CLONING AND EXPRESSION OF ENZYMES FROM GLYCOSYL HYDROLASE FAMILY (XYLANASE AND NEOPULLULANASE) FROM THE SEDIMENTS OF BOR KHLUENG HOT SPRING

VIRIYA NITTERANON

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(MOLECULAR GENETICS AND GENETIC ENGINEERING)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2006

ISBN 974-04-7010-6

Thesis

Entitled

CLONING AND EXPRESSION OF ENZYMES FROM GLYCOSYL HYDROLASE FAMILY (XYLANASE AND NEOPULLULANASE) FROM THE SEDIMENTS OF BOR KHLUENG HOT SPRING

VIRIYA NITTERANON
.....
Miss Viriya Nitteranon
Candidate

.....
Asst. Prof. Kusol Pootanakit, Ph.D.
Major-Advisor

.....
Asst. Prof. Sutipa Tanapongpipat, Ph.D.
Co-Advisor

.....
Asst. Prof. Witoon Tirasophon, Ph.D.
Co-Advisor

.....
Prof. M.R. Jisnuson Svasti, Ph.D.
Dean
Faculty of Graduate Studies

.....
Asst. Prof. Varaporn Akkarapatumwong, Ph.D.
Chair
Master of Science Programme in Molecular Genetics and Genetic Engineering
Institute of Molecular Biology and Genetics

Thesis
Entitled

CLONING AND EXPRESSION OF ENZYMES FROM GLYCOSYL HYDROLASE FAMILY (XYLANASE AND NEOPULLULANASE) FROM THE SEDIMENTS OF BOR KHLUENG HOT SPRING

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science (Molecular Genetics and Genetic Engineering)
on
March 29, 2006

VIRIYA NITTERANON
Miss Viriya Nitteranon
Candidate

Asst. Prof. Kusol Pootanakit, Ph.D.
Chair

Asst. Prof. Sutipa Tanapongpipat, Ph.D.
Member

Mr. Verawat Champreda, Ph.D.
Member

Asst. Prof. Witoon Tirasophon, Ph.D.
Member

Prof. M.R. Jisnuson Svasti, Ph.D.
Dean
Faculty of Graduate Studies
Mahidol University

Asst. Prof. Chartchai Krittanai, Ph.D.
Acting Director
Institute of Molecular Biology and Genetics
Mahidol University

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisors, Asst. Prof. Kusol Pootanakit and Asst. Prof. Sutipa Tanapongpipat for their great advice, helpfulness as well as their attempt to revise this thesis. My appreciation is also expressed to Asst. Prof. Witoon Tirasophon, Dr. Verawat Champreda and Asst. Prof. Varaporn Akkarapatumwong for their helpful discussions and recommendations.

I am extremely thankful to the scholarship from Thailand Graduate Institute of Science and Technology (TGIST), National Science and Technology Development Agency (NSTDA) for the financial support during 2004-2005.

I am also sincerely thankful to Miss Supawadee Chanapan, Miss Katewadee Boonyapakron, Miss Nongluk Plongthongkum and Miss Supattra Treerattrakul for her suggestion and assistance in laboratory techniques.

My appreciation is extended to Miss Benjamaporn Wonganu, Mr. Plaipol Dedvisitsakul, Miss Thitirat Ngaotepprutaram, Miss Varunya Vudtijun, lab members in B203, B205 and enzyme screening laboratory for their friendship and support.

I am also indebted to all my friends, staff, and technicians of the Institute of Molecular Biology and Genetics for their friendships, generosity, and sharing the happy time throughout my studying.

Finally, I would like to express my special deepest appreciation to my beloved parents and my family for their love, encouragement, and understanding throughout my life.

Viriya Nitteranon

CLONING AND EXPRESSION OF ENZYMES FROM GLYCOSYL HYDROLASE FAMILY (XYLANASE AND NEOPULLULANASE) FROM THE SEDIMENTS OF BOR KHLUENG HOT SPRING

VIRIYA NITTERANON 4636864 MBMG/M

M.Sc. (MOLECULAR GENETICS AND GENETIC ENGINEERING)

THESIS ADVISORS: KUSOL POOTANAKIT, Ph.D.,
SUTIPA TANAPONGPIPAT, Ph.D., WITOON TIRASOPHON, Ph.D.

ABSTRACT

Glycosyl hydrolases comprise a huge group of enzymes (xylanases, neopullulanases, amylases, etc.) and are valuable carbohydrate-degrading enzymes in biotechnological industries. In this study, to obtain genes encoding thermostable enzymes that are required for industrial processes, we have turned to directly obtaining the genes from environmental DNA, specifically, the Bor Khlueng hot spring. Consensus primers based on the conserved regions of family 10 xylanases and family 13 glycosyl hydrolases were used to obtain 167 bp and 560 bp PCR fragments of xylanase and neopullulanase genes, respectively. Sequence analysis of the partial xylanase gene exhibited 65% amino acid sequence identities to *Thermobacillus xylanilyticus xynA*. However, the full-length xylanase gene was not successfully obtained. For neopullulanase, the amino acid sequence analysis of one clone, BK44, showed 54% identity to glycosyl hydrolase family 13 of *Deinococcus radiodurans* and 51% identity to *Geobacillus kaustophilus* alpha-cyclodextrinase. This suggested that the obtained partial sequence from Bor Khlueng hot spring encoded a novel enzyme in this family. Using genome walking approaches, the 3' and 5'-end of the gene were obtained. The result showed that the full-length BK44 had an open reading frame of 1,458 bp encoding 485 amino acid residues and exhibited 54% identity to maltogenic amylase of *Thermus* sp., and alpha-cyclodextrinase of *Geobacillus kaustophilus*; and 53% identity to (neo)pullulanase of *Thermus thermophilus*. This suggested that BK44 belonged to the neopullulanase subfamily. Expression of the full-length BK44 in *P. pastoris* and *E. coli* was performed. The expressed protein of approximately 55 kDa was successfully obtained in *E. coli*. However, its enzymatic activities were not detected.

KEY WORDS: XYLANASE/ NEOPULLULANSE/ GLYCOSYL HYDROLASES/ GENOME WALKING/ *PICHIA PASTORIS*

116 P. ISBN 974-04-7010-6

การโคลนและการแสดงออกของยีนที่สร้างเอนไซม์ไซแลนเนสและนีโอพูลูแลนเนสจากสิ่งมีชีวิตในตัวอย่างดินน้ำพุร้อน อำเภอ บ่อคลึง (CLONING AND EXPRESSION OF ENZYMES FROM GLYCOSYL HYDROLASE FAMILY (XYLANASE AND NEOPULLULANASE) FROM THE SEDIMENTS OF BOR KHLUENG HOT SPRING))

วิริยา นิตย์ธีรานนท์ 4636864 MBMG/M

วท.ม. (อณูพันธุศาสตร์และพันธุวิศวกรรมศาสตร์)

คณะกรรมการควบคุมวิทยานิพนธ์ : กุศล ภูธนกิจ, Ph.D., สุทิพา ธนพงศ์พิพัฒน์, Ph.D., วิฑูรย์ ถิระโสภณ, Ph.D.

บทคัดย่อ

กลุ่มเอนไซม์ย่อยแป้ง (ไซแลนเนส, นีโอพูลูแลนเนส, อะไมเลส ฯลฯ) เป็นเอนไซม์ที่มีความสำคัญในอุตสาหกรรมเทคโนโลยีชีวภาพ งานวิจัยนี้ได้ทำการโคลนชิ้นส่วนของยีนไซแลนเนสและนีโอพูลูแลนเนสขึ้นมาจากสิ่งมีชีวิตในตัวอย่างดินของน้ำพุร้อน อ. บ่อคลึง จ. ราชบุรีโดยวิธีพีซีอาร์ ชิ้นส่วนของยีนขนาด 167 และ 560 คู่เบสที่ได้เมื่อนำไปเปรียบเทียบกับฐานข้อมูลพบว่าชิ้นส่วน 167 คู่เบสมีความคล้ายคลึง 65% ในลำดับกรดอะมิโนกับยีน *Thermobacillus xylanilyticus xynA* ส่วนชิ้น 560 คู่เบส โคลน BK44 มีความคล้ายคลึง 54% ในลำดับกรดอะมิโนเมื่อเปรียบเทียบกับเอนไซม์ glycosyl hydrolase family 13 ของ *Deinococcus radiodurans* และ 51% กับเอนไซม์ alpha-cyclodextrinase ของ *Geobacillus kaustophilus* โดยได้นำลำดับเบสจากชิ้นยีนที่เลือกมาทำการโคลนต่อไปโดยวิธีการ genome walking PCR จากผลการศึกษาพบว่าไม่สามารถโคลนยีนไซแลนเนสทั้งชิ้นได้ ในทางตรงกันข้ามชิ้นส่วนของยีนนีโอพูลูแลนเนสถูกโคลนขึ้นโดยได้ open reading frame ที่มีความยาว 1,458 นิวคลีโอไทด์ ซึ่งถอดรหัสให้โปรตีนที่มีความยาว 485 กรดอะมิโน เมื่อเปรียบเทียบในฐานข้อมูลพบว่าโปรตีนมีความคล้ายคลึง 54% ในลำดับกรดอะมิโนกับเอนไซม์ maltogenic amylase ของ *Thermus* sp. และ alpha-cyclodextrinase ของ *Geobacillus kaustophilus* และมีความคล้ายคลึง 53% กับเอนไซม์ (neo)pullulanase ของ *Thermus thermophilus* ซึ่งทั้งหมดเป็นเอนไซม์ในกลุ่มนีโอพูลูแลนเนส ผลการศึกษาการแสดงออกของ BK44 ใน *P. pastoris* และ *E. coli* พบว่าสามารถผลิตโปรตีนขนาดประมาณ 55 kDa ใน *E. coli* ได้ อย่างไรก็ตามจากการศึกษาเบื้องต้นพบว่าโปรตีนนี้ไม่สามารถแสดงคุณสมบัติของเอนไซม์ต่อซับสเตรทที่จำเพาะต่อยีนนีโอพูลูแลนเนส

116 หน้า. ISBN 974-04-7010-6

CONTENTS

Page

ACKNOWLEDGEMENTS iii
ABSTRACT iv
LIST OF TABLES xi
LIST OF FIGURES xii
LIST OF ABBREVIATIONS xv

CHAPTER
I INTRODUCTION
1.1 Xylanases 2
1.1.1 Structure of xylan 2
1.1.1.1 Hemicellulose from hardwood 2
1.1.1.2 Hemicellulose from softwood 2
1.1.2 Xylanolytic enzymes 2
1.1.3 Classification of xylanase family 5
1.1.4 Application of xylanases 5
1.2 Family 13 glycosyl hydrolase (α-amylase family) 7
1.2.1 Classification of α-amylase family 7
1.2.2 The common features of the α-amylase family 8
1.2.3 Pullulan 12
1.2.4 Application of pullulanases 12
1.2.5 Pullulan degrading enzymes 14
1.2.6 CD-/pullulan degrading enzymes (neopullulanase subfamily) 16
1.2.6.1 Cyclomaltodextrinase (EC 3.2.1.54; CDase) 16
1.2.6.2 Maltogenic amylase (EC 3.2.1.133; MAase) 16
1.2.6.3 Neopullulanase (EC 3.2.1.135; NPase) 17
1.2.6.4 3D structure of neopullulanase subfamily 17
1.3 The role of thermostable enzymes 19

CONTENTS (Continued)

	Page
1.4 Direct DNA isolation from environment	19
1.5 Culture-independent approaches (activity-based screening and sequence-based screening approaches)	20
II OBJECTIVES	22
III MATERIALS	
3.1 Bacterial strains	23
3.2 Yeast strain	23
3.3 Plasmid vectors	23
3.4 Sampling site	23
3.5 Synthetic oligonucleotides	24
3.6 Culture media	24
3.6.1 Bacterial culture medium (pGEM®-T Easy vector)	24
3.6.2 Bacterial culture medium (pPICZαA vector)	24
3.6.3 Yeast culture medium	24
3.6.3.1 Culture and selective media	24
3.6.3.2 Expression media	24
3.7 Chemicals	25
3.8 Kits	25
3.9 Enzymes	25
IV METHODS	
4.1 Genomic DNA extraction with SDS-based DNA extraction method	31
4.2 DNA concentration and purity determination	31
4.3 DNA electrophoresis	32
4.4 DNA purification by QIAquick gel extraction kit (QIAGEN kit)	32
4.5 Amplification of bacterial 16S rDNA gene	33
4.6 DNA amplification of xylanase and pullulanase gene	33

CONTENTS (Continued)

	Page
4.6.1 Degenerate primers design	33
4.6.2 PCR amplification	33
4.6.2.1 PCR amplification of xylanase gene	33
4.6.2.2 PCR amplification of pullulanase gene	34
4.7 Cloning of xylanase and pullulanase genes	34
4.7.1 Competent cell preparation using DMSO	34
4.7.2 DNA ligation	34
4.7.3 Transformation of competent *E. coli* DH5α cells	35
4.7.4 Master plate preparation and recombinant clone screening	35
4.8 Plasmid extraction	35
4.8.1 Plasmid extraction by CTAB method	35
4.8.2 Extraction of plasmid DNA using QIAprep spin Miniprep Kit (QIAGEN)	36
4.9 Insert size screening by restriction enzyme digestion	37
4.10 DNA sequencing	37
4.11 Sequence analysis	37
4.12 Amplification of 5' and 3' ends of xylanase and BK44 partial gene sequences	37
4.12.1 Genome walking PCR (Clontech, USA)	37
4.12.2 Semi-random two step PCR	38
4.13 Cloning and amplification of a full length BK44 gene	39
4.14 Construction of BK44 in *P. pastoris* expression vector	41
4.14.1 Amplification of the full-length BK44	41
4.14.2 Preparation of linearized pPICZαA	41
4.15 Transformation of the recombinant plasmid into *P. pastoris*	42
4.15.1 Preparation of *P. pastoris* competent cells	42
4.15.2 Preparation of linearized recombinant plasmids	42

CONTENTS (Continued)

Page

4.15.3 Transformation of linearized plasmid into *P. pastoris* cells by electroporation 42
4.16 Total DNA isolation from *P. pastoris* transformants 42
4.17 PCR analysis of *P. pastoris* integrants 43
4.18 Expression of recombinant BK44 in *P. pastoris* 44
4.19 Cell lysis of *P. pastoris* 44
4.20 Total RNA isolation of *P. pastoris* by TRI REAGENT® 44
4.21 Removal of DNA contamination for total RNA 45
4.22 First-strand cDNA synthesis 45
4.23 Amplification of pullulanase BK44 gene from recombinant *P. pastoris* 46
4.24 Construction of recombinant BK44 in *Escherichia coli* expression vector 46
4.24.1 Construction of recombinant pET-32a(+) plasmid harboring BK44 full length gene 46
4.24.2 Construction of recombinant pET-43.1a(+) plasmid harboring BK44 full length gene 46
4.24.3 Expression of recombinant BK44 in *E. coli* 47
4.25 Protein electrophoresis 47
4.25.1 Sample preparation 47
4.25.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) 48
4.26 Enzyme activity assay 48
4.26.1 Pullulanase activity assay using AZCL-pullulan 48
4.26.2 Cyclodextrinase activity assay using DNS method 48

V RESULTS

5.1 Direct extraction and purification of genomic DNA from the sediment of Bor Khlueng hot spring 51

CONTENTS (Continued)

	Page
5.2 PCR amplification using xylanase degenerate primers	51
5.3 Amplification of the 3' end and 5' end of partial xylanase gene using genomic walking	55
5.4 PCR amplification using α-amylase family degenerate primer	66
5.5 Amplification of the 5' and 3' end of BK44 using genomic waking method	66
5.6 Amplification of the full-length BK 44	72
5.7 Construction of recombinant plasmid harboring BK44 used for expression in *P. pastoris*	80
5.8 Integration of BK44 into *P. pastoris* genome	80
5.9 Expression of BK44 in *P. pastoris*	83
5.10 Investigation of BK44 transcript in *P. pastoris* by RT-PCR	83
5.11 Expression of BK44 in *E. coli*	87
5.11.1 Construction of BK44 in bacterial expression vector (pET-32a)	87
5.11.2 Expression of BK44 using pET-32a in *E. coli*	87
5.11.3 Expression of BK44 using pET-43.1a in *E. coli*	91
VI DISCUSSION	
6.1 Molecular cloning of family 10 and family 13 glycosyl hydrolases	95
6.2 Expression of BK44 in *P. pastoris*	96
6.3 Expression of BK44 in *E. coli*	98
VII CONCLUSIONS	100
REFERENCES	101
BIOGRAPHY	116

LIST OF TABLES

Table		Page
1.	Enzyme applications in industrial processes	1
2.	The four highly conserved regions of family 13 glycosyl hydrolase	11
3.	Primers used in each experiment	30
4.	Components of SDS-PAGE for protein determination	50

LIST OF FIGURES

Figure		Page
1.	Xylanolytic enzyme system	4
2.	Ribbon presentation of the main fold of the catalytic domains of EXs of family 10 and 11	4
3.	The enzymatic activities of the family 13 glycosyl hydrolase	10
4.	Ribbon representation of the conserved domains of α-amylase family	11
5.	The structure of pullulan	13
6.	Action patterns of pullulan-hydrolyzing enzymes on pullulan	15
7.	Ribbon representative of the conserved domain of *Thermoactinomyces vulgaris* R-47 neopullulanase (TVA-II)	18
8.	The physical map of pGEM® -T Easy vector (Promega)	26
9.	A physical map of pPICZαA vector	27
10.	A physical map of pET-32a(+) vector	28
11.	A physical map of pET-43.1a(+) vector	29
12.	Schematic of genomic walking PCR (Clontech kit)	40
13.	Semi-random two step PCR (ST-PCR)	41
14.	Genomic DNA extraction using SDS based DNA extraction method	52
15.	Determination of purified DNA	53
16.	PCR amplification of partial xylanase gene	54
17.	Restriction enzyme analysis with *Eco*RI digestion	56
18.	Alignment of the deduced amino acid sequences of partial xylanase containing 149 bp from the sediment of Bor Khlueng hot spring	57
19.	Alignment of the deduced amino acid sequences of partial xylanase containing 152 bp from the sediment of Bor Khlueng hot spring	57
20.	Alignment of the deduced amino acid sequences of partial xylanase containing 167 bp from the sediment of Bor Khlueng hot spring	58

LIST OF FIGURES (Continued)

Figure **Page**

21. Nucleotide and deduced amino acid sequences of partial xylanase from the sediment of Bor Khlueng hot spring 58
22. PCR amplification of xylanase downstream walking 59
23. Nucleotide and deduced amino acid sequences of 3’ end of xylanase gene 60
24. Amino acid alignment of Bor Khlueng hot spring xylanase gene (3’end) against other bacterial xylanase genes 61
25. PCR amplification of downstream walking using ST-PCR approach 63
26. Nucleotide and deduced amino acid sequences of 3’ end of xylanase gene obtained from ST-PCR 64
27. Amino acid sequence alignment of Bor Khlueng hot spring xylanase gene (3’end) against other bacterial xylanase genes 65
28. PCR amplification of partial BK44 gene 67
29. Alignment of the deduced amino acid sequence of BK44 against bacterial α-amylase family 68
30. Nucleotide and deduced amino acid sequences of BK44 from the sediments of Bor Khlueng hot spring 69
31. PCR amplification to obtain upstream region of BK44 by genome walking 70
32. Nucleotide and deduced amino acid sequences of 5’ end of BK44 71
33. PCR amplification to obtain downstream region of BK44 by genome walking 73
34. Nucleotide and deduced amino acid sequences of 3’ end of BK44 74
35. PCR amplification of the full length BK44 75
36. Nucleotide and deduced amino acid sequences of full-length BK44 from the sediments of Bor Khlueng hot spring 76
37. Amino acid sequence alignment of the full-length BK44 78
38. 3D structure prediction of BK44 79
39. Restriction enzyme analysis of pPIC-BK44 81

LIST OF FIGURES (Continued)

Figure **Page**

40. PCR analysis of *P. pastoris* integrants 82
41. SDS-PAGE analysis of secreted proteins from *P. pastoris* integrants 84
42. SDS-PAGE analysis of intracellular protein from *P. pastoris* 85
43. RT-PCR products of the full-length BK44 from *P. pastoris* 86
44. Restriction enzyme analysis of pET32-BK44 88
45. SDS-PAGE analysis of BK44 expression in *E. coli* at 30 ºC 89
46. SDS-PAGE analysis of BK44 in *E. coli* at 25 ºC at 0-4 h 90
47. Restriction enzyme digestion analysis of pET43.1-BK44 92
48. SDS-PAGE analysis of BK44 soluble and pellet fractions produced from *E. coli* harboring pET43.1-BK44 93
49. Pullulanase activity by AZCL-pullulan 94
50. mRNA secondary structure of BK44 97

LIST OF ABBREVIATIONS

A_{260}	=	absorbance at 260 nm
A_{280}	=	absorbance at 280 nm
AOX	=	alcohol oxidase gene
APS	=	ammonium persulfate
ATP	=	adenosine triphosphate
BLAST	=	Basic Local Alignment Search Tool
BMGY	=	buffered minimal glycerol complex medium
BMMY	=	buffered minimal methanol complex medium
bp	=	base pair(s)
CTAB	=	cetyl trimethyl ammonium bromide
°C	=	degree Celcius
cDNA	=	complementary DNA
dATP	=	deoxyadenosine-5'-triphosphate
dCTP	=	deoxycytosine-5'-triphosphate
dGTP	=	deoxyguanosine-5'-triphosphate
dTTP	=	deoxythymidine-5'-triphosphate
dNTP	=	dATP, dCTP, dGTP, dTTP
DMSO	=	dimethyl sulfoxide
DNA	=	deoxyribonucleic acid
DTT	=	dithiothreitol
EtBr	=	ethidium bromide
E. coli	=	*Escherichia coli*
EDTA	=	ethylenediamine tetraacetic acid
et al.	=	and others
g	=	gram
xg	=	centrifugal force
h	=	hour (s)

LIST OF ABBREVIATIONS (Continued)

IPTG	=	isopropyl-β-D-thiogalactopyranoside
kb	=	kilobase (s)
kDa	=	kilodalton (s)
LB	=	Luria-Bertani (medium)
M	=	molar
mM	=	millimolar
mg	=	milligram
min	=	minute (s)
ml	=	milliliter
µg	=	microgram
µl	=	microliter
mRNA	=	messenger RNA
ng	=	nanogram
nt	=	nucleotide (s)
OD	=	optical density
PAGE	=	polyacrylamide gel electrophoresis
PCR	=	polymerase chain reaction
pmol	=	picomole
P. pastoris	=	*Pichia pastoris*
RNA	=	ribonucleic acid
RNase A	=	ribonuclease A
RT-PCR	=	reverse transcription-polymerase chain reaction
SDS	=	sodium dodecyl sulfate
sec	=	second (s)
ST-PCR	=	semi-random two step PCR
T_m	=	melting temperature
TBE	=	Tris-borate EDTA
TEMED	=	N,N,N',N'-tetramethyl-ethylenediamine

LIST OF ABBREVIATIONS (Continued)

Tris-HCl	=	tris-(hydroxymethyl)-aminoethane hydrochloric acid
Trx	=	thioredoxin
UV	=	ultraviolet
V	=	volt
w/v	=	weight/volume
v/v	=	volume/volume
X-gal	=	5-bromo-4-chloro-3-indodyl-β-D-galactoside
YEPD	=	yeast extract peptone dextrose

CHAPTER I
INTRODUCTION

Enzyme industry is the result of rapid development of biotechnology i.e. recombinant gene technology and protein engineering. With better knowledge and purification of enzymes, a number of applications have increased exponentially (1). Nowadays, the use for industrial enzymes has now extended to almost all industries handling organic compounds.

The glycosyl hydrolases consisted of a large family of enzymes (xylanases, amylases, pullulanases) that are of great significance in industrial processes. Various carbohydrolases including α-amylases family and xylanases were used in various industries such as starch, textile, pulp and paper, detergent and baking (Table 1).

Table 1. Enzyme applications in industrial processes (adapted from: Kirk, Borchert and Fuglsang) (2)

Industry	Enzyme class	Application
Starch and fuel	Amylase	Starch liquefaction and saccharification
	Amyloglucosidase	Saccharification
	Pullulanase	Saccharification
	Glucose isomerase	Glucose to fructose conversion
	Xylanase	Viscosity reduction (fuel and starch)
	Protease	Protease (yeast nutrition – fuel)
Baking	Amylase	Bread softness and volume, flour adjustment
	Xylanase	Dough conditioning
	Lipase	Dough stability and conditioning
	Protease	Biscuits, cookies
Pulp and paper	Lipase	Pitch control, contaminant control
	Protease	Biofilm removal
	Amylase	Starch-coating, de-inking, drainage improvement
	Xylanase	Bleaching
	Cellulase	De-inking, fiber modification

1.1 Xylanases

1.1.1 Structure of xylan

Plant cell walls have three major polymeric constituents: cellulose (insoluble fibers of β-1,4-glucan), hemicellulose (non-cellulolytic polysaccharides including glucans, mannans, and xylans), and lignin (a complex polyphenolic structure). Xylan is a complex polysaccharide comprising a backbone of xylose residues linked by β-1,4-glycosidic bonds (Figure 1). The main chain of xylan is composed of β-xylopyranose residues. Most xylans occur as heteropolysaccharides, containing different substituent groups which are acetyl, arabinosyl and glucuronosyl residues (3).

1.1.1.1 Hemicellulose from hardwood

The xylan of hardwood is *O*-acetyl-4-*O*-methylglucuronoxylan. This polysaccharide consists of at least 70 β-xylopyranose residues (average degree of polymerization [DP] between 150 and 200), linked by β-1,4-glycosidic bonds (4). Every tenth xylose residue carries a 4-*O*-methyglucuronic acid attached to the second position of xylose (5). Hardwood xylans are highly acetylated (e.g., birchwood xylan contains more than 1 mol of acetic acid per 2 mol of xylose (6)). Acetylation is more frequent at the C-3 than at the C-2 position, however, acetylation at both positions have also been reported (7, 8). The presence of these acetyl groups is responsible for the partial solubility of xylan (3). These acetyl groups are readily removed when xylan is subjected to alkali extraction (6).

1.1.1.2 Hemicellulose from softwood

The xylan from softwood is composed of arabino-4-*O*-methyglucuroxylan. This xylan has higher 4-*O*-methylglucuronic acid content than that of hardwood xylan. The 4-*O*-methylglucuronic acid residues are attached to the C-2 position. Softwood xylans are not acetylated. Instead of an acetyl group, they contain α-L-arabinofuranose units linked by α-1,3-glycosidic bonds at the C-3 position of the xylose (9). The arabinosyl substituents occur almost 12% of the xylosyl residues (10).

1.1.2 Xylanolytic enzymes

Xylanolytic enzymes are glycosyl hydrolases which degrade xylan. Xylanases are widespread in nature. They have been reported to be present in marine and terrestrial bacteria, rumen and ruminant bacteria, fungi, marine algae, protozoa, and insects. Xylanases are usually composed of repertoire of hydrolytic enzymes: β-1,4-

endoxylanase, β-xylosidase, α-L-arabinofuranosidase, α-glucosidase, acetyl xylan esterase, and phenolic acid (ferulic and ρ-coumaric acid) esterase (Figure 1) (11, 12).

(1) β-1,4-Endoxylanases (1,4-β-D-xylan xylohydrolase; EC 3.2.1.8) cleave the internal glycosidic linkages of the heteroxylan backbone, resulting in the production of xylooligosaccharides. As hydrolysis proceeds, these oligosaccharides will be further hydrolyzed to xylotriose, xylobiose, and xylose (10). These enzymes have been isolated from several fungi and bacteria such as *Aspergillus niger* (13), *Trichoderma koningii* (14) and *Dictyoglomus thermophilum* Rt46B.1 (15).

(2) β-D-Xylosidases (β-D-xyloside xylohydrolase; EC 3.2.1.37) are exoglycosidases that hydrolyze short xylooligosaccharides and xylobiose from the non-reducing end to liberate xylose. These enzymes have been reported in bacteria and fungi (16, 17).

(3) α-L-Arabinofuranosidases. There are two types of arabinases, the exo-acting α-L-arabinofuranosidase (EC 3.2.1.55), which is active against ρ-nitrophenyl- α-L-arabinofuranoside and on branched arabinans, and endo-1,5-α-L-arabinase (EC 3.2.1.99), which is active only toward linear arabinans. Endoarabinases have been reported in *Bacillus subtilis* (18), *Clostridium felsineum* (19) and various fungi (20, 21).

(4) α-D-Glucosidases (3.2.1.20) hydrolyze the α-1,2 linkages between glucuronic acid and xylose residues in glucuronoxylan. These enzymes have been isolated from *Aspergillus niger* and *Streptomyces flavogriseus* (22).

(5) Acetylxylan esterases (EC 3.1.1.6) remove the *O*-acetyl substituents at the C-2 and C-3 positions of xylose residues in acetylxylan. The production of acetylxylan esterase was found in fungi and bacteria (23).

(6) Ferulic acid esterases (EC 3.1.1.73) cleave the ester linkages between arabinose side chains and ferulic acids in xylan. Similarly, ρ-coumaric acid esterases (EC 3.1.1-) cleave the ester linkage between arabinose and ρ-coumaric acid. Few ferulic and ρ-coumaric acid esterases have been purified and characterized (24).

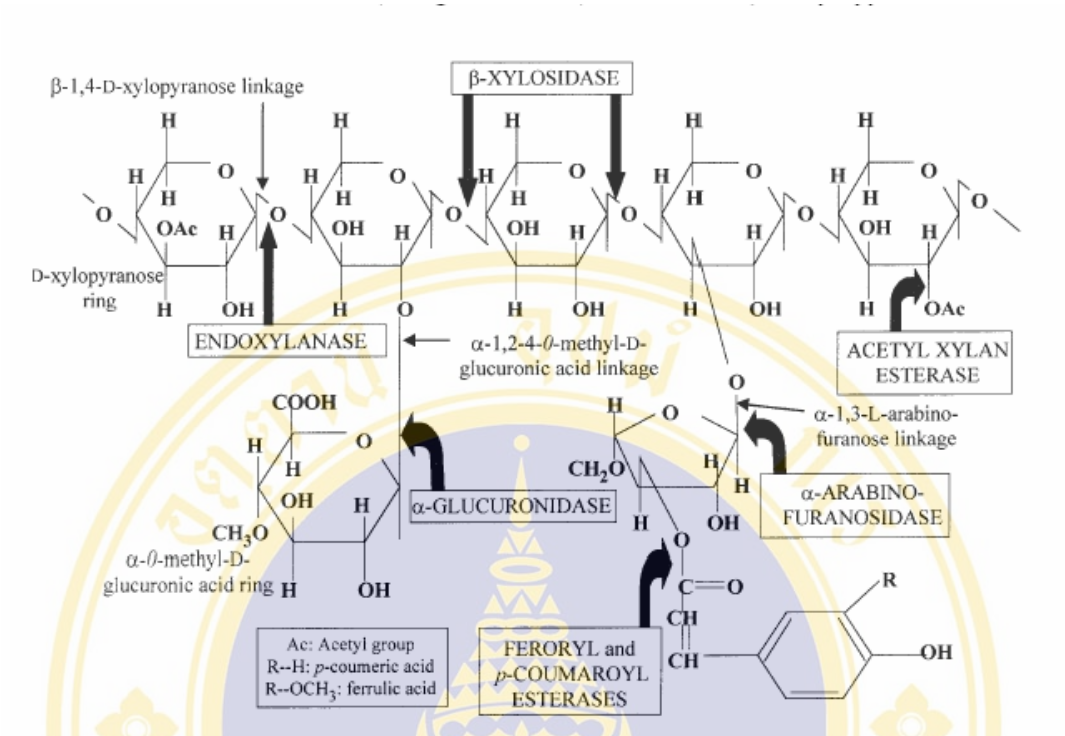
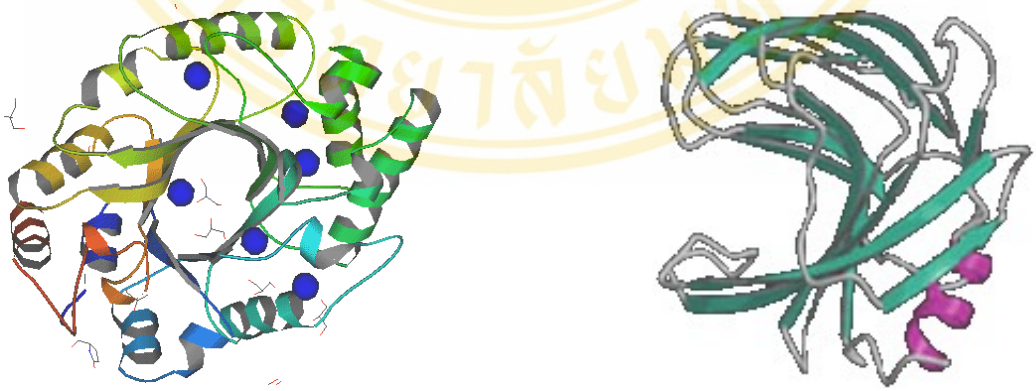


Figure 1. Xylanolytic enzyme system (taken from: Beg *et al*., 2001) (11)



(A) EX 10 **(B) EX 11**

Figure 2. Ribbon presentation of the main fold of the catalytic domains of EXs of family 10 and 11

1.1.3 Classification of xylanase family

Xylanases were first classified into two groups according to their molecular sizes and pH values consisting of (a) the low molecular weight, basic xylanases and (b) high molecular weight, acid xylanases (10). The low molecular weight basic xylanases are usually endo-1,4-β-xylanases which are specifically active only on xylan, whereas the other high molecular weight endoxylanases also have cellulase activity. These groups were classified into β-glycanase families G and F, respectively.

However, later, the classification is based on amino acid sequence similarities and hydrophobic cluster analysis (HCA) (25-28). HCA is a very sensitive tool to uncover distant relatedness between proteins for which standard alignment methods are unreliable (29). Thus, based on amino acid sequence similarities and hydrophobic cluster analysis, endo-β-1,4-xylanases (EXs, EC 3.2.1.8) can be divided into two major families: Acidic high molecular mass EXs (> 30 kDa) were assigned as glycanase family 10 (formerly family F), and the basic low molecular mass EXs, were assigned as glycanase family 11 (formerly family G) (30-32).

Recently, the tertiary structures of EXs have been established by crystallography. EXs of family 10 is a typical 8-fold α/β barrel $(\alpha/\beta)_8$ (33-35) (Figure 2A). In contrast, EXs of family 11 appear to be smaller and well-packed molecules, with mainly β-sheets (36-38) (Figure 2B).

1.1.4 Application of xylanases

Xylanases isolated from microorganisms have attracted a great deal of attention in the last decade, particularly because of their biotechnological potential in various industrial processes (39). Xylanases have shown an immense potential for increasing the production of several useful products in the most economical way. The main possibilities are the production of single cell proteins (SCPs), enzymes, liquid or gaseous fuel, solvents and sugar syrups, which can be used as such or as feed stocks for other microbiological processes (40).

Currently, the most promising application of xylanase is the pre-bleaching of kraft pulps (41). Enzyme application improves pulp fibrillation and water retention, reduction of beating times in virgin pulps, restoration of bonding and increased freeness in recycled fibers, and selective removal of xylans from dissolving pulps.

Xylanases are also useful in yielding cellulose from dissolving pulps from rayon production and biobleaching of wood pulps (42, 43).

In the bread and bakery industries, the efficiency of xylanases in improving the quality of bread has been shown with an increase in specific bread volume. This is further enhanced when amylase is used in combination with xylanase (44).

In the waste industry, xylan is present in large amount in wastes from agricultural and food industries. Hence, xylanases are used for conversion of xylan to xylose in waste water. The development of an efficient process of enzymatic hydrolysis offers new prospects for treating hemicellulosic wastes (45, 46).

In the juice industry, xylanases are used concurrently with cellulase and pectinase for clarifying must and juices, and for liquefying fruits and vegetables (45).

In the fuel industry, xylanase acts in synergism with several other enzymes, such as mannanase, ligninase, xylosidase, glucanase, glucosidase, etc., and can be used for the generation of biological fuel, such as ethanol, from lignocellulosic biomass (47).

In the animal-feed industry, depression in weight gain and feed conversion efficiency in rye-fed broiler chicks has been associated with intestinal viscosity. Incorporation of xylanase into a rye-based diet of broiler chicken results in reduced intestinal viscosity, thus improving both the weight gain of chicks and their feed conversion efficiency (48).

1.2 Family 13 glycosyl hydrolase (α-amylase family)

Starch is one of the most abundant polysaccharides produced by plants and is composed of amylose (15%-20%) and amylopectin (75%-85%). Amylose is a linear molecule consisting of 1,4-linked α-D-glucopyranose residues. Amylopectin is a branched polymer and consists of α-1,6 glycosidic linkages in addition to the α-1,4 glycosidic bonds. Therefore, the starch industry is one of the largest using of starch degrading enzymes. These include α-amylases, glucoamylases, pullulanases, and α-glucosidases (Figure 3).

1.2.1 Classification of family 13 glycosyl hydrolase

A classification system for glycosyl hydrolases based on sequence similarities has led to 85 different families. Most of the starch hydrolyzing enzymes belong to family 13 glycosyl hydrolases or α-amylase family and share common characteristic properties based on amino acid sequence homology according to the classification of Henrissat (25).

Starch degrading enzymes can be classified into four groups, (1) endoamylases, (2) exoamylases, (3) debranching enzymes and (4) transferases.

Endoamylases cleave α,1-4 glycosidic bonds in a random manner present in the inner part (endo-) of amylose or amylopectin chain (starch polymer). α-Amylase (EC 3.2.1.1) is the most well-known enzyme in this group. It is found in a wide variety of microorganisms (49). The end products of α-amylase are linear and branched oligosaccharides of various lengths.

Exoamylases, either cleave α,1-4 glycosidic bonds such as β-amylase (EC 3.2.1.2) or cleave both α,1-4 and α,1-6 glycosidic bonds like amyloglucosidase or glucoamylase (EC 3.2.1.3) and α-glucosidase (EC 3.2.1.20). Exoamylases act on the external glucose residues of amylose or amylopectin and produce only glucose (glucoamylase and α-glucosidase), or maltose and β-limit dextrin (β-amylase). Glucoamylase and α-glucosidase differ in their substrate preference: α-glucosidase acts best on short maltooligosaccharides and liberates glucose while glucoamylase hydrolyzes long-chain polysaccharides. β-amylases and glucoamylases have also been found in a large variety of microorganisms (49). Other exo-acting amylolytic enzymes are cyclodextrin glycosyltransferase (EC 2.4.1.19), an enzyme that additionally has transglycosylation activity, maltogenic α-amylase (EC 3.2.1.133) (50), and

maltooligosaccharide-forming amylases such as the maltotetrahydrolase (EC 3.2.1.60) from *Pseudomonas stutzeri* (51) or the maltohexahydrolase (EC 3.2.1.98) from *Klebsiella pneumoniae* (52).

The third group of starch-converting enzymes are the debranching enzymes that hydrolyze α,1-6 glycosidic bonds: isoamylase (EC 3.2.1.68) and pullulanase (EC 3.2.1.41). Pullulanase hydrolyzes the α,1-6 glycosidic bond in pullulan and amylopectin, while isoamylase can only hydrolyze the α,1-6 bond in amylopectin. These enzymes exclusively degrade amylopectin, thus leaving long linear polysaccharides. There are also a number of pullulanase type enzymes that hydrolyze both α,1-4 and α,1-6 glycosidic bonds, referred to as amylopullulanase (EC 3.2.1.1/41). The main degradation products are maltose and maltotriose. A special enzyme belonging to this group of pullulanases is neopullulanase (EC 3.2.1.35), which can also perform transglycosylation with the formation of a new α,1-4 or α,1-6 glycosidic bond (53).

The fourth group of starch-converting enzymes are the transferases that cleave an α,1-4 glycosidic bond of the donor molecule and transfer part of the donor to a glycosidic acceptor with the formation of a new glycosidic bond. Enzymes such as amylomaltase (EC 2.4.1.25) and cyclodextrin glycosyltransferase (EC 2.4.1.19) form a new α,1-4 glycosidic bond while branching enzyme (EC 2.4.1.18) forms a new α,1-6 glycosidic bond. Amylomaltases are very similar to cyclodextrin glycosyltransferases with respect to the type of enzymatic reaction. The major difference is that amylomaltase performs a transglycosylation reaction resulting in a linear product while the reaction of cyclodextrin glycosyltransferase gives a cyclic product. Amylomaltases have been found in different microorganisms in which they are involved in the utilization of maltose or the degradation of glycogen (54). Glucan branching enzymes are involved in the synthesis of glycogen in many microorganisms. They are responsible for the formation of α,1-6 glycosidic bonds in the side chains of glycogen.

1.2.2 The common features of family 13 glycosyl hydrolase

The enzymes act on the α-glycosidic bonds and hydrolyse this bond to produce α-anomeric mono- or oligosaccharides. They also exhibit transglycosylation activity where α-1,4 or 1,6 glycosidic linkages are formed (55).

In general, enzymes in the α-amylase family contains three domains: A, B and C. Domain A is a $(\beta/\alpha)_8$ barrel structure containing the catalytic domain and four highly conserved regions in their primary sequence (Table 2) (55, 56). Domain B is a small component which protrudes at the third β-strand of the domain A. Domain C is located at the C-terminus of domain A/B which made up of β-strands and is thought to stabilize the catalytic domain by shielding hydrophobic residues of domain A from the solvent (57). It is suggested that domain C may help in substrate binding (58, 59) (Figure 4).

Enzymes in the α-amylase family contain Asp, Glu, and Asp residues corresponding to Asp-206, Glu-230 and Asp-297 of Taka-amylase A which was the first reported structure in the α-amylase family (60, 61) (Table 2). These three residues have been recognized as the catalytic sites in Taka-amylase A (62), *Bacillus subtilis* and *B. stearothermophilus* α-amylase (62, 63), as well as in neopullulanase (64), cyclomaltodextrinase (65), CGTases (66, 67), amylopullulanase (68) and branching enzyme (68, 69). The three essential catalytic residues are strictly conserved in both amino acid sequence and three-dimensional structure (70).

The binding of calcium is also a general common feature of the α-amylase family. Several members are known to bind to one or more calcium ions. Many enzymes have a highly conserved Ca^{2+} binding site located between domain A and B, which preserves the structural integrity of the active site (71). For example, in the *Aspergillus niger* α-amylase, there are two binding sites. The primary site is essential in maintaining proper folding around the active site and contains a tightly bound Ca^{2+} whereas the secondary site is located at the bottom of the substrates binding cleft and involves the catalytic residues (Asp and Glu) (72). The 3D structure of cyclodextrinase (CDase) from *Flavobacterium* sp. revealed that CDase contains two Ca^{2+} binding sites. One is called Ca-I site, a conserved calcium-binding site essential to stabilize the conformation of domain B that is involved in substrate binding. The removal of Ca-I was shown to promote proteolysis (73). The second calcium site of *Flavobacterium* sp. CDase is called Ca-II, which stabilizes a surface region that is far away from the active center (74).

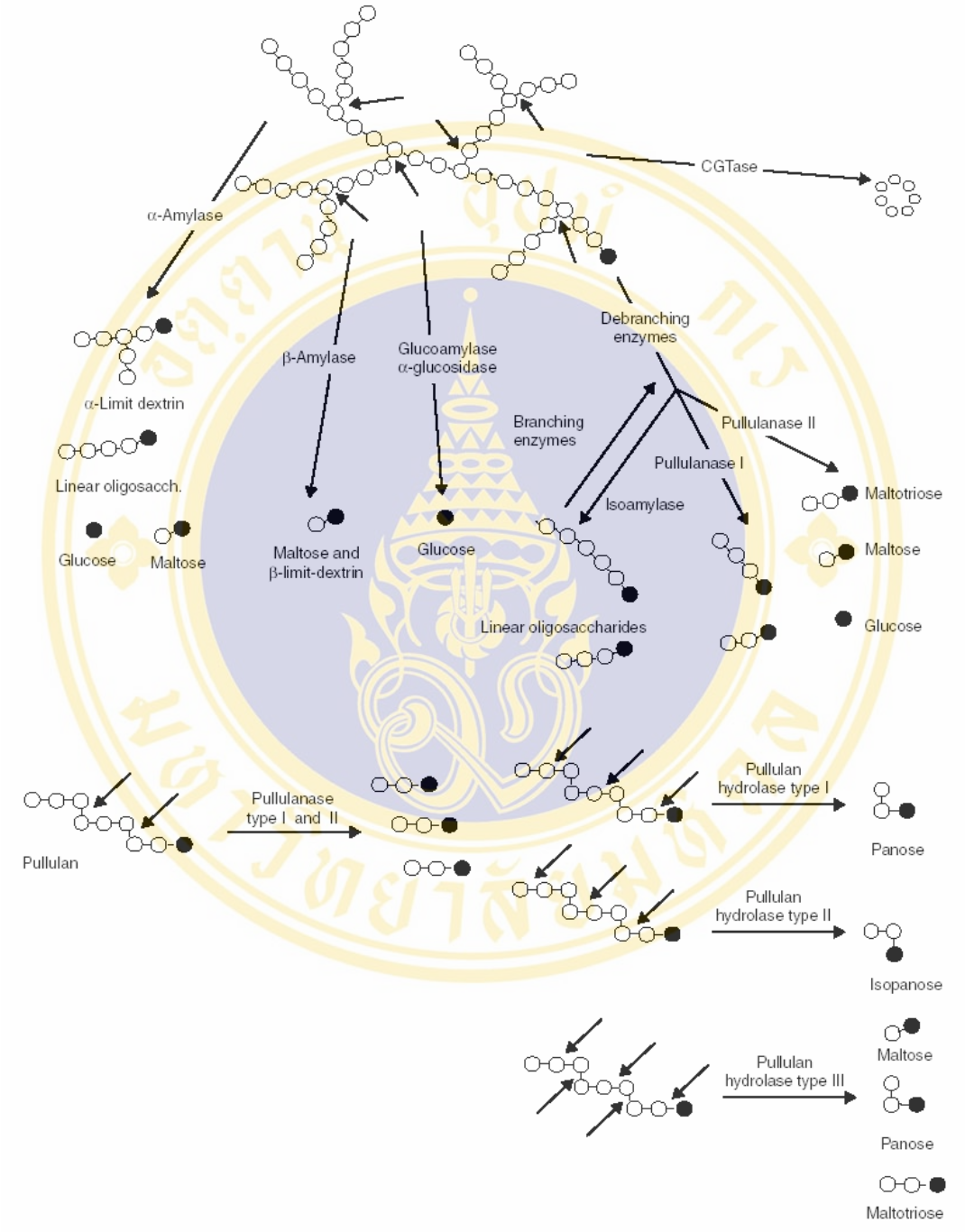


Figure 3. The enzymatic activities of the family 13 glycosyl hydrolase (taken from: Bertoldo and Antranikian, 2002) (75)

Table 2. The four highly conserved regions of family 13 glycosyl hydrolase. Amino acids in catalytic sites are underlined (adapted from: Kuriki and Imanaka, 1999) (56).

Enzyme	Origin	Region 1	Region 2	Region 3	Region 4
α-Amylase	*Aspergillus oryzae*	DVVANH	GLRIDTVKH	EVLD	FVENHD
CGTase	*Bacillus macerans*	DFAPNH	GIRFDAVKH	EWFL	FIDNHD
Pullulanase	*Klebsiella aerogenes*	DVVYNH	GFRFDLMGY	EGWD	YVSKHD
Isoamylase	*Pseudomonas amyloderamosa*	DVVYNH	GFRFDLASV	EPWA	FIDVHD
Branching enzyme	*Escherichia coli*	DWVPGH	ALRVDAVAS	EEST	LPISHD
Neopullulanase	*Bacillus stearothermophilus*	DAVFNH	GWRLDVANE	EIWH	LLGSHD
α-glucosidase	*Saccharomyces calsbergenesis*	DLVINH	GFRIDTAGL	EVAH	YIENHD
Cyclodextrianse	*Thermoanaerobacter ethanolicus*	DAVFNH	GWRLDVANE	EVWH	LIGSHD
Amylopullulanase	*Bacillus sp. KSM-1378*	DGVFNH	GWRLDVANE	EEID	LIGSHD

CGTase = cyclodextrin glucanotransferase (EC 2.4.1.19)

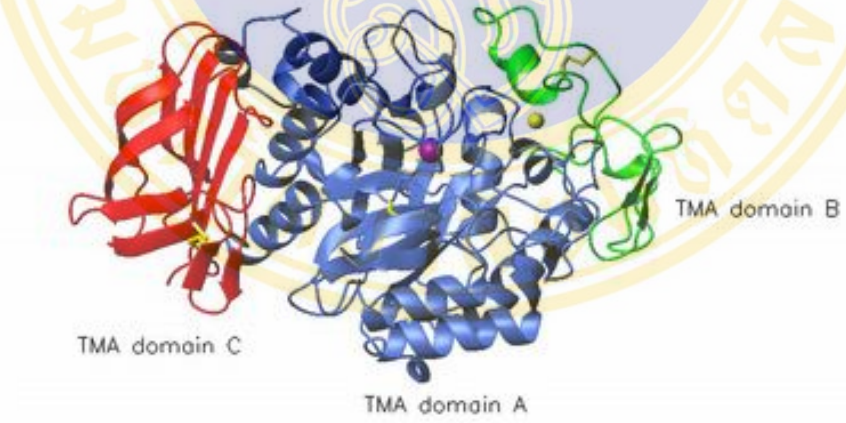


Figure 4. Ribbon representation of the conserved domains of α-amylase family (taken from: www.mol.biol.ethz.ch/groupsglockshuber_group/Research/picture)

1.2.3 Pullulan

Pullulan is a linear homopolysaccharide of glucose which consists of D-glucopyranosyl units alternating between one (1,6)-α-D and two (1,4)-α-D linkages or a linear polymer of α-(1,6) linked maltotriose (Figure 5). However, pullulan can also be considered as a polymer of panose or isopanose subunits. The pullulan molecule is pH neutral and its molecular weight ranges from 1,500 to 810,000 kDa. Pullulan and its derivatives are used in foods, pharmaceuticals, manufacturing, and electronics (76).

1.2.4 Application of pullulanases

Thermostable pullulanases are very useful in industrial applications. They are used in the starch industry, productions of maltose syrup (77), glucose and fructose (78). Furthermore, pullulanases are used to produce useful industrial substrates such as maltose, amylose, and glucose (79). These enzymes are used to produce low carbohydrate "light beer" by adding it together with fungal α-amylase or glucoamylase to the wheat during fermentation (80).

Glucoamylases have wide industrial applications. They are used in the degradation of dextrins to produce glucose, an intermediate step in the production of high-fructose corn syrups and ethanol (80).

Neopullulanases also have various industrial applications. Alkali-resistant neopullulanases can be used as effective additives in dishwashing and laundry detergents at alkaline condition (81). Moreover, panose which is a degradation product of pullulan, might be used as an anti-cariogenic sweetener in foods, because it is mildly sweet, non-fermentable by oral bacteria, and inhibits the synthesis of insoluble glucan from sucrose (64). Therefore, panose might be useful for preventing teeth caries (82).

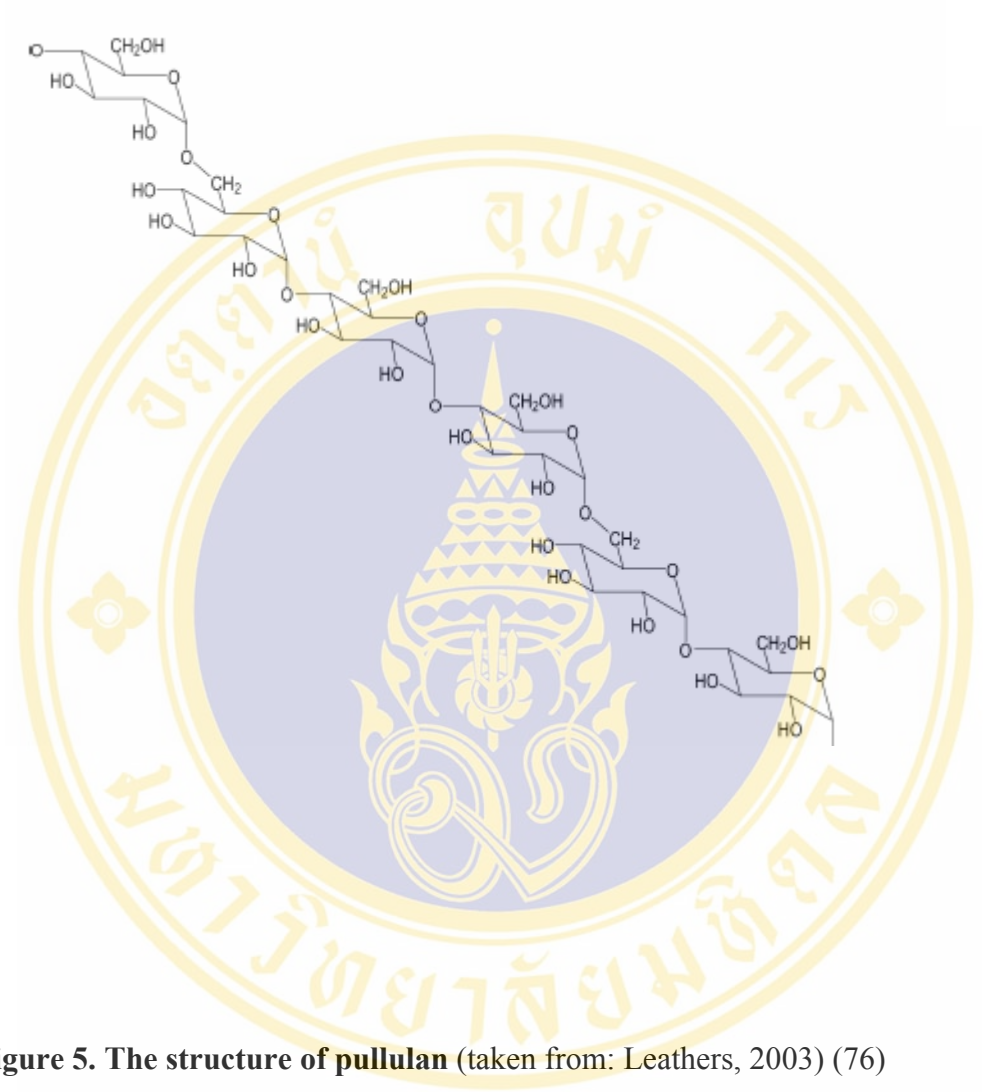


Figure 5. The structure of pullulan (taken from: Leathers, 2003) (76)

1.2.5 Pullulan degrading enzymes

Pullulan degrading enzymes are classified into four types that belong to family 13 glycosyl hydrolase (25, 83) according to the catalytic reaction domain and substrate specificity (56, 84) (Figure 6).

(1) Pullulanases (EC 3.2.1.41), also called true pullulanases (dextrinases, debranching enzymes and amylopectin 6-glucanohydrolases) hydrolyze α-(1,6)-glucosidic linkages of pullulan to produce maltotriose (84, 85).

(2) Glucoamylases (EC 3.4.1.3) are exoglucanases, which hydrolyze pullulan from non-reducing ends to produce glucose. These enzymes have been isolated from thermophilic anaerobic *Clostridium sp*. G0005 (86) and *Thermoanaerobacterium thermosaccharoliticum* (87).

(3) Isopullulanases (EC 3.2.1.57) hydrolyze α-(1,4)-glucosidic linkages of pullulan to produce isopanose (6-0-α-maltosyl-glucose). These enzymes have been isolated from fungi (88).

(4) Neopullulanases (EC 3.2.1.135), a new type of pullulan-hydrolyzing enzyme was first reported from *Bacillus stearothermophilus* TRS40 (82). The enzyme mainly hydrolyzes α-(1,4)-glucosidic linkages of pullulan to produce panose (6-α-D-glucosylmaltose). This type of enzyme includes cyclomaltodextrinase (EC 3.2.1.54) and maltogenic amylase (EC 3.2.1.133). All have the ability to hydrolyze broad range of substrates which are pullulan, cyclodextrins and starch. These three enzymes were named differently according to their substrate preferences. However, due to the similarity of three-dimensional structure of these enzymes, it was proposed that they should be renamed as the same enzyme (89). The details of this group will be given in the next section.

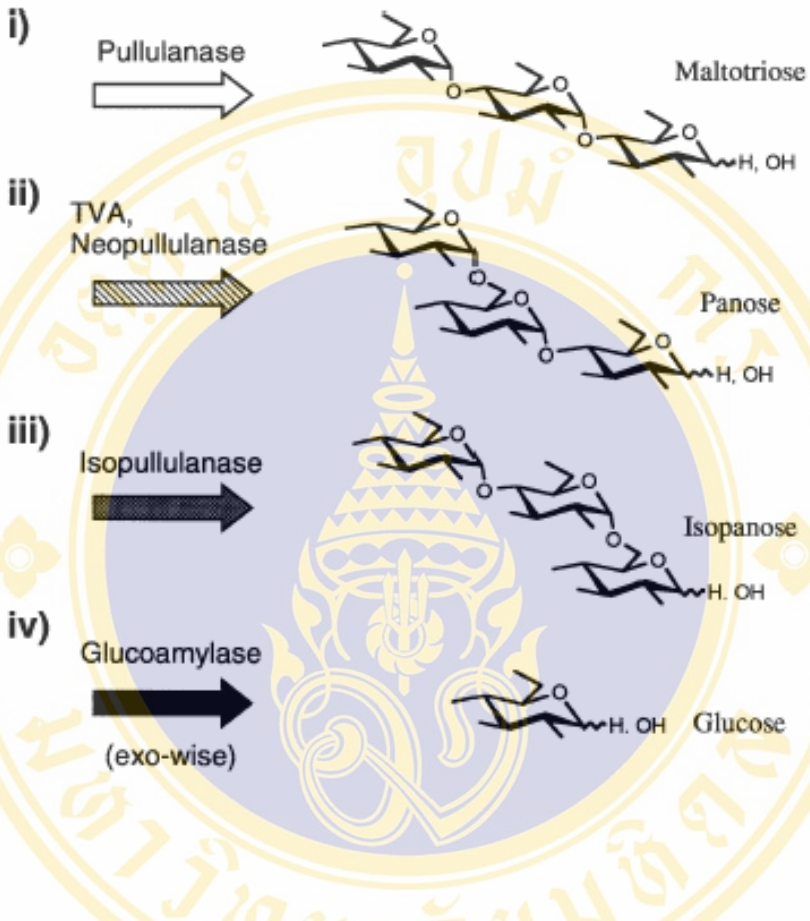


Figure 6. Action patterns of pullulan-hydrolyzing enzymes on pullulan (taken from: Aoki, Yopi and Sakano 1997) (88)

1.2.6 CD-/pullulan hydrolyzing enzymes (neopullulanase subfamily)

The CD-/pullulan hydrolyzing enzymes are composed of cyclomaltodextrinase (CDases), maltogenic amylase (MAases), and neopullulanase (NPases) which hydrolyze cyclodextrins, pullulan and starch (90). These enzymes share similarity in substrate specificity and amino acid identity (40-60%) (89). Enzymes in this group show distinct substrate specificity from other family 13 members (91). Common characteristic found among CD-hydrolyzing enzymes is their additional N-terminal domain, which is the unique addition of approximately 130 residues at the N-terminus. The role of the N-domain has been extensively investigated in neopullulanase of *Thermoactinomyces vulgaris* R-47 (TVA I and TVA II) (57). It was found that domain N is important for the hydrolysis of CDs, and is essential for the activities for all substrates (maltooligosaccharides, pullulan and starch) (57). The unique N-terminal domain of maltogenic amylase from *Thermus* sp. (ThMA) plays a major role in the modification of the common active site structure through dimer formation (90). In CDase of *Bacillus* sp. I-5 (BaCD), domain N participates in dimer formation. In these dimers, the N-terminal domain of one subunit contacts the active site of other subunit and participates in CD binding. Besides the common dimerization, BaCD forms as a hexamer of these dimers (89).

1.2.6.1 Cyclomaltodextrinase (EC 3.2.1.54; CDase)

CDase is defined as an enzyme that hydrolyzes cyclodextrins (CDs) to form linear malto-oligosaccharides. These substrates, α-, β-, and γ-cyclodextrins (CDs) are composed of six to eight D-glucopyranosyl residues linked by α-1,4 glycosidic bonds, respectively. CDase was designated as such due to the fact that the enzymes tend to hydrolyze CDs much faster than other substrates which are pullulan and starch (92). CDases from various sources such as *Bacillus macerans* (93), *B. coagulans* (94), *Thermoanaerobacter ethanolicus* 39E (65), *Bacillus* sp. I-5 (89), *Flavobacterium* sp. (74) and *B. sphaericus* ATCC7055 (95) have been isolated and characterized.

1.2.6.2 Maltogenic amylase (EC 3.2.1.133; MAase)

Maltogenic amylases exhibit unique characteristics that are different from other α-amylase family (96-99) in that they are intracellular enzymes and hydrolyze CDs and starch mainly to maltose thereby named as “maltogenic amylases” (90). They exhibit a dual activity of both α-1,4 and α-1,6 glycosidic bonds and α-1,4, to α-1,3, α-

1,4 or α-1,6-transglycosylation (100). MAases hydrolyze CDs and starch to maltose and pullulan to panose. The enzyme prefers cyclodextrin (CDs) to starch or pullulan. MAases cleave α-1,4 glycosidic bond much more efficiently than α-1,6 glycosidic bond. Several maltogenic amylases have been cloned from bacteria including *Bacillus licheniformis* (96), *Bacillus stearothermophilus* (100) and *Thermus* strain IM6501 (99).

1.2.6.3 Neopullulanase (EC 3.2.1.135; NPase)

Neopullulanases hydrolyze pullulan to panose (84), CDs and amylose to maltose (73). They hydrolyze not only α-1,4-glycosidic linkages but also α-1,6-glycosidic linkages of several branched oligosaccharides (101). Only one active center of the neopullulanase has been reported to participate in the hydrolysis of both glycosidic linkages (64). Furthermore, neopullulanases catalyze all four types of reactions: hydrolysis of α-1,4 and α-1,6 glycosidic linkages, and transglycosylation to form α-1,4 and α-1,6 glycosidic linkages (64, 83). Isolation and characterization of neopullulanases have been reported from *Bacillus stearothermophilus* (83), *Alicyclobacillus acidocaldrius* (102), *Thermoactinomyces vulgaris* R-47 (57), *Bacteroids thetaiotaomicron* 95-1 (103) and *B. polymyxa* CECT155 (104).

1.2.6.4 3D structure of neopullulanase subfamily

Like any other enzymes belonging to glycosyl hydrolase family 13, enzymes in neopullulanase subfamily contain the typical domains A, B and C (Figure 7). As mentioned above, one characteristic of the CD-degrading enzyme is their additional N-terminal domain (57). The crystal structures of two CD-degrading enzymes, maltogenic amylase from a *Thermus* strain (ThMA) (90) and neopullulanase from *T. vulgaris* R-47 (TVA II) (105) revealed that this unique N-terminal domain (composed of β-strands) of approximately 130 residues modifies the active site to achieve distinct specificity properties which contribute the preference for CDs, whereas these enzymes hydrolyze starch or pullulan inefficiently. 3D structure studies revealed that CDase from *Bacillus* I-5, ThMA, and NPase from *T. vulgaris* R-47 (TVA II) are very similar in terms of their structures and catalytic property and that they can be rather easily distinguished from other amylases on the basis of the N-terminal domain and the well conserved residues, Try and Phe (89). Therefore, from previous studies, it has been

suggested that cyclodextrinase, maltogenic amylase and neopullulanase should be renamed into a single enzyme (89).

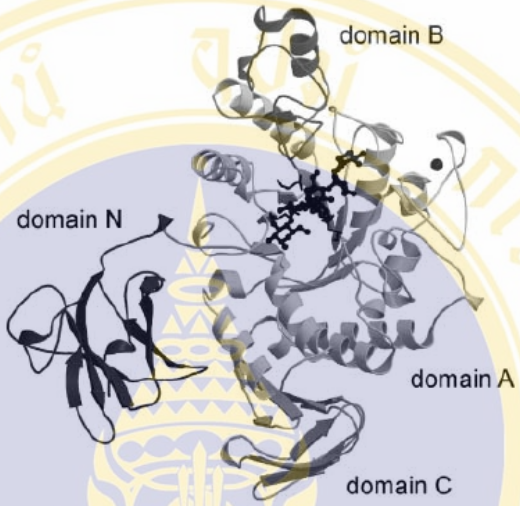


Figure 7. Ribbon representative of the conserved domains of *Thermoactinomyces vulgaris* R-47 neopullulanase (TVA II) (taken from: Mizuno *et al*., 2004) (106)

1.3 The role of thermostable enzymes.

As mention earlier, enzymes are useful in various industrial applications. One characteristic that is required in industrial processes is thermostability. Several substrates can only be degraded or converted by enzymes under extreme conditions. The increase in temperature has a significant influence on the solubility of organic compounds (107). The most widely used thermostable enzymes are amylases and pullulanases in the starch industry. Starch is converted to fructose syrup or modified during the baking process. Thermostable enzymes are required to degrade starch because the substrate must be heated to 60 ℃ to gelanize before it is vulnerable to enzymatic digestion (108). In pulp and paper industry, the wood used for the production of the pulp is treated at high temperature and basic pH, which required thermostable enzymes, such as xylanases (109). Therefore, thermostable enzymes have gained wide industrial and biotechnological interests because they are suitable for harsh industrial processes (110). In addition, high temperature helps eliminate the risk of bacterial or viral contamination. For these reasons, the number of studies on extremophiles from extreme environments have grown rapidly in the past few years (109, 111, 112).

1.4 Direct DNA isolation from environment

Soil is a major source of organic carbons on earth and an important habitat for microorganisms (113). Moreover, soil microorganisms are the main sources for natural products e.g. antibiotics, anticancer drugs, antifungal compounds, and enzymes (114). A great diversity is found in soil microorganisms as revealed by experiments using 16S rRNA sequence analysis (115, 116), *in situ* hybridization (117, 118) or dot blot hybridization (117). Therefore, soil microbial diversity is rich and remains widely unexplored resource for novel industrial enzymes and bioactive compounds (119).

For many years, the diversity of soil microorganisms has been established based on culture-dependent approach. However, this approach has a limitation in that approximately only 1% of soil microorganisms can be cultured using the standard methods. Thus, the remaining 99% of soil microorganisms have yet to be identified (117). Therefore, culture-independent approaches have been developed to overcome such limitation (120, 121).

Thailand, being a tropical country, has been known for the great biodiversity of microorganisms. One of the potential bioresources for obtaining novel enzymes in Thailand is hot spring. Recently, the biodiversity of one of hot springs, Bor Khlueng, has been established (122). Using 16S ribosomal RNA (rRNA) gene-sequence analysis, approximately 80% of the prokaryotic sequences found in Bor Khlueng hot spring were unknown. This suggested the possibility of identifying new potential enzymes from this environment.

1.5 Culture-independent approaches (activity-based screening and sequence-based screening approaches).

To obtain the enzymes or natural products directly from soil without cultivation, two approaches can be used which are activity-based screening and sequence-based screenings.

Activity-based screening has the potential to detect entirely novel genes encoding new types and classes of enzymes or to identify new bioactive compounds. To isolate the novel enzymes from environmental DNA, total soil microbial genomes (metagenome) are extracted to construct metagenomic library and enzymatic reactions with specific substrates can then be screened. Many enzymes of industrial importance have been discovered using this strategy (107, 123). The advantage of this method is the potential of accessing totally unknown sequence and selection for full-length genes with functional gene products. The limitations, however, are that this method requires the expression of the functional protein of interests in the appropriate host cell (e.g. *Escherichia coli*). Therefore, more than one type of hosts might be needed in order to obtain functional proteins. The large size of the library is also required. Furthermore, efficient and economical screening methods for the desired traits must be established to facilitate the screening of vast libraries (124).

Sequence-based screening approach required oligonucleotide primers to identify target genes directly by polymerase chain reaction (PCR). Therefore, the conserved amino acid sequences must be known to design degenerate primers to match unknown target genes followed by genome walking PCR to retrieve the full-length gene. One advantage of the sequence-based screening approach is that this method is not dependent on the heterologous expression of cloned genes. Using this approach, it is

possible to obtain target enzymes with low level of expression which could be undetected when activity-based screening was employed**.** Sequence-based screening approach was successfully used in discovering various enzymes such as xylanases (111, 112), amylases (125) and polyketide synthases (126, 127).

CHAPTER II
OBJECTIVES

The aims of this thesis are to clone and express the enzymes from glycosyl hydrolase family (xylanase and neopullulanase) from the sediments of Bor Khlueng hot spring.

CHAPTER III
MATERIALS

3.1 Bacterial strains

Escherichia coli, DH5α [*sup*E44 ΔlacU169 (ϕ80 *Lac*Z M15) *hsd*R17 *rec*A1 *end*A1 gyrA96 thi-1 *rel*A1] was used as a host for plasmid propagation.

Esherichia coli, Rosetta-gami(DE3)pLysS [(Δ*ara–leu7697* Δ*lacX74* Δ*phoAPvuII phoR araD139 ahpC galE galK rpsL* F'[*lac*+(*lacI*^{q})*pro*] *gor522* ::Tn*10* (Tc^R) *trxB*::kan (DE3) pLysSRARE6 (Cm^R)] was employed as a host cell for expression.

3.2 Yeast strain

Pichia pastoris KM71 (*arg4, his4, aox1::Arg4*) was used as a host cell for protein expression.

3.3 Plasmid vectors

pGEM-T® Easy vector (Promega) was used for all cloning steps. A physical map of pGEM-T® Easy is shown in Figure 8.

pPICZαA vector (Invitrogen) was used as an expression vector in *P. pastoris*. A physical map of pPICZαA vector is shown in Figure 9.

pET-32a(+) and pET-43.1a(+) vectors (Novagen) were used as expression vector in *E. coli*. Physical maps of both vectors are shown in Figure 10 and 11, respectively.

3.4 Sampling site

Soil sample was obtained from Bor Khlueng hot spring in Ratchaburi province, Thailand. The water temperature when the sample was taken was 55 °C with the pH of 6.6.

3.5 Synthetic oligonucleotides

All synthetic oligonucleotides (primers) used in this study (Table 3) were synthesized by Bioscience Unit, BIOTEC (Thailand) or PROLIGO Singapore Pty Ltd.

3.6 Culture media

3.6.1 Bacterial culture medium (for transformation containing pGEM®-T Easy vector)

E. coli strain DH5α was grown in LB [1% (w/v) peptone (Difco), 0.5% (w/v) NaCl, and 0.5% (w/v) yeast extract (Difco)]. The *E. coli* transformants were grown in LB containing 100 μg/ml ampicillin (Sigma, USA). For agar plates, 2% (w/v) of bacteriology agar was added. For X-gal/IPTG agar plates, 40 μl of a stock solution 5-bromo-4-chloro-3-indodyl-β-D-galactoside (X-gal; 20 mg/ml of X-gal in dimethylformamide) and 4 μl of a solution of isopropylthio-β-D-galactoside (IPTG; 200 mg/ml of IPTG in water) were added in agar media.

3.6.2 Bacterial culture medium (for transformants containing pPICZαA vector)

E. coli (strain DH5α) was cultured in low salt LB medium containing 1% (w/v) tryptone or peptone (Gibco BRL, USA), 0.5% (w/v) NaCl and 5% (w/v) yeast extract (Gibco BRL, USA), the pH of the medium was adjusted to 7.5 with NaOH. The *E. coli* transformants were grown in low salt LB medium containing 25 μg/ml Zeocin™ (Invitrogen, Canada) as a selectable marker.

3.6.3 Yeast culture medium

3.6.3.1 Culture and selective media

P. pastoris (strain KM71) was grown in a rich medium YEPD [2% (w/v) peptone, 2% (w/v) glucose (Sigma, USA) and 1% (w/v) yeast extract]. *P. pastoris* transformants were cultured under selective condition in YEPD containing 100 μg/ml Zeocin™.

3.6.3.2 Expression medium

Bufferd minimal glycerol complex medium (BMGY) contained 0.67% (w/v) yeast nitrogen base, 1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate, pH 6.0, 0.00004% biotin (w/v), and 1% (w/v) glycerol.

Buffered minimal methanol medium (BMMY) contained 0.67% (w/v) yeast nitrogen base, 1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate, pH 6.0, 0.00004% biotin, and 3% (w/v) methanol.

3.7 Chemicals

Zeocin™ was purchased from either Invitrogen, Canada or CAYLA, France. Other chemicals and solvents (analytical grade) were purchased from Gibco BRL, USA; Fluka, Switzerland; Sigma, USA or Merck, Germany.

3.8 Kits

TRI REAGENT®	Molecular Research Center, USA
pGEM-T® Easy vector	Promega, USA
QIAGEN gel extraction kit	QIAGEN, Germany
QIAprep Spin Miniprep Kit	QIAGEN, Germany
GenomeWalker™ Kit	CLONTECH, USA
Improp-II™ Reverse Transcription System	Promega, USA
RevertAidTM H Minus First Strand cDNA Synthesis Kit	Fermentas, Lithuania

3.9 Enzymes

All enzymes were purchased from New England Biolabs, USA; Boehringer Mannheim, Germany; Promega, USA; Gibco BRL, USA; Stratagene, USA; Sigma, USA; Fermentas, USA or Invitrogen, USA.

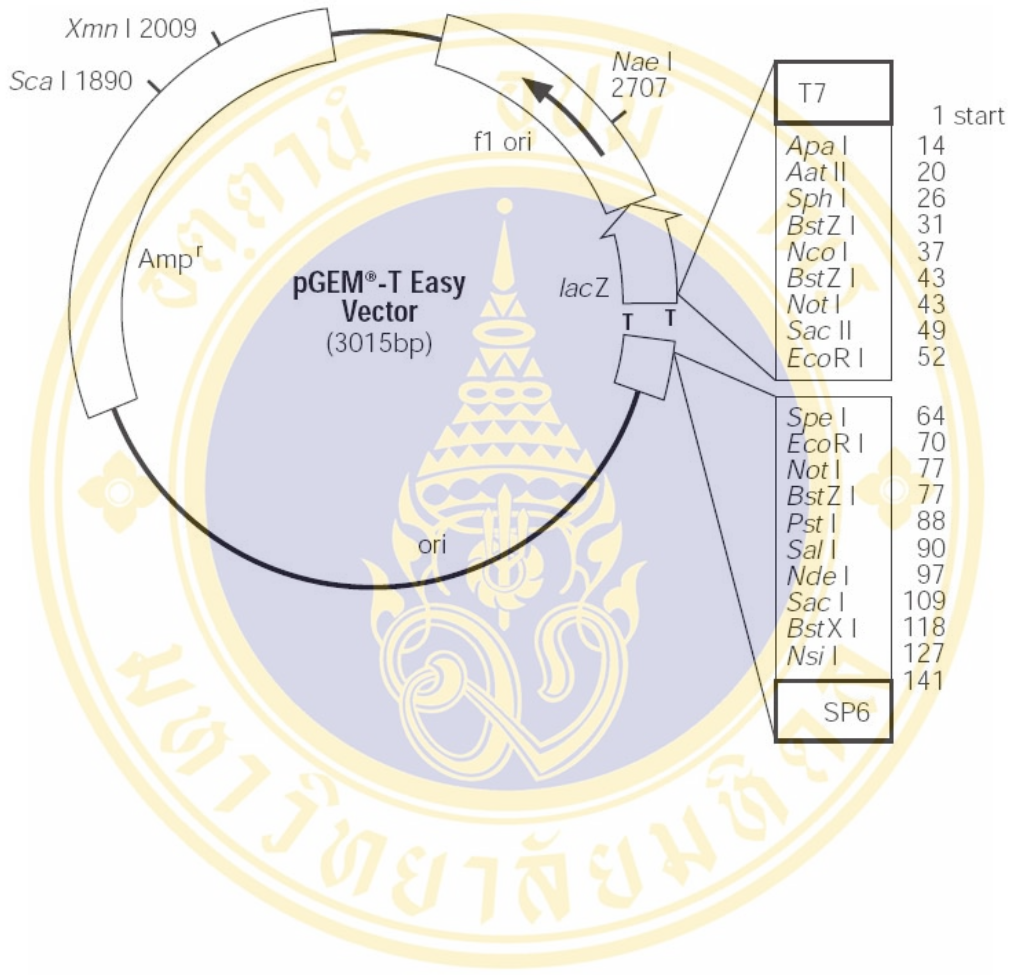


Figure 8. The physical map of pGEM®-T Easy vector (Promega)

The plasmid contains a 3' terminal thymidine added to both ends, which is compatible with a single deoxyadenosine overhang at 3' end of PCR product generated by certain thermostable polymerases.

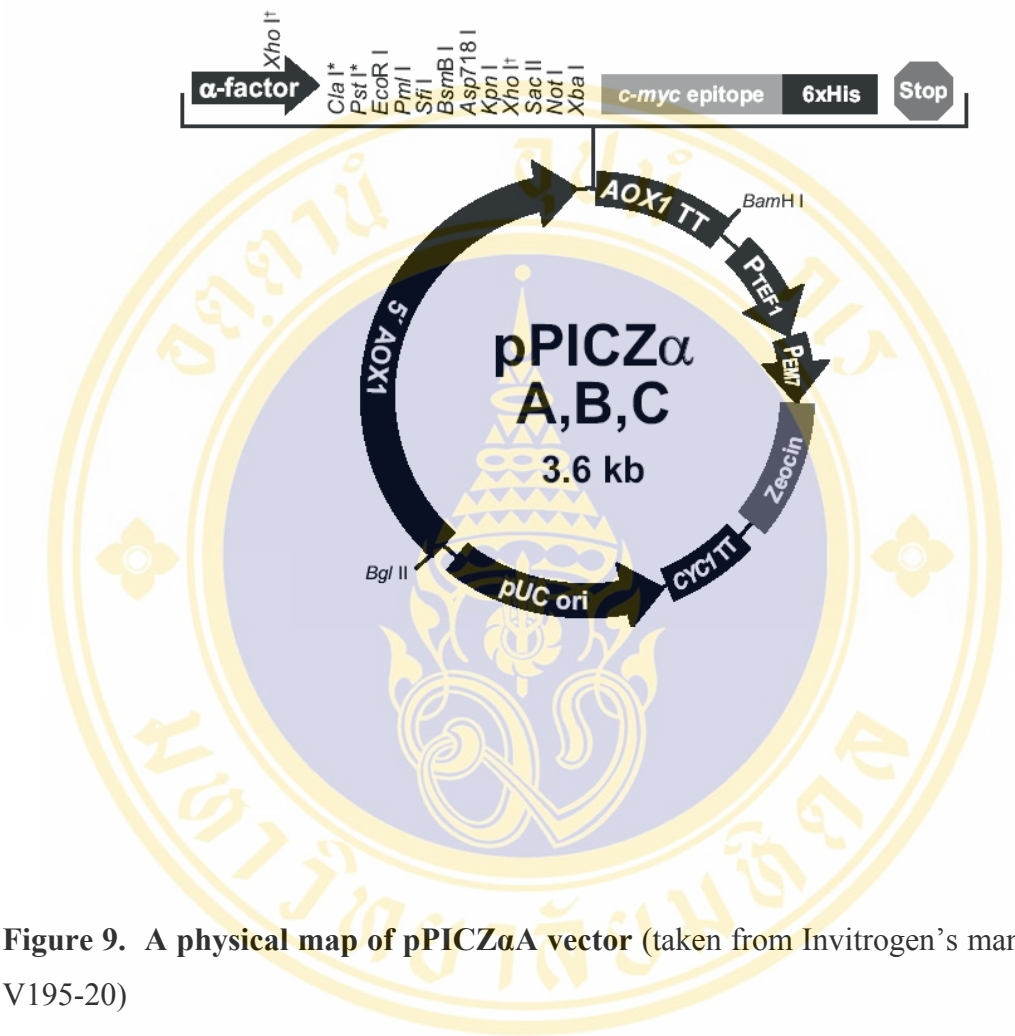


Figure 9. A physical map of pPICZαA vector (taken from Invitrogen's manual No. V195-20)

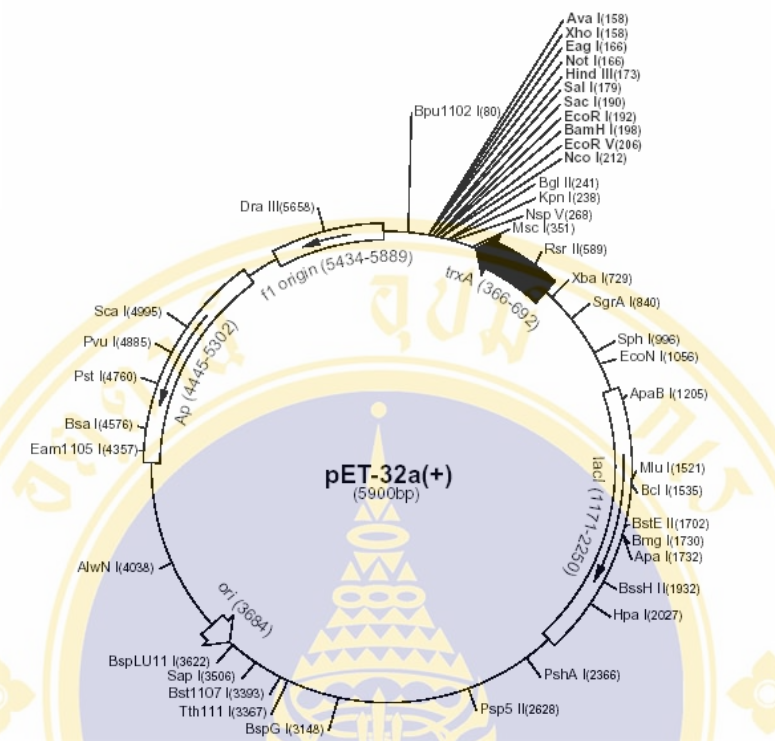


Figure 10. A physical map of pET-32a(+) vector (taken from Novagen's instruction manual TB 122)

pET-32a(+) sequence landmarks:

T7 promoter	764-780
T7 transcription start	763
Trx·Tag coding sequence	366-692
His·Tag coding sequence	327-344
S·Tag coding sequence	249-293
Multiple cloning sites (*Nco*I-*Xho*I)	158-217
His·Tag coding sequence	140-157
T7 terminator	26-72
*lac*I coding sequence	1171-2250
pBR322 origin	3684
bla coding sequence	4445-5302
f1 origin	5434-5889

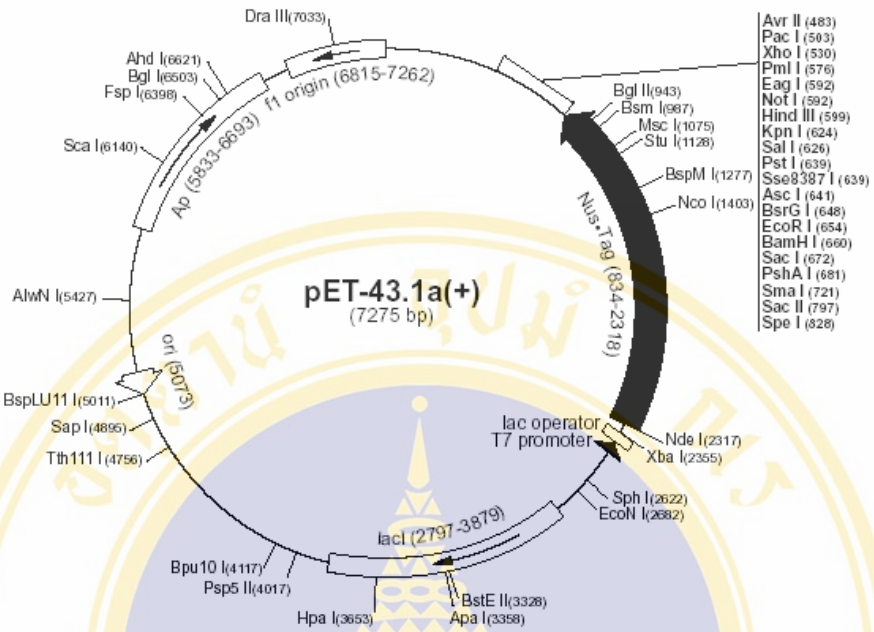


Figure 11. A physical map of pET-43.1a(+) vector (taken from Novagen's instruction manual TB 288)

pET-43.1a(+) sequence landmarks:

T7 promoter	2390-2406
T7 transcription start	2390
Nus·Tag™ coding sequence	834-2318
His·Tag® coding sequence	801-818
S·Tag™ coding sequence	747-791
Multiple cloning sites (*Sma*I-*Xho*I)	721-686
HSV·Tag® coding sequence	537-572
His·Tag® coding sequence	513-530
T7 terminator	27-73
*lac*I coding sequence	2797-3879
pBR322 origin	5073
bla (Ap) coding sequence	5833-6693
f1 origin	6815-7262

Table 3. Primers used in each experiment. For the degenerate oligonucleotides, the following abbreviations are used (I= Inosine; M= A, C; N= A, T, C, G; R= A, G; S= C, G; W= A, T; Y= C, T)

Experiments and Primer names	Sequence (5'-3')	Size (bp)	Tm (ºC)
bacterial 16S forward primer	AGAGTTTGATCCTGGCTCAG	20	60
bacterial 16S reverse primer	GGTACCTTGTTACGACTT	18	52
Adaptor Primer (AP1)	GTAATACGACTCACTATAGGGC	22	59
Nested Adaptor Primer (AP2)	ACTATAGGGCACGCGTGGT	19	71
ST 1 primer	GGCCACGCGTCGACTAGTAC(N_{10})GATAT	35	61
ST 2 primer	$(CUA)_4$ GGCCACGCGTCGACTAGTAC	32	61
Family 10 xylanase gene			
Partial xylanase gene			
XylF23	MGNGGICAYACNYTIGTITGGCA	23	55
XYNFR	T(AC)GTT(GT)AC(AC)AC(AG)TCCCA	17	44
3'-end DNA amplification			
XYL2GW_DW1	GATGCGGGAACACATTTTTAGTGTGG	26	61
XYL2GW_DW2	GGTACATGGGTGGGACGTGGTAAACT	26	62
GWC_DW1	ATTAGACGTGAGCGTGCTCCCGTTTC	26	64
GWC_DW2	GGGCAGATGTGGGCTTGAAGTTTGAAT	27	64
5'-end DNA amplification			
XYL2GW_UP1	CACACTAAAAATGTGTTCCCGCATCC	26	61
XYL2GW_UP2	GTCTTGAAATACCCAAGCCGGTGTCT	26	61
Family 13 glycosyl hydrolase			
Partial amplification			
amyI	YTNGAYGSIGTITTYAAYCA	20	54
amyIV	GTRTCRTGISWICCIARIARRTT	23	44
3'-end DNA amplification			
BK44GW_DW1	GAAGCCTACCTCGTGGGCGAAATCT	26	65
BK44GW_DW2	CGACGCCGTGACGAACTATCTCTTA	25	60
5'-end DNA amplification			
BK44GW_UP1	CGGGTTGTCTGTGTTGAACTTTGGC	25	62
BK44GW_UP2	TTGCCATCGTAAGGATAGAGGGGGA	25	63
full-length DNA amplification			
FLBK44_F	ATGAATAGTACTATGAACGAAAGGATAGCC	30	56
FLBK44_R	TTACACTACCTGAAGAACAACCCCACTTC	29	60
P. pastoris expression			
pICBK44-3F	GCCTCGAGAAAAGAATGAATAGTACTATGA	30	57
pICBK44-3R	GCTCTAGATTACACTACCTGAAGAACAAC	29	54
E. coli expression			
pETBK44_F	GCCCATGGCTATGAATAGTACTATGAACGA	30	62
pETBK44_R	GCGAATTCTTACACTACCTGAAGAACAAC	29	58
BAMpET_F	GCGGATCCATGAATAGTACTATGAACGA	28	60
pET43_1BK44R	GCGGTACCTTACACTACCTGAAGAACAAC	29	60

CHAPTER IV
METHODS

4.1 Genomic DNA extraction with SDS-based DNA extraction method (Zhou *et al*., 1996) (128)

5g of soil sample from Bor Khlueng hot spring was mixed with 13.5 ml of DNA extraction buffer (100 mM Tris-HCl [pH 8.0], 100 mM sodium EDTA [pH 8.0], 100 mM sodium phosphate [pH 8.0], 1.5 M NaCl, 1% (w/v) cetyltrimethylammonium bromide (CTAB)) and 100 µl of proteinase K (10 mg/ml) by horizontal shaking at 225 rpm for 30 min at 37 ºC. Next, 1.5 ml of 20% SDS was added, and the sample was incubated in a 65 ºC water bath for 2 h with gentle end-over-end inversions every 15 to 20 min. The supernatant was collected after centrifugation at 6,000 xg for 10 min at room temperature and transferred to 50-ml centrifuge tubes. The soil pellet was extracted two more times by adding 4.5 ml of the extraction buffer and 0.5 ml of 20% SDS, vortexing for 10 sec, incubating at 65 ºC for 10 min, and centrifuged as before. Supernatants from the three cycles of extraction were combined and mixed with an equal volume of chloroform-isoamyl alcohol (24:1, v/v). The aqueous phase was recovered by centrifugation and precipitated with 0.6 volume of isopropanol at room temperature for 1 h. The pellet of crude nucleic acids was obtained by centrifugation at 16,000 xg for 20 min at room temperature, washed with cold 70% ethanol, and resuspended in sterile distilled water. The DNA was analyzed on 0.8% agarose gel electrophoresis.

4.2 DNA concentration and purity determination

The quantity and quality of DNA were determined by using a spectrophotometer (Hitachi, Japan) at the absorbance of 260 nm and 280 nm. The DNA concentration was calculated by the formula:

$$\text{DNA concentration } (\mu g/\mu l) = \frac{A_{260} \times (\text{dilution factor}) \times 50}{1000}$$

The DNA purity was determined by the absorbance ratio of A_{260}/A_{280}. A ratio should be between 1.6-1.8 which represents a high purity of DNA.

4.3 DNA electrophoresis

The agarose gel was prepared by melting 0.8 % (w/v) of agarose gel in 1X TBE (Tris borate buffer, 89 mM boric acid, 89 mM Tris-HCl, 2 mM EDTA, pH 8.0). The gel was poured into an electrophoretic tray and allowed to set at room temperature for 20-30 min. 1X TBE was used as an electrophoretic buffer. The digested DNA or PCR product was mixed with 30% (v/v) of loading dye (25% glycerol, 60 mM EDTA and 0.25% bromphenol blue), loaded into the prepared gel and run at a constant voltage at 110 volts for 1 h. The gel was then stained in 2.5 g/ml of ethidium bromide solution for 5 min and destained in water for 15-20 min. The DNA was visualized under UV light [Gel Doc model 100 (Bio-Rad, USA) or BioDoc-It™ System (UVP, USA)] and photographed.

4.4 DNA purification by QIAquick gel extraction kit (QIAGEN kit)

A band containing crude extracted DNA was cut from the gel under a UV light. QIAGEN gel extraction kit was used to purify the PCR product. Specifically, three volumes of buffer QX1 were added to the excised gel containing the PCR product. The mixture was incubated at 50 ºC for 10 min or until the gel was completely dissolved. The color of the solution should be yellow to ensure that pH was not less than 7.5 which give high efficiency of DNA binding to QIAquick column. Then, one volume of isopropanol was added to the mixture to increase the yield of DNA fragments that was smaller than 500 bp or larger than 4 Kb. The mixture was, then, applied to the QIAquick column which was placed in a 2 ml collection tube and centrifuged at 14,000 xg for 1 min. After the flow through was discarded, 0.5 ml of buffer QX1 was added to QIAquick column to remove all traces of agarose. The column was then centrifuged as described previously. Next, 0.75 ml of buffer PE was added to the QIAquick column for washing. The column was left at room temperature for 2-5 min and centrifuged twice. The column was placed in a fresh 1.5 ml tube. A volume of 30 µl of elution buffer was added to the QIAquick column and the column was left at room temperature for 1 min to dissolve DNA. The tube was then

centrifuged as above. The step of DNA elution with the elution buffer was repeated to obtain the maximal amount of DNA. Then, 1-2 µl of concentrated DNA was analyzed by gel electrophoresis in order to determine its concentration.

4.5 Amplification of bacterial 16S rDNA

To determine if the purified DNA is of sufficient quality (i.e. no humic acid present) for downstream experiments, PCR, using universal bacterial 16S rDNA primers (Table 3), were used to amplify 16S rDNA from purified DNA. PCR amplification was performed in 50 µl reaction mixtures containing 2 µl of purified DNA, 0.4 µM of each primer, 1X Mg^{2+}-free DyNAzyme EXT buffer, 1.0 mM $MgCl_2$, 1 mM dNTPs and 1 unit DyNAzyme EXT DNA polymerase (Finnzyme, Finland). PCR was performed with a Perkin-Elmer model 2400 GeneAmp apparatus for 35 cycles of denaturation at 94 ºC for 30 sec, annealing at 50 ºC for 30 sec, and extension at 72 ºC for 1 min. The purified DNA was analyzed on 0.8% agarose gel electrophoresis, then visualized under UV light and photographed.

4.6 DNA amplification of xylanase and pullulanase gene

4.6.1 Degenerate primers design

The degenerate oligonucleotide primers were designed based on the conserved regions of family 10 xylanase genes and family 13 α-amylases for xylanase and pullulanase genes, respectively.

4.6.2 PCR amplification

4.6.2.1 PCR amplification of xylanase gene

To amplify the partial sequence of xylanase gene from Bor Khlueng hot spring sediment, PCR was employed using XylF20 and XYNFR primers (Table 3). PCR amplification was performed in 50-µl reaction mixtures containing 2 µl of purified DNA, 1 µM of each primer, 1X thermophilic DNA polymerase buffer [50 mM KCl, 10 mM Tris-HCl (pH 8) and 0.1% Triton® X-100], 1.5 mM $MgCl_2$, 0.5 mM of dNTPs and 1 units of *Taq* DNA polymerase (Promega, USA). The reaction was performed with a Perkin-Elmer model 2400 GeneAmp apparatus for 35 cycles of denaturation at 94 ºC for 30 sec, annealing at 46 ºC for 1 min and extension at 72 ºC for 1 min. PCR

product was analyzed on 1.5 % agarose gel electrophoresis, then visualized under UV light and photographed.

4.6.2.2 PCR amplification of pullulanase gene

To amplify the partial sequence of pullulanase gene from Bor Khlueng hot spring sediments, PCR was employed using amyI and amyIV degenerate primers (Table 3). PCR amplification was performed in 50-µl reaction mixtures containing 2 µl of purified DNA, 1 µM of each primer, 1X thermophilic DNA polymerase buffer [50 mM KCl, 10 mM Tris-HCl (pH 8) and 0.1% Triton® X-100], 2.5 mM $MgCl_2$, 0.4 mM of dNTPs and 1 units of *Taq* DNA polymerase (Promega, USA). The reaction was performed with a Perkin-Elmer model 2400 GeneAmp apparatus for 30 cycles of denaturation at 94 ºC for 30 sec, annealing at 42 ºC for 1 min and extension at 72 ºC for 1 min. PCR product was analyzed on 1% agarose gel electrophoresis, then visualized under UV light and photographed.

4.7 Cloning of xylanase and pullulanase genes

4.7.1 Competent cell preparation using DMSO

A single colony of *E. coli* DH5α was inoculated into 250 ml SOB [2% peptone (Difco), 0.5% yeast extract (Difco), 0.05% NaCl] broth and aerated at 250 rpm 18 ºC until OD_{600} reached 0.6. The cell culture was transferred into five sterile 50 ml polypropylene centrifuge tubes and chilled on ice for 10 min. After centrifugation at 3,000 xg for 7 min at 4 ºC, the cell pellet was resuspended in 20 ml of ice-cold TB buffer [10 mM PIPES, 55 mM $MnCl_2$, 15 mM $CaCl_2$, and 250 mM KCl], and incubated on ice for 10 min. The cell pellets were then resuspended in 20 ml of ice-colded TB buffer and dimethy sulfoxide (DMSO) was added with gently swirling to give the final concentration of 7% (v/v). The cell suspension was finally dispensed into 100 µl aliquots in microcentrifuge tubes and stored at -80 ºC.

4.7.2 DNA ligation

The purified DNA was ligated into the pGEM-T Easy vector (Promega, Wisconsin, USA) in a molar ratio 1:3 (vector:insert). The appropriate amount of DNA used in ligation was calculated from the following formula:

$$\text{ng of insert} = \frac{\text{ng of vector} \times \text{kb size of insert} \times (\text{insert:vector molar ratio})}{\text{Kb size of vector}}$$

In this study, 25 ng of pGEM-T Easy vector was used in a total volume of 10 µl. Ligation mixture contained 1X rapid ligation buffer of T4 DNA ligase [30 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM EDTA, 1 mM ATP and 5% polyethyleneglycol] and 3 units of T4 DNA ligase (Promega). The ligation mixture was mixed and incubated for overnight at 4 ºC.

4.7.3 Transformation of competent *E. coli* DH5α cells

10 ng of ligated product was mixed with 100 µl of competent cells. The mixture was placed on ice for 30 min, immediately heat shocked at 42 ºC for 90 sec and placed back on ice for 5 min. A volume of 900 µl of SOC [2% peptone (Difco), 0.5% yeast extract (Difco), 0.05% NaCl, 20 mM glucose] medium was added to the mixture which was then incubated at 37 ºC for 1:30 h with constant shaking. The cells were spreaded on LB agar plate containing 100 g/ml of ampicillin, 4 µl of 0.8 M IPTG (isopropylthio-β-D-galactoside) and 20 µl of 50 mg/ml X-gal (5-bromo-4-chloro-3-indodyl-β-D-galactoside) (IPTG and X-gal were spreaded over the surface of an LB plate). The agar plate was incubated at 37°C for 12-16 h.

4.7.4 Master plate preparation and recombinant clone screening

After incubation for 12-16 h, blue and white colonies were observed. Approximately 20 white colonies were picked and spotted on a LB agar plate containing ampicillin, so called master plate. The master plate was incubated at 37 ºC for 12-16 h. To screen for the presence of DNA inserts from recombinant clones, simplified rapid size screening was performed. One colony per clone was picked by toothpick from the master plate and lysed in 30 µl of pre-warm lysis buffer [5 mM EDTA, 10% (w/v) sucrose, 0.23% (w/v) SDS, 100 mM NAOH, 60 mM KCl and 0.05% (w/v) bromphenol blue]. The lysed cells were incubated at 37 ºC for 5 min, placed on ice for 5 min and centrifuged at 13,000 xg for 5 min. Then, 20 µl of supernatant was analyzed on 0.8% agarose gel electrophoresis. Clones that contained the plasmid DNA with the larger size than that of the vector alone were selected for plasmid extraction.

4.8 Plasmid extraction

4.8.1 Plasmid extraction by CTAB method

A selected colony was inoculated into 3 ml of LB broth containing 100 mg/ml of ampicillin and incubated with vigorous shaking at 37 ºC, overnight. To collect the cell pellet, the cell suspension was centrifuged at 10,000 xg for 10 sec. The cell pellet was resuspended in 200 µl STET buffer [8% sucrose, 0.1% Triton® X-100, 50 mM EDTA, 50 mM Tris-HCl, pH 8.0] followed by vigorous vortexing. 25 µl of lysozyme solution (10mg/ml) was then added. Cell suspension was incubated at room temperature for 5-10 min and boiled for 45 sec. The lysed cells were centrifuged at 12,000 xg for 15 min before the pellet of cell debris and chromosomal DNA were removed using steriled toothpick. 20 µl of 5%CTAB was added, the mixture was then gently inverted and incubated at room temperature for 5 min. After centrifugation at 12,000 xg for 5 min, the supernatant was removed, the pellet was resuspended in 300 µl of 1.2 M NaCl and 5 µl of 10 mg/ml RNase A. After incubation at 37 ºC for 30 min, 300 µl of chloroform was added, the mixture was vigorously shaked for 30 sec and centrifuged at 12,000 xg for 5 min. The upper aqueous phase was transferred into a new microcentrifuge tube. To precipitate plasmid DNA, 2 volumes of absolute ethanol were added and the mixture was incubated at room temperature for 5 min. The precipitated DNA was obtained by centrifugation at 12,000 xg for 15 min. The plasmid DNA pellet was then washed twice with 70% ethanol and air-dried. Finally plasmid DNA pellet was resuspended in 20-30 µl of sterile water. The quality and concentration of plasmid DNA were analyzed on 0.8% agarose gel electrophoresis.

4.8.2 Extraction of plasmid DNA using QIAprep Spin Miniprep kit (QIAGEN)

The overnight cultured cells were collected by centrifugation at 5,000 xg for 10 min at 4 ºC, and the supernatant was discarded. The pellet was resuspended in 250 µl of buffer P1 and transferred to a new centrifuge tube. A 250 µl of buffer P2 was added and the mixture was gently inverted for 4-6 times. A volume of 350 µl of buffer N3 was then added and the mixture was further inverted for 4-6 times. After centrifugation at 10,000 xg for 10 min the supernatant was transferred to the QIAprep spin column which was inserted in a collection tube. The column was then centrifuged at 10,000 xg for 1 min and the flow-through was discarded. The column was then washed with 500 µl of buffer PB and centrifuged at 10,000 xg for 1 min. After the flow-through was discarded, the QIAprep column was washed with 750 µl

of buffer PE and centrifuged at 10,000 xg for 1 min. After the flow-through was discarded the column was centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep column was then placed in a clean 1.5 ml microcentrifuge tube and the DNA was eluted by adding 50 µl of buffer EB (10 mM Tris-Cl, pH 8.5). The DNA was left standing for 1 min and collected by centrifugation at 10,000 xg for 1 min.

4.9 Insert size screening by restriction enzyme digestion

After recombinant plasmid was purified, the presence of DNA insert was checked by *Eco*RI digestion. Plasmid (1µl) was added into the reaction mixture containing 1X buffer [90 mM Tris-HCl (pH7.5), 50 mM NaCl and 10 mM $MgCl_2$] and 2-4 units of *Eco*RI restriction endonuclease enzyme (Promega) in a final volume of 20 µl. The reaction mixture was incubated at 37 ºC for 3 h. The size of insert was analyzed using appropriate percentage of agarose gel.

4.10 DNA sequencing

After screening with restriction enzyme analysis, the recombinant clones harboring inserted DNA were sent for DNA sequencing (Macrogen, Korea).

4.11 Sequence analysis

DNA sequences obtained were translated into deduced amino acid sequences using Bioedit Sequence Alignment Editor program. The DNA sequences from different clones were aligned by Clustal X program. The homology of nucleotide and the deduced amino acid sequences were searched in Genbank database (http://www.ncbi.nlm.nih.gov) and compared to other related sequences by Clustal X program.

4.12 Amplification of 5' and 3' ends of xylanase and BK44 partial gene sequences

4.12.1 Genome walking PCR (Clontech, USA) (Figure 12)

The construction of DNA libraries was performed by digestion with different restriction enzymes (*Eco*RV, *Dra*I, *Stu*I, and *Pvu*II) that recognize a 6-base site and

generate blunt-ended fragments. Digested genomic DNA was then ligated to the GenomeWalker adaptor. The ligation mixture contained 1 µl of Genome Walker adaptor, 1X ligation buffer and 3 units of T4 DNA ligase (Invitrogen) in the total of 8 µl. After the ligation mixture was incubated for overnight at 16 ºC, two rounds of PCR amplification were performed using different primer pairs for 5'-end amplification. The first PCR amplification was performed using the outer adaptor primer (AP1) and outer gene specific primers (GSP1): XYL2GW_UP1 and BK44_UP1 for xylanase and BK44, respectively. The first PCR mixture was then used as a template for nested PCR amplification using the nested adaptor primer (AP2) and nested gene-specific primers (GSP2): XYL2GW_UP2 and BK44_UP2 for xylanase and BK44, respectively. Both PCR reactions contained 1 µl of digested genomic DNA template, 1X *Tth* PCR reaction buffer, 3 mM $Mg(OAc)_2$, 0.2 mM dNTPs, 0.2 µM of each primer and 1 units Advantage *Tth* polymerase Mix (Clontech) in a final volume of 25 µl. The primary PCR was performed using two-step cycles and the condition for PCR was 7 cycles at 94 ºC for 2 sec and 72 ºC for 3 min, followed by 32 cycles at 94 ºC for 2 sec and 67 ºC for 3 min and an final extension at 67 ºC for 4 min. Then the secondary PCR amplification was performed for 5 cycles at 94 ºC for 2 sec and 72 ºC for 3 min followed by 25 cycles at 94 ºC for 2 sec and 67 ºC for 3 min. The PCR product was analyzed on 1.5% agarose gel electrophoresis. Then they were purified, subcloned and sequenced as described in protocol 4.4, 4.6, 4.8, 4.9, 4.10 and 4.11. For 3'-end amplification, PCR reaction was performed as described above in except that the following primers were used: XYL2GW_DW1 and XYL2GW_DW2 (for xylanase), BK44_DW1 and BK44_DW2 (for BK44).

4.12.2 Semi-random two step PCR (129) (Figure 13)

Semi-random two step PCR (ST-PCR) is another variation of genome walking PCR that can be used to obtain the full-length gene. Like genome walking from Clontech kit, a set of gene specific primers was designed based on specific regions of partial DNA fragments. Two rounds of PCR amplification were performed using different pairs of primers. For 5'-end amplification, the primary PCR was performed with a gene specific primer (GSP1) and ST1 primers. The obtained PCR product was then used as a template for and nested PCR which was performed using a gene specific (GSP2) and ST2 primers. Both PCR reactions contained 2 µl of genomic

DNA template, 1X buffer, 3 mM $MgCl_2$, 0.2 mM of dNTP, 0.2 µM of each primer and 2.5 units of *Taq* DNA polymerase (Promega, USA) in a final volume of 25 µl. To increase the specificity, touchdown PCR was performed. After incubation of the reaction at 94 ºC for 2 min, touchdown PCR was started by decreasing annealing temperature for 1 ºC per cycle from 42 to 37 ºC for six cycles. This was then followed by 30 cycles of denaturation at 94 ºC for 30 sec, annealing at 55 ºC for 30 sec and extension at 72 ºC for 3 min. Finally, a final extension was performed at 72 ºC for 7 min. The nested PCR was then performed for 30 cycles of denaturation 94 ºC for 30 sec, annealing at 55 ºC for 30 sec and extension at 72 ºC for 3 min. The PCR product was analyzed on 1% agarose gel electrophoresis, purified, cloned and sequenced as described in protocol 4.4, 4.6, 4.8, 4.9, 4.10 and 4.11. For 3'-end amplification, PCR was performed as described above except that another set of primers was used as described in 4.12.1

4.13 Cloning and amplification of a full length BK44 gene

To amplify the full-length gene of BK44, two gene specific primers, FLBK44_F and FLBK44_R corresponding to the 5' and 3'ends (Table 3), respectively were designed according to the 5' and 3' sequences obtained from genome walking. The reaction consisted of 4 µl of genomic DNA template, 1X Mg^{2+}-free DyNAzyme EXT buffer, 1 mM $MgCl_2$, 0.5 mM dNTPs, 1 µM of each primer, 1 unit of DyNAzyme EXT DNA polymerase (Finnzyme, Finland), and sterile milliQ water to a final volume of 50 µl. PCR was performed for 35 cycles of denaturation at 94 ºC for 30 sec, annealing at 52 ºC for 30 sec and extension at 72 ºC for 1 min, followed by a final extension at 72 ºC for 10 min. The PCR product was then analyzed on 1% agarose gel electrophoresis, purified, subcloned and sequenced as described in protocol 4.4, 4.6, 4.8, 4.9, 4.10 and 4.11.

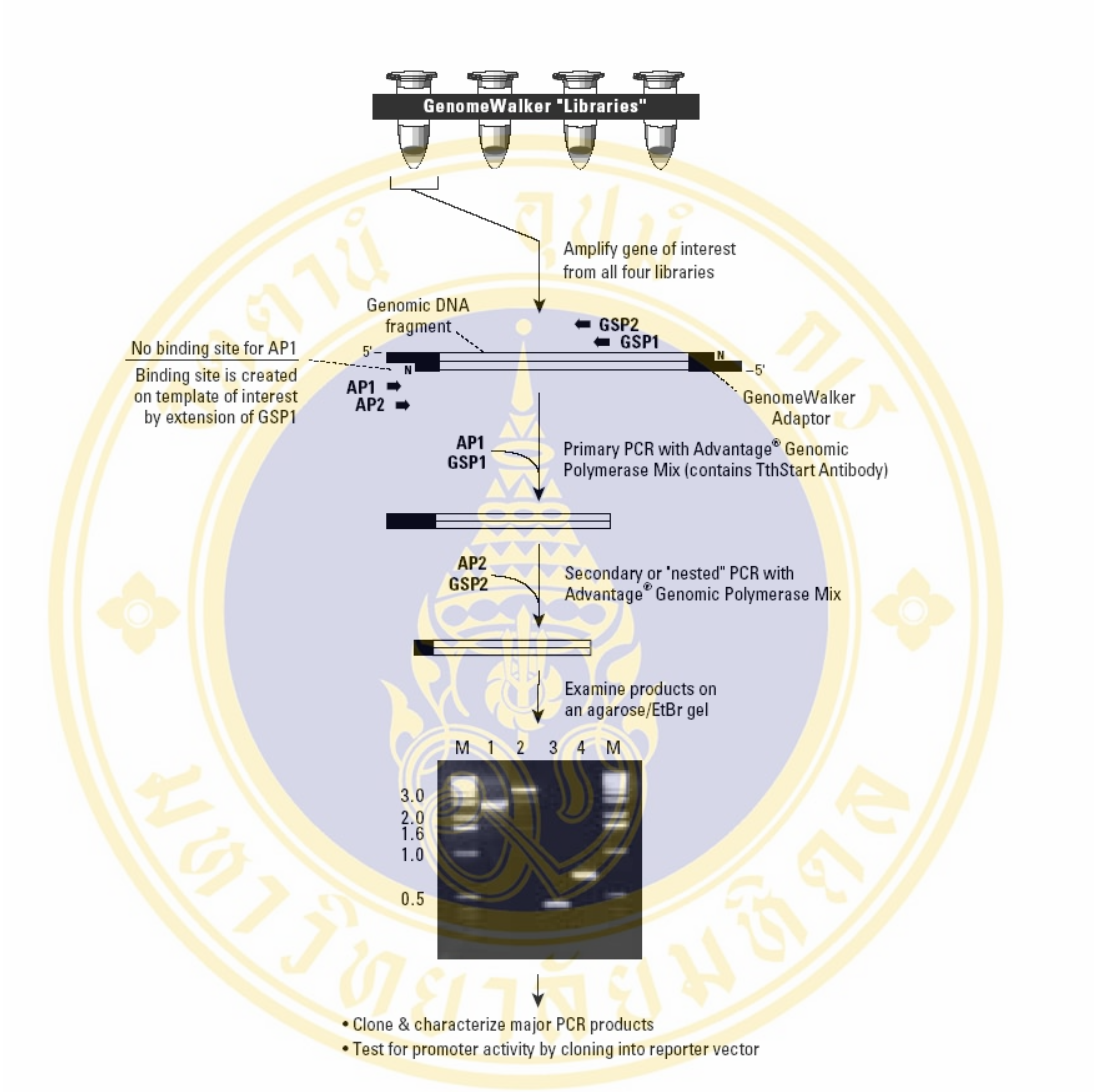


Figure 12. Schematic of genomic walking PCR (Clontech kit)

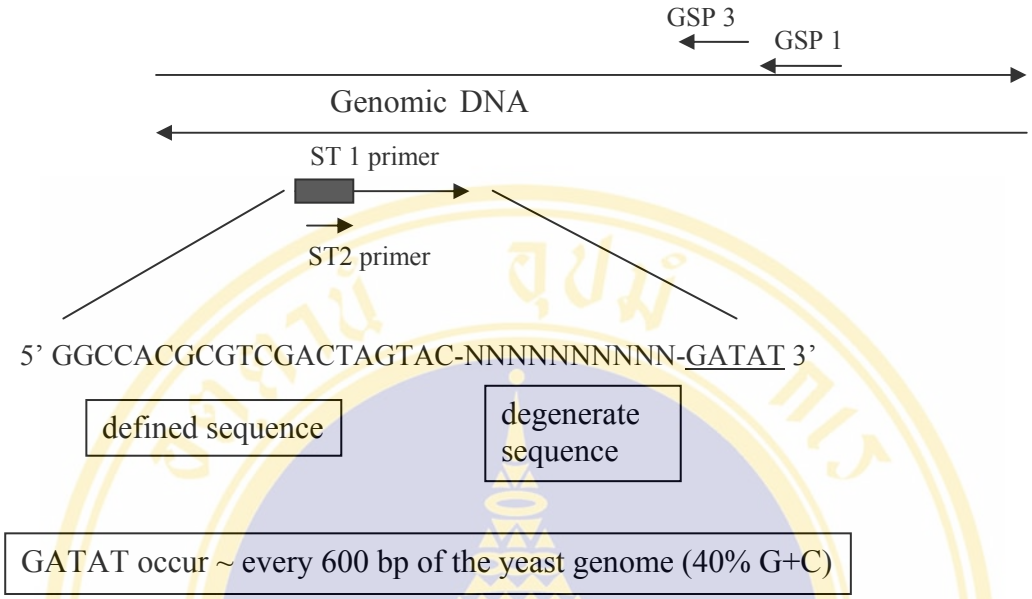


Figure 13. Semi-random two step PCR (ST-PCR) (taken from: Chun *et al*., 1997) (129)

4.14 Construction of BK44 in *P. pastoris* expression vector

4.14.1 Amplification of the full-length BK44

The full length gene of BK44 was amplified by PCR. The reaction was performed in a 50 µl reaction that contained 10 µM of pICBK44_F and pICBK44_R primers (Table 3), 1X DyNAzyme EXT buffer, 10 mM each of dNTPs, 1.5 mM $MgCl_2$, 10 ng of template DNA (full-length gene of BK44 in pGEM-T Easy vector) and 1U of DyNAzyme EXT DNA polymerase (Finnzyme, Finland). This reaction was performed for one cycle at 94 ºC, 35 cycles with 30 sec at 94 ºC, 30 sec at 42 ºC and 1 min at 72 ºC, then 10 min at 72 ºC. The PCR product was analyzed on 0.8% agarose gel electrophoresis and purified using QIAGEN gel extraction kit.

4.14.2 Preparation of linearized pPICZαA

To subclone the full-length BK44 gene into the *P. pastoris* vector, pPICZαA, the linearized pPICZαA vector was prepared by double digestion with *Xho*I and *Xba*I restriction enzymes (Promega, USA). A reaction contained 3 µg of pPICZαA vector, 5 unit of restriction enzyme, 1X reaction buffer and sterile distilled water to give a total volume of 60 µl. The digested products were purified using QIAGEN gel

extraction kit. DNA ligation and transformation were then performed as described in 4.4, 4.72 and 4.73, respectively. The transformants were screened on low salt LB medium [1% peptone (Difco), 0.5% yeast extract (Difco), 0.5% NaCl] containing 25 µg/ml Zeocin™ (Invitrogen).

4.15 Transformation of the recombinant plasmid into *P. pastoris*

4.15.1 Preparation of *P. pastoris* competent cells

A single colony of *P. pastoris* was aerated in 50 ml of YEPD at 30 ºC until OD_{600} reached 1.3-1.5. The cells were harvested and washed with 30 ml cold steriled water and centrifuged at 4,000 xg for 5 min at 4 ºC. The cells were then resuspended in 5 ml of ice-cold 1 M sorbitol and centrifuged at 4,000 xg for 5 min at 4 ºC. Finally, cell pellets were resuspended in 800 µl of ice-cold 1 M sorbitol. 50 µl of the competent cells were used for transformation by electroporation.

4.15.2 Preparation of linearized recombinant plasmids

Recombinant plasmids were linearized with *Dra*I and precipitated with 60 µl of 95% (v/v) ethanol in the presence of 3 M sodium acetate and 1.5 µl of yeast tRNA. After incubation at -80 ºC for 15 min, the mixture was centrifuged at 10,000 xg at 4 ºC for 10 min. The pellet was washed with 70% ethanol, and air-dried. The dried pellet was resuspended in 10 µl sterile distilled water.

4.15.3 Transformation of linearized plasmid into *P. pastoris* cells by electroporation

The reaction mixture was transferred to a pre-chilled 0.2 cm electro-cuvette and left on ice for 5 min. The cells were pulsed by using Bio-Rad Gene Pulser with the following conditions: 1.5 kV, 25 µF and 200 Ω (with time constant of approximate 4.5). 900 µl of YEPD was immediately added to the cells and the solution was transferred to a sterile microcentrifuge tube. The cell mixture was incubated at 30 ºC without shaking for 1 h. The culture was centrifuged at 3,000 xg at 4 ºC for 5 min and the supernatant was discarded. The cell pellets were resuspended in 100 µl sterile distilled water and spreaded on YEPD plate supplemented with Zeocin™, Plates were incubated at 30°C for 2-3 days until colonies were observed.

4.16 Total DNA isolation from *P. pastoris* transformants

Both the recombinant and the wild type *P. pastoris* strains were aerated at 30 ºC until OD_{600} reached 5-10 in YEPD containing 100 µg/ml Zeocin™. Cells were collected by centrifugation at 3,000 xg for 5 min at room temperature. The pellet was washed with 1 ml of sterile water and resuspended in 200 µl of SCED buffer, pH 7.5 (1 M sorbitol, 10 mM sodium citrate, pH 7.5, 10 mM EDTA, 10 mM DTT). A volume of 5 µl of lyticase (25U/ µl) was then added. The mixture was incubated at 30 ºC for 3-4 h and 100 µl of 2% (w/v) SDS was added. After the mixture was gently mixed, it was centrifuged at 10,000 xg for 5-10 min at 4 ºC. The supernatant was collected and transferred into a new microcentrifuge tube. 1 ml of absolute ethanol was added and the mixture was incubated at 4 ºC for overnight. DNA was then centrifuged at 10,000 xg for 20 min at 4 ºC, and resuspended in 500 µl of TE (10 mM Tris HCl, pH 7.4 and 1 mM EDTA, pH 8.0). The DNA solution was incubated with 5 µl of 10 mg/ml RNase A at 37 °C for 1 h. An equal volume of phenol (pH 8) was then added. The mixture was gently inverted and centrifuged at 14,000 xg for 5 min. An aqueous phase was transferred into a new microcentrifuge tube. An equal volume of chloroform: isoamyl alcohol (24:1) was added. The mixture was gently inverted and centrifuged at 14,000 xg for 5 min. An aqueous phase was then transferred to a new microcentrifuge tube. This step was repeated once; then 0.5 volume of 7.5 ammonium acetate, pH 7.5 and two volumes of absolute ethanol were added to the DNA solution which was then placed at -80 ºC for 10 min. The solution was then centrifuged at 10,000 xg for 20 min at 4 ºC. The DNA pellet was washed with 70% (v/v) ethanol and resuspended in 30 µl of sterile distilled water. The obtained genomic DNA was stored at 4 ºC until used.

4.17 PCR analysis of *P. pastoris* integrants

Screening of *P. pastoris* integrants containing integrated recombinant plasmid was performed by PCR. The reaction of 50 µl consisted of 1X buffer, 1.5 mM $MgCl_2$, 10 mM of dNTP, 10 µM of each primer (5' AOX1 forward primer and 3'AOX1 reverse primer), 50-100 ng of genomic DNA and 0.2U of *Taq* DNA polymerase (Promega, USA). The reactions were operated in an automated thermal cycler GeneAmp PCR system model 2400 (Perkin Elmer Cetus, USA). PCR was performed for one cycle at 94 °C for 2 min, followed by 30 cycles at 94 ºC for 30 sec, 50 °C for

30 sec at and 72 °C for 1 min. Finally the reaction was performed at 72 °C for 10 min. A 10 µl of the PCR products was analyzed on 0.8% (w/v) agarose gel electrophoresis.

4.18 Expression of recombinant BK44 in *P. pastoris*

A single colony of *P. pastoris* KM71 recombinant was inoculated into 2 ml of YEPD and incubated at 30 °C with vigorous shaking at 250 rpm for 48 h. The cell culture was then transferred to 5 ml of fresh BMGY medium and grown under the same condition described above until the culture reached an OD_{600} of 5-6. For induction step, the cell pellet was harvested by centrifugation at 4,000 xg for 5 min at room temperature. The supernatant was discarded and the cell pellet was resuspended in BMMY using 1/5 the volume of original culture (1 ml). The cell suspension was transferred in a 20 ml glass tube. Absolute methanol was added every 24 h to give a final concentration of 3% (v/v) to maintain induction. After methanol induction, the supernatant was collected at 0, 24, 48, 72, 96 h. The collected cells were centrifuged at 14,000 xg for 3 min at room temperature. The supernatant was transferred to a new tube. The secreted protein was analyzed by SDS-PAGE and the pullulanase activity was investigated.

4.19 Cell lysis of *P. pastoris*

To investigate the presence of any intracellular proteins in *P. pastoris* cell pellet, yeast cell pellet was thawed on ice. For each 1 ml of sample, 100 µl of Breaking Buffer (50 mM sodium phosphate, pH 7.4, 1 mM PMSF, 1 mM EDTA and 5% (w/v) glycerol) was added followed by the acid-washed glass beads (0.5 mm) to lyse the cells. The mixture was vortexed for 30 sec, then incubated on ice for 30 sec. This step was repeated for 7 more times. The sample was centrifuged at 12,000 xg for 5-10 min at 4 °C. The supernatant was transferred to a new microcentrifuge tube and analyzed by using SDS-PAGE.

4.20 Total RNA isolation of *P. pastoris* by TRI REAGENT®

The cell pellet of *P. pastoris* was mixed with 1 ml of TRI REAGENT® by pipetting for several times and incubated for 5 min at room temperature. Then, 200 µl of chloroform was added. The mixture was mixed vigorously by vortexing followed

by standing for 2-15 min at room temperature. The solution was centrifuged at 12,000 xg for 15 min at 4 ºC. The colorless upper aqueous phase was transferred to a new 1.5 ml microcentrifuge tube. Total RNA in the solution was precipitated by mixing with 0.5 ml of isopropanol and stored at room temperature for 5-10 min. The solution was subsequently centrifuged at 12,000 xg for 8 min at 4 ºC. The RNA pellet was washed by adding 75% (v/v) ethanol followed by centrifugation at 7,500 g for 5 min at 4 ºC. After removal of the supernatant, the RNA pellet was briefly air dried and resuspended in RNase-free (DEPC-treated water) sterile distilled water. The RNA solution was completely dissolved by incubation at 55-60 ºC for 10-15 min. The dissolved RNA was stored at -80 ºC until used.

The concentration of extracted RNA was examined by measuring the absorbance at 260 nm (A_{260}). The total RNA concentration could be calculated by the following equation.

RNA concentration (µg/ml) = 40 x A_{260} x dilution factor

The A_{260}/A_{280} ratio was applied to determine the RNA purity. The RNA sample with the ratio of about 1.8-2.0 was used in the next process.

4.21 Removal of DNA contamination for total RNA

The contamination of DNA in total extracted RNA was avoided by DNase digestion. The reaction consisted of total extracted RNA, 3 units of RNase-free DNase I (Fermentas), 1X DNase I reaction buffer with $MgCl_2$, and DEPC-treated water. The mixture was incubated at 37 ºC for 1 h. To stop the reaction, 1 µl of 25 mM EDTA was added and the mixture was incubated at 65 ºC for 10 min.

4.22 First-strand cDNA synthesis

The cDNAs were synthesized in a mixture containing 1 µg of total RNA with 2µM of random hexamer primer and RNase-free sterile distilled water. The primer was allowed to anneal with mRNA at 70 ºC for 5 min and the solution was quickly cooled down on ice. The following components: 1X ImProm-II™ reaction buffer, 1.0 mM dNTPs, 2 mM $MgCl_2$, 1µl of ImProm-II™ Reverse transcriptase, and RNase-free sterile distilled water, were added into the mixture. The solution was gently mixed before the cDNA synthesis was processed at 25 ºC for 5 min, 42 ºC for 60 min, and

70 ºC for 15 min. Four microlitres of synthesized cDNA were used for PCR amplification.

4.23 Amplification of pullulanase BK44 gene from recombinant *P. pastoris*

To detect transcriptional level of pullulanase BK44 gene, the PCR reaction was performed in a 50 µl reaction that contained 10 µM of FLBK44_F and FLBK44_R primers, 1X DyNAzyme buffer, 10 mM each of dNTPs, 1.5 mM $MgCl_2$, cDNA of BK44 and 1U of DyNAzyme EXT DNA polymerase (Finnzyme, Finland). Amplification was performed as followed: one cycle at 94 ºC, 35 cycles with 30 sec at 94 ºC, 30 sec at 52 ºC and 1 min at 72 ºC, followed by extension for 10 min at 72 ºC. The PCR product was analyzed on 0.8% agarose gel electrophoresis**.**

4.24 Construction of recombinant BK44 in *Escherichia coli* expression vector

4.24.1 Construction of recombinant pET-32a(+) plasmid harboring BK44 full length gene

Two specific primers corresponding to the 5' and 3' ends of BK44 were designed (Table 3). Purified genomic DNA was used as a template in PCR using DyNazyme EXT DNA polymerase. The PCR reaction was prepared in a 50-µl reaction, PCR amplification was performed for 35 cycles of denaturation at 94 ºC for 30 sec, annealing at 42 ºC for 30 sec and extension at 72 ºC for 1 min. The reaction was completed by extension at 72 ºC for 10 min. The PCR product was analyzed on 0.8% agarose gel electrophoresis. The amplified DNA was double digested with *Nco*I and *Eco*RI, purified by QIAGEN gel extraction kit and ligated to pET-32a(+) which had been cut with *Eco*RI and *Nco*I. The ligation reactions were performed in a total volume 20 µl consisting of 1X T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, and 25 µg/ml BSA, pH 7.5) and 2.5 units of T4 DNA ligase (Invitrogen). The reaction was incubated at 16 ºC for overnight. Ligation mixture was then purified, subcloned and sequenced as described in protocol 4.4, 4.7.3, 4.10 and 4.11. Transformation was performed using *E. coli* strain Rosetta-gami as an expression host. The presence of recombinant plasmid harboring BK44 full-length gene was analyzed by *Eco*RI and *Nco*I double digestions.

4.24.2 Construction of recombinant pET-43.1a(+) plasmid harboring BK44 full length gene

Two specific primers corresponding to the 5' and 3' ends of BK44 were designed (Table 3). Purified genomic DNA was used as a template in PCR using DyNAzyme EXT DNA polymerase. The PCR reaction was prepared in a 50-µl reaction. PCR amplification was performed for 35 cycles of denaturation at 94 ºC for 30 sec, annealing at 42 ºC for 30 sec and extension at 72 ºC for 1 min. The reaction was completed by extension at 72 ºC for 10 min. The PCR product was analyzed on 0.8% agarose gel electrophoresis. The amplified DNA was double digested with *Bam*HI and *Kpn*I, purified by QIAGEN gel extraction kit and ligated to pET-43.1a(+) which had been cut with *Bam*HI and *Kpn*I. The ligation reactions were performed in a total volume 20 µl. Ligation mixture was then purified, subcloned and sequenced as described in protocol 4.4, 4.7.3, 4.10 and 4.11. Transformation was performed using *E. coli* strain Rosetta-gami as an expression host. The presence of recombinant plasmid harboring BK44 full-length gene was analyzed by *Bam*HI and *Kpn*I double digestions.

4.24.3 Expression of recombinant BK44 in *E. coli*

A single colony of *E. coli* transformant harboring BK44 gene was inoculated into 3 ml of LB broth containing 50 µg/ml of ampicillin and 34 µg/ml of chloramphenical and the bacterial culture was aerated at 37 ºC for 12-16 h. 3% (v/v) of a starter culture was added into 50 ml flask containing 1/5 volume of LB broth. The culture was incubated at 30 ºC with shaking at 250 rpm. When $O.D._{600}$ reached 0.6, cell culture was induced with 0.1 mM IPTG for 4 h. The culture was placed on ice for 5 min and cells were harvested by centrifugation at 8,000 xg for 5 min at 4 ºC and resuspended in 1 ml of cold 20 mM Tris-HCl (pH8.0). The cell was centrifuged as described above and resuspended in 750 µl of ice cold 20 mM Tris-HCl (pH 8.0). Cells were then lysed by sonication for 1 min. The cell lysate was centrifuged to separate soluble and pellet fractions. All fractions were analyzed by SDS-PAGE to determine the presence of target protein.

4.25 Protein electrophoresis

4.24.1 Sample preparation

Protein samples were mixed with 4X protein sample buffer [60 mM Tris-HCl (pH 7.5), 2% (w/v) SDS, 10% glycerol, 0.025% (w/v) bromphenol blue, 100 mM DTT] at the ratio of 3:1, and boiled for 5 min. The heated samples were kept on ice and centrifuged at 13,000 xg for 5 min. The protein samples were loaded into 12% SDS-polyacrylamide gel.

4.25.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Stacking and separating gels for SDS-PAGE were prepared by mixing various solution as described. The proteins were electrophorised in protein running buffer (25 mM Tris-HCl, pH 8.2, 192 mM glycine and 0.1% (w/v) SDS) using a constant voltage of 150 volts for 60 min or until the dye front reached the bottom of the gel. The gel was stained in staining buffer (45% (v/v) acetic acid and 0.1% (v/v) coomassie blue R-25) for 1-2 h. The excess stain was washed in destaining buffer (10% (v/v) methanol and 10% (v/v) acetic acid) for overnight. Destaining buffer was removed by rinsing several times with distilled water. The gel was dried by placing between two sheets of wet cellophane paper, which later clamped together. Air bubbles were removed. The gel was air-dried overnight at room temperature.

4.26 Enzyme activity assay

4.26.1 Pullulanase activity assay using AZCL-pullulan

0.1% of AZCL-pullulan (Megazyme) was immersed in 96% ethanol for 1 h then mixed with prewarm 1.2% bacteriology agar in 0.1 M phosphate buffer pH 5.8. The mixture was poured into a sterile plate using 20 ml of mixture per plate. Once, the agar was solidified the wells for applying sample were punctured using a sterile 5 mm diameter tube. A volume of 20 µl of the recombinant cell supernatant was applied to each well. The plate was incubated at 50 ºC. The blue zone was developed after overnight incubation. The pullulanase activity was determined by measuring the length of halo from edge of the well to the edge of the blue zone in mm.

4.26.2 Cyclodextrinase activity assay using DNS method

Cyclodextrinase activity was determined by DNS method (130). Enzyme solution was assayed in microcentrifuge tubes that contained 40 µl of 1% w/v α-, β- or γ-cyclodextrin (CD), 50 µl of 0.1M MOPS buffer, pH 7.0 and 2µl of 0.1 M $CaCl_2$ in a

100 µl reaction. The mixture was incubated overnight at 50 ºC. The enzymatic reaction was stopped by adding 300 µl of 3, 5-dinitrosalicylic acid solution and then boiling for 10 min. After centrifugation the mixture was transferred to a microplate and the developed color was determined at the wavelength 540 nm using a microplate reader (SpectraMax 190; Molecular Devices Corporation).

Table 4. Components of SDS-PAGE for protein determination

Preparation of 5% of stacking SDS-PAGE	
Solution	Volume (ml)
30% Acrylamidea	0.83
H2O	3.4
1M Tris-HCl pH 6.8	0.63
10% SDS	0.05
10% APS	0.05
TEMED	0.01
Total volume	4.97
Preparation of 8% separating SDS-PAGE	
Solution	Volume (ml)
30% Acrylamidea	4.6
H2O	2.7
1.5 M Tris-HCl pH 8.8	2.5
10% SDS	0.1
10% APS	0.1
TEMED	0.012
Total volume	10.012
Preparation of 12% separating SDS-PAGE	
Solution	Volume (ml)
30% Acrylamidea	4
H2O	3.3
1.5 M Tris-HCl pH 8.8	2.5
10% SDS	0.1
10% APS	0.1
TEMED	0.007
Total volume	10.007

Acrylamide:N,N'-methylene-bis-acrylamide 29:1 (w/w)

APS = ammonium persulphate

TEMED = N,N,N',N'-tetramethyl-ethylenediamine

CHAPTER V
RESULTS

5.1 Direct extraction and purification of genomic DNA from the sediment of Bor Khlueng hot spring

Genomic DNA from the sediment of Bor Khlueng hot spring was extracted using SDS based DNA extraction method (128). The yield of crude extracted DNA was 11.24 μg of DNA per gram of soil. According to Zhou *et al* (128), 2.5 to 26.9 μg of DNA per gram of soil could be obtained from this method, depending on the sources of the soil. The size of genomic DNA was larger than 23 Kb (Figure 14), which was in agreement with those from previous studies (122, 128). To avoid contamination from humic substances that might interfere in some downstream reactions such as PCR (131, 132) and restriction enzyme digestion (133), DNA purification was performed using QIAGEN gel extraction kit. PCR was then performed with the purified DNA using universal primers specific to bacterial 16S rDNA. This was to investigate if the quality of DNA is sufficiently free of humic acid for downstream experiments. The result showed that a PCR product of 1.5 kb was obtained which was the expected size of bacterial 16S rDNA (Figure 15). This suggested that DNA was of high quality for further experiments.

5.2 PCR amplification using xylanase degenerate primers

To directly obtain partial xylanase gene from unculturable microorganisms, degenerate primers (XylF23 and XYNFR primers) were designed based on the conserved amino acid sequences of glycosyl hydrolase family 10 xylanase. PCR product of 167 bp was obtained (Figure 16) and purified using QIAquick gel extraction kit (QIAGEN). The purified fragment was then ligated to pGEM-T Easy vector and transformed into *E. coli* DH5α. The transformants were selected on LB medium supplemented with ampicillin and recombinant plasmids were extracted. Restriction enzyme analysis with *Eco*RI digestion was performed and the result

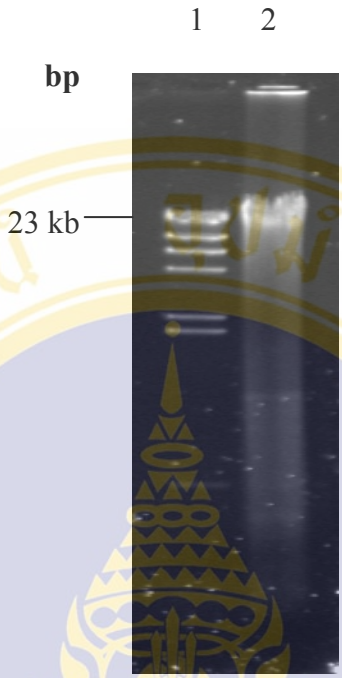


Figure 14. Genomic DNA extraction using SDS based DNA extraction method (128)

Genomic DNA from the sediment of Bor Khlueng hot spring was analyzed on 0.8% agarose gel electophoresis and the gel was stained with ethidium bromide.

Lane 1 : Lamda DNA/*Hin*dIII marker

Lane 2 : Crude extracted DNA

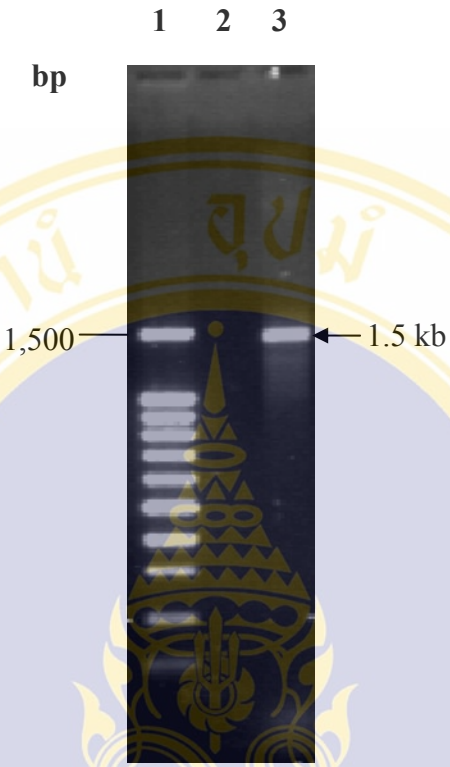


Figure 15. Determination of purified DNA

Crude extracted DNA was purified and PCR was performed using bacterial 16S rDNA primers and analyzed on 1.2% agarose gel electrophoresis and the gel was stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : Negative control for PCR amplification (without template)

Lane 3 : The amplified PCR product of 16S rDNA

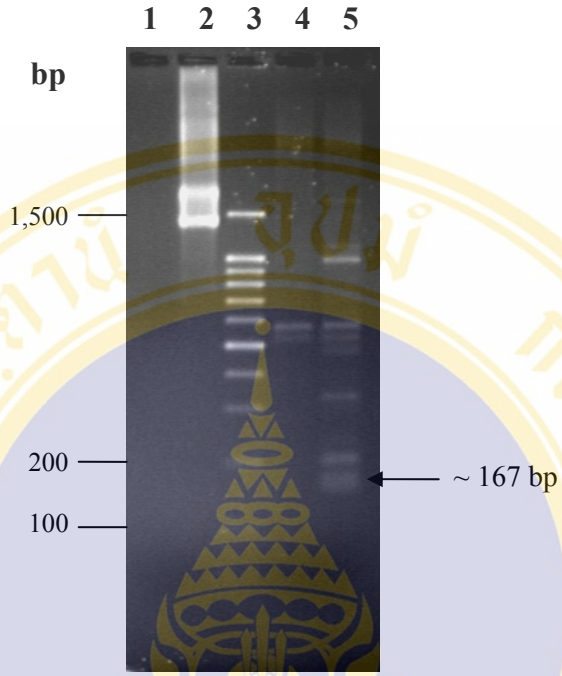


Figure 16. PCR amplification of partial xylanase gene

PCR amplification of partial fragment using XylF23 and XYNFR primers. The size of expected PCR product of approximately 167 bp. PCR products were analyzed on 1.5% agarose gel electrophoresis stained with ethidium bromide.

Lane 1 : Negative control (without template)

Lane 2 : 16S rDNA

Lane 3 : 100 bp ladder with 1.5 kb plus

Lane 4 : The amplified PCR product of XylF23+XylR20 primers

Lane 5 : The amplified PCR product of XylF23+XYNFR primers

showed that each clone harbored various sizes of DNA inserts (Figure 17). The randomly selected recombinants were submitted for DNA sequencing. From DNA sequencing analysis, 3 different partial sequences encoding xylanases were found. These recombinant plasmids contained 149, 152 and 167 nucleotides which exhibited 53% identities to *Streptomyces halstedii* beta-1,4-endoxylanase (*XysA*) gene (Figure 18), 50% identities to xylanase from *Acidobacterium capsulatum* DNA (Figure 19) and 65% identities to *Thermobacillus xylanilyticus xynA* (Figure 20). Accordingly, the longest 167 bp partial xylanase sequence was used for genome walking approach (Clontech) to obtain the full-length sequence of this xylanase gene (Figure 21).

5.3 Amplification of the 3' and 5' ends of partial xylanase gene using genome walking

To obtain the full-length xylanase gene, a set of gene specific primers was designed (Table 3) to retrieve the 3' and 5' ends of xylanase gene directly from the environmental DNA using genome walking PCR (Clontech). To retrieve the 3'-end of xylanase gene, PCR amplification was performed with two rounds of PCR amplification to increase specificity. In the second PCR, the amplified products of approximately 350 and 500 bp were obtained (Figure 22). The fragments were purified and cloned into pGEM-T Easy vector. Restriction enzyme digestion was performed to verify the DNA insert and recombinant clones were subjected to DNA sequencing. DNA sequence analysis revealed that the obtained 350 bp PCR product showed no sequence similarity to family 10 xylanase gene. However, the obtained PCR product of approximately 500 bp revealed the length of 3' end sequence of 460 nucleotides, encoding 153 amino acids (Figure 23). Nucleotide sequence analysis showed that this sequence exhibited 56% identity to endo-1,4-beta xylanase A precursor (Xylanase A) of *Bacteroides ovatus* (Figure 24). Therefore, the downstream region was successfully obtained. To retrieve the 5' end of xylanase gene, the obtained upstream sequence was analyzed. However, no DNA sequence corresponding to any bacterial xylanases in the database was obtained.

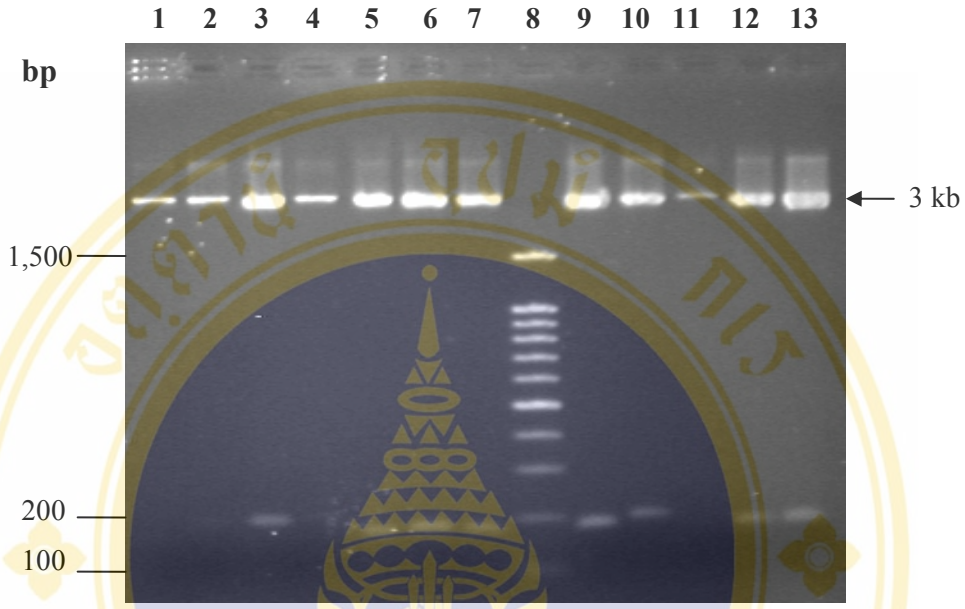


Figure 17. Restriction enzyme analysis with *Eco*RI digestion

Recombinant clones were digested with *Eco*RI restriction enzyme. The digested products were analyzed by 1.2% agarose gel electrophoresis.

Lane 1 : pGEM-T Easy vector

Lanes 2-7, 9-13 : Recombinant clones with *Eco*RI digestion

Lane 8 : 100 bp ladder with 1.5 kb plus

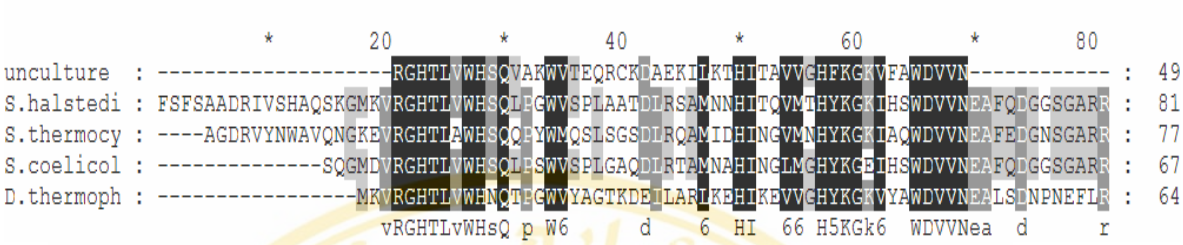


Figure 18. Alignment of the deduced amino acid sequences of partial xylanase containing 149 bp from the sediment of Bor Khlueng hot spring

The deduced amino acid sequence of partial xylanase was aligned by Clustal X with other bacterial xylanase genes. These sequences were xylanases from *Streptomyces halstedii* (AAC45554), *S. thermocyaneoviolaceus* (AAF04600), *S. coelicolor* (CAB61191) and *Dictyoglomus thermophilum* (Q12603).

.

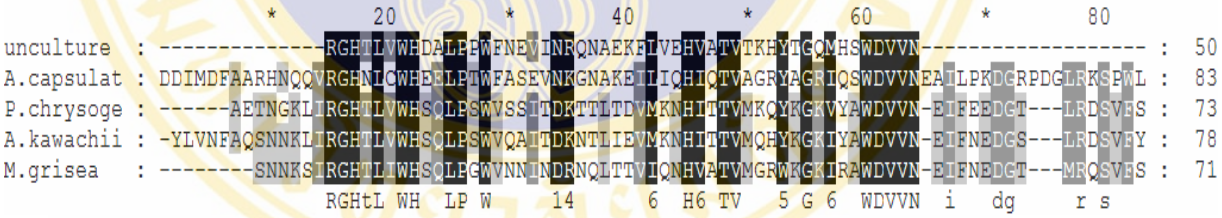


Figure 19. Alignment of the deduced amino acid sequences of partial xylanase containing 152 bp from the sediment of Bor Khlueng hot spring

The deduced amino acid sequence of partial xylanase was aligned by Clustal X with other bacterial xylanase genes. These sequences were xylanases from *Acidobacterium capsulatum* (BAB40957), *Penicillium chrysogenum* (AAS93681), *Aspergillus kawachii* (BAA03575) and *Magnaporthe grisea* (AAC41684).

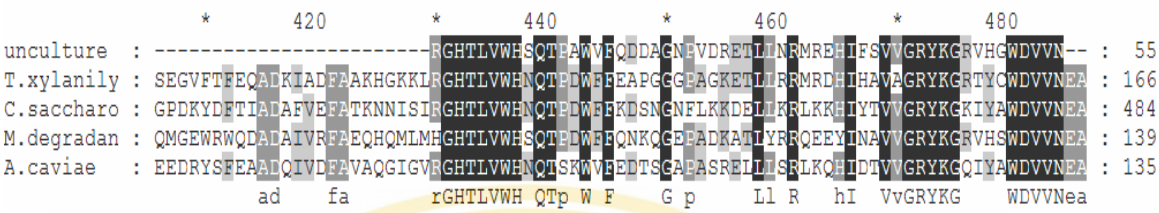


Figure 20. Alignment of the deduced amino acid sequences of partial xylanase containing 167 bp from the sediment of Bor Khlueng hot spring

The deduced amino acid sequence of partial xylanase was aligned by Clustal X with other bacterial xylanase genes. These sequences were xylanases from *Thermobacillus xylanilyticus* (CAA76420), *Caldocellum sacchrolyticum* (ZP_00884870), *Aeromonas punctata* (BAA31551) and *Microbulbifer degradans* (ZP_00315347)

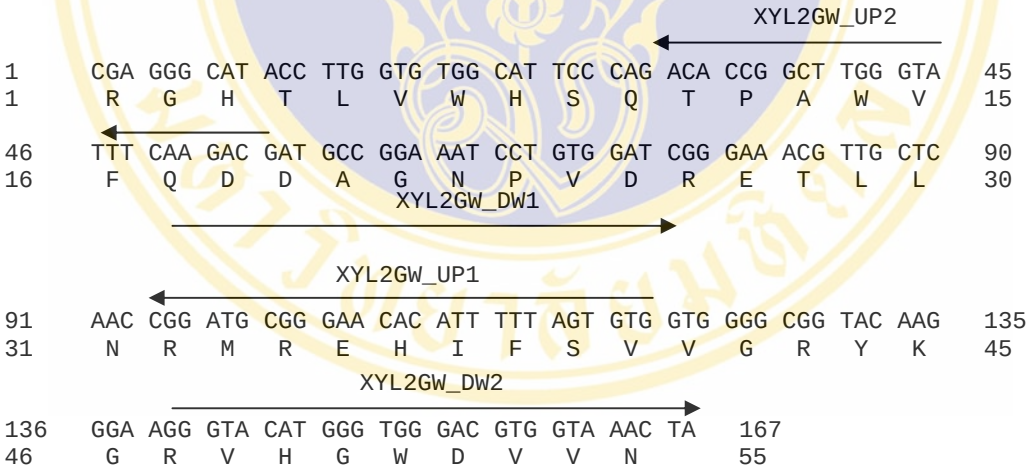


Figure 21. Nucleotide and deduced amino acid sequences of partial xylanase from the sediment of Bor Khlueng hot spring

DNA sequencing result revealed 167 nucleotides of the partial xylanase gene. A set of gene specific primers for 5' end (XYL2GW_UP1 and XYL2GW_UP2 primers) and 3' end (XYL2GW_DW1 and XYL2GW_DW2 primers) were designed and indicated by arrows.

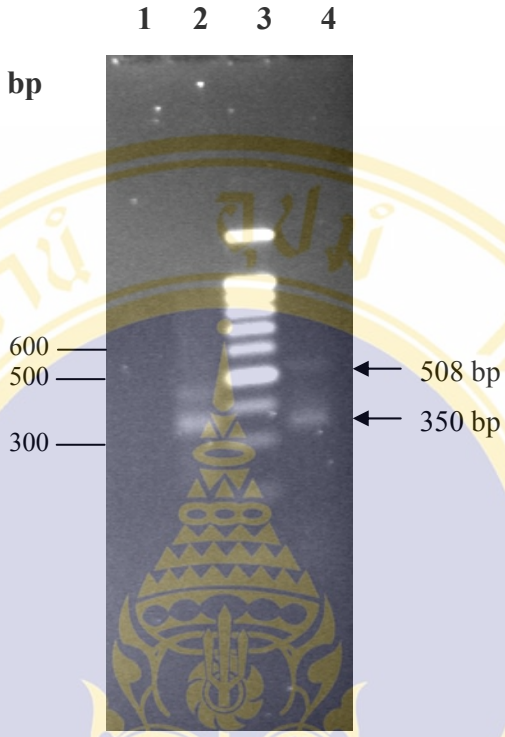


Figure 22. PCR amplification of xylanase downstream walking

PCR amplification of downstream region using a set of gene specific primers (XYL2GW_DW1 and XYL2GW_DW2 primers). PCR product was analyzed on 1.2% agarose gel electrophoresis and stained with ethidium bromide.

Lane1: negative control

Lane2: PCR product obtained from upstream genome walking

Lane3: 100 bp ladder with 1.5 kb plus

Lane4: PCR product obtained from downstream genome walking

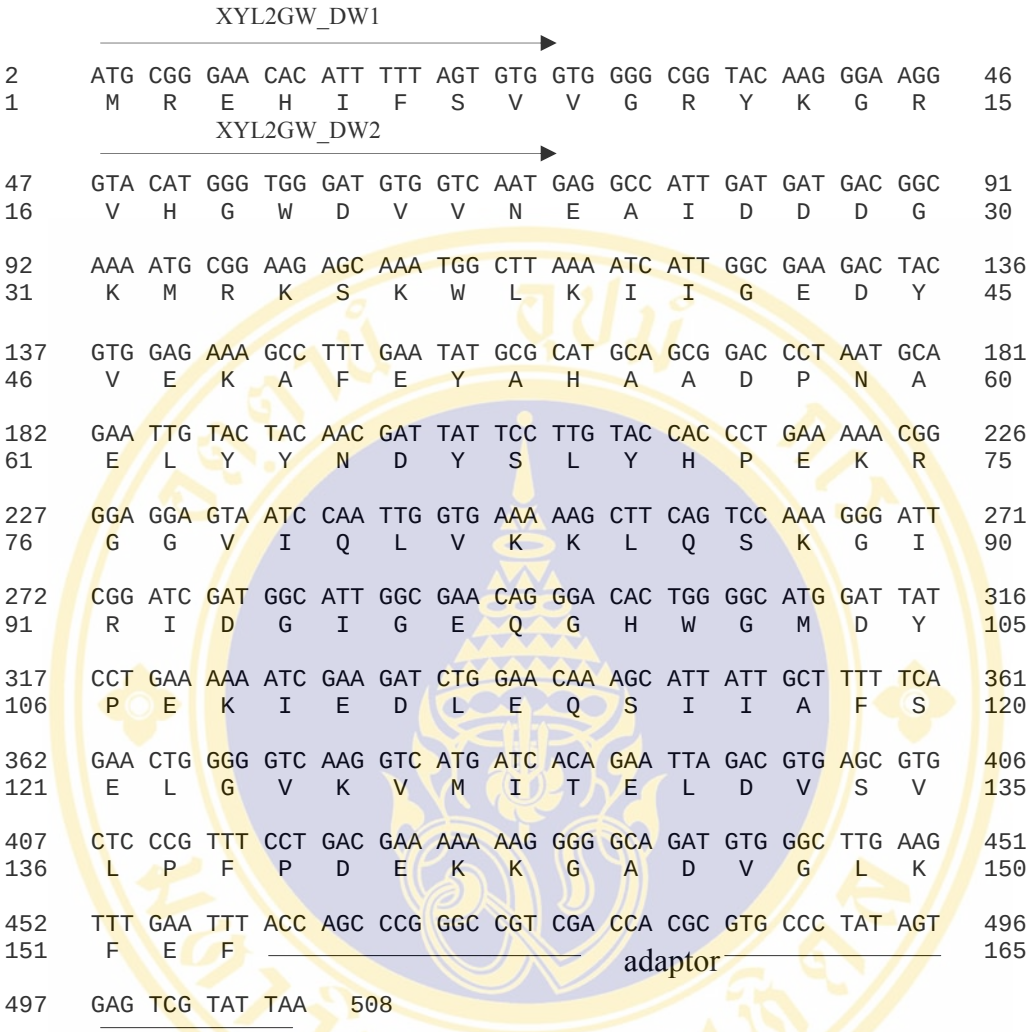


Figure 23. Nucleotide and deduced amino acid sequences of the 3' end of the xylanase gene

The 3' end of the xylanase gene was obtained from unculturable microorganism using genome walking method (Clontech). DNA sequencing revealed 460 nucleotides that encode 153 amino acids. The positions of primers were indicated by arrows.

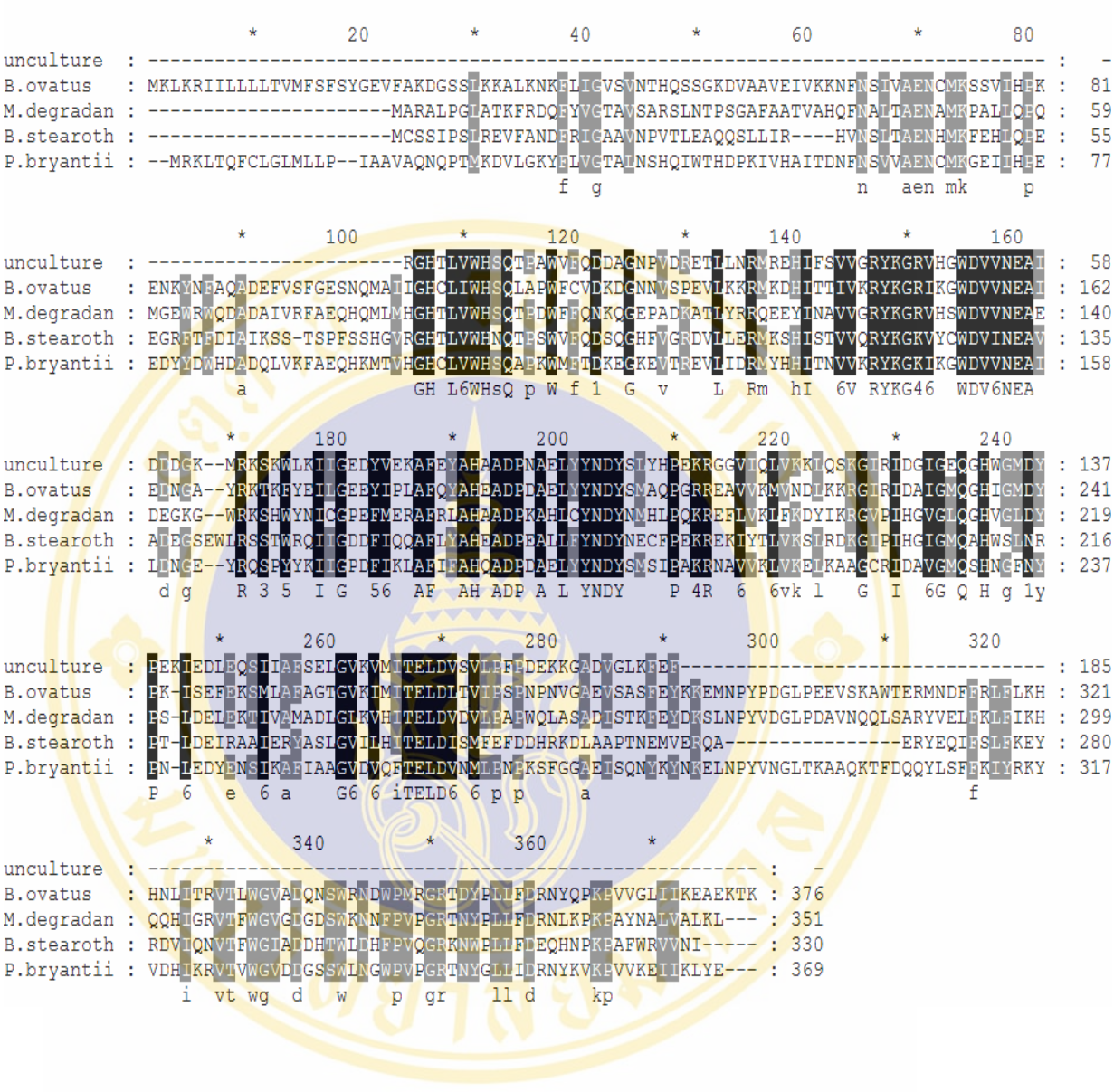


Figure 24. Amino acid alignment of Bor Khlueng hot spring xylanase gene (3'end) against other bacterial xylanase genes

The xylanase sequences used were from *Bacteroides ovatus* (AAB08023), *Microbulbifer degradans* (ZP_00315347), *Bacillus stearothermophilus* (D28121) and *Prevotella bryantii* (CAA89207). Shading of alignment represents the degree of conservation. The black shading shows the most conserved region among all sequences compared.

To solve this problem a new set of gene specific primers were designed and various libraries using different enzymes digestion in genome walking PCR were constructed. However, amplification of the upstream region of approximately 300 bp was unsuccessful.

As the stop codon was not present in the sequence and the upstream region still could not be retrieved, semi-random two-step PCR (ST-PCR) method was performed in both upstream and downstream region. In the 3' end of xylanase gene, additional gene specific primers were designed for another attempt to obtain the complete downstream region of the xylanase. The result showed that PCR product of 700 bp was obtained using this ST-PCR approach (Figure 25). The fragment was purified and cloned into pGEM-T Easy vector. Plasmid extraction and restriction enzyme analysis were performed. The recombinant plasmids harboring target DNA inserts were submitted to DNA sequencing and the result revealed that the obtained sequence of 728 bp included the putative stop codon of the gene (Figure 26). Alignment of amino acid sequences obtained from genome walking with those from other bacterial xylanases suggested that this gene belonged to family 10 xylanase (Figure 27). From these results, it can be concluded that the complete downstream region of directly isolated xylanase was successfully obtained. However, the amplification of 5' end of xylanase gene using ST-PCR did not obtain any PCR product (Figure 25).

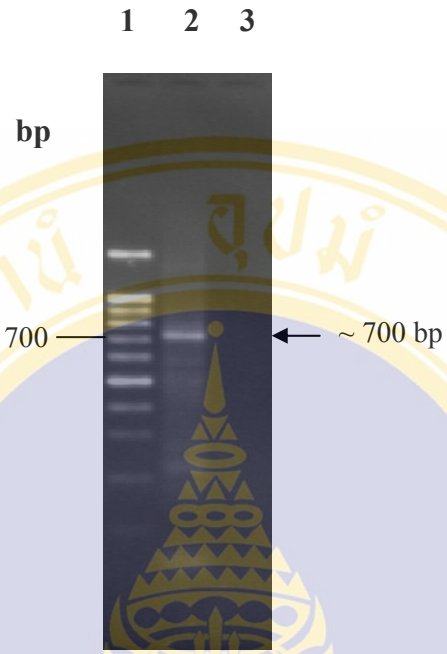


Figure 25. PCR amplification of downstream walking using ST-PCR approach

PCR amplification was performed to obtain the complete downstream region of the xylanase gene. PCR products were analyzed on 1.5% agarose gel electrophoresis stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : PCR amplification from downstream walking

Lane 3 : PCR amplification from upstream walking

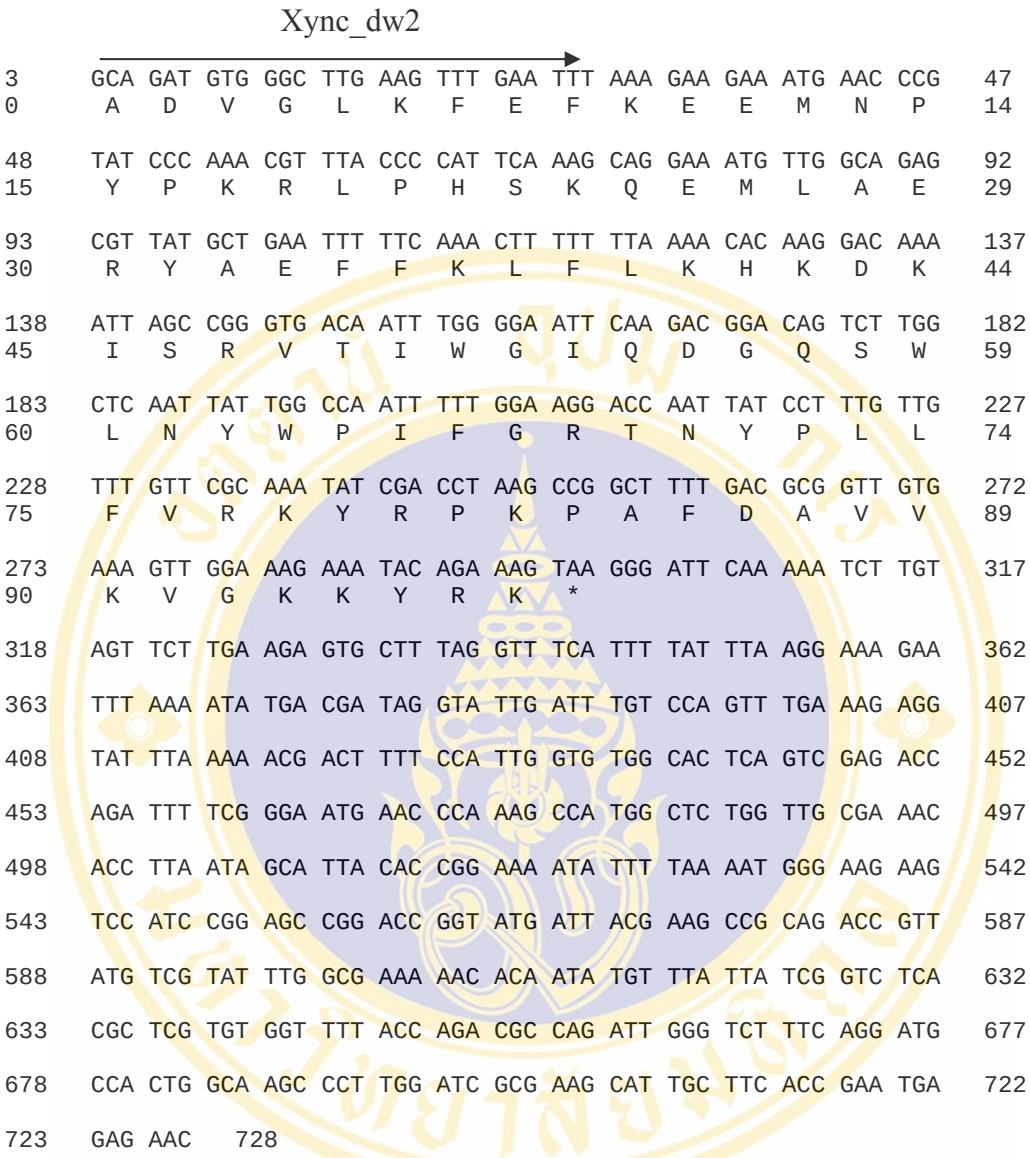


Figure 26. Nucleotide and deduced amino acid sequences of the 3' end of the xylanase gene obtained from ST-PCR

The figure showed the 3' end of the xylanase gene from unculturable microorganism using ST-PCR method. DNA sequencing revealed 723 nucleotides that encoded 97 amino acids. Star represents a stop codon. The position of primer was also shown.

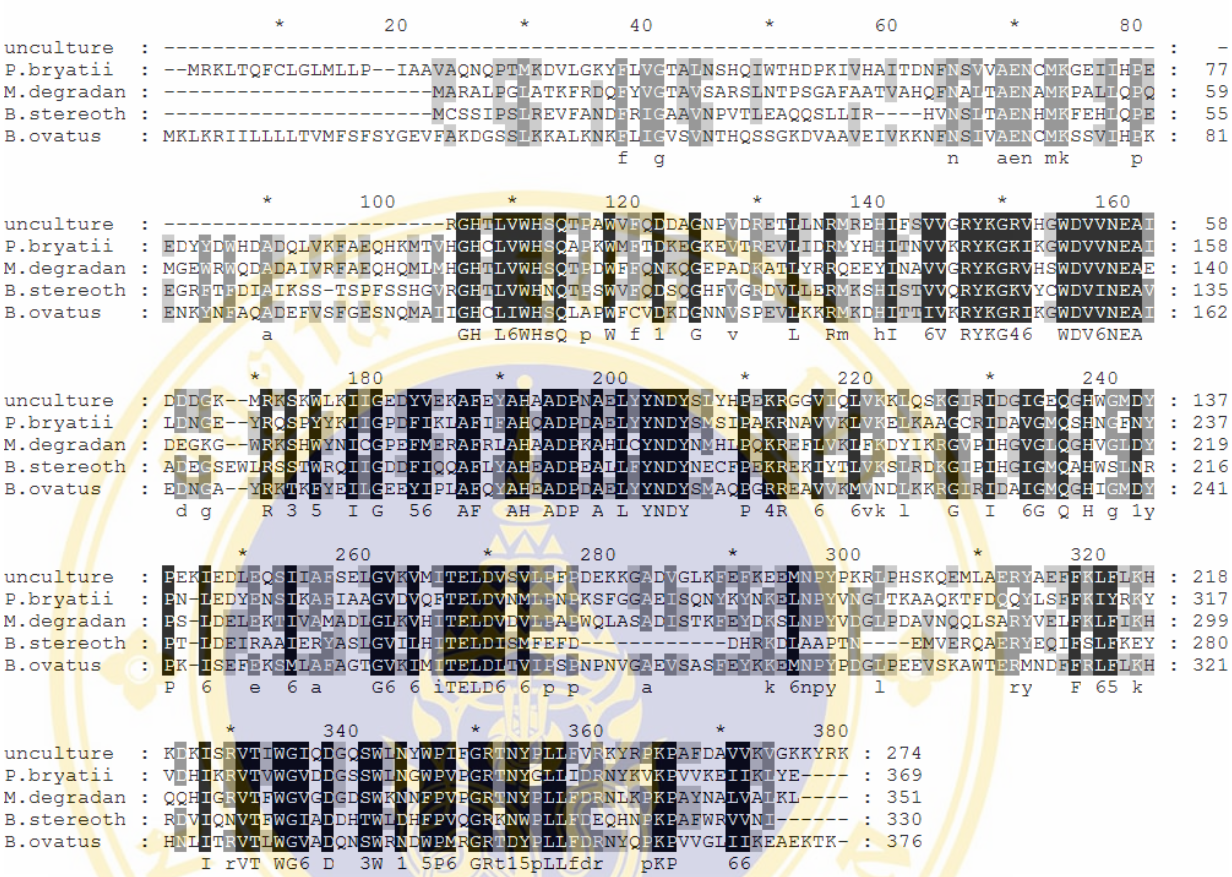


Figure 27. Amino acid sequence alignment of Bor Khlueng hot spring xylanase gene (3'end) against other bacterial xylanase genes

The xylanases used for the alignment were from *Bacteroides ovatus* (AAB08023), *Microbulbifer degradans* (ZP_00315347), *Bacillus stearothermophilus* (D28121) and *Prevotella bryantii* (CAA89207). Shading of alignment represents the degree of conservation. The black shading shows the most conserved region among all sequences compared.

5.4 PCR amplification using α-amylase family degenerate primers

To directly obtain a partial pullulanase gene from environmental DNA, degenerate primers (amyI and amyIV primers) were designed based on the conserved amino acid sequences of α-amylase family (family 13 glycosyl hydrolase). The size of the expected PCR product of approximately 560 bp was obtained (Figure 28). These PCR products were purified from agarose using QIAquick gel extraction kit (QIAGEN). The purified fragment was ligated to pGEM-T Easy vector and transformed into *E. coli* DH5α. The transformants were selected on LB medium supplemented with ampicillin followed by plasmid extraction and restriction enzyme analysis. The chosen recombinant plasmids were subjected to DNA sequencing. BLAST analysis revealed that, one clone, BK44, exhibited 54% amino acid sequence identity to glycosyl hydrolase family 13 of *Deinococcus radiodurans* and 51% identity to *Geobacillus kaustophilus* alpha-cyclodextrinase (Figure 29). Therefore, this BK44 partial sequence was used for obtaining the full-length gene by genome walking approach.

5.5 Amplification of the 5' and 3' ends of BK44 using genome walking method

To obtain the full length gene of BK44, a set of gene specific primers was designed to retrieve the 5' and 3' ends of BK44 directly from the environmental DNA (Figure 30). To retrieve the 5'-end of BK44, PCR amplification was performed using genome walking method with two rounds of PCR amplifications to increase specificity. In the secondary PCR, the amplified product of approximately 1.2 kb was obtained (Figure 31). This fragment was purified and subcloned into pGEM-T Easy vector. Restriction enzyme analysis was performed to verify the size of DNA inserts and four recombinant clones were subjected to DNA sequencing. From DNA sequence analysis, the length of 5' end region obtained from genome walking was 1,114 bp. This region included a putative Shine-Dalgarno sequence (AAGG) and the putative start codon of the target gene (Figure 32). The obtained DNA sequence exhibited 54% amino acid identity to maltogenic amylase of *Thermus* sp., and alpha-cyclodextrinase of *Geobacillus kaustophilus* and 53% amino acid sequence identity to (neo)pullulanase of *Thermus thermophilus*. Thus, this suggested that the BK44 might encode an enzyme belonging to neopullulanase subfamily.

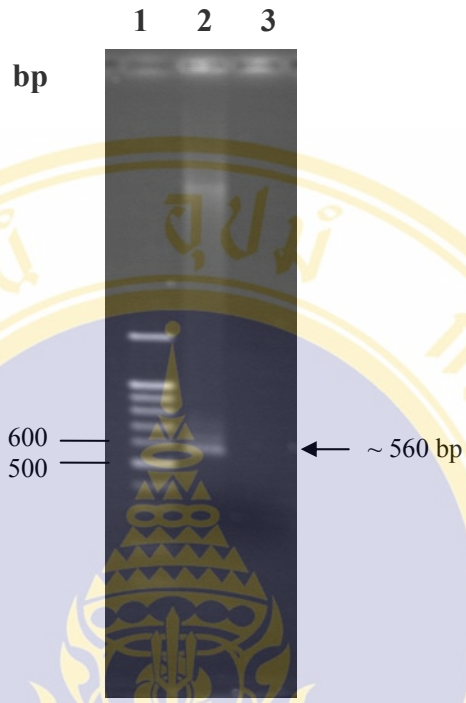


Figure 28. PCR amplification of the partial BK44 gene

PCR amplification of partial fragment using amyI and amyIV primers. PCR products were analyzed on 1% agarose gel electrophoresis and stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : The amplified PCR product of partial BK44 DNA fragment

Lane 3 : Negative control of PCR amplification (without template)

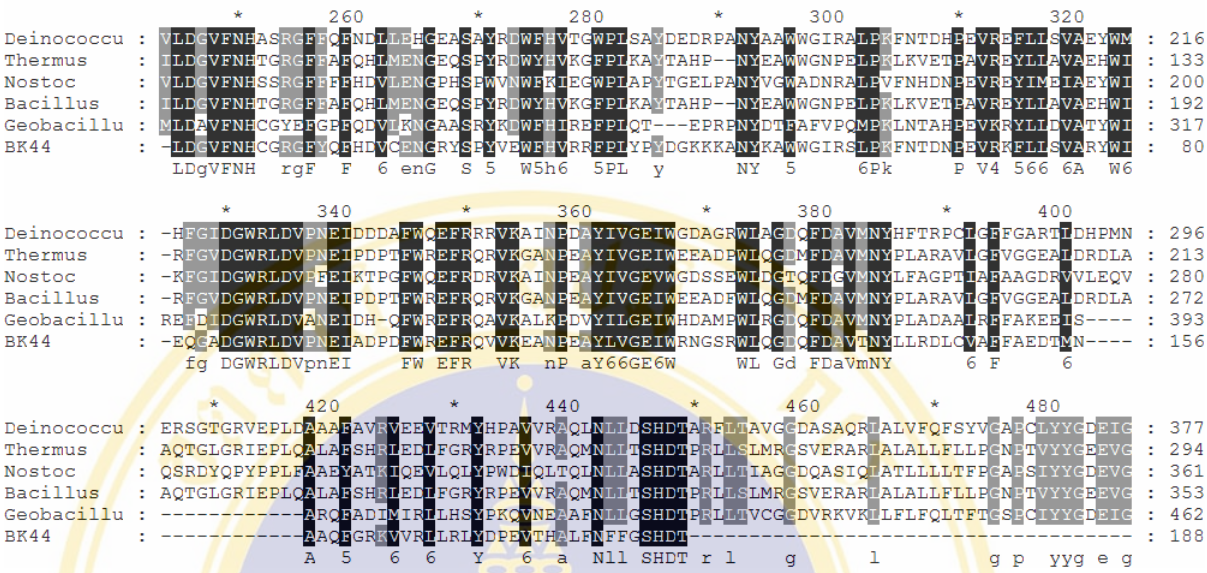


Figure 29. Alignment of the deduced amino acid sequence of BK44 against bacterial α-amylase family

The deduced amino acid sequence of BK44 was aligned by Clustal X with other bacterial enzymes in α-amylase family: *Deinococcus radiodurans* glycosyl hydrolase family 13 (AAF10708), *Thermus thermophilus* HB27 putative pullulanase (YP_005167), *Nostoc sp.* neopullulanase (BA000019), *Bacillus flavocaldarius* pullulanase (AB008764), *Geobacillus kaustophilus* alpha-cyclodextrinase (BA000043). Shading of the alignment represents the degree of conservation. The black shading shows the most conserved region among all sequences compared.

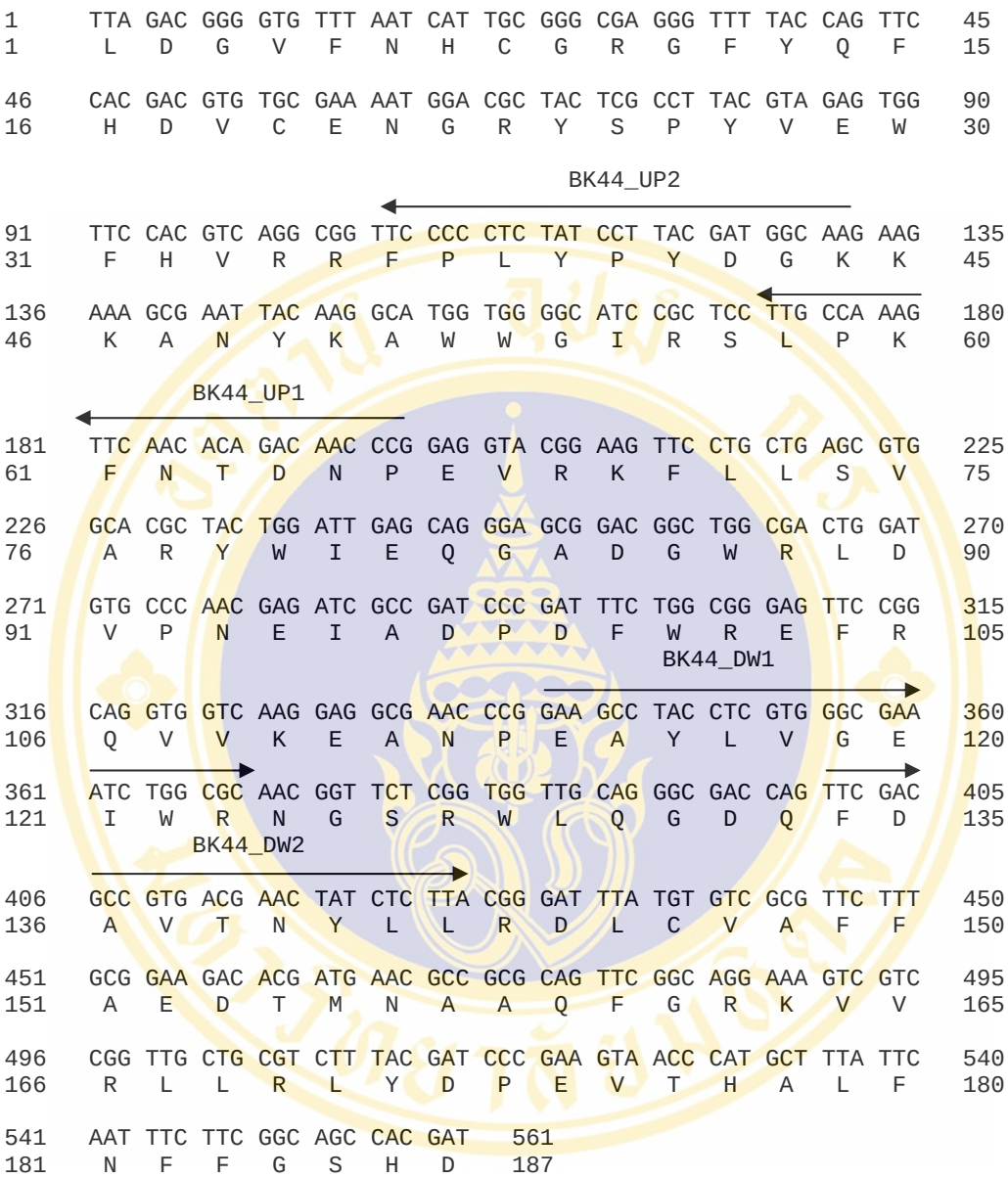


Figure 30. Nucleotide and deduced amino acid sequences of BK44 from the sediments of Bor Khlueng hot spring

DNA sequencing revealed 561 nucleotides of the partial BK44 gene. A set of gene specific primers for 5' end (BK44_UP1 and BK44_UP2 primers) and 3' end (BK44_DW1 and BK44_DW2 primers) were designed and indicated by arrows.

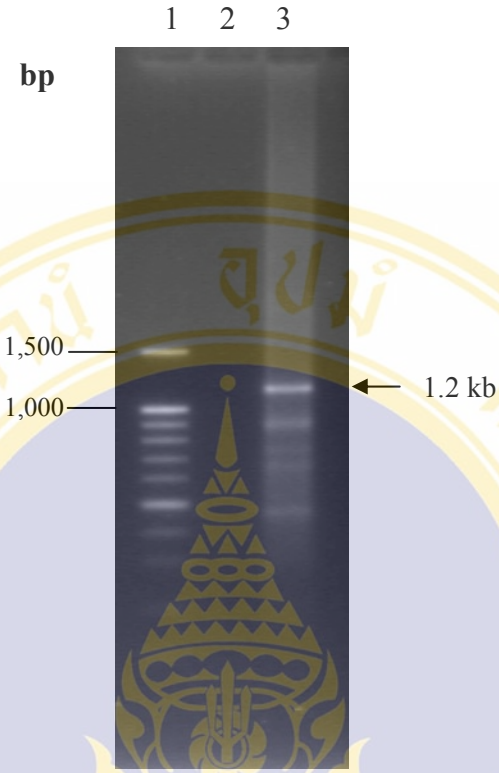


Figure 31. PCR amplification to obtain the upstream region of BK44 by genome walking

PCR amplification was performed using a set of gene specific primers (BK44_UP1 and BK44_UP2 primers) to obtain the upstream region. PCR product was analyzed on 1% gel electrophoresis and stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : Negative control for PCR amplification (without template)

Lane 3 : The amplified PCR product from upstream walking

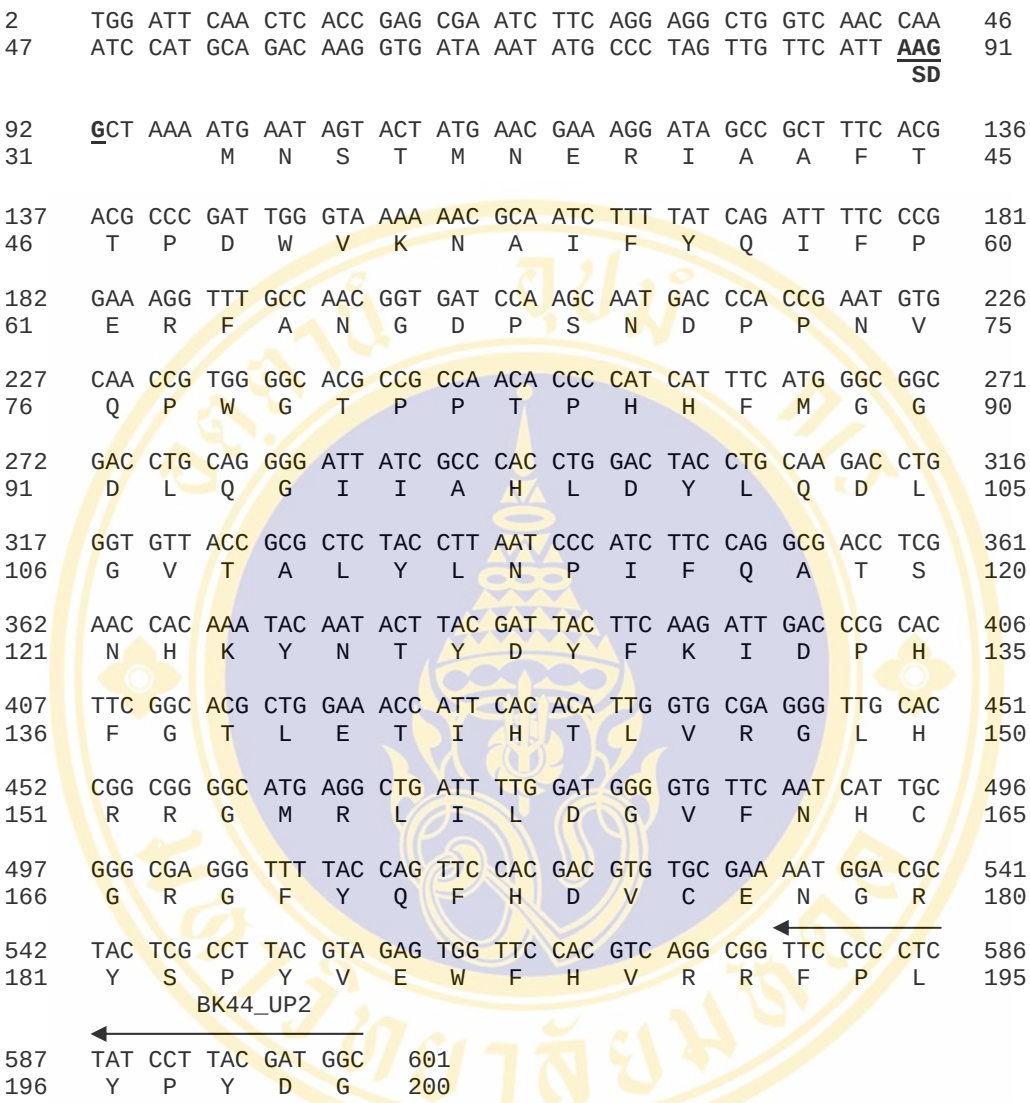


Figure 32. Nucleotide and deduced amino acid sequences of 5' end of BK44

5' end of BK44 was obtained from environmental DNA using genome walking. DNA sequencing revealed 504 nucleotides that encode 168 amino acids including a putative start codon. Positions of primers were indicated by arrows. The putative Shine-Dalgarno sequence (SD) is underlined.

To obtain the 3' end of BK44, genome walking was performed as described above. The PCR product of approximately 1 kb was obtained (Figure 33). This fragment was purified and cloned into pGEM-T Easy vector. Restriction enzyme digestion was performed to verify the size of DNA inserts. Five recombinant plasmids harboring 3' end region of BK44 were subjected to DNA sequencing. The result revealed 920 nucleotides of downstream region obtained from genome walking. The deduced of amino acid sequences included a putative stop codon (Figure 34). BLAST analysis showed similarity of this sequence with those belonging to family 13 glycosyl hydrolase.

5.6 Amplification of the full-length BK 44 gene

After the 5' and 3' ends of BK44 sequence were identified. FLBK44_F and FLBK44_R primers were used for PCR to obtain the full-length gene of BK44. The PCR product of approximately 1,500 bp was obtained (Figure 35). This fragment was purified and cloned into pGEM-T Easy vector. DNA sequencing was performed and the result revealed that the full-length gene of BK44 contained 1,458 bp encoding 485 amino acid sequences with a calculated mass of 55 kDa (Figure 36) and did not contain a signal peptide. Analysis of the sequence revealed that the full length BK44 exhibited 54% amino acid sequence identity to *Thermus sp.* maltogenic amylase, 54% identity to *Bacillus stearothermophilus* alpha-cyclodextrinase and 52% *Bacillus stearothermophilus* neopullulanase (Figure 37).

The four highly conserved regions related to catalytic and substrate binding site among the α-amylase family could be identified in the full-length BK44 sequences as regions I-IV (Figure 37). The 3D structure of BK44 was preliminary predicted from the deduced amino acid sequences using SWISS-MODEL program. The result showed that the monomeric 3D structure of BK44 contained the common feature of α-amylase family in domain A or $(\beta/\alpha)_8$ barrel structure, domain B and domain C (Figure 38B). In addition, three conserved catalytic residues, Asp 215, Glu 245 and Asp 312 like other α-amylase family enzymes were shown in Figure 36 and 38B. However, when compared to the structure of neopullulanase from *Thermoactinomyces vulgaris* R-47 (TVA II) (106), BK44 lacked the additional N-domain which is commonly found among enzymes in the neopullulanase subfamily.

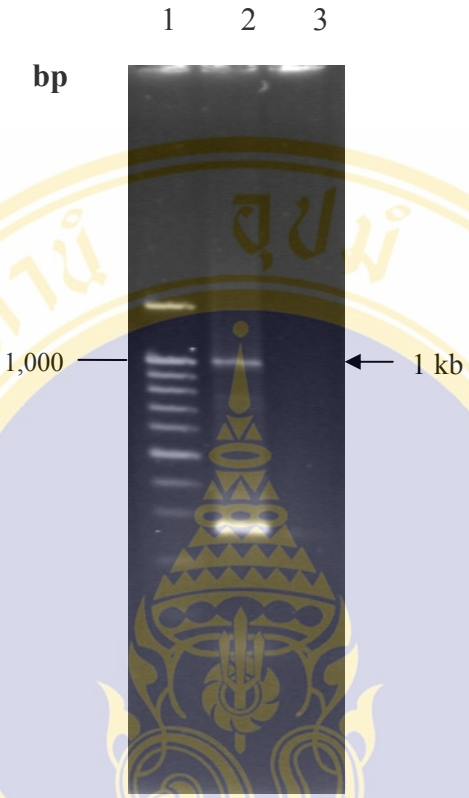


Figure 33. PCR amplification to obtain downstream region of BK44 by genome walking

PCR amplification of downstream region using a set of gene specific primers (BK44_DW1 and BK44_DW2). PCR product was analyzed on 1% agarose gel electrophoresis stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus.

Lane 2 : The amplified PCR product from downstream walking

Lane 3 : Negative control for PCR amplification (without template)

```
                BK44_DW2
                ---------->
2     GAC GCC GTG ACG AAC TAT CTC TTA CGG GAT TTA TGT GTC GCG TTC   46
1      D   A   V   T   N   Y   L   L   R   D   L   C   V   A   F    15

47    TTT GCG GAA GAC ACG ATG AAC GCC GCG CAG TTC GGC AGG AAA GTC   91
16     F   A   E   D   T   M   N   A   A   Q   F   G   R   K   V    30

92    GTC CGG TTG CTG CGT CTT TAC GAT CCC GAA GTA ACC CAT GCT TTA   136
31     V   R   L   L   R   L   Y   D   P   E   V   T   H   A   L    45

137   TTC AAC CTG TTG GGA AGC CAC GAT ACT GCC CGC TTC CTG ACC GTT   181
46     F   N   L   L   G   S   H   D   T   A   R   F   L   T   V    60

182   GCT GGG GAT GAG GTG GAA CGG GTC AAA CTC GCC TTC ACC TTC CTG   226
61     A   G   D   E   V   E   R   V   K   L   A   F   T   F   L    75

227   ATG ACC TAT CCG GGC GCT CCC TGC CTT TAC TAC GGC GAT GAG ATT   271
76     M   T   Y   P   G   A   P   C   L   Y   Y   G   D   E   I    90

272   GGC ATG AAG GGC GCG AAG GAC CCC CAC AAC CGC GCC TGC TTC CCC   316
91     G   M   K   G   A   K   D   P   H   N   R   A   C   F   P   105

317   TGG GAC GAA GGG CAA TGG AAC AAA GAA CTG CAG GCA CAC GTG AAA   361
106    W   D   E   G   Q   W   N   K   E   L   Q   A   H   V   K   120

362   GCC CTG ATC GCC CTG CGC AAG AAG CAC GCT GCC CTC CGC ACC GGG   406
121    A   L   I   A   L   R   K   K   H   A   A   L   R   T   G   135

407   GCT TAC CAA ACC CTG CTG GCT GAT GGA AAG GCA AAT GTG TAC GCC   451
136    A   Y   Q   T   L   L   A   D   G   K   A   N   V   Y   A   150

452   TTC GCC CGC TGG GAT GCA GGC GCA ACG CTC CTG ATC GCT CTG AAC   496
151    F   A   R   W   D   A   G   A   T   L   L   I   A   L   N   165

497   AAC AGC CCG GAG GCT TGG AGC GGT GTC CTC TCT CTC AAA AGC CTT   541
166    N   S   P   E   A   W   S   G   V   L   S   L   K   S   L   180

542   TCT CAA GTT TCA GGC CTC AAG TTT CGG GTT GTT TTC CCT CAA CTC   586
181    S   Q   V   S   G   L   K   F   R   V   V   F   P   Q   L   195

587   GGC ACT CGA AAC CGG AAA CTC GAA ACC GGT CAA CGG ACA GGC CCA   631
196    G   T   R   N   R   K   L   E   T   G   Q   R   T   G   P   210

632   TTT GCC ATC CAA TTG GCT CCG CGA AGT GGG GTT GTT CTT CAG GTA   676
211    F   A   I   Q   L   A   P   R   S   G   V   V   L   Q   V   225

677   GTG TAA TTT CAA CGG GGC TGG AGG ATC GGC AGT GAT TGG ATC GAA   721
226    V   *                                                     240

722   CCT TCA ACA GGA GAC CCT CCG GCT GGC GCG AAT AGG GGA TTA AGG   766
767   TGA AAA AGA ATG CCG GAC TAC                                 787
```

Figure 34. Nucleotide and deduced amino acid sequences of the 3' end of BK44

3' end of BK44 was obtained from environmental DNA using genome walking method. DNA sequencing revealed 681 nucleotides that encoded 227 amino acids. Star represents a stop codon at the 228th position. Arrow indicates position of primer.

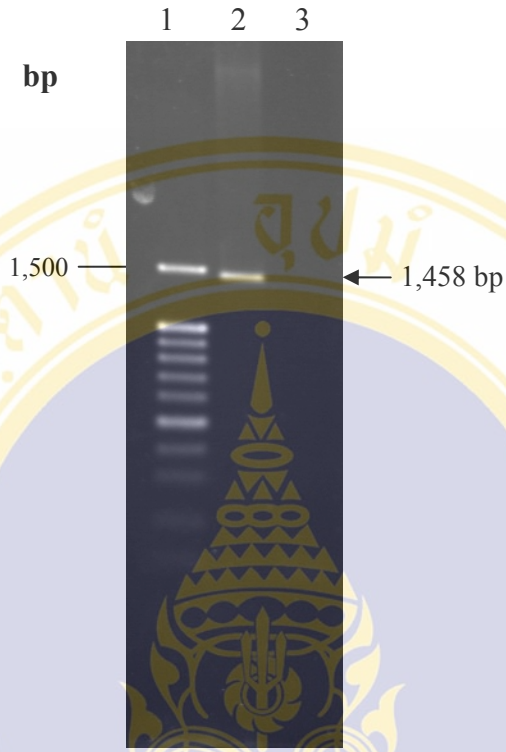


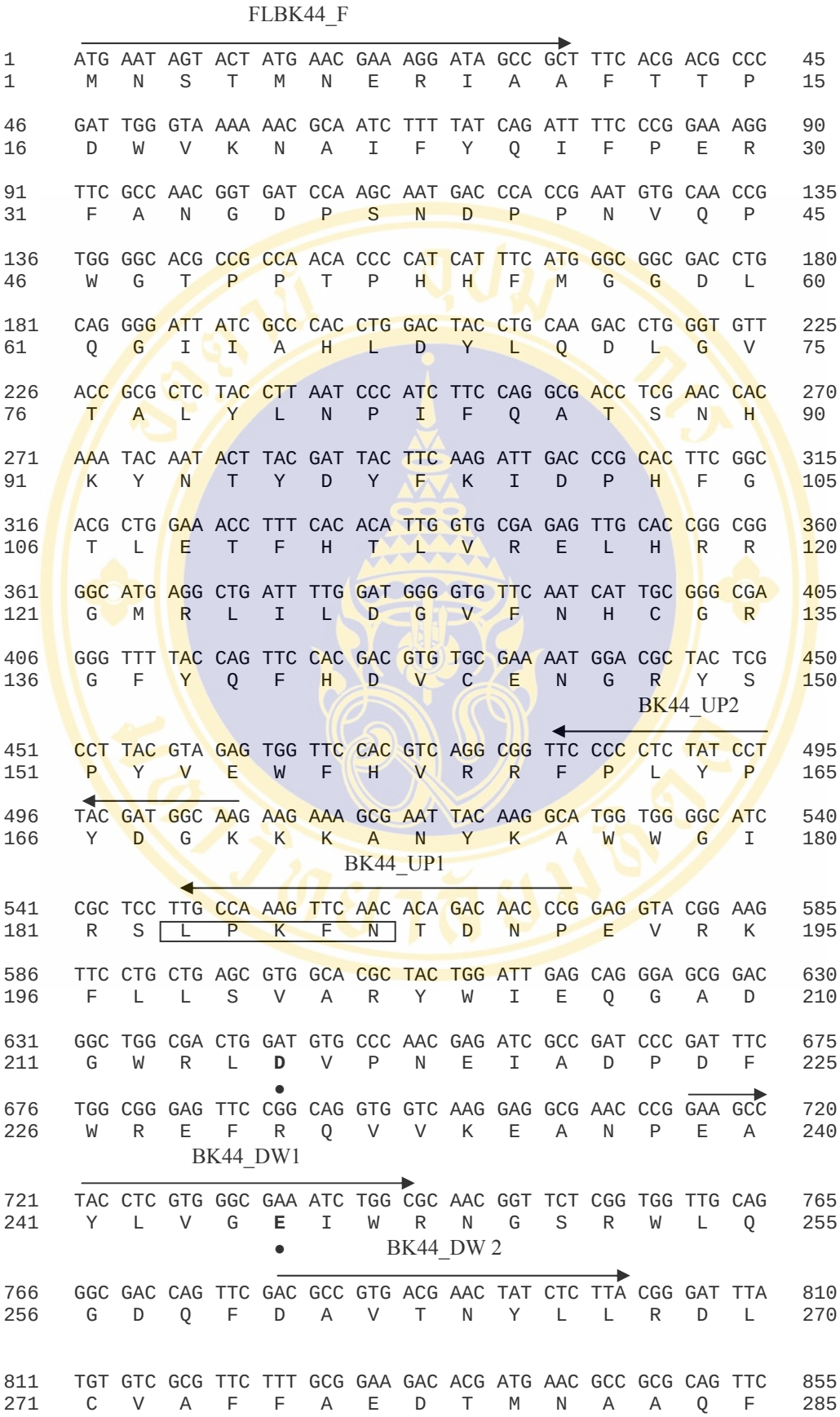
Figure 35. PCR amplification of the full length BK44

The PCR product of BK44 using FLBK44_F and FLBK44_R primers to obtain the full-length gene and analyzed on 1% agarose gel electrophoresis with stained ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : The amplified PCR product of full length BK44

Lane 3 : Negative control of PCR amplification (without template)




```
1     ATG AAT AGT ACT ATG AAC GAA AGG ATA GCC GCT TTC ACG ACG CCC   45
1      M   N   S   T   M   N   E   R   I   A   A   F   T   T   P    15

46    GAT TGG GTA AAA AAC GCA ATC TTT TAT CAG ATT TTC CCG GAA AGG   90
16     D   W   V   K   N   A   I   F   Y   Q   I   F   P   E   R    30

91    TTC GCC AAC GGT GAT CCA AGC AAT GAC CCA CCG AAT GTG CAA CCG   135
31     F   A   N   G   D   P   S   N   D   P   P   N   V   Q   P    45

136   TGG GGC ACG CCG CCA ACA CCC CAT CAT TTC ATG GGC GGC GAC CTG   180
46     W   G   T   P   P   T   P   H   H   F   M   G   G   D   L    60

181   CAG GGG ATT ATC GCC CAC CTG GAC TAC CTG CAA GAC CTG GGT GTT   225
61     Q   G   I   I   A   H   L   D   Y   L   Q   D   L   G   V    75

226   ACC GCG CTC TAC CTT AAT CCC ATC TTC CAG GCG ACC TCG AAC CAC   270
76     T   A   L   Y   L   N   P   I   F   Q   A   T   S   N   H    90

271   AAA TAC AAT ACT TAC GAT TAC TTC AAG ATT GAC CCG CAC TTC GGC   315
91     K   Y   N   T   Y   D   Y   F   K   I   D   P   H   F   G   105

316   ACG CTG GAA ACC TTT CAC ACA TTG GTG CGA GAG TTG CAC CGG CGG   360
106    T   L   E   T   F   H   T   L   V   R   E   L   H   R   R   120

361   GGC ATG AGG CTG ATT TTG GAT GGG GTG TTC AAT CAT TGC GGG CGA   405
121    G   M   R   L   I   L   D   G   V   F   N   H   C   G   R   135

406   GGG TTT TAC CAG TTC CAC GAC GTG TGC GAA AAT GGA CGC TAC TCG   450
136    G   F   Y   Q   F   H   D   V   C   E   N   G   R   Y   S   150

451   CCT TAC GTA GAG TGG TTC CAC GTC AGG CGG TTC CCC CTC TAT CCT   495
151    P   Y   V   E   W   F   H   V   R   R   F   P   L   Y   P   165

496   TAC GAT GGC AAG AAG AAA GCG AAT TAC AAG GCA TGG TGG GGC ATC   540
166    Y   D   G   K   K   K   A   N   Y   K   A   W   W   G   I   180

541   CGC TCC TTG CCA AAG TTC AAC ACA GAC AAC CCG GAG GTA CGG AAG   585
181    R   S   L   P   K   F   N   T   D   N   P   E   V   R   K   195

586   TTC CTG CTG AGC GTG GCA CGC TAC TGG ATT GAG CAG GGA GCG GAC   630
196    F   L   L   S   V   A   R   Y   W   I   E   Q   G   A   D   210

631   GGC TGG CGA CTG GAT GTG CCC AAC GAG ATC GCC GAT CCC GAT TTC   675
211    G   W   R   L   D   V   P   N   E   I   A   D   P   D   F   225

676   TGG CGG GAG TTC CGG CAG GTG GTC AAG GAG GCG AAC CCG GAA GCC   720
226    W   R   E   F   R   Q   V   V   K   E   A   N   P   E   A   240

721   TAC CTC GTG GGC GAA ATC TGG CGC AAC GGT TCT CGG TGG TTG CAG   765
241    Y   L   V   G   E   I   W   R   N   G   S   R   W   L   Q   255

766   GGC GAC CAG TTC GAC GCC GTG ACG AAC TAT CTC TTA CGG GAT TTA   810
256    G   D   Q   F   D   A   V   T   N   Y   L   L   R   D   L   270

811   TGT GTC GCG TTC TTT GCG GAA GAC ACG ATG AAC GCC GCG CAG TTC   855
271    C   V   A   F   F   A   E   D   T   M   N   A   A   Q   F   285
```

```
856   GGC AGG AAA GTC GTC CGG TTG CTG CGT CTT TAC GAT CCC GAA GTA   900
286    G   R   K   V   V   R   L   L   R   L   Y   D   P   E   V    300

901   ACC CAT GCT TTA TTC AAC CTG TTG GGA AGC CAC GAT ACT GCC CGC   945
301    T   H   A   L   F   N   L   L   G   S   H   D   T   A   R    315
                                        ●
946   TTC CTG ACC GTT GCT GGG GAT GAG GTG GAA CGG GTC AAA CTC GCC   990
316    F   L   T   V   A   G   D   E   V   E   R   V   K   L   A    330

991   TTC ACC TTC CTG ATG ACC TAT CCG GGC GCT CCC TGC CTT TAC TAC   1035
331    F   T   F   L   M   T   Y   P   G   A   P   C   L   Y   Y    345

1036  GGC GAT GAG ATT GGC ATG AAG GGC GCG AAG GAC CCC CAC AAC CGC   1080
346    G   D   E   I   G   M   K   G   A   K   D   P   H   N   R    360

1081  GCC TGC TTC CCC TGG GAC GAA GGG CAA TGG AAC AAA GAA CTG CAG   1125
361    A   C   F   P   W   D   E   G   Q   W   N   K   E   L   Q    375

1126  GCA CAC GTG AAA GCC CTG ATC GCC CTG CGC AAG AAG CAC GCT GCC   1170
376    A   H   V   K   A   L   I   A   L   R   K   K   H   A   A    390

1171  CTC CGC ACC GGG GCT TAC CAA ACC CTG CTG GCT GAT GGA AAG GCA   1215
391    L   R   T   G   A   Y   Q   T   L   L   A   D   G   K   A    405

1216  AAT GTG TAC GCC TTC GCC CGC TGG GAT GCA AGC GCA ACG CTC CTG   1260
406    N   V   Y   A   F   A   R   W   D   A   S   A   T   L   L    420

1261  ATC GCT CTG AAC AAC AGC CCG GAG GCT TGG AGC GGT GTC CTC TCT   1305
421    I   A   L   N   N   S   P   E   A   W   S   G   V   L   S    435

1306  CTC AAA AGC CTT TCT CAA GTT TCA GGT CTC AAG CTT CGG GTT GTT   1350
436    L   K   S   L   S   Q   V   S   G   L   K   L   R   V   V    450

1351  TTC CCT CAA CTC GGC ACT CGA AAC CGG AAA CTC GAA ACC GGT CAA   1395
451    F   P   Q   L   G   T   R   N   R   K   L   E   T   G   Q    465
                                                    ◄──────
1396  CGG ACA GGC CCA TTT GCC ATC CAA TTG GCT CCG CGA AGT GGG GTT   1440
466    R   T   G   P   F   A   I   Q   L   A   P   R   S   G   V    480
                FLBK44_R
      ◄──────────────────────
1441  GTT CTT CAG GTA GTG TAA   1458
481    V   L   Q   V   V   *
```

Figure 36. Nucleotide and deduced amino acid sequences of full-length BK44 from the sediments of Bor Khlueng hot spring

The nucleotide sequence of full-length BK44 contained 1,458 bp that encoded 486 amino acid sequences. The deduced amino acid sequence are shown in a single letter code under the respective codon. Positions of primers are indicated by arrows. Asp 215, Glu 245, and Asp 312 (all ●) were catalytic residues of α-amylase family. The fifth conserved region of neopullulanase subfamily is boxed.

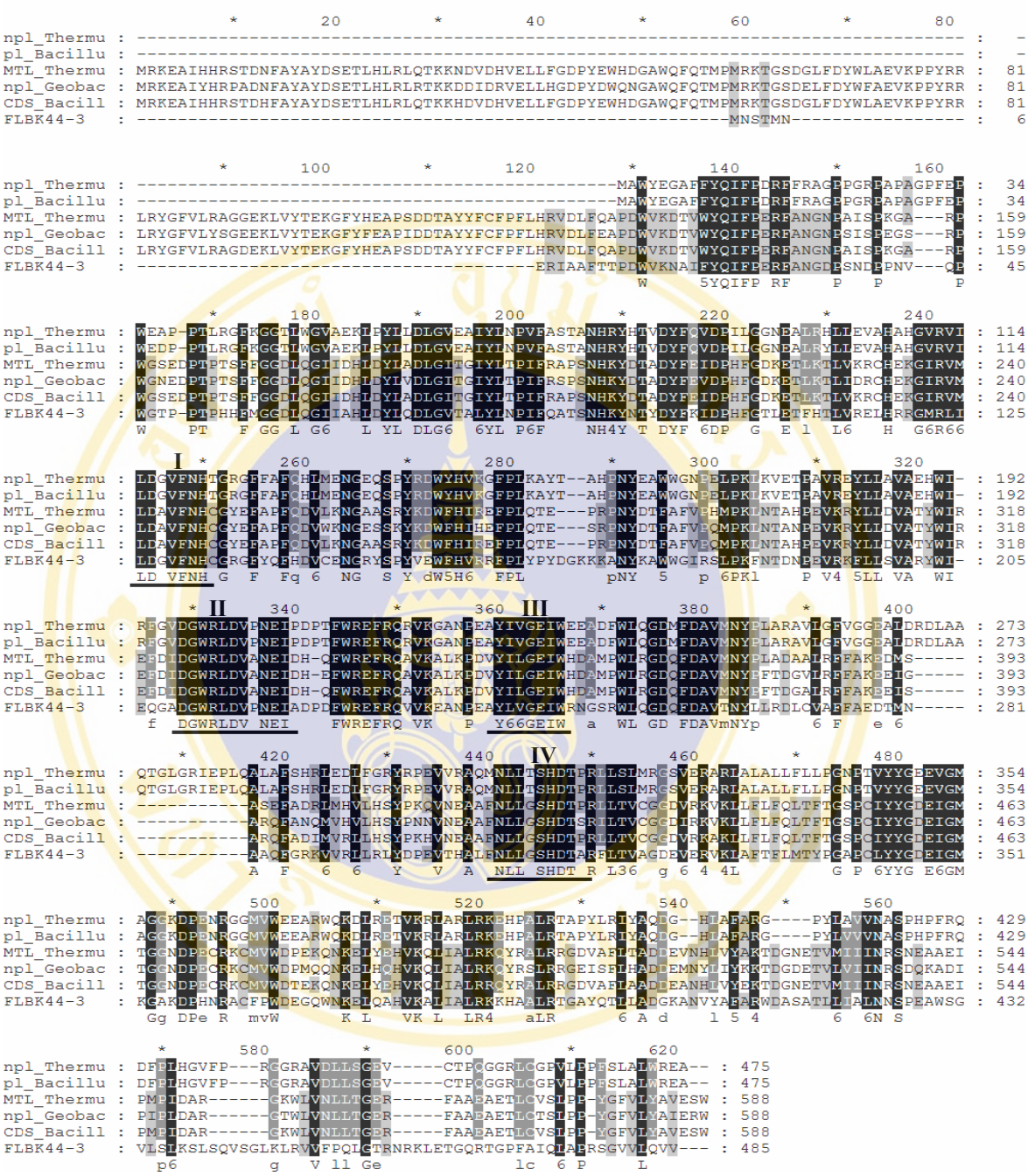


Figure 37. Amino acid sequence alignment of the full-length BK44

The deduced amino acid sequence of BK44 was aligned by Clustal X with other bacterial enzymes in α-amylase family: *Thermus thermophilus* (neo)pullulanase (AP008226), *Bacillus flavocaldarius* pullulanase (AB008764), *Thermus sp.* maltogenic amylase (AF060204), *Bacillus stearothermophilus* neopullulanase (AF233372), *Bacillus stearothermophilus* alpha-cyclodextrinase (AB070710). Four conserved regions of α-amylase family were also shown.

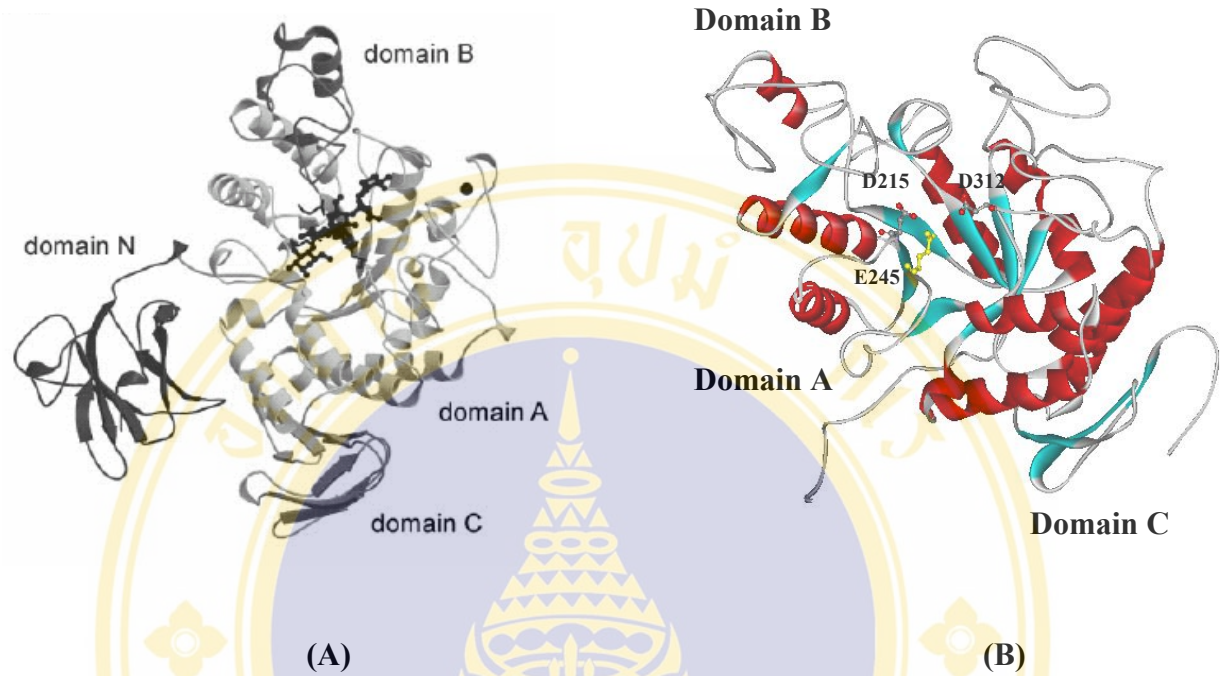


Figure 38. 3D structure prediction of BK44

(A) Representative of a monomeric structure of neopullulanase from *Thermoactinomyces vulgaris* R-47 (TVA II) (106) (B) Predicted tertiary structure of BK44 was performed using An Automated Comparative Protein Modelling Server, SWISS-MODEL in www.expasy.org/swissmod/SWISS-MODEL.html. Three catalytic residues, Asp 215, Glu 245 and Asp 312, are shown as ball and models.

5.7 Construction of a recombinant plasmid harboring BK44 for expression in *P. pastoris*

To express BK44 in *P. pastoris*, pICBK44_F and pICBK44_R primers (Table 3) were used to amplify the target gene. The PCR products were double digested with *Xba*I and *Xho*I, purified and cloned into pPICZαA which had been digested with *Xho*I and *Xba*I. As a result, this BK44 gene was fused in-frame with α-factor secreted signal of pPICZαA vector. Ligation mixture was transferred into *E. coli* for plasmid amplification. Rapid size screening and restriction enzyme analysis were performed to screen for the presence of recombinants clones (Figure 39). DNA sequencing analysis revealed that the target fragment was inserted at the expected site and fused in-frame with the α-factor secretion signal. Moreover, no mutation was observed in the sequence of BK44 within the leader sequence including the fusion region. This recombinant plasmid was designated as pPIC-BK44.

5.8 Integration of BK44 into *P. pastoris* genome

The recombinant plasmids, pPIC-BK44 was linearized with *Dra*I and the linearized plasmid was then electroporated into *P. pastoris* KM71. Transformants were selected on YEPD supplemented with Zeocin™. Genomic DNA was extracted from Zeocin™-resistance transformants. The integration of BK44 was confirmed by PCR using 5' AOX and 3' AOX primers. The result showed the DNA fragment of approximately 2 kb which was obtained from recombinant *P. pastoris* harboring full-length gene. On the other hand, a band of 600 bp was observed from *P. pastoris* transformants harboring only pPICZαA vector (Figure 40). These results suggested that the linearized recombinant plasmid harboring BK44 full-length gene was successfully integrated into the *P. pastoris* genome.

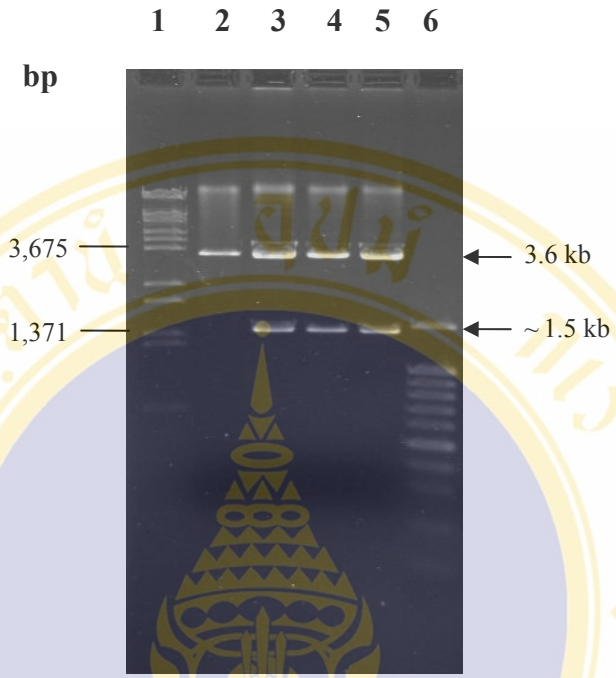


Figure 39. Restriction enzyme analysis of pPIC-BK44

The recombinant plasmids were double digested with *Xho*I and *Xba*I and the digested products were analyzed by 1% agarose gel electrophoresis. DNA fragment of approximately 1.5 kb representing the full length gene of BK44 fragment was observed (clones 1-3).

Lane 1 : Lamda DNA/*Bst*EII marker

Lane 2 : pPICZαA

Lanes 3-5 : Recombinant plasmids with *Xho*I and *Xba*I restriction enzyme digestion.

Lane 6 : 100 bp ladder with 1.5 kb plus

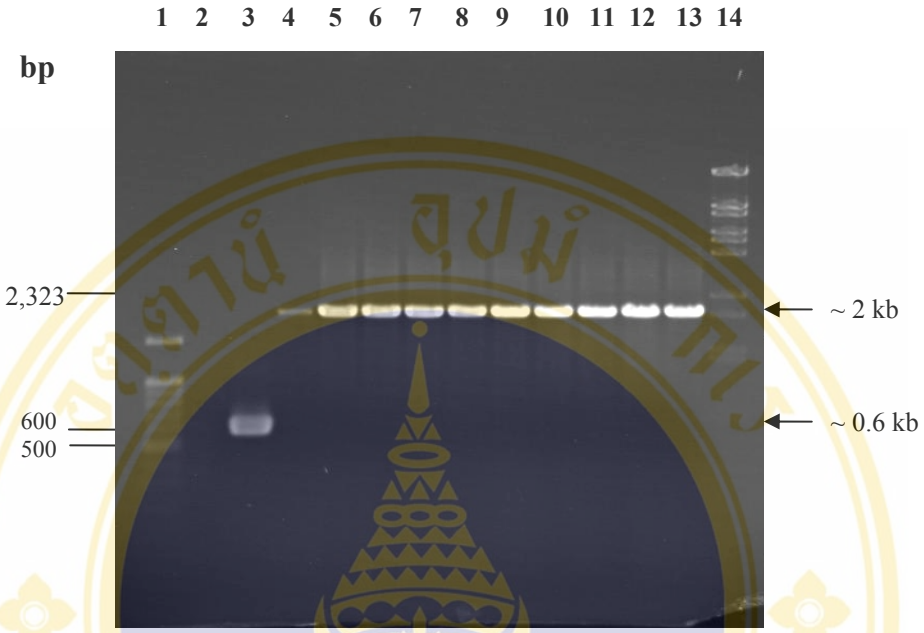


Figure 40. PCR analysis of *P. pastoris* integrants

PCR screening for genome integration using 5'AOX and 3' AOX primers. PCR products were analyzed on 0.8% agarose gel electrophoresis and the gel was stained with ethidium bromide.

Lane 1 : 100 bp ladder with 1.5 kb plus

Lane 2 : Negative control for PCR amplification (without template)

Lane 3 : The amplified product of *Pichia* integrant containing pPICZαA

Lanes 4-13 : The amplified product of *Pichia* integrants containing BK44 recombinant clones 1-10, respectively

Lane 14 : Lamda DNA/*Bst*EII marker

5.9 Expression of BK44 in *P. pastoris*

To investigate the expression of BK44 in *P. pastoris* as a secreted protein, cell cultures were grown in media and induced with methanol as described in Method section. The secreted protein from each day was then analyzed by SDS-PAGE. However, the target protein of approximately 55 kDa was not detected. Therefore, more *Pichia* integrants (clones 2-10) were used to investigate the presence of the secreted protein. The supernatant of each clone was collected after induction for 3 days with methanol. However, the secreted protein of 55 kDa was still not detected from all cultures (Figure 41). Enzymatic activity was also investigated. To determine the cyclodextrinase activity, the supernatant was incubated with α, β, γ cyclodextrin and enzyme activity was determined using DNS method. However, no cyclodextrinase activity was detected from the fractions (data not shown).

One possibility that the protein could not be detected as a secreted protein might be because this protein was located intracellularly. Therefore, to investigate the presence of the intracellular protein, cell lysate of *P. pastoris* was prepared using acid glass bead and the proteins were analyzed by SDS-PAGE. However, the expected protein (55 kDa) was not observed (Figure 42).

5.10 Investigation of BK44 transcript in *P. pastoris* by RT-PCR

As the target protein could not be detected both extracellularly and intracellularly, the transcripts of BK44 gene were then investigated to confirm the presence of mRNA in *P. pastoris*. Total RNA was isolated from *P. pastoris*. The quality and quantity of RNA were determined by spectrophotometry. An absorbance ratio ($A_{260}/_{280}$) of RNA sample was in the range of 1.8-2.0 suggesting that good quality of RNA was obtained. RT-PCR was performed and the contamination of DNA was avoided by treating the total RNA with DNaseI (Fermentas). BK44 transcripts were amplified using FLBK44_F and FLBK44_R primers. The result showed that BK44 transcripts could be detected from culture after induction for 0,1,2,3, and 5 days. This transcript was, however, not detected from *P. pastoris* containing only pPICZαA and non-induced culture (Figure 43). This result confirmed that BK44 was expressed at transcriptional level in *P. pastoris*

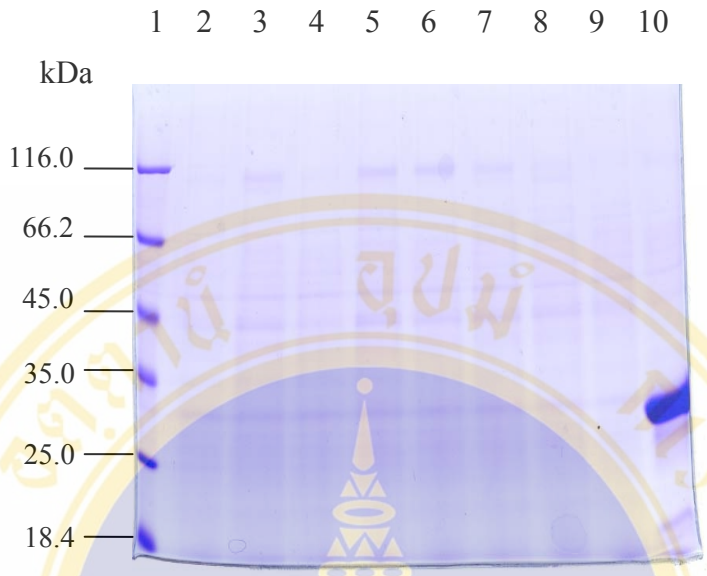


Figure 41. SDS-PAGE analysis of secreted proteins from *P. pastoris* integrants

Cultures were grown in BMGY and induced in BMMY supplemented with 3% methanol (v/v). The culture supernatant was collected after induction for 3 days. 40 µl of supernatant was loaded into each lane and analyzed on 12% SDS-PAGE.

Lane 1 : Protein molecular weight marker (Fermentas).

Lane 2 : *P. pastoris* containing pPICZαA vector after induction for 3 days.

Lanes 3-9 : Culture supernatant from recombinant clones 2-7 and 9.

Lane 10 : Xylanase positive control.

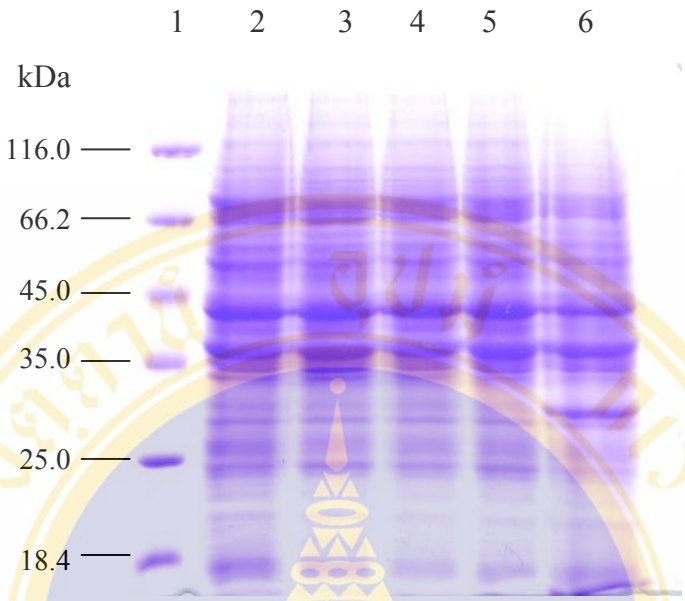


Figure 42. SDS-PAGE analysis of intracellular proteins from *P. pastoris* harboring pPIC-BK44

P. pastoris was lysed by acid glass bead, then the supernatant was analyzed on 12% SDS-PAGE and the gel was stained by Coomassie Blue

Lane 1 : Protein molecular weight marker (Fermentas)

Lane 2 : The cell lysis of *P. pastoris* containing pPICZαA

Lanes 3-5 : The cell lysis of *P. pastoris* integrants #1-3

Lane 6 : The cell lysis of xylanase used as a positive control

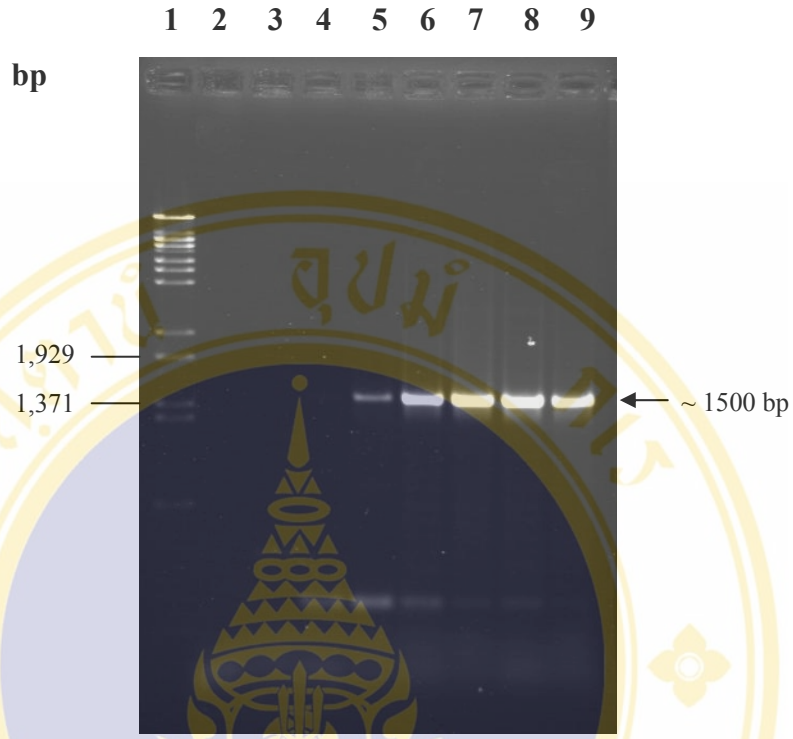


Figure 43. RT-PCR products of the full-length BK44 from *P. pastoris*

Total RNA was extracted from *P. pastoris*. RT-PCR was performed to detect recombinant gene expression at transcriptional level in *P. pastoris* using FLBK44_F and FLBK44_R primers. The PCR product was detected by 0.8% agarose gel electrophoresis and the gel was stained with ethidium bromide.

Lane 1 : Lamda DNA/*Bst*EII marker

Lane 2 : Negative control for amplification (without cDNA).

Lane 3 : The amplification from cDNA of *P. pastoris* containing pPICZαA.

Lane 4 : The amplification from cDNA of non-induced cell.

Lanes 5-9 : The amplification from cDNA of induced cells day 0, 1, 2, 3 and 5

5.11 Expression of BK44 in *E. coli*

5.11.1 Construction of BK44 in bacterial expression vector (pET-32a)

Since BK44 could not be obtained using *P. pastoris* expression system, *E. coli* expression was exploited. BK44 was amplified using pETBK44_F and pETBK44_R primers (Table 3) and the expected DNA fragment of 1.5 kb was obtained. The full-length gene (BK44) was then ligated to pET-32a (Novagen) which had been digested with *Nco*I and *Eco*RI and transformed into *E. coli* DH5α. The recombinant plasmids were analyzed by *Nco*I and *Eco*RI double digestions. Digested product of approximately 1.5 kb was observed from recombinant clones that contain insert fragment (Figure 44). These recombinant clones were subjected to DNA sequencing. Sequence analysis indicated that BK44 was fused in-frame with Trx-Tag in pET-32a. The recombinant plasmid was designated as pET32-BK44.

5.11.2 Expression of BK44 in *E. coli* using pET-32a

To express BK44 in *E. coli*, pET32-BK44 was transformed into *E. coli* strain Rosetta-gami(DE3)pLysS to enhance the protein expression of the gene that contains codons rarely used in *E. coli* (135) and transformants were selected on ampicillin and chloramphenical containing-LB agar plates. To screen for the expressed target protein, recombinant *E. coli* was grown in ampicillin and chloramphenical containing-LB medium. After IPTG induction, crude lysate was analyzed on 12% SDS-PAGE. The expected protein of 55 kDa was detected. Sonication was then performed to separate the soluble and pellet fractions. The result showed the expected protein of BK44, which was mainly present in the pellet fraction (Figure 45). To improve the solubility of the protein, the culture was induced at lower temperature (25 ºC) with 0.1 mM IPTG for 2-4 h. However, the 55 kDa BK44 was still mostly formed as inclusion bodies (Figure 46). Nonetheless, the enzymatic activity was also investigated. Specifically, to determine pullulanase activity, both soluble and pellet fractions were applied into each well of agar plate containing AZCL-pullulan and incubated at 50 ºC for 18 h. The result showed that the blue zone was developed only from the standard pullulanase using as positive control. No blue zone was developed from any test fractions. To determine the cyclodextrinase activity, the soluble fractions were incubated with α, β, γ cyclodextrin and enzyme activity was determined using DNS method. However, no cyclodextrinase activity was detected (data not shown).

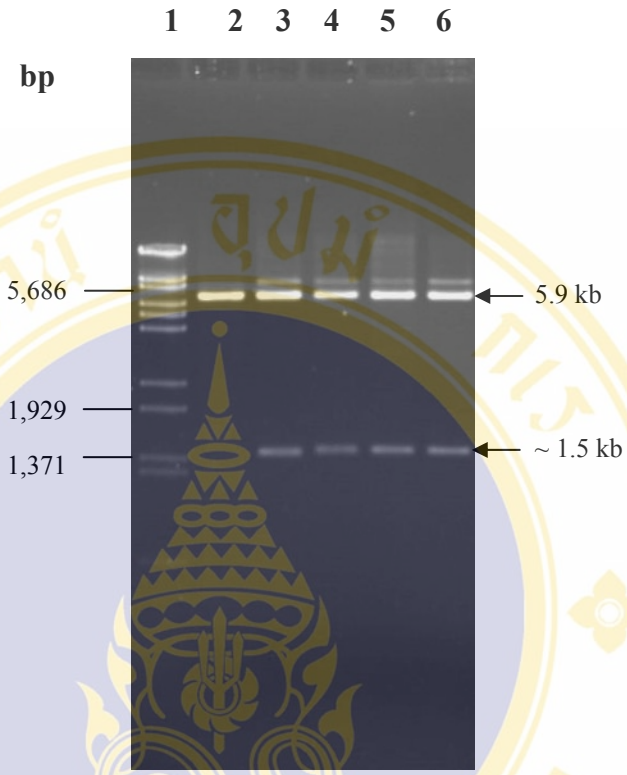


Figure 44. Restriction enzyme analysis of pET32-BK44

The recombinant plasmids were double digested with *Nco*I and *Eco*RI. The digested products were analyzed by 1% agarose gel electrophoresis. DNA fragment of approximately 1.5 kb representing the full length gene of BK44 fragment was observed in recombinant clones (clones 1-4).

Lane 1 : Lamda DNA/*Bst*EII marker

Lane 2 : pET-32a

Lanes 3-6 : Recombinant plasmids with *Nco*I and *Eco*RI restriction enzyme digestion.

S= soluble fraction

P= pellet fraction

Figure 45. SDS-PAGE analysis of BK44 expression in *E. coli* at 30 ºC

BK44 was expressed in *E. coli* at 30 ºC with 0.1 mM IPTG induction for 4 h. The soluble and pellet fractions were analyzed on 12% SDS-PAGE and visualized by Coomassie Blue staining.

Lane 1 : Protein molecular weight marker (Fermentas).

Lane 2 : Soluble fraction from *E. coli* containing only pET-32a.

Lane 3 : Pellet fraction from *E. coli* containing only pET-32a.

Lanes 4, 6, 8 : Soluble fractions from *E. coli* recombinant plasmid, pET32-BK44

Lanes 5, 7, 9 : Pellet fractions from *E. coli* recombinant plasmid, pET32-BK44

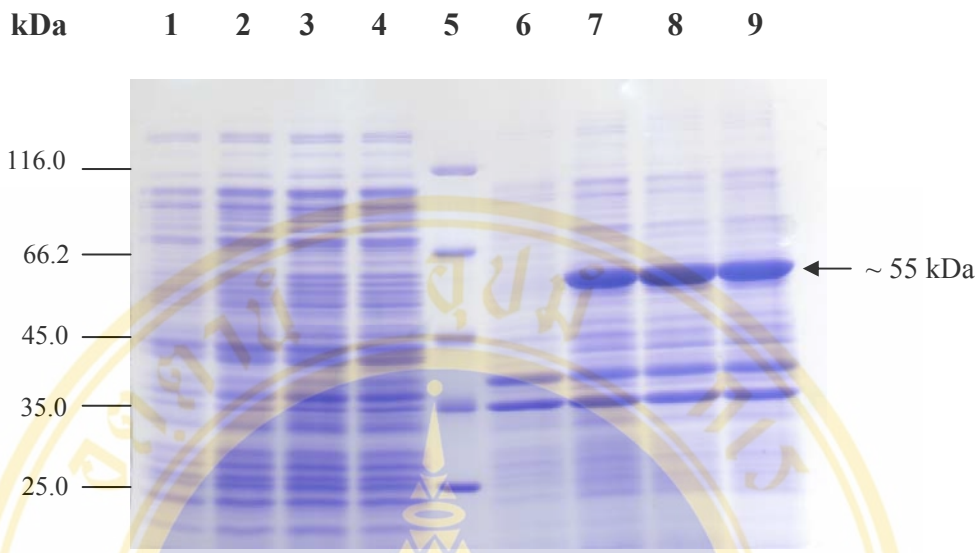


Figure 46. SDS-PAGE analysis of BK44 in *E. coli* at 25 ºC at 0-4 h

BK44 was expressed in *E. coli* at 25°C with 0.1 mM IPTG induction for 0-4 h. The soluble and pellet fractions were analyzed on 12% SDS-PAGE and visualized by Coomassie Blue staining.

Lanes 1-4 : Soluble fractions of BK44 that collected at 0, 2, 3 and 4 h

Lane 5 : Protein molecular weight marker (Fermentas)

Lanes 6-9 : Pellet fractions of BK44 that collected at 0, 2, 3 and 4 h

5.11.3 Expression of BK44 in *E. coli* using pET-43.1a

In order to obtain soluble protein, another expression vector, pET-43.1a was used. This vector contained the NusA gene, which has been shown to increase the solubility of several heterologous proteins (136). As a result, BK44 was amplified using pET43.1BK44_F and pET43.1BK44_R primers (Table 3). The obtained DNA fragment of 1.5 kb was ligated to pET43.1a which had been digested with *Bam*HI and *Kpn*I and transformed into *E. coli* DH5α. Restriction enzyme digestion showed that BK44 was fused in-frame with Nus-Tag in pET-43.1a (Figure 47). *E. coli* transformants harboring recombinant plasmid (pET43.1-BK44) were grown at 30 ºC and induced with IPTG for 4 h. Sonication was performed to separate soluble and pellet fractions. The proteins were then analyzed by SDS-PAGE. The expected protein of approximately 109 kDa was obtained in both soluble and pellet fractions (Figure 48). Next, enzymatic activity was then investigated. To determine pullulanase activity, both soluble and pellet fractions were applied into each well of agar plate containing AZCL-pullulan and incubated at 50 ºC for 18 h. The result showed that the blue zone was developed from only the standard pullulanase using as positive control. No blue zone was developed from any fractions (Figure 49). To determine the cyclodextrinase activity, the soluble fractions were incubated with α, β, γ cyclodextrin and enzyme activity was determined using DNS method. However, no cyclodextrinase activity was detected from the fractions (data not shown).

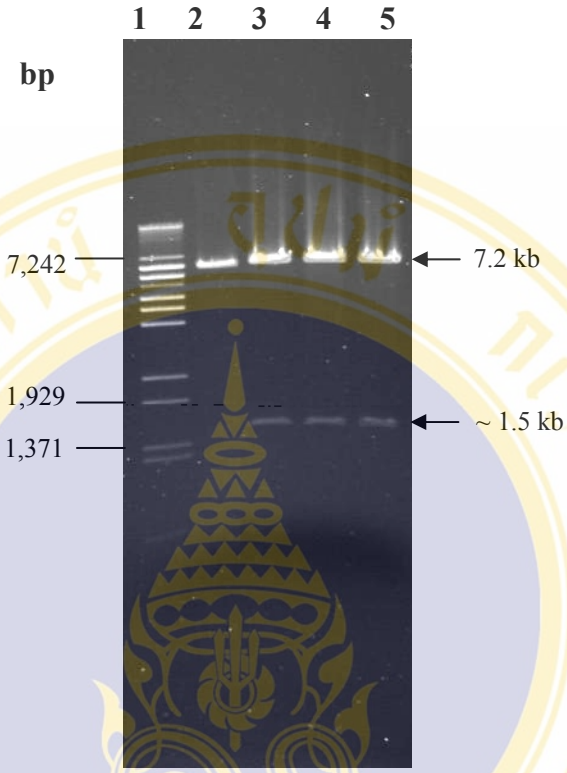


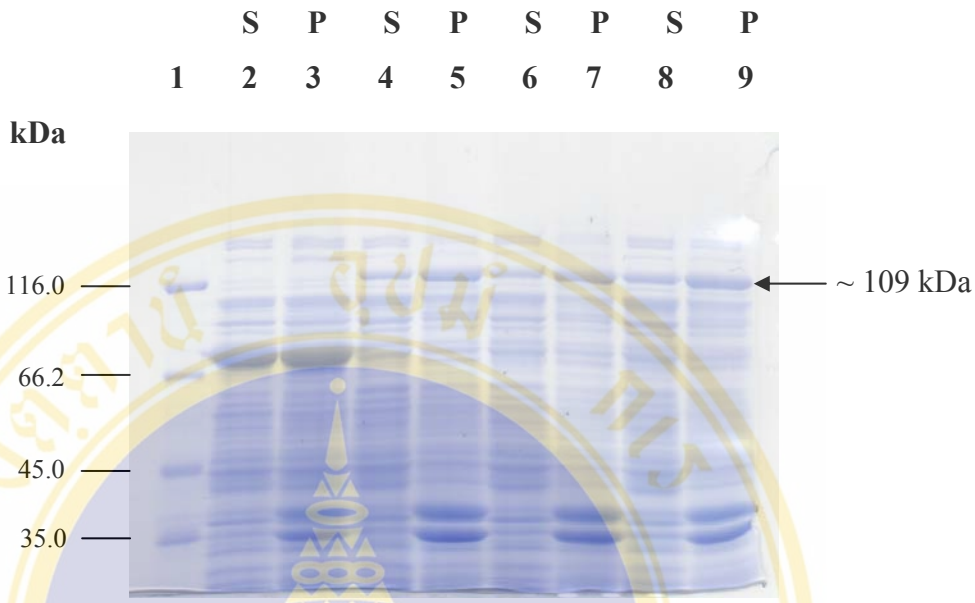
Figure 47. Restriction enzyme digestion analysis of pET43.1-BK44

The recombinant plasmids were double digested with *Bam*HI and *Kpn*I. The digested products were analyzed by 0.8% agarose gel electrophoresis.

Lane 1 : Lamda DNA/*Bst*EII marker

Lane 2 : pET-43.1a digested with *Bam*HI and *Kpn*I

Lanes 3-5 : Recombinant plasmids with *Bam*HI and *Kpn*I restriction enzyme digestion.



S = soluble fraction

P = pellet fraction

Figure 48. SDS-PAGE analysis of BK44 soluble and pellet fractions produced from *E. coli* harboring pET43.1-BK44

BK44 was expressed at 30 ºC with 0.1 mM IPTG induction for 4 h. The soluble and pellet fractions were analyzed on 8% SDS-PAGE and visualized by Coomassie Blue staining.

Lane 1 : Protein molecular weight marker (Fermentas)

Lane 2 : Soluble fraction from *E. coli* containing pET-43.1a

Lane 3 : Pellet fraction from *E. coli* containing pET-43.1a

Lanes 4, 6, 8 : Soluble fractions from *E. coli* recombinant plasmid, pET43.1-BK44

Lanes 5, 7, 9 : Pellet fractions from *E. coli* recombinant plasmid, pET43.1-BK44

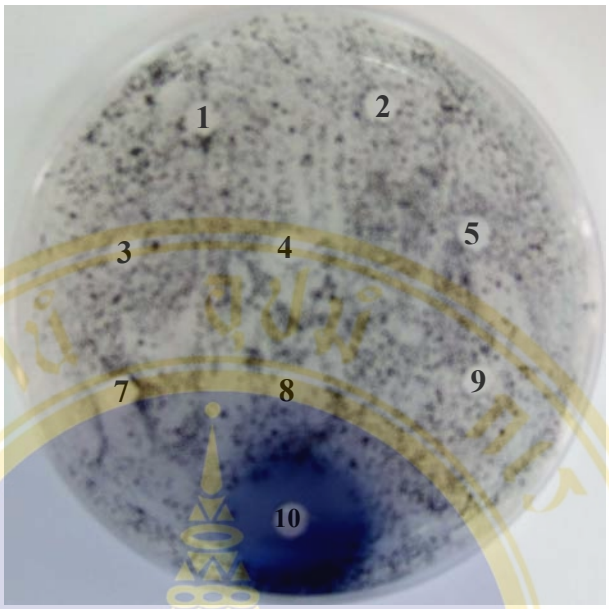


Figure 49. Pullulanase activity by AZCL-pullulan

BK44 was expressed in *E. coli* using pET-43.1a. A 30 µl of soluble and pellet fraction was used to assay the pullulanase activity on AZCL-pullulan plate. Wells 1-2 : the soluble fraction of *E. coli* containing pET-43.1a after expression at 30 and 37 ºC, respectively. Wells 3-5 : the soluble fractions of recombinant clone after expression at 30 and 37 ºC, respectively. Wells 7-9 : the pellet fractions of recombinant clone after expression at 30 and 37 ºC, respectively. Well 10 : the standard pullulanase (Sigma, USA).

CHAPTER VI
DISCUSSION

6.1 Molecular cloning of family 10 and family 13 glycosyl hydrolases

In order to retrieve partial gene of family 10 xylanase and family 13 glycosyl hydrolase directly from the sediment of Bor Khlueng hot spring, degenerate primers based on the conserved regions of both families were designed and genome walking PCR was performed to obtain full-length genes. Genome walking PCR was successfully used to isolate xylanases from both culturable and nonculturable microbes such as *D. thermophilum* XynB (15), *Caldibacillus cellulovorans* XynA (111) and noncultured XynA (112). In this work, only the downstream sequence of xylanase was successfully obtained. One of the reasons that the upstream sequence could not be recovered by this approach may be because the restriction sites used to generate genomic libraries were too far from the site at which the gene specific primer was bound. This resulted in a longer PCR product which could not be amplified. Although xylanase gene could not be obtained, PCR-based cloning was shown to be successfully used for isolation of the gene encoding neopullulanase. This gene, BK44, has an open reading frame of 1,458 bp encoding 458 amino acid residues with a calculated mass of 55 kDa. The primary sequence of BK44 contained the four highly conserved regions similar to other enzymes in the α-amylase family, including three catalytic residues, Asp, Glu, and Asp. This enzyme exhibited 53% identity to neopullulanase subfamily: maltogenic amylase, cyclomaltodextrinase and neopullulanase which share 40-60% sequence identity (89) and distinct substrate specificity from other family 13 members. Furthermore, the amino acid sequence of BK44 showed the fifth conserved region (183-LPKFN) as described for the specific sequence of neopullulanase subfamily (137) (Figure 36).

From the primary amino acid sequence, this novel enzyme did not contain a signal sequence suggesting that it was an intracellular protein. This is similar to most cyclomaltodextrinases, such as *Thermus* maltogenic amylase (90) and *Alicyclobacillus*

acidocaldarius cyclomaltodextrinase (102). Therefore, it is proposed that this enzyme functions in the hydrolyzation of small oligosaccharides including cyclodextrins which can be imported into the cells (90).

3D crystal structure of α-amylase family such as Taka-amylase A (TAA) (60), *Bacillus stearothermophilus* neopullulanase (134), *Thermus* maltogenic amylase (ThMA) (90), *Thermoacitinomyces vulgaris* R-47 neopullulanase-like amylase enzyme (TVA I and TVA II) (138) have been reported. Enzymes belonged to the neopullulanase subfamily contain an additional N-terminal domain, located at the N-terminus of TIM barrel within the sequences of residues 1 to 124 in *Thermus* maltogenic amylase (90), 1 to 123 in *B. stearothermophilus* neopullulanase (134) and 1 to 123 in *Bacillus* sp. I-5 cyclomaltodextrinase (89), respectively. N-domain was shown to participate in dimer formation in *Thermus* maltogenic amylase (90) and play a role in the hydrolysis of cyclodextrins (CDs) (57). Recently, there is a study showing that the dimerization is important for higher activity in CDs hydrolysis of *Anoxybacillus* CDase (*Af*Cda13). This was confirmed by the fact that monomeric enzymes of *Laceyella* CDase (*Ls*Cda13) are capable of cyclodextrin degradation but with lower activity (139). 3D structure prediction of BK44 revealed that this enzyme lacked the N-domain. CDase from *Thermotoga maritima* (TMG) was also found lacking this domain. CDase from TMG exhibited enzymatic activity not only for cyclodextrins but also starch and maltooligosaccharides (140). Therefore, it is possible that BK44 might exhibit similar weak activity against various substrates with cyclodextrins being the most favorable ones. In addition, it seems likely that this BK44 is a monomeric enzyme.

6.2 Expression of BK44 in *P. pastoris*

In this study, *P. pastoris* was chosen as the expression host. BK44 was subcloned into *P. pastoris* expression vector, pPICZαA, which contained the MFα-1 signal sequence that directed the secretion of the heterologous protein. However, protein from BK44 could not be obtained from *P. pastoris* either as intracellular or extracellular protein. This might be the result of gene mutation, mRNA secondary structure or codon usage (141). Sequence analysis of *Pichia* integrants using 5' AOX and 3' AOX primers revealed no mutation in the sequence of BK44. Therefore, BK44

transcript was examined to confirm gene expression at transcriptional level. From RT-PCR result, BK44 transcript could be observed from culture, which was induced for 0,1,2,3, and 5 days. The detection of BK44 transcript from *P. pastoris* induced at day 0 might result from the fact that *P. pastoris* could immediately and rapidly use methanol as an inducer. This transcript was not detected from *P. pastoris* containing only pPICZαA and non-induced culture. In addition, the mRNA secondary structure could also have a major impact on the translational efficiency (142). Therefore, mRNA secondary structure of BK44 which contained many stems-loops at the beginning of mRNA sequence might interfere the translation (Figure 50).

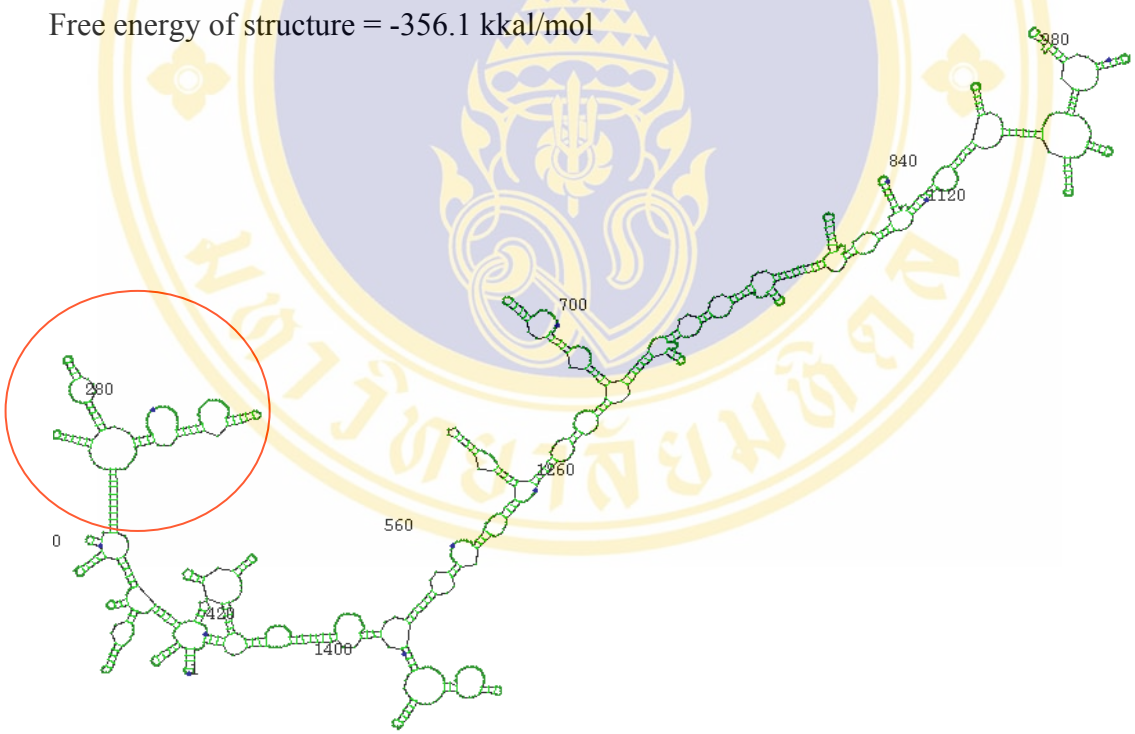


Figure 50. mRNA secondary structure prediction of BK44

Predicted mRNA secondary structure of BK44 was performed using GeneBee-NET server in www.genebee.msu.su. In the circle indicates stems-loops at the beginning of mRNA sequence.

6.3 Expression of BK44 in *E. coli*

As BK44 expression in *P. pastoris* could not be obtained. BK44 was then expressed in *E. coli*. From previous studies, the expression of pullulanase genes in *E. coli* have been reported. These include *Alicyclobacillus acidocalsarius* cyclomaltodextrinase (neopullulanase) (102), *Thermus thermophilus* HB8 pullulanase (143), *Flavobacterium* sp. cyclodextrinase (74) and *Bacillus stearothermopilus* ET1 maltogenic amylase (100). In this study, the expression was performed using pET-32a as an expression vector. BK44 protein of 55 kDa was obtained mostly as inclusion bodies when grown at 25 and 30 ºC but when the temperature was decreased to 18 ºC, no protein of the expected size was detected.

Although *E. coli* is widely used for heterologous protein expression, some proteins are found to have difficulty expressing as a functional soluble form in this host (144). This might be due to mRNA stability, codon bias, protein folding and secretion. CDase from *Thermotoga maritime* or TMG, is highly expressed in *E. coli*; however, it is found that half of the recombinant TMG existed in the insoluble fraction (140). The reason for this is not clear but it might be because of the structure complexity of the enzyme as it contains four different domains (145). There are several methods to overcome these problems i.e. unfolded and refolded protein (146) temperature optimization (147), tuning the inducer concentration (148), adjusting the codon usage (149), or fusing target protein to a highly soluble protein (150). In this study, protein fusion with highly soluble protein was performed. BK44 was fused to NusA in pET43.1a. It is claimed that this protein enhances the solubility of protein because the predicted solubility of NusA was 95% (150). After fusion with NusA, BK44 was obtained both in the soluble and pellet fractions. However, the enzyme activity of BK44 could not be observed when α, β, γ cyclodextrins and pullulan were used as substrates. Turner *et al* (148) claimed that the presence of the fusion protein decreases the activity of enzyme. Consequently, the removal of Nus-Tag with specific enzyme digestion, enterokinase, should be performed to obtain the enzyme activity. NusA, fusion tag was also used in *Anoxybacillus falvithermus* cyclomaltodextrinase (*Afcda13*) (148). When Nus-Tag was removed by digestion with enterokinase, the activity increases two folds when compared to the untreated fusion protein (148).

According to undetected activity of the enzyme from the soluble fraction of pET43.1-BK44, the lack of the additional N domain might affect the enzymatic activity. The deletion of the N domain in neopullulanases from *Thermoactinomyces vulgaris* (TVA I and TVA II) was investigated. The activities of the enzyme lacking N domain against soluble starch and cyclodextrins were drastically decreased about 1,500 to 10,000 fold when compared with those of wild type enzymes (57). For *Thermus* maltogenic amylase (ThMA), the deletion of N-terminal domain resulted in monomeric enzyme and the activity of ThMA to β-cyclodextrin was also decreased (151). Therefore, it is possible that BK44 might exhibit low enzymatic activity against cyclodextrins and pullulan because of the lacking N domain. Another possibility is that the substrates used for assaying the enzyme activity of BK44 was not sensitive enough to detect the low pullulanase or cyclodextrinase activity. Therefore, more sensitive method such as fluorescently labeled pullulan (152) might be needed to assess enzyme activities. In addition, one fact that substrates used in this work might not be the most favorable ones for BK44 should not be excluded. More various substrates could be used to prove this hypothesis.

Lastly, to obtain soluble functional BK44, other alternative heterologous hosts such as *Bacillus subtilis* and yeast, *Saccharomyces cerevisiae* could be used. Successful expression of pullulanase genes have been reported from various microorganisms such as *Bacillus stearothermophilus* neopullulanase (82), *Desulfurococcus mucosus* pullulanase (145) and *Bacillus polymyxa* neopullulanase (104) in such hosts.

CHAPTER VII
CONCLUSIONS

7.1 In this study, an almost complete xylanase gene was obtained from Bor Khlueng hot spring sediments.

7.2 A full-length neopullulanase gene (BK44) was successfully obtained and contained 1,458 bp open reading frame encoding 485 amino acid residues with the calculated mass of 55 kDa.

7.3 The full-length BK44 was classified into family 13 glycosyl hydrolase and belonged to the neopullulanase subfamily.

7.4 The predicted 3D structure of BK44 depicted the $(\beta/\alpha)_8$ fold and C-terminal domain like other enzymes in family 13 glycosyl hydrolase. However, BK44 lacked domain N which is the distinct characteristic for neopullulanase subfamily.

7.5 BK44 transcripts could be detected in recombinant *P. pastoris*, however, BK44 protein could not be obtained.

7.6 BK44 protein was obtained as an inclusion body when expressed in *E. coli* using pET-32a expression vector. On the other hand, both soluble and pellet fractions of the protein could be detected when BK44 was expressed using pET43.1a. However, the pullulanase and cyclodextrinase activities were not detected.

REFERENCES

1. Haki GD, Rakshit SK. Developments in industrially important thermostable enzymes: a review. Biores Technol. 2003; 89: 17-34.
2. Kirk O, Borchert TV, Fuglsang CC. Industrial enzyme applications. Curr Opini Biotechnol. 2002; 13: 345-351.
3. Whistler RL, Richards EL. Hemicellulose. In: The carbohydrates. Academic Press, New York. 1970: 447-469.
4. Timell TE. Wood hemicelluloses. I. Adv Carbohydr Chem. 1964; 19: 247-302.
5. Woodward J. Xylanases: function, properties and application. Top Enzyme Ferment Biotechnol. 1984; 8: 9-30.
6. Dekker RF, Richards GN. Hemicellulases: their occurrence, purification, properties, and mode of action. Adv Carbohydr Chem Biochem. 1976; 32: 277-352.
7. Bouveng HO. Phenylisocyanate derivatives of carbohydrates. II. Location of O-acetyl groups in birch xylan. Acta Chem Scand. 1961; 15: 96-100.
8. Lindberg B, Rosell KG, Svensson S. Position of the O-acetyl groups in birch xylan. Svensk Papperstid. 1973; 76: 30-32.
9. Puls J, Schuseil J. Chemistry of hemicelluloses: relationship between hemicellulose structure and enzyme required for hydrolysis. In: Hemicellulose and hemicellulases. Portland Press, London. 1993: 1-28.
10. Wong, KKY, Tan LUL, Saddler JN. Multiplicity of β-1,4 xylanase in microorganisms: function and applications. Microbiol Rev. 1988; 52: 305-317.
11. Beg QK, Kapoor M, Mahajan L, Hoondal GS. Microbial xylanases and their industrial application: a review. Appl Microbiol Biotechnol. 2001; 56: 326-338.
12. Sunna A, Antranikian G. Xylanolytic enzymes from fungi and bacteria. Crit Rev Biotechnol. 1997; 17: 39-67.

13. Takenishi S, Tsujisaka Y. On the modes of action of three xylanases produced by a strain of *Aspergillus niger* van Tieghem. Agric Biol Chem. 1975; 39: 2315-2323.

14. Wood TM, McCrae SI. Studies of two low-molecular-weight endo-(1,4)-β-D-xylanases constitutively synthesized by the cellulolytic fungus *Trichoderma koningii*. Carbohydr Res. 1986; 148: 321-330.

15. Morris DD, Gibbs MD, Chin CWJ, Koh MH, Wong KKY, Alison RW, Nelson PJ, Bergquist PL. Cloning of the *xynB* gene from *Dictyoglomus thermophilum* strain Rt46B.1 and action of the gene product on kraft pulp. Appl Environ Microbiol. 1998; 64: 1759-1765.

16. John M, Schmidt B, Schmidt J. Purification and some properties of five endo-1,4-β-D-xylanases and a β-D-xylosidase produced by a strain of *Aspergillus niger*. Can J Biochem. 1979; 57: 125-134.

17. Nanmori T, Watanabe T, Shinke R, Kohno A, Kawamura Y. Purification and properties of thermostable xylanase and β-xylosidase produced by a newly isolated *Bacillus stearothermophilus* strain. J Bacteriol. 1990; 172: 6669-6672.

18. Kaji A, Saheki T. Endo-arabinase from *Bacillus subtilis* F-11. Biochim Biophys Acta. 1975; 410: 354-360.

19. Kaji A. L-Arabinosidases. Adv Carbohydr Chem Biochem. 1984; 42: 382-394.

20. Kaji A, Tagawa K. Purification, crystallization, and amino acid composition of α-L-arabinofuranosidase from *Aspergillus niger*. Biochim Biophys Acta. 1970; 207: 456-464.

21. Poutanen K. An α-L-arabinofuranosidase of *Trichoderma reesei*. J Biotechnol. 1988; 7: 271-282.

22. Johnson KG, Harrison BA, Schneider H, Mackenzie CR, Fontana FD. Microbial degradation of hemicellulosic materials. Appl Biochem Biotechnol. 1989; 20/21: 245-258.

23. Biely P, Puls J, Schneider H. Acetyl xylan esterases in fungal xylanolytic systems. FEBS Lett. 1985b; 186: 80-84.

24. Christov LP, Prior BA. Esterases of xylan-degrading microorganisms: production, properties, and significance. Enzyme Microb Technol. 1993; 15: 460-475.

25. Henrissat B. A classification of glycosyl hydrolases based on amino acid sequence similarities. Bioresource Technol. 1991; 89: 17-34.

26. Henrissat B, Davies G. Structural and sequence-based classification of glycosyl hydrolases. Curr Opin Struct Biol. 1997; 7: 637-644.

27. Tomas WJ. Biochemistry and genetics of microbial xylanases. Curr Opin Biotechnol. 1996; 7: 337-342.

28. Gaboriaud C, Bissey V, Benchetrit T, Mornon JP. Hydrophobic cluster analysis: An efficient new way to compare and analyze amino acid sequence. FEBS Lett. 1987; 224: 149-155.

29. Torronen A, Kubicek CP, Herrissat B. Amino acid sequence similarities between low molecular weight endo-1,4-β-xylanase and family H cellulases revealed by clustering analysis. FEBS Lett. 1993; 321: 135-139.

30. Henrissat B, Bairoch A. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J. 1993; 293: 781-788.

31. Henrissat B, Bairoch A. Updating the sequence-based classification of glycosyl hydrolase. Biochem J. 1996; 316: 695-696.

32. Bieley P, Vrsanska M, Tenkanen M, Kluepfel D. Endo-β-1,4 xylanase families: differences in catalytic properties. J Biotechnol. 1997; 57: 151-166.

33. Torronen, A, Harkki A, Rouvinen J. Three dimensional structure of endo-β-1,4-xylanase II from *Trichoderma reesei*: two conformational states in the active site. EMBO J. 1994; 13: 2493-2501.

34. Derewenda U, Swenson L, Green R, Wei Y, Morosoli R, Shareck F, Kluepfel D, Derewenda ZS. Crystal structure, at 2.6 Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of β-1,4-D-glycanases. J Biol Chem. 1994; 269: 20811-20814.

35. Prade RA. Xylanases: from biology to biotechnology. Biotechnol Genet Eng Rev. 1995; 13: 101-131.

36. Katsube U, Hata Y, Yamaguchi H, Moriyama H, Shinmyo A, Okada H. Estimation of xylanase active site from crystalline structure. In: Ikehara M. (Ed). Japan Scientific Societies Press. Japan. 1989: 91-96.

37. Campbell RL, Rose DR, Wakarchuk WW, To R, Sung W, Yaguchi M. A comparison of the structure of the 20 kd xylanases from *Trichoderma*

harzianum and *Bacillus circulans*. In: Suominen P, Reinikainen T. (Eds). Proceedings of the second TRICEL symposium on *Trichoderma* cellulases and other hydrolases. Espoo, Finland, 1993. Foundation for Biotechnical and Industrial Fermentation Research, Helsinki. 1993; 8: 63-72.

38. Wakarchuk W, Campbell RL, Sung WL, Davoodi J, Yaguchi M. Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. Protein Sci. 1994; 3: 467-475.
39. Wong KKY, Saddler JN. *Trichoderma* xylanases, their properties and purification. Crit Rev Biotechnol. 1992; 12: 413-435.
40. Kuhad, RC, Singh A. Lignocellulosic biotechnology: current and future prospects. Crit Rev Biotechnol. 1993; 13: 151-172.
41. Bajpai P. Application of enzymes in the pulp and paper industry. Biotechnol Prog. 1999; 15: 147-157.
42. Bajpai P, Bhardwaj NK, Bajpai PK, Jauhari MB. The impact of xylanases on bleaching of eucalyptus kraft pulp. J Biotechnol. 1994; 38: 1-6.
43. Srinivasan MC, Rele MV. Microbial xylanases for paper industry. Curr Sci. 1999; 77: 137-142.
44. Maat J, Roza M, Verkabel J, Stam H, daSilra MJS, Egmond MR, Hagemans M LD, van Garcom RFM, Hessing JGM, van Derhondel CAMJJ, van Rotterdam C. Xylanases and their application. In: Xylans and xylanases. Elsevier, Amsterdam, 1992: 349-360.
45. Biely P. Microbial xylanolytic systems. Trends Biotechnol. 1985; 3: 286-290.
46. Rani S, Nand K. Development of cellulose-free xylanase producing anaerobic consortia for the use of lignocellulosic wastes. Enzyme MicrobTechnol. 1996; 18: 23-28.
47. Dominguez JM. Xylitol production by free and immobilized *Debaryomyces hansenii*. Biotechnol Lett 1998; 20: 53-56.
48. Bedford MR, Classen HL. The influence of dietary xylanase on intestinal viscosity and molecular weight distribution of carbohydrates in rye-fed broiler chick. In: Xylans and xylanases. Elsevier, Amsterdam. 1992: 361-370.
49. Pandey A, Nigam P, Soccol CR, Soccol VT, Singh D, Mohan R. Advances in microbial amylases. Biotechnol Appl Biochem. 2000; 31: 135-152.

50. Diderichsen B, Christiansen L. Cloning of a maltogenic alpha-amylase from *Bacillus stearothermophilus*. FEMS Microbiol Lett. 1988; 56: 53–60.

51. Robyt JF, Ackerman RJ. Isolation, purification, and characterization of a maltotetraose-producing amylase from *Pseudomonas stuzeri*. Arch. Biochem Biophys. 1971; 145: 105–114.

52. Momma M. Cloning and sequencing of the maltohexaose-producing amylase gene of *Klebsiella pneumoniae*. Biosci Biotechnol Biochem. 2000; 64: 428–431.

53. Takat H, Kuriki T, Okada S, Takesada Y, Iizuka M, Imanaka T. Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at alpha-(1→4) and alpha-(1→6) glucosidic linkages. J Biol Chem. 1992; 267: 18447–18452.

54. Takaha T, Smith SM. The functions of 4-α-glucanotransferases and their use for the production of cyclic glucans. Biotechnol Genet Eng Rev. 1999; 16: 257–280.

55. Reddy NS, Nimmagadda A, Rao KRSS. An overview microbial α-amylase family. Afr J Biotechnol. 2003; 2: 645-648.

56. Kuriki T, Imanaka T. The concept of the α-amylase family: Structural similarity and common catalytic mechanism. J Biosci Bioeng. 1999; 87: 557-565.

57. Yokota T, Tonozuka T, Kamitori S, Sakano Y. The deletion of amino-terminal domain in *Thermoactinomyces vulgaris* R-47 α-amylases: Effects of domain N on activity, specificity, stability and dimerization. Biosci Biotechnol Biochem. 2001; 65: 401-408.

58. Dauter Z, Dauter M, Brzozowski AM, Christensen S, Borchert TV, Berier L, Wilson KS, Davies GJ. X-ray structure of Novamyl, the five-domain "maltogenic" alpha-amylase from *Bacillus stearothermophilus*: maltose and acarbose complexes at 1.7 Å resolution. Biochemistry. 1999; 38: 8385-8392.

59. Lawson CL, Montfort R, Strokopytov B, Rozeboom HJ, Kalk KH, Vries GE, Penninga D, Dijkhuizen L, Dijkstra BW. Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a maltose-dependent crystal form. J Mol Biol. 1994; 236: 590-600.

60. Matsuura Y, Kusunoki M, Date W, Harada S, Bando S, Tanaka N, Kakudo M. Low resolution crystal structures of Taka-amylase A and its complexes with inhibitors. J Biochem. 1979; 86: 1773-1783.

61. Matsuura Y, Kusunoki M, Harada W, Kakudo M. Structure and possible catalytic residues of Taka-amylase A. J Biochem. 1984. 95: 697-702.

62. Nagashima T, Toda S, Kitamoto K, Gomi K, Kumagai C, Toda H. Site-directed mutagenesis of the catalytic active-site residues of Taka-amylase A. Biosci Biotech Biochem. 1992; 56: 207-210.

63. Vihunen M, Ollikka P, Niskanen J, Mantsala P. Site-directed mutagenesis of thermostable α-amylase from *Bacillus stearothermophilus*: putative role of three conserved residues. J Biochem. 1990; 107: 267-272.

64. Kuriki T, Takata H, Okada S, Imanaka T. Analysis of the active center of *Bacillus stearothermophilus* neopullullanase. J Bacteriol 1991; 173: 6147-6152.

65. Podkovyrov SM, Burdette D, Zeikus JG. Analysis of the active center of cyclomaltodextrinase from *Thermoanaerobacter ethanoliccus* 39E. FEBS Lette. 1993; 317: 259-262.

66. Klein C, Hollender J, Bender H, Shulz GE. Catalytic center of cyclodextrin glycosyltransferase derived from X-ray structure analysis combined with site-directed mutagenesis. Biochemistry. 1992; 31: 8740-8746.

67. Nakamura A, Haga K, Ogawa S, Kuwano K, Kimura K, Yamane K. Functional relationships between cyclodextrin glucanotransferase from an alkalophilic *Bacillus* and α-amylases. Site-directed mutagenesis of the conserved two Asp and one Glu residues. FEBS Lett. 1992; 296: 37-40.

68. Mathupala SP, Lowe SE, Podkovyrov SM, Zeikus JG. Sequencing of amylopullulanase (*apu*) gene of the *Thermoanaerobacter ethanolicus* 39E, and identification of the active site by site-directed mutagenesis. J Biol Chem. 1993; 268: 16332-16344.

69. Kuriki T, Guan HP, Sivak M, Preiss J. Analysis of the active center of branching enzyme II from maize endosperm. J Protein Chem. 1996; 15: 305-313.

70. Yoshiki M. A possible mechanism of catalysis involving three essential residues in the enzymes of α-amylase family. Biologia. Bratislava. 2002; 11: 21-27.

71. Buisson G, Duee E, Haser R, Payan F. Three dimensional structure of porcine pancreatic alpha-amylase at 2.9 Å resolution. Role of calcium in structure and activity. EMBO J. 1987; 6: 3909-3916.

72. Boel E, Brady L, Brzozowski AM, Derewenda Z, Dodson GG, Jensen VJ, Petersen SB, Swift H, Thim L, Woldike HF. Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1 Å resolution of two enzymes from *Aspergillus*. Biochemistry. 1990; 29: 6244-6249.

73. Kamasaka H, Sugimoto K, Takata H, Nishimura T, Kuriki T. *Bacillus stearothermophilus* neopullulanase selective hydrolysis of amylose to maltose in the presence of amylopectin. Appl Environ Microbiol. 2002; 68: 1658-1664.

74. Fritzsche HB, Schwede T, Schylz GE. Covalent and three-dimensional structure of the cyclodextrinase from *Flavobacterium* sp. no. 92. Eur J Biochem. 2003; 270: 2332-2341.

75. Bertoldo C, Antranikian G. Starch-hydrolyzing enzymes from thermophilic archea and bacteria. Curr Opin Chem Biol. 2002; 6: 151-160.

76. Leathers TD. Biotechnological production and applications of pullulan. Appl Microbiol Biotechnol. 2003; 62: 468-473.

77. Gantelet H, Duchiron F. Purification and properties of a thermoactive and thermostable pullulanase from *Thermococcus hydrothermalis*, a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. Appl Microbiol Biotechnol. 1998; 49: 770-777.

78. Niehaus F, Bertoldo C, Kahler M, Antranikian G. Extremophiles as a source of novel enzymes for industrial application. Appl Microbiol Biotechnol. 1999; 51: 711-729.

79. Takisawa N, Murooka Y. Cloning of the pullulanase gene and over production of pullulanase in *Escherichia coli* and *Klebsiella aerogenes*. Appl Environ Microbiol. 1985; 49: 294-298.

80. Saha BC, Zeikus JG. Novel highly thermostable pullulanase from thermophiles. Trends in Biotechnol. 1989; 7: 234-238.

81. Yamashita M, Matsumoto D, Murooka Y. Amino acid residues specific for the catalytic action towards α-1,6-glucosidic linkages in *Klebsiella* pullulanase. J Ferment Bioeng. 1997; 84: 283-290.

82. Kuriki T, Okada S, Imanaka T. New type of pullulanase from *Bacillus stearothermophilus* and molecular cloning and expression of the gene in *Bacillus subtilis*. J Bacteriol. 1988; 170: 1554-1559.

83. Kuriki T, Imanaka T. Nucleotide sequence of neopullulanase gene from *Bacillus stearothermophilus*. J Gen Microbiol. 1989; 135: 1521-1528.

84. Webb EC. Enzyme Nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, Academics Press, Orlando. 1992.

85. Saha BC, Shen GJ, Srivastava KC, LeCureux LW, Zeikus JG. New thermostable alpha-amylase-like pullulanase from thermophilic *Bacillus* sp. 3183. Enzyme Microb Technol. 1989; 11: 760-764.

86. Rost B, Schneider RC. Progress in protein structure prediction? Trends Biochem Sci. 1993; 18: 120-123.

87. Ducki A, Grundmann O, Konermann L, Mayer F, Hoppert M. Glucoamylase from *Thermoanaerobacterium thermosaccharolyticum*: sequence studies and analysis of the macromolecular architecture of the enzyme. J Gen Appl Microbiol. 1998; 44: 327-335.

88. Aoki H, Sakano Y. Molecular cloning and heterologous expression of the isopullulanase gene from *Aspergillus niger* A.T.C.C. 9642. Biochem J 1997; 323: 757-764.

89. Lee HS, Kim MS, Cho HS, Kim JI, Kim TJ, Choi JH, Park C, Lee HS, Oh BH, Park KH. Cyclomaltodexrinase, neopullulanase and maltogenic amylase are nearly distinguishable from each other. J Biol Chem. 2002; 277: 21891-21897.

90. Kim JS, Cha SS, Kim HJ, Kim TJ, Ha NC, Oh ST, Cho HS, Cho MJ, Kim MJ, Lee HS, Kim JW, Choi KY, Park KW, Oh BH. Crystal structure of a maltogenic amylase provides insights into catalytic versatility. Biol Chem 1999; 274: 26279-26286.

91. Park KH, Kim TJ, Cheong TK, Kim JW, Oh BH, Svensson B. Structure, specificity and function of cyclomaltodextrinase, a multispecific enzyme of the α-amylase family. Biochim Biophys Acta. 2000; 1478: 165-185.

92. Bender H. Production, characterization, and application of cyclodextrins. Adv Biotechnol Proc. 1986; 6: 31–71.

93. De Pinto JA, Campbell LL Purification and properties of the cyclodextrinase of *Bacillus macerans*. Biochemistry. 1968; 17: 121-125.

94. Kitahata S, Taniguchi M, Beltran SS, Sugimoto T, Okada S. Purification and some properties of cyclodextrinase from *Bacillus coagulans*. Agric Biol Chem. 1983; 47: 1441-1447.

95. Galvin NM, Kelly CT, Fogarty WM. Purification and properties of the cyclodextrinase of *Bacillus sphaericus* ATCC7055. Appl Microbiol Biotechnol. 1994; 42: 46-50.

96. Kim IC, Cha JH, Kim JR, Jang SY, Seo BC, Cheong TK, Lee DS, Choi YD, Park KH. Catalytic properties of the cloned amylase from *Bacillus licheniformis*. J Biol Chem. 1992; 267: 22108-22114.

97. Cha HJ, Yoon HG, Kim YW, Lee HS, Kim JW, Kweon KS, Oh BH, Park KH. Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. Eur J Biochem. 1998; 253: 251-262.

98. Park KH, Kim MJ, Lee HS, Han NS, Kim D, Robyt JF. Transglycosylation reactions of *Bacillus stearothermophilus* maltogenic amylase with acarbose and various acceptors. Carbohydr Res. 1998; 313: 235-246.

99. Kim TJ, Kim MJ, Kim BC, Kim JC, Cheong TK, Kim JW, Park KH. Modes of action of acarbose hydrolysis and transglycosylation catalyzed by a thermostable maltogenic amylase, the gene for which was cloned from a *Thermus* strain. Appl Environ Microbiol. 1999; 65: 1644-1651.

100. Cha HJ, Yoon HG, Kim YW, Lee HS, Kim JW, Kweon KS, Oh BH, Park K H. Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. Eur J Biochem. 1998; 253: 251-262.

101. Lee SP, Morikawa M, Takagi M, Imanaka T. Cloning of the aapT gene and characterization of its product, α-amylase-pullulanase (AapT), from thermophilic *Bacillus* sp. strain XAL601. Appl Environ Microbiol. 1994; 60: 3764-3773.

102. Matzke J, Herrmann A, Schneider E, Bakker EP. Gene cloning, nucleotide sequence and biochemical properties of a cytoplasmic cyclomaltodextrinase (neopullulanase) from *Alicyclobacillus acidocaldarius*, reclassification of a group of enzymes. FEMS Microbiol Lett. 2000; 183: 55-61.

103. Smith KA, Salyers AA. Characterization of a neopullulanase and alpha-glucosidase from *Bacteroides thetaiotaomicron* 95-1, J Bacteriol. 1991; 173: 2962-2968.

104. Yebra MJ, Blasco A, Sanz P. Expression and secretion of *Bacillus polymyxa* neopullulanase in *Saccharomyces cerevisiae*. FEMS Microbiol Lett. 1999; 170: 41-49.

105. Kamitori S, Kondo S, Okuyama K, Yokota T, Shimura Y, Tonozuka T, Sakano T. Crystal structure of *Thermoactinomyces vulgaris* R-47 α amylase II (TVA II) hydrolyzing cyclodextrins and pullulan at 2.6 Å resolution. J Mol Biol. 1999; 287: 907-921.

106. Mizuno M, Tonozuka T, Uechi A, Ohtaki A, Ichikawa K, Kamitori S, Nishikawa A, Sakano Y. The crystal structure of *Thermoactinomyces vulgaris* R-47 α-amylase II (TVA II) complexed with transglycosylated product. Eur. J. Biochem. 2004; 271: 2530-2538.

107. Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S, Ryu S. Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library. Appl Environ Microbiol. 2004; 70: 7229-7235.

108. Fujiwara S. Extremophiles: Developments of their special functions and potential resources. J Biosci Bioeng. 2002; 94: 518-525.

109. Jacques G, Frederic DL, Joste LB, Viviane B, Bart D, Fabrizio G, Benoit G, Jean-marie F. An additional aromatic interaction improves the thermostability and thermophilicity of a thermophilic family 11 xylanase: structural basis and molecular study. Protein Sci. 2000; 9: 466-475.

110. Leuschner C, Antranikan G. Heat stable enzymes from extremely thermophilic and hyperthermophilic microorganisms. World J Microbiol Biotechnol. 1995; 11: 95-114.

111. Sunna A, Gibbs MD, Bergquist PL. A novel thermostable multidomain 1,4-β-xylanase from "*Caldibacillus cellulovorans*" and effect on its xylan-binding domain on enzyme activity. Microbiology. 2000a; 146: 2947-2955.

112. Sunna A, Bergquist P. A gene encoding a novel extremely thermostable 1,4-β-xylanase isolated directly from an environmental sample. Extremeophiles. 2003; 7: 63-70.

113. Whitman W, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. Proc Natl Acad Sci. 1998; 95: 6578-6583.

114. Omura S. Trends in the search for bioactive microbial metabolites. J Ind Microbiol. 1992; 10: 135-156.

115. Stackebrandt E. From strains to domains: measuring the degree of phylogenetic diversity among prokaryotic species and as yet uncultured strains. In Microbial Diversity in Time and Space. Plenum Press, New York. 1996: 19-24.

116. Hugenholtz P., Pace NR. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. Trends Biotechnol. 1996; 14: 190-197.

117. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and *In situ* detection of individual microbial cells without cultivation. Microbiol Rev. 1995; 59: 143-169.

118. Amann R, Kuhl M. *In situ* methods for assessment of microorganisms and their activities. Curr Opin Microbiol. 1998; 1: 352-358.

119. Daniel R. The soil metagenome--a rich resource for the discovery of novel natural products. Curr. Opin. Biotechnol. 2004; 15: 199-204.

120. Lancini G, Parenti G, Gallo GG. Resistance of microorganisms to antibiotics. In Antibiotics: Multidisciplinary Approach. Plenum Press, New York. 1995: 85-106.

121. Zahner H, Fiedler FP. The need for new antibiotics: possible ways forward. In Fifty Years of Antimmicrobials: Past, Present and Future Trends. Cambridge University Press, Cambridge. 1996: 67-84.

122. Kanokratana P, Chanapan S, Pootanakit K, Eurwilaichitr L. Diversity and abundance of *Bacteria* and *Archaea* in the Bor Khlueng Hot Spring in Thailand. J. Basic Microbiol. 2004; 6: 430-444.

123. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles M R, Loiacono KA, Lynch BA, MacNeil LA, Minor C. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultures microorganisms. Appl Environ Microbiol. 2000; 66: 2541-2547.

124. Yun J, Ryu S**.** Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. Microb Cell Fac. 2005; 4: 8-12.

125. Richardson TH, Tan X, Frey G, Callen W, Cabell M, Lam D, Macomber J, Short JM, Robertson DE, Miller C. A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable α-amylase. J Biol Chem. 2002; 277: 26501-26507.

126. Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. Proc Natl Acad Sci USA. 2004; 101: 16222-16227.

127. Piel J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. Proc Nalt Acad Sci USA. 2002; 99: 14002-14007.

128. Zhou J, Bruns MA, Tiedje JM. DNA recovery from soils of diverse composition. Appl Environ Microbiol. 1996; 62: 316-322.

129. Chun KT, Edenberg HJ, Kelley MR, Goebl MG. Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. Yeast. 1997; 13: 233-240.

130. Miller, GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959; 31: 436-428.

131. Smalla K, Cresswell N, Mendonca-Hagler LC, Wolters A, vanElsas JD. Rapid DNA extraction protocol from soil for polymerase chain reaction-mediated amplification. J Appl Bacteriol. 1993; 74: 78-85.

132. Tsai YL, Olson BH. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. Appl Environ Microbiol. 1992; 2292-2295.

133. Porteous LA, Armstrong JL. Recovery of bulk DNA from soil by a rapid, small-scale extraction method. Curr Microbiol. 1991; 22: 345-348.

134. Hondoh H, Kuriki T, Matsuura Y. Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. J Mol Biol 2003; 326: 177-188.

135. Brinkmann U, Mattes RE, Buckel P. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. Gene. 1989; 85: 109-114.

136. Davis GD, Elisee C, Newham DM, Harrison RG. New fusion protein systems designed to give soluble expression in *Esherichia coli*. Biotecnol Bioeng. 1999; 65: 382-388.

137. Oslancova A, Janecek S. Oligo-1,6-glucosidase and neopullulanase enzyme subfamilies from the α-amylase family defined by the fifth conserved sequence region. Cell Mol Life Sci. 2002; 1945-1959.

138. Kamitori S, Abe A, Ohtaki A, Kaji A, Tonozuka T, Sakano Y. Crystal structural and structural comparison of *Thermoactinomyces vulgaris* R-47 α-amylase 1 (TVA I) at 1.6 Å resolution and α-amylase 2 (TVA II) at 2.3 Å resolution. J Mol Biol. 2002; 443-453.

139. Turner P, Labes A, Fridjonsson OH, Hreggvidson GO, Schonheit P, Kristjansson JK, Holst O, Karlsson EN. Two novel cyclodextrin-degrading enzymes isolated from thermophilic bacteria have similar domain structures but differ in oligomeric state and activity profile. J Biosci Bioeng. 2005; 100: 380-390.

140. Lee MH, Kim YW, Kim TJ, Park CS, Kim JW, Moon TW, Park WH. A novel amylolytic enzyme from *Thermotoga maritima*, resembling cyclodextrinase and α-glucosidase, that liberates glucose from the reducing

end of the substrates. Biochem and Biophys Res Commun. 2002; 295: 818-825.

141. Akashi H. Translational selection and yeast proteome evolution. Genetics. 2003; 164: 1291-1303.

142. Kozak M. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc Natl Acad Sci U S A. 1986; 83: 2850-2854.

143. Tomiyasu K, Yato K, Yasuda M, Tonozuka T, Ibuka A, Sakai H. Cloning and nucleotide sequence of the pullulanase gene of *Thermus thermophilus* HB8 and production of the enzyme in *Escherichia coli*. Biosci Biotechnol Biochem. 2001; 65: 2090-2094.

144. Georgiou G, Valax P. Expression of correctly folded proteins in *Escherichia coli*. Curr. Opin. Biotechnol. 1996; 7: 190-197.

145. Duffner F, Bertoldo C, Andersen JT, Wagner K, Antranikian G. A new thermoactive pullulanase from *Desulfurococcus mucosus*: cloning, sequencing, purification, and characterization of the recombinant enzyme after expression in *Bacillus subtilis*. J Bacteriol. 2000; 182: 6331-6338.

146. Rudolph R. Successful protein folding on an industrial scale. In Protein Engineering: Principles and Practices. John Wiley&Sons, New York. 1995: 283-298.

147. Vasina JA, Baneyx F. Expression of aggregation-prone recombinant proteins at low temperature: a comparative study of the *Escherichia coli* cspA and tac promoter system. Protein Expr. Purif. 1997; 9: 211-218.

148. Turner P, Holst O, Karlsson EN. Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration. Protein Expr Purif. 2005; 39: 54-60.

149. Imamura H, Jeon B, Wakagi T, Matsuzawa H. High level expression of *Thermococcus litoralis* 4-alpha glucanotransferase in a soluble form in *Escherichia coli* with a novel expression system involving minor arginine tRNAs and GroELS. FEBS Lett. 1999; 457: 393–396.

150. Davis GD, Elisee C, Newham DM, Harrison RG. New fusion protein systems designed to give soluble expression in *Escherichia coli*. Biotechnol. Bioeng. 1999; 65: 382-388.

151. Kim TJ, Nguyen VD, Lee HS, Kim MJ, Cho HY, Kim YW, Moon TW, Park CS, Kim JW, Oh BH, Lee SB, Svensson B, Park KH. Modulation of the multisubstrate specificity of Thermus maltogenic amylase by truncation of the N-terminal domain and by a salt-induced shift of the monomer/dimer equilibrium. Biochemistry. 2001; 40: 14182-14190.

152. Arnosti C. A new method for measuring polysaccharide hydrolysis rates in marine environments. Org Geochem. 1996; 25:105–115.

BIOGRAPHY

NAME	Miss Viriya Nitteranon
DATE OF BIRTH	March 18, 1981
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTEND	Kasetsart University, 1999-2002 Bachelor of Science (Biology) Mahidol University, 2003-2005 Master of Science (Molecular Genetics and Genetic Engineering)
SCHOLARSHIP	Scholarship from Thailand Graduate Institute of Science and Technology (TGIST), NSTDA 2004-2005
HOME ADDRESS	36/14 Lumtadaeng Rd. Saladaeng, Muang district, Angthong, 14000 E-mail: viriya@yahoo.com