Structural Prediction of Pathogenic Factors of *Neisseria Meningitidis* using Comparative Modeling

Rafia Safdar and Yasmeen Rashid*

Department of Biochemistry, University of Karachi, Karachi-75270, Pakistan

*Author for correspondence (email: yasmeen.rashid@uok.edu.pk)

Abstract:

Neisseria meningitidis, a gram negative diplococcus and the causative agent of meningitis and severe sepsis, is known to occur in five different pathogenic strains including A, B, C, W-135 and Y. Vaccines against all strains are available except strain B due to its poor immunogenicity. Functional genomics studies of *N. meningitidis* serogroup-B revealed some of its proteins to be crucial for its pathogenesis. Here, we report 3D homolog models of NMB1728, NMB1527 and NMB1729 proteins. The homology models were constructed by the MODELLER program. Model structures were shown to have overall good stereochemistry as well as lower internal energy. The lower rmsd values for the models of these proteins showed overall structural resemblance including active site residues with their selected templates.

INTRODUCTION:

Neisseria meningitidis, a Gram-negative diplococcus and a member of Neisseriaceae, has a genome size of 2,272,351 bp. Out of 12, only five serogroups A, B, C, W135 and Y are considered to be the most pathogenic in human [1-2]. In Europe and United States, the major cause of invasive disease is *N. meningitidis* Serogroup B [3]. *N. meningitidis* Serogroup B (MC58) genome sequence provided efficient data related to its detailed molecular characterization [4]. To date there is no effective vaccine available against serogroup B meningococcal disease which is because of homology between the capsule of *N. meningitidis* serogroup B and N-acetylneuraminic acid residue containing polysialosyl glycopeptides in human neural tissue [5, 6].

According to the genome sequence of *N. meningitidis* MC58 strain [7], the bacterium showed to have 2,158 open reading frames. Out of those, only 1158 gene products are known for specific biological function while 1000 gene products remained hypothetical i.e. unidentified [8].

Functional genomics of *N. meningitidis* pathogenic factors have been performed and 73 gene products were found to be involved in septicemic disease [9].

Our aim was to study the structural bioinformatics of known virulence determinants of *N. meningitidis.* A number of virulence factors facilitate *N. meningitidis* survival and proliferation within the host organism. This group contains eight genes [8]. We have performed structural analysis of three known virulence determinants i.e. NMB1728, NMB1527 and NMB1729 genes of this group. We performed the homology modeling of the above-mentioned proteins by searching their best templates through Psi-BLAST [10] against PDB [11]. The best template was chosen on the basis of sequence identity; template with >30% sequence identity was considered as the most acceptable.

MATERIAL AND METHODS:

In present study, homology modeling of a group of different pathogenic proteins of *N. meningitidis* has been performed. The group includes proteins of known virulence factors i.e. NMB1729 (exbB), NMB1728 (exbD) and NMB1527 (rfaF).

Secondary structure prediction and template searches for homology modeling:

Secondary structure prediction of the target sequences i.e. NMB1527, NMB1728 and NMB1729 were carried out using PSIPRED [12] method. For the purpose of homology modeling, protein structures in PDB with the best similarities were used as templates for each of the target proteins. PDB structures complexed with some substrate or inhibitor were preferably chosen.

Multiple sequence alignment:

Multiple sequence alignment was performed using CLUSTALX [13] in order to optimize the gaps present in the target protein alignments, as the presence of gaps in the alignment causes deterioration in regular secondary structure. The homologous sequences were searched against UniProt database using FASTA searching tool. Pair-wise alignment was manually edited so that insertions or deletions should not be there in the regular secondary structures of template and target proteins.

Three-dimensional model building and refinement:

MODELLER 9.10 program [14] was used to build homology models of target proteins using the selected PDB structures on the basis of pair-wise sequence alignments between targets and templates.

The evaluation of the homology models:

Evaluation of model is essential in order to check the possibility of errors in the model. PROCHECK [15] was used for assessing the overall stereochemistry of protein structures. PROSA program [16] was used to calculate energy profiles of the models to evaluate overall fold and side-chain packing of homology models. The variability between 3D model and the template structure was monitored by superposition of backbone atoms onto the template structures to calculate the root mean square value. The SUPERPOSE python file of the MODELLER [14] was used for this function.

Structural analysis:

After building a protein model, analysis of protein three -dimensional structure is one of the most important step. Many software are used for visualizing the structure of proteins and for computing their properties.

DS Visualizer [17] and VMD [18] programs were utilized for protein structure visualization and active site analysis of proteins structures.

RESULTS AND DISCUSSION:

We have carried out structural characterization of the *N. meningitidis* virulence factors. These proteins which have been categorized into a class of known virulence determinants. These proteins include Biopolymer transport protein ExbD (NMB1728), Biopolymer transport protein ExbB (NMB1729) and ADP-heptose-LPS heptosyltransferase II (NMB1527). Analysis of three-dimensional models with detailed active site analyses is given below.

Biopolymer transport protein ExbD (NMB1728):

The gene of NMB1728 encoded Biopolymer transport protein ExbD which consist of 144 amino acids. It is present in cell inner membrane and is a single-pass type II membrane protein. This group of proteins are membrane bound transport proteins essential for ferric ion uptake in bacteria.

Secondary structure prediction of the protein suggested that ExbD protein consisted of 3 helices and 6 beta strands (Fig 1-A)

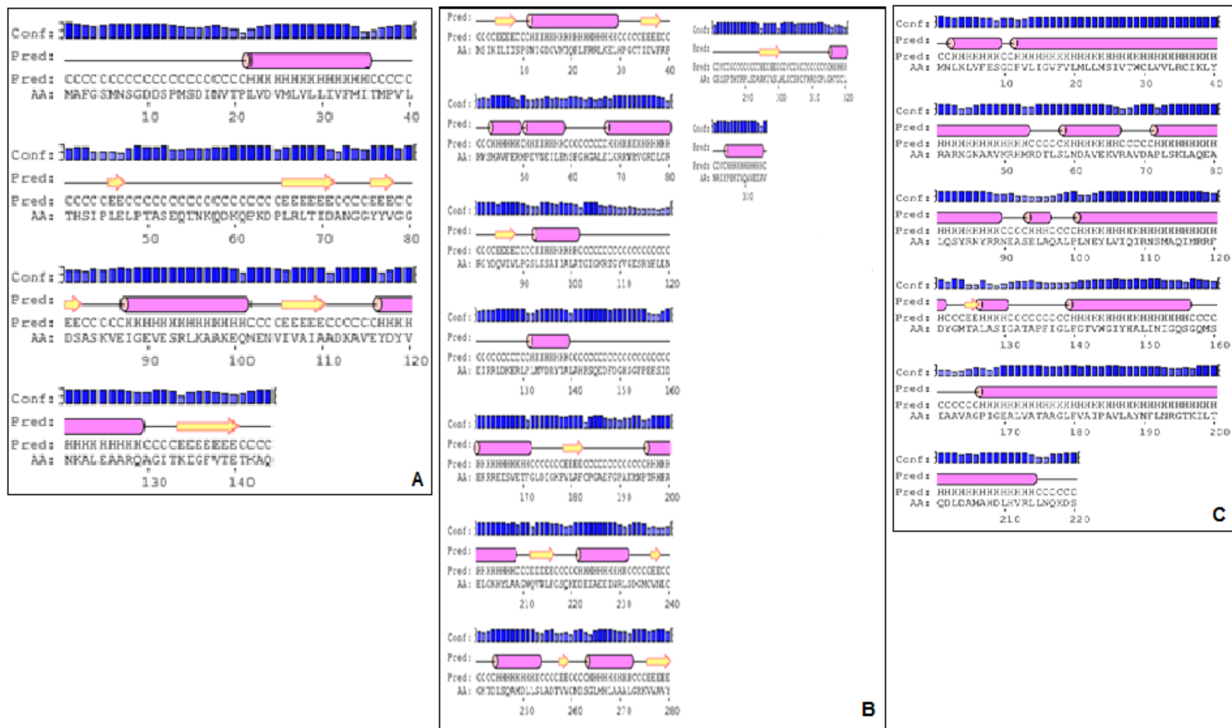

Figure 1: Secondary structure prediction of *N. meningitidis* biopolymer transport protein ExbD (A) and ADP-heptose LPS heptosyltransferaseII (B) and biopolymer transport protein ExbB using PsiPred where alpha helices are shown in cylinders and beta strands are presented by an arrow

---

For the purpose of homology modeling, sequence of NMB1728 (target protein) was searched against PDB (protein data bank) using PSI-BLAST (position specific iterated-BLAST) [10]. The result of BLAST showed a total of 4 hits for our query protein (Fig.2A). ExpD Protein structure with PDB ID 5BY4 having 47% sequence similarity with target protein. Was selected as templet. Template selection was based on sequence similarity and the similarity of protein families and minimum number of gaps in sequence alignment. Secondary structure of the template protein (PDB ID; 5BY4) was obtained from PDBsum to construct structure-based pairwise sequence alignment (Fig. 2B).

Multiple sequence alignment of EXbD protein was carried out in order to optimize the target and template sequences alignment. Homology model of *N. meningitidis* ExbD protein NMB1728 was

shown to have the overall fold similar to that of the *E. coli* ExbD protein structure rmsd = 1.0 A° (Fig.3A).

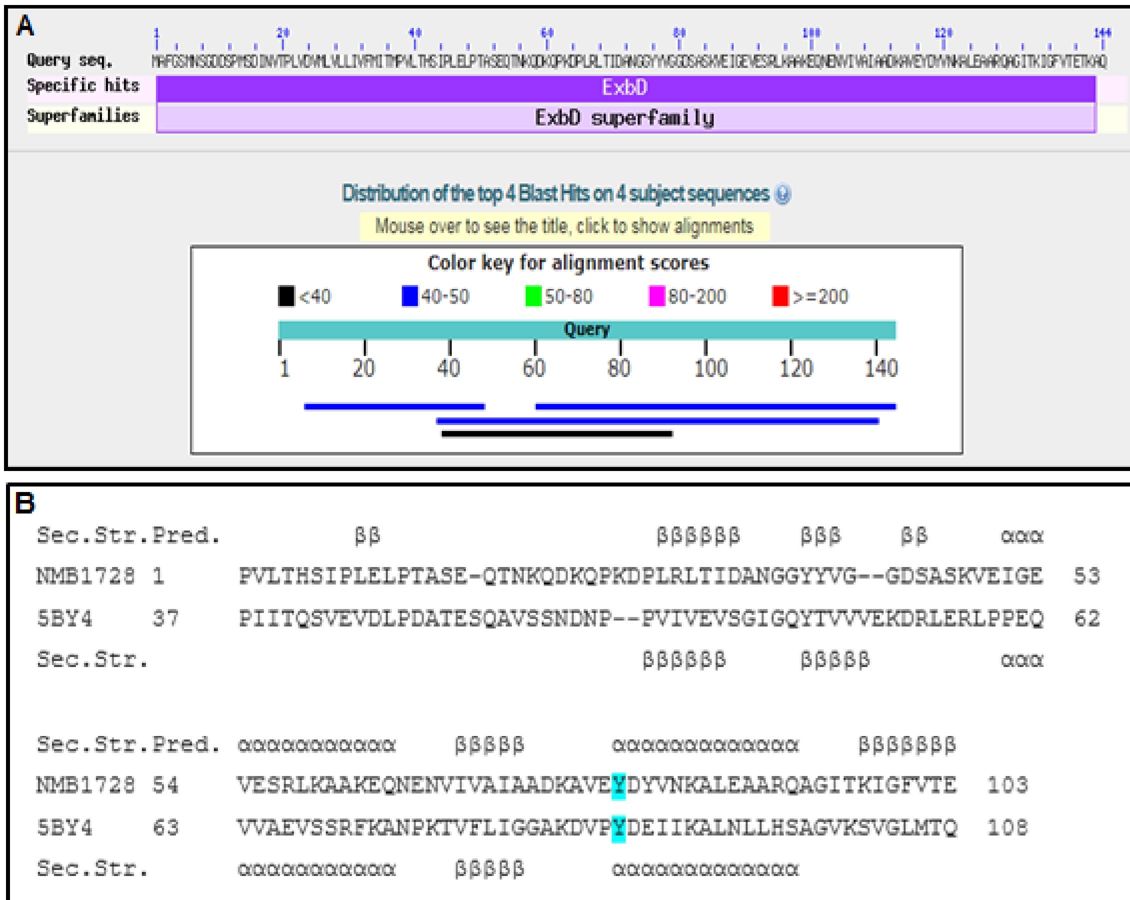

Figure 2: (A) Summary of Psi-BLAST result for NMB1728 protein. (B) Structure-based pairwise sequence alignment of ExbD from *Escherichia Coli* (5BY4) and *N. meningitidis* (NMB1728) used for homology model building.

In order to recognize errors in the model we used PROCHECK [15] and PROSA [16] programs. There were no amino acid residues in the disallowed region while 93.3% residues were present in most favored region. PROSA [16] plot also showed the model to have an overall lower internal energy while the z-score was -4.68. These findings indicated that the constructed homology model of *N. meningitidis* ExbD protein was good (Fig.3B & 3C).

E. coli TolR (ExbD) is a stable dimer which is present at the surface of the phospholipid bilayer because it contains the entire periplasmic domain [19]. We constructed *N. meningitidis* ExbD

monomeric model *using E. coli* TolR (ExbD) crystal structure (PDB ID; 5BY4). It is predicted that *N. meningitidis* ExbD protein would also occur in dimeric form performing the same function. Additionally, Tyr 117 is a crucial residue in *E. coli* TolR (ExbD) protein which is thought to be involved in dimer formation. This residue was found to be conserved in *N. meningitidis* ExbD model which further confirms its dimeric association (Fig. 3D).

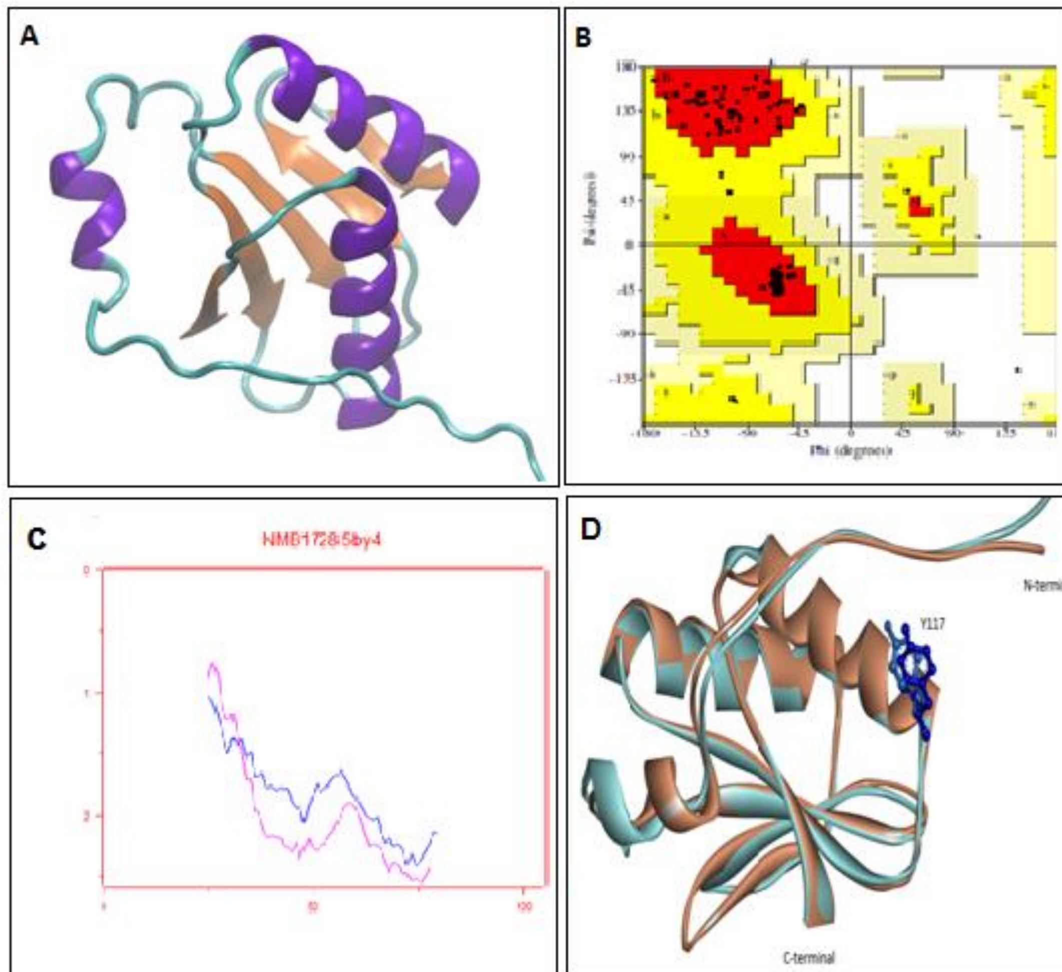

Figure 3: (A) Homology model of *N. meningitidis* biopolymer transport protein exbD along with evaluation of its (B) stereochemistry and (C) energy profile. (D) superposition of *N. meningitidis* (NMB1728) and *E. coli* (5BY4) ExbD protein showing N and C terminal of protein model, with the conserved TYROSINE residue

---

ADP- heptose LPS heptosyl transferase-II (NMB1527):

The complete name of NMB1527 gene product is ADP- heptose LPS heptosyl transferase II rfaF which contains 336 amino acids. It is involved in the transfer of heptose from ADP-heptose to

lipid-(KDO) during lipopolysaccharide biosynthetic process [20]. The predicted secondary structure of NMB1527 protein by Psi-Pred server showed to have 13 helices and 9 beta strands (Fig 1B).

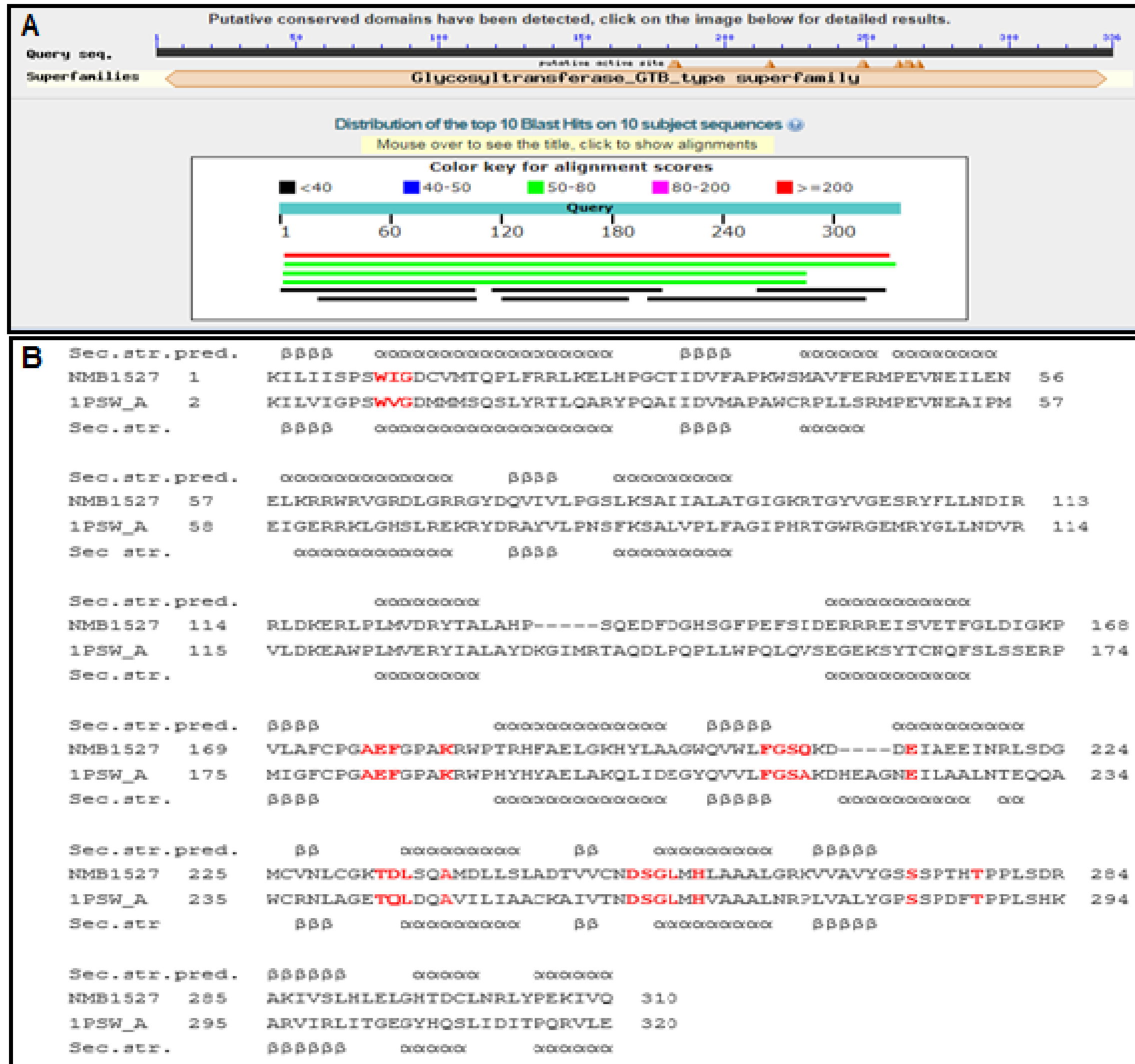

Figure 4: (A) Summary of BLAST result for *N. meningitidis* NMB1527 protein (B) Structure-based pairwise sequence alignment of Adp-Heptose Lps Heptosyltransferase II from *Escherichia Coli* (1PSW) and *N. meningitidis* (NMB1527) used for homology model building.

For homology modeling, template of NMB1527 protein (target protein) was searched in PDB using PSI-BLAST [10]. The BLAST showed 10 hits for our query protein (NMB1527) (Fig.4A). Among all the hits the best hit selected as a template has a PDB ID 1PSW. Secondary structure of the template (PDB ID; 1PSW) was obtained from PDBsum to construct structure-based pairwise

sequence alignment (Fig. 4B). The pair-wise sequence alignment showed conservation of some amino residues which are responsible for transfer of heptose during biosynthetic process of lipopolysaccharide [20].

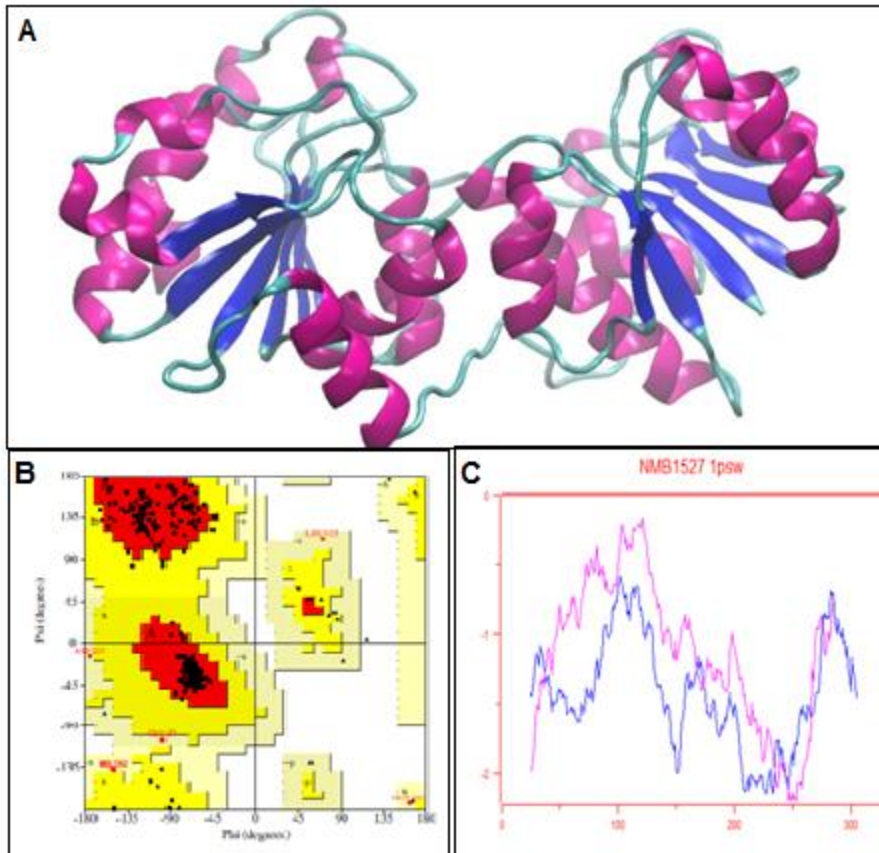

Figure 5: (A) Homology model of *N. meningitidis* Adp-Heptose Lps Heptosyltransferase II along with evaluation of its (B) stereochemistry and (C) energy profile.

Pairwise alignment of targeted template sequence contained 2% gaps in which may cause some deterioration in secondary structure; therefore multiple sequence alignment was made to optimize alignment using CLUSTALX [13]. NMB1527 protein homology model NMB1527 of *N. meningitidis* heptosyltransferase II was built on the basis of alignment between target and template sequences. All steps of homology modeling and refinement were carried out using the program MODELLER [14]. Homology model of *N. meningitidis* heptosyltransferase II comprised of two domains and the fold protein fold was similar to that of the template structure (1PSW) (Fig.5A).

In order to recognize errors in the model we used PROCHECK [15] and PROSA [16] programs were used. PROCHECK [16] tells about the stereochemistry of the model and PROSA [16] provides energy profile. Ramachandran plot showed no amino acid residue in the disallowed region whereas 91.4% residues were present in most favored region indicating that the overall stereochemistry of the model was good. PROSA [17] was used to calculate the overall energy profile of NMB1527 protein model which calculated the value of z-score = -3.98 which represented a good quality model (5B & 5C).

Lipopolysaccharide of gram-negative bacteria consists of three structural units including lipid A the core and an O antigen [21]. In *Escherichia coli* Lipid A is a beta-1,6-linked glucosamine disaccharide which is linked to the inner core (ketodeoxyoctonate [KDO] units) followed by two heptoses. The outer core region and the O antigen are linked to one of these heptoses. The synthesis of inner core region involved several genes including rfaC and rfaF, which have been recognized as heptosyl transferases for the transfer of heptose from ADP-heptose to the inner core. The rfaC and rfaF genes encode heptosyltransferase I and heptosyltransferase II enzymes, respectively. rfaF mutants lacking heptosyltransferase II activity make lipopolysacharides with only one heptose unit attached to KDO. Mutants lacking the O antigen and the outer core components are viable but nonvirulent [22]. However, they are hypersensitive to detergents and hydrophobic antibiotics [23]. Therefore it has been suggested that the NMB1527 protein is a drug target as inhibiting this protein might affect bacterial growth.

We constructed the three-dimensional model of *N. meningitidis* rfaF protein NMB1527 using *E. coli* rfaF crystal structure as a template (PDB ID; 1PSW). Ligand binding-site information from crystal structure of *E. coli* Heptosyltransferase-I was used to deduce the active site conformation. The binding pocket of heptosyltransferase-I is comprised of amino acid residues Ser10, Met11, Gly12, Ala186, Thr187, Thr188, Lys192, Pro216, Trp217, Gly218, Ala219, Glu222, Met242, Ser243, Leu244, Val247, Asp261, Thr262, Glu263, Leu264, His266, Thr282, Ile287. These residues are involved in binding of Adenosine-5'-diphosphate-2-deoxy-2-fluoro heptose. We found equivalent amino acid residues to be conserved in heptosyltransferase-II sequences (Q9JYL5, P37692, D8PHJ0, A0A1W1HZ44, A0A0S4LBX9, A0A1W1HZL4, A0A0S4KXP4, A0A0H7CYR6) from other bacteria as shown in multiple sequence alignment (Fig. 6). Hence *N. meningitidis* heptosyltransferase-II (NMB1527 protein) would also bind with heptose-containing

ligand in a similar mode and will also take part in the transfer of heptose during the biosynthetic process of lipopolysaccharide.

Figure 6: Multiple alignment sequences of ADP-heptose LPS heptosyltransferase from different species using CLUSTAL X. Conserved residues are highlighted in yellow color

---

Biopolymer transport protein ExbB (NMB1729):

The NMB1729 gene encodes is Biopolymer transport protein ExbB which consists of 220 amino acids. It is present in the inner membrane and is a multi-pass membrane protein. These proteins are membrane bound transport proteins essential for the uptake of metals, carbohydrates, cobalamin and many bacteriocins. ExbB uses the proton motive force at the inner membrane to transduce energy for the outer membrane via TonB. The protein family consists of ExbD and TolR which are involved in TonB-dependent transport of various receptor-bound substrates. It helps in protecting ExbD from proteolytic degradation and functionally stabilizes TonB [24]. PsiPred server was used for predicting secondary structure of NMB1729 protein. The protein was shown to have 9 helices and one beta strand (Fig 1C).

Template of NMB1729 (target protein) was searched in PDB using PSI-BLAST [10]. The BLAST result showed 4 hits for query protein sequence (Fig.7A). We have selected Protein structure (PDB ID; 5SV0) as template selection was based on sequence similarity and minimum number of gaps in template-target sequence alignment. The selected template 50% sequence identity with the target sequence. Secondary structure of the template (PDB ID; 5BY4) was obtained from PDBsum which was used to construct structure-based pair-wise alignment of protein sequences (Fig. 7B). The alignment of NMB1729 protein with the selected template contains 7% gaps which may cause some deterioration in regular secondary structure due to which we performed multiple sequence alignment in order to optimize the alignment.

A

Putative conserved domains have been detected, click on the image below for detailed results.

Query seq.
Specific hits: TolQ
Superfamilies: MotA_ExbB superfamily

Distribution of the top 4 Blast Hits on 4 subject sequences

Mouse over to see the title, click to show alignments

Color key for alignment scores: <40, 40-50, 50-80, 80-200, >=200

Query: 1, 40, 80, 120, 160, 200

B

```
NMB1729   1     VFESGDPVLIGVFVLMLLMSIVTWCLVVLRCIKLYRARKGNAAVKRHMRDTLSLNDAVE   60
5SV0      15    MYQHADIVVKCVMIGLILASVVTWAIFFSKSVEFFNQKRRLKREQQLLAEARSLNQAND   74

NMB1729   61    KVRAVDAPLSLAQEALQSYRNYRRNEASELAQALPLNEYLVIQIRNSMAQIMRRF      115
5SV0      75    IAADFGS---SLSLHLLNEAQNELELSEGSDDNEGIXERTSFRLERRVAAVGRQM      116

NMB1729   116   DGMTALASIGATAPFIGLFGTVWGIYHALINIGQSGQMSIAAVAGPIGALVATAAGLF   173
5SV0      117   GGNGYLATIGAISPFVGLFGTVWGIMNSFIGIAQTQTTNLAVVAPGIAALLATAIGLV   174

NMB1729   174   VAIPAVLAYNFLNR---GTKILTQDLDAMAHDLHVRLLNQKDS   213
5SV0      175   AAIPAVVIYNVFARQIGGFKAMLGDVAAQVLLLQSRDLDLEAS   219
```

Figure 7: (A) Summary of BLAST result for *N. meningitidis* NMB1729 protein (B) Structure-based pairwise sequence alignment of ExbB from *E. Coli* (5SV0) and *N. meningitidis* (NMB1729) used for homology model building.

The NMB1729 protein sequence was searched against UniProt database using FASTA search tool. As a result, a number of homologous sequences from different species were obtained. Among these, 15 homologous sequences were used for multiple sequence alignment by CLUSTALX. Manual editing of alignment was carried out by considering the secondary structure of the template as insertions and deletions should not be present in the corresponding regions of target proteins.

Homology model of *N. Meningitidis* NMB1729 was built based on alignment between target and template sequences. The *E. coli* ExbB protein is known to occur in active homopentameric form; Hence, homology model of NMB1729 protein was built in homopentameric state. All steps of homology modeling and refinement were carried out using the program MODELLER [14]. The homopentameric NMB1729 protein model was comprised of five transmembrane regions forming a transmembrane pore whereas their cytoplasmic domains were arranged in the form of a large enclosed cavity (Fig. 8). After homology model building, model was evaluated by PROCHECK [15] and PROSA [16]. Ramachandran plot showed 90.1% residues in core region and PROSA [16] also confirms that NMB1729 protein model has lower internal energy which indicated that the model is of good quality. The overall fold of *N. meningitidis* ExbB homopentameric protein was observed to be of similar three-dimensional fold as that of *E. coli* ExbB protein (PDB ID; 5SV0).

Homology Model of *N. meningitidis* biopolymer transport protein ExbB (NMB1729) which is a part of Ton complex was constructed. ExbB monomer is known to have three transmembrane spanning helices (TMHs) with a large cytoplasmic domain [25-26]. The quaternary structure of ExbB is a homopentamer where the five transmembrane domains form a transmembrane pore. The cytoplasmic domain forms a large enclosed cavity. The cytoplasmic domain of ExbB retains 5-fold symmetry which can be clearly observed with each monomer consisting of very elongated alpha-helices.

To analyze the amino acid residue conservation in alpha-6 and alpha-7 helices in *N. meningitidis* ExbB protein (P64100), its multiple sequence alignment was performed with nine ExbB proteins from other specie including *E. coli K12* (P0ABU7), *Neisseria gonorrhoeae* (Q5F711), *Haemophilus ducreyi* (O51808), *Vibrio harveyi* (D0XEN5), *Yersinia pestis* (D1TTA4), *Methanothermobacter thermautotrophicus* (O27101), *Pseudomonas aeruginosa* (G3XCW0), ExbB1 of *Vibrio cholerae* (O52043) and ExbB2 of *Vibrio cholerae* (AAC69454) using

CLUSTAX. Result showed that those all amino acid residues were quite conserved among all specie (Fig. 8B & 9). In different species this protein is involved in the process of energy transduction. On the basis of sequence and structural conservation in *N. meningitidis* ExbB protein, we predict that this NMB1729 protein model would also have the same function.

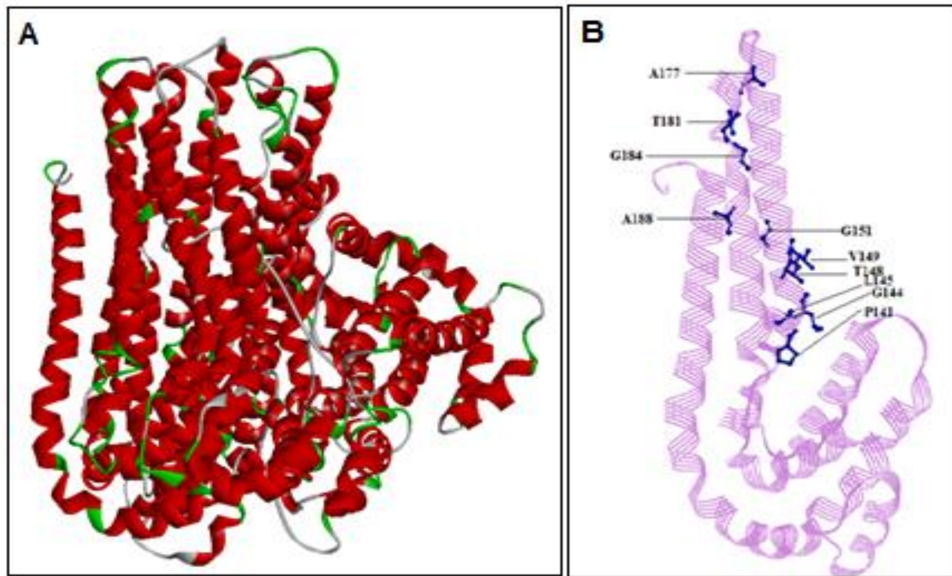

Figure 8: (A) Homology model of *N. meningitidis* biopolymer transport protein exbB (B) Conservation of aminoacid residues in alpha-6 and alpha-7 helices of *N. meningitidis* ExbB model

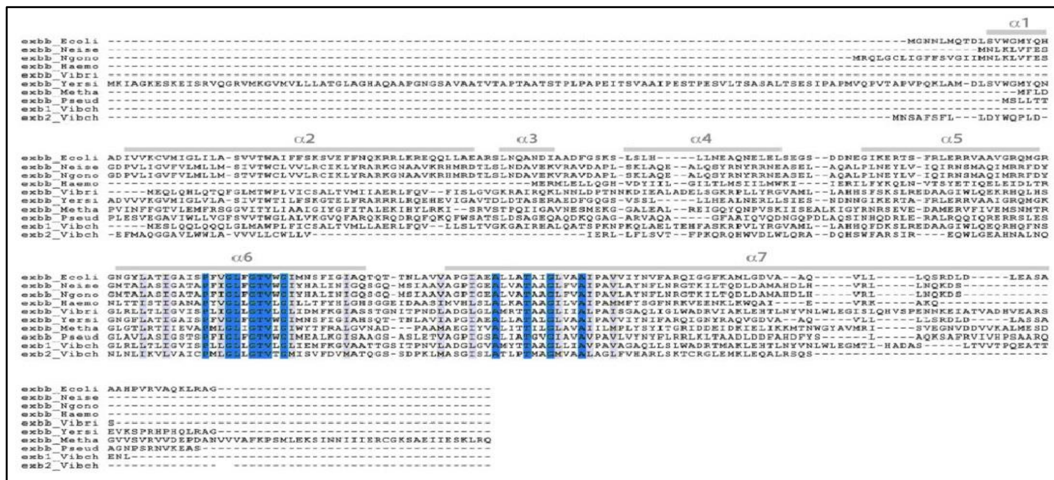

Figure 9: Multiple Sequence alignment of ExbB sequences from different species using CLUSTAL-X. Highlighted area shows conserved amino acid residues among different species.

REFERENCES:

1. Pollard AJ. Global epidemiology of meningococcal disease and vaccine efficacy. Pediatr Infect Dis J. 2004; 23: S274–279
2. Miller E, Salisbury D, Ramsay M. Planning, registration, and implementation of an immunisation campaign against meningococcal serogroup C disease in the UK: a success story. Vaccine. 2001; 20 (Suppl 1): S58–67
3. Jones D., Meningococcal Disease, K. Cartwright, Ed. (Wiley, New York, 1995), p. 147
4. McGuinness BT, Clarke IN, Lambden PR, Barlow AK, Poolman JT, Jones DM, Heckels JE. Point mutation in meningococcal por A gene associated with increased endemic disease. Lancet, 1991; 337(8740):514-7.
5. Invasive Neisseria meningitidis in Europe. London: EU-IBIS Network Health Protection Agency
6. Finne J., Leinonen M., Makela PH. Antigenic similarities between brain components and bacteria causing meningitis: implications for vaccine development and pathogenesis. Lancet. 1983; 322:355–357
7. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, Eisen JA, Ketchum KA, Hood DW, Peden JF, Dodson RJ, Nelson WC, Gwinn ML, Deboy R, Peterson JD, Hickey EK, Haft DH, Salzberg SL, White O, Fleischmann RD, Dougherty BA, Mason T, Ciecko A, Parksey DS, Blair E, Cittone H, Clark EB, Cotton MD, Utterback TR, Khouri H, Qin H, Vamathevan J, Gill J, Scarlato V, Masignani V, Pizza M, Grandi G, Sun L, Smith HO, Fraser CM, Moxon ER, Rappuli R, Venter JC (2000) Science 287:1809–1815
8. Sun YH, Bakshi S, Chalmers R, Tang CM. TANG1. Functional genomics of *Neisseria meningitidis* pathogenesis, Nat Med. 2000; 6(11):1269-73.
9. Altschul, SF., Madden, TL., Schaffer, AA., Zhang, J., Zhang, Z., Miller, W. and Lipman, DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 1997; 25: 3389–3402
10. Berman, HM., Battistuz, T., Bhat, TN., Bluhm, WF., Bourne, PE., Burkhardt, K., Feng, Z., Gilliland, GL., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, JD. and Zardecki, C. The Protein Data Bank. *Acta. Crystallogr. D. Biol. Crystallogr.*,2002; 58: 899-907

11. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T and Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. J. Mol. Biol., 1977; 112(3): 535–542.
12. McGuffin LJ, Bryson K and Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics, 2000;* **16**(4): 404-405.
13. Thompson, JD., Gibson, TJ., Plewniak, F., Jeanmougin, F. and Higgins, DG. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.,* 1997; 25: 4876–4882.
14. Sali, A. 1995. MODELLER: Implementing 3D protein modeling. **2:** 5. Molecular Simulations Inc., Totowa, NJ
15. Laskowski, RA., MacArthur, MW., Moss, DS. and Thornton, JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst*., 1993; 26: 283-291.
16. Sippl, MJ. Recognition of errors in three-dimensional structures of proteins. *Proteins,* 1993; 17: 355–362
17. http://www.biocompare.com/review/ 470/Accelrys-DS-Viewer-50.html/
18. Humphrey, W., Dalke, A. and Schulten, K. VMD-Visual Molecular Dynamics. *J. Mol. Graph.,* 1996; *14.1*: 33-38.
19. Wojdyla J, Kaminska R, Kleanthous C. Structure and function of the Escherichia coli Tol-Pal stator protein TolR. J. Biol. Chem. 2015; 290(44):26675-87
20. Yao Q, Lu Q, Wan X, Song F, Xu Y, Hu M, Zamyatina A, Liu X, Huang N, Zhu P, Shao F. A structural mechanism for bacterial autotransporter glycosylation by a dodecameric heptosyltransferase family. Elife. 2014 Oct 13;3. doi: 10.7554/eLife.03714.
21. Lennox, ES. Transduction of linked genetic characters of the host by bacteriophage P1. Virology, 1955; 1:190-206.
22. Chatterjee, A. K., Sanderson K. E., Ross H., Schlecht S., and Luderitz 0. Influence of temperature on growth of lipopolysaccharide- deficient (rough) mutants of Salmonella typhimurium and S. minnesota. Can. J. Microbiol. 1976; 22:1540-1548.
23. Nikaido H. Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev. 2003; 67(4):593-656.

24. Witty M, et al. Structure of the periplasmic domain of *Pseudomonas aeruginosa* TolA: evidence for an evolutionary relationship with the TonB transporter protein. EMBO J. 2002; 21:4207–4218.

25. Neidhardt, FC., Ingraham, JL., Low, KB., Magasanik, B., Schaechter, MA and Umbarger, HE. *Eschereichia coli* and Salmonella typhimurium: cellular and molecular biology. 1987; American Society for Microbiology, Washington, D.C.

26. Brewer S, et al. Structure and function of X-Pro dipeptide repeats in the TonB proteins of *Salmonella typhimurium* and *Escherichia coli*. J. Mol. Biol. 1990; 216:883–895