Conference Paper

Bioinformatics Analyses of Endo-1,4-beta-xylanase from *Bacillus* sp. SS3.4

Dikson and Reinhard Pinontoan

Biology Department, Universitas Pelita Harapan, Tangerang, Banten, Indonesia

Abstract

Various types of microorganisms particularly bacteria and fungi are capable of producing xylanolytic enzymes that degrade xylan into simple sugars to be used as their carbon sources. Xylanolytic enzymes, particularly endo-1,4-β-xylanase, has been extensively used in various industrial sectors, including biofuel, livestock, paper, and food. Previously, *Bacillus sp*. SS3.4 was observed to have xylanolytic activities on wheat bran, suggesting its capability of producing xylanolytic enzymes, endo-1,4-β-xylanase included. With great numbers of data available in nucleotide and protein sequence databases, bioinformatics analyses of protein structure for function prediction have become attractive to supplement wet laboratory experiments to reduce time and money consumption. Therefore, in this study, bioinformatics analyses were performed on *Bacillus sp.* endo-1,4-β-xylanase sequence in order to further evaluate the enzyme's function by exploring its protein structure in comparison to other known bacterial endo-1,4-β-xylanase. In this study, *Bacillus sp.* endo-1,4-β-xylanase sequence was identified from its whole genome sequence. The sequence was then analyzed by using BLAST to perform a homology search for identifying the protein, MUSCLE to perform multiple sequence alignment for comparing with other enzyme sequences, PHYLOGENY-FR to display its relatedness among *Bacillus*, *Aspergillus niger* and *Trichoderma reesei*, and SWISS-MODEL to generate its three-dimensional structure. Results from BLAST confirmed that the identified sequence was endo-1,4-β-xylanase, with greater relatedness to *Bacillus velezensis* and *Bacillus amyloliquefaciens* endo-1,4-β-xylanase. While results from MUSCLE and SWISS-MODEL suggested that the endo-1,4-β-xylanase belongs to the Glycosyl Hydrolase 11 (GH11) family bearing the distinct shape of a jelly roll with well-conserved binding and catalytic residues. These results strongly suggest that *Bacillus sp.* endo-1,4-β-xylanase is potentially capable of degrading xylan with highly similar mechanism as with other endo-1,4-β-xylanase.

Keywords: *Bacillus*, bioinformatics analysis, *endo-1,4-β-xylanase,* xylan

Corresponding Author:
Reinhard Pinontoan
reinhard.pinontoan@uph.edu

Received: 1 February 2020
Accepted: 8 February 2020
Published: 16 February 2020

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the IC-BIOLIS Conference Committee.

1. Introduction

Xylan is one of the main components of hemicellulose and one of the most abundant polysaccharides found in nature, only second to cellulose. Approximately, xylan makes up 20-35% of terrestrial biomass [1, 2]. Various types of microorganisms particularly

bacteria and fungi are capable of producing various kinds of xylanolytic enzymes to degrade xylan into simple sugars to be used as carbon sources for their growth and development [3]. Among those xylanolytic enzymes is endoxylanase, or better known as endo-1,4-β-xylanase, which belongs to a well-conserved enzyme family known as Glycosyl Hydrolase 11 (GH11) family with Enzyme Commission number 3.2.1.8 [4].

Endo-1,4-β-xylanase has been well-researched and applied in various industrial applications, either to remove xylan in products during manufacturing, or to produce xylose from xylan to be fermented into various useful products. Industrial endo-1,4-β-xylanase has been produced from various sources, primarily fungal and bacterial sources, with varying properties to ideally suit the industry's needs. Bacterial endo-1,4-β-xylanase has a broader effective range of optimal pH than that of fungal endo-1,4-β-xylanase [5]. Few examples of commercial endo-1,4-β-xylanase being used for industrial applications include Novozym's Panzea produced from *Bacillus licheniformis* and used for improving bread dough quality, Alltech's Allzym PT produced from *Aspergillus niger* used for improving animal feed quality, and Iogen's Xylanase GS35 produced from *Trichoderma reesei* used for improving animal feed and paper pulp [6].

In the preliminary results, *Bacillus sp*. SS3.4 was observed to have xylanolytic activities on wheat bran, which served as an affordable substrate for qualitatively screening microorganisms with potential xylanolytic activities. The results suggested the bacterium's capability of producing xylanolytic enzymes to degrade said substrate, including *Bacillus sp.* SS3.4 endo-1,4-β-xylanase, referred to as SS3.4-Xyl. However, isolation, quantitative characterization, and potential engineering of endo-1,4-β-xylanase by conventional means in wet laboratory may consume a large amount of time and money, particularly due to the high cost of high-purity xylan used as substrate in such experiments [7].

With modern advancements in DNA sequencing technologies, a great number of data have become widely available in nucleotide and protein sequence databases. Thus, bioinformatics approaches in analyzing protein structures for predicting its function have been viewed as a time- and cost-effective method to supplement wet laboratory experiments [8]. Therefore, in this study, bioinformatics analyses were performed on SS3.4-Xyl sequence in order to further evaluate the enzyme's function by exploring its protein sequence, tertiary structure, as well as its active site in comparison to other known bacterial endo-1,4-β-xylanase.

2. Method

2.1. Data and Programs

The utilized data during this research was a protein sequence of SS3.4-Xyl in FASTA format taken from annotations of *Bacillus* sp. SS3.4 whole genome sequence from our prior, unpublished research. The utilized programs during this research are publicly available bioinformatics services such as BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) for identification and/or inquiry for protein sequences similar to the uploaded protein sequence [9], MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/) provided on European Bioinformatics Institute website [10] to align and compare protein sequences [11], PHYLOGENY-FR (http://www.phylogeny.fr/simple_phylogeny.cgi) to build a phylogenetic tree based on uploaded protein sequences [12], SWISS-MODEL (https://swissmodel.expasy.org/interactive), a protein structure homology-modelling service, to predict tertiary structure of uploaded protein sequences based on existing protein models in the database [13--17], and VMD (http://www.ks.uiuc.edu/Research/vmd/), a molecular visualization software to display the tertiary structure of proteins in forms of illustrations or animations [18].

2.2. Sequence alignment and phylogenetic tree of SS3.4-Xyl sequence

BLAST search results comprising of *Bacillus*, *Trichoderma reesei*, and *Aspergillus niger* endo-1,4-β-xylanase sequences were downloaded and compared against SS3.4-Xyl sequence through multiple sequence alignment using MUSCLE software. The sequences were then uploaded to PHYLOGENY-FR website set in One-Click mode to build its phylogenetic tree.

2.3. Tertiary structure and active site prediction of SS3.4-Xyl

SS3.4-Xyl enzyme was uploaded to SWISS-MODEL website to predict its tertiary structure based on *Bacillus* endo-1,4-β-xylanase model with PDB accession code 1AXK. The predicted structure was then downloaded in PDB file format to be visualized using VMD software in forms of illustrations and turntable animations of SS3.4-Xyl tertiary structure and its active site.

3. Results and Discussions

3.1. Sequence alignment and phylogenetic tree of SS3.4-Xyl sequence

Whole genome sequencing of *Bacillus* sp. SS3.4 indicating an existence of several xylanolytic enzymes, one of which was identified as endo-1,4-β-xylanase belonging to GH11 enzyme family bearing the distinct shape of a jelly roll with well-conserved binding and catalytic residues. *SS3.4-Xyl* had 213 amino acid residues. Various amino acid sequences from *Bacillus* sp were used to compare with SS3.4-Xyl in this phylogenetic tree analyses while *Trichoderma reesei*, and *Aspergillus niger* endo-1,4-β-xylanase sequences were used for comparison since these fungi were used for producing fungal endo-1,4-β-xylanase for industrial use [6]. Figure 1 shows the phylogenetic tree analyses of SS3.4-Xyl. When it was compared with *B. velezensis* FZB42 and *Bacillus amyloliquefaciens*, SS3.4-Xyl was highly similar. Results from MUSCLE multiple sequence alignment showed that their identity percentage values were 99.53% and 99.06%, respectively.

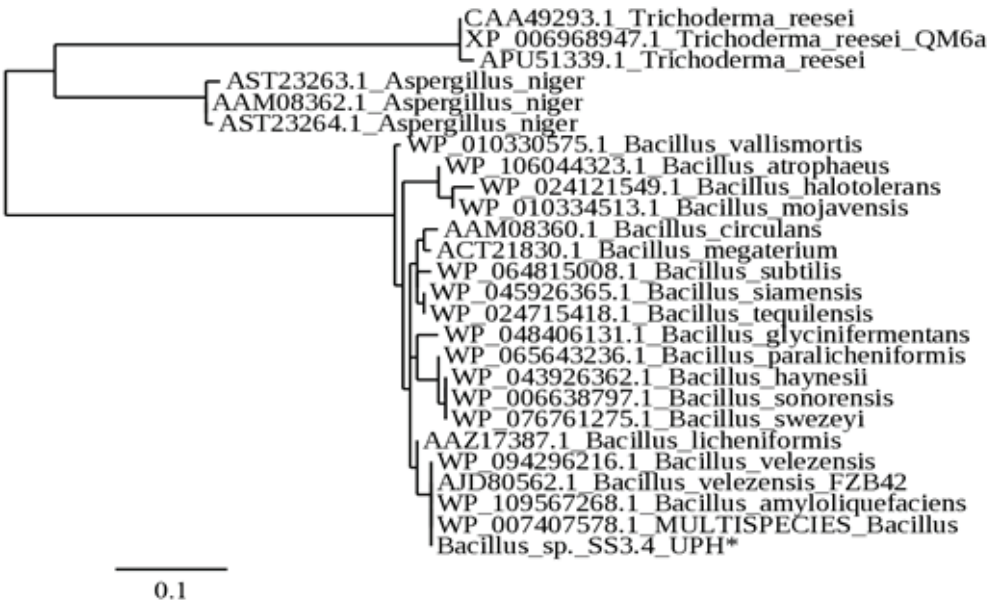

Figure 1: Phylogenetic tree of endo-1,4-β-xylanase enzymes built using PHYLOGENY-FR in One-Click mode. SS3.4-Xyl was denoted with an asterisk (*).

3.2. Prediction of SS3.4-Xyl tertiary structure and active site

SS3.4-Xyl tertiary structure was predicted using SWISS-MODEL program, based on *Bacillus* endo-1,4-β-xylanase model with PDB accession code 1AXK and visualized using VMD software as seen in Figure 2:

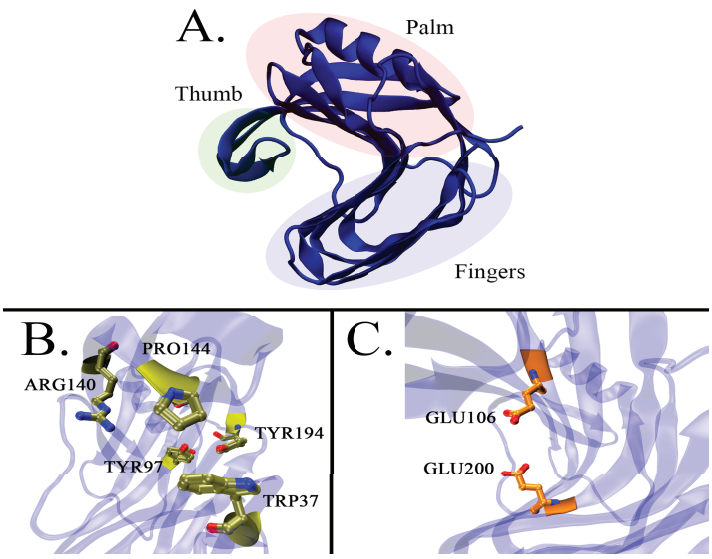

Figure 2: Predicted model of SS3.4-Xyl tertiary structure and active site. Information: A: Predicted tertiary structure; B: Substrate binding residues; C: Catalytic residues.

GH11 family endo-1,4-β-xylanase enzymes possess a distinct shape of a jelly-roll with three known domains as Thumb, Palm, and Fingers domains. A gap formed in between Palm and Fingers domains is an active site, bearing a few well-conserved residues that function as xylan binding and catalysis as suggested in references [4, 19--22] listed in Table 1 as follows:

TABLE 1: Binding and catalytic residues of endo-1,4-β-xylanase.

SS3.4-Xyl	BCX	Tx-xyl*	Function	Reference source
Trp37	Trp9	Trp7	Substrate Binding (-2)	[21]
Tyr97	Tyr69	Tyr67	Substrate Binding (-2)	[21]
Glu106	Glu78	Glu76	Nucleophile	[21]
Arg140	Arg112	Arg110	Substrate Binding (-1)	[17]
Pro144	Pro116	Pro114	Substrate Binding (-1)	[17]
Tyr194	Tyr166	Tyr163	Substrate Binding (-2)	[17]
Glu200	Glu172	Glu169	Acid/Base	[21]

Information: SS3.4-Xyl: *Bacillus* sp. SS3.4; BCX: *Bacillus circulans*; Tx-xyl:*Thermobacillus xylanilyticus*.

*Uses Tx-xyl numbering as of reference [4].

SS3.4-Xyl amino acid sequence was found to contain binding and catalytic residues as commonly found in GH11 family such as of *Bacillus circulans* as suggested in references [20, 21, 23] and of *Thermobacillus xylanilyticus* as suggested in reference [4]. These residues included Trp37, Tyr97, and Tyr194 that might function to bind xylan at subsite (-2), Arg140 and Pro144 that might function to bind xylan at subsite (-1), Glu106

functions as a nucleophile, and Glu200 functions as a general acid/base. Therefore, the existence of these residues as well as high similarities with other GH11 endo-1,4-β-xylanase sequences and tertiary structure strongly suggested that SS3.4-Xyl belongs to the GH11 enzyme family capable of degrading xylan highly similarly as with other endo-1,4-β-xylanase enzymes do.

4. Conclusion

Whole genome sequencing of *Bacillus* sp. SS3.4 indicates the presence of several xylanolytic enzymes, one of which is identified as SS3.4-Xyl belonging to GH11 enzyme family based on sequence annotation, sequence alignment, tertiary structure prediction and conserved binding and catalytic residue identification. This strongly suggests that SS3.4-Xyl is capable of degrading xylan with highly similar mechanism as with other endo-1,4-β-xylanase enzymes of the same family. In future research, however, it is advised to perform X-ray crystallography in order to determine its tertiary structure and molecular dynamics studies in order to determine its catalytic mechanism on xylan.

Acknowledgements

We would like to express our acknowledgement and gratitude for Forrest Faulk and Brendan Burns for the technical advice offered and given during the course of this research, Sebastien Santini - CNRS/AMU IGS UMR7256, for their management of the website PHYLOGENY-FR that has been utilized in this research, and Emily A. Willoughby for her insights and inspiration for the writer to pursue molecular biology and bioinformatics studies.

References

[1] T. Collins, C. Gerday, and G. Feller, "Xylanases, xylanase families and extremophilic xylanases," FEMS Microbiol. Rev., vol. 29, no. 1, pp. 3-23, 2005.

[2] P.J. Smith, H.T Wang, W.S. York, M.J. Peña, and B.R. Urbanowicz, ``Designer biomass for next-generation biorefineries: Leveraging recent insights into xylan structure and biosynthesis,'' Biotechnol. Biofuels, vol. 10, pp. 286, 2017.

[3] A.C.E.P. Benedetti, E.D. da Costa, C.C. Aragon, A.F. dos Santos, A.J. Goulart, D. Attili-Angelis, and R. Monti, "Low-cost carbon sources for the production of a thermostable

xylanase by *Aspergillus niger*," Revista de Ciências Farmacêuticas Básica e Aplicada, vol. 34, no. 1, pp. 25.31, 2013.

[4] G. Paës, J.G. Berrin, and J. Beaugrand, "GH11 xylanases: Structure/function/properties relationships and applications,". Biotechnol. Adv., vol. 30, no. 3, pp. 564-592, 2012.

[5] A. Burlacu, C.P. Cornea, and F. Israel-Roming, "Microbial xylanase: a review," Scientific Bulletin Series F. Biotechnologies, vol. 20, pp. 335-342, 2016.

[6] A.D. Harris, and C. Ramalingam, "Xylanases and its application in food industry: A review," J. Exp. Sci., vol. 1, no. 7, pp. 1-11, 2010.

[7] L.M. Varghese, S. Agrawal, D. Sharma, R.P. Mandhan, and R. Mahajan, "Cost-effective screening and isolation of xylano-cellulolytic positive microbes from termite gut and termitarium," 3 Biotech, vol. 7, no. 2, pp. 108, 2017.

[8] Y.D. Yang, J.Z. Gao, J.H. Wang, R. Hefferman, J. Hanson, K. Paliwal, and Y.Q. Zhou, "Sixty-five years of the long march in protein secondary structure prediction: the final stretch?," Briefings in Bioinformatics, vol. 19, no. 3, pp. 482-494, 2018.

[9] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, "Basic local alignment search tool," J. Mol. Biol., vol. 215, no. 3, pp. 403-410, 1990.

[10] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, and R. Lopez, "The EMBL-EBI search and sequence analysis tools APIs in 2019,". Nucleic Acids Res., vol. 47, no. W1, pp. W636-W641, 2019.

[11] R.C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," Nucleic Acids Res., vol. 32, no. 5, pp. 1792-1797, 2004.

[12] A. Dereeper, V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J.F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J.M. Claverie, and O. Gascuel, "Phylogeny.fr: robust phylogenetic analysis for the non-specialist," Nucleic Acids Res., vol. 36, pp. W465-W469, 2008.

[13] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede, "SWISS-MODEL: homology modelling of protein structures and complexes,". Nucleic Acids Res., vol. 46, no. W1, pp. W296-W303, 2018.

[14] S. Bienert, A. Waterhouse, T.A.P. de Beer, G. Tauriello, G. Studer, L. Bordoli, and T. Schwede, "The SWISS-MODEL Repository - new features and functionality," Nucleic Acids Res., vol. 45, no. D1, pp. D313--D319, 2017.

[15] N. Guex, M.C. Peitsch, and T. Schwede, "Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective," Electrophoresis, vol. 30, no. 1, pp. S162-S173, 2009.

[16] P. Benkert, M. Biasini, and T. Schwede, "Toward the estimation of the absolute quality of individual protein structure models," Bioinformatics, vol. 27, no. 3, pp. 343-350, 2011.

[17] M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, and T. Schwede, "Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology," Sci Rep., vol. 7, pp. 10480, 2017.

[18] W. Humphrey, A. Dalke, and K. Schulten, "VMD - Visual molecular dynamics," J. Mol. Graph., vol. 14, no. 1, pp. 33-38, 1996.

[19] L.A. Plesniak, W.W. Wakarchuk, and L.P. McIntosh, "Secondary structure and NMR assignments of *Bacillus circulans* xylanase," Protein Sci., vol. 5, no. 6, pp. 1118-1135, 1996.

[20] G. Sidhu, S.G. Withers, N.T. Nguyen, L.P. McIntosh, L. Ziser, and G.D. Brayer, "Sugar ring distortion in the glycosyl-enzyme intermediate of a family G/11 xylanase," Biochemistry, vol. 38, no. 17, pp. 5346-5354, 1999.

[21] W.W. Wakarchuk, R.L. Campbell, W.L. Sung, J. Davoodi, and M. Yaguchi, "Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase," Protein Sci., vol. 3, no. 3, pp. 467-475, 1994.

[22] L. Song, B. Siguier, C. Dumon, S. Bozonnet, and M.J. O'Donohue, "Engineering better biomass-degrading ability into a GH11 xylanase using a directed evolution strategy," Biotech. Biofuels, vol. 5, no. 1, pp. 3, 2012.

[23] M.E.S. Soliman, Thesis of: Computational modelling of glycosidase mechanisms: structural and mechanistic aspects. United Kingdom: University of Bath, 2009.