

Original Article

## INSILICO MODELING OF CAPSULAR POLYSACCHARIDE BIOSYNTHESIS PROTEIN AND TYROSINE KINASE OF G54 STRAIN IN *STREPTOCOCCUS PNEUMONIAE* AND THEIR LIGAND IDENTIFICATION

BALASANKAR KARAVADI<sup>1</sup>, M. XAVIER SURESH<sup>2</sup>

<sup>1,2</sup>Department of Bioinformatics, Sathyabama University, Chennai- 600119, INDIA.  
Email: balasankar.sathyabamauniv@gmail.com

Received: 04 Apr 2014 Revised and Accepted: 10 May 2014

### ABSTRACT

**Objective:** Identifying a novel drug target in *Streptococcus pneumoniae* to treat pneumonia as a challenging task in the current era of Post genomics. Hence the ability to predict the drug target by an insilico method is an initiation towards the drug discovery pipeline.

**Methods:** Strain G54 of *Streptococcus pneumoniae* has cpsC and cpsD genes, which is a part of a tyrosine phosphorylation regulatory system involved in modulation of capsule synthesis proteins, which offers a possible target for drug design. The structures of these proteins were modeled using MODELLER 9v7 and validated using Structural Analysis and Verification Server. Ligands were taken from drug port and its analogues were retrieved from pubchem. These ligands were docked to the modeled proteins using DISCOVERY STUDIO and the ADMET toxicity of the ligands was also studied.

**Results:** N, 3-bis (2-chloroethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine and (4-amino-1-hydroxy-1-phosphonobutyl) phosphonic acid were the two ligands which docked with maximum score with the capsular polysaccharide biosynthesis protein and tyrosine kinase respectively.

**Conclusion:** Based on the computational study on target identification and lead molecular discovery of strain G54 in *Streptococcus pneumoniae*, the above mentioned proteins can be considered as the drug targets and the ligand having maximum dock score may be considered as the drug candidate.

**Keywords:** *Streptococcus Pneumoniae*, G54, Polysaccharide Biosynthesis Protein, Tyrosine Kinase.

### INTRODUCTION

*Streptococcus pneumoniae* is a respiratory pathogen which affects human and cause diseases like pneumonia, bacterial meningitis, sepsis, otitis media etc. Strain G54 is an Italian clinical isolate which is resistant to macrolides and tetracycline of serotype 19F [1]. cpsB, cpsC, and cpsD were the genes which are responsible for a part of a regulatory system containing tyrosine phosphorylation and these genes are also involved in modulation of capsule synthesis in *Streptococcus pneumoniae*[2]. These genes codes capsular polysaccharide biosynthesis proteins. Proteins coded by these genes are tyrosine phosphatase, capsular polysaccharide biosynthesis protein and tyrosine kinase respectively [3]. *Streptococcus pneumoniae* is a human pathogen which causes a wide spectrum of diseases which result in significant global morbidity and mortality [4]. The most recently licensed vaccines against pneumococcal disease are made up of a protein-capsular polysaccharide conjugates. However, these have serious limitations in terms of cost, serotype coverage and increases in the incidence of disease caused by non-vaccine serotypes post introduction [5,6,7]. Multilocus sequence typing has identified this strain as serogroup 15, but the capsular locus is serotype 19F, indicating the strain has probably undergone transformation [1].

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%.

### MATERIALS AND METHODS

Sequence of capsular polysaccharide biosynthesis protein (sp\_0316) and tyrosine kinase (sp\_0317) of G54 strain were retrieved from UniProt. Template for the target proteins were identified on the basis of sequence similarity using PDBsum and cross validated with BLAST search [8,9]. Multiple sequence alignment was performed between the template and target protein using CLUSTAL W. Homology modeling of the target proteins were executed by MODELLER 9v7 [10]. The structural confirmation of the modeled proteins were validated using Structural Analysis and verification

Server on the basis of  $\psi$  and  $\phi$  angles of amino acids in maximum favored regions[11].

In order to obtain stable confirmation, amino acids in partially allowed regions of Ramchandran plot were subjected to energy minimization and certain residues of partial helical nature were subjected to loop refinement using Swiss PDB viewer [12]. Ligands of target proteins were obtained using DrugPort and their corresponding analogs were obtained using PubChem. Binding pockets of the target protein were obtained using CASTp server and the binding energy of protein-ligand complex were obtained using ARGUS lab [13]. Various structural confirmations of the protein-ligand complex were subjected to pose based dock score in Ligand fit module of Discovery Studio under CHARMM force field and the energy function is based on pairwise structural analysis between the nonbonding interactions of protein-ligand complex [14,15]. Finally, ADMET properties of ligands were studied through DISCOVERY STUDIO[16,17].

### RESULTS AND DISCUSSION

#### Homology modeling

Homology modeling was performed for Capsular polysaccharide biosynthesis protein (spg\_0316) and Tyrosine- kinase (spg\_0317) using their respective templates (PDB Id:1EG5 and 3B9Q). The modeled protein was validated through SAVS and the validation results are shown in Table 1. From the results of SAVS, it is found that above 80% residues of target proteins are present in the allowed region of Ramchandran Plot. The predicted structures of modeled protein and their corresponding Ramchandran plots are shown in Fig 1 and Fig 2.

#### Ligand Search

Ligands of Capsular polysaccharide biosynthesis protein (spg\_0316) and Tyrosine- kinase (spg\_0317) were retrieved from DrugPort on the basis of their structural identity with related protein of existing drugs. The best analogs for each ligands were obtained from Pubchem. The docking was performed with those analogs using

Discovery Studio 2.0 software. Dock score was calculated for all the analogs.

**Docking**

Capsular polysaccharide biosynthesis protein (spg\_0316)

Amiloride, Triamterene, Decitabine, Ifosfamide, Procainamide and Flucytosine were found to be the best ligands of Capsular polysaccharide biosynthesis protein (spg\_0316). 3,5-diamino-6-chloro-N-(diaminomethylidene)pyrazine-2-carboxamide, 7-phenylpteridine-2,4,6-triamine, 4-amino-1-[(4S)-4-hydroxy-5-(hydroxymethyl)oxolan-2-yl]-1,3,5-triazin-2-one, N,3-bis(2-chloroethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine, 4-amino-N-(2-dimethylaminoethyl) benzamide and 6-amino-5-fluoro-1H-pyrimidin-3-ium-2-one were found to be the best analogs with a

combination of minimum binding energy and maximum dock score respectively as shown in Table 2 and the interaction between the best analog and Capsular polysaccharide biosynthesis protein (spg\_0316) is shown in Fig 3.

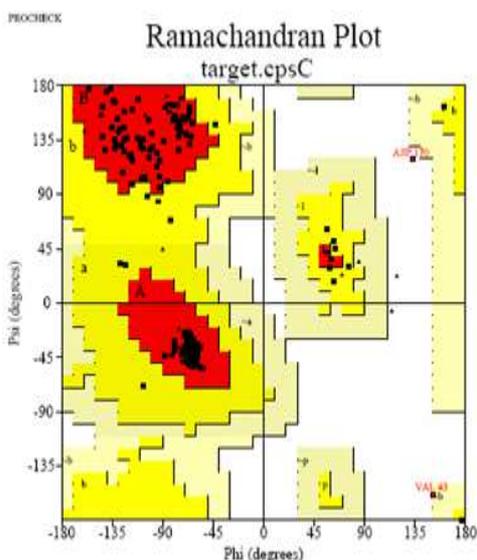
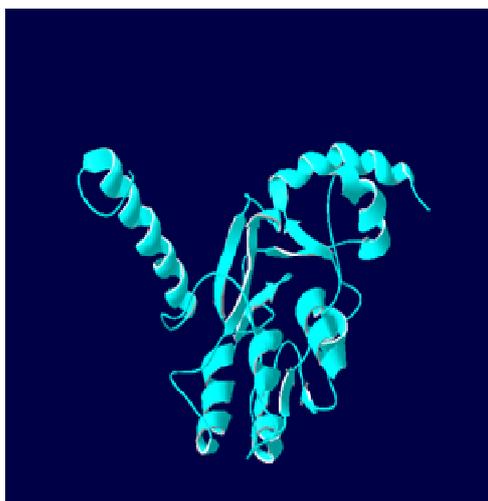
Tyrosine- kinase (spg\_0317)

Alendronate and Choline were found to be the best ligands of Tyrosine- kinase (spg\_0317).

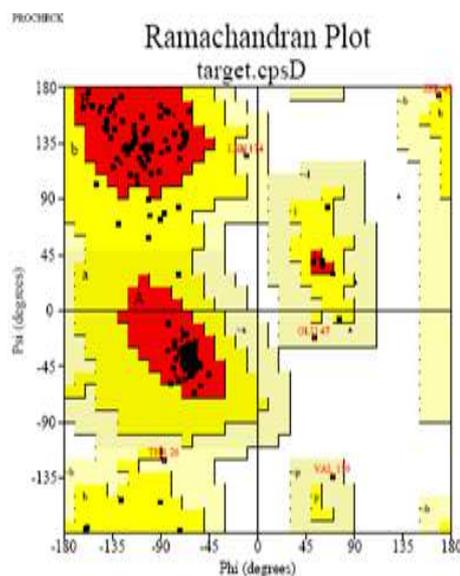
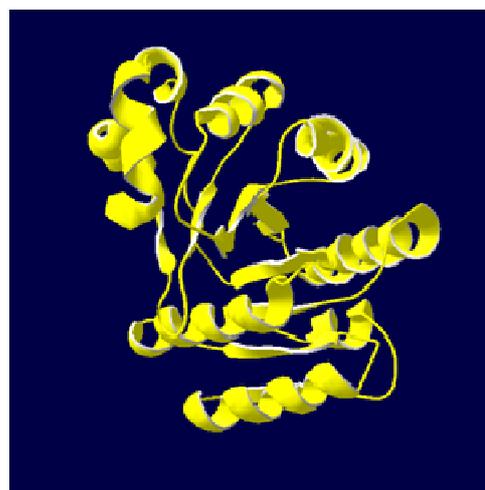
(4-amino-1-hydroxy-1-phosphonobutyl)phosphonic acid and 2-hydroxyethyl(trimethyl)azanium were found to be the best analogs with a combination of minimum binding energy and maximum dock score respectively as shown in Table 3 and the interaction between the best analog and Tyrosine- kinase (spg\_0317) is shown in Fig 4.

**Table 1: 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	Target Length	Template	Template Length	Description of Template	Ramachandran Plot (%)
Capsular polysaccharide biosynthesis protein spg_0316	230	1EG5	364	Thermotoga maritima	85.8%
Tyrosine- kinase spg_0317	227	3B9Q	298	Arabidopsis thaliana	88.6%



**Fig. 1: Structure of Capsular polysaccharide biosynthesis protein (spg\_0316) and its Ramachandran plot.**



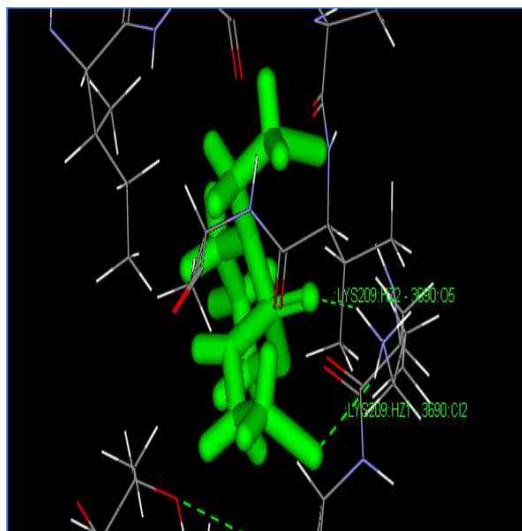
**Fig. 2: Structure of Tyrosine- kinase (spg\_0317) and its Ramachandran plot.**

**Table 2: Non Bonding Interactions for the Analogs with Capsular polysaccharide biosynthesis protein (spg\_0316)**

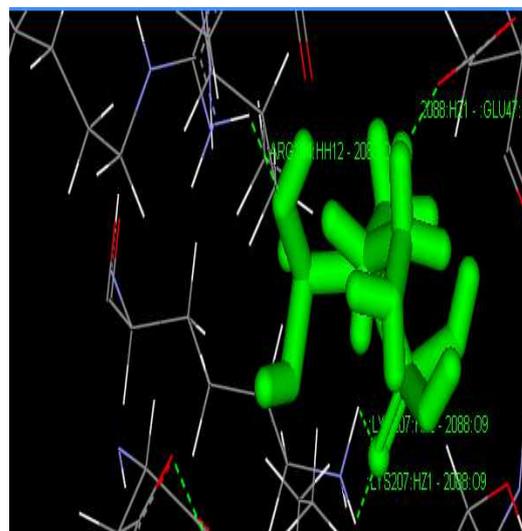
Ligands and analogs		Binding energy	No of hydrogen bonds	Amino acid	Dock score
Triamterene	3,5-diamino-6-chloro-N-(diaminomethylidene)pyrazine-2-carboxamide	-6.89195	1	Thr214	30.705
	(3,5-diamino-6-chloropyrazine-2-carbonyl)-(diaminomethylidene)azanum	-6.79646	4	Gly190 Val145 Thr214 Thr218	28.426
	3,5-diamino-6-chloropyrazine-2-carboxamide	-7.50124	2	Lys209 Thr214	28.004
	6-phenylpteridine-2,4,7-triamine	-6.94288	2	Gly190 Thr214	38.112
	6-(2,3,4,5,6-pentadeuteriophenyl)pteridine-2,4,7-triamine	-6.96206	---	---	---
	7-phenylpteridine-2,4,6-triamine	-6.74708	1	Thr214	40.908
Decitabine	4-amino-1-[(2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)oxolan-2-yl]-1,3,5-triazin-2-one	-6.49535	1	Lys209	36.692
	4-amino-1-[(4S)-4-hydroxy-5-(hydroxymethyl)oxolan-2-yl]-1,3,5-triazin-2-one	-6.69364	4	Val45 Lys209(3)	40.289
	1-[(2R,4S,5R)-4-deuteriooxy-5-(deuteriooxymethyl)oxolan-2-yl]-4-(dideuterioamino)-1,3,5-triazin-2-one	-6.5139	---	---	---
Ifosfamide	N,3-bis(2-chloroethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine	-7.1222	2	Lys209(2)	49.584
	N,3-bis(2-chloroethyl)-4,4,5,5-tetradeuterio-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine	-6.76513	---	---	---
Procainamide	N,3-bis(2-chloro-1,1-dideuterioethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine	-7.1222	---	---	---
	4-amino-N-(2-diethylaminoethyl)benzamide	-7.52498	1	Val45	35.966
	2-[(4-aminobenzoyl)amino]ethyl-dimethylazanium	-6.44734	1	Val145	32.148
	4-amino-N-(2-dimethylaminoethyl)benzamide	-6.86814	1	Val145	38.144
	6-amino-5-fluoro-1H-pyrimidin-2-one	-7.52498	---	---	---
	6-amino-5-fluoranyl-1H-pyrimidin-2-one	-5.44144	1	Thr49	14.778
Flucytosine	6-amino-5-fluoro-1H-pyrimidin-3-ium-2-one	-5.31168	1	Arg207	35.018

**Table 3: Non Bonding Interactions for the Analogs with Tyrosine- kinase (spg\_0317)**

Ligands and analogs		Binding energy	No of hydrogen bonds	Amino acid	Dock score
Alendronate	(4-amino-1-hydroxy-1-phosphonobutyl)phosphonic acid	-7.44909	4	Glu47 arg184 Lys207 (2)	47.46
	deuteriooxy-[1-deuteriooxy-1-[deuteriooxy(hydroxy)phosphoryl]-4-(dideuterioamino)butyl]phosphonic acid	-7.3486	---	---	---
	[4-amino-1-deuteriooxy-1-[deuteriooxy(hydroxy)phosphoryl]butyl]-deuteriooxyphosphonic acid	-6.92454	---	---	---
	2-hydroxyethyl-tri(methyl)azanium	-5.08702	1	Lys207	15.52
Choline	2-hydroxyethyl(trimethyl)azanium	-5.09487	1	Glu47	24.949
	trimethyl-(1,1,2,2-tetradeuterio-2-hydroxyethyl)azanium	-5.10667	---	---	---



**Fig. 3: The protein-ligand interactions of N, 3-bis (2-chloroethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine with Capsular polysaccharide biosynthesis protein (spg\_0316).**



**Fig. 4: The protein-ligand interactions of (4-amino-1-hydroxy-1-phosphonobutyl)phosphonic acid with Tyrosine- kinase (spg\_0317).**

The analog compounds with docking score more than 30.0 were considered to be the best for which ADMET studies were performed. ADMET properties for the analogs of ligands having better dock score and maximum interaction with the active site 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 values as shown Fig 5.

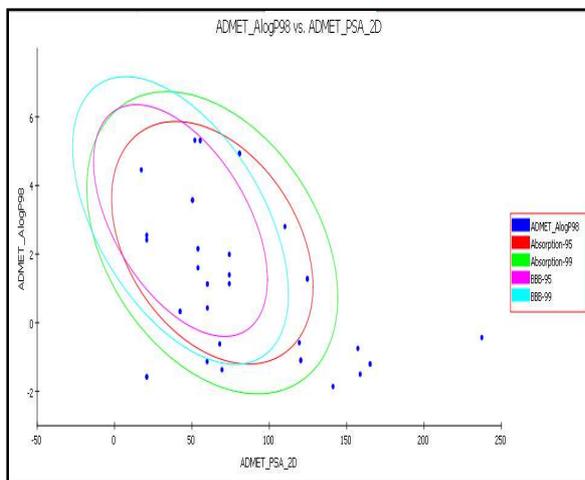


Fig. 5: ADMET properties for the analogs of ligands

## CONCLUSION

The results conclude based on docking studies that N, 3-bis (2-chloroethyl)-2-oxo-1,3,2{5}-oxazaphosphinan-2-amine and (4-amino-1-hydroxy-1-phosphonobutyl)phosphonic acid are the best ligand for Capsular polysaccharide biosynthesis protein (spg\_0316) and Tyrosine- kinase (spg\_0317) with the Dock score of 49.584 and 47.46 with 2 and 4 Hydrogen bond respectively. ADMET descriptors were also analyzed for the drug candidates. Hence, this protein can be considered as the drug targets and the above mentioned ligand having highest dock score may be considered as the drug candidate.

## 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

- Dopazo J, Mendoza A, Herrero J, Caldara F, Humbert Y, Friedli L. et al. Annotated draft genomic sequence from a Streptococcus pneumoniae type 19F clinical isolate. *Microb. Drug Resist.* 2001; 7: 99-125.
- Bender MH, Cartee RT, Yother J. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *J. Bacteriol.* 2003; 185(20):6057-66.
- Yother J. Capsules of *Streptococcus pneumoniae* and Other Bacteria: Paradigms for Polysaccharide Biosynthesis and Regulation. *Annual Review of Microbiology.* 2011; 65: 563-581.
- McCullers J.A, Tuomanen E.I. Molecular pathogenesis of pneumococcal pneumonia. *Front Biosci.* 2001; 6: D877-889.
- Brueggemann AB, Pai R, Crook DW, Beall B. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLoS. Pathog.* 2007; 3(11): e168.
- Hicks LA, Harrison LH, Flannery B, Hadler JL, Schaffner W, Craig AS et al. Incidence of pneumococcal disease due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination. *J. Infect. Dis.* 2007; 196; 1346-1354.
- Altschul S, Gish W, Miller W, Myers E, Lipman D. Basic local alignment search too. *Journal of Molecular Biology.* 1990; 215 (3): 403-410.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25(17):3389-402
- Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* 2002; 45: 1389-1406.
- Mondal UK, Sen A, Bothra AK. Homology modeling of the Cytolethal distending toxin B gene of *Helicobacter hepaticus* ATCC 51449. *International Journal of Interrogative Biology* 2010; 10:35-40.
- 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 (13): 283-291.
- Fiser A, Do RK, Sali A. Modeling of loops in protein structures. *Protein Science.* 2000; 9: 1753-1773.
- Liang J, Edelsbrunner H, Woodward C. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. *Protein Science.* 1998; 7: 1884-1897.
- Evensen E, Joseph-McCarthy D, Weiss GA, Schreiber SL, Karplus M. Ligand design by a combinatorial approach based on modeling and experiment: application to HLA-DR4. *J. Comput. Aided Mol. Des.* 2007; 21: 395-418.
- Zheng CH, Zhou YJ, Zhu J, Ji HT, Chen J, Li YW. et al. Construction of a three-dimensional pharmacophore for Bcl-2 inhibitors by flexible docking and the multiple copy simultaneous search method. *Bio org. Med. Chem.* 2007; 15: 6407-6417.
- Borodina Y, Rudik A, Filimonov D, Kharchevnikova N, Dmitriev A, Blinova V. et al. A new statistical approach to predicting aromatic hydroxylation sites, Comparison with model-based approaches. *J. Chem. Inf. Comput. Sci.* 2004; 44: 1998-2009.
- Venkatesh R, Moudgal CJ, Bruce RM. Assessment of the oral rat chronic lowest observed adverse effect level model in TOPKAT, a QSAR software package for toxicity prediction. *J. Chem. Inf. Comput. Sci.* 2004; 44: 1623-1629