

**FUNCTIONAL AND STRUCTURAL ANALYSIS OF
PLANT GLUCOSYL CERAMIDASE (GBA2)-LIKE
ENZYMES FROM *Arabidopsis thaliana***

Karunambigai Arthanareeswaran



**A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Biochemistry
Suranaree University of Technology
Academic Year 2020**

การวิเคราะห์โครงสร้างและการทำงานของเอ็นไซม์ที่มีลักษณะคล้าย
กลูโคซิลเซรามิเดส (GBA2) จาก *Arabidopsis thaliana*

นางสาวกรุณามภิกาย อาร์ทานริสวารัน



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเคมี
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2563

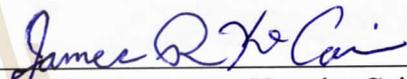
**FUNCTIONAL AND STRUCTURAL ANALYSIS OF
PLANT GLUCOSYL CERAMIDASE (GBA2)-LIKE
ENZYMES FROM *Arabidopsis thaliana***

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. Varodom Charoensawan)

Chairperson


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

Member (Thesis Advisor)


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

Member

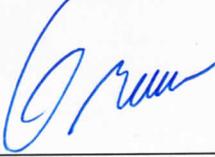

(Dr. Rung-Yi La)

Member


(Dr. Ratana Charoenwattanasatien)

Member


(Assoc. Prof. Dr. Chatchai Jothiyangkoon)


(Assoc. Prof. Dr. Worawat Meevasana)

Vice Rector for Academic Affairs
and Quality Assurance

Dean of Institute of Science

กรุณามบิกาย อาร์ทานริสวารัน : การวิเคราะห์โครงสร้างและการทำงานของเอ็นไซม์ที่มีลักษณะคล้ายกลูโคซิลเซรามิเดส (GBA2) จาก *Arabidopsis thaliana* (FUNCTIONAL AND STRUCTURAL ANALYSIS OF PLANT GLUCOSYL CERAMIDASE (GBA2)-LIKE ENZYMES FROM *Arabidopsis thaliana*) อาจารย์ที่ปรึกษา : ศาสตราจารย์ ดร. เจมส์ เกตุทัต-คาร์นส์, 167 หน้า.

กลูโคซิลเซราไมด์ (Glucosylceramide, GlcCer) เป็นองค์ประกอบสำคัญของไขมันของเยื่อหุ้มเซลล์ทั้งในสัตว์และพืช กลูโคซิลเซราไมด์ในสัตว์ถูกย่อยสลายด้วยกลูโคซิลเซเรโบรซิเดส (glucosylcerebrosidase) 1 และ 2 (GBA1 และ GBA2) เรียกอีกอย่างว่าไลโซโซมอล (lysosomal) และนอนไลโซโซมอล เบตาไกลูโคซิลเซรามิเดส (nonlysosomal β -glucosylceramidase) GBA1 เป็นเอ็นไซม์ในตระกูลไกลโคไซด์ไฮโดรเลส 30 (GH 30) และ GBA2 เป็นเอ็นไซม์ในตระกูลไกลโคไซด์ไฮโดรเลส 116 (GH116) ซึ่งพบในพืช ในปัจจุบันศึกษาเกี่ยวกับเอ็นไซม์ในพืชคือ AtGCD3 ในที่นี้เรียกว่า At4GH116 ซึ่งเป็นหนึ่งในสี่ของเอ็นไซม์ GH116 ที่ถอดรหัสจากจีโนมของ *Arabidopsis thaliana* อาจมีส่วนร่วมในแคแทบอลิซึมของกลูโคซิลเซราไมด์ At4GH116 สามารถไฮโดรไลซ์กลูโคซิลเซราไมด์ที่สกัดจาก *A. thaliana* ได้ดี แต่มีแอกติวิตีเร่งปฏิกิริยาค้ำต่อ fluorescent GlcCer

ในงานวิจัยนี้ ได้ทดสอบการทำงานของ At4GH116 ต่อสารเมแทบอลิซึมอื่น ๆ ของพืช รวมทั้งกลูโคซิลเซราไมด์ ยีนสี่ตัวที่ถอดรหัสเป็น GH116 homologues ในจีโนมของอะราบิโดบซิส ถูกตั้งชื่อเป็น At1GH116 At3GH116 At4GH116 และ At5GH116 ตามโครโมโซมที่พบ homozygous mutant lines (*Atgh116/Atgh116*) ของทั้งสี่ยีน ไม่พบความผิดปกติในการเจริญเติบโตที่เห็นได้ชัดเจนหรือมีความแตกต่างในลักษณะที่ปรากฏในแต่ละ single homozygous mutants เมื่อเปรียบเทียบกับสายพันธุ์ดั้งเดิม (Col-0) ภายใต้สภาวะการเจริญเติบโตปกติ แต่สิ่งที่น่าสนใจคือ มีการสร้างเมล็ดสีม่วงใน single homozygous mutant lines ทั้งสี่ยีนซึ่งเป็นผลมาจากความเครียดจากแสง เพื่อตรวจสอบผลของการสูญเสียหน้าที่ของการกลายพันธุ์ mutant lines ถูกผสมเพื่อสร้าง double mutant lines ความน่าจะเป็นทั้งหมดของการผสมของ single homozygous mutant lines (*Atgh116/Atgh116*) ถูกผสมข้ามเกสร แต่สามารถสร้างได้เพียง double homologous mutant lines *At1gh116* และ *At5gh116*

ในแนวทางอื่นเพื่อวิเคราะห์การทำงาน ของ A4GH116 ผ่านความจำเพาะของสารตั้งต้น โปรตีน A4GH116 แบบรีคอมบิแนนท์ถูกผลิตขึ้นในเชื้อเอสเชอริเชียโคไล (*Escherichia coli*) โปรตีน A4GH116 ถูกแยกให้บริสุทธิ์โดยวิธี immobilized metal affinity chromatography (IMAC) ตามด้วยวิธี size exclusion chromatography (SEC) ชนิด S200 ส่วนของโปรตีนที่ถูกชะออกจาก SEC ถูกตั้งชื่อเป็นพีค I ถึงพีค III ซึ่งแต่ละพีคมีน้ำหนักโมเลกุลแตกต่างกัน การวิเคราะห์ทางจลนศาสตร์พบว่า A4GH116 สามารถไฮโดรไลซ์ *p*-nitrophenyl (*p*NP)- β -D-glucopyranoside (*p*NP-Glc, $k_{cat}/K_m = 17.4 \text{ mM}^{-1}\text{s}^{-1}$) ได้อย่างมีประสิทธิภาพ แต่มีแอกติวิตีเพียงเล็กน้อยต่อ *p*NP-glycosides อื่น ๆ สำหรับไกลโคไซด์ธรรมชาติ A4GH116 สามารถไฮโดรไลซ์ apigenin-7-O- β -D-glucoside ($k_{cat}/K_m = 1133 \text{ mM}^{-1}\text{s}^{-1}$) และ flavonoid-7-O-glucosides อื่น ๆ ได้อย่างมีประสิทธิภาพมากที่สุด การวิเคราะห์ขนาดและรูปร่างโดย small angle X-ray diffraction (SAXS) พบว่า A4GH116 ที่ได้ใน SEC peak II เป็นโอลิโกเมอร์รูปทรงกลมที่มีประมาณสี่หน่วยโมโนเมอร์

จากผลการวิจัยนี้ แสดงให้เห็นว่า A4GH116 สามารถไฮโดรไลซ์ฟลาโวนอยด์กลูโคไซด์ของพืชได้อย่างมีประสิทธิภาพสูง และสามารถรวมตัวเป็นโปรตีนรูปทรงกลมแบบ trimeric หรือ tetrameric ในสารละลาย แสดงว่า A4GH116 อาจมีหน้าที่ทางชีวภาพอื่นอีกนอกเหนือจากเมแทบอลิซึมของกลูโคซิลเซราไมด์

สาขาวิชาเคมี
ปีการศึกษา 2563

ลายมือชื่อนักศึกษา

A. Kaulin

ลายมือชื่ออาจารย์ที่ปรึกษา

James R. McA

KARUNAMBIGAI ARTHANAREESWARAN : FUNCTIONAL AND
STRUCTURAL ANALYSIS OF PLANT GLUCOSYL CERAMIDASE
(GBA2)-LIKE ENZYMES FROM *Arabidopsis thaliana*.

THESIS ADVISOR : PROF. JAMES R. KETUDAT-CAIRNS, Ph.D. 167 PP.

β -GLUCOSYLCERAMIDASE / *ARABIDOPSIS THALIANA*/ GH116 T-DNA
MUTANT PLANT LINES / KINETIC AND STRUCTURAL ANALYSIS

Glucosylceramide (GlcCer) is an important lipid component of plasma membranes in both animals and plants. GlcCer is catabolized by glucosylcerebrosidase 1 and 2 (GBA1 and GBA2), also referred to as lysosomal and nonlysosomal β -glucosylceramidase. GBA1 belongs to the glycoside hydrolase family 30 (GH 30) and GBA2 is a member of glycoside hydrolase family 116 (GH116), in plants. A recent study in plants revealed that AtGCD3, here referred to as At4GH116, one of four GH116 enzymes encoded in the *Arabidopsis thaliana* genome, may participate in GlcCer catabolism. It preferentially hydrolyzed GlcCer extracted from *A. thaliana* and had low catalytic activity towards a fluorescent GlcCer.

Here, I investigated the activity of At4GH116 towards other plant metabolites, in addition to glucosylceramides. The four genes encoding GH116 homologues in the *Arabidopsis* genome were designated as At1GH116, At3GH116, At4GH116 and At5GH116, according to the chromosome on which they were found. Homozygous mutant lines (*Atgh116/Atgh116*) were identified for each of the four genes. Neither obvious growth defects nor differences in appearance were observed in any single homozygous mutants compared to wild type (Col-0) under the normal growth

conditions. Interestingly, purple color pigmentation was observed in all the four single homozygous mutant lines as a result of light stress.

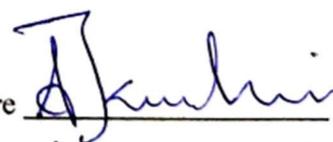
In another approach to analyze the function via the substrate specificity of At4GH116, recombinant At4GH116 protein was produced in *Escherichia coli*. At4GH116 fusion protein was purified by immobilized metal affinity chromatography (IMAC), followed by S200 size exclusion chromatography (SEC). Eluted protein fractions from SEC were named as peak I to peak III, each with a different native molecular weight. Kinetic analysis revealed that At4GH116 hydrolyzed *p*-nitrophenyl (*p*NP)- β -D-glucopyranoside (*p*NPGLc, $k_{cat}/K_m = 17.4 \text{ mM}^{-1}\text{s}^{-1}$) efficiently but had little activity toward other *p*NP-glycosides. Among natural glycosides, At4GH116 hydrolyzed apigenin-7-O- β -D-glucoside ($k_{cat}/K_m = 1133 \text{ mM}^{-1}\text{s}^{-1}$) and other flavonoid 7-O-glucosides most efficiently. Analysis of the size and shape by small angle X-ray diffraction (SAXS) revealed that the At4GH116 found in SEC peak II is a globular shaped oligomer with approximately four monomer units.

Our findings show that At4GH116 hydrolyzes plant flavonoid glucosides with high efficiency, and can form a trimeric or tetrameric globular protein in solution. This suggests that At4GH116 may have additional biological functions, aside from glucosylceramide catabolism.

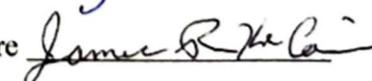
School of Chemistry

Academic Year 2020

Student's Signature



Advisor's Signature



ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Prof. Dr. James R. Ketudat-Cairns for providing me an opportunity to pursue my degree in Doctor of Philosophy in Biochemistry and for his immense support and encouragement throughout my study period.

I am thankful to Assoc. Prof. Dr. Varodom Charoensawan for his training in plant tissue culture lab skills and authorization to access his lab facility at Mahidol University, Thailand. I would like to thank Assoc. Prof. Mariena Ketudat-Cairns, for her effort to establish the plant tissue culture facility in Suranaree University of Technology, specifically for *Arabidopsis thaliana* and allowing me to utilize it.

I would like to thank Dr. Rung-Yi Lai and Ms. Chamaiporn Beagbandee, Centre for Biomolecular Structure, Function and Application (CBSFA), Thailand for their support in molecular cloning. I am grateful to Dr. Nuntaporn Kamonsuttipajit for her guidance in operating 1.3W beam line in Synchrotron Light Research Institute, (SLRI) Thailand.

I sincerely thank Dr. Salila Pengthaisong (CBSFA) and Dr. Ratana Charoenwattanasatien (SLRI) for their support in protein crystallography analysis. I would like to acknowledge all staff at Center for International Affairs (CIA), Suranaree University of Technology, Thailand for their efforts to make my overseas stay comfortable. I wish to express my appreciation to my lab colleagues for their constant support throughout the project.

Finally, I would like to express my gratitude to my family for their faith in my decision and supporting me.

Karunambigai Arthanareeswaran



CONTENTS

	Page
ABSTRACT IN THAI	I
ABSTRACT IN IN ENGLISH	III
ACKNOWLEDGEMNTS	V
CONTENTS.....	VII
LIST OF TABLES	XII
LIST OF FIGURES	XV
LIST OF ABBREVIATIONS.....	XX
CHAPTER	
I INTRODUCTION	
1.1 General introduction.....	1
1.2 Research objectives	4
II REVIEW OF LITERATURE	
2.1 Glycosylation and Deglycosylation.....	5
2.2 Sphingolipids.....	5
2.3 Sphingolipid synthesis	7
2.4 Sphingolipid functions in plants.....	11
2.5 Plant secondary metabolites	18
2.6 Phenolics.....	18
2.7 Flavonoids	19
2.8 Synthesis of flavonoids	20

CONTENTS (Continued)

	Page
2.9 Flavonoid functions in plants.....	23
2.10 Flavonoid breakdown in plants.....	24
2.11 Glycoside hydrolase	25
2.12 Mammalian glucosylceramidase	26
2.13 Plant glucosylceramidase	32
 III MATERIAL AND METHODS	
3.1 Materials	33
3.1.1 Plasmids, bacterial strains and wild type and mutant plant lines	33
3.1.2 Chemicals and reagents	33
3.1.3 Oligonucleotide primers	35
3.2 General methods.....	38
3.2.1 Preparation of <i>E. coli</i> strains NEB5 α , BL21(DE3) and Rosetta-gami(DE3) competent cells	38
3.2.2 Transformation of plasmids into competent cells	38
3.2.3 Plasmid isolation by alkaline lysis method	39
3.2.4 Vivantis plasmid DNA Extraction.....	40
3.2.5 DNA analysis by agarose gel electrophoresis	41
3.2.6 Purification of DNA bands from gels	41
3.2.7 SDS-PAGE electrophoresis	42
3.2.8 Native gel electrophoresis	43
3.2.9 Determination of protein concentration	43

CONTENTS (Continued)

	Page
3.2.10 Extraction of genomic DNA from <i>Arabidopsis thaliana</i> tissues by the CTAB method.....	44
3.2.11 Extraction of genomic DNA by Biofact™ kit	45
3.3 Amplification, ligation and cloning of ATGH16 CDS cDNA	
3.3.1 Amplification of AtGH116 CDS cDNA	46
3.3.2 Ligation of AtGH116 CDS	46
3.3.3 Cloning of AtGH116 CDS into pET30.....	46
3.3.4 Expression of pET30/AtGH116.....	48
3.3.5 Purification of recombinant At4GH116.....	49
3.3.6 Enzyme assay and synthetic and natural substrate specificity.....	52
3.3.7 Thin layer chromatography analysis of glycosyl hydrolysis reaction.....	53
3.3.8 Ultra High performance liquid chromatography (UHPLC)	53
3.3.9 Effect of inhibitors on activity of At4GH116.....	57
3.4 At4GH116 crystallization.....	57
3.4.1 Preliminary screening by vapor diffusion technique	57
3.4.2 Optimization of crystals by hanging drop method.....	58
3.4.3 Data collection and processing	59
3.5 Small Angle X-ray scattering	65
3.6 In <i>planta</i> analysis of effect of GH116 gene knockout in lines of <i>Arabidopsis thaliana</i>	66

CONTENTS (Continued)

	Page
3.6.1 Preparation and growth conditions of wild type (WT) and AtGH116 gene knockout lines.....	66
3.6.2 Identification of AtGH116 mutant lines	66
3.6.3 Generation of double mutant lines	69
3.6.4 Screening of the double mutant lines.....	69
3.6.5 Analysis of effect of nutrients and light duration on phenotypes	69
 IV RESULTS AND DISCUSSION	
4.1 Phylogenetic analysis of GBA2 protein sequences from four loci of <i>A. thaliana</i>	71
4.2 Cloning of AtGH116	78
4.3 Expression of AtGH116	79
4.4 Purification of At4GH116	85
4.5 Substrate specificity of At4GH116.....	91
4.6 Inhibition of At4GH116 by organic inhibitors.....	105
4.7 Structural determination of At4GH116	
4.7.1 Small angle X-ray diffraction	107
4.7.2 Crystallography.....	119
4.7.2.1 Initial screening by the sitting drop vapor diffusion method.....	119
4.7.2.2 Optimization by hanging drop vapor diffusion	120
4.8 Screening of homozygous GH116 mutants in <i>Arabidopsis</i>	123

CONTENTS (Continued)

	Page
4.8.1 Screening of homozygous single GH116 mutants in <i>Arabidopsis</i> .	123
4.9 Generation of homozygous double GH116 mutants in <i>Arabidopsis</i>	128
4.9.1 Generation of heterozygous double GH116 mutants in <i>Arabidopsis</i>	128
4.9.2 Generation of homozygous double GH116 mutants in <i>Arabidopsis</i>	131
4.10 Effect of light stress on the phenotype of wild type and homozygous single mutants of <i>Atgh116</i>	136
V CONCLUSION	140
REFERENCES	143
APPENDICES	161
APPENDIX A DETECTION OF AGLYCONE COMPOUNDS BY UHPLC METHOD	162
APPENDIX B PUBLICATIONS	166
CURRICULUM VITAE.....	167

LIST OF TABLES

Table		Page
3.1	Oligonucleotides primers for cloning of <i>Arabidopsis thaliana</i> glycoside hydrolase family 116 (AtGH116) into the pET30 vector	35
3.2	Oligonucleotides primers for genotyping of <i>Arabidopsis</i> mutant lines plants with respective locus number	37
3.3	PCR cycling parameters for amplification of cDNA encoding mature ATGH116's.....	47
3.4	Specification of fractions pooled together obtained from S200 size exclusion chromatography.....	51
3.5	Substrates used to analyze glucosyl hydrolase reaction	54
3.6	Commercially available crystallization kits used for preliminary screening.....	58
3.7	Optimization of At4GH116 crystals in the conditions 0.15 M - 0.4 M magnesium formate and 0.1 M Bis-Tris, pH 5 - 6.5	60
3.8	Optimization of At4GH116 crystals in the conditions 0.15 M - 0.4 M magnesium formate and 0.1 M Bis-Tris, pH 5 - 6.5 with 25% PEG 3000-6000.....	61
3.9	Optimization of At4GH116 crystals in the conditions 0.15 M - 0.3 M magnesium chloride and 0-15% Glycerol in 0.1M HEPES, pH 7.5, 16% 2-propanol.....	62

LIST OF TABLES (Continued)

Table	Page
3.10 Optimization of At4GH116 crystals in the conditions 0.24 M magnesium chloride, 10-20% 2-propanol and 0-15% glycerol in 0.1M HEPES, pH 7.5	63
3.11 Optimization of At4GH116 crystals in the conditions 0.15 - 0.3 M magnesium chloride and 10 - 17.5% 2 propanol in 0.1M HEPES, pH 7.5.....	64
3.12 PCR cycling parameters for amplification of genomic DNA from wild and mutant lines of <i>Arabidopsis</i>	68
4.1 Relative activities of At4GH116 towards <i>p</i> NP and <i>o</i> NP glycosides	92
4.2 Kinetic parameters of At4GH116 for hydrolysis of <i>para</i> - nitrophenyl glycosides.....	93
4.3 Summary of the activity of At4GH116 toward natural and synthetic glycosides based on TLC analysis	100
4.4 Kinetic parameters of At4GH116 for hydrolysis flavonoid glycosides	104
4.5 A summary of R_g , D_{max} and molecular mass estimated from the Guinier plot.....	110
4.6 Crystallization conditions of obtained crystals of At4GH116.....	119
4.7 Quantitative analysis of genotype of F ₁ progeny.....	135

LIST OF FIGURES

Figure	Page
2.1	Types of sphingolipids in plants 6
2.2	Common structure of a sphingolipid..... 6
2.3	General sphingolipid synthesis in plants..... 8
2.4	Sphingolipid LCB compositions of plants 10
2.5	Phenolic compound and their derivatives 19
2.6	Shikimic acid pathway 20
2.7	Flavonoid biosynthesis by phenylpropanoid pathway 22
2.8	Schematic representation of the glycoside hydrolase reaction 25
3.1	Schematic representations of the <i>Arabidopsis thaliana</i> GH116 genes indicating the locations of the T-DNA insertion..... 67
4.1	Phylogenetic analysis of GH116 protein sequences from four loci of <i>A. thaliana</i> 71
4.2	Amino acid sequence alignment of <i>Arabidopsis</i> GH116, TxGH116, human GBA ₂ 72
4.3	Amplification of a cDNA encoding mature AtGH116 proteins. 78
4.4	SDS-PAGE analysis of pET30/At1GH116 expression in <i>E.coli</i> 81
4.5	SDS-PAGE analysis of pET30/At3GH116 expression in <i>E.coli</i> 82
4.6	SDS-PAGE analysis of pET30/At4GH116 expression in <i>E.coli</i> 83

LIST OF FIGURES (Continued)

Figure	Page
4.7 SDS-PAGE analysis of pET30/At5GH116 expression in <i>E. coli</i>	84
4.8 SDS-PAGE analysis of At4GH116 purified by immobilized metal affinity chromatography (IMAC), followed by S200 size exclusion chromatography (SEC)	86
4.9 S200 Size exclusion chromatogram of At4GH116 expressed in BL21(DE3) host cells.....	87
4.10 <i>p</i> NPGlc hydrolase activity of fractions of At4GH116 from the S200 Size exclusion chromatography	87
4.11 Native-PAGE analysis of At4GH116 purified by immobilized affinity	89
4.12 Super imposition of S200 size exclusion chromatogram of At4GH116 expressed in <i>E. coli</i> BL21(DE3) host cells into chromatogram of different protein standards monitoring A ₂₈₀	90
4.13 TLC analysis of hydrolysis of flavonoid glucosides by At4GH116.....	95
4.14 TLC analyses of hydrolysis of oligosaccharides by At4GH116.....	96
4.15 TLC analyses of hydrolysis of glucose conjugates by At4GH116.....	97
4.16 TLC analyses of hydrolysis of glucose conjugates by At4GH116.....	98
4.17 TLC analyses of hydrolysis of anthocyanin glucosides by At4GH116.....	99
4.18 Effect of inhibitors on the enzyme activity of At4GH116.....	105
4.19 Small angle X-ray scattering profiles of At4GH116 S200 SEC.....	108
4.20 Kratky plot analysis of At4GH116 S200 SEC peak II	109

LIST OF FIGURES (Continued)

Figure	Page
4.21 Guinier plot analysis of At4GH116 S200 SEC peak II	109
4.22 Pair distribution functions $P(r)$ At4GH116 S200 SEC peak II.....	110
4.23 <i>Ab initio</i> model of Peak II.....	111
4.24 Predicted homology protein model of At4GH116 build by Swiss-model based on TxGH116	112
4.25 Predicted homology protein model of At4GH116 Δ 446-492 (with the flexible loop deleted) build by swiss-model based on TxGH116.....	112
4.26 SAXS data of 5 mg/ml of At4GH116 (blue) overlaid with the stimulated scattering oligomer profile (red) calculated by SASREF from the At4GH116 homology model based on the TxGH116 crystal structure	113
4.27 The three monomers of the homology model superimposed onto the SAXS- derived <i>ab initio</i> envelope by SASREF to generate a putative trimer	114
4.28 Four copies of the homology model were superimposed onto the SAXS- derived <i>ab initio</i> envelope by SASREF to generate a putative tetramer	115
4.29 The three monomers of the homology model of At4GH116 Δ 446-492 superimposed onto the SAXS-derived <i>ab initio</i> envelope using SASREF to generate a putative trimer	116
4.30 Four copies of the homology model of At4GH116 Δ 446-492 were superimposed onto the SAXS-derived <i>ab initio</i> envelope by SASREF to generate a putative tetramer	117

LIST OF FIGURES (Continued)

Figure	Page
4.31 At4GH116 crystals obtained from vapor diffusion sitting drop method screening at 15°C	119
4.32 At4GH116 crystals obtained from vapor diffusion hanging drop method screening at 15°C	121
4.33 At4GH116 crystals obtained from vapor diffusion hanging drop method screening at 15°C	122
4.34 Genotyping PCR of the T-DNA insertion knock-out line Salk_143139 of At1GH116.....	124
4.35 Genotyping PCR of the T-DNA insertion A) knock-out line Salk_ 117180, B) Salk_ 039506 of At3GH116	125
4.36 Genotyping PCR of the T-DNA insertion A) knock-out line Salk_099838, B) Salk_010803 of At4GH116	126
4.37 Genotyping PCR of the T-DNA insertion A) knock-out line Salk_ 057433, B) Salk_039511 At5GH116.	127
4.38 Genotyping PCR of the T-DNA insertion At1GH116 knockout lines Salk_143139 and At5GH116 knockout lines Salk_057433 double mutant F ₁ generation	129
4.39 Genotyping PCR of the T-DNA insertion At1GH116 knockout lines Salk_143139 and At4GH116 knockout lines Salk_010803 double mutant F ₁ generation	130

LIST OF FIGURES (Continued)

Figure	Page
4.40	Genotyping PCR of the T-DNA insertion <i>At1gh116</i> knockout lines Salk_143139 and <i>At5gh116</i> knockout lines Salk_057433 double mutant F ₂ generation 132
4.41	Genotyping PCR of the T-DNA insertion <i>At1gh116</i> knockout lines Salk_143139 and <i>At5gh116</i> knockout lines Salk_057433 double mutant F ₂ generation 133
4.42	Genotyping PCR of the T-DNA insertion <i>At1gh116</i> knockout lines Salk_143139 and <i>At5gh116</i> knockout lines Salk_057433 double mutant F ₂ generation 134
4.43	Comparison of phenotypes of wild type <i>A. thaliana</i> , with the other mutants of <i>Atgh116</i> grown under long daylight conditions for 24 days 137
4.44	Comparison of phenotypes of wild type <i>A. thaliana</i> , with the other mutants of <i>Atgh116</i> grown under short day light conditions for 38 days..... 138
4.45	Effect of light duration on phenotypes of wild type <i>A. thaliana</i> , and other mutants of <i>Atgh116</i> on 65th day 139

LIST OF ABBREVIATIONS

3-DHS	3-dehydroshikimic acid
4MUG	4-methylumbelliferyl- β -D-glucosylpyranoside
ABP	Activity-based probe
ACER	Alkaline ceramidase
APS	Ammonium persulfate
ARCA	Autosomal Recessive Cerebellar Ataxia
AtACER	Arabidopsis ceramidase
bis-acrylamide	N,N-methylene-bis-acrylamide
bp	Base pairs
BSA	Bovine serum albumin
C4H	Cinnamate-4-hydroxylase
CAZymes	Carbohydrate active enzymes
CBE	Conduritol B epoxide
CBG	Cytosolic β -glucosidase
CD	Coding sequence
cDNA	Complementary deoxynucleic acid
Cer	Ceramide
CoA	Coenzyme A
CV	Column volume

LIST OF ABBREVIATIONS (Continued)

DAHP	9-deoxy-D-arabino-heptulosonate-7-phosphate
DHS	Dihydrosphingosine
DMSO	Dimethyl sulfoxide
DNA	Deoxynucleic acid
DNase	Deoxyribonuclease
DNJ	Deoxynojirimycin
dNTPs	Deoxynucleoside triphosphates
EC	Enzyme commission
ER	Endoplasmic reticulum
EtOAc	Ethyl acetate
F3'H	Flavonoid 3'-hydroxylase
FB1	Fumonisin B1
FBR11	Fumonisin B1 resistant 11-1
GBA	Glucocerebrosidase
GBA1	Lysosomal β -glucosidase
GBA2	Non-lysosomal β -glucosidase
GD	Gaucher disease
GH	Glycoside hydrolase
GH 116	Glycoside hydrolase family 116
GIPC	Glycosyl inositol phospho ceramide
GlcA	Glucuronic acid

LIST OF ABBREVIATIONS (Continued)

GlcCer	Glucosylceramide
GSL	Glycosphingolipids
GT	Glycosyltransferase
HPLC	High performance liquid chromatography
HSP	Hereditary Spastic Paraplegia
IMAC	Immobilized metal affinity chromatography
IMAC	Immobilized affinity chromatography
IPCS	Inositol phosphorylceramide synthase
IPTG	Isopropyl thio- β -D-galactoside
kDa	Kilo Daltons
LB	Luria-Bertani lysogeny broth
LCB	Long chain bases
LCB-P	Long chain base phosphate
MeOH	Methanol
MIPC	Mannosyl inositol phosphorylceramide
MW	Molecular weight
NaOAc	Sodium acetate
NB-DNJ	N-butyldeoxynojirimycin
NCER	Neutral ceramidase
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PAL	Phenylalanine ammonia lyase

LIST OF ABBREVIATIONS (Continued)

PAR	Photosynthetically active radiation
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PHS	Phytosphingosine / 4-hydroxysphinganine
PL	Polysaccharide lyase
PM	Plasma membrane
PMSF	Phenylmethylsulfonyl fluoride
<i>p</i> NP	para -Nitrophenyl
<i>p</i> NPGlc	<i>p</i> NP- β -D-glucopyranoside
PPP	Pentose phosphate pathway
RNA	Ribonucleic acid
RNase	Ribonuclease
ROI	Reactive oxygen intermediates
ROS	Reactive oxygen species
rpm	Rotations per minute
S1P	Sphingosine 1-phosphate
SAXS	Small angle X-ray diffraction
SB	Sphingoid base
SDS	Sodium dodecyl sulfate
SDS-PAGE	Polyacrylamide gel electrophoresis with SDS
SEC	Size exclusion chromatography
SL	Sphingolipids

LIST OF ABBREVIATIONS (Continued)

SNP	Single nucleotide polymorphism
SPT	Serine palmitoyltransferase
STHL	Short-term high light
TEMED	Tetramethylenediamin
TEV	Tobacco etch virus
TLC	Thin-layer chromatography
Tris	Tris-hydroxymethyl-aminoethane
UHPLC	Ultra high-performance liquid
UV	Ultraviolet radiation
v/v	volume by volume
VLCFA	Very-long chain fatty acids
w/v	weight by volume
WT	Wild type

CHAPTER I

INTRODUCTION

1.1 General Introduction

Glycosidases or glycoside hydrolases (GHs) are a group of enzymes found in all living organisms (Ardevol and Rovira, 2015). GHs catalyze the hydrolysis of glycosidic bonds in glycoconjugates, oligosaccharides and polysaccharides (van Wyk et al., 2017). GHs constitute the most numerous category of carbohydrate active enzymes (CAZymes) listed in the carbohydrate active enzyme database (CAZy, <http://www.cazy.org/>) (Cantarel et al., 2009). In plants, GHs are involved in cell wall metabolism, mobilization of energy, defense, symbiosis, and signaling, as well as in the metabolism of secondary plant metabolites and glycolipids (Minic, 2008). Glucocerebrosidase (GBA) is one of the GH enzymes involved in metabolism of glycolipids.

Inherited deficiencies in GBA activity in humans cause an accumulation of glucosylceramide (GlcCer) within lysosomes, subsequently leading to the most common lysosomal storage disorder, known as Gaucher disease (GD) (Brady et al., 1966; Grabowski and Horowitz, 1997). In 1965 Brady et al. proved the function of GBA in the glycolipid metabolic pathway. In the presence of C^{14} labeled glucocerebroside as a substrate, the percentage of water soluble radioactive C^{14} glucose was drastically lower in the aqueous phase of lysosomal fractions from the

spleens of the Gaucher's patient compared to those from normal spleens. Incubation of membrane suspensions of murine macrophage cell line RAW 264 and human erythroblastoid cell line K562 with 4-methylumbelliferyl- β -D-glucoside (4-MU- β -glucoside) resulted in significant formation of 4-methylumbelliferone (van Weely et al., 1993). These findings suggested the existence of lysosomal and non-lysosomal GBA.

To follow up the previous finding, melanoma cells were preincubated with or without CBE, a potent inhibitor of lysosomal GBA, and deoxynojirimycin (DNJ) analogues and their hydrolysis activity with 4-MU- β -glucoside measured (Overkleeft et al., 1998). The CBE-sensitive activity was certified to the lysosomal GBA and the insensitive activity to the non-lysosomal GBA. This inhibitory study demonstrated the significant difference between the lysosomal and non-lysosomal GBA. Recently, lysosomal and non-lysosomal β -glucosidases that hydrolyze GlcCer have been differentiated as glucocerebrosidase 1 and 2 and abbreviated as GBA1 and GBA2.

Comparison of the activity of GBA1 and GBA2 in the cell lysates and plasma membrane (PM) from the human fibroblast cell lines (as a control) and those from patients with GD revealed that GBA2 was induced in the absence of GBA1 activity in humans (Aureli et al., 2012). GBA1 and GBA2 are members of two different families of sequence-related glycoside hydrolases designated as GH30 and GH116, respectively.

GBA1 (EC: 3.2.1.45) belongs to the GH30 family, members of which are found in a wide range of living organisms, such as animals, fungi, archaea, and eubacteria, but not in plants. In plants, in the absence of GH30, we and others (Dai et al., 2020) have hypothesized that the glucosylceramidase activity may be replaced by

homologues of GBA2 non-lysosomal β -glucosidase, since multiple GH116 family genes are found in plants.

In plants, there are multiple GH116 genes in the genome. *Arabidopsis* and rice each have four genes encoding GH116 enzymes (GBA2-like enzymes). Recently, one of the four *Arabidopsis thaliana* (*A. thaliana*) homologs (At4G10090) of non-lysosomal glucosylceramidase (*AtGCD3*) was expressed in *E. coli* (Dai et al., 2020). Subcellular localization analysis revealed that, the *AtGCD3* localizes in the PM and near the endoplasmic reticulum (ER). Their results demonstrated that *AtGCD3* can hydrolyze long acyl-chain GlcCer purified from *Arabidopsis* wild type leaves and also found to have a low K_m and a very low k_{cat} for the fluorescent C6-NBD GlcCer. In addition, *AtGCD3* is insensitive to conduritol B epoxide (CBE), which is an inhibitor of GBA1, but has a relatively little effect on GBA2.

Based on the findings in this study, we hypothesized *AtGCD3* (At4G10090, designated At4GH116 here) expressed in *E. coli*, may hydrolyze natural substrates other than sphingolipids. So, we expressed the enzyme in *E. coli* and tested its ability to hydrolyze available natural and synthetic substrates. In addition, *in planta* functional characteristics of *A. thaliana* GH116, (*AtGH116*) genes were analyzed in T-DNA insertion mutant lines compared to wild type (WT). Furthermore, additional solution structural of *At4GH116* was analyzed by SAXS to understand the shape and native molecular weight and the protein was crystallized.

1.2 Research objectives

In order to learn the function of GH116 enzymes in plants, I approached the following objectives:

1. Generation of knockout *Arabidopsis thaliana* lines and observes their phenotypes.
2. Expression of the cDNA of *Arabidopsis thaliana* GH116 glucosylceramidase (AtGH116) to produce protein in *E. coli*, optimize for suitable expression conditions and purify the GH116 protein.
3. Characterization of enzyme properties, substrate specificities and their kinetic analysis.
4. Investigation of the structure of AtGH116 in solution by small angle X-ray scattering (SAXS) and X-ray crystallography.

CHAPTER II

REVIEW OF LITERATURE

2.1 Glycosylation and Deglycosylation

Glycosylation and deglycosylation are essential reactions, which occur in all organisms on this earth. Glycoproteins, glycolipids, oligosaccharides, polysaccharides and proteoglycans are products of glycosylation reactions and deglycosylation returns these molecules to their aglycon or monosaccharide forms. Both glycosylation and deglycosylation are enzymatic processes. Among the many glycosylated products, this thesis considers sphingolipids and glycosphingolipids, along with flavonoids and other plant glycosides.

2.2 Sphingolipids

The term “sphingosine” was coined by J. L. W. Thudichum (1829–1901), a German-born physician working in London, who identified sphingolipids (SLs) in brain, and first published these findings in 1884 (Michaelson et al., 2016; Dickson and Lester, 2002). Generally, in all eukaryotes, as well as some prokaryotes, the cell membrane contains three distinct and primary lipid components, glycerophospholipids, sterols and sphingolipids, where their mass presence provides the membranes with chemical and mechanical stability (Rego et al., 2014). Plant sphingolipids can be divided into 4 main classes; these are free long chain bases

(LCB), ceramides, glycosylceramides, and glycosyl inositol phosphoceramides (GIPCs) (Figure 2.1).

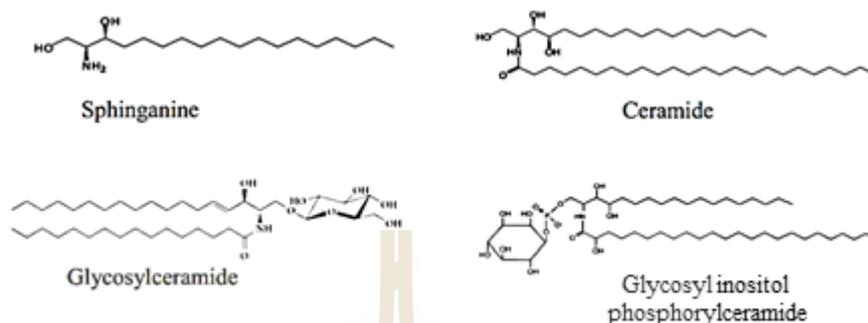


Figure 2.1 Types of sphingolipids in plants.

The predominant glycosphingolipids (GSL) in lipid extracts of plant tissues is GlcCer, although complex GSLs have been isolated from plant tissues. SLs are made by the combination of polar head groups, such as sugar residues and nonpolar ceramides, which are composed of a sphingoid base (SB) carrying an N-acylated fatty acid with 14-26 carbon atoms (Figure 2.2).

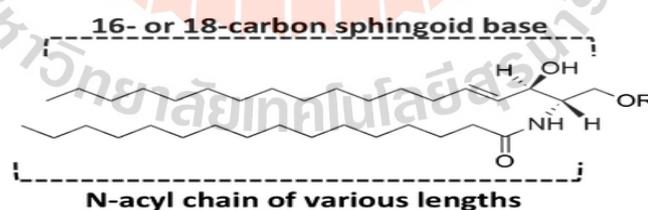


Figure 2.2 Common structure of a sphingolipid (Warnecke and Heinz, 2003).

Complex SLs, such as cerebrosides, are generated by addition of glycosyl residues like glucose and/or mannose and polar phosphate-containing head groups to

the ceramide (Figure 2.1). Up to 2012, 400 different GSLs had been identified that vary in their sugar chain structure (Cowart and Obeid, 2007).

2.3 Sphingolipid synthesis

In all eukaryotic organisms, *de novo* sphingolipid biosynthesis begins in the endoplasmic reticulum (ER) with the production of 3-ketosphinganine, the simplest LCB found in plants. Further sphingolipid biosynthesis starts with the synthesis of LCBs, which is comprised of three enzymatically catalyzed steps (Figure 2.3). The first committed reaction is condensation of a fatty acyl group, (palmitoyl or stearoyl CoA for C16 and C18 bases, respectively) with serine by serine palmitoyltransferase (*SPT*). *SPT* is a heterodimer that consists of *long chain base 1* and *2* from the *LCB1* and *LCB2* genes, respectively (Melser et al., 2010). In the second step, 3-Keto sphinganine is reduced to sphinganine or dihydrosphingosine (DHS) by the NADPH-dependent 3-keto dihydrosphinganine reductase (*TSC10*). The *TSC10* gene encodes this enzyme, an essential enzyme in *S. cerevisiae*, which has been located at the cytosolic side of the ER membrane in mammalian cells.

The structural variability in plant LCBs arises by further modification of LCBs by hydroxylation, phosphorylation and desaturation to form sphingoid base derivatives and acylation to form ceramides (Markham et al., 2013). The final step in the LCB synthesis and/or a prominent modification is conversion of DHS to phytosphingosine / 4-hydroxysphinganine (PHS), which is hydroxylated at C4 by *sur2/syr2* hydroxylase (Suppressor of *rsv161*/syringomycin resistance). The LCB composition of plants is more variable, both DHS and PHS are referred as a sphingoid bases (SB) (Figure 2.4) (Markham et al., 2013).

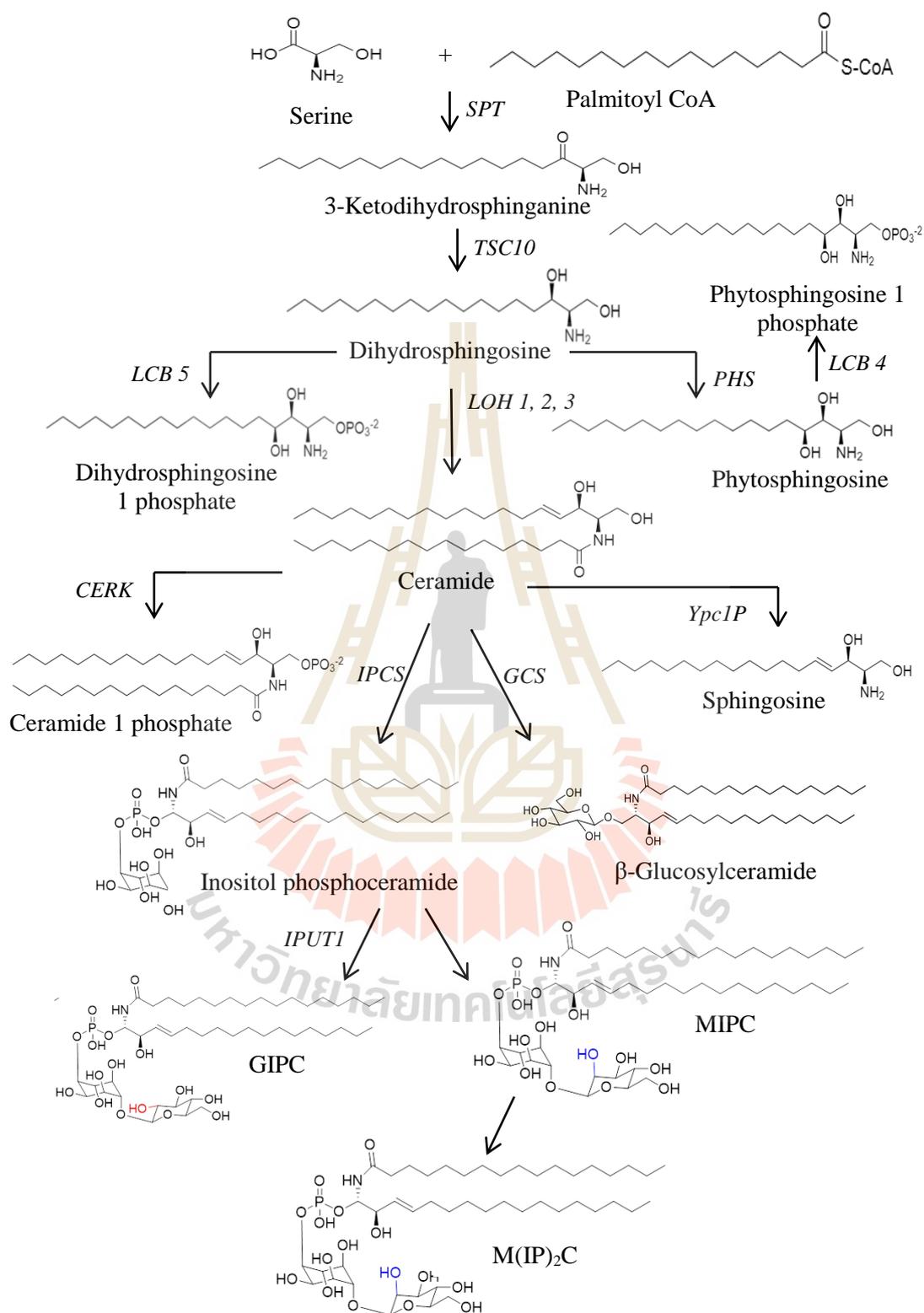


Figure 2.3 General sphingolipid synthesis in plants.

Another modification is phosphorylation of SBs, in which LCB kinases, encoded by *LCB4* and *LCB5* phosphorylate DHS and PHS to dihydrosphingosine-1-phosphate (DHS-1-P) and phytosphingosine-1-phosphate (PHS-1-P), respectively. After acylation of LCBs to form ceramide, additional *cis* or *trans*-bound desaturation occurs at either C4 by $\Delta 4$ desaturase or C8 by $\Delta 8$ desaturase, at the cytosolic face of the ER membrane. This results in, the unsaturated plant LCB's, such as [E/Z] sphing-8-enine [d18:1⁸], [4E, 8E/Z]-sphing-4-8-dienine [d18:2^{4,8}] and [8E/Z]- 4 hydroxy-8-sphingenine [t18:1⁸], where d18:1⁴ is virtually absent and d18:0 and t18:0 are only present in minor proportions. Other LCBs differing in chain length are also present as minor components in plant sphingolipids.

The N-acylation step occurs in the ER membrane and involves the addition of a C16 - C26 fatty acyl group, via an amide bond, by ceramide synthase to SB to yield ceramide. Ceramide synthase are encoded by the *LOH1*, *LOH2* and *LOH3* genes (Luttgeharm et al., 2015). The first choice of LOH2 is C16 acyl chains with dihydroxy LCB, while LOH1 and 3 displayed specificity for 20 to 26 carbon acyl chains in link with a trihydroxylated LCB. The most abundant α -hydroxylated fatty acids are saturated C16, C20, C22 and C24, whereas $\omega 9$ -mono unsaturated very-long chain fatty acids (VLCFA), ranging from C22-C26, occur in small amounts. Dihydroxy LCBs and trihydroxy LCBs that are acylated with C16 FAs and very long chain fatty acyl are primarily for synthesis of glucosylceramide and glucosylceramide and glycosyl inositol phosphoceramide, respectively.

After synthesis of ceramides in the ER, SLs are transported to the Golgi for generation of complex SLs (Lynch and Dunn, 2004). In complex sphingolipids, the polar head group, glucose or inositol phosphate is linked to C-1 of the N-acyl long-

chain base to yield glucosylceramide and derivatives of inositol phosphoryl ceramide (IPC), respectively. The variety of sphingolipids with different length of the fatty acid, hydroxylation, and saturation sites and head group substituents were established within the same organism, as well as across species. GlcCer and IPC, the two complex sphingolipids are reported in higher plants and yeast, and IPC derivatives are found in plants, fungi and kinetoplastid protozoa, but not in mammals.

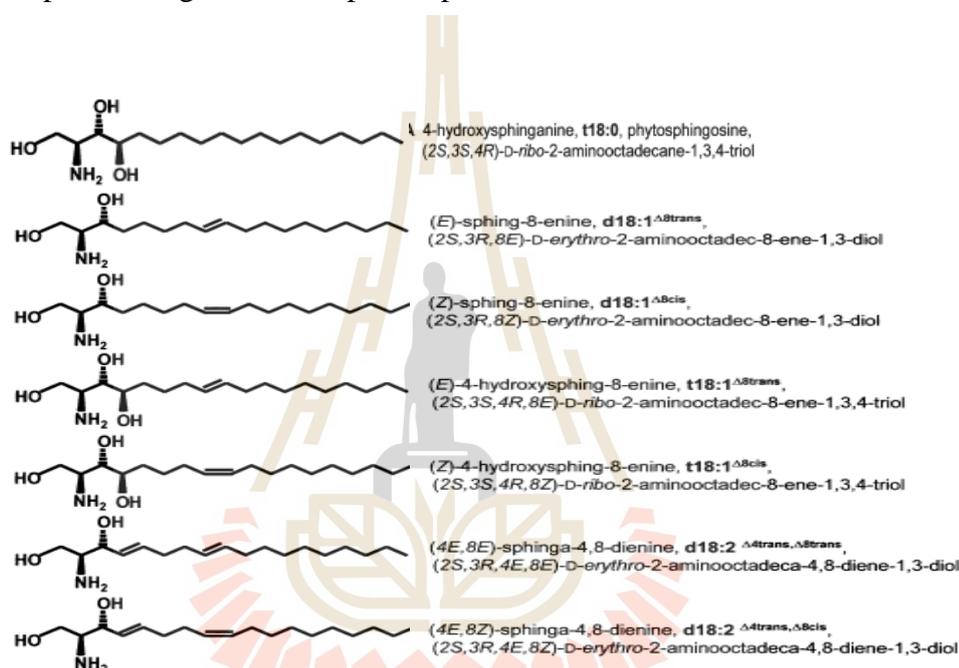


Figure 2.4 Sphingolipid LCB compositions of plants (Warnecke and Heinz, 2003).

Glucosylceramide synthase (*GCS*) catalyzes the transfer of a glucose moiety from UDP-glucose onto the ceramide, forming GlcCer. In complex sphingolipid synthesis, GlcCer is further glycosylated in reactions catalyzed by glycosyltransferases. The glycosylated products can also be deglycosylated by glycoside hydrolases (Lahiri and Futerman, 2007; Markham et al., 2013).

In plants and some species of yeast, ceramides are first converted into IPC by the transfer of a phosphoryl inositol group to ceramide catalyzed by the IPC synthase. Once formed, IPC has two possible fates. First it may be mannosylated to form mannosyl inositol phosphorylceramide (MIPC) via IPC mannosyltransferase, which transfers the mannose from the nucleotide sugar GDP-mannose to the inositol group in IPC. Second, inositol phosphorylceramide glucuronosyltransferase 1 (*IPUT1*), mediates the transfer of glucuronic acid (GlcA) from UDP-GlcA to inositol phosphorylceramides to synthesize glycosyl inositol phosphorylceramides (GIPCs). The concluding step in sphingolipid synthesis is establishment of M(IP)2C, by addition of another inositol phosphate to MIPC, which is catalyzed by inositolphosphotransferase 1 (*IPT1*).

2.4 Sphingolipids functions in plants

Sphingolipids are involved in a diverse range of functions. SLs are the most complex membrane lipids found ubiquitously in eukaryotes (Ali et al., 2018). The mode of actions depends on the distribution and concentration of SLs in the membrane. Relatively few studies focusing on plant SLs function have been reported. Based on the vast number of studies in animals, SLs are known to be signaling molecules that act in cell regulation, growth and development and membrane structural components (Hannich et al., 2011).

Serine palmitoyltransferase (SPT) catalyzes the initial step of sphingolipid synthesis. SPT is a heterodimer that consists of LCB1 and LCB2 subunits. In homozygous Δ AtLCB1, the viability of the plant was reduced and the plant size was reduced, as a result of reduced cell expansion. Plants also showed changed leaf

morphology upon the suppression of expression of AtLCB1 by RNA interference (AtLCB1 RNAi) (Chen et al., 2006). In Arabidopsis fumonisin B1 resistant11-1 (FBR11) mutant (*fbr11-1*) plants, the sphingoid bases dihydrosphingosine, phytosphingosine, phytosphingosine phosphate and dihydrosphingosine phosphate were massively increased in response to fumonisin B1 (FB1). The free sphingoid bases efficiently induce reactive oxygen intermediates (ROI) generation followed by cell death (Shi et al., 2007). In a further study by Teng et al. (2008), a double mutant in two homologous LCB2 genes with T-DNA insertions displayed a phenotype similar to *fbr11-1*. These outcomes recommended that SPT modulated programmed cell death plays significant role in plant morphology. Functional isoforms of Arabidopsis LCB2 subunit has two genes: AtLCB2a and AtLCB2b. Either of a T-DNA single mutant of these genes doesn't exhibit any changes in sphingolipid growth or content. Reciprocal crosses of *Atlcb2a Atlcb2b* (*lcb2*) mutants exhibited lethality and it is related with the incapability to transmit the null genotype of *lcb2* through the haploid pollen (Dietrich et al., 2008).

In another study, the *lcb2a* mutants showed a decline growth, chlorosis and flowering as a stress response, but no tissue death was observed in WT on exposure to FB1. In contrast with WT cells, disturbance of the LCB2a gene exaggerated chloroplast morphology, in particular an increase in the number of rounded plastids and size. On FB1 treatment, in the *lcb2a* mutant, LCB's, especially d18:0, d18:0-P, t18:0 and t18:0-P, displayed foremost upsurges (García et al., 2011). Arabidopsis T-DNA insertion mutants of $\Delta 4$ -were devoid of $\Delta 4$ -unsaturated sphingolipids in floral tissue, and also had condensed accumulation of GlcCer (Michaelson et al., 2009).

Similar results were observed by Luttgeharma, et al. (2015), loss of $\Delta 4$ -unsaturated LCB sphingadiene (d18:2) in the $\Delta 4$ desaturase mutant resulted in GlcCer concentrations up to 50%. Pollen fractions were more highly enriched in GlcCer than leaves and pollen glycosylinositolphosphoceramides (GIPCs) were accumulated with complex array of N-acetyl-glycosylated GIPCs, with three pentose units that were lacking in the leaf GIPCs. Arabidopsis LCB $\Delta 8$ -desaturase mutants (*Atsld1*) revealed huge reductions in $\Delta 8$ -unsaturated LCBs in roots, leaves, flowers, stems and silique, whereas *Atsld2* mutants showed slight change in LCB unsaturation. In *sld1 sld2* double mutant no obvious LCB $\Delta 8$ unsaturation was observed. The *sld1 sld2* mutant lacked apparent growth phenotypes under normal growth conditions, but displayed different changes under prolonged exposure to low temperatures (Chen et al., 2012).

Double LCB C-4 hydroxylase/sphingobase hydroxylase 1 (*sbh1sbh2*) mutants, with no trihydroxy LCBs, exhibited dwarfed statures and lack of development from vegetative to reproductive growth plants, with enhanced programmed cell death related genes (Chen et al., 2008). Biotic stress in *sbh1-1* mutants with virulent *Pseudomonas*, the levels of d18:0 and d18:1 was significantly elevated, while total t18:0 and t18:1 content was about 30% less (Peer et al., 2010). In mutant lines of sphingosine-1-phosphate lyase 1 (DPL1), sphingosine-1-phosphate phosphatase 1 (SPP1), and sphingosine kinase 1 (SPHK1), the degree of FB1-induced cell death was elevated. Treatment with phosphorylated LCBs species did not reduce FB1-induced cell death. In conclusion, the examination of LCB levels during cell death displayed a correlation between nonphosphorylated LCBs and cell death (Glenz et al., 2019).

Inositol phosphorylceramide glucuronosyltransferase1 (IPUT1) moves GlcA from UDP-GlcA to glycosyl inositol phosphorylceramide (GIPC). In *in vitro* studies,

overexpression and silencing of IPUT1 in *Nicotiana benthamiana* resulted in decrease or an increase of IPC glucuronosyltransferase activity, respectively. Accumulation of IPC, ceramides and glucosylceramides were observed in IPUT1 silenced plants.

In plants mutations in IPUT1 are not diffused through pollen, signifying that these sphingolipids are necessary for pollen development (Rennie et al., 2014). These show that, SLs are implicated as signaling molecules participated in cell death. The knockout of the *Arabidopsis* $\Delta 4$ desaturase gene catalyzing a step in the synthesis of sphingosine caused activation of programmed cell death (PCD). Inositol phosphorylceramide synthase (IPCS), the first enzyme of the GIPC synthesis pathway is involved in the regulation of plant PCD and was shown to be related with defense in *Arabidopsis* (Wang et al., 2008; Smith and Fry, 1999; Lynch and Steponkus, 1987).

The important regulators of apoptosis in animal cells are long chain base phosphate (LCB-P), sphingosine 1-phosphate (S1P) and ceramide (Cer). Alden et al. (2011) assessed the role of LCB-Ps in apoptotic-like programmed cell death (AL-PCD) in *Arabidopsis* by exposing *Arabidopsis* cell cultures to cell death-inducing treatments in the occurrence of S1P. The results showed that the inhibitors of sphingosine kinase in the cultured cell suspension, the levels of internal LCB-P were altered and death rates were elevated in heat stress. S1P decreased AL-PCD and stimulated cell survival in heat stress. These results demonstrate that a sphingolipid is involved in controlling cell cycle and regulation of cell death in *Arabidopsis*. Sphingolipids have also developed as important signals for plant response to low temperature. Ceramide kinase activity was estimated in unstressed and cold-stressed WT and *Arabidopsis* ceramide kinase mutant *acd5* plants (Dutilleul et al., 2015). Thin layer chromatography with metabolic radiolabeling revealed that a transient and rapid

creation of Cer-P occurred in Arabidopsis plantlets and cultured cells under cold-stressed WT, which was highly reduced in the *acd5* mutant. *acd5* mutant and WT responded similar at the seedling stage, but seed germination of mutant was hypersensitive to the cold.

Ceramidases catalyze ceramide into fatty acids and sphingosine. Arabidopsis ceramidase (AtACER) was characterized by T-DNA insertion and artificial microRNA-mediated silenced lines (AtACER RNAi). In *acer-1* T-DNA insertion mutant plants, reductions in leaf size, dwarfing and an irregular wax layer, were observed compared to WT. Quantitative SL analysis showed that elevated levels of ceramide and reduced long chain bases in *acer-1* mutants and AtACER RNAi with improved sensitivity to salt stress, on the other hand lines overexpressing AtACER showed increased tolerance to salt stress. Decrease of AtACER also increased plant vulnerability to *P. syringae*. These results emphasize the key biological functions of ceramidases in biotic and abiotic stresses in plants (Wu et al., 2015).

The Arabidopsis gene encoding, sphingosine kinase 1 (SPHK1), an enzyme involved in phosphorylation of sphingosine, phytosphingosine and other sphingoid long-chain bases, was manipulated by T-DNA insertion. The two T-DNA insertion lines are referred as *SPHK1-OE* (over expression) and *SPHK1-KD* (knock-down). The stomata of *SPHK1-KD* were less sensitive, whereas the stomata of *SPHK1-OE* plants were more sensitive than WT to ABA. The germination rate of *SPHK1-KD* was enriched, whereas the rate of germination was reduced for *SPHK1-OE* seeds (Worrall et al., 2008).

Fungal toxins (like fumonisin B and *Alternaria alternata* AAL-toxin) can inhibit *de novo* ceramide biosynthesis and thereby cause accumulation of sphinganine

and 4-hydroxysphinganine, which promotes necrosis and leads to cell death in plant (Melser et al., 2010). Over-expression of the tomato *Asc1* gene, involved in ceramide synthesis, confers resistance to fumonisin B and AAL-toxin and permits SL synthesis and inhibits accumulation of long-chain bases in tissues exposed to the toxin. Sphingosine-1-phosphate is involved in signal transduction, by activating the plant hormone abscisic acid, leading to stomatal pore regulation (Uemura and Steponkus, 1994).

Plants deficient in synthesis of GlcCer caused by functional defects in GCS could not grow further than seedling stage and heterozygous plants possessing a null GCS allele had flawed pollen transmission and also exhibited a less differentiation defect in regenerating callus tissue (Pata et al., 2010). N-5-(adamantane-1-yl-ethoxy)pentyl-L-ido-deoxynojirimycin (L-ido-AEP-DNJ, an N-substituted iminosugar analogue) inhibited root growth of *Arabidopsis* by 92% in agar plates. LC-MS analysis of the inhibitor-treated *Arabidopsis* seedlings showed a reduction in glucosylceramide (Rugen et al., 2018).

In *Arabidopsis*, three genes encoding ceramide synthases (*longevity assurance gene LOH1, 2, and 3*) with distinct substrate specificities are present (Luttgeharm et al., 2015). The cDNA for each gene was transformed and overexpressed in *Arabidopsis* plants. A difference in growth was observed in plants overexpressing *LOH1* and *LOH3* compared to those overexpressing *LOH2*. Overexpression of *LOH2* caused in dwarfing. In contrast, the biomass of plants was enhanced in *LOH1* and *LOH3* overexpression compared to wild type. This results signifying that enhanced synthesis of VLCFA/ trihydroxy LCB ceramides stimulates growth and cell division. Overexpression of *LOH2* also caused the accumulation of sphingolipids with C16

fatty acid/ dihydroxy LCB ceramides, which lead to the accumulation of salicylic acid and induction of programmed cell. Overall, the findings demonstrated that overexpression of Arabidopsis ceramide synthases results in different physiological and metabolic phenotypes.

In other studies on ceramide synthases, *loh1* knockout plants had spontaneous cell death apparently initiated by improved expression of pathogenesis-related gene (PR-1) under short-day conditions. Free trihydroxy sphingoid bases in ceramide and GlcCer with C16 fatty acid were significantly raised while species with C20–C28 fatty acids were decreased. These data recommend that spontaneous cell death in the *loh1* line is activated either by the accumulation of free trihydroxy sphingoid bases or ceramide species with C16 fatty acid (Ternes et al., 2011).

Zheng et al. (2017) reported that under the nutrient starvation condition alkaline ceramidase (ACER) mutant, *acer-1* plants exhibited considerable phenotypic changes, such as higher accumulation of anthocyanin and the growth of primary root length was reduced. ACEROX-1-GFP-ATG8e, the plant overexpressing alkaline ceramidase, which is producing higher content of LCB's, exhibited higher amount GFP fluorescent autophagosomes under the stress conditions than WT, and the *acer-1* mutant exposed significantly fewer autophagosome-like structures compared with WT. *A. thaliana* neutral ceramidase (NCER) mutants *ncer-1*, *ncer-2* and *ncer-3* were obtained by T-DNA insertion. The *ncer-1*, mutant exhibited early leaf senescence and an increase in hydroxyceramides. Plants with *ncer-2* showed spontaneous cell death accompanied by an increase in LCB t18:0 when compared with Col-0 and *ncer-1* (Zienkiewicz et al., 2019).

2.5 Plant secondary metabolites

Plants produce a great number of organic compounds as intermediates and products, which do not participate directly in development, growth or the reproduction of the organism. These substances are conventionally referred as secondary metabolites (Tiwari and Rana, 2015). Secondary metabolites have various functions, including stimulatory and inhibitory effects on enzymes as coenzymes, signaling, defense and interactions with other organisms (Jain et al., 2019). Terpenes, phenolics and nitrogen-containing compounds are three chemically distinct groups of plant secondary metabolites (Pagare et al., 2015).

2.6 Phenolics

A plant contains phenolics, phenol with a hydroxyl group on an aromatic ring as one of the types of secondary metabolites (Figure 2.5 A) (Lin et al., 2016). According to their chemical structures, they are divided into sub groups, such as phenolic acids (hydroxybenzoic and hydroxycinnamic acids), flavonoids (flavanone, flavanol, flavonols, isoflavones, anthocyanidins, tannins), stilbenes (resveratrol) and lignans (Pandey and Rizvi, 2009).

2.7 Flavonoids

The flavonoids are an enormous class of phenolic natural products. The structural skeleton of it includes a chroman ring possess an aromatic ring on position 2, 3 or 4 (Figure 2.5 B). It can be divided into six groups according to differences in the pyran ring (ring C) such as flavone, isoflavone, flavan-3-ol, flavanone, anthocyanidin and flavonols. In each family, pattern of hydroxylation and methylation of rings A and B differ (Figure 2.5 C) (Panche et al., 2016; Jakimiuk et al., 2021).

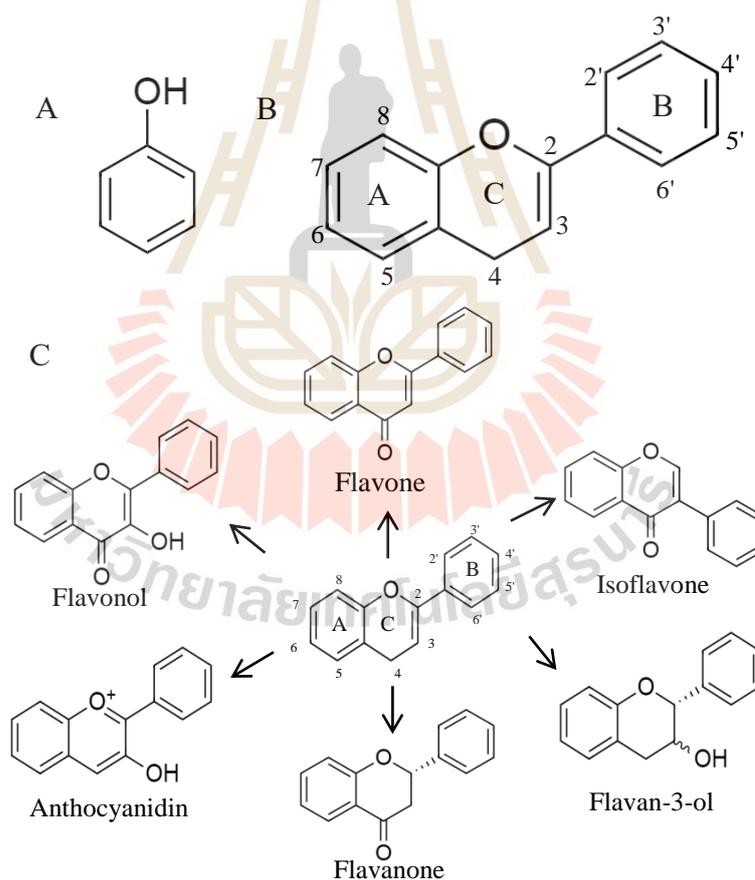


Figure 2.5 Phenolic compound and their derivatives. A. Phenolic ring, B. Structural skeleton of flavonoids, C. Basic structure of flavonoids (Nabavia et al., 2020).

2.8 Synthesis of flavonoids

There are two distinct biosynthetic pathways in plants for the synthesis of flavonoid-based compounds, includes the shikimic acid pathway and phenylpropanoid pathway (Figure 2.6).

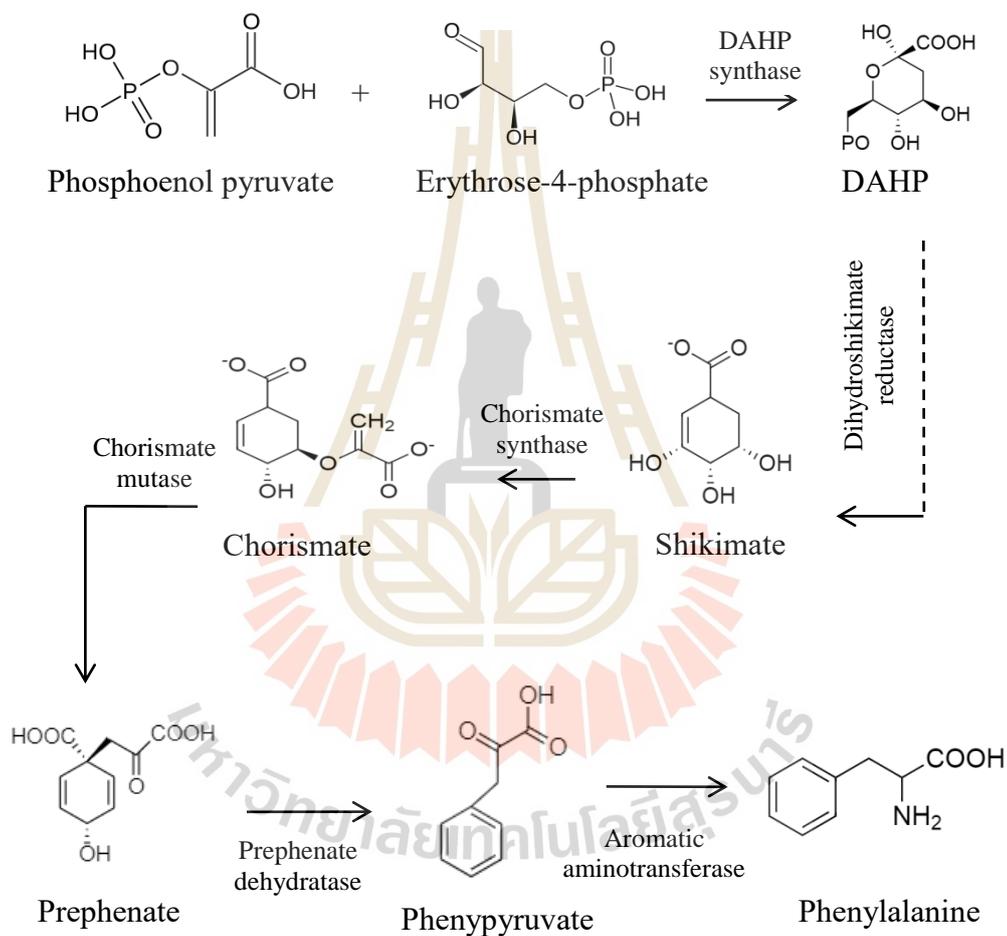


Figure 2.6 Shikimic acid pathway.

Initially, erythrose-4-phosphate from the pentose phosphate pathway (PPP) along with phosphoenol pyruvate from the glycolysis, are used to synthesize 9-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) by DAHP synthase, from which 3-dehydroshikimic acid (3-DHS) is formed by removal of phosphate and water by successive enzymatic reactions. Shikimate dehydrogenase catalyzes shikimic acid by a hydrogenation reaction (Tohge et al., 2013).

The next key intermediate in the shikimic acid pathway is chorismate. Further, phenylalanine is produced by rearrangement of the aliphatic chain, oxidative decarboxylation, dehydration and transamination of chorismate. In phenylpropanoid pathway the primary substrate for the synthesis of phenolic compound is phenylalanine (Herrmann and Weaver 1999; and Herrmann 1995).

In the phenylpropanoid pathway (Figure 2.7) L-phenylalanine is converted into *p*-coumaric acid by two subsequent reactions. First, phenylalanine ammonia lyase (PAL) is involved in deamination of L-phenylalanine and hydroxylated at the 4th position by cinnamate-4-hydroxylase (C4H). *p*-coumaric acid is combined to coenzyme A (CoA) by 4-coumarate CoA ligase to form a *p*-coumaroyl CoA. The formation of the flavonoid skeleton is the initial step in the biosynthesis of flavonoids, which occurs by condensing three molecules of malonyl-CoA and one molecule of *p*-coumaroyl-CoA in the presence of chalcone synthase to give chalcone.

Chalcone is subsequently isomerized by the enzyme chalcone flavanone isomerase (CHI) to yield the flavanone. Flavanone is converted to flavone by flavone synthase (FNS). Further, isoflavanone is synthesized by two subsequent reactions, first migration of aryl group followed by dehydration of flavanone by isoflavanone synthase and 2-hydroxy isoflavanone dehydratase (HID), respectively.

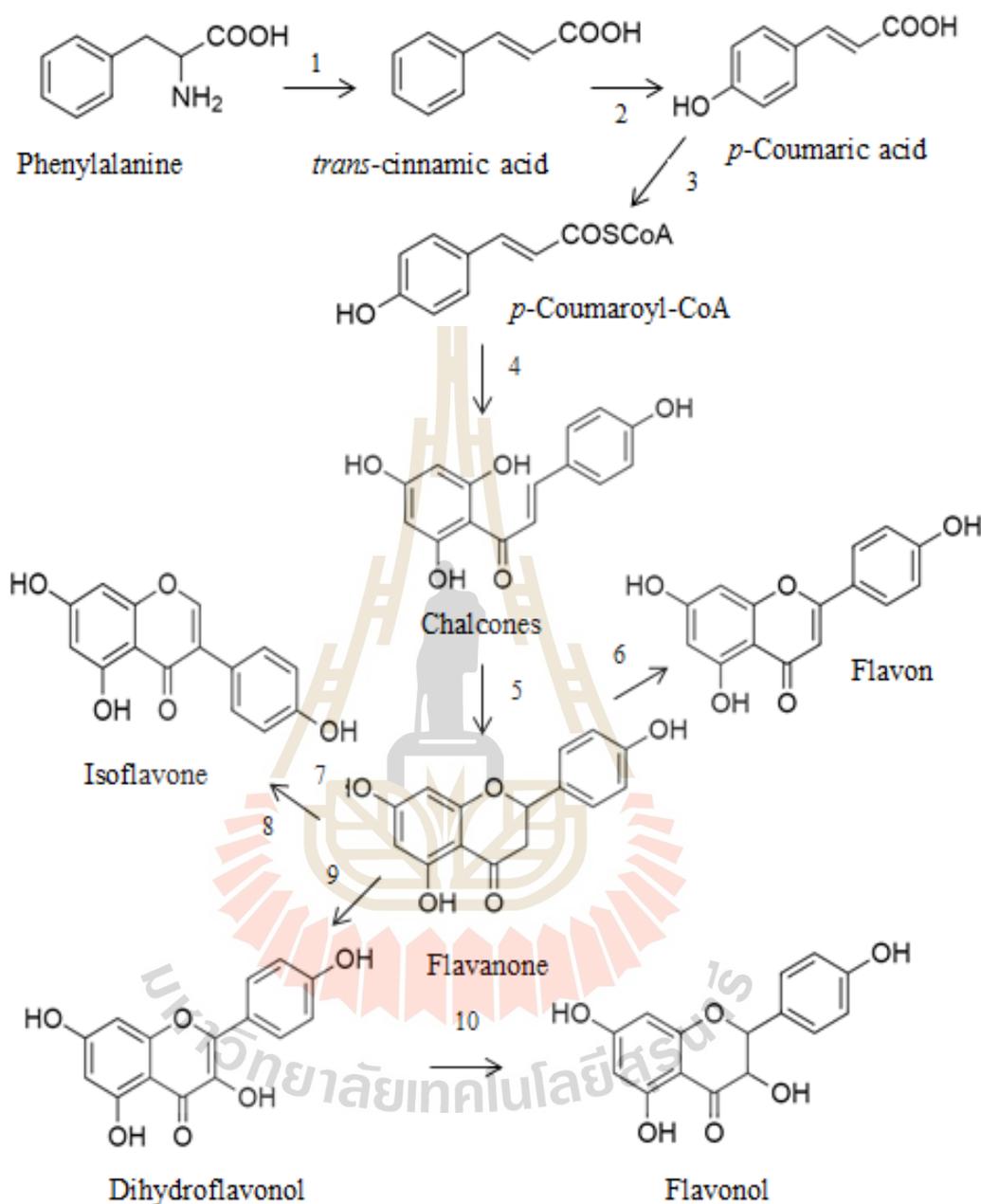


Figure 2.7 Flavonoid biosynthesis by the phenylpropanoid pathway. The enzymes catalyzing the reactions shown are: 1. phenylalanine aminolyase, 2. Cinnamate 4 hydroxylase, 3. 4-coumaroyl-coenzyme A ligase, 4. chalcone synthase, 5. chalcone flavanone isomerase, 6. flavone synthase, 7. isoflavanone synthase, 8. 2-hydroxy isoflavanone dehydratase 9. Flavanone 3 hydroxylase, 10. flavonol synthase.

A hydroxyl group is added to the ring C of flavanone at C-3 by flavanone-3-hydroxylase (F3H) to synthesize dihydroflavonols. Dihydroflavonol is oxidized by flavonol synthase to produce flavonols (Santos et al., 2017; and Nanda et al., 2016).

Glycosylation of flavonoids is occurs at the C-3, C-5 and C-7 positions of flavonoid aglycones. Flavonoid glycosyltransferase (FGT) catalyzes the glycosylation of flavonoid skeletons by attachment of Sugar moieties to flavonoid aglycones (Saito et al., 2013).

2.9 Flavonoids functions in plants

A variety of flavonoid derivatives aid various and important roles in plant reproduction, structural integrity, internal regulation of plant cell physiology and signaling, and UV photoprotection. Furthermore, it act as key chemical modulators of plant communication with insects and microbes, either as attractants or repellants, as phytoalexins against pathogens and herbivores, and as attractants to pollinators via flower color (Mandal et al., 2010). The biological functions of flavonoids are linked to their potential cytotoxicity and their capacity to interact with enzymes through protein complexation. Some flavonoids provide stress protection, for example, chelating metals that generate ROS via the Fenton reaction as well as acting as scavengers of free radicals such as reactive oxygen species (ROS) (Williams et al., 2004).

Two mutants lines of transparent testa 7 (tt7-1 and tt7-2) which is defected in flavonoid 3'-hydroxylase (F3'H) involved in quercetin biosynthesis, formed less numbers of lateral roots and tt7-2 had higher levels of kaempferol within lateral root than WT. Less fluorescence of a superoxide-selective probe was found within the

primordia of *tt7-2* compared with WT. These results support the role of kaempferol in quenching of ROS (Chapman and Muday, 2021). Studies in chalcone isomerase deletion mutants (*chi*) and anthocyanin synthase deletion mutant (*ans*) under the short-term high light (STHL) reveal the major photoprotection mechanism. After STHL treatment, photosynthetic pigment damage was detected in *chi* plants. The highest degree of Rubisco damage was observed in *chi*, followed by Col, and with least damage in *ans*. The results showed that *ans* presented advanced resistance to STHL treatment than Col, while *chi* with no anthocyanin accumulation had the lowest resistance (Yu et al., 2021).

Photooxidative stress caused by photosynthetically active radiation (PAR) and UV radiation in leaves of *Tilia platyphyllos*, showed that sun leaves contained 4.2 times more flavonoids than shade leaves as well the ratio of quercetin: kaempferol was greater in the former analysis. Sun leaves also contained significant amounts of myricetin, and only traces of amount was detected in shade leaves. In addition glycosylated quercetin and myricetin were much better antioxidants than kaempferol glycosides (Majer et al., 2014).

2.10 Flavonoid breakdown in plants

Since, most flavonoids are glycosylated; the first step in their breakdown is often deglycosylation. Recombinant BGLU15 (At2g44450), an *Arabidopsis* GH1 enzyme, displayed the highest catalytic efficiency for kaempferol and quercetin 3-O- β -glucoside-7-O- α -rhamnoside (K3G7R and Q3G7R) yielding their respective 7-O-rhamnosides as products. Expression of BGLU15 was 300% higher within 1 day of the recovery from nitrogen deficiency and low temperature (NDLT), a condition that

increases flavonoid levels, relative to control plants (Roepke and Bozzo, 2015). Shoots of *bglu15* mutants contained negligible K3G7R and Q3G7R hydrolase activity, whereas this activity increased by 223% within 2 days of NDLT recovery in wildtype plants. Thus, BGLU15 is essential for catabolism of flavonol 3-O- β -glucoside-7-O- α -rhamnosides in *Arabidopsis* (Roepke et al., 2017).

2.11 Glycoside hydrolases

Carbohydrates are the most abundant biomolecules on Earth. They play various roles in living organisms, ranging from structural elements in cell walls (e.g., cellulose, xyloglucans and arabinoxylans) to cell–cell recognition processes of non-photosynthetic cells (Vigerust and Shepherd, 2007). A large group of enzymes is accountable for the synthesis, degradation, and modification of carbohydrates (Larsbrink et al., 2014). Those enzymes are categorized into glycoside hydrolases (GHs), polysaccharide lyases (PLs) or glycosyltransferases (GTs) and esterase which constitute approximately 1-2% of the genome of any organism. (Lombard, 2014; Hart and Copeland., 2010; and Davies, et al., 2005).

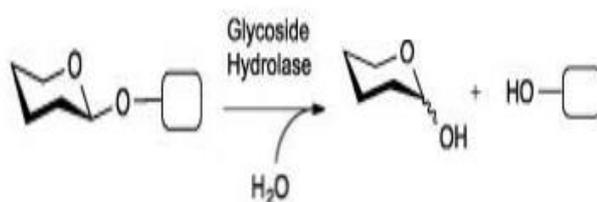


Figure 2.8 Schematic representation of the glycoside hydrolase reaction (O'Neill and Field, 2015).

Glycoside hydrolases (GH – 3.2.X.Y) are pervasive enzymes that catalyze the hydrolysis of β -glycosidic bonds in glycosides, glycans and glycoconjugates (Vuong and Wilson, 2010). The simplest classification of GHs is that based on their substrate specificities, i.e., their enzyme commission (EC) numbers.

Of interest, EC.3.2.1.X cleaves O-glycosidic bonds to yield monosaccharide or other reduced size carbohydrate from their polymers or glycosides, where X represents the substrate specificity (and, in some cases, also the molecular mechanism or the type of linkage e.g. β -1,4 or β -1,3; etc) (Coutinho et al., 2003). O-GHs (EC 3.2.1.X) are found in all domains of living organisms, in which they play a diverse range of biologically significant functions (Vuong and Wilson, 2010). Especially, selective hydrolysis of glycosidic bonds plays significant and vital roles in the regulation of cells. For example, hydrolysis reactions are crucial for energy uptake, turnover of signaling molecules, and cell wall expansion and degradation. In humans, heritable deficiencies in glycoside hydrolases cause diseases, for example, lactose intolerance and a number of lysosomal storage diseases, which are among the most frequent genetically based syndromes in humans (Henrissat and Davies, 2000).

2.12 Mammalian Glucosylceramidases

GlcCer is degraded into glucose and ceramide by glucosylceramidases, also designated glucosylcerebrosidases, (GBA) EC 3.2.1.45 and also EC 3.2.1.21 to indicate their general β -glucosidase activity, which are found in animals, plants, fungi, archaea, and eubacteria. During initial studies of mammalian β -glucosidases, GBA1 and GBA2 are referred as glucosyl cerebrosidase (GBA) and bile acid β -glucosidase, but have recently been differentiated as glucosylceramidases GBA1 and GBA2.

GBA1 and GBA2 are also known as lysosomal acid β -glucosidase and non-lysosomal β -glucosidase. A third enzyme, sometimes called GBA3, is known as cytosolic β -glucosidase. GBA1, GBA2 and GBA3 are members of different families of sequence-related glycoside hydrolases designated GH30, GH116 and GH1, respectively (Charoenwattanasatien et al., 2016; Körschen et al., 2013; Lombard et al., 2014; Matern et al., 2001; Cobucci-Ponzano et al., 2010). These three enzymes have similar enzymatic activity, but little or no structural or sequence homology and their subcellular locations are different from each other (Dekker et al., 2011; Hayashi et al., 2007).

As noted above, GBA1 belongs to the GH30 family, which has been reported only in bacteria, fungi and animals and not in plants (<http://www.cazy.org/>), which suggests that the function of GBA1 may be replaced by GBA2 homologues or other GH enzymes in plants. In humans, mutations in GBA1 have been identified as the cause of Gaucher disease (GD) and mutations in GBA2 may cause Hereditary Spastic Paraplegia (HSP), Autosomal Recessive Cerebellar Ataxia (ARCA) with spasticity and Marinesco-Sjogren like syndrome (Hruska et al., 2008; Votsi et al., 2014; Citterio et al., 2014). The pathological features of HSP, ARCA and Marinesco-Sjogren-like syndrome include cataracts, peripheral neuropathy, and mild to moderate mental impairment, hearing loss, hypotonia, muscle weakness, short stature and various skeletal abnormalities, including scoliosis (Haugarvoll et al., 2017).

GD, the most common lysosomal storage disorder, is characterized by accumulation of glucosylceramide in lysosomes of monocytes (Gaucher cells) (Yildiz et al., 2006). β -Glucosidase activities (GBA1 and GBA2) were analyzed between the PM and total cell lysates of human fibroblast cell lines from control and the patients

with GD clinical types 1, 2, and 3. The GBA1 activities in both PM and total lysate of GD fibroblasts were lower than control cells. GBA2 activities were increased in GD cells than in control cells, and the activity differed among the three GD types. These results suggested that GBA2 was induced in the absence of GBA1 activity in humans (Aureli et al., 2012). In contrast, Schonauer et al. (2017) reported that the GBA1-dependent down regulation of GBA2 activity in patients with Gaucher disease. They proved a negative feed-back loop, in which sphingosine, the cytotoxic metabolite accumulating in Gaucher cells through the action of GBA2, directly binds to GBA2 and inhibits its activity, preventing further sphingosine accumulation and cytotoxicity.

In general, breakdown of GlcCer mainly occurs in the lysosomes by GBA1. Reduction in GBA1 activity in lysosomes, leaks its substrates outside the lysosomes, where GBA2 located at cytosolic face of the ER and cis-Golgi may take over the reaction (Schonauer et al, 2017). ER and Golgi GBA2 have been proven to be a non-integral membrane protein associated with the membrane, with the bulk of the protein in the cytoplasm (Körschen et al., 2013). Accumulation of GlcCer was observed in GBA2 knockout mice, which exhibit impaired male fertility due to aberrant sperm formation and delayed liver regeneration after partial hepatectomy (Yildiz et al., 2006; Gonzalez-Carmona et al., 2012).

The inhibitory studies on the non-lysosomal glucocerebrosidase, demonstrated that the GBA2 was not inhibited by low concentrations of conduritol B-epoxide, which is a potent inhibitor of GBA1, but it was markedly inhibited by taurocholate (van Weely et al., 1993). In order to characterize a potential role of GBA2 in GD, functional and genetic approaches were proposed (Yildiz et al., 2013). GlcCer levels in the liver, spleen and brain of GBA2-deficient mice and protein and mRNA

expression of GBA2 in GBA1-deficient murine fibroblasts were analyzed. A significant accumulation of GlcCer was found compared to wild-type controls in all three organs studied. In addition, in GBA1-deficient murine fibroblasts, protein and mRNA levels of GBA2 were significantly increased. Genome-wide single nucleotide polymorphism (SNP) genotyping was performed to investigate autosomal recessive cerebellar ataxia with cataracts and mental retardation in two Norwegian families with a Genome Wide Human SNP array and potentially pathogenic mutations were seen in the GBA2 gene (Haugarvoll et al., 2017). The GBA2 characterization in the patients showed that the less activity of GBA2 is sufficient to raise the levels of glucosylceramide to similar levels as those seen in GD.

Both GBAs thus play a vital role in human sphingolipid metabolism. A gene (GenBank accession No. KM677956) encoding a GH116 protein from *Thermoanaerobacterium xylanolyticum* (GenBank accession No. AEF18218.1), which is 37% amino-acid sequence identity with human glucosylceramidase 2 (GBA 2), was identified and the protein expressed in *E.coli* (Sansenya et al., 2015). The purified TxGH116 enzyme hydrolyzed the *p*-nitrophenyl (*p*NP) glycosides, *p*NP- β -D-glucoside, *p*NP- β -D-galactoside and *p*NP-N-acetyl- β -D-glucopyranoside, as well as cellobiose and cellotriose. Recombinant TxGH116 protein was also crystallized. The structural determination of TxGH116 by X-ray crystallography allowed the identification of the glucoside binding and active site residues of both TxGH116 and GBA2 (Charoenwattanasatien et al., 2016). The structural solution of TxGH116 also allowed a structural explanation of the effects of human GBA2 mutations. Mutagenic analysis of TxGH116 and structural modeling of GBA2 was employed to provide the structural and functional rationale for the effects of mutations of GBA2. TxGH116

was subjected to the following missense mutations F419V, D594H, R630W, G683R and R873H, which correspond to known pathogenic human GBA2 mutations.

In type 1 Gaucher disease (GD1), GBA deficiency causes the accumulation of two key sphingolipids, GlcCer and glucosylsphingosine (LysoGlcCer), within the lysosomes of mononuclear phagocytes, leading to a bone-formation defect. Conditional deletion of GBA2 in GD1 mice enhanced elevation of GlcCer and LysoGlcCer, but normalized the bone volume and bone formation rate. Nanomolar concentrations of sphingosine revealed a strong inhibition of osteoblast viability. The results from this mouse model of GD1 suggested that sphingosine can act as a osteoblast toxin, and GBA2 may be a useful target for the development of therapeutic inhibitors to certain consequences of GD1 (Mistry et al., 2014).

Sorli et al. (2012) investigated whether GBA2 can modulate human melanoma cell growth and death. GBA2 expression was quantified in melanoma cells by RT-qPCR. The results suggested that GBA2 is down-regulated in melanoma. Over expression of GBA2 affected endogenous sphingolipid metabolism by promoting glucosylceramide degradation (which decreased by 78%) and ceramide generation. Following this, decreased anchorage-independent cell growth, and reduced *in vivo* tumor growth (by 40%) was observed. All these events were not observed when a catalytically inactive GBA2 was expressed. This study demonstrated that overexpression of GBA2 may lead to an increase in growth inhibiting products or decrease in growth promoting compounds, thereby resulting in antitumor activity.

Mammalian cytosolic β -glucosidase is sometimes referred to as GBA3, non-specific β -glucosidase, or broad-specificity β -glycosidase, and more recently as Klotho-related protein (Hayashi et al., 2007). GBA3 efficiently hydrolyzes

glycosylated flavonoids. The cell-free extracts from human liver and small intestine were analyzed for β -glucosidase activity towards (iso)flavonoid glycosides. Those cell free extracts deglycosylated various (iso)flavonoid glycosides, such as quercetin-4-glucoside, naringenin-7-glucoside, apigenin-7-glucoside, genistein-7-glucoside and daidzein-7-glucoside (Day et al., 1998). Németh et al. (2003) isolated flavonoid glycoside hydrolyzing enzymes from human liver and small intestine using chromatographic techniques. The results revealed that cytosolic β -glucosidase (CBG), from human small intestine mucosa is efficient in hydrolyzing flavonoid glycosides. GBA3 is able to degrade a variety of substrates with a β -D-glucose, β -D-galactose, β -D-xylose, or α -L-arabinose moiety linked to a hydrophobic group. Although GBA3 has been demonstrated to hydrolyze GlcCer-like substrates (Hayashi et al., 2007), this activity is extremely low compared to its activity on the plant-derived glycosides, and may not be physiologically relevant.

The mechanisms of GBA1, GBA2 and GBA3 were investigated by use of cyclophellitol-epoxide and aziridine activity-based probes (ABPs), which covalently modify the nucleophile of retaining glucosidases (Kallemeijn et al., 2014). The catalytic residues of retaining-exoglucosidases, were subjected to site-directed mutagenesis and substituted with nonionizable amino acids, such as glycine or glutamine, and sodium azide mediated rescue was employed to verify the acid/base and nucleophile. Sodium azide-mediated rescue results in retention of stereochemistry in the case of acid/base mutant (azidoglucoside) and inversion of stereochemistry when the nucleophile is lacking. By these means, the catalytic acid/base and nucleophiles were verified for all three human intracellular β -glucosidases and the catalytically active enzymes could be fluorescently labeled with the covalent

inhibitors. Human GBA2 isoform 1 showed high activity to hydrolyze 4-methylumbelliferyl- β -D-glucosylpyranoside (4MUG) in COS-7 cell extracts (Jatoorathawichot et al., 2020).

2.13 Plant Glucosylceramidases

In *Arabidopsis thaliana*, GH116 enzymes are encoded by four genes: At1G33700, At3G24180, At4g10060 and At5G49960; these proteins share sequence similarity with the human GBA2-type glucosylceramidase. Among the four, one of the genes, At4g10060 (AtGCD3), which is located in the plasma membrane and ER was expressed in E.coli (Dai et al., 2020). The TLC and HPLC analysis confirmed the hydrolysis of C6-NBD-GlcCer by recombinant AtGCD3 but with a low k_{cat} and K_m value. Inhibitory analysis with conduritol-B-epoxide (CBE) and N-butyldeoxynojirimycin (NB-DNJ) showed an IC_{50} of 1 mM, and 5 μ M respectively. This result indicated like GBA2, AtGCD3 was relatively insensitive to CBE.

When AtGCD3 was tested on glucosylceramides from WT *Arabidopsis*, it preferred to hydrolyze Glc-t18:1-h16:0 and Glc-d18:1-h16:0 most and Glc-d18:0-h16:0 less. By contrast, the amounts of GlcCer with very long acyl-chain (fatty acids >20, VLACFA) incubated with AtGCD3 showed no significant changes. Sphingolipids analysis by electrospray ionization-MS/MS in the T-DNA insertion mutants of AtGCD3, SALK_099838 (*gcd3-1*) and SALK_019663 (*gcd3-2*) mutants showed no significant changes compared with wild type, indicating that other redundant enzymes might hydrolyze GlcCer in *Arabidopsis*.

CHAPTER III

MATERIALS AND METHODS

3.1 Materials

3.1.1 Plasmids, bacterial strains and wild type and mutant plant lines

The plasmids containing cDNA encoding the AtGH116 proteins including clones RAFL22-05-G12 for At1GH116, RAFL09-37-G13 for At3H116, RAFL09-27-E09 for At4GH116 and RAFL09-66-C11 for At5GH116, were acquired from the RIKEN Biological Resource Center, Ibaraki, Japan. A modified pET30 expression vector from Dr. Rung-Yi Lai was used for the AtGH116 proteins. *E. coli* strain NEB5 α as used for cloning, while the BL21(DE3) and Rosetta-gami(DE3) strains were used for protein expression. *Arabidopsis* GH116 mutant seeds for the lines SALK_143139, SALK_039506, SALK_117180, SALK_099838, SALK_010803, SALK_039511, and SALK_057433, along with wild type Columbia-0 were generously provided by Prof. Jong-Seong Jeon, Kyung Hee University, Korea and were originally purchased from the purchased from the Salk Institute Genomic Analysis Laboratory, USA. Another Columbia-0 seeds were provided by Assoc. Prof. Dr.Varodom Charoensawan of Mahidol University.

3.1.2 Chemicals and reagents

2-Deoxy-2-fluoro- β -D-glucoside, 2-mercaptoethanol, *p*-coumaric acid, acetic acid, acetonitrile, acrylamide, agarose, ammonium chloride, ammonium persulfate, ampicillin, apigenin-7-O- β -D-glucoside, arbutin, bacto-agar, Bis-Tris, bovine serum

albumin, bromophenol blue, C16 β -D-glucosylceramide, calcium chloride dihydrate ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$), cellobiose, chloramphenicol, chloroform/isoamyl alcohol (24:1), Coomassie brilliant blue R250, daidzin, D-amygdalin, sodium chloride (NaCl), deoxyribonucleotide triphosphate (dATP, dGTP, dCTP, and dTTP), dhurrin, dichlorodimethylsilane, digalactosyldiacylglycerol, dipotassium hydrogen phosphate (K_2HPO_4), di-sodium hydrogen phosphate, dithiothreitol, DNase I, esculin, ethanol, ethidium bromide, ethylenediaminetetraacetic acid disodium salt (EDTA), apigenin, formic acid, genistin, gentibiose, glycine, glycitin, glycerol, gossypin, helicin, hesperitin, HPLC grade water, hydrochloric acid, imidazole, indoxyl β -D-glucoside, isopropanol, isopropyl- β -D-thiogalactopyranosid (IPTG), kaempferol, kaempferol-3-O- β -D-glucoside, kaempferol-7-O- β -D glucosides, kanamycin, luteolin-7-O- β -D glucosides, lysozyme, magnesium sulfate heptahydrate ($\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$), mangiferin, melibiose, methanol, methyl β -D-glucoside, sodium hydroxide, ethyl acetate, sophorose, monogalactosyldiacylglycerol, sucrose, sulfuric acid, N,N'-methylene-bis-acrylamide, N,N',N'',N'''- tetramethylethylenediamine (TEMED), glucose, naringin, syringin, *n*-heptyl β -D-glucoside, acetic acid, *n*-octyl- β -D-thio-glucopyranoside, T4-DNA ligase, *ortho*-nitrophenyl (*o*NP)- β -D galactopyranoside, *o*NP- β -D-glucopyranoside, *p*-coumarol glucoside, peptone, *Pfu* DNA polymerase, phenylmethylsulfonylfluoride, phlorizin, dimethyl sulfoxide, *para*-nitrophenyl(*p*NP)- β -D-glucopyranoside, *p*NP-N-acetyl- β -D-glucosamine, *p*NP- α -D-galactopyranoside, *p*NP- α -D-glucopyranoside, *p*NP- α -D-mannopyranoside, *p*NP- α -L-arabinofuranoside, *p*NP- α -L-arabinopyranoside, *p*NP- β -D-cellobioside, *p*NP- β -D-fucopyranoside, *p*NP- β -D-galactopyranoside, *p*NP- β -D-mannopyranoside, *p*NP- β -D-xylopyranoside, potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), protein MW

makers, quercetin-3-O- β -D-glucoside, quercetin-7-O- β -D glucoside, salicin, sodium acetate anhydrous, sodium carbonate (Na₂CO₃), sodium dihydrogen phosphate dihydrate (NaH₂PO₄·2H₂O), sodium dodecyl sulfate (SDS), soybean trypsin inhibitor, *Taq* DNA polymerase, thin-layer chromatography silica gel 60 aluminum F₂₅₄ plates, Tris base, Triton X-100, tryptone, yeast extract were purchased from a variety of suppliers.

3.1.3 Oligonucleotides primers

Oligonucleotides were synthesized by Bio Basic Inc. (Canada) and are shown in Tables 3.1 and 3.2.

Table 3.1 Oligonucleotides primers for cloning of *Arabidopsis thaliana* glycoside hydrolase family 116 (AtGH116) into the pET30 vector. Overlapping region specific to the vector is shown in lowercase and nucleotides specific to coding regions are in uppercase.

Primer name	Sequence (5'-3')
AT1GpET30For	gagaacctgtacttccagggcggtggcATGCAAACGTTCTGAAGAC
AT1GpET30Rev	gctttgtagcagccggatcCTATAGGCGAATCGTCTTGAGG
AT3GpET30For	gagaacctgtacttccagggcggtggcATGGTTGGAGCAACTTTGTTTC AC
AT3GpET30Rev	gctttgtagcagccggatcTTAGCAAGAACAGCTTAATGCTGAA
AT4GpET30For	gagaacctgtacttccagggcggtggcATGGAGAAGAATGGTCACACG G
AT4GpET30Rev	gctttgtagcagccggatcTTACAAGCGGAGAGTCTTGAGAAC

Table 3.1 Oligonucleotides primers for cloning of *Arabidopsis thaliana* glycoside hydrolase family 116 (AtGH116) into the pET30 vector. Overlapping region specific to the vector is shown in lowercase and nucleotides specific to coding regions are in uppercase (Continued)

Primer name	Sequence (5'-3')
AT5GpET30For	gagaacctgtacttccagggcggtggcATGTTTGAGGAAAAGATCATG GATATTG
AT5GpET30Rev	gctttgtagcagccggatcCTATGACATCATTCTCCTGCAAGTATA G
pET30For	gatccggctgctaacaag
pET30Rev	ccctggaagtacaggttctc
T7 promoter	CGACTCACTATAGGGGAATTG
T7 terminator	CGTTTAGAGGCCCAAGG

Table 3.2 Oligonucleotides primers for genotyping of *Arabidopsis* mutant lines plants with the respective locus number.

Locus number	Primer name	Sequence (5'-3')
At1G33700	SALK143139LP	TCCTGGAGGTGTGAAACACTC
	SALK143139RP	ACAAGGACATGAGGAGACACG
At3G24180	SALK039506LP	GCGCATGTGAATAGATAAGCC
	SALK039506RP	GCTTATTTGGTGGGAGACTCC
	SALK117180LP	TGACAATTCAGCACAAAGAGC
	SALK117180RP	CATGTGAATGAGCCATTCATG
At4G10060	SALK099838LP	AGTTTCCCTTTGACCG
	SALK099838RP	TCTCTCGAACAAAATGGATGG
	SALK010803LP	CATGTTCCAGTCCCATGATTC
	SALK010803RP	TTAATTCATTATGCGGAACGG
At5G49900	SALK039511LP	GGGTGAAACTTGACGACAAAC
	SALK039511RP	TTCCAGCTGGTACACCTATGG
	SALK057433LP	AATTTGTGCTGCAGCATCAC
	SALK057433RP	AAAGCAGCTTTCCAGTGTCTG
T-DNA insert	T-DNA line BP	ATTTTGCCGATTTCCGGAAC

3.2 General methods

3.2.1 Preparation of *E. coli* strains NEB5 α , BL21(DE3) and Rosetta-gami(DE3) competent cells

Glycerol stocks of NEB5 α , BL21(DE3) and Rosetta-gami(DE3) strains were streaked on Luria-Bertani (LB) agar plates and incubated at 37°C for 16-18 h. A single colony was picked and inoculated into 5 ml of LB broth and incubated with shaking at 37°C, with 200 rpm, rotary shaking for 16-18 h as the starter culture. One milliliter of starter culture was added to 100 ml of LB broth and shaken at 37°C, 200 rpm, until the optical density at 600 nm (OD₆₀₀) reached 0.4-0.6. The cell culture was chilled on ice for 10 min in a 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 CaCl₂ and centrifuged to collect the cell pellets again. Finally, the pellets were resuspended with 1 ml of 0.1 M CaCl₂ containing 15% glycerol and 50 μ l aliquots were stored at - 80°C

3.2.2 Transformation of plasmids into competent cells

An aliquot of frozen competent cells was thawed 5 min on ice, then 1 μ l of DNA plasmid (20-100 ng) was mixed with the thawed competent cells. The reaction was incubated on ice for 30 min. The plasmid was transformed into the competent *E. coli* cells by heating at 42°C for 45 s and quickly chilling on ice for 5 min. Two hundred microliters of LB broth were 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 antibiotic and incubated at 37°C overnight.

3.2.3 Plasmid isolation by the alkaline lysis method

A single colony containing plasmid on the plate was picked into 3-5 ml of LB with appropriate antibiotic and incubated at 37°C with shaking at 200 rpm for 16-18 h. The cultured cells were collected by centrifugation at 12,000 rpm for 1 min. The supernatant was discarded 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 mixed gently, chilled on ice for 3 min. 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 one volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added into the tube and thoroughly shaken by hand for approximately 20 s. The mixture was centrifuged at room temperature for 5 min at 13,000 rpm. The upper aqueous phase was carefully removed into a fresh tube and 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 ethanol was removed by speed vacuum. Then, the DNA pellet was re-suspended 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 chilling 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 the ethanol solution was removed and the tube was dried by speed vacuum. Finally, the DNA was re-dissolved with 30 μ l of TE buffer.

3.2.4 Vivantis Plasmid DNA Extraction

The Vivantis plasmid DNA extraction kit was used to purify recombinant plasmid DNA according to the manufacturer's instructions. A single colony was picked in 5-10 ml LB broth and grown with appropriate antibiotics for overnight (12-16 hours) at 37°C with shaking. The cultured cells were collected by centrifugation at 10,000 rpm for 1 min. The supernatant was decanted completely. Cells were resuspended in 250 μ l S1 buffer (RNase A) by vortexing or pipetting. Two hundred fifty microliters of S2 (0.2 N NaOH, 1% SDS) was added to the resuspended cells and they were gently mixed by inverting tube several times (4-6 times) to obtain a clear lysate. The alkaline lysate was incubated on ice or at room temperature for 5 min. The lysate was neutralized by addition of 400 μ l of Buffer NB and gentle mixing by inverting the tube several times (6-10 times) until a white precipitate forms. The solution was centrifuged at 14,000-16,000 rpm for 10 min. After centrifugation, 650 μ l of supernatant was transferred into a column assembled in a clean collection tube and centrifuged at 10,000 rpm for 1 min. The flow-through solution was discarded. The column was washed with 650 μ l Wash Buffer and centrifuged at 10,000 rpm for 1 min. The flow-through solution was discarded and the column was dried by centrifugation at 10,000 rpm for 1 min. Lastly, the column was placed in a new 1.5 ml microtube and 50 μ 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. The DNA was stored at 4°C or -20°C.

3.2.5 DNA analysis by agarose gel electrophoresis

The purified plasmids, amplified PCR products and genomic DNA were analyzed by gel electrophoresis. A 1% agarose gel was prepared in 1 X TAE buffer (40 mM Tris-HCl, 40 mM acetic acid, 1 mM EDTA, pH 8.0). DNA samples were mixed 5:1 with 6X loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol) and applied to the gel wells. Electrophoresis was performed at a constant 100 V for 30-40 min. After electrophoresis, the gel was stained in 0.1 mg/ml ethidium bromide solution for 30 s, and destained with distilled water for 20 min. The DNA bands in the gel were visualized by UV light transillumination with a Fluoro-S™ MultiImager (Bio-Rad). The size of DNA fragments were estimated by comparison with 1 kb DNA markers (Thermo Fisher).

3.2.6 Purification of DNA bands from gels

The desired DNA bands separated on agarose gel electrophoresis were purified with the Thermo Scientific GeneJET Gel Extraction Kit. The agarose gel containing the target DNA band was excised with a blade cutter and transferred to a micro centrifuge tube. An equal volume of binding buffer was added to the sliced gel. The gel mixture was incubated at 50-60°C for 10 min or until the gel slice was completely dissolved. Tubes were mixed and vortexed every few minutes to enhance the melting process. Up to 800 µL of the solubilized gel solution was transferred to the GeneJET purification column. The purification column was centrifuged for 1 min. The flow-through solution was discarded and column was placed back into the same collection tube. Then 100 µL of binding buffer was added to the GeneJET purification column and centrifuged for 1 min. The flow-through solution was discarded and the column was placed back in the tube. After 700 µL of wash buffer was added, the tubes was

again centrifuged for 1 min. The flow-through solution was discarded and the column was placed back into the same collection tube. The GeneJET purification column was emptied by centrifuging for an additional 1 min. The GeneJET purification column was transferred into a clean 1.5 mL micro centrifuge tube. Fifty microliters of elution buffer was added to the center of the purification column membrane, and it was centrifuged for 1 min, then the DNA was stored at 4°C or -20°C until use for further experiments.

3.2.7 SDS-PAGE electrophoresis

The protein profiles and the apparent molecular weights of proteins in various fractions were determined by polyacrylamide gel electrophoresis with sodium dodecyl sulfate (SDS-PAGE), as described by Laemmli (1970). The SDS-PAGE 10% separating gel consisted of 10% (w/v) acrylamide, 375 mM Tris-HCl, pH 8.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED, while the 5% stacking gel consisted of 5% (w/v) acrylamide, 125 mM Tris-HCl, pH 6.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED. Protein samples were denatured by boiling for 5 mins with 5:1 ratio with 6X loading buffer (50 mM Tris-HCl, pH 6.8, 10% SDS, 0.2 mg/ml bromophenol blue, 50% glycerol, 20% β -mercaptoethanol). Ten microliters of protein samples were loaded into sample wells, and electrophoresed through the polymerized gel at 150 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 with staining solution containing 0.1% (w/v) Coomassie Brilliant Blue R250, 40% (v/v) methanol, and 10% (v/v) acetic acid in water for 30-60 mins 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 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 α -lactalbumin (14.0 kDa).

3.2.8 Native gel electrophoresis

The native 10% separating gel consisted of 10% (w/v) acrylamide, 375 mM Tris-HCl, pH 8.8, 0.05% ammonium per sulfate and 0.05% TEMED. Protein samples were mixed with 5:1 ratio of 6X loading buffer (50 mM Tris-HCl, pH 6.8, 0.2 mg/ml bromophenol blue, 50% glycerol). Ten microliters of protein samples were loaded into sample wells, and electrophoresed through the polymerized gel at 150 V with Tris-glycine electrode buffer (50 mM Tris base, 125 mM glycine, pH 8.8) until the dye front reached the bottom of the gel plate. The gels were subsequently stained with 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 mins and destained with destaining solution [40% (v/v) methanol and 10% (v/v) acetic acid] for 1-2 h.

3.2.9 Determination of protein concentration

The absorbance of the protein fractions and purified proteins were measured by the method A_{280} spectrophotometrically by NanoDrop 2000 spectrophotometer (Thermo Scientific, MA, USA) and calculating with the following equation:

$$\text{Protein conc. (mg/ml)} = [\text{OD}_{280} / \text{extinction coefficient}] \times \text{dilution-fold} \times 1/\text{path length}$$

The protein concentration calculated from this absorbance with the extinction coefficient, which was calculated with the ProtParam program on the EXPASY website (www.expasy.org).

3.2.10 Extraction of genomic DNA from *Arabidopsis thaliana* tissues by the CTAB method

Two hundred milligrams fresh weight of leaf tissue was weighed and transferred into precooled mortar and pestle. An appropriate amount of liquid nitrogen was poured into it, the material was allowed to freeze and ground into powder. The extraction buffer (1 M Tris, pH 8.0, 5 M NaCl, 0.5 M EDTA and 5 g of CTAB) was preheated in a 65°C water bath and 1.2 ml of preheated extraction buffer was added to the tissue powder and vortexed for 5-10 s to mix thoroughly. The extraction mixture was incubated at 65°C for 30 min and mixed every 5-10 min by inverting the tubes. Insoluble debris was removed by centrifuging at 13,500 g for 10 min at room temperature (RT). The supernatant was transferred and 800 µl of phenol:chloroform:isoamyl alcohol (25:24:1,v/v/v) was added. The mixture was incubated for 20 min at RT and mixed gently by inverting. Phases were separated by centrifuging the mixture at RT for 10 min at 13,000 g. The upper aqueous phase was carefully transferred into a fresh tube containing 800 µl cold isopropanol. DNA was allowed to precipitate by mixing the tube by inverting and incubating at RT for 10 min. DNA was collected by centrifuging the mixture at 13,500 g for 10 min. The supernatant was removed and then the pellet was resuspended in 250 µl TE. Then, 2.5 µl of DNase - free RNase was pipetted into the sample mixture and it was incubated at 37°C for 30 min. Following that, 25 µl of 3 M sodium acetate was added and mixed well. Six hundred microliters of precooled ethanol was added, mixed and incubated at 20°C for 20 min to precipitate the DNA. The mixture was centrifuged at 13,500 g to collect the DNA. The pellet was washed with 500 µl of cold 70% (v/v)

ethanol and 70% ethanol was removed after centrifuging at 13,500 *g* for 10 min at RT. The pellet was resuspended in 25 μ l of water (Allen et al., 2006).

3.2.11 Extraction of genomic DNA with the BioFact™ kit

Ten to eighty milligrams of fresh plant tissue was ground in liquid nitrogen in a precooled mortar and pestle. The tissue sample was transferred into a 1.5 ml microfuge tube, and 350 μ l GD1 solution was added to the prepared sample. Five microliters of Proteinase K (20 mg/ml) and 2 μ l of RNase A (4 mg/ml) were added. The sample was vortexed for 1 min and incubated for 10 mins at 65°C, then 100 μ l PPB solution was added and the tube vortexed for 10 s. Plant tissue debris was removed by centrifuging at 13,000 *g* for 5 min. The supernatant was transferred into a new microfuge tube and 200 μ l of GB solution was added and mixed well by inverting 10-20 times. A column was prepared by assembling the spin column into a 2 ml collection tube. The column was activated by addition of 200 μ l Help B Buffer and centrifuged at 10,000 rpm for 30 s. The flow-through solution was discarded. The sample lysate was transferred into a spin column. Excess materials were removed by centrifuging at 7,000 rpm for 1 min, and the flow-through solution was discarded. The column was washed twice with addition 500 μ l WB (80% ethanol). The spin column was centrifuged at 13,000 *g* for 30 s for first time and 3 min for the second time. The flow-through solution was discarded. Residual liquid was removed from column by centrifuging at 13,000 *g*. A new 1.5 ml microfuge tube was attached to the column and DNA was eluted by addition of 50 - 100 μ l DNA Hydration Solution. The column was left for 1 min at RT. The DNA was collected by centrifuging at 13,000 rpm for 2 min. The concentration and purity of DNA was analyzed by measuring the 260nm and

280nm absorbance in a Nanodrop spectrophotometer and by electrophoresis. The DNA was stored at 4°C or - 20°C for further use.

3.3 Amplification, ligation and cloning of AtGH116 CDS cDNA

3.3.1 Amplification of AtGH116 CDS cDNA

The genes encoding At1GH116, AtGH116, At4GH116 and At5GH116 were amplified from the plasmid cDNA clones RAFL22-05-G12 (GenBank Accession number BP813105.1), RAFL09-37-G13 (GenBank Accession number AK226971.1), RAFL09-27-E09 (GenBank Accession number AV801203.1) and RAFL09-66-C11 (GenBank Accession number BT002482.1) respectively, which were purchased from RIKEN Biological Resource Center, Japan. The cDNA encoding the AtGH116 CDS were amplified individually with specific primers which contain nucleotide sequence overlapping the pET30 vector sequence, as shown in Table 3.1. The amplification reactions of coding sequence and linearization of vector were carried out with *Phusion* polymerase with the temperature cycling parameters shown in Table 3.3. The PCR products of 2.8 Kb to 3 Kb were checked by electrophoresis on a 1% agarose gel.

3.3.2 Ligation of ATGH116 CDS

The PCR products were purified with the Thermo Scientific GeneJET Gel Extraction Kit. The purified DNA fragments were incubated with linearized pET30 vector in the ratio of 3:1 along with 1X Gibson Assembly[®] mixture (New England, Biolabs) which possesses 5' exonuclease, DNA ligase and DNA *Polymerase* with appropriate buffer at 50°C for 1 hour.

Table 3.3 PCR cycling parameters for amplification of cDNA encoding mature ATGH116's.

Segment	Cycles	Temperature (°C)	Duration
Initial Denaturation	1	98	5 min
Denaturation	35	98	30 s
Annealing	35	58	30 s
Extension	35	72	1.30 min
Final Extension	1	72	10 min

3.3.3 Cloning of AtGH116 CDS into pET30

Three to four microliters of ligation mixture was transformed into NEB5-alpha *E. coli* cells as mentioned in the Section 3.2.2. The transformed cells were spread on LB agar plate with 30 µg/ml of kanamycin, and then incubated at 37°C overnight. The colonies that had grown were selected and cultured on the LB broth with 30 µg/ml of kanamycin overnight at 37°C. Plasmids were prepared from the grown liquid cell cultures by the Vivantis plasmid DNA extraction method (Section 3.2.4). The presence of recombinant gene insert was checked by amplifying with T7 promotor and T7 terminator primers with the PCR conditions shown in the Table 3.2, followed by agarose gel electrophoresis. The orientation and reading frame of the inserted gene sequences of the plasmids were verified by automated DNA sequencing at ATGC Co., Ltd, (Bangkok, Thailand).

3.3.4 Expression of pET30/AtGH116 CDS

The pET30 vectors containing the AtGH116 CDS were transformed and expressed in *E. coli* BL21(DE3) and Rosetta-gami(DE3). Transformed cells possessing pET30/ATGH116CDS were spread onto LB-agar containing 30 µg/ml of kanamycin for BL21(DE3) and 50 µg/ml of ampicillin, 12.5 µg/ml of tetracycline, 34 µg/ml of chloramphenicol, and 30 µg/ml of kanamycin for Rosetta-gami(DE3). The transformed cell LB-agar plates were incubated at 37°C overnight. The colonies that grew overnight were picked and inoculated into Luria-Bertani (LB) Broth containing the same set of antibiotics, and the cultures were incubated at 37°C with 200 rpm rotary shaking overnight, to make a starter culture. To express a AtGH116 β -glucosidase, a 1% final concentration of starter was transformed into the desired volume of same type of LB Broth with antibiotics, and incubated at 37°C with 200 rpm, shaking until it reached the absorbance 0.4 at OD₆₀₀. The protein expression was induced by addition of IPTG. The optimum expression conditions were determined by varying the final concentration of IPTG from 0 to 0.6 mM at 20°C for 16 to 18 h. The cell pellets were collected by centrifugation at 4500 rpm for 30 mins at 20°C. Cells were stored at -80°C until protein extraction.

The IPTG-induced expression cells were thawed by incubating at room temperature. Cells were resuspended in freshly prepared lysis buffer (50 mM potassium phosphate buffer pH 8.0, 300 mM KCl, 10 mM imidazole, 10% glycerol, 1% Triton X-100, 100 mM PMSF, 100 mM aminocarbonic acid, 100 mM benzamidine with 1 mg/ml lysozyme) at the ratio of 15 volume of lysis buffer per liter of cell culture, mixed by pipetting, and then incubated for 30 min on ice. The insoluble proteins were separated by centrifugation. Aliquots of both soluble and

insoluble protein fractions were used to test the hydrolysis activity on 5 mM *p*NPGlc at 30°C for 30 min (section 3.3.6) and also analyzed by loading 10 µl into SDS-PAGE (Section 3.2.7).

3.3.5 Purification of recombinant At4GH116

The pET30 vectors containing the At4GH116 CDS were transformed and expressed in *E.coli* BL21(DE3). Transformed cells possessing pET30/At4G were spread onto LB-agar containing 30 µg/ml of kanamycin for BL21(DE3). The transformed cell LB-agar plates were incubated at 37°C overnight. The colonies that grew overnight were picked and inoculated into Luria-Bertani (LB) Broth containing the same set of antibiotics, and the cultures were incubated at 37°C with 200 rpm rotary shaking overnight, to make a starter culture. To express an At4GH116, a 1% final concentration of starter was transformed into the desired volume of same type of LB Broth with antibiotics, and incubated at 37°C with 200 rpm, shaking until it reached the absorbance 0.4 at OD₆₀₀.

The final concentration of 0.4 mM isopropyl-D-thiogalactopyranoside (IPTG) was added to the cell culture and incubated at 20°C for 16-18 hrs to induce transcription. Cells were harvested by centrifugation and stored at -80°C overnight. Frozen cells were thawed by incubating at room temperature and resuspended in lysis buffer (50 mM potassium phosphate buffer, pH 8.0, 300 mM KCl, 10 mM imidazole, 10% glycerol, 1% Triton X-100, 1 mM PMSF, 1 mM aminocarbonic acid, 1 mM benzamidin) then incubated for 30 min on ice with 1 mg/ml lysozyme. The resuspended cells were lysed via sonication for 20×10 s pulses at 50% power of a Bandelin sonoplus GM 2200. Cell debris was removed by centrifugation at 12,000 rpm for 60 min at 4°C. At4GH116 was purified by supplying the supernatant into the

column filled with immobilized affinity chromatography (IMAC) resins activated with Ni^{2+} , which was equilibrated with equilibration buffer (50 mM potassium phosphate buffer, pH 8.0, 300 mM KCl, 10 mM imidazole, 10% glycerol). The flow-through solution was collected and stored for analysis. Then, the column was washed with 5 column volumes of wash buffer (50 mM potassium phosphate buffer, pH 8.0, 300 mM KCl, 20 mM imidazole, 10% glycerol), and eluted with 5 column volumes of elution buffer (500 mM imidazole, 50 mM potassium phosphate buffer, pH 8.0, 300 mM KCl, 10% glycerol).

Proteins from the elution fractions were concentrated, and the buffer was changed to 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 0.5 mM EDTA and 1 mM DTT in a 50-kDa cutoff ultrafiltration membrane. The N-terminal S-tag and His₆-tag, were then excised with tobacco etch virus (TEV) protease, at a ratio of 1:50 (w/w) TEV protease to fusion protein overnight at 4°C. The cleaved N-terminal fusion tag and TEV protease were removed by adsorption to Ni^{2+} IMAC resin. The flow-through and the active elution fractions with 0 mM imidazole elution buffer containing *p*NPGlc hydrolysis activity were pooled together and concentrated in a 50-kDa cutoff ultrafiltration membrane.

The enzyme was further purified via Superdex 200 gel filtration size exclusion chromatography (SEC). Before injecting into the column, the protein solution was clarified by centrifugation at 13,000 g at 4°C for 10 min, to remove the dust. Fractions were eluted with 50 mM potassium phosphate buffer, pH 7.5, 150 mM KCl, 1 mM DTT, 1 mM EDTA, 10% glycerol. Protein fractions were pooled together based on their native molecular weight, as shown in the Table 3.4, and concentrated. The

enzymes were stored in storage buffer 50 mM potassium phosphate buffer, pH 7.5, 150 mM KCl, 1 mM DTT, 1 mM EDTA, 10% glycerol at -80°C for further usage.

Table 3.4 Specification of fractions pooled together obtained from S200 size exclusion chromatography.

Fraction numbers	Peak number
16 - 17	I
20 - 21	II
23 - 24	III
26	IV
29 - 30	V
34 - 35	VI
39 - 40	VII
42 - 43	VIII

3.3.6 Enzyme assay and synthetic and natural substrate specificity

The substrate specificity of At4GH116 was evaluated towards natural and synthetic substrates. To compare the activity on nitrophenol glycosides of different sugars, 0.25 μ g of enzyme with 1 mg/ml BSA was incubated with 1 mM final concentration of different synthetic substrates such as *p*NP- α -D-glucopyranoside, *p*NP- β -D-glucopyranoside, *o*NP- β -D-glucopyranoside, *p*NP- α -D-galactopyranoside, *p*NP- β -D-galactopyranoside, *o*NP- β -D-galactopyranoside, *p*NP- β -D-fucopyranoside, *p*NP- β -D-xylopyranoside, *p*NP- α -L-arabinofuranoside, *p*NP- α -L-arabinopyranoside, *p*NP- β -D-cellobioside, *p*NP-N-acetyl- β -D-glucosamine, *p*NP- α -D-mannopyranoside, *p*NP- β -D-mannopyranoside at 30°C for 30 min in 140 μ L of 50 mM MES buffer, pH 6. The hydrolysis reactions were stopped by addition of 70 μ L of 2 M sodium carbonate. The absorbance was measured at 405 nm and, to determine the amount of *p*NP released, this absorbance was compared to that of a *p*NP standard curve.

A set of natural substrates listed in Table 3.5 were used to determine their potential hydrolysis by At4GH116. The reactions contained 1 mM final concentration of natural glycoconjugates and 1 μ g of At4GH116 in 50 mM MES buffer, pH 6. The reaction mixtures were incubated at 30°C overnight. Reactions with natural glycosides were stopped by boiling for 5min for TLC analysis and adding 1:3 volumes of methanol for quantitative analysis. The amount of glucose/galactose released was determined qualitatively by TLC, while the respective aglycones of flavonoid glucosides were quantitatively measured by ultra-high performance liquid chromatography (UHPLC).

3.3.7 Thin layer chromatography analysis of glycosyl hydrolysis reaction

The product from the hydrolysis reaction was determined by spotting 10 μL of reaction mixture onto thin-layer chromatography silica gel 60 F₂₅₄ plates (Merck, Darmstadt, Germany). In separate lane, 10 mM glucose was spotted as a standard. TLC plates were developed with a mobile phase of ethyl acetate, acetic acid and water (2:1:1, v/v/v) for all the glycosides, and (7:2.5:0.5, v/v/v) used for anthocyanin glucosides and detected by spreading 10% sulfuric acid in ethanol on the plate and heating until the purple spots were visible.

3.3.8 Ultra High performance liquid chromatography (UHPLC)

Ten microliters of 10 to 1000 μM aglycone standards prepared in 99% methanol with 1% formic acid were used to plot the standard calibration curve. Aglycones, including kaempferol, quercetin, apigenin, and luteolin, in the enzyme reactions were quantified on an Agilent 1290 Infinity ultra-high performance liquid chromatography system. Aglycones were separated by a ZORBAX SB C18 (1.8 μm , 2.1 \times 150 mm) column (Agilent, USA) eluted with mobile phases that contained 0.2% formic acid in water (solvent A) and acetonitrile (solvent B). In order to elute the compounds, a linear gradient was set up from 5% to 50% B (v/v) from 0 to 13 min, 50% to 100% B (v/v) from 13 to 16 min, and 100% to 5% B (v/v) from 16 to 25 min at a flow rate of 0.2 ml/min and compared with calibration curve of the aglycone standards injected on the same system.

Table 3.5 Substrates used to analyze glucosyl hydrolase reaction.

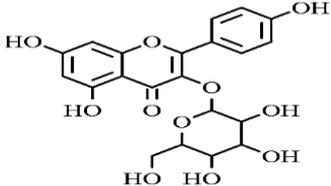
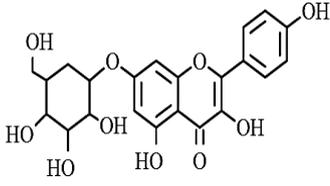
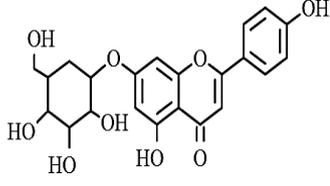
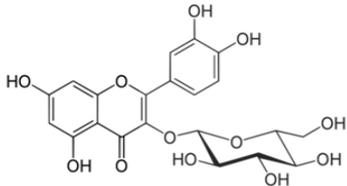
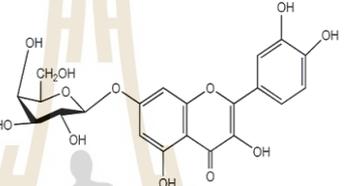
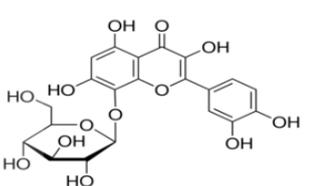
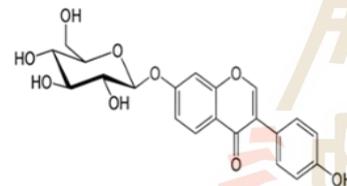
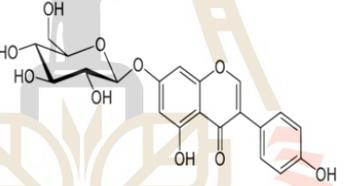
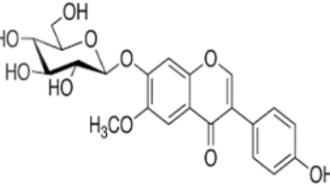
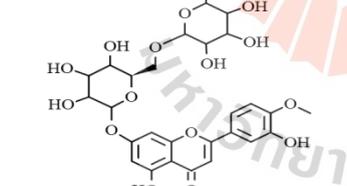
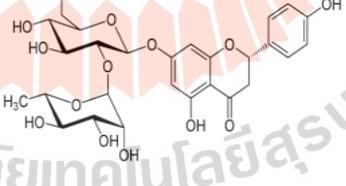
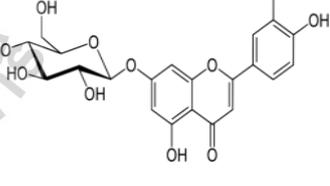
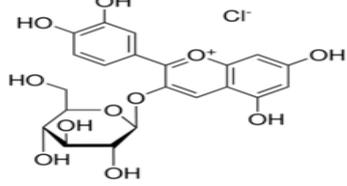
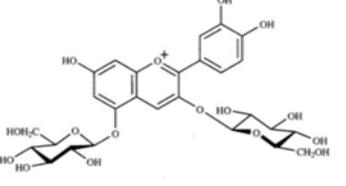
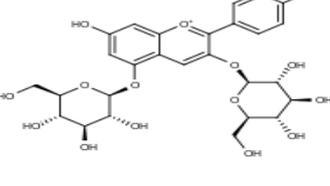
Flavonoids		
		
Kaempferol 3-O- β -D-glucoside	Kaempferol 7-O- β -D-glucoside	Apigenin 7-O- β -D-glucoside
		
Quercetin 3-O- β -D-glucoside	Quercetin 7-O- β -D-glucoside	Gossypin
		
Diazidin	Genistin	Glycitin
		
Hesperidin	Narigin	Luteolin 7-O- β -D-glucoside
		
Cyanidin-3-O- β -D-glucoside	Cyanidin-3,5-O- β -D-glucoside	Peonidin-3,5-O- β -D-glucoside

Table 3.5 Substrates used to analyze glucosyl hydrolase reaction (Continued).

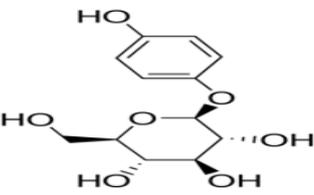
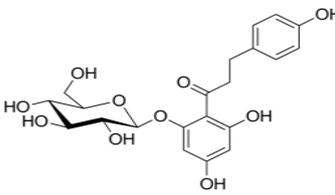
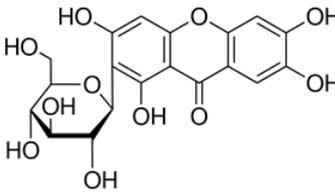
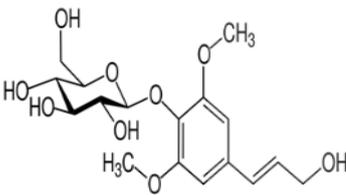
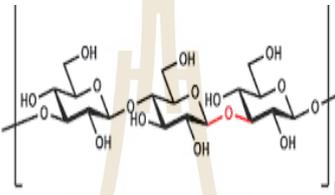
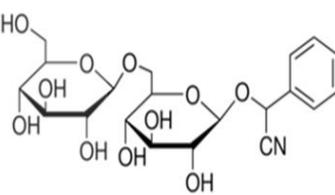
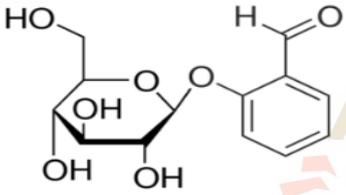
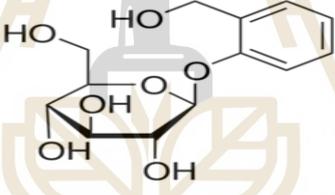
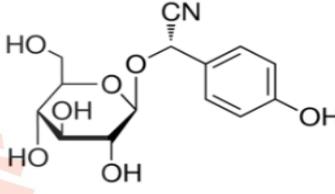
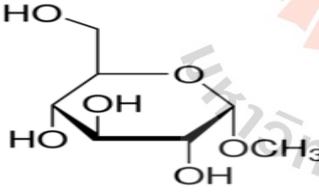
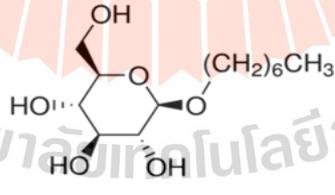
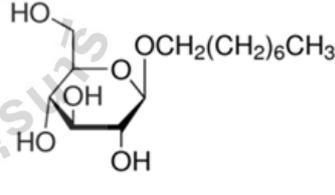
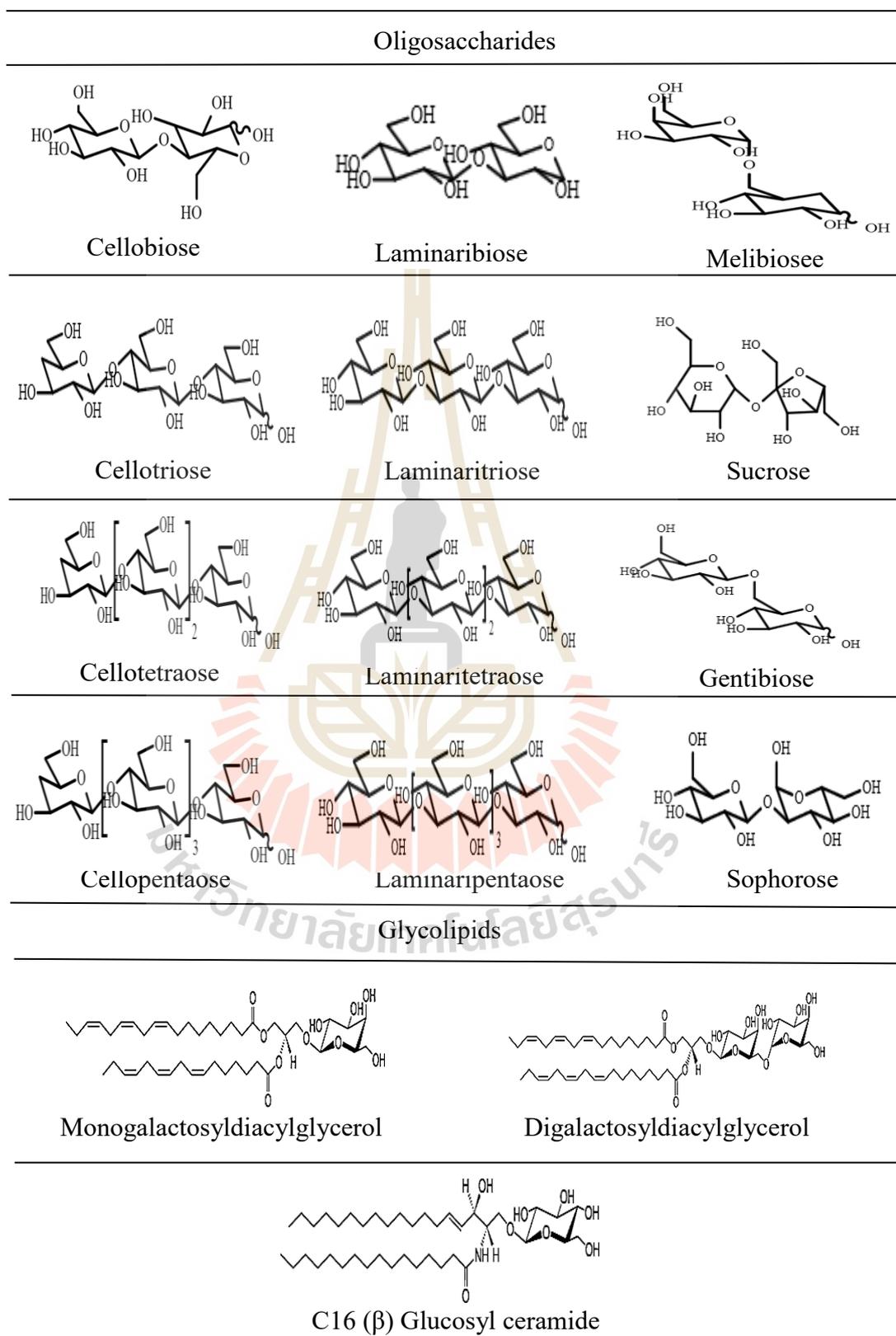
Natural substrates		
		
Arbutin	Phlorizin	Mangiferin (C-glucoside)
		
Syringin	Lichenan	D-Amygdalin
		
Helicin	Salicin	Dhurrin
		
<i>methyl</i> β -D-glucopyranoside	<i>Heptyl</i> β -D-glucopyranoside	<i>octyl</i> β -D-glucopyranoside

Table 3.5 Substrates used to analyze glucosyl hydrolase reaction (continued).

3.3.9 Effect of inhibitors on activity of At4GH116

Potential inhibitors were pre-incubated at final concentrations of 1 and 10 μM for 30 mins with 1 μg of At4GH116 at room temperature. Then, enzyme activity was then assayed with 1 mM *p*NP- β -D-glucopyranoside, as mentioned in the Section 3.3.6.

3.4 At4GH116 crystallization

3.4.1 Preliminary screening by vapor diffusion technique

Solutions of purified At4GH116 (SEC peak I or II) in 20 mM Tris-HCl and 300 mM NaCl, pH 8.0, and 50 mM potassium phosphate buffer, pH 7, 150 mM KCl, 1 mM DTT, 1 mM EDTA, 10% glycerol were used to screen for crystallization conditions by the sitting drop vapor diffusion method. Prior to usage, the protein solution was filtered with an Ultrafree MC 0.22 μm filter by centrifugation at 13,000 rpm at 4°C for 5 min. The crystallization conditions for At4GH116 were screened by the sitting drop vapor diffusion method in a 60 well plastic plates, with the crystallization kits listed in Table 3.6. An Oryx 8 robot from Douglas Instruments was employed for screening. Fifty microliters of precipitant was distributed in each reservoir of the plate and 0.2 μl of precipitant was dropped in both the adjacent wells. Then 0.2 μl of 5-10 mg/ml of protein was mixed with precipitants in one well and 0.4 μl with that in the other. The plate was covered with plastic wrap and incubated at 15°C. The growth of the crystals was monitored by observing the drops under a stereo microscope at various times. Crystal conditions obtained from the robotic screening were optimized further to obtain good crystals.

Table 3.6 Commercially available crystallization kits used for preliminary screening.

Screening Kits	Company
Crystals Screen High Throughput HR2-130 and HR2-134 Crystal Screen Cryo™ HR2-122	Hampton Research, Aliso Viejo, CA, USA
JBScreen Classic HTS I & HTS II	Jena Bioscience, Max-Planck-Institute of Molecular Physiology, Dortmund, Germany
PACT premier™ HT-96 / FX-96 MD 1-36 JCSG-plus™ HT-96/FX-96 MD1-40	Molecular Dimensions, Maumee, OH, USA

3.4.2 Optimization of crystals by hanging drop method

Crystallization conditions were optimized by hanging drop vapor diffusion method in sterile 24 well plates to obtain single crystals in good shape by varying the protein concentration at 2 and 5 mg/ml. High vacuum grease was applied to the edges of each well of the plate. The reservoirs were filled with 500 μ l of the different precipitants shown in Tables 3.7 – 3.11. Then, 1 μ l of precipitant and 2 μ l of protein (2 and 5 mg/ml in separate drops) were pipetted and mixed upon a siliconized cover slip. The cover slip was inverted and placed with drops facing downwards over the well. The cover slip was sealed to the grease layer. The plates were labeled and incubated at 15°C. Every few days, the crystallization drops were checked under a stereomicroscope.

3.4.3 Data collection and processing

Crystals of At4GH116 were soaked briefly in 10% higher concentration of precipitant containing 20% glycerol. For ligand complexes, the soaking solution also contained either 10 mM 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside (DNPG2F) or 100 mM D-glucose. Crystals were used to diffract 1.0 Å wavelength X-rays on the BL13B1 beamline at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan), and reflections were recorded with a MX-300HS detector (Rayonix) with the crystal to detector distance of 300 mm.

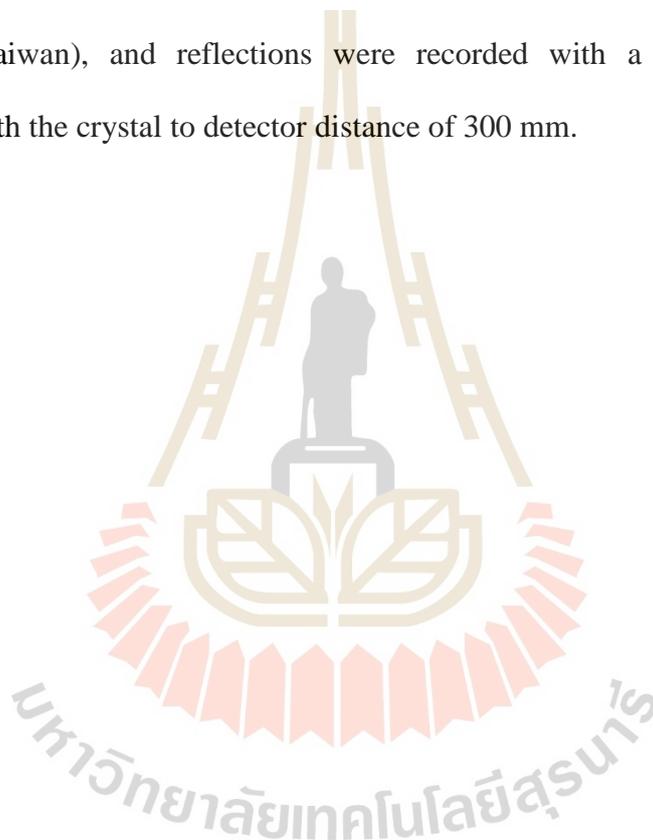


Table 3.7 Optimization of At4GH116 crystals in the conditions 0.15 M - 0.4 M magnesium formate and 0.1 M Bis-Tris, pH 5 - 6.5.

	1	2	3	4	5	6
A	0.15 M Magnesium formate	0.2 M Magnesium formate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate,	0.4 M Magnesium formate,
	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5
B	0.15 M Magnesium formate	0.2 M Magnesium formate,	0.25 M Magnesium formate,	0.3 M Magnesium formate,	0.35 M Magnesium formate,	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5
C	0.15 M Magnesium formate	0.2 M Magnesium formate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6
D	0.15 M Magnesium formate dehydrate	0.2 M Magnesium formate dehydrate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5

Table 3.8 Optimization of At4GH116 crystals in the conditions 0.15 M - 0.4 M magnesium formate and 0.1 M Bis-Tris, pH 5 - 6.5 with 25% PEG 3000-6000.

	1	2	3	4	5	6
A	0.15 M Magnesium formate	0.2 M Magnesium formate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5	0.1 M Bis-Tris, pH 5
	25% PEG 3000	25% PEG 3000	25% PEG 3000	25% PEG 3000	25% PEG 3000	25% PEG 3000
B	0.15 M Magnesium formate	0.2 M Magnesium formate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis-Tris, pH 5.5	0.1 M Bis -Tris, pH 5.5	0.1 M Bis -Tris, pH 5.5	0.1 M Bis -Tris, pH 5.5,
	25% PEG 4000	25% PEG 4000	25% PEG 4000	25% PEG 4000	25% PEG 4000	25% PEG 4000
C	0.15 M Magnesium formate	0.2 M Magnesium formate,	0.25 M Magnesium formate,	0.3 M Magnesium formate,	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6	0.1 M Bis-Tris, pH 6
	25% PEG 5000	25% PEG 5000	25% PEG 5000	25% PEG 5000	25% PEG 5000	25% PEG 5000
D	0.15 M Magnesium formate	0.2 M Magnesium formate	0.25 M Magnesium formate	0.3 M Magnesium formate	0.35 M Magnesium formate	0.4 M Magnesium formate
	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5	0.1 M Bis-Tris, pH 6.5
	25% PEG 6000	25% PEG 6000	25% PEG 6000	25% PEG 6000	25% PEG 6000	25% PEG 6000

Table 3.9 Optimization of At4GH116 crystals in the conditions 0.15 M - 0.3 M magnesium chloride and 0-15% Glycerol in 0.1M HEPES, pH 7.5, 16% 2-propanol.

	1	2	3	4	5	6
A	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	0% Glycerol	0% Glycerol				
	0.1 M HEPES, pH 7.5	0.1 M HEPES, pH 7.5				
	16% 2-propanol	16% 2-propanol				
B	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	5% Glycerol	5% Glycerol				
	0.1 M HEPES, pH 7.5	0.1 M HEPES, pH 7.5				
	16% 2-propanol	16% 2-propanol				
C	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	10% Glycerol	10% Glycerol				
	0.1 M HEPES pH 7.5	0.1 M HEPES pH 7.5	0.1 M HEPES, pH 7.5	0.1 M HEPES pH 7.5	0.1 M HEPES pH 7.5	0.1 M HEPES pH 7.5
	16% 2-propanol	16% 2-propanol				
D	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	15% Glycerol	15% Glycerol				
	0.1 M HEPES pH 7.5	0.1 M HEPES pH 7.5				
	16% 2-propanol	16% 2-propanol				

Table 3.10 Optimization of At4GH116 crystals in the conditions 0.24 M magnesium chloride, 10-20% 2-propanol and 0-15% Glycerol in 0.1M HEPES, pH7.5.

	1	2	3	4	5	6
	0.24 M Magnesium chloride					
A	0.1 M HEPES, pH 7.5 10% 2-propanol 0% Glycerol	0.1 M HEPES, pH 7.5 12% 2-propanol 0% Glycerol	0.1 M HEPES, pH 7.5 14% 2-propanol 0% Glycerol	0.1 M HEPES, pH 7.5 16% 2-propanol 0% Glycerol	0.1 M HEPES, pH 7.5 18% 2-propanol 0% Glycerol	0.1 M HEPES, pH 7.5 20% 2-propanol 0% Glycerol
	0.24 M Magnesium chloride					
B	0.1 M HEPES, pH 7.5 10% 2-propanol 5% Glycerol	0.1 M HEPES, pH 7.5 12% 2-propanol 5% Glycerol	0.1 M HEPES, pH 7.5 14% 2-propanol 5% Glycerol	0.1 M HEPES, pH 7.5 16% 2-propanol 5% Glycerol	0.1 M HEPES, pH 7.5 18% 2-propanol 5% Glycerol	0.1 M HEPES, pH 7.5 20% 2-propanol 5% Glycerol
	0.24 M Magnesium chloride					
C	0.1 M HEPES, pH 7.5 10% 2-propanol 10% Glycerol	0.1 M HEPES, pH 7.5 12% 2-propanol 10% Glycerol	0.1 M HEPES, pH 7.5 14% 2-propanol 10% Glycerol	0.1 M HEPES, pH 7.5 16% 2-propanol 10% Glycerol	0.1 M HEPES, pH 7.5 18% 2-propanol 10% Glycerol	0.1 M HEPES, pH 7.5 20% 2-propanol 10% Glycerol
	0.24 M Magnesium chloride					
D	0.1 M HEPES, pH 7.5 10% 2-propanol 15% Glycerol	0.1 M HEPES, pH 7.5 12% 2-propanol 15% Glycerol	0.1 M HEPES, pH 7.5 14% 2-propanol 15% Glycerol	0.1 M HEPES, pH 7.5 16% 2-propanol 15% Glycerol	0.1 M HEPES, pH 7.5 18% 2-propanol 15% Glycerol	0.1 M HEPES, pH 7.5 20% 2-propanol 15% Glycerol

Table 3.11 Optimization of At4GH116 crystals in the conditions 0.15 - 0.3 M magnesium chloride and 10 - 17.5% 2-propanol in 0.1M HEPES, pH 7.

	1	2	3	4	5	6
A	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	0.1 M HEPES	0.1 M HEPES				
	pH 7.5	pH 7.5				
	10% 2-propanol	10% 2-propanol				
B	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	0.1 M HEPES	0.1 M HEPES				
	pH 7.5	pH 7.5				
	12.5% 2-propanol	12.5% 2-propanol				
C	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	0.1 M HEPES	0.1 M HEPES				
	pH 7.5	pH 7.5				
	15% 2-propanol	15% 2-propanol				
D	0.15 M Magnesium chloride	0.18 M Magnesium chloride	0.21 M Magnesium chloride	0.24 M Magnesium chloride	0.27 M Magnesium chloride	0.3 M Magnesium chloride
	0.1 M HEPES	0.1 M HEPES				
	pH 7.5	pH 7.5				
	17.5% 2-propanol	17.5% 2-propanol				

3.5 Small Angle X-ray scattering

To determine the three-dimensional structure and other characterization of At4GH116, small angle x-ray scattering (SAXS) experiments were carried out at the beam line 1.3W of the Synchrotron Light Research Institute (Public Organization) (SLRI), Thailand. Sixty microliters of 1, 2, 5 or 10 mg/ml of At4GH116 and a buffer blank for comparison were injected into copper sample holder. The copper sample holder was prepared by sealing the inner sides of the two copper cell plates with krypton tape and the plastic inner sample holder with loading holes was placed in between and all three parts of the sample holder were screwed together. Further, it was placed in the radiation chamber and exposed to X-ray (9 keV) at 16°C for 10 minutes, with the sample to detector distance of 3770 mm. SAXS data capture covered the q range of $0.0122 < q < 0.1816 \text{ \AA}^{-1}$. The matched reference buffer was measured first, followed by the protein sample in the same sample cell. The scattering of buffer was subtracted for background. The raw data were processed by the program SAXSIT, which was developed by SLRI staff, and BioXTAS RAW (Hopkins et al., 2017) to obtain scattering curves. Guinier analysis, $P(r)$ analysis, and Kratky plots, and maximum particle diameter (D_{\max}) were calculated from the scattering curves using the Primus program (Konarev et al., 2003), which is part of the ATSAS package (Franke et al., 2017). The *ab initio* envelope reconstructions were carried out by the DAMMIF server (Franke and Svergun, 2009) and DAMAVER (Volkov, and Svergun, 2003) with ten independent bead models without imposed symmetry (P1) of each protein.

3.6 In planta analysis of effect of GH116 gene knockout in lines of *Arabidopsis thaliana*

3.6.1. Preparation and growth conditions of wild type (WT) and AtGH116 gene knockout lines

Arabidopsis WT and GH116 gene mutant seeds were sown individually in separate pots with a soil mixture containing peat moss, vermiculate and perlite in the ratio of 3:1:1 (v/v/v). Both WT and mutant seeds were surface sterilized with 70% ethanol for 1 min, then with 1% sodium hypochlorite solution 1 min, and subsequently rinsed by sterile deionized water 6 times, and then kept in the dark for 3 days at 4° for stratification. Then, seeds were sown on half-strength Murashige and Skoog (MS) plates containing 0.8% agar. Plants were grown in a plant growth room maintained at 25°C under a long day light (16 hr light/8 hr dark) conditions with the intensity of 75 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

3.6.2 Identifications of AtGH116 mutant lines

Seeds of *Arabidopsis thaliana* with T-DNA insertions in AtGH116 genes, including At1G33700-Salk_143139, At3G24180-SALK_039506, SALK_117180, At4G10060 SALK_099838, SALK_010803 and At5G49900- SALK_039511, and SALK_057433, were obtained from the Salk Institute Genomic Analysis Laboratory, San Diego, CA, USA. Genomic DNA was extracted from the individual mutant lines (Sections 3.2.10 and 3.2.11). Plants with the T-DNA insertion were identified by PCR using the primers shown in Table 3.2.

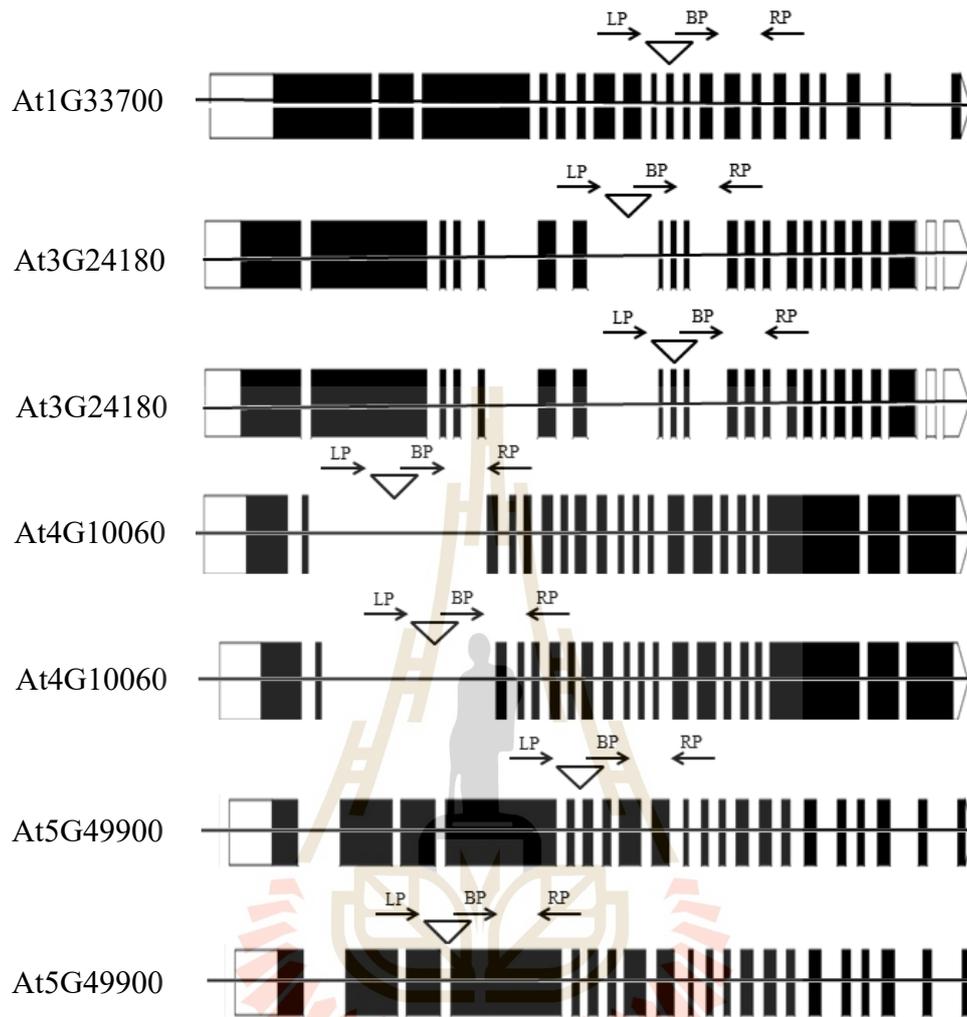


Figure 3.1 Schematic representations of the *Arabidopsis thaliana* GH116 genes indicating the locations of the T-DNA insertions; solid boxes represent exons; thin lines introns; white boxes untranslated regions, and triangle indicate the location of T-DNA insertions. Positions of primers used for PCR-genotyping are shown. LP - left border primer, RP - right border primer, BP – T-DNA insert border primer.

Gene specific fragments were amplified with pairs of LP and RP primers and for T-DNA inserted fragments the BP primer from the T-DNA sequence was paired with the gene-specific RP primer as illustrated in Figure 3.1. The DNA fragments were amplified with *Taq* polymerase. Temperature cycling parameters for amplifications are shown in Table 3.12. The PCR products of gene specific fragment and T-DNA inserted fragments of approximately 1000 bp and 500 bp, respectively were checked by electrophoresis on a 1% agarose gel.

Table 3.12 PCR cycling parameters for amplification of genomic DNA from wild and mutant lines of *Arabidopsis*.

Segment	Cycles	Temperature (°C)	Duration
Initial Denaturation	1	98	5 min
Denaturation	35	98	30 s
Annealing	35	55	30 s
Extension	35	72	1.30 min
Final Extension	1	72	10 min

3.6.3 Generation of double mutant lines

Screened and selected single homozygous mutant lines were grown up to inflorescence stage. In the selected inflorescence, anthesis flowers and inflorescence buds were removed. Flowers in late pre-anthesis were selected for cross germination. One of the homozygous mutant lines was considered as female and another one as male. In flowers of the homozygous mutant line that was considered as the female, anthers were removed along with some sepals and petals. The pistil consisting of the ovary, style and stigma was covered with remaining sepals and petal, and left undisturbed for a day. The next day, the anthers from the other homozygous mutant line were removed along with filament and rubbed over the top of the designated female line stigma. Plants were allowed to grow and germinate under the normal growth conditions, as shown in Section 3.6.1. Individual seed pots were collected and designated as F₁ generation once it's started to turn brown and used for double mutant screenings.

3.6.4 Screening of the double mutant lines

Seeds from the heterozygous mutant lines F₁ were prepared as described in Section 3.6.1 and allowed to self-germinate. Self-germinated seeds of F₁ were grown to rosette stage of plants and DNA was extracted as described in the Section 3.2.11 and followed as described in the Section 3.6.2.

3.6.5 Analysis of effect of nutrients and light duration on phenotypes

A. thaliana (Columbia-0) and seeds of *Atgh116* single homozygous mutant lines were sown on soil and half-MS agar plate, as described in Section 3.6.1 on day 0. Plants grown on the soil under long day and short days are referred to as SL and SS respectively. Plants grown on the half MS agar plate under long day and short day

conditions are referred to as HL and HS, respectively. After 24 days, all the plants from HS and HL conditions were moved to the SS and SL conditions respectively. Plants were left in this condition for another 14 days. On day 38, all the plants from the SS conditions were moved to SL conditions and grown until the plants died.



CHAPTER IV

RESULTS AND DISCUSSION

4.1 Phylogenetic analysis of GBA2 protein sequences from four loci of *A. thaliana*

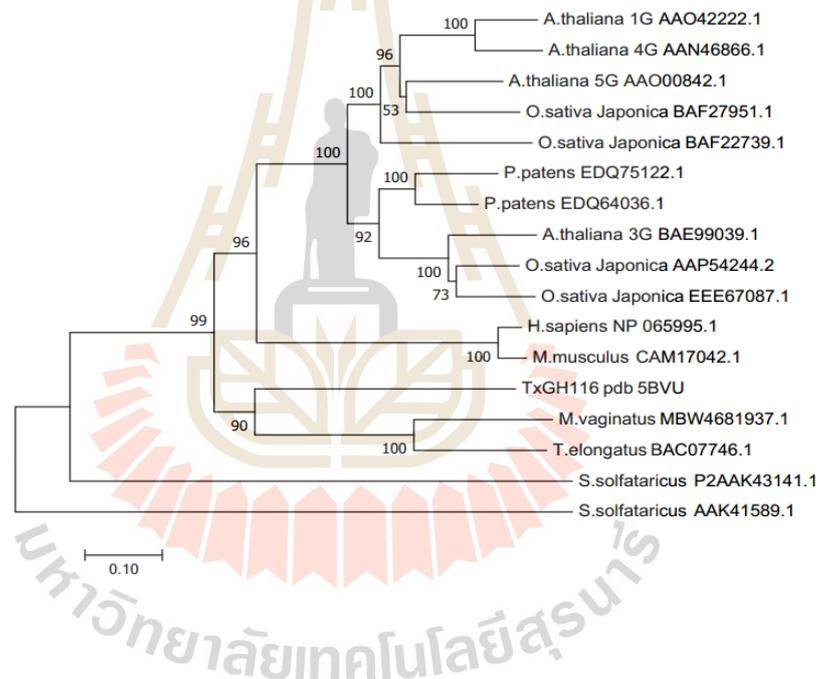


Figure 4.1 Phylogenetic analysis of GH116 protein sequences encoded by four gene loci of *A. thaliana*. The protein sequences were aligned by MUSCLE method in MEGA 6 program. The phylogenetic tree was built using the neighbor-joining method with AtGH116 and other characterized GH116 family enzymes sequences from plants, animals and bacteria. Bootstrap reproducibility percentages are shown on the internal branches.

```

*      20      *      40      *      60      *
AAO42222.1 : -----MQNVPEDQ-----N : 9
BAE99039.1 : -----MVGATLFHRRKHSWPTTEEFISRSTLQ : 26
AAN46866.1 : -----MEKNGHTESELQ-----T : 13
AAO00842.1 : -----MFEEKIMDIGEDVKPLNS-----S : 19
pdb|5BVU|A : -----AMALTGCSEKINI-----N : 14
NP_065995. : MGTQDPGNMGTGVPASEQISCAKEDPQVYCP EETGGTKDVQVTDCKSPEDSRPPKETDCCNPEDSGQLMVSYE : 73

      80      *      100      *      120      *      140
AAO42222.1 : KMVIDDKLP-P-FSWE----RKLNSQAKTPSEFKLSKRDMHLFPLGYRL--WRHTKDEAAKGRASIFDI FRK : 74
BAE99039.1 : LLDFDSAAPP-P-HAWR----RRLNCHANILKEFTITFREAIK MVR LGIRL--WSYVREEASHGRKAPIDPFTK : 92
AAN46866.1 : QMVENERLP-Q-VTWQ----RKLNSKVKNPSEFKMSIRDVLHLFPLGYRL--WRYTKEEAKKGRFSMYDIFFK : 78
AAO00842.1 : DTKVDPAVPAS-LTWQ----RKIDSDVKAPREFNLSVKEIFQLAPVGI RL--WFLCREEAAKGRLAFTDPFSK : 85
pdb|5BVU|A : EDKISHKIDIPDSAWT-----IGIGKFK-----KNAGHPNVKYPMIDDS-- : 53
NP_065995. : GKAMGYQVP-P-FGWRICLAHEFTKRRKPFQANNVSLSNMIKHIGMGLRYLQWYRKTTHVEK-KTPFI DMINS : 143

      p      W
*      160      *      180      *      200      *      22
AAO42222.1 : HHIT--GDHGVPLGGIGSGSISGRSYKGEFQQFKLFPKICEEAPILTNQFSAFVSRPVG--VKHSTVLCPSKQOV : 144
BAE99039.1 : ENCKPSASQGVPLGGMGSGSISGRGFRGEFKQWQITPGTCDPSPMMSNQFSIFISRDCGGHKYASVLSPGQHGS : 165
AAN46866.1 : RHVR--GDHGVPLGGIGSGSISGRSYKGEFQQFKLFPKICEEAPILTNQFSVSVSRPVG--LSYSTVLCPTKPKS : 148
AAO00842.1 : HSVT--SSHGVPLGGIGAGSISGRSFKGEFQRWQLFPKCEDEFVLANQFSAFVSRANG--KKYSSVLCPRNPKL : 155
pdb|5BVU|A : -----YVQCAPLGGFGAGTIGRTYNGGFSRWHLIEIGKNKYTTVYANQFSVQKVEGNKDGVAQVLYAGEEEN : 120
NP_065995. : VPLR--QIYGCPLGGIGGGTITRGRWGFQCRWQLNPGMYQHRTVIADQFTVCLRRECQ--TVYQQVLSLERESV : 213

      G PLGG G G3I R 5 G F 5 6 p 6 1QF3 f r g VL p
*      240      *      260      *      280      *
AAO42222.1 : VKDNGGYLCQGQAPNMGIESWDNMTGEKSTYHALYPRSWTVYDGE PDPELRIVSRQVSPFI PHNYEESLPLV : 217
BAE99039.1 : LG-----KSRDKGLSSWGNNLNGQHSTYHALFPRAWITTYDGE PDPELKRISCRQLSPFI PHNNYRDSSELP : 229
AAN46866.1 : VK-----GKTEDLGIESWDNMEGDKSTYHALYPRSWTVYN-EPDPELRIVSRQVSPFI PHNYKESLPLV : 212
AAO00842.1 : DK-----QDSESGIGSWDNLKGDKSTYHALYPRSWTVYEGEPDPELRIVCRQVSPFI PHNYKESSEFPV : 219
pdb|5BVU|A : GY-----LSSWKWDYPKESGMYYALYPNWYTYT-NKDLPLVQLAVKQFSPTI PHNYKETSYPV : 177
NP_065995. : LR-----SWNWGLCGYFAFYHALYPRAWTVYQ-LEGQNVTLTCRQITPILPHDYQDSSLPV : 268

      SW W g YhAL5Pr Wt Y pd 6 6 4Q 3P 6P 1Y 3S Pv

```

Figure 4.2 Amino acid sequence alignment of *Arabidopsis* GH116, *Thermoanaerobacterium xylanolyticum* GH116 (TxGH116), Human GBA₂. Conserved nucleophilic amino acids glutamic acid (E499) and acid-base aspartic acid (D647) residues are boxed.

```

      300          *          320          *          340          *          360
AAO42222.1 : SVENFTVNTCAEPAIVTLLFTWNSV---GCASGLTGCQHINSTMKAK---DGVHAVALQKKTANGHP-PVSY : 283
BAE99039.1 : AVFVYTLVNTCKERAKVSLLEFTWANSM---GCTSHMSGGHVNEPFIGE---DGVSGVILHFKTGKGNP-PVTF : 295
AAN46866.1 : SVFEFTMSNLCKEEATVTLLEFTWNSV---GCASGLTGEHINSTIMER---DGVHAIVLHFKTGNGHP-PVTY : 278
AAO00842.1 : SVFTFTLHNLGNTTADVTLLEFTWANSV---GDESEFSGGHVNSKITMN---DGVQGVILHFKTANGLP-SLSY : 285
pdb|5BVU|A : AVFKWTAYNPTNKNVDVSIIMFTWQNMIGFFGKQVNVNSGNENKI IKDKSKDSEIIVAAMGNIISNDNEEWNGEY : 250
NP_065995. : GVFVVDVENEDEALDVSIMFSMRNGL---GGDDAPGGLWNEPFCLESGETVVRGLLHHPPT---LNPYTM : 335
      VF 5t N g V366F3w N 6 Gg g N 6 6 h 3 p
      *          380          *          400          *          420          *          4
AAO42222.1 : AIAAKETEDVRVSSPCPFIVSGTTPNQITAGDMWDEIKKNASEDKLT--SNACSPSKPETSIGAAIAAKVKVP : 354
BAE99039.1 : AIAASETQNVNVTVLPCEGLSEDSS--FTAKDMWDTMEODGKFDQENFNSGSPSPSLAGDTICAAVSASAWVE : 366
AAN46866.1 : AIAAQETEDVHVECPCEFLVSGHSPKEITAKEMWDEIKKNKSFDELN--SEFGSPSRPETSIGAAIAAKVKVP : 349
AAO00842.1 : AISAQATDGVSVSACPFFIVSGKQD-GITAKDMWQAVKENGSEDLHK-ASEASMQSDHSSIGAAVAASVTVL : 356
pdb|5BVU|A : SIGVKKVPVDISYKAKFVTIGD-----GSDLWHEFSKNGILDNDK---DETPTKQ-DGIGSAIAVNFKLQ : 312
NP_065995. : AVAARVTAATTVTHITAE-----DP-DSIGQVWQDLDLQGLDSPT---GQSTPTQKCVGIAGAVCVSSKLR : 399
      a6 a t v 63 F t 6W 1 D p3 g I A6 6
      40          *          460          *          480          *          500          *
AAO42222.1 : PGCDRTVTFSLSWDCPEARF--DEKTYHRRYTRFYGSLGNAAVAMAHDALLNPEWETOIEEWQAPVLADTTL : 425
BAE99039.1 : AHGKCTVVSFALSWSSPKVKFS--KGSTYDRRYTKFYGTSPRAALDLVHDALTNKRWEEDIEAWQNPILRDERL : 438
AAN46866.1 : PGCDRTVTFSLSWDCPEVRF--NEKTYHRRYTKFYGNLGDAAVKMARDALLNYVDWESQIEAWQSPILSDTTL : 420
AAO00842.1 : PGESRIVTFSLAWDCPEVQFP--SGKIYSRRYTKFYGNNGDAAAQIAHDAILGHSQWESWIEDWQRPILEDKRL : 428
pdb|5BVU|A : PGQTIEVPEFALSWDLPIMKFG--GGDKWYKMYTKYFGKNGKNSFAILKEALNNYQKWEKMIDDWQKPILSNKS : 384
NP_065995. : ERGQCRLEFSLAWDMPRIMEGAKQOVHYRRYTRFFGQDGDAAPALSHYALCRMAEWEERTSAWQSPVLDDRSI : 472
      p 6 F L Wd P F 4rYT455G g aa 6 A6 WE I WQ P6L 1 1
      520          *          540          *          560          *          580
AAO42222.1 : PEWYRVTLFNELYFNSGGTWTWDTGLPPKQLDSIGRRKI----SLGLSTIDKNDQDQNNVALDILGRID--- : 491
BAE99039.1 : PEWYKFTLFNELYFLVAGGTWIDSSSLNANGNSQHQQS-----GLGNSDGKVGGLDINDQRNDLGNGNSVGV : 506
AAN46866.1 : PDWYRVTLFNELYFNSGGTWTWDTGLPPKESIE-----RSKVTNTEQNDIVIDLFQKIN--- : 474
AAO00842.1 : PAWYPVTLFNELYLNSGGTWTWDTGSSPVHSLAGVREKKFSLDKSQLGLKNDIDVPHQNDTAVSVLEKMA--- : 498
pdb|5BVU|A : PDWYKTALFNELYFLADGGTAWENG-----KVGEKDK--- : 416
NP_065995. : PAWYKSALFNELYFLADGGTWWLEVLE-----SLPEELGRNMC--- : 511
      P WY LFNELY5 GGT W 6

```

Figure 4.2 Amino acid sequence alignment of Arabidopsis GH116, Thermoanaerobacterium xylanolyticum GH116 (TxGH116), Human GBA2. Conserved nucleophilic amino acids glutamic acid (E499) and acid-base aspartic acid (D647) residues are boxed (Continued.).

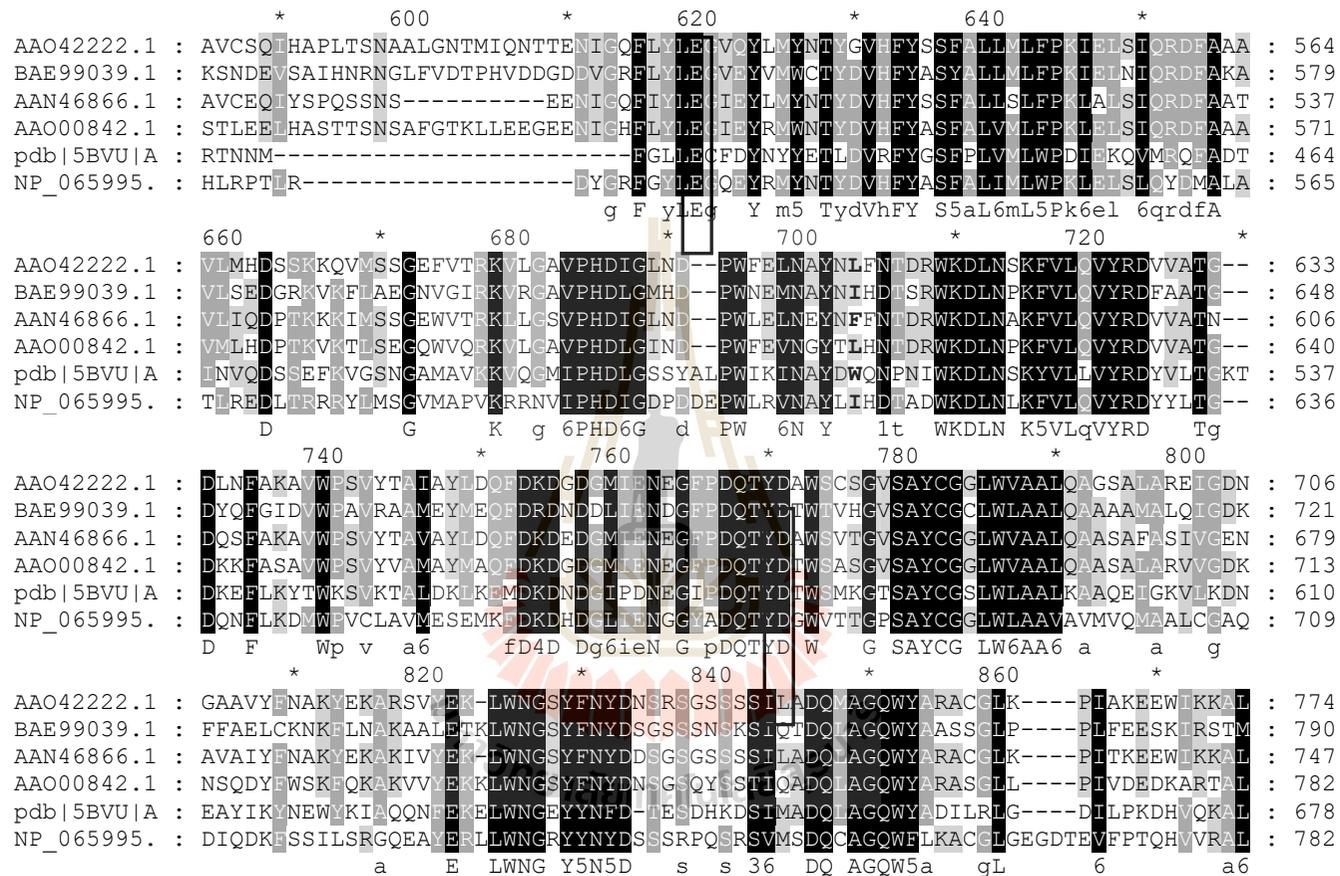


Figure 4.2 Amino acid sequence alignment of *Arabidopsis* GH116, *Thermoanaerobacterium xylanolyticum* GH116 (TxGH116), Human GBA₂. Conserved nucleophilic amino acids glutamic acid (E499) and acid-base aspartic acid (D647) residues are boxed (Continued.).

```

      880          *          900          *          920          *          940
AAO42222.1 : ETVYDFNVMRVRDGRGAVNGMLFDGRVDTSTMVSREVWAGTYSVAACMIQEGGLADKGFRTASGIYEAAWSD : 847
BAE99039.1 : QKTFDFNVMTKGGKMGAVNGMHPDGVDDTCMQSREIWTGVTYAAAAATMLSGMEEQGFRTAEGIFTAGWSE : 863
AAN46866.1 : ETIYEFNVMKVKGGRGAVNGMSTEGKVDNTSLVSKVWAGTYSVAACMIQEGQREKGFRTASGIYEAIWSD : 820
AAO00842.1 : EKVYNYNVMKIKDKGRGAVNGMHENGKVDTASMQSREIWSGVTYALSATMIQEGLVEMAFRTASGIYEAAWSE : 855
pdb|5BVU|A : KKIYEFNVMKFENGKMGAVNGMRPFGIVDESDFQAQEVVWTGVTYALASFMKYRGMTEEAANTAYGVYKMTYDK : 751
NP_065995. : QTIFELNVQAFAGGAMGAVNGMOPHGVDPDKSSVQSDDEVVWGVVYGLAATMIQEGLTWEGFRTAEGCYRTVWER : 855
      65  NVm  G  GAVNGM  p  G  vD  6  s  E6W  G  tY  aa  Mi  G  5  TA  G  5  5

      *          960          *          980          *          1000          *          1020
AAO42222.1 : RGLGCAFOTPEAWTTNDEYRSICYMRPLAIWGIQWAHTMPKPNREQE--VSLRPQEE DATSVLFQQHAGFIKV : 918
BAE99039.1 : ECFQYWFOTPEGWTMDGHYRSIIYMRPLAIWGMQWALS LPKAILD-----APQINMMDRVHLSPR : 923
AAN46866.1 : RCLSCSFOTPEAWNMNDEYRSICYMRPLAIWAIQWALTRTQSFGEEK-QKLVAGDEEEESNLLLRQHKGFKDV : 892
AAO00842.1 : TCLGYSFOTPEAWNTVDEYRSIYMRPLAIWAMQWALTKTSQKQEQGLGLEPEQPELEPSSSMKHDIGFSRV : 928
pdb|5BVU|A : SCKCYWFRTPEAWTKDGNYRASM YMRPLSLIWSMEVNYNE-----VLEHHHHHH-- : 799
NP_065995. : --LGLAFOTPEAYCQQRVFRSLIAYMRPLSIWAMQLALQQQHQKASW-----PKVKQGTGLRTGPMFGPK : 918
      g  g  FqTPE  5      5Rs1  YMRPL  IW  62  a

      *          1040          *
AAO42222.1 : AHYLNK-TKGKDHNRNLQTAYETFLKTIRL : 947
BAE99039.1 : SRRFSNNFKVVKHKAKCFGNSALSCSC--- : 950
AAN46866.1 : ARFVKIVPTSNVHRSRLQHTYETVLKTLRL : 922
AAO00842.1 : SRLLSL-PNEASAKSTLQTLFDYTCRRMMS : 957
pdb|5BVU|A : ----- : -
NP_065995. : EAMANLSPE----- : 927

```

Figure 4.2 Amino acid sequence alignment of Arabidopsis GH116, *Thermoanaerobacterium xylanolyticum* GH116 (TxGH116), Human GBA2. Conserved nucleophilic amino acids glutamic acid (E499) and acid-base aspartic acid (D647) residues are boxed (Continued.).

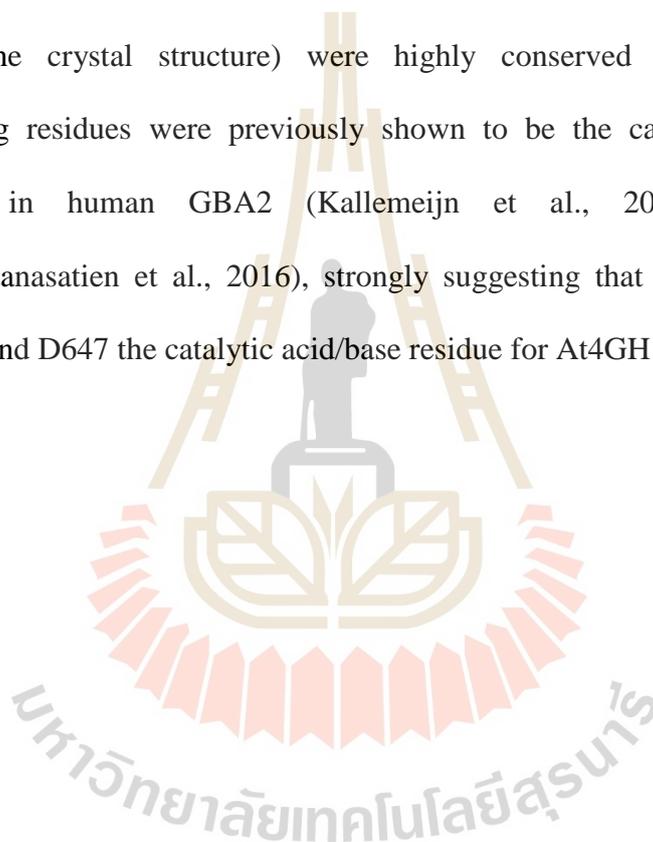
The phylogenetic analysis based on amino acid sequence similarity shown in Figure 4.1 separated the plant eukaryotic β -glucosidases from the human and bacterial β -glucosidases, which indicated that these genes are diverged from a shared ancestor of all plant GH116 genes.

The tree is based on protein sequence alignment that compares the putative *A. thaliana* non-lysosomal β -glucosidases At1GH116, At3GH116, At4GH116 and At5GH116 with the eukaryotic plant β -glucosidases from rice, earth moss, and human and a prokaryotic enzyme from a thermophilic bacteria (TxGH116) (Charoenwattanasatien et al., 2016). At3GH116 gene is the most conserved to moss and the other AtGH116 enzymes group together with a few rice genes. Also, it makes it clear that AtGH116 all have common ancestor with the animal enzyme, as well as the animals and plants have one common ancestor with the bacteria.

Cyanobacteria group with other bacteria (TxGH116), not with plant sequences. It shows that the sequences of plants and animals diverged from bacteria at the same time and did not come from the cyanobacterium that became the chloroplast. At1g33700, At4g10060 and At5g49900 are closely related to OsBAF27951, while At3GH116 clustered with other rice β -glucosidases, such as OsAAP54244.2 and OsEEC81952.1. At1GH116, At3GH116, At4GH116 and At5GH116 shared 44-71% amino acid sequence identity with each other, with maximum sequence identity of 71% between At4GH116 and At1GH116, suggest that these two genes might have redundant functions. At4GH116 shares approximately 35% and 32% sequence identity with human and TxGH116 GBA, respectively.

Generally GH116 follows a two-step double displacement mechanism, which involves the formation and hydrolysis of a glycosyl intermediate. In the active site of

GH, conserved glutamic or aspartic acid residues generally serve as nucleophilic amino acids in acid-base catalysis. Alignment of the amino acid sequences of AtGH116, TxGH116 and HsGH116 has been useful in defining the basic catalytic machinery, including the general acid/base and catalytic nucleophile. Glutamic acid 499 (E499) and aspartic acid 647 (D647) of At4GH116, which is equivalent to human GBA2 E527 and D677 and TxGH116 E441 (E426 in the crystal structure) and D593 (D578 in the crystal structure) were highly conserved (Figure 4.2). These corresponding residues were previously shown to be the catalytic acid/base and nucleophile in human GBA2 (Kallemeijn et al., 2014) and TxGH116 (Charoenwattanasatien et al., 2016), strongly suggesting that E499 is the catalytic nucleophile and D647 the catalytic acid/base residue for At4GH116.



4.2 Cloning of AtGH116

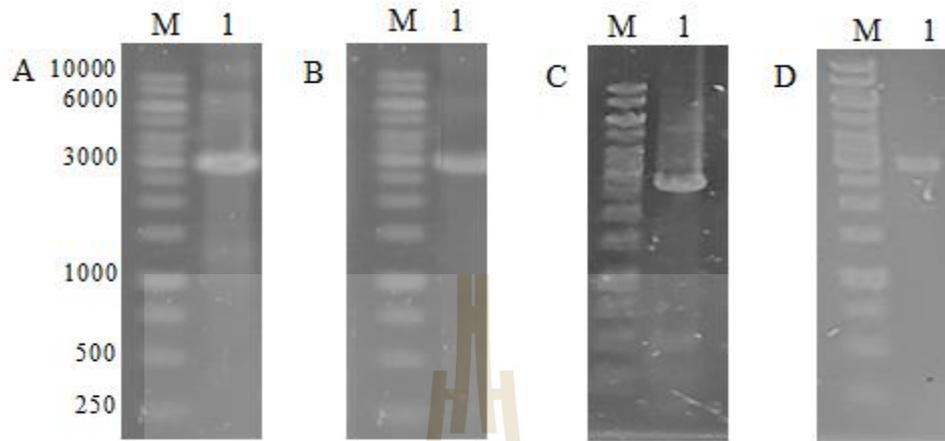


Figure 4.3 Amplification of a cDNA encoding mature AtGH116 proteins including A) At1GH116, B) At3GH116, C) At4GH116, and D) At5GH116. The PCR products were separated on 1% agarose gel electrophoresis and stained with ethidium bromide. Lane M, Thermo Scientific Gene ruler 1kb DNA ladder and Lane 1, in A, B, C and D, a PCR products of a cDNA encoding mature AtGH116s, At1GH116, At3GH116, At4GH116 and At5GH116 respectively,

Preliminary efforts to clone the AtGH116's genes from the *Arabidopsis* seedling RNA were unsuccessful. The cDNA clones of RAFL22-05-G12, RAFL09-37-G13, RAFL09-27-E09 and RAFL09-66-C11 from the RIKEN Biological Resource Center, Japan, and, used for amplification of coding sequence. The PCR products were obtained with sizes around 3 kb from cDNA clones (Figure 4.3) (TAIR Accession number BP813105.1, AK226971.1, AV801203.1 and BT002482.1) were ligated into pET30 expression vector and designated as pET30/At1GH116, pET30/At3GH116, pET30/At4GH116, and pET30/At5GH116, respectively.

Nucleotide and protein blast of sequencing results of pET30/At1GH116, pET30/At3GH116, pET30/At4GH116, and pET30/At5GH116 exhibited the maximum percent of similarity to the NCBI nucleotide Genbank accession number of BT004204 (RAFL16-03-J18), AK226971.1, BT001085.1 and BT002482.1 respectively, as well as to their respective protein accession number AAO42222.1, BAE99039.1, AAN46866.1, and AAO00842. Closer inspection indicated the correct sequences were obtained.

4.3 Expression of AtGH116

In an attempt to characterize the AtGH116 proteins all the four recombinant vectors were separately transformed into BL21(DE3). Protein expression was induced by addition of 0 to 0.6 mM of IPTG at 20°C. SDS-PAGE analysis revealed that only At4GH116 was expressed in BL21(DE3), with the best expression when it was induced by 0.4 mM IPTG at 20°C for 16 hr. The remaining three unexpressed clones were transformed into Rosetta-gami(DE3), but none of the clones were expressed in either BL21(DE3) or Rosetta-gami(DE3) bacterial strains under the conditions tested (Figure 4.4-4.7).

Among the available heterologous protein expression systems *E. coli* is popular, due to many useful characteristics, such as growing rapidly at high density on inexpensive substrates, and the availability of an increasingly large number of cloning and expression vectors and mutant host strains (Makrides, S.C. 1996). Many plant GH family enzymes were expressed in *E. coli* and soluble, active forms of the proteins were obtained (Luang et al., 2013, Opassiri et al, 2006, Hua et al., 2013). Heterologous expressions of At1GH116, At3GH116 and At5GH116 were not

successful. Possible explanations for this issue are that the expressed protein may be toxic after the induction and codon bias may cause incomplete translation (Rosano and Ceccarelli, 2014).

Stability of the recombinant protein is very important in case of production, purification, formulation and storage. Various aspects affecting protein instability are like amino acid sequences of the protein, protein construction, host cell strain, expression and purification conditions will all affect the stability of the protein (Joseph et al., 2015).

There are instances where the amino acid sequence of a protein itself is prone to degradation. Certain amino acids like arginine, lysine, leucine, phenylalanine, tyrosine and tryptophan residues at the N-terminus region can lead to protein degradation. Replacing these amino acids with compatible amino acids can greatly enhance the protein stability. Expression of recombinant proteins with tags or fusion partners prevents proteolytic degradation and increase the stability. Medium pH should also be neutral to improve the stability of the protein. Induction of protein at lower temperature and for shorter duration will enhance the protein stability. Sometimes, changing the expression host also can result in getting the stable protein (Hosur et al., 2014).

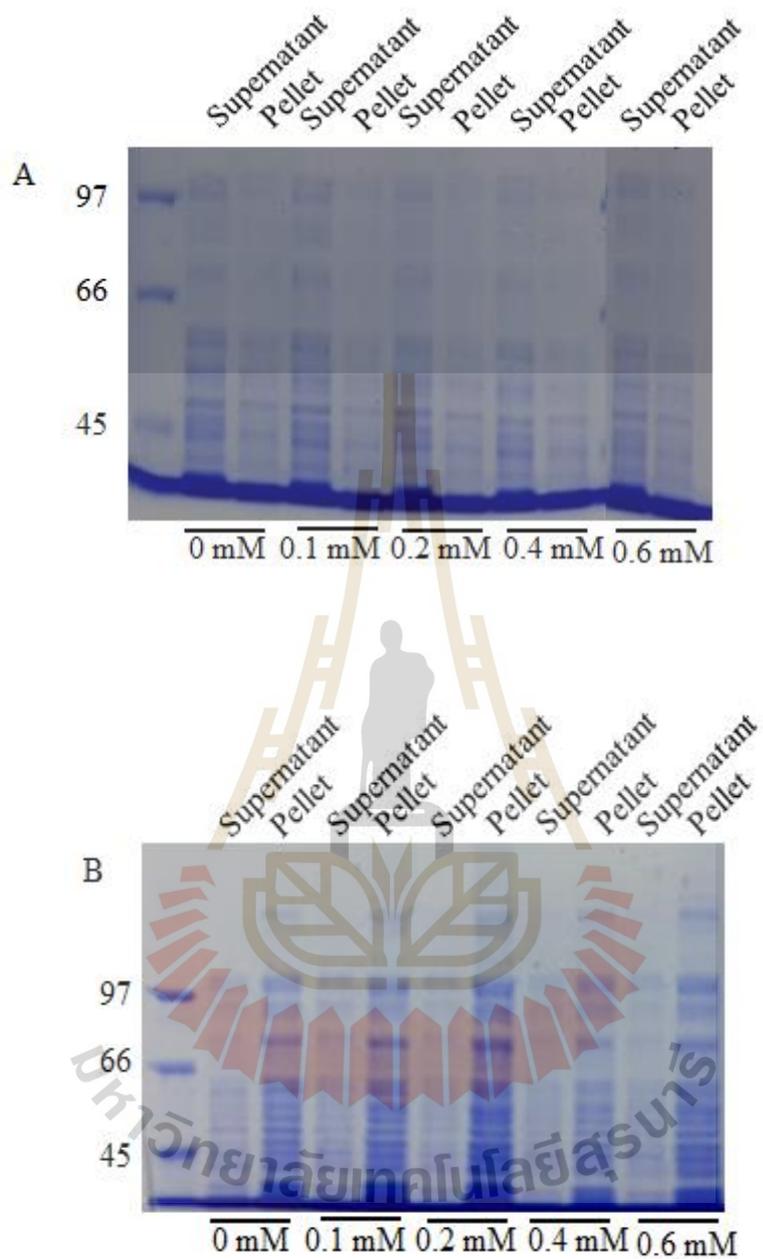


Figure 4.4 SDS-PAGE analysis of pET30/At1GH116 expression in *E. coli*. A) BL21(DE3) and B) Rosetta-gami(DE3) host cells were used for expression. The cells were induced with 0-0.6 mM IPTG at 20°C for 16 hr.

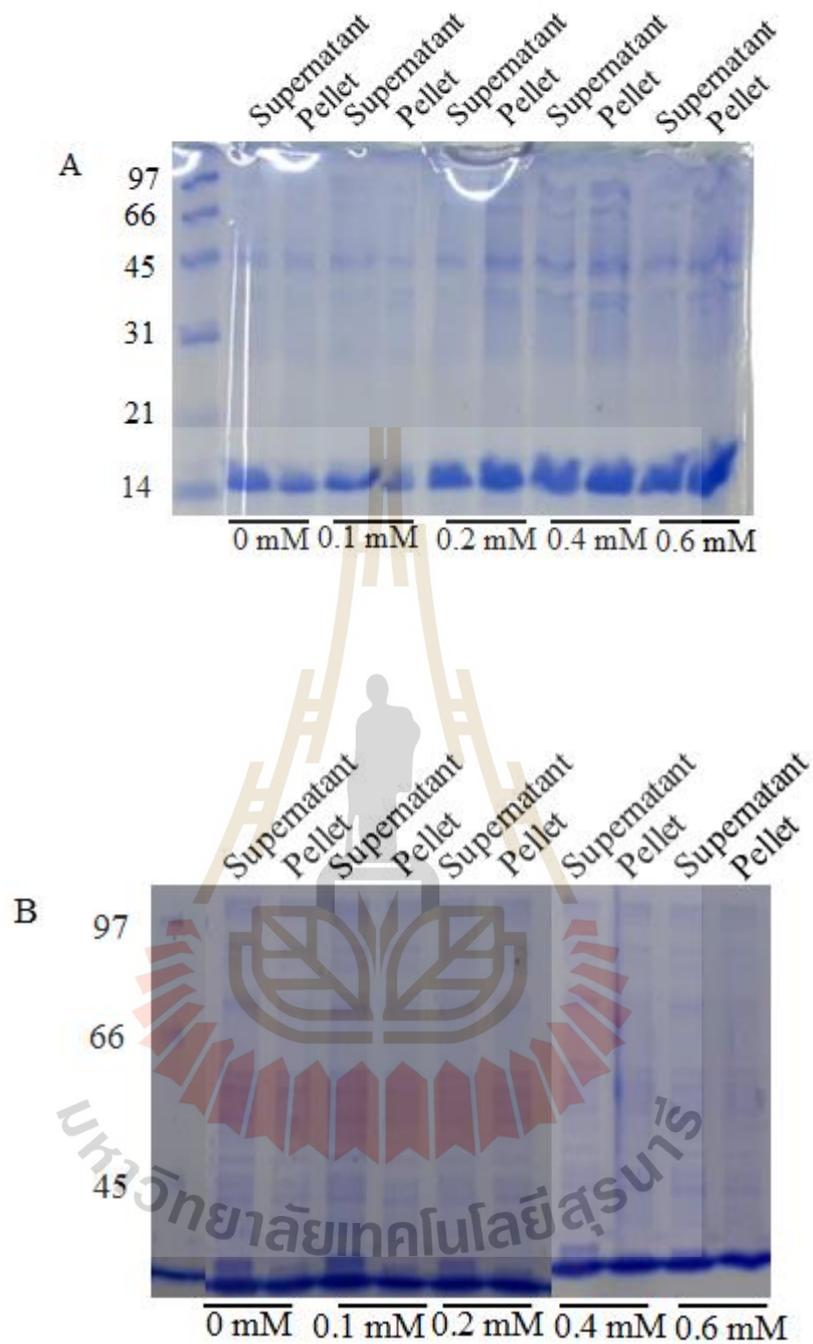


Figure 4.5 SDS-PAGE analysis of pET30/At3GH116 expression in *E. coli*. A) BL21(DE3) and B) Rosetta-gami(DE3) host cells were used for expression. The cells were induced with 0-0.6 mM IPTG at 20°C for 16 hr.

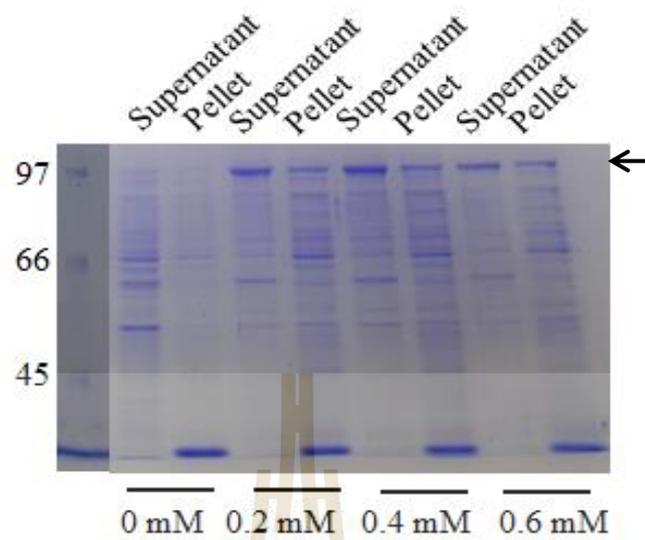


Figure 4.6 SDS-PAGE analysis of pET30/At4GH116 expression in *E. coli*. BL21(DE3) host cell was used for expression. The cells were induced with 0-0.6 mM IPTG at 20°C for 16 hr. The band of expressed protein is indicated by the arrow.

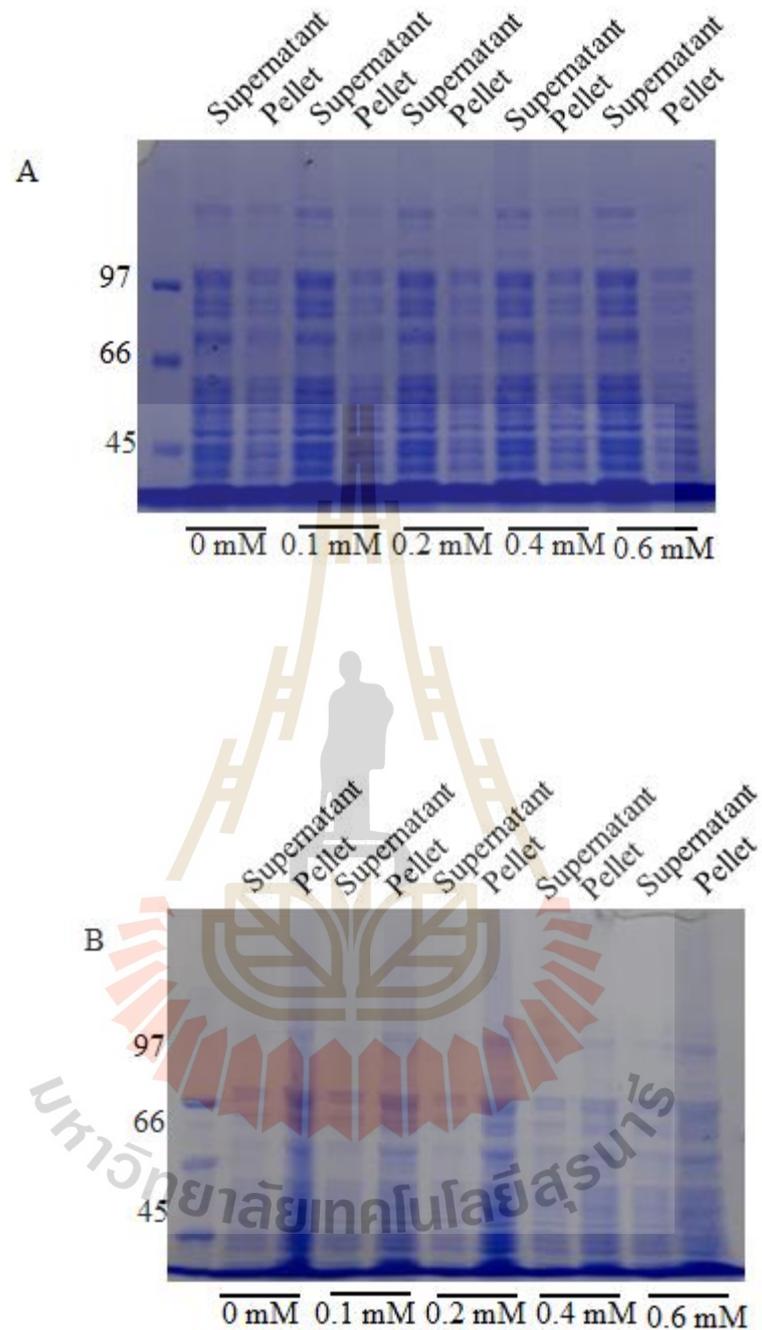


Figure 4.7 SDS-PAGE analysis of pET30/At5GH116 expression in *E. coli*. A) BL21(DE3) and B) Rosetta-gami(DE3) host cells were used for expression. The cells were induced with 0-0.6 mM IPTG at 20°C for 16 hr.

4.4 Purification of At4GH116

When expressed in *E. coli* strain BL21(DE3) cells, At4GH116 was produced as a soluble fusion protein with a Tobacco Etch Virus nuclear-inclusion-a endopeptidase (TEV) cleavage site following the S-tag by a six histidine tag (His-tag) at the N-terminus, with an expected molecular weight (MW) of 104 kDa, similar to that seen on the Coomassie-stained SDS-PAGE gel (Figure 4.8). During extraction, the protein was protected from proteolysis, by the addition of the protease inhibitors PMSF, aminocarbonic acid and benzamidine to the extraction buffer. Recombinant At4GH116 fusion protein was eluted from immobilized metal affinity chromatography (IMAC) with Ni²⁺ resin with 500 mM imidazole. The cleaved N-terminal fusion tag and TEV protease were removed by adsorption to Ni²⁺ IMAC resin. The flow-through solution and the active elution fractions with 0 mM imidazole, containing pNP-β-D-glucopyranoside hydrolysis activity were pooled and concentrated.

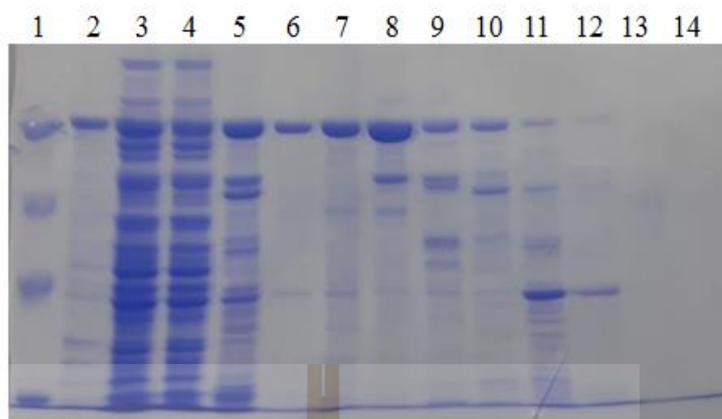


Figure 4.8 SDS-PAGE analysis of At4GH116 purification by immobilized metal affinity chromatography (IMAC), followed by S200 size exclusion chromatography (SEC). Two μg of protein was loaded in the Lane 5-14. Lane 1) Bio-Rad low molecular weight markers, Lane 2) insoluble protein from BL21(DE3) expression, Lane 3) soluble protein from BL21(DE3) expression, Lane 4) flow-through fraction after passing pET30/At4GH116 cell lysate through the IMAC column, Lane 5) pooled fractions from IMAC elution with 500 mM imidazole, Lane 6) flow-through fraction after the TEV digest was passed through the IMAC column, Lane 7) pooled fractions 16-17 from S200 SEC, Lane 8) pooled fractions 20-21 from S200 SEC, Lane 9) pooled fractions 23-24 from S200 SEC, Lane 10) pooled fractions 26 from S200 SEC, Lane 11) pooled fractions 29-30 from S200 SEC, Lane 12) pooled fractions 34-35 from S200 SEC, Lane 13) pooled fractions from 39-40 S200 SEC, Lane 14) pooled fractions from 42-43 S200 SEC.

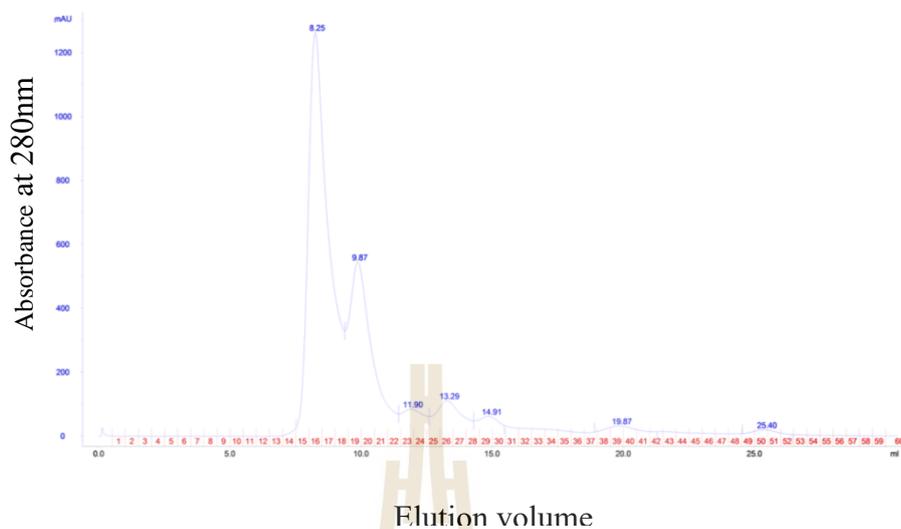


Figure 4.9 S200 Size exclusion chromatogram of At4GH116 expressed in BL21(DE3) host cells. The protein fraction numbers are mentioned in red.

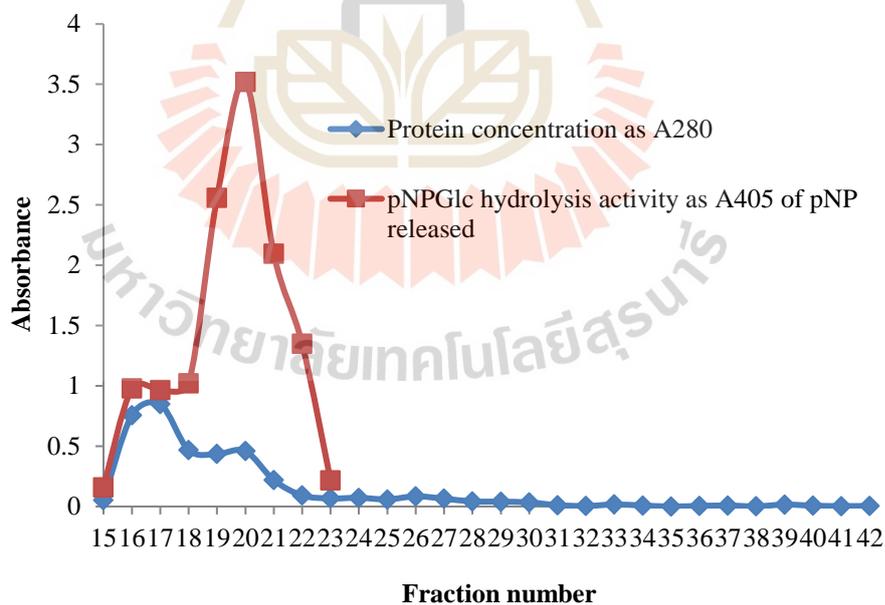


Figure 4.10 pNPGlc hydrolase activity of fractions of At4GH116 from the S200 Size exclusion chromatography. Ten microl of fractions were used to analyses the pNPGlc hydrolase activity.

Concentrated proteins from the 2nd IMAC purification were further purified through Superdex 200 size exclusion chromatography (SEC) (Figure 4.9). Fractions were separated and pooled together based on their native size, as shown in the Table 3.3, and concentrated separately. Concentrated protein fractions were stored at -80°C in aliquots. The protein stable was during storage, as the activity of stored the protein was consistent over the time period.

The elution of peaks I and II from S200 SEC were consistent throughout all the SEC purifications of individual injection and yield was adequate for further experiments with high purity. They were used for all the kinetic analysis. Peak III was inconsistent and little throughout the gel filtration, so it has been omitted from the further studies and pooled fraction after this one was named as peak III, because it contained increased β -glucosidase activity (Figure 4.10). Along with peaks I and II, peak III was used for small angle X-ray scattering (SAXS). Either peak I or II was used for substrate specificity and kinetic analysis. Both peaks I and II were used for crystallization.

The molecular weight differences in the peaks of the At4GH116 due to oligomerization were observed in the native gel electrophoresis (Figure 4.11). The At4GH116 purified by 1st IMAC and peak I of S200 SEC of At4GH116 stayed inside the well of gel itself and couldn't move at all, but the from the 2nd IMAC migrated little below the well. In contrast, peak II of S200 SEC was moved to the middle of the gel. This result was further emphasized by overlaying the S200 SEC chromatogram of At4GH116 onto the chromatograms of various protein MW standards (Figure 4.12).

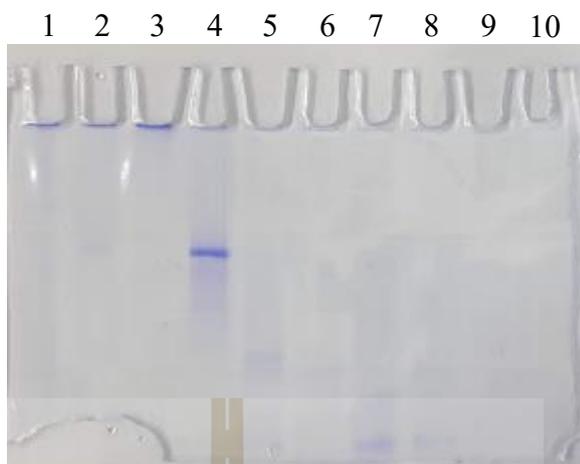


Figure 4.11 Native-PAGE analysis of At4GH116 purified by immobilized affinity chromatography (IMAC) followed by S200 SEC. Lane 1) fractions from IMAC elution with 500 mM imidazole, 2) flow-through fraction after TEV digests through the IMAC column, 3) pooled fractions 16-17 from S200 SEC, 4) pooled fractions 20-21 from S200 SEC, 5) pooled fractions 23-24 from S200 SEC, 6) fraction 26 from S200 SEC, 7) pooled fractions 29-30 from S200 SEC, 8) pooled fractions 34-35 from S200 SEC, 9) pooled fractions 39-40 from S200 SEC, 10) pooled fractions 42-43 from S200 SEC.

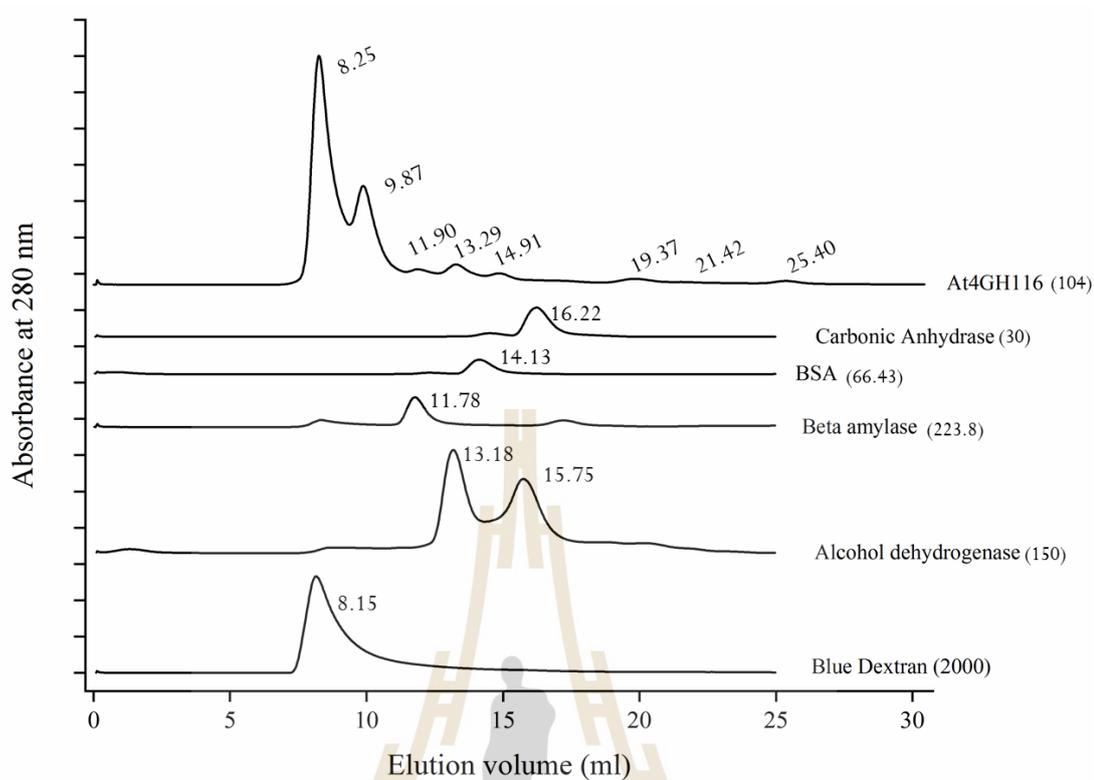


Figure 4.12 Super imposition of S200 size exclusion chromatogram of At4GH116 expressed in *E. coli* BL21(DE3) host cells into chromatogram of different protein standards monitoring A_{280} . Chromatogram of individual protein standards were shown with molecular weight (KDa) in bracket.

Sonication of the cells increases the yield of protein. At the same time, plenty of insoluble protein was observed in the pellets. The recombinant protein expressed in the environment of *E. coli*, may differ from their original source in terms of pH, osmolarity, redox potential, cofactors, and folding mechanisms. All of these factors lead to protein instability and aggregation (Rosano and Ceccarelli, 2014). Soluble His-tagged At4GH116 appeared to elute in the flow-through (FT) of 1st IMAC. Reloading the FT twice again into the column and elution of protein helped to recover

the protein and decreased loss. This was confirmed by the loss of pNPGlc hydrolysis activity from the flow-through solution after reloading. Five milligrams of purified proteins per liter of *E. coli* culture was obtained. Protein purification by IMAC and S200 SEC produced a protein with approximately 90% purity, as judged by SDS-PAGE.

4.5 Substrate specificity of At4GH116

The substrate specificities of the purified recombinant At4GH116 enzyme was analyzed with different *para* and *ortho* nitrophenyl (NP) glycosides as shown in Table 4.1. At4GH116 hydrolysis of β -D-glucoside was designated 100%. At4GH116 had lowest activity to *p*NP- β -D-fucopyranoside and no activity for *p*NP-N-acetyl- β -D-glucosamine and *p*NP- β -D-mannoside. Interestingly, At4GH116 had a maximum activity towards *p*NP- β -D-galactopyranoside when the substrates were compared at 1 mM concentration.

Table 4.1 Relative activities of At4GH16 towards *p*NP and *o*NP glucosides

Substrates	Relative activity (%)
<i>p</i> NP- β -D-galactopyranoside	246 \pm 0.008
<i>p</i> NP- β -D-glucopyranoside	100 \pm 0.014
<i>o</i> NP- β -D-galactopyranoside	54.7 \pm 0.004
<i>p</i> NP- α -L-arabinofuranoside	33.4 \pm 0.002
<i>o</i> NP- β -D-glucopyranoside	16.8 \pm 0.012
<i>p</i> NP- β -D-xylopyranoside	11.0 \pm 0.002
<i>p</i> NP- β -D-cellobioside	5.29 \pm 0.004
<i>p</i> NP- α -D-glucopyranoside	3.47 \pm 0.002
<i>p</i> NP- α -L-arabinopyranoside	3.10 \pm 0.005
<i>p</i> NP- β -D-fucopyranoside	2.74 \pm 0.002
<i>p</i> NP- α -D-galactopyranoside	1.09 \pm 0.001
<i>p</i> NP- α -D-mannopyranoside	0.36 \pm 0.003
<i>p</i> NP- β -D-mannopyranoside	0.00 \pm 0.002
<i>p</i> NP-N-acetyl- β -D-glucosamine	0.00 \pm 0.00

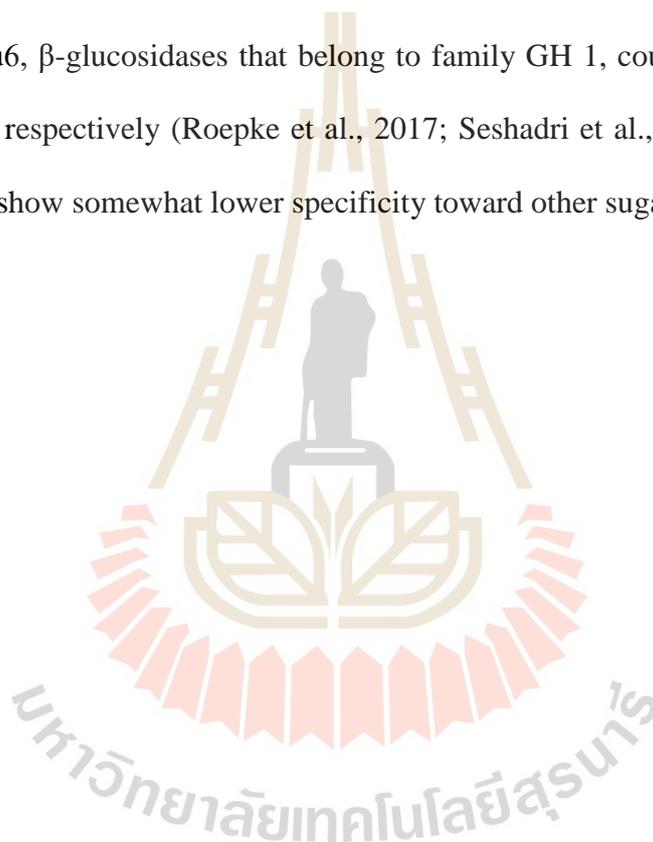
1 mM final concentrations of substrates were used and assayed at 30°C for 30 min.

Table 4.2 Kinetic parameters of At4GH116 for hydrolysis of *para*-nitrophenyl glycosides.

Substrates	Kinetic Parameters		
	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (mM ⁻¹ s ⁻¹)
<i>p</i> NP- β -D-glucopyranoside	0.077 \pm 0.005	1.35 \pm 0.008	17.4
<i>p</i> NP- β -D-galactopyranoside	0.264 \pm 0.016	2.43 \pm 0.027	9.2
<i>p</i> NP- α -L-arabinofuranoside	0.795 \pm 0.058	0.37 \pm 0.010	0.48
<i>p</i> NP- β -D-xylopyranoside	0.516 \pm 0.045	0.043 \pm 0.001	0.084

Primarily, glycone specificity was tested by assaying towards different α and β -nitrophenyl glycosides. Among all the tested nitrophenyl glycosides, At4GH116 hydrolyzed *p*NP- β -D-galactopyranoside (*p*NPGal), prominently followed by *p*NP- β -D-glucopyranoside (*p*NPGlc), *p*NP- α -L arabinofuranoside (*p*NP α Araf) and *p*NP- β -D-xylopyranoside (*p*NPXylp), so these four substrates were used for kinetic analysis. Kinetic analysis revealed At4GH116 hydrolyzed *p*NPGlc ($k_{cat}/K_m = 17.41$ mM⁻¹s⁻¹) and *p*NPGal ($k_{cat}/K_m = 9.22$ mM⁻¹s⁻¹) efficiently and had relatively little specificity toward other *p*NP-glycosides, such as *p*NP α Araf and *p*NPXylp, as shown in Table 4.3.

At4GH116 showed higher catalytic efficiency with *p*NPGlc and *p*NPGal, which were hydrolyzed with over 30-fold higher specificity constants (k_{cat}/K_m) than *p*NP- α -L arabinofuranoside and *p*NP- β -D-xylopyranoside. Similarly, other known *Tx*GH116 β -glucosidases, and *GH5BG*, a stress-induced GH5 (glycosyl hydrolase family 5) β -glucosidase from rice, are able to hydrolyze *p*NPGlc, *p*NPGal, *p*NP α Araf and *p*NPXylo (Charoenwattanasatien et al., 2016; Opassiri et al., 2007). AtBGLU15 and Os3BGlu6, β -glucosidases that belong to family GH 1, could hydrolyse *p*NPGlc and *p*NPGal, respectively (Roepke et al., 2017; Seshadri et al., 2006). So, several β -glucosidases show somewhat lower specificity toward other sugars.



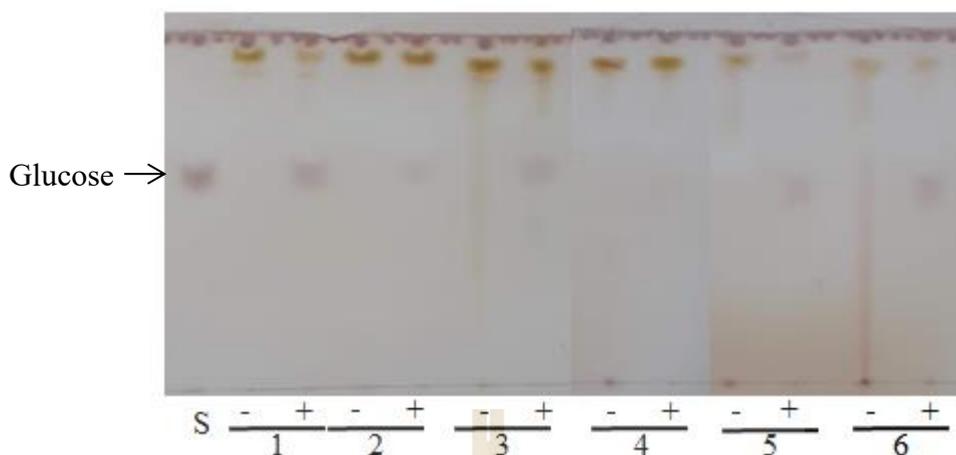


Figure 4.13 TLC analysis of hydrolysis of flavonoid glucosides by At4GH116.

S) glucose, 1) kaempferol-7-O- β -D-glucoside, 2) kaempferol-3-O- β -D-glucoside, 3) quercetin-7-O- β -D-glucoside, 4) quercetin-3-O- β -D-glucoside, 5) apigenin-7-O- β -D-glucoside, 6) luteolin-7-O- β -D-glucoside. Hydrolysis reactions were assessed by overnight incubation of 1 μ g of purified enzyme with 1 mM substrates in 50 mM MES buffer, pH 6.0, at 30°C. The reactions were analyzed on silica gel 60 F₂₅₄ plates with a solvent system of ethyl acetate:acetic acid:water (2:1:1). '+' and '-' indicate the presence and absence of At4GH116 enzyme.

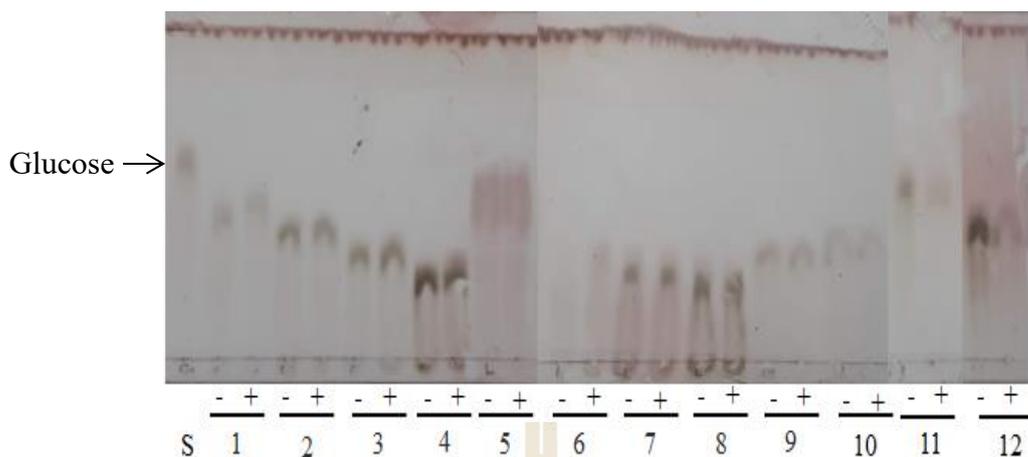


Figure 4.14 TLC analyses of hydrolysis of oligosaccharides by At4GH116. S) glucose, 1) cellobiose, 2) cellotriose, 3) cellotetraose, 4) cellopentaose, 5) laminaribiose, 6) laminaritriose, 7) laminaritetraose, 8) laminaripentaose, 9) melibiose, 10) sophorose, 11) sucrose, 12) gentibiose. Hydrolysis reactions were assessed by overnight incubation of 1 μ g of purified enzyme with 1 mM substrates in 50 mM MES buffer, pH 6.0, at 30°C. The reactions were analyzed on silica gel 60 F₂₅₄ plates with a solvent system of ethyl acetate:acetic acid:water (2:1:1). ‘+’ and ‘-’ indicate the presence and absence of At4GH116 enzyme.

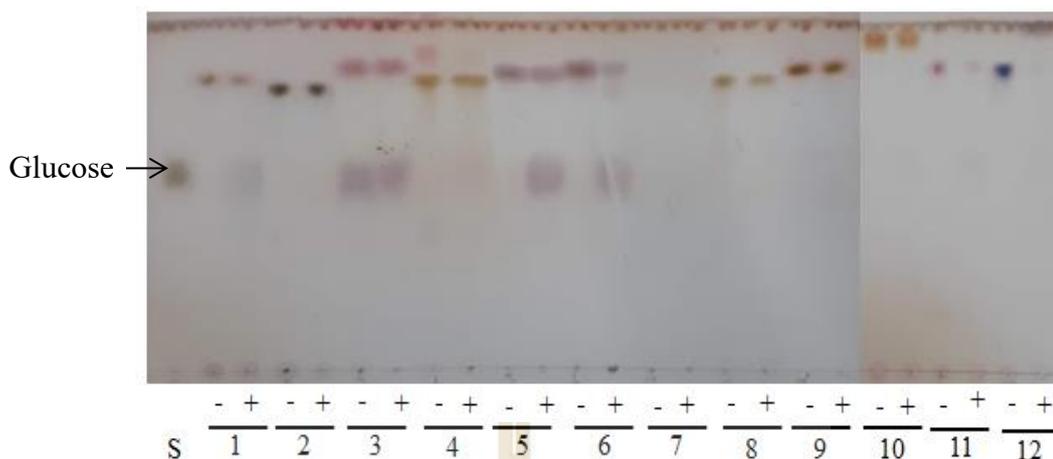


Figure 4.15 TLC analyses of hydrolysis of glucose conjugates by At4GH116.

S) glucose, 1) arbutin 2) D - amygdalin, 3) dhurrin, 4) gossypin, 5) helicin, 6) indoxy β -D-glucosides, 7) lichenan, 8) mangiferin, 9) narigin, 10) phlorizin, 11) salicin, 12) syringin. Hydrolysis reactions were assessed by overnight incubation of 1 μ g of purified enzyme with 1 mM substrates in 50 mM MES buffer, pH 6.0, at 30°C. The reactions were analyzed on silica gel 60 F₂₅₄ plates with a solvent system of ethyl acetate:acetic acid:water (2:1:1). '+' and '-' indicate the presence and absence of At4GH116 enzyme.

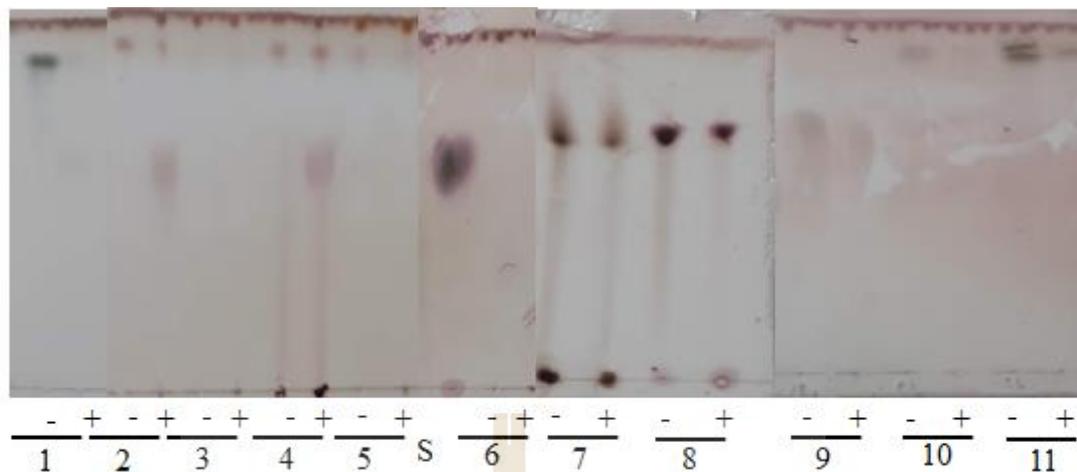


Figure 4.16 TLC analyses of hydrolysis of glucose conjugates by At4GH116.

1) *p*-Coumarol glucoside, 2) daidzin 3) genistin, 4) glycitin 5) hesperitin, S) glucose ,
 6) C16 (β) glucosylceramide, 7) monogalactosyldiacylglycerol,
 8) digalactosyldiacylglycerol, 9) methyl β-D-glucoside, 10) *n*-heptyl-β-D-glucoside,
 11) *n*-octyl-β-D-glucoside. Hydrolysis reactions were assessed by overnight incubation of 1 μg of purified enzyme with 1 mM substrates in 50 mM MES buffer, pH 6.0, at 30°C. The reactions were analyzed on silica gel 60 F₂₅₄ plates with a solvent system of ethyl acetate:acetic acid:water (2:1:1). '+' and '-' indicate the presence and absence of At4GH116 enzyme.

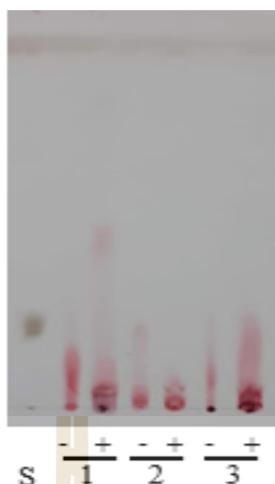


Figure 4.17 TLC analyses of hydrolysis of anthocyanin glucosides by At4GH116. 1) cyanidin-3-O- β -D-glucoside, 2) cyanidin-3,5-O- β -D-glucoside, 3) peonidin-3,5-O- β -D-glucoside. Hydrolysis reactions were assessed by overnight incubation of 1 μ g of purified enzyme with 1 mM substrates in 50 mM MES buffer, pH 6.0, at 30°C. The reactions were analyzed on silica gel 60 F₂₅₄ plates with a solvent system of ethyl acetate:acetic acid:water (7:2.5:0.5). '+' and '-' indicate the presence and absence of At4GH116 enzyme.

Table 4.3 Summary of the activity of At4GH116 toward natural and synthetic glycosides based on TLC analysis.

Natural and synthetic glucosides	Activity
Monolignol glucosides	
<i>p</i> -Coumarol glucoside	-
Syringin	-
Flavonols glucosides	
Kaempherol-3-O- β -D-glucoside	+
Kaempherol-7-O- β -D-glucoside	+
Quercetin-3-O- β -D-glucoside	+
Quercetin-7-O- β -D-glucoside	+
Flavonone glucosides	
Naringin	-
Hesperitin	-
Phlorizin	-
Flavone glucosides	
Luteolin-7-O- β -D-glucoside	+
Apigenin-7-O- β -D-glucoside	+
Isoflavonols glucosides	
Genistin	-
Glycitin	+
Daidzin	+
Chalcone glucosides	
Arbutin	+

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

Table 4.3 Summary of the activity of At4GH116 toward natural and synthetic glycosides based on TLC analysis (Continued).

Natural and synthetic glycosides	Activity
Salicylaldehyde derivatives	
Helicin	+
Salicin	-
Alkyl glycosides	
Methyl- β -D-glucoside	-
<i>n</i> -Octyl- β -D-glucoside	-
<i>n</i> -Heptyl- β -D-glucoside	-
Oligosaccharides	
Sucrose	-
Sophorose	-
Melibiose	-
Gentibiose	-
Cello oligosaccharides	-
Laminari oligosaccharides	-
Lichenan (polysaccharide)	-
Anthocyanin glucoside	
Cyanidin-3-O- β -D glucoside	-
Cyanidin-3,5-O- β -D glucoside	-
Peonidin-3,5-O- β -D glucoside	-

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

Table 4.3 Summary of the activity of ATGH116 toward natural and synthetic glucosides based on TLC analysis (Continued).

Natural and synthetic glucosides	Activity
Other natural glucosides	
Indoxyl β -D-glucoside	+
D-Amygdalin	-
Dhurrin	-
Esculin	-
Mangiferin	-
Gossypin	-
Glycolipids	
Monogalactosyldiacylglycerol	-
Digalactosyldiacylglycerol	-
C16- β -glucosylceramide	-

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

Further aglycones specificities were investigated by testing the At4GH116 activity towards different natural and synthetic glucosides by TLC (Figures 4.12 - Figure 4.16). At4GH116 exhibited hydrolyzing activity towards flavonoid glucosides. Qualitative analysis revealed that flavonols glucosides, such as kaempferol-7-O- β -D glucoside, kaempferol-3-O- β -D glucoside, and quercetin-7-O- β -D glucoside and isoflavonol glucosides, such as glycitin and daidzin, the flavone glucosides luteolin-7-O- β -D glucoside and apigenin-7-O- β -D glucoside were hydrolyzed, along with other natural glucosides: helicin, arbutin and indoxyl β -D-glucoside.

So far, no hydrolyzing activity of AtGH116 towards natural glucosides except glycosphingolipids was reported. However other commercial and recombinant β -glucosidases, such as that from the hyperthermophilic archaeon *Pyrococcus furiosus* and that from the tree *Dalbergia nigrescens* hydrolyzed isoflavone glucosides of glycitin, daidzin and genistin (Richelle et al., 2002; Chuankhayan et al., 2007; Yeom et al., 2012). Same as like hydrolyze activity of recombinant Os4BGlu13 and β -glucosidase, from *Polygonum tinctorium* towards helicin and indoxyl- β -D-glucoside respectively (Hua et al., 2015; Minami et al., 2014).



Table 4.4 Kinetic parameters of At4GH116 for hydrolysis of flavonoid glycosides.

Substrates	Kinetic Parameters		
	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (mM ⁻¹ s ⁻¹)
Apigenin-7-O- β -D-glucoside	0.049 \pm 0.003	56.5 \pm 0.76	1133
Quercetin-7-O- β -D-glucoside	0.048 \pm 0.003	34.9 \pm 0.53	735
Kaempherol-7-O- β -D-glucoside	0.077 \pm 0.008	17 \pm 0.39	220
Luteolin-7-O- β -D-glucoside	0.125 \pm 0.009	13.3 \pm 0.18	106
Kaempherol-3-O- β -D-glucoside	0.037 \pm 0.002	1.98 \pm 0.02	54.3

Kinetic analysis showed that At4GH116 has high specificity for apigenin-7-O- β -D-glucoside ($k_{cat}/K_m = 1133 \text{ mM}^{-1}\text{s}^{-1}$), quercetin-7-O- β -D-glucoside ($k_{cat}/K_m = 735 \text{ mM}^{-1}\text{s}^{-1}$), kaempherol-7-O- β -D-glucoside ($k_{cat}/K_m = 220 \text{ mM}^{-1}\text{s}^{-1}$), and luteolin-7-O- β -D-glucoside ($k_{cat}/K_m = 106 \text{ mM}^{-1}\text{s}^{-1}$) and had less activity toward kaempherol-3-O- β -D-glucoside. Similarly, other β -glucosidases from almonds and propolis exhibited hydrolysis activity towards 7-O- β -glucosides of apigenin and luteolin (Pekić et al., 1994; Zhang et al., 2012).

4.6 Inhibition of At4GH116 by organic inhibitors

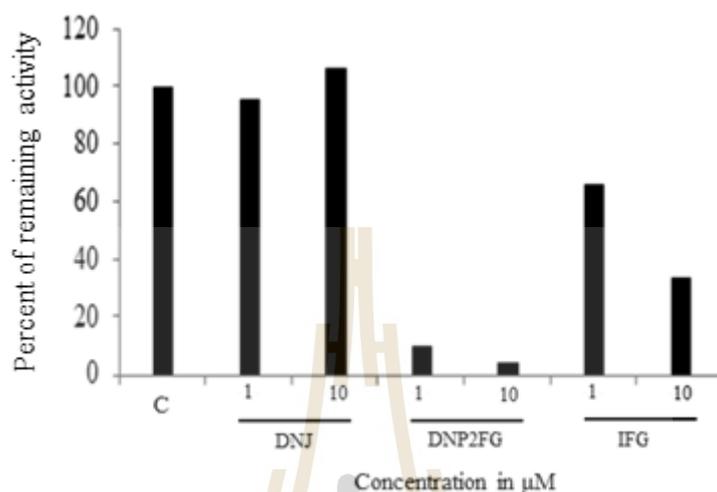


Figure 4.18 Effect of inhibitors on the enzyme activity of At4GH116. DNJ represents 1-Deoxynojirimycin, DNP2FG represents 2, 4-Dinitrophenyl- β -D-2-deoxy-2-fluoroglucopyranoside, IFG represents isofagamine. One μ g of enzyme was preincubated for 30 mins with inhibitors at RT. Then hydrolysis reaction of At4GH116 was assessed with 1mM of *p*NPGlc in 50mM MES buffer, pH 6.0 at 30°C.

DNP2FG at 1 and 10 μ M could inhibit the activity of At4GH116 up to 95% and 97%, respectively, when pre-incubated for 30 mins. At4GH116 was partially inhibited by isofagamine at 1 and 10 μ M and deoxynojirimycin at 1 μ M exhibited poor inhibitory activity (Figure 4.18).

These results contrast to the finding of Overkleeft et al. (1998), which showed that human GBA2 was highly inhibited by an analogue of DNM, N-(5-adamantane-1-yl-methoxy) pentyl)-DNM (AMP-DNM) and N-(5-cholesteroxypentyl)-DNM (CP-DNM) more than DNM. Similarly, *Tx*GH116 is also inhibited by the covalent

inactivator DNP2FG, and by the noncovalent inhibitors IFG and DNJ (Charoenwattanasatien et al., 2016). Dai et al. (2020) also reported the inhibitory activity of N-butyldeoxynojirimycin (NB-DNJ) and CBE towards At4GH116. Their results showed that NB -DNJ inhibits AtGCD3, more effectively than CBE. NB-DNJ inhibits GBA2 but not GBA (Ridley et al., 2013) and CBE is a potent inhibitor of GBA.



4.7 Structural determination of At4GH116

4.7.1 Small angle X-ray diffraction

In order to gain a better understanding of molecular behavior and to determine the shape of the At4GH116 in solution, we performed Small angle X-ray scattering (SAXS). The scattering intensities $I(q)$ of S200 SEC peaks I, II and III of At4GH116 were obtained in the concentration of 1, 2, 5 and 10 mg/mL (Figure 4.19). The scattering patterns of peak I (2 mg/ml) and peak III (1 mg/ml) exhibited a polydispersity either by aggregation and/or oligomerization of protein at lower q . On the other hand, the scattering pattern of peak II at all the concentrations showed a well-defined diffraction pattern. Therefore, for the further studies peak II was used. A Kratky plot of the scattering data of peak II showed a well-defined, bell-shaped peak, suggesting that At4GH116 is a compact globular protein (Figure 4.20). This is established by a bell-shaped peak at low q . The linearity of the Guinier plot indicated that the peak II of protein solutions used in the SAXS analysis were monodispersed and did not contain a large amount of larger aggregates (Figure 4.21). The pair-distance distribution function, $P(r)$, is characteristic of globular particles. The $P(r)$, calculated from the scattering data, has a maximum distance (D_{\max}) range of 251 and 168 Å for lower and higher concentrations of protein, respectively (Figure 4.22). The R_g calculated on the basis of the Guinier plot was 66.71 to 71.26 Å (Figure 4.20). The estimated molecular mass of At4GH116 was approximately range ~ 327.8 to ~421.9 kDa, indicating that At4GH116 exists as a trimer or tetramer in peak II (Table 4.5). Further, the molecular weight obtained by porod volume (V_p) method was between ~ 254.7 to ~ 339.4 kDa, which is consistent with the estimate from the volume of correlation.

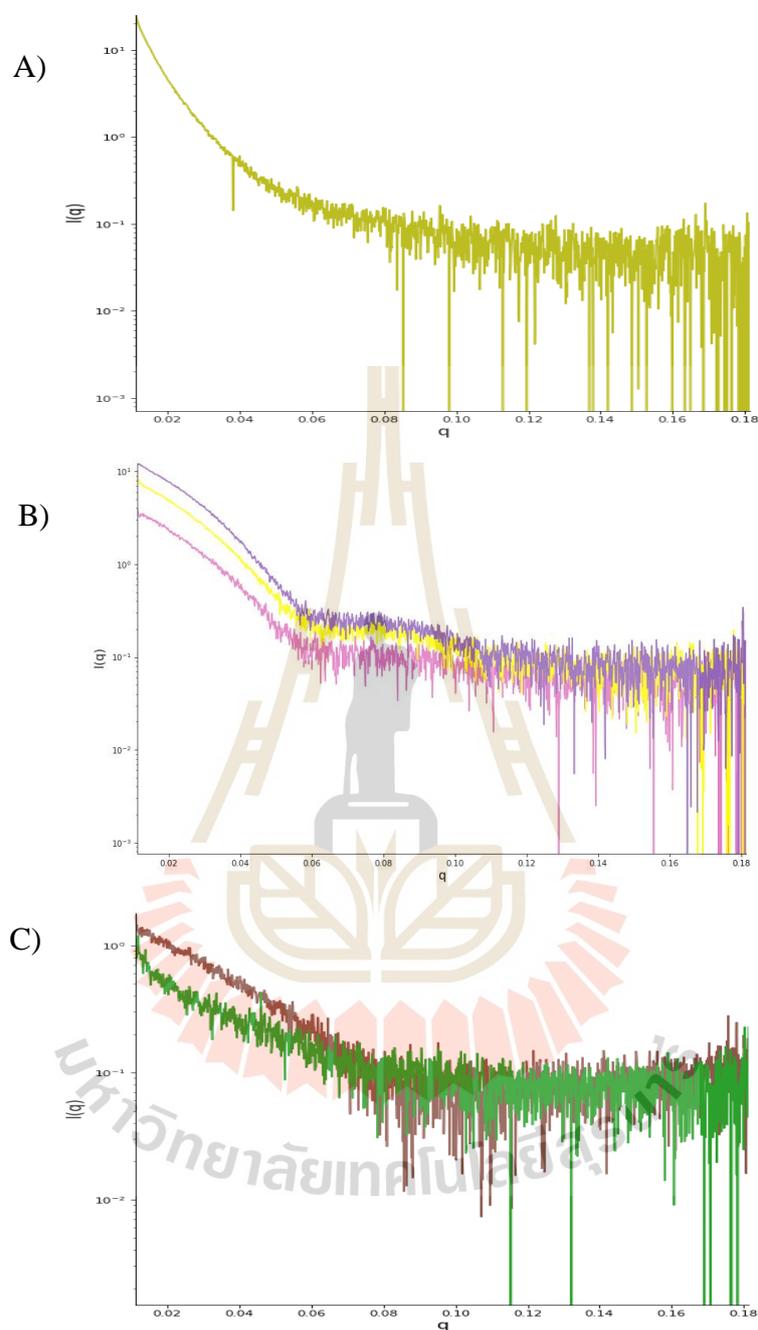


Figure 4.19 Small angle X-ray scattering profiles of At4GH116 S200 SEC. A) Peak I - 2 mg/ml, B) Peak II - purple 2 mg/ml, yellow 5 mg/ml, violet 10 mg/ml C) Peak III - brown - 2 mg/ml, green - 1 mg/ml.

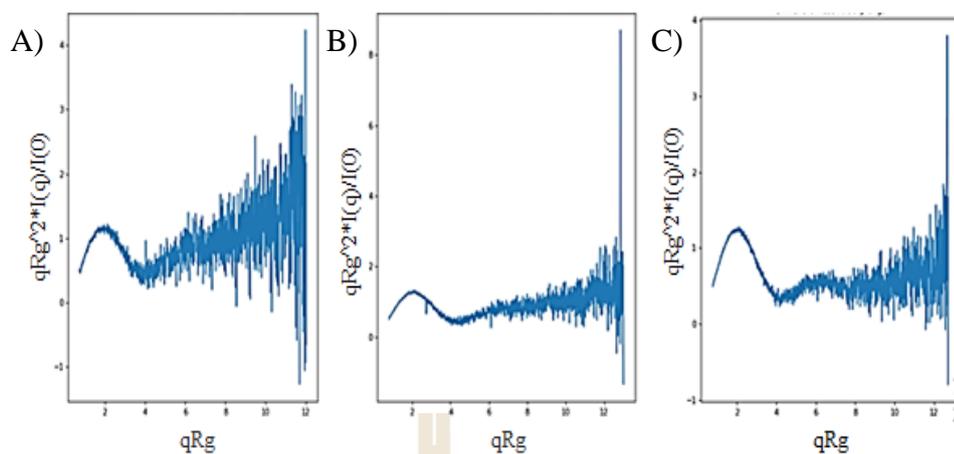


Figure 4.20 Kratky plot analysis of At4GH116 S200 SEC peak II. Protein concentrations are A) 2 mg/ml B) 5 mg/ml and C) 10 mg/ml.

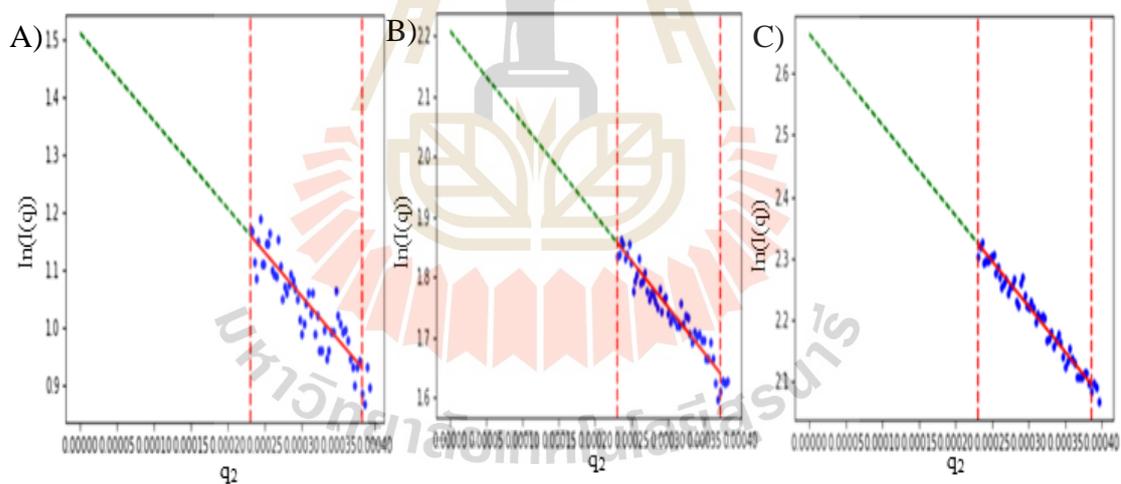


Figure 4.21 Guinier plot analysis of At4GH116 S200 SEC peak II. Protein concentrations are A) 2 mg/ml B) 5 mg/ml and C) 10 mg/ml.

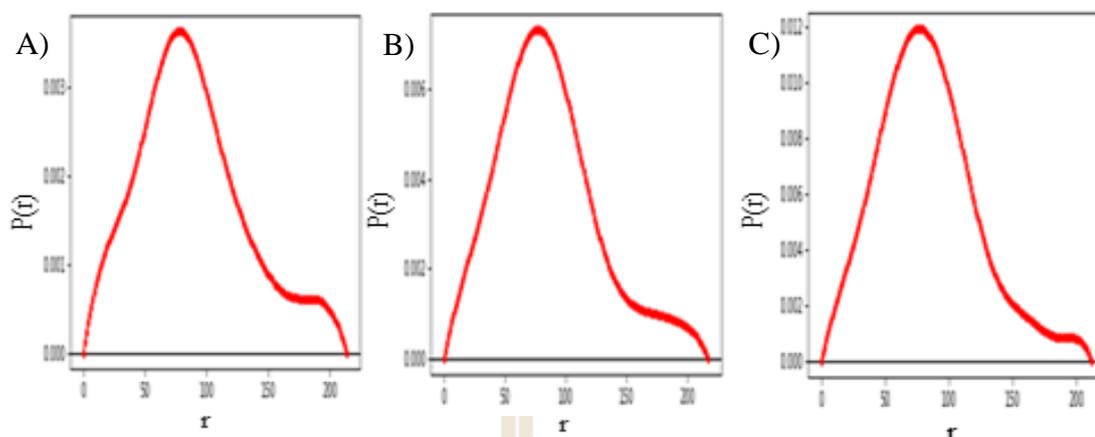


Figure 4.22 Pair distribution functions $P(r)$ At4GH116 S200 SEC peak II. Protein concentrations are A) 2, B) 5 and C) 10 mg/ml.

Table 4.5 A summary of R_g , D_{max} and molecular mass estimated from the Guinier plot.

Protein	Concentration (mg/ml)	Radius of gyration R_g (Å)	Molecular weight (MW)		Maximal dimension D_{max} (Å)
			Volume of correlation V_c	Porod volume V_p	
At4GH116	2	66.71	327.8	254.7	251
S200 SEC	5	68.20	421.9	339.4	202
Peak II	10	71.26	388.8	311.1	168

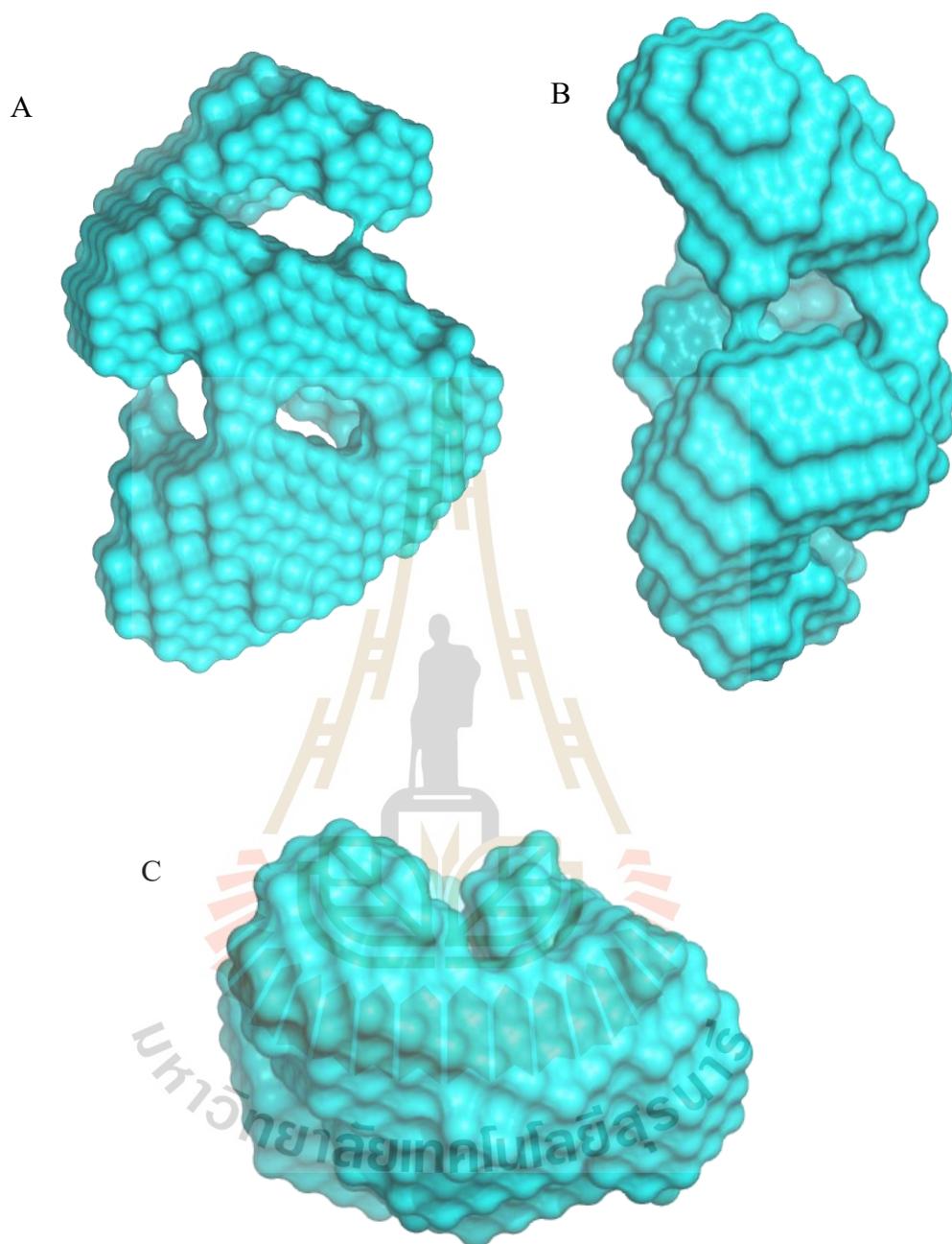


Figure 4.23 *Ab initio* model of Peak II. A) *Ab initio* model from 10 mg/mL of Peak II by Dammif with P1 symmetry (10 models). B) The modeling is rotated 90° around the y-axis compared to A and C) rotated 90° around the x-axis compared to A.

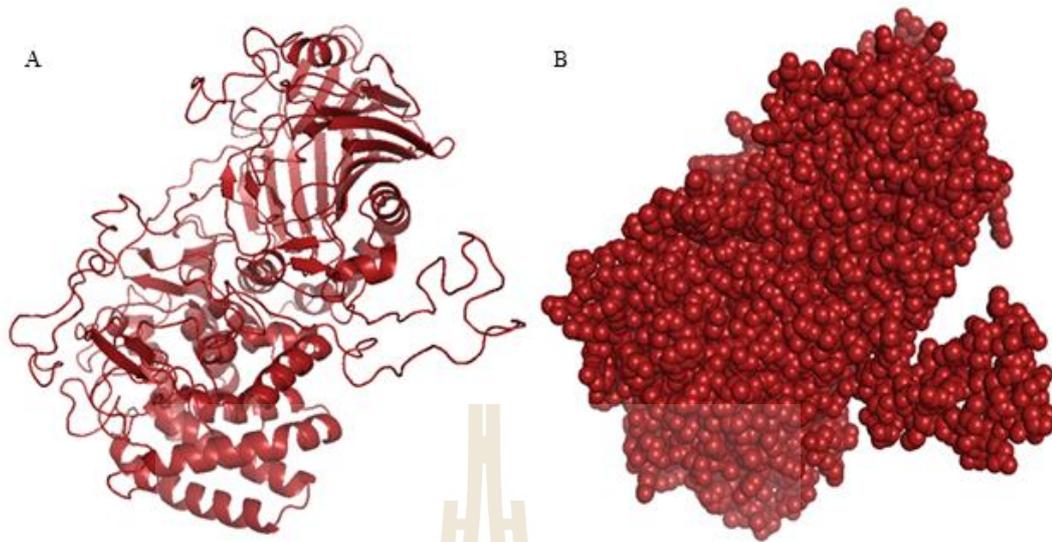


Figure 4.24 Predicted homology protein model of At4GH116 build by Swiss-model based on *TxGH116*. A) Model represented in cartoon B) model represented in spheres.

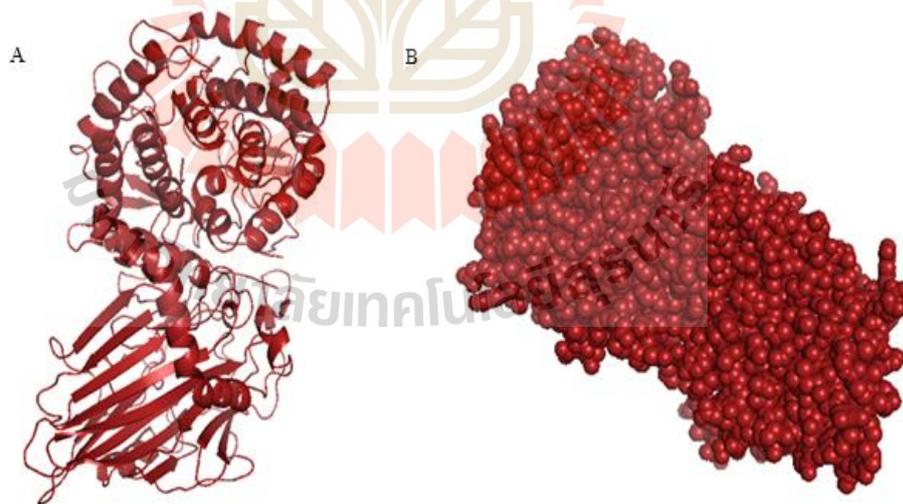


Figure 4.25 Predicted homology protein model of At4GH116 Δ 446-492 (with the flexible loop deleted) build by Swiss-model, based on *TxGH116*. A) Model represented in cartoon B) model represented in spheres.

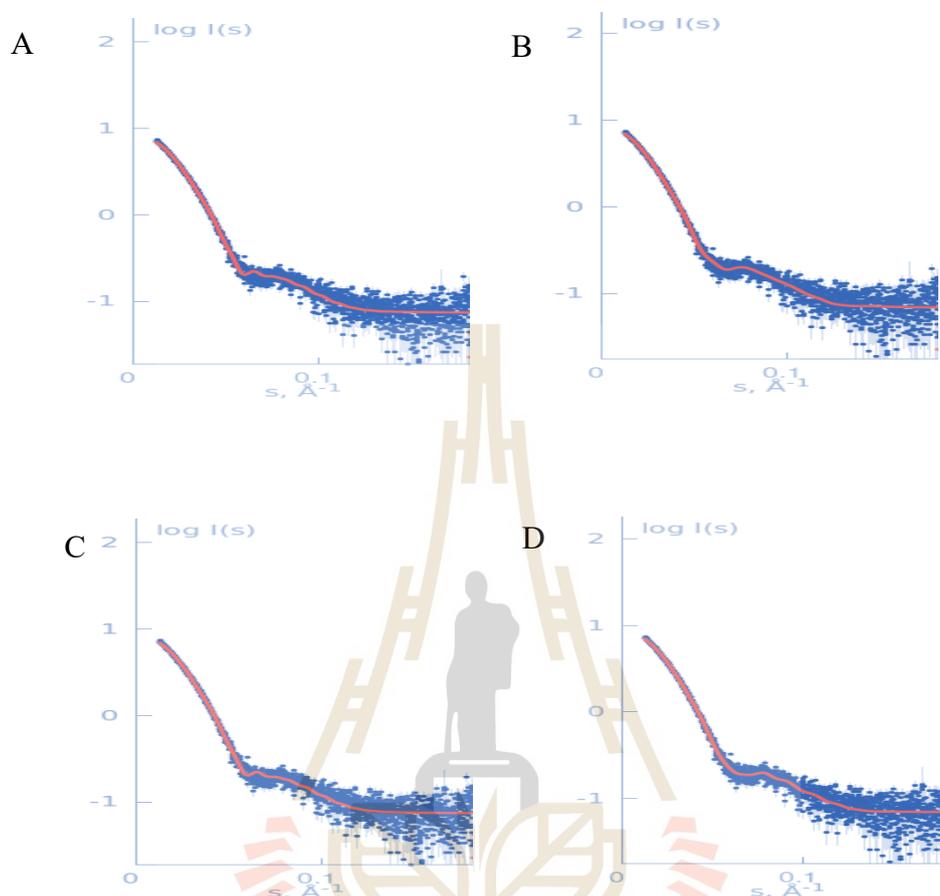


Figure 4.26 SAXS data of 5 mg/ml of At4GH116 (blue) overlaid with the stimulated scattering oligomer profile (red) calculated by SASREF from the At4GH116 homology model based on the TxGH116 crystal structure. A) Trimer with At4GH116 B) Tetramer with At4GH116 C) Trimer with At4GH116 Δ 446-492 D) Tetramer with At4GH116 Δ 446-492

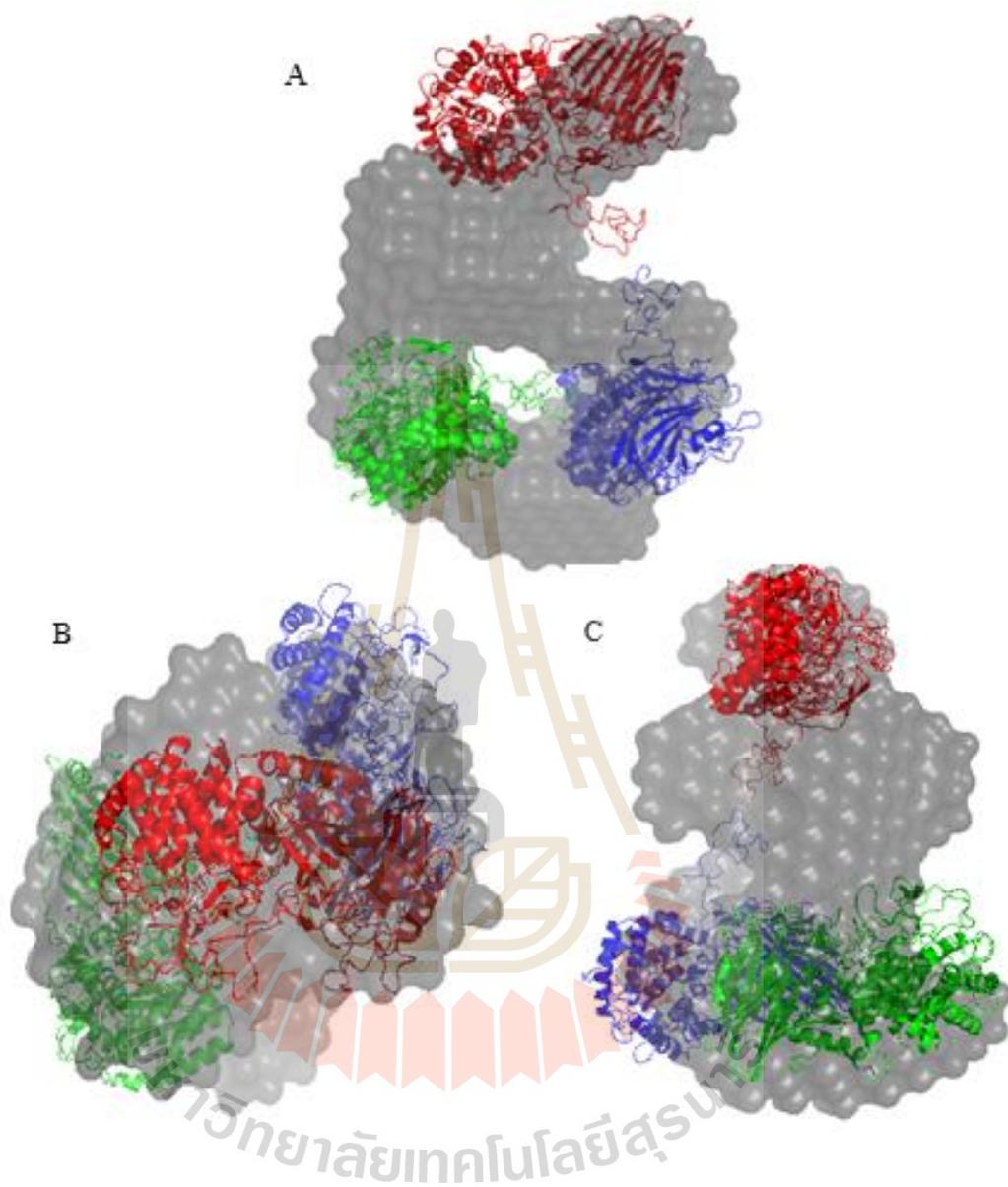


Figure 4.27 The three monomers of the homology model superimposed onto the SAXS-derived *ab initio* envelope by SASREF to generate a putative trimer. Envelopes are shown in a gray 20% transparent surface representation. Models are represented in cartoon and individual monomers are represented with different colors: monomer A - red, monomer B – green, and monomer C – blue. The model in B is rotated 90° around the x-axis relative to A. The model in C is rotated 90° around the y-axis compared to A.

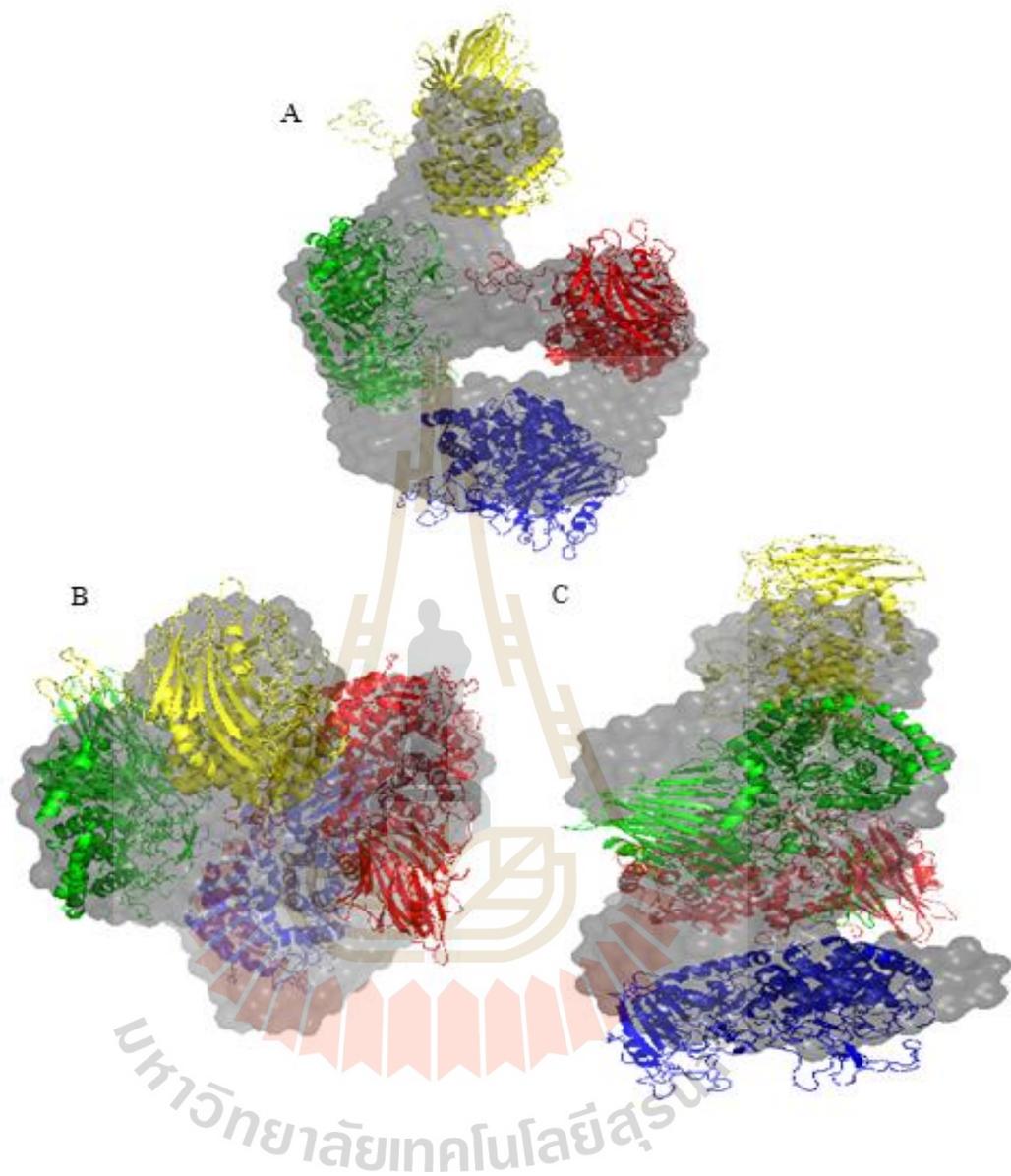


Figure 4.28 Four copies of the homology model were superimposed onto the SAXS-derived *ab initio* envelope by SASREF to generate a putative tetramer. Envelopes are shown in a gray 20% transparent surface representation. Models are represented in cartoon and individual monomers are represented with different colors: monomer A - red, monomer B - green, monomer C - blue and monomer D - yellow. The model in B rotated 90° around the x-axis with respect to A. The model in C rotated 90° around the y – axis with respect to that in A.

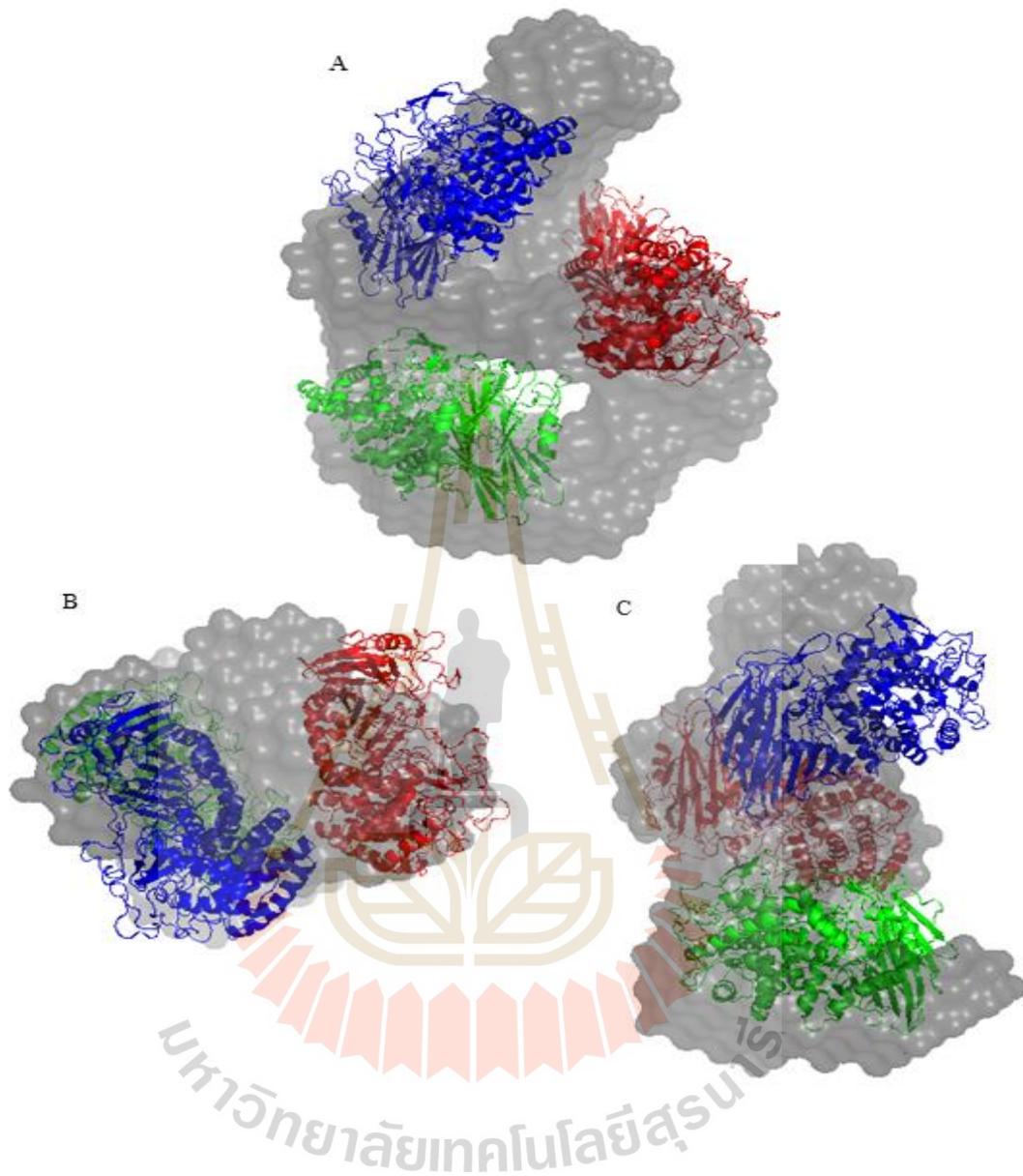


Figure 4.29 The three monomers of the homology model of At4GH116 Δ 446-492 superimposed onto the SAXS-derived *ab initio* envelope using SASREF to generate a putative trimer. Envelopes are shown in a gray 20% transparent surface representation. Models are represented in cartoon and individual monomers are represented with different colors: monomer A - red, monomer B - green, and monomer C - blue. The model in B is rotated 90° around the x-axis relative to A. That model in C is rotated 90° around the Y- axis compared to A.

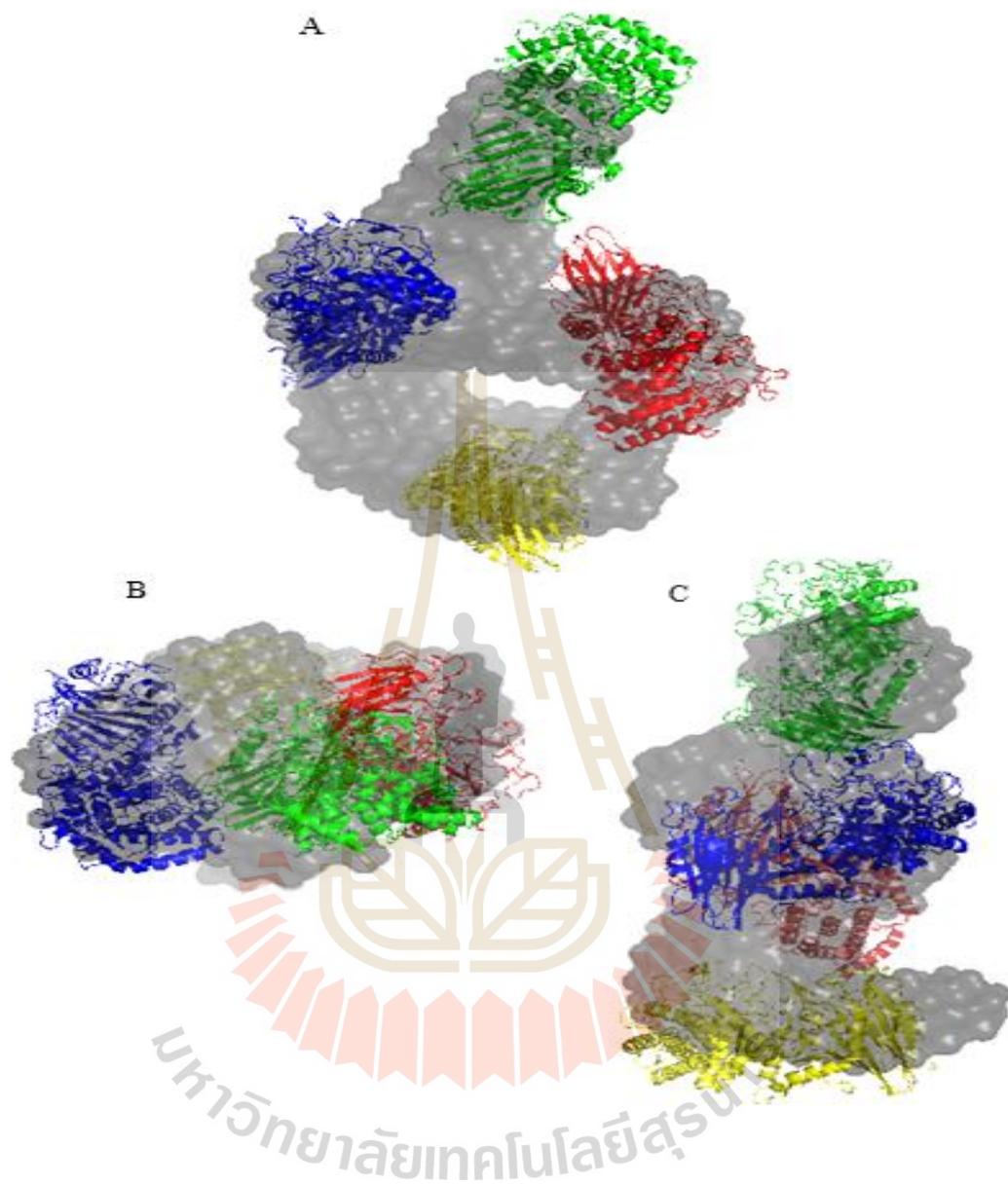


Figure 4.30 Four copies of the homology model of At4GH116 Δ 446-492 were superimposed onto the SAXS-derived *ab initio* envelope by SASREF to generate a putative tetramer. Envelopes are shown in a gray 20% transparent surface representation. Models are represented in cartoon and individual monomers are represented with different colors monomer A - red, monomer B - green, monomer C - blue, and monomer D - yellow. The model in B is rotated 90° around the x-axis with respect to A. the model in C is rotated 90° around the y- axis with respect to that in A.

Finally, *ab initio* shape predicted using DAMMIF modeling revealed the molecular envelope structure of At4GH116 in solution which exhibits a triangular shape with a middle hole (Figure 4.23). The stimulated scattering oligomer profile calculated by SASREF was overlaid well with SAXS data of 5 mg/ml of At4GH116 (Figure 4.26). The rigid-body fit of the homology models of the At4GH116 (Figure 4.24) and At4GH116 Δ 446-492 (Figure 4.25) was superimposed into the *ab initio* SAXS molecular envelope (Figure 4.27 to Figure 4.34). Superimposition of rigid-body fit generated with trimer/tetramer of At4GH116 Δ 446-492 exhibited a best fitting (Figure 4.31).

Each of the monomers appeared to touch at least one other monomer, but the distance between the individual monomers in the superimposition of rigid-body fit generated with tetramer of At4GH116 Δ 446-492 is in the range of 7.2 to 8.5 Å. In the other superimposition of rigid-body fit generated with trimer/tetramer of At4GH116 and trimer of At4GH116 Δ 446-492 exhibited more distance between the individual monomers (Data not shown).

4.7.2 Crystallography

To try to determine the 3-dimensional structure of At4GH116, the protein was crystallized for X-ray diffraction. Upon screening, various sizes of hexagonal shaped crystal clusters appeared after approximately 20 days (Figure 4.31) under conditions, as summarized in Table 5.1.

4.7.2.1 Initial screening by the sitting drop vapor diffusion method

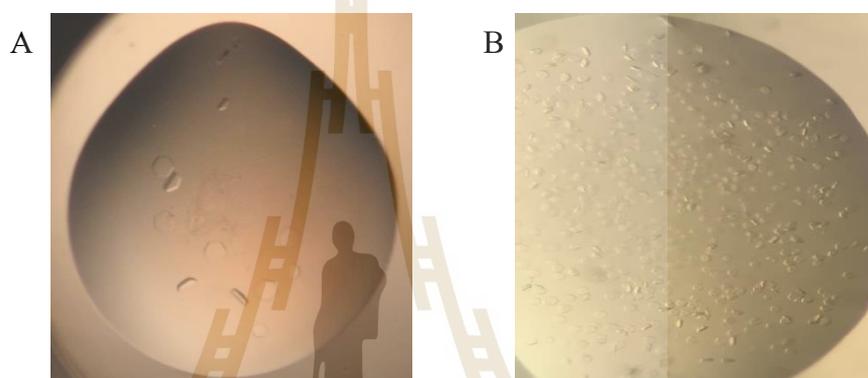


Figure 4.31 At4GH116 crystals obtained from vapor diffusion sitting drop method screening at 15°C. (A) 0.3 M magnesium formate, 0.1 M Bis-Tris, pH 5.5. (B) 15% - 2-propanol, 100 mM HEPES, pH 7.5, 200 mM magnesium chloride.

Table 4.6 Crystallization conditions of obtained crystals of At4GH116.

Screening Kit	Precipitant	Crystal morphology
JBScreen Classic HTS II	15% 2 propanol, 100 mM HEPS; pH7.5, 0.2 M Magnesium chloride	Single, hexagonal disks
JCSG-plus™ HT-96/FX-96	0.3 M Magnesium formate, 0.1 M BIS-Tris pH 5.5	Single

4.7.2.2 Optimization by hanging drop vapor diffusion

The crystallization conditions shown in Table 4.6 were chosen for further optimization with both the S200 SEC peaks I and II of At4GH116. Optimization conditions were drafted as mentioned on Tables 3.7-3.11. Mostly peak II protein gave a better crystals in the conditions mentioned in the above tables. Well-formed, single hexagonal shaped crystals were obtained earlier on 0.2 M to 0.4 M magnesium formate, 0.1 M Bis-Tris pH 5 to 6 (Table 3.7). The presences of 25% PEG 3000 to 6000 along with 0.15 M to 0.4 M magnesium formate, 0.1 M Bis-Tris, pH 5 to 6.5, did not helped in the formation of crystals (Figure 4.32).

In contrast, sharp edged, elongated hexagonal shaped crystals were obtained on conditions 0.15 M to 0.3 M magnesium chloride and 10 to 20% 2-propanol in 0.1M HEPES, pH 7.5 with and without 0-15% glycerol as mentioned in the Tables 3.9-3.11. Fewer smooth surfaced crystals were formed with 0.15 M to 0.3 M magnesium chloride, 0.1 M HEPES pH 7.5, 16% propanol with 0-15% glycerol (Table 3.9). Nearly all the precipitant 0.24M magnesium chloride, 10-20% 2-propanol and 0-15% glycerol in 0.1M HEPES, pH 7.5 (Table 3.10) were formed smooth surfaced crystals conditions shown in Figure 4.33. The quality and quantity of crystals were lower in the crystals obtained in the conditions mentioned in the Table 3.11.

At4GH116 crystals obtained in the above mentioned conditions were soaked in cryosolution alone or one that contained either 10 mM of DNPG2F, glucose or IFG. Native and crystals with inhibitors flash vitrified in liquid nitrogen and used for x-ray diffraction at the National Synchrotron Radiation Research Center, (NSRRC) Taiwan. Crystals formed in 0.2 M to 0.4 M magnesium formate with 0.1 M Bis-Tris, pH 5 to 6 diffracted X-rays to between 7 to 17 Å. Crystals of At4GH116 from peak I

and II were formed in 0.15 to 0.3 M MgCl_2 , 10%-17.5% propanol, 0.1M HEPES buffer pH 7.5, with and without 0-15% glycerol at 15°C diffracted poorly up to 9 Å.

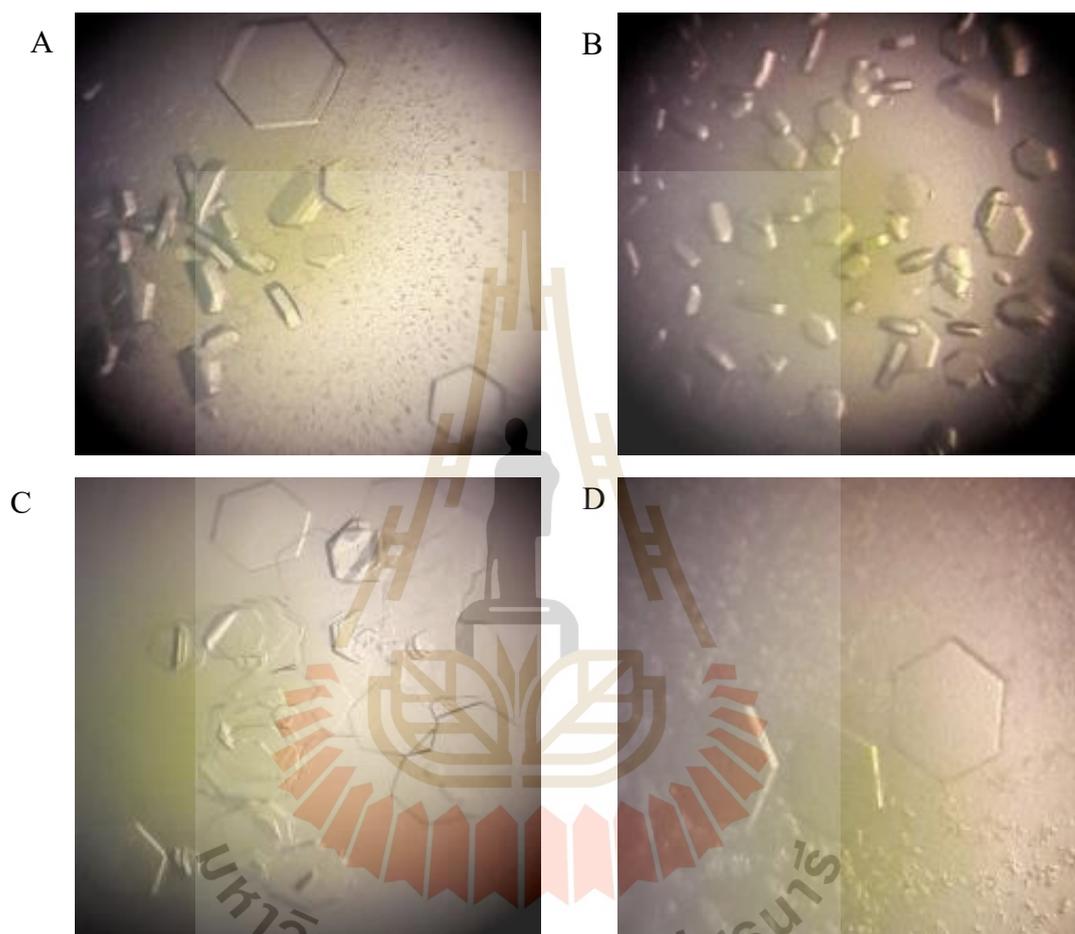


Figure 4.32 At4GH116 crystals obtained from vapor diffusion hanging drop method screening at 15°C. A) 0.15 M magnesium formate, 0.1 M Bis-Tris, pH 5. B) 0.15 M magnesium formate, 0.1 M Bis-Tris, pH 5.5. C) 0.2 M magnesium formate, 0.1 M Bis-Tris, pH 5.5. D) 0.2 M magnesium formate, 0.1 M Bis-Tris, pH 5.

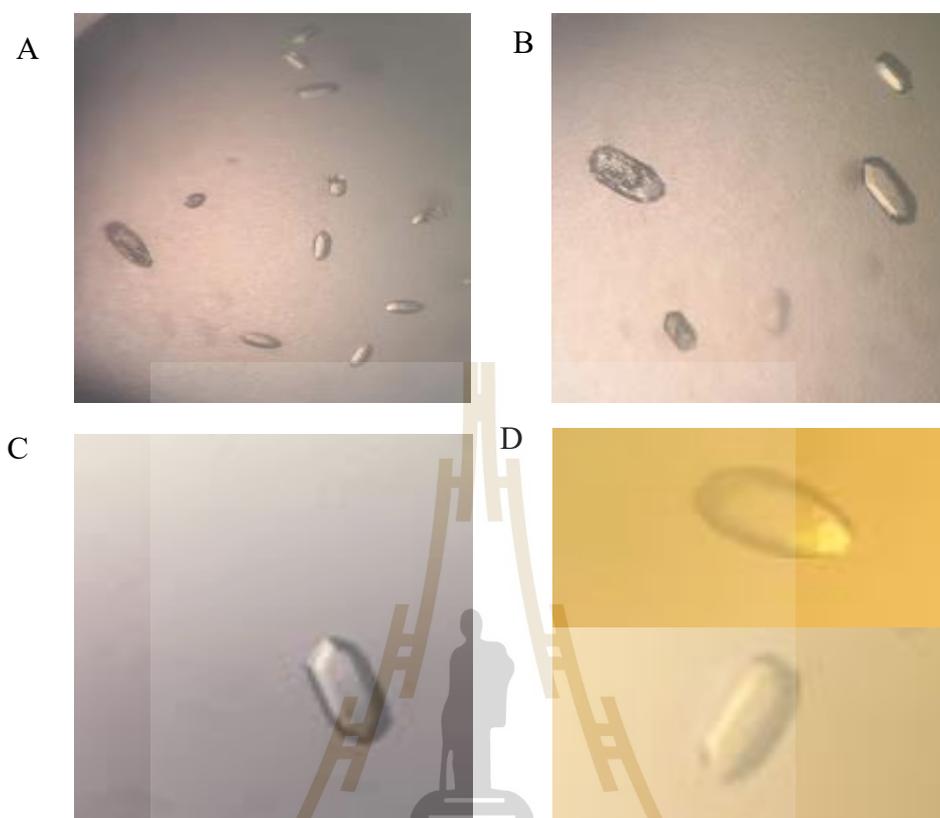


Figure 4.33 At4GH116 crystals obtained from vapor diffusion hanging drop method screening at 15°C. A) 0.18 M magnesium chloride, 0.1 M HEPES, pH 7.5, 10% 2-propanol, B) 0.15 M magnesium chloride, 0.1 M HEPES, pH 7.5, 16% 2-propanol, C) 0.24 M magnesium chloride, 0.1 M HEPES, pH 7.5, 16% 2-propanol, 10% glycerol, D) 0.18 M magnesium chloride, 0.1 M HEPES, pH 7.5, 16% 2-propanol. in the presence and absence of 0-15% glycerol.

4.8 Screening of homozygous GH116 mutants in *Arabidopsis*

4.8.1 Screening of homozygous single GH116 mutants in *Arabidopsis*

In order to test the function of AtGH116 and analyze the effect of mutations in differences in the plant phenotype and metabolic profiles between wild type and mutants, seven T-DNA insertion mutant lines, two per each of At3GH116, At4GH116 and At5GH116 and one for At1GH116 were bred for homozygosity and identified by PCR genotyping. T-DNA insertions in both the alleles are designated as homozygous single mutant lines (*Atgh116/Atgh116*). The plants from those seeds were designated as F₀ generation homozygous single mutant lines of *At1gh116*, *At3gh116*, *At4gh116* and *At5gh116* (Figure 4.34-4.37). T-DNA insertions in either one of the alleles are designated as heterozygous single mutant lines (*AtGH116/Atgh116*).

Among the plants genotyped for Salk_143139 knock-out line in Figure 4.34 plant one is considered as a homozygous single mutant line, because of the presence of PCR product observed only in the TP lane and plant four is considered as heterozygous mutant line. All the plants genotyped for the knock-out line Salk_117180 of At3GH116 in Figure 4.35 A were considered as homozygous single mutant lines. In the case of the other mutant lines of At3GH116, Salk_039506 only one homozygous mutant line shown in Figure 4.35 B was observed by genotyping, plant two, and plant one and three are heterozygous mutants. The screening of At4GH116 knocks out lines of Salk_099838 and Salk_010803 in Figure 4.36 exhibited two homozygous mutant lines for plant 2 and 4 and plant 3 and 4, respectively. In the case of the knock out lines of Salk_057433 shown in Figure 4.37 only plant one is a homozygous single mutant and remaining plants are considered as

homozygous wild type. In the case of Salk_039511 mutant lines, an equal number of homozygous single mutant lines and heterozygous mutant lines were observed. Plant one and three in Figure 4.37 B are homozygous single mutant lines and two and four are heterozygous mutant lines.

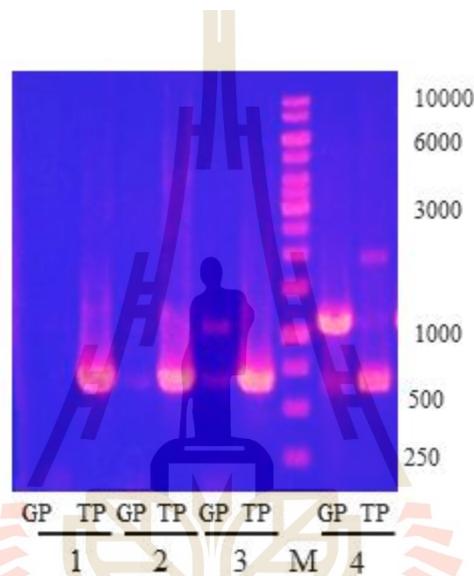


Figure 4.34 Genotyping PCR of the T-DNA insertion knock-out line Salk_143139 of At1GH116. GP-amplification with gene-specific primers; TP- amplification with T-DNA specific primers, the product of which indicated an insertions mutations, Lane M, Thermo Scientific Gene ruler 1KB DNA ladder and the numbers are those designated to individual plants. With T-DNA insert DNA, the band is expected at around 600 bp in the TP lane; without insert it is expected at around 1200 bp in the GP lane.

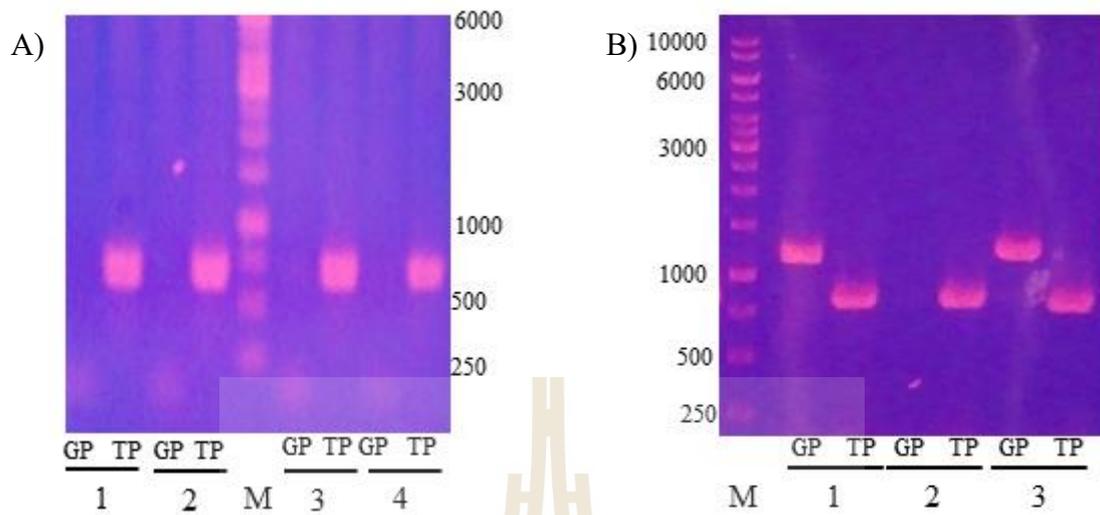


Figure 4.35 Genotyping PCR of the T-DNA insertion in knock-out lines Salk_117180 (A), and Salk_039506 (B) of At3GH116. GP-amplification with gene-specific primers; TP-amplification with T-DNA specific primers and gene right primer. Lane M, Thermo Scientific Gene ruler 1kb DNA ladder and the numbers are those designated to individual plants. With T-DNA inserts DNA band is expected at around 600 bp without insert, it is expected at around 1200 bp in the GP lane.

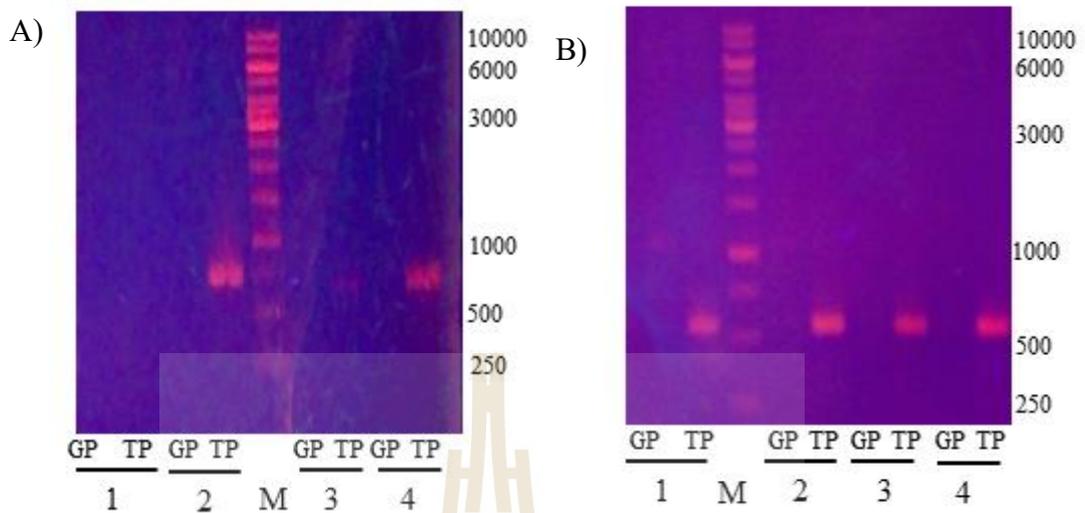


Figure 4.36 Genotyping PCR of the T-DNA insertions in knock-out lines Salk_099838 (A) and Salk_010803 (B) of At4GH116. GP-amplification with gene-specific primers; TP- amplification with T-DNA specific primer and one gene-specific primer. Lane M, Thermo Scientific Gene ruler 1kb DNA ladder the numbers are those designated to individual plants. With T-DNA insert DNA band is expected at around 600 bp in the TP lane, while without insert, the band is expected at around 1200 bp in the GP lane.

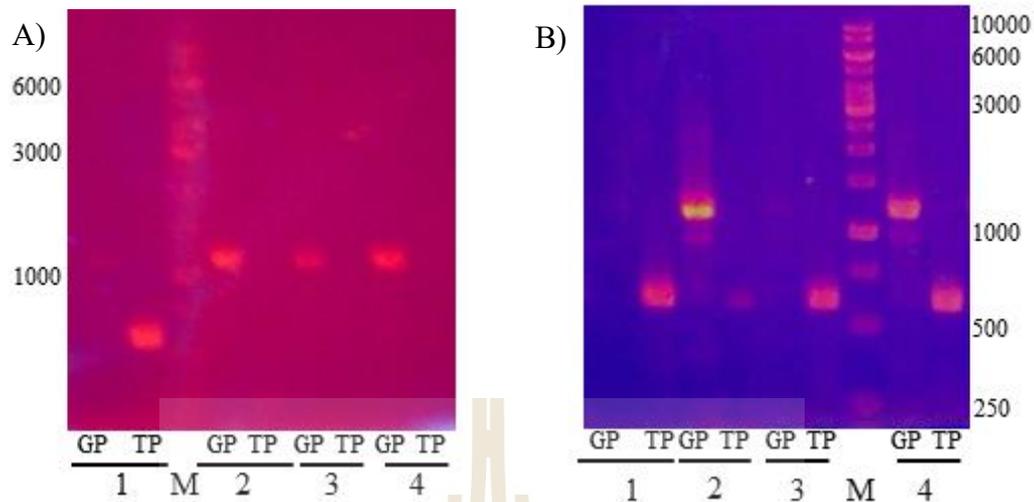


Figure 4.37 Genotyping PCR of the T-DNA insertions in knock-out line of Salk_057433 A), and Salk_039511 (B) of At5GH116. GP-amplification with gene-specific primers; TP-amplification with the T-DNA specific primer and one gene-specific primer. Lane M, Thermo Scientific Gene ruler 1kb DNA ladder the numbers are those designated to individual plants. With T-DNA insert, the DNA band is expected at around 600 bp in the TP lane, while without insert it is expected at around 1200 kb in the GP lane.

None of the single mutants showed an obvious phenotype compared with WT under normal growth conditions (Data not shown). These findings suggest that in the absence of any of the individual genes among the four, the functions of the particular gene might be replaced by one or more of the other three genes. These results are similar to findings of Dai et al., 2020 reported the same phenomena in mutants of AtGCD3 (At4GH116). It is also similar to T-DNA insertion mutant lines of β -glucosidase homologs, from other GH families like BG2, BGLU45 and BGLU46

belonging to GH family 1 that exhibited no obvious phenotypes (Xu et al., 2012; Chapelle et al., 2012).

4.9 Generation of homozygous double GH116 mutants in *Arabidopsis*

4.9.1 Generation of heterozygous double GH116 mutants in *Arabidopsis*

In order to analyze the phenotype and metabolic profile changes in the double homozygous knockout lines, single homozygous mutant lines (F_0) were cross-pollinated and progeny were designated as F_1 generation. In F_1 generation plants, either one of the alleles of both the parental chromosome possesses T-DNA insertions and these were considered as double heterozygous mutant lines (*Atgh116/AtGH116XAtgh116/AtGH116*). The double heterozygous mutant lines were obtained for *At1gh116XAt5gh116* (Figure 4.38) and *At1gh116XAt4gh116* (Figure 4.39). Seeds of those mutant lines were allowed to self-pollinate.

Screening of the F_1 generation by genotyping plant number three exhibited a PCR product in GP and TP lanes of the both the *1g* and *5g* chromosomes, as well as plants number 2 and 5 for the *1g* and *4g* were considered as double heterozygous mutants.

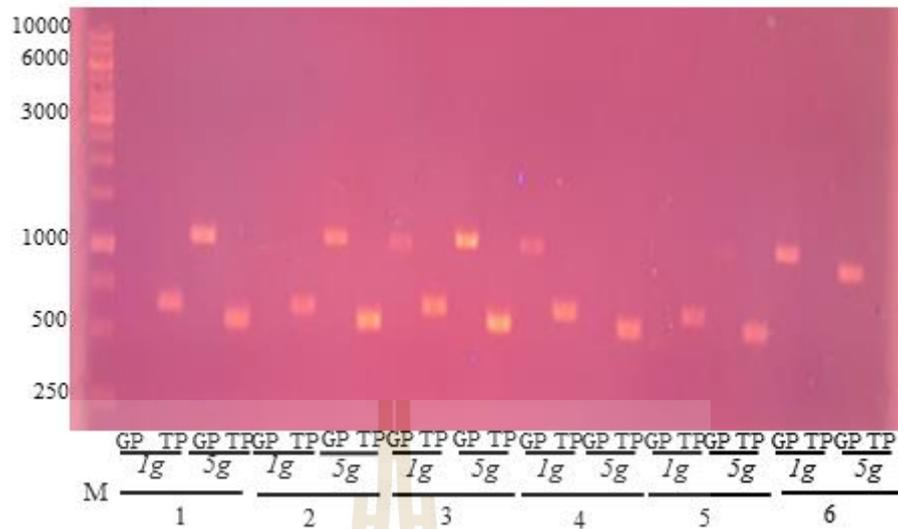


Figure 4.38 Genotyping PCR of the T-DNA insertion At1GH116 knockout lines Salk_143139 and At5GH116 knockout lines Salk_057433 double mutant F1 generation. *1g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_143139 and *5g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_057433. With T-DNA insert the DNA band was expected around 600 bp in TP lanes without insert around 1200 bp in GP lanes.

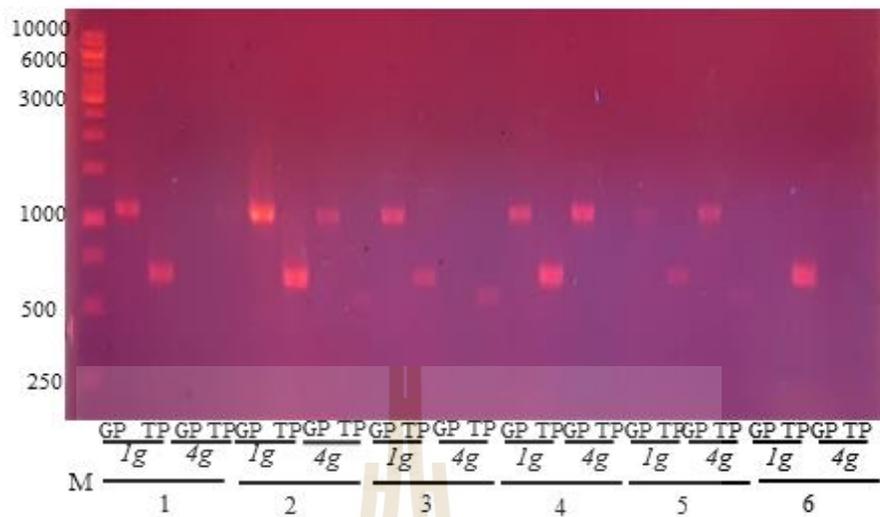
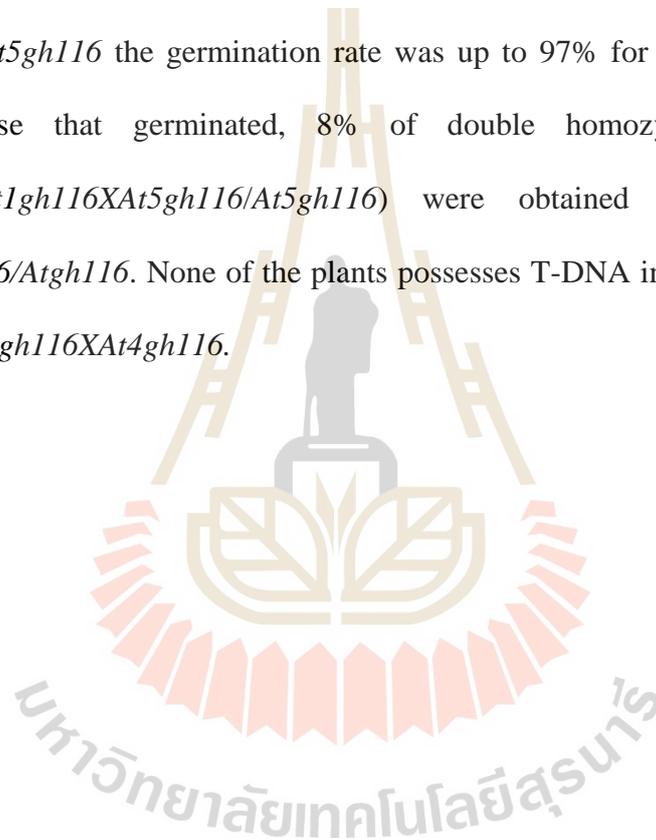


Figure 4.39 Genotyping PCR of the T-DNA insertion At1GH116 knockout lines Salk_143139 and At4GH116 knockout line Salk_010803 double mutant F1 generation. *1g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_143139 and *4g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_010803. With T-DNA insert, the DNA band is expected at around 600 bp in the TP lane, while without insert, it is expected around 1200 bp in the GP lane.

4.9.2 Generation of homozygous double GH116 mutant lines in *Arabidopsis*

The genotype of F₂ generation, progeny of F₁ heterozygous mutant lines are summarized in the Table 5.2. Based on the PCR products amplified with gene specific (GP) and T-DNA specific primer (TP) respective to At1G and At5G revealed that plant numbers 13, 28 and 34 were considered as double homozygous mutants. Only 65% of the seeds germinated and were genotyped for *At1gh116XAt4gh116*. But for *At1gh116XAt5gh116* the germination rate was up to 97% for *At1gh116XAt4gh116*. Among those that germinated, 8% of double homozygous mutant lines (*At1gh116/At1gh116XAt5gh116/At5gh116*) were obtained and designated as $\Delta 1\Delta 5Atgh116/Atgh116$. None of the plants possesses T-DNA insertion in all the four alleles of *At1gh116XAt4gh116*.



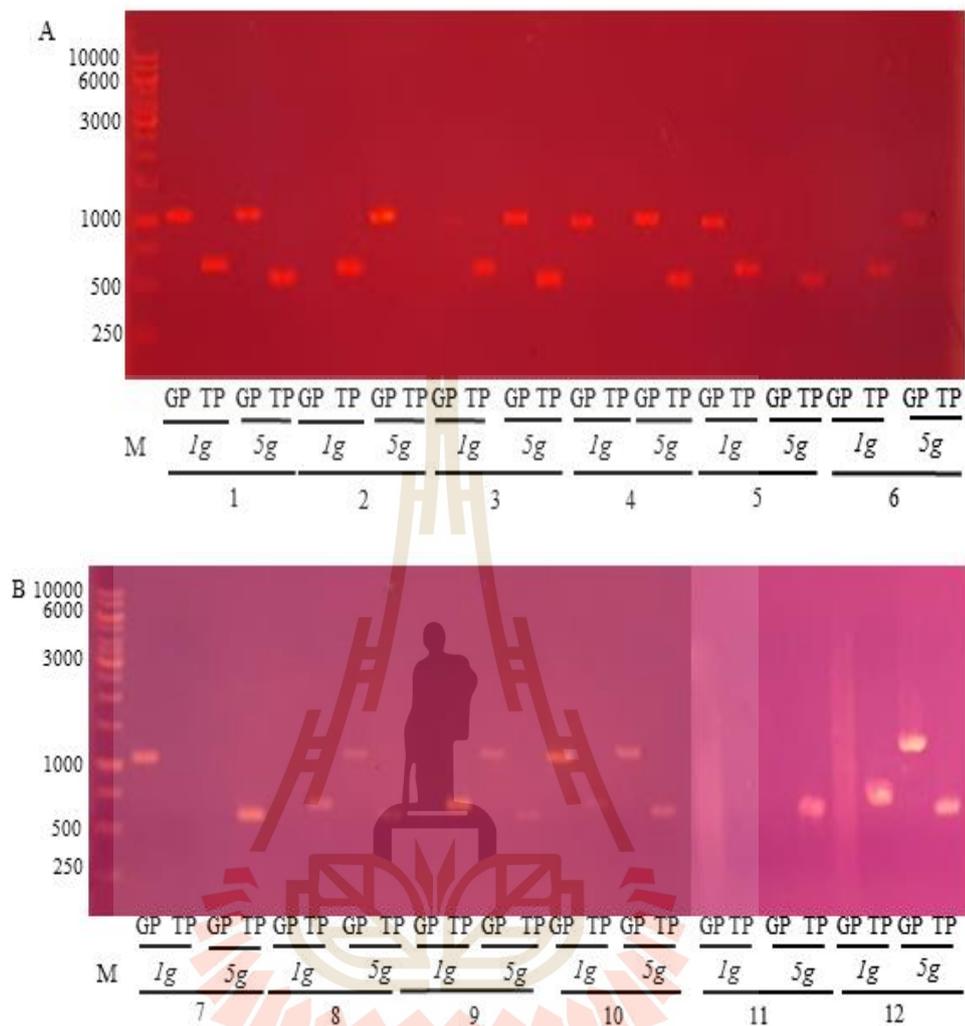


Figure 4.40 Genotyping PCR of the T-DNA insertion *At1gh116* knockout lines Salk_143139 and *At5gh116* knockout lines Salk_057433 double mutant F2 generation. *1g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_143139 and *5g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_057433. With T-DNA insert, the DNA bands were expected at around 600 bp, while without insert they were expected at around 1200 bp.

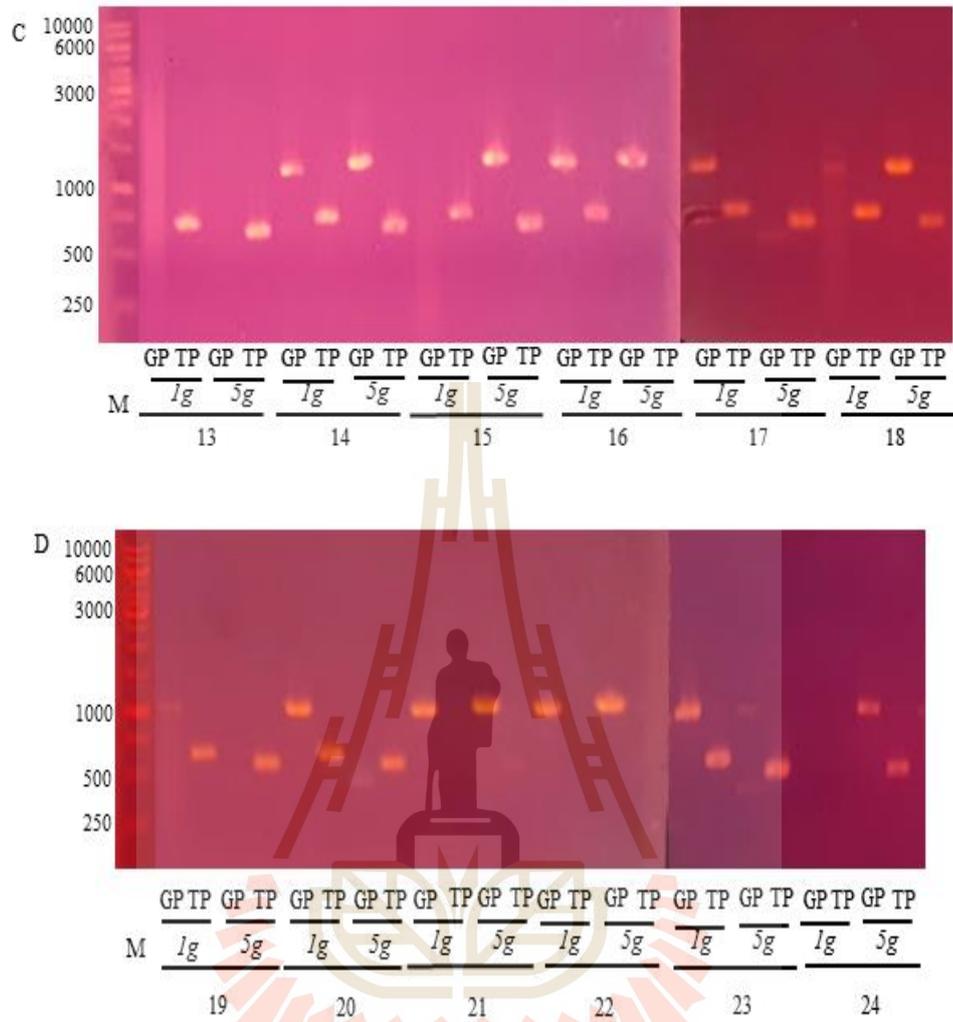


Figure 4.41 Genotyping PCR of the T-DNA insertion *At1gh116* knockout lines Salk_143139 and *At5gh116* knockout lines Salk_057433 double mutant F2 generation. *1g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_143139 and *5g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_057433. With T-DNA, the insert DNA band is expected at around 600 bp without insert, it is expected at around 1200 bp.

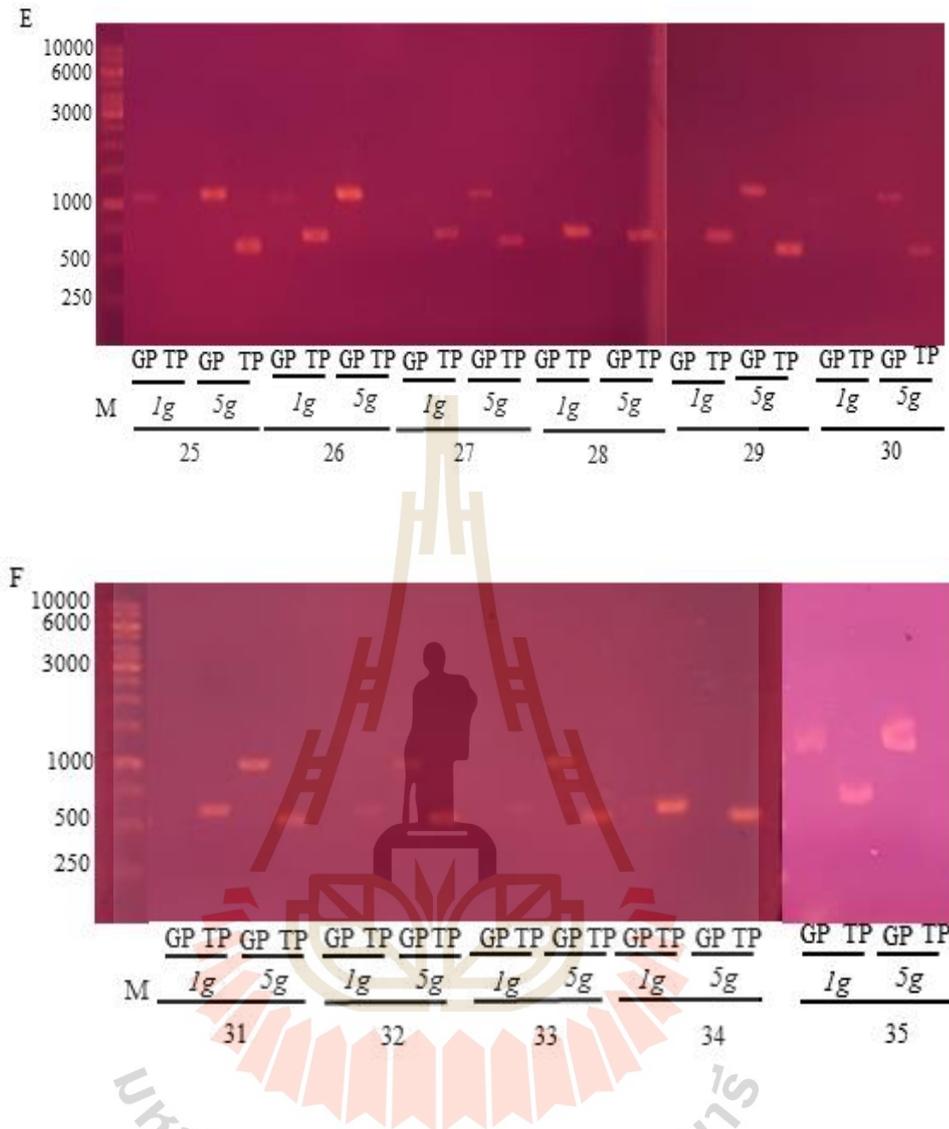


Figure 4.42 Genotyping PCR of the T-DNA insertion *At1gh116* knockout lines Salk_143139 and *At5gh116* knockout lines Salk_057433 double mutant F2 generation. *1g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_143139 and *5g* represents amplification with gene-specific primers (GP) and T-DNA specific primers (TP) respective to Salk_057433. With T-DNA, the insert DNA band is expected at around 600 bp; while without insert, it is expected at around 1200 bp.

Table 4.7 Quantitative analysis of genotype of F₁ progeny. G and g represents dominate and recessive alleles present in the chromosomes of *Arabidopsis* respective to the GH116 genes. Genotypes were grouped based on the Mendel's law of inheritance (9:3:3:1).

Genotype of progeny of F ₁ lines	No of plants obtained per genotype	
	<i>At1gh116</i> X <i>At4gh116</i>	<i>At1gh116</i> X <i>At5gh116</i> .
GGGG		
gGGG	7	15
gGgG		
Gggg	9	6
gGgg		
ggGG		
ggGg	3	9
gggG		
gggg		3
Planted	32	36
Genotyped	21	35
Considered	19	33

4.10 Effect of light stress on the phenotype of wild type and homozygous single mutants of *Atgh116*.

The enrichments from half MS media and the benefit of light source were observed in the phenotype of all plants grown on half MS agar plate under long day (HL) conditions on the 24th day (Figure 4.43). On the 38th day, due to extended darkness, obvious dwarfism was observed in the plants grown in short day conditions. On the same day, all the plants grown on half MS agar plate under short day conditions (HS) for 24 days and then moved to grow on soil under short day (SS) conditions for another 14 days were exhibited enlarged rosette compared to the one grown on SS conditions since day 0 (Figure 4.44).

After 10 days of growing plants in soil under long day (SL) conditions, on day 49, slight purple color pigmentations on the leaves of mutant plants was initially noticed (data not shown). The dwarfism that occurred in the plants grown under short day conditions were also rescued by growing the plants under long day conditions, in less than 15 days. Extreme pigmentations were also noticed on the 65th day for all the mutant lines except Salk_010803, one of the single homozygous mutant lines of *At4gh116*, and wild type (Figure 4.45).

After 14 days of high light stress conditions, plants overexpressing UDP-glucosyl transferase from citrus (Cs-UGT78D3), a homolog of *Arabidopsis* flavonoid 3-O-glucosyltransferase in *A. thaliana* exhibited purple color pigmentation more than wild type. As an effect of all overexpressing lines showed high expression of many genes related to anthocyanin biosynthesis pathway such as Transparent testa 8 and 9 (TT8, TT9), BANYULS (BAN), Dihydroflavonol 4-reductase (DFR). These results showed the potential function step of Cs-UGT78D3 in anthocyanins biosynthesis

pathway (Rao et al., 2019). In T-DNA insertional inactivation lines (*bglu15-1* and *bglu15-2*) of BGLU15, a *Arabidopsis* β -glucosidase, apparent purple coloration appeared on the leaf surface of plants sampled at the end of the nitrogen deficiency and low temperature (NDLT) stress period (Roepke et al., 2017).



Figure 4.43 Comparison of phenotypes of wild type *A. thaliana*, with the other mutants of *Atgh116* grown under long daylight conditions for 24 days. A) Col-0-VC B) Col-0-JKC, C) Salk_143139, D) Salk_117180, E) Salk_010803, F) Salk_099838 G) Salk_057433, H) Salk_039511. Plants initially grown on half MS are below the red lines and plants grown on soil are above the red line.

These findings suggest that, as an effect of higher accumulation of flavonoid glucosides under the stress conditions, the plants turned to purple. This suggests that either our *AtGH116* mutant plants are more stressed than wildtype or they are less able to turnover the flavonoids, including anthocyanins that build up under the stress of increased light.



Figure 4.44 Comparison of phenotypes of wild type *A. thaliana*, with the other mutants of *Atgh116* grown under short day light conditions for 38 days. A) Col-0-VC, B) Col-0-JKC, C) Salk_143139, D) Salk_117180, E) Salk_010803, F) Salk_099838, G) Salk_057433, H) Salk_039511. Plants initially grown on half MS are below the red line and plants grown on soil are above the red line.

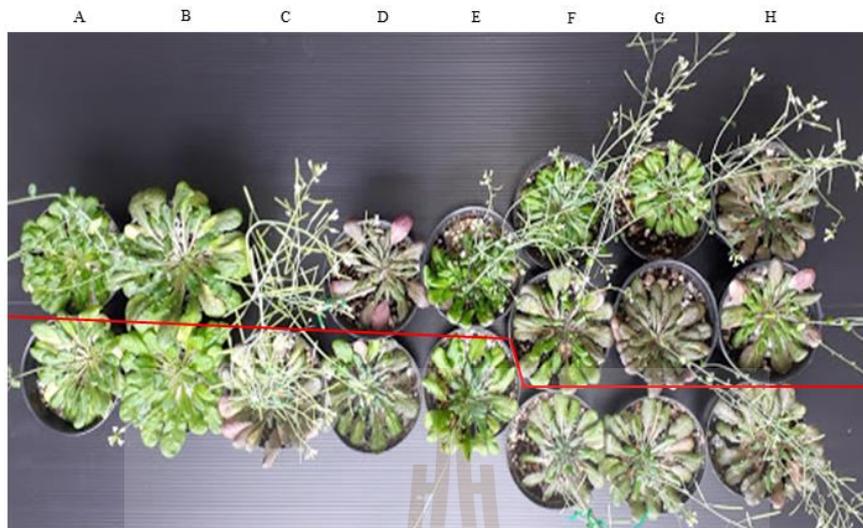


Figure 4.45 Effect of light duration on phenotypes of wild type *A.thaliana*, and other mutants of *Atgh116* on the 65th day.) Col-0-VC, B) Col-0-JKC, C) Salk_143139, D) Salk_117180, E) Salk_010803, F) Salk_099838, G) Salk_057433, H) Salk_039511. Plants grown on half MS are below the red lines and plants grown on soil were on above the red line.

CHAPTER V

CONCLUSION

In *Arabidopsis*, four GH116 genes, At1g33700, At3g24180, At4g10060 and At5g49900, are present and phylogenetic analysis indicated that these genes diverged from a single shared ancestor in plants. The plasmids containing the cDNA provided by the RIKEN Biological Resource Center, Japan, were used to amplify the AtGH116 CDS with the Gibson assembly sequence to attain a PCR products with sizes around 3 kb. The purified plasmid sequences established that the AtGH116's were cloned into the pET30 expression vector. Only At4GH116 was produced as active protein, while the other proteins could not be expressed in either *E. coli* strain BL21(DE3) or Rosetta-gami(DE3) as an active β -glucosidase.

At4GH116 could be expressed in *E. coli* BL2(DE3) by induction with IPTG at 20°C for 16 hours, so it was purified and characterized. The target protein was present in the soluble cell lysate, as well as in the insoluble pellet. The At4GH116 protein was purified from cell lysate by IMAC and gel filtration size exclusion chromatography over an S200 column. At4GH116 enzyme purified on S200 gel filtration column chromatography was separated based on its native size, dependent on its oligomerization, into 3 peaks.

Among *pNP* α - and β -glycosides, At4GH116 hydrolyzed *pNP*- β -D-glucopyranoside (*pNPGlc*) with greatest specificity and *pNP*- β -D-galactopyranoside with the highest k_{cat} . Among available natural substrates, At4GH116 could hydrolyze

the flavonoid glucosides, apigenin-7-O- β -glucoside, kaempferol-7-O- β -glucoside, luteolin-7-O- β -glucoside, quercetin-7-O- β -glucoside and kaempferol-3-O- β -glucoside and other phenolic glucosides, such as helicin, arbutin, indoxy β -D-glucoside, diazidin, and glycitin.

The kinetic parameters for hydrolysis of flavonoid glucosides substrates by At4GH116 indicated high catalytic efficiency towards apigenin-7-O- β -glucoside ($k_{cat}/K_m = 1133 \text{ mM}^{-1}\text{s}^{-1}$) quercetin-7-O- β -glucoside ($k_{cat}/K_m = 735 \text{ mM}^{-1}\text{s}^{-1}$), kaempferol-7-O- β -glucoside ($k_{cat}/K_m = 220 \text{ mM}^{-1}\text{s}^{-1}$) and luteolin-7-O- β -glucoside ($k_{cat}/K_m = 106 \text{ mM}^{-1}\text{s}^{-1}$) and a little activity towards kaempferol-3-O- β -glucoside. This study supports, that At4GH116 could also hydrolyze flavonoid glucosides along with sphingolipid glucosides, since it exhibits low K_m ($1.52 \pm 0.1 \mu\text{M}$) and k_{cat} ($(8.0 \pm 0.9) \times 10^{-3}$) values with C6-NBD GlcCer as a substrate, thereby giving this GlcCer a specificity constant (k_{cat}/K_m) about 200-fold lower than apigenin-7-O- β -glucoside.

A SAXS experiment with 5 and 10 mg/ml of peak II At4GH116 from S200 gel filtration column chromatography, revealed that peak II At4GH116 is a globular shaped trimer or tetramer with a molecular weight of 328-422 and 255-339 from V_c and V_p respectively. The Guinier plot showed no pattern of aggregation with R_g of 71.26 Å and D_{max} 168 Å. In structural analysis of At4GH116, hexagonal protein crystals were obtained by vapor diffusion method but diffracted X-rays very poorly.

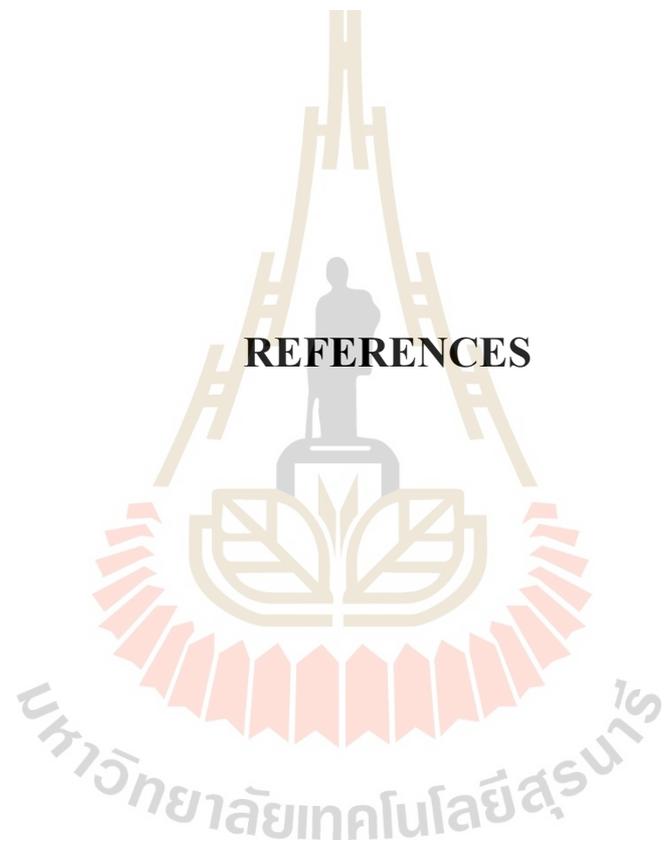
In order to test the function of AtGH116 and analyze the effect of loss of function mutations on the plant phenotypes and metabolite profiles, homozygous single mutant lines (F_0) (*Atgh116/Atlgh116*) of all the four GH116: *Atlgh116*, *At3gh116*, *At4gh116* and *At5gh116* genes were identified by PCR genotyping with gene and T-DNA specific primers. None of the homozygous mutant lines exhibited obvious

changes in growth of the plants under normal growth conditions. But all the single homozygous mutant lines exhibited purple-colored leaves under the light stress of changing from short day to long day conditions.

Among, all the possible attempts made in cross pollination between the single homozygous T-DNA mutant lines, double homozygous mutant line of *At1gh116/At5gh116* was obtained (*At1gh116/At1gh116* X *At5gh116/At5gh116*) and designated as $\Delta 1\Delta 5Atgh116/Atgh116$.

These findings suggest that At4GH116 has better catalytic activity towards flavonoid glucosides. Pigmentation on the *Atgh116* double knock out lines under the stress conditions further supports a possible role of *Atgh116* in flavonoid metabolism. Additional experiments on *in planta* comparative metabolite analysis between the wild type and *Atgh116* mutant lines under the normal and stress growth conditions could provide more evidence to understand the role of GH116 in *Arabidopsis thaliana* and plants in general.

REFERENCES



REFERENCES

- Alden, K.P., Cordelier, S.D., McDonald, K.L., Reape, T.J., Ng, C.K.Y., McCabe, P. F., and Leaver, C.J. (2011) Sphingolipid long chain base phosphates can regulate apoptotic-like programmed cell death in plants. **Biochem. Biophys. Res. Commun.** 410: 574-580.
- Ali, U., Li, H., Wang, X., and Guo, L. (2018) Emerging roles of sphingolipid signaling in plant response to biotic and abiotic stresses. **Mol. Plant.** 1(11): 328-1343.
- Ardèvol, A., and Rovira, C. (2015) Reaction mechanisms in carbohydrate-active enzymes: glycoside hydrolases and glycosyltransferases. Insights from *ab initio* quantum mechanics/molecular mechanics dynamic simulations. **J. Am. Chem. Soc.** 137(24): 7528-7547.
- Aureli, M., Bassi, R., Loberto, N., Regis, S., Prinetti, A., Chigorno, V., Johannes, M. A., Boot, R.G., Filocamo, M., and Sonnino, S. (2012) Cell surface associated glycohydrolases in normal and Gaucher disease fibroblasts. **J. Inherit. Metab. Dis.** 35: 1081-1091.
- Brady, R.O., Kanfer, J., and Shapiro, D. (1965) The metabolism of glucocerebrosides I purification and properties of a glucocerebroside-cleaving enzyme from spleen tissue. **J. Biol. Chem.** 240: 39-43.

- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active enzymes database (cazy): an expert resource for glycomics. **Nucleic Acids Res.** 37: 233-238.
- Chapman, J.M., and Muday G.K. (2021) Flavonols modulate lateral root emergence by scavenging reactive oxygen species in *Arabidopsis thaliana*. **J. Biol. Chem.** 296: 1-18
- Charoenwattanasatien, R., Pengthaisong, S., Breen, I., Mutoh, R., Sansenya, S., Hua, Y., Tankrathok, A., Wu, L., Songsiriritthigul, C., Tanaka, H., Williams, S.J., Davies, G.J., Kurisu, G., and Cairns, J.R.K. (2016) Bacterial β -glucosidase reveals the structural and functional basis of genetic defects in human glucocerebrosidase 2 (GBA2). **ACS Chem. Biol.** 11: 1891-1900.
- Chen, M., Han, G., Dietrich, R.C. Dunn, M.T., and Cahoon, B.E. (2006) The essential nature of sphingolipids in plants as revealed by the functional identification and characterization of the *Arabidopsis* LCB1 subunit of serine palmitoyltransferase. **Plant Cell.** 18: 3576-3593.
- Chen, M., Markham, E.J., Dietrich, R.C., Jaworski, G.J., and Cahoon B.E. (2008) Sphingolipid Long-Chain Base hydroxylation is important for growth and regulation of sphingolipid content and composition in *Arabidopsis*. **Plant Cell.** 20: 1862-1878.
- Chen, M., Markham, J.E., and Cahoon, E.B. (2012) Sphingolipid D8 unsaturation is important for glucosylceramide biosynthesis and low-temperature performance in *Arabidopsis*. **Plant J.** 69: 769-781.
- Chuankhayan, P., Rimlumduan, T., Tantanuch, W., Mothong, N., Kongsaree, P.T., Methenukul, P., Svasti, J., Jensen, O.N., and Cairns J.R.K. (2007) Functional

and structural differences between isoflavonoid β -glycosidases from *Dalbergia* *sp.* **Arch. Biochem. Biophys.** 468: 205-216.

Citterio, A., Arnoldi, A., Panzeri, E., D'Angelo, M.G., Filosto, M., Dilella, R., Arrigoni, F., Castelli, M., Maghini, C., Germiniasi, C., Menni, F., Martinuzzi, A., Bresolin, N. and Bassi, M.T. (2014) Mutations in CYP2U1, DDHD2 and GBA2 genes are rare causes of complicated forms of hereditary spastic paraparesis. **J. Neurol.** 26: 373-381.

Cobucci-Ponzano, B., Aurilia, V., Riccio, G., Henrissat, B., Coutinho, P.M., Strazzulli, A., Padula, A., Corsaro, M.M., Pieretti, G., Pocsfalvi, G., Fiume, I., Cannio, R., Rossi, M., and Moracci, M. (2010) A new archaeal β -glycosidase from *Sulfolobus solfataricus*: seeding a novel retaining β -glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA2. **J. Biol. Chem.** 285: 20691-20703.

Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. **J. Mol. Biol.** 328: 307-317.

Cowart, L.A., and Obeid, L.M. (2007) Yeast sphingolipids: Recent developments in understanding biosynthesis, regulation, and function. **Biochim. Biophys. Acta.** 1771: 421-431.

Dai, G. Y., Yin, J., Li, K.E., Chen, D.K., Liu, Z., Bi, F.C., Rong, C., and Yao, X.N. (2020) The *Arabidopsis* AtGCD3 protein is a glucosylceramidase that preferentially hydrolyzes long-acyl-chain glucosylceramides. **J. Biol. Chem.** 295(3): 717-728.

- Davies, G.J., Gloster, T.M., and Henrissat, B. (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. **Curr. Opin. Struct. Biol.** 15:637-645.
- Day, A.J., DuPont, M.S., Ridley, S., Rhodes, M., Rhodes, M.J., Morgan, M.R., and Williamson, G. (1998) Deglycosylation of flavonoid and isoflavonoid glycosides by human small intestine and liver beta-glucosidase activity. **FEBS Lett.** 436: 71-75.
- Dekker, N., Brouwer, T.V., Verhoek, M., Wennekes, T., Narayan, R.S., Speijer, D., Hollak, C.E., Overkleeft, H.S., Boot, R.G., and Aerts, J.M. (2011) The cytosolic β -glucosidase GBA3 does not influence type 1 Gaucher disease manifestation. **Blood. Cells. Mol. Dis.** 46: 19-26.
- Dickson, R.C., Nagiec, E.E., Wells, G.B., Nagiec, M.M., and Lester R.L. (1997) Synthesis of Mannose-(inositol-P)₂-ceramide, the major sphingolipid in *Saccharomyces cerevisiae*, requires the IPT1 (YDR072c) Gene. **J. Biol. Chem.** 272(47): 29620-29625.
- Dickson, R.C., and Lester, R.L. (2002) Sphingolipid functions in *Saccharomyces cerevisiae*. **Biochim. Biophys. Acta.** 1583: 13-25.
- Dietrich, C.R., Han, G., Chen, M.R., Berg, H., Dunn, T.M., and Cahoon, E.B. (2008) Loss-of-function mutations and inducible RNAi suppression of Arabidopsis LCB2 genes reveal the critical role of sphingolipids in gametophytic and sporophytic cell viability. **Plant. J.** 54: 284-298.
- Dutilleul, C., Chavarria, H., Rézé, N., Sotta, B., Baudouin, E., and Guillas, I. (2015) Evidence for ACD5 ceramide kinase activity involvement in Arabidopsis response to cold stress. **Plant Cell Environ.** 38(12): 2688-97.

- García, S.M., García, G.A., Solís, G.A., García, C.F., Santana, V.S., Markham, E. J., Rosas, L.M.G., Dietrich, R.C., Vega, R.M., Cahoon, B.E., and Ruiz, G.M. (2011) MPK6, sphinganine and the LCB2a gene from serine palmitoyltransferase are required in the signaling pathway that mediates cell death induced by long chain bases in *Arabidopsis*. **New Phytol.** 191: 943-957.
- Glenz, R., Schmalhaus, D., Krischke, M., Mueller, J.M., and Waller, F. (2019) Elevated levels of phosphorylated sphingobases do not antagonize sphingobase or Fumonisin B1-induced plant cell death. **Plant Cell Physiol.** 60(5): 1109-1119.
- Gonzalez-Carmona, M.A., Sandhoff, R., Tacke, F., Vogt, A., Weber, S., Canbay, A. E., Rogler, G., Sauerbruch, T., Lammert, F., and Yildiz, Y. (2012) Beta-glucosidase 2 knockout mice with increased glucosylceramide show impaired liver regeneration. **Liver Int.** 32: 1354-62.
- Grabowski, G.A., and Horowitz, M. (1997) Gaucher's disease: molecular, genetic and enzymological aspects. Baillieres. **Clin. Haematol.** 10(4): 635-56.
- Hannich, J.T., Umebayashi, K., and Riezman, H. (2011) Distribution and functions of sterols and sphingolipids. **Cold Spring Harb. Perspect. Biol.** 3(5): 1-14.
- Hart, G.W., Copeland, R.J. (2010) Glycomics hits the big time. **Cell.** 143(5): 672-676.
- Haugarvoll, K., Johansso, S., Rodriguez, C.E., Boman, H., Haukanes, B.I., and Bruland, O. (2017) GBA2 Mutations cause a Marinesco Sjogren-Like syndrome: genetic and biochemical studies. **PLoS ONE.** 12: 1-10.
- Hayashi, Y., Okino, N., Kakuta, Y., Shikanai, T., Tani, M., Narimatsu, H., and Ito, M. (2007) Klotho-related protein is a novel cytosolic neutral-glycosylceramidase. **J. Biol. Chem.** 282: 30889-30900.

- Herrmann, K.M. (1995) The Shikimate pathway: Early steps in the biosynthesis of aromatic compounds. **Plant Cell**. 7(7): 907-919.
- Herrmann, K.M., and Weaver, L.M. (1999) The shikimate pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50: 473-503.
- Hosur, V., Johnson, K.R., Burzenski, L.M, Stearns, T.M., Maser, R.S., and Shultz, L.D. (2014) Rhbdf2 mutations increase its protein stability and drive EGFR hyperactivation through enhanced secretion of amphiregulin. **Proc Natl Acad Sci USA**. 111: E2200-2209.
- Hruska, K.S., LaMarca, M.E., Scott, C.R. and Sidransky, E. (2008) Gaucher disease: Mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). **Hum. Mutat.** 29: 567-583.
- Hua, Y., Ekkhara, W., Sansenya, S., Srisomsap, C., Roytrakul, S., Saburi, W., Takeda, R., Matsuura, H., Mori, H., and Cairns, J.R.K. (2015) Identification of rice Os4BGlu13 as a β -glucosidase which hydrolyzes gibberellin A4-1-O- β -D-glucosyl ester, in addition to tuberonic acid glucoside and salicylic acid derivative glucosides. **Arch. Biochem. Biophys.** 583: 36-46.
- Jain, C., Khatana, S., and Vijayvergia, R. (2019). Bioactivity of secondary metabolites of various plants: A review. **IJPSR**. 10(2): 494-504.
- Jatoorathawichot, P., Talabnin, C., Ngiwsara, L., Rustam, Y.H., Svasti J., Reid, G.E., and Cairns, J.R.K. (2020) Effect of Expression of Human Glucosylceramidase 2 Isoforms on Lipid Profiles in COS-7 Cells. **Metabolites**. 10(12): 488.
- Joseph B.C., Pichaimuthu, S., Srimeenakshi, S., Murthy, M., Selvakumar, K., Ganesan M., and Manjunath, S.R. (2015) An overview of the parameters for recombinant protein expression in *Escherichia coli*. **J. Cell. Sci. Ther.** 6: 5

- Kallemeijn, W.W., Witte, M.D., Voorn-Brouwer, T.M., Walvoort, M.T.C., Li, K.Y., Codée, J.D.C., van der Marel, G.A., Boot, R.G., Overkleeft, H.S., and Aerts, J.M. F.G. (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes for the identification of acid/ base residues in human retaining-glucosidases. **J. Biol. Chem.** 289: 35351-35362.
- 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 - glucosidase GBA2 is a non-integral membrane-associated protein at the endoplasmic reticulum (ER) and golgi. **J. Biol. Chem.** 288: 3381-3393.
- Lahiri, S., and Futerman, A.H. (2007) The metabolism and function of sphingolipids and glycosphingolipids. **Cell. Mol. Life. Sci.** 64: 2270-2284.
- Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tauzin, A.S., Spadiut, O., Klintner, S., Pudlo, N.A., Urs, K., Koropatkin, N.M., Creagh, A.L., Haynes, C. A., Kelly, A.G., Cederholm, S.N., Davies, G.J., Martens, E.C., and Brumer, H. (2014) A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. **Nature.** 506: 498-502.
- Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic. Acids. Res.** 42: 490-495.
- Luang, S., Cho, J.I., Mahong, B., Opassiri, R., Akiyama, T., Phasai K., Komvongsa, J., Sasaki, N., Hua, Y., Matsuba, Y., Ozeki, Y., Jeon, J.S., and Cairns, J.R.K. (2013) Rice Os9BGlu31 is a transglucosidase with the capacity to equilibrate phenylpropanoid, flavonoid, and phytohormones glycoconjugates. **J. Biol. Chem.** 288(14): 10111-10123.

- Luttgeharm, K.D., Chen, M., Mehra, A., Cahoon, R.E., Markham, J.E., and Cahoon, E.B. (2015) Overexpression of *Arabidopsis* ceramide synthases differentially affects growth, sphingolipid metabolism, programmed cell death, and mycotoxin resistance. **Plant. Physiol.** 169: 1108-1117.
- Lynch, D.V., and Dunn, T.M. (2004) An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. **New. Phytol.** 161: 677-702.
- Lynch, D.V., and Steponkus, P.L. (1987) Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). **Plant Physiol.** 83: 761-767.
- Majer, P., Neugart, S., Krumbein, A., Schreiner, M., and Hideg, É. (2014) Singlet oxygen scavenging by leaf flavonoids contributes to sunlight acclimation in *Tilia platyphyllos*. **Environ. Exp. Bot.** 100: 1-9.
- Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. **Microbiol. Rev.** 60(3): 512-38.
- Mandal, S.M., Chakraborty, D., and Dey, S. (2010) Phenolic acids act as signaling molecules in plant microbe symbioses. **Plant. Signal. Behav.** 5: 359-368.
- Markham, J.E., Lynch, D.V., Napier, J.A., Dunn, T.M., and Cahoon, B.E. (2013) Plant sphingolipids: function follows form. **Curr. Opin. Plant Biol.** 16: 350-357.
- Matern, H., Boermans, H., Lottspeich, F., and Matern, S. (2001) Molecular cloning and expression of human bile acid β -glucosidase. **J. Biol. Chem.** 276: 37929-37933.

- Melser, S., Bataille, B., Peypelut, M., Poujol, C., Bellec, Y., Boyer, V.W., Peyret, L., M., Faure, J.D., and Moreau, P. (2010) Glucosylceramide biosynthesis is involved in golgi morphology and protein secretion in plant cells. **Traffic**. 11: 479-490.
- Michaelson, L.V., Napier, J.A., Molino, D., and Faure, J.D. (2016) Plant sphingolipids: Their importance in cellular organization and adaptation. **Biochim. Biophys. Acta**. 1861: 1329-1335.
- Michaelson, V.L., Zäuner, S., Markham, E.J., Haslam, P.R., Desikan, R., Mugford, S., Albrecht, S., Warnecke, D., Sperling, P., Heinz, E., and Napier A.J. (2009) Functional characterization of a higher plant sphingolipid Δ^4 -Desaturase: defining the role of sphingosine and sphingosine-1-phosphate in *Arabidopsis*. **Plant Physiol**. 149:487-498.
- Minic, Z. (2008) Physiological roles of plant glycoside hydrolases. **Planta**. 227(4): 723-40.
- Mistrya, P.K., Liu, J., Sun, L., Chuang, W.L., Yuen, T., Yang, R., Lu, P., Zhang, K., Li, J., Keutzer, J., Stachnik, A., Mennone, A., Boyer, J.L., Jain, D., Brady, R.O., New, M.I. and Zaidi, M., (2014) Glucocerebrosidase 2 gene deletion rescues type 1 Gaucher disease. **PNAS USA**. 111(13): 4934-4939.
- Nabavia, S.M., Šamecb, D., Tomczykc, M., Milellad, L., Russod, D., Habtemariame, S., Suntarf, I., Rastrellig, L., Daglah, M., Xiaoi, J., Giampierij, F., Battinoj, M., Sanchezk E. S., Nabavim, S.F., Yousefin, B., Jeandeto, P., Xup, S., and Shirooieq, S. (2020) Flavonoid biosynthetic pathways in plants: Versatile targets for metabolic engineering. **Biotechnol. Adv**. 38:107316.

- Nanda, S., Mohanty, J.N., Mishra, R., and Joshi, R.K. (2016) Chapter 1: Metabolic engineering of phenylpropanoids in plants. **Transgenesis and secondary metabolism**. Edited by Jha. S., 485-510. Springer, Cham.
- Németh, K., Plumb, G.W., Berrin, J.G., Juge, N., Jacob, R., Naim, H.Y., Williamson, G., Swallow, D.M., and Kroon, P.A. (2003) Deglycosylation by small intestinal epithelial cell β -glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans. **Eur. J. Nutr.** 42: 29-42.
- O'Neill, E.C., and Field, R. (2015) Enzymatic synthesis using glycoside phosphorylases. **Carbohydr. Res.** 403: 23-37.
- Opassiri, R., Pomthong, B., Akiyama, T., Nakphaichit, M, Onkoksoong, T., Cairns M.K., and Cairns, J.R.K. (2007) A stress-induced rice (*Oryza sativa* L.) β -glucosidase represents a new subfamily of glycosyl hydrolase family 5 containing a fascin-like domain. **Biochem. J.** 408(2): 241-249.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Cairns J.R.K. (2006) Analysis of rice glycosyl hydrolase family 1 and expression of Os4Bglu12 β -glucosidase. **BMC Plant Biol.** 6: 33.
- Overkleeft, H.S., Renkema, G.H., Neele, J., Vianelloi, P., Hung, O., Strijland, A., van der Burg, A.M., Koomen, G.J., Pandit, U.K., and Aerts, J.M.F.G. (1998) Generation of specific deoxynojirimycin-type inhibitors of the non-lysosomal glucosylceramidase. **J. Biol. Chem.** 273(41): 26522-26527.
- Panche, A.N., Diwan, A.D., and Chandra, S.R. (2016) Flavonoids: an overview. **J. Nutr. Sci.** 5(e47): 1-15.
- Pandey, K.B., and Rizvi, S.I. (2009) Plant polyphenols as dietary antioxidants in human health and disease. **Oxid. Med. Cell. Longev.** 2(5): 270-278.

- Pata, M.O., Hannun, Y.A., and Ng, C.K. (2010) Plant sphingolipids: decoding the enigma of the sphinx. **New Phytol.** 185: 611-630.
- Pekić, B., Zeković, Z., and Lepojević, Ž. (1994) Investigation of apigenin-7-O- β -glucoside hydrolysis by β -glucosidase from almonds. **Biotechnol. Lett.** 16(3): 229-234.
- Rego, A., Trindade, D., Chaves, S.R., Manon, S., Costa, V., Sousa, M.J., and Real, M.C. (2014) The yeast model system as a tool towards the understanding of apoptosis regulation by sphingolipids. **FEMS Yeast Res.** 14: 160-178.
- Rennie, E.A., Ebert, B., Miles, G.P., Cahoon, R.E., Christiansen, K.M., Stonebloom, S., Hoda, K., Twell, D., Petzold, C.J., Adams, P.D., Dupree, P., Heazlewood, J.L., Cahoon, E.B., and Schellera, H.V. (2014) Identification of a sphingolipid α -glucuronosyltransferase that is essential for pollen function in *Arabidopsis*. **Plant Cell.** 26:3314-3325.
- Richelle, M., Merten, P.S., Bodenstab, S., Enslin, M., and Offord, E.A. (2002) Hydrolysis of isoflavone glycosides to aglycones by β -glucosidase does not alter plasma and urine isoflavone pharmacokinetics in postmenopausal women. **J. Nutr.** 132(9): 2587-2592.
- Robert, M.M., Le Bourse, D., Markham, J., Dorey, S., Clément, C., Baillieux, F., and Cordelier D.S. (2015) Modifications of sphingolipid content affect tolerance to hemibiotrophic and necrotrophic pathogens by modulating plant defense responses in *Arabidopsis*. **Plant Physiol.** 169: 2255-2274.
- Roepke, J., and Bozzo, G.G. (2015) *Arabidopsis thaliana* β -glucosidase BGLU15 attacks flavonols 3-O- β -glucoside-7-O- α -rhamnosides. **Phytochem.** 109: 14-24.

- Roepke, J., Gordon H.O.W., Neil, K.J.A., Gidda, S., Mullen, R.T., Coutin, J.A.F., Stone, D.B., and Bozzo, G.G. (2017) An apoplastic β -glucosidase is essential for the degradation of flavonol 3-O- β -glucoside-7-O- α -rhamnosides in *Arabidopsis*. **Plant Cell Physiol.** 58(6): 1030-1047.
- Rosano, G.L., and Ceccarelli, E.A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. **Front Microbiol.** 5(172): 1-17.
- Rugen, M.D., Vernet, M.J.L.M., Hantouti, L., Amalia, S., Andriotis M.E.V., Rejzek, M., Brett, P., Berg, J.B.H.N.R., Aerts, M.F.G.J., Overkleeft, S.H., and Field, A.R. (2018) A chemical genetic screen reveals that iminosugar inhibitors of plant glucosylceramide synthase inhibitor root growth in *Arabidopsis* and cereals. **Sci. Rep.** 8:16421
- Saito, K., Sakakibara, Y.K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., and Fernie, A.R. (2013) The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. **Plant Physiol. Biochem.** 72(21).
- Sansanya, S., Mutoh, R., Charoenwattanasatien, R., Kurisu, G. and Cairns, J.R.K. (2015) Expression and crystallization of a bacterial glycoside hydrolase family 116 β -glucosidase from *Thermoanaerobacterium xylanolyticum*. *Acta Crystallogr. F. Struct. Biol. Commun.* 71: 41-44.
- Santos, E.L., Sales Maia, B.H.L.N, Ferriani, A.P., and Teixeira, S.D. (2017) Chapter 1: Flavonoids: classification, biosynthesis and chemical ecology. **Flavonoids- from biosynthesis to human health.** Edited by Justino. G., Intech Open Limited, United Kingdom.
- Schonauer, S., Körschen, H.G., Penno, A., Rennhack, A., Breiden, B., Sandhoff, K., Gutbrod, K., Dörmann, P., Raju, D.N. , Haberkant, P., Gerl, M.J., Brüggel, B.,

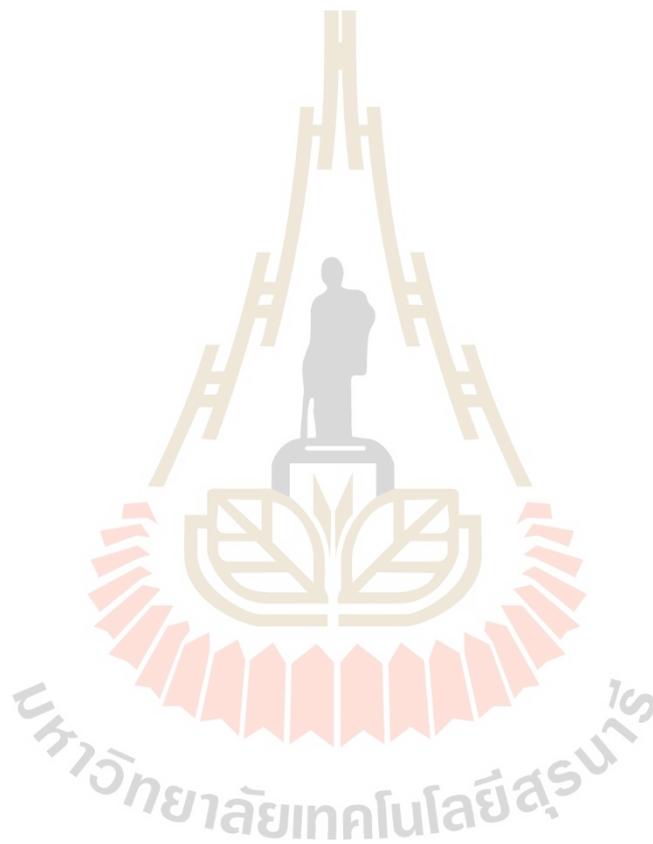
- Zigdon, H., Vardi, A., Futerman, A.H., Thiele, C., and Wachten, D. (2017) Identification of a feedback loop involving beta-glucosidase 2 and its product sphingosine sheds light on the molecular mechanisms in Gaucher disease. **J. Biol. Chem.** 292: 6177-6189.
- Seshadri, S., Akiyama, T., Opassiri R., Kuaprasert, B., and Cairns, J.K. (2009) Structural and enzymatic characterization of Os3BGlu6, a rice β -glucosidase hydrolyzing hydrophobic glycosides and (1 \rightarrow 3)-and (1 \rightarrow 2)-linked disaccharides. **Plant Physiol.** 151(1): 47-58.
- Shi, L., Bielawski, J., Mu, J., Dong, H., Teng, C., Zhang, J., Yang, X., Tomishige, N., Hanada, K., Hannun, A.Y., and Zuo, J. (2007) Involvement of sphingoid bases in mediating reactive oxygen intermediate production and programmed cell death in *Arabidopsis*. **Cell Res.** 17: 1030-1040.
- Smith, C.K., and Fry, S.C. (1999) Biosynthetic origin and longevity in vivo of alpha-D-mannopyranosyl-(1 \rightarrow 4)-alpha-D-glucuronopyranosyl-(1 \rightarrow 2)-myo-inositol, an unusual extracellular oligosaccharide produced by cultured rose cells. **Planta.** 210: 150-156.
- Sorli, S.C., Colié, S., Albinet, V., Dubrac, A., Touriol, C., Guilbaud, N., Bedia, C., Fabriàs, G., Casas, J., Ségui, B., Levade, T., and Abadie, A.N. (2012) The nonlysosomal β -glucosidase GBA2 promotes endoplasmic reticulum stress and impairs tumorigenicity of human melanoma cells. **FASEB. J.** 27: 489-498.
- Ternes, P., Feussner, K., Werner, S., Lerche, J., Iven, T., Heilmann, I., Riezman, H., and Feussner, I. (2011) Disruption of the ceramide synthase LOH1 causes spontaneous cell death in *Arabidopsis thaliana*. **New Phytol.** 192: 841-854.

- Tiwari, R., and Rana, C.S. (2015) Plant secondary metabolites: a review. **Int. J.Eng.** 3(5).
- Tohge, T., Watanabe, M., Hoefgen, R., Fernie, R.A. (2013) Shikimate and phenylalanine biosynthesis in the green lineage. **Front. Plant Sci.** 4(62): 1-13.
- Treviño, L.L.E., Chen, W., Card, M.L., Shih, M.C., Cheng, C.L., and Poulton, J.E. (2006) *Arabidopsis thaliana* beta-Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. **Phytochem.** 67(15): 1651-60.
- Uemura, M., and Steponkus, P.L. (1994) A contrast of the plasma membrane lipid composition of oat and rye leaves in relation to freezing tolerance. **Plant Physiol.** 104: 479-496.
- van Weely, S., Brandsma, M., Strijland, A., Tager, J.M., and Aerts, J.M. (1993) Demonstration of the existence of a second, non-lysosomal glucocerebrosidase that is not deficient in Gaucher disease. **Biochim. Biophys. Acta.** 118: 55-62.
- van Wyk, N., Drancourt, M., Henrissat, B., and Kremer, L. (2017) Current perspectives on the families of glycoside hydrolases of *Mycobacterium tuberculosis*: their importance and prospects for assigning function to unknowns. **Glycobiology.** 27: 112-122.
- Vigerust, D.J., and Shepherd, V.L. (2007) Virus glycosylation: role in virulence and immune interactions. **Trends Microbiol.** 15:211-218.
- Votsi, C., Papanicolaou, Z.E., Middleton, L.T., Pantzaris, M., and Christodoulou, K.A. (2014) Novel GBA2 gene missense mutation in spastic ataxia. **Ann. Hum. Genet.** 78: 13-22.
- Vuong, T.V., and Wilson, D.B. (2010) Glycoside Hydrolases: Catalytic base/nucleophile diversity. **Biotechnol. Bioeng.** 107: 195-205.

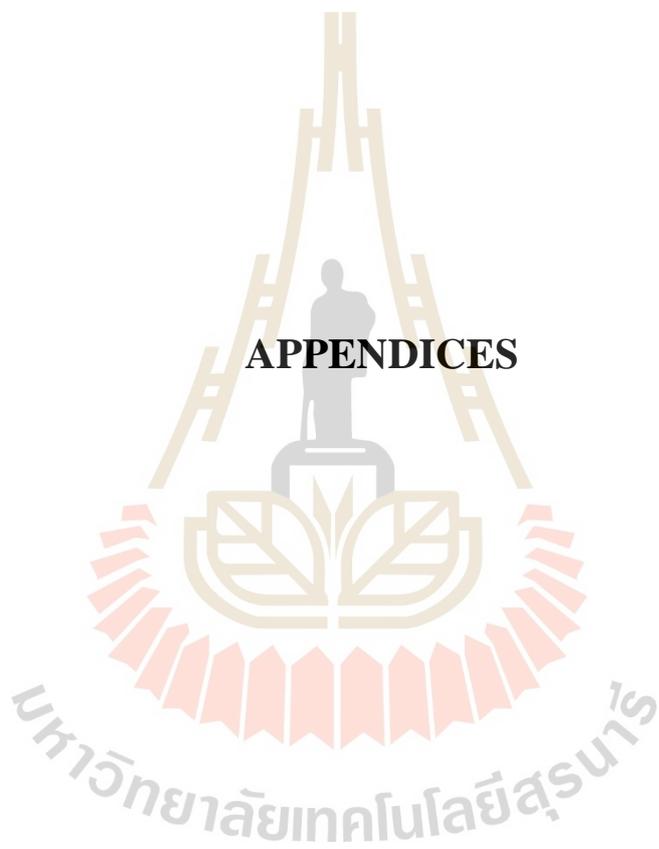
- Wang, W., Yang, X., Tangchaiburana, S., Ndeh, R., Markham, J.E., Tsegaye, Y., Dunn, T.M., Wang, G.L., Bellizzi, M., and Parsons, J.F. (2008) An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in *Arabidopsis*. **Plant Cell**. 20: 3163-79.
- Wang, W., Yang, X., Tangchaiburana, S., Ndeh, R., Markham, J.E., Tsegaye, Y., Dunn, T.M., Wang, G.L., Bellizzi, M., and Parsons, J.F. (2008) An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in *Arabidopsis*. **Plant Cell**. 20: 3163-3179.
- Warnecke, D., and Heinz, E. (2003) Recently discovered functions of glucosylceramides in plants and fungi. *Cell. Mol. Life Sci.* 60: 919-941.
- Williams, R., Spencer, J., and Rice Evans, C. (2004) Flavonoids: antioxidants or signaling molecules?. **free radic. Biol. Med.** 36: 838-849.
- Worrall, D., Liang, Y.K., Alvarez, S., Holroyd, G.H., Spiegel, S., Panagopoulos, M., Gray, J.E., and Hetherington, A.M. (2008) Involvement of sphingosine kinase in plant cell signaling. **Plant J.** 56: 64-72.
- Wu, J. X., Li, J., Liu, Z., Yin, J., Chang, Z.Y., Rong, C., Wu, J.L., Bi, F.C., and Yao, N. (2015) The *Arabidopsis* ceramidase AtACER functions in disease resistance and salt tolerance. **Plant J.** 81:767-80.
- Yeom, S.J., Kim, B.N., Kim, Y.S., and Oh, D.K. (2012) Hydrolysis of isoflavone glycosides by a thermostable β -glucosidase from *Pyrococcus furiosus*. **J. Agric. Food. Chem.** 60(6): 1535-41.

- Yildiz, Y., Hoffmann, P., Vom Dahl, S., Breiden, B., Sandhoff, R., Niederau, C., Horwitz, M., Karlsson, S., Filocamo, M., Elstein, D., Beck, M., Sandhoff, K., Mengel, E., Gonzalez, M. C., Nöthen, M. M., Sidransky, E., Zimran, A., and Mattheisen, M. (2013) Functional and genetic characterization of the non-lysosomal glucosylceramidase 2 as a modifier for Gaucher disease. **Orphanet. J. Rare. Dis.** 8: 1-8.
- Yildiz, Y., Matern, H., Thompson, B., Allegood, J.C., Warren, R. L., Ramirez, D.M.O., Hammer, R. E., Matern, F.K.H.S., and Russell, D.W. (2006) Mutation of β -glucosidase 2 causes glycolipid storage disease and impaired male fertility. **J. Clin. Invest.** 116: 2985-2994.
- Yu, Z.C., Zhenga, X.T., Lin, W., He, W., Shao, L., and Penga, C.L. (2021) Photoprotection of *Arabidopsis* leaves under short-term high light treatment: The antioxidant capacity is more important than the anthocyanin shielding effect. **Plant Physiol. Biochem.** 166: 258-269.
- Zhang, C.P., Liu, G., and Hu, F.L. (2012) Hydrolysis of flavonoid glycosides by *propolis* β -glycosidase. **Nat. Prod. Res.** 26(3): 270-273.
- Zhang, K., Sun, Y., Li, M., and Ruicai, L. (2021) CrUGT87A1, a UDP-sugar glycosyltransferases (UGTs) gene from *Carex rigescens*, increases salt tolerance by accumulating flavonoids for antioxidation in *Arabidopsis thaliana*. **Plant Physiol. Biochem.** 159: 28-36.
- Zheng, P., Wu, J., Sahu, S.K., Zeng, H.Y., Huang, L.Q., Liu, Z., Xiao, S., and Yao, N., (2018) Loss of alkaline ceramidase inhibits autophagy in *Arabidopsis* and plays an important role during environmental stress response. **Plant. Cell. Environ.** 41(4): 837-849.

Zienkiewicz, A., Jasmin, G., Konig, S., Herrfurth, C., Liu, Y.T., Meldau, D., and Feussner, I. (2019) Disruption of *Arabidopsis* neutral ceramidases 1 and 2 results in specific sphingolipid imbalances triggering different phytohormone-dependent plant cell death programmes. **New. Phytol.** 226: 170-188.



APPENDICES



APPENDIX A

DETECTION OF AGLYCONE COMPOUNDS BY UHPLC

METHOD

1. Methodology

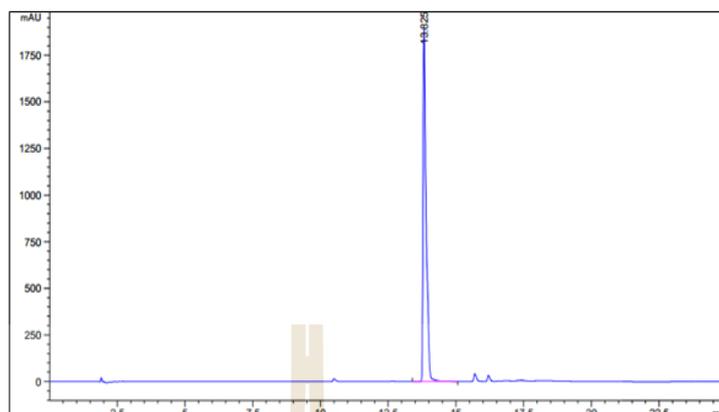
An Agilent 1290 Infinity ultra-high performance liquid chromatography system was used to separate and detect flavonoids aglycones as shown in the section 3.3.8.

2. Standard chromatograms

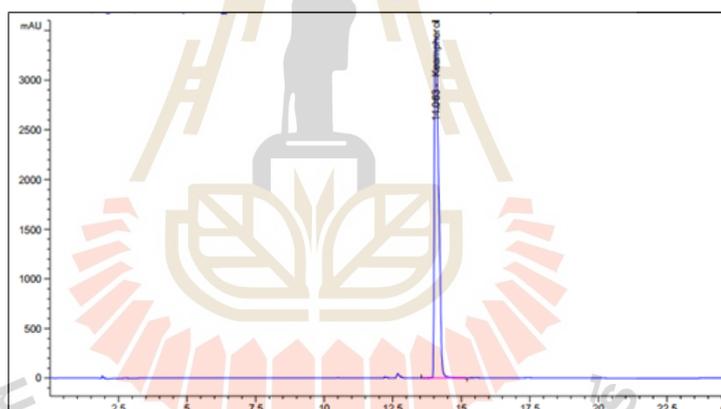
Table AI 1 Retention time of measured compounds

Flavonoid compounds	Time (min)
Quercetin	12.648
Kaempferol	14.086
Luteolin	12.59
Apigenin	13.825

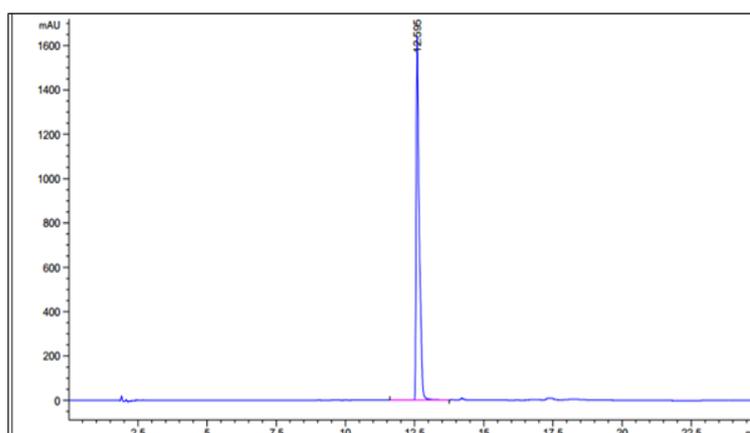
2.1 Single injection of apigenin standard



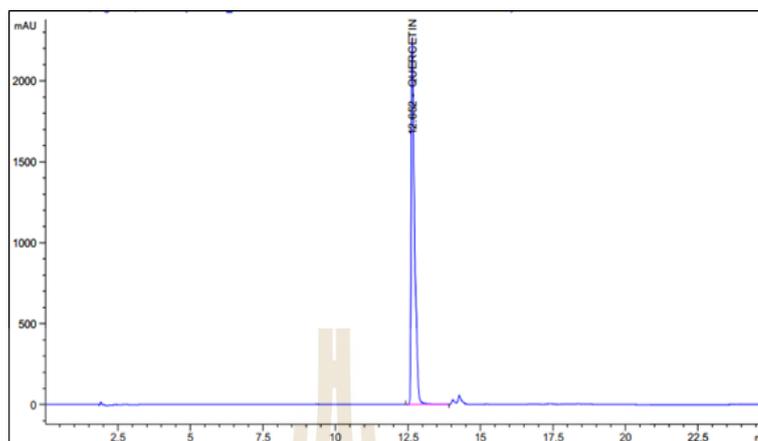
2.2 Single injection of kaempherol standard



2.3 Single injection of luteolin standard

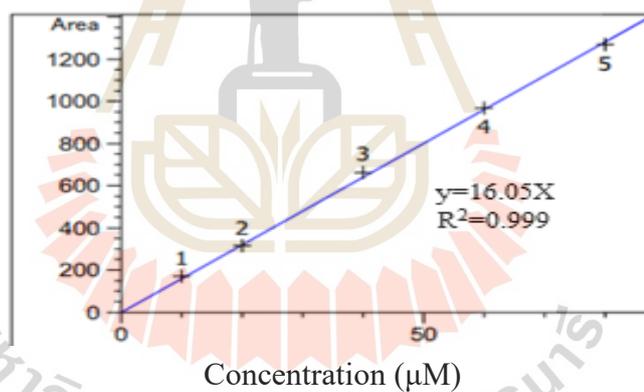


2.4 Single injection of quercetin standard

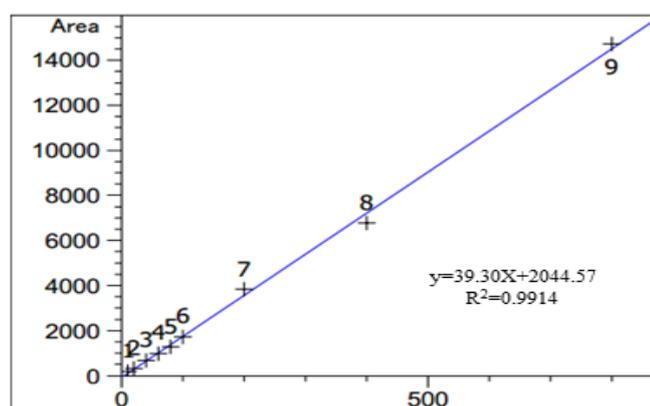


3. Standard curves

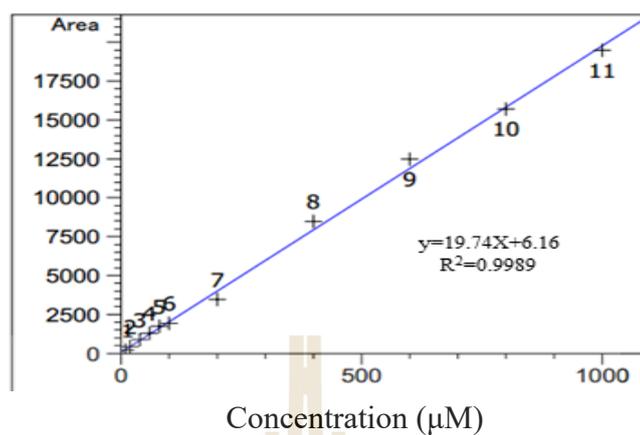
3.1 Standard curve for apigenin



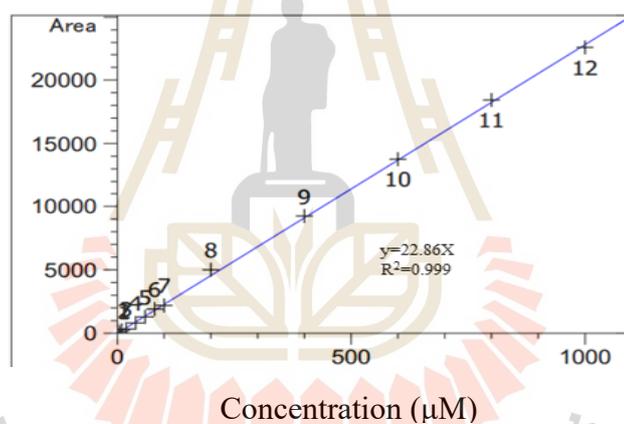
3.2 Standard curve for kaempferol



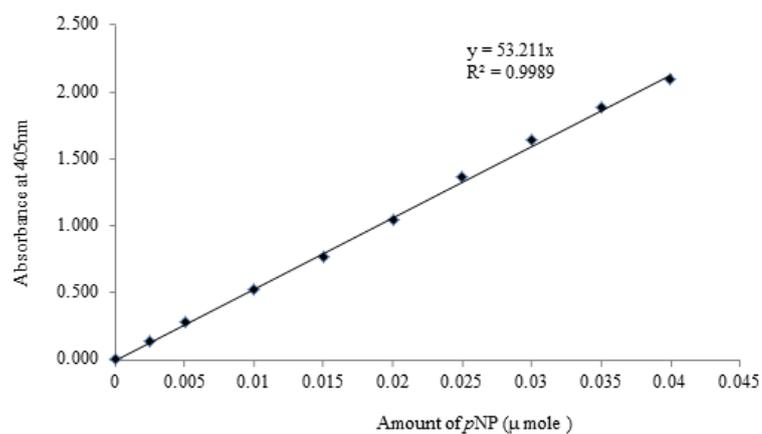
3.3 Standard curve for luteolin



3.4 Standard curve for quercetin



3.5 Standard curve for pNP



APPENDIX B

PUBLICATIONS

- Sharma, R., Karunambigai, A., Gupta, S., Arumugan, N. (2021) Evaluation of biologically active secondary metabolites isolated from the toothache plant *Acmella ciliata* (Asteraceae). Adv. Tradit. Med. <https://doi.org/10.1007/s13596-021-00584-5>.
- Karunambigai, A., Devi, S.G. (2015) Assessment of enzymic, non-enzymic antioxidants and in vitro free radical scavenging activities of different extracts of leaves and roots of *Eclipta alba*. J. adv. biol. 4(4): 1-10.
- Karunambigai, A., Devi, S.G. (2013) Antibacterial activity of leaves and roots of *Eclipta alba*. Int J Pharm Sci. 6(1):454-456.
- Devi, S.G., Tamilarasi, K., Priyadharshini, R., Karunambigai, A. (2010) Antioxidant activity of *Acalypha indica*. Adv.Appl.Res.3(1):36-41.
- Priyadarsini, R, Tamilarasi, K., Karunambigai, A., Gayathri, D.S. (2010) Antioxidant potential of the leaves and roots of *Solanum surattense*. Plant Arch.10(2):815–818

CURRICULUM VITAE

NAME Miss Karunambigai Arthanareeswaran

DATE OF BIRTH 26 November, 1987

PLACE OF BIRTH Erode, Tamilnadu, India.

INSTITUTIONS ATTENDED

2008-2010 Master of Science, Department of
Biotechnology, Avinashilingam University for
Women, Coimbatore. Tamilnadu, India.

2005-2008 Bachelor of Science, Department of
Biotechnology, Hindusthan College of Arts &
Science, Coimbatore. Tamilnadu, India.

มหาวิทยาลัยเทคโนโลยีสุรนารี