



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS TERESINA CENTRAL
CURSO DE ANÁLISE E DESENVOLVIMENTO DE SISTEMAS**

WILDRIMAK DE SOUZA PEREIRA

**GLUCOSAMINE-6-PHOSPHATE SYNTHASE de *Mycobacterium tuberculosis*: um
estudo *in silico* para predição de um modelo tridimensional refinado**

**TERESINA
2021**

WILDRIMAK DE SOUZA PEREIRA

GLUCOSAMINE-6-PHOSPHATE SYNTHASE de *Mycobacterium tuberculosis*: um estudo *in silico* para predição de um modelo tridimensional refinado

Trabalho de Conclusão de Curso (artigo) apresentado como exigência parcial para obtenção do diploma do Curso de Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central.

Orientador: Prof. Dr. Ricardo Martins Ramos.

GLUCOSAMINE-6-PHOSPHATE SYNTHASE de *Mycobacterium tuberculosis*: um estudo *in silico* para predição de um modelo tridimensional refinado

W. S. Pereira*
J. M. S. Costa**
F. L. C. C. Junior***
R. O. Barros****
R. M. Ramos*****

RESUMO

A Tuberculose é uma das principais causas de morte por um agente infeccioso no mundo, segundo a Organização Mundial de Saúde. Estudos indicam que enzimas envolvidas na biossíntese de uridine diphospho-N-acetylglucosamine são essenciais para o ciclo de vida da bactéria. Uma dessas enzimas é a glucosamine-6-phosphate synthase (GlmS), que não possui estrutura tridimensional disponibilizada nos banco de dados de proteínas na internet. Neste trabalho, métodos de bioinformática estrutural (modelagem comparativa e refinamento molecular) foram utilizados para construir um modelo tridimensional refinado para a enzima GlmS de *Mycobacterium tuberculosis*. O modelo foi gerado utilizando quatro estruturas moldes (cristalográficas). Os resultados obtidos para os parâmetros estereoquímicos e gerais do modelo refinado foram melhores que o modelo original e similares aos das estruturas moldes, validando o modelo refinado.

Palavras-chave: Glucosamine-6-phosphate synthase; *Mycobacterium tuberculosis*; Modelagem comparativa; Refinamento molecular.

ABSTRACT

Tuberculosis is one of the main causes of death by an infectious agent in the world, according to the World Health Organization. Studies indicate that enzymes involved in the biosynthesis of uridine diphosphate-N-acetylglucosamine are essential for the life cycle of the bacterium. One of these enzymes is glucosamine-6-phosphate synthase (GlmS), which does not have a three-dimensional structure available in the protein database on the internet. In this work, structural bioinformatics methods (comparative modeling and molecular refinement) were used to build a refined three-dimensional model for the GlmS enzyme of *Mycobacterium tuberculosis*. The model was generated using four template structures (crystallographic). The results obtained for the stereochemical and general parameters of the refined model were better than the original model and similar to those template structures, validating the refined model.

Keywords: Glucosamine-6-phosphate synthase; *Mycobacterium tuberculosis*; Comparative modeling; Molecular refinement.

Data de aprovação: 05/02/2021 (data de apresentação do TCC).

* Discente do Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI). E-mail: wildrimak@gmail.com. Currículo: <https://orcid.org/0000-0003-3801-8022>.

** Discente do Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI). E-mail: jhonmscosta@gmail.com. Currículo: <https://orcid.org/0000-0003-2235-3475>.

*** Discente do Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI). E-mail: fabio.cardosojr@gmail.com. Currículo: <https://orcid.org/0000-0002-2859-3982>.

**** Servidor no Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI). E-mail: romulobarros@ifpi.edu.br. Currículo: <https://orcid.org/0000-0002-0000-033X>.

***** Docente no Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI). E-mail: ricardo@ifpi.edu.br. Currículo: <https://orcid.org/0000-0003-2016-3344>.

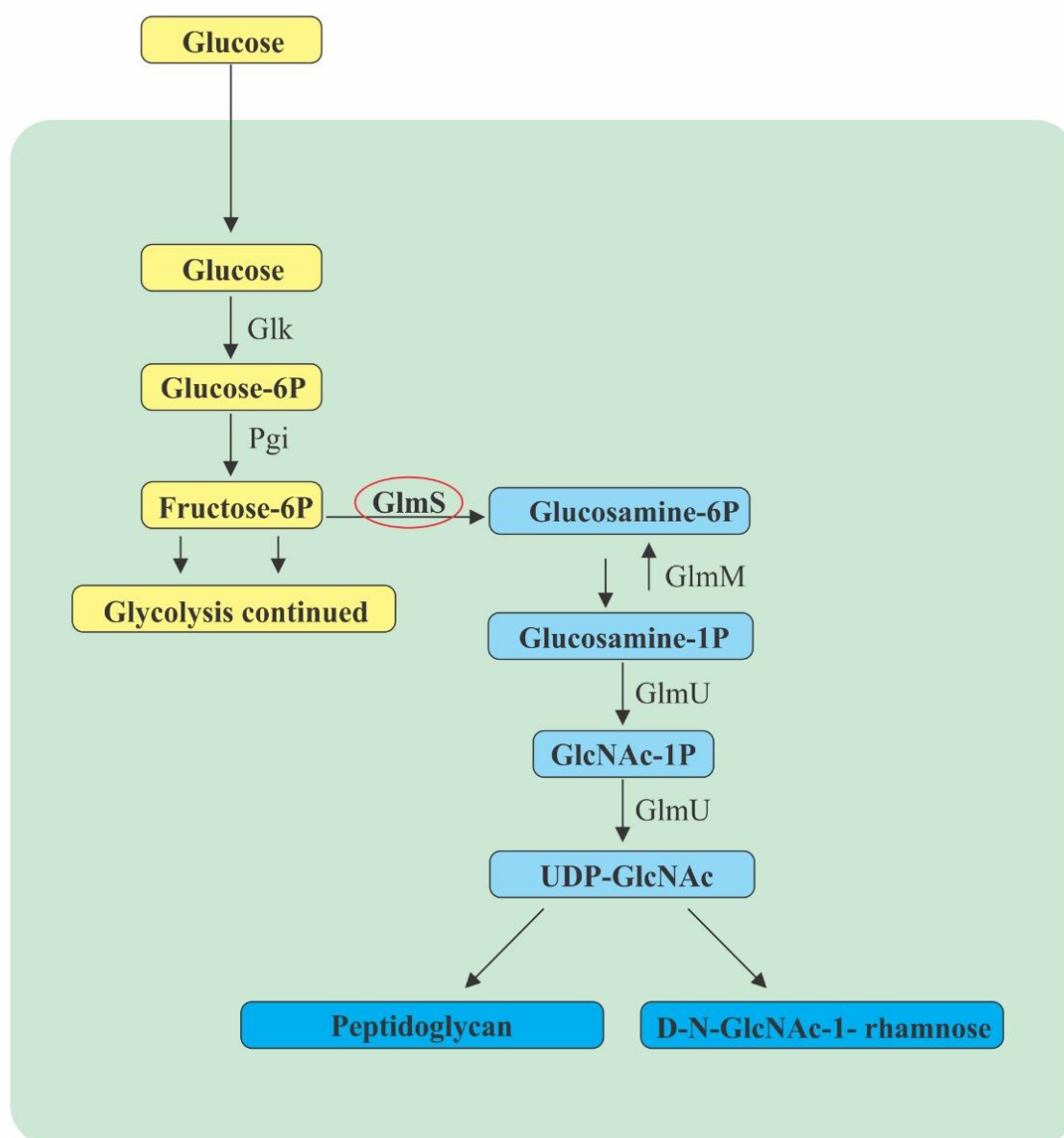
1 INTRODUÇÃO

Em 1882, Robert Koch identificou o *Mycobacterium tuberculosis* (MTB) como o agente causador da tuberculose (TB), mas desde a sua descoberta esta epidemia global parece inabalável (KOUL *et al.*, 2011). Segundo dados da Organização Mundial de Saúde (OMS), no ano de 2019 estima-se que 8,9 a 11,0 milhões de pessoas tenham adoecido com tuberculose em 2019, um número que tem diminuído lentamente nos últimos anos. Estima-se que tenha havido de 1,1 a 1,3 milhões de mortes por TB entre as pessoas soronegativas em 2019 (uma redução de 1,7 milhões em 2000), e mais 208.000 mortes entre as pessoas soropositivas (uma redução de 678.000 em 2000). Os homens (com idade maior que 15 anos) representavam 56% das pessoas que desenvolveram tuberculose em 2019; as mulheres representavam 32% e as crianças 12% (com menos de 15 anos). Entre todas as pessoas afetadas, 8,2% viviam com o HIV. Geograficamente, segundo a OMS, a maioria das pessoas que desenvolveram tuberculose em 2019 encontravam-se no Sudeste Asiático (44%), na África (25%) e no Pacífico Ocidental (18%), com percentagens menores no Mediterrâneo Oriental (8,2%), nas Américas (2,9%) e na Europa (2,5%). Oito países representaram dois terços do total global: Índia (26%), Indonésia (8,5%), China (8,4%), o Filipinas (6,0%), Paquistão (5,7%), Nigéria (4,4%), Bangladesh (3,6%) e África do Sul (3,6%). Os outros 22 países, da lista da OMS de 30 países, com elevada incidência de TB representam 21% do total global. (WHO, 2020).

Nos últimos anos, a epidemia de TB tem se fortalecido pelo surgimento de cepas Multi e extensivamente Resistente a Drogas (MDR-TB e XDR-TB), diminuindo as opções de tratamento que são utilizadas há décadas (por exemplo, com isoniazida e rifampicina). Para complicar ainda mais a situação, interações medicamentosas impedem a coadministração de alguns fármacos para TB com certos fármacos de tratamentos anti-HIV e outros usados por diabéticos (KOUL *et al.*, 2011).

Estudos indicam que enzimas envolvidas na biossíntese de uridine diphospho-N-acetylglucosamine (UDP-GlcNAc), sejam essenciais para o crescimento de *M. tuberculosis in vitro* (RANI; KHAN, 2016). Uma dessas enzimas é a glucosamine-6-phosphate synthase (GlmS) como mostra a Figura 1. A inibição desta enzima impede a produção de UDP-GlcNAc e leva a bactéria a morte.

Figura 1 - Via enzimática necessária para o crescimento bacteriano de *Mycobacterium tuberculosis*



Fonte: Adaptado de (RANI; KHAN, 2016).

A determinação da estrutura 3D de uma dada proteína representa etapa fundamental para o entendimento de sua função e de seu mecanismo de ação. O modelo estrutural de uma proteína pode ser obtido através de técnicas experimentais tais como cristalografia de raios-X,

ressonância magnética nuclear (RMN) e microscopia eletrônica (HUANG *et al.*, 2010). Tais métodos, contudo, normalmente requerem recursos estruturais e financeiros elevados. O emprego de métodos computacionais (*in silico*) no estudo e simulação de propriedades bioquímicas e farmacológicas de moléculas bioativas e macromoléculas biológicas tem crescido a cada ano (KORTAGERE, EKINS, 2010; MOROY *et al.*, 2012). Comparando os métodos tradicionais usados no estudo de propriedades farmacológicas e toxicológicas de compostos bioativos com os métodos computacionais, estes apresentam a vantagem de exigirem tempo e custo financeiro consideravelmente menores, além de possuírem em muitos casos, capacidade preditiva.

Uma vez que não há estruturas cristalográficas disponíveis da GlmS de *M. tuberculosis* nos bancos de dados de proteínas, o objetivo do presente trabalho foi gerar (prever) um modelo 3D refinado para essa enzima utilizando ferramentas de Bioinformática Estrutural (modelagem comparativa e refinamento molecular).

2 MATERIAIS E MÉTODOS

2.1 Modelagem comparativa (por homologia)

O modelo 3D da enzima GlmS de *M. tuberculosis* foi gerado por modelagem comparativa usando o software Modeller versão 9.22 (ESWAR *et al.*, 2008) através da sua interface gráfica integrada ao software UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/>) (PETTERSEN *et al.*, 2004). No processo, a sequência primária de aminoácidos da GlmS, previamente encontrada (com código de acesso GI 614087054) no National Center for Biotechnology Information (NCBI) (SAYERS *et al.*, 2009), foi usada para procurar sequências homólogas disponíveis no Protein Data Bank (PDB) (BERMAN *et al.*, 2003). Quatro estruturas templates (moldes) foram escolhidas, todas GlmS do organismo da *Escherichia coli*, com códigos PDB, resolução e identidade de sequência respectivamente de: (1MOQ, 1,57Å, 45,8%); (1XFF, 1,8Å, 46%); (2VF5, 2,9Å, 45,1%) e (4AMV 2,05Å, 45,2%). Foram gerados cem modelos e o modelo com o menor valor de Discrete Optimized Proteins Energy (DOPE) foi escolhido para refinamento molecular (PLÁCIDO *et al.*, 2017).

2.2 Refinamento molecular

O modelo gerado foi refinado usando o módulo GalaxyRefine do servidor GalaxyWeb (SHIN *et al.*, 2014). Logo após esta etapa os modelos original e refinado foram avaliados. Os parâmetros estereoquímicos foram analisados utilizando os softwares PROCHECK (gráfico de Ramachandran) (LASKOWSKI *et al.*, 1993) e ERRAT (COLOVOS; YEATES, 1993), implementados no servidor SAVES (Structural Analysis and Verification Server - (<https://servicesn.mbi.ucla.edu/SAVES/>)). Além disso, a qualidade geral das estruturas dos modelos original e refinado foram analisadas pelo serviço Z-score (<https://prosa.services.came.sbg.ac.at/prosa.php>) (SIPPL, 1993).

3 RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os valores dos parâmetros estereoquímicos (estrutura 3D) para os templates (estruturas cristalográficas), o modelo original (obtido por modelagem comparativa)

e o modelo refinado. Com relação aos parâmetros do Procheck é esperado que os valores tenham o maior percentual de resíduos nas regiões favoráveis (Core) e permitidas (Allow). Valores positivos são esperados para o parâmetro G-Factor. Espera-se que os modelos apresentem um alto valor para o parâmetro Errat e um baixo valor para Z-Score.

Um modelo de qualidade apresenta valores próximos das suas estruturas templates. Os valores apresentados para o modelo refinado da GlmS de MTB apresentaram-se similares aos das estruturas templates. Observa-se na Tabela 1 que todos os valores dos parâmetros do modelo refinado são melhores comparados ao modelo original. O modelo refinado apresentou 98,9% de resíduos nas regiões Core e Allow, enquanto o modelo original apresentou um valor de 98,2%. O modelo refinado diminuiu a porcentagem de resíduos nas regiões desfavoráveis (Gener e Disall) 1,1% enquanto o modelo original apresenta 1,8%. O modelo refinado apresenta melhores resultados para os parâmetros Errat e Z-Scores comparados ao modelo original (89,92 e 75,97; -11,11 e -10,78), respectivamente. Esses resultados mostram a importância do refinamento molecular para a qualidade final do modelo refinado. O modelo refinado utilizado na Tabela 1 corresponde ao modelo refinado 1 (modelo escolhido após o refinamento) especificado na Tabela 2.

Tabela 1 - Comparação da qualidade entre os templates, o modelo original e o modelo refinado da GlmS.

Model	Procheck					Errat	Z-Score
	Core	Allow	Gener	Disall	G-factor		
1MOQ_A	92,9%	7,1%	0,0%	0,0%	0,09	99,71	-8,56
2VF5_X	80,7%	18,4%	0,6%	0,3%	-0,33	93,53	-8,54
1XFF_A	93,7%	5,8%	0,0%	0,5%	-0,03	96,94	-8,72
4AMV_A	90,0%	9,1%	0,5%	0,3%	-0,10	97,82	-11,82
Modelo original	90,1%	8,1%	1,1%	0,7%	-0,10	75,97	-10,78
Modelo Refinado	92,5%	6,4%	0,9%	0,2%	0,04	89,92	-11,11

A Tabela 2 apresenta os parâmetros de refinamento obtidos do Servidor GalaxyRefine. O modelo original foi submetido ao Servidor GalaxyRefine que retorna cinco modelos refinados. Dentre os parâmetros listados na Tabela 2, “Poor rotamers” (menores valores são desejados) e “Rama favored” (maiores valores são desejados) foram analisados para a escolha do modelo refinado. O modelo refinado 1 foi escolhido por apresentar menor valor (1,0) para “Poor rotamers” e maior valor (96,3) para “Rama favored”. O valor de RMSD igual a 0,321 mostra que os dois modelos (original e refinado 1) são similares.

Tabela 2 - Parâmetros de refinamento para o modelo original e refinados.

Model	GDT- HA^a	RMSD^b	MolProbity^c	Clash score^d	Poor rotamers^e	Rama favored^f
Modelo original	1,0000	0,000	3,293	114,5	2,9	93,6
Modelo refinado 1	0,9796	0,321	2,063	20,6	1,0	96,3
Modelo refinado 2	0,9848	0,308	2,174	21,8	1,0	95,2
Modelo refinado 3	0,9820	0,310	2,181	18,3	1,4	95,7
Modelo refinado 4	0,9804	0,320	2,115	18,2	1,2	95,8
Modelo refinado 5	0,9808	0,314	2,195	18,4	1,4	95,5

^aGlobal Distance Test - High Accuracy; ^bRoot-mean-square deviation; ^cMolProbity, ^dClashscore; ^ePoor rotamers (CHEN *et al.*, 2010); ^fRamachandran favored

Moraes *et. al.* (2012) criaram um modelo para a GlmS de MTB por modelagem comparativa usando a estrutura cristalográfica PDB ID: 4AMV como molde. Neste trabalho além da 4AMV foram utilizadas mais três estruturas moldes para a criação do modelo. Logo após, o modelo escolhido, entre os cem modelos gerados, foi refinado. O modelo refinado é composto por 624 resíduos de aminoácidos, 23 alfa-hélices e 27 folhas-beta (Figura 2).

Rani e Khan (2016) apresentaram uma visão geral de múltiplas enzimas (entre elas a GlmS) catalisadoras de processos químicos que são fundamentais para o ciclo de vida da MTB. Além disso, apresentaram informações (de estudos *in vitro*) sobre inibidores para estas enzimas. Estudos *in silico* (ancoramento e dinâmica molecular) com estes inibidores estão sendo realizados com o modelo refinado da enzima GlmS de MTB por este grupo de pesquisa.

Figura 2 - Estrutura global do modelo refinado da GlmS de MTB. (A) representação em fitas e cores em estrutura secundária (B) representação em superfície e cores em hidrofobicidade.



Fonte: Imagem gerada utilizando o software UCSF Chimera.

4 CONCLUSÃO

O modelo obtido por modelagem comparativa (por homologia) para a enzima Glucosamine-6-phosphate Synthase (GlmS) de *M. tuberculosis* foi submetido a refinamento molecular através do Servidor GalaxyWeb. O modelo refinado (01) com melhores resultados para os parâmetros “Poor rotamers” e “Rama favored” foi escolhido. O modelo refinado apresentou resultados melhores que o modelo gerado por modelagem comparativa em todos os parâmetros (qualidade estereoquímica e estrutura geral) analisados neste estudo, além disso, apresentou resultados similares as estruturas cristalográficas. Estes resultados validam a estrutura tridimensional do modelo refinado. Estudos *in silico* podem ser realizados com este modelo para descoberta de novos candidatos a fármaco para a tuberculose resistente.

REFERÊNCIAS

- BERMAN, H.; HENRICK, K.; NAKAMURA, H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology*, [S.L.], v. 10, n. 12, p. 980-980, dez. 2003. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/nsb1203-980>
- CHEN, V. B.; ARENDALL, W. B.; HEADD, J. J.; KEEDY, D. A.; IMMORMINO, R. M.; KAPRAL, G. J.; RICHARDSON, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, v. 66, n. 1, p. 12-21, jan. 2010.
- COLOVOS, C.; YEATES, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science*, [S.L.], v. 2, n. 9, p. 1511-1519, set. 1993. Wiley. <http://dx.doi.org/10.1002/pro.5560020916>
- ESWAR, N.; ERAMIAN, D.; WEBB, B.; SHEN, M.; SALI, A. Protein Structure Modeling with MODELLER. *Methods in Molecular Biology*, [S.L.], p. 145-159, mai, 2008. Humana Press. http://dx.doi.org/10.1007/978-1-60327-058-8_8
- HUANG, H. J.; YU, H. W.; CHEN, C. Y.; HSU, C. H.; CHEN, H. Y.; LEE, K. J.; CHEN, C. Y. C. Current developments of computer-aided drug design. *Journal of the Taiwan Institute of Chemical Engineers*, v. 41, n. 6, p. 623-635, mar. 2010.
- KORTAGERE, S.; EKINS, S. Troubleshooting computational methods in drug discovery. *Journal of Pharmacological And Toxicological Methods*, [S.L.], v. 61, n. 2, p. 67-75, mar. 2010. Elsevier BV. <http://dx.doi.org/10.1016/j.vascn.2010.02.005>
- KOUL, A.; ARNOULT, E.; LOUNIS, N.; GUILLEMONT, J.; ANDRIES, K. The challenge of new drug discovery for tuberculosis. *Nature*, [S.L.], v. 469, n. 7331, p. 483-490, jan. 2011. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/nature09657>
- LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, v. 26, n. 2, p. 283-291, abr. 1993.
- MORAES, G. L.; GOMES, G. C.; SOUSA, P. R. M.; ALVES, C. N.; GOVENDER, T.; KRUGER, H. G.; MAGUIRE, G. E. M.; LAMICHHANE, G.; LAMEIRA, J. Structural and functional features of enzymes of *Mycobacterium tuberculosis* peptidoglycan biosynthesis as targets for drug development. *Tuberculosis*, [S.L.], v. 95, n. 2, p. 95-111, mar. 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.tube.2015.01.006>
- MOROY, G.; MARTINY, V. Y.; VAYER, P.; VILLOUTREIX, B. O.; MITEVA, M. A. Toward in silico structure-based ADMET prediction in drug discovery. *Drug Discovery Today*, [S.L.], v. 17, n. 1-2, p. 44-55, jan. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.drudis.2011.10.023>
- PETTERSEN, E. F.; GODDARD, T. D.; HUANG, C. C.; COUCH, G. S.; GREENBLATT, D. M.; MENG, E. C.; FERRIN, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, v. 25, n. 13, p. 1605-1612, jul.

2004.

PLÁCIDO, A.; COELHO, A.; NASCIMENTO, L. A.; VASCONCELOS, A. G.; BARROSO, M. F.; RAMOS-JESUS, J.; COSTA, V.; LIMA, F. C. A.; DELERUE-MATOS, C.; RAMOS, R. M. Cry1A(b)16 toxin from *Bacillus thuringiensis*: theoretical refinement of threedimensional structure and prediction of peptides as molecular markers for detection of genetically modified organisms. *Proteins: Structure, Function, and Bioinformatics*, [S.L.], v. 85, n. 7, p. 1248-1257, 18 abr. 2017. Wiley. <http://dx.doi.org/10.1002/prot.25285>

RANI, C.; KHAN, I. A. UDP-GlcNAc pathway: Potential target for inhibitor discovery against *M. tuberculosis*. *European Journal of Pharmaceutical Sciences*, v. 83, p. 62-70, fev. 2016.

SAYERS, E. W.; BARRETT, T.; BENSON, D. A.; BOLTON, E.; BRYANT, S. H.; CANESE, K.; CHETVERNIN, V.; CHURCH, D. M.; DICUCCIO, M.; FEDERHEN, S. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, [S.L.], v. 39, p. D38-D51, 21 nov. 2010. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkq1172>

SHIN, W. H.; LEE, G. R.; HEO, L.; LEE, H.; SEOK, C. Prediction of protein structure and interaction by GALAXY protein modeling programs. *Biodesign*, v. 2, p. 1-11, mar. 2014.

SIPPL, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins: Structure, Function, and Genetics*, [S.L.], v. 17, n. 4, p. 355-362, dez. 1993. Wiley. <http://dx.doi.org/10.1002/prot.340170404>

WORLD HEALTH ORGANIZATION (WHO). Global tuberculosis report. 2020. Disponível em: <https://apps.who.int/iris/handle/10665/336069>. Acesso em: 11 jan. 2020.