Journal of Biochemistry, Biotechnology and Biomaterials
Journal Home Page: http://molcare.org/jbcbb

Structural and Functional Characterization of a Hypothetical protein of *Streptococcus Pyrogenes*: An In-Silico Approach

Muhammad Naveed*[1], Muhammad Matloob[1], Ubair Aziz[2], Muhammad Usman[1], Muhammad Zubair Mehboob[1], Zoma chaudhry[1], Nadia Zeeshan[1], Muhammad Irshad[1], Hira Munir[1]

[1]*Department of Biochemistry & Biotechnology, University of Gujrat, Punjab, Pakistan*

[2]*Institute of Biochemistry & Biotechnology, University of Veterinary & Animal Sciences, Lahore*

Abstract

Streptococcus pyogenes is a human pathogenic bacteria that involved in infectious diseases like scarlet fever, strep throat, glomerulonephritis and necrotizing fasciitis. The *S. pyogenes* infections affect almost 700 million people worldwide and 650,000 patients have severe or invasive infections. Although the mortality rate of these infections is not very high (0.1%) except patients with severe and invasive infection. The antibiotic drugs like *penicillin*, *tetracyclines* and *clindamycin* are considered as the medicines of choice for the control of *S. pyogenes* infection but there are several reports against the development of resistant of *S. pyogenes* strains. In this study, a therapeutically important hypothetical protein from the genomic data of *S. pyogenes* was retrieved and analyzed using different bioinformatics tools such as BLAST, TMHMM, I-TASSER etc. This study provides basic information about structure and function of the hypothetical protein (MATE like multidrug efflux transporter) that involved to efflux the chemicals from bacterial cells thus helping bacteria to survive and would be helpful to design an inhibitor for this transporter. Designed inhibitor can be used with bacterial drugs to enhance drugs retention time in bacterial cells thus helps to control diseases caused by this pathogen.

Keywords: *Streptococcus pyogenes*; hypothetical proteins; scarlet fever; BLAST; MATE like multidrug efflux transporter.

Full length article: Received: 20 December, 2016 Revised: 28 December, 2016 Accepted:29 December, 2016
***Corresponding Author* e-mail**: naveed.quaidian@gmail.com; dr.naveed@uog.edu.pk

1. Introduction

Over the past many years, a huge amount of genomic and transcriptomic data has been deposited in online databases. A large portion of this genomic data has been proposed to encode a protein based on the computational analysis of gene structure and nucleic acid sequences but there is no evidence for their in vivo expression. These proteins are termed as hypothetical proteins (HPs). For some organisms, these HPs might represent up to half of their reported proteome (Minion *et al.,* 2004). The structural and functional characterization of these HPs of unknown function is a challenge for functional genomics as well as for general biology. If we can cope with this challenge, this might fill the gaps between genomic sequence data and structural and functional genomics as well as proteomics. HPs functional annotation is of crucial importance for development of potential antibacterial agents against pathogens as well as for our better understandings for the drug resistance, infection and other biological pathways (Naveed *et al.,* 2016b).

Streptococcus pyogenes is a gram-positive beta-hemolytic bacterium that causes multiple infectious diseases in humans ranging from strep throat to scarlet fever or some other serious diseases including necrotizing fasciitis or glomerulonephritis (Galperin and Koonin 1999). These infections affect round about 700 million people worldwide with a mortality rate of only 0.1%. Out of these 700 million people, 650,000 people have a severe or invasive infection of *S. pyogenes* with a mortality rate of ~25 % (Aziz *et al.*, 2010).

The mechanism of infection of *S. pyogenes* starts with the adhesion of bacterial cells to the epithelial cells of the skin, nasal or buckle cavity of humans. For this purpose, different strains of *S. pyogenes* express different proteins for the binding to human extracellular matrix proteins including collagen (Bisno *et al.*, 2003; Ryan and Ray 2004), laminin (Terao *et al.*, 2002) and fibronectin (Podbielski *et al.*, 1999). After the invasion into the human body, *S. pyogenes* has the ability to spread rapidly to different organs. This ability of *S. pyogenes* in conferred to its resistance against the human immune response. It has been reported that fewer neutrophils migrate toward the *S. pyogenes* infection site (Hidalgo-Grass *et al.*, 2004). It has also been suggested that *S. pyogenes* escapes complement system of the human immune system by targeting proteins of the complement system (Ryan and Ray 2004) including C5a protein (Wexler and Cleary 1985). Other virulence factors of *S. pyogenes* include the presence of hyaluronic acid capsule around its body provide protection against the phagocytosis (Ryan and Ray 2004).

Antibiotic treatments have been used against the *S. pyogenes* infections for a long time. But with time, these bacteria have gained resistance to many antibiotic drugs. The treatment of choice has been penicillin (Falagas *et al.*, 2008) but many reports of penicillin tolerance have been made since 1985 (Kim and Kaplan 1985). In different studies, many other antibiotics have also been tested and their resistance has been reported. In Italy 32% of *S. pyogenes* showed resistance to macrolides while in another study 11% of *S. pyogenes* were reported to be resistant to macrolides in Portugal. In France, 23% of *S. pyogenes* were reported to be resistant to the action of erythromycin (Lamagni *et al.*, 2008). So, there is need to find new targets against which bacterial resistance is still not developed. Many hypothetical proteins are used as a new hotspot for therapetical targets (Barragán-Osorio *et al.*, 2016). so, this study aims to highlight structural and functional characteristics of the therapeutical important hypothetical protein of *Streptococcus pyrogenes*.

2. Material and Methods

2.1- Selection of Hypothetical Protein

GenBank database of NCBI was used to search the hypothetical proteins of *S. pyrogenes*. The whole genome was explored through GenBank and all the hypothetical protein from the whole genome was searched to find out therapeutically important hypothetical protein and selected hypothetical protein had Accession ID of NP_268458.1 in GenBank.

2.2- Conserved Domain Database

Conserved Domain Database or CDD is an interface provided by NCBI for searching the domains present in the query sequence. The query could be a protein or nucleotide FASTA sequence. The output of submitted query is shown as a multiple sequence alignment with conserved domains annotated graphically (Marchler-Bauer and Bryant 2004; Naveed *et al.*, 2014).

2.3- BLAST

Basic Local Alignment Search Tool or BLAST is another online service provided by the NCBI. This service is based on the local alignment of protein and DNA sequences to find a similar protein or DNA sequences respectively using heuristic methods for quick results (Madden 2013). Firstly, BLASTP was used to find out proteins with similar sequences in non-redundant databases with respect to that of selected HP. This analysis provided us with the information about the proteins having similar sequence and possibly similar function to

that of selected HP. Secondly, BLASTP used against PDB database to find out any protein with similar sequence whom structure has been reported in any previous research.

2.4- ProtParam

ProtParam is a bioinformatics tool for the prediction of various physiochemical parameters for the protein structure (Gasteiger *et al.,* 2005). These parameters include molecular weight of protein, theoretical pI, ratio of different amino acids in protein, atomic composition, chemical formula, the total number of atoms in the protein, estimated half-life in different cellular environments, stability, hydropathicity, extinction coefficient, Instability Index and aliphatic index.

2.5- TMHMM (v2.0)

TMHMM is transmembrane protein topology prediction software based on the hidden Markov model with a high level of accuracy. TMHMM is an online tool for the prediction of secondary structure prediction of proteins with high accuracy, sensitivity, and specificity. The accuracy of TMHMM was calculated to be 97-98% and its sensitivity and specificity to discriminate between membrane proteins and soluble proteins in the absence of signal peptides were more than 99% (Krogh *et al.,* 2001; Naveed *et al.,* 2014).

2.6- I-TASSER

I-TASSER (Iterative Threading Assembly Refinement) is a hierarchical structure modeling method for the prediction of full-length atomic models from primary protein sequence of amino acids. I-TASSER identifies the templates for query sequence through multiple threading approach and then constructs 3-D model based on the information from these templates (Naveed *et al.,* 2016a; Zhang 2008; Zhang and Skolnick 2013)

2.7- ERRAT

ERRAT is a bioinformatics tool for differentiation of correctly determined protein regions from that of incorrectly registered. Atoms that are non-randomly distributed throughout the protein might be randomly distributed during protein modeling. ERRAT analyses the query model for these distributions of atoms. ERRAT provides with an overall quality factor evaluating the nonbonding atomic interactions whereas higher scores signifying higher quality. Normally accepted a range of value is >50 (Li and Wang 2007; Naveed M *et al.,* 2016b).

2.8- STRING

STRING (search tool for recurring instances of neighboring genes) is a bioinformatics tool for the retrieval and display of functionally related genes. For this purpose, STRING searches in the clusters in which the query gene repeatedly occurs. The principle of STRING is based on the previous studies that suggest genes repeatedly occurring in close proximity in genomes tend to express functionally related proteins that are involved in same metabolic pathways (Mushegian and Koonin 1996; Naveed *et al.,* 2016a; Overbeek *et al.,* 1999a; Overbeek *et al.,* 1999b; Peer Bork *et al* ; Tamames *et al.,* 1997; Watanabe *et al.,* 1997).

3. Results

All information of therapeutically important hypothetical protein was retrieved from GenBank of NCBI. The FASTA entry of hypothetical protein used for the query is given below.

>gi|15674285|ref|NP_268458.1| conserved hypothetical protein [*Streptococcus pyogenes* M1 GAS]

MIYNNRRKIFSLALPSMIENILQMLMGMVDNYVAQIGVAVSGVSIANNIISIYQSL-FIALG
AAVSSLIARSIGENNQNKQLNMAGVLQVTLLLSVGLGLLSVAGHHQVL EWLGAEASVT
LVGGQYLSIVGGMIVSLGLLTSLGAIVRAQGYPKIP MQVSL LINVLNAI FSALSI YVWG
FGLLGVAWATVLSRLVGVFLLCQFIPIKQVAKRL-MRPLDKIIFDLSLPAAGERLMMRAG

DVLIIGIVVRFGTTALAGNAIGE-TLTQFNYMPGLAMATATIILVARQLG GGKVTEIRYIIR
EAFILSTLMMLVM-GALTYLLGPSLLPLFTQNTDAQRSAMIVLLFSLLG APATAGTLVYT
AVWQGLGKAKLPFYATTIGMWVIRIGLGYVIGVVWQYGLIGVWMA TVLDNTSRW-FIL
SKHFKKYQEITFH

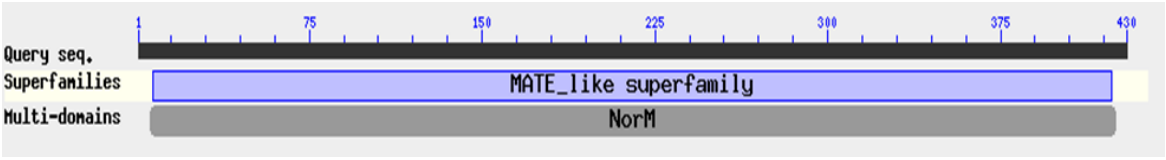

Figure 1: conserve domain of hypothetical protein predicted by CDD

In Analysis of selected HP from Streptococcus pyogenes, Conserved Domain Database of NCBI suggested the presence of MATE_like multi-drug efflux transporter domain 'NorM' throughout the protein sequence with an e-value of 1.70e-74 as shown in Figure 1. In BLASTP, the query was submitted to the non-redundant database and resulted in hits also showed the MATE_like multidrug efflux pump in different Streptococcus species i.e. *Streptococcus equinus, Streptococcus dysgalactiae*, with high score identity and good e-value confirming the results of CDD. Few hits of BLASTP was given in Table 1 with their e-values and identity scores. BLASTP results against PDB database to find out similar proteins whom 3D structures have already been charac-terized, provided only three candidate proteins with adequate e-value < 1e-6. All of these three proteins were MATE transporters in *Pyrococcus furiosus* (e-value = 2e-17, 6e-17, 1e-15) as shown in Table 2. These predicted hits also supported the results of our previous analysis. Fur-thermore, the availability of 3D-structures of proteins with sufficient sequence similarity (query coverage = 94%, identity 24%) provided us with a template for 3D-structure prediction process.

ProtParam provided us with some important physiochemical features about the query protein. The extinction coefficient of selected HP was 66350 M-1 cm-1 and predicted absorbance was 0.1% at 280 nm wavelength in 1g/L of selected HP. This coefficient is of crucial importance while following a protein during protein purification. The half-life of query sequence was pre-dicted to be 30 hours, 20 hours and 10 hours in mammalian reticulocyte, yeast, and E. coli respectively based upon the Methionine being the N-terminal residue. Proteins with instability in-dex of <40 are assumed to be stable instability index of our query protein was calculated to be 27.95 hence the protein was predicted to be stable.

The aliphatic index of HP was calculated to be 132.65 that suggest that this protein is very stable at higher temperatures (Atsushi 1980). The GRAVY score represents the calculated hydropathy of a protein that suggests the cellular localization of a protein based upon the hydrophilic or hydrophobic nature of amino acids present in it. GRAVY score of our protein was calculated to be 0.813. The positive value suggests the overall abundance of hydrophobic amino acids, thus supporting our previous prediction of NorM_like protein that is to be present in the lipid bilayer of a cell. The molecular weight and theoretical pI values of HP sequence were calculated to be 46532.7 and 9.84 respectively. pI value is of grave importance while purifying or separating a protein from a mixture of proteins on pH gradient gels. Summary of all the physiochemical properties of HP are shown in Table 3.

Table 1: BLASTP hits against non-redundant database

Description	Max score	Total score	Query cover	E-value	Ident	Accession
MATE family efflux transporter [*Streptococcus pyogenes*]	846	846	100%	0	99%	WP_028796381.1
MATE family efflux transporter [*Streptococcus pyogenes*]	845	845	100%	0	99%	WP_063631702.1
MULTISPECIES: MATE family efflux transporter [*Streptococcus*]	835	835	100%	0	99%	WP_002987736.1
MATE family efflux transporter [*Streptococcus* dysgalactiae]	635	635	99%	0	77%	WP_046177506.1

Table 2: BLASTP hits against pdb database

Description	Max score	Total score	Query cover	E value	Ident	Accession
A Mop Superfamily Transporter from *Pyrococcus furiosus* Vc1 (dsm3638)	85.1	85.1	94%	2.00E-17	24%	4MLB_A
Chain A, Crystal Structure of Mate in The Straight Conformation	83.2	83.2	94%	6.00E-17	24%	3VVN_A
Chain A, Crystal Structure of Mate P26a Mutant	79.3	79.3	94%	1.00E-15	24%	3W4T_A

Cellular location of proteins is very crucial for its stability and function according to its amino acids contents. TMHMM, a tool for the secondary structure prediction of transmembrane proteins, predicted 11 transmembrane helices in query HP with N-terminal in the inside region and C-terminal in the outside of the cell membrane as shown in Figure 2, confirming the HP as a transporter.

As the structure of HPs is unknown so the 3D structure was predicted by I-TASSER with a C-score of 1.56, TM-score of 0.93±0.06 and RMSD value of 3.8±2.6Å. The structure pre-dicted contained 12 transmembrane helices as present in NorM efflux pump (van Veen 2010). This structure was also verified by the ERRAT for the misconfiguration of atoms. ERRAT results calculated the overall quality factor of 97.867 thus suggesting the high quality of 3D-model shown in Figure 3. Interaction of selected HP with other protein was checked through STRING tool that predicted its interaction with three proteins.These proteins are rpsJ, that encodes 30S ribosomal protein S10, adhA, that encodes an alcohol dehydrogenase, and cinA, that encodes a competent damage-inducible protein shown in Figure 4.

Table 1: Physiochemical properties predicted by Protparam server

Molecular weight	Theoretical pI	Extinction coefficients	Estimated half-life	Instability index	Aliphatic index	GRAVY
46532.7	9.84	66350	30 hours (mammalian reticulocytes, in vitro).	27.95	132.65	0.813
			>20 hours (yeast, in vivo).			
			>10 hours (Escherichia coli, in vivo).			

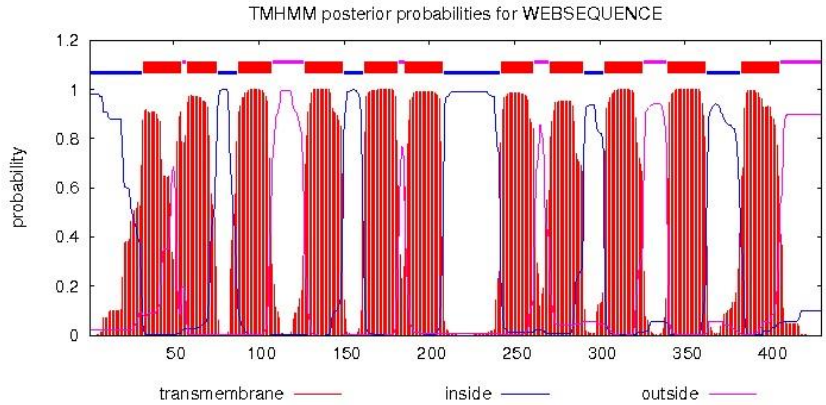

Figure 2: Alpha helices with C and N terminal of HP

4. Discussion

Today a lot of studies are carried out to use hypothetical proteins for their functional and structure characterization. Some studies are focusing on targeting them in several diseases (Barragán-Osorio *et al.*, 2016) and more specifically, these protein have also been used as a vac-cine antigen (Lage *et al.*, 2016; Martins *et al.*, 2016). The different predicted functions such as involvement as a pathogenicity factor (Conrad *et al.*, 2016), as a histone acetyltransferase (Jose *et al.*, 2016)and chemotaxis (Zhang *et al.*, 2016) make them essential for targeting and controlling microorganisms. By seeming out the previous data, this study also aims to characterize structural-ly and functionally a hypothetical protein of S.pyrogenes and highlight its importance to use it as a target to control bacteria.

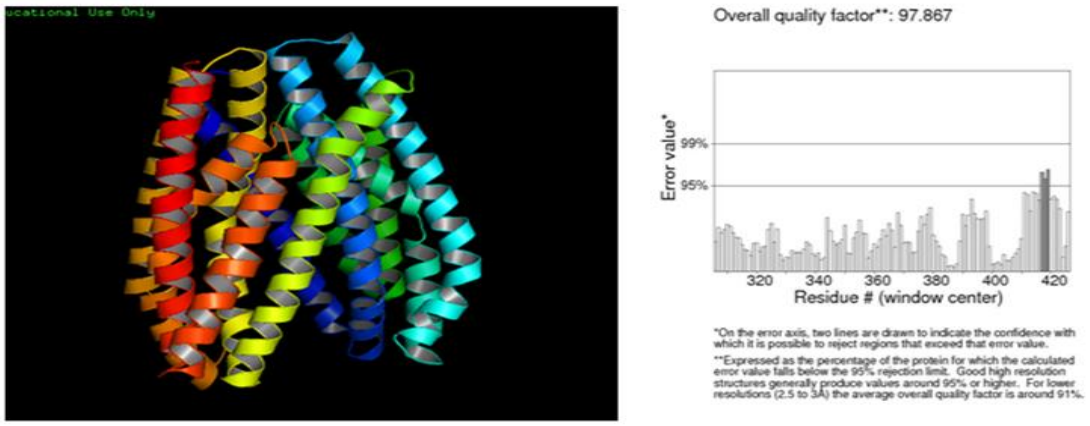

Figure 3: 3D structure predicted by I-TASSER and verified table by ERRAT

MATE is one of the five superfamilies of multidrug efflux transporters (Brown *et al.*, 1999) that are present in all kingdoms of life (van Veen 2010). MATE efflux transporters have been found in many bacterial species including Acinetobacter baumannii (Su *et al.*, 2005), Clos-tridium difficile(Dridi *et al.*, 2004), Brucella melitensis(Braibant *et al.*, 2002), Haemophilus influ-enza (Xu *et al.*, 2003), Neisseria meningitides (Rouquette-Loughlin *et al.*, 2003), Neisseria gonor-rhoeae(Singh *et al.*, 2006), Pseudomonas aeruginosa(Huda *et al.*, 2003), Vibrio cholera(Begum *et al.*, 2005; Kaatz *et al.*, 2005; Moreno *et al.*, 2013) and Staphylococcus aureus(Tettelin *et al.*, 2000). These microbes are involved in causing diverse infectious diseases in humans including Pseudomembranous colitis (Ryan and Ray 2004), meningitis, septicemia (Morita *et al.*, 1998) and gonorrhea (Hvorup *et al.*, 2003). NorM, a member of MATE superfamily of multidrug efflux transporters (Brown *et al.*, 1999) was firstly found in Vibrio parahaemolyticus causing efflux of norfloxacin, ciprofloxacin, kanamycin, ethidium and streptomycin from the intracellular environ-ment to the outside decreasing the cellular concentrations of these chemicals to subtoxic levels(Morita *et al.*, 1998) thus helps the bacteria to survive in host. These multidrug efflux trans-porters are considered as main intrinsic antibiotic resistance in gram-negative bacteria (Webber and Piddock 2003).

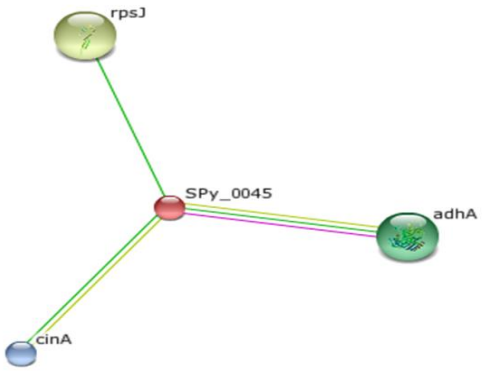

Figure 2: Interaction of HP with other proteins predicted by STRING

The recent studies suggest that multidrug efflux transporters are considered as the main determinant of the efficacy of both old and new antibiotics (Lomovskaya and Bostian 2006). As a consequence of resistance and more antibiotic development, the scientist also focuses on to overcome the antibiotic resistance. Studies that mainly focus on the structure of multidrug efflux pumps not only provide information about the mechanism of transporter but also highlights the discovery of structure-based efflux pump inhibitor. The present study also provides insight in Na+ dependent multi-drug efflux pump by designing its structure thus suggesting to develop its inhibitor that can be used with antibacterial drugs to stops their efflux from bacterial cells.

5- Conclusion

Hypothetical proteins are considered as orphaned proteins because of their unknown structure and function. In this study, different bioinformatics tools are applied to one of thera-peutically important hypothetical protein to find its structure and function. we found that select-ed hypothetical protein is a transporter that involved to efflux the chemicals from bacterial cells thus helping bacteria to survive. This study helps design an inhibitor that can be used with bac-terial drugs to inhibit their efflux from bacterial cell thus improving the drugs results.

6. References

1. Atsushi, IKAI 1980. Thermostability and aliphatic index of globular proteins. Journal of biochemistry 88(6):1895-1898.
2. Aziz, Ramy K *et al.*, 2010 Microevolution of group A streptococci in vivo: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. PloS one 5(4):e9798.
3. Barragán-Osorio, Laline *et al.*, 2016. Computational Analysis and functional prediction of ubiquitin hypothetical protein: a possible target in Parkinson disease. Central Nervous System Agents in

Medicinal Chemistry (Formerly Current Medicinal Chemistry-Central Nervous System Agents) 16(1):4-11.

4. Begum, Anowara *et al.,* 2005. Gene Cloning and Characterization of Four MATE Family Multidrug Efflux Pumps from Vibrio cholerae Non-O1. Microbiology and immunology 49(11):949-957.
5. Bisno, AL, MO Brito, and CM Collins. 2003Molecular basis of group A streptococcal virulence. The Lancet infectious diseases 3(4):191-200.
6. Braibant, Martine, Laurence Guilloteau, and Michel S Zygmunt. 2002 Functional characterization of Brucella melitensis NorMI, an efflux pump belonging to the multidrug and toxic compound extrusion family. Antimicrobial agents and chemotherapy 46(9):3050-3053.
7. Brown, Melissa H, Ian T Paulsen, and Ronald A Skurray. 1999. The multidrug efflux protein NorM is a prototype of a new family of transporters. Molecular microbiology 31(1):394-395.
8. Conrad, Turner Allen *et al.,* 2016The chromosome-encoded hypothetical protein TC0668 is an upper genital tract pathogenicity factor of Chlamydia muridarum. Infection and immunity 84(2):467-479.
9. Dridi, L, J Tankovic, and J-C Petit. 2004. CdeA of Clostridium difficile, a new multidrug efflux transporter of the MATE family. Microbial Drug Resistance 10(3):191-196.
10. Falagas, Matthew E *et al.,* 2008. Effectiveness and safety of short-course vs long-course antibiotic therapy for group A β-hemolytic streptococcal tonsillopharyngitis: a meta-analysis of randomized trials. Mayo Clinic Proceedings, 2008. Vol. 83, pp. 880-889. Elsevier.
11. Galperin, Michael Y, and Eugene V Koonin. 1999. Searching for drug targets in microbial genomes. Current opinion in biotechnology 10(6):571-578.
12. Gasteiger, Elisabeth *et al.,* 2005. Protein identification and analysis tools on the ExPASy server: Springer.
13. Hidalgo-Grass, Carlos *et al.,* 2004. Effect of a bacterial pheromone peptide on host chemokine degradation in group A streptococcal necrotising soft-tissue infections. The Lancet 363(9410):696-703.
14. Huda, Md *et al.,* 2003. Gene Cloning and Characterization of VcrM, a Na+-Coupled Multidrug Efflux Pump, from Vibrio cholerae Non-O1. Microbiology and immunology 47(6):419-427.
15. Hvorup, Rikki N *et al.,* 2003. The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. European Journal of Biochemistry 270(5):799-813.
16. Jose, Leny *et al.,* 2016. Hypothetical protein Rv3423. 1 of Mycobacterium tuberculosis is a histone acetyltransferase. FEBS journal 283(2):265-281.
17. Kaatz, Glenn W, Fionnuala McAleese, and Susan M Seo. 2005. Multidrug resistance in Staphylococcus aureus due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. Antimicrobial Agents and Chemotherapy 49(5):1857-1864.
18. Kim, Kwang Sik, and Edward L Kaplan. 1985. Association of penicillin tolerance with failure to eradicate group A streptococci from patients with pharyngitis. The Journal of pediatrics 107(5):681-684.
19. Krogh, Anders *et al.,* 2001Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of molecular biology 305(3):567-580.
20. Lage, Daniela P *et al.,* 2016. A new Leishmania-specific hypothetical protein and its non-described specific B cell conformational epitope applied in the serodiagnosis of canine visceral leishmaniasis. Parasitology research 115(4):1649-1658.
21. Lamagni, Theresa L *et al.,* 2008. Epidemiology of severe *Streptococcus pyogenes* disease in Europe. Journal of clinical microbiology 46(7):2359-2367.
22. Li, Minyong, and Binghe Wang 2007. Homology modeling and examination of the effect of the D92E mutation on the H5N1 nonstructural protein NS1 effector domain. Journal of molecular modeling 13(12):1237-1244.
23. Lomovskaya, Olga, and Keith A Bostian. 2006. Practical applications and feasibility of efflux pump inhibitors in the clinic—a vision for applied use. Biochemical pharmacology 71(7):910-918.
24. Madden, Thomas. 2013. The BLAST sequence analysis tool.
25. Marchler-Bauer, Aron, and Stephen H Bryant. 2004. CD-Search: protein domain annotations on the fly. Nucleic acids research 32(suppl 2):W327-W331.

26. Martins, Vívian Tamietti *et al.*, 2016. A new Leishmania-specific hypothetical protein, LiHyT, used as a vaccine antigen against visceral leishmaniasis. Acta tropica 154:73-81.
27. Minion, F Chris *et al.*, 2004. The genome sequence of Mycoplasma hyopneumoniae strain 232, the agent of swine mycoplasmosis. Journal of bacteriology 186(21):7123-7133.
28. Moreno, Megan A, Fred Furtner, and Frederick P Rivara. 2013Clostridium difficile: a cause of diarrhea in children. JAMA pediatrics 167(6):592-592.
29. Morita, Yuji *et al.*, 1998. NorM, a putative multidrug efflux protein, of Vibrio parahaemolyticus and its homolog in Escherichia coli. Antimicrobial agents and chemotherapy 42(7):1778-1782.
30. Mushegian, Arcady R, and Eugene V Koonin. 1996. Gene order is not conserved in bacterial evolution. Trends in Genetics 12(8):289-290.
31. Naveed M *et al.*, 2016b. Computational Analysis and Polymorphism study of Tumor Suppressor Candidate Gene-3 for Non Syndromic Autosomal Recessive Mental Retardation. Journal of Applied Bioinformatics & Computational Biology 5(2).
32. Naveed, Muhammad *et al.*, 2014. Bioinformatics based structural characterization of glucose dehydrogenase (gdh) gene and growth promoting activity of Leclercia sp. QAU-66. Brazilian Journal of Microbiology 45:603-611.
33. Naveed, Muhammad *et al.*, 2016a. In-silico analysis of non-synonymous-SNPs of STEAP2: To provoke the progression of prostate cancer. *In* Open Life Sciences. Pp. 402, Vol. 11.2016b In-silico analysis of non-synonymous-SNPs of STEAP2: To provoke the progression of prostate cancer. Open Life Sciences 11(1):402-416.
34. Overbeek, Ross *et al.*, 1999a Use of contiguity on the chromosome to predict functional coupling. In silico biology 1(2):93-108.
35. Overbeek, Ross *et al.*, 1999b. The use of gene clusters to infer functional coupling. Proceedings of the National Academy of Sciences 96(6):2896-2901.
36. Peer Bork, Thomas Dandekar *et al.*, Predicting Function: From Genes to Genomes and Back. Podbielski, Andreas *et al.*, 1999. Characterization of nra, a global negative regulator gene in group A streptococci. Molecular microbiology 31(4):1051-1064.
37. Rouquette-Loughlin, Corinne *et al.*, 2003. The NorM efflux pump of Neisseria gonorrhoeae and Neisseria meningitidis recognizes antimicrobial cationic compounds. Journal of bacteriology 185(3):1101-1106.
38. Ryan, KENNETH J, and C GEORGE Ray. 2004. Medical microbiology: McGraw Hill.
39. Singh, Anil Kumar *et al.*, 2006. Analysis of the topology of Vibrio cholerae NorM and identification of amino acid residues involved in norfloxacin resistance. Antimicrobial agents and chemotherapy 50(11):3717-3723.
40. Su, Xian-Zhong *et al.*, 2005. AbeM, an H+-coupled *Acinetobacter baumannii* multidrug efflux pump belonging to the MATE family of transporters. Antimicrobial agents and chemotherapy 49(10):4362-4364.
41. Tamames, Javier *et al.*, 1997. Conserved clusters of functionally related genes in two bacterial genomes. Journal of molecular evolution 44(1):66-73.
42. Terao, Yutaka *et al.*, 2002 Novel laminin-binding protein of *Streptococcus pyogenes*, Lbp, is involved in adhesion to epithelial cells. Infection and immunity 70(2):993-997.
43. Tettelin, Herve *et al.*, 2000. Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science 287(5459):1809-1815.
44. van Veen, Hendrik W. 2010. Structural biology: Last of the multidrug transporters. Nature 467(7318):926-927.
45. Watanabe, Hidemi *et al.*, 1997. Genome plasticity as a paradigm of eubacteria evolution. Journal of molecular evolution 44(1):S57-S64.
46. Webber, MA, and LJV Piddock. 2003. The importance of efflux pumps in bacterial antibiotic resistance. Journal of Antimicrobial Chemotherapy 51(1):9-11.
47. Wexler, DANIEL E, and P PATRICK Cleary. 1985. Purification and characteristics of the streptococcal chemotactic factor inactivator. Infection and immunity 50(3):757-764.

48. Xu, Xing-Jue *et al.,* 2003. Molecular cloning and characterization of the HmrM multidrug efflux pump from Haemophilus influenzae Rd. Microbiology and immunology 47(12):937-943.
49. Zhang, Kai *et al.,* 2016. Hypothetical Protein BB0569 Is Essential for Chemotaxis of the Lyme Disease Spirochete Borrelia burgdorferi. Journal of bacteriology 198(4):664-672.
50. Zhang, Yang. 2008. I-TASSER server for protein 3D structure prediction. BMC bioinformatics 9(1):40.
51. Zhang, Yang, and Jeffrey Skolnick. 2013. Segment assembly, structure alignment and iterative simulation in protein structure prediction. BMC biology 11(1):44.