Interaction of Enzyme-Substrate from Indigenous Cellulolytic Bacteria by Bioinformatics

Yanty Maryanty, Tri adryati, Suharjono, Sutiman Bambang Sumitro
Department of Industrial Chemical Industry, State Polytechnic of Malang, Jl. Soekarno Hatta No.9, Jatimulyo, Kec. Lowokwaru, Kota Malang, Jawa Timur 65141, Indonesia
Email: yanty.maryanty@polinema.ac.id

Abstract- This study aims to predict the degradation mechanism of cellulose substrate in silico from the results of isolation and test the potential of cellulolytic bacteria from rice fields in Greater Malang in producing cellulolytic enzymes, among others: β-1, 4 exoglunase, β-1, 4 endoglucase and β-glucosidase, identified isolates that have high potential in producing cellulolytic enzymes based on 16S rDNA ie isolates A, B, D and F. Only Isolate B was successfully predicted by its substrate enzyme interaction both homologically and based on the results of the isolation of the cellulite gene. Homology B isolate analysis results showed that the hydrogen bonds that occur in glutamic acid GLU257, tryptophan 207, serine 264 and glutamic acid 169, while hydrophobic interactions occur in tryptophan 207 bonds. While the results of the analysis based on isolation of cellulase encoding genes from isolate B were predicted in-silico showed the interaction between hydrogen bonds in tyrosin (TRY 299), glutamine (GLU201 and GLU 342.72), Asparagine (ASN 200) and glycine (GLY 384). This interaction is slightly different from the insilico results obtained homologically i.e there is no interaction with serine and glutamate acid, but with asparagine. Isolate B predicted that its homology in silico can degrade cellulose substrate with the binding affinity of -7.0 Kcal/mol while based on the results of isolation of cellulase encoding genes shows the degradation ability of cellulose substrate with the binding affinity of -6.5 Kcal/mol.

Keywords: Cellulase, in silico, hydrophobic interaction, the interaction of hydrogen bonds, binding affinity

1. Introduction
Cellulose biomass can be degraded by cellulase enzymes produced by cellulolytic bacteria. These enzymes have a variety of industrial applications and are currently referred to as the main group in the enzyme industry. Cellulase contributes eight percent of the needs of the enzyme industry worldwide (Alam et al., 2004; Meryandini et al., 2009; Sadhu & Maiti, 2013). In the past decade, cellulose-degrading bacteria have been isolated from various sources such as soil, weathered plants, hot springs, organic sources, ruminant manure, compost, and rice straw to obtain a more effective source of cellulase enzymes. Enzymes derived from cellulolytic bacteria are generally more resistant to alkali and thermophilic conditions with high cellulase activity (Immanuel et al., 2006; Doi, 2008; Irfan et al., 2012; Sadhu and Maiti, 2013; Feng et al., 2013; Abdel-Rahman, 2016). Enzymatic hydrolysis of cellulose depends on many factors including physical properties of the substrate (composition, crystallinity, and degree of polymerization), enzyme synergy (origin, composition), mass transfer (adsorption of the substrate, bulk and pore diffusion) and intrinsic kinetics. Previous research has succeeded in isolating some cellulolytic bacteria from agricultural...
land in the Greater Malang area. Interaction of substrate enzymes from cellulolytic bacteria can be predicted by bioinformatics.

Bioinformatics is a study of science that combines the disciplines of molecular biology, mathematics and information technology. Bioinformatics can be defined as the application or use of tools and computational methods (in silico methods) that aim to handle molecular biology data or information. (Thompson et al, 2000). Molecular docking is a computerized method used to describe interactions between a molecule as a ligand and a receptor or protein. Receptors or targets in the docking process can be obtained from experimental results or modelling homology results. Molecular docking as a methodology in structure-based virtual screening began in the early 1980s. In general, the purpose of docking studies is modelling accurate structures and predictions of appropriate activities (Thompson et al., 2000). In the docking process, the molecules involved are considered as atoms, surfaces or grids that represent the molecules. The docking process consists of several complex stages. This process begins with the application of a docking algorithm that positions the ligand on the active site with certain conformations and specific conformational search sequences, then the scoring function that completes the docking algorithm will evaluate the conformation by calculating based on physicochemical properties to obtain optimal molecular structure (Kitchen et al., 2004). Based on this process, a molecular docking program is a combination of algorithm docking and scoring functions (Thompson et al., 2000). Utilization and understanding of the interaction of substrate enzymes in bioinformatics in silico are carried out to help identify and measure the factors that influence the system for the hydrolysis process. This study aims to predict the mechanism of degradation of cellulolytic enzymes from isolates resulting from the isolation of Malang Raya rice fields in degrading cellulose substrate to cellulose substrate in silico.

2. Methodology

2.1. In Silico Analysis of Cellulase and Substrate Interactions
This study predicts the interaction of substrate enzymes from the isolation of cellulolytic bacteria from agricultural land in the Greater Malang area. The selected isolates were isolates A, B, D and F. These four bacteria were tested for cellulolytic activity and showed the highest activity compared to other isolates that had been successfully isolated. The DNA sequences of the four isolates from the sequenced cellulase gene sequences were validated and translated and then modelled using a homology modelling approach with the SWISS-MODEL webservice. The results of modelling simulated its interaction with the glucose substrate. 3D models of glucose substrate were obtained from the PubChem database. Structures in the SDF format are saved and used for further analysis. To find out the binding position of the enzyme and substrate so that the mechanism of action is known, a molecular docking study is performed. The interaction simulation between enzymes and substrate was studied using the Molecular Docking PyRx 0.8 program. The docking process uses autodock vina and is carried out specifically on the active site of the target protein.

Molecular docking can predict the binding affinity of a compound to a specific target protein (Trott & Olson, 2010). The higher the binding affinity of the substrate, the more accurate the interaction prediction will be. The molecular docking process is also based on control of the ligands that have previously interacted. Interactions between substrates and enzymes were analyzed with the Discovery Studio V.4.0 program. The purpose of this analysis is to understand the types of amino acids involved in interactions, so it can be concluded the active side and mechanism of action of the enzyme (Trott & Olson, 2010).

3. Results and Discussions

3.1. In silico analysis of cellulase and substrate interactions
The purification results of the selected cellulase encoding genes A, B, D, and F have been sequenced to ensure that the gene is well amplified. The results of this sequencing are useful for knowing the sites of cutting restriction enzymes contained in cellulase genes. These restriction enzyme sites are very useful for genetic engineering.
Figure 1: Electrophoresis results of cellulase-encoding genes selected isolates A, B, D, and F, along with Bacillus subtilis isolates; 1. Bacillus subtilis, 2 & 3 isolates B, 4 & 5 isolates D, 6 isolates D (with endoglucanase genes), 7 isolates A and 8 isolates F

The results of the sequencing were BLAST to analyze the level of similarity with cellulase-encoding nucleotides from other bacteria. The results of this study only succeeded in isolating endoglucanase-coding genes from isolate B. The results showed that the sequence of nucleotide sequences from isolate B had a high degree of similarity with Bacillus subtilis and Bacillus amyloliquefaciens bacteria by 99%. This shows that the amplified genes are cellulase genes and primers that are designed can be used to isolate and classify these cellulase genes. The sequence of cellulase genes can also be used to predict amino acid sequences and tertiary structure of cellulase isolate B. Prediction of the structure of tertiary cellulase isolate B is carried out by the comparative protein modelling method that is homology modeling. In this method, the structure of the target protein is determined based on other proteins namely the insolate protein that has been recognized by its tertiary structure. The protein template chosen is the one that has the greatest similarity to the target protein sequence. The template used is cellulase from Bacillus amyloliquefaciens (KF240848.1. Protein ID: AGW99978.1) because the sequence has the highest alignment score compared to cellulase for other bacteria in the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) The modeling process is carried out using the help of a program on the site http://swissmodel.expasy.org The protein model selected has the best QMEAN Z value or quality value. The modeling results are downloaded in the form of a file with PDB type so Structural analysis can be done Analytical structure of cellulase enzymes is carried out with the VMD program.
Determination of docking enzyme-ligation complex conformation by selecting one ligation conformation with the best or smallest value of Go binding affinity (kcal/mol). The tendency or strength of a bond or can be called the affinity of a ligand for a receptor or enzyme can be determined by looking at the value of binding affinity (kcal/mol) produced when the ligand enzyme complex is formed. The highest affinity of a ligand for enzymes results from the large intermolecular force between the ligand and the enzyme. The higher the affinity of a ligand for enzymes, the value of binding affinity (kcal/mol) is getting smaller, conversely, if the affinity is lower, the value of ΔG° binding affinity (kcal/mol) will be greater.

The results of interactions between the substrate cellulose and enzymes are predicted using molecular docking. The bond between the ligand with the receptor is said to be good if the ligand is chemically bound to amino acids in the binding region of the receptor, such as hydrogen bonds, van der Waals, hydrophobic region, and so on. The ligand conformation can affect the affinity of the ligand with the receptor. The analysis showed that the hydrogen bonds that occur in glutamate acid GLU257, tryptophan 207, serine 264 and glutamic acid 169, while hydrophobic interactions occur in tryptophan 207 bonds. The binding affinity of -7.0 Kcal/mol. A further process is to isolate the cellulase-encoding gene from isolate B and predict in silico the possible interactions of the substrate and enzymes produced by the selected isolate and shown in Figure 2.

These results indicate that there are interactions of hydrogen bonds in tyrosin (TRY 299), glutamine (GLU201 and GLU 342.72), Asparagine (ASN 200) and glycine (GLY 384). This interaction is slightly different from the in-silico results obtained homologically i.e there is no interaction with serine and glutamate acid, but with asparagine. The isolation of *Bacillus amyloliquefaciens* showed the degradation ability of cellulose substrate with a binding affinity of -6.5 Kcal/mol.
Figure 3: Interaction between the cellobiose substrate and the endoglucanase enzyme in silico

Compared with the results of research conducted by Manhar 2016 that docking CMC bonds with Bacillus amyloliquefaciens AMS1 cellulase (endoglucanase). Hydrophobic interactions that occur in Asn134, Lys33, His131, Gln297, Lys296 and Trp69 and the hydrogen concentration on Ala98, Asp 99, Thr97, Ala36, His65 with an affinity binding of - 7.97648 Kcal/mol. Studies show that the interaction of AMS1 with carbohydrate-binding surfaces is mainly by charged and polar residues such as aspartate, lysine, glutamine, threonine, histidine, asparagine and tryptophan. The only aromatic residue that interacts with CMC is Trp 69. Polar residues are not only responsible for changes in carbohydrate conformation, but can exclusively replace inter-chain hydrogen bonds which in turn damage the complex cellulose structure. Threonine and aspartate residues are capable of protonating histidine, which allows AMS1 to break down glycosyl hydrolase according to the donated proton histidine donor.

Figure 4: Interaction between cellulose substrate with selected isolate endoglucanase enzyme in silico
For isolate D which was predicted as *Acinetobacter baumannii* was also analyzed as isolate B. Prediction of tertiary cellulase D isolate structure was carried out by the comparative protein modeling method, namely homology modeling. The template used is cellulase from *Acinetobacter baumanni* (Cellulase Synthase (predicted) (JEXJ01000380.1, Protein ID: A0A009RP65 Whole-genome Seq).

![Cellulase Synthase](image)

**Figure 5:** Interaction between cellulose substrate and Cellulase Synthase (predicted) isolate of *Acinetobacter baumanni* homology in silico

The analysis shows that the hydrogen bonds that occur in asparagine ASN97, proline PRO190, Histidine HIS193, and Glycine GLY194, while hydrophobic interactions occur in Leucine LEU196. The binding affinity of -5.7 Kcal/mol. In isolate, D was unable to isolate the encoding gene after isolation was carried out using primary primers based on homology prediction in silico. This is possible because in the isolate D the gene reference used is Cellulase Synthase (predicted) so that when sequencing is not successful.

4. **CONCLUSION**

The isolates of B and D are predicted to have homologous degradation ability in cellulose substrate with the binding affinity of -7.0 Kcal/mol for isolate B, namely *Bacillus amyloliquefaciens* and isolate D, namely *Acinetobacter baumannii* of -5.7 Kcal/mol. The isolation of *Bacillus amyloliquefaciens* showed the degradation ability of cellulose substrate with a binding affinity of -6.5 Kcal/mol.

**ACKNOWLEDGMENT**

The research was funded by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia through the BPDN scholarship.
REFERENCES

[1] Abdel-Rahman, M.A. 2016. Establishment of efficient cellulolytic bacterial consortium potential for designed composting of rice straw. *International Journal of Advanced Research in Biological Sciences*. 3(4):211-228.

[2] Alam, M.Z., M.A. Manchur & M.N. Anwar. 2004. Isolation, purification, characterization of cellulolytic enzymes produced by the isolat *Streptomyces omiyaensis*. *Journal Bioscience*. 7: 1647-1653.

[3] Feng, H.W., Y.E. Zhi., W.W. Shi., L. Mao & P. Zhou. 2013. Isolation, identification and characterization of a straw degrading *Streptomyces griseorubens* JSD-1. *African Journal of Microbiology Research*. 7(22): 2730-2735.

[4] Immanuel, G., Dhanusha ., R. Premap & A. Palavesam. 2006. Effect of different growth parameters on endoglucanase activity by bacteria isolated from coir retting effluents of estuarine environment. *International journal environmental science technology*. 3:25-34.

[5] Irfan, M., S. Asma., S. Quratulain & N. Muhammad. 2012. Isolation and screening of cellulolytic bacteria from soil and optimization of cellulase production and activity. *Turkish Journal of Biochemistry*. 37(3):287-293.

[6] Kitchen, D.B., H. Decornez., J.R. Furr & J. Bajorath. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. *Drug Discovery*. 3: 935-949.

[7] Meryandini, A., W. Wahyu., M. Besty., C . Titi., Sunarti,. R. Nisa & S. Hasrul. 2009. Isolasi bakteri selulolitik dan karakterisasi enzimnya. *Makara Sains*.13(1):33-38.

[8] Sadhu, S. & M.T. Maiti. 2013. Cellulase production by bacteria; a review. *British Microbiology Research Journal*. 3(3):235-258.

[9] Thompson, J.D., F. Plewniak., J.C. Thierry & O. Poch. 2000. DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. *Nucleic Acids Research*. 28(15):2919-2926.

[10] Trott, O. & A.J. Olson. 2010. Auto Dock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *Journal of Computational Chemistry*. 31:455-461.