Asian Journal of Biomedical and Pharmaceutical Sciences
ISSN : 2249 - 622X

RESEARCH ARTICLE

**Received on: 22-04-14**
**Accepted on: 20-05-2014**
**Published on: 15-06-2014**

**M. Xavier Suresh**
Department of Bioinformatics, Sathyabama University, Chennai – 600 119

**QR Code for Mobile users**

Conflict of Interest: None Declared !

**DOI: 10.15272/ajbps.v4i32.489**

# *In silico* studies on structure prediction and inhibitory action of selected natural compounds on cell invasion protein SipB from Salmonella Typhi

M. Xavier Suresh* and Ophilia B. Pushpa
Department of Bioinformatics, Sathyabama University, Chennai – 600 119.

**Abstract**
Salmonellosis is one of the most common and widely distributed food borne diseases caused by Salmonella enterica serovar Typhi (S.Typhi). Powerful strategies of drug design such as targeting unique effectors of this pathogen are required for combating the emergence of multi drug resistant strains which have become a serious public health problem. The cell invasion protein, SipB has been identified as a potential target which is known to be involved in important functions like host cell entry, transfer of other effector proteins into the host cell, inducing macrophage apoptosis, activating proapoptotic enzyme caspase I for inducing autophagy. This study reports the structure of SipB determined by the combination of homology modelling and comparative modelling techniques. Further 75 natural compounds have been identified as inhibitors and allowed to dock at the binding site, 10 shown to have better dock score and interactions with amino acid residues Glu394, Asn489, Gln517 and Asn562. Further investigations into the antipathogenic potential of caffic acid and phloretin analog compounds may open new avenues for drug development in the control of antibiotic resistant pathogens.

**Keywords:** Salmonella, drug resistance, cell invasion protein, SipB, docking, drug design.

**Cite this article as:**
M. Xavier Suresh* and Ophilia B. Pushpa In silico studies on structure prediction and inhibitory action of selected natural compounds on cell invasion protein SipB from Salmonella Typhi. Asian Journal of Biomedical and Pharmaceutical Sciences; 04 (32); 2014. 20-24.

## Introduction:

Salmonella enterica serovar Typhi (hereafter referred to as S. Typhi) is the aetiological agent of typhoid fever, a serious invasive bacterial disease of human. Salmonellosis is one of the most common and widely distributed food borne diseases. In people at risk such as infants, small children, the elderly, Salmonella infections can become very serious, leading to complications. If these are not treated, HIV patients and those with suppressed immunity can become seriously ill. Children with sickle cell anaemia who are infected with Salmonella may develop osteomyelitis [1-2].

The genus Salmonella is divided by serology into over 2,500 serovars, many S. enterica serovars actively invade the mucosal surface of the intestine and S. Typhi has evolved the ability to spread to the deeper tissues of human, including liver, spleen and bone marrow [3]. Currently, 107 strains of this organism have been isolated, many containing varying metabolic characteristics, levels of virulence, and multi-drug resistance genes. Resistance to all first line antimicrobials- ampicillin, trimethoprim-sulfamethoxazole and chloramphenicol is defined as multi drug resistance (MDR) [4]. S. Typhi Strain CT18 is resistant to multiple drugs which is a serious emerging threat to the treatment of infectious diseases. The emergence of multi drug resistant strains has become a threatening public health problem with the record of affecting 21 million people worldwide and kills 200,000 every year [5-6].

An important event in salmonella infection is the invasion of non-phagocytic intestinal epithelial cells. The pathogen is taken up by macropinocytosis, induced by contact dependent delivery of bacterial proteins that subvert signaling pathways and promote cytoskeletal rearrangement. The salmonella pathogenicity island 1 is required for invasion, it encodes a group of invasion effectors SipA-D, including SipB that upon contact with a target cell undergo 'type III' export from the bacterial cytoplasm and translocate into the cell membrane. This invasion protein (SipB) is proven to induce macrophage apoptosis by binding to caspase-1. Caspase-1 activity is essential for the cytotoxicity, functional inhibition of caspase-1 blocks macrophage cytotoxicity. Therefore, targeting unique effectors of this pathogen can be considered as a powerful strategy for drug design against bacterial variations to drug resistance [7-9].

Strengthening the reliable information is the fastest way in drug discovery process for development and optimization of potential inhibitors to find better drug candidates. Though a number of current drug targets for all classes of approved therapeutic drugs have been proposed, the rate of developing drugs against new target families is significantly lower [10] . The availability of complete genomes and the need for tertiary structures of proteins in drug discovery has initiated several computational approaches [11-13]. In this present study we report the homology modeling methodologies is applied to obtain a reasonable structure which helps to understand the behaviour of SipB, cell invasion protein. Further, computational inhibitor screening and docking strategies were employed to find new potential inhibitors against SipB [14-15].

## Materials and Methods

The amino acid sequence of cell invasion protein SipB from Salmonella typhi was obtained from the protein sequence database of UNIPROT (Accession no: Q56134). FASTA format of the SipB sequence was used for the structure prediction analysis. The instability index and the aliphatic index were predicted using PROTPARAM tool [16]. The secondary structure was predicted using NNPREDICT [17]. In order to determine the location of transmembrane helices, several tools were employed. TM prediction tools such as HMMTOP [18] and SOSUI [19] were used. The protein Blast was carried out against protein data bank (PDB). A combination of homology and comparative modeling methodologies were applied to obtain a reasonable structure for SipB. We exploited I-TASSER server, an automated structure prediction tool based on the sequence to structure prediction paradigm, to predict the tertiary structure of SipB protein. Three dimensional atomic models were generated from multiple threading alignment and iterative structural assembly simulations [20]. Hydrogens were added and the models were subjected to several tests to make sure that it was of reasonable quality for analyzing ligand binding and to check its ability for virtual screening. The stereochemical quality of the models was assessed by PROCHECK which indicate the amino acids with unusual backbone conformation [21]. Similarly the non-bonded interactions between different atom types were calculated by ERRAT [22].

Then the ligand binding sites from the modelled structure were predicted with the help of binding pocket detection server tools such as pocket finder and Q-site finder (www.modelling.leeds.ac.uk/qsitefinder) [23-24]. Additionally, the binding pockets of the receptor were also determined by using Accelrys Discovery Studio 2 (DS). Molecular docking is a method to evaluate the feasible binding geometries of a

putative ligand with a target whose target site is known. The binding geometries is often known as binding poses, includes, in principle, both the position of the ligand relative to the receptor and conformational state of the ligand and the receptor. Through literature search 75 herbal compounds showing anti-salmonella activity have been identified. In addition to this dataset a few known antibiotic compounds and their similar compounds from PubChem were also included. The selected compounds were docked into the binding site of the receptor using Ligand fit protocol implemented in Accelrys Discovery Studio [25]. The docking run generated 10 poses for each of the analog compounds. The ligscore, Jain, PLP and PMF scoring functions were used to identify the best docked pose. Prior to docking CHARMm forcefield is applied to prepare both ligands and proteins [26].

**Results & Discussion**

Sequence analysis and topology prediction:

The primary structure of SipB consists of a polypeptide chain with 593 amino acid residues. Sequence similarity searches across various databases show that SipB is the novel member of the invasion protein family which belongs to the type III secretion system [27]. The protein is predicted as a stable protein with instability index as 23.35 and the aliphatic index as 93.59 using PROTPARAM. Also, analysis of the similarity search showed that the subunit structure of the protein is a homotrimer. Analysis of the secondary structure predicted using NNPREDICT prediction results reveals that it has several alpha helices (around 90%) and few (10%) regions of beta sheets and coils. Protein analysis shows the regions 151- 208 and 287 – 314 are found to be potential coils. Transmembrane segments predicted using several transmembrane region prediction tools were the regions 320-354 and 408-429. Further spectroscopic analysis of these domains in isolation showed that the hydrophobic regions insert obliquely into the bilayer, whereas the C-terminal domain associates with the bilayer surface, tilted parallel to the membrane. The combined data suggest a topological model for membrane-inserted SipB [28].

3-D Structure prediction and model validation:

A reliable tertiary structure of SipB was neither available in PDB nor obtained through homology based methods as there is no template structures with good similarity are available. Thus we exploited I-TASSER server to predict the tertiary structure of SipB protein which is an automated structure prediction tool based on the sequence to structure prediction paradigm. PROCHECK was used to validate the structure, in which Ramachandran plot statistics shows the most favored region, additional allowed region, generously allowed region, and disallowed regions has 79.4%, 16.5%, 2.7% and 1.4% residues respectively. The modelled structure of SipB is shown in Figure 1.

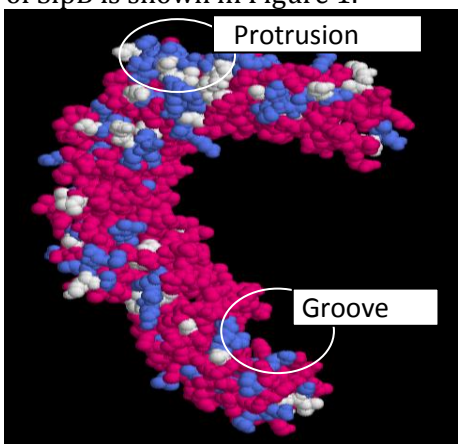


**Figure 1. 3-D structure of SipB model predicted by I-TASSER.**

It is interesting to note that the subunit structure of the protein is a homotrimer. An exhaustive analysis of the protein surface through visual inspection reveals that at one end there is a groove appears whereas at the other end there observed a small protrusion, which indicate that these could be the protein-protein interaction sites with in the protein which can facilitate the oligomerization in to homotrimer. Moreover these groove also has the binding site like properties, inhibitor binding at this site may probably alter the mechanism through which the trimmer formation would have been stopped. This model can serve as a valuable reference for characterizing the protein and could be explored for drug targeting by design of suitable inhibitors to control the disease caused by the pathogen.

**Binding site analysis of SipB:**

The binding sites identified and the residues in the binding site predicted by pocket finder and Q-site finder shows common ligand binding residues which belongs to site 1 where site 1 span over 379-564 whereas site 2 span over 34-197 (Figure 2). In addition to that, the binding pockets of the receptor determined by using Accelrys Discovery Studio 2 also provide similar results. And hence, the site 1 with larger volume is designated as the best binding site and used further in the studies.

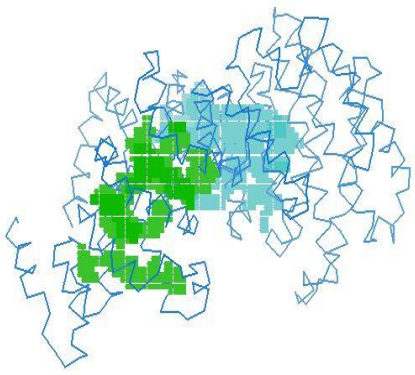

**Figure 2. Predicted ligand binding sites of SipB.**

Molecular Docking of natural compounds as ligands:
Docking predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using scoring functions. We performed the molecular docking using Discovery Studio against the 75 natural compounds and a few known antibiotic compounds and similar compounds from PubChem. The selected compounds were allowed to dock into the binding site of the receptor using Ligand fit protocol implemented in Accelrys Discovery Studio. The docking run generated 10 poses for each of the analog compounds. The ligscore, Jain, PLP and PMF scoring functions were used to identify the best docked pose. The computed dock scores along with the interactions at the binding site for the top ten compounds are tabulated in Table 1. The docked poses of the top 2 compounds and their interacting residues are illustrated in Figure 3.

| S.No | Compound Name | Dock Score | Number of H-Bonds |
|---|---|---|---|
| 1. | Caffic Acid | 56.778 | 02 |
| 2. | Phloretin | 55.572 | 02 |
| 3. | Berberine | 50.428 | 03 |
| 4. | Curcumin | 49.788 | 01 |
| 5. | Hypericin | 48.778 | 02 |
| 6. | Eugenol benzoate | 47.148 | 02 |
| 7. | Eluthrol | 45.754 | 03 |
| 8. | Eugenol oxide | 43.848 | 02 |
| 9. | Colchine | 43.36 | 03 |
| 10. | Eugenol | 42.529 | 02 |

**Table 1. The top 10 herbal compounds along with the dock scores number of hydrogen bonds with the receptor SipB.**

The stability of the docked poses was evaluated by determining the hydrogen bonding between the receptor and ligand. Docking results suggested that compounds caffic acid and phloretin had a good docking score. Caffic acid has highest dock score of 56.778 and had good hydrogen bonds with the amino acid residues Glu394 and Asn489. Docking result showed that phloretin has second highest dock score of 55.572 and two H-bonds were formed between ligand and protein with amino acid residues Gln517 and Asn562. Apart from the hydrogen bonds there are hydrophobic interactions that also contribute for the stability of the compounds at the binding site. The overall binding of caffic acid and phloretin is illustrated in Figure 3 & 4. The docking results indicate that these ligands were amenable for designing new drug candidates.

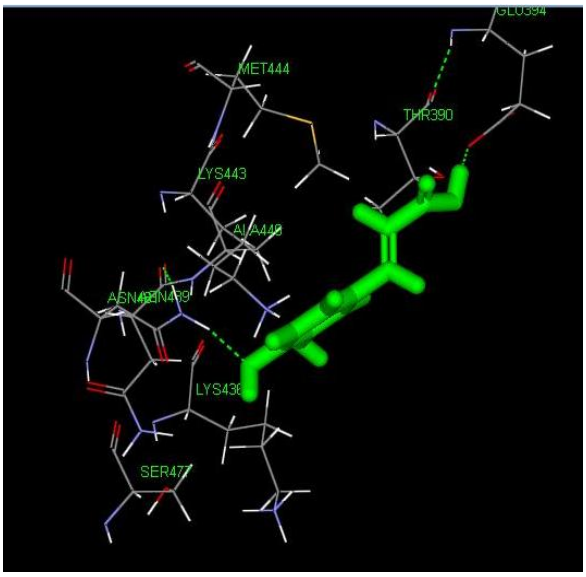


**Figure 3. Binding mode of caffic acid at receptor showing interactions with Glu394 and Asn489.**

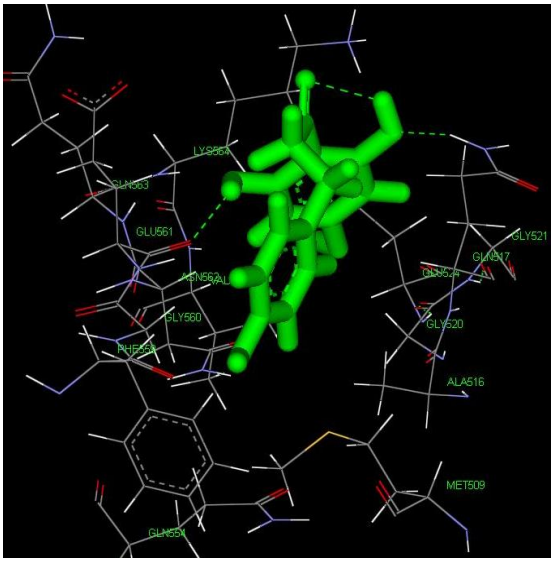


**Figure 4. Binding mode of phloretin at receptor showing interactions with Gln517 and Asn562.**

## *In silico* ADMET analysis:

The action of a drug is dependent on a sufficient amount of it being able to get into the body, find its way to the correct site of action, and for it to remain there unchanged for long enough to elicit a pharmacological response. In vivo pharmacokinetics of a potential drug molecule in man, whilst it exists as only a virtual structure is analysed by in silico prediction of ADMET properties. The plot of polar surface area (PSA) vs AlogP is shown in Figure 5. One of the compounds not satisfying the ADMET properties is hypericin found to be having high molecular weight. The two compounds shown out of pink ellipse are curcumin and phloretin however the levels are agreeable for these compounds when compared to the other molecules suggesting that the analogs of these compounds can evolve as potential antibiotics.

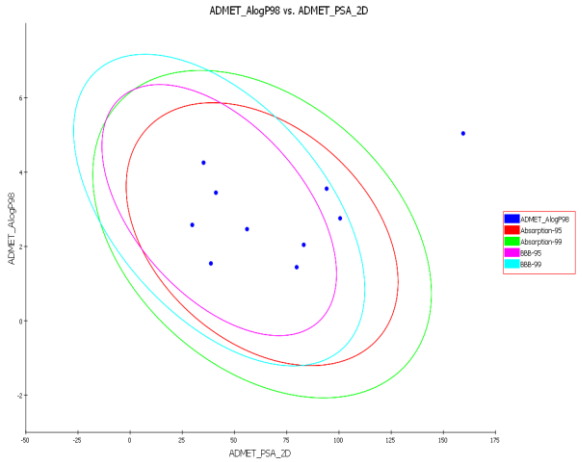


**Figure 5. ADMET plot of the ten natural compounds**

## CONCLUSIONS

In conclusion, we have created homology models of the the structure of SipB which in turn agrees with literature data and can serve as a valuable reference for characterizing the protein and could be explored for drug targeting. Also studied its complexes with some of the known active ligands, further invitro functional assays have to be performed for the compound identified to prove its effectiveness. Further investigations into the antipathogenic potential of these compounds may open new avenues for drug development in the control of antibiotic resistant pathogens.

## Acknowledgements

The authors thank the management of Sathyabama University for the cluster computing facilities installed in the department of Bioinformatics and thank the anonymous reviewers for their valuable comments and suggestions.

## References

1. Thong KL, Passey M, Clegg A, Combs BG, Yassin RM and Pang T. Molecular analysis of isolates of Salmonella typhi obtained from patients with fatal and nonfatal typhoid fever. J. Clin. Microbiol. 1996; 34:1029-1033.
2. Malik AS. Complications of bacteriologically confirmed typhoid fever in children. J Trop Pediatr. 2002; 48(2):102-108.
3. Le Minor L, Popoff MY, Bockemühl J. Supplement 1989 (n. 33) to the Kauffmann-White scheme. Res Microbiol. 1990; 141(9):1173-1177.
4. Harish BN, Menezes GA. Antimicrobial resistance in typhoidal salmonellae. Indian J Med Microbiol. 2011; 29(3):223-229.
5. Alós JI, Gómez-Garcés JL, Cogollos R, Amor E, Pérez-Rivilla A. Susceptibilities of ampicillin-resistant strains of Salmonella other than S. typhi to 10 antimicrobial agents. Antimicrob. Agents Chemother. 1992; 36:1794-1796.
6. Dromigny J-A, Perrier-Gros-Claude J-D. Antimicrobial Resistance of Salmonella enterica Serotype Typhi in Dakar, Senegal. Clin Infect Dis. 2003; 37 (3): 465-466.
7. Galán JE, Collmer A. Type III secretion machines: bacterial devices for protein delivery into host cells. Science 1999; 284: 1322 – 1328.
8. Ong SY, Ng FL, Badai SS, Yuryev A, Alam M. Analysis and construction of pathogenicity island regulatory pathways in Salmonella enterica serovar Typhi. J Integr Bioinform, 2010; 7(1):145.
9. Kaur J, Jain SK. Role of antigens and virulence factors of Salmonella enterica serovar Typhi in its pathogenesis. Microbiol Res. 2012; 167(4):199-210.
10. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov. 2006; 5:993–996.
11. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F. Complete genome sequence of Salmonella enterica serovar Typhimurium LT2, Genomics & Informatics. 2006; 4(4):182-187.
12. Pushpa OB, Suresh MX. SALMONELLABASE - An online database of druggable targets of Salmonella species. Bioinformation 2012; 8(15): 734-737.
13. Christopher F, Thangam EB, Suresh MX. A Bioinformatics search for selective Histamine H4 Receptor antagonists through structure-based virtual screening strategies. Chem. Biol. Drug Des. 2012; 79: 749-759.
14. SharmaR, Kumar BSV. In Silico Interaction Studies on Inhibitory Action of endophytic fungal Diketopiperazine and its related compounds on Heat-Shock Protein 90 (Hsp90). Asian J Biomed and Pharmaceu Sci. 2014; 04 (28): 25-29.
15. Patel SK, Jasrai YT., George L-B, Highland H N. Virtual Screening of Xanthones in Combating Malaria Targeting Plasmodium Falciparum Erythrocyte Membrane Protein1(PfEMP1) Asian J Biomed and Pharmaceu Sci 2012; 1(6): 01-05.
16. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press 2005; pp. 571-607
17. Kneller DG, Cohen FE, Langridge R. Improvements in Protein Secondary Structure Prediction by an Enhanced Neural Network. J. Mol. Biol. 1990; 214: 171-182.
18. Tusnady GE, Simon I. The HMMTOP transmembrane topology prediction server. Bioinformatics. 2001; 17:849-850.
19. Hirokawa T, Boon-Chieng S, Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics. 1998; 14:378-379.
20. Yang Zhang. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008; 9:40.
21. Laskowski RA, Mac Arthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst 1993; 26: 283-291.
22. Colovos C, Yeates TO. Verification Of Protein Structures: Patterns of nonbonded atomic interactions. Protein Sci. 1993; 2: 1511-1519.
23. Weisel M, Proschak E, Schneider G. PocketPicker: analysis of ligand binding-sites with shape descriptors. Chem. Cent. J 2007;1:7.
24. Laurie AT, Jackson RM. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics 2005; 21:1908–1916.
25. Venkatachalam CM, Jiang X, Oldfield T, Waldman M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, J. Mol. Graph. Model. 2003; 21(4) :289-307.
26. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, J. Comp.Chem. 1983; 4:187–217.
27. Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, Behrens S, Niinikoski A, Mewes HW, Horn M, Rattei T. Sequence-based prediction of type III secreted proteins. PLoS Pathog. 2009; 5(4): e1000376.
28. McGhie EJ., Hume PJ, Hayward RD, Torres J, Koronakis V. Topology of the Salmonella invasion protein SipB in a model bilayer. Mol. Microbiol. 2002; 44(5): 1309–132.