SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Development of Tertiary Structure of Laccase from *Geobacillus Uzenensis* Using Homology Modelling

Thurgaah Balakrishnan[1], Ragheed Hussam Youseif[2], Nurulbahiyah Ahmad Khairudin[1,*]

[1] Chemical Energy Conversion and Applications (ChECA) iKohza, Department of Chemical and Environmental Engineering (ChEE), Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Malaysia
[2] Al-Farahidi University College of Medical Technology, Baghdad, Iraq

ABSTRACT

Laccases are a versatile enzyme with immense potential in various biotechnological applications. In this study, the protein structure prediction of laccase from Geobacillus uzenensis, a bacteria known for its robust enzymatic activity is focused. The protein structural elucidation of this laccase is crucial to understand its substrate binding and overall functionality. Protein structure can be determined by computational prediction, which is commonly known as homology modelling. The three-dimensional (3D) model was built based on the available crystal structures of related laccase enzymes and aligned with the primary sequence of the target protein. The resulting model exhibited a high degree of structural similarity and conservation of key catalytic residues, supporting its reliability for further analyses. Firstly, the suitable template was identified with 85.61% of sequence identity determined by sequence alignment. Then, MODELLER program was used to predict the model using the method of satisfaction of spatial restraints. The model was then analyzed for its quality by computational analysis tools such as Ramachandran's Plot. Finally, the binding site of the protein was identified using the GRaSP computational tool. These findings provide significant insight on comprehensive analysis of the laccase from G. uzenensis, by providing a structural framework to unravel its functional properties and substrate specificity.

Keywords:
Laccase; homology modelling; protein structure prediction

1. Introduction

Laccase, a versatile enzyme can be found in fungi, plants, bacteria, and insects. They are classed as benzenediol oxygen reductases (EC 1.10.3.2). Unlike the other oxidases there is no toxic chemical such as hydrogen peroxide involved in laccase enzymatic reaction [1]. Due to their low substrate specificity, they are regarded versatile enzymes capable of oxidizing a vast number of phenolic and non-phenolic compounds [1, 2]. The enzyme is regarded environmentally friendly since it requires molecular oxygen as a co-substrate for catalysis and produces only water as a by-product [1, 2]. One of the functions of laccase is in removing environmental pollutants through a process called bioremediation due to its structural and functional properties [3-5]. A significant increase in demand for commercial laccases is predicted

* *Corresponding author.*
E-mail address: r-bahiah@utm.my

https://doi.org/10.37934/araset.62.1.151157

in the upcoming years as laccase has been examined and employed as a green catalytic agent in production lines [4]. Laccase from different sources might have different function and characteristic. In this research, laccase from the bacteria named Geobacillus uzenensis was chosen as the subject of study. Analyzing the protein's tertiary structure of laccase helps identify its function. A protein's biological function is dictated by the arrangement of the atoms in the 3D structure. Having a protein structure provides a greater level of understanding of how a protein works, which can allow us to create hypotheses about how to affect it, control it, or modify it.

As of now, there was still no exact study on the three- dimensional (3D) tertiary protein structure of laccase from G. uzenensis. In other words, the 3D tertiary structure of this protein is still unsolved either experimentally or computationally.

Homology modeling is an effective method when experimental techniques like Nuclear Magnetic Resonance and X-ray crystallography are laborious and cost prohibitive. Homology modeling is a reliable technique for predicting protein structures based on known sequences [6, 7]. Therefore, homology modelling was chosen to carry out this research which is to develop the 3D tertiary structure of laccase from G. uzenensis. Binding site is one of the important sites in a protein structure. The location of the binding site that might be important in the enzyme's function will then be identified using pocket detection algorithms. There are three objectives in this research that will be focused on such as development of 3D tertiary structure of G. uzenensis, evaluating the quality of the developed 3D model and identifying the binding sites in the developed model by using several computational tools.

2. Methodology

The amino acid sequence of the laccase from G. uzenensis, known as the protein target, was retrieved from the database of National Centre for Biotechnology Information (NCBI). Protein structure prediction is one of the crucial steps that was carried out in this research. It consists of four key steps which are template identification and selection, sequence alignment, model building, and model evaluation as shown in Figure 1.

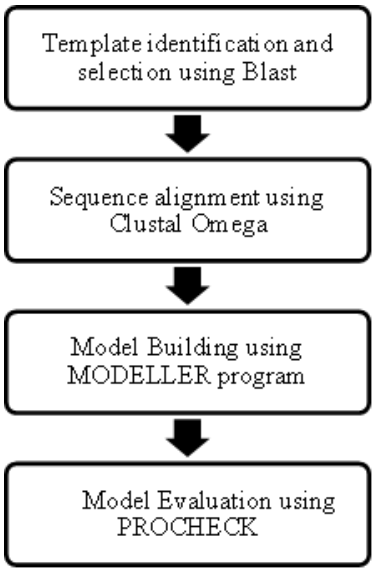

Fig. 1. Flowchart of key steps in homology modelling

The first step in the process was to upload the amino acid sequence of laccase from G. uzenensis into the BLAST program [8]. The Clustal Omega program was employed for sequence alignment [9]. The MODELLER program was used to generate protein models using the method of satisfaction of spatial restraints [10]. Either distances or optimization techniques were used to satisfy the spatial constraints. MODELLER requires 3 input files which were the sequence alignment file between the target and template, crystal structure of template protein in PDB format and a script file. To acquire more precise results one hundred models were obtained from the MODELLER where each model contains a different energy value.

The 3D structure of the model with the lowest energy will then be created using BIOVIA Discovery Studio Visualizer software where the PDB of the lowest energy value model was used as an input file. PROCHECK program was used to evaluate the quality of developed model [11]. The PDB file of G. uzenensis was uploaded to the PROCHECK to obtain the Ramachandran plot, from which the comparison could be carried out between the generated model with its original structure. The binding sites of the developed tertiary model were predicted using the program GRaSP (Graph- based Searching of Protein) [12]. GRaSP is a computational method specifically designed for predicting ligand binding sites in protein structures. It utilizes a graph-based approach to analyse the protein structure and identify potential binding pockets or sites.

3. Results

The suitable template for protein G. uzenesis was selected from the result of BLAST program which was hydrolase enzyme from Geobacillus stearothermophilus (PDB ID: 6T0Y) with sequence identity of 85.61% compared to that of the sequence of G. uzenensis. Figure 2 shows the sequence alignment between these two proteins.

```
target      -MPDIFQQMARGWLRCEASPFAGAIAGMTTKQGGESKGPFASLNMGLHVGDDRTAVVNNR    59
template    GMPDIFQQEARGWLRCGAPPFAGAVAGLTTKHGGESKGPFASLNMGLHVGDDRTDVVNNR    60
             ******* ******* * *****:**:***:********************** *****

target      RRLAEWLAFPLDDWVCCEQVHGAVIRKVTKSDRGSGAHDFAAAIRGADGLYTDEAGVLLA    119
template    RRLAEWLAFPLERWVCCEQVHGADIQKVTKSDRGNGAQDFATAVPGVDGLYTDEAGVLLA    120
            **********: ********* *:********.**:***:*: *.*************

target      LCFADCVPVYFLAPSAGLVGLAHAGWRGTAGGIAKNMVRLWQEQERIAPTDMYAAIGPAI    179
template    LCFADCVPIYFVAPSAGLVGLAHAGWRGTAGGIAGHMVWLWQTREHIAPSDIYVAIGPAI    180
            ********:**:********************** :** *** :*:***:*:*.******

target      GPCCYTVDDRVINGLRSILPAGSPLPWRETSPGQYALDLKEANRLQLIAAGVPDRHIYVS    239
template    GPCCYTVDDRVVDSLRPTLPPESPLPWRETSPGQYALDLKEANRLQLLAAGVPNSHIYVS    240
            **********::.**  **  ************************:*****: *****

target      ERCTSCEETLFFSHRRDRGTTGRMLAFIGRREG--       272
template    ERCTSCEEALFFSHRRDRGTTGRMLAFIGRREEWT       275
            ********:**********************
```

Fig. 2. The sequence alignment of target and template

The three-dimensional tertiary structure was built using MODELLER applying the methods of satisfying the spatial restraints. A total of 100 models were generated in which the model with the lowest energy (1428.533 kcal/mol) was selected for further analysis as it has high stability and no bumping effect [13, 14]. The energy value refers to the measure of free energy within the generated model structure. High energy value means that it has higher free energy which cause a bumping effect that happens between molecules in the model and consequently lowered the stability of a model. So, it was more preferable to choose a model with lower energy value for further analysis as it has high stability and no bumping effect.

The 3D tertiary structure of the selected model was generated using the BIOVIA Discovery Studio Visualizer as shown in Figure 3. Since the structure of the template protein is known, it is assumed that the spatial arrangement of the target protein's residues resembles that of the aligned residues in the template. As a result, the target protein receives the spatial restrictions from the template during the sequence alignment [15, 16]. The general folding pattern and the spatial arrangement of its secondary structure elements which are the alpha-helices, beta-sheets, and coil can be found in Figure 2. It is the structure that determines the protein's unique shape and function [17].

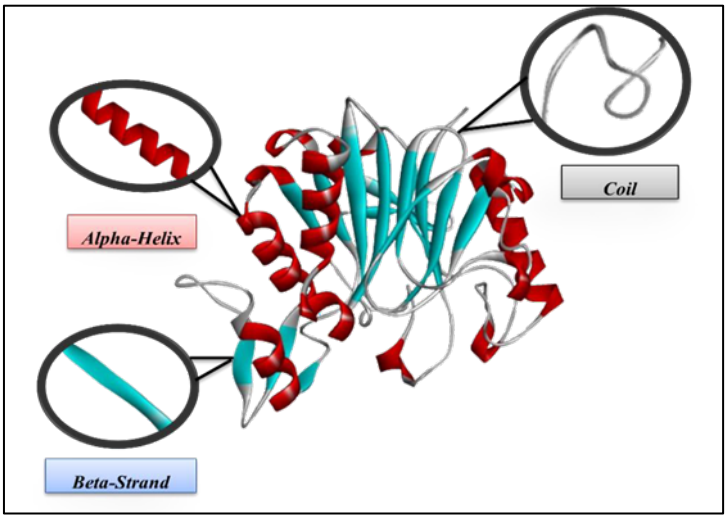

Fig. 3. 3D tertiary structure of G. uzenensis

The quality of the developed model must be taken into consideration. Evaluation of the developed model was performed using PROCHECK program. This program focuses on the backbone dihedral angles of the structure φ (phi angle) against ψ (psi angle) of selected as shown in Figure 4.

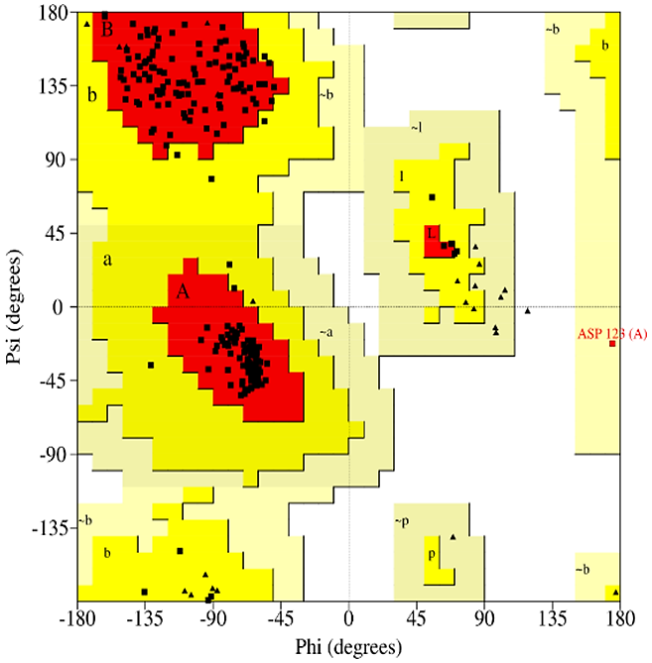

Fig. 4. Ramachandran plot of developed tertiary structure of G. uzenensis

Around 93.8 % (252 residues) of the amino acid residues are located in the most favorable core region, which is the red region, 5.8 % (17 residues) are located in the allowed region which is the yellow region. 0.4 % (2 residues) are located in generously allowed region, and 0.2% (1 residues) are located in disallowed region which is the white region. Mostly, the amino acid residues were found in the core regions of Ramachandran Plot which confirms that the final structure is highly reliable in terms of stereochemistry of the amino acids. This result was notable because of the large percentage of residues in the preferred location (>90%) [18]. Thus, the Ramachandran Plot for the model confirms that the tertiary structure of the target protein, G. uzenensis, is high in quality.

Figure 5 shows the 3D tertiary structure of laccase *with* its binding site identified in green and red color of surface presentation. The binding site plays a crucial role in understanding protein function and designing targeted therapies. The binding site refers to the region on a protein's surface where it interacts with other molecules, such as ligands, substrates, or other proteins. Predicting the binding site in protein structures can serve several purposes such as understanding protein structure, drug discovery and design, protein-protein interaction, and others.

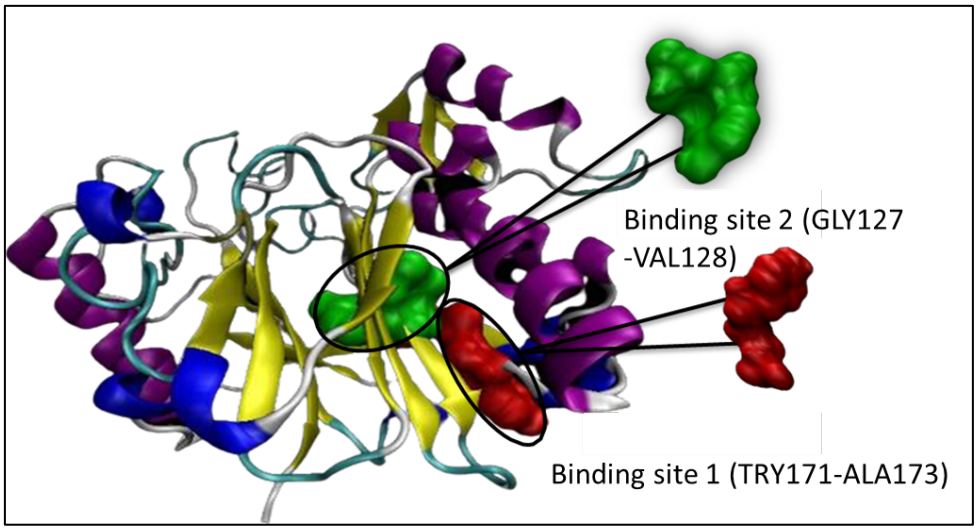

Fig. 5. G. uzenensis binding sites in tertiary structure

The binding site analysis in GRaSP predicted that residues Tyrosine (TYR) 171, Alanine (ALA) 172, and Alanine (ALA) 173 and Glycine (GLY) 127 and Valine (VAL) 128 were found to be involved in the binding site of G. uzenensis as shown in Figure 5. Since only a few residues actually participate in binding the ligand, the other residues usually act as a framework to provide correct conformation and orientation [19]. Thus, the predicted binding site analysis has supplied useful information about probable binding sites within the protein of interest. The binding site prediction approach used yielded encouraging results in identifying locations likely to interact with ligands or other compounds [20, 21].

4. Conclusions

As a conclusion, the protein structure prediction by using computational methods is a reliable technique that can be conducted in a short time of operation and the result is as reliable as the experimental methods. Moreover, the cost of operation is not as costly as the experimental method which needs to use several equipment to get the result. Thus, homology modelling is the best method to predict the protein structure as the alternative of experimental method. The 3D Tertiary structure of laccase from G. uzenensis was successfully predicted using this method. As a recommendation,

the utilisation of hybrid modelling such as ab initio modelling, molecular dynamic simulations, or machine learning algorithms can be considered for future studies. Integrating several approaches can improve model reliability and precision while providing a more comprehensive understanding of protein structure and its function.

References

[1] Mäkelä, Miia R., Marja Tuomela, Annele Hatakka, and Kristiina Hildén. "Fungal laccases and their potential in bioremediation applications." *Laccases in bioremediation and waste valorisation* (2020): 1-25. https://doi.org/10.1007/978-3-030-47906-0_1

[2] Prajapati, Hiren V., Farida P. Minocheherhomji, and R. Scholar. "Laccase-a wonder molecule: A review of its properties and applications." *International. Journal Pure & Applied Bioscience* 6 (2018): 766-773. https://doi.org/10.18782/2320-7051.6233

[3] Janusz, Grzegorz, Anna Pawlik, Urszula Świderska-Burek, Jolanta Polak, Justyna Sulej, Anna Jarosz-Wilkołazka, and Andrzej Paszczyński. "Laccase properties, physiological functions, and evolution." *International journal of molecular sciences* 21, no. 3 (2020): 966. https://doi.org/10.3390/ijms21030966

[4] Hussain, Asim, Muhammad Bilal, Hamza Rafeeq, Zara Jabeen, Nadia Afsheen, Farooq Sher, Vineet Kumar, Ram Naresh Bharagava, Luiz Fernando Romanholo Ferreira, and Hafiz MN Iqbal. "Role of laccase in the pulp and paper industry." In *Nanotechnology in Paper and Wood Engineering*, pp. 35-60. Elsevier, 2022. https://doi.org/10.1016/B978-0-323-85835-9.00006-4

[5] Osma, Johann F., José L. Toca-Herrera, and Susana Rodríguez-Couto. "Uses of laccases in the food industry." *Enzyme Research* 2010, no. 1 (2010): 918761. https://doi.org/10.4061/2010/918761

[6] Kufareva, Irina, and Ruben Abagyan. "Methods of protein structure comparison." *Homology Modeling: Methods and Protocols* (2012): 231-257. https://doi.org/10.1007/978-1-61779-588-6_10

[7] Fiser, Andras. "Template-based protein structure modeling." *Computational biology* (2010): 73-94. https://doi.org/10.1007/978-1-60761-842-3_6

[8] Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. "Basic local alignment search tool." *Journal of molecular biology* 215, no. 3 (1990): 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

[9] Sievers, Fabian, Andreas Wilm, David Dineen, Toby J. Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez et al. "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." *Molecular Systems Biology* 7, no. 1 (2011): 539. https://doi.org/10.1038/msb.2011.75

[10] Šali, Andrej, and Tom L. Blundell. "Comparative protein modelling by satisfaction of spatial restraints." *Journal of molecular biology* 234, no. 3 (1993): 779-815. https://doi.org/10.1006/jmbi.1993.1626

[11] Laskowski, Roman A., Malcolm W. MacArthur, David S. Moss, and Janet M. Thornton. "PROCHECK: A program to check the stereochemical quality of protein structures." *Journal of applied crystallography* 26, no. 2 (1993): 283-291. https://doi.org/10.1107/S0021889892009944

[12] Santana, Charles A., Sabrina de A. Silveira, João PA Moraes, Sandro C. Izidoro, Raquel C. de Melo-Minardi, Antonio JM Ribeiro, Jonathan D. Tyzack, Neera Borkakoti, and Janet M. Thornton. "GRaSP: a graph-based residue neighborhood strategy to predict binding sites." *Bioinformatics* 36, no. Supplement_2 (2020): i726-i734. https://doi.org/10.1093/bioinformatics/btaa805

[13] Valanciute, Audrone, Lasse Nygaard, Henrike Zschach, Michael Maglegaard Jepsen, Kresten Lindorff-Larsen, and Amelie Stein. "Accurate protein stability predictions from homology models." *Computational and Structural Biotechnology Journal* 21 (2023): 66-73. https://doi.org/10.1016/j.csbj.2022.11.048

[14] Makigaki, Shuichiro, and Takashi Ishida. "Sequence alignment generation using intermediate sequence search for homology modeling." *Computational and Structural Biotechnology Journal* 18 (2020): 2043-2050. https://doi.org/10.1016/j.csbj.2020.07.012

[15] Bongirwar, Vrushali, and A. S. Mokhade. "Different methods, techniques and their limitations in protein structure prediction: A review." *Progress in Biophysics and Molecular Biology* 173 (2022): 72-82. https://doi.org/10.1016/j.pbiomolbio.2022.05.002

[16] Peng, Zhenling, Wenkai Wang, Renmin Han, Fa Zhang, and Jianyi Yang. "Protein structure prediction in the deep learning era." *Current Opinion in Structural Biology* 77 (2022): 102495. https://doi.org/10.1016/j.sbi.2022.102495

[17] Schleif, Robert, and Manuel Espinosa. "Where to from here?." *Frontiers in Molecular Biosciences* 9 (2022): 848444. https://doi.org/10.3389/fmolb.2022.848444

[18] Hwang, Howook, Fabian Dey, Donald Petrey, and Barry Honig. "Structure-based prediction of ligand–protein interactions on a genome-wide scale." *Proceedings of the National Academy of Sciences* 114, no. 52 (2017): 13685-13690. https://doi.org/10.1073/pnas.1705381114

[19] Wang, Lijuan, Wenmei Zhang, Yunlong Shao, Dongtang Zhang, Guangsheng Guo, and Xiayan Wang. "Analytical methods for obtaining binding parameters of drug–protein interactions: A review." *Analytica Chimica Acta* 1219 (2022): 340012. https://doi.org/10.1016/j.aca.2022.340012

[20] Zhao, Jingtian, Yang Cao, and Le Zhang. "Exploring the computational methods for protein-ligand binding site prediction." *Computational and Structural Biotechnology Journal* 18 (2020): 417-426. https://doi.org/10.1016/j.csbj.2020.02.008

[21] Rahim, Nasimah, Siti Zalita Talib, Nur Ainun Mokhtar, Nurulbahiyah Ahmad Khairudin, and Ragheed Hussam Yousif. "In-Silico Search Analysis of Potential Inhibitors for 3-Chymotrypsin-Like Protease Of Sars-Cov-2 (Covid-19)." *Journal of Research in Nanoscience and Nanotechnology* 4, no. 1 (2021): 49-56. https://doi.org/10.37934/jrnn.4.1.4956