BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, August 2014. Vol. **11**(2), 517-522

Homology Modeling and Molecular Drug Design Approach in Identifying Drug Targets of TIGR4 in *Streptococcus pneumonia*

Balasankar Karavadi and M. Xavier Suresh

Department of Bioinformatics, Sathyabama University, Chennai - 600119, India.

doi: http://dx.doi.org/10.13005/bbra/1301

(Received: 10 March 2014; accepted: 25 April 2014)

Streptococcus pneumoniae **is a respiratory pathogen which is responsible for causing various invasive diseases in humans. TIGR4 is a highly virulent strain under capsular serotype 4. It contains hypothetical proteins synthesized by various coding genes .The efficient degradation of host proteins is an integral aspect of pneumococcal Virulence. In this article we focus on the computational modeling of virulent proteins (Sp_0372, Sp_0192 and Sp_0311) and validating the nature of the proteins as a future drug target of TIGR4 in *Streptococcus pneumonia*. We have also focused on identifying specific ligands for the above mentioned proteins by using the statistical method of high throughput screening of lead molecules on the basis of structure activity relationship. Finally the lead molecules were validated using ADMET descriptors.**

Key words: *Pneumonia, Streptococcus pneumonia,* TIGR4, hypothetical proteins.

Pneumonia is an inflammation in lung which is often caused by an infection with bacteria, viruses or other pathogens, *Streptococcus pneumoniae* is one among the most significant microbe to cause bacterial disease in humans[1,2]. *S.pneumoniae* plays a vital role in integrating the pathogenic genetic material in humans. During the period of 2001-2003, scientists found that *Streptococcus pneumoniae* were resistant to usual drug treatments and their samples had come from children attending 13 day care centers in the city of Lisbon. General vaccination with the 7-valent pneumococcal conjugate vaccine was recommended in Germany during July 2006 for children with more than 2 years of age[3,4]. It is estimated that more than 1 million people die each year from pneumococcal infections worldwide[5,6]. In the United States and elsewhere, resistance to a range of antibiotics is increasing among clinical isolates of *S.Pneumoniae*[7,8]. As part of its life cycle, pneumococcus exists as a commensal bacterium that inhabits and colonizes the nasopharynx of 50% adults and children[9, 10]. The transition from commensal bacterium to opportunistic pathogen often occurs after another infection in respiratory tract eg., pneumococcal pneumonia had been a leading secondary infection which is responsible for causing of death in humans during the pandemics of influenza. Till date only few strains of *S.pneumoniae* contain their complete genome sequence[11]. Strains TIGR4 of *Streptococcus pneumonia* is a clinical isolate which was obtained from the blood sample of 30 year old male patient in Kongsvinger, Norway. TIGR 4 was subjected various genetic tests and it has been found that the

* To whom all correspondence should be addressed.
E-mail: balasankar.sathyabamauniv@gmail.com
xaviersuresh@gmail.com

strain is a virulent capsular strain of serotype 4. Genome of TIGR4 contains a sequence length of 2.34 Mb which includes 2106 proteins and 2302 genes with a GC content of 39.7%.

MATERIALS AND METHODS

Sequences of hypothetical protein (Sp_0372, Sp_0192 and Sp_0311) in TIGR4 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[12, 13]. 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[14].The structural confirmation of the modeled proteins were validated using Structural Analysis and verification Server on the basis of ø and ö angles of amino acids in maximum favored regions[15].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[16]. 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[17]. 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[18,19]. Finally, ADMET properties of ligands were studied through Discovery Studio[20, 21].

RESULTS AND DISCUSSION

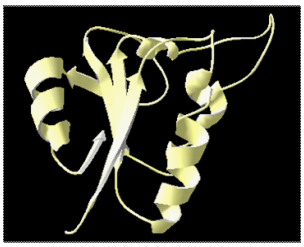
2D1P(B), 2GUY(A), 2RBD(A) were found to be the suitable template for modeling the hypothetical proteins (Sp_0372, Sp_0192 and Sp_0311) with a sequence identity of 35.9%, 40%, 41% respectively(Table 1).The modeled proteins has about 86% of residues in the most favored region of the Ramachandran Plot. Ligands were searched for the above mentioned protein and the best analogs are chosen from pubchem. Analogs of Famciclovir , Miconazole and Riboflavin molecule were docked with hypothetical proteins (Sp_0372, Sp_0192 and Sp_0311) respectively and the dock score is calculated by Discovery Studio 2.0 software suite. Among various analogs best 2 hig scoring compounds were considered to be the drug candidates. After docking the analogs were further subjected to ADMET analysis. In ADMET analysis, comparative graph of plot of polar surface area (PSA) vs logP suggest that analog has optimal concentration of absorption in the blood brain barrier. Homology Modeling was performed for hypothetical proteins (Sp_0372, Sp_0192 and Sp_0311) using the respective templates structures 2D1P (B), 2GUY(A), 2RBD(A) and finally modeled proteins in Figure 1 were validated

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	Sequence length	Template	Description of template	Length	Similarity	Ramachandran Plot
Sp_0372	109	2D1P(B)	Crystal structure of heterohexamerictusbcd proteins, which are crucial for the tRNA modification.	119	35.9%	88%
Sp_0192	88	2GUY(A)	Monoclinic crystal form of *Aspergillus niger* alpha-amylasein complex with maltose.	476	40%	88.3%
Sp_0311	58	2RBD(A)	Crystal structure of a domain with unknown function and a ferritin-like fold from *Bacillus*.	159	41%	86.5%

Table 2. The top scoring ligands obtained from drug port for each protein with their best analogs, interacting residues, Dock score and Binding energy

Target protein	Drug name	Ligand identity	Best Analog (Selected on the basis of Dock Score)	Interacting Residues	Dock score
Sp_0372	Famciclovir	35.7%	Carbonic acid 4-(2-amino-purin-9-yl)-2-hydromethyl-butylester	Arg20(2)	31.329
			4-(2-acetamidopurin-7-yl)butyl acetate(c13h17n503)	Ser 80(2)	28.691
Sp_0192	Miconazole	42.3%	1-[2-(2,4-dichlorophenyl)-2-[(2,4-dichlorophenyl)methoxy]ethyl]-3-[(E)-3-phenylprop-2-enyl]imidazol-3-ium	SER80, 81(2)	64.704
			1-[4-(4-chlorophenyl)-2-[(2,4-dichlorophenyl) methoxy]butyl]imidazole	Met1	52.615
Sp_0311	Riboflavin	40.9%	7,8-dimethyl-10-(3,4,5-trihydroxypentyl) benzo[g]pteridine-2,4-dione	Arg20	22.762
			6,7-dimethyl-10-[(2S,3S,4R)-2,3,4,5-tetrahydroxypentyl]benzo[g]pteridine-2,4-dione	Leu48	14.949



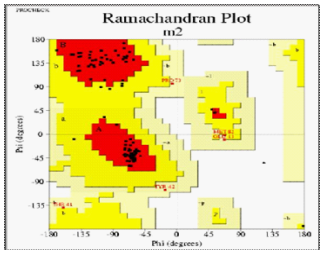


Fig. 1(a). Structure of modeled hypothetical protein (Sp_0372) and its Ramachandran plot

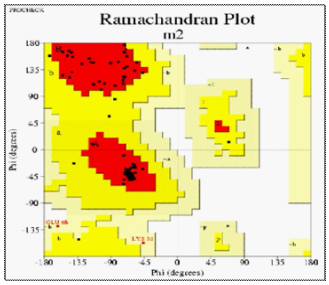
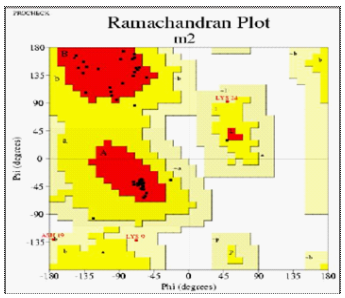
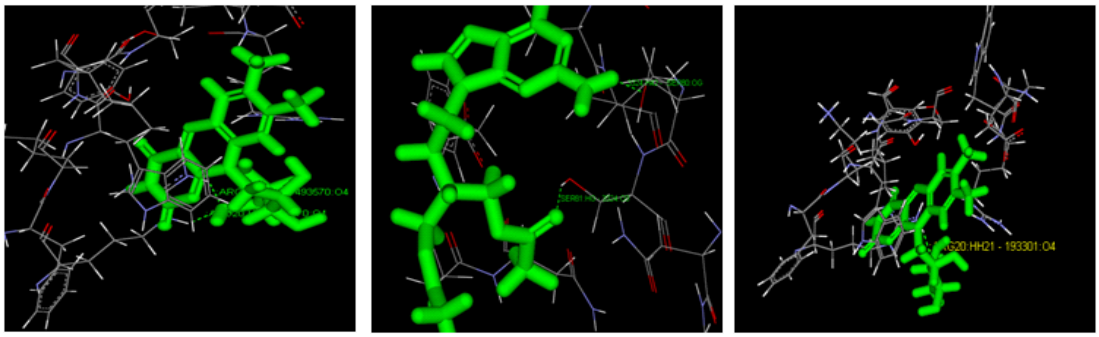


Fig. 1(b). Structure of modeled hypothetical protein (Sp_0192) and its Ramachandran plot.



1(c). Structure of modeled hypothetical protein (Sp_0311) and its Ramachandran plot

Fig. 1. The final modeled protein structures with their Ramachandran plots



(a) Docking of Carbonic acid 4-(2-amino-purin-9-yl)-2-hydromethyl-butylester with Sp_0372
(b) Docking of 1-[2-(2,4-dichlorophenyl)- 2-[(2,4-dichlorophenyl)methoxy]ethyl]-3-[(E)-3-phenylprop-2-enyl]imidazol-3-ium with Sp_0192 .
(c) Docking of 7, 8-dimethyl-10-(3,4,5-trihydroxypentyl) benzo[g]pteridine-2,4-dione with Sp_0311

Fig. 2. Docking results of best analogs with their respective protein

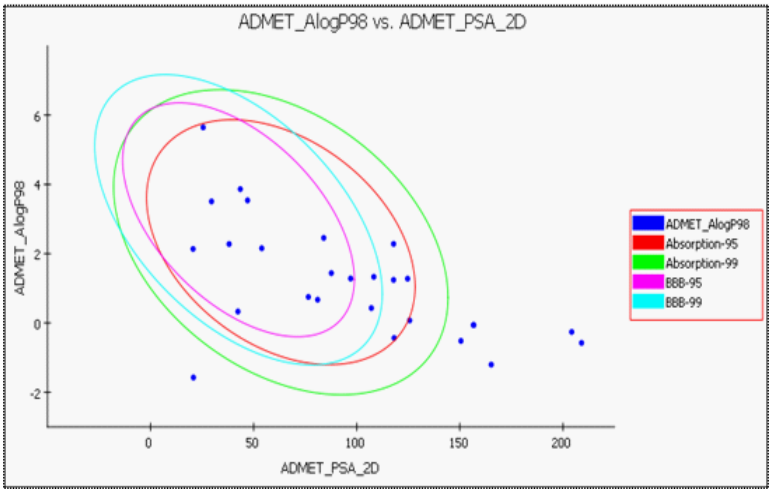


Fig. 3. ADMET plot of the analog compounds

through SAVS and 86% residues in target proteins are present in the allowed region of Ramachandran Plot.

Ligand search

Ligands for hypothetical proteins (Sp_0372, Sp_0192 and Sp_0311) were retrieved from DrugPort sharing more than 40% identity with related protein sequence for which already a drug exists. The analogs for those ligands were obtained from PubChem and for each protein 10 best analogs were chosen from the hit. The docking was performed with those analogs using Discover studio 2.0 software. Docking score was calculated for all the analogs with their respective proteins and the affinity was more when the dock score is high.

Docking

Famciclovir was the best ligand for the protein (Sp_0372) which had 35.7% of identity with target protein and the Carbonic acid 4-(2-amino-purin-9-yl)-2-hydromethyl-butylester was the analog based on the dock score of 31.329.

Miconazole was the best ligand for the protein (Sp_0192) which had 42.3%of identity with target protein and the 1-[2-(2,4-dichlorophenyl)-2-[(2,4-dichlorophenyl)methoxy]ethyl]-3-[(E)-3-phenylprop-2-enyl]imidazol-3-ium was the analog based on the dock score of 64.704.

Riboflavin was the best ligand for the protein (Sp_0311) which had 40.9%of identity with target protein and the 7,8-dimethyl-10-(3,4,5-

trihydroxypentyl) benzo[g]pteridine-2,4-dione was the analog based on the dock score of 22.762.The results are shown in Table 2 and their interactions are shown in Figure 2. 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 lopP, Absorption and Blood Brain Barrier values are shown in figure 3.

CONCLUSIONS

Based on docking studies it has been concluded that Carbonic acid 4-(2-amino-purin-9-yl)-2-hydromethyl-butylester, 1-[2-(2,4-dichloro phenyl)- 2-[(2,4-dichlorophenyl)methoxy] ethyl]-3-[(E)-3-phenylprop-2-enyl]imidazol-3-ium,7,8-dimethyl-10-(3,4,5-trihydroxypentyl) benzo[g] pteridine- 2,4-dione were the three ligands interacting analogs with maximum dock score with the proteins Sp_0372, Sp_0192 and Sp_0311 respectively and ADMET descriptors were also analyzed for the drug candidates. Hence, these proteins can be considered as the drug targets and the above mentioned ligands having higher dock score may be considered as the drug candidates.

REFERENCES

1. Avery,O.T., MacLeod,C.M., McCarty,M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J.Exp. Med.*, 1944; **79**:137-158.
2. Klein,D.L. Pneumococcal disease and the role of conjugate vaccines. *Microb. Drug Resist.*, 1999; **5**:147-157.
3. Blue, C. E. , Paterson, G.K., Kerrm, A.R., Berge, M., Claverys, J.P. , Mitchell, T.J. ZmpB, a novel virulence factor of Streptococcus pneumoniae that induces tumor necrosis factor alpha production in the respiratory tract. *Infect. Immun.*, 2003; **71**: 4925-35.
4. LeMessurier, K.S., Ogunniyi, A.D., Paton, J.C. Differential expression of key pneumococcal virulence genes in vivo. *Microbiology*, 2006; **152**: 305-11.
5. Orihuela, C.J, Gao, G., McGee, M., Yu, J., Francis, K.P., Tuomanes, E. Tissue-specific contribution of pneumococcal virulence factors to pathogenesis. *J. Infect. Dis.*, 2004; **190**:1661-1669.
6. Orihuela, C.J, Gao, G., McGee, M., Yu, J., Francis, K.P., Tuomanes, E. Organ-specific models of Streptococcus pneumoniae disease. *Scand. J. Infect.*, 2003; **35**: 647-652.
7. Robertson, G.T., Ng, W.L. Foley, J., Gilmour, R., Winkler, M.E. Global transcriptional analysis of clpP mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence. *J. Bacteriol.*, 2002; **184**: 3508-3520.
8. Blue, C.E., Mitchell, T.J. Contribution of a response regulator to the virulence of Streptococcus pnuensoniae is strain dependent. *Infect. Immun.*, 2003; **71**: 4405-4413.
9. Cardoso, T.C., Lopes, L.M. Carneiro AH: A case-control study on risk factors for early-onset respiratory tract infection in patients admitted in ICU. *BMC Pulm. Med.*, 2007; **7**:12-17.
10. Del Toro, M.D., Rodriguez-Bano, J., Martinez-Matinez, L., Pasculal, A., Perez-Canoa, R., Perea, E.J., Muniain, M.A. Epidemiology, clinical features and prognosis of infections due to Stenotrophomonas lmaltophilia. *Clin. Microbiol.*, 2006; **24**: 4-9.
11. Kuo, L.C., Lai, C.C., Liao, C.H., Hsu, C.K., Chang, Y.L., Change C.Y., Hsueh, R.R. Multidrug-resistant Acinetobacter baumannii bacteremia: clinical feature, antimicrobial therapy and outcome. *Clin. Microbiol. Infect.*, 2007; **13**: 196-198.
12. Altschul, S.F., Madden, T.L., Schäffer, A.A., 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):3389-3402.
13. Hava, D.L., Camilli, A., Large-scale identification of serotype 4 Streptococcus pneumoniae virulence factors. *Mol. Microbiol.*, 2002; 45: 1389-1406.
14. Mondal, U.K., Sen, A., Bothra, A.K. Homology modeling of the Cytolethaldistending toxin B gene of Helicobacter hepaticus ATCC 51449. *International Journal of Integrative Biology*, 2010; **10**:35-40.
15. Laskoswki, R.A., MacArthur, M.W., Moss, D.S., Thorton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 1993; **26**(13): 283-291.
16. Fiser, A., Do, R.K., Sali, A. Modeling of loops in protein structures. *Protein Science*, 2000; **9**: 1753-1773.
17. 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.
18. Evensen, E., Joseph-McCarthy, D., Weiss, G.A., Schreiber, S.L., 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.

19. Zheng, C.H., Zhou, Y.J., Zhu, J., Ji, H.T., Chen, J., Li, Y.W. 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.
20. 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.
21. Venkatapathy, R., Moudgal, C.J., Bruce, R.M. 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.