



Virtual screening of Threonine synthase as a target for antimicrobial resistance in *Toxoplasma gondii*

Budhayash Gautam¹, Gurmit Singh² and Satendra Singh^{1,*}

¹Department of Computational Biology and Bioinformatics, Jacob School of Biotechnology and Bio-engineering, Sam Higgin bottom Institute of Agricultural, Technology and Sciences, Allahabad 211007, U.P, India.

²Department of Computer Science and Information Technology, Shepherd School of Engineering and Technology, Sam Higgin bottom Institute of Agricultural, Technology and Sciences, Allahabad 211007, U.P, India.

ARTICLE INFO

Article history:

Received: 2 April 2012;

Received in revised form:

28 June 2012;

Accepted: 17 July 2012;

Keywords

Homology modeling,
Molecular dynamics,
Threonine synthase,
T. gondii,
Antimicrobial resistance,
Virtual screening,
Drug targets.

ABSTRACT

Toxoplasma gondii is an obligate intracellular apicomplexan parasite that can infect a wide range of warm-blooded animals including humans. In humans and other intermediate hosts, toxoplasma develops into chronic infection that cannot be eliminated by host's immune response or by currently used drugs. In most cases, chronic infections are largely asymptomatic unless the host becomes immune compromised. Thus, toxoplasma is a global health problem and the situation has become more precarious due to the advent of HIV infections and poor toleration of drugs used to treat toxoplasma infection, having severe side effects and also resistance have been developed to the current generation of drugs. The emergence of these drug resistant varieties of *T. gondii* has led to a search for novel drug target. Threonine synthase is one such protein. It is a member of Serine / Threonine protein family which is a very important metabolic enzyme especially in signal transduction mechanisms of a cell. It is an active part of Glycine, Serine, Threonine Metabolism, Vitamin B6 Metabolism pathways. The PDB and Modebase 3D structure of Threonine synthase was not available; therefore homology model of Threonine synthase was constructed using MODELLER and model was validated using PROCHECK and Verify3D programs and SAVS and PROSA server to obtain a stable structure. The model was further explored for the molecular dynamics simulation study with GROMACS. Binding site was characterized by CASTp and PASS. Virtual screening of Threonine synthase was performed, against the NCI diversity subset II molecules (retrieved from ZINC database), through molecular docking studies using Autodock- vina. The Python scripts in MGL tools package were used to analyze the docking results. The top ten ligands were selected based on AutoDock energy score. Further the Complexes were analyzed for their interaction through LIGPLOT. On the basis of Complex scoring and binding ability it is deciphered that these top ten NCI diversity set II compounds could be promising inhibitors for *T. gondii* using, Threonine synthase as drug target, yet pharmacological studies have to confirm it.

© 2012 Elixir All rights reserved.

Introduction:

Toxoplasma gondii is an obligate intracellular Apicomplexan parasite that can infect a wide range of warm-blooded animals including humans [1]. *Toxoplasma gondii* was initially discovered by accident, in 1908, by a scientist, Charles Nicolle, who was working in North Africa and searching for reservoir of Leishmania in a native rodent, *Ctenodactylus gundi*. The gundis live in the foothills and mountains of Southern Tunisia and were commonly used to study Leishmania at the Pasteur Institute in Tunis. The name Toxoplasma means 'arc-form' in Greek and was named according to the crescent-shaped morphology of the tachyzoite and bradyzoite stages of the organism observed by the scientists. At about the same time, Alfonso Splendore working in Sao Paulo discovered a similar parasite in rabbits. This pathogen is one of the most common in humans due to many contributing factors that include: (1) its complex life cycle allows it to be transmitted both sexually via felid fecal matter and asexually via carnivorousism. (2) *Toxoplasma* has an extremely wide host cell

tropism that includes most nucleated cells. (3) In humans and other intermediate hosts, *Toxoplasma* develops into a chronic infection that cannot be eliminated by the host's immune response or by currently used drugs. In most cases, chronic infections are largely asymptomatic unless the host becomes immune compromised. Together, these and other properties have allowed *Toxoplasma* to achieve infection rates that range from ~23% in the USA (2) to 50-70% in France (3).

There are two major reasons that new drugs are needed to treat *Toxoplasma* infections. First, the drugs currently used to treat *Toxoplasma* infections are poorly tolerated, have severe side effects, and cannot act against bradyzoites (4). Second, there are reports that *Toxoplasma* is developing resistance to the current generation of drugs (5, 6). How resistance to these drugs has developed is not known but is critical to understand because it will lead to improved drug design and will increase our understanding of the biological functions of these drug targets. One way to understanding mechanisms of resistance is to compare the transcriptional profiles of wild-type and resistant

parasites grown in the absence or presence of the drug. Such studies in bacterial resistance have demonstrated that pathogen responses to antibiotics are multifactorial and complex (7). Whether the same will be true in *Toxoplasma* is unclear, but data from these types of experiments will likely impact new anti-*Toxoplasma* drug design.

Threonine synthase is one such protein (7). It is a member of Serine / Threonine protein family, a very important metabolic enzyme especially in signal transduction mechanisms of a cell (7). If a drug molecule can inhibit this protein, it can kill the pathogen while not affecting the host. As this is also present in the targets of tropical diseases of TDR (Tropical disease related Drug) target database and no PDB and MODBASE 3D structural model is available. Thus, as a case study, we have built homology models and validated by using various online servers. The model was further explored for the molecular dynamics simulation study, virtual screening and molecular docking studies with suitable inhibitors. The three dimensional model of Threonine synthase presented here would be helpful in guiding both enzymatic studies as well as design of specific inhibitors.

Materials and Methods:

Template selection is a critical step in homology modeling. The amino acid sequence of Threonine synthase of *T. gondii* (target) (Genbank Accession I.D. EEB02794.1) was retrieved from the protein database of National Center of Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>). The three-dimensional structure of the protein was not yet available in Protein Data Bank and Modbase Database. Also there is no expression information available for this protein. The 3D structure of the threonine synthase of *T. gondii*, was predicted by comparative modelling approach. BLASTP search was performed against Brookhaven Protein Data Bank (PDB) with the default parameters to find suitable templates for homology modeling (8). The academic version of MODELLER9v6 (<http://www.salilab.org/modeller/>), was used for 3D structure generation based on the information obtained from sequence alignment of target and template/s. The sequence alignment of threonine synthase of *T. gondii* and template sequence was carried out using the CLUSTALW (<http://www.ebi.ac.uk/clustalw/>) program. Models generated by MODELLER, were validated using PROCHECK and Verify3D programs and SAVS and PROSA server to obtain a stable structure. The one with the best G-score of PROCHECK and with the best VERIFY3D profile was selected as best model and subjected to energy minimization (8).

With the aim of evaluating the stability and folding, conformational changes and getting insights into the natural dynamics on different timescale of protein in solution, 10 nanoseconds (ns) molecular dynamics simulation were performed. The molecular dynamics (MD) simulations of modeled threonine synthase protein were carried out with the GROMACS 4.0.6 software package by employing GROMOS 96 force field and the flexible SPC (Simple Point Charge) water model. The initial structure was immersed in a periodic water box of cubic shape (0.5 nm thick). Electrostatic energy was calculated using the particle mesh Ewald method. Cutoff distance for the calculation of the coulomb and van der Waals interaction was 1.0 Å. After energy minimization using a steepest descent for 1000 steps, the system was subjected to equilibration at 300K and normal pressure for 100 ps under the conditions of position restraints for heavy atoms. LINCS

constraints were performed for all bonds, keeping the whole protein molecule fixed and allowing only the water molecule to move to equilibrate with respect to the protein structure. The system was coupled to the external bath by the Berendsen pressure and temperature coupling. The final MD calculations were performed for 10 ns under the same conditions except that the position restraints were removed. The results were analyzed using the standard software provided by the GROMACS package. An average structure was refined further using a steepest descent energy minimization (8).

Then, Active site was characterized by CASTp and PASS and compared by using the information of binding sites. By comparing prediction of above two algorithms, best active site was selected. The docking of threonine synthase was performed, against the NCI subset II molecules retrieved from ZINC database by using Autodock- vina (<http://vina.scripps.edu/>), where 1,364 molecules from the NCI diversity subset II (<http://zinc.docking.org/index.shtml>) were screened. The Python scripts in MGL tools package were used to analyze the docking results. The search for the best ways to fit ligand molecules from the NCI diversity subset II, into threonine synthase modeled structure, using Autodock- vina resulted in docking files that contained detailed records of docking. These log files were read into ADT (Auto Dock Tool) to analyze the results of docking. The top ten ligands were selected based on the energy score after virtual screening and were further analyzed by the program LIGPLOT. The Ligplot represents the hydrogen and hydrophobic interactions between ligand and active site residues (9).

All the visualization of structures was performed using Pymol molecular graphic system (<http://www.pymol.org>).

Results and Discussion:

Based on the maximum identity with high score and lower e-value the Crystal Structure of Threonine Synthase from Yeast (*Saccharomyces cerevisiae*) with PDB-ID "1KL7[A]" was selected as template. The sequence identity and similarity between the target and template are 58% and 73%, respectively. E- Value was 0.0 (8).

By comparison of the results for all the models second model (Figure 1 A) was judged as best among five. In the case of second model, it was found that the phi/psi angles of 85.2% of the residues fell in the most favored regions, 13.2% of the residues fell in the additional allowed regions, 1.7% fell in the generously allowed regions, and only 0.0 % of the residues fell in the disallowed regions (Figure 1 B). All these findings suggest a stereo-chemically very good model.

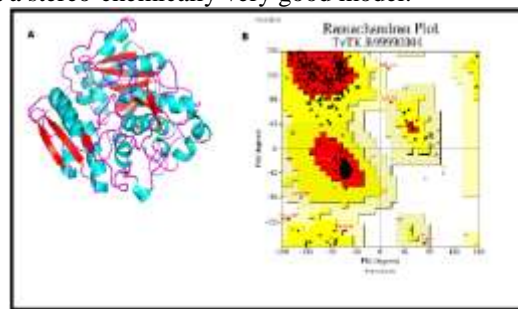


Figure 1: (A) Ribbon representation of the final model of threonine synthase protein; (B) Ramachandran plot computed with PROCHECK program with 85.2 % of the residues in the most favored regions.

The overall PROCHECK G factor for the homology-modeled structure was 0.00. The energetic architecture as

predicted by PROSA score was negative (-9.62) for the modeled protein. The value is close to that of template (-9.60), which indicates the reliability of the model. The final structure was validated by an ERRAT graph. The quality factor of 77.106 indicated good quality, as scores of >50 are acceptable for a reasonable model.

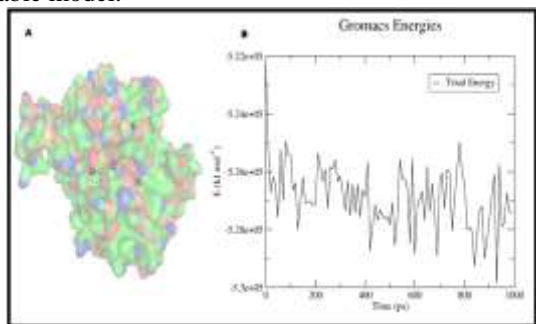


Figure 2: (A) Active sites are shown; spheres show the coordinates of the center of the active site predicted by CastP program. (B) Gromacs energies plot shows total energy of the molecule throughout the time period of molecular dynamics simulation.

High quality of model is also confirmed from VERIFY 3D server as 76.00% of residues of modeled protein showed a score higher than 0.02 thus the model showed satisfactory 3D-1D scores for all the residues in the sequence. Results of WHAT_CHECK also indicate about the correctness of the modeled structure. These results indicate that the obtained structure has reasonably good quality (8, 10).

The predicted Model of threonine synthase has shown good accuracy of the structure. But in order to use this model for virtual screening predicted model should have stable molecular dynamic behavior. The molecular dynamic stability has been established by performing the molecular dynamic simulation study using GROMACS 4.0.6 software package. The root mean square deviation (RMSD) during the simulation was increasing in the beginning but after 220 ps it becomes almost constant for rest of the duration of the simulation. This suggests that the threonine synthase model has very less RMSD for the backbone and it also has less flexibility, indicating that model has a stable dynamic behavior structure. Molecular dynamic simulation study showed that the energy of the molecule was found to be constant throughout the time period of simulation (Figure 2 B). This suggests that the molecule has a stable structure as required for the virtual screening and drug designing. The root mean square (RMS) fluctuations were very less. Most the atoms were free from the RMS fluctuations. Very few atoms have shown RMS fluctuation at C and N terminal due to the loop region. This suggests the threonine synthase model has an accurate and stable structure which can be used for virtual screening. The simulation studies also indicated that radius of gyration was increasing in the beginning but after 400 ps it decreases upto 420 ps and finally became almost constant for rest of the duration of the simulation. This suggests that the threonine synthase model has a compact structure which provides the required stability to the molecule (8, 10).

After getting the final stable model, the possible binding sites of modeled structure were searched using the CASTp server and PASS program. Ten possible binding sites were obtained. These sites were compared with active site of the template and it was found that pocket1 is highly conserved with the template. Since the threonine synthase proteins of *T. gondii*

and *saccharomyces cerevisiae* are well conserved in both sequence and structure, it is predicted that their biological function may be identical. From the structure-structure comparison of template and final refined models of threonine synthase, it was found that the residues in the site1 Thr203, Arg112, Ala137, Asn136, Ser134, Ser318, Tyr182, Ile260, Leu109 and Thr395 are highly conserved within the active site of template. In this study, site1 is chosen as the binding site to dock with the NCI diversity set molecules (8, 10) (Figure 2 A).

The top ten ligand molecules having minimum energy were screened out as the possible inhibitors for threonine synthase (Table 1). Protein-ligand interactions of top four ligand molecules have been shown in Figure 3. Correlation coefficient analysis was performed between energy score calculated for all selected ligands molecules, log P value and molecular weight (Table 1). The correlation coefficient is a statistical calculation that is used to examine the relationship between two sets of value. The Correlation coefficient between energy score and log P value was -0.101603358 and between energy score and molecular weight was -0.236816724. These values suggest a clear negative correlation between energy score of the ligands and their molecular weight and log P value. Thus the energy score of the ligands is independent of their molecular weight and log P value. The value of the correlation coefficient tells us about the strength and the nature of the relationship. Correlation coefficient values can range between +1.00 to -1.00. The negative correlation suggests that energy score may depend on interactions or the conformation of ligands and active site residues (8, 10).

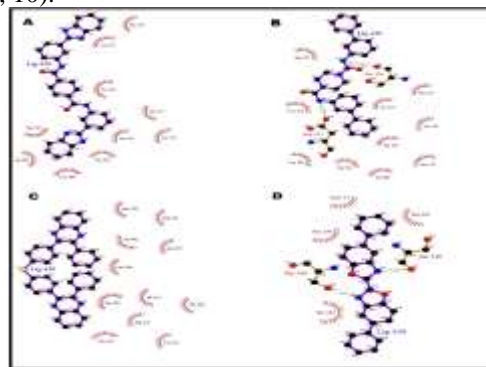


Figure 3: Ligplots showing the protein-ligand interactions of top four ligands, based on energy score (hydrogen bonding and hydrophobic), generated by Ligplot program.

(a) ZINC01690699 (b) ZINC04783229 (c) ZINC05492794 & (d) ZINC13152284

It was found that the inhibitor ZINC01690699 has minimum energy score which reveals higher binding affinity towards the threonine synthase and the inhibitor was also showing one of the best interactions with residues of the active site. The other important drug like properties like molecular weight and logP value were also found within the limits of drug like molecules (8, 10).

Conclusion:

Hence, in present study, based on virtual screening, it can be concluded that the molecule ZINC01690699 1-N, 4-N-bis [3-(1H-benzimidazol-2-yl) phenyl] benzene-1, 4-dicarboxamide has the potential to inhibit the growth of *T. gondii* and can act as remedy for the treatment of *T. gondii* infection. For the wet laboratory validation of present study, Growth inhibition of *T. gondii* by inhibitor ZINC01690699 can be established by the qualitative and quantities microbiological techniques. Thus, the

three dimensional model of threonine synthase presented here would be helpful in guiding both enzymatic studies as well as design of specific inhibitors. Yet to confirm it to be promising target, pharmacological studies need to be performed.

Acknowledgment:

The authors would like to acknowledge the support and facilities provided by the Department of Computational Biology and Bioinformatics, Sam Higginbottom Institute of Agriculture, Technology & Sciences, Deemed to be University, Allahabad, U.P., India.

References:

1. Montova JG & Liesenfeld O, *Lancet*. 2004 363: 1965 [PMID: 15194258].
2. Jones JL et al. *Am J Epidemiol*. 2001. 154: 357 [PMID: 11495859].
3. Tenter AM et al. *Int J Parasitol*. 2000 30: 1217 [PMID: 11113252].
4. McCabe R. Cambridge University Press. 2001 p. 319-59.
5. Baatz H et al. *Ocul Immunol Inflamm*. 2006 14: 185 [PMID: 16766403].

16766403]

6. Aspinall TV et al. *J Infect Dis*. 2002 185: 1637 [PMID: 12023770].
7. Brazas MD & Hancock RE. *Drug Discov Today* 2005 10: 1245 [PMID: 16213417].
8. Gautam B., Singh G., Wadhwa G., Farmer R., Singh S., Singh A. K., Jain P.A., Yadav P.K. Metabolic pathway analysis and molecular docking analysis for identification of putative drug targets in *Toxoplasma gondii*: novel approach. *Bioinformation* 8(3): 134-141 (2012).
9. Wallace AC, Laskowski RA, and Thornton JM: **LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions**. *Prot. Eng.* (1995) 8:127-134.
10. Farmer R., Gautam B., Singh S., Yadav P.K., Jain P.A. Virtual screening of AmpC/β-lactamase as target for antimicrobial resistance in *Pseudomonas aeruginosa* *Bioinformation* 2010 4(7): 290-294.

Table 1: Top ten ligand molecules identified from NCI diversity subset II of the ZINC database after virtual screening

S. No.	ZINC ID of the screened molecules	Energy score (Kcal/Mol)	No. of Hydrogen Interaction of ligand with active site residues	No. Hydrophobic Interaction of ligand with active site residues	Molecular Weight (g/mol)	Log-P value
1.	ZINC01690699	-9.1	2	10	548.59336	6.0
2.	ZINC04783229	-8.1	5	10	546.52144	5.9
3.	ZINC05492794	-7.9	4	11	546.52144	5.9
4.	ZINC13152284	-7.9	3	4	453.47264	5.3
5.	ZINC13152284_03	-7.9	4	6	546.52144	5.9
6.	ZINC01572309	-7.9	3	5	474.55802	5.0
7.	ZINC18055497_01	-7.8	1	7	479.67574	4.7
8.	ZINC18141294	-7.8	1	12	453.47264	5.3
9.	ZINC05462674	-7.8	4	6	546.52144	5.9
10.	ZINC13152284_01	-7.8	1	7	392.4492	6.7