Thaumatin protein predicted from oil palm and rubber using PHYRE\(^2\) and SWISS-MODEL software

R Hayati\(^1\), M Basyuni\(^2\)*, Lisnawita\(^3\) and E Munir\(^4\)

\(^1\) Doctoral Program of Agricultural Sciences, Faculty of Agriculture, Universitas Sumatera Utara, Jl. Prof. A. Sofyan No. 3, Padang Bulan 20155, Indonesia
\(^2\) Department of Forestry, Faculty of Forestry, Universitas Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera, Padang Bulan 20155, Indonesia
\(^3\) Program Study of Agrotechnology, Faculty of Agriculture, Universitas Sumatera Utara, Jl. Prof. A. Sofyan No. 3, Padang Bulan 20155, Indonesia
\(^4\) Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara, Jl. Bioteknologi No. 1, Padang Bulan 20155, Indonesia

*E-mail: m.basyuni@usu.ac.id

Abstract. Thaumatin-like proteins play important roles in plant defense responses. The protein sequence of a thaumatin gene was found from the NCBI database and analyzed using PHYRE\(^2\) and SWISS-MODEL software. The oil palm showed that the total different protein modeling determines by SWISS-MODEL. In addition, the binding site predicted in oil palm has three model proteins CYS, GLN, and TYR using PHYRE\(^2\). Another case, distinct in rubber model protein which just has two kind proteins (THR and TYR). The image of the structural proteins revealed that c2ahnA differences were performed using PHYRE\(^2\). While, the 2ahn.1.A structure protein had the closest homology image effects in the oil palm and rubber using SWISS-MODEL. The extracellular-cytoplasmic from oil palm was constructed in 19-34 membranes. But different helix, an extracellular-cytoplasmic from rubber was constructed in 22-37 membranes (N-C terminal).

1. Introduction

Thaumatin protein was found as a transcript level that can be increased when *Ganoderma boninense* infected oil palm [1]. On the other hand, the PR5 thaumatin with eight peptides has been reported as a defense protein that increased and a role when the callus development in *E. guineensis* and *E. oleifera* [2]. Thaumatin-like protein was found a long time ago as a tremendous sweet-taste protein from the plant of African [3]. That protein has a potent eight disulfides as an antifungal [4-5].

Recently, the thaumatin has been represented as a defense of plant protein (PR5) against the attack of the pathogens, mainly the fungal such as glycoproteins (20–30 kDa molecular masses) [6]. In other cases, such as canola plants, the thaumatin protein delivers increased to resistance when against *Sclerotinia sclerotiorum* disease [7]. Besides, the protein structures occurrences with a helix of transmembrane ~23 kDa of thaumatin belong to proteins as antimicrobial [8]. The main secondary structures (β-sheet) 24 kDa of thaumatin-like protein reported in pure of litchi [9].

The N-terminal sequence of the thaumatin 24 kDa protein in natural rubber latex was reported as an osmotic closest homologous to the cherry thaumatin protein [10]. For the first case, the genetic transformation of the gene CsTL1 encoding chestnut as thaumatin-like protein (pathogenesis-related
protein) achieved wherein explant was pre-cultured in 2 weeks efficiency 2% transformation [11]. Other than that, the thaumatin II (miRNAs regulatory) transformation with cucumber also has been reported using bioinformatic methods in the transgenic study [12]. Our predicted aims to approach of protein modelling of oil palm and rubber in different software which evaluated based template and structure the thaumatin protein.

2. Materials and Method

2.1. Oil palm and rubber protein

Thaumatin proteins from National Center for Biotechnology Information (NCBI) database genes (cDNA) and amino acid (https://www.ncbi.nlm.nih.gov/) from oil palm (Elaeis guineensis) and rubber (Hevea brasiliensis) were collected. The Accession numbers of GenBank: XM_010921872; XM_029264864 (oil palm), and XM_021800321; XM_021806920 (rubber).

2.2. Template of protein and modelling

The template thaumatin protein and builds model of the 3-D structure was performed by software the PHYRE2 server (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) [13], but compared the SWISS-MODEL server (https://swissmodel.expasy.org/) [14]. The highest template quality selected to protein structure. The 3D ligand-binding site prediction of the oil palm and rubber protein build structures analyzed using online software http://www.sbg.bio.ic.ac.uk/3dligandsite/ [15].

2.3. Transmembrane helix

Transmembrane helices were predicted from the topology sequence analyzed by PHYRE2 server (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index). Thus, the cytoplasmic and extracellular sides from membrane have been labelled publication [14]. The process of START until END for each Transmembrane helix was established to indicate the number and residue index.

3. Results and Discussion

3.1. Thaumatin proteins modelling and 3D ligand-binding site predicted

The thaumatin protein modelling analyzed with PHYRE2 with ID: c2ahnA for oil palm and rubber template was shown in Table 1. Both of protein have different results besides in confidence value 100%. The oil palm has a sequence aligned showed 72.00, but rubber showed 87.00. On the other hand, the highest coverage and residue were found in rubber protein (0.87 and 819, respectively).

| Species   | ID protein | Sequence Aligned | Coverage | Residue | Confidence |
|-----------|------------|------------------|----------|---------|------------|
| Oil palm  | c2ahnA     | 72.00            | 0.71     | 221     | 100.00     |
| Rubber    | c2ahnA     | 87.00            | 0.87     | 819     | 100.00     |

On the other hand, the SWISS-MODEL describes in protein ID: 2ahn.1.A (Table 2). The oil palm showed the highest value in similarity 0.49, identity 57.47, range 25-252, and Q-mean 0.65. Otherwise, the rubber protein was shown in similarity 0.47, identity 52.94, range 26-249, and Q-mean 0.67. Recently, the TgPR8 genes 3D structure PR-5 family of thaumatin-like protein (TLPs), was constructed using homology modeling with Hevea brasiliensis predicted in 99% of coverage and 4% of identity [16].

| Species   | ID protein | Similarity | Identity | Coverage | Range | Q-mean |
|-----------|------------|------------|----------|----------|-------|--------|
| Oil palm  | 2ahn.1.A   | 0.49       | 57.47    | 0.71     | 25-252| 0.65   |
| Rubber    | 2ahn.1.A   | 0.47       | 52.94    | 0.76     | 26-249| 0.67   |
The thaumatin ligand-binding site prediction from oil palm and rubber was shown in Table 3. Three residues from 33; 34; 35 confirmed the several amino acids from oil palm (CYS, GLN, TYR). A variant of contact was performed 1-2-1. Furthermore, the longest residue was found in TYR of rubber (121-131). Other case, the cycloartenol synthase protein from KcMS have been reported in the lowest contact 3D-Ligand binding site from 14-22 in mangrove [17].

3.2. Protein imaging based on PHYRE$^2$ and SWISS-MODEL

Homology models showed in 221 and 819 residues (72% and 87% within the sequences) and 100.0% confidence was shown the image coloured by rainbow N → C terminus protein models from oil palm $\hat{A}$; X: 48.753 Y:45.296 Z:49.307, and rubber $\hat{A}$; X:49.540 Y:44.872 Z:50.117, respectively by PHYRE$^2$ online software (Figure 1A and 1B). The coverage values of protein ID: 2ahn.1.A was the distinct value compared the SWISS-MODEL predicted 81% and 76% in Figure 2A and 2B.

| Species | Amino acid | Average distance | Contact | Residue |
|---------|------------|------------------|---------|---------|
| Oil palm | CYS        | 0.01             | 1       | 33      |
|         | GLN        | 0.00             | 2       | 34      |
|         | TYR        | 0.00             | 1       | 35      |
| Rubber  | THR        | 0.00             | 1       | 121     |
|         | TYR        | 0.00             | 1       | 131     |

Figure 1. The image coloured by rainbow N → C terminus protein models from oil palm $\hat{A}$; X: 48.753 Y:45.296 Z:49.307 (A) and rubber $\hat{A}$; X:49.540 Y:44.872 Z:50.117 (B)

Figure 2. The protein modelling from oil palm (A), and rubber (B) by SWISS-MODEL
3.3. Transmembrane helix

Figure 3 was shown the transmembrane of helices in oil palm and rubber. Both genes have one (S1) the transmembrane helix. However, the different value between oil palm was on S1 (19-34), and rubber S1 (22-37) for extracellular and cytoplasmic in N→C terminal.

![Figure 3. Transmembrane helices by PHYRE² from oil palm (A) and rubber (B)](image)

4. Conclusions

In this paper, we have collected information regarding modeling the thaumatin-like proteins and their contribution to oil palm and rubber plants. We described the structure and template of different defense-related proteins found in total binding proteins. The last section describes transmembrane helices distinct occurring in both the plants.

Acknowledgment

The PMDSU project supported this study number: 11/E1/ KP.PTNB/2019, from Directorate of Research and Technology and Higher Education, Republic of Indonesia.

References

[1] Ho CL., Tan YC., Yeoh KA., Lee WK., Ghazali AK., Yee WY. and Hoh CC. 2019. Leaf transcriptome of oil palm (Elaeis guineensis Jacq.) infected by Ganoderma boninense. *Trees*, 33, 943-950.

[2] Ribeiro DG., Almeida RF., Fontes W., Souza CM., Sousa MV., Ricart CAO. and Mehta A. 2019. Stress and cell cycle regulation during somatic embryogenesis plays a key role in oil palm callus development. *J. Proteom.* 192 137-146.

[3] Edens L., and Van DWH. 1985. Microbial synthesis of the sweet-tasting plant protein thaumatin. *Trends Biotechnol.* 3 61-64.

[4] Lee JH., Weickmann JL., Koduri RK., Ghosh-Dastidar P., Saito K., Blair LC. and Kendall RL. 1988. Expression of synthetic thaumatin genes in yeast. *Biochemistry.* 27 5101-5107.

[5] Koiwa H., Kato H., Nakatsu T., Oda JL., Yamada Y and Sato F. 1999. Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *J Mol. Biol.* 4 1137-1145.

[6] Palacín A, Quirce S, Sánchez-Monge R, Fernández-Nieto M, Varela J. 2008. Allergy to kiwi in patients with baker's asthma: identification of potential cross-reactive allergens. *Ann. Allergy Asthma Immunol.* 101 200–205.

[7] Aghazadeh R., Zamani M., Motallebi M., Moradyar M., Jahromi ZM. 2016. Co-transformation of canola by chimeric chitinase and tlp genes towards improving resistance to *Sclerotinia*
sclerotiorum World. J. Microbiol. Biotechnol. 32 p. 144

[8] Sharma A., Tyagi S., Alok A., Singh K. and Upadhyay SK. 2020. Thaumatin-like protein kinases: Molecular characterization and transcriptional profiling in five cereal crops. Plant Sci. 290 110317.

[9] Chen H., Wang K., Xiao H., Hu Z. and Zhao L. 2020. Structural Characterization and Pro-inflammatory Activity of a Thaumatin-like Protein from the Pulp Tissues of Litchi chinensis. J. Agric. Food Chem. 10 1021.

[10] Pastorello EA., Farioli L., Pravettoni V., Ortolani C., Fortunato D., Giuffrida MG. and Conti A. 2003. Identification of grape and wine allergens as an endochitinase 4, a lipid-transfer protein, and a thaumatin. J. Allergy Clin. Immunol. 111 350-359.

[11] Cano V., Martinez MT., San José MC., Cousel JL., Varas E., Bouza-Morcillo L. and Corredoira, E. 2020. Regeneration of transgenic plants by Agrobacterium-mediated transformation of Quercus ilex L. somatic embryos with the gene CsTL1. New Forest. 10 1-19.

[12] Pawelkowicz ME., Skarzyńska A., Sroka M., Szwacka M., Pniewski T. and Pląder W. 2020. Effect of Transgenesis on mRNA and miRNA Profiles in Cucumber Fruits Expressing Thaumatin II. Genes, 11 334.

[13] Kelley LA., Mezulis S.,Yates CM., Wass MN., Sternberg MJE. 2015. The Phyre² web portal for protein modeling, prediction and analysis. Nat. Protocols 10 845–858.

[14] Arnold K., Bordoli L., Kopp J. and Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22 195–201.

[15] Benkert P., Biasini M., Schwede, T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27 343–350.

[16] Silva RJS., Alves, RM., Gramacho KP., Marcellino LH. and Micheli F. 2020. Involvement of structurally distinct cupuassu chitinases and osmotin in plant resistance to the fungus Moniliophthora perniciosa. Plant Physiol. Bioch. 148 142-151.

[17] Basyuni M., Wati R., Sulistiyono N., Hayati R., Oku H., Baba S. and Sagami H. 2018. Protein modelling of triterpene synthase genes from mangrove plants using Phyre2 and Swiss-model. J. Phys. Conf. Ser. 978 012095.