on the pneumococci, thereby reducing the effectiveness of the complement-receptor-mediated pathways of clearance and suggests a significance role for PspA in the invasive pneumococcal infection (Cheng et al. 2000, Ren et al. 2004, Tu et al. 1999). In addition, PspA has lactoferrin-binding activity (Tai et al. 1993). Unlike PsaA, Ply, BVH-3 and BVH-11, PspA shows inter strain variability at amino acid sequence

Key words:
Pneumococcal surface protein A;
Multiple sequence alignment;
Secondary structure analysis;
Antibody generation

Copyright © 2014 Naeem Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
level. Immunization of PspA elicits antibodies that has been proven protective against invasive infection and nasopharyngeal carriage (Wu et al. 1997). Structural analysis of PspA from various pneumococcal strains suggests that it consists of four distinct structural domains: an N-terminal α-helical structure, a proline rich region, a stretch of ten highly conserved 20-amino acid repeats, and a slightly hydrophobic sequence of seventeen amino acids at the C-terminus (Yother and White 1994). Based on the amino acid sequence of α-helical region, PspA from various pneumococcal strains have been classified into three families and six clades. Family 1 is comprised of clade 1 and 2; family 2 consist clade 3, 4 and 5 and family 3 has only clade 6 (Hollingshead et al. 2000). The purpose of this study was to identify the conserved region that could be immunogenic and confer protective immunity. Here, we report the cloning, expression and purification of N-terminal domain of PspA belongs to the all six clades. Multiple sequence alignment and in silico secondary structure prediction suggested that PspA from different clades showed very high sequence variability while it evolutionary conserved at secondary structure level and immunogenicity of recombinants PspAs was further confirmed by raising polyclonal sera in mice.

MATERIALS AND METHODS

Mice

Six to eight week old female BALB/c mice were maintained in the Small Animal Facility of the National Institute of Immunology, New Delhi. All experimental procedures involving use of recombinant DNA technology and mice were conducted in accordance with the Institutional Biosafety Committee and Institutional Animals Bioethics Committee guidelines, respectively.

Bacterial strains, plasmids, media and growth condition

Pneumococcal strains and plasmids used in this study are listed in Table 1. Pneumococci were grown on tryptic soy agar supplemented with 5% sheep blood (TSA) plates and Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) (Becton Dickinson, USA) at 37°C in the presence of 5% CO₂ (Rohatgi et al. 2009). Escherichia coli was maintained in Luria-Bertani broth containing appropriate antibiotic(s), were required.

Molecular cloning, overexpression and purification of recombinant PspAs

The sub-fragments encoding the N-terminal (surface exposed) region of PspA for clade 3 and 5 were amplified by PCR using the genomic DNA from the pneumococcal strains TIGR4 and ATCC 6303. The cloning strategy and PCR conditions were adapted from Rohatgi et al. The ORF of pspA genes were PCR amplified using the primers listed in Table 2. A HindIII recognition site (a/agctt) was included in the antisense primer to facilitate cloning. Hot start PCR was performed using Programmable Thermal Cycler (Applied Biosystems, USA). The resultant PCR amplified fragment was digested with HindIII and ligated into pQE-30Xa vector (Qiagen, Germany).

The constructs were transformed into E. coli strain XL1 Blue. The right constructs were confirmed by restriction and sequencing analysis. For expression purposes the construct was transformed in E. coli K-12 expression strain SG13009 (Qiagen), Isopropyl β-D-1-thiogalactopyranoside (1 mM) (IPTG) (Sigma, USA) was added to mid-logarithmic phase E. coli culture (A₆₀₀ ~ 0.6) to induce high-level expression of recombinant proteins at 37°C for 2 hr with vigorous shaking. Cells were pelleted by centrifuging at 6000 x g for 10 min and lysed in lysis buffer (50 mM NaH₂PO₄/ 300 mM NaCl/ 10 mM imidazole, pH 8.0) and subsequently by sonication. The recombinant proteins were purified by using Ni-NTA affinity chromatography following the manufacturer instructions (Sigma, USA). Bound protein was eluted using buffer containing 250 mM imidazole. Eluted fractions containing proteins were pooled and dialyzed extensively against PBS (pH 7.4) to remove imidazole. Purity of the recombinant PspAs were assessed by SDS-PAGE.

SDS-PAGE and immunoblot analysis

Purified proteins were resolved on a 12% SDS-PAGE (run for 90 min at 100 V) and transferred onto a nitrocellulose membrane (15 V for 30 min) using semidry transfer apparatus, (Trans-Blot SD Cell, Bio-Rad, USA). Membrane was blocked with 2% skim milk in PBS for 1 hr. Membrane was probed with anti-histidine tag antibodies (dilution 1 in 2000) for 1 hr at room temperature. The membrane was washed 3 times with PBS-0.5% Tween 20 for 10 min each. Bound antibodies were detected by using horseradish peroxidase-conjugate goat-anti-mouse Ig antibodies (Becton, Dickinson and Company, USA) (dilution 1 in 10,000) and 3, 3'-diaminobenzidine/H₂O₂ as substrate.

Bioinfrmatic analysis

The multiple sequence alignment corresponding to α-helical domain of all six clade of PspA were aligned and analyzed by the ClustalW algorithm using MacVector version 10 (MacVector Inc, USA). The amino acid similarity / identity percentages were calculated for all six clades of PspA showed across the clade. The secondary structure prediction was carried out by the online computational algorithm / package using PHD software (Rost and Sander 1993).

Generation of polyclonal sera against PspA fragments

The groups of 3 female BALB/c mice were immunized subcutaneously with 25 µg of recombinant PspAs (in 25 µg alum, Pierce, USA) in PBS. Mice were boosted at day 14 and 28 with the same amount of antigen and route of immunization. One week after the third immunization, mice were bleed retro-orbitally and sera were collected and stored at -20°C.

Measurement of antibody response by ELISA

Antigen-specific total immunoglobulin was quantitated by ELISA. Briefly, 96-well ELISA plate (Greiner Bio-one, Germany) was coated with different rPspA (50 µl per well of 5 µg/ml) in carbonate-bicarbonate buffer (pH 9.6) overnight at

5858 Naeeem Khan et al. Cloning, expression, purification and in silico analysis of n-terminal domain of pneumococcal surface protein a from all six clades
RESULTS

Expression and Purification of His-tagged recombinant PspA fragments (N-terminal) from all six clades

We have successfully constructed the expression vectors (pQE-30Xa) containing the N-terminal region of PspA. PspA genes were amplified by PCR from the genomic DNA of *S. pneumoniae* TIGR4 (clade 3), and ATCC6303 (clade 5) using the sets of primers (Fig. 1A and B). The constructs were transformed into *E.coli* XL-1 Blue. The presence of right construct was than confirmed by restriction digestion with different restriction enzymes and final confirmation was done by sequencing. The constructs encoding the corresponding PspA sub-fragments for clade 1 (strain L82016), clade 4 (strain JCP#56) and clade 6 (strain BG9300) were obtained from Prof. Susan K. Hollingshead, USA. The right constructs were than transformed into expression strains *E.coli* SG13009 and BL-21 (DE3). For further analysis we expressed and purified surface exposed domain of from all six clades of PspA with histidine tagged recombinant fusion protein from *E. coli* by using Ni-NTA (Ni- Nitriolacetic acid resin) chromatography. The purity and size of the recombinant proteins were visualized by 12% sodium dodecyl sulfate –polyacrylamide gel electrophoresis (SDS-PAGE) and observed > 95% pure when stained with Commassie Brilliant Blue R. The recombinant proteins were confirmed by western blotting by using anti-histidine tag antibodies (Fig. 1C and D).

Multiple sequence alignment and secondary structure prediction

Previous reports have shown that the greatest amount of variability was seen in their α-helical region between the different proteins, which are surface exposed and evolutionary conserved. The primary sequences of all six clades of PspA (N-

![Table 1. Pneumococcal strains and plasmids used in this study](image1)

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Pneumococcal strain (serotype)</th>
<th>PspA family (clade)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUAB069&lt;sup&gt;a&lt;/sup&gt;</td>
<td>LS2016 (6B)</td>
<td>1 (1)</td>
<td>University of Alabama</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;R36A&lt;/sub&gt;</td>
<td>R36A&lt;sup&gt;a&lt;/sup&gt; (Nonencapsulated strain)</td>
<td>1 (2)</td>
<td>Rohatgi et al&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>PspATIGR4&lt;sub&gt;TIGR4&lt;/sub&gt;</td>
<td>TIGR4&lt;sup&gt;a&lt;/sup&gt; (4)</td>
<td>2 (3)</td>
<td>This study</td>
</tr>
<tr>
<td>pUAB100&lt;sub&gt;BG9300&lt;/sub&gt;</td>
<td>JCP#56 (8)</td>
<td>2 (4)</td>
<td>University of Alabama</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;ATCC6303&lt;/sub&gt;</td>
<td>ATCC 6303&lt;sup&gt;a&lt;/sup&gt; (3)</td>
<td>2 (5)</td>
<td>This study</td>
</tr>
<tr>
<td>pUAB104&lt;sub&gt;BG9300&lt;/sub&gt;</td>
<td>BG9300 (37)</td>
<td>3 (6)</td>
<td>University of Alabama</td>
</tr>
</tbody>
</table>

<sup>a</sup>Pneumococcal strains procured from American Type Culture Collection, USA  
<sup>b</sup>Constructs kindly provided by Prof. Susan K. Hollingshead, University of Alabama, USA  
<sup>c</sup>Rohatgi et al. 2009

![Table 2. Primers used for PCR amplification and expression cloning of PspA fragments (N-terminal only)<sup>b</sup>](image2)

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence of Primer (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS-951 (sense)</td>
<td>GAA GAA TCT CCA CAA GTT GTC</td>
</tr>
<tr>
<td>DS-952 (antisense)</td>
<td>CCC CCC AAG CTT TTA GCC TAA CTC ATT AAG AGC TGC</td>
</tr>
<tr>
<td>DS-953 (antisense)</td>
<td>CCC CCC AAG CTT CTA TTC TTC TTC TTC ACC GTC AGG</td>
</tr>
</tbody>
</table>

<sup>b</sup>HindIII (A/AGCTT) site was incorporated in the antisense primers (boldface) to facilitate directional cloning.

![Table 3. Bioinformatics analysis of secondary structure prediction in N-terminal domain of PspA from all six clades<sup>a</sup>](image3)

<table>
<thead>
<tr>
<th>Pneumococcal strain (PspA family / clade)</th>
<th>Number of amino acid analyzed</th>
<th>α-helix (%)</th>
<th>β-bridge (%)</th>
<th>Extended strand (%)</th>
<th>β-turn (%)</th>
<th>Random coil (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS2016 (1 / 1)</td>
<td>1-276</td>
<td>87.68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12.32</td>
</tr>
<tr>
<td>R36A (1 / 2)</td>
<td>1-288</td>
<td>85.76</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14.24</td>
</tr>
<tr>
<td>TIGR4 (2 / 3)</td>
<td>1-419</td>
<td>84.01</td>
<td>0</td>
<td>1.43</td>
<td>0</td>
<td>14.56</td>
</tr>
<tr>
<td>JCP#56 (2 / 4)</td>
<td>1-352</td>
<td>88.35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11.65</td>
</tr>
<tr>
<td>ATCC 6303 (2 / 5)</td>
<td>1-360</td>
<td>84.44</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15.56</td>
</tr>
<tr>
<td>BG9300 (3 / 6)</td>
<td>1-388</td>
<td>70.62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29.38</td>
</tr>
</tbody>
</table>

<sup>a</sup>Secondary structure analysis was carried out by PHD program of all clades of PspA (N-terminal only) using MacVector software which confirmed the expected high α-helical content of all PspA.

![Table 4. Total antibody response elicited in mice immunized with recombinant PspAs<sup>a</sup>](image4)

<table>
<thead>
<tr>
<th>Recombinant PspAs</th>
<th>PspA family (clade)</th>
<th>Number of mice&lt;sup&gt;b&lt;/sup&gt;</th>
<th>End point titer&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PspA&lt;sub&gt;L82016&lt;/sub&gt;</td>
<td>1 (1)</td>
<td>3</td>
<td>8.03 X 10^7</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;R36A&lt;/sub&gt;</td>
<td>1 (2)</td>
<td>3</td>
<td>2.1 x 10^7</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;TIGR4&lt;/sub&gt;</td>
<td>2 (3)</td>
<td>3</td>
<td>1.47 x 10^6</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;JCP#56&lt;/sub&gt;</td>
<td>2 (4)</td>
<td>3</td>
<td>3.4 x 10^6</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;ATCC6303&lt;/sub&gt;</td>
<td>2 (5)</td>
<td>3</td>
<td>1.7 x 10^7</td>
</tr>
<tr>
<td>PspA&lt;sub&gt;BG9300&lt;/sub&gt;</td>
<td>3 (6)</td>
<td>3</td>
<td>3.1 x 10^7</td>
</tr>
</tbody>
</table>

<sup>a</sup>Mice were immunized subcutaneously with 25 μg/mouse recombinant PspA belongs to different clades at day 0, 14 and 28. After one wk mice were bleed retro-orbitally and total antibody response were analyzed by ELISA in pooled serum.  
<sup>b</sup>Female, inbred Balb/c mouse strain  
<sup>c</sup>Endpoint titer was calculated as the reciprocal of the highest serum dilution that gave an absorbance reading greater than twice the value obtained with the preimmune sera (diluted 1:100).

4°C. The plate was washed with PBS containing 0.5% Tween-20 (PBST) and blocked with 2% BSA-PBST at 37°C for 1 hr. The plate was washed 3 times with PBST, serial dilutions of anti-PspAs sera was added and the plate incubated at 37°C for 1 hr. Bound antibodies were detected using using horseradish peroxidase-conjugate goat-anti-mouse Ig antibodies (Becton, Dickinson and Company, USA) (dilution 1 in 2000). Endpoint titer was plotted as the reciprocal of the highest serum dilution that gave an absorbance reading greater than twice the value obtained with the preimmune sera (diluted 1:100).
terminal only) constructs (L82016, R36A, TIGR4, JCP#56, ATCC6303, and BG9300) were compared and aligned the sequence data analyzed by the ClustalW algorithm using the MacVector version 10 software (Fig. 2). The multiple sequence alignment data revealed that the N-terminal region of PspA from all six clades highly variable and shares 42 - 79% amino acid sequence identity. The amino acids showed identities across the clades are highlighted (in pink) and similarity (in orange). The amino acids in green colour revealed the variability and dashed mentioned the gap in the region (Fig. 2). The secondary structures was analyzed by online PHD program to detect predominant structure for PspA which included seven-residue amino acid pattern (a-b-c-d-e-f-g) of the coiled-coil conformation in all PspA (Table 3). No other secondary structures were detected except small random coil between the α-helical coiled coil structures. The length of such coils varies for different strains of PspA by several amino acids. PspA from strain JCP#56 consist highest percentage (88.35%) of α-helical content, while PspA from strain BG9300 consist lowest (70.62%). However, the random coil content in all clades of PspA ranging from 11.65% to 15.56%, except PspA from BG9300 have highest percentage of random coil (29.38%). PspA from TIGR4 showed extended strand (1.43%) but other PspA not (Table 3).

**PspA fragments induce high titer antibodies in mice**

To assess the immunogenicity of recombinant PspA fragments, the groups of 3 mice were immunized with 3 dose of 25 µg of purified PspA fragments on day 0, 14 and 28 subcutaneously with aluminum hydroxide as an adjuvant, and the antibody responses were measured 7 days after the last immunization. The PspA specific endpoint antibody titer of the pooled hyperimmune sera was determined by ELISA. The data indicated that the immunization with purified PspA elicited high antibody titers in serum (Table 4).

**DISCUSSION AND CONCLUSION**

Protein based vaccine would be the alternate to provide the protection against the pneumococcal infection. PspA is one of the potential vaccine candidates and has been studied at great extent. One of the major challenge with PspA as a vaccine is the sequence variability and the antibody elicited to one PspA cannot be cross-reactive and cross-protective when challenge with heterologous strains (Miyaji et al. 2002). Hollingshead et al has been classified the complete PspA into 3 family and six clades based on amino acid sequence (Hollingshead et al. 2000). In this study, we performed cloning, expression, purification and in silico analysis of PspA N-terminal fragments to check the conservation label of PspA across the clades. The recombinant proteins were expressed well and easy to purify in highly pure form from soluble fraction (Fig. 1). In contrast, our bioinformatics results indicate that the PspA are highly variable and consists of highly charged residues at primary sequence of N-terminal domain in all 6 clades of PspA analyzed (Fig 2), while the conservation at secondary label was greater extent and predicted PspA to be predominantly α-helical coiled coil structure (Table 3). Humoral immune response is the major protective mechanism against pneumococcal infections. Protein antigen capable of eliciting high antibody titers can be a good protective antigen. In mice, PspA has been shown to elicit a good antibody response that protects against fatal challenge with encapsulated S.
pneumoniae (Briles et al. 2000, Nabors et al. 2000), and the protection-eliciting residues have been found to the α-helical N-terminal half of the protein (Briles et al. 1998). Thus, antibody induces in response to immunization with PspA fragments would unlikely to cross react and cross-protect. In conclusion, its structure, function, immunogenicity and method of antigenic variation are unique among known virulence molecules and may serve as a model for certain microbial systems.

Acknowledgments

This work was supported by a research grant from the Department of Biotechnology (DBT), Government of India. Naeeem Khan is a recipient of a Senior Research Fellowship from DBT.

Conflict of interest

We declare that we have no competing interests.

REFERENCES


*******