## In Silico Analysis of rbcl Protein and D-Ribulose 1,5-bisphosphate Bond

Ihwan<sup>1,2\*</sup>, Dewi Ratih Tirto Sari<sup>1,3</sup>, Luchman Hakim<sup>1</sup>, Catur Retnaningdyah<sup>1</sup>, Estri Laras Arumingtiyas<sup>1</sup>

<sup>1</sup>Department of Biology, Faculty of Mathematics and Natural Sciences, Brawijaya University, Jl. Veteran, Malang <sup>2</sup>Department of Biology Education, University of Muhammadiyah Kupang, Kupang - NTT) <sup>3</sup>Research Center for Smart Molecules and Natural Genetic Resources, Brawijaya University, Jl. Veteran, Malang, Jawa

Timur, Indonesia 65145

Submission: January 2021; Revised: January 2021; Accepted: March 2021

\*Corresponding author: Ihwan; e-mail: ihwan.fkipbio@yahoo.com; telp.: 081339292655

**ABSTRACT.** The aim of this study was to determine the bond between the protein rbcL and D-Ribulose 1,5-bisphosphate by insilico. DNA sequences of the Mangrove Rhizophora mucronate rbcL gene DNA sequences from 7 different sources were obtained from NCBI for further alignment analysis using BioEdit software, phylogeny analysis using Mega6 software, molecular docking using PyRx software, preparation and visualization of docking results using Biovia Discovery Studio Visualizer software and analysis of the protein model quality based on the number of amino acid residues (Ramachandran plot analysis). The results of the docking molecular analysis showed interaction of 9 hydrogen bonds namely Asp203, Thr173, His294, Glu204, His327, Ser379, His298, Arg295, and Asn123 and 2 unfavorable bonds namely Lys177 and Lys175. This ligand and protein interaction complex was of good quality because the amino acid residues were in the most favored regions was greater than the amino acid residues in the disallowed regions outcomes.

Keywords: Insilico, rbcL protein, ribulose, biphosphate

#### **INTRODUCTION**

Molecular identification of genetic diversity could be conducted by using data sources from nuclear DNA (nDNA), mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA), with the data source widely used in plant researches is chloroplast DNA (cpDNA) [1]. It has a stable structure and a small genome [2]. The genome is more conservative with a low rate of nucleotide substitution [3], and the genome does not undergo recombination and inherited uniparentally [1]. Chloroplast DNA (cpDNA) is circular with a size range of 85-2000 kilobases (kb). cpDNA controls the transfer RNA (tRNA), ribosomal RNA (rRNA), and most of the proteins contained in chloroplast organelles. There are at least 19 known chloroplast DNA sequences, with hundreds of genes contained in the cpDNA genome. There are several codes for the subunits that make up the photosynthetic protein complex, namely for photosystem I, photosystem II, ribulose-1,5-biphosphate carboxylase-oxygenase, cytochrome b6-f complex, and ATP synthase [2]. Vascular plants generally have similar content and function of genes in cpDNA [3].

DNA barcode is a short sequence of DNA from a gene used to identify specimens of living things, both plants and animals [4]. There are four coding genes in plants: matK, rbcL, rpoB, and rpoC1), and 3 noncoding spacer genes, namely atpF - atpH, trnH psbA, and psbK - psbI [5]. The DNA barcode sequence described in this article is the rbcL gene as the standard barcode for plant DNA.

One of the key proteins encoded by cpDNA is ribulose-1,5biphosphate carboxylase-oxygenase (abbreviated as RuBisCO), which participates in carbon fixation in the photosynthesis process. RuBisCO makes up about 50% of the protein found in green plants which are thought to be the most abundant protein on earth [3]. The RuBisCO protein complex is composed of eight identical large protein subunits and eight identical small protein subunits. Large subunit proteins are encoded by cpDNA, whereas nuclear DNA encodes the small subunit proteins.

The rbcL gene measures about 1400 bp in length so it provides many characters for phylogenetic studies [5]. The role of the rbcL gene encoding the RuBisCO protein is thought to cause this gene sequence to have a low mutation rate compared to other barcode genes in cpDNA and therefore, the level of similarity between species is quite high [6]. This low mutation rate provides the advantage for in-depth study of intraspecies genetic and phylogenetic variation.

The rbcL gene can be amplified with high success rates with one or two kinds of universal primers [5]. It was further argued that when compared with other barcode gene candidates, the rbcL gene had a high success rate of bidirectional sequencing (a two-way sequencing process with forward and reverse primers). That the success rate of the rbcL gene could reach 100% in 251 plant species with only two kinds of primers [7]. That the database on the rbcL gene was owned by many species, making it easier to compare in data analysis [8].

### **METHODS**

D-Ribulose 1,5-bisphosphate and protein was prepared using the Discovery Studio Visualizer software. 3D structure (3 dimensions) of D-1.5-bisphosphate Ribulose compound was from the website downloaded https://pubchem.ncbi.nlm.nih.gov/ in SDF file format and then stored in the pdb file format. The structure of the compounds to be docked is shown in Figure 1.



Figure 1. Structure of D-Ribulose 1,5-bisphosphate Compound; a. 3D, b. 2D Structure of D-Ribulose 1,5-bisphosphate Compound (PubChem ID : 123658)

Protein structure of rbcL with pdb document format (protein data bank) was downloaded via the macromolecular RCSB biology structure information portal (http://www.rcsb.org/pdb/). The protein downloaded was the rbcL protein (ID: 5IU0). Protein preparation was conducted by cleaning the protein from the water molecule bonds and ligands that contaminate the protein structure using the Discovery Studio Visualizer and stored in the pdb file format. Furthermore, the preparation was carried out by minimizing the energy of the compound and then molecular docking was carried out using the PyRx application. The protein structure to be docked is shown in Figure 2 below.



Figure 2. The 3D Structure of rbcL Protein in Arabidopsis thaliana (ID: 5IU0)

Visualization of the docking results was done using Discovery Studio Visualizer and PyMol software. The result visualized was the interaction between D-Ribulose 1,5-bisphosphate and rbcL. The resulting complexes were also analyzed with a Ramachandran plot.

## **RESULTS AND DISCUSSIONS**

The of ribulose characteristics the bisphosphate carboxylase was illustrated in Figure 4. The grey color represented the small chain, while the yellow color was the large chain Ribulose bisphosphate carboxylase (Figure 4a). Ramachandran plot was used to determine the intrinsic quality of protein 3D structures [9]. The three-dimensional structure model's feasibility could be seen in the Ramachandran plot divided into four quadrants, namely most favored regions, additional allowed regions, generously allowed disallowed regions. and regions. The Ramachandran Plot model above shows that the rbcL protein (ID: 5IU0) has different amino acid residues in each region, the green color (Figure 4b) shows the amino acids of Ribulose bisphosphate carboxylase protein. In quadrant I, most favored regions have a greater number of amino acid residues than other quadrants, namely quadrant II additional regions, quadrant III generously allowed regions, and quadrant IV disallowed regions.

# Docking results of protein interactions

To show that the complex between D-Ribulose 1,5-bisphosphate and rbcL is a complex that has a good quality is to describe the D-Ribulose 1.5bisphosphate complex with rbcL into а Ramachandran plot. Through the Ramachandran plot, it could be seen whether a protein structure has good quality or not. It is by looking at the plot of non-glycine residues located in the disallowed regions of the dihedral angle. Glycine is an amino acid that has no side chains so that the angles  $\Phi$  and  $\psi$  can be in the four quadrants of the Ramachandran plot. A protein structure is declared good if the presence of amino acid residues in the most favored regions is greater than the amino acid residues in disallowed regions [10]. To make a complex Ramachandran plot between D-Ribulose 1,5bisphosphate and rbcL stored in .pdb form then illustrated with the RAMPAGE server. Α Ramachandran plot between D-Ribulose 1.5bisphosphate and rbcL is shown in Figure 5d.



**Figure 5.** Complex interaction between D-Ribulose 1,5-bisphosphate and Ribulose bisphosphate carboxylase, a. 3D overview and interaction view of D-Ribulose 1,5-bisphosphate + Ribulose bisphosphate carboxylase in a and b, 2D view of interaction showed as c, Ramachandran plot in d, and the interaction table was described in e.

The ligands and 9 amino acid residues' interaction was a hydrogen bond with a conventional type of hydrogen bond. The hydrogen bond in the amino acid residue His294 has a stronger bond compared to His298 and His327. Hydrogen bonds made up the interaction very strong if the distance generated from hydrogen to the receptor gets smaller [11-13]. **D**-Ribulose Ramachandran plot of 1.5 bisphosphate + ribulose bisphosphate carboxylase revealed that the number of amino acid residues in the favour region was greater than the number of amino acid residues in other sites. Furthermore, amino acid residues in the disallowed region showed at least one amino acid. These results indicate that this ligand and protein interaction complex is of good quality. The greater the amino acid residue in the most favored region, suggested that the model's quality is proper and stable [10].

#### CONCLUSION

The protein model of ribulose bisphosphate carboxylase (rbcL, ID: 5IU0) has 9 hydrogen bonds Asp203, Thr173, His294, Glu204, His327, Ser379, His298, Arg295, and Asn123 and 2 unfavorable bonds namely Lys177 and Lys175.

#### REFERENCES

- Provan J, Powell W, Hollingsworth PM. 2001. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends Ecol Evol*,16(3):142–7.
- [2] Snustad, D.P and M.J. Simmons. 2015. Principles of Genetiks (Seventh Edition). Massachusetts: The McGrawHill Companies.
- [3] Pierce, B. A. 2017. Genetics: A Conceptual Approach (seventh edition). New York: W. H. Freeman and Company.
- [4] Hebert PDN, Cywinska A, Ball SL, DeWaard JR. 2003. Biological identifications through DNA barcodes. Proc R Soc B Biol Sci.270(1512):313–21.
- [5] P Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, and R. Ratnasingham. 2012. A DNA barcode for land plants area. 106(31):12794–7.
- [6] Kellogg EA, Juliano ND. 1997. The structure and function of RuBisCO and their implications for systematic studies. *Am J Bot*.84(3):413–28.
- [7] Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, et al. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One*, 3(7).

- [8] Cummings MP, Nugent JM, Olmstead RG, Palmer JD. 2003. Phylogenetic analysis reveals five independent transfers of the chloroplast gene rbcL to the mitochondrial genome in angiosperms. *Curr Genet.*, 43(2):131–8.
- [9] Petsko, G.A., & D. Ringe. 2004. Protein structure and function: from sequence to consequence. New Science Press Ltd, London.
- [10] Ho BK, Brasseur R. 2005. The Ramachandran plots of glycine and preproline. *BMC Struct Biol.*,5:1–11.
- [11] Y Bare, S Mansur, SSND Tiring, DRT Sari, A Maulidi. 2020. Virtual screening: prediksi

potensi 8-shogaol terhadap c-Jun N-Terminal Kinase (JNK). Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika, 4 (1), 1-6

- [12] DRT Sari, JRK Cairns, A Safitri, F Fatchiyah. 2019. Virtual Prediction of the Delphinidin-3-O-glucoside and Peonidin-3-O-glucoside as Anti-inflammatory of TNF-α Signaling. Acta Informatica Medica, 27 (3), 152-157.
- [13] DRT Sari, A Safitri, JRK Cairns, F Fatchiyah. 2020. Anti-apoptotic activity of anthocyanins has potential to inhibit caspase-3 signaling. *Journal of Tropical Life Science*, 10 (1), 15-25.