

**CHARACTERIZATION OF A
THERMOSTABLE CELLULASES FROM
THERMOFLAVIFILUM AGGREGANS SP1
ISOLATED FROM PORING HOT SPRING
SABAH, MALAYSIA**



NURSHAFRINA AIDA BINTI YAHYA

UMS
UNIVERSITI MALAYSIA SABAH

**BIOTECHNOLOGY RESEARCH
INSTITUTE
UNIVERSITI MALAYSIA SABAH
2023**

**CHARACTERIZATION OF A
THERMOSTABLE CELLULASES FROM
THERMOFLAVIFILUM AGGREGANS SP1
ISOLATED FROM PORING HOT SPRING
SABAH, MALAYSIA**

NURSHAFRINA AIDA BINTI YAHYA

**THIS IS SUBMITTED IN FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE**



**BIOTECHNOLOGY RESEARCH
INSTITUTE**

UNIVERSITI MALAYSIA SABAH

2023

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL : **CHARACTERIZATION OF A THERMOSTABLE CELLULASES FROM THERMOFLAVIFILUM AGGREGANS SP1 ISOLATED FROM PORING HOT SPRING SABAH, MALAYSIA**

IJAZAH : **SARJANA SAINS**

BIDANG : **BIOTEKNOLOGI**

Saya **NURSHAFRINA AIDA BINTI YAHYA**, Sesi **2018-2023**, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)

TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD



**NURSHAFRINA AIDA BINTI
YAHYA
MZ1821104T**

Disahkan Oleh,



**ANITA BINTI ARSAD
PUSTAKAWAN KANAN**

UNIVERSITI MALAYSIA SABAH

(Tandatangan Pustakawan)



(Dr. Cahyo Budiman)
Penyelia Utama

Tarikh : 11 Mei 2023

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries and references, which have been duly acknowledged.

14 December 2022



Nurshafarina Aida binti Yahya

Nurshafarina Aida binti Yahya

MZ1821104T

UMMS
UNIVERSITI MALAYSIA SABAH

CERTIFICATION

NAME : **NURSHAFRINA AIDA BINTI YAHYA**

MATRIC NUM. : **MZ1821104T**

TITLE : **CHARACTERIZATION OF A THERMOSTABLE CELLULASES FROM THERMOFLAVIFILUM AGGREGANS SP1 ISOLATED FROM PORING HOT SPRING SABAH, MALAYSIA**

DEGREE : **MASTER OF SCIENCE**

FIELD : **BIOTECHNOLOGY**

VIVA DATE : **14 DECEMBER 2022**



CERTIFIED BY;

UMS
UNIVERSITI MALAYSIA SABAH

Signature

1. MAIN SUPERVISOR

Dr. Cahyo Budiman

2. CO - SUPERVISOR

Professor Dr. Clemente Michael Wong Vui Ling

3. CO - SUPERVISOR

Dr. Zarina Amin

ACKNOWLEDGEMENTS

In The Name of Allah, Most Gracious, Most Merciful

Assalamualaikum W.B.T

Alhamdulillah to Allah S.W.T that finally, with His blessings and permission as well as the cooperation from those who are involved in this study, I am able to complete it. I would like to say thank you, especially my supervisor, Dr. Cahyo Budiman and also my co supervisor Prof. Dr. Clemente Michael Wong Vui Ling and Dr. Zarina Amin for all the help, support, comments and guidance in helping me to finish this thesis. Furthermore, I would like to say thank you to all other lecturers that provided me with comments and advices to improve my work.

Also not to be forgotten, I would also like to thank you to my parents for their encouragement and financial assistance which had helped me to go through all the obstacles and challenges throughout this research study.

Last but not least, thank you for all of my friends from our team (Protein group) and others for the continuous help and support in completing my thesis.

Nurshafrina Aida binti Yahya

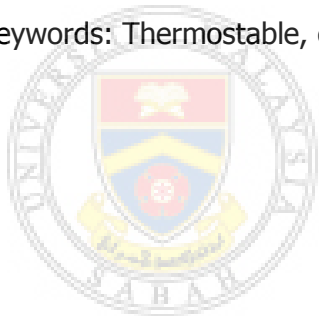
14 December 2022

ABSTRACT

Thermophilic microorganisms and their enzymes have been utilized in various industrial applications. Cellulases are essential enzymes in various industries and have great significance in present-day biotechnology. Currently, thermostable cellulases have been steadily increasing in demand due to their versatile applications under harsh conditions. Nevertheless, the availability of thermostable cellulases remains limited. Sabah houses hot springs and mud volcanoes that are promising as the sources of indigenous thermophilic bacteria producing thermostable cellulases. This study is aimed to isolate and characterize thermophilic cellulose-degrading bacteria from Sabah hot springs and mud volcanoes. To address, the samples were collected from Poring hot spring and Tawau mud volcanoes and spread onto the carboxymethylcellulose (CMC) agar medium for the screening of cellulose-degrading bacteria. Based on the biochemical test, SP1 was chosen and undergone whole genome sequencing, assembly and annotation. The gene was then cloned into pET-28a (+) and transformed into *Escherichia coli* BL21(DE3). The expression of this protein was successfully performed by induction of 0.2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 25 °C for overnight. The expressed protein was then successfully purified using Ni-NTA affinity chromatography and size-exclusion chromatography. As results, out of 6 isolated colonies exhibiting cellulose degradation activities, colony SP1 from Poring hot spring displayed the highest cellulolytic activity on crude enzyme at 60 °C. Further, 16S rRNA sequence analysis of SP1 showed closest similarity to *Thermoflavifilum aggregans* (accession no: AM749771), a thermostable bacterium isolated from New Zealand hot spring, with 99.74 % homology. Accordingly, SP1 is designated as *T. aggregans* SP1 strain. To note, the strain SP1 is the first strain from Poring hot spring known to exhibit thermostable cellulolytic activity. Whole genome sequence (WGS) of SP1 strain was then decoded using the Pacific Biosciences Single Molecule, Real-Time sequencing platform which revealed the genome size is 2,874,051 bp with the presence of three (3) genes encoding cellulose-degrading enzymes which might responsible for its cellulolytic activity. From the three genes, one gene is from glycosyl hydrolase (GH)-5 and another two genes from GH-9 family members. Due

to the unique properties of GH-5, one gene of GH-5 (designated as CePH4) was selected for further characterization. CePH4 is 1053 bp in size, which encodes a 347-residue polypeptide with theoretical molecular mass of 38 kDa. The sequence analysis of CePH4 indicated that this protein is organized into a catalytic domain of GH5 and an aryl-phospho-beta-D-glucosidase (BglC), with canonical active sites of Glu151 and Glu271. CePH4 was successfully expressed and purified with the yield of 24.22 mg per 1000 mL culture, with apparent size of 38 kDa on SDS. Size exclusion chromatography showed that CePH4 is a monomeric protein. Circular dichroism spectroscopy revealed that CePH4 has melting temperature (T_m) of 80 °C, indicating that this protein is a thermostable enzyme. Further, the specific activity of purified CePH4 against CMC substrate was 7.46 U/mg with the optimum activity at 70 °C and pH 7. Interestingly, up to 80°C, the activity of CePH4 decreased by up to 40 % only. Our findings suggest promising applications of these thermoaerobic bacteria and their potent enzymes for industrial purposes.

Keywords: Thermostable, cellulases, hot springs, whole genome sequence



UMS
UNIVERSITI MALAYSIA SABAH

ABSTRAK

CIRI-CIRI SEL TERMOSTABIL DARIPADA THERMOFLAVIFILUM AGGREGANS SP1 DIASINGKAN DARIPADA PORING HOT SPRING SABAH, MALAYSIA

Mikroorganisma termofilik dan enzimnya telah digunakan dalam pelbagai aplikasi perindustrian. Selulase adalah enzim penting dalam pelbagai industri dan mempunyai kepentingan yang besar dalam bioteknologi masa kini. Pada masa ini, selulase termostabil telah semakin meningkat dalam permintaan kerana aplikasi serba bolehnya di dalam keadaan yang melampau. Namun begitu, ketersediaan selulase termostabil kekal terhad. Sabah menempatkan mata air panas dan gunung berapi lumpur yang menjanjikan sebagai sumber bakteria termofilik asli yang menghasilkan selulase termostabil. Kajian ini bertujuan untuk mengasingkan dan mencirikan bakteria pengurai selulosa termofilik daripada mata air panas Sabah dan gunung berapi lumpur. Untuk menjawabnya, sampel telah dikumpulkan dari mata air panas Poring dan gunung berapi lumpur Tawau dan disebar ke medium agar carboxymethylcellulose (CMC) untuk saringan bakteria yang menguraikan selulosa. Berdasarkan ujian biokimia, SP1 telah dipilih dan menjalani penjujukan, pemasangan dan anotasi genom keseluruhan. Gen tersebut kemudiannya diklonkan kepada pET-28a (+) dan dimasukkan ke dalam *Escherichia coli* BL21(DE3). Ekspresi protein ini berjaya dilakukan dengan induksi 0.2 mM isopropil β -D-1-thiogalactopyranoside (IPTG) pada suhu 25 °C untuk semalaman. Protein kemudiannya berjaya ditulenkan menggunakan kromatografi afiniti Ni-NTA dan kromatografi pengecualian saiz. Hasilnya, daripada 6 koloni terpencil yang mempamerkan aktiviti degradasi selulosa, koloni SP1 dari mata air panas Poring menunjukkan aktiviti selulolitik tertinggi untuk enzim mentah pada 60 °C. Selanjutnya, analisis jujukan 16S rRNA SP1 menunjukkan persamaan yang paling hampir dengan *Thermoflavifilum aggregans* (nombor kesertaan: AM749771), bakteria termostabil yang diasingkan daripada mata air panas New Zealand, dengan 99.74 % homologi. Sehubungan itu, SP1 ditetapkan sebagai *T. aggregans* SP1 strain. Untuk makluman, terikan SP1 ialah terikan pertama dari mata air panas Poring yang diketahui mempamerkan aktiviti selulolitik termostabil. Keseluruhan penjujukan genom strain SP1 kemudiannya dinyahkodkan menggunakan Pacific

Biosciences Single Molecule, platform penjujukan Masa Nyata yang mempunyai saiz genom 2,874,051 bp dengan kehadiran tiga (3) gen yang mengkod enzim pengurai selulosa yang mungkin bertanggungjawab untuk aktiviti selulolitiknya. Daripada tiga gen tersebut, satu gen adalah daripada glikosil hidrolase (GH)-5 dan dua lagi gen daripada ahli keluarga GH-9. Disebabkan sifat unik GH-5, satu gen GH-5 (ditetapkan sebagai CePH4) telah dipilih untuk pencirian selanjutnya. CePH4 bersaiz 1053 bp, yang mengkodkan polipeptida 347-residu dengan jisim molekul teori 38 kDa. Analisis jujukan CePH4 menunjukkan bahawa protein ini disusun ke dalam domain pemangkin GH5 dan aril-phospho-beta-D-glucosidase (BglC), dengan tapak aktif kanonik Glu151 dan Glu271. CePH4 berjaya diekspresikan dan dituliskan dengan hasil 24.22 mg setiap 1000 mL kultur, dengan saiz ketara 38 kDa pada SDS. Kromatografi pengecualian saiz menunjukkan bahawa CePH4 ialah protein monomerik. Spektroskopi dichroism bulat mendedahkan bahawa CePH4 mempunyai suhu lebur (T_m) 80 °C, menunjukkan bahawa protein ini adalah enzim termostabil. Selanjutnya, aktiviti spesifik CePH4 yang telah dituliskan terhadap substrat CMC ialah 7.46 U/mg dengan aktiviti optimum pada 70 °C dan pH 7. Menariknya, sehingga 80°C, aktiviti CePH4 berkurangan sehingga 40 % sahaja. Penemuan kami mencadangkan aplikasi menjanjikan bakteria termoaerobik ini dan enzim kuatnya untuk tujuan perindustrian.

Kata kunci: Thermostabil, selulase, air panas, keseluruhan jujukan genom

LIST OF CONTENTS

	Page
TITLE	i
DECLARATION	ii
CERTIFICATION	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
<i>ABSTRAK</i>	vii
LIST OF CONTENTS	ix
LIST OF TABLES	xiii
LIST OF FIGURES	xv
LIST OF SYMBOL	xvii
LIST OF ABBREVIATIONS	xviii
LIST OF APPENDICES	xx
CHAPTER 1: INTRODUCTION	
1.1 Background Study	1
1.2 Problem Statement	5
1.3 Research Question	6
1.4 Research Hypothesis	6
1.5 Aim and Objectives	6
1.6 Significance of the Study	7
1.7 Scope and Limitation of Study	8
CHAPTER 2 : LITERATURE REVIEW	
2.1 Cellulose	9
2.2 Cellulases	10

2.2.1	Classification and Structure Of Cellulases	10
2.2.2	Catalytic Mechanism of Cellulases	11
2.3	Thermophilic Bacteria	16
2.3.1	Classification of Psychrophiles, Mesophiles and Thermophiles	16
2.3.2	Mechanism Adaptation of Thermophilic Bacteria	17
2.4	Thermostable Enzymes	18
2.5	Thermostable Cellulases	20
2.5.1	Discovery of Thermostable Cellulase	20
2.5.2	Adaptation of Thermostable enzymes to High Temperature	22
2.6	Hot Environments in Sabah: Hot Spring and Mud Volcano	23
2.7	Type of Application of Cellulases	24
2.7.1	Pulp and Paper	24
2.7.2	Textile Industry	24
2.7.3	Bioethanol	25
2.7.4	Laundry	25
2.7.5	Food and Feed	26
2.7.6	Agriculture	26
2.7.7	Brewing	27
2.8	Whole Genome Sequencing to Reveal Functional Genes	27
2.9	Heterologous Production of Enzymes for Large Scale Production	29
2.10	Local Cellulases for Local Needs in Industry	30

CHAPTER 3 : METHODOLOGY

3.1	Screening and Isolation Bacteria Producing Thermostable Cellulose Degrading Enzyme	32
3.1.1	Study Area	31
3.1.2	Culture Enrichment and Isolation	33
3.1.3	Screening of Cellulolytic Activity of Bacterial Isolates	34
3.1.4	Molecular Identification of Bacteria Producing Cellulose Degrading Bacteria	34
3.1.5	Bacterial Cell Morphology of Cellulose Degrading Bacteria	37
3.1.6	Growth Profile Analysis	38
3.1.7	Enzymatic Assay	38

3.1.8	Crude Enzyme Characterization	39
3.2	Whole Genome Sequencing, Assembly, and Annotation	40
3.2.1	Genomic DNA Extraction	40
3.2.2	Library Preparation and Sequencing	40
3.2.3	Sequence Assembly and Annotation	40
3.3	Protein Structural Analysis	41
3.3.1	Sequence Retrieval and Alignment	41
3.3.2	Primary Sequence Analysis	42
3.3.3	Secondary Structure and Signal Peptide	42
3.3.4	3D Structure Prediction Using Homology Modeling, Model Evaluation, and Refinement	42
3.3.5	Molecular Docking	42
3.3.6	Molecular Docking Simulation	43
3.4	Expression and Purification of the Enzyme	44
3.4.1	Gene Synthesis and Expression System Construction	44
3.4.2	Expression of Enzyme	45
3.4.3	Purification of the Enzyme	45
3.5	Biophysical Characterization	46
3.5.1	Oligomerization	46
3.5.2	Circular Dichroism Spectra Analysis	47
3.5.3	Thermal Stability Analysis	47
3.6	Biochemical Characterization	47
3.6.1	The Optimum PH and Stability	48
3.6.2	The Optimum Temperature and Stability	48
3.6.3	Effect of Various Additives on Enzyme Activity	48
3.6.4	Specific and Kinetics Activity	48
 CHAPTER 4 : RESULTS AND DISCUSSION		
4.1	Isolation and Screening of Cellulase Producing Bacteria	49
4.1.1	Sampling Locations	49
4.1.2	Isolation and Screening of Thermophilic Cellulase Producing Bacteria	53
4.1.3	Molecular Identification of Bacteria Strains	56

4.1.4	Morphological and Biochemical Examination of the Isolates	60
4.1.5	Optimal Growth Condition of SP1 Strain	66
4.1.6	Cellulolytic Activity of Crude Enzyme Characterization	68
4.2	Whole Genome Sequencing, Assembly and Annotation	75
4.2.1	Genomic Profile	75
4.2.2	Gene Involves in Production of Cellulose-Degrading Enzymes	79
4.2.3	Gene Involved in Thermal Adaptation	85
4.3	Amino Acid Sequence Analysis of CePH4	87
4.3.1	Primary Structure	87
4.3.2	Theoretical Physicochemical Properties	88
4.3.3	Active Site Identification	94
4.3.4	Secondary Structure Profile	95
4.3.5	Three-Dimensional Model Structures	98
4.3.6	Molecular Docking	102
4.3.7	Molecular Dynamic Simulation	104
4.4	Heterologous Production of CePH4	106
4.4.1	Expression System Construction of CePH4	106
4.4.2	Protein Expression	108
4.4.3	Purification of CePH4	111
4.5	Characterization of the Enzyme	113
4.5.1	Oligomerization	113
4.5.2	Circular Dichroism Spectra Analysis	114
4.5.3	Catalytic Properties of CePH4	117
CHAPTER 5 : CONCLUSIONS AND FUTURE PROSPECT		126
5.1	Conclusion	126
5.2	Future Prospect	128
REFERENCES		130
APPENDICES		160

LIST OF TABLES

	Page
Table 2.1 : The various groups of bacteria producing cellulase	16
Table 2.2 : Thermostable enzymes from different organisms	19
Table 2.3 : Sources of thermostable cellulases	21
Table 3.1 : Cellulolytic index formula	34
Table 3.2 : The 16s rRNA primers	35
Table 3.3 : PCR amplification reagents used per reaction for 16s rRNA amplification of the isolates	36
Table 3.4 : PCR cycling condition	37
Table 3.5 : Root mean square fluctuation (RMSF) formula	44
Table 4.1 : The environmental characteristics in Poring hot spring, Tawau hot spring and Kalabakan mud volcano	50
Table 4.2 : Number of isolates from all sample sites and four bacterial isolates shows halo zone on CMC agar plate	54
Table 4.3 : Screening for cellulolytic activity of four bacterial isolates on CMC agar plate	55
Table 4.4 : Identity of SP1 strain based on NCBI BLAST and SILVA search	58
Table 4.5 : Differential characteristics of SP1 strain and selected type strains of species of the family <i>Chitinophagaceae</i>	62
Table 4.6 : Physiological and biochemical properties of SP1 strain	63
Table 4.7 : Growth condition comparison of <i>Thermoflavifilum aggregans</i> SP1 with selected type strains of species of the family <i>Chitinophagaceae</i>	68
Table 4.8 : Comparison of enzymatic activity of <i>Thermoflavifilum aggregans</i> SP1 with other thermophilic cellulases	70
Table 4.9 : Strain and genome description of <i>Thermoflavifilum aggregans</i> SP1 strain	76
Table 4.10 : Sequencing statistics of <i>Thermoflavifilum aggregans</i> SP1 strain	77

Table 4.11	:	Distribution of subsystem annotation of <i>Thermoflavifilum aggregans</i> SP1 strain at level 1	80
Table 4.12	:	Distribution of subsystem annotation of <i>Thermoflavifilum aggregans</i> SP1 strain at carbohydrate subsystem level 2	81
Table 4.13	:	Carbohydrate active enzymes count of SP1	84
Table 4.14	:	Heat shock protein count of <i>Thermoflavifilum aggregans</i> SP1 strain	86
Table 4.15	:	Comparison of physicochemical properties of thermophilic and mesophilic bacteria within GH5 family	92
Table 4.16	:	Comparable percentage of α -helix, β -turns, extended strand and random coil within all GH5-cellulase sequences	97
Table 4.17	:	3D structure prediction and validation	100
Table 4.18	:	The structural alignment comparison between CePH4 with another thermophilic and mesophilic cellulases	102
Table 4.19	:	The purification fold of CEPH4	113
Table 4.20	:	The comparison of kinetic parameters between CePH4 and other cellulases	119
Table 4.21	:	Differential temperature, temperature stability, pH and pH stability of CePH4 and other cellulase protein of selected type strains of species from the GH5 family	122

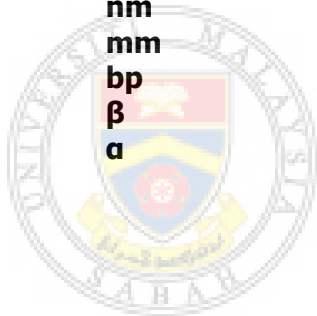
LIST OF FIGURES

	Page
Figure 2.1 : Cellulose structure	10
Figure 2.2 : Structural conservation and diversity of cellulases	11
Figure 2.3 : Complete cellulose hydrolysis	13
Figure 2.4 : Catalytic mechanisms of cellulases, inverting mechanism (a) and retaining mechanism (b)	14
Figure 2.5 : The pipeline of genome assembly and annotation by long reads	29
Figure 3.1 : The study area in Sabah	33
Figure 3.2 : Expression system of pET-28a-CePH4	45
Figure 4.1 : Genome extraction (a) and PCR result of SP1 (b)	56
Figure 4.2 : Phylogenetic tree based on 16S rRNA gene sequences of SP1 strain with related species was using MEGA X software	59
Figure 4.3 : Gram staining (a) and scanning electron microscope (b) of isolate SP1	61
Figure 4.4 : Effects of temperature (a) and pH (b) on the growth of SP1 strain	67
Figure 4.5 : Effects of temperature on enzyme activity	71
Figure 4.6 : Effects of pH on enzyme activity	72
Figure 4.7 : Effects of additives on enzyme activity	74
Figure 4.8 : Single genome map overview of the <i>Thermoflavifilum aggregans</i> SP1	78
Figure 4.9 : PCR amplification of 5 putative cellulases of SP1	84
Figure 4.10 : Molecular architecture of CePH4 amino acid sequence	88
Figure 4.11 : Multiple sequence alignment (MSA) of CePH4 from the first row and another reference gene from GH5 family	94
Figure 4.12 : Schematic wiring diagrams of different cellulase structures	96
Figure 4.13 : Predicted 3D structure of CePH4 with both active sites	99

Figure 4.14 :	Ramachandran plot of CePH4 obtained through modelling revealed that most of the residue of <i>Thermoflavifilum aggregans</i> endoglucanase were located in favored region (93 %)	101
Figure 4.15 :	2D diagram of CePH4 docked with cellobiose	103
Figure 4.16 :	RMSF plot evidence of specific residue positions at 37 °C, 60 °C, and 90 °C	106
Figure 4.17 :	Colonies of successful transformant of pET28a-CePH4 into <i>E. coli</i> BL21 (DE3)	109
Figure 4.18 :	Validation of expression using SDS-PAGE gel at 25 C	110
Figure 4.19 :	Purification results of CePH4	112
Figure 4.20 :	Elution profile for size exclusion chromatography	114
Figure 4.21 :	Far-UV CD spectra of CePH4	115
Figure 4.22 :	Melting temperature (T_m) of CePH4	117
Figure 4.23 :	Kinetic parameters of CePH4	118
Figure 4.24 :	Effect of (a) pH on enzyme activity, (b) pH on enzyme stability	120
Figure 4.25 :	Effect of (a) temperature on enzyme activity, (b) temperature on enzyme stability	121
Figure 4.26 :	Effects of additives on enzyme activity	124

LIST OF SYMBOLS

°C	-	Degree celcius
%	-	Percentage
μL	-	Microliter
μmol	-	Micromol
w/v	-	Weight over volume
mM	-	Millimolar
U	-	Unit
ml	-	Mililiter
μg	-	Microgram
kb	-	Kilobase pair
Å	-	Angstrom
kcal/mol	-	kilocalorie per mole
pM	-	picomole
K	-	Kelvin
mg/ml	-	Miligram per milliliter
ml/min	-	Milliliter per minute
kDa	-	Kilodaltons
nm	-	Nanometer
mm	-	Millimeter
bp	-	Base pair
β	-	Beta
α	-	Alpha



UMS
UNIVERSITI MALAYSIA SABAH

LIST OF ABBREVIATIONS

AI	-	Aliphatic index
BglC	-	Aryl-phospho-beta-D-glucosidase
CaCl₂	-	Calcium chloride
CBB	-	Coomassie brilliant blue
CBM	-	Carbohydrate-binding module
CDS	-	Protein-coding sequences
CMC	-	Carboxymethyl cellulose
CI	-	Cellulolytic index
CuSO₄	-	Copper(II) sulfate
COG	-	Clusters of orthologous groups
GH	-	Glycoside Hydrolase
gDNA	-	Genomic DNA
GO	-	Gene ontology
GRAVY	-	Grand Average Hydropathy
GMQE	-	Global Model Quality Estimation
HGAP4	-	Hierarchical genome-assembly process 4
HGT	-	Horizontal gene transfer
HMF	-	5-hydroxymethylfurfural
HSP	-	Heat shock protein
IPTG	-	Isopropyl -D-1-thiogalactopyranoside
KCl	-	Potassium chloride
KEGG	-	Kyoto Encyclopedia of Genes and Genomes
MD	-	Molecular docking
NaCl	-	Sodium chloride
NADH	-	Nicotinamide adenine dinucleotide
NCBI	-	National Center for Biotechnology Information
OD600	-	Optical density at 600 nm
PDB	-	Protein information bank
PEG	-	Protein encoding gene
<i>pI</i>	-	Isoelectric point
QMEAN	-	Qualitative Model Energy Analysis
Rpm	-	Revolutions per minute
RAST	-	Rapid Annotations using subsystems Technology
RMSF	-	Root means square fluctuation
Sdn Bhd	-	Sendirian Berhad

SEM	-	Scanning electron microscope
SDS-PAGE	-	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
sHSP	-	Small heat shock proteins
SMRTbell	-	Single-molecule, real-time bell
TIM	-	Triose-phosphate isomerase
T_m	-	Transition temperature
UV CD	-	Ultraviolet circular dichroism
YASARA	-	Yet Another Scientific Artificial Reality Application
ZnSO₄	-	Zinc sulphate
3D	-	Three-dimensional

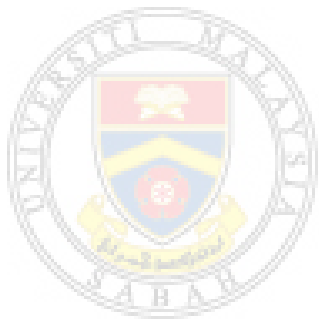


UMS
UNIVERSITI MALAYSIA SABAH

LIST OF APPENDICES

Appendix A : pET-28a Vector Map

Page
160



UMS
UNIVERSITI MALAYSIA SABAH

CHAPTER 1

INTRODUCTION

1.1 Background

Lignocellulosic biomass is the world's most economical and highly renewable natural resource. The development of renewable energy converted from lignocellulosic biomass as an alternative to fossil fuel is ultimately essential for the survival of humans (Zoghalmi & Paes, 2019). Apart from its application as a material for producing renewable energy, there has also been increased interest in developing lignocellulosic biomass-derived platform chemicals for biobased polyurethane applications. Technologies have been developed to produce various platform chemicals, such as sugar alcohols, organic acids, furfural, and 5-hydroxymethylfurfural (HMF), from lignocellulosic biomass via biorefining technologies (Kohli *et al.*, 2019). For these purposes, cellulose degrading enzyme or cellulases (E.C. 3.2.1.4) is required to process lignocellulosic biomass. The enzymes work through hydrolysis of cellulose to its simple sugar for further conversion to bioproducts through fermentation or other chemical processes. In nature, complete hydrolysis of cellulose is done by the three main cellulases types: endoglucanase, exoglucanase, and beta-glucosidase, which converts cellulose into sugar (glucose) (Jayasekara & Ratnayeke, 2019).

Nowadays, cellulases have been gaining broad interest in the industry due to their ability to replace inorganic catalysts for more environmentally friendly industrial processes and minimize carbon emissions. In Malaysia, cellulase enzymes are majorly

utilized in the industry related to edible oil and palm oil products, the oleochemical industry, the detergent industry, and the food and beverages, manufacturing, animal feed, and baking industries (Kiew *et al.*, 2012). Notably, Sabah is known to be the biggest producer of palm oil products. In 2021, national production reached 26.48 million tonnes, with more than 6.29 million tonnes being contributed by Sabah (Malaysian Palm Oil Board, 2021). This huge industry leads to a huge number of by-products or cellulosic wastes, which can be converted into value-added products such as bioethanol and sustainable bio-based products (Kunasundari *et al.*, 2016; Goh *et al.*, 2010). Such processes are usually triggered by cellulose degrading enzymes produced by specific bacteria and microorganisms.

A critical bottleneck to the wide industrial use of hydrolytic enzymes is that they are often required to perform under harsh conditions, including temperatures above 40 °C. As a result, the enzyme stability against elevated temperatures is a crucial prerequisite before the broad industrial use of this type of enzyme is realized (Santos & da Costa, 2002). Accordingly, thermostable enzymes and microorganisms have been topics of much research during the last two decades (Singh *et al.*, 2016). In industrial applications, isolated thermostable enzymes are today dominating over thermostable microorganisms. In this respect, the thermostable enzymes undergo some purification process from target microorganisms. An enzyme or protein is classified as thermostable when it exhibits optimum activity between 60 - 80 °C (Vieille & Zeikus, 2001). In some cases, enzymes also exhibited optimum activity at temperatures higher than 80 °C, which is then classified as hyperthermophilic (Viele *et al.*, 1996). Active at high temperatures, thermophilic and hyperthermophilic enzymes typically do not function below 40 °C. Meanwhile, as producers of thermostable enzymes, thermostable microorganisms are usually characterized by their optimum growth temperature range of 50 - 80 °C (Vieille & Zeikus, 2001).

In nature, fungi tend to produce more cellulases than bacteria, but cellulases produced by bacteria are better catalysts as they encounter less feedback inhibition. Moreover, bacteria not only grow rapidly compared to fungi, allowing for higher