Bioinformation Volume 5

open access

www.bioinformation.net Issue 4

Hypothesis

Prediction and characterization of T-cell epitopes for epitope vaccine design from outer membrane protein of *Neisseria meningitidis* serogroup B

Sharat Chandra[1], Digvijay Singh[1], Tiratha Raj Singh[2,3*]

[1]School of Biotechnology, Devi Ahilya University, Indore, India ; [2]Bioinformatics Centre, School of Biotechnology, Devi Ahilya University, Indore, India; [3]Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghata, Solan, India; Tiratha Raj Singh - E-mail : tirathraj@gmail.com; *corresponding author

Received July 25, 2010; Accepted August 25, 2010; Published September 20, 2010

Abstract:
Neisseria meningitidis serogroup B (MC58) is a leading cause of meningitis and septicaemia, principally infects the infants and adolescents. No vaccine is available for the prevention of these infections because the serogroup B capsular polysaccharide is unable to stimulate an immune response, due to its similarity with polysialic acid. To overcome these obstacles, we proposed to develop a peptide based epitope vaccine from outer membrane protein contained in outer membrane vesicles (OMV) based on our computational analysis. In OMV a total of 236 proteins were identified, only 15 (6.4%) of which were predicted to be located in outer membrane. The major requirement is the identification and selection of T-cell epitopes that act as a vaccine target. We have selected 13 out of 15 outer membrane proteins from OMV proteins. Due to similarity of the *fkp*A and *omp*85 with the human FKBP2 and SAMM50 protein, we removed these two sequences from the analysis as their presence in the vaccine is likely to elicit an autoimmune response. ProPred and ProPred1 were used to predict promiscuous helper T Lymphocytes (HTL) and cytotoxic T Lymphocytes (CTL) epitopes and MHCPred for their binding affinity in *N. meningitidis* serogroup B (MC58), respectively. Binding peptides (epitopes) are distinguished from nonbinding peptides in properties such as amino acid preference on the basis of amino acid composition. By using this dataset, we compared physico-chemical and structural properties at amino acid level through amino acid composition, computed from ProtParam server. Results indicate that porA, porB, opc, rmpM, mtrE and nspA are more suitable vaccine candidates. The predicted peptides are expected to be useful in the design of multi-epitope vaccines without compromising the human population coverage.

Keywords: Outer membrane vesicles, epitope vaccine, epitopes, *Neisseria meningitides* serogroup B.

Background:
Meningococcal disease is a major health problem worldwide. *Neisseria meningitidis* is a major cause of meningitis and septicaemia globally, predominantly affecting children and adolescents. Disease develops rapidly and often difficult to distinguish from other febrile illnesses. Meningococcal meningitis and sepsis are devastating diseases that kills children and young adults within hours despite the availability of effective antibiotics. Mortality and permanent disability rates are high, even under optimal health care conditions **[1].** Therefore, any further reduction in morbidity and mortality is most likely to be achieved through prophylactic vaccination. The sole ecological niche of *Neisseria meningitidis* is mucosa of the oropharynx of humans. Meningococcal colonization of the respiratory tract, a phenomenon commonly referred to as carriage, represents a successful commensal relationship between the host and the bacterium, with the host experiencing no detectable pathology. On the other hand disease represents a failed or dysfunctional relationship with the host. Acquistion of *Neisseria meningitidis* demands person to person transmission via direct contact or through dispersion of respiratory droplets from an infected to a susceptible individual. Although, often protected by a polysaccharides capsule, meningococci are particularly sensitive to desiccation; thus spread from one individual to another requires close contact **[2].**

The causative agent, *Neisseria meningitidis*, is a Gram- negative encapsulated bacterium classified into different groups according to the chemical composition and immunogenic properties of the capsular polysaccharide. Serogroups A, B, C, W135, and Y account for >95% of infections. Effective vaccines consisting of capsular polysaccharide or capsular polysaccheride-cojugates are available for the prevention of infections caused by serogroup A, C, Y, and W135 strains **[3, 4]**. However, no capsule-based vaccine is available for the prevention of infections caused by *N. meningitidis* serogroup B, which is highly prevalent in industrialized countries. This problem is caused by the poor immunogenicity of the serogroup B capsular polysaccharide. It is likely that the immune system tolerates the serogroup B capsular polysaccharide because of its similarity or mimicry to the widely distributed human carbohydrate α(2→8)N-acetyl neuraminic acid or polysialic acid, both consisting of repeating units of two to eight linked sialic acid **[5].** Therefore, for the development of an effective vaccine against serogroup B meningococci, researchers have focused on the proteins of the outer membrane (OM).

Alternative vaccine candidates have been sought, which are based on protein components being at the most advanced stage of development. These are commonly presented as outer membrane vesicle (OMV) formulations prepared by detergent extraction of meningococcal broth cultures **[6].** OMV vaccines have been used with some success in Norway and Cuba **[7, 8]** and have recently been granted a provisional license in New Zealand **[9].** Outer membrane vesicles (OMVs) are released from the outer membrane of *N. meningitidis,* which is a characteristic of many

Bioinformation Volume 5
www.bioinformation.net Issue 4
open access
Hypothesis

Gram-negative bacteria **[10].** It has been demonstrated that they contain outer membrane and periplasmic proteins and in some cases DNA or cell-cell signaling molecules **[11, 12]**. This makes OMVs an ideal structure to transport hydrophobic compounds like membrane proteins into the host. This clearly indicates the potential of OMVs to deliver membrane active virulence factors into the host. OMV vaccine is not only composed simply of OM proteins but contain a number of periplasmic, membrane-associated and cytoplasmic proteins. OMV vaccines present the immune system with a complex mixture of antigens, many of which are highly variable. In OMV 236 non-redundant proteins have been identified out of which 15 proteins were predicted as to be located in the outer membrane and represents only 6.4% of the total number of proteins **[13]**. Present study against serogroup B *Neisseria meningitidis* MC58 are based on meningococcal outer membrane proteins present in OMV. We have selected 13 out of 15 outer membrane proteins for the vaccine design, by identifying its HTL and CTL epitopes from the protein sequence of these proteins. This is because, two proteins *fkp*A and *omp*85 shows similarity with the human protein, and their presence in the vaccine is likely to elicit an autoimmune response.

Major obstacles to the development of an effective vaccine against these serogroup B are the complex mixture of OMV antigens and many of which are highly variable as well as the HLA allelic diversity in human. Activation of both Helper T-Lymphocytes (HTL) and Cytotoxic T-Lymphocytes (CTL) requires recognition of specific peptides bound to Major Histocompatibility Complex (MHC) molecules on the antigen presenting cells (APCs) / target cell. Protein antigens inside the APCs / target cell are degraded into small peptide fragments by the intracellular proteases. After partial proteolysis some of the peptides bind to MHC molecules and are transported to cell surface for recognition by the antigen specific T-cell receptors called T-cell Epitopes. Thus, MHC binding is a pre-requisite for a peptide to be a T-cell epitope. Bioinformatics approaches allow for the design of peptide vaccines starting from the prediction of all antigens *in silico* from the genome sequence of pathogens, independent of their abundance and without the need to grow the microorganism *in vitro* **[14-16]**.

Bioinformatics tools can be used to screen MHC-peptide complexes that become potential T-cell epitopes used for the development of peptide-based epitope vaccines **[17]**. A peptide which has a proteosomal recognition site may not be useful as epitope vaccine candidate because it is degraded during antigen processing **[18]**. Therefore methods for predicting the proteasomal cleavage sites were incorporated in algorithms used for predicting T-cell epitopes **[19].** Four fundamental properties of proteasomes are: (i) the broad cleavage-site specificity; (ii) the enhanced preference for hydrophobic residues in a hydrophobic sequence context; (iii) the excision of large numbers of peptide $\geq$ 8 amino acids; and (iv) the versatility of proteasomal processing resulting in overlapping peptide patterns. The aim of this study is to predict epitope peptides that bind with class I and class II MHC molecules, and will be useful for the prediction of new peptide molecules for vaccine design.

Methodology:

Target protein sequence retrieval:

A set of the 15 OM protein complements have been selected from the OMV 236 non-redundant proteins and represents only 6.4% the total number of proteins detected **[13].** The complete genome and protein sequences of *N. meningitidis* serogroup B (MC58) were taken from Genbank (NCBI) and UniProtKB. The selected protein sequences were retrieved in FASTA format and used for further analyses.

Vaccine candidate characterization:

Theoretical isoelectric point (pI), molecular weight and amino acid compositions were computed by using the Expasy's ProtParam server **[20].** The antigenicity of the proteins is identified by the VaxiJen server **[21].** The prediction of the non-allergenicity is done by Algpred sever **[22].** The similarity of human protein with OM proteins of OMV are searched through BlastP.

Epitope prediction:

The web servers ProPred and ProPred1 were used to predict HTL and CTL epitopes respectively **[17, 19].** Here, ProPred allow to predict MHC class II binding peptides (HTL epitopes) for 51 alleles and ProPred1 to predict MHC class I binding peptides (CTL epitopes) for 47 alleles. The ProPred and ProPred1 implements matrix based prediction algorithm. The obtained matrices are multiplication matrices, where the scores are calculated by multiplying and summing the score of each position. For example, the score for the peptide 'PACDPGRAA" can be calculated using the following equations:Score = P(1) x A(2) x C(3) x D(4) x P(5) x G(6) x R(7) x A(8) x A(9);Score P(1)= P(1)+A(2)+C(3)+D(4)+P(5)+G(6)+R(7)+A(8)+A(9)
Where 'P (1)' is score of amino acid 'P' at position 1.

For the prediction of the MHC binders the score of these 9mer peptides were calculated using quantitative matrix of selected MHC alleles. The all peptides having scored greater than selected threshold score (at 4%) were assigned as predicted binders for selected MHC alleles. For prediction of proteasome cleavage site overlapping 12mer peptides were calculated using weight matrix of proteasome. All peptides having score greater than selected threshold score (at 5%) are considered as peptides having proteasome cleavage site. Similar approach is used for prediction of peptides having immunoproteasome cleavage site **[17, 19].** In order to calculate threshold score for each allele/matrix, the steps used were:

1. All proteins were obtained from SWISSPROT database for creating the overlapping peptides of length nine. For example, a protein of length n will have (n+1 – 9) overlapping peptides.

2. The score of all natural 9-mer peptides was calculated using the weighed matrix of that allele. These peptides have been sorted on the basis of score in descending order and top 1% natural peptides have been extracted. The minimum score that we called threshold score was determined from these selected peptides. Similarly, threshold scores at 2%, 3% … 10% were calculated. The following threshold dependent parameters were used for the evaluation of tools ProPred and ProPred1: Sensitivity= TP/(TP+FN); Specificity= TN/(TN+FP); Accuracy= (TP+TN)/(TP+FN+TN+FP). Where, TP and TN are correctly predicted binders and non-binders respectively, whereas FP and FN are wrongly predicted binders and non binders respectively from the database MHCBN **[23].**

The IC_{50} value of corresponding peptides is deduced by MHCPred **[24-25].** IC_{50} value of allele HLA_A*0101 for MHC class I peptides and IC_{50} value of HLA_DRB1*0101 for MHC class II peptides. The predicted output is given in units of IC_{50} nM.

Discussion:

The physiochemical properties of proteins are given in **(Table 1.see supplementary material.)** The pI values of all these OM proteins indicated the stability of protein at that particular pI. The theoretical pI of the proteins identified ranged from 4.79 to 9.71 and theoretical MW ranged from 18.4 to 159.9. This information is useful for the proteomics study especially for the 2D PAGE and is also useful for cataloguing OMV proteins with emphasis on 15 OM proteins used in this study. Outer membrane proteins are notoriously difficult to resolve by 2D PAGE. By their nature, hydrophobic proteins are often difficult to solubilise and are least soluble when focused at their pI. The antigenicity of all the OM proteins predicted using VaxiJen server is given in **(Figure 1** and **Table 2 see supplementary material).** The allergenicity of all OM proteins are predicted by AlgPred server and all 15 OM proteins found non-allergens. The antigenicity and non-allerginicity of OM proteins shows that they are protective antigens. The BlastP result shows the similarity of two proteins with human proteins, when OM proteins searched against translated human genome through BlastP. The fkpA (NMB1567) has limited similarity to the

Bioinformation Volume 5

www.bioinformation.net Issue 4

open access

Hypothesis

human FK506 binding proteins 2 (FKBP2) and AIP aryl hydrocarbon receptor interacting proteins, and omp85 (NMB0182) has limited similarity with human SAMM50 sorting and assembly machinery component 50 homolog, this give rises the possibility of immunological cross reactivity between vaccine and host cell proteins, which could lead to autoimmunity. Therefore 13 out of 15 OM proteins are being selected for the epitope prediction.

The summary of the results of the prediction tools ProPred and ProPred1 used for the HTL and CTL epitopes identification respectively in the present study are being made available at (http://www.bioinfoindia.org/epitopes/HTL_CTL_Table.pdf). According to the prediction results of MHCPred, peptides with the best predicted binding affinities of HLA_A*0101 for MHC class I and HLA_DRB*0101 for MHC class II molecules are also presented therein. The predicted output is given in units of IC_{50} nM. A lower value of peptide IC_{50} indicates higher affinity towards MHC molecules. Peptides with IC_{50} values < 50nM, < 500nM, and <5000nM are considered having high affinity, intermediate affinity and low affinity towards MHC molecules respectively (HTL_CTL_Table.pdf). In OMV 236 nonredundent proteins have been identified out of which 15 (6.4%) proteins were predicted to be located in the outer membrane. From BlastP results it is established that 13 out of 15 OM proteins are the most suitable for the vaccines and they contain maximum promiscuous epitopes at many HLA alleles as these 13 OM proteins didn't show any significant similarity with human proteins and neglect the chances of autoimmunity response. A multiple promiscuous target peptide epitopes would overcome the genetic restriction, low immunological responsiveness and parasite evasion of the immune response and provide support for developing a multistage, multivalent, universal, safe, stable and effective meningococcal vaccine.

The OM proteins contain a porins, porA and porB, which are used for serotyping and serosubtyping of meningococci. The most abundant of these included proteins with important biological function, including the porA and porB porins, opc is involved in the adhesion of the meningococci to host cell-surface proteoglycan and therefore plays a part in colonisation and invasion by bacterium. This protein associated with invasion of epithelial and endothelial cells **[26].** The rmpM protein, which may protect against complement-mediated bactericidal attack **[27],** mtrE is one of four proteins encoded by the multiple transferable resistance mtrE locus and forms outer membrane component of multidrug efflux pump [28], outer membrane protein nspA and pilQ which is essential for pilus assembly **[29]**. Two hypothetical proteins NMB0345 and NMB0088 are identified in the outer membrane protein constituent of OMV. Some other major OM proteins are also present like hap for adhesion and penetration, lptD involved in the assembly of LPS in the outer leaflet of the outer membrane, determines N-hexane tolerance and is involved in outer membrane permeability, essential for envelope biogenesis, thpB acts as a transferrin receptor and is required for transferrin utilization.

The peptides of 13 antigenic proteins are classified into binding and non-binding groups. The binding peptides signify the predicted epitopes and non-binding peptide signifies the protein sequence of OM protein minus binding sequence. Both dataset constructed for the MHC class I and class II individually. Then amino acid composition is computed and results are evaluated in three category of amino acid: hydrophobic, charged and polar and other small amino acid. Our data illustrate **(Figure 2, 3)** that hydrophobic amino acids are found more frequently in binding group (epitope) than in the non-binding group. The non binding group have more polar amino acid compared to other amino acids. This condition is found in both CTL as well as in the HTL binding and non binding dataset **(Table 3 see supplementary material).** Based on our analysis a computational protocol **(Figure. 4)** has been proposed which will work as a model for such further studies. The predicted epitopes in OMV proteins would be useful for the earlier identification of meningitis and septicaemia, and will be helpful for secure vaccine development against meningococcal diseases.

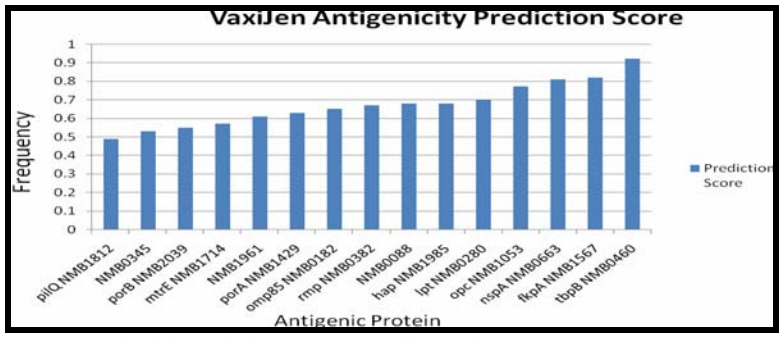

Figure 1: Antigenicity prediction score computed by VaxiJen server.

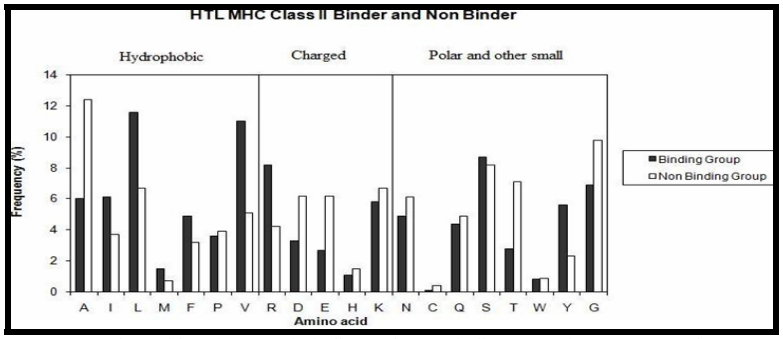

Figure 2: Amino acid preference of Binding and Non Binding group for HTL MHC class II.

Bioinformation Volume 5

www.bioinformation.net Issue 4

open access

Hypothesis

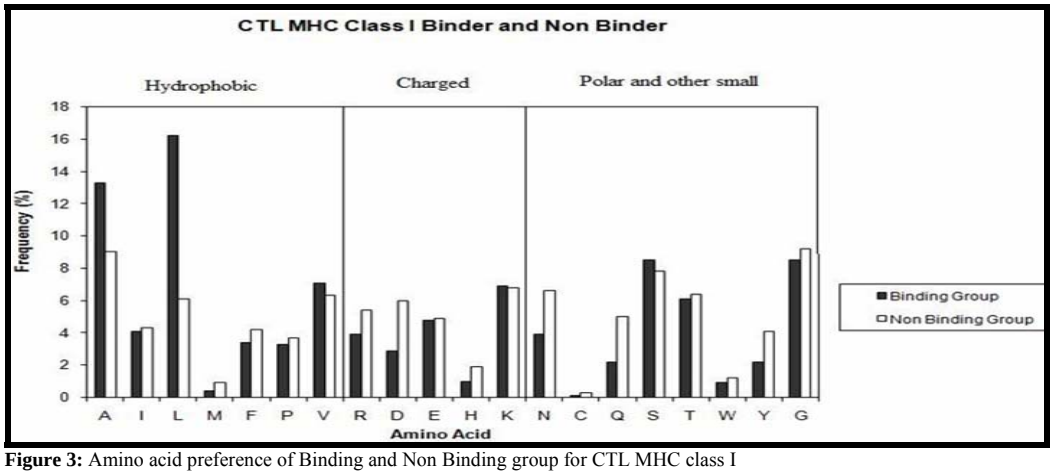

Figure 3: Amino acid preference of Binding and Non Binding group for CTL MHC class I

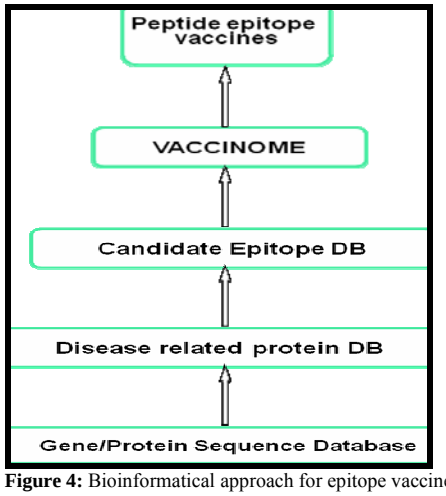

Figure 4: Bioinformatical approach for epitope vaccine developmemt.

Conclusion:

Predicted promiscuous HTL and CTL epitopes from the genome/proteome sequences of the pathogens would greatly reduced the time as well as cost and is useful for experiment planning in development of peptide-based epitope vaccines. We have predicted numerous epitopes in OMV proteins which would be useful for the earlier identification of meningitis and septicaemia, the predicted epitopes may be used for safe vaccine development against meningococcal diseases. Results indicate that porA, porB, opc, rmpM, mtrE and nspA are more suitable vaccine candidates. These six proteins are selected due to their smaller size and most important thing is that they are present most abundantly in the OMV proteins as an outer membrane proteins. We frequently found large number of hydrophobic amino acid, certain polar and charged amino acids in binding group. Hydrophobic protein regions may tend to be highly conserved because of their importance for correct folding they may show epitopes particularly effective for protective host immunity. Also they have variable pI which is important for 2D study for separation of proteins. These results have important propositions for the development and use of vaccines based on epitopes of OM proteins after *in vitro* study and experimental validation.

Acknowledgements:

Authors acknowledges computational facility provided by Bioinformatics Sub-DIC (funded by DBT, India), School of Biotechnology, DAVV, Indore, India and bioinfoindia.org (http://www.bioinfoindia.org) for providing web space for storing epitope data.

References:

[1] M van Deuren *et al. Clin Microbiol Rev* **13**:144 (2000) [PMID:10627495]
[2] IW DeVoe *et al. Microbiol Rev* **46**:162 (1982) [PMID:6126800]
[3] L Jódar *et al. Lancet* **359** :1499 (2002) [PMID;11988262]
[4] Morley SL *et al. Vaccine* **20**:666 (2001) [PMID:11738731]
[5] J Finne *et al. Lancet* **2** :355 (1983) [PMID:6135869]
[6] JH Fredriksen *et al. NIPH Ann* **67** (1991) [PMID:1812438]
[7] GV Sierra, ***et*** *al. NIPH Ann.* **14**:195 (1991) [PMID;1812432]
[8] G Bjune *et al.* **338**:1093 (1991) [PMID:1682541]
[9] M Thomas *et al. N Z Med J* **117**(1200):U1016. (2004) [PMID:15475986]
[10] Beveridge TJ *et al. J Bacteriol* **181**:4725 (1999) [PMID:10438737]
[11] LM Mashburn M *et al. Nature*. **437**:422 (2005) [PMID:16163359]
[12] DW Dorward Garon *et al. J Bacteriol.***171**(5):2499 (1989) [PMID:2496108]
[13] JN Williams *et al. Infect Immun* **75** :1364 (2007) [PMID:17158897]
[14] R Rappuoli *et al. Vaccine* **19**:2688 (2001) [PMID:11257410]
[15] DL Doolan *et al. Parasitol Today* **13**:171 (1997) [PMID:15275087]
[16] AS De Groot *et al. Expert Rev Vaccines* **3**:59 (2004) [PMID:14761244]
[17] H Singh *et al. Bioinformatics* **19**:1009 (2003) [PMID:12761064]
[18] RE Toes *et al. J Exp Med* **194**:1 (2001) [PMID:11435468]
[19] H Singh *et al. Bioinformatics* **17**:1236 (2001) [PMID:11751237]
[20] www.expasy.ch/tools/protparam.html
[21] IA Doytchinova *et al. BMC Bioinformatics* **8**:4 (2007) [PMID:17207271]

[22] S Saha *et al. Nucleic Acids Res* **34** (2006) [PMID:16844994]
[23] M Bhasin Singh *et al. Bioinformatics* **19**:665 (2003) [PMID:12651731]
[24] P Guan *et al. Nucleic Acids Res* **31**:3621 (2003) PMID:12824380]
[25] Guan P ***et al.*** *Appl Bioinformatics.* **5**:55 (2006) [PMID:16539539]
[26] M Virji Makepeace *et al. Mol Microbiol* **6**:2785 (1992) [PMID:1435257]
[27] A Munkley *et al. Microb Pathog* **11**(6):447 (1991) [PMID:1795633]
[28] RM Delahay *et al. Microbiology* **143**:2127 (1997) [PMID:9245802]
[29] SL Drake *et al. Mol Microbiol* **18**:975 (1995) [PMID:8825101]

Edited by P. Kangueane
Citation: Chandra *et al.* Bioinformation 5(4): 155-161 (2010)

Supplementary materials:

Table 1: Physicochemical properties of proteins

Protein designation	Gene name	Accession No.	Size aa	Expected MW kDA	pI
Major outer membrane protein P.IA (Protein IA) (PIA) (Class 1 protein)	porA NMB1429	Q52140	392	42.1	9.13
Major outer membrane protein P.IB (Protein IB) (PIB) (Porin) (Class 3 protein)	porB NMB2039	P30690	331	35.7	7.14
Class 5 outer membrane protein	Opc NMB1053	Q7DDI3	272	29.9	9.71
Type IV pilus biogenesis and competence protein pilQ	pilQ NMB1812	Q70M91	769	82.5	9.43
Probable FKBP-type peptidyl-prolyl cis-trans isomerase fkpA (PPIase) (Rotamase)	fkpA NMB1567	Q9JYI8	272	28.9	5.72
Outer membrane protein class 4	rmpM NMB0382	P070A3	242	26.2	6.97
Putative cell-binding factor	NMB0345	Q7DDR0	288	31.5	9.23
Outer membrane protein omp85	omp85 NMB0182	QPK10	797	88.4	8.65
Multidrug efflux pump channel protein	mtrE NMB1714	Q9JY68	467	50.4	8.34
Putative outer membrane protein NMB0088	NMB0088	Q9K1M2	466	50.5	9.37
Adhesion and penetration protein	Hap NMB1985	Q9JXL6	1457	159.9	9.19
LPS-assembly protein lptD	lptD NMB0280	Q9K187	802	88.7	8.50
Outer membrane protein NspA	nspA NMB0663	Q7DDM2	174	18.4	9.54
VacJ-related protein	NMB1961	Q9JXN3	275	29.5	4.79
Transferrin-binding protein 2 (TBP-2)	tbpB NMB0460	Q9K0V0	712	77.4	5.79

Table 2: Computational prediction of Antigenicity (in ascending order) of the proteins by VaxiJen server at threshold 0.4

Proteins	Probable Antigenicity	Prediction score
pilQ NMB1812	Anitigen	0.49
NMB0345	Anitigen	0.53
porB NMB2039	Anitigen	0.55
mtrE NMB1714	Anitigen	0.57
NMB1961	Anitigen	0.61
porA NMB1429	Anitigen	0.63
omp85 NMB0182	Anitigen	0.65
rmpM NMB0382	Anitigen	0.67
NMB0088	Anitigen	0.68
Hap NMB1985	Anitigen	0.68
lptD NMB0280	Anitigen	0.70
Opc NMB1053	Anitigen	0.77
nspA NMB0663	Anitigen	0.81
fkpA NMB1567	Anitigen	0.82
tbpB NMB0460	Anitigen	0.90

Table 3: Amino acid preference of Binding and Non Binding group for CTL MHC class I and HTL MHC class II.

	Amino Acid	CTL MHC class I (%)		HTL MHC class II (%)	
		Binding	Non Binding	Binding	Non Binding
Hydrophobic	Ala, A	13.3	9	6	12.4
	Ile, I	4.1	4.3	6.1	3.7
	Leu, L	16.2	6.1	11.6	6.7
	Met, M	0.4	0.9	1.5	0.7
	Phe, F	3.4	4.2	4.9	3.2
	Pro, P	3.3	3.7	3.6	3.9
	Val, V	7.1	6.3	11	5.1
Charged	Arg, R	3.9	5.4	8.2	4.2
	Asp, D	2.9	6	3.3	6.2
	Glu, E	4.8	4.9	2.7	6.2
	His, H	1	1.9	1.1	1.5
	Lys, K	6.9	6.8	5.8	6.7
Polar & other small	Asn, N	3.9	6.6	4.9	6.1
	Cys, C	0.1	0.3	0.1	0.4
	Gln, Q	2.2	5	4.4	4.9
	Ser, S	8.5	7.8	8.7	8.2
	Thr, T	6.1	6.4	2.8	7.1
	Trp, W	0.9	1.2	0.8	0.9
	Tyr, Y	2.2	4.1	5.6	2.3
	Gly, G	8.5	9.2	6.9	9.8