ACTA SCIENTIFIC PHARMACEUTICAL SCIENCES (ISSN: 2581-5423)

Volume 3 Issue 6 June 2019

Research Article

Virtual Screening of Compounds Against *Mycobacterium tuberculosis* Maltosyltransferase GlgE

Junie B Billones*, Juriel Kristan F Ramirez, Yves Ira A ReyeS, Abdul-Rashid B Sampaco III

Department of Physical Sciences and Mathematics, College of Arts and Sciences, University of The Philippines Manila, Padre Faura, Ermita, Philippines

***Corresponding Author:** Junie B Billones, Department of Physical Sciences and Mathematics, College of Arts and Sciences, University of the Philippines Manila, Padre Faura, Ermita, Philippines.

Received: March 29, 2019; **Published:** May 13, 2019

DOI: 10.31080/ASPS.2019.03.0281

Abstract

The prevalence of tuberculosis (TB) caused by Mycobacterium tuberculosis (Mtb) particularly the occurrence of multi-drug resistant and extensively-drug resistant strains of Mtb has prompted intense efforts to develop new anti-TB drugs. The enzyme maltosyltransferase GlgE of M. tuberculosis was determined to be a potential drug target. In this work, homology model of Mtb maltosyltransferase GlgE was generated based on eight reported protein structures with similar amino acid sequence. A pharmacophore based on the model was used to screen two databases of natural products. The virtual screening top hit, F_2.726 with binding energy (BE) of -322.83 kcal/mol, and the natural substrate maltose (BE = -322.85 kcal/mol), have comparable binding affinity. The top hit, a phenylcyclohexenyl carbamimidic acid, was then subjected to in silico structure optimization by De Novo Evolution method and yielded five variants with better binding affinities. The study also indicates that the GlgE structures from Streptomyces coelicolor can be used as templates for designing the GlgE inhibitors of Mycobacterium tuberculosis.

Keywords: *Mycobacterium tuberculosis*; GlgE Inhibitors; Homology Modeling; Streptomyces *coelicolor*; Maltosyltransferase

Introduction

Tuberculosis, commonly abbreviated as TB, has been a perennial global health problem. It impairs the health of approximately 10 million people each year and is one of the ten leading causes of death worldwide. In the past five years, it has overtaken HIV/AIDS as the number one cause of death by a single infectious agent [1]. In 2017, 6.4 million new cases of TB with 1.3 million deaths were reported to WHO. The United Nations (UN) meeting in September 2018, a first high-level meeting on TB, underscores the need for immediate response towards the global goal of eradicating the TB epidemic by 2030.

TB is an infectious disease caused by *Mycobacterium tuberculosis* which affects primarily the lungs and also other extrapulmonary sites. It is easily transmitted through airborne droplets that proceed to infect mainly the tracheal pathways. A number of drug treatments (i.e. Isoniazid, Rifampicin, Ethambutol, etc.) are already available but the emergence of resistance to available drugs has made tuberculosis a continuous health threat worldwide [2]. Multi-drug resistant TB (MDR-TB), defined as TB resistant to the two most potent first-line anti-TB drugs (i.e. Rifampicin and Isoniazid), and extremely drug-resistant TB (XDR-TB), defined as MDR-TB resistant to any fluoroquinolone and at least one second-line injectable drug such as capreomycin, kanamycin, or amikacin [3,4], are not only prevalent in highly populated countries [5,6] but also pose additional challenges for effective control of TB in many parts of the world [1]. As resistance to anti-TB drugs continue to rise, there is a growing need to develop new classes of antitubercular agents.

A genetically validated target in Mtb is maltosyltransferase GlgE, an important enzyme in the α-glucan pathway of *M. tuberculosis.* α-glucan in mycobacteria is exclusively synthesized intracellularly using α-maltose-1-phosphate as the substrate for the maltosyltransferase GlgE [7]. The inhibition of GlgE does not only lead to failure of biosynthesis of α-glucan, the main constituent of the mycobacterial capsule that is thought to be involved in immune evasion and virulence, but also the toxic accumulation of maltose-1-phosphate, that eventually leads to cell death [8].

In our previous work, we used the structure of GlgE enzyme in *Streptomyces coelicolor* (Sco), a homologue of *M. tuberculosis* GlgE as the drug target in virtually screening several databases of natural products [9]. These two enzymes having the same catalytic and very similar kinetic properties were found to have conserved donor pocket capable of binding maltose [10]. However, some inhibitors of the *S. coelicolor* GlgE like cyclodextrin does not inhibit the *M. tuberculosis* GlgE because a hydrophobic patch near the maltose-binding site is not completely conserved in Mtb GlgE. We previously employed pharmacophore-based screening and molecular docking to search for new potential leads against Mtb GlgE, based on the structure of Sco GlgE [9].

Computational methods have been proven useful in the discovery and development of new drugs [11,12]. As an essential step in computer-aided drug discovery specially in the absence of crystal data, homology modeling has been very successful in predicting protein structures using sequence similarities particularly with the use of multiple templates [13]. In particular, compound R207910, now known as Bedaquiline and the first anti-TB drug introduced in 60 years, was discovered by molecular docking of series of diarylquinolines on the homology model of the binding site of Mycobacterium tuberculosis ATPase [14]. We have also applied these computational approaches in discovering leads against other druggable targets in Mtb [9,15-21].

In this study, homology modeling was employed to generate a 3-dimensional structure model for the *M. tuberculosis* GlgE enzyme. A pharmacophore was subsequently generated based on the features of the binding site. The compounds from Ambinter (www.ambinter.com) and Analyticon (https://ac-discovery.com/screening-libraries/) databases were screened against the Mtb GlgE target with the use of the generated pharmacophore. The top hits were subjected to molecular docking studies and the top-binding ligand were further elaborated *in silico* using De Novo Evolution technique.

Materials and Methods

All computational procedures were done on Accelrys Discovery Studio (DS) Client v2.5.0.9164 installed on a computer running on Microsoft® Windows 7 Home Premium 64-bit Operating System using a processor of Intel® Core™ i3 CPU @ 2.27 GHz and installed memory (RAM) of 4.00 GB.

Homology Modelling of *M. tuberculosis* GlgE Enzyme

The amino acid sequence of the *M. tuberculosis* GlgE enzyme was retrieved from www.uniprot.org. The sequence was then inputted in RCSB Protein Databank to search for GlgE crystal data. All PDB structures yielding E value of 0 was retrieved. A subunit was copied from each PDB structure and pooled with the rest. The subunits were aligned using Align Structures (MODELER) protocol. The sequence of the *M. tuberculosis* GlgE was aligned with that of the templates. The aligned sequences were then used as template for the homology modeling. The model with the best PDF score was used in the subsequent procedure.

Ligand preparation

Isomers, tautomers, different ionization states and 3D conformations of the molecules were generated using the *Prepare Ligands* protocol. This protocol generated several variants of each ligand. All structures were saved and stored in 3D databases using the Build 3D Database protocol.

Pharmacophore generation and virtual screening

The active site of the protein was modeled using a site sphere based on experimentally identified amino acid residues in the binding site. That is, the amino acid residues reported to participate in binding of substrate to the protein was listed and selected. The site sphere was generated using the Define Sphere from Selection protocol. A pharmacophore based on the active site was created using the Interaction Generation protocol. The resulting pharmacophore was used to screen the molecules in the constructed database of 3D structures. Virtual screening process was done using the Screen Library protocol and was divided into two parts using two different fitting methods. The first process involved rigid fitting of the molecules onto the pharmacophore, while the second process involved flexible fitting. A fit-value cut-off of 2.6 was set for rigid fitting. All molecules that had a fit value greater than or equal to 2.6 was reprepared for flexible fitting. The molecule with high fit values in the flexible fitting run were forwarded for molecular docking studies.

Molecular docking

The top hits, along with our previously reported hits [9], were docked using the *CDOCKER* docking protocol. The binding energies of the top docking poses were calculated using the *Calculate Binding Energy* protocol. Finally, the ligand interaction diagrams were generated for each enzyme-ligand complex.

De Novo evolution

To further explore the inhibitory potential of the top hits, *in silico* structure optimization was performed using the De Novo Evolution protocol in DS, which contains orgnaicx.str or fragment_link.inp libraries.

Results and Discussion

Our previous work on virtual screening against the GlgE drug target was based on the structure of the enzyme in *S. coelicolor* [9] because Mtb GlgE was not yet available at that time. The recent isolation of *M tuberculosis* GlgE allows us to perform virtual screening studies using the GlgE enzyme from the target pathogen itself. Thus, in this study we generated a homology model of the *M. tuberculosis* GlgE enzyme using the sequence of the GlgE from Mtb (strain ATCC 25618/H37Rv) [8] that was made available in UniProtKB (https://www.uniprot.org/uniprot/P9WQ17).

The sequence was used as a query to perform a blast search in RCSB Protein databank [22] for protein structures with similar sequence, the E value being the primary criterion of sequence similarity. An E value of 0 means there is exact match between the query sequence and the retrieved sequence from the database. Thus, only hits with 0 E value were retrieved and were used as reference structures for the homology modeling. Accordingly, eight structures were retrieved with PDB codes of: 3ZSS, 3ZST, 3ZT7, 3ZT6, 3ZT5, 4CN6, 4CN4, 4CN1. These were GlgE structures of *S. coelicolor* [10]. This shows that *S. coelicolor* and *M. tuberculosis* indeed have very related GlgE enzymes. Subsequently, homology modeling was done vide supra.

The results of the sequence alignment of the *M. tuberculosis* GlgE enzyme with that of the templates is shown in Figure 1. The amino acid residues were highlighted at every position according to consensus with the other sequences. The color ranges from blue to white in order of decreasing similarity. Although the GlgE in *M. tuberculosis* and *S. coelicolor* shared a number of conserved residues, they differ in lengths, with Mtb GlgE being longer.

Figure 1: Alignment of the *M. tuberculosis* GlgE protein sequence with the templates.

Figure 2A shows the structure of Sco GlgE that we used in our previous work while Figure 2B depicts the outcome of the homology modeling of Mtb GlgE. Figure 2C shows the comparison of the generated homology model to the homologue (PDB code = 3ZT5) we used previously [9], a crystal structure of the *S. coelicolor* GlgE enzyme. Careful inspection of Figure 2C will show that some part of the generated model (yellow) has no overlap with that of the 3ZT5 (red). As pointed above, the *M. tuberculosis* GlgE has a longer amino acid sequence and some of its amino acids have no counterpart in the *S. coelicolor* homologue. These parts of non-overlap represent the gaps in the alignment shown in Figure 1.

Amino acids found in the binding site of the *S. coelicolor* GlgE was identified using the Show Ligand Binding Site Atoms menu option. This information was used to define the binding sphere of the *M. tuberculosis* GlgE model. The overlay image shows that the relevant residues are located in the same spot in both models, validating the coordinates that we used to define the binding site in Mtb GlgE. The resulting binding sphere, shown in Figure 3, was used to generate a structure-based pharmacophore of the active site. The pharmacophore contained features that represent acceptors, donors, and hydrophobes. It was then used to screen databases of natural products.

Figure 2: A) Solid ribbon diagram of the structure of *S. coelicolor* GlgE isoform used in [9], (PDB code = 3ZT5).
B) Generated solid ribbon diagram of the 3D structure of the *M. tuberculosis* GIgE enzyme through Homology Modelling.
C) Overlay of the two structures using Carbon stick diagram.

Figure 3: Binding sphere (red) in Mtb GlgE that directs the docking of the substrate maltose

The databases built from the 74,673 molecules from Ambinter and Analyticon databases were used as input for the screening processes. Virtual screening by rigid fitting searches for molecules with favorable interactions with the acceptors, donors, and hydrophobes in the binding site as represented by the complementary features in the generated pharmacophore model. A higher fit value means better fit. An arbitrary cut-off value of 2.6 was set to trim down the hits to manageable number. All molecules that passed

were re-prepared and re-screened using flexible fitting. Flexible fitting method performs the same algorithm as rigid fitting but treats the molecule as a flexible structure. The top scoring hit from each database was docked to the Mtb GlgE model.

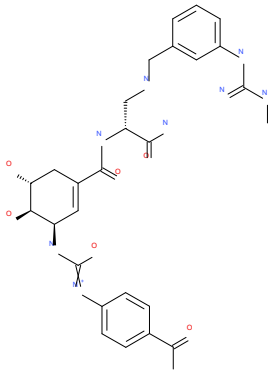
The binding sphere (Figure 3) was used to direct the docking of the natural substrate of the enzyme, maltose, as well as the two top hits from virtual screening (Figure 4). The two-dimensional (2D) Ligand Interaction Diagrams (Figure 5 and 6) were generated to examine the binding of ligand to the protein. The Binding Energies (BE) were calculated to determine the ligands that form more favorable interactions with the receptor. A more negative binding energy denotes stronger or more favorable binding. Maltose binding energy was calculated to be -322.85 kcal/mol. Among the top hits, only F_2.726 (N'-(4-acetylphenyl)-N-((1R,5R,6R)-3-(((R)-1-amino-1-oxopropan-2-yl)carbamoyl)-5,6-dihydroxycyclohex-2-en-1-yl)carbamimidic acid) from Analyticon, with BE of -322.83 kcal/mol) has comparable binding energy with maltose.

Figure 4: Virtual screening top hit from A) AnalytiCon and B) Ambinter Natural Products databases.

Figure 5: Maltose docked to the defined active site of the *M. tuberculosis* GlgE model.

Figure 6: F_2.726 docked to the defined active site of the *M. tuberculosis* GlgE model.

De Novo Evolution technique was implemented to generate derivatives of the F_2.726. De Novo Evolution protocol attaches chemical fragments, from a predefined fragment library, onto a ligand scaffold to generate more favorable interaction with the receptor. The process generated structures with even better binding energies than F_2.726 as detailed in Table 1. Ligand 1 features a 1-methyl-3-(3-((methylamino)methyl)phenyl) guanidine moiety at the amido end of parent F_2.726. Ligand 2 is a slight modification of Ligand 1 in having a 1,1'-(1,3-phenylene) bis(N-methylmethanamine) group on the same amide terminal. Ligand 3 has much simpler pendant, an N-methyl-1-phenylmethanamine, although an additional hydroxyl group was added to the cyclohexene core. Ligand 4 is quite different with a pyrroleaminopropanol terminal on the amido functionality.

ID	Structure	Binding Energy (kcal/mol)	Fragment Library
Ligand 1		-374.342	fragment_link.inp

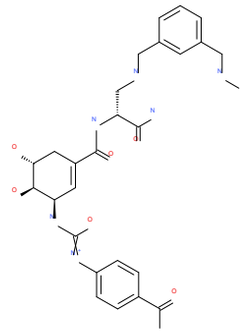
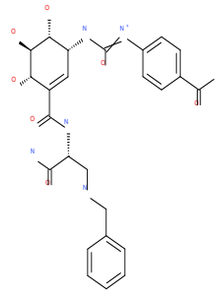
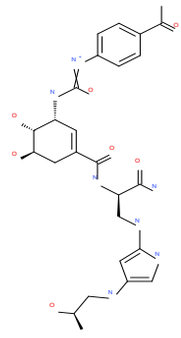
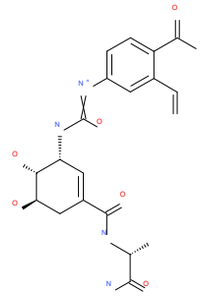
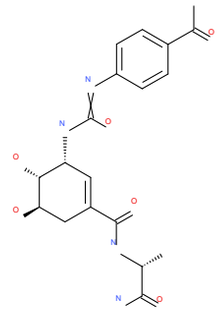
Ligand 2		-368.867	fragment_link.inp
Ligand 3		-363.448	fragment_link.inp
Ligand 4		-360.493	fragment_link.inp
Ligand 5		-336.609	orgx
F_2.726		-322.825	Unmodified

Table 1: Structure and binding energies of the F_2.726-derived derivatives, which were generated by *De Novo Evolution* method.

Lastly, Ligand 5 has the slightest modification, as it only involves the addition of a vinyl substituent on the acetophenone end of F_2.726.

The top hits from our previous study [9] were also docked to see any difference between screening using a homology model of a complete and correct sequence as target, and a crystal structure of a homologue, as was done before. Interestingly, of the four structures [9] only the unmodified structure, ZINC39010596 (5,7-dihydroxy-2-propan-2-yl-8-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H- pyran-2-yl]oxychromen-4-one), generated poses with the use of CDOCKER protocol in DS. Maltose was consistent in returning a more negative binding energy relative to ZINC39010596, whose binding energy is roughly a 100 kcal/mol positive than that of maltose in both cases. This means that this present work involving a homology model of Mtb GlgE predicts the same relative potency of ZINC39010596 vis-à-vis the natural substrate of GlgE. This also supports the use of *S. coelicolor* GlgE homologue [9] as basis for designing *M. tuberculosis* GlgE inhibitors.

Conclusion

A homology model of the *M. tuberculosis* GlgE enzyme based on the template structures with PDB codes of 3ZSS, 3ZST, 3ZT7, 3ZT6, 3ZT5, 4CN6, 4CN4, 4CN1, was generated. A pharmacophore was generated based on the protein's binding site. Virtual screening was done to search for potential inhibitors of GlgE, which is a highly druggable target in Mtb. The high scoring hits were docked to the homology model of Mtb GlgE. Only the molecule F_2.726 from Analyticon exhibited a binding affinity that is closer to that of the substrate maltose. *In silico* derivatization of this ligand by De Novo Evolution technique yielded five more ligands that have potentially superior binding abilities with Mtb GlgE target. It is also noteworthy, that the substrate-enzyme complex is consistently a 100-kcal/mol more stable than that with a top hit (ZINC39010596) justifying the use of a *S. coelicolor* GlgE homologue in designing *M. tuberculosis* GlgE inhibitors.

Acknowledgement

We are grateful to the Enhanced Creative Work and Research Grant (ECWRG 2017-1-032) and Emerging Interdisciplinary Research (EIDR) Program of the University of the Philippines System (OVPAA-EIDR 12-001-121102) for the use of the computing facility.

Bibliography

1. World Health Organization (WHO). Global Tuberculosis Report 2018.
2. Ahmad S and Mokaddas E. "Current status and future trends in the diagnosis and treatment of drug-susceptible and multidrug-resistant tuberculosis". *Journal of Infection and Public Health* 7.2 (2014): 75-91.
3. World Health Organization (WHO). Multidrug-resistant tuberculosis (MDR-TB). In: 2015 UPDATE. Geneva, (2015).
4. World Health Organization (WHO). Definitions and reporting framework for tuberculosis-2013 revision. Geneva, (2013).
5. Alene KA., *et al*. "Treatment outcomes of patients with multidrug-resistant and extensively drug resistant tuberculosis in Hunan Province, China". *BMC Infection Disease* 17 (2017): 573.
6. Isaakidis P., *et al*. "Alarming levels of drug-resistant tuberculosis in HIV-Infected patients in metropolitan Mumbai, India". *PLoS One* 9.10 (2014): e110461.
7. Koliwer-Brandl H., *et al*. "Metabolic Network for the Biosynthesis of Intra- and Extracellular α-Glucans Required for Virulence of Mycobacterium tuberculosis." *PLoS Pathogen* 12.8 (2016): e1005768.
8. Kalscheuer R., *et al*. "Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an alpha-glucan pathway". *Nature Chemical Biology* 6.5 (2010): 376-384.
9. Billones JB and Valle AMF. "Structure-based Design of Inhibitors Against Maltosyltransferase GlgE". *Oriental Journal of Chemistry* 30.3 (2014): 1137-1145.
10. Syson K., *et al*. "Structure of Streptomyces maltosyltransferase GlgE, a homologue of a genetically validated anti-tuberculosis target". *Journal of Biological Chemistry* 286 (2011): 38298-38310.
11. Sliwoski G., *et al*. "Computational methods in drug discovery". *Pharmacology Review* 66.1 (2013): 334-395.
12. Ou-Yang SS., *et al*. "Computational Drug Discovery". *Acta Pharmacologica Sinica* 33.9 (2012): 1131-1140.
13. Mobarec JC. "Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use?". *Journal of Medicinal Chemistry* 52.16 (2009): 5207-5216.

14. de Jonge MR., *et al*. "A computational model of the inhibition of Mycobacterium tuberculosis ATPase by a new drug candidate R207910". *Proteins* 67.4 (2007): 971-980.

15. Billones JB *et al*. "In silico discovery and in vitro activity of inhibitors against Mycobacterium tuberculosis 7,8-diaminopelargonic acid synthase (Mtb BioA)". *Drug Design, Development and Therapy* 11 (2017): 563-574.

16. Billones JB., *et al*. "Toward antituberculosis drugs: in silico screening of synthetic compounds against Mycobacterium tuberculosis l,d-transpeptidase 2". *Drug Design, Development and Therapy* 10 (2016): 1147-1157.

17. Sampaco III AB, Billones JB. "Virtual Screening of Natural Products, Molecular Docking and Dynamics Simulations on M. tuberculosis S-adenosyl-L-homocysteine Hydrolase". *Oriental Journal of Chemistry* 31.4 (2015): 1859-1865.

18. Billones JB *et al*. "Virtual Screening against M. tuberculosis 7, 8-Diaminopelargonic Acid Synthase (MtbBioA) and In Silico Toxicity Evaluation of Top Hits". *Current Enzymes Inhibition* 10.2 (2014): 105-112.

19. Billones JB., *et al*. "Virtual Screening against Mycobacterium tuberculosis Lipoate Protein Ligase B (MtbLipB) and In Silico ADMET Evaluation of Top Hits". *Oriental Journal of Chemistry* 29.4 (2013): 1457-1468.

20. Yang CTM and Billones JB. "Towards Antituberculosis Drugs: Molecular Docking of Curcumin and Its Analogues to Pantothenate Synthetase". *Philippine Journal of Science* 141.2 (2012): 187-196.

21. Uy VCC and Billones JB. "Towards Antituberculosis Drugs: Virtual Screening for Potential Inhibitors of Pantothenate Synthetase of Mycobacterium tuberculosis". *Philippine Science Letters* 5.2 (2012): 122-130.

22. Berman HM., *et al*. "The Protein Data Bank". *Nucleic Acids Research* 28 (2000): 235-242.

Volume 3 Issue 6 June 2019
© All rights are reserved by Junie B Billones., *et al.*