Identification of B-cell Epitope Regions in Cell Surface Proteins of Streptococcus pneumoniae Serotype 19F Using Bioinformatic Tools

Shirin Tarahomjoo*

Division of Genomics and Genetic Engineering, Department of Biotechnology and Central Laboratory, Razi Vaccine and Serum Research Institute, Karaj 31975/148, Iran

Email: starahomjoo@gmail.com

Abstract

Pneumococcal conjugate vaccines (PCVs) were developed through chemical coupling of polysaccharide capsules of pneumococci to immunogenic carrier proteins and World Health Organization recommends inclusion of these vaccines in national immunization programs for children. However, the PCVs implementation in developing countries is prevented by the high vaccine manufacturing costs. This issue can be resolved by construction of protein based vaccines against pneumococci. We already identified three pneumococcal surface proteins including autolysin, zinc binding lipoprotein (ZBL), and plasmid stabilization protein (PSP) as appropriate protein candidates for eliciting protection against S. pneumoniae serotype 19F. The protein protective antigenicity and the absence of autoimmunity induction were used as selection criteria. However, regarding the necessity of the antibody response for protection against pneumococci, analysis of protective B-cell epitopes of these proteins is required to elucidate their usefulness in new vaccine formulations. In the present study, therefore, we aim to identify protective B-cell epitope regions of these proteins via widely used bioinformatic tools. Both of the Bepipred and BCPreds programs were used for identification of linear B-cell epitopes. The conformational B-cell epitopes were predicted using the CBTope program. The immunoprotective abilities of epitopes were evaluated using VaxiJen. We determined the linear B-cell epitope regions, which were predicted by both Bepipred and BCPreds and have common amino acids with conformational B-cell epitopes. Our results showed that all of the three studied proteins included such protective overlapped linear B-cell epitope regions. However, a truncated form of PSP had the greatest number of the protective overlapped B-cell epitope regions.

----

* Corresponding author.

Moreover, the most probable protective epitope region resides in this protein and this epitope region is completely conserved in PSPs of serotype 19F pneumococcal strains. Therefore, the truncated PSP was an appropriate candidate for development of protein based vaccines against S. pneumoniae serotype 19F.

Keywords: B-cell Epitopes; Pneumococcal serotypes; Pneumococcal vaccines; Streptococcus pneumoniae.

1. Introduction

Streptococcus pneumoniae is a major pathogen that causes diseases such as pneumoniae, meningitis and sepsis [1]. World Health Organization (WHO) estimated that 476000 annual deaths among children less than 5 years of age were caused by pneumococcal infections [2]. S. pneumoniae serotype 19F is among the main pneumococcal serotypes that cause most invasive pneumococcal disease in children less than 5 years of age globally [3]. Pneumococcal vaccines have been used for protection against pneumococcal infections. The polysaccharide capsules of pneumococci are main antigenic components of these vaccines. However, the capsules are poorly immunogenic in children and are not able to induce anamnestic antibody responses upon revaccination. In order to circumvent these issues, the capsules are chemically conjugated to immunogenic carrier proteins. Therefore, the capsules are converted from T-cell independent antigens to T-cell dependent antigens, which enhance antibody responses and induce the immune memory. WHO recommends inclusion of pneumococcal conjugate vaccines (PCVs) in national immunization programs for children [2]. However, high manufacturing costs of PCVs limit their implementation in developing countries [1,2]. Development of protein based vaccines against pneumococci offers a more affordable protective strategy against pneumococcal infections. Cell surface proteins are key factors in infectious processes of pathogens and have extensively been evaluated as vaccine candidates [4,5]. We already identified three pneumococcal surface proteins including autolysin, zinc binding lipoprotein (ZBL), and plasmid stabilization protein (PSP) as protein candidates for eliciting protection against S. pneumoniae serotype 19F. The candidate selection was based on the protective antigenicity, and the absence of autoimmunity induction [6]. Antibody responses are required to protect against pneumococcal infections [7]. Thus, analysis of protective B-cell epitopes in the candidates is necessary to evaluate the effectiveness of these proteins as new vaccine constituents. Compared with conventional laboratory methods, computational approaches offer the ability to undertake rapid and comprehensive epitope assessments for vaccine candidates at much lower costs [8]. In the present study, therefore, we analyzed autolysin, ZBL, and PSP via widely used bioinformatic tools for identification of immunoprotective B-cell epitope regions. To our knowledge, there is no report concerning recognition of B-cell epitope regions of these proteins for the purpose of vaccine development against S. pneumoniae serotype 19F.

2. Materials and Methods

The NCBI accession numbers of autolysin, ZBL, and PSP were WP_000405240, WP_000724068, and WP_00034749 respectively. PSP was indicated in our previous work [6] as cell wall surface anchor family protein under accession number YP_002743140. Conserved domains of proteins were specified using Pfam. Protein sequences were submitted to the Bepipred and BCPreds programs for identification of linear B-cell epitopes. The immunoprotective abilities of selected B-cell epitopes were evaluated using VaxiJen. A higher

VaxiJen score refers to a higher probability for protective ability [9]. The CBTope program was used to predict conformational B-cell epitopes using amino acid sequences of the protein.

COBALT alignment tool was used for the epitope region conservancy analysis in PSPs of serotype 19F pneumococcal strains. Accession numbers of PSPs were KKW85405 (S. pneumoniae CU-SPNE1-05), KKW83972 (S. pneumoniae CU-SPNE32-06), AFC95386 (S. pneumoniae ST556), and WP_000873519 (S. pneumoniae G54). PHD program was used for the secondary structure prediction. SCRATH was used for predicting the protein solubility upon overexpression in Escherichia coli. The default settings were applied to all the tools used.

3. Results and Discussion

3.1. Identification of protective B- cell epitope regions of autolysin

Antibodies bind specifically to a continuous amino acid sequence of a protein known as the linear B-cell epitope or to a folded structure formed by discontinuous amino acids known as the conformational B-cell epitopes. The majority of B-cell epitopes are conformational [10]. Nevertheless, the identification of linear B-cell epitopes has demonstrated promising results for selection of the vaccine antigens [11,12]. Therefore, in this study, we determined both types of the B-cell epitopes in the proteins. Bacterial autolysins are enzymes capable of hydrolyzing peptidoglycan and have a major role in concealing this inflammatory molecule from the immune system [13].

Table 1: Protective linear B-cell epitopes of autolysin identified by Bepipred

Position[1]	Epitope sequence	VaxiJen score
13	PQVGVQPY	1.3051
24	HAHSTGNPHSTVQNEADYHWRKDPE	1.0542
66	PVDNGAWDVGGGWNAETY	0.4215
135	YCTNNQPNNHSDHVDPYPY	0.5565
285	QSADGTGW	1.4423
294	YLKPDGTLADKPEFTVEPD	1.3907

1: amino acid residue number at the epitope beginning is indicated.

Bepipred and BCPreds are widely used bioinformatic tools for the linear B-cell epitope identification. The Simultaneous application of different linear B-cell epitope prediction tools results in greater accuracy of the epitope identification. Therefore, we used both of these tools to determine the B-cell epitope regions in autolysin. Nine linear B-cell epitopes were identified in the protein using the Bepipred program. VaxiJen results indicated that 6 epitopes among them were immunoprotective (Table 1). In addition, BCPreds identified 7 linear B-cell epitopes. Considering VaxiJen results, 5 epitopes among them were protective (Table 2).

Table 2: Protective linear B-cell epitopes of autolysin identified by BCPreds

Position[1]	Epitope sequence	Vaxijen score
12	LPQVGVQPYRQVHAHSTGNP	1.161
134	EYCTNNQPNNHSDHVDPYPY	0.4469
221	RKHTDGNWYWFDNSGEMATG	0.6321
251	FNEEGAMKTGWVKYKDTWYY	0.7268
283	FIQSADGTGWYYLKPDGTLA	1.0224

1: The amino acid residue number at the epitope beginning is indicated.

The CBTope program uses the protein amino acid sequence for the prediction of conformational B-cell epitopes. Using this program, 8 conformational B-cell epitope regions were identified in the protein (Table 3). The epitope regions beginning at the amino acid residue 97 and 172 were exclusively conformational, whereas the rest of the conformational B-cell epitope regions overlapped with the identified linear B-cell epitopes.

Table 3: Conformational B-cell epitope regions of autolysin

Position[1]	No. of epitope residues	Epitope region sequence [2]
1	20	MEINVSKLRT DLPQVGVQPYRQ
36	7	QNEADYH
97	7	EFMT DYRLYI
131	22	KTHEYCTNNQPNNHSDHVDPYPYLAKWG
172	13	LTIETGWQK NDTGYW
207	31	YYFDSSGYMLADRWRKHTDGNWYWFDNSGEMATGW
254	21	EGAMKTGWVKYKDTWYYLDAKEGAMVSNAFIQSA
300	4	TLAD

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: Amino acid residues of conformational B-cell epitopes in each region are shown underlined.

The linear B-cell epitopes determined by Bepipred and BCPreds were overlapped at three regions, which included amino acids common with the conformational B-cell epitopes (Table 4). Regarding the VaxiJen scores, the epitope region beginning at the amino acid 283 was the most probable protective linear B-cell epitope region of autolysin.

Table 4: Protective overlapped linear B-cell epitope regions of autolysin

Position [1]	B-cell epitope region [2,3]	VaxiJen score
12	*LPQVGVQPYRQ*VHAHSTGNP	1.1610
134	*EYCTNNQPNNHSDHVDP*YPY	0.4469
283	**FIQSADGTGWYYLKPDG*TLAD*KPEFTVEP**	1.2326

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: Amino acids common with the conformational B-cell epitope regions are in italics.

3: The most probable linear B-cell epitope region is in bold.

3.2. Identification of protective B- cell epitope regions of ZBL

ZBL is a lipoprotein, which is involved in the bacterial zinc uptake. The signal peptide of ZBL is cleaved between amino acids 18 and 19 [14]. The mature protein obtained following cleavage of the signal peptide was used in the B-cell epitope analysis. Fifteen linear B-cell epitopes were identified in the protein using Bepipred. Thirteen epitopes among them were immunoprotective as indicated by VaxiJen results (Table 5). Eight linear B-cell epitopes were recognized in the protein using BCPreds. Considering VaxiJen scores, four epitopes among them were protective (Table 6).

Table 5: Immunoprotective linear B-cell epitopes of mature ZBL identified by Bepipred

Position [1]	Epitope sequence	VaxiJen score
1	CSNQKQADG	2.2763
26	QVAGDTA	1.3983
41	GTEPHEYEPSAKAVA	0..6642
100	GGEEEEGDHDHGEEGHHHEF	2.0596
168	KAYTEGLSQAKQ	0.5478
205	GLSPDAEPSAA	0.8943
247	SKEAGVK	2.3193
262	SLTEEDTKAGE	2.0618
287	QTTDQEGPAIEPEKAEDTKTVQNGYFEDAAVK	1.1307
354	KMTQAEYKAY	0.6471
365	TKGYQTDV	1.4374
385	VQGGQSKKY	2.6108
461	DNWPTYYPDNLSGQEI	1.0051

1: The amino acid residue number at the epitope start is indicated.

In the CBTope program, 11 conformational B-cell epitope regions were recognized in the mature ZBL (Table 7). The epitope region beginning at the amino acid residue 430 was exclusively conformational. The rest of the conformational epitope regions were overlapped with the identified linear B-cell epitopes.

Table 6: Immunoprotective linear B-cell epitopes of mature ZBL identified by BCPreds

Position[1]	Epitope sequence	VaxiJen score
58	QDADTFVYENENMETWVPKL	0.5112
99	PGGEEEEGDHDHGEEGHHHE	2.4164
285	LKQTTDQEGPAIEPEKAEDT	1.4277
348	KAKLTGKMTQAEYKAYYTKG	0.7859

1: The amino acid residue number at the epitope start is indicated.

Table 7: Conformational B-cell epitope regions of mature ZBL

Position[1]	No. of epitope residues	Epitope region sequence[2]
1	13	CSNQKQADGKLNI
26	12	QVAGDTANVELLIG
49	20	PSAKAVAKIQDADTFVYENENMETWV
141	23	LSADYPDKKE TFEKNAAAYIEKLQALDKAYTEGL
196	7	YGLKQVAISG
214	6	AARLAELTEYVKKN
249	8	EAGVKTDVLNPLESLT
286	40	KQTTDQEGPAIEPEKAEDTKTVQNGYFEDAAVKDRTLSDYAGNWQSVYPF
369	19	QTDVTKINITDNTMEFVQGGQ
430	5	QFSDH
459	20	EMDNWPTYYPDNLSGQEIAQEMLA

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: Amino acid residues of conformational B-cell epitopes in each region are shown underlined.

The linear B-cell epitopes identified in the mature ZBL by Bepipred and BCPreds were overlapped at two regions, which included amino acids common with the identified conformational B-cell epitopes (Table 8). The epitope region beginning at the amino acid 285 showed the highest VaxiJen score and was therefore the most probable protective linear B-cell epitope region of the protein.

Table 8: Overlapped linear B-cell epitope regions of mature ZBL

Position[1]	Epitope region[2,3]	Vaxijen score
285	***LKQTTDQEGPAIEPEKAEDTKTVQNGYFEDAAV***	1.0701
348	KAKLTGKMTQAEYKAYYTKGY*QTDV*	0.8959

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: Amino acids common with the conformational B-cell epitope regions are in italics.

3: The most probable protective B-cell epitope region is in bold.

3.3. Identification of protective B- cell epitope regions of PSP

PSP contains an YSIRK type family signal peptide, which is cleaved between the amino acids 36 and 37. Moreover, it contains an LPxTG motif beginning at the amino acid residue 658 [6]. We used the truncated PSP containing the amino acids 37-657 for the B-cell epitope analysis. The linear B-cell epitope analysis using Bepipred showed the presence of 15 linear B- cell epitopes in the protein. VaxiJen results indicated that all of the epitopes were protective (Table 9). BCPreds analysis identified 17 linear B-cell epitopes in the truncated PSP. Fourteen epitopes among them were protective as indicated by VaxiJen results (Table 10).

Table 9: Protective linear B-cell epitopes of truncated PSP identified by Bepipred

Position[1]	Sequence	VaxiJen score
2	EVVGETSSSKDNM	1.1265
24	STSLSAEPHSDTLSLD	0.9694
41	NSSLSTDNNGGVSINQNKPDITND	1.0703
94	YGAIGDGVNDDRQAIQ	0.7286
112	IDAAAQGLGGGNVYFP	1.1389
173	TDDGAQVEWGPTEDISYSGGTID	0.9767
199	ALNEEGTKAK	2.1738
290	YALNDDGKKSENVT	1.6011
309	FGKSDKSGE	3.0686
328	TLSTQNPS	1.0665
367	FDKKVKGESVHYRE	1.7546
444	INNNSKETEQP	1.7294
509	NANEKEPVIRDSDGN	0.9024
540	LSDKNEKEKNKEEKQSNSNNVIDSNQKNGEF NSSKDNRQMNDKIDNKQDNKTEEV	1.6322
599	VGDGRETENHIN	2.1816

1: The amino acid residue number at the epitope beginning is indicated.

Table 10: Protective linear B-cell epitopes of truncated PSP identified by BCPreds

Position[1]	Sequence	VaxiJen score
1	NEVVGETSSSKDNMQMNSLL	1.062
22	STSTSLSAEPHSDTLSLDNN	0.8067
55	NQNKPDITNDTMMGNISGME	0.7343
90	NVKDYGAIGDGVNDDRQAIQ	0.499
113	DAAAQGLGGGNVYFPEGTYL	0.4168
173	TDDGAQVEWGPTEDISYSGG	0.9974
217	GAFAIGNSNNVTIKNVTFKD	0.9995
293	NDDGKKSENVTIQNSYFGKS	1.2743
364	GNRFDKKVKGESVHYRESGA	1.3528
439	VTKNVINNNSKETEQPNIEL	1.0637
511	NEKEPVIRDSDGNFNIVTEN	0.6281
534	KIVTNNLSDKNEKEKNKEEK	1.6349
555	SNSNNVIDSNQKNGEFNSSK	0.7679
585	NKQDNKTEEVNYKIVGDGRE	1.5427

1: The amino acid residue number at the epitope beginning is indicated.

CBTope analysis revealed the presence of 10 conformational B-cell epitope regions in the protein (Table 11). The epitope region beginning at the amino acid 615 was exclusively conformational. However, the rest of the conformational B-cell epitope regions were overlapped with the linear B-cell epitopes.

Table 11: Conformational B-cell epitope regions of truncated PSP

Position[1]	No. of epitope residues	Epitope region sequence[2]
51	11	GVSINQNKPDIT
89	4	LNVKD
126	5	FPEGTYLVK
187	5	ISYSGGTIDMNGA
224	13	SNNVTIKNVTFKDSYQGHAIQIAGSKNVLVDNSR
272	45	IISKESIQIEPLTRKGFPYALNDDGKKSENVTIQNSYFGKSD KSGELVTAIGTHYQTLSTQNPSNIKIL NNHF
375	8	SVHYRESGAALVN
457	27	ELLRVSDNLVVSENSIFGGKEGIVIEDSKGKITVLNNQF
520	10	SDGNFNIVTENGLYKIV
587	8	QDNKTEEVNYKIVGDGRE
615	7	IVDVKQK

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: The amino acid residues of conformational B-cell epitopes in each region are shown underlined.

The linear B-cell epitopes identified in the truncated PSP by Bepipred and BCPreds were overlapped at 8 regions, which included amino acids common with the identified conformational B-cell epitopes (Table 12). Considering VaxiJen scores, the epitope region beginning at the amino acid 534 was the most probable protective linear B-cell epitope region of the protein. Moreover, the VaxiJen score of this overlapped B-cell epitope region was higher than those identified in autolysin and the mature ZBL. The conservancy of this region in PSPs of pneumococcal serotype19F strains was analyzed using the COBALT multiple protein sequences alignment tool. The region was fully conserved in the pneumococcal strains. This result indicated that the immunological response elicited against this region is protective against serotype 19F pneumococcal strains. The secondary structures predicted for the truncated PSP using the PHD program were α helix, extended strand and random coil with the predominance of random coil. The overlapped B-cell epitope region beginning at the amino acid residue 534 was located in random coil and extended strand. SCRATCH predicted 80% solubility for the truncated PSP upon overexpression in *E. coli.* Therefore, the protein can be recovered from the cellular soluble fraction.

Table 12: Overlapped linear B-cell epitope regions of truncated PSP

Position[1]	Epitope sequence[2,3]	VaxiJen score
41	NSSLSTDNNG*GVSINQNKPDIT*NDTMMGNISGME	0.9445
112	IDAAAQGLGGGNVY*FPEGTYL*	0.5299
173	TDDGAQVEWGPTED*ISYSGGTID*	0.9767
290	YAL*NDDGKKSENVTIQNSYFGKSDKSGE*	1.4416
364	GNRFDKKVKGE*SVHYRESGA*	1.3528
439	VTKNVINNNSKETEQPNI*EL*	1.0637
509	NANEKEPVIR*DSDGNFNIVTEN*	0.6279
534	***KIV*TNNLSDKNEKEKNKEEKSDKNEKEKNKEEKQSNSNNV IDSNQKNGEFNSSKDNRQMNDKIDNK*QDNKTEEVNYKIVGDGRE*TENHIN**	1.6046

1: Amino acid residue numbers at the beginning of epitope regions are indicated.

2: Amino acids common with the conformational B-cell epitope regions are in italics.

3: The most probable protective B-cell epitope region is in bold.

4. Conclusions and Recommendations

Among three proteins including autolysin, the mature ZBL, and the truncated PSP of S. pneumoniae serotype 19F, the truncated PSP possesses the greatest number of protective overlapped linear B-cell epitope regions,

which were predicted by both Bepipred and BCPreds programs and have amino acids common with the protein conformational B-cell epitopes. In addition, the overlapped B--cell epitope region with the most probability of eliciting protection resides in this protein and this region is fully conserved in PSPs of serotype 19F pneumococcal strains. Therefore, the truncated PSP is considered as an appropriate candidate for development of protein based vaccines against S. pneumoniae serotype 19F. In future, the protein ability in eliciting protection against pneumococci will be assessed experimentally. The limit of this study was the availabilities of the PSP amino acid sequences from serotype 19F S. pneumoniae for the analysis of the epitope conservation. Therefore, the experimental assessing of the PSP encoding gene presence in serotype 19F S. pneumonia strains and obtaining the amino acid sequence are recommended.

References

[1] B. B. Mook-Kanamori, M. Geldhoff, T. van der Poll and D. van de Beek. "Pathogenesis and pathophysiology of pneumococcal meningitis." Clin Microbiol Rev. vol. 24. pp. 557-591. Jul. 2011.

[2] WHO. "Pneumococcal vaccines, WHO position paper." Wkly Epidemiol Rec. vol. 87. pp. 129-144. Apr. 2012.

[3] H. L. Johnson, M. Deloria-Knoll, O. S. Levine, S. K. Stoszek, L. F. Hance, R. Reithinger, et al. "Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project." PLoS Med. vol. 7. pp. 1-13. Oct. 2010.

[4] T. J. Foster, J. A. Geoghegan, V. K. Ganesh, and M. Hook. "Adhesion, invasion and evasion: the many functions of surface proteins of Staphylococcus aureus." Nature Rev Microbiol. vol. 12. pp. 49-62. Dec. 2014.

[5] S. Bergmann and S. Hammerschmidt. "Versatility of pneumococcal surface proteins." Microbiol. vol. 152. pp. 295-303. Feb. 2006.

[6] S. Tarahomjoo. "Bioinformatic analysis of surface proteins of Streptococcus pneumoniae serotype 19F for identification of vaccine candidates." Am J Microbiol Research. vol. 2. pp. 174-177. Oct. 2014.

[7] S. Tarahomjoo. "Recent approaches in vaccine development against Streptococcus pneumonia." J Mol Microbiol Biotechnol. vol. 24. pp. 215-227. August 2014.

[8] D. F. Kelly and R. Rappuoli. Hot topics in infection and immunity in children II. New York: Springer. 2005. pp. 217-223.

[9] I. A. Doytchinova and D. R. Flower. "Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines." Open Vaccine J. vol. 1. pp. 22-26. Sep. 2008.

 [10] J. Novotny,. M. Handschumacher, E. Haber, R. E. Bruccoleri, W. B. Carlson, D. W. Fanning, et al. "Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody

domains).” Proc Natl Acad Sci USA vol. 83. pp. 226-230. Jan. 1986.

[11] L. Quijada, M. J. Requema. M. Soto, L. C. Gomez, F. Guzman, M. E. Patarroyo, et al. “Mapping of the linear antigenic determinants of the Leishmania infantum Hsp 70 recognized by leishmaniasis sera.” Immunol Lett. vol. 52. pp. 73-79. Sep. 1996.

[12] A. R. Faria, M. M. Costa, M. S. Giusta, G. G. Jr, M. L. O. Penido, R. T. Gazzinelli, et al. “High throughput analysis of synthetic peptides for the immunodiagnosis of canine visceral leishmaniasis.” PLos Negl Trop dis. vol. 5. pp. 1-9. Sep. 2011.

[13] H. L. Atilano, P. M. Pereira, F. Vaz et al. “Bacterial autolysins trim cell surface peptidoglycan to prevent detection by the Drosophila innate immune system.” eLife, vol. 3, pp. 1-23, Apr. 2014.

[14] R. Gabbianelli, R. Scotti, S. Ammendola, P. Petrarca, L. Nicolini and A. Battistoni. “Role of ZnuABC and ZinT in Escherichia coli O157: H7 zinc acquisition and interaction with epithelial cells.” BMC Microbiol. vol. 11. pp. 1-11. Feb. 2011.