Comparative modeling and molecular refinement of the enzyme *Phosphoglucosamine mutase* from *Mycobacterium tuberculosis*

Jhonatan M. S. Costa[1], Wildrimak S. Pereira[2], Romulo O. Barros[1], Diego S. Mendes[3], Fábio L. C. C. Junior[4], Ricardo M. Ramos[5]

[1]LaPeSI – Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI)
Cep 64.000-040 – Teresina – PI – Brazil

{romulobarros, ricardo}@ifpi.edu.br

Abstract. *Tuberculosis (TB) is the leading cause of death from an infectious agent in the world according to the World Health Organization (WHO) and the epidemic is being strengthened by the emergence of Multi and eXtensively Drug Resistant strains (MDR-TB and XDR-TB), decreasing the treatment options that have been used for decades. The GlmM catalyzes the interconversion of glucosamine-6-phosphate (GlcN6P) to glucosamine-1-phosphate (GlcN1P), catalyzing the second step of the biosynthesis process of UDPGlcNAc (uridine diphospho-N-acetylglucosamine) which is essential for the life cycle of Mycobacterium Tuberculosis (MTB). In the present study, comparative modeling and molecular refinement techniques were used to generate a three-dimensional model of MTB GlmM enzyme.*

1. Introduction

The bacterium *Mycobacterium tuberculosis* (MTB) is also called as Koch bacillus being discovered by Robert Koch in 1882 as a causative agent of tuberculosis, that has become a global pandemic according to World Health Organization [WHO et al. 2018], overcoming HIV. In year 2017, TB could be related to: a) 10 million new cases, 9% of cases of people living with HIV (72% in Africa); b) an estimated 1.3 million deaths and c) an additional 300,000 people with HIV deaths [WHO et al. 2018].

Since the rise of Multi and eXtensively Drug Resistant strains (MDR-TB and XDR-TB), the treatment options used for decades against TB (for example, isoniazid and rifampicin) have decreased in effectiveness, thus strengthening the epidemic TB. With the strengthening of tuberculosis and the efficiency of conventional therapies reduces, new approaches to fight resistant tuberculosis become emphatically necessary [Koul et al. 2011].

In vitro research using *Mycobacterium tuberculosis* indicates that enzymes involved in the *UDP-GlcNAc* (*uridine diphospho-N-acetylglucosamine*) (figure 1) biosynthesis process are essential for the life cycle of MTB. These enzymes include GlmM that catalyses the interconversion of glucosamine-6-phosphate (GlcN6P) to glucosamine-1-phosphate (GlcN1P), catalyzing the second step of the process [Moraes et al. 2015].

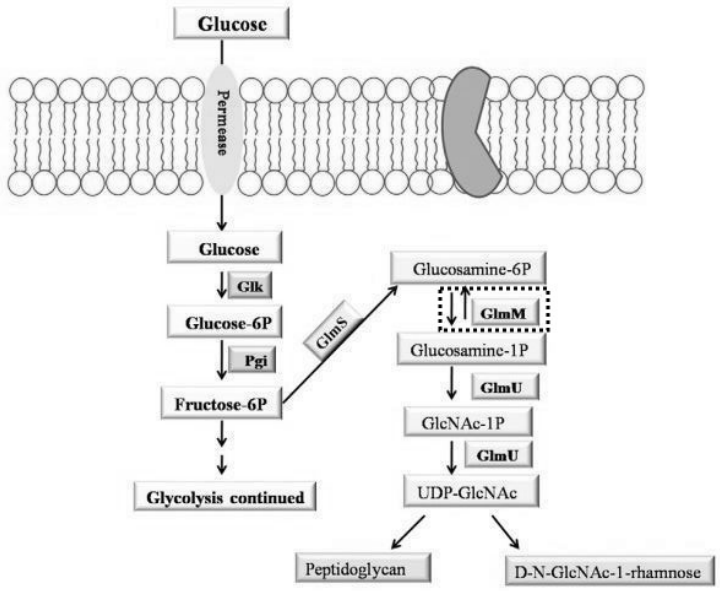

Figure 1. Enzymatic route needed for bacterial growth with the GlmM enzyme highlighted (Adapted from [Rani and Khan 2016]).

The use of computational methods (in silico) in the study and simulation of biochemical and pharmacological properties of bioactive molecules and biological macromolecules has been growing every year [Ekins et al. 2007, kor , mor]. Comparing the traditional methods used in the study of pharmacological and toxicological properties of bioactive compounds with the computational methods, these have the advantage that they require considerably less time and financial cost and have predictive capacity. This work will make an in silico study aiming to generate a new three-dimensional model of GlmM enzyme fundamental for the correct functioning of MTB metabolism, making it possible to perform simulations with inhibitors aiming to inhibit the GlmM enzyme.

2. Methods

2.1. Comparative Modeling

Due to the nonexistence of a three-dimensional model of the enzyme in biological databases, the three-dimensional model of GlmM enzyme was generated by comparative modeling (homology modeling) using modeller, version 9.22 [Eswar et al. 2008]. In the process, the amino acid sequence, previously found (with access code WP_052630404.1) in the National Center for Biotechnology Information database [Sayers et al. 2009], is used to search for homologous sequences available from the Protein Data Bank (PDB) [Berman et al. 2003].

Two template structures were chosen, GlmM from Bacillus anthracis organisms, PDB code 3PDK (2.7 Å resolution, 47.96% sequence identity) and GlmM from Staphylococcus aureus organisms, PDB code 6GYZ (3.0 Å resolution, 47.52% identity) of sequence. Similarity between models and target was quantified with sequence identity and resolution. as well as the statistical measure of E-value. One hundred models were generated, but only the one with the highest negative value for Discrete Optimized Proteins Energy (DOPE) [Plácido et al. 2017].

2.2. Refinement and evaluation of predicted structures

The generated model has been refined using module GalaxyRefine from GalaxyWeb server [Shin et al. 2014]. To determine the predicted structure reliability, the model evaluation was performed. The strategy was the calculation of important parameters of the modeled structure using specific tools. Thus, the Ramachandran graph of the modeled structure was made and validated in PROCHECK [Laskowski et al. 1993]. In addition, stereochemical excellence and overall quality of refined structure was analyzed by Z-scores [Sippl 1993] and ERRAT [Colovos and Yeates 1993], respectively.

3. Results

Table 1 shows structural parameters of selected modeled molecules before (original) and after (refined) refinement as well as for X-ray solved structures used as a template. With respect to the Procheck parameter it is expected that the models have the highest percentage of residues in the favorable and allowed regions (columns Core and Allow, respectively) and G-factor with positive values. Models are expected to have a higher value with respect to the Errat parameter and a lower value with respect to Z-Score. High quality models present parameter values close to those of the template structure.

Table 1. Quality comparison between the template, original e refined models (data obtained from Saves server [Doreleijers et al. 1998]).

Template/Model/Refined Model	Procheck					Errat	Z-Score
	Core[a]	Allow.[b]	Gener.[c]	Disall.[d]	G-factor[e]		
3PDK:A (PDB)	85.0%	13.4%	1.0%	0.5%	-0.18	92.43	**-11.00**
6GYZ:A (PDB)	91.9%	7.6%	0.5%	0.0%	0.20	100	**-10.43**
GlmM:A (original)	90.3%	8.6%	0.5%	0.5%	-0.08	84.77	**-10.36**
GlmM:A (refined)	93.7%	5.5%	0.8%	0.0%	0.13	98.34	**-10.75**

Table 2 shows additional data obtained from the GalaxyRefine server for the same modeled molecules of the table 1. Refined models with values less than or equal to the Poor rotamers parameter and greater or equal to the Rama favored parameter were selected.

Table 2. GalaxyRefine [Shin et al. 2014] server parameters for original and refined models.

Model	GDT-HA[a]	RMSD[b]	MolProbity[c]	Clash score[d]	Poor rotamers[e]	Rama favored[f]
GlmM:A (original)	1.0000	0.000	3.155	74.9	3.7	94.4
GlmM modelo 1:A (refined)	0.9851	0.288	1.692	13.6	0.3	97.8
GlmM:A (original)	1.0000	0.000	3.155	74.9	3.7	94.4
GlmM modelo 2:A (refined)	0.9829	0.319	1.642	13.6	0.3	98.0
GlmM:A (original)	1.0000	0.000	3.155	74.9	3.7	94.4
GlmM modelo 3:A (refined)	0.9867	0.289	1.855	12.3	1.9	97.8
GlmM:A (original)	1.0000	0.000	3.155	74.9	3.7	94.4
GlmM modelo 4:A (refined)	0.9856	0.300	1.811	15.1	0.6	97.3
GlmM:A (original)	1.0000	0.000	3.155	74.9	3.7	94.4
GlmM modelo 5:A (refined)	0.9856	0.308	1.788	13.2	0.9	97.1

In Figure 2, we can see the best three-dimensional model created from table 2 model 2:A.

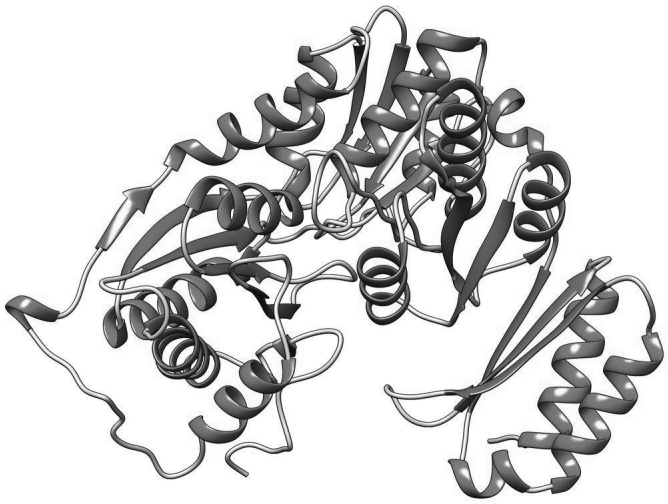

Figure 2. 3D structure of the model obtained by GlmM comparative modeling of MTB in ribbons representation (model generated with UCSF Chimera software [Pettersen et al. 2004]).

4. Discussion

The generation of a new three-dimensional model of MTB GlmM enzyme represents a considerably interesting structure for molecular docking and simulations aiming to neutralizing the *Mycobacterium Tuberculosis*.

Rani et al [Rani et al. 2015] in their work was indicated that a very fruitful way to identify novel anti-tubercular agents is the development of compounds that target the enzymes essentially required for the biosynthesis and assembly of the mycobacterial cell wall. Biosynthesis of uridine diphospho-N-acetylglucosamine (UDP-GlcNAc) represents one such pathway. showing that enzymes involved in UDPGlcNAc biosynthesis are essential for *Mycobacterium tuberculosis* growth.

Moraes [Moraes et al. 2015] indentified that the innermost segment from *Mycobacterium Tuberculosis* of the cell wall is comprised of peptidoglycan, a layer that is required for survival and growth of the pathogen. Enzymes that catalyse biosynthesis of the peptidoglycan are essential and are therefore attractive targets for discovery of novel antibiotics as humans lack similar enzymes making it possible to selectively target bacteria only. In their paper, was have reviewed the structures and functions of enzymes GlmS, GlmM, GlmU, MurA, MurB, MurC, MurD, MurE and MurF from M. tuberculosis that are involved in peptidoglycan biosynthesis. In their paper was created a GlmM homology model of M. Tuberculosis with a single crystallographic mold [Morris and Lim-Wilby 2008]. The model generated in this work was based on two mold structures and refined with better quality and greater sequence identity, thus making possible future studies with molecular docking and molecular dynamic and based in this results be able to identify new possible inhibitors.

Tosi et al [Tosi et al. 2019] performed *in vitro* and *in silico* tests to inhibit the GlmM enzyme of the organism *Staphylococcus aureus*. The strategy was identification of c-di-AMP as important molecule that plays a pivotal role in regulating fundamental cellular processes, including osmotic and cell wall homeostasis. In the opportunis-

tic human pathogen Staphylococcus aureus, c-di-AMP is produced by the membrane-anchored DacA enzyme. The gene coding for DacA is part of the conserved three-gene dacA/ybbR/glmM operon that also encodes the phosphoglucosamine mutase GlmM.

References

Berman, H., Henrick, K., and Nakamura, H. (2003). Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology*, 10(12):980.

Colovos, C. and Yeates, T. O. (1993). Verification of protein structures: patterns of non-bonded atomic interactions. *Protein science*, 2(9):1511–1519.

Doreleijers, J. F., Rullmann, J. A., and Kaptein, R. (1998). Quality assessment of nmr structures: a statistical survey. *Journal of molecular biology*, 281(1):149–164.

Ekins, S., Mestres, J., and Testa, B. (2007). In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British journal of pharmacology*, 152(1):9–20.

Eswar, N., Eramian, D., Webb, B., Shen, M.-Y., and Sali, A. (2008). Protein structure modeling with modeller. In *Structural proteomics*, pages 145–159. Springer.

Koul, A., Arnoult, E., Lounis, N., Guillemont, J., and Andries, K. (2011). The challenge of new drug discovery for tuberculosis. *Nature*, 469(7331):483.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). Procheck: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, 26(2):283–291.

Moraes, G. L., Gomes, G. C., De Sousa, P. R. M., Alves, C. N., Govender, T., Kruger, H. G., Maguire, G. E., Lamichhane, G., and Lameira, J. (2015). Structural and functional features of enzymes of mycobacterium tuberculosis peptidoglycan biosynthesis as targets for drug development. *Tuberculosis*, 95(2):95–111.

Morris, G. M. and Lim-Wilby, M. (2008). Molecular docking. In *Molecular modeling of proteins*, pages 365–382. Springer.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612.

Plácido, A., Coelho, A., Abreu Nascimento, L., Gomes Vasconcelos, A., Fátima Barroso, M., Ramos-Jesus, J., Costa, V., das Chagas Alves Lima, F., Delerue-Matos, C., Martins Ramos, R., et al. (2017). Cry1a (b) 16 toxin from bacillus thuringiensis: Theoretical refinement of three-dimensional structure and prediction of peptides as molecular markers for detection of genetically modified organisms. *Proteins: Structure, Function, and Bioinformatics*, 85(7):1248–1257.

Rani, C. and Khan, I. A. (2016). Udp-glcnac pathway: Potential target for inhibitor discovery against m. tuberculosis. *European Journal of Pharmaceutical Sciences*, 83:62–70.

Rani, C., Mehra, R., Sharma, R., Chib, R., Wazir, P., Nargotra, A., and Khan, I. A. (2015). High-throughput screen identifies small molecule inhibitors targeting acetyltransferase activity of mycobacterium tuberculosis glmu. *Tuberculosis*, 95(6):664–677.

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Federhen, S., et al. (2009). Database resources of the national center for biotechnology information. *Nucleic acids research*, 38(suppl_1):D5–D16.

Shin, W.-H., Lee, G. R., Heo, L., Lee, H., and Seok, C. (2014). Prediction of protein structure and interaction by galaxy protein modeling programs. *Bio Design*, 2(1):1–11.

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins: Structure, Function, and Bioinformatics*, 17(4):355–362.

Tosi, T., Hoshiga, F., Millership, C., Singh, R., Eldrid, C., Patin, D., Mengin-Lecreulx, D., Thalassinos, K., Freemont, P., and Gründling, A. (2019). Inhibition of the staphylococcus aureus c-di-amp cyclase daca by direct interaction with the phosphoglucosamine mutase glmm. *PLoS pathogens*, 15(1):e1007537.

WHO et al. (2018). *Global tuberculosis report 2018*. World Health Organization.