Predicted cycloartenol synthase protein from *Kandelia obovata* and *Rhizophora stylosa* using online software of Phyre2 and Swiss-model

M Basyuni¹*, N Sulistiyono¹, R Wati¹, Sumardi², H Oku³, S Baba⁴, H Sagami⁵

¹Department of Forestry, Faculty of Forestry, Universitas Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera, 20155, Indonesia
²Faculty of Pharmacy, Universitas Tjut Nyak Dhien, Jl. Rasmi No. 28 Medan 20123, Indonesia
³Molecular Biotechnology Group, Tropical Biosphere Research Center, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
⁴International Society for Mangrove Ecosystems, Faculty of Agriculture, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
⁵Institute of Multidisciplinary Research for Advanced Material, Tohoku University, Sendai, 980-8577, Japan

*Email: m.basyuni@usu.ac.id*

**Abstract.** Cloning of *Kandelia obovata* *KcCAS* gene (previously known as *Kandelia candel*) and *Rhizophora stylosa* *RsCAS* have already been reported and encoded cycloartenol synthases. In this study, the predicted *KcCAS* and *RsCAS* protein were analyzed using online software of Phyre2 and Swiss-model. The protein modelling for *KcCAS* and *RsCAS* cycloartenol synthases was determined using Pyre2 had similar results with slightly different in sequence identity. By contrast, the Swiss-model for *KcCAS* slightly had higher sequence identity (47.31%) and Qmean (0.70) compared to *RsCAS*. No difference of ligands binding site which is considered as modulators for both cycloartenol synthases. The range of predicted protein derived from 91-757 amino acid residues with coverage sequence similarities 0.86, respectively from template model of lanosterol synthase from the human. Homology modelling revealed that 706 residues (93% of the amino acid sequence) had been modelled with 100.0% confidence by the single highest scoring template for both *KcCAS* and *RsCAS* using Phyre2. This coverage was more elevated than swiss-model predicted (86%). The present study suggested that both genes are responsible for the genesis of cycloartenol in these mangrove plants.

1. Introduction

Phytosterols and pentacyclic triterpenoids are widespread in higher plants including mangrove forests [1]. Molecular cloning of two mangrove plants namely *Kandelia obovata* *KcCAS* gene (previously known as *Kandelia candel*) and *Rhizophora stylosa* *RsCAS* had been cloned, characterized, and encoded cycloartenol synthases [2]. The *KcCAS* and *RsCAS*, therefore, are involved in phytosterol biosynthesis, a member of oxidosqualene cyclases (OSCs) gene. The open reading frame (ORF) of both genes consist of 2277 bp that encode 58 amino acid polypeptide, respectively [2]. Salt tolerance mechanism in two mangrove plants *K. obovata* and *R. stylosa* showed variation responses: mRNA
level of \( \text{KcCAS} \) was not modulated by salinity concentration in the roots and was lowered after removal the salinity and transfer to freshwater [3-4]. By contrast, \( \text{KcCAS} \) decreased in the leaves of \( K. \ obovata \) [3]. In case of \( \text{RsCAS} \), the mRNA level was enhanced with salinity only in the roots but not in the leaves [5].

Particular prominence was on the phytosterols because these compounds were significant membrane permeability for steroid hormone and biomarkers for marine organic matter [6-7] and might give more insight into the physiological roles of phytosterols in mangrove plant species. It has been reported that the highest mitochondrial target peptide in the possibility of the potential transit peptide was in \( \text{KcCAS} \) [8]. Regardless of the critical part of the protein in the plant kingdom, the predicted protein model of cycloartenol synthases from mangrove plants have rarely been studied. Thus, the present study aimed to analyze protein structure from two cycloartenol synthases of \( K. \ obovata \) and \( R. \ stylosa \).

2. Materials and Method

2.1. Material

Two cycloartenol synthase genes from NCBI database (https://www.ncbi.nlm.nih.gov/) were collected. The GenBank accession numbers of the DNA and amino acid sequences of using this analysis is as follows: AB292609, BAF73930 (\( K. \ obovata \ \text{KcCAS} \)) and AB292608, BAF73929 (\( R. \ stylosa \ \text{RsCAS} \)).

2.2. Template search and model building

Template search and the modelling of the three-dimensional structure of the protein was performed using Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) [9] and Swiss-model (https://swissmodel.expasy.org/) [10]. The templates with the highest quality have then been selected for model building.

2.3. Prediction of ligand binding site

The ligand binding site prediction in the modelled protein structure was carried out using 3DLigandSite server (http://www.sbg.bio.ic.ac.uk/3dligandsite/) [11].

2.4. Transmembrane helix prediction

Transmembrane helices were predicted the sequence using phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) [9]. The extracellular and cytoplasmic sides of the membrane were labelled [9]. The start and end of each transmembrane helix described a number indicating the residue index.

3. Results and Discussion

3.1. Protein modelling using Phyre2 and Swiss-model

Table 1 showed protein modelling using Pyre2 of \( \text{RsCAS} \) and \( \text{KcCAS} \) cycloartenol synthases. Both genes had similar results with slightly different in sequence identity. The protein model used human lanosterol synthase with protein database (PDB) ID: c1w6kA as a template structure with confidence 100% and coverage 93%.

Similarly, results of protein modelling using Swiss-model was listed in Table 2. \( \text{KcCAS} \) slightly had higher in sequence identity (47.31%) and Qmean (0.70, estimated model reliability between 0-1) compared to \( \text{RsCAS} \). The range of predicted protein derived from 91-757 amino acid residues with coverage sequence similarities 0.86, respectively from template model of lanosterol synthase from human (Table 2).
### Table 1. Results of protein modelling using Phyre2

| Gene  | Template ID | Sequence Identity | Residue | Confidence | Coverage (%) |
|-------|-------------|-------------------|---------|------------|--------------|
| KcCAS | c1w6kA      | 46                | 706     | 100        | 93           |
| RsCAS | c1w6kA      | 45                | 706     | 100        | 93           |

### Table 2. Results of protein modelling using Swiss-model

| Gene  | Template   | Sequence Identity | Sequence similarity | Range | Coverage | QMEAN |
|-------|------------|-------------------|---------------------|-------|----------|-------|
| KcCAS | 1w6k.1.A   | 47.31             | 0.43                | 91-757| 0.86     | 0.70  |
| RsCAS | 1w6k.1.A   | 46.85             | 0.43                | 91-757| 0.86     | 0.69  |

#### 3.2. Predicted binding site

Ligands which are considered in the current study as modulators for both cycloartenol synthases are displayed in Table 3. The ligands bind particular residue, amino acid, contact, and average distance. 

KcCAS had seven residues from 677, 678, 679, 680, 683, and 74 which corresponded to amino acids of His, Val, Val, Asn, Trp, and Met. The longest distance belongs to Val678.

### Table 3. Predicted binding site of two cycloartenol synthase

| Gene  | Residue | Amino Acid | Contact | Average distance |
|-------|---------|------------|---------|------------------|
| KcCAS | 677     | HIS        | 24      | 0.00             |
|       | 678     | VAL        | 11      | 0.69             |
|       | 679     | VAL        | 24      | 0.00             |
|       | 680     | ASN        | 21      | 0.29             |
|       | 683     | TRP        | 15      | 0.46             |
|       | 724     | MET        | 24      | 0.00             |
| RsCAS | 678     | ILE        | 11      | 0.69             |
|       | 679     | VAL        | 24      | 0.00             |
|       | 680     | ASN        | 21      | 0.29             |
|       | 683     | TRP        | 15      | 0.46             |
|       | 724     | MET        | 24      | 0.00             |

In contrast to this observation, RsCAS possed five residues Ile678, Val679, Asn680, Trp683, and Met724 with contact value was 11-24. A similar result was obtained for average distance for Ile678 (0.69) of RsCAS cycloartenol synthase.

It is noteworthy that the phytosterol composition from leaves and roots of K. obovata and R. stylosa was diversity. β-sitosterol is the major phytosterol component of both species in leaves and roots, as could be expected for higher plants including mangrove species [1, 5]. The concentration of cycloartenol, a precursor for phytosterol biosynthesis, is usually low in the plant tissue [12]. Our previous study also found the low level of cycloartenol in the leaf of both of R. stylosa and K. candel [1], suggesting that both genes are responsible for the genesis of cycloartenol in these mangrove plants. It is yet to come to demonstrate the involvement of cycloartenol synthase gene in the phytosterol biosynthesis of the plant kingdom.

#### 3.3. Phyre2 and Swiss-protein model

Homology modelling revealed that 706 residues (93% of the amino acid sequence) had been modelled with 100.0% confidence by the single highest scoring template for both KcCAS and RsCAS using Phyre2 (Figure 1 A and B). This coverage was higher than the Swiss-model predicted (86%) as showed in Figure 2A and B. Furthermore the number aligned proteins both genes were 93 with some matched PDB structures were 31. This results supported therefore by the predict molecular weight for
both $Kc$CAS and $Rs$CAS were 83 kD, and approximately theoretical isoelectric residues were 4.91, respectively.

This slight difference model may be due to the difference zonation of mangrove between $K. obovata$ and $R. stylosa$. $R. stylosa$ usually grows close to the sea or coast, and $K. obovata$ distributes more inland ward [5]. These results suggested the importance of understanding the diversity of phytosterol genes in mangroves. Therefore further investigation is needed to clarify this finding.

![Phyre2 protein model for KcCAS (A) and RsCAS (B).](image1)

**Figure 1.** Phyre2 protein model for $Kc$CAS (A) and $Rs$CAS (B).

![Swiss protein model for KcCAS (A) and RsCAS (B).](image2)

**Figure 2.** Swiss protein model for $Kc$CAS (A) and $Rs$CAS (B).

### 3.4. Transmembrane helices

Figure 3 showed transmembrane (TM) helices in both $Kc$CAS and $Rs$CAS. Both genes had four TM helices (S1-S4). The different position between $Kc$CAS and $Rs$CAS were on S2 (257-272 and 256-271) and S3 (330-315 and 334-319) both in the extracellular and cytoplasmic. Several lines of the study demonstrated that the sterol is a modulator of membrane permeability [1, 12]. Triterpenoid alcohols have been structurally differentiated from phytosterols, although they share common biosynthetic pathway and hence the similar chemical structure [2]. However, there has been no scientific rationale to distinguish these compounds functionally as the structural membrane lipid. In this context, our *in*
vitro study showed that triterpene, as well as phytosterols, was incorporated into the lipid bilayer of the liposome, suggesting that triterpenoid is also a modulator of membrane permeability [4-5].

Figure 3. Transmembrane helices of KcCAS (A) and RsCAS (B)

4. Conclusions
The predict protein modelling for KcCAS and RsCAS cycloartenol syntheses showed similar results with slightly different in sequence identity using Phyre2 and Swiss-model. In a case for the Swiss-model for KcCAS fairly had higher sequence identity and Qmean compared to RsCAS. The present study suggested that both genes are responsible for the origin of cycloartenol in these mangrove plants.

Acknowledgment
A part of this work was supported by an International Research Collaboration and Scientific Publication Grant 2017 (No. 003/SP2H/LT/DRPM/IV/2017 to MB) from the Directorate for Research
and Community Service, Ministry of Research, Technology and Higher Education, Republic of Indonesia.

References
[1] Basyuni M, Oku H, Baba S, Takara K, Iwasaki H 2007 Isoprenoids of Okinawan mangroves as lipid input into estuarine ecosystem J Oceanography 63 601–608.
[2] Basyuni M, Oku H, Tsujimoto E, Baba S 2007 Cloning and functional expression of cycloartenol synthases from mangrove species Rhizophora stylosa Griff. and Kandelia candel (L.) Druce Biosci. Biotechnol. Biochem. 71 1788–1792.
[3] Basyuni M, Baba S, Inafuku M, Iwasaki H, Kinjo K, Oku H 2009 Expression of terpenoid synthase mRNA and terpenoid content in salt-stressed mangrove J. Plant Physiol. 166 1786–1800.
[4] Basyuni M, Baba S, Kinjo Y, Putri LAP, Hakim L, Oku H 2012 Salt-dependent increase in triterpenoids is reversible upon transfer to fresh water in mangrove plants Kandelia candel and Bruguiera gymnorhiza J. Plant Physiol. 169 1903–1908.
[5] Basyuni M, Baba S, Kinjo Y, Oku H 2012 Salinity increases the triterpenoid content of a salt secretor and a non-salt secretor mangrove Aquat. Bot. 97 17–23.
[6] Rontani JF, Vaultier F, Bonin P 2014 Biotic and abiotic degradation of marine and terrestrial higher plant material in intertidal surface sediments from Arcachon Bay (France): A lipid approach Mar. Chem. 158 69–79.
[7] He D, Mead RN, Belicka L, Pisani O, Jaffé R 2014 Assessing source contributions to particulate organic matter in a subtropical estuary: a biomarker approach Org. Geochem. 75 129–139.
[8] Basyuni M, R Wati 2017 Bioinformatics analysis of the oxidosqualene gene and the amino acid sequence in mangrove plants J. Phys. : Conf. Ser. 801 012011.
[9] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg, MJ 2015 The Phyre2 web portal for protein modeling, prediction and analysis Nat. Protocols 10 845–858.
[10] 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.
[11] Wass M.N., Kelley LA, Sternberg MJ 2010 3DLigandSite: predicting ligand-binding sites using similar structures Nucleic Acid Res. 38 W469–473.
[12] Uddin M, Sarker M, Islam Z, Ferdosh S, Akanda M, Haque J, Easmin M, Shamsudin B, Hadijah S, Yunus KB 2015 Phytosterols and their extraction from various plant matrices using supercritical carbon dioxide: a review J. Sci. Food Agric. 95 1385–1394.