CMMB

ORIGINAL PAPERS

https://doi.org/10.31689/rmm.2025.32.4.403

Determination of Certain Immunological and Physicochemical Features of Pneumococcal Surface Protein A (PspA): *In-silico* Cloning of *pspA*

Enass Waad AL-HADIDI[1], Ali Adel DAWOOD[2*]

Abstract

Objective: Pneumococcal surface protein A (PspA) is present in almost all strains of pneumococci and is highly immunogenic, stimulating the production of antibodies. It has a role in pneumococcal pathogenicity. This study aimed to determine some immunological and physicochemical properties of PspA using bioinformatics software. Also, an *in silico* study was applied to clone the *pspA* gene by ligating it into pCRII-TOPO and transferring it into *E. coli* Bl21.Materials and Methods: PspA (WP_368082102) and *pspA* (U89711) were extracted from NCBI. The programs performed in this work included: PSIPRED, MemBrain 3.1, XtalPred, FUpred, CoCoPRED, Signal-CF, EzyPred, OSML, C-IMMSIM, Geneious Prime, AMUSER 1.0, PyMOL, and CCP4MG. Results: PspA has 5 distinct boundary domains and 8 coiled-coil domain (CCD) network. FU-score values were found between residues 430-540 and 780-810. Antibody production (IgM+ IgG) commences on the fifth day of infection. The IFN-g reached high peak concentration and stabilized between days 12 and 17, followed by IL-23. The pspA was successfully ligated in pCRII-TOPO and cloned in *E. coli* BL21. Conclusions: The PspA lacks enzymatic activity, which may pave the path for future research into developing an anti-PspA vaccination or discovering a medication that may negate its effects.

Keywords: Cloning; Domain; Immunological; PspA; Streptococcus

[1] M.Sc. Microbiology, Dept. of Biology, College of Science, University of Mosul, Mosul, Iraq
[2] Ph.D. Microbiology, Dept. of Anatomy, College of medicine, University of Mosul, Mosul, Iraq

***Corresponding author:**
Ali Adel DAWOOD, Ph.D. Microbiology, Dept. of Anatomy, College of medicine, University of Mosul, Mosul, Iraq
E-mail: aad@uomosul.edu.iq

INTRODUCTION

Pneumococci produce several proteins, including pneumococcal surface protein A (PspA), which helps them evade C3 complement deposition, making them virulent[1]. An important fact is that this protein is presented in almost strains of pneumococci and is a highly immunogenic, stimulating the production of antibodies. There are two main families of PspA, families 1-3. Family 2 is further divided into subgroups[2,3]. PspA consists of five domains: an amino-terminal signal peptide, an α-helical domain, a proline-rich region, a choline-binding domain, and a carboxyl-terminal tail. PspA's choline-binding domain anchors to the surface of lung cells via lipoteichoic acids, whereas its alpha-helical domain binds to host cell surfaces via an unknown receptor[4,5].

A previous study has revealed that antibodies against PspA and PspC variations carried by bacteria provide protection against infection. This shows that antibody-based infection protection may be genotype-specific, implying that adaptive responses targeting previously carried isolates may not impact newly acquired pneumococci[6,7].

Allele 17 of PspA is frequent enough to indicate that IgG levels plateau after a few doses. IgG levels to PspA were shown to be significantly linked in newborn samples, indicating that all women had a rich repertoire of IgG antibodies against PspA[8,9].

A previous study proposed a PspA-engineered vaccine, however additional studies and development are required. It revealed that the PspA protein has the potential to be turned into a human peptide vaccine. The majority of the pneumococcal strains isolated in Pakistan include genes for the PspA protein, which belongs to the family 2[1,10,11].

This study aimed to determine some immunological and physicochemical properties of PspA using bioinformatics programs. Also, an *in silico* study was applied to clone the *pspA* gene by ligating it into pCRII-TOPO and transferring it into *E. coli* Bl21.

MATERIALS AND METHODS

PspA (825aa) with the accession number (WP_368082102) was extracted from NCBI.

Predict the 2D and 3D structure of PspA

The secondary structure was established utilizing the PSIPRED tool. The three-dimensional structure underwent an enhancement process through the application of MemBrain version 3.1.

Estimate bio-physiochemical properties

Analysis of the physicochemical properties was conducted via the XtalPred server. In order to elucidate the binding domain, the FUpred algorithm was employed. CoCoPRED was utilized to ascertain a comprehensive prediction of the three specific features, oligomeric state, log network, and coiled-coil domain (CCD) network, which are indicative of the structural characteristics of the helix. The Protein Peeling 3.0 was engaged to decompose the protein into a finite array of modules that do not yield a singular protein module. Generally, for the analyzed protein fragments, the density of connections within the protein module is maximized, whereas the inter-module connections are minimized. The employed algorithm serves to detect a series of overlapping sequence partitions and constructs a hierarchical tree that delineates the sequence partitioning of the protein modules, thereby facilitating the identification of the secondary and tertiary structures in relation to the protein. The following equation was used for the measurement:

$$PI = \frac{AB - C^2}{(A+C)\,(B+C)}$$

PI: Partition index, *A*, *B*: Intra contact, *C*: Inter contact.

Protein-Sol was used to predict the measured solubility value of PspA when the mean of the experimental data set at 0.45. Thus, any measured solubility value greater than 0.45 is expected to have a higher solubility than the mean of the *E. coli* protein standard within the experimental solubility data set. Signal-CF server, an automated technique is used to predict signal peptide sequences and their cleavage sites in bacterial and eukaryotic sequences. EzyPred was performed to ascertain if PspA had enzyme qualities or not. However, OSML was examined in an attempt to find a ligand that may attach to PspA.

Estimate immunological properties

To determine molecular simulation of the immune system to PspA, the C-IMMSIM server is adopted.

***In silico* cloning**

pspA was selected from GenBank with ID (U89711) and cloned using Geneious Prime. The pCRII-TOPO

plasmid (3973bp) was chosen because it involves a single step in which DNA fragments are inserted into linear vectors using topoisomerase I. This plasmid has an ampicillin and kanamycin resistance marker. Primers were created with AMUSER 1.0, and commercial restriction enzymes were selected based on the cutting site.

RESULTS

PSIPRED gives the outcomes of a secondary structure predictor when the structure is built on a neural network and executed by a single sequence structure in the same formats. Figure (1) depicts the query sequence with secondary structure animations and calculates the confidence value at each place of alignment sequences using a series of blue bar graphs for each protein fragment. The 3D structure of PspA was constructed using MeMBrian methods, and the connection between residues was computed to rather weak degrees, Figure (2). PyMOL was used to visualize the cartoon in the mesh of PspA, and CCP4MG was used to determine the correlation between residues.

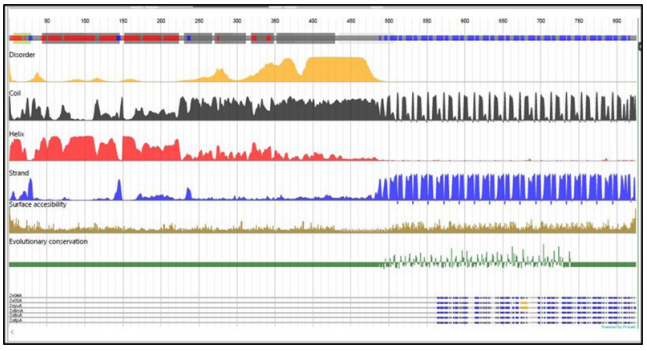


Fig 1. Description of parts of PspA constructed in 2D structure.

The expected physical parameters of PspA are provided along with matching crystallization probability distributions derived from the target database. Figure (3) depicts the distributions of crystallization failures and successes in learning sets obtained from the algorithm. The blue box calculates the right observed values, whereas the gray box estimates the unsuccessful observations. In each graph, a number in red represents the best-computed value for each parameter, which referes protein length, isoelectric point, gravity, instability index, coil structure percentage, insertion score, surface entropy, hydrophobicity, and ruggedness. The schematic depiction of PspA showed 5 distinct boundary domains, with the lowest FU-score values found between residues 430-540 and 780-810, Figure (4-A and B). Figure (4-C) depicts a heatmap distributing bounding domains.

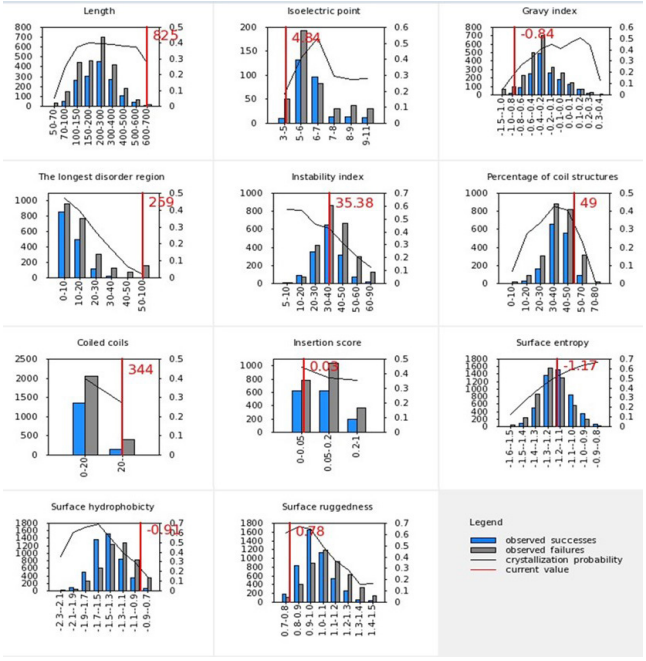


Fig 3. Estimation of physical parameters of PspA.

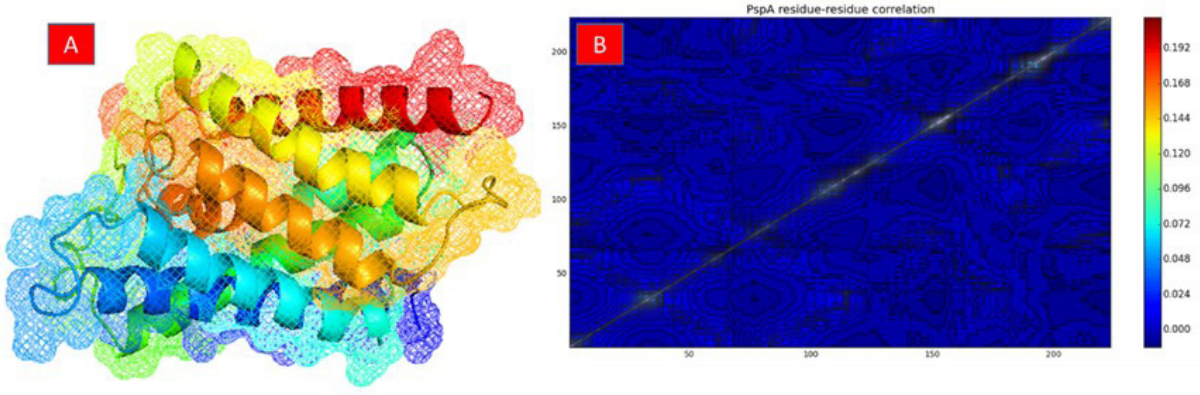


Fig 2. 3D structure of PspA. A: Cartoon in the mesh model of constructed protein. B: The degree of PspA reside-residue correlation.

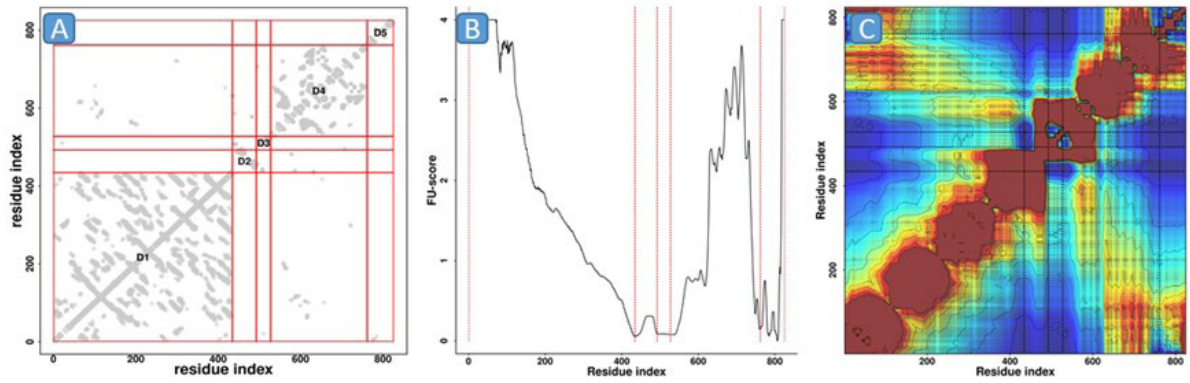
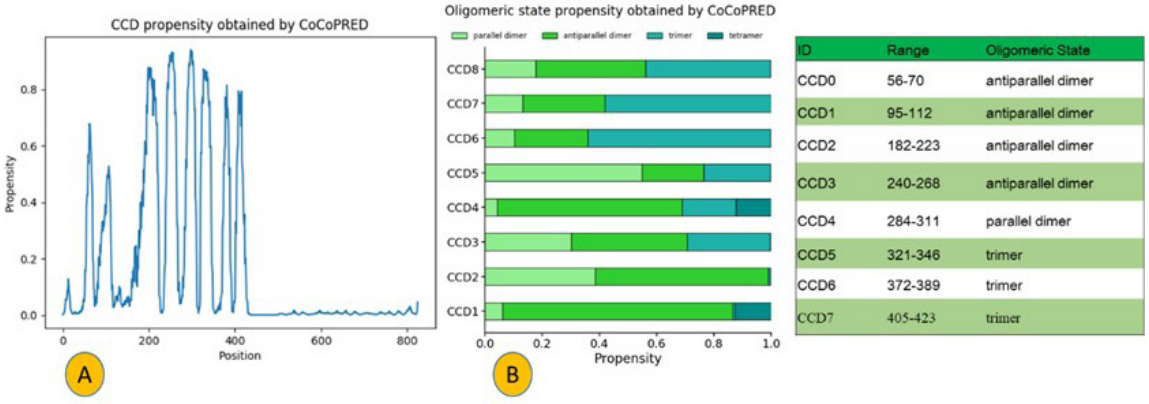


Fig 4. 3A: Distribution of boundary domains in contact map. B: FUscore curve of the total PspA. C: Heat map of PspA.

The 3D structure of PspA is made up of four helices that are twisted around each other. This relationship is important in biological functions. The CCD propensity curve of PspA begins at residue 5, increases to a peak at residue 318, and then stabilizes at residue 428 (Figure 5-A). The presence of 8 convolutional layers in PspA's CCD network was found and classified into four types based on their placement regions: parallel dimer, antiparallel dimer, trimer, and tetramer Figure 5-B. Other biological traits revealed are the minimal size of the secondary structure is 8. The minimum size of a protein unit is 16, whereas the maximum size is 0. The total protein cut-off value is 0.2, with an atom cut-off distance of 8.0. The delta value for a logistic function is 1.5. Figure 6 represents the distribution of PspA units as well as the hierarchical splitting process. The solubility of the protein sequence algorithm was estimated. The values of deviation from the population average range from -2 to 2. The windowed charge score for each amino acid ranges between -0.4 and 2. The windowed fold propensity per amino acid ranges between -0.5 and 0.5, Figure 7. In this study, no ligand was identified that could bind to PspA and this protein does not have any enzymatic properties.

The antigenicity of PspA may be evaluated, and the immune system can be induced to generate cytokines. The antigen level was shown to be high during the first 5 days of infection with *S. pneumoniae*. Antibody production (IgM + IgG) commences on the fifth day of infection, peaks on the sixteenth day, and subsequently declines. Figure (8-A) shows that IgM and IgG1 are elevated less than (IgM+IgG). Interleukin production begins in the early stages of infection, with IFN-g reaching its peak concentration and stabilizing between days 12 and 17, followed by IL-23. TGF-b is the most poorly produced cytokine, with a concentration of zero (Figure 7-B). Figure (8-C) shows that the number of natural killer cells increases from the first day of infection, peaks on day 12, then declines to zero on day 30.

The nucleotide sequence of the *pspA* gene (2046bp) was obtained from NCBI under the identification (GenBank: U89711). AMUSER 1.0 was used to choose the proper PCR primers, as shown in Table (1). In vitro investigations rely heavy on *In-silico* cloning. Several restriction enzymes were determined that can cut the *pspA* at different points along the TOPO



ID	Range	Oligomeric State
CCD0	56-70	antiparallel dimer
CCD1	95-112	antiparallel dimer
CCD2	182-223	antiparallel dimer
CCD3	240-268	antiparallel dimer
CCD4	284-311	parallel dimer
CCD5	321-346	trimer
CCD6	372-389	trimer
CCD7	405-423	trimer

Fig 5. Biological properties of PspA. A: CCD propensity curve. B: Oligomeric propensity state degree.

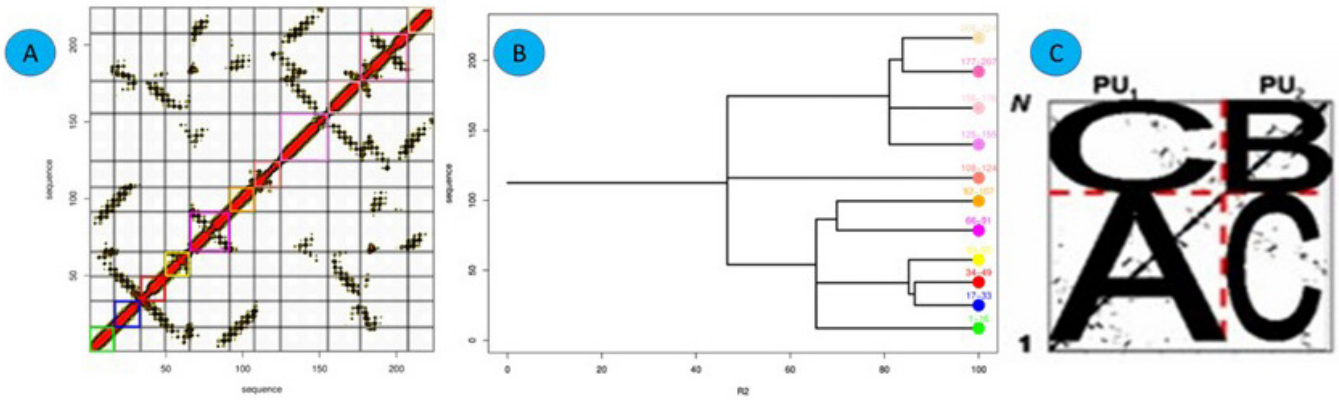


Fig 6. Illustrate the principle of PspA peeling. A: Probability contact matrix of PspA modules. B: Hierarchical process of splitting units of PspA. C: Distribution of Intra and inter-contact modules in the matrix map.

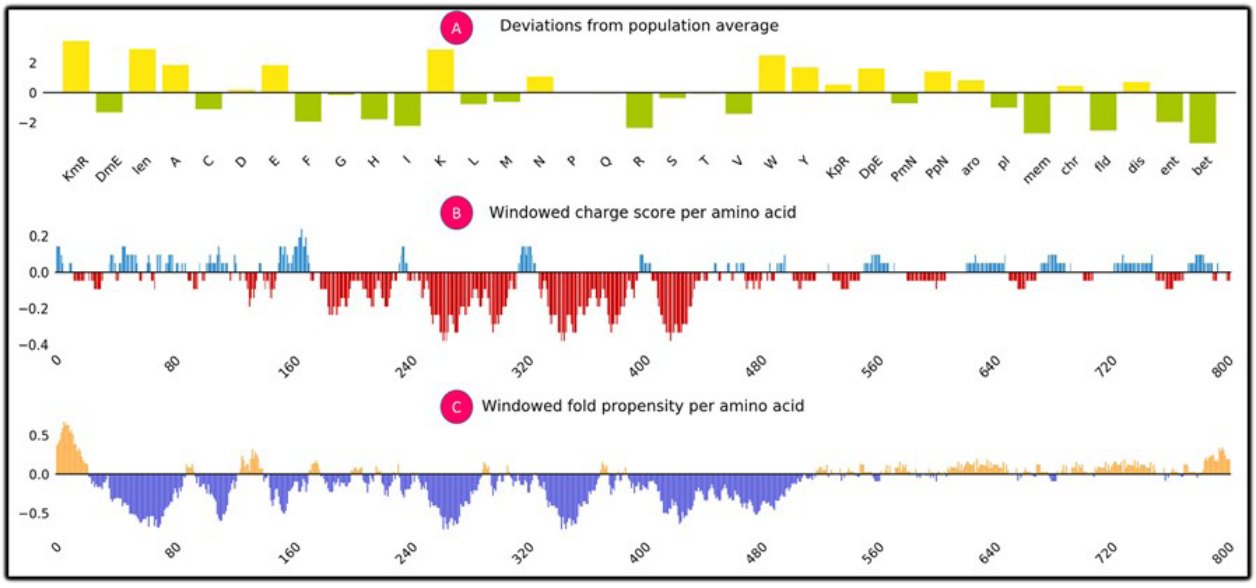


Fig 7. Solubility parameters of PspA. A: Deviation of population average. B: Charge score per a.a. C: Fold propensity per a.a. Description of amino acids: KmR = K minus R, DmE=D minus E, KpR = K plus R, DpE = D+E, PmN = K+R-D-E, PpN = K+R+D+E, aro = F + W + Y.

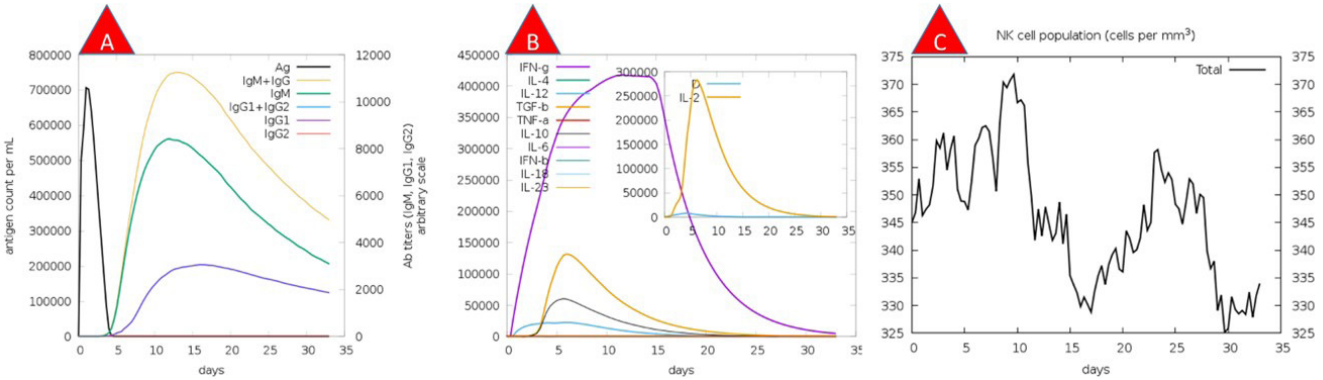


Fig 8. Response of immune system parameters. A: Following infection, antigen and antibody levels peak. B: Different types of interleukins produced in response to infection. C: Level of NK cell production following infection.

sequence. The two best restriction enzymes (BseRI and BsgI) were selected because they contained the highest proportion of *pspA* that could be transferred into *E. coli* BL21 in gel electrophoresis test, Figure (9). Figure (10) illustrates the clone pspA-pCRII TOPO map.

Table 1. Designing primers for PCR test

Strand	TM	Primer	GC%
Forward	50.0C	AGCCGTUGACAAATATTTACGGAGGAG	44.4
Reverse	46.6C	AACGGCUTAAACCCATTCACCATTG	44

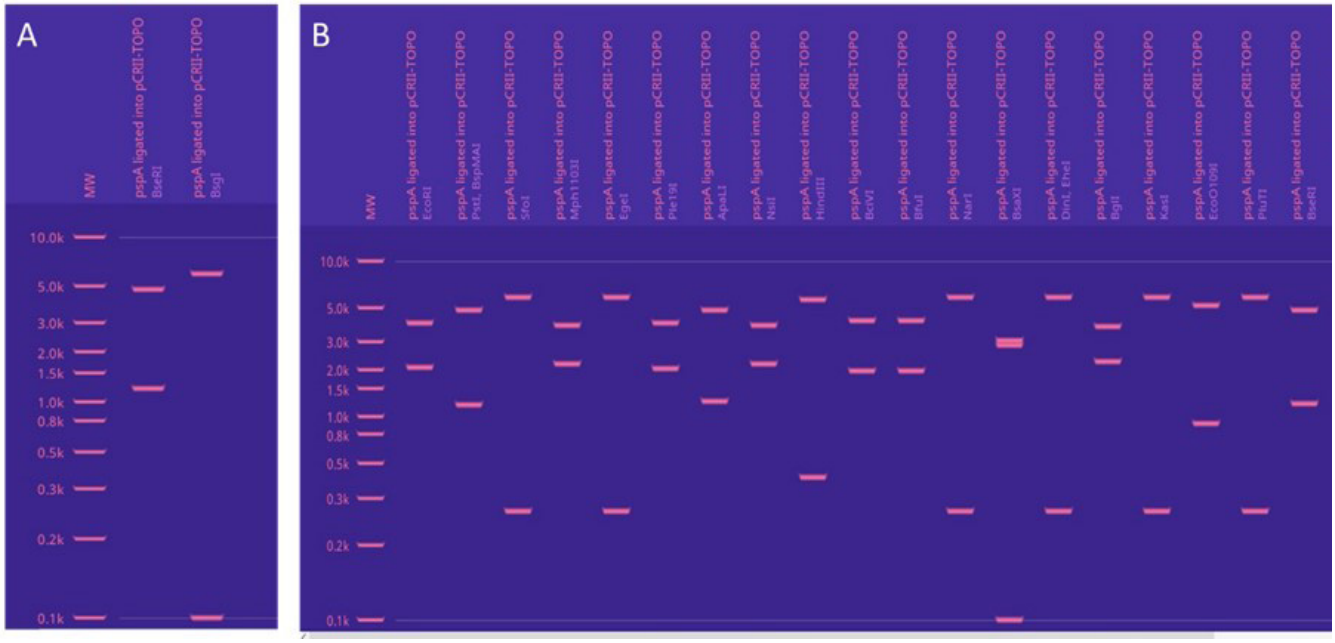


Fig 9. Gel electrophoresis of pspA-pCRII TOPO clone. A: the best restriction enzymes (BseRI and BsgI) for getting a high rate of pspA. B: Types of restriction enzymes that can cut different distances.

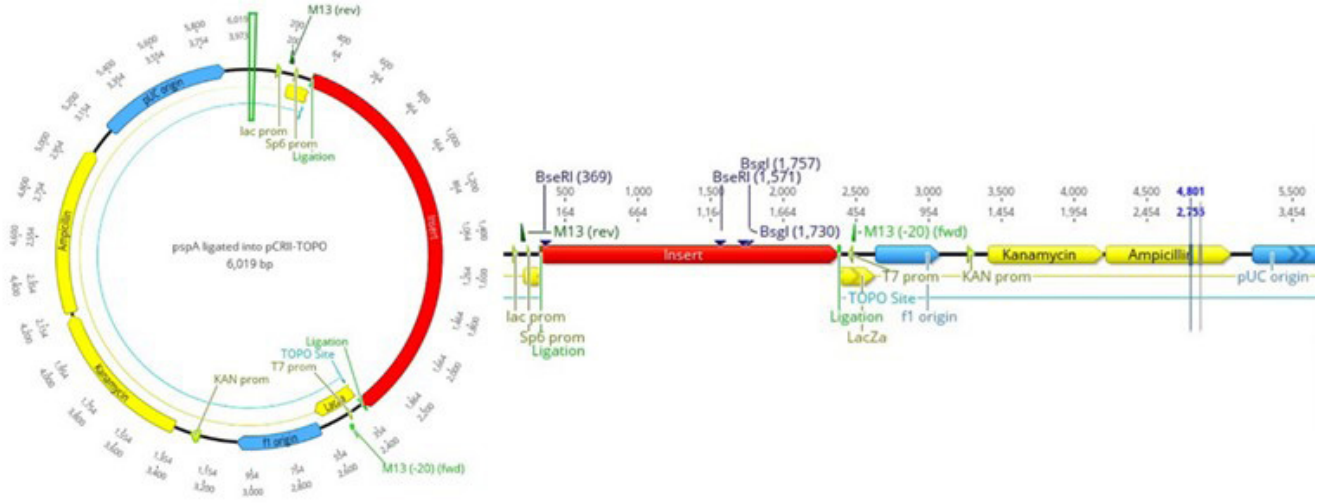


Fig 10. Schematic diagram of pspA-pCRII TOPO clone (6019bp). The insertion part of the red color in the TOPO site is the pspA.

DISCUSSION

PspA is a protein virulence factor for *S. pneumoniae* that is exposed to the surface. Several studies have used animals lacking in complement C5, C3, or factor B to study PspA's role in pneumococcal pathogenicity. Significant serum complement activation was seen in mice infected with a homozygous strain of pneumococci lacking PspA within 30 minutes of infection. PspA-negative pneumococci were very infectious in animals lacking C3 or factor B. PspA-deficient pneumococci. These contradictory findings are due to PspA interfering with complement-dependent host defense mechanisms mediated by factor B in immunocompetent mice infected with pneumococci[12,13].

Furthermore, PspA promotes virulence by interfering with the deposition of C3b on pneumococci and/or preventing the formation of a fully functional alternate pathway of the C3 convertases. PspA reduces the amount of C3b deposited on pneumococci by blocking the recruitment of the alternative route, hence affecting the effectiveness of complement receptor-mediated clearance routes[14,15].

PspA inhibits host cell recognition by C-reactive protein (CRP). This is accomplished by attaching to phosphorylcholine residues on the bacterial surface that are recognized by CRP. PspA also binds to human lactoferrin and neutralizes lactoferricin's antibacterial action through an unknown mechanism. PspA may potentially function as a bacterial adhesin on dying host cells, where it binds to inverted phosphatidylserine residues on the cell membrane of apoptosis-prone lung epithelial cells[16-18]. A previous study found that PspA and GAPDH interactions in host cells during pneumococcal infection influenced bacterial location in the lower airways[17,19]. The GAPDH-dependent assembly of PspA that occurs during bacterial colonization of host cells may shield them from nutrients released by the related dead host cells[20].

Although there are key physicochemical qualities referred to as boundary domain in this study. There are properties that altered the protein's properties, such as isoelectric points, gravity, and degree of insertion. Furthermore, our findings revealed that the CCD was high for over 400 PspA residues. Protein peeling plays a crucial role in analyzing the impact of PspA fragments on secondary structure and converting contact matrix data into contact probability when enabled. These parameters were utilized to generate the contact probability matrix used in the peeling process.

Another aspect that causes PspA to lose functionality and influences its selection as an immunostimulator or vaccine candidate is a lack of ligand binding. In addition, this study revealed that PspA had no enzymatic activity.

Given the relevance of PspA in pathogenesis, this study opted to clone the pspA gene in pCRII-TOPO and analyze its chemical, biophysical, and immunological characteristics as a first step toward developing an effective vaccine against *S. pneumoniae*.

Due to the importance of PspA and the reasons indicated above, our study opted to do a thorough examination of the entire protein to determine its physical, chemical, and immunological characteristics. This is significant because it paves the path for future research into developing an anti-PspA vaccination or identifying a medication that may negate its effects. It may also offer up new paths for battling *S. pneumonia* or developing strategies to prevent antibiotic resistance.

CONCLUSIONS

Although PspA has important physicochemical features known as the boundary domains, there are other properties that can change the protein's properties. When enabled, PspA peeling is critical for studying the influence of fragments on secondary structure and turning contact matrix data into contact probability. Another factor that causes PspA to lose activity and influences its selection as an immunostimulant or vaccine candidate is a loss of ligand binding. PspA possesses no enzymatic activity. This discovery might pave the path for future research into developing an anti-PspA vaccination or discovering a medication that can neutralize its effects.

Ethics Statement and Conflict of Interest Disclosures

Financial support and sponsorship: All authors have declared that no financial support was received from any organization for the submitted work.

Ethics Consideration: The authors declare that all the procedures and experiments of this study respect the ethical standards in the Helsinki Declaration of 1975, as revised in 2008(5), as well as the national laws. Written informed consent was provided by the patient participant in this study.

This study was approved by the Institutional Research Board and Ethics Committee.

Conflict of interest: No known conflict of interest correlated with this publication.

Availability of data and materials: The data used and/ or analyzed throughout this study are available

from the corresponding authors upon reasonable request.

References

1. Khan F, Khan MA, Ahmed N, Khan MI, Bashir H, and Tahir S, et al. Molecular Characterization of Pneumococcal Surface Protein A (PspA), Serotype Distribution and Antibiotic Susceptibility of Streptococcus pneumoniae Strains Isolated from Pakistan. *Infect Dis Ther*. 2018 J;7(2):277-289. doi: 10.1007/s40121-018-0195-0.
2. Briles DE, Hollingshead S, Brooks-Walter A, et al. The potential to use PspA and other pneumococcal proteins to elicit protection against pneumococcal infection. Vaccine. 2000;18(16):1707–11.
3. Hollingshead SK, Becker R, Briles DE. Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae. Infect Immun. 2000;68(10):5889–900.
4. Holmlund E, Quiambao B, Ollgren J, Nohynek H, and Käyhty H. Development of natural antibodies to pneumococcal surface protein A, pneumococcal surface adhesin A and pneumolysin in Filipino pregnant women and their infants in relation to pneumococcal carriage. *Vaccine*. 2006; 24, 57–65.
5. Croucher NJ, Campo JJ, Le TQ, Pablo JV, Hung C, and Teng AA, et al. Genomic and panproteomic analysis of the development of infant immune responses to antigenically-diverse pneumococci. *Nat Commun*. 2024;15(1):355. doi: 10.1038/s41467-023-44584-2.
6. Francis JP, Richmond PC, Pomat WS, Michael A, Keno H, and Phuanukoonnon S, et al. Maternal antibodies to pneumolysin but not to pneumococcal surface protein A delay early pneumococcal carriage in high-risk Papua New Guinean infants. *Clin Vaccine Immunol*. 2009;16(11):1633-8. doi: 10.1128/CVI.00247-09.
7. Moreno AT, Oliveira ML, Ferreira DM, Ho PL, Darrieux M, and Leite LC, et al. Immunization of mice with single PspA fragments induces antibodies capable of mediating complement deposition on different pneumococcal strains and cross-protection. *Clin Vaccine Immunol*. 2010 ;17(3):439-46. doi: 10.1128/CVI.00430-09.
8. Li J, Glover DT, Szalai AJ, Hollingshead SK, Briles DE. PspA and PspC minimize immune adherence and transfer of pneumococci from erythrocytes to macrophages through their effects on complement activation. *Infect Immun*. 2007;75(12):5877-85. doi: 10.1128/IAI.00839-07.
9. Ferreira DM, Darrieux M, Silva DA, Leite LC, Ferreira JM Jr, and Ho PL, et al. Characterization of protective mucosal and systemic immune responses elicited by pneumococcal surface protein PspA and PspC nasal vaccines against a respiratory pneumococcal challenge in mice. *Clin Vaccine Immunol*. 2009;16(5):636-45. doi: 10.1128/CVI.00395-08.
10. Darrieux M, Goulart C, Briles D, Leite LC. Current status and perspectives on protein-based pneumococcal vaccines. *Crit Rev Microbiol*. 2015;41(2):190-200. doi: 10.3109/1040841X.2013.813902.
11. Haleem KS, Ali YM, Yesilkaya H, Kohler T, Hammerschmidt S, and Andrew PW, et al. The Pneumococcal Surface Proteins PspA and PspC Sequester Host C4-Binding Protein to Inactivate Complement C4b on the Bacterial Surface. *Infect Immun*. 2018;87(1): e00742-18. doi: 10.1128/IAI.00742-18.
12. Hendriaaningtyas M, Noer E. Product and Insulin Resistance Insulin in Diabetes Mellitus Type 2. *Moderna Med*. 2025; 32(1): 73-79. https://doi.org/10.31689/rmm.2025.32.1.73
13. Al-Waely H, Al-Ahmer A. Evaluation of the Role of Some microRNAs in Corona Virus-19 Infection in Sample of Iraqi Patients. *Moderna Med*. 2024; 31(3): 237-242. https://doi.org/10.31689/rmm.2024.31.3.233
14. Florescu C, Vasiliu O, Prelipceanu. Psychiatric-Oncologic Interferences in the Respiratory Pathology of Lung Cancer. *Moderna Med*. 2024; 31(3): 203-208. https://doi.org/10.31689/rmm.2024.31.3.203.
15. Chang B, Kinjo Y, Morita M, Tamura K, Watanabe H, and Tanabe Y, et al. Distribution and Variation of Serotypes and Pneumococcal Surface Protein A Clades of Streptococcus pneumoniae Strains Isolated From Adult Patients With Invasive Pneumococcal Disease in Japan. *Front Cell Infect Microbiol*. 2021; 11:617573. doi: 10.3389/fcimb.2021.617573.
16. Amani A. Alayash; Firas M. Abed. "The novel genomic DNA integrity in neutral buffer formalin, Bouin's fluid, and zinc formalin fixatives: A comparative study". Iraqi Journal of Veterinary Sciences, 38, 3, 2024, 715-719. https://doi.org/10.33899/ijvs.2024.148257.3571.
17. Dawood A.A. Designing an immuno-epitope candidate vaccine from (Opa, ProA, ProB, RmpM and BamD) proteins against Neisseria gonorrhoeae and Neisseria meningitides. Vacunas. 2024;25(4): 481-491. https://www.sciencedirect.com/science/article/abs/pii/S1576988724000736
18. Dawood A.A. A Method Utilizing an Image Visibility Graph to Portray the Arrangement of Genomic Data Sequencing, Gene Frequencies for The Peptidoglycan-Associated Lipoprotein (Pal) Gene in Brucella Spp., and Prevalence of Brucellosis in Nineveh. Moderna Med. 2024; 31(2): 155- 165. https://doi.org/10.31689/rmm.2024.31.2.155.
19. Muñoz-Elías EJ, Marcano J, Camilli A. Isolation of Streptococcus pneumoniae biofilm mutants and their characterization during nasopharyngeal colonization. *Infect Immun*. 2008;76(11):5049-61. doi: 10.1128/IAI.00425-08.
20. Brissac T, Martínez E, Kruckow KL, Riegler AN, Ganaie F, and Im H, et al. Capsule Promotes Intracellular Survival and Vascular Endothelial Cell Translocation during Invasive Pneumococcal Disease. *mBio*. 2021;12(5):e0251621. doi: 10.1128/mBio.02516-21.