IOP Conf. Series: Materials Science and Engineering **335** (2018) 012026 doi:10.1088/1757-899X/335/1/012026

Partial Gene Cloning and Enzyme Structure Modeling of Exolevanase Fragment from *Bacillus subtilis*

M Azhar[1*], D Natalia[2], S Syukur[3], N Andriani[1], J Jamsari[4]

[1]Biochemistry Laboratory, Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Jl. Prof. Hamka, Air Tawar, Padang 25131, Indonesia.
[2]Biochemistry Research Division, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Jl. Ganesha10, Bandung 40132, Indonesia
[3]Biochemistry Laboratory, Faculty of Mathematics and Natural Sciences, Universitas Andalas, Padang 25163, Indonesia
[4]Biotechnology Laboratory, Faculty of Agriculture, Universitas Andalas, Padang 25163, Indonesia

*minda@fmipa.unp.ac.id

Abstract. Inulin hydrolysis thermophilic and thermotolerant bacteria are potential sources of inulin hydrolysis enzymes. Partial gene that encodes inulin hydrolysis enzymes had been isolated from *Bacillus subtilis* using polymerase chain reaction (PCR) method with the DPE.slFandDPE.eR degenerative primers. The partial gene was cloned into pGEM-T Easy vector with *E. coli* as host cells and analyzed using BLASTx, CrustalW2, and Phyre2 programs. Size of thepartial gene had been found539 bp that encoded 179aminoacid residues of protein fragment. The sequences of protein fragment was more similar to exolevanase than exoinulinase. The protein fragment had conserved motif FSGS, and specific hits GH32 β-fructosidase. It had three residues of active site and five residues of substrate binding. The active site on the protein fragment were D (1-WLNDP-5), D (125-FRDPK-129) and E (177-WEC-179). Substrate binding on the protein fragment were ND (1-*WLNDP*-5), Q (18- FYQY-21), FS (60-FSGS-63) RD (125-FRDPK-129) and E (177-WEC-179).

1. Introduction

Inulin hydrolysis thermophilic bacteria are potential sources of inulin hydrolysis enzymes. Inulin hydrolysis thermophilic bacteria has been screened and characterized as *Bacillus*genus[1]. Most of the thermophilic bacteria belong to the *Bacillus*genus [2].Nucleotide sequences of 16S rRNA partial gene of inulin hydrolysis thermophilic bacteria that expressed thermostable endoinulinase had been reported in 2008 by Gao [3]. The bacteria was identified as *Bacillus smithii* T7.

Thermophilic and thermotolerant bacteria can produce inulinase or levanase thermostable enzymes. Inulinase and levanase are belong to glycoside hydrolase family 32. The enzymes are very important in industrial aplications. The enzymes aremost appropriate choice as catalyst of inulin hydrolysis reaction toproduce fructose and Fructo-Oligosaccharides (FOS). The enzyme can also be used for production of various metabolites such as citric acid, lactic acid, ethanol, biofuels, butanediol [4].Another advantage, the use of the bacteria is the enzyme that expressed more easily purified due to heat treament. In addition, the enzymatic reaction at high temperature enhances inulin solubility, lowering the viscosity, reduces the risk of contaminants and accelerates reaction [5].Furthermore, inulin solubility was greater in hot water [6].

Inulinase and levanase gene has been isolated from bacteria, fungi and yeast. Inulinase gene that had Open Reading Frame (ORF) 1542 bphas been isolated from genomic DNA *Pichia guilliermondii* yeastusing PCR technique [7].The gene encodes 514 amino acid residues of the inulinase. Levanase gene has been isolated from genomic library *Gluconacetobacter diazotrophicus* SRT4[8].Tertiary structure of exoinulinase and endoinulinase has been found in PDB (Protein Data Bank, http://www.rcsb.org/pdb/ home/home.do), but tertiary structure of exolevanase was not available.

Inulin hydrolysis partial gene has been isolated from genomic DNA *Bacillus licheniformis* by using PCR technique with DPE.slF and DPE.eR degenerative primers[9]. This paper described isolation of inulin hydrolysis partial gene from genomic DNA *Bacillus subtilis* UBCT-030 by using the primer. The partial gene was compared with protein sequences in GenBank database and determined structure modeling of enzyme fragment.

2. Materials and Methods

2.1. Bacteria

Bacillus subtilis UBCT-030 that has been identified in phenotype and genotype was used as source of genomic DNA [10].The bacteria was cultured in the medium that contained inulin as the sole carbon source[11].

2.2. Isolation of Genomic DNA Bacteria

Isolation and purification of genomic DNA bacteria was carried out according to the procedure in the Wizard Genomic DNA Purification Kit.

2.3. Isolation of partial gene

The partial gene that encodes inulin hydrolysis enzyme was isolated from the genomic DNA *Bacillus subtilis* UBCT-030 by PCR method using DPE.slFandDPE.eR primers. The composition of the master mix as follows: 2.5 µL10xDreamTaqbuffer, 0.5 µL10 mMdNTPmix, $MgCl_2$ 3mM, 0.5 µL forward primer DPE.slF(TGG MTR AAY GAY CCM AAC GGA C)20 µM, 0.5 µL reverse primer DPE.eR (GGR CAT TCC CAY WCV CCR TC) 20 µM, 0.5 µL samples(75ng/µL), 0.125 µL DreamTaqDNA polymerase. ddH2O was added to volume of 25 µL. The PCR process was done at initial denaturation 94°C for 2 minutes, denaturation 94°C for 45 seconds, annealing 56°C for 30 minutes, elongation 72°C for 45 seconds, final elongation 72°C for 5minutes. PCR cycles were done29times. Amplicons were electrophoresed using 0.8% agarosegel in TAE buffer and DNA band were visualized by UV transillumination. DNA marker was used 1kb DNA ladder (Fermentas).

2.4. Partial gene cloning and sequencing

Partial gene that encodes inulin hydrolysis enzyme was ligatedtopGEM-T Easyvector[12]. Mole ratio ofpartial geneto thevectorwas 3:1. Mixture of ligationreaction were 5 µL of ligationbuffer 2 xrapidT4DNA ligase, 1 µL of pGEM-T Easy50ng, 1 µL of T4DNA ligase (3 Weissunits/µL), 1 µL of insert 20ng/µL, andddH_2O was added to final volume of 10 µL. There action mixture was incubated at room temperature at4°C for overnight. DNA recombinants (5 µL) were transformed in *E. coli*TOP10F'[13]. Recombinant DNA that carried the insert DNA was isolated from white colony transformant according to the procedure in the High-Speed PlasmidMiniKit(Geneaid). Recombinant DNA (1000ng)were sequenced by using dideoxy-Sanger method with T7 universal primer in Macrogen Korea.

2.5. Bioinformatics

Similarity with another protein in the GenBank database was determined by using the BLASTx program (http://www.ncbi.nlm.nih.gov). Determination of enzyme structure modeling was used Phyre2 program at http://www.sbg.bio.an.uk/phyre2 [14].

2.6. Accession number of nucleotide sequence

Nucleotide sequence of 16S rRNA gene from *Bacillus subtilis* UBCT-030 has assesion number KU358720. The sequences of partial genes from *Bacillus subtilis* UBCT-030 has been deposited in GenBank with accession number KU377724.

3. Results and Discussion

3.1.Amplification of the partial gene

Amplification of the partial gene encoding inulin hydrolysis enzyme from *Bacillus subtilis* UBCT-030 was performed at the optimum condition of PCR using DPE.slF and DPE.eR primers[9]. The PCR process was carried out at annealing temperature 56ºC, 3 mM $MgCl_2$. The each cycle of PCR was set up denaturation94°C for 45seconds, annealing 56°C for 30 minutes, elongation 72°C for 45seconds. Finalelongation of PCR was 72°C for 5 minutes. Amplicon was purified and cloned in pGEM-T Easy vector. In the condition was produced amplicons about 550 bp (Figure 1).

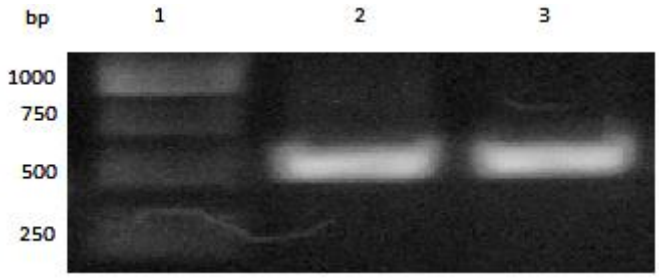


Figure 1. Electrophoresis of amplicon. Lane 1 was 1 kb DNA ladder.Lane 2 and 3 were amplicons of partial gene

3.2.Cloning of partial gene in pGEM-T Easy

The partial gene (about 550 bp) was ligated to the pGEM-T Easy vector to produce recombinant DNA. Recombinant DNA was transformed in *E. coli* TOP10F' cells by using heat shock method. Transforman cells were grown in LB medium containing ampicillin, IPTG and X-gal. There were white and blue transformant colonies on the surface of the media after 16 hours. White transformant colonies contained pGEM-T Easy with insert DNA, while blue colonies contained pGEM-T Easy without insert DNA (Figure2).

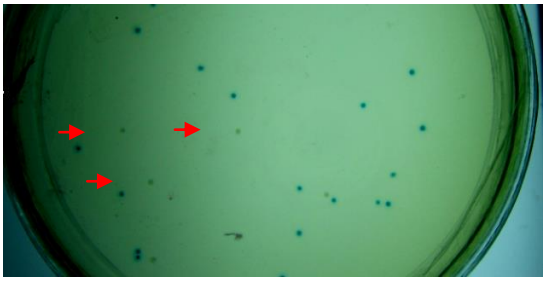


Figure 2. *E.coli* TOP10F' transformant colonies.

White colonies were transformants containingpGEM-T Easy with insert DNA (red arrows). Blue colonies were transformants containingpGEM-T Easywithout insert DNA.

The difference of transformants color were caused composition of bacteria media. Bacteria media was LB medium that contained IPTG (isopropyl-thiogalactoside), X-gal (5-bromo-4-chloro-3-indolyl-β-D galactopyranoside), and ampicillin. X-gal is β-galactosidase substrate, while IPTG is an inducer of lacZ gene. LacZ gene express β-galactosidase that catalyze breaking of bond on the X-gal produces

galactose and 5-bromo-4-chloro-3-hydroxyindole which further oxidized to 5,5'-dibromo-4.4'-dichloro-indigo. The compound gives blue color in the transformant cell.

Expression of β-galactosidase will be disrupted if the pGEM-T Easy carry insert DNA. The insert DNA was ligated to the lacZ gene (on the cutting side of the EcoRV restriction enzyme) in the pGEM-T vector. Thymidine were added at the both ends of the vector [12]. In the pGEM-T Easy vector also contain gene that expressed ampicillin degrading enzymes. Therefore, the transformants that contained pGEM-T Easy vector can live in LB medium that contained ampicillin. Thus, white transformant colonies contained insert DNA.

White transformant colonies that contained desired insert DNA were screened by using colonyPCR method. The method is very fast and easy to screen transformant colonies that contained specific DNA sequences using PCR without DNA purification [15]. Products of colonyPCR that aligned with the insert DNA fragment was transformant colonies that contained desired insert DNA (Figure 3). Recombinant DNA from the colonies were isolated and purified. The DNA was used as a template for sequencing process of the insert DNA.

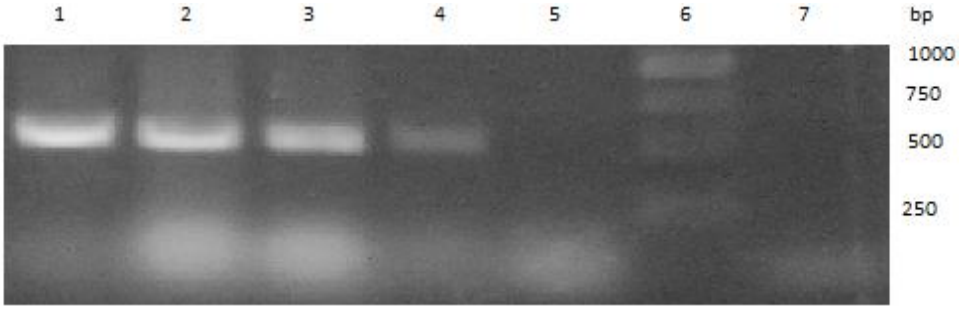


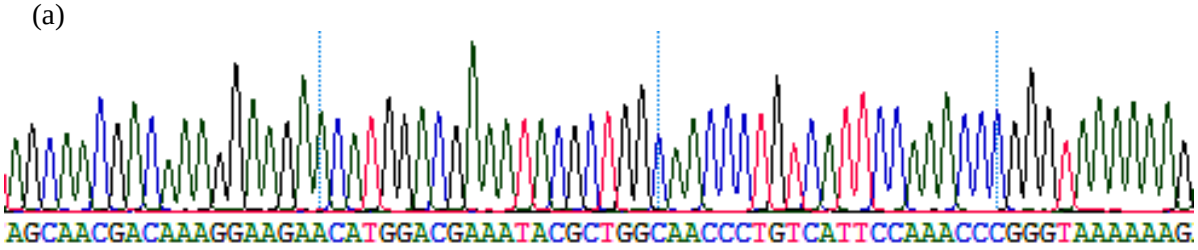
Figure 3. Electrophoresis of amplicon from white transformant colonies using colonyPCR**.** Lane 1 to 5 were amplicon colonyPCR using DPE.slFand DPE.eR primers. Lane 6 was the 1 kb DNA ladder marker. Lane 7 was primers (negative control).

3.3.The nucleotide sequences of partial gene and deduced amino acid sequences

The partial gene of *Bacillus subtilis* UBCT-030 was found 539 bp using the DPE.slF and DPE.eR primers (Figure 4). Deduction 539 bp of the partial gene produce protein fragment along the 179 amino acid residues (Figure 4b). In the protein fragment was found conserved domain A, B, B1, C, D and E. The domains were found in the group of Glycoside Hydrolase family (GH)32 enzyme [16]. Thus, partial gene is belong to GH32 enzymes. Domain A and E were position of DPE.slF, and DPE.eR primers [9]. The nucleotide sequences of the 539 bp was compared to the protein sequence in the GenBank database using BLASTx program. The protein fragment had specific hits GH32 β-fructosidase with three residues of the active and five residues of substrate binding (Figure 5).

Protein fragment along the 179 amino acid residues had high similarity with levanase from *Bacillus subtilis* (GenBank accession number WP041053668.1), and *Bacillus licheniformis* (GenBank accession number AGU89675.1) 98% and 97% respectively (Table 1). The protein fragment also had some similarities with the other GH32 enzymes, namely glycoside hydrolase from *Bacillus* sp. UNC41MFS5 (GenBank accession number WP026568421.1), SacC2 *Bacillusbataviensis* (GenBank accession number WP007084638.1), exoinulinase from *Aspergillus awamori* (PDB accession number 1Y9M|A), β-fructosidase from *Thermotoga maritima* (PDB accession number 1W2T|A), endoinulinase from *Aspergillus ficuum* (PDB accession number 3SC7|X), and levan fructotransferase from *Arthrobacter ureafaciens* (PDB accession number 4FFH|A) 65%, 64%, 49%, 41%, 39% and 34% respectively. Thus, the partial gene was isolated from genomic DNA *Bacillus subtilis* UBCT-030 was group of family GH32 enzymes. The protein fragment had closest resemblance to levanase than exoinulinase.

(a)



b)

```
TGGCTGAATG ATCCAAACGG ACTGGTATAT TATGCCGGGG AATATCACTT GTTCTATCAA   60
TACCATCCAT ACGGGCTTCA GTGGGGGCCC ATGCATTGGG GGCATGCCGT CAGCAAAGAT  120
TTGGTTACAT GGGAACACCT TCCTGTTGCG CTTTATCCGG ATGAAAAAGG CACGATCTTT  180
TCTGGAAGCG CAGTTGTAGA TAAAAATAAC ACAAGCGGTT TTCAAACAGG CAAAGAGAAG  240
CCGCTTGTGG CCATTTATAC ACAGGATCGG GAAGGCCATC AAGTGCAAAG TATTGCCTAT  300
AGCAACGACA AAGGAAGAAC ATGGACGAAA TACGCTGGCA ACCCTGTCAT TCCAAACCCG  360
GGTAAAAAAG ATTTTCGCGA TCCAAAAGTC TTTTGGTATG AGAAAGAAAA AAAGTGGGTG  420
ATGGTGCTTG CGGCCGGTGA CCGAATCCTC ATTTATACAT CAAAAAATCT GAAGCAGTGG  480
ACGTATGCAA GTGAATTTGG ACAGGATCAA GGAAGCCACG ACGGAGAATG GGAATGCCC    539
```

c)

```
WLNDPNGLVY YAGEYHLFYQ YHPYGLQWGP MHWGHAVSKD LVTWEHLPVA LYPDEKGTIF   60
SGSAVVDKNN TSGFQTGKEK PLVAIYTQDR EGHQVQSIAY SNDKGRTWTK YAGNPVIPNP  120
GKKDFRDPKV FWYEKEKKWV MVLAAGDRIL IYTSKNLKQW TYASEFGQDQ GSHDGEWEC   179
```

Figure 4.The nucleotide sequences of partial gene (a, b) and the amino acid residue of protein fragment (c). Electropherogram (a) is blue nucleotide sequences in (b). Red nucleotide sequences in (b) were DPE.slF and DPE.eR primers. Blue amino acid residues were conserved domain A, D and E respectively (c).

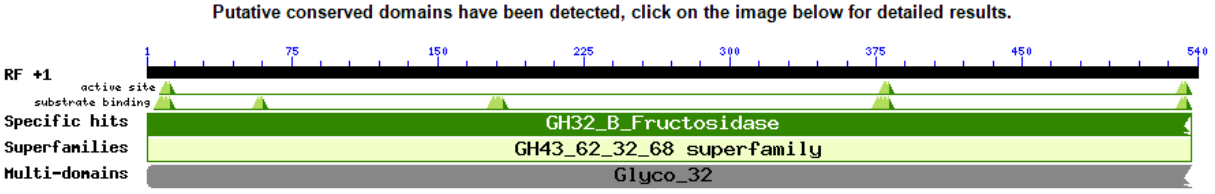


Figure 5. Group of the partial gene of *Bacillus subtilis* UBCT-030

Type action exo- of inulinase on conserved domain C in fungi have conserved motif Phe-Ser-Gly-Ser (FSGS), while the type action endo- had conserved motif FTGT[17,18]. In *Aspergillus niger* CBS13.88 was found exoinulinase that have conserved motif FSGS, while endoinulinase FTGT [16]. Conserved motif FSGS also was found in the protein fragment of *Bacillus subtilis* UBCT-030. Thus, the protein fragment of *Bacillus subtilis* UBCT-030 was a levanase that had type action exo-.

Table 1. Similarity of partial gene 539 bp with inulin hydrolysis enzymes from bacteria and yeast

GenBank Accession number	Description	Similarity (%)
WP041053668.1	levanase *Bacillus subtilis*	98
AGU89675.1	Lev1*Bacillus licheniformis*	97
WP026568421.1	glycoside hydrolase *Bacillus sp6*. UNC41MFS5	65
WP007084638.1	SacC2 *Bacillus bataviensis*	64
1Y9M|A	Exo-Inulinase from *Aspergillus awamori*	49
1W2T|A	Beta-fructosidase from *Thermotoga maritima*	41
3SC7|X	Endo-inulinase from *Aspergillus Ficuum*	39
4FFH|A	Levan fructotransferase from *Arthrobacter ureafaciens*	34

3.4. Structure modeling of exolevanase fragment

Tertiary structure of exolevanase was not available in the Protein Data Bank (PDB) database. Therefore, template for structure modeling of protein fragment from *Bacillus subtilis* UBCT-030 was used tertiary structure of exoinulinase from *Aspergillus awamori* in the PDB database at http://www.rcsb.org/pdb/home/ home.do. Identical the protein fragment with exoinulinase from *Aspergillus awamori* were 52%. Structure modeling of the protein fragment was built from 179 amino acid residues. Alignment of protein fragment from *Bacillus subtilis* UBCT-030 with exoinulinase from *Aspergillus awamori* was showed in Figure 6

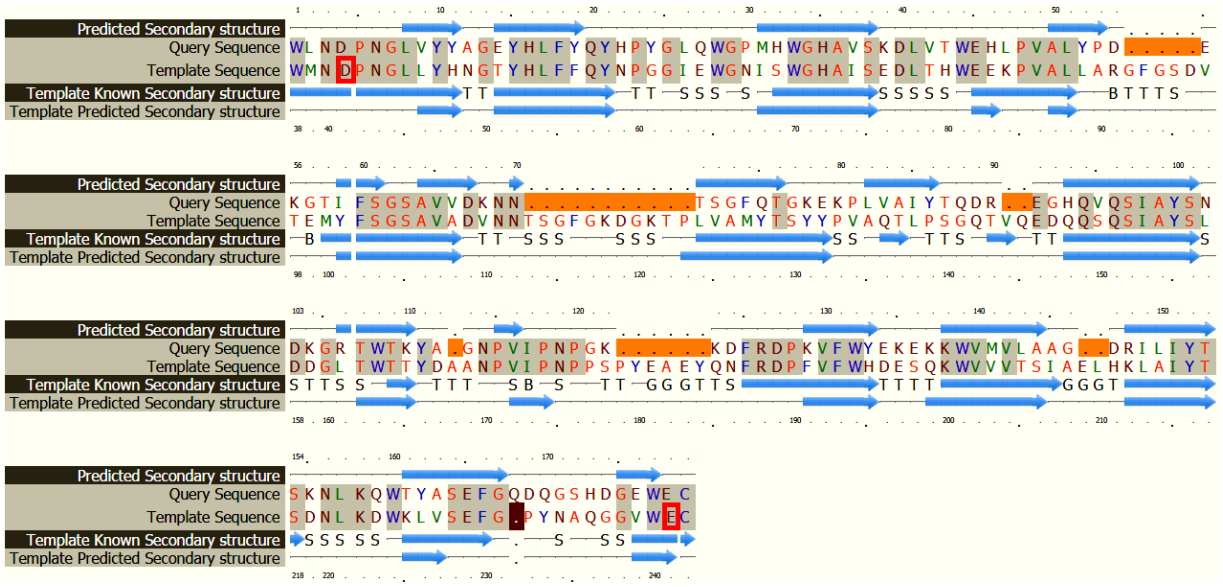
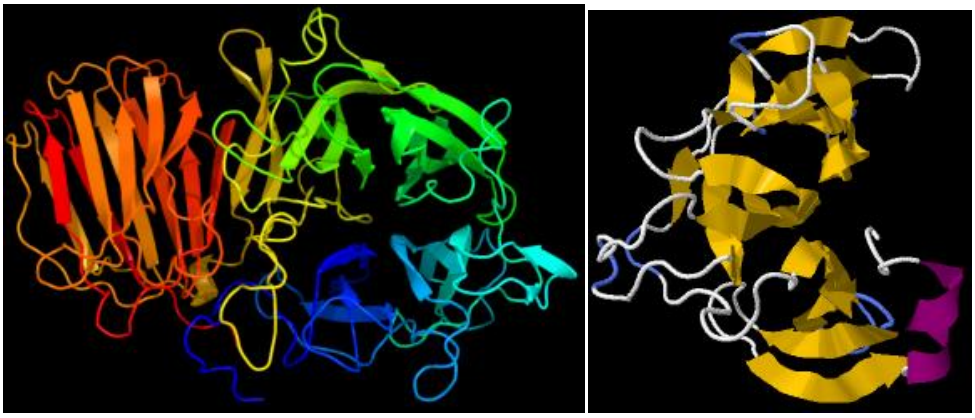


Figure 6. Alignment of exoinulinase with protein fragment of *Bacillus subtilis* UBCT-030 Catalytic sites were red box

Overall folding of the tertiary structure of the protein fragment was as folding of family GH32 enzyme. The family GH32 enzyme is composed two domains, namely β-propeller domain at the N terminal domain (Domain 1) and β-sandwich domain at the C terminal domain (Domain 2). Domain β-propeller consists of five blades. They are blade 1 (blue), blade 2 (light blue), blade 3 (light green), blade 4 (green), blade 5 (yellow) (Figure 8a). Domain β-propeller on exoinulinase *Aspergillus awamori* consists of 353 amino acid residues, from Phe20 to Gln372[19]. Protein fragment was

domain β-propeller was formed from amino acid residues 179. The protein fragment consist of three blades are blade 1 (yellow), blade 2 (yellow) and blade 3 (yellow) (Figure 7b).



a. Exoinulinase of *Aspergilllusawamori* (PDB accession number c1y9gA)

b. Exolevanase fragment of *Bacillus subtilis* UBCT-030

Figure 7. Structure modeling of exoinulinase and exolevanase fragment

The catalytic residues on the protein fragment were Asp4 (1-WMNDPNG-7), and Glu178 (177-WEC-179). The amino acid residues located on blade 1, and the blade 3 resvectively in the β-propeller domain of the protein fragment. The position of catalytic residues is equivalent position in the family GH32 enzyme. Catalytic residues of exoinulinase from *Aspergillus awamori* is Asp41 on blade 1 and Glu241 on blade 4. Position of Asp 41 was in motif 38-WMNDPNG-44, while position of Glu241 was in the motif 241-ECPGL-245 [19]. Active site on the protein fragment of *Bacillus subtilis* UBCT-030 were D (1-WLNDP-5), D (125-FRDPK-129) and E (177-WEC-179). Substrate binding were ND (1-WLNDP-5), Q (18- FYQY-21), FS (60-FSGS-63) RD (125-FRDPK-129) and E (177-WEC-179). Position of the residues in the blade was in Figure 8.

a. Catalytic site (red) b. Active site (red) c. Substrate binding (red)

Figure 8. Enzyme structure modeling of exolevanase fragment from *Bacillus subtilis* UBCT-030

4. Conclusions

Thepartial genehad been found539 bp from *Bacillus subtilis* UBCT-030 that encodes 179aminoacid residues of exolevanase fragment. The protein fragment had *specific hits* GH32 β-fructosidase. It had two residues catalytic site, three residues active site and five residues substrate binding

Acknowlegment

This research was funded by Hibah Doktor, from Directorate General for Higher Education, Ministry of National Education, Indonesia under contract number of 486/SP2H/PP/DP2M/VI/2010.

References

[1] Allais JJ, Hoyos-Lopez G, Kammoun S, Baratti JC. 1987. Isolation and characterization of the thermophilic bacterial strain with inulinase activity. *Appl. Enviromental Microbiol*.**53:**942–945.
[2] Souza AN, Martins MLL. 2001. Isolation, properties, and kinetics of growth of a thermophilic Bacillus. *Brazillian J. Mirobiology***32:** 271–275.
[3] Gao W, Bao Y, Liu Y, Zhang X. 2008. Characterization of thermostable endo-inulinase from a new strain *Bacillus Smithii* T7. *Appl. Biochem Biotechnol***157:**498–506.
[4] Singh RH, Chauhan K, Kennedy JE. 2016. A panorama of bacterial inulinase: production, purification, characterization and industrial applications. *International Journal of Biological Macromolecules.* **96:** 312-322.
[5] Vieille C, Zeikus GJ. 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev*.**65:** 1–43.
[6] Phelps CF. 1965.The physical properties of inulin solution. *Biochem.J***95:** 41–47.
[7] Zhang T, Gong F, Chi Z, Liu G, Chi Z, Sheng J, Li J and Wang X. 2009. Cloning and characterization of the inulinase gene from a marine yeast *Pichia guilliermondii* and its expression in *Pichia pastoris*. *Antonic van Leeuwenhoek***95:** 13–22.
[8] Menendez C, Hernandez L, Selman G, Mendoza MF, Hevia P, Sotolongo M, A. J. 2002.Molecular cloning and expression in Escherichia coli of an exo-levanase gene from the endophytic bacterium Gluconaceto-bacter diazotrophicus SRT4. *Curr. Microbiol*.**45:** 5–12.
[9] Azhar M, Natalia D, Syukur S, Vovien, Jamsari. 2015. The gene fragment that encodes inulin hydrolysis enzymes from genomic *Bacillus licheniformis* by PCR technique using new primers. *Int. J. Biol. Chem*.**9:** 59–69.
[10] Azhar M. 2013. Characterization of extracellular enzymes on inulin substrate and molecular levanase gene from thermophilic *Bacillus licheniformis* UBCT-007, local isolate of West Sumatra. (Universitas Andalas. Indonesia).
[11] Castro GR, Baigori MD, Sineriz F. 1995. A plate technique for screening of inulin degrading microorganisms. *J. Microbiol. Methods***22:** 51–56.
[12] PROMEGA. 2010. pGEM(R)-T and pGEM(R)-T Easy Vector Systems Technical Manual TM042 - pgem-t and pgem-t easy vector systems protocol.pdf.
[13] Sambrook J, Russell DW. 2001. *Molecular Cloning a Laboratory Manual*. pp.1.105-1.111. 3th Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
[14] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 Web Portal for Protein Modeling, Prediction, and Analysis. *Nature Protocols*. **10:** 845–858.
[15] Woodmen, M.E. 2008.*Current Protocols in Microbiology* A.3D.pp.1-6 (John Wiley & Sons, Inc.)
[16] Goosen C. 2007. Identification and characterization of glycoside hydrolase family 32 enzymes from *Aspergillus niger*. (Rijksuniversitet Groningen).
[17] Ohta K, Akimoto H, Moriyama S. 2004. Fungal inulinases: enzymology, molecular biology and biotechnology. *J.Appl.Glyosci***51:** 247–254.
[18] Adawiyah SR, Shuhaimi M, Yazid AMM, Manaf AA, Rosli N, S. 2011. Molecular cloning and sequence analysis of an inulinase gene from an *Aspergillus* sp. *World J Microbiol Biotechnol***27:** 2173–2185.
[19] Nagem RAP, Rojas AL, Golubev AM, Korneeva OS. 2004. Crystal structure of exo-inulinase from *Aspergillus awamori*: the enzyme fold and structural determinant of substrate recognition. *J.Mol.Biol***344:** 471–480.