HIBISCUS PUBLISHER

JOURNAL OF BIOCHEMISTRY, MICROBIOLOGY AND BIOTECHNOLOGY

Website: https://journal.hibiscuspublisher.com/index.php/JOBIMB

JOBIMB VOL 8 NO 1 2020
SARS-Cov-2 protease

Molecular Modelling of *Oryza sativa* Starch-branching Enzyme 1

Sulaiman Mohammed[1,5*], Mohammed Jibrin Ndejiko[2,5], Mohammed Maikudi Usman[3], Yilmaz Kaya[4] and Fahrul Zaman Huyop[5]

[1]Department of Biological Sciences, Faculty of Science, Gombe State University, P.M.B 0127 Gombe, Nigeria, Orcid: 0000 0003 3094 4769.
[2]Department of Microbiology, Faculty of Natural Sciences, Ibrahim Badamasi Babangida University Lapai, Nigeria. Orcid: 0000-0003-0790-833X.
[3]Department of Biotechnology, School of Life Sciences, Modibbo Adama University of Technology Yola, Nigeria Orcid: 0000-0002-5699-2673.
[4]Department of Plant Biotechnology, Faculty of Agriculture, Ondokuz Mayis University, Samsun, Turkey Orcid: 0000-0003-1506-7913.
[5]Department of Bioscience, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru. Orcid:0000-0003-3979-4087.

*Corresponding author:
Sulaiman Mohammed
Department of Biological Sciences,
Faculty of Science,
Gombe State University,
P.M.B 0127 Gombe, Nigeria
Orcid: 0000 0003 3094 4769
Email: sumulsu@yahoo.com

HISTORY

Received: 15th Jan 2020
Received in revised form: 14th of March 2020
Accepted: 18th of April 2020

KEYWORDS

rice *SBE1*
secondary structure
domains interaction
3D model
C-score

ABSTRACT

Starch-branching enzymes (SBE) serves as the only enzyme generating glucan branches in green plants and consequently plays a significant role on the resulting starch final structure. Research on rice (*Oryza sativa*) *SBE1* (*OsSBE1*) structural biology remain untapped. Therefore, there is a necessity for research on the enzyme molecular structure which could lead to the protein function annotation, starch production and energy booster drug design. Analysis of *OsSBE1* secondary structure, domains and their interactions, enzyme 3D structure prediction and validation based on C-score were carried out. The *OsSBE1* primary sequence was retrieved from GenBank and its secondary structure was predicted to be; α-helix (27.68%), extended strand (22.78%) and higher random coil (949.54%). Enzyme domains were found to be carbohydrate-binding module (CBM) 48 (isoamylase N-terminal domain), α-amylase catalytic domain and α-amylase C-terminal all-beta domain with active sites important amino acids asparagine and glutamic acid. From the five 3D models generated, model 3 displayed best prediction. The Ramachandran refinement has 97.3 amino acids residues in favoured region and 0.4 C-score. This bioinformatics study has elucidated on the *OsSBE1* molecular model and first to report on its domain interaction.

INTRODUCTION

Rice (*Oryza sativa*) production is mostly in tropical and sub-tropical environments as dominant cereal. The importance of rice as diet can never be overemphasized and varies among nations [1]. The cereal is a starch source of food, serving as a staple diet for numerous communities more especially in Asia and Africa nations. Kole [2] reported that rice seeds provide more than 20% of the world calories in form of starch. Thus, its accounts for over 70% of the daily calories' intake in many Asian countries. The rice seeds bran is a raw material for oil extraction by cosmetics and healthcare industries. Interestingly, the crop is measured as the most essential crop with close relation to other cereals and prototypical species for enzyme analysis through molecular studies such as protein engineering, mutagenesis and protein function [3.4].

Naturally, starch is produced in the plastids of higher plants which represents the most abundant storage polyglucan, then, functioning as a short- and long-term reserve carbohydrate. It is synthesized via coordinated acts of multiple starch-branching enzymes (*SBEs*) isoforms which are shown to have strong influence in glucan structure generation [5]. This starch is made up of linear α-1,4-linked glucans synthesized by starch synthases (SS, E.C. 2.4.1.21), a group of ADP-glucose-dependent transferases and α-1,6-linked branches. *SBEs* stimulates starch structure formation by catalysing the α-1,6-branch points at different frequency and branch chain length [6]. Many have reported on *SBE* IIa and b as most abundant protein in endosperm

amyloplast which influence the starch structure toward facilitating nocturnal degradation in maize [7], barley [8], *Arabidopsis* [9] and expression in *E. coli* [10].

The existing studies on *Oryza sativa SBE1* (*OsSBE1*) has failed to elucidate on the phenotypic and molecular nature of the enzyme structure, leaving behind its role as starch synthesizing protein a rather open question [11]. Moreover, the analysis of the protein deficiency with inefficient α-amylase digestion of the starch reserves in the seed is also incompletely understood. Therefore, the selection of the *OsSBE1* in this bioinformatics study appears to impact on the protein molecular nature, functional sites and future transformation to overexpress as advantage associated with seedling vitality, potency and fitness. In fact, this is among the major reasons that makes rice a model monocot plant for seed biology studies and genomics analysis. The plant is currently studied at transformed levels using histological, cytological, physiological, biochemical, genetic and molecular biology tools, for example bioinformatics software. Until now, several reports have been shown molecular modelling of some enzymes from *Oryza sativa*. To the best of our knowledge, no reports are available in literature for molecular modelling of starch-branching enzyme from *Oryza sativa.* Therefore, the purpose of this research constitutes the first ever successful and complete molecular modelling of *Oryza sativa* starch-branching enzyme.

EXPERIMENTAL

Rice *Starch-branching Enzyme1* (*OsSBE1*) sequence retrieval

Rice *SBE1* open reading frame (ORFs) was obtained from the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). The full-length sequence was further analyzed using BLASTp algorithm and predicted the nucleotides from the amino acid sequence using tool provided by ExPASy (http://web.expasy.org/translate/).

Secondary structure prediction and analysis of domain

The secondary structure of the *OsSBE1* was predicted using NPS@: GOR4 tool by information theory algorithm. To compare and confirm the presence of the enzyme domains, Pfam (http://pfam.sanger.ac.uk/) tool was used for the query protein domains and position of important amino acids. The relationship and interaction(s) between the domains were also determined using NPS@: GOR4.

Tertiary structure 3D modelling and validation

The 3-dimensional (3D) model structure of the enzyme was generated using I-TASSER (Iterative-Threading / ASSembly / Refinement). The modeller gave a prediction of the structures based on protein sequence. Important amino acids of the enzyme were detected using UniProt (https://www.uniprot.org/) software from NCBI. The best 3D model was used for quality validation analysis and assessment of stereochemistry in RAMPAGE SAVES for Ramachandran plot.

RESULTS AND DISCUSSION

OsSBE1 secondary structure and domains

In this bioinformatics analyses, the *OsSBE1* primary sequence was retrieved from GenBank, while its secondary structure prediction was performed using NPS@: GOR4 tool as shown in **Fig. 1**. The secondary structure results showed alpha-helix (27.68%), extended strand (22.78%) and higher random coil 49.54%. Enzyme domains were identified by using Pfam tool (**Fig. 2**). The protein N terminal consists of domain belongs to carbohydrate-binding module (CBM) 48 (isoamylase N-terminal domain) like class (alignment region 57-140). Followed by alpha-amylase (α-amylase) catalytic domain (alignment region 228-301) found in the middle and alpha-amylase (α-amylase) C-terminal all-beta domain (alignment region 592-692). Equally, active sites important amino acids of the enzyme were further predicted as shown in **Fig. 2** (black line with pink headed). They were at position 344-asparagine as nucleophiles and 399-glutaric acid as proton donor.

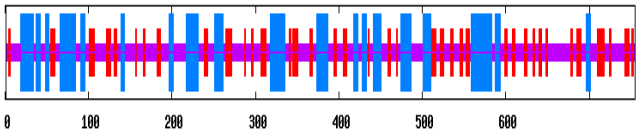


Fig. 1. The result of *OsSBE I* secondary structure prediction using NPS@: GOR4

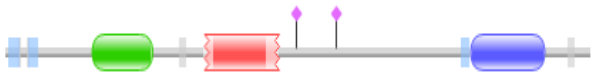


Fig. 2. Domains found on *OsSBE1*. The Pfam tool algorithm predicted the 3 main enzyme domains as; carbohydrate-binding module (CBM) 48 (isoamylase N-terminal domain) (first in green colour), α-amylase catalytic domain (middle in red colour) and α-amylase C-terminal all-beta domain (last in blue colour) with interfering important amino acids (black line with pink head).

The protein structures were comprehensively scrutinized for relationship and quality selection. Examined toward determining its secondary structures includes alpha, beta and loops. As well, the diverse enzyme domain and their functions were analysed and elucidated [12]. The relationships between family domains in a clan are determined as shown in **Fig. 3** (denote the domains in red colour). They are deemed to be closely located as the E-value is less than 10^{-3}. The strong relationships are indicated with a solid line, whereas less closely related family pairs (E-value of between 10^{-3} and 10^{-1}) are shown with a traced line (**Fig. 3**). Using the Pfam family box, the E-value for each pair either closely or partially related families were shown succeeding to the line connecting the families.

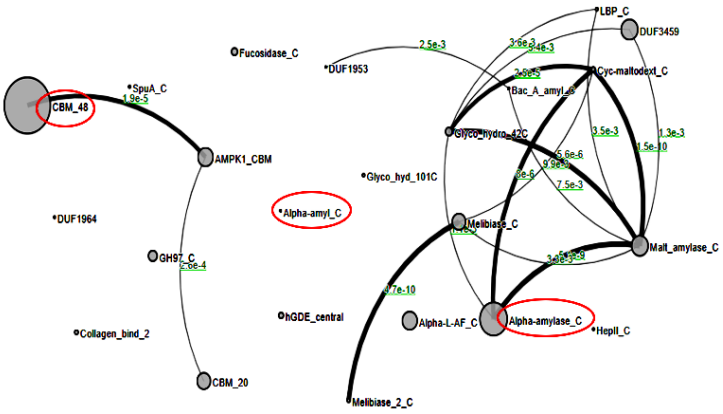


Fig. 3. The relationship between *OsSBE1* domains; solid lines indicated the strong relationship, while the lighter lines shown the partial relationship.

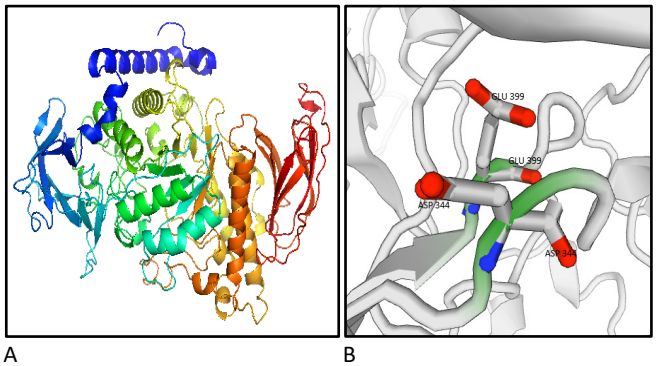
Carbohydrate-binding module (CBM) domain is found in carbohydrate-active enzymes. Majority of such domains can perform carbohydrate-binding activity in starch producing plants including rice, while some were found on cellulosomal scaffolding proteins. The diverse CBM were

earlier known as cellulose-binding domains and classified into many families based on their amino acid similarity(s) and precise role. Recently, about 64 families of CBM were confirmed using CAZy database [13,14]. For example, microbial glycoside hydrolases CBMs plays a vital role in photosynthetic fixed carbon recycling pathway through binding to a specific plant structural polysaccharides. It equally recognises both amorphous and crystalline cellulose forms, associated with enzymes active in plant cell wall hydrolysis. Some CBMs were identified by amino acid multiple sequence alignment, nonetheless, only few have experimentally shown carbohydrate-binding function [15].

The CBM48 of *SBE1* was specifically found in enzymes containing glycosyl hydrolase family 13 as catalytic domain. It was found in a range of enzymes function on branched substrates including branching enzyme, isoamylase and pullulanase. Isoamylase hydrolyses 1,6-α-D-glucosidic branch linkages in amylopectin, glycogen and dextrin; 1,4-α-glucan branching enzyme, then play a role in the formation of 1,6-glucosidic linkages of glycogen. The CBM 48 binds glycogen which serves as system in plant species. Evenly, α-amylase is present in the plant different organs serving as food preservative. The enzyme functions by hydrolysing the bonds of larger α-linked polysaccharides such as glycogen and starch yielding maltose, fructose and glucose [16,17]. It also catalyses hydrolysis of 1-4-α-D-glucosidic linked polysaccharides in order to remove successive α-maltose residues from non-reducing ends of starch to maltose conversion [18,19]. Likewise, α-amylase, C-terminal all-beta domain was classified as glycosyl hydrolases mostly containing the enzyme active site. It is a calcium binding domain bulging between α-helix and β-strand [20,21].

OsSBE1 3D structure and validation of refinement model

The *OsSBE1* 3D structure prediction was done by using I-TASSER of threading method. Five (5) molecular models was generated by the server, all having good quality and resolution. **Fig. 4a** showed the best model which was selected based on C-score. The C-score measurement determine the quality of resulting models and demonstration of their correlation quality [22]. From the results, model 3 display best predicted *OsSBE1* 3D structure at 97.3 model quality and 0.4 C-score as tabulated in **Table 1**. The active sites important amino acids, their respective names (asparagine and glutaric acid) and position (344 and 399) was determine as shown in **Fig. 4b**.



A B

Fig. 4. (A) *OsSBE1* predicted 3D (model 3) model. (B) The enzyme model showing the 2 important amino acids.

Table 1. C-score and quality of the *OsSBE1* 3D models predicted using I-TASSER.

Model	C-Score	Model Quality
Model I	3.5	83.1
Model II	0.4	97.2
Model III	0.4	97.3
Model IV	6.3	81.1
Model V	6.0	80.4

Interestingly, the I-TASSER was found to be a better server for predicting protein(s) 3D model, albeit, a complex approach used in the last two CASP experiments. It is an integrated platform for automated protein 3D prediction, modelling and determining its function based on sequence-to-structure to function paradigm [23]. As the biological function of protein molecules is determined by their 3D shape, which prescribe how the protein – protein or protein – ligand interacts. Hence, the most common motivation for protein 3D molecular structure prediction is to use the model information to gain insight into the protein's genetic activities. The *OsSBE1* prediction study could assist in understanding the molecular function of enzyme as well as encouraging further investigation into other rice enzymes, proteins or genes molecular analysis, interaction and function.

OsSBE1 enzyme best 3D (model 3) model was evaluated by using the Ramachandran plot (**Fig. 5**) via RAMPAGE server and indicated the predicted models high-quality. The refinement model was 97.3% that is slightly above that of model 2 (97.2%) favoured residues and each protein residue is represented in blue dot (**Table 1**). Previous finding indicated that 80-90% favoured regions in the protein measured as high-quality model [24]. The validation was for the correctness of fold and structure, error over localized regions and stereochemical parameters. In fact, the model must satisfy as much restrains as possible prior to any validation analysis. The Ramachandran plot discloses the backbone dihedral angles ψ (psi) against φ (phi) of amino acid residues in the 3D model.

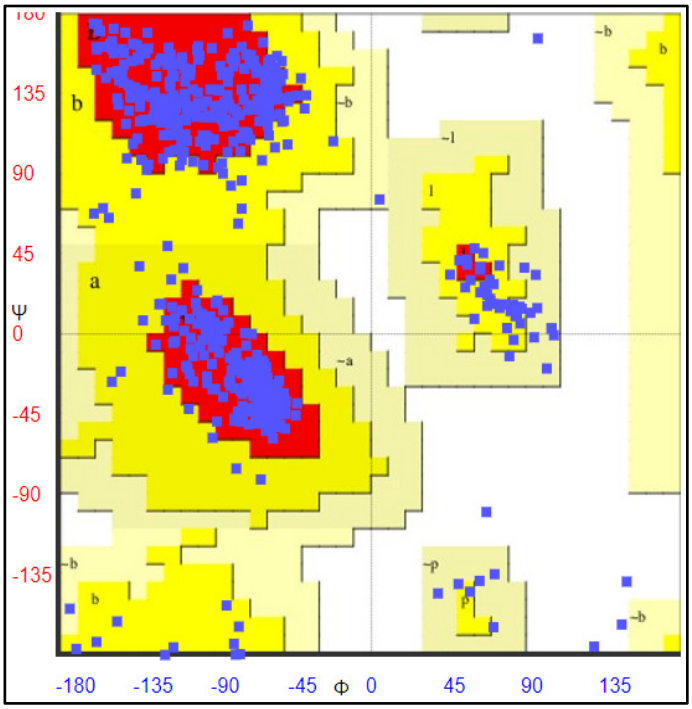


Fig. 5. *OsSBE1* 3D model validation result using Ramachandran plot.

The Ramachandran plot shows the empirical distribution of information experiential in any structure which favours appearance of theoretical regions. The angles from the kind of conformation can possibly be assumed to correlate it to specific secondary structures. Therefore, β-sheets and α-helices plus loops (to some extent) adopt a limited set of phi-psi angle interactions. Whereas, some precise amino acids involving proline and glycine have unique plots as they differ from others in the structure. Consequently, with the development of more technical modeller software, for example ESyPred3D and (PS)2 (protein structure prediction server) it is now possible to obtain the tri-dimensional structure and calculate its disulfide bridge. It is two points calculation from the right side of the molecule, serving as very important characteristic of protein function. The model validation software used in this study ensures that the structure obtained is presentable prior to begin a more robust analysis.

CONCLUSION

OsSBE1 secondary structure determination, plus elucidation on its domain's interactions and tertiary model molecular nature is of vital importance towards improving the plant starch production and reduction. Hence, the assortment of the enzyme for bioinformatics study appears to impart on the protein molecular nature, domain and their interactions, how to engineer and transform soon for overexpression as advantage associated with seedling vitality, potency and fitness. The *OsSBE1* structural modelling conducted in this study revealed the information about secondary structure, domains and their interaction and enzyme active sites. As well, the enzyme tertiary structure model was constructed by using I-TASSER as the first report. The findings confirmed the 3-main domains and validated the 3D model quality based on C-score. This modelled structure may favour conduction of protein engineering or mutagenesis and energy booster drug design towards enhancing productivity of human. As well, wet lab experiment should be conducted to verify the in silico of the enzyme.

ACKNOWLEDGEMENT

We appreciate the support of the collaborative universities in assisting the wet lab research on the enzyme, as well as the respective departments.

CONFLICT OF INTEREST

All authors consented to this submission and declare no conflict of interest

REFERENCES

1. Deng ZY, Gong CY, Wang T. Use of proteomics to understand seed development in rice. Proteomics, 2013;13: 1784-1800
2. Kole PC, Hasib KM. Mutant heterosis in aromatic rice. Trop Agric. 2005;82(3–4):319–24.
3. Shrawat AK, Carroll RT, DePauw M, Taylor GJ, Good AG. Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. Plant Biotechnol J, 2008;6: 722-732
4. Brauer EK, Rochon A, Bi YM, Bozzo GG, Rothstein SJ, Shelp BJ. Reappraisal of nitrogen use efficiency in rice overexpressing glutamine synthetase1. Physiol Plant, 2011;141: 361-372
5. Pfister B, Zeeman SC. Formation of starch in plant cells. Cell Mol Life Sci, 2016;73: 2781-2807
6. Tetlow IJ, Emes MJ. A review of starch-branching enzymes and their role in amylopectin biosynthesis. IUBMB Life,2014;66: 546-558
7. Regina A et al.. Starch-branching enzyme IIb in wheat is expressed at low levels in the endosperm compared to other cereals and encoded at a non-syntenic locus. Planta, 2005;222: 899
8. Ahmed Z, Tetlow IJ, Ahmed R, Morell MK, Emes MJ. Protein–protein interactions among enzymes of starch biosynthesis in high-amylose barley genotypes reveal differential roles of heteromeric enzyme complexes in the synthesis of A and B granules. Plant Sci, 2015;233: 95-106
9. Fisher DK, Gao M, Kim K-N, Boyer CD, Guiltinan MJ. Two closely related cDNAs encoding starch-branching enzyme from *Arabidopsis thaliana*. Plant Mol Biol, 1996;30: 97-108
10. Feng L, Fawaz R, Hovde S, Gilbert L, Chiou J, Geiger JH. Crystal structures of *Escherichia coli* branching enzyme in complex with linear oligosaccharides. Biochemistry, 2015;54: 6207-6218
11. Han Y, Xu M, Liu X, Yan C, Korban SS, Chen X, Gu M. Genes coding for starch-branching enzymes are major contributors to starch viscosity characteristics in waxy rice (Oryza sativa L.). Plant Sci, 2004;166: 357-364
12. Bomediano L, Braz A, Costa R, Simoes A. Molecular modeling of *Leptospira biflexa* OxyR regulator. Cong Brasil Eng Bioméd, 2014; 781 - 784.
13. Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology, 2010;20: 1574-1584
14. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels, 2013;6: 41
15. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J, 2004;382: 769-781
16. Sundarram A, Murthy TPK. α-amylase production and applications: a review. J Appl Environ Microbiol, 2014;2: 166-175
17. Suriya J, Bharathiraja S, Krishnan M, Manivasagan P, Kim S-K. Marine Microbial Amylases: Properties and Applications. In: Advances in food and nutrition research, vol 79. Elsevier, 2016;pp 161-177
18. Li Y et al.. Modification by α-d-glucan branching enzyme lowers the in vitro digestibility of starch from different sources. International J Biological Macromolecules, 2018;107: 1758-1764
19. Meshram A, Singhal G, Bhagyawant SS, Srivastava N. Plant-Derived Enzymes: A Treasure for Food Biotechnol. In: Enzymes in Food Biotechnol. Elsevier, 2019;pp 483-502
20. Sefidbakht Y, Ranaei Siadat O, Taheri F. Homology modeling and molecular dynamics study on Schwanniomyces occidentalis alpha-amylase. J Biomol Struct Dyna, 2017;35: 574-584
21. Shaw AK et al.. Profiling of BABA-induced differentially expressed genes of Zea mays using suppression subtractive hybridization. RSC Adv, 2017;7: 43849-43865
22. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protocol, 2010;5: 725
23. Mohammed S, Mohammed Z, Galadima AI, Abubakar A, Sadiq A. Rice second florigen is crucial for expansion of its' cultivation areas: A Bioinformatics. Nig J Technol Res, 2018;13: 62-67
24. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol, 1963;7: 95-99