

Homology Modeling and Structural Dynamics of the Glucose Oxidase

Farhan Azhwin Maulana¹, Laksmi Ambarsari^{2,*}, and Setyanto Tri Wahyudi³

¹Master of Biochemistry Program, Postgraduate School, Bogor Agricultural University, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia

²Molecular Biology Division, Department of Biochemistry, Bogor Agricultural University, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia

³Computational Biophysics and Molecular Modeling Research Group, Department of Biophysics, Bogor Agricultural University, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia

* **Corresponding author:**

tel: +62-251-8423267

email: laksmi@apps.ipb.ac.id

Received: September 27, 2018

Accepted: December 4, 2018

DOI: 10.22146/ijc.39135

Abstract: Glucose oxidase from *Aspergillus niger* IPBCC.08.610 (GOD_IPBCC) is a locally sourced flavoenzyme from Indonesia that can potentially be developed in a variety of industrial processes. Although this enzyme has a high activity in catalyzing the redox reactions, the use of this enzyme was still limited to be applied as glucose biosensor. Using information from the amino acid sequences, a computational structure of GOD_IPBCC was therefore designed by homology modeling method using two homologous structures of GOD from protein data bank (1CF3 and 5NIT) as the templates. The quality of the resulting structures was evaluated geometrically for selection of the best model, and subsequently, 50 ns of MD simulations were carried out for the selected model as well as the corresponding template. Results obtained from the validation analysis showed that the 1CF3 template-built structure was selected as the best reliable model. The structural comparison exhibited that the best-modeled structure consisted of two functional domains and three catalytic residues similarly to the corresponding experimental structure. The overall dynamic behavior of the 50 ns of the structure was structurally stable and comparable with that of the positive control both from globally and locally observations. Implications of these stable nature within the best-modeled structure unfold the possibilities in search of notable residues and their roles to enhance enzyme thermostability.

Keywords: glucose oxidase; homology modeling; molecular dynamics; three-dimensional structure

■ INTRODUCTION

Glucose oxidase (GOD), or β -D-glucose, oxygen 1-oxidoreductase is a flavoprotein oxidase (EC 1.1.3.4) catalyzing the redox reaction of β -D-glucose and molecular oxygen to generate δ -gluconolactone and hydrogen peroxide, respectively. In the initial step catalysis (reductive-half reaction), GOD uses flavin adenine dinucleotide (FAD) as a redox carrier through a mechanism of hydride abstraction describing the removal of a proton at hydroxyl group in C1 atom of glucose by His516 and thus facilitates hydride transfer from the anomeric carbon to N5-isoalloxazine ring of FAD,

creating negative charge around N1 of the reduced coenzyme. Subsequently, in the oxidative-half reaction, the reduced enzyme was therefore re-oxidized by oxygen via stepwise single-electron transfers [1-2]. GOD has an important application in the fields of chemical, pharmaceutical, and food industries, which reach far beyond its typical use as blood glucose biosensor in preliminary diagnostics. Economically, this versatile enzyme has gained valuable importance [3].

The source of GOD is extracted mainly from *Aspergillus* and *Penicillium* that have been studied in detail, especially for biotechnological applications. The

GOD produced from *Aspergillus* is more stable but less affinity constant for β -D-glucose than from *Penicillium* [3-5]. In Indonesia, GOD had successfully been produced from a strain of *Aspergillus niger* IPBCC.08.610 (GOD_IPBCC). A study regarding the development of this enzyme concluded that at the enzymatic level, the isolated GOD_IPBCC was ideal only to be utilized for blood glucose measurement both in normal and hyperglycemia situations [5]. These conditions pave the path of developing it into a more diverse utilization, such as glucose-based biofuel cells for example by designing the enzyme with modified kinetics parameters such as low-affinity constant value (K_m) for β -D-glucose but large maximum velocity (V_{max}) to boost power output and improved in thermal resistance [6]. Several mutagenesis studies elucidated that a highly-evolved GOD could merely achieve minor enhancement of its kinetic properties (approximately 5 times lower K_m) [2,7-8]. In the present study, we used the three-dimensional structure-based understanding derived from its encoding genes, as an alternative way to provide extensive knowledge regarding the design of this enzyme suitable for biofuel cell, before the protein experiment in the lab. Unfortunately, the experimental structure of GOD_IPBCC has not been determined.

Experimentally, the structure of a protein can be determined by relatively expensive and time-consuming technics such as crystallography and nuclear magnetic resonance spectroscopy. Moreover, these technics also require high protein purity that needs to be solved. As an alternative, in-silico protein prediction has provided advances in knowledge to compute and predict protein structures based on their amino acid sequences. Among the other computational protein prediction methods, the most reliable approach is by using homology modeling, in which a target sequence is modeled using known structures of candidate proteins where the similarity between their sequences are judged to be similar [9-10]. This methods also have the important benefits in the sense of gaining insightful information about the structure-function relationship in GOD, as previously reported [11].

Exploration of conformational dynamics and stability

in proteins can be observed computationally at the atomic level and complements the analysis of experimental structures by also delineating the underlying dynamics of the protein. For example, the observation of residual rigidity located in the active site using atomistic (MD) simulations has often been found to be prominent than for noncatalytic. This preorganized and rigid active site makes the chemical step more efficient, which positively affects enzymatic rates and minimizes the futile enzyme-substrate encounters. This method is also useful in determining how mutations affect the active-site preorganization, leading to an enhancement in their catalytic efficiency [2].

Therefore, the aim of this study was to predict the three-dimensional structure of GOD_IPBCC by a homology modeling procedure based on the most homologous crystal structures of *Aspergillus niger* GOD, deposited in Protein Data Bank (PDB ID 1CF3 [12] and 5NIT [2] with 97 and 96% identity, respectively). The generated models were screened for geometrically favorable structure and reliability. The Molecular dynamics (MD) simulation was employed to the selected model with the positive control in the hope of revealing the importance of certain elements of the GOD_IPBCC structure in the overall enzymatic thermostability.

■ EXPERIMENTAL SECTION

Computational Section

Hardware

The entire simulations were conducted by a single CPU with specifications of a Quad-core 3.4-GHz Intel® i7 processor, RAM of 16 GB, and Ubuntu version 16.04 operating system.

Software

The multiple sequence alignment was performed using Esript program [13]. All structure visualizations were generated by PYMOL version 2.0 [14]. For minimization and MD simulations, both model and template structures were applied to the AMBER system version 16 [15]. The resulting trajectories were analyzed using cpptraj in the AmberTools16 module to generate RMSD and RMSF profiles using the initial structure of each simulation as a reference. The B-Factors analysis of

wild-type GOD (PDB ID 1CF3) was performed using B-Fitter program [16]. The visual molecular dynamics (VMD) [17] was used to produce electrostatic energy and solvent accessible area data. The generated plots except in RMSF were smoothed with moving average method in Microsoft Excel before displayed into the plots.

Procedure

Model construction and evaluation

The protein sequence of the GOD_IPBCC with GB (GenBank) accession number MH593586 was obtained from NCBI (<https://www.ncbi.nlm.nih.gov/>). The amino acid properties of GOD_IPBCC was computed by ProtParam such as molecular weight, aliphatic index, grand average of hydropathy (GRAVY), and instability index [18]. The protein was modeled by SWISS-MODEL [19] using two known crystallographic structures (PDB ID 1CF3 [12] and 5NIT [2]). The evaluation of these structures was performed such as the residual percentage of Ramachandran plot and overall G-factor value using PROCHECK [20]. Another evaluation tools (molProbity and proQ) were also carried out [21]. Moreover, the functional regions of the best protein were then predicted using the Pfam web server [22]. The secondary structural elements (α -helices and β -strands) were analyzed by PDBsum and other components present in the protein tertiary structure [23].

Molecular dynamics simulation

MD simulations of the best model were carried out in apo-monomer form. The preparation step was performed by firstly removing the hydrogen atoms and other organic molecules within the best model and its corresponding template. Subsequently, the new hydrogen atoms were added and the protonation state of the titratable side chain within the proteins was adjusted by the Virginia Tech H++server (<http://biophysics.cs.vt.edu/H++>) [24-25]. These structures were then prepared in an explicit solvent solvated box implementing TIP3P water molecules with the box distance of 16 Å. The sodium ions were added to neutralize the system.

The six series of energy minimizations were run with a total of 60000 steps and 50000 steps for final stage by using constrained steepest descent followed by conjugate

gradient algorithm methods. After which the system was gradually heated from 0 to 300 K with 50 K interval using the implementation of the Langevin dynamics. After the desired temperature was attained, the equilibration phase was performed in a total of 6 stages to ensure the stability of structural properties concerning time. In the initial equilibration, the constraint was first applied in a fixed volume run (NVT) for 50 ps. In the subsequent phase, under constant pressure (NPT), the constraint was then removed slowly for 300 ps in total. This allows the system to obtain a proper density and is likewise to avoid aggregation [26]. The final step of the simulation was a long production run, where the solvated protein was run for 50 ns of simulation in 25 separate stages of 2.0 ns each (2.0-fs timestep). The non-bonded cutoff value used was 10 Å and the long-range electrostatic energy was calculated using the Particle Mesh Ewald (PME) algorithm.

RESULTS AND DISCUSSION

Protein Modeling and Validation

A number of 605 amino acid sequences of the GOD_IPBCC was successfully obtained for further analysis. The physicochemical parameters from ProtParam calculation revealed the majority of its primary structure comprised non-polar residues with rich alanine content and had the monomeric average molecular weight of 63 kDa, whereas the aliphatic index of protein (83.63) showed the protein could withstand for a wide range of temperature. Moreover, this enzyme also had negative GRAVY score (-0.208) which signifies a better interaction with the solvent due to its greater hydrophilicity. These results along with instability index (28.88) rendered information that the properties of the GOD_IPBCC primary structure was categorized good and needed to be discussed even further.

The model of GOD_IPBCC was constructed using two templates that shared over 95% homologous sequences. The first template was a structure of *A. niger* glucose oxidase (PDB ID 1CF3 [12]), whereas the second was the mutant form (PDB ID 5NIT [2]). The final resolution of these structures was resolved at 1.9 and 1.87 Å, respectively. The residual alignment of

GOD_IPBCC with the templates is displayed in Fig. 1. Most of GOD_IPBCC amino acids were homologous (*) with 16 distinct residues found when compared with both 5NIT and 1CF3 experimental structures. Other than that, the first 22 amino acid residues of GOD_IPBCC were

observed as peptide signals and therefore excluded in the three-dimensional structure as depicted in Fig. 2.

The model generated by SWISS-MODEL results in the QMEAN statistical value that provides how native the structure in a global scope [27]. The global QMEAN



Fig 1. Multiple sequence alignment of GOD_IPBCC and the selected templates

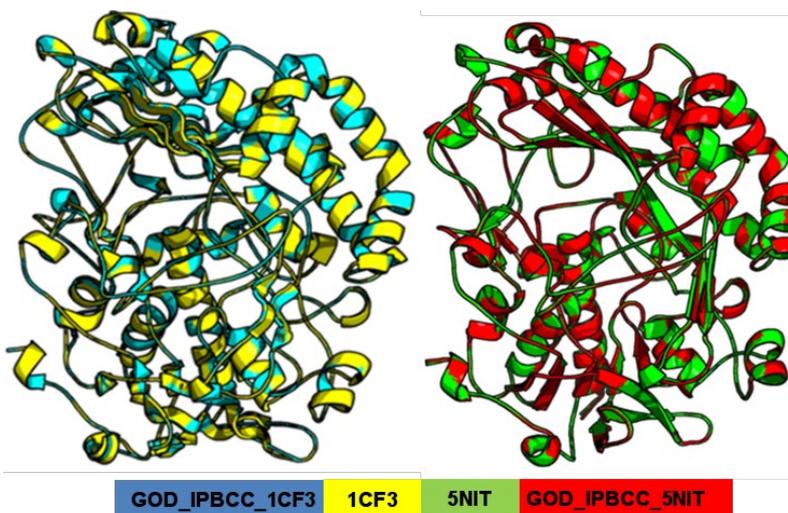


Fig 2. Structural alignment of GOD_IPBCC homology models and corresponding templates in cartoon representation

profile of the 1CF3 template-built model (0.04) rendered a high degree of nativeness compared with the 5NIT template-generated model (-0.48). Nevertheless, the initial constructed protein model from a comparative modeling study does not ensure a valid 3D structure since it contains the possibility of mistakenly folded from the native conformation. Thus, the quality of the computationally determined protein structure needed to be analyzed.

The quality estimation of GOD_IPBCC models was evaluated by PROCHECK consisting of Ramachandran plot and overall G-factor value. As shown in Table 1, the highest percentage of residues (89.4%) were found similar in the favored region for 5NIT experimental structure and GOD_IPBCC_1CF3 protein model, whereas 1CF3 template showed the lowest score (88.4%). The Ramachandran percentage for GOD_IPBCC_5NIT showed one residue less (89.2%) in the favored area compared to the homology model from the 1CF3 crystal structure. None of the residues were located in the outlier region (0%) for all structures. In general, a protein with good backbone dihedral angles (ϕ and ψ) of the amino acid residues are expected to have 90% of the residues in the core or

favored region and also has an overall geometrical factor (G-factor) above -0.5 and vice versa as an indication of having a stereochemical agreement [28-29].

Moreover, the successful attempt in constructing the structure of GOD_IPBCC was indicated when comparing the quality of each model generated by two different templates (1CF3 and 5NIT). The models which were built from both templates respectively were good and reliable since nearly 90% residues fell into the favored regions. However, the overall geometry score (G-factor) found in GOD_IPBCC_1CF3 (-0.08) structure was greater than that of GOD_IPBCC_5NIT (-0.16), indicated the protein is in agreement with defined values.

In addition, the overall molProbity scores for 1CF3 template-built model structure exhibited the lowest (the highest crystallographic resolution) than that of GOD_IPBCC_5NIT and even those two templates (Table 2), whereas results from the proQ (LGscore and Maxsub) showed all input structures were categorized as the very good model. Nonetheless, this validation analysis clearly explained that the modeled structure of GOD_IPBCC_1CF3 was superior in overall quality. As

Table 1. Summary of the protein quality assessed by PROCHECK

Protein	Ramachandran plot (%) ^a		Overall G-factor ^b
	Favored region	Disallowed region	
5NIT	89.4	0.0	-0.08
GOD_IPBCC_5NIT	89.2	0.0	-0.16
1CF3	88.6	0.0	0.09
GOD_IPBCC_1CF3	89.4	0.0	-0.08

^aRamachandran plot consists of the percentage of residues plotted into one internal coordinate (i.e., favored, allowed, generally allowed, and disallowed region)

^bGeometrical factors (G-factors) divided into the quality of covalent and overall dihedral distances. For a reliable model, the score should be greater than -0.5

Table 2. Summary of the protein quality assessed by PROCHECK

Protein	proQ ^a		MolProbity score ^b
	LGscore	Disallowed region	
5NIT	6.834	0.533	1.51
GOD-IPBCC-5NIT	6.815	0.526	1.38
1CF3	7.009	0.546	1.16
GOD-IPBCC-1CF3	6.944	0.531	0.91

^aProQ is calculated based on a number of structural characteristics and optimized to uncover native structures comprising LGscore and Maxsub. A protein with good quality should have LGscore above 5 and Maxsub ranges from 0 to 1, where 0 is insignificant and vice versa

^bMolProbity score incorporates Ramachandran outliers, clashscore, and bad rotamer into a single score, normalized to be on the same scale as X-ray resolution

the best model, further characterization of the secondary structural information was therefore conducted by PDBsum.

Structural Features of the *A. niger* 1CF3 Template-built Structure

Analysis of the GOD_IPBCC_1CF3 secondary structure was found to have 28 α -helices, 23 β -strands, 23 helix-helix interactions, 1 disulfide bridge, 45 beta, and 18 gamma turns, respectively. This results correlated well with the domain features of *A. niger* glucose oxidase crystal structure (PDB ID 1CF3), where consists of two following functional regions; FAD-binding domain and C-terminal domain (substrate binding) as is apparent from Fig. 3(a). These domains are characterized by two separate and distinctive β -sheet systems, one of which is a sandwich form comprising five-stranded sheet incorporated into FAD-binding domain, the other one contains a large antiparallel β -sheet bolstered by four α -helices that subsequently form one side of the active site [30].

Moreover, three catalytic residues Glutamate-410 (E410), Histidine-514 (H514), and Histidine-557 (H557) (equivalent to E412, H516, and H559 respectively in experimental structure) were also positioned in accordance with the same conformation in the 1CF3 experimental structure [12]. The hydrogen bond formed between E412 and H559 in 2.7 Å atomic distance plays a vital role in maintaining the reactivity on the catalytic site. The catalytically important H516 involves as proton acceptor from the anomeric carbon of glucose and has found to be flexible in the wild-type [2]. A single conserved disulfide bridge by Cysteine-162 (C162) and Cysteine-204 (C204) (equivalent to C164 and C206 in the crystal structure) were located in the protein surface area. This region facilitates the path for electron transfer from the flavin oxygen-4 (O4) [12]. In respect to the sulfur atoms in the *A. niger* GOD crystallographic structure, Marín-Navarro et al. [11] elucidated a new sulfur- π interaction formed by residual mutation of threonine to methionine in position 554

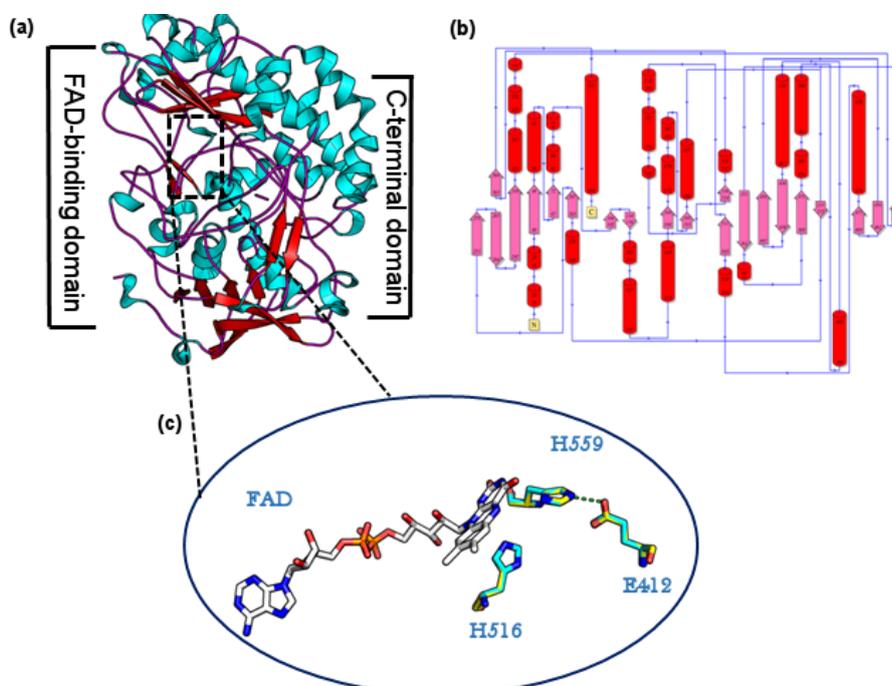


Fig 3. (a) The domain view of the GOD_IPBCC_1CF3 secondary structural model. (b) Topology diagram of the GOD_IPBCC_1CF3 protein with helix, sheet, beta turn, and gamma turn simulated by PDBsum. (c) Stick representation of FAD and three amino acids responsible for catalytic activity of GOD_IPBCC_1CF3 as according to Hecht et al. [30]. The hydrogen bond is shown in green dashed line

(T554M) that has greater thermally stabilization energy (4.2–12.6 kJ/mol) when compared with that of hydrogen bond-associated (1.3–6.3 kJ/mol). On the whole observation, the biochemical characterization within the GOD_IPBCC_1CF3 is correlated well to its experimental structure.

Stabilization of the Structure by MD Simulations

Validation of simulation was performed using comparative analysis of the calculated average B-factor for each residue between the experimental and simulation structure (Fig. 4(a)). The simulation resulted in higher B-factor compared with x-ray but showed a similar trend, indicates the MD performance was reasonably good agreement in total with the crystallographic data.

The 50 ns of simulation was applied to investigate the stability of the positive control (PDB ID 1CF3) and its homology model. The global stability of these structures was monitored by RMSD backbone. The modeled structure was found to be generally stable during the simulation, comparable to RMSD value of 1CF3 template (Fig 4(b)). In the absence of FAD (apo form), this enzyme conformation may result in the inactive form that leads to dissociation of its tertiary structure (less stable) either experimentally or computationally [31]. The overall structural stability of the single monomer of GOD_IPBCC_1CF3 was considered to be good as indicated by $C\alpha$ RMSD $< 2 \text{ \AA}$.

Furthermore, the residual fluctuation of all systems was analyzed through RMSF calculation (Fig. 5). The greatest fluctuation was found in the C-terminal because this region was unrestrained, whereas the fluctuation around the catalytic residues was relatively high. This might be due to the absence of the substrate that altered all catalytic systems from open to closed conformation and their secondary structure as well, as previously reported [29,31]. The difference in RMSF values between two structures was observed from residue numbers 160 to 162; these residues adopted a turn shape in 1CF3 which fluctuates more in nature than the same residues in its homology model as 3(10)-helical conformation. These results were due to the formation of two additional intermolecular hydrogen bonds between Phenylalanine-158 (F158) and Asparagine-159 (N159) and between Alanine-160 (A160) and Histidine-163 (H163) in GOD_IPBCC_1CF3 which were not found in its template (Fig. 5). These interactions can reversibly switch the two protonation forms of histidine (and pKa value) [32] and are regarded to be important in many biological systems [29,33]. The other natural, flexible residues ranging from the number 257 to 262 formed loops in both 1CF3 and GOD_IPBCC_1CF3 structures were also high in RMSF. These fluctuating regions can be targeted for improved enzyme thermostability [34].

The stability of the protein conformation during MD was determined by calculating its non-bonded energy

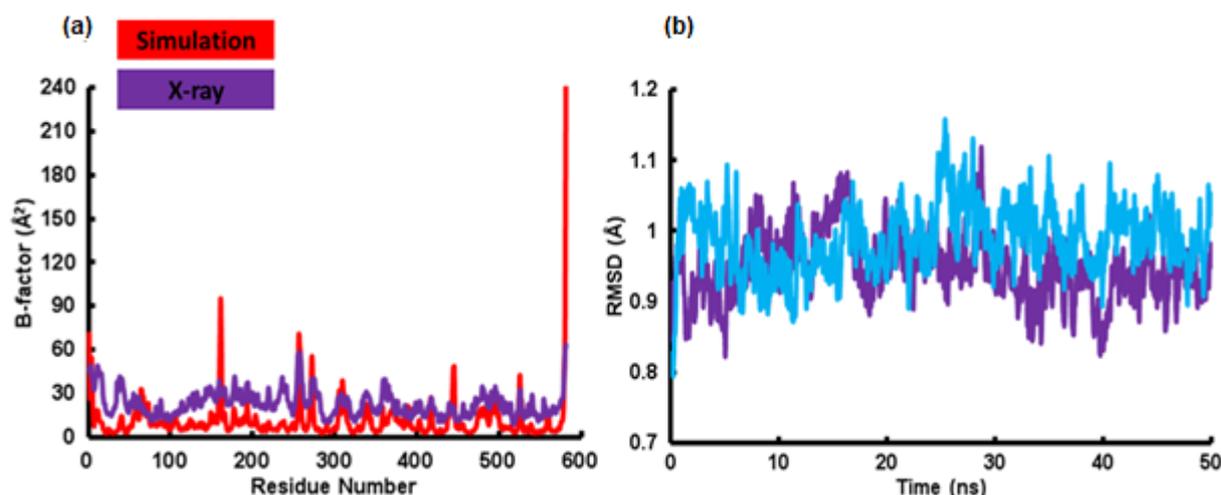


Fig 4. (a) Comparison of amino acid B-factor in GOD. (b) The RMSD trajectory for polypeptide chain backbone throughout 50 ns of simulation

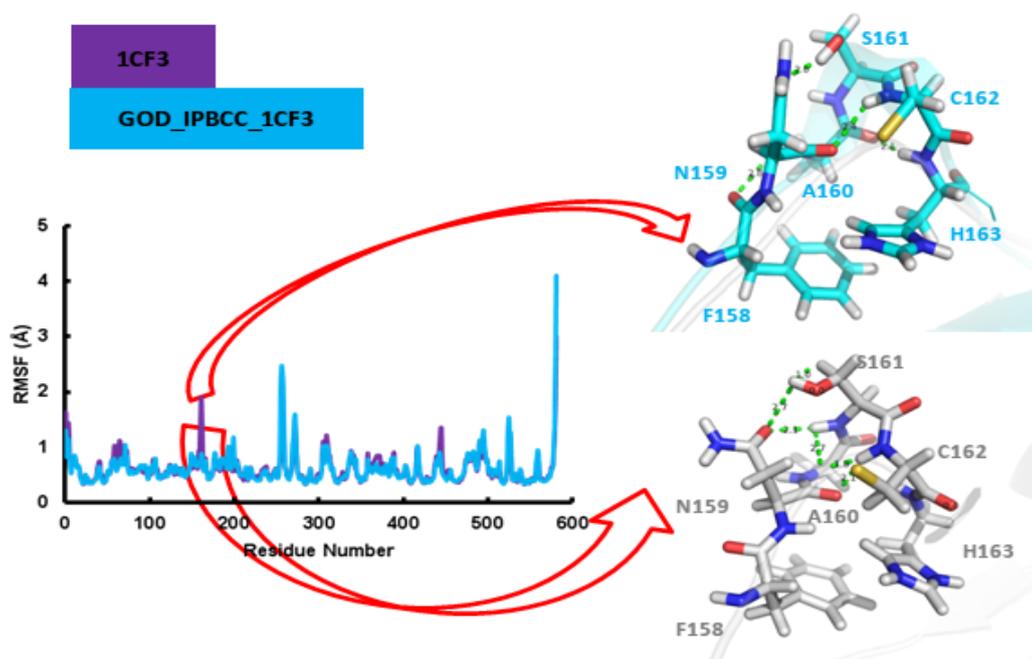


Fig 5. Residual RMSF profile of all systems during 50 ns of simulation. Panel on the right display details of fluctuating residues in the 50 ns structure of GOD_IPBCC_1C3 (blue stick) and 1CF3 (gray stick). The hydrogen bond is represented in green dashed line

which is plotted in Fig. 6(a). These interactions play the ubiquitous role in steering the folding route and modulating stability. The electrostatic strength has two distinctive effects, direct charge-charge interactions, and long-range effects, which together contribute to funneling the binding landscape of predicted structure [33].

In the positive control, the total average of non-bonded energy (coulomb electrostatic and van der Waals) over the course of MD simulation can be observed higher

than GOD_IPBCC_1CF3, the lowest fluctuation was observed at 27 ns (approximately -18700 kcal/mol) in the latter structure. The hydrophobicity content between these structures is not so different since their structural homology is high. However, the 1CF3 template-built structure conformation consistently exposed a decreased total surface area (Fig 6(b)) than that of the template during the first ns of trajectory, meaning the content of hydrophobic and hydrophilic surfaces in the structure

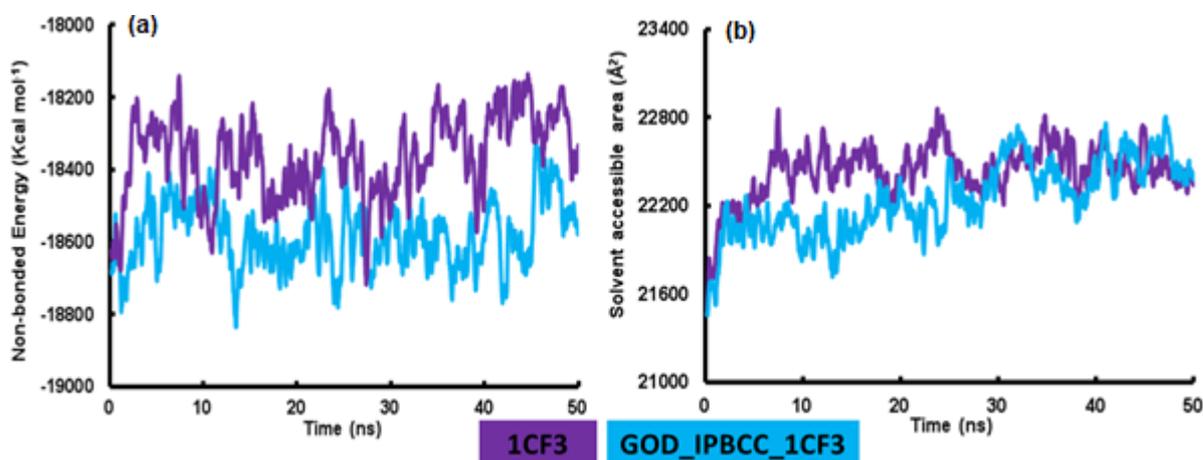


Fig 6. (a) Total non-bonded energy (in kcal/mol) and (b) protein SASA profile during 50 ns of simulation

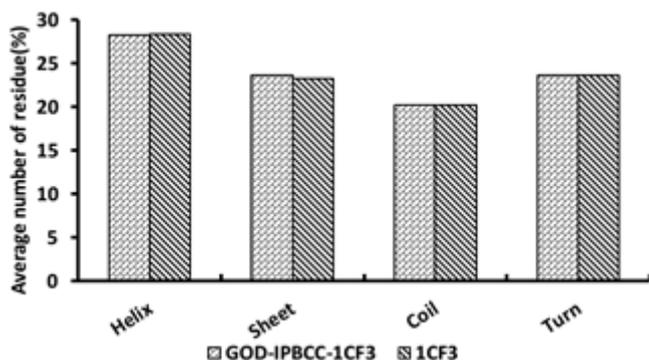


Fig 7. Average number of residual occurrence forming secondary structure during 50 ns of simulation

were in the equilibrium state and finally comparable at the end of simulation to around 22300 Å. Together, these trends imply that the modeled structure is modulated by increasing the non-bonded strength and having a stable hydrophobic context leading to the integrity of the tertiary structure.

An investigation related to a change in secondary structure (SS) is essential for enzymes to carry out their biological activity [31,35]. On overall observation, we found that the percentage of secondary structure for both structures was nearly similar where α -helices occupied the major proportion (approximately 28%) among the other structures (Fig. 7). Interestingly, these structures were constant over the course of a simulation, in agreement with the previously reported simulation of GOD [36]. An experimental study about a slight change in GOD secondary structure upon inactivation has been explained under ambient condition, using analysis of the circular dichroic (CD) spectra at 470 nm confirmed that the loss of its secondary structure is minimal [37].

The MD simulations performed revealed that the 50 ns of trajectory within GOD_IPBCC_1CF3 did not show any peculiar features and could retain its conformation on the basis of analyzed parameters. On the long run, the further computational experiment is necessary to figure out the dynamics of this enzyme together with the substrate docked and performs a potential residual mutation having a significant improvement in terms of the thermostability. Structure prediction along with MD simulation will provide more fruitful information regarding the role of each residue components, especially

around the catalytic site to rationalize the design of an engineered enzyme before conducting the laboratory protein experiment.

■ CONCLUSION

In conclusion, the limited use of GOD_IPBCC as only for biosensor prompted the understanding of the structure-function relationship corresponding to the encoding gene. Since it shared high homology level with the GOD from protein data bank (1CF3 [2] and 5NIT [12]), the structural comparison between the constructed structures resulting from their respective templates showed that the GOD_IPBCC_1CF3 can be considered reliable structure not only due to the in-silico methods but the conserved structural features associated with fungal glucose oxidase enzymes. Molecular dynamics simulations denoted that the overall conformation in 50 ns of trajectory did not undergo any significant conformational changes both from the secondary and tertiary structure and form a well-packed structure. This structure will later be used as an initial structure to investigate the thermostability in more details.

■ ACKNOWLEDGMENTS

The support for this research was provided by University Consortium Thesis Grant (Ref. No. GCS17-2758) in part of the scholarship of Southeast Asian Regional Center for Graduate Study and Research in Agriculture (SEARCA).

■ REFERENCES

- [1] Leskovic, V., Trivić, S., Wohlfahrt, G., Kandrač, J., and Peričin, D., 2005, Glucose oxidase from *Aspergillus niger*: The mechanism of action with molecular oxygen, quinones, and one-electron acceptors, *Int. J. Biochem. Cell Biol.*, 37 (4), 731–750.
- [2] Petrović, D., Frank, D., Kamerlin, S.C.L., Hoffmann, K., and Strodel, B., 2017, Shuffling active site substrate populations affects catalytic activity: The case of glucose oxidase, *ACS Catal.*, 7 (9), 6188–6197.
- [3] Bankar, S.B., Bule, M.V., Singhal, R.S., and Ananthanarayan, L., 2009, Glucose oxidase-An overview, *Biotechnol. Adv.*, 27 (4), 489-501.

- [4] Kiess, M., Hecht, H.J., and Kalisz, H.M., 1998, Glucose oxidase from *Penicillium amagasakiense*. Primary structure and comparison with other glucose-methanol-choline (GMC) oxidoreductases, *Eur. J. Biochem.*, 252 (1), 90–99.
- [5] Rohmayanti, T., Ambarsari, L., and Maddu, A., 2017, Enzymatic activity of glucose oxidase from *Aspergillus niger* IPBCC.08.610 on modified carbon paste electrode as glucose biosensor, *IOP Conf. Ser.: Earth Environ. Sci.*, 58 (1), 12046.
- [6] Zhu, Z., Momeu, C., Zakhartsev, M., and Schwaneberg, U., 2006, Making glucose oxidase fit for biofuel cell applications by directed protein evolution, *Biosens. Bioelectron.*, 21 (11), 2046–2051.
- [7] Holland, J.T., Lau, C., Brozik, S., Atanassov, P., and Banta, S., 2011, Engineering of glucose oxidase for direct electron transfer via site-specific gold nanoparticle conjugation, *J. Am. Chem. Soc.*, 133 (48), 19262–19265.
- [8] Holland, J.T., Harper, J.C., Dolan, P.L., Manginell, M.M., Arango, D.C., Rawlings, J.A., Aplett, C.A., and Brozik, S.M., 2012, Rational redesign of Glucose oxidase for improved catalytic function and stability, *PLoS One*, 7 (6), e37924.
- [9] Altschul, S.F., Boguski, M.S., Gish, W., and Wootton, J.C., 1994, Issues in searching molecular sequence databases, *Nat. Genet.*, 6 (2), 119–129.
- [10] He, Y., Rackovsky, S., Yin, Y., and Scheraga, H.A., 2015, Alternative approach to protein structure prediction based on sequential similarity of physical properties, *Proc. Natl. Acad. Sci. U.S.A.*, 112 (16), 5029–5032.
- [11] Marín-Navarro, J., Roupain, N., Talens-Perales, D., and Polaina, J., 2015, Identification and structural analysis of amino acid substitutions that increase the stability and activity of *Aspergillus niger* glucose oxidase, *PLoS One*, 10 (12), e0144289.
- [12] Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H.M., and Hecht, H.J., 1999, 1.8 and 1.9 Å resolution structures of the *Penicillium amagasakiense* and *Aspergillus niger* glucose oxidases as a basis for modelling substrate complexes, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 55 (Pt 5), 969–977.
- [13] Robert, X., and Gouet, P., 2014, Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.*, 42 (Web Server issue), W320–W324.
- [14] DeLano, W.L., 2002, The PyMOL molecular graphics system, *Proteins*, 30, 442–454.
- [15] Case, D.A., Betz, R.M., Botello-Smith, W., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T.S., LeGrand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, B., Mermelstein, D., Merz, K.M., Monard, G., Nguyen, H., Nguyen, H.T., Omelyan, I., Onufriev, A., Roe, D.R., Roitberg, A., Sagui, C., Simmerling, C.L., Botello-Smith, W.M., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu, X., Xiao, L., and Kollman P.A., 2016, *AMBER 2016*, University of California, San Francisco.
- [16] Reetz, M.T., Carballeira, J.D., and Vogel, A., 2006, Iterative saturation mutagenesis on the basis of b factors as a strategy for increasing protein thermostability, *Angew. Chem. Int. Ed.*, 45 (46), 7745–7751.
- [17] Humphrey, W., Dalke, A., and Schulten, K., 1996, VMD: Visual molecular dynamics, *J. Mol. Graphics*, 14 (1), 33–38.
- [18] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A., 2005, “Protein Identification and Analysis Tools on the ExPASy Server” in *The Proteomics Protocols Handbook*, Eds., Walker J.M., Humana Press, New York City.
- [19] Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., and Swede, T., 2014, SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Res.*, 42 (Web Server issue), W252–W258.
- [20] Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., and Thornton, J.M., 2018, PDBsum: Structural summaries of PDB entries, *Protein Sci.*, 27 (1), 129–134.

- [21] Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C., 2010, MolProbity: All-atom structure validation for macromolecular crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 66 (Pt 1), 12–21.
- [22] Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., and Bateman, A., 2016, The Pfam protein families database: Towards a more sustainable future, *Nucleic Acids Res.*, 44 (D1), D279–D285.
- [23] Laskowski, R.A., 2001, PDBsum: Summaries and analyses of PDB structures, *Nucleic Acids Res.*, 29 (1), 221–222.
- [24] Anandakrishnan, R., Aguilar, B., and Onufriev, A.V., 2012, H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations, *Nucleic Acids Res.*, 40 (Web Server issue), W537–W541.
- [25] Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A., 2006, A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules, *Proteins*, 63 (4), 928–938.
- [26] Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J., 2005, The Amber biomolecular simulation programs, *J. Comput. Chem.*, 26 (16), 1668–1688.
- [27] Benkert, P., Biasini, M., and Schwede, T., 2011, Toward the estimation of the absolute quality of individual protein structure models, *Bioinformatics*, 27 (3), 343–350.
- [28] Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M., 1993, PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 26, 283–291.
- [29] Yusuf, M., Baroroh, U., Hasan, K., Rachman, S.D., Ishmayana, S., and Subroto, T., 2017, Computational model of the effect of a surface-binding site on the *Saccharomycopsis fibuligera* R64 α -amylase to the substrate adsorption, *Bioinf. Biol. Insights*, 11, 1177932217738764.
- [30] Hecht, H.J., Kalisz, H.M., Hendle, J., Schmid, R.D., and Schomburg, D., 1993, Crystal structure of glucose oxidase from *Aspergillus niger* refined at 2.3 Å resolution, *J. Mol. Biol.*, 229 (1), 153–172.
- [31] Janati-Fard, F., Housaindokht, M.R., and Monhemi, H., 2016, Investigation of structural stability and enzymatic activity of glucose oxidase and its subunits, *J. Mol. Catal. B: Enzym.*, 134 (Part A), 16–24.
- [32] Liao, S.M., Du, Q.S., Meng, J.Z., Pang, Z.W., and Huang, R.B., 2013, The multiple roles of histidine in protein interactions, *Chem. Cent. J.*, 7 (1), 44.
- [33] Meyer, E.A., Castellano, R.K., and Diederich, F., 2003, Interactions with aromatic rings in chemical and biological recognition, *Angew. Chem. Int. Ed.*, 42 (11), 1210–1250.
- [34] Yu, H., Yan, Y., Zhang, C., and Dalby, P.A., 2017, Two strategies to engineer flexible loops for improved enzyme thermostability, *Sci. Rep.*, 7, 41212.
- [35] Tsai, M.Y., Zheng, W., Balamurugan, D., Schafer, N.P., Kim, B.L., Cheung, M.S., and Wolynes, P.G., 2016, Electrostatics, structure prediction, and the energy landscapes for protein folding and binding, *Protein Sci.*, 25 (1), 255–269.
- [36] Todde, G., Hovmöller, S., Laaksonen, A., and Mocci, F., 2014, Glucose oxidase from *Penicillium amagasakiense*: Characterization of the transition state of its denaturation from molecular dynamics simulations, *Proteins*, 82 (10), 2353–2363.
- [37] Caves, M.S., Derham, B.K., Jezek, J., and Freedman, R.B., 2011, The mechanism of inactivation of glucose oxidase from *Penicillium amagasakiense* under ambient storage conditions, *Enzyme Microb. Technol.*, 49 (1), 79–87.