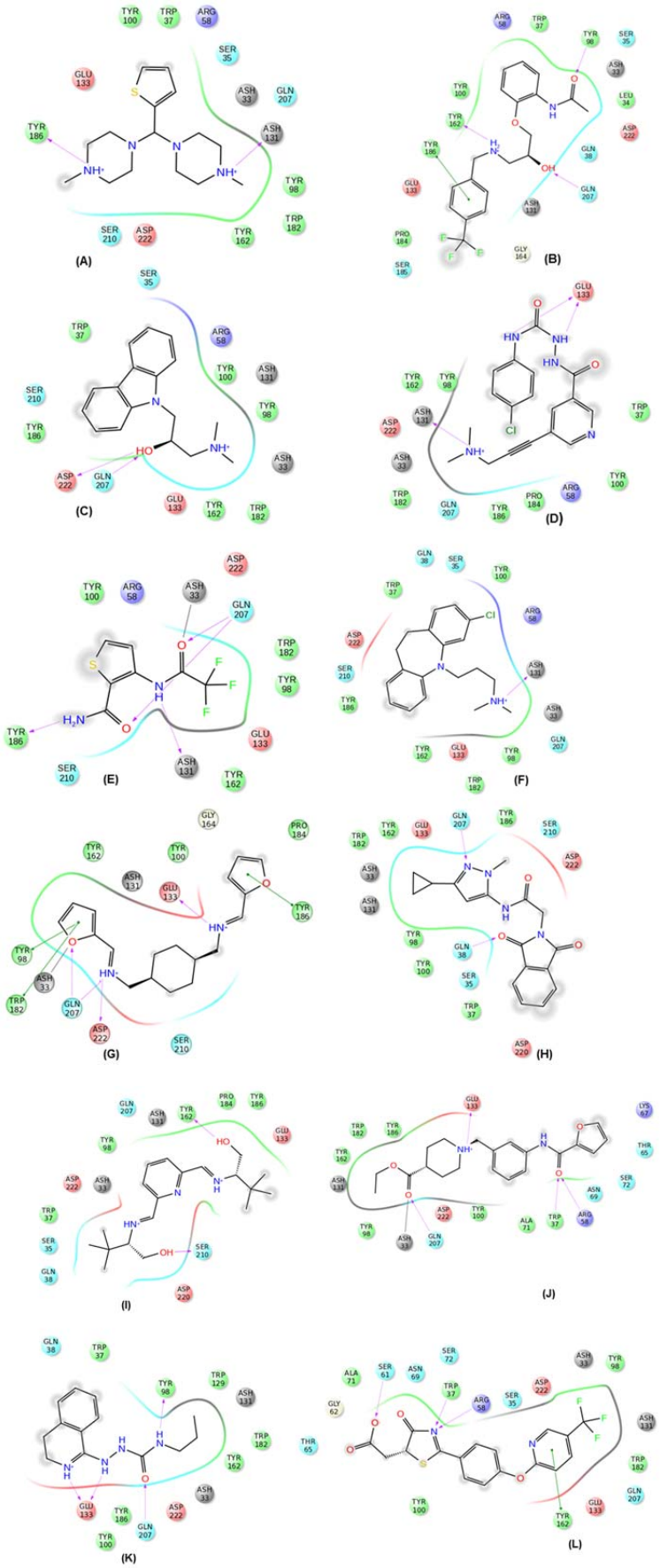


Fig. 4. Representation of the interactions between the ligand molecules and the receptor. The hydrogen bond interactions are shown in pink colored arrows and π-π interactions are shown in green colored arrows. (A) PD00612 (B) SEW02675 (C) BTB12226 (D) KM08436 (E) GK00487 (F) RJC01223 (G) S01517 (H) SCR00967 (I) JFD02837 (J) SPB08437 (K) MWP01096 (L) HTS04925.

### 3.2. Free energy calculation of the docked complexes using Prime MM-GBSA

The binding free energies of the protein-ligand complexes were estimated to further consolidate the ability of the identified ligands to bind with the target protein. Along with the binding site residues, the water molecules present in the binding site also play a crucial role in protein-ligand recognition. This recognition could either be by displacement upon ligand binding or by forming water bridges and thereby stabilizing the complex [15]. However, it is difficult to rigorously treat explicit binding-site waters, which requires to completely sample ensembles of water molecules and to consider the free energy cost of replacing waters. The MM-GBSA approach is a computationally efficient method, which employs molecular mechanics, generalized Born model and solvent accessibility method to elicit free energy [16].

The observed binding free energy values of all the 12 complexes are given in Table 4, along with the ligand and receptor strain energy. It could be seen that, apart from the compound **GK00487** (binding free energy: -28.13 kcal/mol), all the other complexes had nearly equal binding free energies, suggesting equal binding affinities. The ligand and receptor strain energies were also found to be lower and equivalent except for the compounds **SEW02675** and **JFD02837**. However, their corresponding observed binding free energy of these compounds were found to be -54.9 and -58.6 kcal/mol, respectively, suggesting compensation by other favourable energy terms. It should also be noted that the compound **KM08436** which was ranked fourth according to glide XP score was predicted to have the highest binding affinity among the 12 compounds with a binding free energy value of -65.9 kcal/mol.

### 3.3. ADME Screening

The drug like abilities of the identified inhibitors were further emphasized by analyzing their ADME properties which are presented in Table 5. Further description and discussion of the obtained results are given below.

#### 3.3.1. Comparison of logP (o/w)

Partition coefficient (log P) is used to predict the hydrophobicity and hydrophilicity of a drug in the body, where (o/w) represents octanol/water [17]. For an ideal drug, the logP value should be within a range of -2.0 to 6.5. If a compound has high logP value, it refers to high hydrophobicity and if it is less, the compound is highly water soluble. A compound should have both hydrophobic and hydrophilic properties in equal proportion in order to reach the target site. Here, all the obtained ligands are within a range of -0.1 and 4.6 and thus have equal probability of reaching the target site.

#### 3.3.2. Comparison of log HERG

HERG refers to Human Ether-à-go-go-Related Gene which codes for potassium ion channel. This is best known for its contribution to the electrical activity of the heart. The log HERG value gives the predicted $IC_{50}$ value for blockage of HERG K+ channels. There is a risk of sudden death when this channel's ability to conduct electrical current across the cell membrane is inhibited or compromised [18]. Generally, it is a concern if the value is lesser than -5. The values obtained for the identified 12 compounds were in the range of -3.6 to -7.1. Additionally, it should also be noted that the compound **RJC01223** which is an already FDI approved drug and available in the market has a value of -6.1.

#### 3.3.3. Comparison of Caco and Oral Absorption Capability

Caco-2 cells are human epithelial colorectal adeno carcinoma cells. Pharmaceutical industries use Caco-2 monolayers as an in vitro model of the human small intestinal mucosa to predict the absorption of orally administered drugs. They are a model for the gut-blood barrier. These predictions are only for non-active transport. It is considered that values of Caco-2 cell permeability below 25 nm/sec are poor and values above 100 are better [19]. It could be noted from the Table 5 that all the lead molecules have a good oral absorption rate and in particular, the compounds **BTB12226**, **RJC01223**, **S01517** and **JFD02837** have 100% oral absorption rate.

#### 3.3.4. Comparison of MDCK

MDCK stands for Madin-Darby Canine Kidney Cells. It helps to gain a greater understanding of the mechanism of drug efflux and highlights early potential issues with drug permeability. MDCK cells are considered to be a good mimic for the blood-brain barrier [20]. A value lesser than 25 nm/sec is considered to be poor and values greater than 500 is considered to be very good. It could be noted that all the 12 compounds obtained by our screening procedure had admissible values.

Finally, it has to be stated that all the identified 12 compounds fits well within the Lipinski's rule of five, which states that an orally active drug has no more than one violation of the rules such as not having more than 5 hydrogen bond donors, not having more than 10 hydrogen bond acceptors, having a molecular mass less than 500 Daltons and octanol-water partition coefficient (log P) being not greater than 5 [10]. This substantiates that the obtained lead compounds by our screening study have potential pharmacological properties and could be used for further experimental phases in drug discovery.

## 4. CONCLUSION

In this study, a high throughput structure based virtual screening approach was used to find potential inhibitors against GHIP a virulence factor of *S.pneumonia*, which is a Glycosyl Hydrolase 25 Related Invasion Protein which is involved in host cell invasion. Twelve potential lead compounds were identified by the screening procedure. The favourable ADME properties confirm the drug likeness of the identified compounds. Estimation of binding free energies of the protein-ligand complexes using Prime MM-GBSA calculation showed that the identified compounds have similar binding affinity towards the receptor. The conserved Asp 131 and Glu 133 residues located at the active site made consistent hydrogen bonding interactions with the ligands. One of the lead compounds (**RJC01223** - Clompiramine) identified in the present study was also found to be a potential drug candidate against *Salmonella typhi*, *Plasmodium falciparum*, Marburg Virus, Human colon cancer and prostate cancer.

## ACKNOWLEDGMENTS

We thank SASTRA University for infrastructural and financial support (TRR Fund for ST). Financial support from Department of AYUSH (Z.15075/1/2010-COE), Govt of India for purchase of Schrödinger suite is gratefully acknowledged.

## REFERENCES

[1] Sternberg, G., A fatal form of septicaemia in the rabbit, produced by the subcutaneous injection of human saliva. *Annual Reports of the National Board of Health* **1881,** *3*, 87-108.

[2] Pasteur, L.; Chamberland, M.; Roux, Sur une maladie nouvelle, pro-voquke par la salive d' un enfant mart de la rage. *Compt Rend Acad d Sci* **1881,** *92*, 159-165.

[3] Verma, R.; Khanna, P., Pneumococcal conjugate vaccine: A newer vaccine available in India. *Human Vaccines & Immunotherapeutics,* **2012,** *8* (9), 1317-1320.

[4] Siemieniuk, R.; Gregson, D.; Gill, M., The persisting burden of invasive pneumococcal disease in HIV patients: an observational cohort study. *BMC Infect Dis,* **2011,** *11* (1), 1-8.

[5] Donkor, E. S.; Dayie, N. T. K. D.; Badoe, E. V., Vaccination against pneumococcus in West Africa: perspectives and prospects. *Int J General Med,* **2013,** *6*, 757-764.

[6] Ladhani, S. N.; Slack, M. P. E.; Andrews, N. J.; Waight, P. A.; Borrow, R.; Miller, E., Invasive Pneumococcal Disease after Routine Pneumococcal Conjugate Vaccination in Children, England and Wales. *Emerging Infectious Diseases,* **2013,** *19* (1), 61-68.

[7] Kries RV, S. A., Schmitt HJ, Rene R Proportion of Invasive Pneumococcal Infections in German Children Preventable by pneumococcal Conjugate Vaccines. *Clinical Infectious Diseases* **2000,** *31*, 482–487.

[8] BG, S.; BM, G., Prevention of pneumococcal disease by vaccination: does serotype replacement matter? *Lancet,* **2000,** *356*, 1210-1211.

[9] Niu, S.; Luo, M.; Tang, J.; Zhou, H.; Zhang, Y.; Min, X.; Cai, X.; Zhang, W.; Xu, W.; Li, D.; Ding, J.; Hu, Y.; Wang, D.; Huang, A.; Yin, Y.; Wang, D., Structural Basis of the Novel <italic>S. pneumoniae</italic> Virulence Factor, GHIP, a Glycosyl Hydrolase 25 Participating in Host-Cell Invasion. *PLoS ONE,* **2013,** *8* (7), e68647.

[10] Lipinski, C. A.; Lombardo F Fau - Dominy, B. W.; Dominy Bw Fau - Feeney, P. J.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev,* **2001,** *46*, 3-26.

[11] Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J Med Chem,* **2006,** *49* (21), 6177-6196.

[12] Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A., A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics,* **2004,** *55* (2), 351-367.

[13] Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B., On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *J Mol Biol,* **2002,** *320* (3), 597-608.

[14] Lyne, P. D.; Lamb Ml Fau - Saeh, J. C.; Saeh, J. C., Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J Med Chem,* **2006,** *49*, 4805-4808.

[15] Sun, H.; Zhao, L.; Peng, S.; Huang, N., Incorporating replacement free energy of binding-site waters in molecular docking. *Proteins: Structure, Function, and Bioinformatics,* **2014,** *82* (9), 1765-1776.

[16] Dill, K. A., Additivity Principles in Biochemistry. *J Biol Chem,* **1997,** *272* (2), 701-704.

[17] Ghose, A.; Crippen, G., Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. *J Chem Inf Comput Sci,* **1987,** *27*, 21-35.

[18] Cavalli, A.; Poluzzi, E.; De Ponti, F.; Recanatini, M., Toward a Pharmacophore for Drugs Inducing the Long QT Syndrome: Insights from a CoMFA Study of HERG K+ Channel Blockers. *J Med Chem,* **2002,** *45* (18), 3844-3853.

[19] Yazdanian, M.; Glynn Sl Fau - Wright, J. L.; Wright Jl Fau - Hawi, A.; Hawi, A., Correlating partitioning and caco-2 cell permeability of structurally diverse small molecular weight compounds. *Pharm Res, 15*, 1490-1494.

[20] Irvine, J. D.; Takahashi L Fau - Lockhart, K.; Lockhart K Fau - Cheong, J.; Cheong J Fau - Tolan, J. W.; Tolan Jw Fau - Selick, H. E.; Selick He Fau - Grove, J. R.; Grove, J. R., MDCK (Madin-Darby canine kidney) cells: A tool for membrane permeability screening. *J Pharm Sci, 88*, 28-33.