

**FUNCTIONAL ANALYSIS OF RICE  
MONOLIGNOL BETA-GLUCOSIDASES**



**Supaporn Baiya**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the**

**Degree of Doctor of Philosophy in Biochemistry**

**Suranaree University of Technology**

**Academic Year 2014**

# การวิเคราะห์หน้าที่ของเอนไซม์โมโนลิกลนอลเบตา-กลูโคซิเดสในข้าว



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต  
สาขาวิชาชีวเคมี  
มหาวิทยาลัยเทคโนโลยีสุรนารี  
ปีการศึกษา 2557

**FUNCTIONAL ANALYSIS OF RICE  
MONOLIGNOL BETA-GLUCOSIDASES**

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

---

(Assoc. Prof. Dr. Jatuporn Wittayakun)

Chairperson

---

(Prof. Dr. James R. Ketudat-Cairns)

Member (Thesis Advisor)

---

(Prof. Dr. Thomas W. Okita)

Member

---

(Prof. Dr. Jong-Seong Jeon)

Member

---

(Assoc. Prof. Dr. Mariena Ketudat-Cairns)

Member

---

(Assoc. Prof. Dr. Jaruwan Siritapetawee)

Member

---

(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs  
and Innovation

---

(Assoc. Prof. Dr. Prapun Manyum)

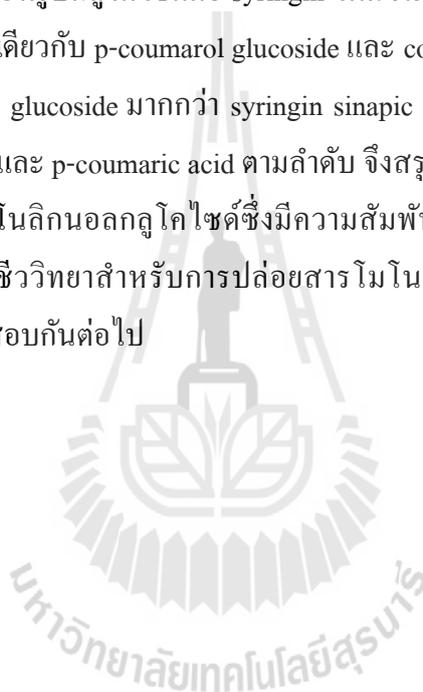
Dean of Institute of Science

สุภาพรณั์ ไบยา : การวิเคราะห์หน้าที่ของเอนไซม์โมโนลิกนอลเบตา-กลูโคซิเดสในข้าว  
(FUNCTIONAL ANALYSIS OF RICE MONOLIGNOL  $\beta$ -GLUCOSIDASES)  
อาจารย์ที่ปรึกษา : ศาสตราจารย์ ดร.เจมส์ เกตุทัต-คาร์นส์, 140 หน้า.

ในพืชชั้นสูงเอนไซม์เบตา-กลูโคซิเดสจัดอยู่ในตระกูลไกลโคไซโกลิโกลิโคเลส กลุ่มที่ 1 มีหน้าที่สำคัญหลายอย่าง รวมไปถึงกระบวนการสร้างลิกนิน จากการวิเคราะห์ phylogenetic ของเอนไซม์เบตา-กลูโคซิเดสในข้าว พบว่า Os4BGlu14 Os4BGlu16 และ Os4BGlu18 มีลำดับกรดอะมิโนที่มีความใกล้เคียงกับเอนไซม์ในกลุ่มโมโนลิกนอลเบตา-กลูโคซิเดสมากที่สุด จึงนำมาสู่สมมติฐานว่าเอนไซม์ทั้ง 3 ตัวนี้เกี่ยวข้องกับกระบวนการสร้างลิกนิน cDNA ของโมโนลิกนอลเบตา-กลูโคซิเดส ยีน Os4BGlu14 และ Os4BGlu18 ถูกโคลน ตรวจสอบความถูกต้องของลำดับเบส นิวคลีโอไทด์ และตัดต่อเข้าไปในพลาสมิด pET32a จากนั้นพลาสมิดที่ถูกตัดต่อยีนเหล่านี้ ถูกนำเข้าไปใน Escherichia coli เพื่อนำไปใช้ในการผลิตเอนไซม์ โดยมี thioredoxin และ His<sub>6</sub> tags ต่อที่ปลาย N-terminal เนื่องจากตำแหน่ง acid/base ของ Os4BGlu14 ถูกแทนที่ด้วยกรดอะมิโนกลูตามีน ซึ่งกรดอะมิโนนี้อาจทำให้เอนไซม์ Os4BGlu14 ไม่สามารถทำงานได้เหมือนเอนไซม์เบตา-กลูโคซิเดสโดยทั่วไป ดังนั้น การตัดแปลงกรดอะมิโนกลูตามีนไปเป็นกลูตามีนจึงถูกทำขึ้นด้วยเทคนิค QuikChange site-directed mutagenesis อย่างไรก็ตาม ไม่สามารถตรวจวัดการทำงานของเอนไซม์ Os4BGlu14 ได้ทั้งในเอนไซม์ดั้งเดิมและเอนไซม์ที่ถูกตัดแปลงกรดอะมิโน และเนื่องจากมีปัญหาในการโคลน Os4BGlu16 จึงทำการสังเคราะห์ยีนของ Os4BGlu16 ขึ้นเพื่อใช้ในการผลิตเอนไซม์ตัวนี้ในยีสต์ เอนไซม์ Os4BGlu16 ที่ถูกผลิตและปล่อยออกมาจากผนังเซลล์ของยีสต์ถูกนำมาแยกให้บริสุทธิ์ด้วยวิธี immobilized metal affinity chromatography (IMAC) ส่วนเอนไซม์ Os4BGlu18 ถูกสกัดมาจาก E. coli และนำมาแยกให้บริสุทธิ์ด้วย anion exchange chromatography hydrophobic interaction chromatography และ IMAC column ตามลำดับ เอนไซม์ Os4BGlu16 และ Os4BGlu18 สามารถย่อยโมโนลิกนอลกลูโคไซด์สับสเตรท coniferin (ด้วย  $k_{cat}/K_M$ , 21.6  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu16 และ 31.9  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu18) syringin (ด้วย  $k_{cat}/K_M$ , 22.8  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu16 และ 24.0  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu18) และ p-coumarol glucoside (ด้วย  $k_{cat}/K_M$ , 6.2  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu16 และ 1.4  $\text{mM}^{-1}\text{s}^{-1}$  สำหรับ Os4BGlu18) ได้อย่างมีประสิทธิภาพมากกว่าสับสเตรทตัวอื่น

ในการตรวจหาการแสดงออกของยีนโมโนลิกนอลเบตา-กลูโคซิเดสในข้าว พบว่า mRNA ของ Os4BGlu14 มีการตรวจพบมากที่สุดในส่วนของเมล็ด รวง และเกสร ส่วน mRNA ของ

Os4BGlu16 พบมากที่สุดในส่วนของใบจากต้นข้าวสัปดาห์ที่ 4 ถึงสัปดาห์ที่ 10 เอนโดสเปิร์ม และเปลือกเมล็ดส่วนนอก และ mRNA ของ Os4BGlu18 ส่วนใหญ่ถูกพบในช่วงแรกของการเจริญเติบโตตั้งแต่สัปดาห์แรกถึงสัปดาห์ที่ 4 และยังคงพบในเกสรและเปลือกเมล็ดส่วนนอกอีกด้วย ข้อมูลเหล่านี้บ่งบอกถึงการทำงานของเอนไซม์โมโนลิคนอลเบตา-กลูโคซิเดสเกิดขึ้นทั้งในช่วงเจริญเติบโตจนถึงช่วงออกผลผลิต การศึกษาปริมาณของสารโมโนลิคนอล โมโนลิคนอลกลูโคไซด์ และสารที่เกี่ยวข้องในข้าวโดยใช้เครื่อง UPLC-MS เพื่อหาความสัมพันธ์ระหว่างปริมาณเอนไซม์และสับสเตรทที่ถูกผลิต พบว่า ปริมาณ sinapyl alcohol ถูกพบในส่วนของรากและกาบใบอายุ 10 ถึง 40 วัน ในขณะที่รูปกลูโคไซด์คือ syringin เพิ่มขึ้นในช่วงอายุ 2-3 เดือนทั้งในราก ใบ กาบใบ และข้อปล้อง เช่นเดียวกับ p-coumarol glucoside และ coniferin ในส่วนของดอกข้าว พบว่ามีปริมาณของ p-coumarol glucoside มากกว่า syringin sinapic acid coniferin caffeic acid sinapyl alcohol coniferyl alcohol และ p-coumaric acid ตามลำดับ จึงสรุปได้ว่าเอนไซม์นี้จะถูกสร้างขึ้นในระหว่างการสร้างสารโมโนลิคนอลกลูโคไซด์ซึ่งมีความสัมพันธ์กันทางด้านกระบวนการชีวเคมีและการศึกษาสภาพทางชีววิทยาสำหรับการปล่อยสารโมโนลิคนอลสำหรับกระบวนการสร้างลิกนินยังคงต้องมีการทดสอบกันต่อไป



SUPAPORN BAIYA : FUNCTIONAL ANALYSIS OF RICE

MONOLIGNOL BETA-GLUCOSIDASES. THESIS ADVISOR :

PROF. JAMES R. KETUDAT-CAIRNS, Ph.D. 140 PP.

GLYCOSIDE HYDROLASE/CHARACTERIZATION/RICE/MONOLIGNOL  
 $\beta$ -GLUCOSIDASE/MONOLIGNOL COMPOUNDS

In higher plants,  $\beta$ -glucosidases belonging to glycoside hydrolase family 1 (GH1) have been implicated in several fundamental processes, including lignification. Phylogenetic analysis of rice (*Oryza sativa* L.) GH1  $\beta$ -glucosidases indicated that Os4BGlu14, Os4BGlu16, and Os4BGlu18 are closely related to known monolignol  $\beta$ -glucosidases, leading to the hypothesis that they may release monolignols from their inactive glucosides. The cDNAs for Os4BGlu14 and Os4BGlu18 genes were cloned, sequenced, and ligated into pET32a, and the resulting recombinant plasmids were used to express fusion proteins with N-terminal thioredoxin and His<sub>6</sub> tags in *Escherichia coli*. Because the conserved acid/base glutamate position of Os4BGlu14 is replaced with glutamine residue 191, the cDNA of Os4BGlu14 was mutated to glutamate at this position to produce Os4BGlu14 Q191E. However, no activity could be detected for proteins expressed from pET32a/Os4BGlu14 and pET32a/Os4BGlu14Q191E. Due to difficulty in cloning a functional cDNA for *E. coli* expression, an optimized Os4BGlu16 cDNA was synthesized for expression of the protein in *Pichia pastoris*. The secreted Os4BGlu16 fusion protein was purified from induced *P. pastoris* culture media by immobilized metal affinity chromatography (IMAC) to yield a single prominent protein band on SDS-PAGE analysis and strong  $\beta$ -glucosidase activity. In contrast, active Os4BGlu18  $\beta$ -glucosidase fusion protein with N-terminal thioredoxin and His<sub>6</sub> tags was successfully

expressed and extracted from *E. coli* cells, and was purified by anion exchange chromatography, hydrophobic interaction chromatography and IMAC. Os4BGlu16 and Os4BGlu18 hydrolyzed the monolignol glucosides coniferin ( $k_{cat}/K_M$ , 21.6  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu16 and 31.9  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu18), syringin ( $k_{cat}/K_M$ , 22.8  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu16 and 24.0  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu18), and *p*-coumarol glucoside ( $k_{cat}/K_M$ , 6.2  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu16 and 1.4  $\text{mM}^{-1}\text{s}^{-1}$  for Os4BGlu18) with much higher catalytic efficiencies than other substrates.

By quantitative RT-PCR, highest Os4BGlu14 mRNA levels were detected in seed, panicle and pollen. Os4BGlu16 was detected at highest levels in leaf from 4 to 10 weeks, endosperm and lemma, while Os4BGlu18 mRNA was most abundant in vegetative tissues from 1 week to 4 weeks old and in pollen and lemma. These data suggest a role for monolignol  $\beta$ -glucosidases in both vegetative and reproductive rice tissues. The relative amounts of monolignols, their glycosides and related compounds in rice tissues were analyzed by UPLC-MS to relate the enzyme expression levels to levels of putative substrates and products. Sinapyl alcohol was detected in root and leaf sheath from 10-40 days, whereas its glucoside, syringin, was dramatically increased in 2-3 month-old roots, leaf, leaf sheath, and stem, as were the levels of *p*-coumarol glucoside and coniferin. In rice flower extracts, *p*-coumarol glucoside was highest, followed by syringin, sinapic acid, coniferin, caffeic acid, sinapyl alcohol, coniferyl alcohol, and *p*-coumaric acid, respectively. Thus, the enzymes appear to be present during the build-up of monolignol glucosides, although the exact biological context for the release of the monolignols from these compounds remains to be determined.

School of Biochemistry

Student's signature\_\_\_\_\_

Academic Year 2014

Advisor's signature\_\_\_\_\_

## **ACKNOWLEDGEMENTS**

I am thankful to my thesis advisor, Prof. Dr. James R. Ketudat-Cairns, for providing me the opportunity to study toward my Ph.D. degree in Biochemistry, kind supervision, valuable advice, and encouragement when I had problems with the experiments. I am grateful to my co-advisor, Prof. Dr. Thomas W. Okita, for giving me an opportunity to work on rice transformation, kindness and suggestions of my work when I was a visiting graduate student at Washington State University, USA, and Prof. Dr. Jong-Seon Jeon, for inviting me to work on protein localization in Kyung Hee University, Korea. I would also like to thank Assoc. Prof. Dr. Mariena Ketudat-Cairns for advice about molecular cloning and encouraging throughout my project work, and Assoc. Prof. Dr. Jaruwat Siritapetawee for her suggestions about protein expression and purification. I also would like to thank my colleagues and all friends for their help in the Biochemistry Lab.

I thank the Royal Golden Jubilee Ph.D. Program of the Thailand Research Fund, (Grant No. PHD/0106/2551), Suranaree University of Technology, the National Research University Project of the Commission on Higher Education and the National Science and Technology Development Agency for funding.

Finally, I would like to express my deepest gratitude to my family for their unconditional love, care, support, and encouragement.

Supaporn Baiya

## CONTENTS

	<b>Page</b>
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH .....	III
ACKNOWLEDGEMENTS.....	V
CONTENTS.....	VI
LIST OF TABLES .....	X
LIST OF FIGURES .....	XI
LIST OF ABBREVIATIONS.....	XIV
<b>CHAPTER</b>	
<b>I INTRODUCTION.....</b>	<b>1</b>
1.1 Overview of $\beta$ -glucosidases.....	1
1.2 Glycoside hydrolases .....	4
1.3 Plant glycoside hydrolase family 1 .....	7
1.4 Rice $\beta$ -glucosidases.....	8
1.5 Plant lignification.....	13
1.6 Lignans, structure and functions .....	16
1.7 Monolignol $\beta$ -glucosidases .....	18
1.8 Research objectives.....	22
<b>II MATERIALS AND METHODS .....</b>	<b>23</b>
2.1 Materials .....	23
2.1.1 Plant, plasmids, bacterial and yeast stains .....	23

## CONTENTS (Continued)

	<b>Page</b>
2.1.2 Oligonucleotides primers .....	24
2.1.3 Chemicals and reagents.....	26
2.2 General methods .....	27
2.2.1 RNA extraction .....	27
2.2.2 First-strand cDNA synthesis .....	27
2.2.3 Preparation of <i>E. coli</i> competent cells .....	28
2.2.4 Transformation of plasmids into competent cells .....	28
2.2.5 Plasmid isolation by alkaline lysis method .....	29
2.2.6 QIAGEN plasmid miniprep .....	30
2.2.7 Agarose gel electrophoresis for DNA.....	31
2.2.8 Purification of DNA bands from gels .....	31
2.2.9 SDS-PAGE electrophoresis .....	31
2.2.10 Determination of protein concentration .....	32
2.2.11 Preparation of <i>P. pastoris</i> SMD1168H competent cells.....	33
2.3 Amplification and cloning of Os4BGlu14.....	33
2.3.1 Amplification of gene encoding mature Os4BGlu14 .....	33
2.3.2 Cloning of mature Os4BGlu14 into pET32a vector .....	34
2.3.3 Mutagenesis of pET32a/Os4BGlu14 .....	35
2.4 Cloning of Os4BGlu16 .....	37
2.4.1 Cloning of optimized Os4BGlu16 into pPICZ $\alpha$ B(NH <sub>8</sub> ) vector.....	37
2.5 Amplification and cloning of Os4BGlu18.....	37
2.5.1 Amplification of a cDNA encoding mature Os4BGlu18.....	37

## CONTENTS (Continued)

	<b>Page</b>
2.5.2 Cloning of a cDNA encoding mature Os4BGlu18 into pET32a .....	38
2.6 Expression of Os4BGlu14 .....	39
2.7 Expression and purification of Os4BGlu16.....	40
2.8 Expression and purification of Os4BGlu18.....	41
2.9 Enzyme assay, pH and temperature optimum and stability studies.....	43
2.10 Substrate specificity and enzyme kinetics .....	44
2.11 Inhibition study .....	46
2.12 <i>In planta</i> expression analysis of rice monolignol $\beta$ -glucosidase .....	46
2.13 Detection of monolignol compounds in rice KDML105 by UPLC-MS analysis .....	47
<b>III RESULTS</b> .....	49
3.1 Cloning and expression of Os4BGlu14 .....	49
3.2 Cloning and expression of Os4BGlu16 .....	56
3.3 Cloning and expression of Os4BGlu18 .....	58
3.4 Purification of Os4BGlu16 .....	63
3.5 Effect of pH and temperature on the activity and stability of Os4BGlu16....	64
3.6 Purification of Os4BGlu18 .....	69
3.7 Effect of pH and temperature on the activity and stability of Os4BGlu18....	70
3.8 Substrate specificity and kinetic analysis of Os4BGlu16 and Os4BGlu18 ...	74
3.9 Inhibition of Os4BGlu16 and Os4BGlu18 by metal-salt inhibitors .....	82
3.10 <i>In planta</i> expression analysis .....	84

## CONTENTS (Continued)

	<b>Page</b>
3.11 Detection of monolignol compounds in rice strain KDML105 by UPLC-MS analysis .....	88
<b>IV DISCUSSION</b> .....	94
4.1 Rice monolignol $\beta$ -glucosidase sequence identification and analysis .....	94
4.2 Expression and purification of Os4BGlu14 .....	95
4.3 Expression and purification of Os4BGlu16 .....	97
4.4 Expression and purification of Os4BGlu18 .....	98
4.5 Substrate specificity of Os4BGlu16 and Os4BGlu18 .....	99
4.6 Inhibition of Os4BGlu16 and Os4BGlu18 .....	101
4.7 mRNA expression of monolignol $\beta$ -glucosidase genes .....	102
4.8 Level of monolignol compounds and its precursor in rice KDML105 plants .....	104
<b>V CONCLUSION</b> .....	106
<b>REFERENCES</b> .....	112
<b>APPENDICES</b> .....	128
<b>APPENDIX A DETECTION OF MONOLIGNOL COMPOUNDS BY UPLC-MS METHOD</b> .....	129
<b>APPENDIX B PUBLICATIONS</b> .....	139
<b>CURRICULUM VITAE</b> .....	140

## LIST OF TABLES

Table	Page
2.1 Oligonucleotide primers for monolignol $\beta$ -glucosidase cloning, mutagenic primers of Os4BGlu14, and unique primers for RT-PCR.....	24
2.2 Cycling parameters for mutation of pET32a/Os4BGlu14 by the QuikChange <sup>®</sup> Site-Directed Mutagenesis method.....	36
2.3 Cycling parameters for amplification of cDNA encoding mature Os4BGlu18 ...	38
3.1 Relative activities of Os4BGlu16 and Os4BGlu18 toward nitrophenyl glycosides.....	75
3.2 Activities of Os4BGlu16 and Os4BGlu18 toward natural and synthetic glycoside .....	76
3.3 Kinetic parameters of Os4BGlu16 and Os4BGlu18 for hydrolysis of nitrophenyl glycosides .....	79
3.4 Kinetic parameters of Os4BGlu16 and Os4BGlu18 for hydrolysis of monolignol glucosides and other aryl and alkyl glucosides .....	80
3.5 Effects of EDTA, metal salts and inhibitors on Os4BGlu16 and Os4BGlu18 activity .....	83

## LIST OF FIGURES

Figure	Page
1.1 The two major mechanisms of glycoside hydrolase enzyme.....	6
1.2 Phylogenetic tree of the predicted protein sequences of rice and <i>Arabidopsis</i> glycoside hydrolase family 1 genes.....	10
1.3 Lignin monomers and structures in the polymer .....	17
3.1 Amplification of a cDNA encoding mature Os4BGlu14 gene using AK067841 cDNA clone plasmid as a template .....	50
3.2 Amino acid sequence alignment of predicted rice ( <i>Oryza sativa</i> ) monolignol $\beta$ -glucosidases (Os4BGlu14, Os4BGlu16, Os4BGlu18), rice Os3BGlu7 (BGlu1), <i>Arabidopsis thaliana</i> $\beta$ -glucosidases (BGLU45, BGLU46, BGLU47) and <i>Pinus contorta</i> coniferin $\beta$ -glucosidase ...	51
3.3 SDS-PAGE analysis of pET32a/Os4BGlu14 expression in various <i>E. coli</i> host cells with 0.4 mM IPTG and induced at 20°C for 16 h.....	52
3.4 SDS-PAGE analysis of pET32a/Os4BGlu14 expression in Origami(DE3) by varying the concentration of IPTG from 0-0.8 mM and induced at 20°C for 16 h.....	53
3.5 SDS-PAGE analysis of pET32a/Os4BGlu14 expression in Origami(DE3) by varying the temperature from 10-22°C and induced at 0.4 mM IPTG for 16 h.....	54

## LIST OF FIGURES (Continued)

Figure	Page
3.6 SDS-PAGE analysis of pET32a/Os4BGlu14 expression in Origami(DE3) induced with 0.4 mM IPTG at 20°C for 16 h and purified by IMAC column .....	55
3.7 An optimized Os4BGlu16 plasmid cut with <i>Pst</i> I and <i>Xba</i> I.....	56
3.8 The activity toward <i>p</i> NPGlc of Os4Glu16 in pichia media over 7 days of expression .....	57
3.9 Amplification of a cDNA encoding mature Os4BGlu18 gene from 7 days old rice roots and shoots cDNA as a template .....	58
3.10 SDS-PAGE analysis of pET32a/Os4BGlu18 expression in various <i>E. coli</i> host cells.....	60
3.11 SDS-PAGE analysis of pET32a/Os4BGlu18 expression in Origami(DE3) at temperatures varying from 10-22°C and induced at 0.4 mM IPTG for 16 h ...	61
3.12 SDS-PAGE analysis of pET32a/Os4BGlu18 expression in Origami(DE3) induced at IPTG concentrations from 0 to 0.8 mM. Cultures were induced at 18°C for 16 h.....	62
3.13 SDS-PAGE analysis of Os4BGlu16 protein expressed in <i>Pichia pastoris</i> .....	63
3.14 Activity versus pH profiles of Os4BGlu16.....	65
3.15 pH stability of Os4BGlu16 activity .....	66
3.16 Temperature optima of Os4BGlu16.....	67
3.17 Thermostability of Os4BGlu16.....	68
3.18 SDS-PAGE analysis of Os4BGlu18 production in <i>Escherichia coli</i> .....	69
3.19 Activity versus pH profiles of Os4BGlu18.....	70

## LIST OF FIGURES (Continued)

<b>Figure</b>	<b>Page</b>
3.20 pH stability of Os4BGlu18 activity .....	71
3.21 Temperature optimum of Os4BGlu18 .....	72
3.22 Thermostability of Os4BGlu18.....	73
3.23 Gene expression analysis of rice Os4BGlu14 from qRT-PCR.....	85
3.24 Gene expression analysis of rice Os4BGlu16 from qRT-PCR.....	86
3.25 Gene expression analysis of rice Os4BGlu18 from qRT-PCR.....	87
3.26 Chemical structures of monolignol compounds .....	89
3.27 Relative abundance of monolignol compounds and intermediates in their biosynthesis in root extracts from rice (cv. KDML105) from 10 days to 120 days .....	90
3.28 Relative abundance of monolignol compounds and intermediates in their biosynthesis in leaf extracts from rice (cv. KDML105) from 10 days to 120 days .....	91
3.29 Relative abundance of monolignol compounds and intermediates in their biosynthesis in leaf sheath extracts from rice (cv. KDML105) from 20 days to 120 days.....	92
3.30 Relative abundance of monolignol compounds and intermediates in their biosynthesis in stem extracts from rice (cv. KDML105) from 60 days to 120 days, flower, and seed.....	93

## LIST OF ABBREVIATIONS

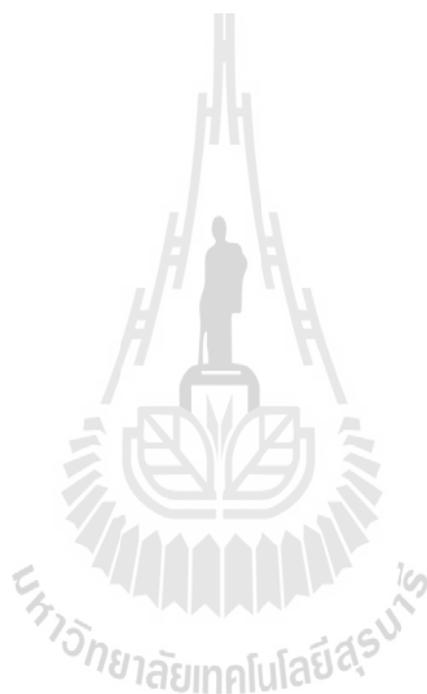
ABTS	2,2'-azinobis-3-ethylbenthaiazolinesulfonic acid
APS	Ammonium persulfate
bis-acryalmide	N,N-methylene-bis-acrylamide
bp	Base pairs
BSA	Bovine serum albumin
cDNA	Complementary deoxynucleic acid
CV	Column volume
DMSO	Dimethyl sufoxide
DNA	Deoxynucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxynucleoside triphosphates
EDTA	Ethylenediamine tetraacetate
ESI-MS	Electrospray ionization-mass spectrometry
EtOAc	Ethyl acetate
GH	Glycoside hydrolase
GH1	Glycoside hydrolase family 1
Glc	Glucose
HPLC	High performance liquid chromatography
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl thio- $\beta$ -D-galactoside

**LIST OF ABBREVIATIONS (Continued)**

kb	Kilo base pair(s)
kDa	Kilo Dalton(s)
LB	Luria-Bertani lysogeny broth
UPLC-MS	Ultra performance liquid chromatography-mass spectrometry
MeOH	Methanol
MW	Molecular weight
NaOAc	Sodium acetate
NMR	Nuclear magnetic resonance
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonylfluoride
<i>p</i> NP	<i>para</i> -Nitrophenyl
<i>p</i> NPGlc	<i>p</i> NP- $\beta$ -D-glucopyranoside
PVDF	Polyvinylidene flouride
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rotations per minute
SDS	Sodium dodecyl sulfate
TEMED	Tetramethylenediamine

**LIST OF ABBREVIATIONS (Continued)**

TLC	Thin-layer chromatography
Tris	Tris-(hydroxymethyl)-aminoethane
UV	Ultraviolet



# CHAPTER I

## INTRODUCTION

### 1.1 Overview of $\beta$ -glucosidases

$\beta$ -Glucosidases (EC. 3.2.1.21) are hydrolases that catalyze the hydrolysis of  $\beta$ -glycosidic linkages in aryl and alkyl  $\beta$ -O-D-glucosides and gluco-oligosaccharides to release D-glucose and an aglycone. These enzymes are found widely in the living organisms, including plants, fungi, animals, archaea and bacteria.  $\beta$ -Glucosidases have several important roles, including acting in cell wall remodeling, chemical defense, plant-microbe interactions, phytohormone activation, activation of metabolic intermediates, release of volatiles from their glycosides in plants, conversion of biomass in micro-organisms and breakdown of glycolipids and exogenous glucosides in animals (Ketudat Cairns and Esen, 2010). Many  $\beta$ -glucosidases from different sources have similarity in substrate specificity for glycone and some non-physiological aglycones, although they may have different physiological glucosidic substrates with different aglycone moieties.

In mammals, there are four known  $\beta$ -glucosidases, including the glycoside hydrolase (GH) family 1 (GH1) enzymes lactase-phloridzin hydrolase and cytoplasmic  $\beta$ -glucosidase, the GH family 30 (GH30) enzyme human acid  $\beta$ -glucosidase and the GH family 116 (GH116) bile acid  $\beta$ -glucosidase (Ketudat Cairns and Esen, 2010). The most studied and best characterized is the human acid  $\beta$ -

glucosidase or lysosomal glucocerebrosidase, the deficiency of which causes Gaucher disease. Such defects are usually the result of mutations in the glucocerebrosidase gene located in the q21 region of chromosome 1. More than 100 different mutations have been identified in this gene, with the two point mutations N370S and L444P being most frequently observed in patients with Gaucher disease (Koprivica et al., 2000). Human cytosolic  $\beta$ -glucosidase, also known as klotho-related protein and glucoceramide 3 (KLRP, GBA3), is an enzyme that hydrolyzes a wide variety of  $\beta$ -D-glucosides, such as synthetic aryl glycosides (4-nitrophenyl and 4-methylumbelliferyl monoglycosides), dietary flavonoid and isoflavone glucosides, and glucosyl ceramide (Berrin et al., 2003; Tribolo et al., 2007; Noguchi et al., 2008; Hayashi et al., 2007). The bile acid  $\beta$ -glucosidase GBA2 is predicted to be involved in the metabolic pathway of glucosylceramide synthesized on the cytosolic faces of the ER/Golgi membranes (Körschen et al., 2013).

In fungi,  $\beta$ -glucosidases play an important role in the cellulase system. This system requires three cellulolytic enzymes for complete degradation of cellulose. Two of these that catalyze hydrolysis of internal bonds to produce cellobiose and cello-oligosaccharides are endo-glucanases (EC 3.2.1.4) and cellobiohydrolases (EC 3.2.1.91), while the other is  $\beta$ -glucosidase that hydrolyzes cellobiose and cello-oligosaccharides to produce glucose (Coughlan and Ljungdahl, 1988). Fungi have also been reported to produce saponin-hydrolyzing enzymes, such as the avenacinases, which are a subset of  $\beta$ -glucosidases that hydrolyze antimicrobial saponins to suppress induced defense responses and signal transduction processes leading to disease resistance in plants (Bouarab et al., 2002; Osbourn et al., 1995).

$\beta$ -Glucosidases are also found in the insect midgut or specialized tissues such as defensive glands (Boonclarm et al., 2006). Insect  $\beta$ -glucosidases have been divided into two classes based on their relative catalytic efficiency toward several substrates. Class A  $\beta$ -glucosidases hydrolyze substrates with hydrophilic aglycones, such as disaccharides and oligosaccharides with  $\beta$ -1,3-,  $\beta$ -1,4-, and  $\beta$ -1,6- glycoside bonds, while class B  $\beta$ -glucosidases are active toward substrates with hydrophobic aglycones, such as alkyl, 4-nitrophenyl-, and methylumbelliferyl-glycosides, as well as plant glycosides (Terra and Ferreira, 2005). From their substrate specificity,  $\beta$ -glucosidases may release toxic aglycones from plant glycosides, which might lead to autotoxicity because  $\beta$ -glucosidases act as activation enzymes (Ferreira et al., 1997; Yu, 1989).

Plant  $\beta$ -glucosidases play important roles in defense, symbiosis, cell wall catabolism and lignification, signaling, and plant secondary metabolism.  $\beta$ -Glucosidases help defend against herbivores and invasive fungi by hydrolyzing relatively inert glycosides to produce toxic compounds, such as hydrogen cyanide, saponins, coumarins, quinones, hydroxamic acid, rotenoids, etc. (Poulton, 1990, Nisius, 1988, Duroux et al., 1998, Babcock and Esen, 1994, and Svasti et al., 1999). Moreover,  $\beta$ -glucosidases are thought to play roles in cell wall metabolism by the degradation of oligosaccharides, such as  $\beta$ -1,3- and  $\beta$ -1,4- linked oligosaccharides from plant cell walls (Hrmova et al., 1998), and release of monolignols by removing  $\beta$ -glucosyl residues from monolignol glycosides, like cinnamyl alcohol  $\beta$ -glucoside (Hösel et al., 1978).  $\beta$ -Glucosidases control the biological activity of phytohormones, including cytokinins, gibberellins, auxins, and abscisic acids by releasing active forms from inactive glucoside conjugates (Gaskin and MacMillian, 1975; Ganguly et al.,

1974; Brzobohatý et al., 1994; Millborrow, 1970).  $\beta$ -Glucosidases are also important for release of plant volatiles, and metabolism of many other significant natural products (Ketudat Cairns and Esen, 2010).

## 1.2 Glycoside hydrolases

$\beta$ -Glucosidases are classified as glycoside hydrolases (EC 3.2.1-3.2.3) which are an extensive group of enzymes that hydrolyze the glycosidic bond between two carbohydrates or between a carbohydrate and a noncarbohydrate moiety. The biological functions of oligo- and polysaccharides and glycosides are various, so glycoside hydrolases act in many essential steps of life, including hydrolysis of structural or storage polysaccharides, defense against pathogens, invasion of certain pathogens into cells, turnover of cell surface carbohydrates, etc (Henrissat et al., 1995).

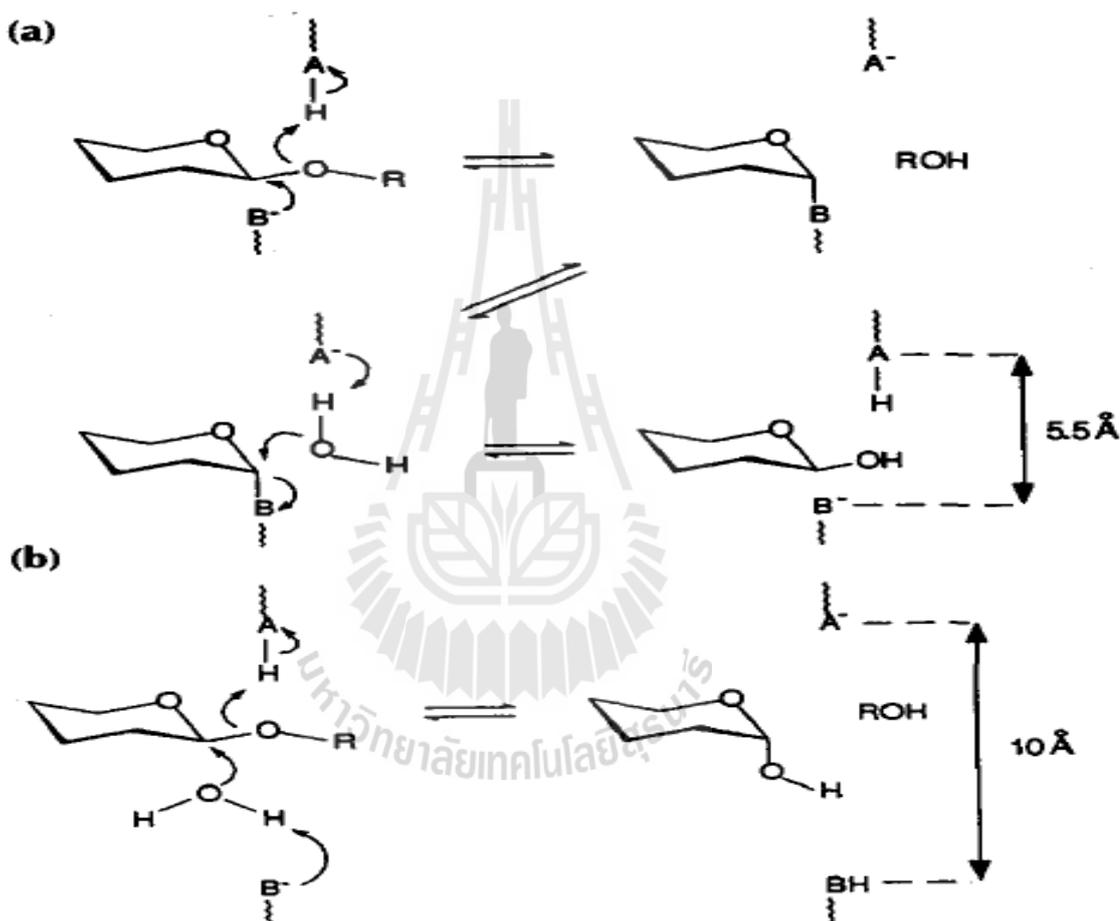
The large diversity in the stereochemistry of carbohydrates and the astronomical number of their possible combinations are paralleled by a large multiplicity of the enzymes designed for their selective hydrolysis. Since the number of protein folds has been determined to be not more than a few thousand (Chothia, 1992 and Orengo et al., 1994), this suggests that glycoside hydrolases of different substrate specificity could well have similar folds and may have evolved from a common ancestor.

Glycoside hydrolases have been grouped and classified in families based on amino acid sequence similarities (Henrissat, 1991 and Henrissat and Bairoch, 1993). If the sequences of two or more glycoside hydrolases can be aligned significantly over an entire domain, they are assigned to the same family. Cellulases belonging to

different glycoside hydrolase families have been found to have different folds (Davies et al., 1993 and Divne et al., 1994), as do chitinases from GH18 and GH19, but GH19 chitinases and lysozymes from GH22, GH23, and GH24 have related folds (Holm and Sander, 1994). The three-dimensional structures of two plant  $\beta$ -glucanases with distinct substrate specificities have been found to be strongly related (Varghese et al., 1994). Given that the fold of proteins is better conserved than the sequence, it is possible that several families share similar folds. Therefore, glycoside hydrolases have been assigned to 14 structurally related clans, comprising 133 families ([http://www.cazy.org/fam/acc\\_GH.html](http://www.cazy.org/fam/acc_GH.html)). The largest of the glycoside hydrolase clans is the GH-A clan, in which the proton donor and the nucleophile are found on  $\beta$ -strands 4 and 7 of a  $(\beta/\alpha)_8$  barrel, respectively (Jenkins et al., 1995). It includes GH families 1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53, 59, 72, 79, 86, 113 and 128, which contain enzymes that possess different substrate specificities.

Catalysis by most glycoside hydrolases can be divided into two mechanisms, retaining and inverting, which differ in whether the anomeric configuration of the product is the same as the substrate or not. Both types of mechanisms result from the action of a catalytic acid/base and nucleophile (Koshland, 1953; Sinnott, 1990; and Davies and Henrissat, 1995). The retaining mechanism occurs in two steps. First, the glycosidic oxygen is protonated by the acid catalyst to facilitate aglycone departure and the nucleophile forms a covalent bond with the glycone anomeric center producing a glycosyl-enzyme intermediate (Figure 1.1a). Second, a water molecule hydrolyzes the glycosyl-enzyme intermediate with basic assistance from the catalytic acid/base to generate a product which has the same stereochemistry as the substrate. For the inverting mechanism, the hydrolysis activity occurs when water attacks the

anomeric carbon and plays the role of the nucleophile. The water is activated by the carboxylate of an acidic amino acid residue acting as a catalytic base (Figure 1.1b). The carboxylate side chain of another acidic residue facilitates departure of the aglycone part by protonating the oxygen atom at the glycosidic bond (Davies and Henrissat, 1995).



**Figure 1.1** The two major mechanisms of glycoside hydrolase enzymes. (a) Retaining mechanism and (b) inverting mechanism (Davies and Henrissat, 1995).

### 1.3 Plant glycoside hydrolase family 1

$\beta$ -Glucosidases are grouped into GH families 1, 3, 5, 9, 30, and 116, and most plant  $\beta$ -glucosidases that have been characterized fall in GH family 1 (GH1) (Ketudat Cairns and Esen, 2010). Besides  $\beta$ -glucosidases, plant GH1 members include myrosinases (thio- $\beta$ -glucosidases) hydrolyzing the *S*-glycosidic bonds of plant 1-thio- $\beta$ -D-glucosides (glucosinolates), (Burmeister et al., 1997),  $\beta$ -mannosidases,  $\beta$ -galactosidases,  $\beta$ -glucuronidases,  $\beta$ -fucosidases, diglycosidases like primeverosidase (Mizutani et al., 2002), furcatin hydrolase (Ahn et al., 2004) and isoflavone 7-*O*- $\beta$ -apiosyl -1,6- $\beta$ -glucosidase (Chuankhayan et al., 2005), hydroxyisourate hydrolase, which hydrolyzes an internal bond in a purine ring, rather than a glycosidic bond (Raychaudhuri and Tipton, 2002), and galactosyl and glucosyl transferases (Moellering et al., 2010; Matsuba et al., 2010). GH1 members catalyze their reactions with a molecular mechanism leading to overall retention of the anomeric configuration, which involves the formation and breakdown of a covalent glycosyl enzyme intermediate as described above (Figure 1). All of the enzymes display a common  $(\beta/\alpha)_8$  TIM barrel structure. Apart from plant myrosinases and animal Klotho (KL) subfamily members, all GH1  $\beta$ -glucosidases contain two conserved catalytic glutamate residues located at the C-terminal end of  $\beta$ -strands 4 and 7 (Jenkins et al., 1995).

The *Agrobacterium* sp.  $\beta$ -glucosidase (Abg) was the first to have the catalytic nucleophile identified, as Glu358 in the sequence YITENG, through trapping of the 2-deoxy-2-fluoroglucosyl-enzyme intermediate and subsequent peptide mapping (Withers et al., 1990). The general acid/base catalyst was identified as Glu170 in this same enzyme through detailed mechanistic analysis of mutants at that position, which

included azide rescue experiments (Wang et al., 1995). Interestingly, the plant myrosinases, which catalyze hydrolysis of thioglycosides to release an anionic aglycone (glucosinolates), have evolved an active site in which the acid/base glutamate is replaced by glutamine. The activity of these enzymes with substrates does not require the acid catalyst, while the base catalyst is provided by ascorbate acting as a cofactor, which binds to the glycosyl enzyme after aglycone departure (Burmeister et al., 2000).

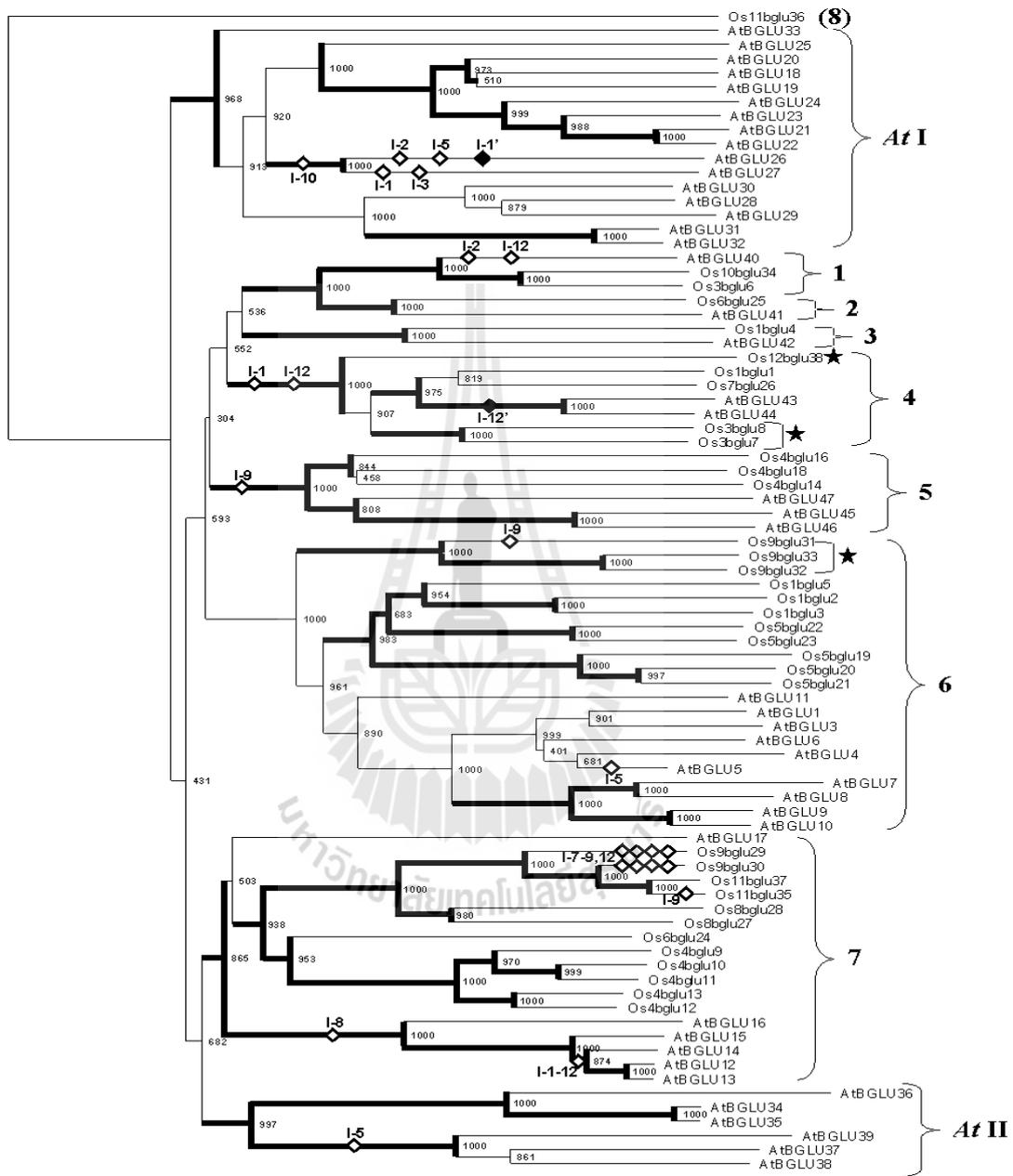
#### 1.4 Rice $\beta$ -glucosidases

Opassiri et al. (2006) identified GH1 genes from the rice genome, and their structures, predicted protein products and evidence of expression were evaluated. They found forty putative  $\beta$ -glucosidase genes, including 34 full-length genes, 2 pseudogenes, 2 gene fragments, and 2 intronless genes, likely to come from endophytes. Thirty-six out of 40 genes are found in both *japonica* and *indica* rice with 98-100% sequence identity.

Phylogenetic analysis of predicted protein sequences of rice and *Arabidopsis* GH1 genes showed that there are eight clusters containing both rice and *Arabidopsis* proteins that are more closely related to each other than they are to enzymes from the same plants outside the clusters, and two clusters (AtI and AtII) are found only in *Arabidopsis* (Figure 1.2). Interestingly, Os4BGlu14, Os4BGlu16 and Os4BGlu18 are clustered with *Arabidopsis* BGLU45, 46 and 47, which have been shown to hydrolyze glucosides of lignin precursors (Escamilla-Treviño et al., 2006), suggesting that Os4BGlu14, Os4BGlu16 and Os4BGlu18 may play roles in lignification. In addition, thirty-seven of the rice GH1 genes showed high protein sequence similarity to each

other and other known plant  $\beta$ -glucosidases. However, all  $\beta$ -glucosidases contain conserved acid/base and nucleophile residues that are glutamate residues. Normally, GH1 enzymes have W-XT/I-F/L/I/V/S/M-N/A/L/I/D/G-E/Q-P/I/Q and V/I/L-X-EN-G as conserved sequences around the catalytic acid/base and nucleophile, respectively (Czjzek et al., 2000; Hoffman et al., 1999).

In Os4BGlu14 and Os9BGlu33, the acid/base is replaced by glutamine, as seen in thioglucosidases (Opassiri et al., 2006). Therefore, Os4BGlu14 and Os9BGlu33 may be inactive as  $\beta$ -glucosidases, because the catalytic acid/base is converted to a nonionizable residue.  $\beta$ -Glucosidases with glutamate replaced by glutamine at the acid/base residue have been shown to have transferase activity in the presence of good leaving group aglycone and a nucleophilic acceptor, however (Müllegger et al., 2005; Chuenchor et al., 2011). In terms of the genome organization, Os4BGlu14, Os4BGlu15, Os4BGlu16, Os4BGlu17, and Os4BGlu18 are located in the same region of chromosome 4. However, Os4BGlu15 and Os4BGlu17 were identified as gene fragments, into which additional DNA had been inserted into Os4BGlu16 gene sequences. Os4BGlu15 and Os4BGlu17 lack half of the exons and are inactive as they do not have EST/full-length cDNA representation in the public databases (Opassiri et al., 2006).



**Figure 1.2** Phylogenetic tree of the predicted protein sequences of rice and *Arabidopsis* glycoside hydrolase family 1 genes (Opassiri et al., 2006). The clusters supported by a maximum parsimony analysis are shown as bold lines, and the loss and gain of introns are shown as open and closed diamonds, respectively.

Several rice  $\beta$ -glucosidase isoenzymes have been characterized for the possible function. For example, a cell wall-bound  $\beta$ -glucosidase (Os4BGlu12) and BGlu1 (Os3BGlu7) and BGlu2 (Os9BGlu30)  $\beta$ -glucosidases have been characterized and cloned from rice seedlings (Akiyama et al., 1998; Opassiri et al., 2003, 2006). Akiyama et al. (1998) determined the N-terminal sequence of the cell wall-bound  $\beta$ -glucosidase, this enzyme that is able to hydrolyze laminari- and cello-oligosaccharides. Opassiri et al. (2003) reported that BGlu1 (Os3BGlu7) and BGlu2 (Os9BGlu30) were both highly expressed in germinating shoots; the BGlu1 gene was also highly expressed in flower. The BGlu1 protein was expressed in *Escherichia coli*, purified and characterized. It showed activity with a variety of *p*-nitrophenyl  $\beta$ -D-glycosides, demonstrating some flexibility in sugar binding. It also hydrolyzed a variety of natural glycosides at low levels and showed strong hydrolysis of laminaribiose, laminaritriose, and cello-oligosaccharides. These two reports suggested that rice  $\beta$ -glucosidases not only have hydrolytic activity, but can also catalyze tranglycosylation of  $\beta$ -(1,3)- and  $\beta$ -(1,4)-linked oligosaccharides.

Another rice  $\beta$ -glucosidase that has been characterized is Os4BGlu12, which was found to be induced by herbivore attack and salinity stress (Opassiri et al., 2006). It has recently been reported that the transcription of Os4BGlu12 is up-regulated by wounding, methyl jasmonate and ethephon responses (Opassiri et al., 2010). Moreover, the sequence of Os4BGlu12 is most closely related to the cell wall bound  $\beta$ -glucosidase (Akiyama et al., 1998), among the GH1 sequences identified in the rice genome (Opassiri et al., 2006). Os4BGlu12 may be involved in remodeling of damaged cell wall, because it can hydrolyze mixed  $\beta$ -(1,3)- and  $\beta$ -(1,4)-linked oligosaccharides generated by wounding-induced rice endo-(1,3)-,-(1,4)- $\beta$ -glucanase.

Among natural glycosides, Os4BGlu12 hydrolyzed deoxycorticosterone 21-glucoside (an animal steroid glucoside) and apigenin 7-O- $\beta$ -D-glucoside, which suggests that this enzyme's function may also be related to defense (Opassiri et al., 2010). Recently, it has been shown that Os4BGlu12 has high activity toward salicylic acid (SA) glucoside, suggesting a different role in signaling in response to plant wounding as well (Himeno et al., 2013).

Rice Os3BGlu6 has also been characterized and was found to hydrolyze *p*-nitrophenyl  $\beta$ -D-glycosides, *n*-octyl- $\beta$ -D-glucoside, and  $\beta$ -(1,2)- and  $\beta$ -(1,3)-linked disaccharides. The crystal structures of Os3BGlu6 alone and in complex with 2-deoxy-2-fluoroglucoside and *n*-octyl- $\beta$ -D-thioglucoopyranoside suggested that methionine residue 251 located in the mouth of the active site may block the binding of  $\beta$ -(1,4)-linked oligosaccharides and, therefore Os3BGlu6 cannot hydrolyze long chain cellooligosaccharides (Seshadri et al., 2009). Hua et al., 2013 found that Os3BGlu6 hydrolyzed gibberellin A4  $\beta$ -D-glucosyl ester (GA<sub>4</sub>-GE) at a higher rate than Os3BGlu7, Os4BG12, Os4BGlu18, and Os9BGlu31, which are associated with different phylogenetic clusters within GH1. The crystal structure of Os3BGlu6 and the rates of hydrolysis of *p*NPGlc and GA<sub>4</sub>-GE by Os3BGlu6 and its mutated form E178Q, E178A, E394D, E394Q, and M251N revealed that M251 plays a role in binding to the hydrophobic aglycones. The crystal structure of Os3BGlu6 E178Q in complex with glucose released from GA<sub>4</sub>-GE or *p*NPGlc showed that the  $\alpha$ -D-glucosyl moiety is attached to the catalytic nucleophile, E394, suggesting that the hydrolysis of 1-O-acyl glucose ester and glucoside had the same retaining mechanism.

Kuntothom et al. (2009) reported that Os3BGlu8, which is closely related to Os3BGlu7, hydrolyzed *p*NPGlc much better than *p*-nitrophenyl  $\beta$ -D-mannopyranoside

(*p*NPMan), while Os7BGlu26, which is closely related to barley rHvBII  $\beta$ -D-mannosidase can hydrolyze *p*NPMan better than *p*NPGlc. However, both enzymes can hydrolyze  $\beta$ -(1,2)-, -(1,3)-, and -(1,4)-linked gluco-oligosaccharides, while only Os4BGlu8 can hydrolyzed  $\beta$ -(1,6)-linked disaccharide (gentiobiose).

OsTAGG1 (Os4BGlu13) is a tuberonic acid glucoside (TAG) hydrolyzing  $\beta$ -glucosidase found to hydrolyze a variety of other natural substrates as well, including methyl tuberonic acid glucoside, jasmonoyl-1- $\beta$ -glucoside, and salicylic acid glucoside, as well as *p*-nitrophenyl  $\beta$ -D-glycosides (Wakuta et al., 2010). Tuberonic acid (TA) and TAG are produced from jasmonic acid (JA) by hydroxylation and glycosylation, respectively. Since TAG was hydrolyzed most efficiently, the authors concluded that OsTAGG1 is a specific  $\beta$ -glucosidase hydrolyzing tuberonic acid glucoside to release the active TA, but this TA cannot be converted into jasmonic acid (Wakuta et al., 2010). Moreover, Himeno et al., 2013 reported that OsTAGG2 (Os4BGlu12) is upregulated by wounding and methyl jasmonate, and that recombinant OsTAGG2 produced in *P.pastoris* (rOsTAGG2P) hydrolyzed salicylic acid  $\beta$ -D-glucoside (SAG), another inactive phytohormone, around 4.5-fold more rapidly than TAG.

## 1.5 Plant lignification

Many plants have a variety of biochemical solutions to reinforce their cell walls and fulfill different adaptive strategies. Lignification is an important component of the cell wall thickening process (Lewis and Davin, 1994). Some genes for lignification are similar to the flavonoid synthesis pathway, suggesting that the lignification genes may have evolved from the flavonoid synthesis genes, or have

diverged from a common ancestral gene associated with primary metabolism has occurred (Lacombe et al., 1997). The lignin polymer is a major component of many plant cell walls, such as cell walls of tracheids, vessels and fibers, and it has significant impact on the pulp and paper industry, for which removal of lignin is a major process. The functions of lignin were found to include contributing to the strength of woody stems and water proofing of the liquid-conducting elements within the xylem (Donaldson, 2001).

Lignins are complex, three-dimensional aromatic polymers derived from three hydroxycinnamyl alcohol monomers, which differ only in their degree of methoxylation, *p*-coumaryl (M1H), coniferyl (M1G), and sinapyl (M1S) alcohols (Boerjan et al., 2003) (Figure 1.3). These monolignols produce the *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) phenylpropanoid units that are found in the lignin polymer. The amount and composition of lignins vary among taxa, cell types, and individual cell wall layers and are influenced by developmental and environmental cues (Campbell and Sederoff, 1996). Although there are exceptions, dicotyledonous angiosperm (hardwood) lignins consist principally of G and S units and traces of H units, whereas gymnosperm (softwood) lignins are composed mostly of G units with low levels of H units.

Lignins from grasses (monocots) contain G and S units at comparable levels, and more H units than dicots (Baucher et al., 1998). In wheat straw, triticale straw, rye straw, and maize stalk, the levels of H units are three to fifteen times higher than *Arabidopsis* stem, tobacco stem, and poplar wood (Barrière et al., 2007). The significant amount of H units is a special characteristic for monocotyledonous lignins that separates them from constitutive dicotyledonous. She et al. (2011) reported that

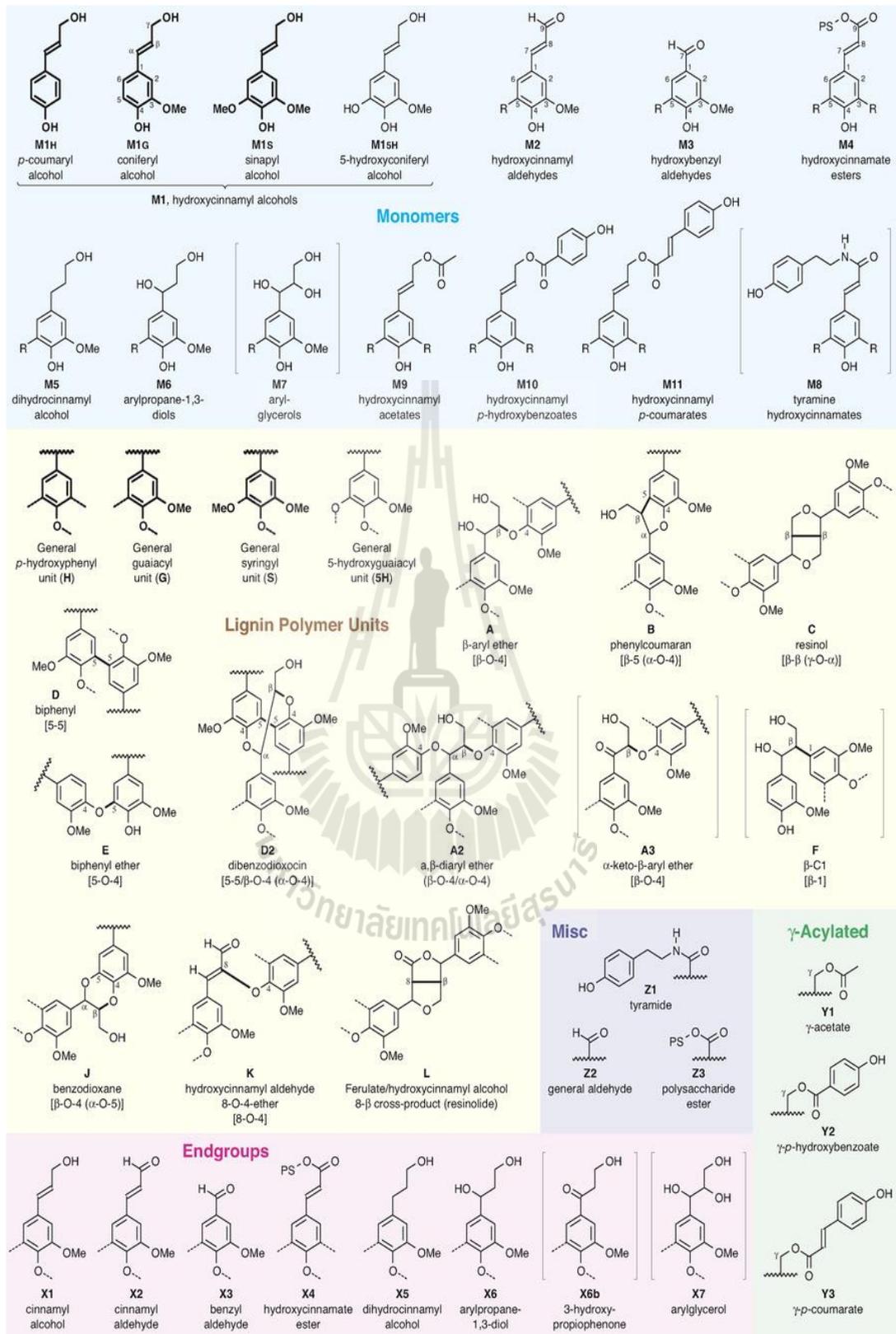
dewaxed rice straw with different alcohol treatments was composed of phenolic acids and aldehydes in the lignin fractions. The ratios of G (vanillin, vanillic acid, and acetovanillin), S (syringaldehyde, syringic acid, and acetosyringone), and H (*p*-hydroxybenzaldehyde and *p*-hydroxybenzoic) were found to change as large amounts of non-condensed guaiacyl, syringyl, and *p*-hydroxyphenyl units were produced. The composition suggest that the lignin preparations from rice straw can be considered as GSH lignin.

Lignification is the process by which the monolignol units are linked together via radical coupling reactions (Freudenberg and Neish, 1968; Sarkanen and Ludwig, 1971). The main “end-wise” reaction couples a new monomer to the growing polymer, giving rise to structures A, B, and D2 shown in Figure 1.3 (Boerjan et al., 2003). Coupling between preformed lignin oligomers results in units linked 5-5 (D) and 5-O-4 (E). The coupling of two monolignols is a minor event, with resinol ( $\beta$ - $\beta$ ) units (C) or cinnamyl alcohol end groups (X1) as the outcome. Monolignol dimerization and lignin production are substantially different processes (Adler, 1977). Normally, almost all of the inter-unit linkages of lignin are  $\beta$ -O-4 ( $\beta$ -aryl ether), that are easy to chemically hydrolyze for industrial processes. However, lignin can have other linkages, which are difficult to degrade by chemical means, such as  $\beta$ -5 (B),  $\beta$ - $\beta$  (C), 5-5 (D), 5-O-4 (E), and  $\beta$ -1 (F). The properties of different lignin polymers depend on the relative abundance of particular monomer units. For example, conifer lignin consists mainly of G units linked  $\beta$ -5 (B), 5-5 (D), and 5-O-4 (E), and are more difficult to hydrolyze than lignins with S units because the C5 position are available for coupling (Boerjan et al., 2003).

## 1.6 Lignans, structure and functions

The phenylpropanoid pathway is the initial pathway of lignin biosynthesis and also of another hydrophobic polymer, the lignans. Lignans are a large group of naturally occurring phenols, widely spread within the plant kingdom, that are derived from the shikimic acid biosynthetic pathway (Ayres and Loike, 1990). Normally, the plant lignans are polymers of phenolic compounds mainly derived from cinnamyl units, such as by dimerization of cinnamic alcohols to dibenzylpropane. The structures formed vary by  $\beta,\beta'$ -linkage between two phenyl propane units with a different degrees of oxidation in the side-chains and by a different substitution pattern in the aromatic moieties (Ayres and Loike, 1990). Because of their activities, such as antioxidant, antimicrobial, antitumor, anti-inflammatory and antiviral properties, lignans have been used for a long time both in ethnic and conventional medicine (Osawa, 1992). In addition, most researchers have been focused on their antioxidant activity, which is mainly due to the radical scavenging properties of these compounds.

The study of dirigent proteins implicated in lignin biosynthesis found that lignans are dehydro-dimers of monolignols and are typically optically active. Pinoresinol formation is catalyzed by the first dirigent protein discovered by the dimerization of coniferyl alcohol radicals to generate an optically active lignan (Davin and Lewis, 2000; Davin et al., 1997).



**Figure 1.3** Lignin monomers and structures in the polymer (Boerjan et al., 2003).

Although monolignol glycosides could be involved in lignan formation, available evidence suggests that lignans are glycosylated after they form. The study of *Sesamum indicum* uridine diphosphate (UDP) glucose:lignan glucosyltransferases revealed that these enzyme could glucosylate at the 2-hydroxyl group of (+)-sesaminol (Noguchi et al., 2008). Moreover, a lignan glucosyltransferase from *Forsythia koreana*, which contains the majority of accumulated lignans in glucoside forms, including pinoresinol 4-*O*-glucoside, epipinoresinol 4-*O*-glucoside, phillyrin, matairesinoside, and arctiin, (Guo et al., 2007; Ono et al., 2010) was described and biochemical characterized. The recombinant UDP-sugar dependent-glycosyltransferases (UGT) 71A17 and UGT71A18 encoded by *Forsythia* cDNA could glycosylate (+)-pinoresinol (Ono et al., 2010). More recently, Okazawa et al. (2014) reported that the expressed *Arabidopsis thaliana* UGT71C1 showed glucosyltransferase activity towards pinoresinol and lariciresinol lignan, suggesting that UGT is involved in lignan glucosylation. It is interesting to determine the localization of UGT. Monolignol dimerization is catalyzed by peroxidase enzyme that is localized in the cell wall and vacuole, which led to the proposal that lignan should be localized around these organelles (Wang et al., 2013).

## 1.7 Monolignol $\beta$ -glucosidases

Lignin production could be regulated by both monolignol synthesis and by the transport of monolignol precursors to the cell wall and their release from inactive forms (Terashima, 1989). The 4-*O*- $\beta$ -D-glucosides of monolignols, namely *p*-coumaryl alcohol glucoside, coniferin and syringin have been suggested to serve as inactive transport or storage forms of monolignols. They are synthesized by uridine

diphosphate-glucose (UDPGlc)-utilizing glucosyltransferases and subsequently hydrolyzed by monolignol-specific  $\beta$ -glucosidases (Dharmawardhana and Ellis, 1998). A coniferin  $\beta$ -glucosidase was described in *Pinus banksiana* (Leinhos et al., 1994). Subsequent characterization of a coniferin  $\beta$ -glucosidase (CBG) from *Pinus contorta* allowed cDNA cloning and the predicted amino acid sequence suggested that CBG is an extracellular glycoprotein belonging to GH1 (Dharmawardhana et al., 1995; Dharmawardhana and Ellis, 1998). CBG was also found to be localized in the lignification zone in the tree stem. Although monolignol glucosides are found in all gymnosperms, not all angiosperms appear to have them. However, in poplar, coniferin  $\beta$ -glucosidase activity has been histochemically localized to lignifying cells (Dharmawardhana and Ellis, 1998) and radiolabeled monolignol glucosides are efficiently incorporated into lignin (Fukushima and Terashima, 1990). These data suggest that the monolignol glucosides may be hydrolyzed by  $\beta$ -glucosidases in lignifying tissues of angiosperms. In fact, monolignol-specific glycosyltransferase activity was detected in all angiosperm species tested, and a  $\beta$ -glucosidase that could hydrolyze monolignol glucosides were originally isolated from cell walls of chick pea cell suspension cultures (Hösel et al., 1978). While the chick pea monolignol  $\beta$ -glucosidase had higher hydrolysis activity against coniferin than syringin, a  $\beta$ -glucosidase purified from cell cultures of soybean (*Glycine max*) hypocotyls and roots had identical  $V_{max}$  values for the two substrates and a  $K_m$  for coniferin two times higher than that for syringin (Hösel and Todenhagen, 1980; Hösel et al., 1982).

In a more recent study of monolignol glucosidases in *Arabidopsis*, the proteins encoded by the loci At1g61810 (BGLU45), At1g61820 (BGLU46), and At4g21760 (BGLU47) were found to cluster with *P. contorta* coniferin  $\beta$ -glucosidase in protein

sequence-based phylogenetic analysis, leading to the hypothesis that the respective gene products may hydrolyze monolignol glucosides (Escamilla-Treviño et al., 2006). The cDNA encoding mature BGLU45 and BGLU46 were cloned and expressed in a yeast system. The recombinant BGLU45 and BGLU46 proteins were purified to 8.8 and 13.7-fold, respectively. Among the natural substrates tested, BGLU45 exhibited specific activity toward the monolignol glucosides syringin, coniferin, and *p*-coumarol glucoside, while BGLU46 exhibited broader substrate specificity, cleaving salicin, *p*-coumarol glucoside, phenyl- $\beta$ -D-glucoside, coniferin, syringin, and arbutin. In addition, among nitrophenol (NP) glycoside substrates, BGLU45 hydrolyzed *p*NPGlc, *p*NPGal, and *o*-nitrophenyl  $\beta$ -D-glucoside (*o*NPGlc). BGLU46 hydrolyzed *p*NPGlc, *p*NPGal, *p*NPXyl, and *o*NPGlc. RT-PCR showed that BGLU45 and BGLU46 expression was highest in *Arabidopsis* organs that are major sites of lignin deposition. Both genes showed to increase the levels of gene expression from apex to base. Moreover, BGLU45 is expressed in siliques, while BGLU46 is expressed in roots.

Recently, Chapelle et al. (2012) reported that T-DNA insertions in the *Arabidopsis thaliana* BGLU45 and BGLU46 genes resulted in a significant increase in coniferin content in stem extracts, while syringin content was not changed. Other compounds of the phenylpropanoid pathway such as ferulic acid hexoside, sinapyl alcohol hexose, coniferaldehyde derivate, and syringaresinol were also detected in these mutants but at low levels. The knockout BGLU45 and BGLU46 plant lines do not exhibit a lignin-deficient phenotype. Immunolocalization showed that BGLU45 protein is mainly located in the interfascicular fibers, whereas BGLU46 protein is located in the protoxylem. Moreover, no change was observed in plants with a T-

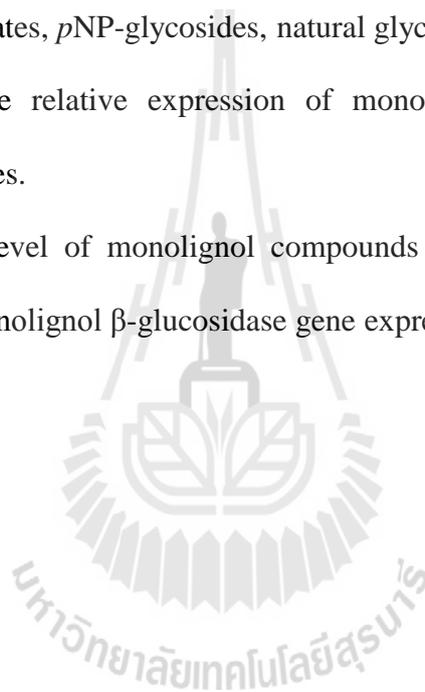
DNA insert in the BGLU47 gene, suggesting it is not involved in the phenylpropanoid pathway or lignification in the stem (Chapelle et al., 2012). The lack of large changes in the lignin content and composition in these *Arabidopsis* gene knockout lines suggested that the monolignol glucosides may be storage forms, but are not required transport forms of monolignols or direct precursors of lignin.

Little has been reported about monolignol  $\beta$ -glucosidases in monocots. Opassiri et al. (2006) identified the GH1 genes from the rice (*Oryza sativa* L.) genome. Phylogenetic analysis of predicted protein sequences of rice and *Arabidopsis* GH1 genes showed that Os4BGlu14, Os4BGlu16, and Os4BGlu18 are clustered with *Arabidopsis* BGLU45 and BGLU46  $\beta$ -glucosidases. Since BGLU45 and BGLU46 have been shown to hydrolyze lignin precursors (Escamilla-Treviño et al., 2006), Os4BGlu14, Os4BGlu16 and Os4BGlu18 were hypothesized to have similar activities. However, Os4BGlu14 may be inactive as a  $\beta$ -glucosidase, since it is one of the 2 putative active rice  $\beta$ -glucosidase genes identified by Opassiri et al. (2006) in which the conserved catalytic acid/base glutamate residue is replaced by a nonionizable glutamine residue.

## 1.8 Research objectives

The objectives of this study included:

1. To clone and express the putative rice monolignol  $\beta$ -glucosidases, optimize for suitable expression conditions, and purify proteins.
2. To characterize the enzymatic properties and substrate specificities of recombinant Os4BGlu16 and Os4BGlu18 enzymes, including activities toward monolignol substrates, *p*NP-glycosides, natural glycosides and oligosaccharides.
3. To investigate the relative expression of monolignol  $\beta$ -glucosidase gene in different rice tissues.
4. To measure the level of monolignol compounds by UPLC-MS to see how it corresponds to monolignol  $\beta$ -glucosidase gene expression patterns.



## CHAPTER II

### MATERIALS AND METHODS

#### 2.1 Materials

##### 2.1.1 Plant, plasmids, bacterial and yeast strains

The accession AK067841 cDNA clone plasmid that encodes full length precursor of Os4BGlu14 was provided by the Rice Genome Resource full-length cDNA project (Kikuchi et al., 2003). A cDNA optimized for Os4BGlu16 expression in *Pichia pastoris* (Genbank accession number KJ579205) was synthesized and inserted into the pUC57 vector by GenScript Corporation (Piscataway, NJ, USA). Seven-day-old rice (*Oryza sativa* L. ssp. *Indica* cv. KDML105) seedlings were used to extract RNA for cloning of the cDNA encoding the mature Os4BGlu18 protein. The vectors used for cloning and expression of monolignol  $\beta$ -glucosidase included pENTR<sup>TM</sup>/D-TOPO (Invitrogen, CA, USA), pET32a(+) (Novagen, WI, USA), and pPICZ $\alpha$ B(NH<sub>8</sub>) (Toonkool et al., 2006). *Escherichia coli* stains DH5 $\alpha$  and XL1-Blue were used for cloning and Origami(DE3) was used for protein expression. *P. pastoris* strain SMD1168H was used to produce the Os4BGlu16 protein.

### 2.1.2 Oligonucleotides primers

Oligonucleotides were synthesized by BioDesign Co., Ltd. (Thailand) and Bio Basic Inc. (Canada) and are shown in Table 2.1.

**Table 2.1** Oligonucleotides primers for monolignol  $\beta$ -glucosidase cloning, mutagenic primers of Os4BGlu14, and unique primers for RT-PCR. The extra CACC bases used for directional cloning in the pENTR<sup>TM</sup>/D-TOPO<sup>®</sup> vector are shown in bold and the sequences of the restriction enzyme sites for cloning are underlined.

Primer name	Sequence (5'-3')
Os4BGlu14Matstrf	<u>CACCATGGCCGTCGACCGCAGCCAG</u>
Os4BGlu14Stop	<u>GAGGATCCTTAATAAGAGTTAACTTATGAAGC</u>
BG14OxNotIF	<u>GCGGCCGCATGGCGGCGGCGTGGCTCG</u>
BG14OxAscISpeIR	<u>ACTAGTGGCGCGCCAATAAGAGTTAACTTATGA</u> AGCTGAGATTC
Os4BGlu14AB2Ef	CGAATAAAATTCTGGACAACATTTAATGAGCCGA ATTTGTCCATAAAGTTCAGTTAC
Os4BGlu14AB2Er	GTAAGTGAACCTTTATGGACAAATTCGGCTCATTAA ATGTTGTCCAGAATTTTATTCG
BG16OxNotIF	<u>GCGGCCGCATGGCCGTGGCGGCGGCGA</u>
BG16OxAscISpeIR	<u>ACTAGTGGCGCGCCACGAATCTGCTCTACGTGACC</u> GCACTTGG
Os4BGlu18Startf	<u>CACCATGGCAGGGAGGCAGTAAGACG</u>

**Table 2.1** Oligonucleotides primers for monolignol  $\beta$ -glucosidase cloning, mutagenic primers of Os4BGlu14, and unique primers for RT-PCR. The extra CACC bases used for directional cloning in the pENTR<sup>TM</sup>/D-TOPO<sup>®</sup> vector are shown in bold and the sequences of the restriction enzyme sites for cloning are underlined. (Continued).

Primer name	Sequence (5'-3')
Os4BGlu18MatStf	<b>CACCATGGCGATCCACAGGAGCGACTTC</b>
Os4BGlu18StopR	<b>GAGGATCCTCATGCAGATTTTGGAGGAATTC</b>
BG18OxNotIF	<u>GCGGCCGCATGGCAGGAGGCAGTAAGACG</u>
BG18OxAscISpeIR	<u>ACTAGTGGCGCGCCATTGATTTTCATGCAGATTTTG</u> GAGG
eGFPNotIF	<u>GCGGCCGCATGGTGAGCAAGGGCGAGGAGCTGTT</u>
eGFPAscIF	<u>GGCGCGCCATGGTGAGCAAGGGCGAGGAG</u>
eGFPSpeIR	<u>GGACTAGTCTACTATTACTTGTACAGCTCGTCCATG</u> CCGAGAGT
Os4BGlu14RNAiF	CACCGAATTACCTACATTCAAGGCTACC
Os4BGlu14RNAiR	CAGTACATATTATACTCTATGCATATGG
Os4BGlu16RNAiF	CACCGAGGGGAATGCTCTCG
Os4BGlu16RNAiR	ACCACTTGGCAGACATC
Os4BGlu18RNAiF	CACCTGCCAATGTTTCATCAC
Os4BGlu18RNAiR	TAGGTGGATATCATGTATCC
OsActinF	ACTCTGGTGATGGTGTCAGCC
OsActinR	GTCAGCAATGCCAGGGAACATA

### 2.1.3 Chemicals and reagents

Tryptone, yeast extract, sodium dodecyl sulfate (SDS), acrylamide, N, N', N'', N'''- tetramethylethylenediamine (TEMED), ammonium persulfate (APS), Triton X-100, lysozyme, sodium hydroxide (NaOH), sodium acetate (NaAc), sodium chloride (NaCl), sodium carbonate (Na<sub>2</sub>CO<sub>3</sub>), ethylene diamine tetraacetic acid (EDTA), glacial acetic acid, methanol, acetonitrile, HPLC-grade water, ethanol, 2,2'-azinobis (3-ethylbenthaiazolinesulfonic acid) (ABTS), 3,3',5,5'-tetramethyl benzidine (TMB), isopropyl thio-β-D-galactoside (IPTG), β-mercaptoethanol (BME), *p*-nitrophenol-β-D-glucoside (*p*NPGlc), *p*NP-β-D-fucoside, *p*NP-α-L-arabinoside, *p*NP-β-D-galactoside, *p*NP-β-D-xyloside, *p*NP-β-D-mannoside, *p*NP-β-D-cellobioside, Coomassie brilliant blue R250, phenylmethylsulfonyl fluoride (PMSF), calcium chloride (CaCl<sub>2</sub>), ethidium bromide (EB), *Pfu* DNA polymerase, *Taq* DNA polymerase, ammonium sulfate ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), potassium dihydrogen phosphate (KH<sub>2</sub>PO<sub>4</sub>), manganese sulfate (MnSO<sub>4</sub>), zinc sulfate heptahydrate (ZnSO<sub>4</sub>·7H<sub>2</sub>O), potassium iodide (KI), boric acid (H<sub>3</sub>BO<sub>3</sub>), cobalt chloride hexahydrate (CoCl<sub>2</sub>·6H<sub>2</sub>O), copper sulfate pentahydrate (CuSO<sub>4</sub>·5H<sub>2</sub>O), disodium molybdate 45-hydrate (Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O), calcium chloride dihydrate (CaCl<sub>2</sub>·2H<sub>2</sub>O), magnesium sulfate heptahydrate (MgSO<sub>4</sub>·7H<sub>2</sub>O), iron sulfate heptahydrate (FeSO<sub>4</sub>·7H<sub>2</sub>O), dipotassium hydrogen phosphate (K<sub>2</sub>HPO<sub>4</sub>), sodium dihydrogen phosphate dihydrate (NaH<sub>2</sub>PO<sub>4</sub>·2H<sub>2</sub>O), ammonium chloride (NH<sub>4</sub>Cl), and potassium chloride (KCl) were purchased from a variety of suppliers. Cellooligosaccharides of degree of polymerization (DP) 3-6 and laminari-oligosaccharides of DP 2-6 were purchased from Segu-gaku Corp. (Tokyo, Japan) or Megazyme (Bray, Ireland), SYBR Green PCR master mix was purchased from Kapa Biosystems, Inc (Boston, MA, USA), optical tubes and cap strips were purchased

from Bio-Rad (Hercules, CA, USA), and Superscript first-strand cDNA synthesis system for RT-PCR and 100 bp DNA ladder were purchased from Invitrogen (Carlsbad, CA, USA).

## **2.2 General methods**

### **2.2.1 RNA extraction**

The rice (*Oryza sativa* L. ssp. *Indica* cv. KDML105) seeds were washed with 5% Chlorox bleach for 20 min, washed with distilled water twice, and soaked in distilled water overnight. The rice seeds were spread onto a polystyrene box with wet tissue paper at the bottom. The rice was germinated in the dark from day 0 to day 3 and in 12 h light 12 h dark from day 4 to day 7 at room temperature (25-30 °C) on paper moistened with distilled water. The shoots from 7 days rice seedlings were harvested and used in total RNA purification. The total RNA was isolated according to the Spectrum™ plant total RNA kit (Sigma-Adrich, St. Louis, MO, USA). The total RNA was used immediately for cDNA synthesis or stored at -80 °C.

### **2.2.2 First-strand cDNA synthesis**

Total cDNA was reverse-transcribed from the total RNA template. Total RNA at a concentration of 100 ng/μl was mixed with 2.5 μM oligo(dT)<sub>20</sub> and 0.5 mM dNTP, adjusted to a final volume of 10 μl with DEPC-treated water and incubated at 65 °C for 5 min and immediately placed on ice at least 1 min. After that, 1X RT buffer, 5 mM MgCl<sub>2</sub>, 10 mM DTT, and 40 U RNase OUT™ were added into the reaction mixture and it was incubated at 50 °C for 2 min. Then, 200 U Superscript™ III reverse transcriptase (Invitrogen, CA, USA) was added to the mixture and it was incubated at 50 °C for an additional 50 min. The reaction was stopped by incubating it

at 85 °C for 5 min. The mRNA was then hydrolyzed by 1 µl of Rnase-H (Invitrogen, CA, USA) at 37 °C for 20 min. The single stranded cDNA was ready for PCR.

### **2.2.3 Preparation of *E. coli* strains DH5α, XL1-Blue and Origami(DE3) for competent cells**

Glycerol stocks of DH5α and XL1-Blue were streaked on LB plates without antibiotic. Origami(DE3) was streaked on an LB plate containing 15 µg/ml kanamycin and 12.5 µg/ml tetracycline and incubated at 37 °C for 16-18 h. A single colony was picked and inoculated into 5 ml of LB broth with shaking at 37 °C, 200 rpm for 16-18 h. One hundred microliters of starter culture was transferred to 100 ml of LB broth and shaken at 37 °C, 200 rpm until the optical density at 600 nm (OD<sub>600</sub>) reached to 0.4-0.6. The cell culture was chilled on ice for 5 min in sterile polypropylene tube and collected at 4,000 rpm at 4 °C for 10 min. The cell pellets were resuspended in 10 ml ice-cold sterile 0.1 M CaCl<sub>2</sub> and centrifuged to collect the cell pellets again. Finally, the pellets were resuspended with 1 ml of 0.1 M CaCl<sub>2</sub> containing 15% glycerol and 50 µl aliquots were stored at -80 °C.

### **2.2.4 Transformation of plasmids into competent cells**

An aliquot of frozen competent cells was thawed 5 min on ice, then 1 µl of cloning or expression plasmids (20-100 ng) or ligation reactions were mixed with fresh or thawed competent cells. The reaction was incubated on ice for 30 min. The plasmid was transformed by heat shocking the cells at 42 °C for 45 s and quickly chilling on ice for 5 min. Two hundred microliters of LB broth was added to the transformed competent cells, which were then incubated at 37 °C for 1 h. The transformed cells were spread on LB agar containing appropriate antibiotics and incubated at 37 °C overnight.

### 2.2.5 Plasmid isolation by alkaline lysis method

A single colony of transformed recombinant bacteria was picked into 5 ml of LB broth and incubated at 37 °C with shaking at 200 rpm for 16-18 h. The cultured cells were collected by centrifugation at 10,000 rpm, 1 min. The supernatant was removed and the cells were resuspended in 100 µl of lysis buffer I (50 mM glucose, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0). Then, 200 µl of freshly prepared lysis buffer II (0.2 N NaOH, 1% (w/v) SDS) was added and the tube was inverted 4-6 times. After that, 150 µl of ice-cold lysis buffer III (3 M potassium acetate, pH 4.8) was added and the tube was mixed by inverting 4-6 times. The alkaline lysis reaction was incubated on ice for 5 min and the clear solution containing the plasmids was separated from the cell debris by centrifugation at 13,000 rpm, 10 min. The supernatant was transferred to a new tube and the DNA was precipitated with 2 volumes absolute ethanol for 10 min at 4 °C. The precipitated DNA was collected by centrifugation at 13,000 rpm for 10 min. The left over absolute ethanol was removed by speed vacuum. Then, the DNA pellet was resuspended in 100 µl TE buffer containing 2 µg RNase A and incubated at 37 °C for 10 min. The RNase A-treated plasmids were further purified by adding 70 µl of ice-cold precipitation solution (20% PEG 6000, 2.5 M NaCl) and chilled on ice for 1 h. The precipitated DNA was collected by centrifugation at 13,000 rpm for 10 min. The supernatant was removed and the pellet was washed by adding 0.5 ml of 70% ethanol and inverting the tube twice, after that the ethanol solution was removed and the tube dried by speed vacuum. Finally, the DNA was redissolved with 30 µl of TE buffer or sterile water.

### 2.2.6 QIAGEN plasmid miniprep

The QIAprep<sup>®</sup> spin miniprep kit (QIAGEN) was used to purify recombinant plasmid DNA according to the manufacturer's instructions. A single colony was picked and inoculated in 5 ml LB broth with appropriate antibiotics, as described in sections 2.2.3 and 2.2.5. The cultured cells were pelleted by centrifugation at 10,000 rpm for 1 min. The cell pelleted was resuspended completely in 250  $\mu$ l P1 buffer (100 mg/ml RNAs A in 10 mM EDTA, 50 mM Tris-HCl, pH 8.0). Two hundred fifty microliters of P2 buffer (200 mM NaOH, 1% (v/v) SDS) was added to the resuspended cells, and mixed by inverting the tube gently 4-6 times until the solution became viscous and slightly clear. After that, 350  $\mu$ l of P3 buffer (3 M potassium acetate, pH 5.5) was added and mixed immediately, to avoid localized precipitation, by inverting the tube gently 4-6 times. The solution was centrifuged at 12,000 rpm for 10 min to compact the white pellet. The supernatant was applied to a QIA prep column by pipetting and centrifuging at 12,000 rpm for 1 min, and then the flow-through was discarded. To protect nuclease activity or carbohydrate content, 0.5 ml of PB buffer (1.0 M potassium acetate, pH 5.0) was added to the column and centrifuged at 12,000 rpm for 1 min. The column was washed 2 times by applying 0.75 ml PE buffer (1.0 M NaCl, 50 mM MOPS, pH 7.0, 15% (v/v) isopropanol) and centrifuging at 12,000 rpm for 1 min. The flow-through was discarded, and the column was centrifuged for an additional 1 min to remove residual wash buffer. Lastly, the column was placed in a new 1.5 ml microtube and 50  $\mu$ l distilled water was added to the center of column. The column was allowed to stand for 1 min, and centrifuged at 12,000 rpm for 1 min to elute the plasmid DNA.

### **2.2.7 Agarose gel electrophoresis for DNA**

The purified plasmids and PCR products were checked by agarose gel electrophoresis. One percent agarose gels were prepared in TAE buffer (40 mM Tris-HCl, pH 8.0, 40 mM acetic acid, 1 mM EDTA, pH 8.0) or in TBE buffer (90 mM Tris-HCl, pH 8.0, 89 mM boric acid, 2.5 mM EDTA, pH 8.0). The DNA samples were mixed 5:1 with 6X loading dye (0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol, 30% (v/v) sterilized glycerol). Agarose gel electrophoresis was performed in a Pharmacia GNA-100 Gel Electrophoresis Apparatus (GE Healthcare, Buckinghamshire, UK) at a constant voltage of 120 V for 30 min. The DNA bands on the agarose gel were detected by staining with ethidium bromide (0.1 µg/ml) 30 s and destained in distilled water for 5 min. The DNA bands were visualized by UV light transillumination with a Fluoro-S™ MultiImager (Bio-Rad). The sizes of the DNA bands were estimated by comparing their migration with those of 1 kb or 100 bp ladder (Fermentas, Burlington, ON, Canada).

### **2.2.8 Purification of DNA bands from gels**

The correct size DNA bands that had been separated on agarose gel electrophoresis were purified with a HiYield™ Gel/PCR DNA fragments extraction kit (RBC Bioscience Corp., Taiwan). The agarose gel containing the target DNA band was sliced with a blade cutter and not more than 300 mg was transferred to a microtube. The agarose gel purification was done according to the manufacturer's instructions.

### **2.2.9 SDS-PAGE electrophoresis**

The protein profile and the apparent molecular weights of proteins in various fractions were determined by SDS-PAGE, as described by Laemmli (1970). The

SDS-PAGE 12% separating gel consisted of 12% (w/v) acrylamide, 375 mM Tris-HCl, pH 8.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED, while the 4% stacking gel consisted of 4% (w/v) acrylamide, 125 mM Tris-HCl, pH 6.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED. Protein samples were mixed 5:1 with 6X loading buffer (50 mM Tris-HCl, pH 6.8, 10% SDS, 0.2 mg/ml bromophenol blue, 50% glycerol, 20%  $\beta$ -mercaptoethanol) and boiled for 5 min to denature proteins. Twenty microliters of protein samples were loaded into sample wells, and electrophoresed through the polymerized gel at 170 V with Tris-glycine electrode buffer (50 mM Tris base, 125 mM glycine and 0.1% SDS, pH 8.3) until the dye front reached the bottom of the gel plate. The gels were subsequently stained in staining solution containing 0.1% (w/v) Coomassie Brilliant Blue R250, 40% (v/v) methanol, and 10% (v/v) acetic acid in water for 30 min and destained with destaining solution [40% (v/v) methanol and 10% (v/v) acetic acid] for 1-2 h. The molecular masses of protein bands were determined by comparing to standard low molecular weight protein markers (GE Healthcare, Uppsala, Sweden), which consist of phosphorylase b (97.4 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa) and bovine  $\alpha$ -lactalbumin (14.0 kDa).

#### **2.2.10 Determination of protein concentration**

The protein concentration was determined by the Bio-Rad assay (Hercules, CA, USA). The bovine serum albumin (BSA) was used as a standard ranging from 0-5  $\mu$ g. Each concentration containing 200  $\mu$ l of Bio-Rad protein assay solution and made up to 1 ml with sterile water. The mixture was incubated at room temperature for 10 min. The absorbance was measured at a wavelength of 595 nm ( $A_{595}$ ) with the

Protein Bradford program of a NanoDrop 2000 spectrophotometer (Thermo Scientific, MA, USA).

### **2.2.11 Preparation of *P. pastoris* SMD1168H competent cells**

A glycerol stock of *P. pastoris* strain SMD1168H was streaked on a yeast extract peptone dextrose (YPD) plate without antibiotic, which was then incubated at 28 °C for 2-3 days. A single colony was inoculated into 5 ml YPD broth and grown at 28 °C with 220 rpm overnight. The starter culture was transferred into 500 ml YPD broth and grown until the OD<sub>600</sub> reached 1.3-1.5. The cells were collected by centrifugation at 1,500 rpm for 5 min at 4 °C. The pellet was washed 2 times in 500 ml and 250 ml of ice-cold sterile water and collected by centrifugation at 1,500 rpm for 5 min at 4 °C each time. Next, the pellet was resuspended with 20 ml of ice-cold 1 M sorbitol and centrifuged at 1,500 rpm for 5 min at 4 °C. Finally, the pellet was resuspended and kept in 1 ml of ice-cold 1 M sorbitol and 80 µl aliquots per tube was used for transformation.

## **2.3 Amplification and cloning of Os4BGlu14**

### **2.3.1 Amplification of gene encoding mature Os4BGlu14**

Os4BGlu14 was cloned from the Genbank accession number AK067841 cDNA plasmid clone name [J013128A11](#) provided by the Rice Genome Resource full-length cDNA project. A fragment of the cDNA that encoded the predicted mature rice Os4BGlu14 gene was amplified with Os4BGlu14Matstrt primer, which contains a CACC sequence and *Nco*I site at the 5' end, and Os4BGlu14StopR primer, which contains a *Bam*HI site at 5' end, respectively (Table 2.1), and *Pfu* DNA polymerase

(Promega, WI, USA). The PCR product around 1.5 kb was run checked on 1% agarose gel.

### 2.3.2 Cloning of mature Os4BGlu14 into the pET32a vector

The PCR product of mature Os4BGlu14 was combined with pENTR-D-TOPO from a pENTR™ Directional TOPO® Cloning Kit, in a 2:1 molar ratio of purified PCR product:TOPO vector, to which were added 1 µl of salt solution and sterile water to 6 µl final volume. The topoisomerase reaction was incubated at room temperature overnight, and the mixture was used to transform chemically competent XL1-Blue or DH5α *E. coli*. The transformed XL1-Blue competent cells were spread onto an LB plate containing 15 µg/ml kanamycin, then incubated at 37 °C overnight. The colonies that had grown overnight were picked, and the plasmids prepared by the QIAGEN miniprep method (section 2.2.6) were checked by digestion with *EcoRV*, followed by agarose gel electrophoresis. The cDNA insert sequences of the plasmids with proper sized inserts were determined by automated DNA sequencing at Macrogen Corp. (Seoul, South Korea).

The recombinant pENTR-D-TOPO plasmid containing the cDNA encoding the mature Os4BGlu14 was recombined with pET32a/DEST (Opassiri et al., 2006) by an LR clonase recombination reaction, as described by the supplier (Invitrogen, CA, USA), to insert the Os4BGlu14 cDNA into pET32a/DEST. Each LR recombination reaction was transformed into competent cells by the CaCl<sub>2</sub> heat shock method (section 2.2.4) and selected on LB-agar containing 50 µg/ml ampicillin at 37 °C overnight. The colonies that grew overnight were picked and inoculated into LB media containing 50 µg/ml ampicillin and incubated at 37 °C, shaking at 200 rpm

overnight. The recombinant plasmids were extracted from the cultures by the alkaline lysis method.

In addition, the PCR mixture that contained mature Os4BGlu14 gene was cloned into the pET32a expression vector. The vector and insert gene were digested with *NcoI* and *BamHI*, then incubated at 37 °C for 2 h. The digested Os4BGlu14 gene and pET32a vector were purified with a HiYield™ Gel/PCR DNA fragments extraction kit (RBC Bioscience Corp., Taiwan). The purified Os4BGlu14 gene and pET32a vector were ligated by mixing vector:DNA insert (1:3) at 45 °C for 5 min to melt any cohesive termini that had reannealed, then immediately chilled on ice and 1X ligation buffer and 1 µl of T4 DNA ligase (Promega, WI, USA) were added and mixed in well but gently, and the reaction mixture was incubated at 15 °C for 18 h. Then, 1 µl of the reaction was transformed into competent cells and selected on a 50 µg/ml ampicillin LB agar plate. The colonies were cultured and plasmids extracted as described above. The presence of the gene insert was checked by cutting the prepared plasmids with *NcoI* and *BamHI* and evaluating the digested DNA bands by agarose gel electrophoresis, and the plasmid insert sequence was verified by automated DNA sequencing at Macrogen Corp.

### **2.3.3 Mutagenesis of pET32a/Os4BGlu14**

To try to produce active Os4BGlu14 protein, the mutation of Os4BGlu14 Q191 to E was done with the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, CA, USA). The pET32(a)/Os4BGlu14 plasmid was used as a template to amplify full-length plasmid with specific Os4BGlu14AB2Ef and Os4BGlu14AB2Er primers (Table 2.1) designed according to the criteria of QuikChange user manual. The PCR reaction was set up to contain 5-50 ng template, 125 ng forward and reverse

primers, 0.2  $\mu$ M dNTP mix, 1 $\times$  *Pfu* Ultra HF reaction buffer (20 mM Tris-HCl (pH 8.8), 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml BSA and 2 mM MgSO<sub>4</sub>), and 2.5 U of *Pfu* Ultra HF DNA polymerase and amplified with the temperature cycling parameters shown in Table 2.2. To eliminate the methylated and hemimethylated DNA of the DNA template, 10 U of *DpnI* endonuclease was used to treat the PCR products overnight at 37 °C. One microliter of the products was transformed into XL1-Blue competent cells and selected on 50  $\mu$ g/ml ampicillin LB agar plate. The recombinant insert gene was checked by automated DNA sequencing at Macrogen Corp. (Seoul, Korea).

**Table 2.2** Cycling parameters for mutation of pET32a/Os4BGlu14 by the QuikChange<sup>®</sup> Site-Directed Mutagenesis method.

Segment	Cycles	Temperature (°C)	Time
1	1	95	30 s
		95	30s
2	16	55	1 min
		68	6 min
3	1	68	10 min

## 2.4 Cloning of Os4BGlu16

### 2.4.1 Cloning of optimized Os4BGlu16 cDNA into the pPICZ $\alpha$ B(NH<sub>8</sub>) vector

An optimized Os4BGlu16 cDNA (protein accession number Q7XSK2) was synthesized and inserted into the pUC57 vector by GenScript Corporation (Piscataway, NJ, USA). The optimized Os4BGlu16 cDNA was transformed into XL1-Blue competent cells and spread onto LB agar containing 50  $\mu$ g/ml ampicillin. The colonies were picked and inoculated into LB media containing 50  $\mu$ g/ml ampicillin and incubated at 37 °C overnight. The plasmid was extracted with a QIAprep Spin miniprep Kit (QIAGEN, CA, USA). The optimized Os4BGlu16 and pPICZ $\alpha$ B(NH<sub>8</sub>) plasmid (Toonkool et al., 2006) were cut with *Pst*I and *Xba*I. The gel purified Os4BGlu16 insert and pPICZ $\alpha$ B(NH<sub>8</sub>) plasmid were ligated as described for the cloning of Os4BGlu14 and colonies selected on 25  $\mu$ g/ml zeocin. Colonies were picked and plasmid DNA extracted as described in section 2.2.6. The recombinant gene insert was checked by cutting the plasmid clones with *Pst*I and *Xba*I and evaluating the products by agarose gel electrophoresis, and the sequence confirmed by automated DNA sequencing at Macrogen Corp. (Seoul, Korea).

## 2.5 Amplification and cloning of Os4BGlu18

### 2.5.1 Amplification of cDNA encoding mature Os4BGlu18

The gene encoding full-length rice Os4BGlu18 was amplified from the cDNA synthesis product with the Os4BGlu18Startf and Os4BGlu18Stop primers (Table 2.1). Then, the gene encoding the predicted mature protein was also amplified from the initial PCR product with the Os4BGlu18MatStf and Os4BGlu18Stop primers. Both

reactions were carried out with *Pfu* DNA polymerase with the temperature cycling parameters shown in Table 2.3. The PCR product (~1.5 kb) was checked by electrophoresis on a 1% agarose gel.

**Table 2.3** Cycling parameters for amplification of cDNA encoding mature Os4BGlu18.

Segment	Cycles	Temperature (°C)	Time
1	1	95	2 min
		95	30 s
2	30	58	30 s
		72	3 min
3	1	72	10 min

### 2.5.2 Cloning of mature Os4BGlu18 into pET32a

The PCR mixture that contained the mature Os4BGlu18 gene and pET32a expression vector were digested with *Nco*I and *Bam*HI at 37 °C for 2 h. The digested Os4BGlu18 gene and pET32a vector were purified with the HiYield™ Gel/PCR DNA fragments extraction kit. The purified Os4BGlu18 gene and pET32a vector were ligated together by mixing 1:3 vector:DNA insert and heating at 45 °C for 5 min to melt any cohesive termini that have reannealed, then immediately chilled on ice. Then, 1X ligation buffer and T4 DNA ligase were added and the tube was mixed well but gently and incubated at 15 °C for 18 h. One microliter of the reaction was transformed into DH5α competent cells and selected on a 50 µg/ml ampicillin LB agar plate. The colonies were cultured and plasmids prepared as described for the

pET32/Os4BGlu14 plasmid (section 2.2.6). The recombinant gene insert was checked by cutting with *NcoI* and *BamHI* and the sequence verified by automated DNA sequencing at Macrogen Corp. (Seoul, South Korea).

## 2.6 Expression of Os4BGlu14

The recombinant pET32a/Os4BGlu14 wild type and pET32a/Os4BGlu14 Q191E mutant plasmids were transformed into Origami(DE3), Origami B(DE3), Rosetta-gami(DE3), and BL-21(DE3) competent cells and spread onto LB-agar containing 50 µg/ml ampicillin, 15 µg/ml kanamycin and 12.5 µg/ml tetracycline for Origami(DE3) and Origami B(DE3), the same antibiotics plus 34 µg/ml chloramphenicol for Rosetta-gami(DE3), and 50 µg/ml ampicillin alone for BL-21(DE3). The transformed cells were incubated at 37 °C overnight. The colonies that grew overnight were picked and inoculated into LB media containing the same antibiotics to make a starter culture. To express recombinant β-glucosidases, 1% final concentration of starter culture was added into the same type of media and cultured at 37 °C with rotary shaking at 200 rpm. Protein expression was induced when the OD<sub>600</sub> of the culture reached 0.4. The optimum expression conditions were determined by varying the final concentration of IPTG from 0 to 0.8 mM, and the temperature at 10 °C, 15 °C, 18 °C, 20 °C, and 22 °C for 16 h. The cell pellets were collected by centrifugation at 4,500 rpm for 15 min at 4 °C. The cell pellets was kept at -80 °C to allow freeze-thaw breakage before use.

The IPTG-induced bacterial cell pellets were thawed on ice and then resuspended in freshly prepared extraction buffer (20 mM Tris-HCl buffer, pH 8.0, 200 µg/ml lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride (PMSF),

25 µg/ml DNase I and 0.1 mg/ml soy bean trypsin inhibitor) in a ratio of 5 ml extraction buffer per gram fresh weight of cell pellets. The resuspended cells were incubated at room temperature for 30 min. Then, the insoluble proteins were removed by centrifugation. An aliquot of 5 µl of the supernatant fraction was used to test hydrolysis activity with 2 mM *p*NPGlc at 30 °C for 30 min.

## 2.7 Expression and purification of Os4BGlu16

The optimized pPICZαBNH<sub>8</sub>/Os4BGlu16 plasmid was linearized with *Sac*I. Linearization of the plasmid was checked by electrophoresis on a 1% agarose gel. The restriction enzyme was inactivated by heating at 65 °C for 10 min. Linear DNA was precipitated with PEG6000/2.5 M NaCl solution. The DNA pellet was dissolved in 5-10 µl of sterile de-ionized water. The linearized recombinant pPICZαBNH<sub>8</sub>/Os4BGlu16 vector was transformed into *P. pastoris* competent cells by electroporation (Bio-Rad) with the parameters of 1.5 kV, 25 µF and 400 Ω (*Pichia* manual, Invitrogen). The transformed cells were selected on Yeast Extract Peptone Dextrose medium with Sorbitol (YPDS) plates containing 100 µg/ml zeocin. The YPDS plate was incubated at 28 °C for 3-5 days. The transformed cells were selected again on a YPDS plate containing 500 µg/ml zeocin. The colonies were screened for protein production in small scale cultures. For protein production, a single colony that has been selected on a 500 µg/ml zeocin YPDS plate was inoculated into 500 ml of buffered glycerol-complex medium (BMGY) medium containing 100 µg/ml zeocin and grown in a shaking incubator (220 rpm) at 28 °C until the cell culture optical density at 600 nm (OD<sub>600</sub>) reached 2-3. The cells were harvested by centrifugation and resuspended in 1000 ml buffered methanol-complex medium (BMMY) at the final

OD<sub>600</sub> of 1. Protein expression was induced by adding methanol to 1% (v/v) final concentration every 24 h for 4 days at 20 °C with shaking, and the media was checked each day for activity (Luang et al., 2010).

The protein was purified from the culture broth after removal of the cells by centrifugation. The pH of the culture broth with secreted protein was adjusted to 7.5 with 2 M K<sub>2</sub>HPO<sub>4</sub> and it was loaded onto immobilized metal ion affinity chromatography (IMAC) column (GE Healthcare, Buckinghamshire, United Kingdom) charged with Co<sup>2+</sup>, and the column was washed with 5 column volumes of 5 mM and 10 mM of imidazole in 50 mM sodium phosphate buffer, pH 7.5, then the protein was eluted with 250 mM imidazole in 50 mM sodium phosphate buffer, pH 7.5. The active fractions were reconstituted by centrifugal filtration.

To test for the significance of glycosylation, the purified protein was deglycosylated with endoglycosidase H (New England BioLabs, MA, USA). The mixture of 90 µg Os4BGlu16 enzyme and 500 U endoglycosidase H in 50 mM sodium acetate, pH 5.5, was incubated at 4 °C for 3-4 days with gentle shaking, until deglycosylation was completed, based on inspection on SDS-PAGE.

## **2.8 Expression and purification of Os4BGlu18**

The recombinant pET32a/Os4BGlu18 plasmids was transformed into Origami(DE3) competent cells and spread onto LB-agar containing 50 µg/ml ampicillin, 15 µg/ml kanamycin and 12.5 µg/ml tetracycline. The plate was incubated at 37 °C overnight. The colonies that grew overnight were picked and inoculated into LB media containing the same antibiotics to make a starter culture. To express the recombinant Os4BGlu18 β-glucosidases, 1% final concentration of starter culture was

added into the same type of media and cultured at 37 °C with rotary shaking at 200 rpm. Protein expression was induced when the OD<sub>600</sub> of the culture reached 0.6. The protein expression was induced with 0.1 mM IPTG at 18 °C for 16-18 h. The cell pellets were collected by centrifugation at 4,500 rpm 15 min at 4 °C. The cell pellets was kept at -80 °C to allow freeze-thaw breakage before use.

The IPTG-induced bacterial cell pellets were thawed on ice and then resuspended in freshly prepared extraction buffer (20 mM Tris-HCl buffer, pH 8.0, 200 µg/ml lysozyme, 1% Triton-X 100, 1 mM PMSF, 25 µg/ml DNase I and 0.1 mg/ml soy bean trypsin inhibitor) in a ratio of 5 ml extraction buffer per gram fresh weight of cell pellets. The resuspended cells were incubated at room temperature for 30 min. Then, the insoluble proteins and cell debris were removed by centrifugation.

The supernatant fraction was loaded onto a Q sepharose chromatography (GE Healthcare) column, which had been pre-equilibrated with 50 mM Tris-HCl buffer, pH 8.0, and run at a flow rate of 2 ml/min. Unbound protein was washed out with 2 CV of 50 mM Tris-HCl buffer, pH 8.0. The protein was eluted by a gradient from 0-1 M NaCl in 50 mM Tris-HCl buffer, pH 8.0. The active protein was pooled and NaCl was added to 2 M final concentration. The protein solution was then loaded onto a phenyl sepharose chromatography (GE Healthcare) column, which was pre-equilibrated with 2 M NaCl in 50 mM Tris-HCl, pH 8.0. The column was washed with 2 CV of 2 M NaCl in 50 mM Tris-HCl, pH 8.0, and eluted with gradient from 2-0 M NaCl in 50 mM Tris-HCl, pH 8.0, and further eluted with 0-60% ethylene glycol in 50 mM Tris-HCl, pH 8.0. The fractions that contained β-glucosidase activity were pooled and concentrated, and the buffer was changed to 50 mM Tris-HCl buffer, pH 8.0, in 30 kDa molecular weight cut-off (MWCO) Amicon® Ultra-15 centrifugal

filter. Finally, active Os4BGlu18 protein from the previous column was loaded into an immobilized metal affinity chromatography (IMAC) column charged with  $\text{Co}^{2+}$ . After loading the active protein, the IMAC column was washed twice with 10 CV of equilibration/wash buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl) to remove unbound protein and washed again with 5 CV each of equilibration/wash buffer containing 5 mM and 10 mM of imidazole, respectively. Bound protein was eluted with 3 CV of elution buffer (50 mM Tris-HCl, pH 8.0, containing 250 mM imidazole). The eluted protein was checked by assaying activity with *p*NPGlc and the presence and purity of protein of appropriate size evaluated by SDS-PAGE. The fractions that contained  $\beta$ -glucosidase were pooled and imidazole removed by buffer exchange in 30 kDa MWCO Amicon® Ultra-15 centrifugal filters.

## **2.9 Enzyme assay, pH and temperature optimum and stability studies**

Initially, 1  $\mu\text{g}$  of purified protein or 5–50  $\mu\text{l}$  of purification fraction was incubated in 140  $\mu\text{l}$  total volume with 1 mM *p*NPGlc in 50 mM sodium acetate, pH 5.0, at 30 °C for 30 min, and then 70  $\mu\text{l}$  of 2 M  $\text{Na}_2\text{CO}_3$  was added to stop the reaction. The absorbance was measured at 405 nm and compared to a *p*-nitrophenolate (*p*NP) standard curve in the same buffers to determine the amount of *p*NP released. Upon determination of the pH optima, Os4BGlu16 was assayed in 50 mM sodium phosphate, pH 6.5, and Os4BGlu18 in 50 mM sodium acetate, pH 5.0.

To determine the pH optima of the purified protein, activity was assayed in 100 mM universal buffers (citric acid–disodium hydrogen phosphate) ranging from

pH 2.0 to 9.0, at 0.5 pH increments. The pH stability for the purified enzyme was determined by incubating the enzymes in 100 mM buffers ranging from pH 2.0 to 9.0 for 15, 30, 60 min and 2 h at room temperature (~25 °C). After incubation, the enzyme was assayed in 50 mM buffer at the optimum pH and 1 mM pNPGlc, as described above.

The optimum temperature for enzyme activity was determined by pre-incubating the purified enzyme at temperatures ranging from 5 °C to 70 °C at 5 °C increments for 10 min, and then pre-incubated enzyme was incubated with 1 mM pNPGlc in 50 mM buffer at the optimum pH, in a reaction volume of 140 µl at the same temperatures for 30 min, and then 70 µl of 2 M Na<sub>2</sub>CO<sub>3</sub> was added to stop the reaction, the enzyme activity was measured, as described above. Thermostability of the purified enzyme was measured by incubating the enzyme in 50 mM buffer at different temperatures in the range of 20 °C to 60 °C at 10 °C increments for 15, 30, 45, 60 min, then the enzyme samples was assayed with 1 mM pNPG in 50 mM sodium acetate, pH 5.0, at 30 °C for 30 min.

## 2.10 Substrate specificity and enzyme kinetics

The substrate specificity toward natural and synthetic substrates was evaluated by incubating 1 µg of enzyme with 1 mM final concentration of substrates at 30 °C for 1 h in 100 mM sodium phosphate, pH 6.5, for Os4BGlu16 or 100 mM sodium acetate, pH 5.0, for Os4BGlu18. Hydrolysis of pNP- and *ortho*-nitrophenyl (*o*NP) glycosides was detected as described for pNPGlc above. Reactions with other synthetic and natural glucosides were stopped by boiling for 5 min and the amount of

glucose released was determined by peroxidase/glucose oxidase coupled reactions (PGO assay, Sigma Aldrich) or the product from the hydrolysis reaction was spotted onto thin-layer chromatography silica gel 60 F254 plates (Merck, Darmstadt, Germany). TLC plates were developed with a solvent of ethyl acetate, methanol and water (7:2.5:1, v/v/v) and detected by spraying with 10% sulfuric acid in methanol, drying and heating at 120 °C.

Apparent kinetic parameters,  $K_M$  and  $V_{max}$ , of purified protein with *p*NP-glycosides and natural and synthetic substrates were determined in triplicate reactions. The optimum time point, at which the velocity of hydrolysis gives a first order rate constant, was determined. The rates were determined over a substrate concentration range of at least  $0.2 K_M$  to  $3 K_M$ . Kinetic parameters were calculated by fitting the rate of product formation and substrate concentrations by nonlinear regression of the Michaelis-Menten curves with Grafit 5.0 (Erithacus Software, Horley, Surrey, U.K.). The apparent  $k_{cat}$  values were calculated by dividing the  $V_{max}$  by the total amount of enzyme in the reaction.

For monoglucosyl substrates, glucose released was measured by high performance anion exchange chromatography on a Dionex ion chromatography system (ICS-3000). The boiled reactions were diluted 100-fold with HPLC grade water and loaded onto a Dionex CarboPac<sup>TM</sup> PA20 carbohydrate column (3 × 150 mm), which was then eluted with 10 mM NaOH in water. The peak areas were compared to a glucose standard curve to determine the amount of glucose released. Since the substrate concentrations in these assays did not bracket the  $K_M$  for accurate nonlinear regression, the kinetic parameters were calculated by linear regression of

the Hanes-Woolf ( $[S]/v$  vs.  $[S]$ ) plot, which produced a straight line with slope  $1/V_{\max}$  and a y-intercept of  $K_M/V_{\max}$  and an x-intercept of  $-K_M$  (Haldane, 1957).

### **2.11 Inhibition study**

To evaluate inhibition of Os4BGlu16 and Os4BGlu18, one microgram of enzyme was pre-incubated with 1 mM of potential inhibitor at 30 °C for 1 h and then pNPGlc was added to 3 mM for Os4BGlu16 protein and to 5 mM for Os4BGlu18 protein, and the reactions were incubated for another 20 min. The released pNP was then measured as described above and the absorbance was compared to identical reactions pre-incubated without inhibitor.

### **2.12 *In planta* expression analysis of rice monoglucosyl $\beta$ -glucosidase**

Total RNA was isolated from tissues of various age rice plants with the Spectrum™ plant total RNA kit (Sigma-Aldrich, MO, USA). Subsequently, 100 ng/ $\mu$ l of total RNA was mixed with 2.5  $\mu$ M oligo(dT)<sub>20</sub> and 200 U Superscript™III reverse transcriptase (Invitrogen, CA, USA), and the mixture was reverse transcribed at 50 °C for 50 min. Real-time PCR was done with the gene-specific primers listed in Table 2.1. The PCR reaction included SYBR Green Mix (KapaBiosystems, Boston, MA, USA), 1<sup>st</sup> strand cDNA, the gene specific primer pair and *Taq* polymerase, along with deoxyribonucleotides and buffer. After the PCR finished, the products were examined by 2% agarose gel to verify specificity. Relative gene expression ratios (rERs) were calculated from the  $C_T$  values and efficiency,  $E$ , with a 7-day-old shoot cDNA as the reference sample and actin as the reference gene (Scheffe et al., 2006).

## 2.13 Detection of monolignol compounds in rice KDML105 by UPLC-MS analysis

The rice (*Oryza sativa* L. ssp. *Indica* cv. KDML105) seeds were soaked in water until coleoptiles were generated, then they were planted in soil in pots. The rice samples were collected every 10 days and the tissues separated, then frozen at -40 °C. Fifty-milligram samples of frozen samples of rice tissues at various ages from vegetative until ripening stage were ground in liquid nitrogen and extracted in 500 µl of 80% methanol. The samples were mixed by vortexing 5 min and sonicating 15 min. Then, the supernatant was collected by centrifugation at 13,000 rpm for 15 min. Five microliters of each sample was injected onto an Agilent SB-C18 RRHD 1.8 µm, 2.1x150 mm column (Agilent Technologies, CA, USA) on an Agilent 1290 UPLC system inline with an Agilent 6490 triple quadrupole mass spectrometer. A gradient of buffer A (100:1:0.1, water:acetonitrile:2 M ammonium acetate, pH 5.0) and buffer B (100:1:0.1, acetonitrile:water:2 M ammonium acetate, pH 5.0) was used. The gradient was begun at 95% buffer A, 5% buffer B at 0.1 min and increased to 45% buffer B in 36.9 min at the flow rate of 0.2 ml/min (Chapelle et al., 2012). The temperature of the column was set at 40 °C. A UV visible spectrum was measured between 190-600 nm on the inline diode array detector (DAD). Electrospray ionization (ESI) was used in the negative ion mode with MS2 selected ion monitoring (SIM). Ionic masses of *p*-coumaric acid ( $m/z$  163, retention time [Rt] 10.8 min), caffeic acid ( $m/z$  179, Rt 7.6 min), coniferyl alcohol ( $m/z$  179, Rt 13.6 min), sinapyl alcohol ( $m/z$  209, Rt 14.1 min), sinapic acid ( $m/z$  223, Rt 12.9 min), *p*-coumaryl alcohol glucoside [*p*CAG] ( $m/z$  371, Rt 6.8 min), coniferin ( $m/z$  401, Rt 8.2 min), and syringin ( $m/z$  431, Rt 9.2 min) were monitored. The abundance values were compared

to standard curves of the specific compounds to determine the amount of each monolignol compound in the extract. Five replicates of each sample were analyzed to carry out statistical analyses.



## CHAPTER III

### RESULTS

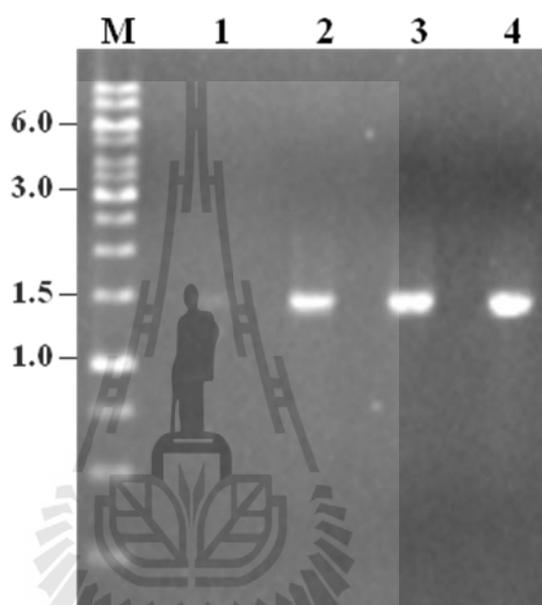
#### 3.1 Cloning and expression of Os4BGlu14

The cDNA encoding the mature Os4BGlu14 protein was amplified from the Genbank accession number AK067841 cDNA clone plasmid provided by the Rice Genome Resource full-length cDNA project (Kikuchi et al., 2003). The PCR product around 1.5 kb (Figure 3.1) was ligated into the pET32a expression vector. Sequencing of the recombinant pET32a/Os4BGlu14 plasmid showed that it matched the expected sequence from the *Oryza sativa* Japonica group cDNA clone: J013128A11 (GenBank: AK067841.1) in the National Center for Biotechnology Information (NCBI).

Because the catalytic acid/base of Os4BGlu14 is replaced by glutamine residue 191 (Figure 3.2), the mutation of pET32a/Os4BGlu14Q191E was made by site-directed mutagenesis to test if regenerating the catalytic acid/base would rescue catalytic activity.

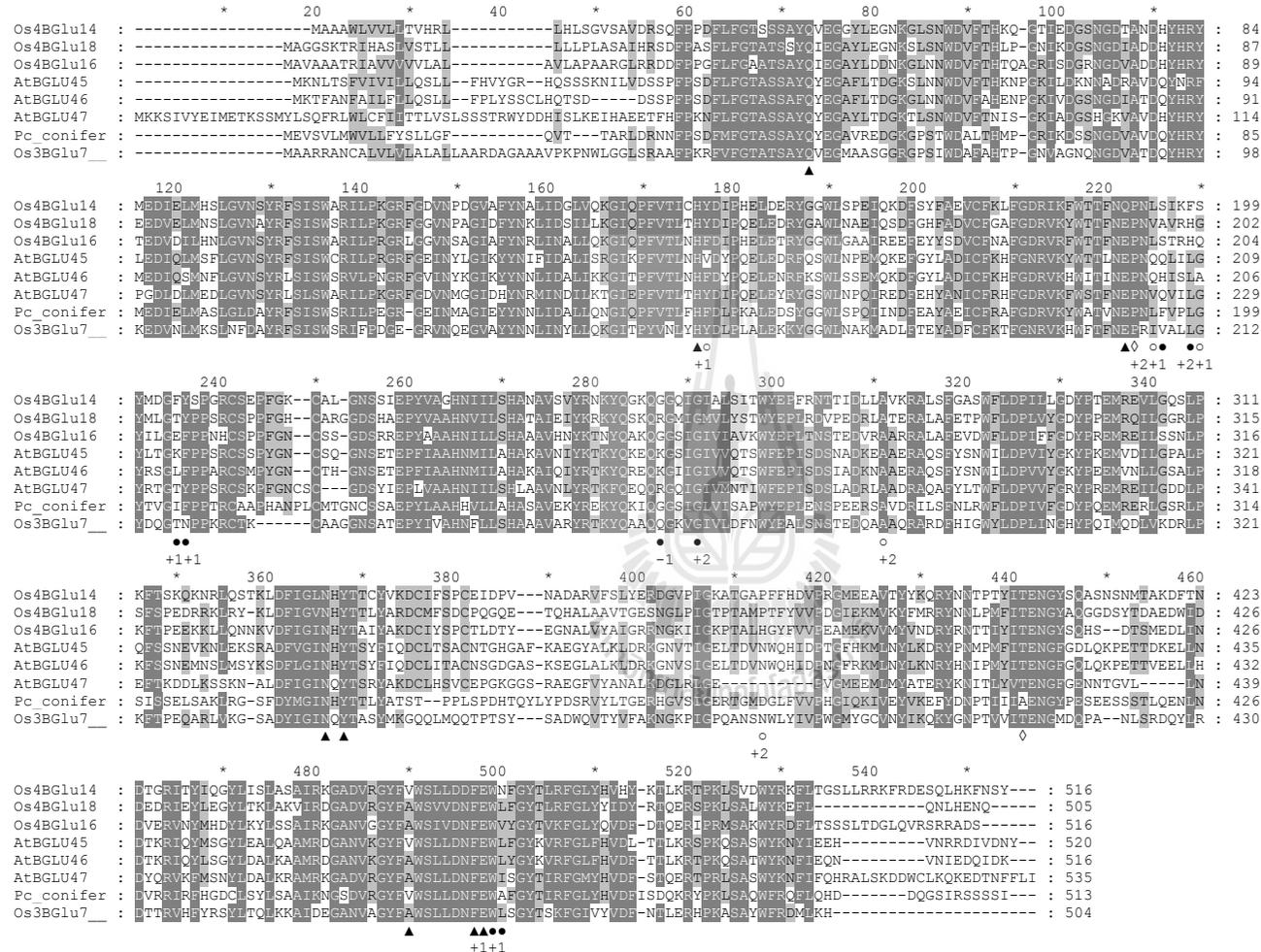
Expression of Os4BGlu14 was attempted with the recombinant pET32a/Os4BGlu14 wild-type and pET32a/Os4BGlu14Q191E mutant plasmids in *E.coli* strains Origami(DE3), Origami B(DE3), and Rosetta-gami(DE3) (Figure 3.3). The concentration of IPTG used for induction was varied from 0 to 0.8 mM (Figure 3.4) and the temperature of induction from 10-37 °C (Figure 3.5). SDS-PAGE analysis of fractions which induced at 0.4 mM IPTG at 20 °C for 16 h from an attempt

to purify the expressed protein by IMAC is shown in Figure 3.6. No correct-size protein band could be observed in the elution fractions and no activity could be detected in assays of the soluble cell lysate with pNPGlc, oNPGlc, and 2,4-dNPGlc substrates.

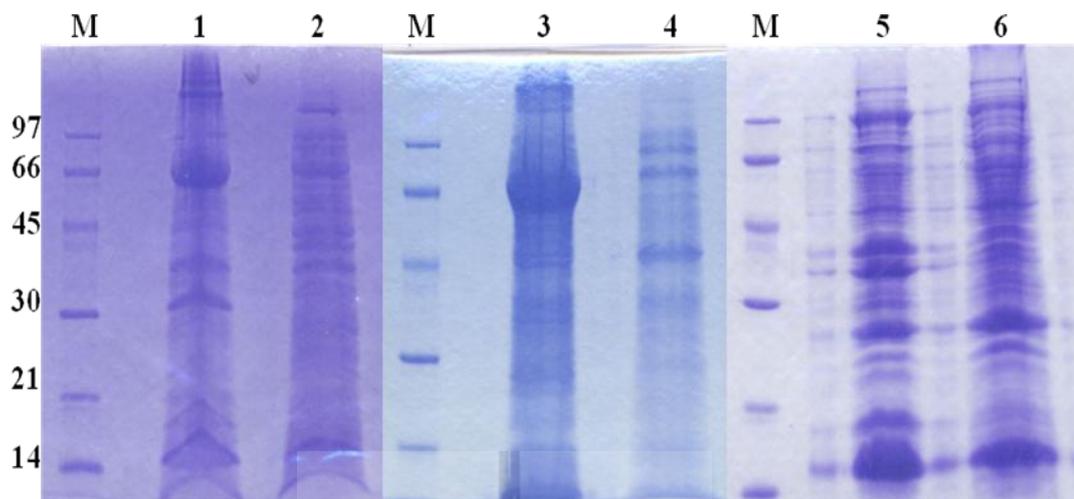


**Figure 3.1** Amplification of a cDNA encoding mature Os4BGlu14 gene. The AK067841 cDNA clone plasmid (J013128A11) was used as a template. The PCR products were separated on 1% agarose gel electrophoresis and stained with ethidium bromide.

Lane M, Thermo Scientific GeneRuler 1kb DNA ladder and Lanes 1-4, PCR products of a cDNA encoding mature Os4BGlu14 which differ in annealing temperature at 54, 56, 58, and 60 °C, respectively.

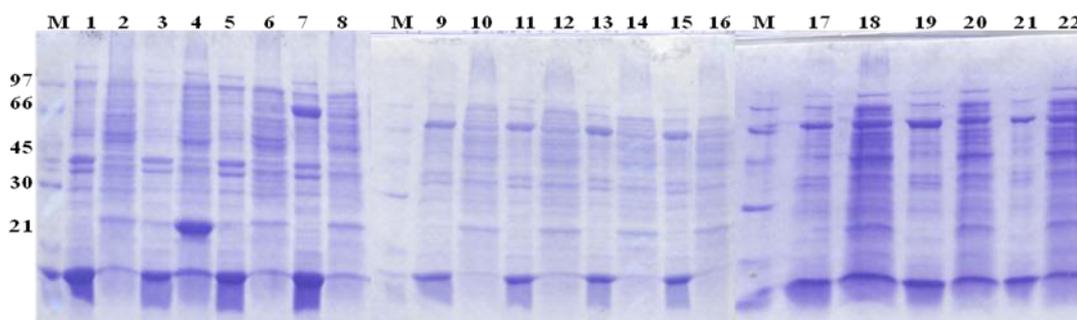


**Figure 3.2** Amino acid sequence alignment of predicted rice monolignol  $\beta$ -glucosidases, rice Os3BGLu7 (BGLu1), *Arabidopsis thaliana*  $\beta$ -glucosidases (BGLU45, BGLU46, BGLU47) and *Pinus contorta* coniferin  $\beta$ -glucosidase.



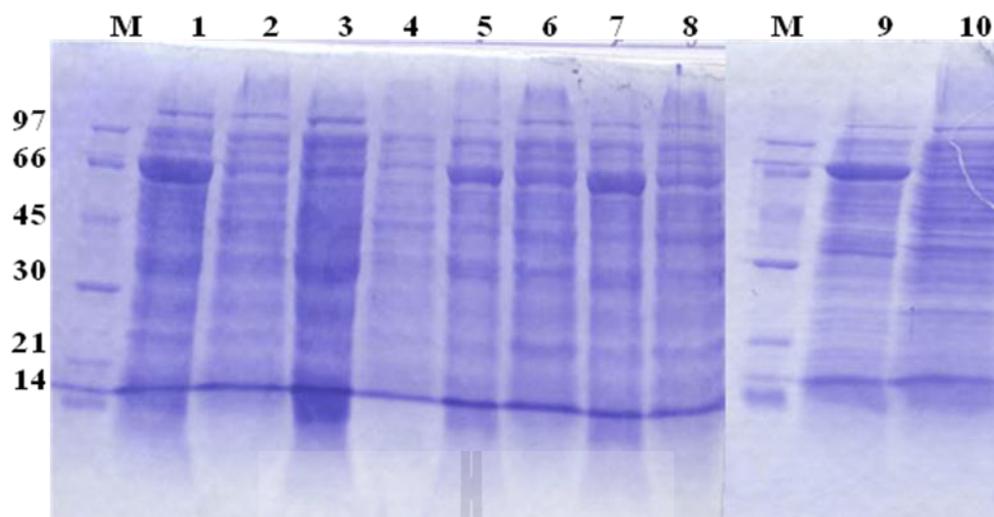
**Figure 3.3** SDS-PAGE analysis of pET32a/Os4BGlu14 expression in various *E. coli* host cells. The cells were induced with 0.4 mM IPTG at 20 °C for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, insoluble protein from Origami(DE3); Lane 2, soluble protein from Origami(DE3); Lane 3, insoluble protein from Origami B(DE3); Lane 4, soluble protein from Origami B(DE3); Lane 5, insoluble protein from Rosetta-gami(DE3); Lane 6, soluble protein from Rosetta-gami(DE3).



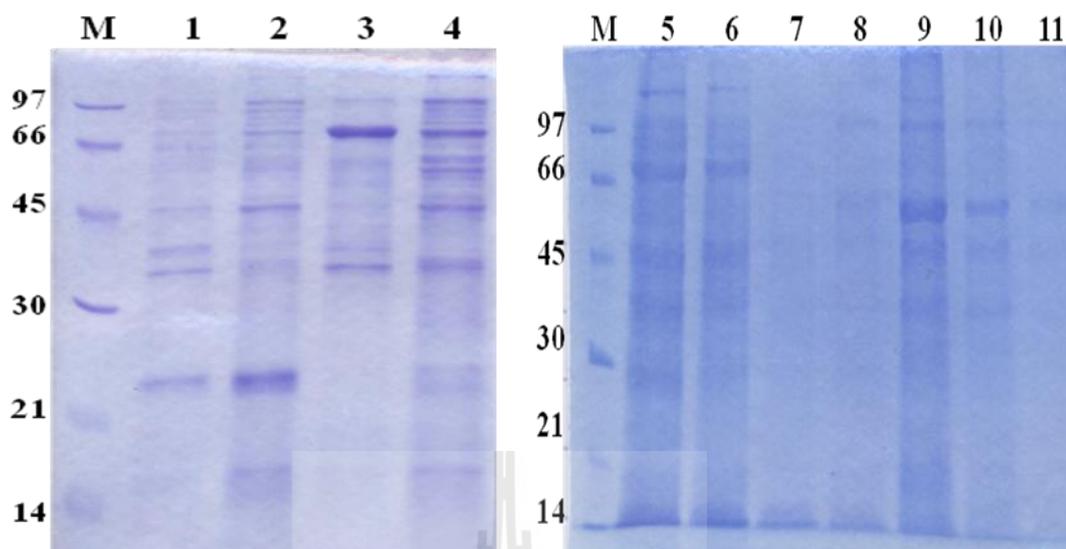
**Figure 3.4** SDS-PAGE analysis of pET32a/Os4BGlu14 (BGlu14) expression in Origami(DE3) induced at 0 to 0.8 mM of IPTG. Cultures were induced at 20 °C for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, 0 mM IPTG pET32a culture insoluble protein; Lane 2, 0 mM IPTG pET32a culture soluble protein; Lane 3, 0.4 mM IPTG pET32a culture insoluble protein; Lane 4, 0.4 mM IPTG pET32a culture soluble protein; Lane 5, 0 mM IPTG BGlu14 culture insoluble protein; Lane 6, 0 mM IPTG BGlu14 culture soluble protein; Lane 7, 0.1 mM IPTG BGlu14 culture insoluble protein; Lane 8, 0.1 mM IPTG BGlu14 culture soluble protein; Lane 9, 0.2 mM IPTG BGlu14 culture insoluble protein; Lane 10, 0.2 mM IPTG BGlu14 culture soluble protein; Lane 11, 0.3 mM IPTG BGlu14 culture insoluble protein; Lane 12, 0.3 mM IPTG BGlu14 culture soluble protein; Lane 13, 0.4 mM IPTG BGlu14 culture insoluble protein; Lane 14, 0.4 mM IPTG BGlu14 culture soluble protein; Lane 15, 0.5 mM IPTG BGlu14 culture insoluble protein; Lane 16, 0.5 mM BGlu14 culture soluble protein; Lane 17, 0.6 mM IPTG BGlu14 culture insoluble protein; Lane 18, 0.6 mM IPTG BGlu14 culture soluble protein; Lane 19, 0.7 mM IPTG BGlu14 culture insoluble protein; Lane 20, 0.7 mM IPTG BGlu14 soluble protein; Lane 21, 0.8 mM IPTG BGlu14 culture insoluble protein; Lane 22, 0.8 mM IPTG BGlu14 culture soluble protein.



**Figure 3.5** SDS-PAGE analysis of pET32a/Os4BGlu14 expression in Origami(DE3) at temperatures varying from 10-22 °C and induced at 0.4 mM IPTG for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, insoluble protein from pET32a/Os4BGlu14 culture at 10 °C ; Lane 2, pET32a/Os4BGlu14 soluble protein at 10 °C; Lane 3, pET32a/Os4BGlu14 insoluble protein at 15 °C; Lane 4, pET32a/Os4BGlu14 soluble protein at 15 °C; Lane 5, pET32a/Os4BGlu14 insoluble protein at 18 °C; Lane 6, pET32a/Os4BGlu14 soluble protein at 18 °C; Lane 7, pET32a/Os4BGlu14 insoluble protein at 20 °C; Lane 8, pET32a/Os4BGlu14 soluble protein at 20 °C; Lane 9, pET32a/Os4BGlu14 insoluble protein at 22 °C; Lane 10, pET32a/Os4BGlu14 soluble protein at 22 °C.

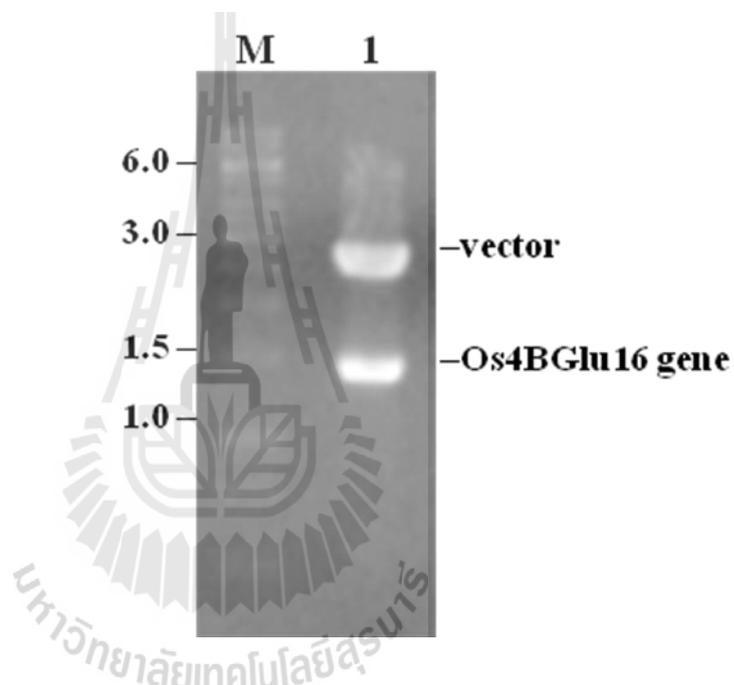


**Figure 3.6** SDS-PAGE analysis of pET32a/Os4BGlu14 expression in Origami(DE3) induced with 0.4 mM IPTG at 20 °C for 16 h and purified by IMAC column.

Lane M, Bio-Rad low molecular weight markers; Lane 1, insoluble protein from induced cells carrying empty pET32a; Lane 2, soluble protein from induced cells with empty pET32a; Lane 3, insoluble protein from induced cells with pET32a/Os4BGlu14; Lane 4, soluble protein from induced cells with pET32a/Os4BGlu14; Lane 5, flow-through fraction after passing pET32a/Os4BGlu14 cell lysate through the IMAC column; Lane 6, IMAC wash with equilibration buffer; Lane 7, wash of IMAC column with 5 mM imidazole; Lane 8, wash of IMAC column with 10 mM imidazole; Lanes 9-11, fractions from IMAC elution with 250 mM imidazole.

### 3.2 Cloning and expression of Os4BGlu16

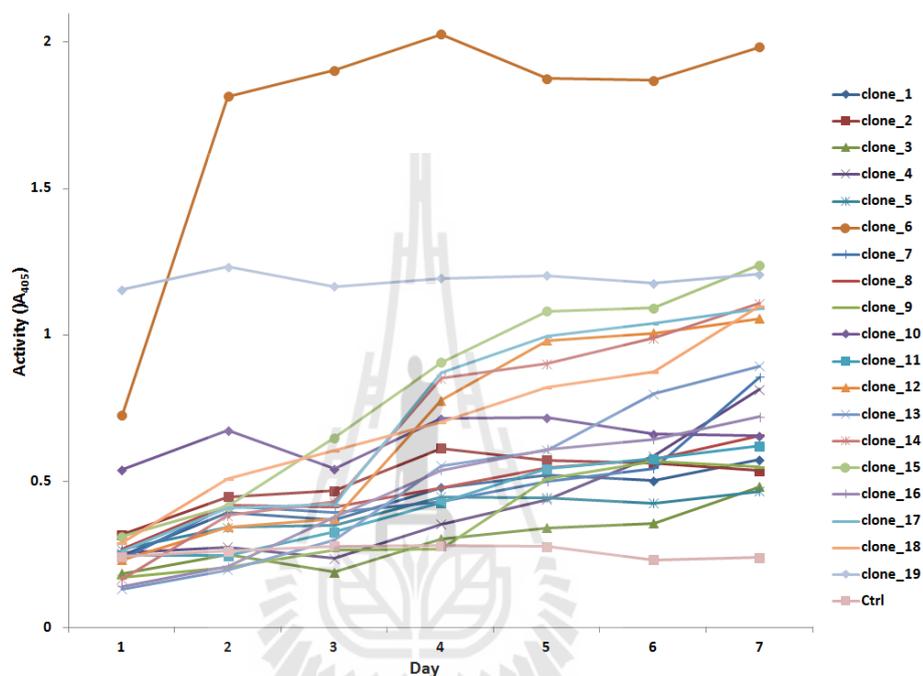
A gene optimized for Os4BGlu16 (Genbank accession number KJ579205) expression in *Pichia pastoris* was synthesized and inserted into the pUC57 vector by GenScript Corporation (Piscataway, NJ, USA) (Figure 3.7). The optimized Os4BGlu16 cDNA was inserted into the pPICZ $\alpha$ BNH<sub>8</sub> plasmid (Toonkool et al., 2006).



**Figure 3.7** The optimized Os4BGlu16 in pUC57 plasmid cut with *Pst*I and *Xba*I. The digestion reaction was separated on 1% agarose gel electrophoresis and stained with ethidium bromide.

Lane M, Thermo Scientific GeneRuler 1kb DNA ladder and Lane 1, the digestion product of the optimized Os4BGlu16.

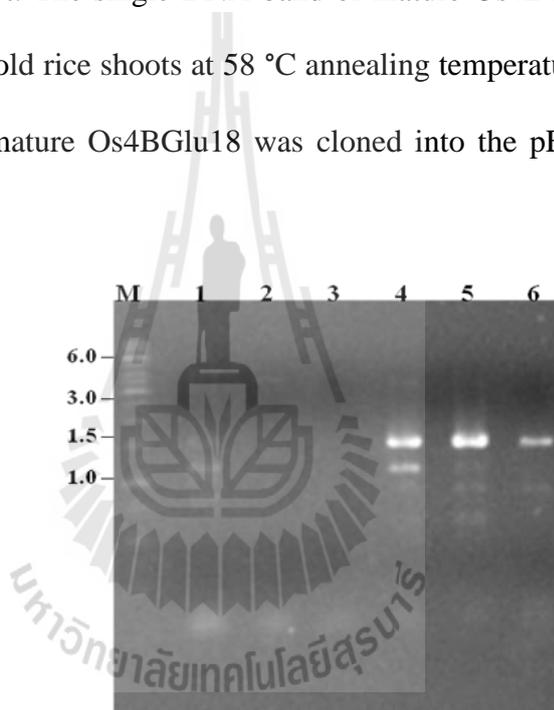
The pPICZ $\alpha$ BNH<sub>8</sub>/Os4BGlu16 plasmid was cloned into *P. pastoris* strain SMD1168H, and protein expression was induced with 1% (v/v) methanol for 7 days at 20 °C. Of the twenty clones examined for pNPGlc hydrolysis activity in the media, Clone 6 had the highest activity, which was maximal at 4 days induction (Figure 3.8).



**Figure 3.8** The pNPGlc hydrolysis activity in *pichia* media over 7 days of expression of clones expressing Os4Glu16. The activity was determined by incubating 50  $\mu$ l of induced cells media with 1 mM pNPGlc, in 50 mM NaOAc, pH 5.0, at 30 °C for 30 min. The OD<sub>405</sub> is plotted versus the day.

### Cloning and expression of Os4BGlu18

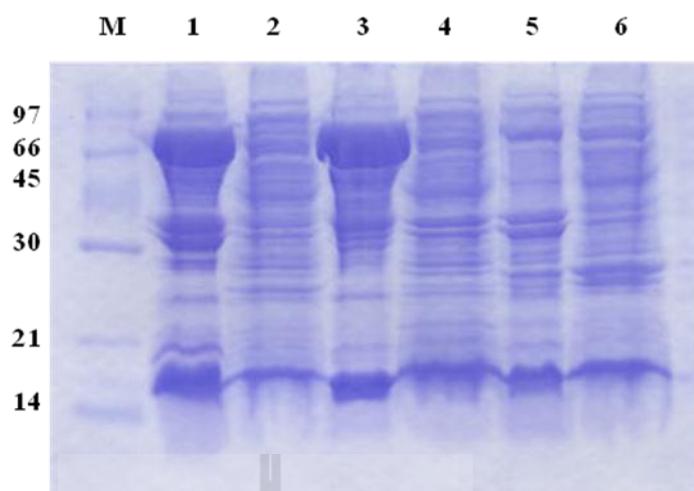
The cDNA encoding mature rice Os4BGlu18 was amplified by reverse transcription and nested PCR with total mRNA of 7-day-old rice (cv.KDML105) roots and shoots as template. First, the cDNA encoding the full-length rice Os4BGlu18 precursor protein was amplified by varying annealing temperature at 50, 54, 58 °C and then the fragment encoding the mature protein was also amplified from the initial PCR product. The single DNA band of mature Os4BGlu18 gene could be amplified from 7-day-old rice shoots at 58 °C annealing temperature (Figure 3.9). The cDNA encoding the mature Os4BGlu18 was cloned into the pET32a(+) expression vector.



**Figure 3.9** Amplification of a cDNA encoding mature Os4BGlu18 gene from 7-day-old rice roots and shoots cDNA as a template. The PCR products were separated on 1% agarose gel electrophoresis and stained with ethidium bromide.

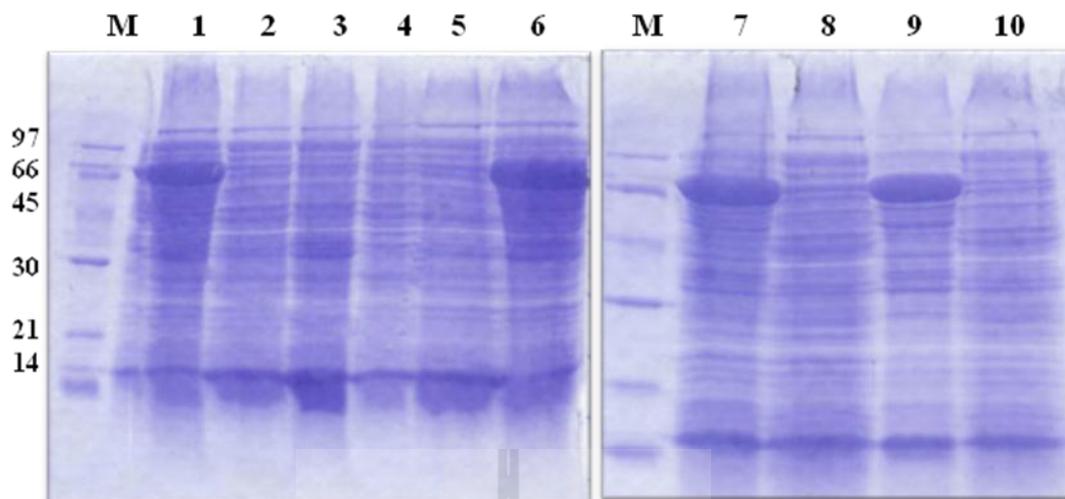
Lane M, Thermo Scientific GeneRuler 1kb DNA ladder, Lane 1-3, PCR product when used 7-day-old rice roots as template and varied annealing temperature at 50, 54, 58 °C, respectively, Lane 4-6 , PCR product when used 7-day-old rice shoots as template and varied annealing temperature at 50, 54, 58 °C, respectively.

Expression of Os4BGlu18 was attempted with the recombinant pET32a/Os4BGlu18 plasmids in *E.coli* strains Origami(DE3), Origami B(DE3), and Rosetta-gami(DE3) by inducing with 0.4 mM IPTG at 20 °C for 16 h (Figure 3.10). The activity of this enzyme could be detected in soluble cell lysates of Origami(DE3) and Origami B(DE3) with pNPGlc substrate as the same pNP release value; thus, Origami(DE3) was used as host cells for Os4BGlu18 expressions. The optimization of expression temperature was varied from 10 °C to 22 °C (Figure 3.11) with 18 °C found to give highest activity with pNPGlc substrate. Then, the concentration of IPTG used for induction was varied from 0 to 0.8 mM at 18 °C for 16 h (Figure 3.12). Os4BGlu18 enzyme could be expressed at all concentrations of IPTG. However, cells induced with 0.1 mM IPTG had the highest activity when Os4BGlu18 enzyme activity was tested with pNPGlc. In the absence of IPTG, the activity detected was only 44% compared with 0.1 mM IPTG induced cells. At 0.2-0.8 mM IPTG induction, the activity decreased 10-30%. Os4BGlu18 required at least 16 h to 24 h for induction. After induction for 12 h, the activity detected was only 28% of that at 16 h. Finally, the *E.coli* strains Origami(DE3) was used to express Os4BGlu18 enzyme and induced with 0.1 mM IPTG at 18 °C.



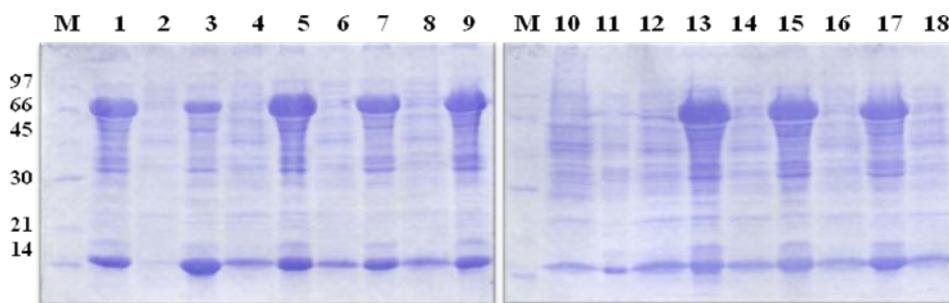
**Figure 3.10** SDS-PAGE analysis of pET32a/Os4BGlu18 expression in various *E. coli* host cells. The cells were induced with 0.4 mM IPTG at 20 °C for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, insoluble protein from Origami(DE3); Lane 2, soluble protein from Origami(DE3); Lane 3, insoluble protein from Origami B(DE3); Lane 4, soluble protein from Origami B(DE3); Lane 5, insoluble protein from Rosetta-gami(DE3); Lane 6, soluble protein from Rosetta-gami(DE3).



**Figure 3.11** SDS-PAGE analysis of pET32a/Os4BGlu18 expression in Origami(DE3) at temperatures varying from 10-22 °C and induced at 0.4 mM IPTG for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, insoluble protein from pET32a/Os4BGlu18 culture at 10 °C; Lane 2, pET32a/Os4BGlu18 soluble protein at 10 °C; Lane 3, pET32a/Os4BGlu18 insoluble protein at 15 °C; Lane 4, pET32a/Os4BGlu18 soluble protein at 15 °C; Lane 5, pET32a/Os4BGlu18 insoluble protein at 18 °C; Lane 6, pET32a/Os4BGlu18 soluble protein at 18 °C; Lane 7, pET32a/Os4BGlu18 insoluble protein at 20 °C; Lane 8, pET32a/Os4BGlu18 soluble protein at 20 °C; Lane 9, pET32a/Os4BGlu18 insoluble protein at 22 °C; Lane 10, pET32a/Os4BGlu18 soluble protein at 22 °C.

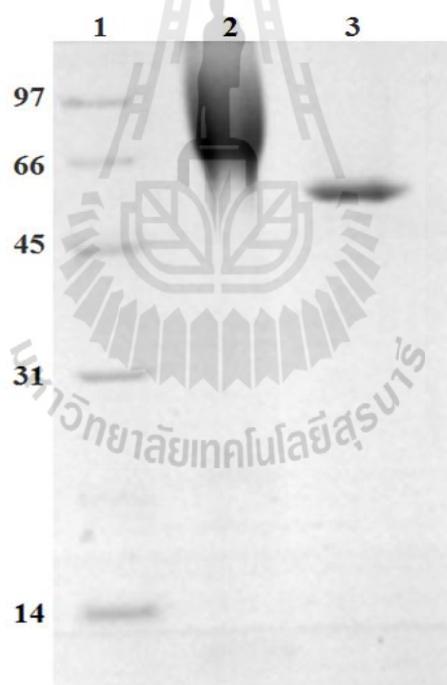


**Figure 3.12** SDS-PAGE analysis of pET32a/Os4BGlu18 expression in Origami(DE3) induced at 0 to 0.8 mM of IPTG. Cultures were induced at 18 °C for 16 h.

Lane M, Bio-Rad low molecular weight markers; Lane 1, 0 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 2, 0 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 3, 0.1 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 4, 0.1 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 5, 0.2 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 6, 0.2 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 7, 0.3 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 8, 0.3 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 9, 0.4 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 10, 0.4 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 11, 0.5 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 12, 0.5 mM pET32a/Os4BGlu18 culture insoluble protein; Lane 13, 0.6 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 14, 0.6 mM IPTG pET32a/Os4BGlu18 culture insoluble protein; Lane 15, 0.7 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 16, 0.7 mM IPTG pET32a/Os4BGlu18 insoluble protein; Lane 17, 0.8 mM IPTG pET32a/Os4BGlu18 culture soluble protein; Lane 18, 0.8 mM IPTG pET32a/Os4BGlu18 culture insoluble protein.

### 3.3 Purification of Os4BGlu16

An optimized Os4BGlu16 cDNA was used to produce a secreted, N-terminally His<sub>8</sub>-tagged protein in *P. pastoris*, the recombinant protein was induced and purified from the pichia media by IMAC.  $\beta$ -Glucosidase activity was detected in the media, as shown in Figure 3.8, and a broad protein band was detected above 67 kDa on the SDS-PAGE gel (Figure 3.13). To test whether the broadness of the protein band was due to glycosylation, deglycosylation with endoglycosidase H was done and a single band of approximately 60 kDa was detected on a Coomassie-stained SDS-PAGE gel (Figure 3.13).



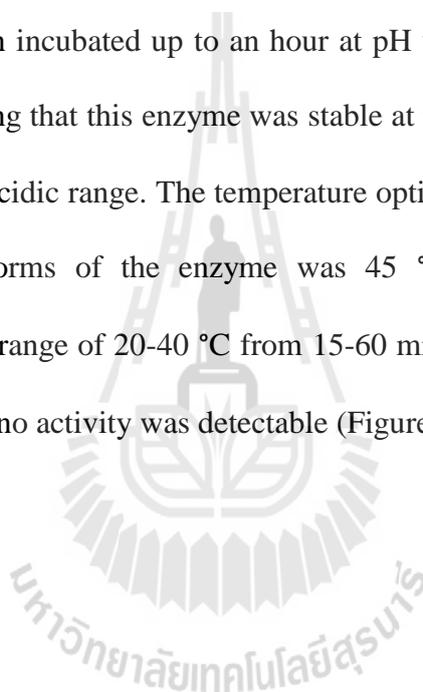
**Figure 3.13** SDS-PAGE analysis of Os4BGlu16 protein expressed in *Pichia pastoris*.

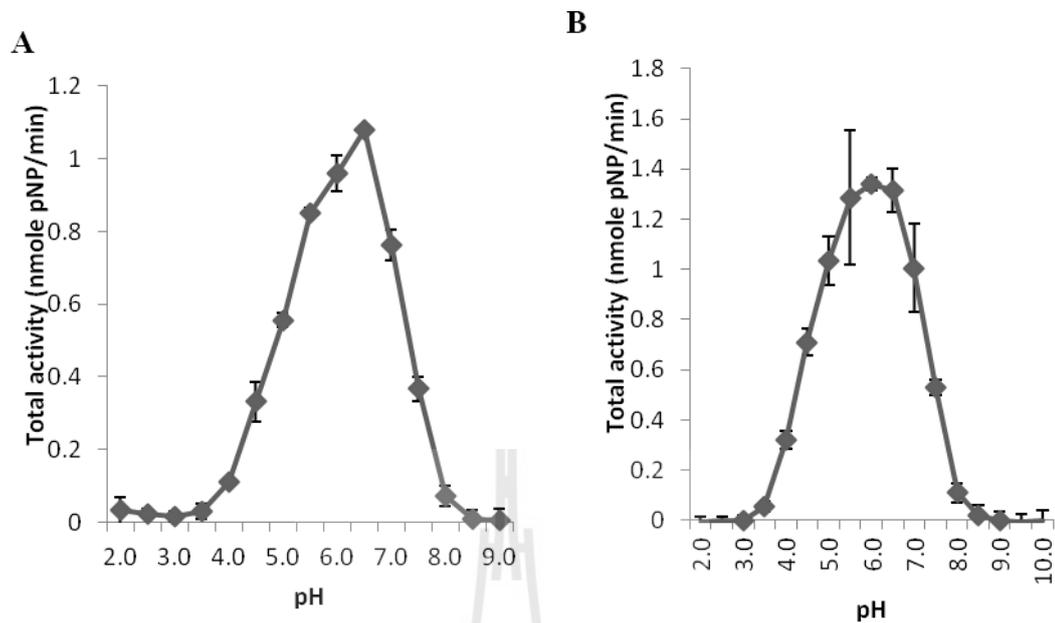
Lane 1, standard protein marker; Lane 2, Os4BGlu16 protein before deglycosylation; Lane 3, Os4BGlu16 protein after deglycosylation with endoglycosidase H.

### **3.4 Effect of pH and temperature on the activity and stability of**

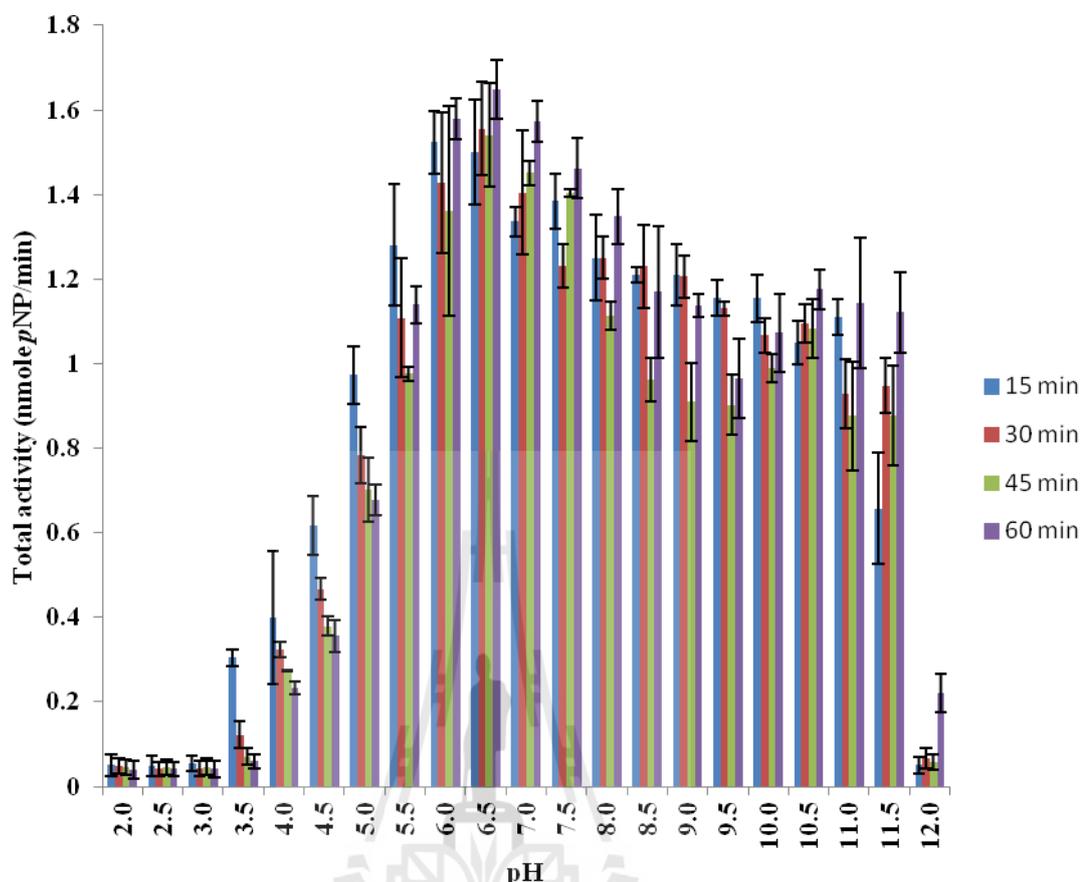
#### **Os4BGlu16**

As shown in Figure 3.14, the optimum pH for glycosylated Os4BGlu16 was found to be 6.5 and after deglycosylation the optimum pH for Os4BGlu16 was 6.0, but at pH 5.5 and 6.5 the activity was within error of this (only 4% and 2% lower, respectively). Figure 3.15 shows that the activity of Os4BGlu16 maintained >50% maximal activity when incubated up to an hour at pH values ranging from pH 5.5 to 11.5 at 25 °C, indicating that this enzyme was stable at neutral to basic pH, but lost its activity in the highly acidic range. The temperature optimum for both the glycosylated and deglycosylated forms of the enzyme was 45 °C (Figure 3.16), but it was thermostable over the range of 20-40 °C from 15-60 min and began to lose activity at 50 °C, while, at 60 °C no activity was detectable (Figure 3.17).

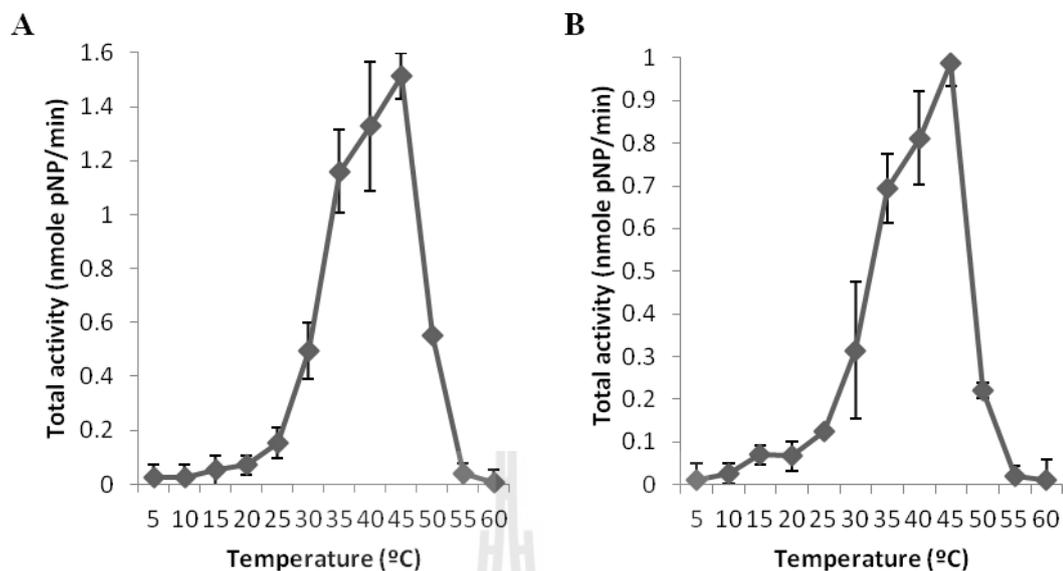




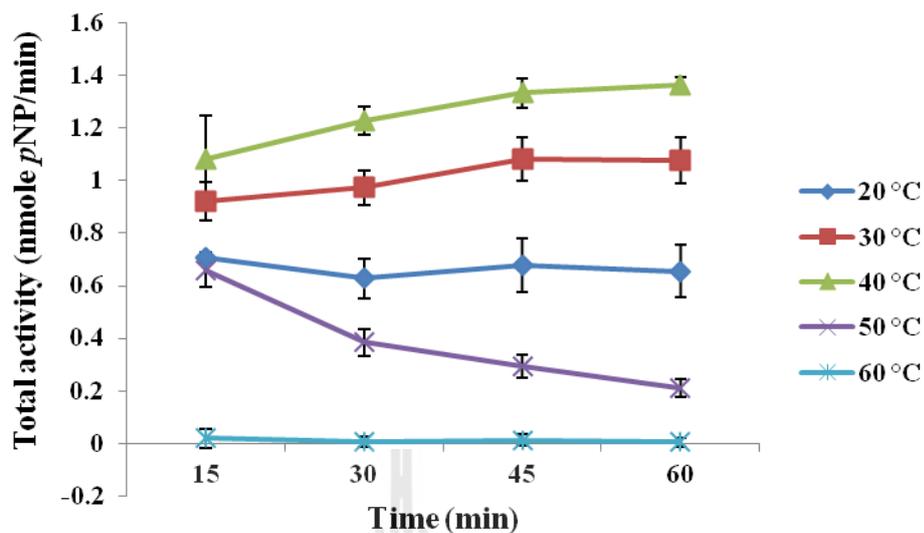
**Figure 3.14** Activity versus pH profiles of Os4BGlu16. The purified enzyme was incubated with 1 mM *p*NPGlc in 100 mM McIlvaine's universal buffers (citric acid-disodium hydrogen phosphate) ranging from pH 2.0 to 9.0 at 30 °C for 30 min. A is pH optimum of recombinant Os4BGlu16 before deglycosylation and B is pH optimum of recombinant Os4BGlu16 after deglycosylation.



**Figure 3.15** pH stability of Os4BGlu16 activity. The Os4BGlu16 was incubated in universal buffer (pH 2-12) for time periods of 15, 30, 45, and 60 min at 25 °C, then enzyme was diluted 100 fold into 100 mM sodium phosphate buffer, pH 6.5, and assayed with 1 mM *p*NPGlc for 30 min at 30 °C.



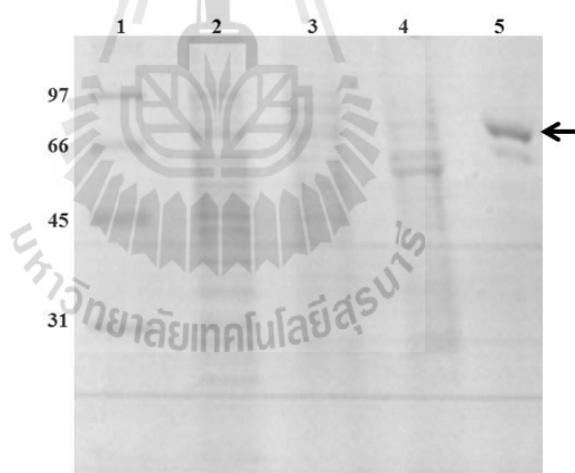
**Figure 3.16** Temperature optima of Os4BGlu16. The activity was assayed by pre-incubating 1  $\mu$ g of enzyme in 50 mM sodium phosphate buffer, pH 6.5, at the specified temperatures for 10 min to bring the reactions up to temperature, then the enzymes were assayed with 1 mM *p*NPGlc for 10 min at the same temperature as the preincubation. Panel A depicts the temperature optimum of recombinant Os4BGlu16 before deglycosylation while panel B shows the temperature optimum of recombinant Os4BGlu16 after deglycosylation.



**Figure 3.17** Thermostability of Os4BGlu16. The Os4BGlu6 was incubated at various temperatures (20-60 °C) from 15 to 60 min in 50 mM sodium phosphate buffer, pH 6.5, and enzyme activity was determined at 30 °C by adding 1 mM final concentration of *p*NPGlc and incubating for 30 min. Pre means preincubation time at 0 min.

### 3.5 Purification of Os4BGlu18

The Os4BGlu18 fusion protein with N-terminal thioredoxin, His<sub>6</sub> and S-tags was highly expressed in *E. coli* strain Origami(DE3) and induced with 0.1 mM IPTG at 18 °C for 16-18 h. The soluble protein of recombinant Os4BGlu18 was purified by anion-exchange chromatography, and the active protein started to elute at 0.3-0.5 M NaCl. Then, the active protein was loaded on hydrophobic interaction chromatography and eluted at 0.3-0.5 M NaCl concentration. Finally, the protein was purified by IMAC column as described in section 2.8. Os4BGlu18 was produced at approximately 75 kDa and approximately 90% pure, as judged by SDS-PAGE (Figure 3.18).



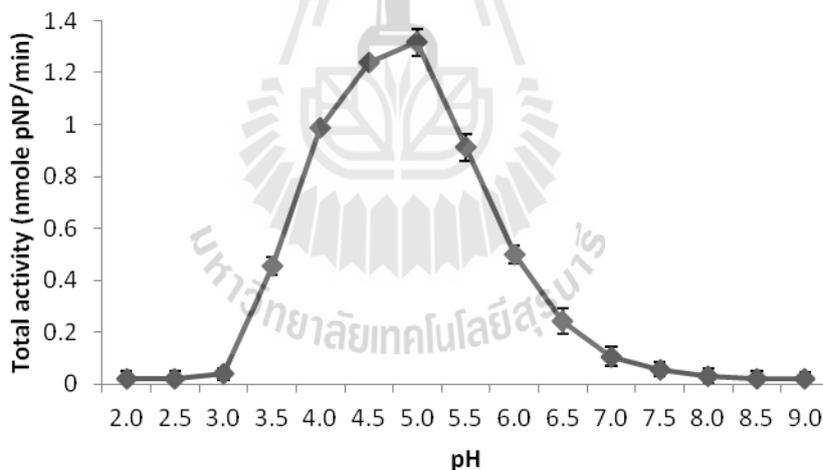
**Figure 3.18** SDS-PAGE analysis of Os4BGlu18 production in *Escherichia coli*.

Lane 1, standard protein marker; Lane 2, crude protein extract of cells expressing Os4BGlu18; Lane 3, Os4BGlu18 purified by Q-sepharose ion exchange chromatography; Lane 4, Os4BGlu18 after purification by phenyl sepharose chromatography; Lane 5, Os4BGlu18 after purification by IMAC.

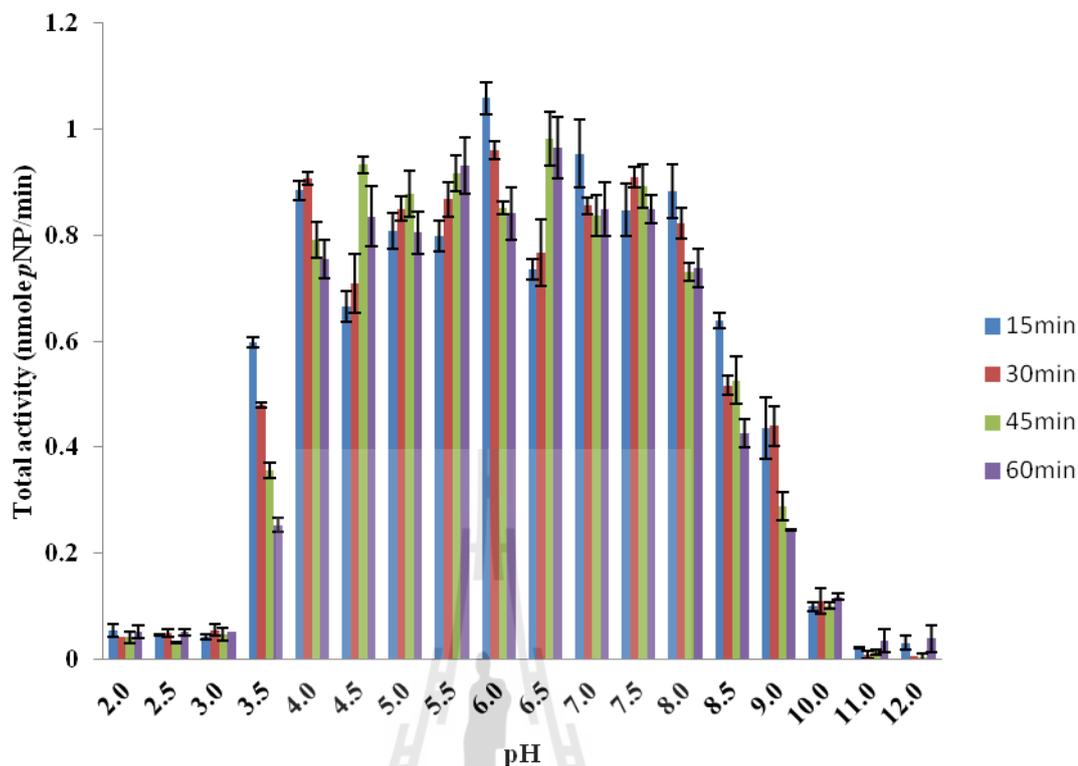
### 3.6 Effect of pH and temperature on the activity and stability of

#### Os4BGlu18

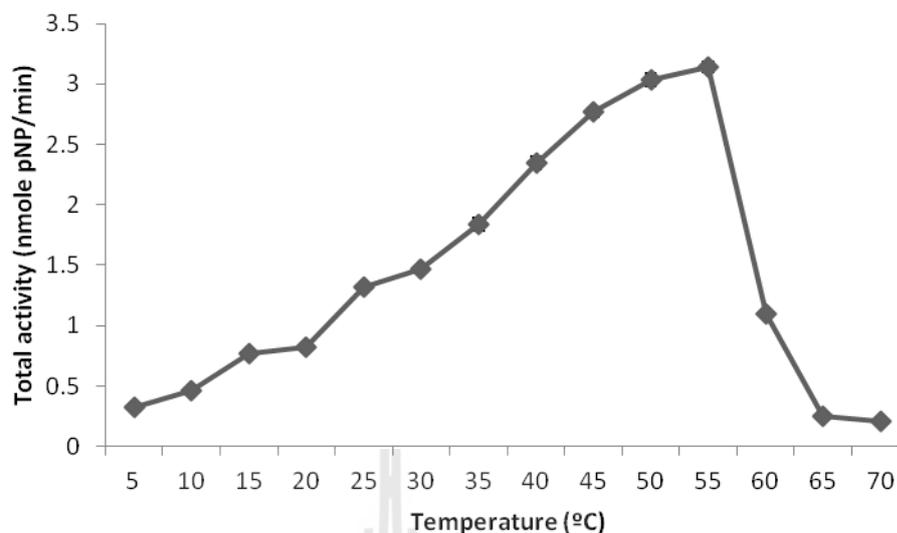
The optimum pH for Os4BGlu18 was found to be 5.0 and the activity dropped by 40% and 60% at pH 3.5 and 6.5, respectively (Figure 3.19). Figure 3.20 shows that Os4BGlu18 maintained similar activity when it was incubated 15-60 min at pH values from pH 4.0 to 8.0 at 25 °C, indicating that this enzyme was stable from pH 4.0 to 8.0, but unstable at pH  $\leq 3.5$  and  $\geq 8.5$ . The temperature optimum for this enzyme is 55 °C (Figure 3.21), but it was only thermostable for 60 min between 20-40 °C as it began to lose activity at 50 °C in 30 min, and only a little activity could be detected at 60 °C (Figure 3.22).



**Figure 3.19** Activity versus pH profile of Os4BGlu18 purified from *E. coli*. The purified enzyme was incubated with 1 mM pNPGlc in 100 mM McIlvaine's universal buffers (citric acid-disodium hydrogen phosphate) ranging from pH 2.0 to 9.0 at 30 °C for 30 min.

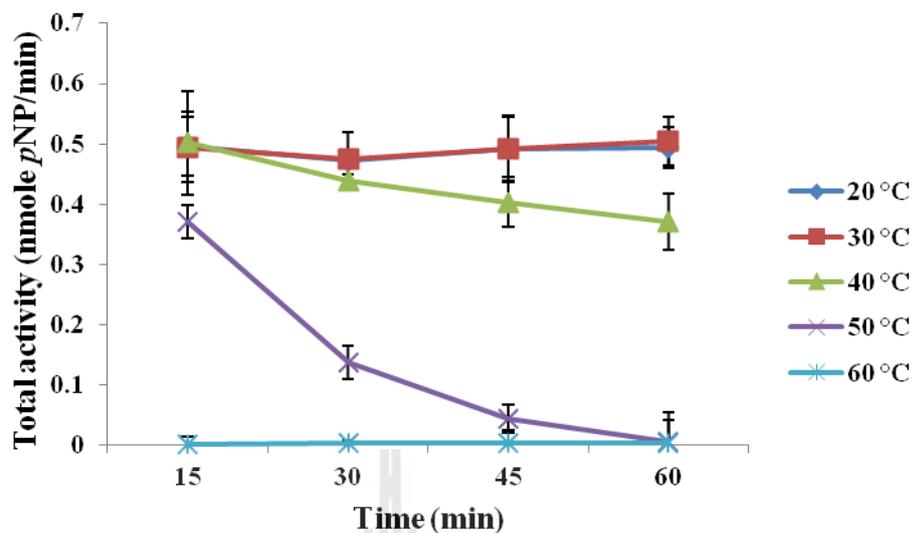


**Figure 3.20** pH stability of Os4BGlu18 activity. The Os4BGlu18 was incubated in universal buffer (pH 2-12) for time periods of 15, 30, 45, and 60 min at 25 °C, then was diluted 100 fold into 100 mM sodium acetate buffer, pH 5.0, and assayed with 1 mM *p*NPGlc for 30 min at 30 °C.



**Figure 3.21** Temperature optimum of Os4BGlu18 fusion protein expressed in *E. coli*.

The activity was assayed by pre-incubating 1  $\mu$ g of enzyme in 50 mM buffer (sodium phosphate, pH 6.5, for Os4BGlu16 or sodium acetate, pH 5.0, for Os4BGlu18) at the specified temperatures for 10 min to bring the reactions up to temperature, then the enzymes were assayed with 1 mM *p*NPGlc for 10 min at the same temperature as the preincubation.



**Figure 3.22** Thermostability of Os4BGlu18. The Os4BGlu18 was incubated at various temperatures (20-60 °C) from 15 to 60 min in 50 mM sodium acetate buffer, pH 5.0, and enzyme activity was determined at 30 °C by adding 1 mM final concentration of *p*NPGlc and incubating for 30 min. Pre means preincubation time at 0 min.

### 3.7 Substrate specificity and kinetic analysis of Os4BGlu16 and

#### Os4BGlu18

The substrate specificities of the purified recombinant Os4BGlu16 and Os4BGlu18 enzymes were determined by testing different *p*NP glycosides and natural and synthetic glucosides, as shown in Table 3.1. First, the glycone specificities of both enzymes were investigated by assaying their activities toward  $\beta$ - and  $\alpha$ -linked *p*-nitrophenyl (*p*NP)-sugars. Among all the *p*NP glycosides tested, Os4BGlu16 hydrolyzed *p*NP- $\beta$ -D-fucopyranoside best, followed by *p*NPGlc, *o*NP- $\beta$ -D-glucopyranoside, *p*NP- $\beta$ -D-galactopyranoside and *p*NP- $\beta$ -D-xylopyranoside, respectively. Among the *p*NP-glycosides, Os4BGlu18 hydrolyzed the same substrates as Os4BGlu16. Os4BGlu16 hydrolyzed *o*NP- $\beta$ -D-glucopyranoside with higher activity than *p*NP- $\beta$ -D-galactopyranoside, while Os4BGlu18 displayed low activity toward *o*NP- $\beta$ -D-glucopyranoside, suggesting a difference in aglycone specificity. To evaluate further their aglycone specificities, hydrolysis of natural and synthetic glycosides and oligosaccharides was tested by TLC analysis. Os4BGlu16 and Os4BGlu18 could hydrolyze the monolignol glucosides *p*-coumarol  $\beta$ -D-glucoside, coniferin and syringin, along with daidzin, esculin, helicin, salicin, indoxyl  $\beta$ -D-glucoside, 4-methylumbelliferyl  $\beta$ -D-glucoside and 4-methylumbelliferyl  $\beta$ -D-fucoside. Moreover, Os4BGlu18 also hydrolyzed arbutin, methyl  $\beta$ -D-glucoside, *n*-octyl  $\beta$ -D-glucoside and *n*-heptyl  $\beta$ -D-glucoside. Neither enzyme could hydrolyze  $\beta$ -1,3-,  $\beta$ -1,4-, or  $\beta$ -1,6-linked gluco-oligosaccharides.

**Table 3.1** Relative activities of Os4BGlu16 and Os4BGlu18 toward nitrophenyl glycosides.

Substrates	Relative activity (%)*	
	Os4BGlu16	Os4BGlu18
<b><i>p</i>NP glycosides</b>		
<i>p</i> NP-β-D-glucopyranoside	100.0±1.9	100.0±4.8
<i>p</i> NP-β-D-mannopyranoside	NA	NA
<i>p</i> NP-β-D-fucopyranoside	297.3±5.9	214.2±2.8
<i>p</i> NP-β-D-galactopyranoside	31.8±2.5	38.4±3.9
<i>p</i> NP-β-D-maltopyranoside	NA	NA
<i>p</i> NP-β-D-cellobiopyranoside	NA	NA
<i>p</i> NP-β-D-xylopyranoside	25.8±0.4	23.1±0.8
<i>p</i> NP-1-thio-β-D-glucopyranoside	NA	NA
<i>p</i> NP-β-L-arabinopyranoside	NA	NA
<i>o</i> NP-β-D-glucopyranoside	46.2±2.5	3.7±1.6
<i>p</i> NP-α-D-glucopyranoside	NA	NA
<i>p</i> NP-α-D-mannopyranoside	NA	NA
<i>p</i> NP-α-D-galactopyranoside	NA	NA
<i>p</i> NP-α-D-glucopyranoside	NA	NA
<i>p</i> NP-α-L-fucopyranoside	NA	NA
<i>p</i> NP-N-acetyl-β-D-glucosaminide	NA	NA

\*The substrates were assayed at 1 mM substrate concentrations at 30 °C for 30 min.

\*\*NA means no activity detected.

**Table 3.2** Activities of Os4BGlu16 and Os4BGlu18 toward natural and synthetic glycosides based on TLC analysis.

Natural and synthetic glucosides	Activity	
	Os4BGlu16	Os4BGlu18
<b>Monolignol glucosides</b>		
<i>p</i> -Coumarol glucoside	+	+
Coniferin	+	+
Syringin	+	+
<b>Natural glucosides</b>		
D-Amygdalin	NA	NA
Quercetin-3-glucoside	NA	NA
Phlorizin	NA	NA
Daidzin	+	+
Gossypin	NA	NA
Mangiferin	NA	NA
Esculin	+	+
Arbutin	NA	+
Helicin	+	+
Salicin	+	+
Naringin	NA	NA
Dhurrin	NA	NA
Apigenin 7-glucoside	NA	NA

The “+” sign means activity detected, while “NA” means no activity detected.

**Table 3.2** Activities of Os4BGlu16 and Os4BGlu18 toward natural and synthetic glycosides based on TLC analysis (Continued).

Substrates	Activity	
	Os4BGlu16	Os4BGlu18
Pyridoxyl $\beta$ -D-glucoside	NA	NA
Trans-zeatin-glucoside	NA	NA
Indoxyl $\beta$ -D-glucoside	+	+
<b>Alkyl glycosides</b>		
Methyl $\beta$ -D-glucoside	NA	+
<i>n</i> -octyl $\beta$ -D-glucoside	NA	+
<i>n</i> -heptyl $\beta$ -D-glucoside	NA	+
4-methylumbelliferyl $\beta$ -D-glucoside	+	+
4-methylumbelliferyl $\beta$ -D-fucoside	+	+
<b>Thioglycosides</b>		
Octyl $\beta$ -D-thio-glucoside	NA	NA
Isopropyl $\beta$ -D-thiogalactoside	NA	NA
<b>Oligosaccharides</b>		
Sophorose	NA	NA
Laminari oligosaccharides	NA	NA
Cello oligosaccharides	NA	NA
Gentiobiose	NA	NA

The “+” sign means activity detected, while “NA” means no activity detected.

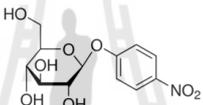
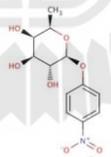
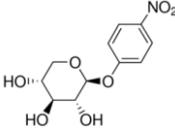
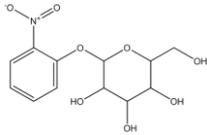
The kinetic parameters for hydrolysis of *p*NP- $\beta$ -D-glucopyranoside, *p*NP- $\beta$ -D-fucopyranoside, *p*NP- $\beta$ -D-xylopyranoside, *o*NP- $\beta$ -D-glucopyranoside, *p*-coumarol  $\beta$ -D-glucoside, coniferin, syringin, salicin, helicin, *n*-octyl  $\beta$ -D-glucoside, *n*-heptyl  $\beta$ -D-glucoside and 4-methylumbelliferyl  $\beta$ -D-glucoside were determined for Os4BGlu16 and Os4BGlu18 (Table 3.3). Among *p*NP-glycosides, Os4BGlu16 has highest catalytic efficiency with *p*NPGlc ( $k_{cat}/K_M=1.062$ ), followed by *p*NP- $\beta$ -D-fucopyranoside ( $k_{cat}/K_M=0.896$ ), *p*NP- $\beta$ -D-xylopyranoside ( $k_{cat}/K_M=0.263$ ) and *o*NP- $\beta$ -D-glucopyranoside ( $k_{cat}/K_M=0.275$ ). Os4BGlu18 preferred *p*NP- $\beta$ -D-fucopyranoside ( $k_{cat}/K_M=4.06$ ), followed by *p*NPGlc ( $k_{cat}/K_M=0.789$ ) and *p*NP- $\beta$ -D-xylopyranoside ( $k_{cat}/K_M=0.302$ ) (Table 3.3).

Among monolignol glucoside substrates, Os4BGlu16 hydrolyzed syringin with a  $k_{cat}/K_M$  of  $22.8 \text{ mM}^{-1}\text{s}^{-1}$ , followed by coniferin and *p*-coumarol glucoside, with  $k_{cat}/K_M$  values of 21.6 and  $6.2 \text{ mM}^{-1}\text{s}^{-1}$ , respectively. In comparison, Os4BGlu18 hydrolyzed coniferin with a  $k_{cat}/K_M$  of  $31.9 \text{ mM}^{-1}\text{s}^{-1}$ , followed by syringin and *p*-coumarol glucoside, with  $k_{cat}/K_M$  values of 24.0 and  $1.41 \text{ mM}^{-1}\text{s}^{-1}$ , respectively. However, these high  $k_{cat}/K_M$  values were driven by the high  $k_{cat}$  values, with rather high  $K_M$  for these substrates, and both enzymes have lower  $K_M$  values for syringin than for the other two monolignol glucosides (Table 3.4). The high  $K_M$  and  $k_{cat}/K_M$  values for these substrates would make the rates of monolignol glucoside hydrolysis by Os4BGlu16 and Os4BGlu18 highly sensitive to the substrate concentrations in the plant.

Hydrolysis of other synthetic and natural glycosides was tested to assess the aglycone specificity of Os4BGlu16 and Os4BGlu18. Os4BGlu16 hydrolyzed helicin, 4-methylumbelliferyl  $\beta$ -D-glucopyranoside, and salicin with  $k_{cat}/K_M$  values of 0.145,

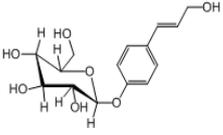
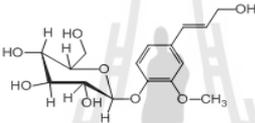
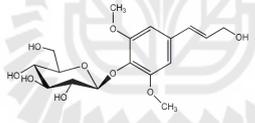
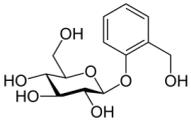
0.139, and 0.035  $\text{mM}^{-1}\text{s}^{-1}$ , respectively. Os4BGlu18 showed broader substrate specificity than that of Os4BGlu16, in that it could hydrolyze 4-methylumbelliferyl  $\beta$ -D-glucopyranoside, *n*-octyl  $\beta$ -D-glucopyranoside, *n*-heptyl  $\beta$ -D-glucopyranoside, and salicin with  $k_{cat}/K_M$  values of 1.054, 0.998, 0.452, and 0.0321  $\text{mM}^{-1}\text{s}^{-1}$ , respectively (Table 3.4).

**Table 3.3** Kinetic parameters of Os4BGlu16 and Os4BGlu18 for hydrolysis of nitrophenyl glycosides.

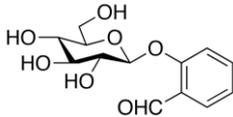
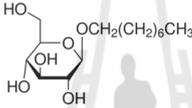
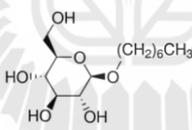
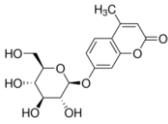
Substrate	Chemical structure	Os4BGlu16	Os4BGlu18
<i>p</i> NP- $\beta$ -D-glucopyranoside			
$K_M$ (mM)		2.88±0.26	4.84±0.15
$k_{cat}$ ( $\text{s}^{-1}$ )		3.05±0.22	3.82±0.16
$k_{cat}/K_M$ ( $\text{mM}^{-1}\text{s}^{-1}$ )		1.062±0.077	0.789±0.012
<i>p</i> NP- $\beta$ -D-fucopyranoside			
$K_M$ (mM)		7.72±0.45	5.86±0.58
$k_{cat}$ ( $\text{s}^{-1}$ )		6.92±0.15	23.8±2.3
$k_{cat}/K_M$ ( $\text{mM}^{-1}\text{s}^{-1}$ )		0.896±0.035	4.06±0.07
<i>p</i> NP- $\beta$ -D-xylopyranoside			
$K_M$ (mM)		2.99±0.27	1.26±0.064
$k_{cat}$ ( $\text{s}^{-1}$ )		0.788±0.028	0.381±0.009
$k_{cat}/K_M$ ( $\text{mM}^{-1}\text{s}^{-1}$ )		0.263±0.034	0.302±0.018
<i>o</i> NP- $\beta$ -D-glucopyranoside			
$K_M$ (mM)		23.1±1.1	ND
$k_{cat}$ ( $\text{s}^{-1}$ )		6.34±0.52	ND
$k_{cat}/K_M$ ( $\text{mM}^{-1}\text{s}^{-1}$ )		0.275±0.014	ND

ND means not determined due to low activity.

**Table 3.4** Kinetic parameters of Os4BGlu16 and Os4BGlu18 for hydrolysis of monolignol glucosides and other aryl and alkyl glucosides.

Substrate	Chemical structure	Os4BGlu16	Os4BGlu18
<b><i>p</i>-coumarol glucoside</b>			
$K_M$ (mM)		13.10±0.88	8.15±0.20
$k_{cat}$ (s <sup>-1</sup> )		81.5±3.0	11.49±0.08
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		6.22±0.19	1.411±0.024
<b>coniferin</b>			
$K_M$ (mM)		19.9±2.2	8.02±0.19
$k_{cat}$ (s <sup>-1</sup> )		429±31	255.8±4.1
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		21.6±0.89	31.9±0.3
<b>syringin</b>			
$K_M$ (mM)		4.66±0.13	5.34±0.06
$k_{cat}$ (s <sup>-1</sup> )		106.3±2.1	127.9±1.7
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		22.8±0.2	24.0±0.1
<b>salicin</b>			
$K_M$ (mM)		0.622±0.058	5.41±0.21
$k_{cat}$ (s <sup>-1</sup> )		0.0218±0.0027	0.1736±0.0009
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		0.0350±0.0010	0.0321±0.0011

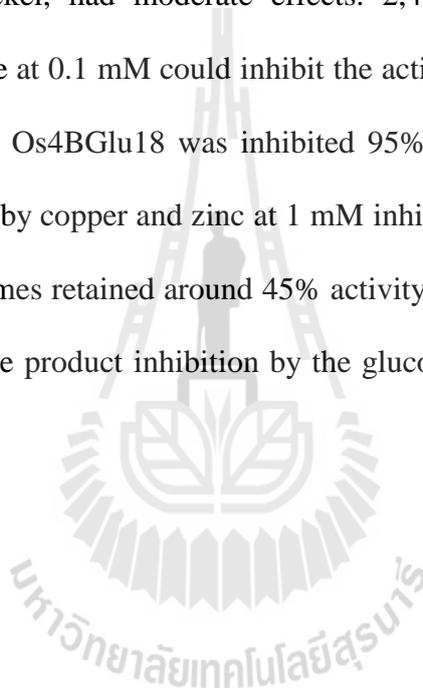
**Table 3.4** Kinetic parameters of Os4BGlu16 and Os4BGlu18 for hydrolysis of monoglucosides and other aryl and alkyl glucosides (Continued).

Substrate	Chemical structure	Os4BGlu16	Os4BGlu18
<b>Helicin</b>			
$K_M$ (mM)		18.4±1.4	ND
$k_{cat}$ (s <sup>-1</sup> )		2.66±0.12	ND
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		0.145±0.005	ND
<b><i>n</i>-octyl β-D-glucopyranoside</b>			
$K_M$ (mM)		ND	8.05±0.64
$k_{cat}$ (s <sup>-1</sup> )		ND	8.03±0.15
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		ND	0.998±0.067
<b><i>n</i>-heptyl β-D-glucopyranoside</b>			
$K_M$ (mM)		ND	24.5±2.0
$k_{cat}$ (s <sup>-1</sup> )		ND	11.05±0.66
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		ND	0.452±0.010
<b>4-methyl umbelliferyl β-D-glucopyranoside</b>			
$K_M$ (mM)		4.62±0.38	9.39±0.48
$k_{cat}$ (s <sup>-1</sup> )		0.641±0.027	9.89±0.22
$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )		0.139±0.005	1.054±0.049

ND means not determined, due to low activity.

### **3.8 Inhibition of Os4BGlu16 and Os4BGlu18 by metal salts and organic inhibitors**

The effects of various chemicals that have been reported to affect  $\beta$ -glucosidase activity were tested, as shown in Table 3.5. Mercury ion and glucono  $\delta$ -lactone caused nearly complete inhibition of Os4BGlu16 activity at 1 mM, while copper, zinc, and nickel, had moderate effects. 2,4-Dinitrophenyl- $\beta$ -D-2-deoxy-2-fluoro-glucopyranoside at 0.1 mM could inhibit the activity of Os4BGlu16 65% when pre-incubated for 1 h. Os4BGlu18 was inhibited 95% by 1 mM glucono  $\delta$ -lactone, and partially inhibited by copper and zinc at 1 mM inhibitor and conduritol B epoxide at 0.1 mM. Both enzymes retained around 45% activity with 500 mM glucose present in the reaction, so little product inhibition by the glucose is expected to occur in the plant.

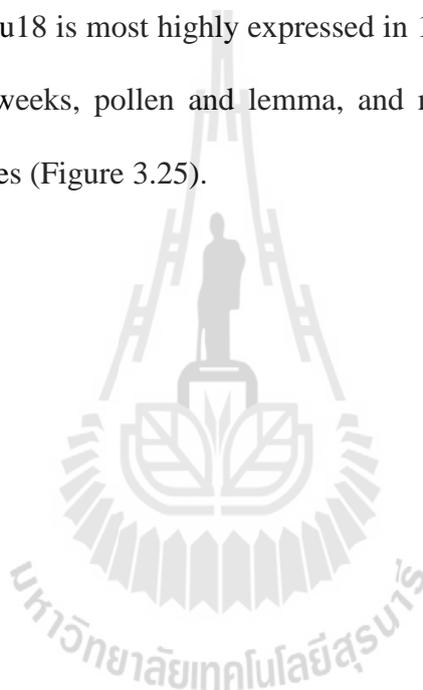


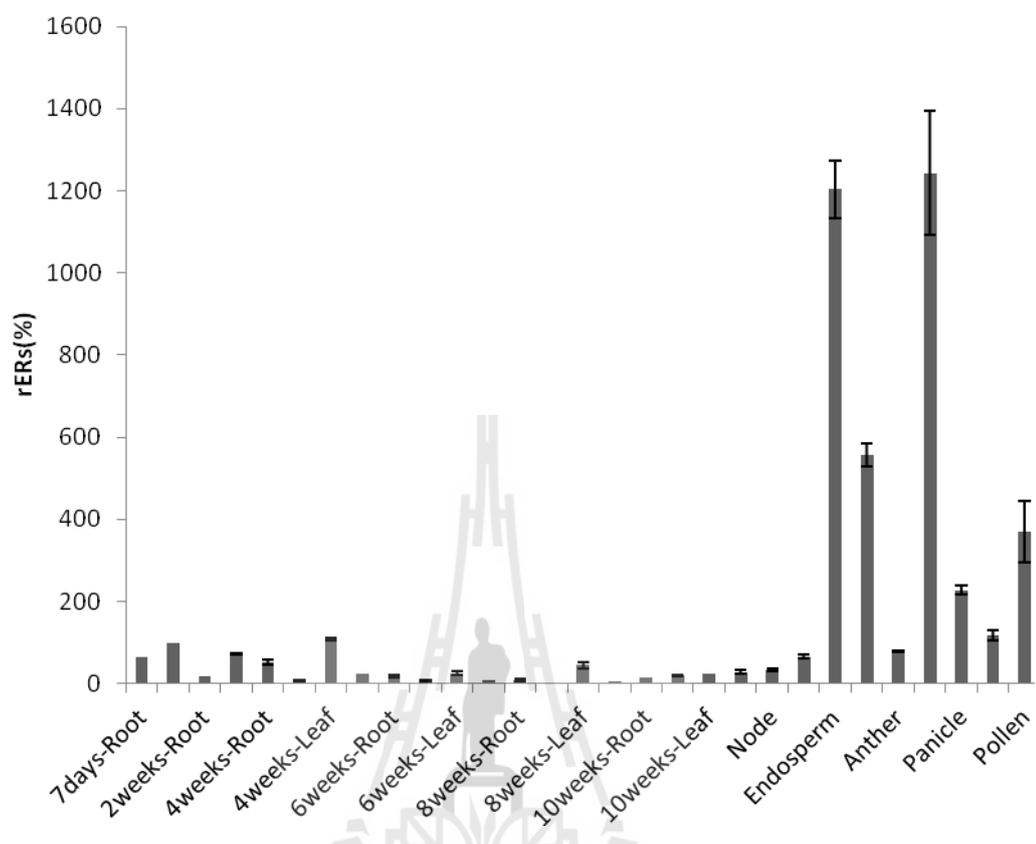
**Table 3.5** Effects of EDTA, metal salts and inhibitors on Os4BGlu16 and Os4BGlu18 activity.

Potential inhibitor	Concentration (mM)	Relative activity of Os4BGlu16 (%)	Relative activity of Os4BGlu18 (%)
None	0	100.0±6.9	100.0±6.5
EDTA	1	102.1±0.9	97.9±3.3
CaCl <sub>2</sub>	1	102.6±2.0	97.8±2.9
CoCl <sub>2</sub>	1	93.5±2.3	93.7±2.8
HgCl <sub>2</sub>	1	3.7±0.7	89.6±8.6
MgCl <sub>2</sub>	1	89.4±4.2	96.2±5.2
MnCl <sub>2</sub>	1	106.8±5.4	96.7±3.6
FeCl <sub>3</sub>	1	101.1±2.7	101.7±2.0
NiSO <sub>4</sub>	1	68.8±1.1	94.7±0.6
ZnSO <sub>4</sub>	1	66.6±4.2	71.6±1.7
CuSO <sub>4</sub>	1	51.4±0.9	53.5±1.0
Glucono δ-lactone	1	4.2±0.35	5.3±0.5
2,4-Dinitrophenyl-β-D-2-deoxy-2-fluoro-glucopyranoside	0.1	35.1±1.1	101.9±1.3
Conduritol B epoxide	0.1	98.2±0.9	52.5±4.5
D-Glucose	500	45.4±0.4	45.04±0.05

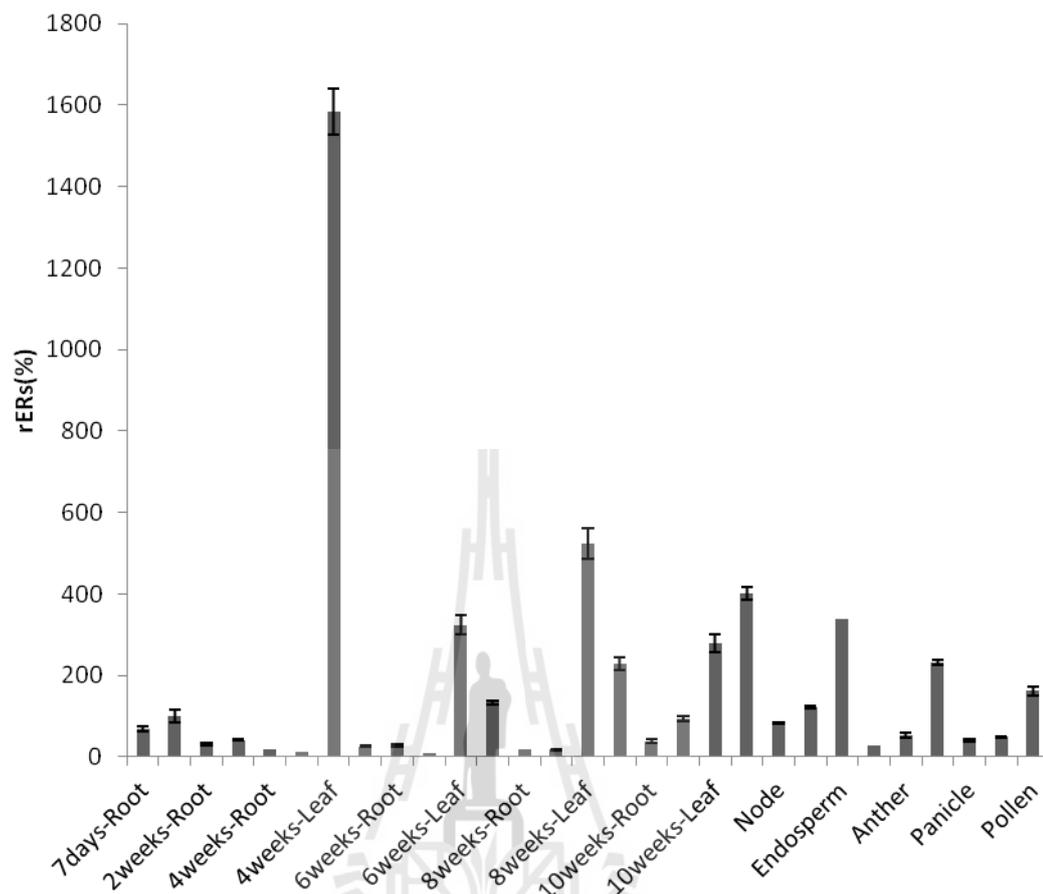
### 3.9 *In planta* expression analysis

Quantitative real-time RT-PCR was used to measure the level of the putative monolignol  $\beta$ -glucosidases gene expression in various rice organs. The Os4BGlu14 gene is most highly expressed in reproductive tissues, especially in endosperm, lemma, embryo, pollen, panicle and flower (Figure 3.23). Os4BGlu16 is most highly expressed in 4 weeks to 10 weeks-old leaves, endosperm and lemma (Figure 3.24). Os4BGlu18 is most highly expressed in 1-week-old seedling, stem, leaf and leaf sheath at 4 weeks, pollen and lemma, and moderately expressed in other young vegetative tissues (Figure 3.25).

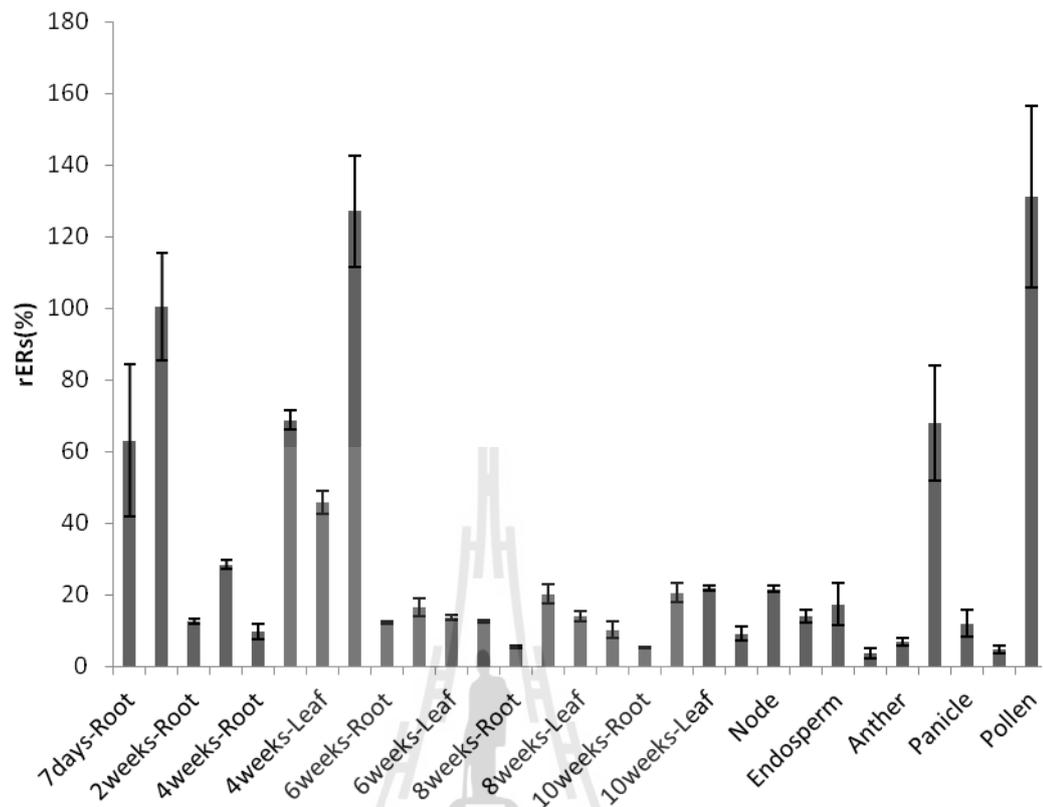




**Figure 3.23** Gene expression analysis of rice Os4BGlu14 from qRT-PCR. RNA was extracted from tissues from 7 days to 10 weeks and endosperm and embryo of mature, dried seed, and the relative mRNA amounts for each gene assessed by qRT-PCR with gene-specific primers. Actin was used as a reference gene and 7-day shoots as the reference tissue, which is 100% in the graphs.



**Figure 3.24** Gene expression analysis of rice Os4BGlu16 from qRT-PCR. RNA was extracted from tissues from 7 days to 10 weeks and endosperm and embryo of mature, dried seed, and the relative mRNA amounts for each gene assessed by qRT-PCR with gene-specific primers. Actin was used as a reference gene and 7-day shoots as the reference tissue, which is 100% in the graphs.

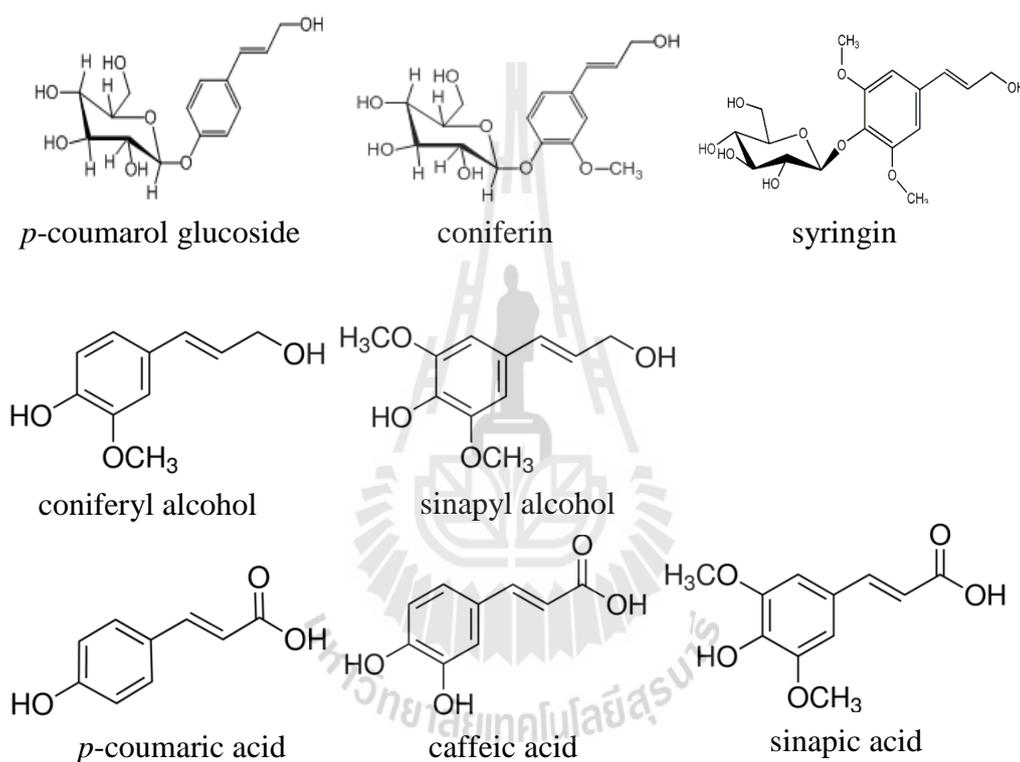


**Figure 3.25** Gene expression analysis of rice Os4BGlu18 from qRT-PCR. RNA was extracted from tissues from 7 days to 10 weeks and endosperm and embryo of mature, dried seed, and the relative mRNA amounts for each gene assessed by qRT-PCR with gene-specific primers. Actin was used as a reference gene and 7-day shoots as the reference tissue, which is 100% in the graphs.

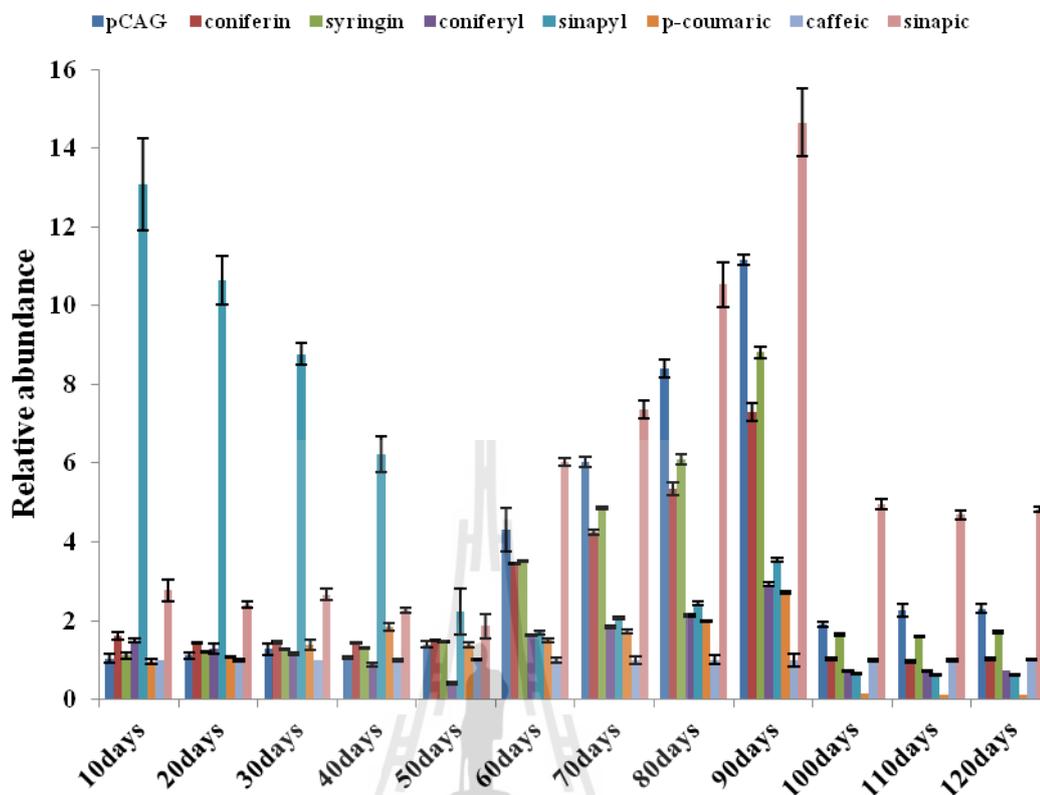
### 3.10 Detection of monolignol compounds in rice KDML105 by UPLC-MS

The content of monolignol compounds in rice tissues, as analyzed by UPLC-MS, is shown in Figure 3.26. The compounds measured included the monolignol glucosides (*p*-coumarol glucoside, coniferin, syringin), monolignols (coniferyl and sinapyl alcohol), and the lignin precursor metabolic intermediate hydroxycinnamic acids (*p*-coumaric, caffeic, and sinapic acids). In the early stages of developing roots, sinapyl alcohol levels were relatively higher than other compounds. In 2-3 month-old roots, the levels of *p*-coumarol glucoside, coniferin, syringin, and sinapic acid were dramatically increased. The level of sinapic acid was still higher than the other compounds from 100 to 120 days (Figure 3.27). In the rice leaf extracts, *p*-coumarol glucoside, coniferin, syringin, and sinapic acid were higher than coniferyl alcohol and the lignin precursor metabolic intermediates in all ages of leaf. Their levels increased at 60 days until 90 days, then decreased at 100 days (Figure 3.28). The monolignol compounds in the leaf sheath showed the same patterns as in the leaf, but the level of sinapyl alcohol at 20 days to 40 days was highest compared to the other compounds (Figure 3.29). In extracts from 60-90 day-old stems, *p*-coumarol glucoside, coniferin, syringin, and sinapic acid were high compared with other monolignol compounds, and in extracts from 100-120 day-old sinapic acid was highest. In the rice flower extract, *p*-coumarol glucoside was highest, followed by syringin, sinapic acid, coniferin, caffeic acid, sinapyl alcohol, coniferyl alcohol, and *p*-coumaric acid, respectively. Sinapic acid was higher than other compounds in seed at all ages (Figure 3.30). The level of *p*-coumaric acid in every stage and tissue was lower than the other

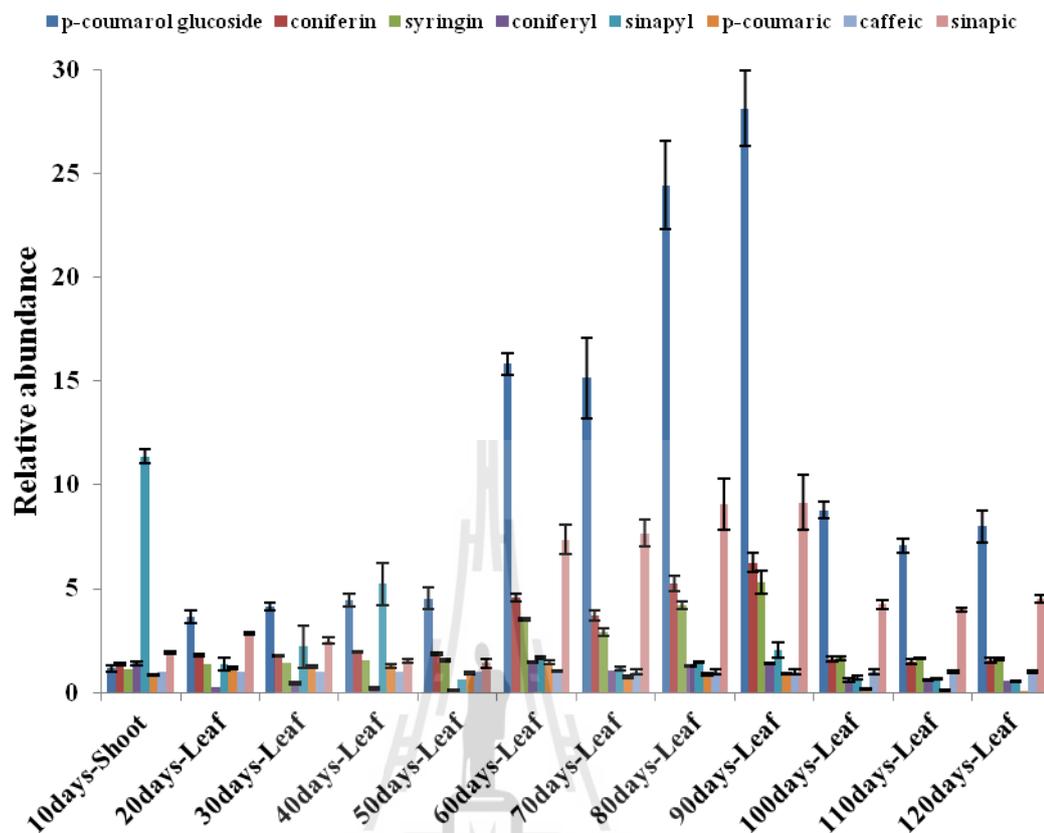
compounds and after 90 days it was undetectable. These values give relative concentrations under the assumption that similar losses occurred for extraction of all compounds from all tissues, since no internal standard was available to generate truly quantitative data.



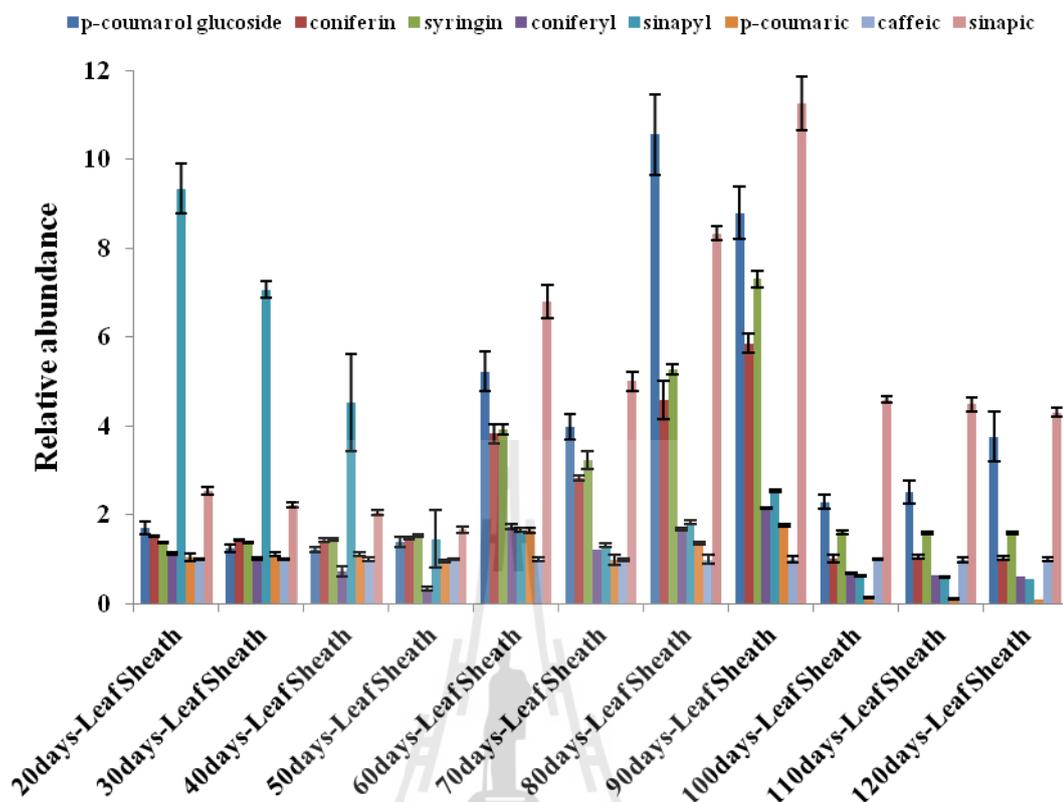
**Figure 3.26** Chemical structures of monolignol compounds. The monolignol glucoside substrates *p*-coumarol glucoside, coniferin, and syringin. Monolignol alcohol monomer coniferyl and sinapyl alcohol, and lignin precursor metabolic intermediates *p*-coumaric acid, caffeic acid, and sinapic acid.



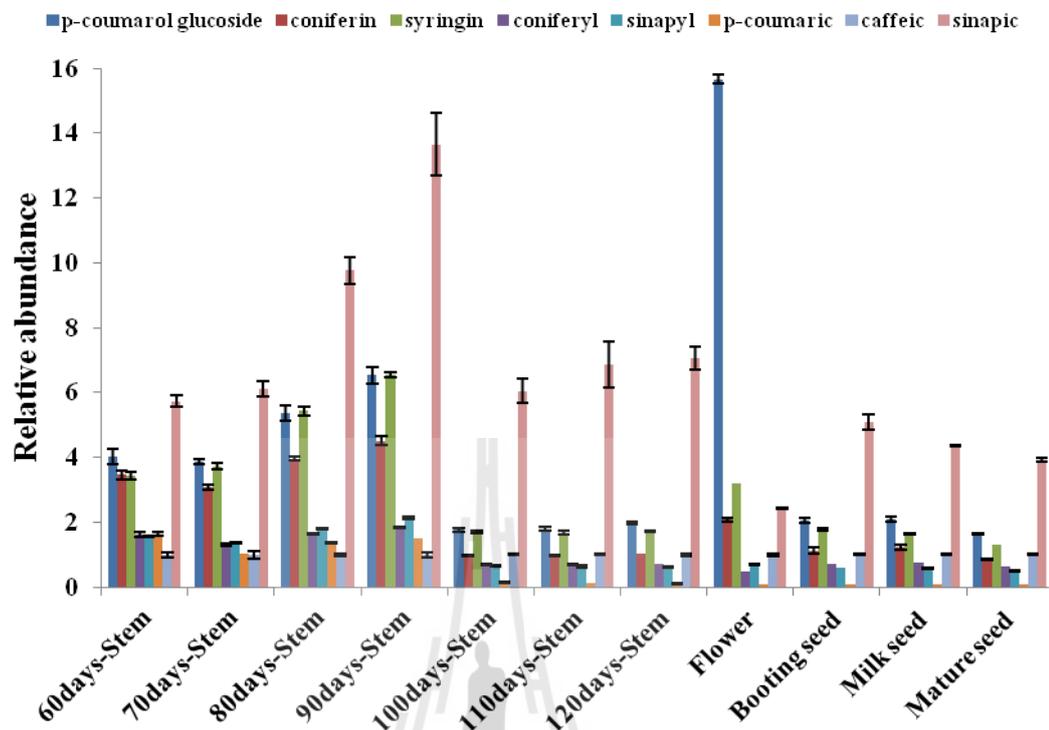
**Figure 3.27** Relative abundance of monolignol compounds and intermediates in their biosynthesis in root extracts from rice (cv. KDML105) from 10 days to 120 days, as determined by UPLC-MS analysis. Caffeic acid was used as a reference compound with the abundance shown relative to the measured caffeic acid concentration. The abundance values of each compound based on the abundance of their selected ions at in the peak at their specific elution time relative to a standard curve of the same compound. The labels coniferyl and sinapyl refer to the corresponding alcohols, while the labels *p*-coumaric, caffeic and sinapic refer to the corresponding acids.



**Figure 3.28** Relative abundance of monolignol compounds and intermediates in their biosynthesis in leaf extracts from rice (cv. KDML105) from 10 days to 120 days, as determined by UPLC-MS analysis. Caffeic acid was used as a reference compound with the abundance shown relative to the measured caffeic acid concentration. The abundance values of each compound based on the abundance of their selected ions at in the peak at their specific elution time relative to a standard curve of the same compound. The labels coniferyl and sinapyl refer to the corresponding alcohols, while the labels *p*-coumaric, caffeic and sinapic refer to the corresponding acids.



**Figure 3.29** Relative abundance of monolignol compounds and intermediates in their biosynthesis in leaf sheath extracts from rice (cv. KDML105) from 20 days to 120 days, as determined by UPLC-MS analysis. Caffeic acid was used as a reference compound with the abundance shown relative to the measured caffeic acid concentration. The abundance values of each compound based on the abundance of their selected ions at in the peak at their specific elution time relative to a standard curve of the same compound. The labels coniferyl and sinapyl refer to the corresponding alcohols, while the labels *p*-coumaric, caffeic and sinapic refer to the corresponding acids.



**Figure 3.30** Relative abundance of monolignol compounds and intermediates in their biosynthesis in stem extracts from rice (cv. KDML105) from 60 days to 120 days, flower, and seed, as determined by UPLC-MS analysis. Caffeic acid was used as a reference compound with the abundance shown relative to the measured caffeic acid concentration. The abundance values of each compound based on the abundance of their selected ions at in the peak at their specific elution time relative to a standard curve of the same compound. The labels coniferyl and sinapyl refer to the corresponding alcohols, while the labels *p*-coumaric, caffeic and sinapic refer to the corresponding acids.

## CHAPTER IV

### DISCUSSION

#### 4.1 Rice monolignol $\beta$ -glucosidase sequence identification and analysis

The protein sequence alignment in Figure 3.2 compares the putative rice monolignol  $\beta$ -glucosidases, Os4BGlu14, Os4BGlu16, and Os4BGlu18, with the related monolignol  $\beta$ -glucosidases from *A. thaliana*, AtBGLU45 and AtBGLU46, their closely related homologue AtBGLU47, and pine coniferin  $\beta$ -glucosidase (Escamilla-Treviño et al., 2006; Dharmawardhana et al., 1995; Dharmawardhana and Ellis, 1998). Os3BGlu7 (also called rice BGlu1) is included as a non monolignol  $\beta$ -glucosidase with a known X-ray crystallographic structure (Chuenchor et al., 2008). Os4BGlu14, Os4BGlu16 and Os4BGlu18 shared 58–60% amino acid sequence identity with each other, 47–55% identity with Arabidopsis BGLU45, BGLU46, and BGLU47 and 45–47% with Os3BGlu7 (rice BGlu1). It is also notable that residue W358 of Os3BGlu7, which has been shown to stack aglycones in plant GH1  $\beta$ -glucosidases (Czjzek et al., 2000; Verdoucq et al., 2004; Barleben et al., 2007; Chuenchor et al., 2008; Chuenchor et al., 2011) is replaced by smaller residues in the rice and pine monolignol  $\beta$ -glucosidase group, similar to the disaccharidase primeverosidase (Saino et al., 2014). The loop containing this interacting tryptophan

is greatly shortened in AtBGLU47, although AtBGLU45 and AtBGLU46 maintain tryptophan in this position.

## 4.2 Expression and purification of Os4BGlu14

The cloned Os4BGlu14 cDNA sequence had 1,551 bases which match 100% with *Oryza sativa* Japonica Group Os04g0513100 mRNA and cDNA clone J013128A11 accession number AK067841. The longest Os4BGlu14 cDNA ORF translates to a protein containing 516 amino acid residues and mature protein 493 amino acid residues. The SignalP program predicted Os4BGlu14 to have an N-terminal secretory signal sequences with a cleavage site between residues 23 and 24 (VSA-VDR). To produce mature Os4BGlu14 in *E.coli*, the forward and reverse primers were designed to amplify a cDNA encoding Os4BGlu14 without the signal sequence. The cDNA encoding the mature Os4BGlu14 was cloned into the pET32a+ expression vector, as was previously successful with other rice  $\beta$ -glucosidase (Opassiri et al., 2003, 2006; Kuntothom et al., 2009; Seshadri et al., 2009). However, Opassiri et al. (2006) reported that most putative family GH1  $\beta$ -glucosidase enzymes encoded in the rice genome have the catalytic acid/base glutamate in the conserved sequence “TXNEP”, except for Os4BGlu14 and Os9BGlu33, in which the glutamate (E) is replaced with glutamine (Q). So, Os4BGlu14 and Os9BGlu33 were predicted to be inactive  $\beta$ -glucosidases. To produce active Os4BGlu14, pET32a/Os4BGlu14 was used as a template to mutant Q191 to E. The predicted molecular weight and isoelectric point (pI) of the Os4BGlu14 fusion protein were approximately 73 kDa and 6.35, respectively. I attempted to express the wild type Os4BGlu14 from pET32a/Os4BGlu14 and mutant from pET32a/Os4BGlu14Q191E with N-terminal

thioredoxin and His<sub>6</sub> fusion tags in *E.coli* strains Origami(DE3), Origami B(DE3), RosettaGami(DE3), and BL21(DE3). Expression from both plasmid clones was attempted at various temperatures (10-37 °C) and concentrations of IPTG (0-1 mM) to find the appropriate expression condition. However, most of the expressed protein was observed in the insoluble fraction of the cell lysates, and no activity could be detected in the soluble cell lysate, suggesting that this Os4BGlu14  $\beta$ -glucosidase was not expressed in soluble, active form. Since Os4BGlu14 was predicted to be inactive (Opassiri et al., 2006), expression was not pursued further. Nevertheless, the corresponding *Agrobacterium* sp.  $\beta$ -glucosidase E170Q mutant (Abg E170Q) has thioglycoligase activity which can transglycosylate the sugar moiety onto a thio sugar acceptor. Abg E170Q showed the highest catalytic efficiency compared to other mutants selected for this activity (E170G, E170N, E170S, E170A, E170T) when  $\beta$ -D-glucopyranosyl azide (Glc-N<sub>3</sub>) was used as a donor and 4-methylumbelliferyl 4-deoxy-4-thio- $\beta$ -D-glucopyranoside (4SGlcMU) as acceptor (Müllegger et al., 2005). The Gln residue could assist binding at the deglycosylation step by hydrogen bonding with the incoming thio sugar acceptor. Additionally, Chuenchor et al. (2011) reported that the activity of acid/base mutant of rice BGlu1 E176Q can be rescued by anionic nucleophiles such as ascorbate, azide, acetate, formate, KF, TFA, and cyanate in 50 mM MES buffer. Testing of more expression systems and activity assays may be required to understand the function of Os4BGlu14.

### 4.3 Expression and purification of Os4BGlu16

The predicted Os4BGlu16 cDNA sequence also had 1,551 bases, with 100% match between the *Oryza sativa* Japonica group Os04g0513400 mRNA and cDNA clones J013081O13 and J013164L07, corresponding to NCBI Genbank accession numbers AK066850 and AK068772, respectively. The Os4BGlu16 cDNA ORF translates to a protein containing 516 amino acid residues and mature protein 489 amino acid residues. The SignalP program predicted that Os4BGlu16 has a secretory signal sequence with a cleavage site between amino acid residues 27 and 28 (ARG-LRR). To produce mature Os4BGlu16 in *E.coli*, the forward and reverse primers were designed as for Os4BGlu14, but several attempts resulted in no correct PCR product clones for mature Os4BGlu16, with various deletions, insertions and mutations occurring in the clones obtained. To facilitate cloning and expression in yeast, an *Os4BGlu16* cDNA (Genbank accession number KJ579205) optimized for expression in *P. pastoris* was synthesized.

The optimized *Os4BGlu16* cDNA was used to produce a secreted, N-terminally H-tagged protein in *P. pastoris*, with an expected molecular weight (MW) of 58.2 kDa and pI of 6.4. After purification from the pichia media by IMAC,  $\beta$ -glucosidase activity was detected with pNPGlc substrate and a broad band was detected above 67 kDa on the SDS-PAGE gel (Figure 3.7). After deglycosylation with endoglycosidase H, a single band of approximately 61 kDa was detected on a Coomassie-stained SDS-PAGE gel (Figure 3.7). Thus, a heterogeneously glycosylated, active Os4BGlu16 was expressed in pichia.

#### 4.4 Expression and purification of Os4BGlu18

The Os4BGlu18 mRNA sequence contains 1,518 bases, corresponding to that of the *Oryza sativa* Japonica Group Os04g0513900 mRNA and cDNA clone J090046N08, accession number AK058333. The longest *Os4BGlu18* cDNA ORF translates to a protein containing 505 amino acid residues and mature protein 479 amino acid residues. The SignalP program predicted that the Os4BGlu18 precursor contains an N-terminal signal sequence prepeptide with a cleavage site between amino acid residues 26 and 27 (in the sequence ASA-IHR). The cDNA encoding mature Os4BGlu18 was amplified from 7-day-old rice shoot and cloned into the pET32a+ expression vector, as was done with Os4BGlu14. The Os4BGlu18 fusion protein with N-terminal thioredoxin, His<sub>6</sub> and S-tags, which was predicted to have a molecular weight of 73 kDa and pI of 5.3, was optimally expressed in *E. coli* strain Origami(DE3) induced with 0.1 mM IPTG at 18 °C for 16 h, and significant pNPGlc hydrolysis activity was detected in the crude cell extract. Protein purification by ion-exchange chromatography, hydrophobic interaction chromatography and IMAC produced a protein of approximately 75 kDa that was approximately 90% pure, as judged by SDS-PAGE (Figure 3.8). Although the IMAC step provided most of the purification, it was done last, since purification with IMAC first resulted in less pure Os4BGlu18 fusion protein and more degradation.

#### 4.5 Substrate specificity of Os4BGlu16 and Os4BGlu18

The hydrolysis specificity of Os4BGlu18 is much broader than that observed for Os4BGlu16. Both enzymes hydrolyzed several *p*NP-glycosides, including *p*NP- $\beta$ -D-glucopyranoside (*p*NPGlc), *p*NP- $\beta$ -D-fucopyranoside, *o*NP- $\beta$ -D-glucopyranoside (*o*NPGlc), *p*NP- $\beta$ -D-galactopyranoside (*p*NPGal) and *p*NP- $\beta$ -D-xylopyranoside (*p*NPXyl). Similarly, other known monolignol  $\beta$ -glucosidases, including native pine tree coniferin  $\beta$ -glucosidase (CBG) and *E. coli*-expressed CBG, crude cell wall preparations from *Cicer urietinum* L. cell suspension cultures, and *Arabidopsis* BGlu45 and BGlu46 recombinant proteins are able to hydrolyze *p*NPGlc, *p*NPGal, *p*NPXyl, *o*NPGlc, and *o*NPGal (Dharmawardhana et al., 1995; Dharmawardhana and Ellis, 1998; Hösel et al., 1978; Escamilla-Treviño et al., 2006).

Os4BGlu16 showed higher catalytic efficiency with *p*NPGlc and *p*NP- $\beta$ -D-fucopyranoside, which were hydrolyzed at around 4-fold higher rates than *p*NP- $\beta$ -D-xylopyranoside and *o*NP- $\beta$ -D-glucopyranoside. Os4BGlu18 preferred *p*NP- $\beta$ -D-fucopyranoside, with catalytic efficiency ( $k_{cat}/K_m$ ) 5-fold higher than that for *p*NPGlc and 13-fold higher than that for *p*NP- $\beta$ -D-xylopyranoside. Since  $\beta$ -D-fucopyranosides have rarely been reported in nature and many GH1  $\beta$ -glucosidases have high  $\beta$ -D-fucosidase activity (Ketudat Cairns and Esen, 2010), these data support the designation of these enzymes as  $\beta$ -glucosidases.

Among natural and synthetic glycosides, Os4BGlu16 and Os4BGlu18 could hydrolyze monolignol glucosides (*p*-coumarol glucoside, coniferin and syringin), isoflavonoid glucosides (daidzin), phenolic alcohol glucoside (salicin and esculin), plant glycosides with similarity to phytohormones (helicin and indoxyl  $\beta$ -D-glucoside), and alkyl glycosides (4-methylumbelliferyl  $\beta$ -D-glucoside and 4-

methylumbelliferyl  $\beta$ -D-fucoside). Neither enzyme could hydrolyze oligosaccharide substrates.

By comparing the kinetic parameters of Os4BGlu16 and Os4BGlu18, it was found that Os4BGlu16 could hydrolyze coniferin and syringin at approximately the same catalytic efficiencies ( $k_{\text{cat}}/K_m$ ), which were higher than that for *p*-coumarol glucoside by around 3.5 fold. Os4BGlu18 hydrolyzed coniferin at around 1.3- and 22.6-fold higher rates than that of syringin and *p*-coumarol glucoside, respectively. In addition, Os4BGlu18 could hydrolyze coniferin and syringin at higher efficiencies than Os4BGlu16, but Os4BGlu16 hydrolyzed *p*-coumarol glucoside about 4.4-fold more efficiently than Os4BGlu18. Characterization of the relative activities of *E. coli*-expressed CBG and native CBG revealed that these enzymes hydrolyzed coniferin at a rate 1.5-2-fold higher than syringin around and 10-fold higher than salicin (Dharmawardhana et al., 1995). Hösel et al. (1978) reported that the  $\beta$ -glucosidase from a crude particulate fraction of pine xylem show high specificity activity to coniferin ( $4.2 \mu\text{molmin}^{-1}\text{mg}^{-1}$ ) which was approximately 7-fold higher than syringin and 35-fold higher than that for salicin. CBG and pine cell wall extract isoenzyme exhibited activity similar to rice Os4BGlu16, which prefers coniferin to other substrates. Furthermore, *Arabidopsis* BGlu45 has same relative activity toward coniferin and syringin as Os4BGlu18 and this was about 12.6-14.5-fold higher than *p*-coumarol glucoside. On the other hand, *Arabidopsis* BGlu46 showed high activity against salicin, which was 1.4-fold higher than *p*-coumarol glucoside and 12.5- and 16.6-fold, higher than coniferin and syringin, respectively (Escamilla-Treviño et al., 2006).

#### 4.6 Inhibition of Os4BGlu16 and Os4BGlu18

The hydrolysis activity of Os4BGlu16 was near completely inhibited by 1 mM mercury ion and glucono  $\delta$ -lactone, which is a strong inhibitor for other rice  $\beta$ -glucosidases such as Os3BGlu7 (rice BGlu1), and Os1BGlu4 (Opassiri et al., 2003; Rouyi et al., 2014). Mercury ion and glucono  $\delta$ -lactone have also been reported to be strong inhibitors of *Dalbergia cochinchinensis*  $\beta$ -glucosidase and *Cicer arietinum* isoflavonoid  $\beta$ -glucosidases, whereas 1 mM glucono  $\delta$ -lactone inhibited *D. nigescens* isoflavonoid 7-O- $\beta$ -apiosyl-glucoside  $\beta$ -glucosidase only 34% (Srisomsap et al., 1996; Hösel and Barz 1975; Chuankhayan et al., 2005). On the other hand, mercury ion could not inhibit the activity of Os4BGlu18, even though  $\text{Hg}^+$  is often a potent GH1  $\beta$ -glucosidase inhibitor.

The activities of Os4BGlu16 and Os4BGlu18 were also inhibited by other metal salts, including copper, zinc and nickel, in the presence of which these enzymes retained 51% to 69% of their activities without inhibitor. In addition, 2,4-dinitrophenyl- $\beta$ -D-2-deoxy-2-fluoro-glucopyranoside inhibited the activity of Os4BGlu16, but not Os4BGlu18. The lack of sensitivity of Os4BGlu18 to this inhibitor was also seen for Os9BGlu31 GH1 transglucosidase (Luang et al., 2013). The activity of Os4BGlu18 was, however, affected by the covalent inhibitor conduritol B epoxide at 0.1 mM. These data suggested that Os4BGlu16 and Os4BGlu18 showed metal sensitivities comparable to other GH1  $\beta$ -glucosidases, but were relatively resistant to the mechanism-based inhibitor 2,4-dinitrophenyl- $\beta$ -D-2-deoxy-2-fluoro-glucopyranoside (Withers et al., 1990; Ketudat Cairns and Esen, 2010).

#### 4.7 mRNA expression of monolignol $\beta$ -glucosidase genes

The levels of the Os4BGlu16 and Os4BGlu18 gene expression are relatively high in the organs that undergo lignification. Fukushima and Terashima (1990) found that the differentiating xylem of magnolia, beech, lilac, and poplar contain H, G, and S units of lignin. H units were found in the initial stage of secondary cell wall and middle lamella of vessels and fibers. G units were also found in the initial and late stages of secondary cell wall and middle lamella. S units were mostly found at the late stage of secondary cell wall. In monocots (wheat straw, triticale straw, rye straw, and maize stalk), Barrière et al. (2007) demonstrated that the levels of H units are three to fifteen times higher than dicot plants, which when compared with monocots had S units at higher relative frequencies than G and H units, respectively. Moreover, She et al. (2011) reported that dewaxed rice straw after different alcohol treatments contained phenolic acids and aldehydes in the lignin fractions, and the ratios of G, S, and H were found to include large amounts of non-condensed guaiacyl, syringyl, and *p*-hydroxyphenyl units, suggesting the lignin preparations from rice straw can be considered as GSH lignin.

The preference of Os4BGlu16 and Os4BGlu18 for coniferin and syringin rather than *p*-coumarol glucoside is in line with the lower levels of H units than G and S units in grass lignin. The Os4BGlu16 gene is highly expressed in maturing leaves, which are at the stage of secondary cell wall production, and Os4BGlu16 protein could hydrolyze coniferin with the same catalytic efficiency as syringin. Although Os4BGlu16 had a pH optimum of 6.5, it had roughly 50% maximal activity at pH 5, suggesting it could act in either neutral or moderately acidic compartments, such as the apoplast. The Os4BGlu18 gene was expressed in young vegetative tissues at an

early stage of secondary cell wall production. In these rapidly growing tissues, the middle lamella is being produced with more G than S units (Terashima et al., 1989; Fukushima and Terashima, 1990), which is consistent with Os4BGlu18  $\beta$ -glucosidase having higher catalytic efficiency for coniferin than syringin hydrolysis. Moreover, the Os4BGlu14 and Os4BGlu18 genes were expressed in the seed, similar to *Arabidopsis* BGLU45 which was suggested to be involved in the lignification of valve margin layers of siliques (Escamilla-Treviño et al., 2006). The pH optimum of Os4BGlu18 is in the acidic range, consistent with a role in the apoplast, as suggested for BGLU45. The specific expression of Os4BGlu14 in reproductive tissues suggests it may play an active role, although whether this involves  $\beta$ -glucosidase activity is unclear, given the substitution of its catalytic acid/base position with glutamine.

BGLU45 and BGLU46 have been shown to be expendable for lignin production in *Arabidopsis*, despite their activities on monolignol  $\beta$ -glucosidases, protoxylem and tracheal localization and the build-up of coniferin in knockout lines (Escamilla-Treviño et al., 2006; Chapelle et al., 2012). Only small changes in lignification were seen in certain BGLU45 knockout lines (Chapelle et al., 2012). The biochemical functions of Os4BGlu16 and Os4BGlu18 appear to be similar, but given the differences in lignin structure between grasses and *Arabidopsis*, the biological function of the rice enzymes requires further assessment. As noted earlier lignin in grasses is composed mainly of G and S units but contains more H units than dicot plants. The kinetics results of rice Os4BGlu16 and Os4BGlu18, however, revealed that these enzymes are able to hydrolyze coniferin and syringin more efficiently than *p*-coumarol glucoside. In comparison, substrate specificity of chick pea and *Arabidopsis* BGLU45 showed to hydrolyze coniferin rather than syringin and *p*-

coumarol glucoside, respectively. For each of these enzymes, the biological function will depend on to which of these substrates they have access in the plant.

#### **4.8 Level of monolignol compounds and its precursor in rice KDML105 plants**

In the initial stage of root and leaf sheath development (10-40 days) higher levels of sinapyl alcohol were detected than coniferyl alcohol and other compounds. Analysis of the levels of monolignol glucosides, coniferin, syringin, and *p*-coumarol glucoside, showed that all three compounds dramatically increase from 60-90 days of all tissues tested, then decreased until harvesting stage.

Generally, lignin is generated by radical coupling of coniferyl alcohol, sinapyl alcohol, and *p*-coumaryl alcohol in vascular plants (Martone et al., 2009). Although it was at one point thought that monolignol glucosides played a role in transport of monolignols to the cell wall, it has recently been shown that the monolignols can be directly transported across the plasma membrane. Miao and Liu (2010) noted that in poplar and *Arabidopsis*, tonoplast-derived vesicles transported coniferin, but not coniferyl alcohol. In contrast, plasma membrane-derived vesicles transported coniferyl alcohol but not coniferin suggesting that aglycone forms of monolignol unit could be transported into the cell wall across the plasma membrane in an ATP-dependant manner by an ABC-transporter and consequently polymerized by laccases and/or peroxidases (Wang et al., 2013). An ABC transporter in *Arabidopsis* (AtABCG29) that transports monolignols to the cell wall was localized in the plasma membrane and this protein could transport *p*-coumaryl alcohol and had minor activity with sinapyl alcohol, but not coniferyl alcohol (Alejandro et al., 2012). An *Abcg29* mutant showed

reduced root growth and decreases in the amounts of H, G, and S units compared to WT root, which was an unexpected result since AtABCG29 is highly specific for *p*-coumaryl alcohol so only H unit should be reduced.

Irrespective of their biological roles, the levels of monolignol glucosides must be regulated by their production by glycosyltransferases and break-down by  $\beta$ -glucosidases. The over-expression of *Arabidopsis* UGT72E2 gene resulted in a 10-fold increase in coniferin and a lower increase in syringin in roots, suggesting the involvement of this glycosyltransferase gene in monolignol glycosylation (Lanot et al., 2006). They noted that rosette leaves have more syringin than coniferin and suggested this reflects tissue-specific variations in glycosylation mechanisms. If monolignol glucosides release monolignols for building up of lignin polymers, then monolignol  $\beta$ -glucosidase should be co-localized with coniferin and syringin (Wang et al., 2013). Further study on the localization of monolignol  $\beta$ -glucosidase is important to gain further insights on the control of lignin biosynthesis.

## CHAPTER V

### CONCLUSION

Based on their amino acid sequences, Os4BGlu14, Os4BGlu16, and Os4BGlu18 cluster with characterized monolignol  $\beta$ -glucosidases, including *Arabidopsis* BGLU45 and BGLU46 and *Pinus contorta* coniferin  $\beta$ -glucosidase. The cDNA sequence of Os4BGlu14 contains an open reading frame of 1551 nucleotides encoding 516 amino acids, which is predicted to include an N-terminal secretory signal peptide of 23 residues and a mature protein of 493 residues. However, the catalytic acid/base glutamate residue of Os4BGlu14 is replaced by glutamine, which led to the prediction that it may be an inactive  $\beta$ -glucosidase. The cDNA sequence of Os4BGlu16 also contained an open reading frame of 1551 nucleotides encoding 516 amino acids, including a predicted N-terminal secretory signal peptide of 27 residues and mature protein of 489 residues. Similarly, the cDNA sequence of Os4BGlu18 contained an open reading frame of 1518 nucleotides encoding 505 amino acids, which were predicted to include an N-terminal secretory signal peptide of 26 residues and a mature protein of 479 residues.

The plasmid containing the AK067841 cDNA provided by the Rice Genome Resource full-length cDNA project was used to amplify the mature gene of Os4BGlu14 to get PCR product size around 1.5 kb. The purified gene encoding mature Os4BGlu14 was cloned into a pET32a expression vector to express an Os4BGlu14 fusion protein with N-terminal thioredoxin and His<sub>6</sub> tags. The

recombinant pET32a/Os4BGlu14 plasmid was used as a template for mutation of pET32a/Os4BGlu14Q191E by site-directed mutagenesis to test if regenerating the catalytic acid/base would rescue the activity. Nevertheless, Os4BGlu14 protein could not be expressed in *E. coli* as an active  $\beta$ -glucosidase, suggesting that Os4BGlu14 may indeed be an inactive  $\beta$ -glucosidase as described by Opassiri et al., 2006.

The gene optimized for Os4BGlu16 expression in *P. pastoris* was inserted in the recombinant pPICZ $\alpha$ B(NH<sub>8</sub>)/Os4BGlu16 plasmid, which was used to express Os4BGlu16 as a secreted protein from *P. pastoris* strain SMD1168 at 20 °C for 4 days of inducing with 1% methanol. Os4BGlu16 enzyme was successfully purified to produce an apparently homogeneous protein with only one IMAC step from the *Pichia* media.

The cDNA encoding the full-length Os4BGlu18 protein was amplified from 7 days rice shoot cDNA and a cDNA encoding the predicted mature Os4BGlu18 gene was amplified from that initial amplicon. The gene encoding mature Os4BGlu18 was cloned into the pET32a expression vector and the thioredoxin-Os4BGlu18 fusion protein expressed in *E. coli* Origami(DE3) at 18 °C for 16-18 h. The catalytically active Os4BGlu18 enzyme was purified to get a single prominent protein with ion-exchange chromatography, hydrophobic interaction chromatography and IMAC, respectively.

The optimum pH of Os4BGlu16 was found to be 6.5 and after deglycosylation the optimum pH for Os4BGlu16 was 6.0, but its activity was stable after incubation from pH 5.5 to 11.5 over a period of 15 min up to 60 min at 25 °C, indicating that this enzyme was stable at neutral to basic pH, but it lost its activity in the highly acidic range. The temperature optimum for both the glycosylated and deglycosylated forms

of the Os4BGlu16 enzyme is 45 °C, but it was thermostable only over the range of 20-40 °C from 15-60 min. Meanwhile, the optimum pH for Os4BGlu18 was found to be 5.0. The activity of Os4BGlu18 maintained similar activity when incubated 15-60 min at pH values from pH 4.0 to 8.0 at 25 °C. The temperature optimum for Os4BGlu18 enzyme is 55 °C, but it was only thermostable over the range of 20-40 °C from 15-60 min.

Among *p*NP glycosides, Os4BGlu16 and Os4BGlu18 hydrolyzed *p*NP- $\beta$ -D-fucopyranoside, *p*NP- $\beta$ -D-glucopyranoside (*p*NPGLc), *o*NP- $\beta$ -D-glucopyranoside, *p*NP- $\beta$ -D-galactopyranoside and *p*NP- $\beta$ -D-xylopyranoside, respectively, in overall order of activity. However, Os4BGlu16 hydrolyzed *o*NP- $\beta$ -D-glucopyranoside with higher activity than *p*NP- $\beta$ -D-galactopyranoside, while Os4BGlu18 displayed low activity toward *o*NP- $\beta$ -D-glucopyranoside. For hydrolysis of natural and synthetic glycosides and oligosaccharides substrates, Os4BGlu16 and Os4BGlu18 could hydrolyze the monolignol glucosides *p*-coumarol  $\beta$ -D-glucoside, coniferin and syringin, along with daidzin, esculin, helicin, salicin, indoxyl  $\beta$ -D-glucoside, 4-methylumbelliferyl  $\beta$ -D-glucoside and 4-methylumbelliferyl  $\beta$ -D-fucoside. In addition, Os4BGlu18 hydrolyzed arbutin, methyl  $\beta$ -D-glucoside, *n*-octyl  $\beta$ -D-glucoside and *n*-heptyl  $\beta$ -D-glucoside. Neither enzyme could hydrolyze  $\beta$ -1,3-,  $\beta$ -1,4-, or  $\beta$ -1,6-linked gluco-oligosaccharides, which suggests neither enzyme plays a role in cell wall degradation.

Os4BGlu16 and Os4BGlu18 have more efficiency to hydrolyze monolignol glucoside substrates than other natural and synthetic substrates. The kinetic parameters for hydrolysis of monolignol glucoside substrates, Os4BGlu16 has high catalytic efficiency toward syringin ( $k_{cat}/K_M$  of 22.8 mM<sup>-1</sup>s<sup>-1</sup>), followed by coniferin

( $k_{cat}/K_M$  21.6  $\text{mM}^{-1}\text{s}^{-1}$ ) and *p*-coumarol glucoside ( $k_{cat}/K_M$  6.2  $\text{mM}^{-1}\text{s}^{-1}$ ). In comparison, Os4BGlu18 hydrolyzed coniferin best ( $k_{cat}/K_M$  31.9  $\text{mM}^{-1}\text{s}^{-1}$ ), followed by syringin ( $k_{cat}/K_M$  24.0  $\text{mM}^{-1}\text{s}^{-1}$ ) and *p*-coumarol glucoside ( $k_{cat}/K_M$  1.41  $\text{mM}^{-1}\text{s}^{-1}$ ), respectively. These high  $k_{cat}/K_M$  values were driven by the high  $k_{cat}$  values, with rather high  $K_M$  for these substrates, and both enzymes have lower  $K_M$  values for syringin than for the other two monolignol glucosides. This study supports the hypothesis that Os4BGlu16 and Os4BGlu18 are monolignol  $\beta$ -glucosidase enzyme.

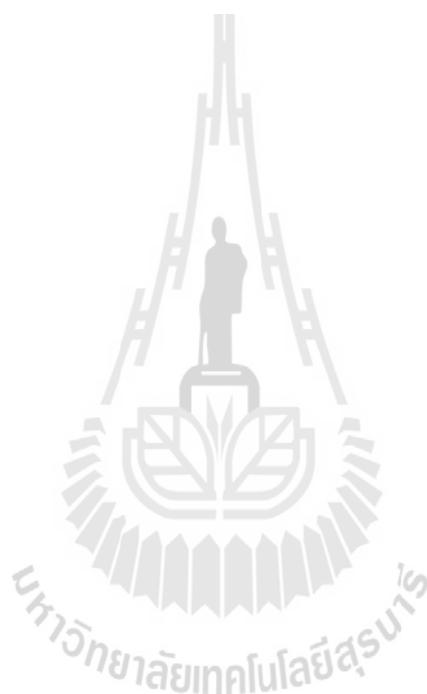
The inhibition study showed that Os4BGlu16 activity was inhibited nearly completely by mercuric ion and glucono  $\delta$ -lactone at 1 mM, while copper, zinc, and nickel, had partial effects. 2,4-Dinitrophenyl- $\beta$ -D-2-deoxy-2-fluoro-glucopyranoside, which is a strong inhibitor for other  $\beta$ -glucosidases could inhibit the activity of Os4BGlu16. Os4BGlu18 was inhibited by 1 mM glucono  $\delta$ -lactone, and moderately inhibited by copper and zinc at 1 mM inhibitor and conduritol B epoxide at 0.1 mM. Both enzymes retained around 45% activity with 500 mM glucose present in the reaction, so little product inhibition by the glucose is expected to occur in the plant.

Analysis of the gene expression level of the putative monolignol  $\beta$ -glucosidases in various rice organs showed that the *Os4BGlu14* gene is highly expressed in reproductive tissues, especially in endosperm, lemma, embryo, pollen, panicle and flower. *Os4BGlu16* is highly expressed in leaf from 4 weeks to 10 weeks, endosperm and lemma. *Os4BGlu18* is highly expressed in 1-week-old seedling, stem, leaf and leaf sheath at 4 weeks, pollen and lemma, and moderately expressed in other young vegetative tissues. This result suggested that Os4BGlu14 enzyme may primarily function in the reproductive tissue, while Os4BGlu16 and Os4BGlu18 apparently function in both vegetative and reproductive tissues.

The monolignol glucosides and related compounds in rice tissues were analyzed by UPLC-MS. Sinapyl alcohol levels were relatively high in root and leaf sheath from 10-40 days. In 2-3 month-old roots, leaf, leaf sheath, and stem, the levels of *p*-coumarol glucoside, coniferin, and syringin were dramatically increased. Their levels increased at 60 days until 90 days, then decreased at 100 days. In the rice flower extract, *p*-coumarol glucoside was highest, followed by syringin, sinapic acid, coniferin, caffeic acid, sinapyl alcohol, coniferyl alcohol, and *p*-coumaric acid, respectively. Sinapic acid was higher than other compounds in seed at all ages. The level of *p*-coumaric acid in every stage and tissue was lower than the other compounds and after 90 days it was undetectable.

In summary, several lines of evidence support the idea that Os4Bglu14, Os4Bglu16, and Os4Bglu18 are monolignol  $\beta$ -glucosidase that are involved in the lignification system in rice plants. First, successful recombinant expression and purification allowed us to see that Os4Bglu16 and Os4Bglu18 enzymes have higher efficiency to hydrolyze monolignol glucoside substrates than other substrates. Second, the pH optimum of Os4Bglu16 and Os4Bglu18 is in the acidic range, consistent with a role in the apoplast or acidic vacuole. The synthesis of monolignols from phenylalanine and shikimate involves in cytosolic and ER membrane-anchored cytosolic enzymes. Monolignols and lignans may then be conjugated by UGTs and then transported to the vacuole or the monolignols may be directly transported to the cell wall for oxidative cross-linking by apoplastic peroxidases and laccases into lignins (Wang et al., 2013), so a role in further processing of monoglucosides in the vacuoles or upon release to the apoplast is possible for Os4Bglu16 and Os4Bglu18. Last, the Os4Bglu14, Os4Bglu16, and Os4Bglu18 expression was found to occur in

the tissues that undergo lignification from vegetative to reproductive stage, and were in line with and the build-up of monolignol compounds in this study.



## REFERENCES



## REFERENCES

- Adler, E. (1977). Lignin chemistry-past present and future. **Wood Sci. Technol.** 11: 169-218.
- Ahn, Y.O., Mizutani, M., Saino, H., and Sakata, K. (2004). Furcatin hydrolase from *Viburnum furcatum* blume is a novel disaccharide-specific acuminosidase in glycosyl hydrolase family 1. **J. Biol. Chem.** 279: 23405-23414.
- Akiyama, T., Kaku, H., and Shibuya, N. (1998). A cell wall-bound  $\beta$ -glucosidase from germinated rice: purification and properties. **Phytochemistry.** 48: 49-54.
- Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., Bovet, L., Lee, Y., Geldner, N., Fernie, A.R. and Martinoia, E. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. **Curr.Biol.** 22: 1207-1212.
- Ayres, D.C, and Loike, J.D. (1990). In Lignans: Chemical, Biological and Clinical Properties. **Cambridge University Press.** 402 pp.
- Babcock, G.D, and Esen, A. (1994). Substrate specificity of maize  $\beta$ -glucosidase. **Plant Sci.** 101: 31-39.
- Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., and Stöckigt, J. (2007). Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpenoid indole alkaloid family. **Plant Cell.** 19: 2886-2897.
- Barrière, Y., Riboulet, C., Méchin, V., Stéphane, M., Maltese, S., Pichon, M., Cardinal, A., Lapierre, C., Lübberstedt, T., and Martinant, J. (2007). Genetics

- and genomics of lignification in grass cell walls based on maize as model species. **G3**. 1(2): 133-156.
- Baucher, M., Monties, B., Van Montagu, M., and Boerjan, W. (1998). Biosynthesis and genetic engineering of lignin. **Crit. Rev. Plant Sci.** 17: 125-97.
- Bendtsen, J.D., Nielsen, H., Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: Signal P3.0. **J Mol Biol.** 340: 783-795.
- Berrin, J.G., Czjzek, M., Kroon, P.A., McLauchlan, W.R., Puigserver, A., Williamson, G., and Juge, N. (2003). Substrate (aglycone) specificity of human cytosolic beta-glucosidase. **Biochem. J.** 373: 41-48.
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. **Annu. Rev. Plant Biol.** 54: 519-546.
- Boonclarm, D., Sornwatana, T., Arthan, D., Kongsaree, P., and Svasti, J. (2006).  $\beta$ -Glucosidase catalyzing specific hydrolysis of an iridoid  $\beta$ -glucoside from *Plumeria obtusa*. **Acta Biochim Biophys Sin.** 38: 563-570.
- Bouarab, K., Melton, R., Peart, J., Baulcombe, D., and Osbourn, A. (2002). A saponin-detoxifying enzyme mediates suppression of plant defenses. **Nature.** 418: 889-892.
- Brzobohatý, B., Moore, I., and Palme, K. (1994). Cytokinin metabolism: implications for regulation of plant growth and development. **Plant Mol. Biol.** 26: 1483-1497.
- Burmeister, W.P., Cottaz, S., Driguez, H., Palmieri, S., and Henrissat, B. (1997). The crystal structures of *Sinapis alba* myrosinase and of a covalent glycosyl enzyme intermediate provide insights into the substrate recognition and active site machinery of an *S*-glycosidase. **Structure.** 5: 663-675.

- Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A., and Henrissat, B. (2000). High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. **J. Biol. Chem.** 275(50): 39385-39393.
- Campbell, M.M., and Sederoff, R.R. (1996). Variation in lignin content and composition. Mechanisms of control and implications for the genetic improvement of plants. **Plant Physiol.** 110: 3-13.
- Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Lapierre, C., Boerjan, W., Jouanin, L., and Demont-Caulet, N. (2012). Impact of absence of stem-specific  $\beta$ -glucosidase on lignin and monolignols. **Plant Physiol.** 160: 1204-1217.
- Chothia, C. (1992). One thousand families for the molecular biologist. **Nature.** 357: 543-544.
- Chuankhayan, P., Hua, Y., Svasti, J., Sakdarat, S., Sullivan, P.A., and Ketudat Cairns, J.R. (2005). Purification of an isoflavonoid 7-O- $\beta$ -apiosyl-glucoside  $\beta$ -glycosidase and its substrates from *Dalbergia nigrescens* Kurz. **Phytochemistry.** 16(66): 1880-1889.
- Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Oonant, W., Bevan, D.R., Esen, A., Chen, C.J., Opassiri, R., Svasti, J., and Ketudat Cairns, J.R. (2008). Structural insights into rice BGlu1  $\beta$ -glucosidase oligosaccharide hydrolysis and transglycosylation. **J. Mol. Biol.** 377: 1200-1215.
- Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Svasti, J., and Ketudat Cairns, J.R. (2011). The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase. **J. Struct. Biol.** 173: 169-179.

- Coughlan, M.P., and Ljungdahl, L.G. (1988). Comparative biochemistry of fungal and bacterial cellulolytic enzyme systems. I. In: Aubert JP, Beguin P, Millet J (eds), **Biochemistry and Genetics of Cellulose Degradation. Academic Press, London.** 11-30 pp.
- Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W.P., Bevan, D.R., Henrissat, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in  $\beta$ -glucosidases is revealed by crystal structures of mutant maize  $\beta$ -glucosidase-DIMBOA, DIMBOAGlc, and -dhurrin complexes. **Proc. Natl. Acad. Sci. USA.** 97: 13555-13560.
- Davies, G.J., Dodson, G.G., Hubbard, R.E., Tolley, S.P., Dauter, Z., Wilson, K.S., Hjort, C., Mikkelsen, J.M., Rasmussen, G., and Schulein, M. (1993). Structure and Function of Endoglucanase-V. **Nature.** 365: 362-364.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. **Structure.** 3: 853-859.
- Davin, L.B., Bedgar, D.L., Katayama, T., and Lewis N.G. (1992). On the stereoselective synthesis of (+)-pinoresinol in *Forsythia suspense* from its achiral precursor, coniferyl alcohol. **Phytochemistry.** 31: 3869-3874.
- Davin, L.B., Wang, H.B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., and Lewis, N.G. (1997). Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. **Science.** 275: 362-366.
- Davin, L.B., and Lewis, N.G. (2000). Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignin and lignin biosynthesis. **Plant Physiol.** 123: 453-461.

- Dharmawardhana, D.P., Ellis, B.E., and Carlson, J.E. (1995). A  $\beta$ -glucosidase from lodgepole pine specific for the lignin precursor coniferin. **Plant Physiol.** 107: 331-339.
- Dharmawardhana, D.P., and Ellis, B.E. (1998).  $\beta$ -Glucosidases and glucosyltransferase in lignifying tissues. **J. Am. Chem. Soc.** 697: 76-83.
- Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J.K.C., Teeri, T.T., and Jones, A. (1994). The three-dimensional structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. **Science.** 265: 524-528.
- Donaldson, L.A. (2001). Lignification and lignin topochemistry-an ultrastructural view. **Phytochemistry.** 57: 859-873.
- Duroux, L., Delmotte, F.M., Lancelin, J.-M., Keravis, G., and Jay-Alleand, C. (1998). Insight into naphthoquinone metabolism:  $\beta$ -glucosidase catalysed hydrolysis of hydrojuglone  $\beta$ -D-glucopyranoside, **Biochem. J.** 333: 275-283.
- Escamilla-Treviño, L.L., Chen, W., Card, M.L., Shih, M-C., Cheng, C.L., and Poulton, J.E. (2006). *Arabidopsis thailiana*  $\beta$ -glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. **Phytochemistry.** 67: 1651-1660.
- Ferreira, C., Parra, J.R., and Terra, W.R. (1997). The effect of dietary plant glycosides on larval midgut beta-glucosidases from *Spodoptera frugiperda* and *Diatraea saccharalis*. **Insect Biochem. Mol. Biol.** 27: 55-59.
- Freudenberg, K., and Neish, A.C. (1968). **Constitution and Biosynthesis of Lignin.** Berlin: Springer-Verlag. 129 pp.

- Fukushima, K., and Terashima, N. (1990). Heterogeneity in formation of lignin. XIII. Formation of *p*-hydroxyphenyl lignin in various hardwood visualized by microautoradiography. **J. Wood Chem. Technol.** 10: 413-433.
- Ganguly, T., Ganguly, S., Sircar, P., and Sircar, S. (1974). Rhamnose bound indole-3-acetic acid in the floral parts of *Peltophorum Ferrugineum*. **Physiol. Plant.** 31: 330-334.
- Gaskin, P., and MacMillian, J. (1975). Polyoxygenated ent-kauranes and water soluble conjugates in seed of *Phaseolus cocineus*. **Phytochemistry.** 14: 1575-1578.
- Guo, H., Liu, A.H., Ye, M., Yang, M., and Guo, D.A. (2007). Characterization of phenolic compounds in the fruits of *Forsythia suspense* by high-performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry. **Rapid Commun. Mass Spectrom.** 21: 715-729.
- Haldane, J.B.S. (1957). Graphical methods in enzyme chemistry. **Nature.** 179(832).
- Hayashi, Y., Okino, N., Kakuta, Y., Shikanai, T., Tani, M., Narimatsu, H., and Ito, M. (2007). Klotho-related protein is a novel cytosolic neutral beta-glycosylceramidase, **J. Biol. Chem.** 282: 30889-30900.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 280: 309-316.
- Henrissat, B., and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 293: 781-788.
- Henrissat, B., Callebaut, I., Mornon, J.P., Fabrega, S., Lehn, P., and Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold

- for many families of glycosyl hydrolases. **Proc. Natl. Acad. Sci. U.S.A.** 92(15): 7090-7094.
- Himeno, N., Saburi, W., Wakuta, S., Takeda, R., Matsuura, H., Nabeta, K., Sansenya, S., Ketadat Cairns, J.R., Mori, H., Imai, R., and Matsui, H. (2013). Identification of rice  $\beta$ -glucosidase with high hydrolytic activity towards salicylic acid  $\beta$ -D-glucoside. **Biosci. Biotechnol. Biochem.** 77(5): 934-939.
- Hoffman, K., Bucher, P., Falquet, L., and Bairoch, A. (1999). The PROSITE database. **Nucl. Acids Res.** 27: 215-219.
- Holm, L., and Sander, C. (1994). Structural similarity of plant chitinase and lysozymes from animals and phage: An evolutionary connection. **FEBS Lett.** 340: 129-132.
- Hösel, W., and Barz, W. (1975).  $\beta$ -Glucosidases from *Cicer arietinum* L. purification and properties of isoflavone-7-O-glucosides-specific  $\beta$ -glucosidases. **Eur. J. Biochem.** 57: 607-616.
- Hösel, W., Surholt, E., and Borgmann, E. (1978). Characterization of  $\beta$ -glucosidase isoenzymes possibly involved in lignifications from chick pea (*Cicer arietinum* L.) cell suspension culture. **Eur. J. Biochem.** 84: 487-492.
- Hösel, W., and Todenhagen, R. (1980). Characterization of a  $\beta$ -glucosidase from *Glycine max* which hydrolyses coniferin and syringin. **Phytochemistry.** 19: 1349-1353.
- Hösel, W., Fiedler-Preiss, A., and Borgmann, E. (1982). Relationship of coniferin  $\beta$ -glucosidase to lignification in various plant cell suspension cultures. **Plant Cell Tissue Organ Culture.** 1: 137-148.

- Hrmova, M., MacGregor, E.A., Biely, P., Stewart, R.J., and Fincher, G.B. (1998). Substrate binding and catalytic mechanism of a barley  $\beta$ -D-glucosidase/(1,4)- $\beta$ -D-glucan exohydrolase. **J. Biol. Chem.** 273: 11134-11143.
- Hua, Y., Sansenya, S., Saetand, C., Wakuta, S., and Ketudat Cairns, J.R. (2013). Enzymatic and structural characterization of hydrolysis of gibberellins A4 glucosyl ester by a rice  $\beta$ -D-glucosidase. **Arch. Biochem. Biophys.** 537: 39-48.
- Jenkins, J., Lo Leggio, L., Harris, G., and Pickersgill, R. (1995).  $\beta$ -glucosidase,  $\beta$ -galactosidase, family A cellulases, family F xylanases and two barley glucanases from a superfamily of enzymes with 8-fold  $\beta/\alpha$  architecture and with two conserved glutamates near the carboxy-terminal ends of  $\beta$ -strands four and seven. **FEBS Lett.** 362: 281-285.
- Ketudat Cairns, J.R., and Esen, A. (2010).  $\beta$ -Glucosidases. **Cell. Mol. Life Sci.** 67: 3389-3405.
- Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.J., Ohtsuki, K., Shishiki, T., Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato,

- N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., Hayashizaki, Y., and Yasunishi, A. (2003). Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. **Science**. 301: 376-379.
- Körschen, H.G., Yildiz, Y., Raju, D.N., Schonauer, S., Bönigk, W., Jansen, V., Kremmer, E., Kaupp, U.B., and Wachten, D. (2013). The non-lysosomal  $\beta$ -glucosidase GBA2 is a non-integral membrane-associated protein at the endoplasmic reticulum (ER) and Golgi. *J. Biol. Chem.* 288: 3381-3393.
- Koprivica, V., Stone, D.L., and Park, J.K. (2000). Analysis and classification of 304 mutants alleles in patients with type I and type III Gaucher disease. **Am. J. Med. Gen.** 66: 1777-1786.
- Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. **Biol. Rev. Camb. Philos. Soc.** 28: 416-436.
- Kuntothom, T., Luang, S., Harvey, A.J., Fincher, G.B., Opassiri, R., Hrmova, M., and Ketudat Cairns, J.R. (2009). Rice family GH1 glycoside hydrolases with  $\beta$ -D-glucosidase and  $\beta$ -D-mannosidase activities. **Arch. Biochem. Biophys.** 491: 85-95.
- Lacombe, E., Hawkins, S., Van Doorselaere, J., Piguemal, J., Goffner, D., Poeydomenge, O., Boudet, A.M., and Grima-Pettenati, J. (1997). Cinnamoyl-CoA reductase, the first committed enzyme of lignin branch biosynthetic pathway: cloning, expression and phylogenic relationships. **Plant J.** 11: 429-441.
- Lanot, A., Hodge, D., Jackson, R.G., George, G.L., Elias, L., Lim, E.K., Vaistij, F.E., and Bowles, D.J. (2006). The glucosyltransferase UGT72E2 is responsible for

- monolignol 4-*O*-glucoside production in *Arabidopsis thaliana*. **Plant J.** 48: 286-295.
- Laemmli, U.K. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T4. **Nature.** 227: 680-685.
- Leinhos, V., Udagama-Randeniya, P.V., and Savidge, R.A. (1994). Purification of an acid coniferin-hydrolysing  $\beta$ -glucosidase from developing xylem of *Pinus banksiana*. **Phytochemistry.** 37: 311-315.
- Lewis, N.G., and Davin, L.B. (1994). Evolution of lignan and neolignan biochemical pathways, in **Evolution of Natural Products** (Nes. W.D. ed.), American Chemical Society Washington, D.C. 202-246 pp.
- Luang, S., Hrmova, M., and Ketudat Cairns, J.R. (2010). High-level expression of barley  $\beta$ -D-glucan exohydrolase HvExoI from a codon-optimized cDNA in *Pichia pastoris*. **Prot. Express. Purif.** 73: 90-98.
- Martone, P.T., Estevez, J.M., Lu, F., Ruel, K., Denny, M.W., Somerville, C., and Ralph, J. (2009). Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. **Curr. Biol.** 19: 169-175.
- Matsuba, Y., Sasaki, N., Tera, M., Okamura, M., Abe, Y., Okamoto, E., Nakamura, H., Funabashi, H., Takatsu, M., Saito, M., Matsuoka, H., Nagasawa, K., and Ozeki, Y. (2010). A novel glucosylation reaction on anthocyanins catalyzed by acyl-glucose-dependent glucosyltransferase in the petals of carnation and delphinium. **Plant Cell.** 22:3374-3389.
- Miao, Y.C., and Liu, C.J. (2010). ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. **Proc. Natl. Acad. Sci. U.S.A.** 107: 22728-22733.

- Millborrow, B. (1970). The metabolism of abscisic acid. **J. Exp. Bot.** 21: 17-31.
- Mizutani, F.M., Nakanishi, H., Ema, J., Ma, S.J., Noguchi, N., Ochiai, M.I., Mizutani, M.F., Nakao, M., and Sakata, K. (2002). Cloning of  $\beta$ -primeverosidase from tea leaves, a key enzyme in tea aroma. **Plant Physiol.** 130: 2164-2176.
- Moellering, E.R., Muthan, B., and Benning, C. (2010). Freezing tolerance in plants requires lipid remodeling at the outer chloroplast membrane. *Science*. 330: 226-228.
- Müllegger, J., Jahn, J., Chen, H-M., Warren, R.A.J., and Withers, S.G. (2005). Engineering of a thioglycoligase: randomized mutagenesis of the acid-base residue leads to the identification of improved catalysts. **Protein Eng. Des. Sel.** 18: 33-40.
- Nisius, A. (1988). The stromacenter in *Avena* plastids: an aggregation of  $\beta$ -glucosidase responsible for the activation of oat-leaf saponins. **Planta.** 173: 474-481.
- Noguchi, A., Fukui, Y., Iuchi-Okada, A., Kakutani, S., Satake, H., Iwashita, T., Nakao, M., Umezawa, T., and Ono, E. (2008). Sequential glucosylation of a furofuran lignan, (+)-sesaminol, by *Sesamum indicum* UGT71A9 and UGT94D1 glucosyltransferases. **Plant J.** 54: 415-427.
- Noguchi, J., Hayashi, Y., Baba, Y., Okino, N., Kimura, M., Ito, M., and Kakuta, Y. (2008). Crystal structure of the covalent intermediate of human cytosolic  $\beta$ -glucosidase. **Biochem. Biophys. Res. Comm.** 374: 549-552.
- Okazawa, A., Kusunose, T., Ono, E., Kim, H.J., Satake, H., Shimizu, B., Mizutani, M., Seki, H., and Muranaka, T. (2014). Glucosyltransferase activity of

- Arabidopsis* UGT71C1 towards pinoresinol and lariciresinol. **Plant Biotech J.** shot communication 1-6.
- Ono, E., Kim, H.J., Murata, J., Morimoto, K., Okazawa, A., Kobayashi, A., Umezawa, T., and Satake, H. (2010). Molecular and functional characterization of novel furofuran-class lignan glucosyltransferases from *Forsythia*. **Plant Biotech J.** 27: 317-324.
- Opassiri, R., Ketudat Cairns, J.R., Akiyama, T., Wara-Aswapati, O., Svasti, J., and Esen, A. (2003). Characterization of a rice  $\beta$ -glucosidase genes highly expressed in flower and germinating shoot. **Plant Sci.** 165: 627-638.
- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., and Ketudat Cairns, J.R. (2004).  $\beta$ -Glucosidase, exo- $\beta$ -glucanase and pyridoxine transglucosylase activities of rice BGlu1. **Biochem. J.** 379: 125-131.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J.R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12  $\beta$ -glucosidase. **BMC Plant Biol.** 6: 33-51.
- Opassiri, R., Maneesan, J., Akiyama, T., Pomthong, B., Jin, S., Kimura, A., and Ketudat Cairns, J.R. (2010). Rice Os4BGlu12 is a wound-induced  $\beta$ -glucosidase that hydrolyzes cell wall- $\beta$ -glucan-derived oligosaccharides and glycosides. **Plant Sci.** 179: 273-280.
- Orengo, C.A., Jones, D.T., and Thornton, J.M. (1994). Protein domain superfolds and superfamilies. **Nature.** 372: 631-634.
- Osawa, T. (1992). Phenolic antioxidants in dietary plants as antimutagens. **J. Am. Chem. Soc.** 507: 135-149.

- Osbourn, A., Bowyer, P., Lunness, P., Clarke, B., and Daniels, M. (1995). Fungal pathogens of oat roots and tomato leaves employ closely related enzymes to detoxify different host plant saponins. **Mol. Plant Microbe Interact.** 8: 971-978.
- Poulton, J.E. (1990). Cyanogenesis in plants. **Plant Physiol.** 94: 401-405.
- Raychaudhuri, A., and Tipton, P.A. (2002). Cloning and expression of the gene for soybean hydroxyisourate hydrolase. Localization and implications for function and mechanism. **Plant Physiol.** 130: 2061-2068.
- Rouyi, C., Baiya, S., Lee, S., Mahong, B., Jeon, J., Ketudat Cairns, J.R., and Ketudat Cairns, M. (2014). Recombinant expression and characterization of the cytoplasmic rice  $\beta$ -glucosidase Os1BGlu4. **PLoS One.** 9: 1-13.
- Saino, H., Shimuzu, T., Hiratake, J., Nakatsu, T., Kato, H., Sakata, K., and Mizutani, M. (2014). Crystal structures of  $\beta$ -primeverosidase in complex with disaccharide amidine inhibitors. **J. Biol. Chem.** 289: 16826-16834.
- Sarkanen, K.V., and Ludwig, C.H. (1971). Lignins: Occurrence, Formation, Structure, and Reactions. New York: **Wiley-Interscience.** 916 pp.
- Scheffe, J.H., Lehmann, K.E., Buschmann, I.R., Unger, T., and Funke-Kaiser, H. Quantitative real-time RT-PCR data analysis: current concepts and the novel gene expression's  $C_T$  difference formula. (2006). **J. Mol. Med.** 84: 901-910.
- Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., and Ketudat Cairns, J.R. (2009). Structural and enzymatic characterization of Os3BGlu6, a rice  $\beta$ -glucosidase hydrolyzing hydrophobic glycosides and (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 2)-linked disaccharides. **Plant Physiol.** 151: 47-58.

- She, D., Nie, X.N., Xu, F., Geng, Z.C., Jia, H.T., Jones, G.L., and Baird, M.S. (2011). Physico-chemical characterization of different alcohol-soluble lignins from rice straw. **Cellulose Chem. Technol.** 46(3-4): 207-2019.
- Sinnott, M.L. (1990). Catalytic mechanisms of enzymic glycosyltransfer. **Chem. Rev.** 90: 1171-1202.
- Srisomsap, C., Svasti, M.R.J., Surarit, R., Chammpattanachai, V., Boonpuan, K., Sawangaretrakul, P., Subhasitanont, P., and Chokchaichamnankit, D. (1996). Isolation and characterization of an enzyme with beta-glucosidase/beta-fucosidase activities from *Dalbergia cochinchinensis* Pierre. **J. Biochem.** 119: 585-590.
- Svasti, J., Srisomsap, C., Techasakul, S., and Surarit, R. (1999). Dalcochinin-8'-O- $\beta$ -D-glucoside and its  $\beta$ -glucosidase enzyme from *Dalbergia cochinchinensis*. **Phytochemistry.** 50: 739-743.
- Terashima, N. (1989). An improved radiotracer method for studying formation and structure of lignin. **J. Am. Chem. Soc.** 399: 148-159.
- Terra, W.R., and Ferreira, C. (2005). Biochemistry of digestion. In: **Comprehensive Molecular Insect Science, 1<sup>st</sup> ed, vol. 4.** Elsevier, Amsterdam. 171-224 pp.
- Toonkool, P., Metheenukul, P., Sujiwattanarat, P., Paiboon, P., Tongtubtim, N., Ketudat Cairns, M., Ketudat Cairns, J.R., and Svasti, J. (2006). Expression and purification of dalcochinase, a  $\beta$ -glucosidase from *Dalbergia cochinchinensis* Pierre, in yeast and bacterial hosts. **Prot. Express. Purif.** 48: 195-204.
- Tribolo, S., Berrin, J.G., Kroon, P.A., Czjzek, M., and Juge, N. (2007). The crystal structure of human cytosolic beta-glucosidase unravels the substrate aglycone specificity of a family 1 glycoside hydrolase. **J. Mol. Biol.** 370: 964-975.

- Verdoucq, L., Morinière, J., Bevan, D.R., Esen, A., Vasella, A., Henrissat, B., and Czjzek, M. (2004). Structural determinants of substrate specificity in family 1  $\beta$ -glucosidases: novel insights from the crystal structure of sorghum dhurrinase-1, a plant  $\beta$ -glucosidase with strict specificity, in complex with its natural substrate. **J. Biol.Chem.** 279: 31796-31803.
- Varghese, J.N., Garrett, T.P.J., Colman, P.M., Chen, L., Høj, P.B., and Fincher, G.B. (1994). Three-dimensional structures of two plant  $\beta$ -glucan endohydrolases with distinct substrate specificities. **Proc. Natl. Acad. Sci. USA.** 91: 2785-2789.
- Wakuta, S., Hamada, S., Ito, H., Matsuura, H., Nabeta, K., and Matsui, H. (2010). Identification of a  $\beta$ -glucosidase hydrolyzing tuberonic acid glucoside in rice (*Oryza sativa* L.). **Phytochemistry.** 71: 1280-1288.
- Wang, Q., Trimbur, D., Graham, R., Warren, R.A., and Withers, S.G. (1995). Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants. **Biochemistry.** 44: 14554-14562.
- Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013). Plant cell wall lignification and monolignol metabolism. **Front Plant Sci.** 4(220): 1-8.
- Withers, S.G., Warren, R.A.J., Street, I.P., Rupitz, K., Kempton, J.B., and Aebersold, R. (1990). Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a "retaining" glycosidase. **J. Am. Chem. Soc.** 112: 5887-5889.
- Yu, S.J. (1989).  $\beta$ -Glucosidase in four phytophagous Lepidoptera. **Insect Biochem.** 19: 103-110.



**APPENDICES**

# APPENDIX A

## DETECTION OF MONOLIGNOL COMPOUNDS

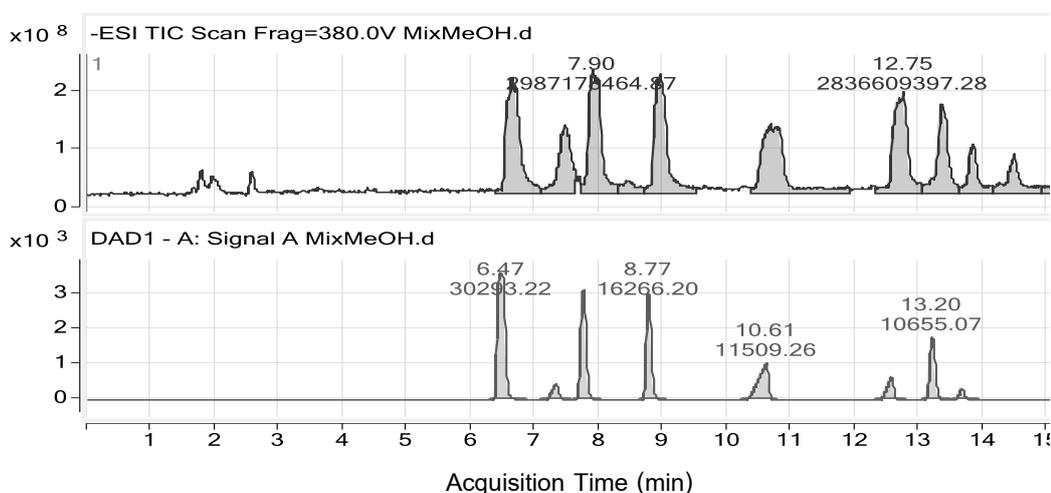
### BY UPLC-MS METHOD

#### 1. Methodology

An Agilent 1290 UPLC system inline with an Agilent 6490 triple quadrupole mass spectrometer with an Agilent SB-C18 RRHD 1.8  $\mu\text{m}$ , 2.1x150 mm column (Agilent Technologies, CA, USA) was used to detect monolignol compounds. Ionic masses of *p*-coumaric acid ( $m/z$  163, retention time [Rt] 10.8 min), caffeic acid ( $m/z$  179, Rt 7.6 min), coniferyl alcohol ( $m/z$  179, Rt 13.6 min), sinapyl alcohol ( $m/z$  209, Rt 14.1 min), sinapic acid ( $m/z$  223, Rt 12.9 min), *p*-coumarol glucoside [*p*CAG] ( $m/z$  371, Rt 6.8 min), coniferin ( $m/z$  401, Rt 8.2 min), and syringin ( $m/z$  431, Rt 9.2 min) were monitored.

#### 2. Standard chromatograms.

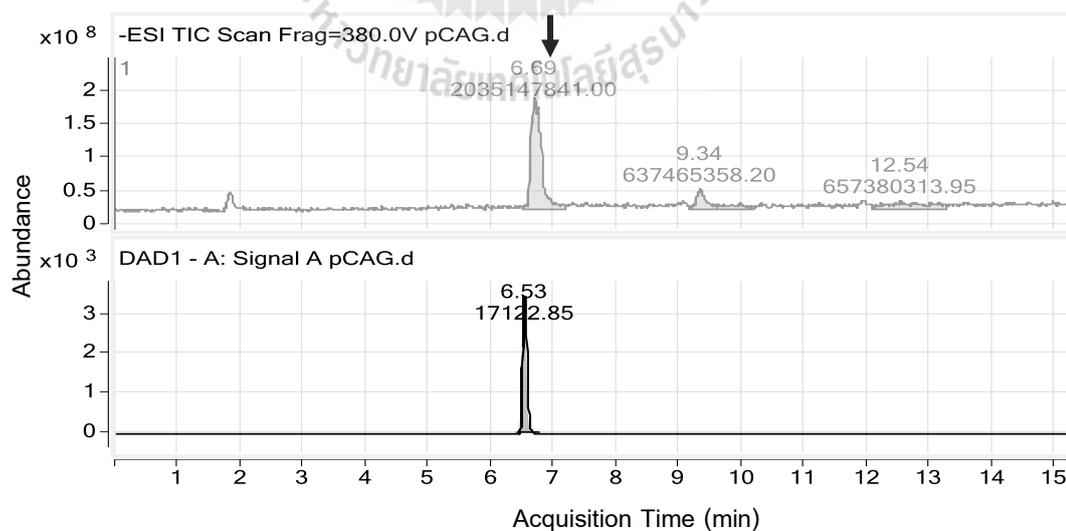
##### 2.1 Mixed standard of monolignol compounds



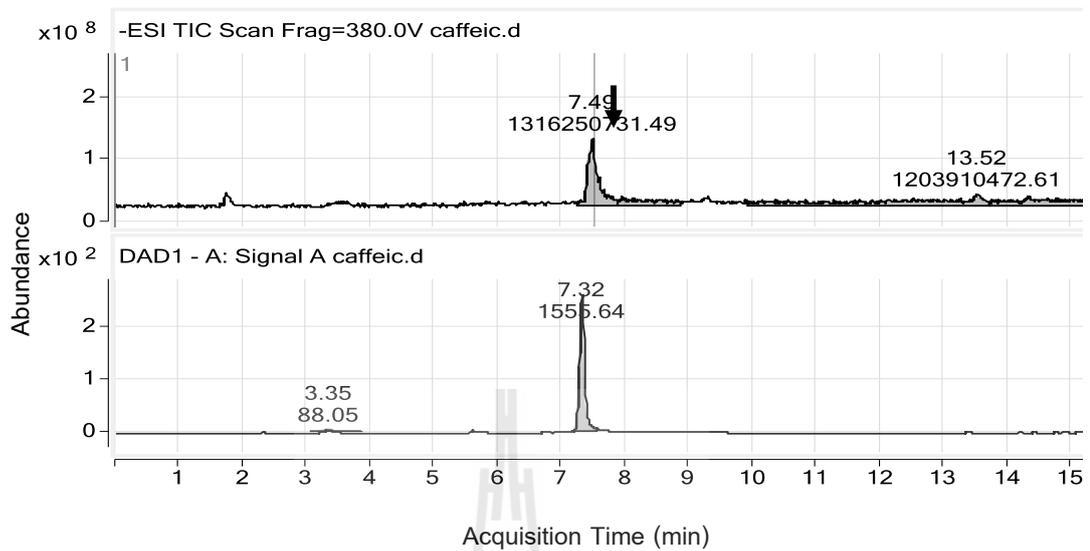
**Table AI 1 Mass and its retention time of each compounds.**

Monolignol compounds	m/z	Time (min)
<i>p</i> -coumarol glucoside	371.1	6.6
Caffeic acid	179.0	7.4
Coniferin	401.1	7.9
Syringin	431.1	8.9
<i>p</i> -coumaric acid	163.1	10.8
Sinapic acid	223.0	12.7
Coniferyl alcohol	179.1	13.4
Sinapyl alcohol	209.0	13.8

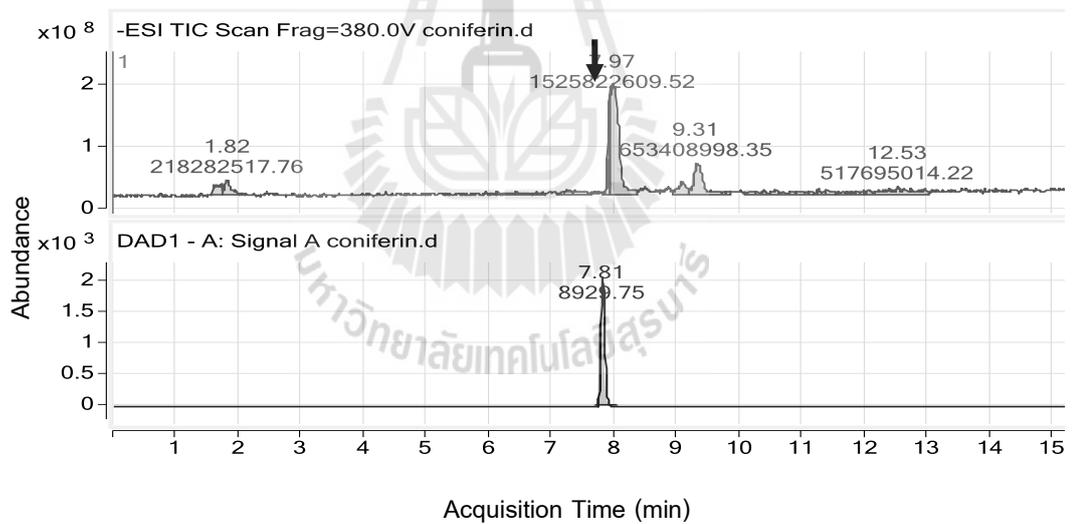
## 2.2 Single injection of *p*-coumarol glucoside standard



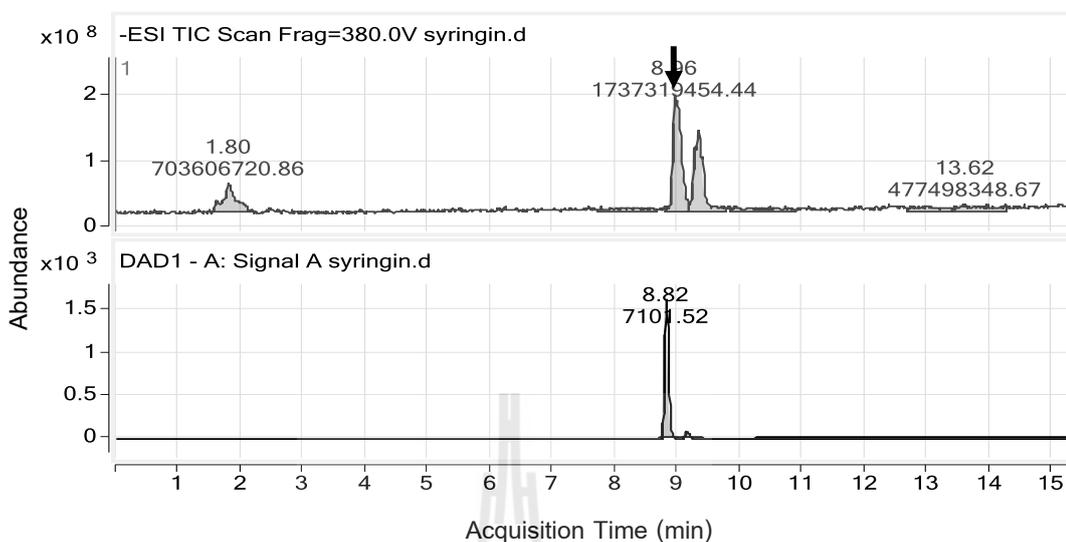
### 2.3 Single injection of caffeic acid standard



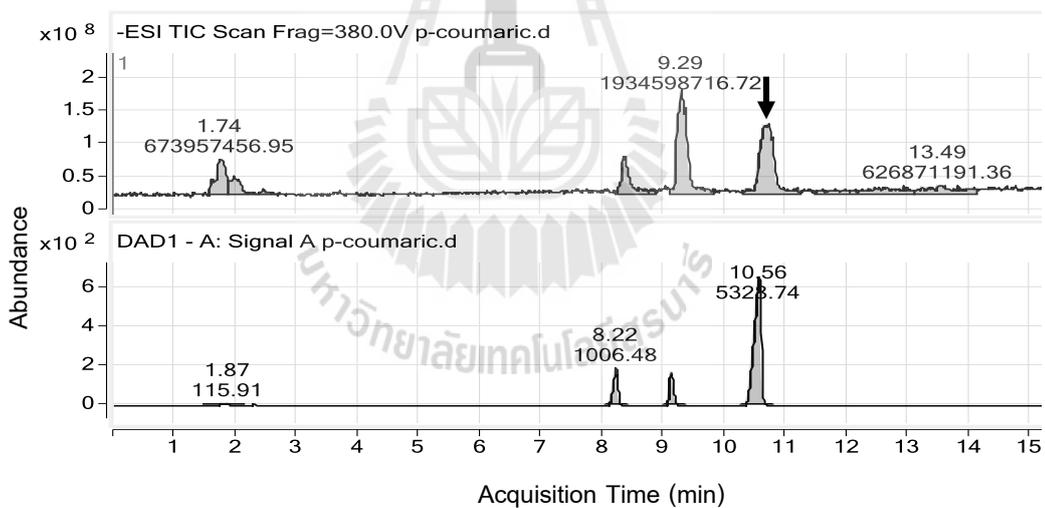
### 2.4 Single injection of coniferin standard



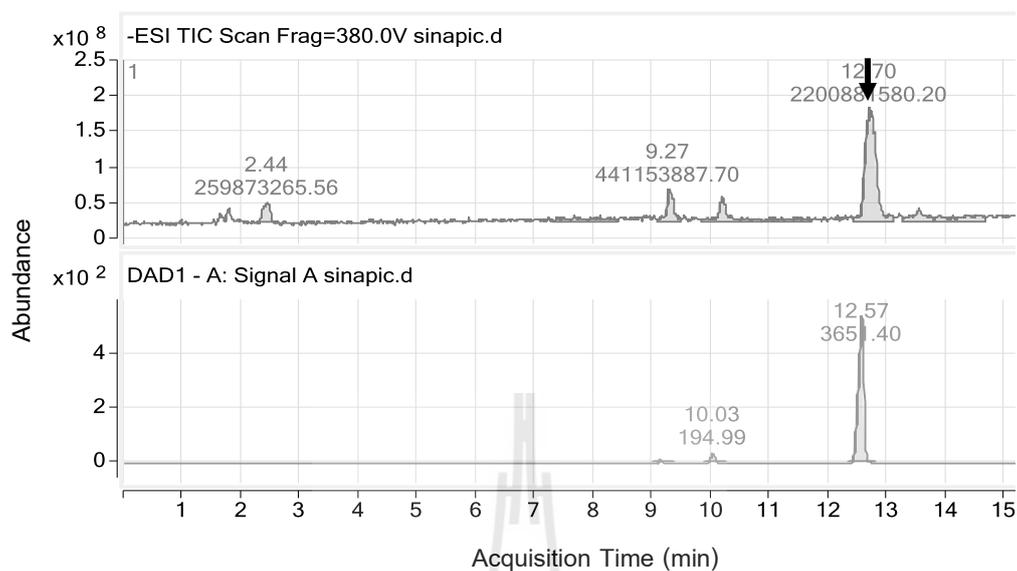
## 2.5 Single injection of syringin standard



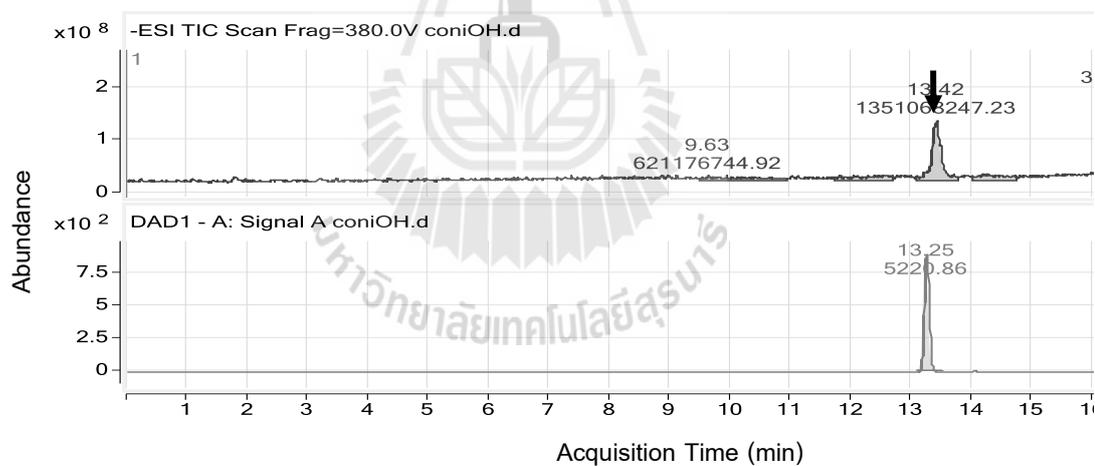
## 2.6 Single injection of *p*-coumaric acid standard



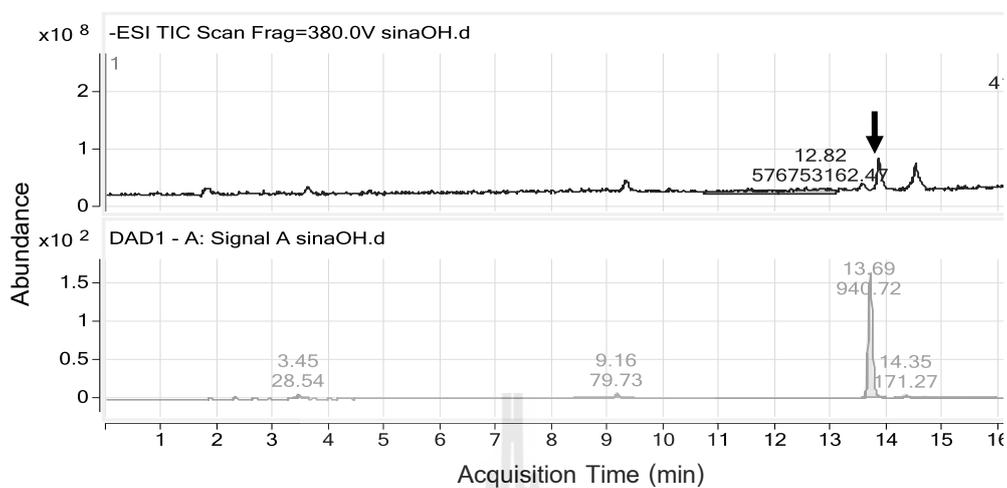
## 2.7 Single injection of sinapic acid standard



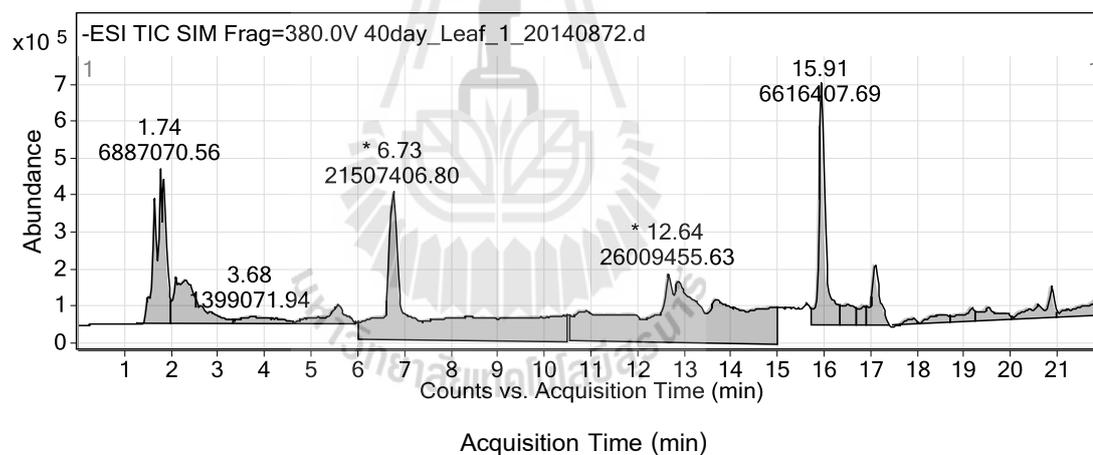
## 2.8 Single injection of coniferyl alcohol standard



## 2.9 Single injection of sinapyl alcohol standard

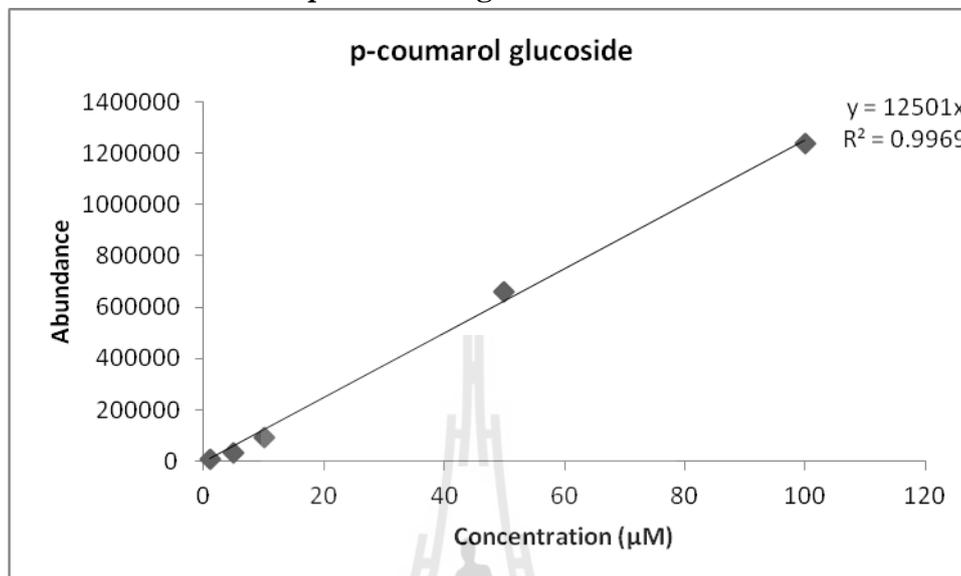


## 2.10 An example total ion chromatogram of 40 days rice leaf

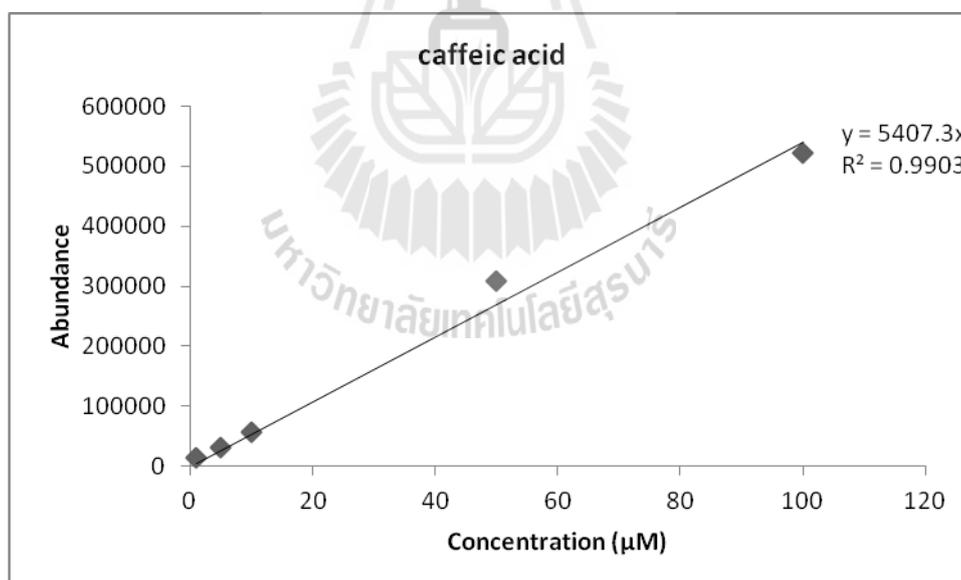


## II. Standard curves

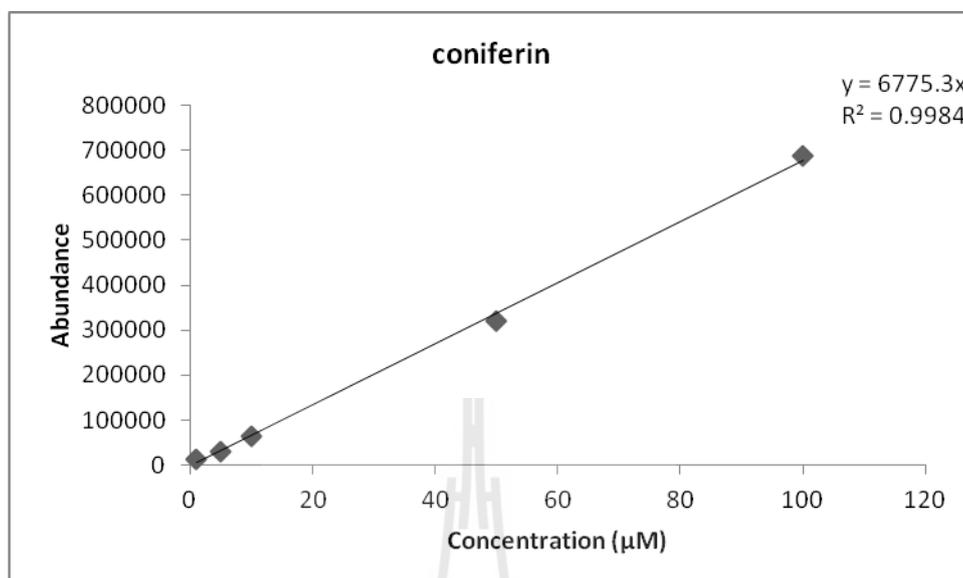
### 1. Standard curve of *p*-coumarol glucoside



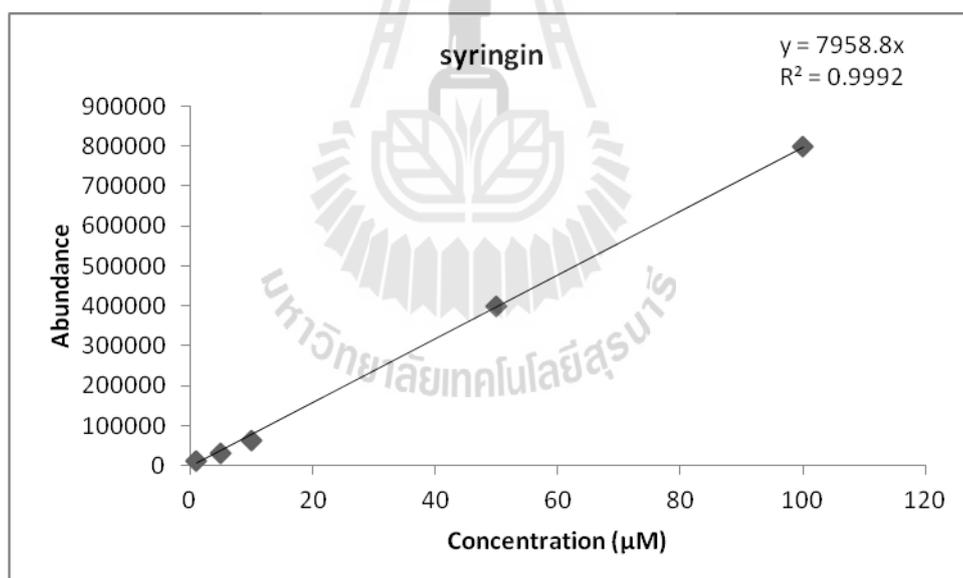
### 2. Standard curve of caffeic acid



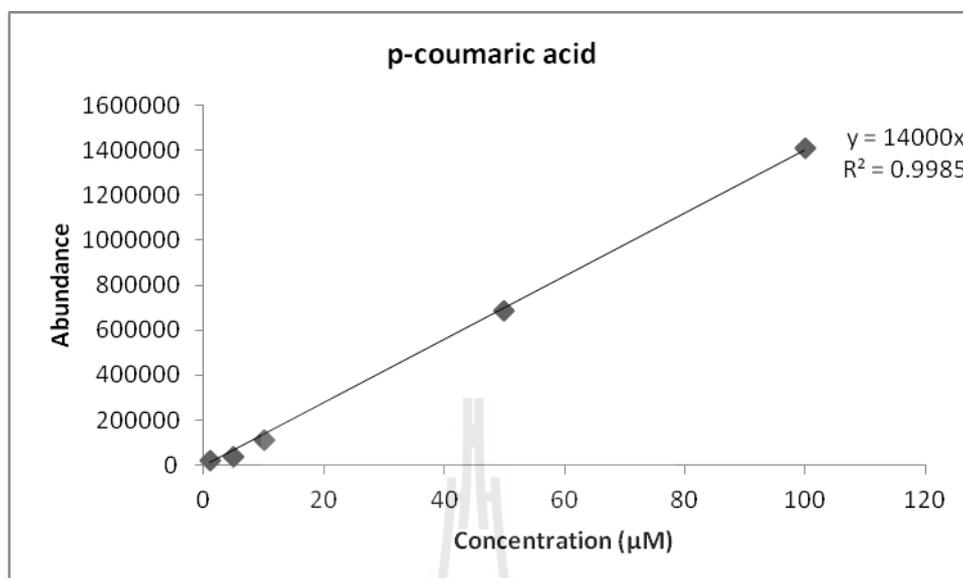
### 3. Standard curve of coniferin



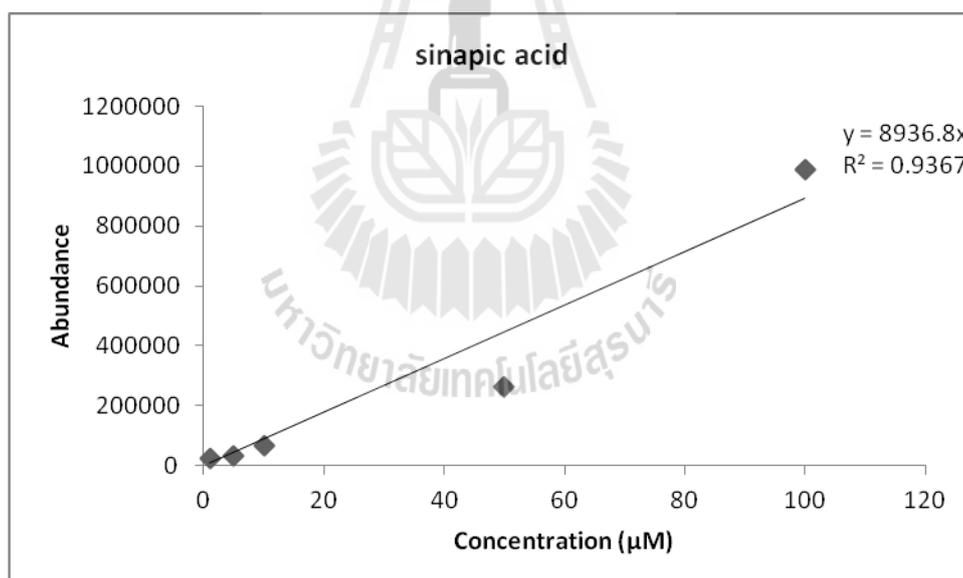
### 4. Standard curve of syringin



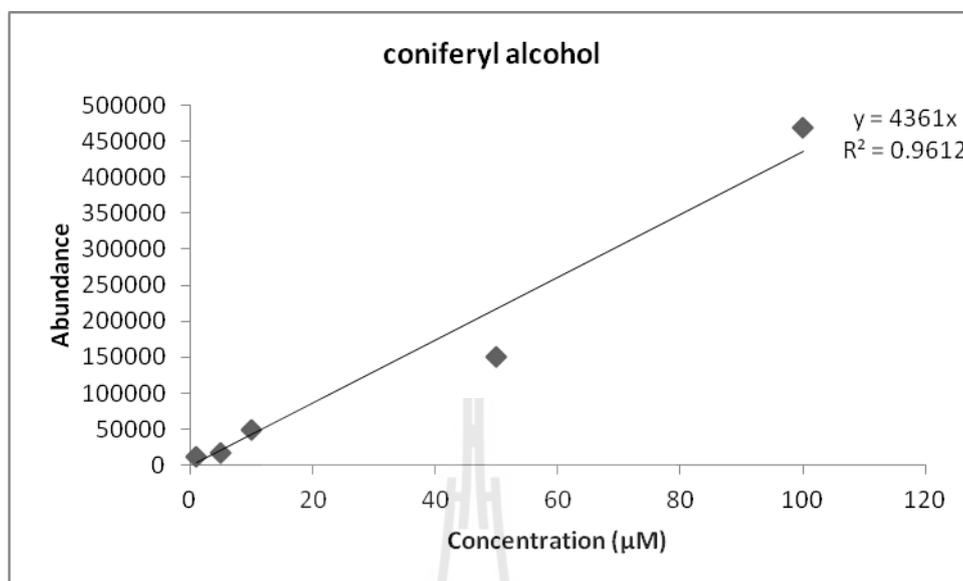
### 5. Standard curve of *p*-coumaric acid



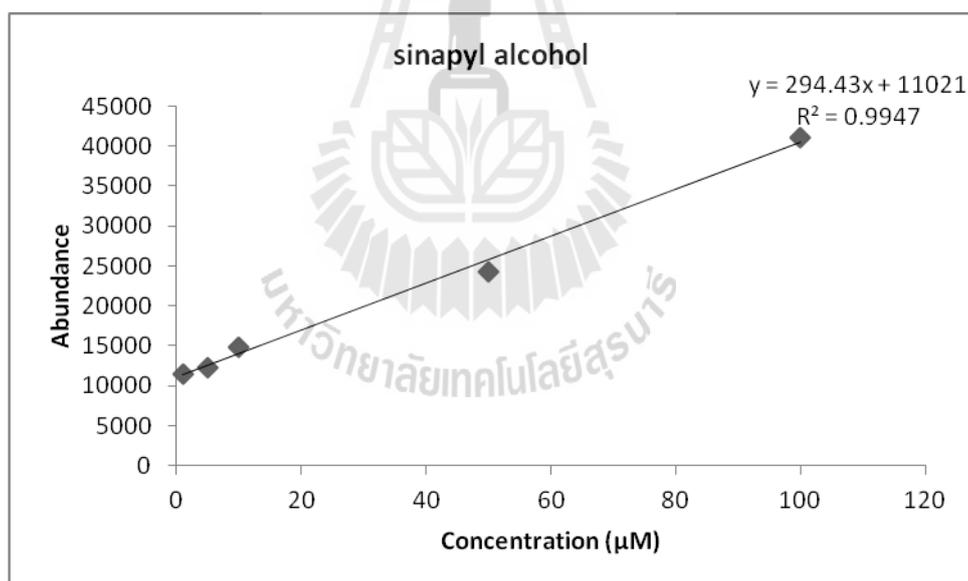
### 6. Standard curve of sinapic acid



### 7. Standard curve of coniferyl alcohol



### 8. Standard curve of sinapyl alcohol

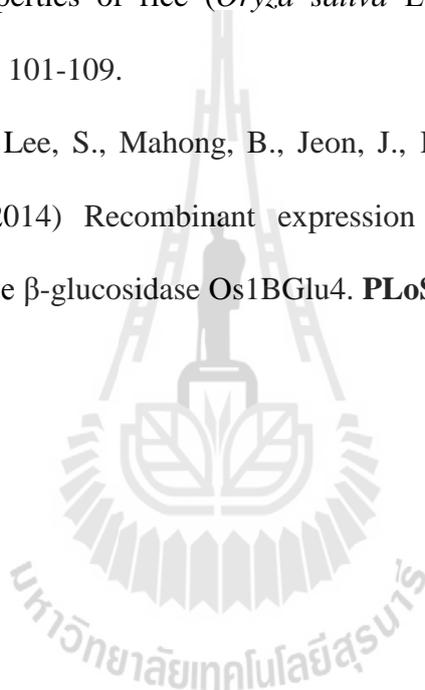


## APPENDIX B

### PUBLICATIONS

**Baiya, S.**, Hua, Y., Ekkhara, W., Ketudat Cairns., J.R. (2014) Expression and enzymatic properties of rice (*Oryza sativa* L.) monolignol  $\beta$ -glucosidases. **Plant Sci.** 227: 101-109.

Rouyi, C., **Baiya, S.**, Lee, S., Mahong, B., Jeon, J., Ketudat-Cairns, J.R., Ketudat-Cairns, M. (2014) Recombinant expression and characterization of the cytoplasmic rice  $\beta$ -glucosidase Os1BGlu4. **PLoS One.** 9(5):1-13.



## **CURRICULUM VITAE**

**NAME** Miss Supaporn Baiya  
**DATE OF BIRTH** December 31, 1986  
**PLACE OF BIRTH** Nan, Thailand

### **EDUCATION**

2009-Present School of Biochemistry, Institute of Science , Suranaree University of Technology, Nakhon Ratchasima, Thailand  
**Ph.D. (Biochemistry)**

2005-2009 School of Crop Production Technology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand  
**B.Sc. (Crop Production Technology)**

### **AWARDS AND SCHOLARSHIPS**

ChumBhot-Pantip Scholarship, 2006-2009

First Class Honors, Suranaree University of Technology, 2009

SUT Outstanding Academic Performance Scholarship, 2009-2010

The Royal Golden Jubilee Ph.D. Scholarship, from Thailand Research Fund, Thailand, 2010-2014

Outstanding Abstract Award, 7<sup>th</sup> AOHUPO Congress and 9<sup>th</sup> International Symposium of the Protein Society of Thailand, 2014