Journal of Genetic Engineering and Biotechnology

RESEARCH Open Access

In silico characterization of the GH5-cellulase family from uncultured microorganisms: physicochemical and structural studies

Rahmat Eko Sanjaya[1,2,3], Kartika Dwi Asni Putri[2], Anita Kurniati[1,2,4], Ali Rohman[2,5] and Ni Nyoman Tri Puspaningsih[2,5*]

Abstract

Background: Hydrolysis of cellulose-based biomass by cellulases produce fermented sugar for making biofuels, such as bioethanol. Cellulases hydrolyze the β-1,4-glycosidic linkage of cellulose and can be obtained from cultured and uncultured microorganisms. Uncultured microorganisms are a source for exploring novel cellulase genes through the metagenomic approach. Metagenomics concerns the extraction, cloning, and analysis of the entire genetic complement of a habitat without cultivating microbes. The glycoside hydrolase 5 family (GH5) is a cellulase family, as the largest group of glycoside hydrolases. Numerous variants of GH5-cellulase family have been identified through the metagenomic approach, including CelGH5 in this study. University-CoE-Research Center for Biomolecule Engineering, Universitas Airlangga successfully isolated CelGH5 from waste decomposition of oil palm empty fruit bunches (OPEFB) soil by metagenomics approach. The properties and structural characteristics of GH5-cellulases from uncultured microorganisms can be studied using computational tools and software.

Results: The GH5-cellulase family from uncultured microorganisms was characterized using standard computational-based tools. The amino acid sequences and 3D-protein structures were retrieved from the GenBank Database and Protein Data Bank. The physicochemical analysis revealed the sequence length was roughly 332–751 amino acids, with the molecular weight range around 37–83 kDa, dominantly negative charges with pI values below 7. Alanine was the most abundant amino acid making up the GH5-cellulase family and the percentage of hydrophobic amino acids was more than hydrophilic. Interestingly, ten endopeptidases with the highest average number of cleavage sites were found. Another uniqueness demonstrated that there was also a difference in stability between in silico and wet lab. The *II* values indicated CelGH5 and ACA61162.1 as unstable enzymes, while the wet lab showed they were stable at broad pH range. The program of SOPMA, PDBsum, ProSA, and SAVES provided the secondary and tertiary structure analysis. The predominant secondary structure was the random coil, and tertiary structure has fulfilled the structure quality of QMEAN4, ERRAT, Ramachandran plot, and Z score.

* Correspondence: ni-nyoman-t-p@fst.unair.ac.id
[2]University-CoE-Research Centre for Bio-Molecule Engineering, 2nd Floor ITD Building, Kampus C Universitas Airlangga, Mulyorejo, Surabaya, East Java 60115, Indonesia
[5]Department of Chemistry, Faculty of Science and Technology, Kampus C Universitas Airlangga, Mulyorejo, Surabaya, East Java 60115, Indonesia
Full list of author information is available at the end of the article