

Computational modeling of tyrosine protein phosphatase WZH in G54 Strain of *Streptococcus pneumoniae* and its ligand identification

Balasankar Karavadi* and M Xavier Suresh²

¹Department of Bioinformatics, Sathyabama University, Chennai- 600119, INDIA

²Department of Physics, Sathyabama University, Chennai- 600119, INDIA

*Corresponding author: E-Mail: balasankar.sathyabamauniv@gmail.com

ABSTRACT

In G54 strain of *Streptococcus pneumoniae* Cps B is an element in tyrosine phosphorylation regulatory system which is involved in modulation of capsule synthesis. Protein coded by this gene is tyrosine protein phosphatase wzh which offers a possible target identification using computational drug design. Identifying Drug target makes an important molecule involved in infectivity or survival of microbial pathogen, tyrosine protein phosphatase wzh was taken as the drug target and the structure of this protein was homology modeled using Modeller 9v7 and validated through SAV Server. Ligands for tyrosine protein phosphatase wzh were taken from drug port and its analogues were retrieved from pubchem. Discovery studio was used to dock these ligands with modeled proteins and the ADMET properties of the analogs were also studied.

KEY WORDS: Pneumonia, Tyrosine phosphorylation, Homology Modeling and Drug design.

1. INTRODUCTION

Streptococcus pneumoniae is one of the human pathogen that causes pneumonia, bacterial meningitis, sepsis, otitis media etc. Strain G54 is an Italian clinically isolated which is resistant to macrolides and tetracycline belongs to serotype 19F (Dopazo, 2014). In *Streptococcus pneumoniae*, capsular synthesis in G54 strain is modulated through tyrosine phosphorylation regulatory system involved by cpsB, cpsC, and cpsD genes (Brueggemann, 2007). These genes codes capsular polysaccharide synthesis proteins.

Proteins coded by these genes are tyrosine protein phosphatase, capsular polysaccharide synthesis protein, tyrosine protein kinase respectively (Bender, 2003). Diseases such as pneumonia, otitis media, sepsis and meningitis causes in wide range due to *Streptococcus pneumoniae* is a human, and this result in significant global deaths and illness (Hava and Camilli, 2002).

The protein-capsular polysaccharide conjugates were involved in discovering licensed vaccines for pneumococcal disease (Hicks, 2007). However, these have severe limits in terms of price, serotype coverage and increases in the frequency of disease caused by non-vaccine serotypes post introduction. (McCullers and Tuomanen 2001; Yother, 2011).

Multilocus sequence typing has recognized this G54 strain as a serogroup 15, even though capsular locus is serotype 19F, which indicates that strain has gone through transformation (Dopazo, 2014). The strain of G54 has length of 2,078,953 nt, Genes of 2186, Protein coding 2114, Structural RNA's 71, coding of 85% and GC content of 39%.

2. MATERIALS AND METHODS

Sequence of the identified target protein SPG_0315 of *Streptococcus pneumoniae* G54 were retrieved from UniProtKB/Swiss-Prot. Template for the target protein was identified and retrieved using PDBsum based on the identity of the target sequence. 3D coordinate of the template which is used for modeling the target protein was retrieved from PDB. The alignment of raw sequence and template sequences was performed with Clustal W and modeling of target protein on homology basis to fine the coordinated was performed by Modeller 9v7. SAVS was used to visualize conformational parameters of the modeled protein and their amino acid residues through ramachandran plot. To view the protein structure and to carry out energy minimization for making the protein stable we used Swiss PDB viewer. Drugs which can be used as the ligands designed for the protein molecule were retrieved from Drug port and the ligands were retrieved from PubChem from which the 3D coordinates of the ligands was obtained, structural pockets and cavities were predicted from Castp, Argus lab is performed for finding the binding energy required for each ligand to get docked with the respective protein. Discovery Studio 2.0 was used for docking, Receptor-ligand interactions were performed and analyzed to find interactions between ligand and protein receptors. This allow to carry out structure-based design or even to examine possible interactions with theoretical structures such as homology models. The Ligand Fit docking was performed to specify the regions of the receptor's binding site. The Discovery studio 2.0 was used to perform ADMET properties for all the small molecules.

3. RESULTS

Homology modeling was performed for tyrosine protein phosphatase wzh (SPG_0315) and was modeled using the template structure (PDB Ids: 2WJE:A). The modeled protein was analyzed and verified using SAVS and the validation results are given in Table 1. From the result we infer that 93.3% residues in target proteins are present in the Ramachandran Plot's allowed region.

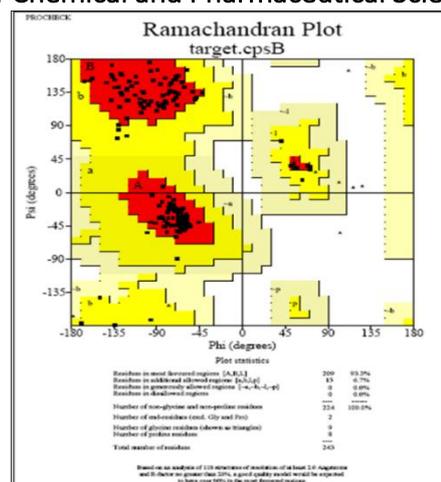
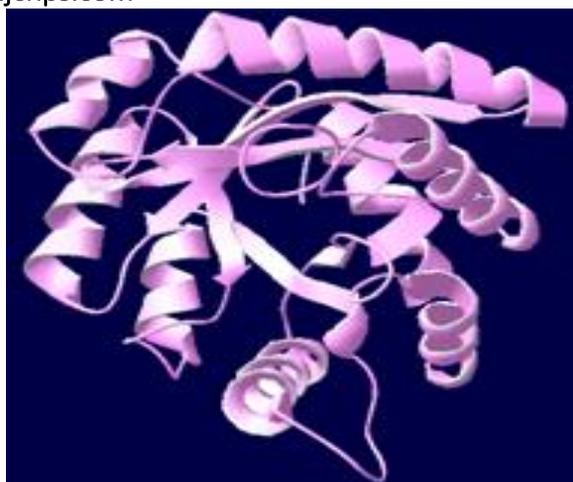


Figure.1. Structure of tyrosine protein phosphatase wzh (SPG_0315) and its Ramachandran plot

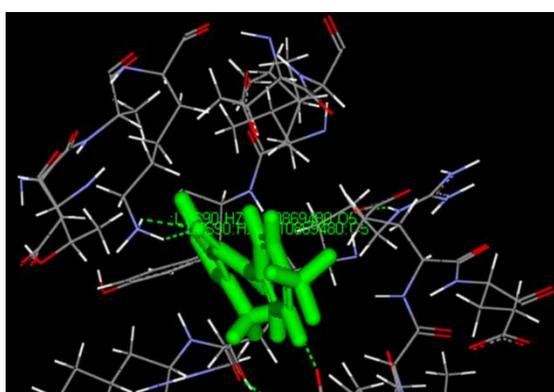


Figure.2. The protein-ligand interactions of 5-methyl-N-[4-(trifluoromethyl) phenyl]-1, 2-oxazole-4-carboxamide with Tyrosine protein phosphatase wzh

Table.1. List of analogues for the receptor SPG_0315 and their binding energy and dock score

Ligands and analogues		Binding energy	Site	H-bond	Amino acid	Dock score
Olopatadine	2-[(11Z)-11-[3-(dimethylamino)propylidene]-6H-benzo[c][1]benzoxepin-2-yl]acetic acid.	-9.03515	Site4	1	Arg139	21.15
	2-[(11Z)-6,6-dideuterio-11-[3-(dimethylamino)propylidene]benzo[c][1]benzoxepin-2-yl]acetic acid.	-9.10643	---	---	---	---
Leflunomide	5-methyl-N-[4-(trifluoromethyl)phenyl]-1,2-oxazole-4-carboxamide.	-7.86362	Site1	2	Lys90	39.76
	N-[4-(1,1-difluoroethyl)phenyl]-5-methyl-1,2-oxazole-4-carboxamide.	-8.18458	Site4	1	Arg139	38.88

The Ramachandran plots and its respective final modeled protein structure is shown in Figure 1. Ligands for Tyrosine protein phosphatase wzh (SPG_0315) were retrieved from Drug Port sharing more identity with related protein sequence for which already a drug exists. PubChem database was used to find the best analogs based on best hits for each ligands.

Table.2. The percentage of residues of modeled structure present in the allowed region of Ramachandran plot as predicted by SAVS with its similarity and template description

Target Protein	length	Template	Description of template	Ramachandran Plot (%)
Tyrosine protein phosphatase wzh (SPG_0315)	243	2WJE:A	Tyrosine phosphatase cps4b from <i>S.pneumoniae</i> tigr4	93.3

The docking was calculated using Discovery Studio 2.0 software with the chosen analogs. Dock score was calculated for all best chosen analogs. The Leflunomide ligand molecule had the best analog compound 2-[(11Z)-11-[3-(dimethylamino) propylidene] -6H-benzo[c][1]benzoxepin-2-yl]acetic acid with dock score having 21.15. The Cefazolin ligand molecule had a best analog compounds 5-methyl-N-[4-(trifluoromethyl)phenyl]-1,2-oxazole-4-

carboxamide, N-[4-(1,1-difluoroethyl)phenyl]-5-methyl-1,2-oxazole-4-carboxamide with dock score having 39.79,38.88 respectively. The analog compounds with docking score more than 30.0 were considered to be the best for which ADMET studies were performed.

The top scoring analogs, binding energy, H-bonds, interacting amino acid and Dock score are shown in Table 2. The protein-ligand interactions at the binding site are illustrated in Figure 2. ADMET properties for the best analogs having good dock score and maximum interaction with the active pocket residues were analyzed. Based on our analysis, it has been found that the analogs which had maximum dock score have proper $logP$, Absorption and Blood Brain barrier value.

DISCUSSION

There had been a lot of research on G54 and many other strains of *Streptococcus pneumoniae*, though there is a lot of worldwide mortality due to pneumonia. So we have concentrated on G54 strain for the analysis due to its virulent nature. We found were 2114 proteins coded by this strain in which many proteins are not crystallized and our objective is to identify the efficient drug target in these proteins. As we were analyzing these proteins, we analyzed that the protein tyrosine protein phosphatase (SPG_0315) may act as drug target and hence to understand the role of this receptor on the basis of 3D structure in disease pathway we have performed homology modeling. Based on the docking studies we came to a conclusion that the analog molecules having higher dock score can be considered as leads and the receptor as their target.

4. CONCLUSION

The outcome of docking studies conclude that 5-methyl-N-[4-(trifluoromethyl)phenyl]-1,2-oxazole-4-carboxamide is the finest ligand for Tyrosine protein phosphatase wzh (SPG_0315) with the Dock score of 39.76 with 1 Hydrogen bond. ADMET descriptors were also analyzed for the drug candidates. Hence, this protein can be considered as the drug target and the above analyzed ligand having highest dock score may be considered as the drug candidate.

5. ACKNOWLEDGEMENT

The bioinformatics computational facilities available at Department of Bioinformatics, Sathyabama University, and the support by the management of Sathyabama University (Dr. Marie Johnson & Dr. Mariazeena Johnson, Directors) are greatly acknowledged. The authors also thank the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

- Bender MH, Cartee RT, Yother J, Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*, *J. Bacteriol.*, 185, 2003, 6057.
- Brueggemann AB, Pai R, Crook DW, Beall B, Vaccine escape recombinants emerge after pneumococcal vaccination in the United States, *PLoS. Pathog.*, 3, 2007, 168.
- Dopazo J, Mendoza A, Herrero J, Caldara F, Humbert Y, Friedli L, Guerrier M, Grand-Schenkm E, Gandin C, Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate, *Microb. Drug Resist.*, 7, 2001, 99-125.
- Hava DL, Camilli A, Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors, *Mol. Microbiol.*, 45, 2002, 1389-1406.
- Hicks LA, Harrison LH, Flannery B, Hadler JL, Schaffner W, Incidence of pneumococcal disease due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998–2004. *J. Infect. Dis.* 196, 2007, 1346-1354.
- McCullers JA, Tuomanen EI, Molecular pathogenesis of pneumococcal pneumonia. *Front Biosci.*, 6, 2001, 877-889.
- Yother J, Capsules of *Streptococcus pneumoniae* and Other Bacteria: Paradigms for Polysaccharide Biosynthesis and Regulation. *Annual Review of Microbiology*, 65, 2011, 563-581.